Is Spin-Component Scaled Second-Order Møller−Plesset Perturbation Theory an Appropriate Method for the Study of Noncovalent Interactions in Molecules?

Jens Antony and Stefan Grimme*

Theoretische Organische Chemie, Organisch-Chemisches Institut der Universität Münster, Corrensstrasse 40, D-48149 Münster, Germany

Received: January 23, 2007; In Final Form: March 21, 2007

Testing of the spin-component scaled second-order Møller−Plesset (SCS-MP2) method for the computation of noncovalent interaction energies is done with a database of 165 biologically relevant complexes. The effects of the spin-scaling procedure (i.e., MP2 vs SCS-MP2), the basis set size, and the corrections for basis set superposition error (BSSE) are systematically examined. When using two-point basis set extrapolations for the correlation energy, augmentation of the atomic orbital basis with computationally costly diffuse functions is found to be obsolete. In general, SCS-MP2 also improves results for noncovalent interactions statistically on MP2, and significant outliers are removed. Moreover, it is shown that effects of BSSE and one-particle basis set incompleteness almost cancel each other in the case of triple-ζ sets (SCS-MP2/TZVPP or SCS-MP2/cc-pVTZ without counterpoise correction), which opens a practical route to efficient computations for large systems. We recommend SCS-MP2 as the preferred quantum chemical wave function based method for the noncovalent interactions in large biologically relevant systems when reasonable coupled-cluster with single and double and perturbative triple excitations (CCSD(T)) calculations cannot be performed anymore. A comparison to MP2 and CCSD(T) interaction energies for *n*-alkane dimers, however, indicates (and this also holds to a lesser extent for hydrogen-bonded systems) limitations of SCS-MP2 when treating chemically “saturated” interactions. The different behavior of second-order perturbation theory for saturated and for stacked π-systems is discussed.

1. Introduction

The fundamental importance of noncovalent (intermolecular) interactions to many fields of science cannot be overestimated. An understanding of real systems and their properties and of all condensed-phase matter must be rooted in the knowledge of noncovalent interactions. Consequently, they are nowadays considered as cornerstones in supramolecular chemistry, materials science, and biochemistry.[1−4] Although a very detailed understanding on an atomic or a molecular level is still lacking, important progress has been achieved in recent years in the quantum mechanical description of the relevant forces.[5] However, still problematic is the accurate account of the dispersion (van der Waals, vdW) part of the interactions for large systems. Dispersion interactions are ubiquitous, long-range attractive forces that act between separated molecules or fragments 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.[5] Because of their relatively long-ranged (R^{-6} dependence with intermolecular distance) character (as compared to other quantum mechanical effects), dispersion contributions to inter- and intramolecular energies become increasingly significant for larger systems that are relevant in supramolecular or nanochemistry.[6,7]

Coupled-cluster theory[8] is currently the most accurate wave-function-based method for calculating dispersion interactions. Because of the N^7 scaling of the computational effort for the coupled-cluster with single and double and perturbative triple excitations (CCSD(T)) method, where N characterizes the size of the electronic system, calculations on realistic complexes have to be performed with relatively small one-particle basis sets. Usually the basis set dependence of the correlation energy is estimated with lower-level electronic structure methods such as second-order Møller−Plesset perturbation theory (MP2).[9−11] Within this standard procedure, systems containing up to 50 atoms have been treated. Beyond this limit, reasonable CCSD(T) calculations cannot be executed with present computational resources, and MP2 alone is not accurate enough. Furthermore, small basis sets make a treatment of the basis set superposition error (BSSE) necessary, for example, with the laborious counterpoise correction.[12] The goal of describing macromolecular systems with, for example, several hundreds of atoms and with a consistent treatment of intramolecular dispersion requires alternative approaches.

Recently,[13] the capability of density functional theory (DFT) including dispersion corrections (DFT−D) to calculate intermolecular interaction energies was tested by comparison with CCSD(T) complete basis set (CBS) limit estimates for an extensive benchmark set provided by Jurecka et al.[14] It consists mainly of DNA base pairs and amino acid pairs in a wide variety of realistic geometries and is expected to yield conclusive insight into the performance of quantum chemical methods. It was shown that the DFT−D interaction energies deviate on average by less than 1 kcal·mol^{-1}, or 10%, from the reference values, which thereby substantially improves the results of pure DFT.

Despite this and other successes (see ref 6 for a recent review), the description of dispersion interactions in DFT is mostly

---

* To whom correspondence should be addressed. E-mail: grimmes@uni-muenster.de; fax: (+49)251-83-36515.

10.1021/jp070589p CCC: $37.00

empirical (similar to many density functionals themselves), which may limit the accuracy of unconventional systems. The DFT−D approach cannot, for example, account for the influence of changes in the molecular polarizability on the strength of the vdW interaction, because the C_6 coefficients are fixed (system independent) parameters.[15] If in a series of structurally similar molecules of increasing size the HOMO−LUMO gap gets smaller, then DFT will account for the increase of the fragment polarizability (induction effects), but the dispersion correction to the intermolecular interaction energies remains (pairwise) additive.

To take care of system-dependent vdW effects, we currently have to resort to wavefunction-based methods (for a recent DFT approach to this problem see refs 16 and 17). Because dispersion is an electron correlation effect, a level of at least second-order perturbation theory (e.g., MP2) has to be applied. Whereas it is rather successful in the case of saturated or hydrogen-bonded molecules, MP2 quite strongly overestimates dispersion in π-stacked systems (see, e.g., refs 18 and 19).

