

# Complete Basis Set Extrapolation and Hybrid Schemes for Geometry Gradients of Noncovalent Complexes

Jiří Černý,<sup>\*,†</sup> Michal Pitoňák,<sup>‡</sup> Kevin E. Riley,<sup>§</sup> and Pavel Hobza<sup>§,||</sup>

<sup>†</sup>Institute of Biotechnology, Academy of Sciences of the Czech Republic, v. v. i., Vídeňská 1083, 142 20 Praha 4, Czech Republic

<sup>‡</sup>Department of Physical and Theoretical Chemistry, Faculty of Natural Sciences, Comenius University, Mlynská Dolina, 842 15 Bratislava 4, Slovak Republic

<sup>§</sup>Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, v. v. i., Flemingovo nám. 2, 166 10 Praha 6, Czech Republic

<sup>||</sup>Department of Physical Chemistry, Palacký University, 771 46 Olomouc, Czech Republic

**ABSTRACT:** In this paper, we focus on the performance of popular WFT (MP2, MP2.5, MP3, SCS(MI)-MP2, CCSD(T)) and DFT (M06-2X, TPSS-D) methods in optimizations of geometries of noncovalent complexes. Apart from the straightforward comparison of the accuracy of the resulting geometries with respect to the most accurate, computationally affordable, reference method, we have also attempted to determine the most efficient utilization of the information contained in the gradient of a particular method and basis set. Essentially, we have transferred the ideas successfully used for noncovalent interaction energy calculations to geometry optimizations. We have assessed the performance of the hybrid gradients (for instance, MP2 and CCSD(T) calculated in different basis sets), investigated the possibility of extrapolating gradients calculated with a particular method in a series of systematically built basis sets, and finally compared the extrapolated gradients with the counterpoise(CP)-corrected optimizations, in order to determine which of these approaches is more efficient, in terms of their convergence toward the CBS geometry for the respective calculation cost. Further, we compared the efficiency of the CP-corrected, extrapolated, and hybrid gradients in terms of the rate of convergence with respect to basis set size. We have found that CCSD(T) geometries are most faithfully reproduced by the MP2.5 and MP3 methods, followed by the comparably well performing SCS(MI)-MP2 and MP2 methods, and finally by the worst performing DFT-D and M06 methods. Basis set extrapolation of gradients was shown to improve the results and can be considered as a low-cost alternative to the use of CP-corrected gradients. A hybrid gradient scheme was shown to deliver geometries close to the regular gradient reference. Analogously to a similar hybrid scheme, which nowadays is routinely used for the calculation of interaction energies, such a hybrid gradient scheme can save a huge amount of computer time, when high accuracy is desired.

## 1. INTRODUCTION

The crucial role of noncovalent interactions in biology, nanoscience, catalysis, and other scientific disciplines<sup>1</sup> has been clearly stressed in many scientific discussions and studies during past decades. The necessity of using highly sophisticated *ab initio* quantum chemical methods, such as the coupled-cluster (CC) methods at the complete basis set (CBS) limit, for calculation of noncovalent interactions has also been recognized.<sup>2–4</sup> The failures of otherwise rather successful methods, such as second-order Møller-Plesset perturbation theory (MP2) or the entire family of DFT methods (except for the recently developed functionals by Truhlar and Zhao<sup>5</sup> and Scuseria et al.<sup>6</sup>) to describe, for instance, the  $\pi$ – $\pi$  stacking interactions of aromatic systems, has motivated the community of quantum chemists to develop novel, more efficient (both in terms of accuracy and computational cost) approximate methods. A more detailed discussion of different improved DFT approaches can be found in our review.<sup>4</sup> Most research efforts have focused solely on interaction energies, and attempts to develop and assess methods that accurately describe noncovalent complex geometries have been largely pushed aside. Except for a few studies on the geometries of hydrogen-bonded (H-bonded) and stacked DNA base pairs<sup>7–10</sup>

and other complexes,<sup>11,12</sup> most of the publications in the field have regarded the geometry problem as being less important.