A modification of MP2, in which the total MP2 correlation energy is partitioned into parallel- and antiparallel-spin components that are separately scaled, is denoted as SCS-MP2.[20] In general, this method improves in almost all respects on standard MP2, and many successful applications in various areas of chemistry are documented.[7,21−24] Preliminary studies on systems with important dispersive interactions already indicated significant improvements as compared to standard MP2.[20,25,26] This has been recently confirmed by local SCS-LMP2 calculations of the benzene dimer.[19] The SCS-LMP2 interaction energies for the parallel-displaced and sandwich structures are in excellent agreement with the best available literature values along the entire potential energy curves, whereas for the T-shaped structure SCS-LMP2 slightly underestimates the binding energy.[19] Therefore, SCS-MP2 appears to be a suitable method for investigating larger π-systems.

A variant of SCS-MP2 where only the (scaled) opposite-spin correlation energy is retained (termed SOS-MP2) has been proposed by Jung et al.[27] This approach has been further developed to become asymptotically correct for vdW interactions (MOS-MP2).[28] Although promising applications of these OS-MP2 methods have recently been published for very large vdW systems,[29] a comprehensive study of their performance for a wide variety of vdW complexes is still missing. Very recently, specially adjusted scaling parameters (denoted as SCSN) for weak and stacking interactions have been proposed by Hill and Platts.[30]

To get a more conclusive picture about the performance of MP2 and related SCS methods, the current work tests the SCS-MP2 method with the database of 165 noncovalent complexes[14] that has already been used in the aforementioned DFT−D study.[13] The effects of the spin-scaling procedure (i.e., MP2 vs SCS-MP2), the basis set size, and the BSSE are systematically examined. We also tested SOS-MP2[27], but because it behaves qualitatively very similar to SCS-MP2 (in fact slightly worse on average), we find it inconvenient and not very enlightening to further expand the data presentation and discussion. Therefore, we concentrate on SCS-MP2 and MP2 and briefly present only a few representative statistical data for SOS-MP2 (and SCSN-MP2) in section 3.4 (additionally, all data can be found in Supporting Information).

It is shown here comprehensively, that for noncovalent interactions SCS-MP2 improves, in general (statistically), on MP2. Moreover, we find that effects of the BSSE and of the one-particle basis set almost cancel each other in the case of triple-ζ basis sets (SCS-MP2/TZVPP or SCS-MP2/cc-pVTZ without counterpoise correction), which opens a practical route to efficient computations for large systems. A comparison to MP2 and CCSD(T) interaction energies for n-alkane dimers,[31] however, indicates limitations of SCS-MP2 when treating saturated molecules.

2. Technical Details

All calculations were performed with the TURBOMOLE suite of programs,[32,33] and the resolution of identity (RI) approach[34,35] was used for the correlation energy. The Hartree−Fock calculations were done without approximation. Because the error introduced by the RI treatment is insignificant (<0.1−0.3% of the interaction energy ΔE) as compared to other (e.g., basis set) effects, we use the terms SCS-MP2 and MP2 instead of RI-SCS-MP2 and RI-MP2, respectively.

In the SCS-MP2 method, the second-order correlation correction is scaled according to the following equation:

$$E^{\mathrm{corr}}(\mathrm{SCS} - \mathrm{MP2}) = p_{\mathrm{S}} E_{\uparrow\downarrow} + p_{\mathrm{T}} E_{\uparrow\uparrow+\downarrow\downarrow}$$

where $E_{\uparrow\uparrow+\downarrow\downarrow}$ and $E_{\uparrow\downarrow}$ are the second-order perturbation contributions from double excitations of electron pairs with parallel- and antiparallel-spin, respectively, and $p_{\mathrm{S}} = (6/5)$ and $p_{\mathrm{T}} = (1/3)$ are the default scaling parameters.[20]

Basis sets were taken from the TURBOMOLE library.[36] The BSSE is treated in the usual way by the counterpoise correction (CP).[12] The CBS limit is determined by a standard two-point extrapolation, which assumes an exponential dependence of the Hartree−Fock energy and an inverse cubic dependence of the MP2 correction on the cardinal number X of the basis set.[37]

$$E_X^{\ \mathrm{HF}} = E_{\mathrm{CBS}}^{\mathrm{HF}} + A \exp(-\alpha X)$$

and,

$$E_X^{\ \mathrm{corr}} = E_{\mathrm{CBS}}^{\mathrm{corr}} + BX^{-3}$$

The exponent was chosen as $\alpha = 1.63$.[38] The geometries of the molecules in the benchmark set were taken from ref 14. In the case of the alkane dimers, we optimized structures according to ref 31 at the MP2/6-311G** level with fixed MP2/6-31G* geometries of the monomers to allow for direct comparison with the literature data.

3. Results and Discussion

3.1 Benchmark Dataset of 22 Small Model Complexes. Results for individual systems in the CBS limit are given in Table 1. Table 2 collects the average absolute deviations (MAD) and root-mean-square deviations (RMSD) of SCS-MP2 and MP2 ΔE values with and without the CP from estimated CCSD(T)/CBS results[14] for the 22 complexes listed as a function of the basis set. Figure 1 graphically compares SCS-MP2 and MP2 ΔE values with the cc-pVTZ sets and at the CBS limit to the reference values from Jurecka et al.[14]

For the aug-cc-pVTZ basis, our CP-corrected MP2 and SCS-MP2 values can be directly compared to those from ref 19 that were obtained with the same atomic orbital set and used the (CP-uncorrected) local correlation approach. The deviations between the results of the two studies are in general very small, and the MAD agree quite well (1.13 vs 1.26 (SCS-MP2) and 0.75 vs 0.81 (MP2) kcal·mol^{-1}), which demonstrates the reliability of RI and local approximations for our problem. For comparison, the MAD and RMSD for the 22 complexes at the DFT-D level

TABLE 1: SCS-MP2 and MP2 ΔE Values[a] for 22 Small Model Complexes (S22) in the Complete Basis Set Limit[b]

	SCS-MP2				MP2				
	no CP[d]		CP[d]		no CP[d]		CP[d]		
complex (symmetry)	ΔE	dev[e]	ΔE	dev[e]	ΔE	dev[e]	ΔE	dev[e]	ΔE_{ref}[c]
Hydrogen-Bonded Complexes (7)									
$(NH_3)_2$ (C_{2h})	−2.77	0.40	−2.73	0.44	−3.19	−0.02	−3.16	0.01	−3.17
$(H_2O)_2$(C_s)	−4.60	0.42	−4.50	0.52	−5.04	−0.02	−4.96	0.06	−5.02
formic acid dimer (C_{2h})	−17.31	1.30	−16.86	1.75	−18.91	−0.30	−18.52	0.09	−18.61
formamide dimer (C_{2h})	−14.63	1.33	−14.40	1.56	−15.98	−0.02	−15.79	0.17	−15.96
uracil dimer (C_{2h})	−18.66	1.99	−18.46	2.19	−20.52	0.13	−20.37	0.28	−20.65
2-pyridoxine•2-aminopyridine	−15.49	1.22	−15.31	1.40	−17.47	−0.76	−17.35	−0.64	−16.71
adenine•thymine WC	−14.79	1.58	−14.62	1.75	−16.64	−0.27	−16.52	−0.15	−16.37
Complexes with Predominant Dispersion Contribution (8)									
$(CH_4)_2$ (D_{3d})	−0.31	0.22	−0.31	0.22	−0.49	0.04	−0.49	0.04	−0.53
$(C_2H_4)_2$(D_{2d})	−1.05	0.46	−1.03	0.48	−1.59	−0.08	−1.58	−0.07	−1.51
benzene•CH_4(C_3)	−1.12	0.38	−1.16	0.34	−1.77	−0.27	−1.82	−0.32	−1.50
benzene dimer (C_{2h})	−2.82	−0.09	−2.83	−0.10	−4.92	−2.19	−4.95	−2.22	−2.73
pyrazine dimer (C_s)	−4.67	−0.25	−4.60	−0.18	−6.94	−2.52	−6.90	−2.48	−4.42
uracil dimer (C_2)	−8.53	1.59	−8.29	1.83	−11.27	−1.15	−11.09	−0.97	−10.12
indole•benzene	−4.87	0.35	−4.88	0.34	−8.03	−2.81	−8.08	−2.86	−5.22
adenine•thymine stack	−10.86	1.37	−10.62	1.61	−14.99	−2.76	−14.84	−2.61	−12.23
Mixed Complexes (7)									
ethene•ethyne (C_{2v})	−1.33	0.20	−1.32	0.21	−1.66	−0.13	−1.66	−0.13	−1.53
benzene•H_2O(C_s)	−2.93	0.35	−2.89	0.39	−3.56	−0.28	−3.54	−0.26	−3.28
benzene•NH_3(C_s)	−1.99	0.36	−2.01	0.34	−2.63	−0.28	−2.66	−0.31	−2.35
benzene•HCN (C_s)	−4.27	0.19	−4.24	0.22	−5.16	−0.70	−5.16	−0.70	−4.46
benzene dimer (C_{2v})	−2.40	0.34	−2.46	0.28	−3.55	−0.81	−3.63	−0.89	−2.74
indole•benzene T−shape	−5.24	0.49	−5.29	0.44	−6.89	−1.16	−6.97	−1.24	−5.73
phenol dimer	−6.37	0.68	−6.30	0.75	−7.79	−0.74	−7.74	−0.69	−7.05

[a] Measured in kcal•mol^{-1}. [b] aug-cc-pVTZ/aug-cc-pVQZ. [c] Estimated CCSD(T)/CBS.[14] [d] Counterpoise correction. [e] $\Delta E - \Delta E_{\mathrm{ref}}$.

TABLE 2: Statistics of the Deviation of SCS-MP2 and MP2 ΔE Values[a] from the Estimated CCSD(T)/CBS Reference Data[14] for 22 Small Complexes (S22 Set)[b]

	SCS-MP2		MP2	
	no CP[c]	CP[c]	no CP[c]	CP[c]
CBS(aT-aQ)[d]	0.71/0.90/2.24	0.79/1.03/2.37	0.79/1.21/2.94	0.78/1.19/3.13
aug-cc-pVQZ	0.51/0.67/2.32	0.92/1.21/2.56	1.13/1.61/3.81	0.75/1.11/3.30
aug-cc-pVTZ	0.72/0.90/2.93	1.13/1.47/3.00	1.72/2.31/5.50	0.75/1.06/3.58
aug-cc-pVDZ	1.69/2.27/7.02	1.69/2.12/4.23	2.73/3.73/9.42	0.88/1.19/4.54
CBS(T−Q)[d]	0.52/0.74/2.10	0.78/1.03/2.31	1.05/1.46/3.36	0.80/1.21/3.12
cc-pVQZ	0.47/0.67/2.06	1.15/1.44/2.80	1.08/1.46/3.52	0.63/0.90/3.10
cc-pVTZ	0.40/0.50/2.04	1.70/2.07/3.77	1.35/1.72/4.35	0.70/0.91/3.12
cc-pVDZ	0.61/0.82/2.51	3.12/3.69/7.12	1.54/1.88/4.08	2.03/2.49/4.55
CBS(T−Q)[d]	0.61/0.81/2.26	0.79/1.03/2.45	0.94/1.38/3.27	0.82/1.25/3.26
QZVP	0.60/0.81/2.36	1.10/1.40/2.77	0.95/1.43/3.43	0.68/0.98/3.29
TZV(2df,2pd)	0.53/0.76/2.55	1.57/1.94/3.56	1.10/1.60/4.24	0.70/0.90/3.30
TZV(2d,2p)	0.58/0.81/2.80	1.74/2.15/3.93	1.16/1.68/4.58	0.77/1.02/3.55
SV(d,p)	0.71/0.95/2.53	3.07/3.60/7.36	1.80/2.15/4.89	2.02/2.41/4.48