The truth is that the amount of attention paid to the quality of the description of noncovalent complex geometries, compared to interaction energies, to some extent reflects the relative importance of these issues in real life applications. The sensitivity of the geometry of a noncovalent complex to a particular method (and basis set size) applied for geometry optimization is, with no doubt, lower than that of the “single-point” interaction energy. Presuming that a qualitatively correct method (i.e., capable of describing a particular noncovalent interaction type) is used, convergence toward the (at least qualitatively) correct geometry with respect to the basis set size is rather fast. As the most straightforward example, the Hartree–Fock (HF) method with a very small basis set, such as Pople STO-3G, yields at least semiquantitatively correct geometries for “standard” H-bonded complexes (such as the water dimer) in all cases known to us.<sup>13</sup> The opposite is clearly true if optimization of, for instance, the benzene dimer is attempted using the same methodology. In this

Received: July 11, 2011

Published: October 17, 2011

case, use of a method capable of properly describing the dispersion interaction (i.e., a correlated method, unlike HF) in combination with a basis set containing diffuse functions is critical.

This raises another question, which concerns the convergence of the geometry with respect to the quality of the method (in terms of being closer to the full configuration interaction result), presuming (as mentioned above) that the investigated method is capable of describing the particular interaction type. As an example, we can ask how accurate the geometry of the benzene dimer  $\pi$ – $\pi$  stacked complex is when it is optimized with a series of MPX (X being 2, 3, or 4) or CC (CCSD, CCSD(T), ...) methods. This certainly is meaningful, since all of the MPX and CC methods are capable of describing dispersion interactions. Furthermore, gaining the knowledge that, for instance, MP2 (being  $\sim N^5$  scaling method with system size “N”) delivers similar accuracy to that of the MP4 or the CCSD(T) methods (both scaling as  $\sim N^7$ ) could save a tremendous amount of computer time.

The basis set dependence of optimized geometries has been studied in several works, of which the most relevant to this study are the papers of Salvador and co-workers<sup>14,15</sup> and Hobza and co-workers<sup>7,16–18</sup> and also the papers from the groups of Klopper,<sup>19</sup> Pulay and Janowski,<sup>20</sup> Kim,<sup>11</sup> and others. The common issue discussed in these papers is the role of Basis Set Superposition Error (BSSE) on the molecular (complex) geometry. As was pointed out for the first time in the paper of Simon et al.<sup>14</sup> and further analyzed in the papers of Salvador and Duran<sup>15</sup> and Paizs et al.,<sup>21</sup> BSSE affects not only interaction energies but also geometrical parameters. In a similar way to how the BSSE leads to an artificial increase in the stabilization of complexes, it also leads to the shortening of, for example, the length of H bonds<sup>14</sup> and the displacement of stacked aromatic rings<sup>7,20</sup> etc.

We have several goals in this work: first, to assess the performance of the hybrid (for instance, MP2 and CCSD(T) calculated in different basis sets) gradients; second, to investigate the possibility of extrapolating gradients calculated with a particular method in a series of systematically built basis sets; and finally, to compare the extrapolated gradients with the CP-corrected optimizations, to determine which of these approaches is more efficient, in terms of their convergence toward the CBS geometry for the respective calculation cost.

## 2. METHODS

**2.1. CBS Gradient Optimization.** In this work, basis set extrapolations are carried out using the scheme of Helgaker et al.<sup>22,23</sup> (eq 1), where  $\alpha$  is a basis set dependent parameter,  $E_X$  and  $E_{\text{CBS}}$  are the HF and correlation energies for the basis set with largest angular momentum X and for the complete basis set, respectively.

$$E_X^{\text{HF}} = E_{\text{CBS}}^{\text{HF}} + A e^{-\alpha X} \text{ and } E_X^{\text{corr}} = E_{\text{CBS}}^{\text{corr}} + B X^{-3} \quad (1)$$

For the RI-MP2/CBS optimization, the HF and MP2 correlation gradients were, as in the case of energies, extrapolated separately due to the different (exponential vs power) convergence profiles of the procedure. The necessary “correlation energy gradient” is obtained as a difference between the total MP2 gradient and the HF gradient. Technically, each element of the HF and correlation Cartesian gradients is extrapolated (using its value instead of energy  $E$  in eq 1), and the final CBS gradient is obtained as a sum of the HF/CBS and correlation/CBS gradients. For small complexes, a similar approach was also applied for

higher order correlation methods, directly extrapolating the MP2.5,<sup>24</sup> MP3, and CCSD(T) gradients.

**2.2. Hybrid Gradient Optimization.** The proposed CCSD(T)/CBS gradient is constructed as a sum of the MP2/CBS gradient and a correction gradient, in a way similar to that of the energy calculation. This correction gradient is obtained as a difference between the CCSD(T) and MP2 gradients calculated in a smaller basis set. The 6-31G\*\*(0.25,0.15) and (when applicable) larger cc-pVDZ or cc-pVTZ basis sets were used for the correction gradient calculations. The MP2.5 and MP3 CBS gradients are constructed in a way similar to the one for CCSD(T), employing the (scaled) MP3 correlation gradient correction.

## 3. COMPUTATIONAL DETAILS

The HF and MP2 CBS calculations were performed with two different consecutive basis sets, cc-pVDZ/cc-pVTZ and cc-pVTZ/cc-pVQZ, using the recommended coefficients for both the energy and gradient element extrapolations.

The optimizations were carried out using our newly developed scripts, which call (depending on the level of theory) the rimp2 module of the TurboMole 5.10 package<sup>25</sup> and/or the ACESII/ACESIII<sup>26</sup> program for the MP2, MP3 and CCSD(T) energy and gradient calculations. The collected gradients as well as energies are extrapolated, and the relax module of the TurboMole package is called to obtain the new coordinates.

The SCS(MI)-MP2<sup>27</sup> optimizations were performed using the recommended parameters of the scs keyword for the ricc2 module of TurboMole.

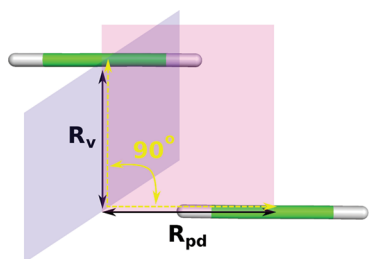
For a comparison of the performance, we also included the results of DFT calculations, namely, the M06-2X (as implemented in Gaussian 03<sup>28</sup>) and TPSS-D<sup>29</sup> both using the 6-311++G(3df,3pd) basis set. The TPSS-D method is implemented using a combination of the ridft module of TurboMole and our tm\_dis program.<sup>29</sup>

These DFT methods were selected from the wide range of available methods based on our previous results, showing that these methods provide consistently highly accurate interaction energy as well as PES scan results.

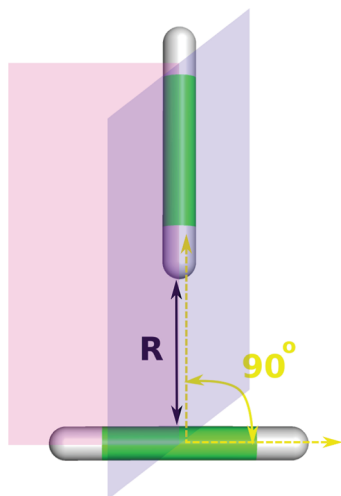
All of the optimizations were performed employing the C1 symmetry, even for starting structures with higher symmetry. The same initial geometry was used for all of the optimization methods. The gradients on all atoms of the complex were included, optimizing fully the geometry of monomers within the complex as well as their separation.