[a] Measured in kcal•mol^{-1}. [b] Given are the mean absolute deviation/root-mean-square deviation/error spread (largest positive minus largest negative deviation). [c] Counterpoise correction. [d] Extrapolated to complete basis set limit.

are also given: 0.47 and 0.58 kcal•mol^{-1} with B-LYP-D, 0.82 and 1.05 kcal•mol^{-1} with PBE-D, and 0.35 and 0.46 kcal•mol^{-1} with B97-D, using the TZV(2df, 2pd) basis.[13]

In general, the CP-correction drops by roughly an order of magnitude upon the two point extrapolation, from an average value of −0.6 kcal•mol^{-1} for the aug-cc-pVQZ basis to −0.09 and −0.06 kcal•mol^{-1} for SCS-MP2 and MP2 in the extrapolated CBS limit, respectively. Consequently, the uncorrected and corrected ΔE values almost coincide (see Table 1 and Figure 1). The CBS limit intermolecular interaction energies are practically free from BSSE and finite basis error and are dominated by the methodical (correlation energy) error. These values are therefore mainly discussed in the following. We note that in the CBS limit the results with the cc-pVXZ or TZV-QZV valence sets and those with the augmented basis sets (aug-cc-pVXZ) are very similar. This shows that the effect of the computationally costly diffuse functions on the interaction energies is accounted for by the extrapolation procedure.

A closer inspection of the data in Table 1 reveals that SCS-MP2 underestimates the ΔE values for the hydrogen-bonded complexes whereas MP2 slightly overestimates them. For the second group of complexes that are mostly stacked, unsaturated π-systems dominated by dispersion, MP2 shows the well-known overbinding behavior. SCS-MP2 performs much better for these and also for the complexes in group three (mixed type). For the entire set, the SCS-MP2 values systematically deviate from the CCSD(T)/CBS reference to smaller values, with the deviation being roughly proportional to the magnitude of the interaction energy (see Figure 1). The slope of the linear regression line through the CP-uncorrected values is 0.915, and it crosses the ordinate at a ΔE value of 5×10^{-2} kcal•mol^{-1}. The MP2 ΔE values, in contrast, overestimate the reference,

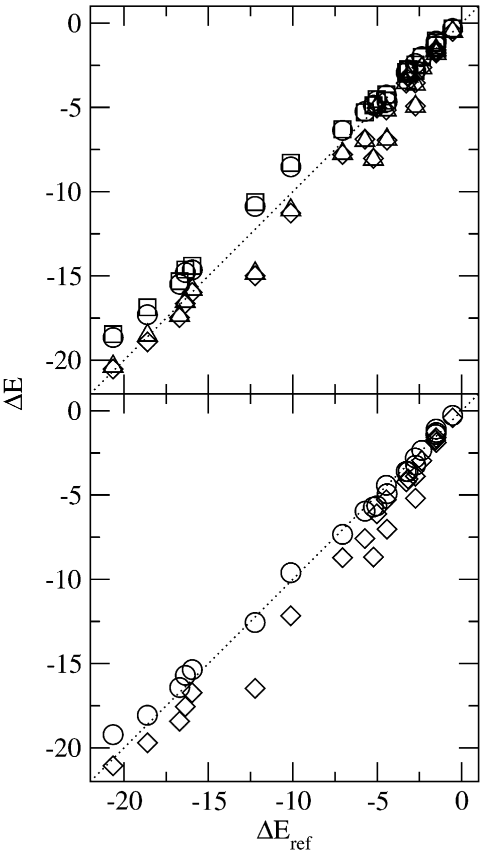

Figure 1. ΔE values (kcal·mol^{-1}) for 22 small complexes (S22) vs estimated CCSD(T)/CBS reference data[14]: Top cc-pVTZ, bottom aug-cc-pVTZ/aug-cc-pVQZ extrapolated to complete basis set limit. Circles and squares: SCS-MP2 without and with CP, respectively. Diamonds and triangles: MP2 without and with CP, respectively.

and the largest deviations occur for the stacked complexes with ΔE values around −5 to −10 kcal·mol^{-1}, whereas the reference is closely reproduced by MP2 for the hydrogen-bonded dimers with values between −15 and −20 kcal·mol^{-1} (see Table 1). These trends are also present in the results with the TZV(2df, 2pd) and cc-pVTZ basis sets, where the linear regression line through the CP-uncorrected SCS-MP2 results have slopes of 0.913 and 0.950, and ordinate ΔE values of −0.318 and −0.304 kcal·mol^{-1}, respectively. Quite importantly, the SCS-MP2 values obviously scatter less in Figure 1 than in the case of MP2 as is also indicated by the linear regression correlation coefficients of 0.9989 (SCS-MP2) and 0.9891 (MP2), respectively, with the CP-uncorrected data at the extrapolated (aT-aQ) level.

The overall slightly better performance of SCS-MP2 is also reflected in the statistical data provided in Table 2. The MAD of the CP-uncorrected SCS-MP2 and MP2 ΔE values in the CBS limit from the reference are 0.7 and 0.8 kcal·mol^{-1}, respectively. This improvement over standard MP2 holds for all three groups of basis sets (cc-pVXZ, aug-cc-pVXZ, and SV-TZV-QZV) where SCS-MP2 outperforms MP2 with MAD values of 0.5−0.7 kcal·mol^{-1} as compared to 0.8−1.1 kcal·mol^{-1}. Also the error spread is consistently smaller with SCS-MP2.