Throughout the paper, we are using following notation: D, T, aD, aT, ... for Dunning (aug)-cc-pVXZ basis sets; DT and TQ stand for the (X→Y) Helgaker type of basis set extrapolation. “025015” stands for the above-mentioned 6-31G\*\*(0.25, 0.15) modified Pople’s basis set, where the exponent of the most diffuse d functions on heavy (C, N, O, ...) atoms is changed from 0.8 to 0.25, and the exponent of the p function of hydrogens is changed from 1.1 to 0.15. Further, for a particular method “X”, the correction term  $\Delta$  is defined as  $\Delta = X - \text{MP2}$ .

The computational cost of optimizations is reduced significantly, especially by employing the hybrid scheme. For example, one cycle of methanol dimer optimization at the CP-CCSD(T)/cc-pVTZ level takes about 60 h, while the hybrid MP2(TQ)+ $\Delta$ (CCSD(T)/025015) offers better final geometry with about 2 h for one optimization cycle.



**Figure 1.** Acetylene dimer, parallel-displaced conformation: structure and the analyzed geometrical parameters.



**Figure 2.** Acetylene dimer, T-shaped conformation: structure and the analyzed geometrical parameters.

## 4. RESULTS AND DISCUSSION

**4.1. Acetylene Dimer.** The acetylene dimer was chosen as the model system for  $\pi$ – $\pi$  stacked, dispersion dominated noncovalent complexes. In analogy with the benzene dimer,<sup>20,30</sup> two characteristic minima were studied, i.e., the “parallel-displaced”, PD (see Figure 1), and the “T-shaped” structure, T (see Figure 2). Due to the spatial symmetry restrictions applied, the two intermolecular distances, vertical displacement,  $R_v$ , and parallel displacement,  $R_{pd}$ , define unambiguously the PD structure of the acetylene dimer. In the case of the T structure, it is only the vertical displacement,  $R$ , which defines the structure, for the reasons mentioned above. We will repeat here that the monomer geometries were in all cases optimized as well.

The proof of concept for the hybrid MP2/CCSD(T) gradient is demonstrated in Table 1.

Comparing the first two rows of the table, it is clear that MP2 delivers a tighter conformation of the complex compared to CCSD(T), i.e., monomers are vertically shifted by about 0.03 Å further apart at the CCSD(T) level, due to the well-known overestimation of the  $\pi$ – $\pi$  stacked interaction at the MP2 level. The vertical displacement decreases upon the extension of basis set size (i.e., cc-pVDZ (“D”) to cc-pVTZ (“T”)) at both levels, by 0.26 Å at the MP2 and 0.37 Å at the CCSD(T) level. To demonstrate the success of using the “hybrid gradient”, i.e., mixing of the MP2 gradient calculated in a larger basis set, T, and the CCSD(T) correction to the MP2 gradient (i.e.,  $\Delta$ CCSD(T)) in a smaller basis set, D, we can compare the vertical

**Table 1.** Acetylene Dimer–PD Conformation<sup>a</sup>

MP2/X	$\Delta$ CCSD(T)/X	$R_v$ [Å]	$R_{pd}$ [Å]	RMSD [Å]
D		2.8577	3.1947	0.0234
D	D	2.8900	3.2020	0.0311
T		2.8317	3.1492	0.0191
T	D <sup>b</sup>	2.8710	3.1684	0.0095
T	T	2.8534	3.1791	—

<sup>a</sup> Comparison of the geometrical parameters obtained on the MP2 level (i.e., no  $\Delta$ CCSD(T)/X), CCSD(T) level (i.e., MP2/X +  $\Delta$ CCSD(T)/X), and hybrid MP2/CCSD(T) level (i.e., MP2/X +  $\Delta$ CCSD(T)/(X-1)). “D” and “T” stand for Dunning’s cc-pVDZ and cc-pVTZ basis sets. RMSD stands for the root mean square deviation of the optimized geometry with regard to the reference geometry (i.e., RMSD of “—”).

<sup>b</sup> The 025015 basis set for the  $\Delta$ CCSD(T)/X calculation performs slightly worse than the D basis set:  $R_v$ , 2.8626 [Å];  $R_{pd}$ , 3.1908 [Å]; RMSD, 0.0164 [Å].

displacement of 2.8710 Å, as obtained by the hybrid scheme, with the “pure” MP2 and CCSD(T) values. This value is clearly smaller than that of CCSD(T)/D, in agreement with the trends mentioned above, but also clearly larger than that of the MP2/T value, in accordance with the effect of the higher-order correlation on the gradient. At the same time, this value is in between the CCSD(T)/D and CCSD(T)/T ones, which is a clear improvement, taking into consideration that the computational demands of the hybrid scheme are practically equal to that of pure CCSD(T)/D optimization.

Table 2 summarizes the optimized geometry parameters and the CCSD(T)/CBS interaction energies (calculated as MP2/(aTaQ) +  $\Delta$ CCSD(T)/cc-pVTZ), calculated for the most accurate geometries obtained by a particular method.

Since obtaining the CCSD(T)/CBS interaction energies for all geometries would be too time-consuming, the discussion on the accuracy of the obtained geometries with respect to the quality of the basis set will be limited only to the MP2 method and can be found in the last paragraph of this section. Let us start first with the T-shaped acetylene dimer. When the MP2/(DT) +  $\Delta$ CCSD(T)/025015 (or MP2/(DT) +  $\Delta$ CCSD(T)/D) is compared with MP2/(DT) values of the vertical displacement ( $R$ ), a similar increase of  $R$  by about 0.04 (0.05) Å due to the higher-order correlation is observed. Comparing the value of  $R$  calculated on the hybrid MP2/(TQ) level corrected with the one  $\Delta$ CCSD(T) calculated in 025015, D, and T basis sets with the one pure MP2/(TQ) value, almost constant shifts of 0.036, 0.048, or 0.041 Å are obtained. This indicates that, similar to calculation of the interaction energies, MP2 gradients converge more slowly with basis set size than the higher order correlation contributions, as calculated at the CCSD(T) level. It is also interesting to analyze the convergence of the MP2 gradient extrapolation. The MP2/(DT)  $R$  value repeats the trend of its decreasing value with basis set size, being significantly lower than the MP2/Q value, i.e., 2.6257 vs 2.6537 Å. At the same time, the MP2/(DT) value is closer to the most accurate MP2  $R$  value, obtained either from the MP2/(TQ), 2.6486 Å, or from the two-point “1/X<sup>3</sup>”  $R$  MP2 values extrapolated from T and Q basis sets, 2.6497 Å, by about 0.01 Å.

In analogy with the interaction energy calculations, we can also inspect the CP-corrected gradient optimization results as the concurrent approach to the extrapolated gradients. Both approaches converge to the same geometries, since in the CBS there is no BSSE. However, complexes of a rather small number of atoms can be optimized in basis sets large enough to approach

**Table 2.** Acetylene Dimer—Selected Geometrical Parameters Obtained from the Unconstrained Geometry Optimization on Various Theoretical Levels and Basis Sets<sup>a</sup>