The data in Table 2 also very clearly demonstrate the special behavior of double-ζ type basis sets (SVP and VDZ). These are very incomplete and always yield the largest deviation in one column of the table. In particular, augmentation of the cc-pVDZ basis leads to serious BSSE effects. Because these cannot be corrected for in intramolecular cases, computations with aug-cc-pVDZ or similar sets should be avoided.

Very small MADs (0.4−0.6 kcal·mol^{-1}) and error ranges (<3 kcal·mol^{-1}) are obtained at the CP-uncorrected SCS-MP2 level. Very interestingly, the MAD is minimal for SCS-MP2/cc-pVTZ (0.4 kcal·mol^{-1}), and for cc-pVQZ, aug-cc-pVQZ, and TZV-(2df,2pd) the MADs are also only about 0.5 kcal·mol^{-1}. It seems that the BSSE, one-particle basis set deficiencies, and the correlation error almost cancel for SCS-MP2/TZV(2df, 2pd) (or SCS-MP2/cc-pVTZ) without the CP-correction, and the error increases again for larger basis sets and in the CBS limit. For MP2 we could not observe a similar effective error compensation as the non-CP-corrected MAD values are for triple-ζ and bigger sets that are larger than the corresponding CP-corrected MAD values. The MAD of the MP2/CBS ΔE values (0.8 kcal·mol^{-1}) is equal to the average absolute CCSD(T) correction of the 22 reference values (see Table 3 of ref 14). The slightly better performance of SCS-MP2/cc-pVTZ versus SCS-MP2/TZV(2df, 2pd) is attributed to the, in general, larger BSSE of the Dunning as compared to the Ahlrichs type basis sets.

We also investigated a truncated triple-ζ basis set which misses the most costly part of the polarization functions (TZV-(2d, 2p)). As can be seen from the comparison with the corresponding TZV(2df, 2pd) or cc-pVTZ data, this causes overall only a small increase of the error (MAD(no CP) of 0.53 and 0.58 kcal·mol^{-1}, respectively) but saves a considerable amount of computation time, which is important for much larger molecules.

3.2 Full Benchmark Dataset. As a first test for the range of validity of the observed error compensation between BSSE, finite basis, and incomplete correlation treatment, the ΔE values for the entire JSCH2005 benchmark dataset were calculated at the SCS-MP2/TZV(2df, 2pd), MP2/TZV(2df, 2pd) (see Figure 2), and SCS-MP2/TZV(2d,2p) levels.

The MAD of 161 ΔE values from the estimated CCSD(T)/CBS reference are 0.79 and 0.84 kcal·mol^{-1} for SCS-MP2/TZV-(2df, 2pd) and SCS-MP2/TZV(2d, 2p), respectively. This corresponds to an increase of only 0.27 and 0.26 kcal·mol^{-1}, respectively, as compared to the result for the 22 complexes in the small dataset (see Table 2). This increase of the MAD, although significant on a percentage basis (50% for TZV(2df, 2pd) and 45% for TZV(2d, 2p)), is likely also due to the quality of the reference data, which in general were derived with smaller basis sets for the large benchmark dataset than for the set of 22 small complexes.[14] As compared to the average absolute ΔE value of 12 kcal·mol^{-1}, the aforementioned MADs correspond to relative errors of 6 and 7% for TZV(2df, 2pd) and TZV(2d, 2p), respectively. On average, the deviation of the SCS-MP2 intermolecular interaction energies from the estimated CCSD-(T) reference is well below 10%. These results can be directly compared to those with the DFT−D method for the same set[13] that are of similar quality, i.e., a MAD of about 0.6 kcal·mol^{-1} with B97-D or B-LYP density functionals.

The degree of deviation of the SCS-MP2 results from the reference is not the same for the entire dataset but is dependent upon the bonding mode (see Figure 2). For hydrogen-bonded DNA base pairs, SCS-MP2 deviates systematically to smaller ΔE values by (TZV(2d, 2p) values in parentheses) 1.5 (1.5) kcal·mol^{-1}, whereas for the interstrand and the stacked base pairs the SCS-MP2 results are roughly centered around the reference with mean deviations of −0.1 (−0.1) and −0.2 (−0.4) kcal·mol^{-1}, respectively. The MAD amounts to 1.58 (1.58) kcal·mol^{-1} for hydrogen bonded, 0.17 (0.22) kcal·mol^{-1} for interstrand, and 0.44 (0.57) kcal·mol^{-1} for stacked base pairs, with the respective error spreads equal to 4.25 (4.29), 1.61 (1.88), and 2.27 (2.50) kcal·mol^{-1}.