method	hybrid/basis	T			PD			
		<i>R</i> [Å]	RMSD [Å]	<i>E</i> <sub>int</sub>	<i>R</i> <sub>v</sub> [Å]	<i>R</i> <sub>pd</sub> [Å]	RMSD [Å]	<i>E</i> <sub>int</sub>
CC	MP2(DT <sup>b</sup> )+Δ <sup>c</sup> (025015)	2.6675	0.0126		2.8440	3.1710	0.0274	
	MP2(TQ)+Δ(025015)	2.6847	0.0029		2.8320	3.1074	0.0058	
	MP2(DT)+Δ(D)	2.6803	0.0071		2.8531	3.1545	0.0237	
	MP2(TQ)+Δ(D)	2.6962	0.0045		2.8358	3.0858	0.0122	
	MP2(TQ)+Δ(T)	2.6899	—	−1.500	2.8357	3.1114	—	−1.382
	CP D	2.8211	0.0816		<i>d</i>		0.9746	
	CP T	2.7225	0.0147		2.8200	3.2597	0.0556	
	CP MP2(D)+Δ(025015) <sup>c</sup>	2.8654	0.1047		2.9884	3.3286	0.1347	
	CP MP2(T)+Δ(025015)	2.7584	0.0375		2.8937	3.2413	0.0690	
MP2.5	D	2.7473	0.0424		2.8782	3.1874	0.0460	
	T	2.6520	0.0238		2.8342	3.1635	0.0222	
	MP2(DT)+Δ(025015)	2.6563	0.0222		2.8449	3.1419	0.0171	
	MP2(TQ)+Δ(025015)	2.6760	0.0089		2.8241	3.0845	0.0146	
	MP2(DT)+Δ(D)	2.6576	0.0219		2.8434	3.1349	0.0147	
	MP2(TQ)+Δ(D)	2.6759	0.0093		2.8323	3.0647	0.0208	
	MP2(TQ)+Δ(T)	2.6740	0.0109	−1.484	2.8323	3.0774	0.0163	−1.369
	CP D	2.7972	0.0660		2.8903	3.2956	0.0909	
	CP T	2.6988	0.0061		2.8532	3.1919	0.0379	
MP3	D	2.7719	0.0519		2.9068	3.1962	0.0573	
	T	2.6794	0.0143		2.8689	3.1709	0.0353	
	MP2(DT)+Δ(025015)	2.6894	0.0103		2.8726	3.1626	0.0450	
	MP2(TQ)+Δ(025015)	2.7084	0.0082		2.8526	3.1043	0.0226	
	MP2(DT)+Δ(D)	2.6886	0.0114		2.8744	3.1402	0.0279	
	MP2(TQ)+Δ(D)	2.7029	0.0064		2.8557	3.0766	0.0164	
	MP2(TQ)+Δ(T)	2.6996	0.0069	−1.470	2.8527	3.0797	0.0140	−1.364
	CP D	2.8232	0.0764		2.9120	3.3187	0.1055	
	CP T	2.7281	0.0160		2.8909	3.2080	0.0559	
SCS(MI)-MP2	Q	2.7097	0.0090		2.8491	3.1539	0.0247	
	CP T	2.6897	0.0045		2.8462	3.1871	0.0351	
	CP Q	2.6662	0.0120	−1.506	2.8557	3.1597	0.0285	−1.351
MP2	D	2.7257	0.0363		2.8577	3.1947	0.0454	
	T	2.6592	0.0146		2.8317	3.1492	0.0170	
	Q	2.6537	0.0181		2.8084	3.1080	0.0115	
	DT <sup>f</sup>	2.6257	0.0357		2.8078	3.1320	0.0120	
	TQ <sup>g</sup>	2.6486	0.0211	−1.497	2.8168	3.0444	0.0330	−1.373
	CP D	2.7719	0.0570		2.8636	3.2924	0.0838	
	CP T	2.6864	0.0027		2.8368	3.1848	0.0310	
	CP Q	2.6658	0.0122		2.8105	3.1216	0.0074	
M06-2X	6-311++G(3df,3pd)	2.6700	0.0238	−1.473	2.8167	2.9372	0.0765	−1.323
DFT-D	TPSS/6-311++G(3df,3pd)	2.8498	0.0801	−1.446	2.9277	3.1255	0.0414	−1.368

<sup>a</sup> *E*<sub>int</sub> corresponds to the CCSD(T)/CBS interaction energy calculated at the particular complex geometry. RMSD stands for the root mean square deviation of the optimized geometry with regard to the reference geometry (i.e., RMSD of “—”). <sup>b</sup> D, T, ... stand for Dunning's cc-pVXZ basis sets. DT and TQ stand for the (X→Y) Helgaker type of basis set extrapolation. “025015” stands for the 6-31G\*\*(0.25, 0.15) modified Pople's basis set, where the exponent of the most diffuse d functions on heavy (C, N, O, ...) atoms is changed from 0.8 to 0.25, and the exponent of the p function of hydrogens is changed from 1.1 to 0.15. <sup>c</sup> For a particular method “X”, the Δ term is defined as Δ = X − MP2. <sup>d</sup> Geometry optimization converged to the “T”-shaped structure. <sup>e</sup> CP corrected MP2 gradient with added CP uncorrected Δ term. <sup>f</sup> Extrapolated *R*, *R*<sub>v</sub>, and *R*<sub>pd</sub> values according to the two-point “1/*X*<sup>3</sup>” extrapolation for D and T basis sets are 2.6312, 2.8208, and 3.1300 Å. <sup>g</sup> Extrapolated *R*, *R*<sub>v</sub> and *R*<sub>pd</sub> values according to the two-point “1/*X*<sup>3</sup>” extrapolation for T and Q basis sets are 2.6497, 2.7914, and 3.0779 Å.