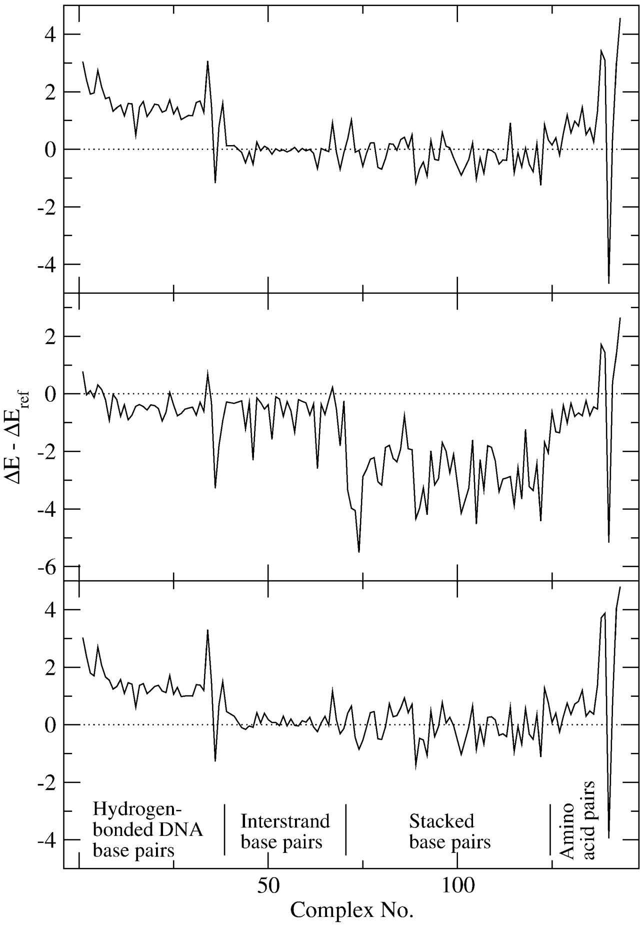

Figure 2. Deviation of ΔE values (kcal•mol^{-1}) for DNA base and amino acid pairs (JSCH2005) from the reference data[14]: Top SCS-MP2/TZV(2df,2pd), no CP-correction, middle and bottom MP2/TZV-(2df, 2pd), without and with CP-correction, respectively.

Whereas the MP2 method without CP-correction generally overestimates the ΔE values considerably (mean deviations of -1.3 and -1.4 kcal•mol^{-1} for TZV(2df, 2pd) and TZV(2d, 2p), respectively), the CP-corrected interaction energies closely resemble the uncorrected SCS-MP2 results for these two basis sets (see Figure 2). This means that, at this level, the BSSE and the effect of the scaling procedure (that has *not* been derived specifically for the here studied systems) almost exactly cancel. At present we have no deeper physical reasoning for this surprising conincidence.

3.3 Alkane Dimers. The above-described studies have shown that SCS-MP2 provides very accurate interaction energies for π-systems in stacked or mixed arrangements but tends to underbind saturated complexes. To investigate this issue in more detail, we consider as an extreme case alkane dimers that are bound dominantly by dispersion forces. These systems have been studied recently by Tsuzuki et al.[31]

Table 3 shows SCS-MP2 and MP2 ΔE values for methane, propane, and pentane dimers for various basis sets and the estimated MP2 limit from ref 31. In general, our MP2/CBS values agree within 0.05 kcal•mol^{-1} to the corresponding values in ref 31. The ΔCCSD(T) correction is 6% of the MP2 result for the methane dimer and is 5% for the propane dimer. Opposed to π-stacked complexes, MP2 is essentially correct for these saturated systems, and the coupled cluster correction is small and can be neglected for practical purposes.

For the TZV(2df, 2pd) basis, SCS-MP2 without the CP-correction roughly yields only two-thirds of the MP2/CBS intermolecular interaction energy. The relative errors of ΔE amount to 37, 34, and 29% for methane, propane, and pentane dimers, respectively. The relative errors of SCS-MP2 without CP-correction in the CBS limit are still 32, 36, and 35% for (methane)$_2$, (propane)$_2$, and (pentane)$_2$, respectively. This shows that the above proposed error compensation at the uncorrected SCS-MP2/TZV(2df, 2pd) or SCS-MP2/TZV(2d, 2p) level for alkane dimers only partly works. Although at this level the CBS limit is quite well reproduced, the correlation energy error is not compensated for.

For MP2, the situation is different. Although the CP-corrected MP2 results for the TZV(2df, 2pd) basis are off by 31, 20, and 16% for increasing chain lengths, the relative errors are reduced to 13, 9, and 6% with the QZVP basis. The TZV(2df, 2pd)-QZVP extrapolated CP-corrected MP2 intermolecular interaction energies (and the same holds for cc-pVXZ) are within 1% of the MP2 values from ref 31.

In summary, assuming an always positive ΔCCSD(T) correction of about 5% of ΔE for larger alkane dimers, the SCS-MP2 error with respect the "real" interaction energies is about 25%. This is much larger than any error in the above considered benchmark sets and is probably the upper limit for the SCS-MP2 error of noncovalently bound nonmetallic complexes of main-group molecules.

3.4 Modified Scaling Factors. As has been mentioned in the introduction, variants of SCS-MP2 where only the (scaled) opposite-spin correlation energy is retained (SOS-MP2)[27] and with specially adjusted scaling parameters (denoted as SCSN) have been proposed.[30] It was suggested by a reviewer to also include results from these approaches as a comparison to those from MP2 and standard (all purpose) SCS-MP2. Table 4 presents the statistical data of all four methods for the S22, the full JSCH2005, and the alkane dimer set. Detailed results for SOS-MP2 and SCSN-MP2 can be found in Supporting Information.

For the S22 set (on which the SCSN-MP2 parametrization is more or less based) and at the CBS limit, the SCSN-MP2 variant clearly outperforms the other methods with a very small MAD of 0.34 kcal•mol^{-1}. SOS-MP2 is significantly worse than MP2 and SCS-MP2 because the neglected same-spin correlation energy contribution to binding in vdW complexes is usually larger than the opposite-spin component, and this is not compensated for by the larger p_S factor as compared to SCS-MP2. At the TZV(2df, 2pd)/no CP level, however, SCS-MP2 benefits most from the error compensation, and SCS-MP2 and SCSN-MP2 perform similarly and much better than the other two methods. This conclusion also remains valid for the full JSCH2005 set. Although the performance of SCSN-MP2 is striking, one should keep in mind that this is a very empirical special purpose approach that does not fulfill basic physical requirements (e.g., it yields zero electron correlation energy for (closed-shell) systems with two electrons). On the contrary, the SCS-MP2 method can be used in general chemistry applications for covalent thermodynamics and simultaneously for vdW complexes. For alkane dimers, SCSN-MP2 significantly underbinds and performs only slightly better than SCS-MP2, but again, both are better than SOS-MP2.

4. Summary and Conclusion

It is a very reasonable conjecture that the general problems in the description of vdW interactions with MP2 are *not* solved by any of the SCS procedures. It seems important to mention here, however, that the SCS-MP2 method has not been developed specifically for weakly bound vdW complexes but

TABLE 3: SCS-MP2 and MP2 Interaction Energies[a] for Alkane Dimers with and without the Counterpoise Correction[b] for BSSE

	methane$_2$ (D_{3d})				propane$_2$ (C_{2h})				pentane$_2$ (C_{2h})			
	SCS-MP2		MP2		SCS-MP2		MP2		SCS-MP2		MP2	
	noCP	CP	no CP	CP	no CP	CP	no CP	CP	no CP	CP	no CP	CP
CBS(aT-aQ)[c]	0.31	0.31	0.47	0.47	1.27	1.27	2.03	2.04	2.44	2.46	3.88	3.91
aug-cc-pVQZ	0.36	0.30	0.51	0.46	1.52	1.24	2.26	2.00	2.96	2.40	4.36	3.85
aug-cc-pVTZ	0.43	0.29	0.57	0.45	1.88	1.19	2.58	1.95	3.70	2.31	5.04	3.76
aug-cc-pVDZ	0.74	0.25	0.87	0.39	2.76	1.00	3.44	1.72	5.40	1.99	6.73	3.39
CBS(T−Q)[c]	0.31	0.31	0.47	0.47	1.33	1.29	2.10	2.06	2.58	2.49	4.04	3.97
cc-pVQZ	0.29	0.26	0.44	0.41	1.28	1.14	2.02	1.88	2.51	2.24	3.93	3.66
cc-pVTZ	0.28	0.19	0.41	0.32	1.24	0.94	1.94	1.65	2.48	1.89	3.83	3.25
cc-pVDZ	0.22	0.03	0.31	0.12	0.92	0.32	1.49	0.87	1.87	0.76	2.98	1.84
CBS(T−Q)[c]	0.32	0.32	0.48	0.48	1.34	1.29	2.10	2.06	2.54	2.48	4.00	3.95
QZVP	0.31	0.27	0.45	0.41	1.34	1.15	2.08	1.90	2.62	2.25	4.03	3.68
TZV(2df,2pd)	0.30	0.20	0.43	0.33	1.37	0.96	2.06	1.67	2.77	1.93	4.11	3.30
TZV(2d,2p)	0.28	0.18	0.41	0.31	1.37	0.90	2.06	1.60	2.79	1.82	4.12	3.18
SVP	0.19	0.00	0.28	0.08	1.01	0.22	1.55	0.73	2.08	0.59	3.16	1.60
ref[d]	0.48 (0.03)				2.08 (−0.10)				3.92			

[a] $-\Delta E$ is measured in kcal·mol^{-1}. [b] With correction (CP), without correction (no CP). [c] Extrapolated to complete basis set limit. [d] MP2/CBS. ΔCCSD(T) correction in parentheses.[31]

TABLE 4: Statistics of the Deviation of MP2, SCS-MP2, SCSN-MP2, and SOS-MP2 ΔE Values[a] from the estimated CCSD(T)/CBS Reference Data[14] for 22 Small Complexes (S22 Set), 132 Neutral DNA Base and Amino Acid Pairs (JSCH2005), and from the MP2/CBS Reference Data[31] for the Six Alkane Dimers[b]

scaling factor		MP2	SCS-MP2	SOS-MP2	SCSN-MP2
p_S		1	6/5	1.3	0.0
p_T		1	1/3	0.0	1.76
S22 set	CBS(aT-aQ)[c], no CP[d]	0.79 (2.94)	0.71 (2.24)	1.40 (3.44)	0.34 (1.67)
JSCH2005	TZV(2df,2pd), no CP[d]	1.48 (6.29)	0.72 (4.32)	1.35 (4.39)	0.49 (4.34)
alkane dimers[e]	CBS(aT-aQ)[c], no CP[d]	0.04 (0.10)	0.99 (1.91)	1.47 (2.81)	0.68 (1.32)

[a] Measured in kcal·mol^{-1}. [b] The mean absolute deviation and (in brackets) the error spread (largest positive minus largest negative deviation) are given. [c] aug-cc-pVTZ/aug-cc-pVQZ extrapolated to complete basis set limit. [d] Counterpoise correction. [e] Dimers of methane to *n*-hexane.

that it was meant as a general improvement of MP2. Our finding that SCS-MP2 is on average still superior to MP2 for noncovalent interactions therefore supports the physical significance of the SCS modifications. At the (extrapolated) complete basis set limit, SCS-MP2 outperforms standard MP2, and the interaction energies are very close to CCSD(T) reference data for most unsaturated (π-stacked) systems. Our study fully confirms previous findings based on detailed investigations of the benzene dimer potential energy surface. Hydrogen-bond strengths are underestimated, but the errors are generally less than 10%, which is considered as acceptable in typical applications. The tendency to underbind saturated complexes has been investigated in detail for alkane dimers as model systems, and typical SCS-MP2 errors of about 25% for the interaction energies have been observed. This, however, is less than the typical MP2 errors for unsaturated (π-stacked) systems that are often 50−100%. On the basis of these observations and the good results obtained for the large benchmark set of biologically relevant systems, it is concluded that SCS-MP2 is the preferred wavefunction method in biochemical applications when meaningful CCSD(T) computations could not be performed. At the non-CP corrected SCS-MP2/TZV(2d, 2p) (or SCS-MP2/cc-pVTZ) level we found a very good error compensation between the BSSE, the incomplete correlation treatment, and the one-particle basis deficiencies for many systems studied. Together with the efficient RI (density fitting) procedures, this opens a way to routine applications for very large systems.

From the theoretical point of view, our results point to directions for future research. First, it seems important to understand in more detail why saturated and unsaturated molecules behave so differently in a second-order perturbation treatment. This is more-or-less independent of the SCS procedure as, for example, also with specially adjusted (extreme) SCS scaling parameters ($p_S = 0$, $p_T = 1.76$, see ref 30), alkane dimers are still underbound by about 25%. Our experiences with many other systems[6] indicate that the problems are not predominantly related to the degree of unsaturated character in the monomers but rather the “unsaturatedness” of the binding mode, for example, hydrogen-bonded vs π-stacked systems. This is already evident from the thoroughly investigated T-shaped and parallel-displaced benzene dimers with MP2 (SCS-MP2) errors of 32 (10) and 81 (4)% of ΔE, respectively, but it is also reflected in our results for the large benchmark set. Our point of view in this context is that the behavior of standard MP2 for π-stacked systems is quite typical (exaggeration of correlation effects, e.g., as also found for normal thermochemistry[20]) and that the good performance of MP2 for “saturated” noncovalent interactions is more unusual (fortuitous). In summary, the SCS modifications of MP2 can also be considered for the here investigated problem as removing the “outliers” and providing a more balanced description for systems with different electronic structure.

Acknowledgment. Financial support by the German research foundation (SFB 424 “Molecular Orientation and its Function in Chemical Systems”) is gratefully acknowledged.

Supporting Information Available: Tables with SCSN-MP2 and SOS-MP2 results for the S22 set and the alkane dimers and a statistical summary for the different basis sets. Two figures showing deviations from reference values for SCSN-MP2 and SOS-MP2 for the S22 and the JSCH2005 set. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Lehn, J.-M. *Supramolecular Chemistry: Concepts and Perspectives*; VCH: Weinheim, 1995.

(2) Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem., Int. Ed.* **2003**, *42*, 1210−1250.

(3) Kannan, N.; Vishveshwara, S. *Protein Eng.* **2000**, *13*, 753−761.

(4) Müller-Dethlefs, K.; Hobza, P. *Chem. Rev.* **2000**, *100*, 143−168.

(5) Stone, A. J. *The Theory of Intermolecular Forces;* Oxford University Press: Oxford, 1997.

(6) Grimme, S.; Antony, J.; Schwabe, T.; Mück-Lichtenfeld, C. *Org. Biomol. Chem*. **2007**, *5*, 741−758.

(7) Grimme, S.; Diedrich, C.; Korth, M. *Angew. Chem., Int. Ed.* **2006**, *45*, 625−629.

(8) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory;* J. Wiley: New York, 2000.

(9) Hobza, P.; Zahradnik, R.; Müller-Dethlefs, K. *Collect. Czech. Chem. Commun*. **2006**, *71*, 443−531.

(10) Tsuzuki, S.; Uchimaru, T. *Curr. Org. Chem.* **2006**, *10*, 745−762.

(11) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2006**, *110*, 10656−10668.

(12) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553−566.

(13) Antony, J.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5287−5293.

(14) Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985−1993.

(15) Grimme, S. *J. Comput. Chem.* **2006**, *25*, 1787−1799.

(16) Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2005**, *122*, 154104.

(17) Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2005**, *123*, 154101.

(18) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2004**, *108*, 10200−10207.

(19) Hill, J. G.; Platts, J. A.; Werner, H.-J. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4072−4078.

(20) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095−9102.

(21) Goumans, T. P. M.; Ehlers, A. W.; Lammertsma, K.; Würthwein, E.-U.; Grimme, S. *Chem.*−*Eur. J.* **2004**, *10*, 6468−6475.

(22) Grimme, S. *J. Phys. Chem. A* **2005**, *109*, 3067−3077.

(23) Hyla-Kryspin, I.; Grimme, S. *Organometallics* **2005**, *23*, 5581−5592.

(24) Grimme, S. *Angew. Chem., Int. Ed.* **2006**, *45*, 4460−4464.

(25) Grimme, S. *Chem.*−*Eur. J.* **2004**, *10*, 3423−3429.

(26) Piacenza, M.; Grimme, S. *ChemPhysChem* **2005**, *6*, 1554−1558.

(27) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. *J. Chem. Phys.* **2004**, *121*, 9793−9802.

(28) Lochan, R. C.; Jung, Y.; Head-Gordon, M. *J. Phys. Chem. A.* **2005**, *109*, 7598−7605.

(29) Jung, Y.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2831−2840.

(30) Hill, J. G.; Platts, J. A. *J. Chem. Theory Comput*. **2007**, *3*, 80−85.

(31) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M. *J. Chem. Phys.* **2006**, *124*, 114304.

(32) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett*. **1989**, *162*, 165−169.

(33) R. Ahlrichs et al. TURBOMOLE, version 5.7; Universität Karlsruhe, 2003. See http://www.turbomole.com.

(34) Weigend, F.; Häser, M. *Theor. Chem. Acc.* **1997**, *97*, 331−340.

(35) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett*. **1998**, *294*, 143−152.

(36) The basis sets are available from the TURBOMOLE homepage via the FTP server button (in the subdirectories basen, jbasen, and cbasen). See http://www.turbomole.com.

(37) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639−9646.

(38) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Olsen, J. *Chem. Phys. Lett*. **1999**, *302*, 437−446.