the CBS; thus, it is natural to ask whether the CP-corrected gradients in particular basis sets are systematically more accurate than the extrapolated ones. Comparing, for instance, the MP2/(DT)+ΔCCSD(T)/D value of *R* with the CP-corrected

one calculated in the D basis set, i.e., 2.6803 vs 2.8211 Å, it is clear that the first one is much closer to the most accurate, MP2/(TQ)+ΔCCSD(T)/T, value of 2.6899 Å. It is important to realize that the CP-corrected optimization requires the calculation



of five gradients, one for the dimer (as in the hybrid MP2/CCSD(T) scheme) and also four additional gradients of monomers in supermolecular and monomer AO basis sets. After the two-point “ $1/X^3$ ” extrapolation of the CP-corrected  $R$  values from the D and T basis sets is done, the obtained value of 2.6810 Å is close to the reference. But the computational overhead of the CP-corrected CCSD(T) calculation in the T basis set is certainly not worth the effort, at least not in this case. The unsatisfactory accuracy of the CP-corrected gradients is even more obvious at the “pure” MP2 level, where the CP-corrected  $R$  value in the Q basis set is even less accurate than the plain MP2/T one. On the other hand, two-point “ $1/X^3$ ” extrapolation of CP-corrected MP2 gradients either from D to T or from T to Q basis sets leads (probably by chance) to about the same values, 2.6504 and 2.6508 Å, which are in a good agreement with the best reference MP2 values.

Concerning the performance of methods other than MP2 or CCSD(T), we can compare the “best” geometries for each method, by either taking the MP2/(DT)+ $\Delta X$ /T gradients ( $\Delta X$  being  $\Delta$ MP2.5 or  $\Delta$ MP3) and the two-point extrapolation of the CP-corrected gradients in D and T basis sets (or T and Q, in the case of the SCS(MI)-MP2 method) or just by taking only the plain  $R$  value (M06-2X and DFT-D). MP2.5 underestimates the reference MP2/(TQ)+ $\Delta$ CCSD(T)/T value of  $R$  by only 0.016 Å on the extrapolated-hybrid level, but by 0.033 Å ( $R$  being 2.6574 Å) on the extrapolated CP-corrected level. The plain CP-corrected MP2.5 value in the T basis set is in much better agreement, with an error of 0.009 Å, but this is most likely the consequence of error compensation due to a lack of basis set saturation. MP3 clearly delivers more accurate results, the agreement of the extrapolated-hybrid and extrapolated CP-corrected scheme (2.6881 Å) being within  $-0.01$  and  $0.002$  Å. The minus sign means that the  $R$  value is overestimated (larger) compared to the reference. The SCS(MI)-MP2  $R$  values match the reference within 0.02 and 0.03 Å at the CP-corrected Q and extrapolated CP-corrected levels. From the DFT methods, it is clearly M06-2X that is more accurate, the error being 0.02 Å vs 0.16 Å for DFT-D.

In the PD acetylene dimer, we observe very similar trends. Higher-order correlation increases the vertical displacement, affecting the parallel displacement only marginally. Favorable error compensation takes place at the MP2 level, since the optimized  $R_v$  converges with the basis set from above. As might be anticipated, the best extrapolated-hybrid MP2.5 results reproduce the CCSD(T) reference more closely than those of MP3, the errors being 0.003/0.03 and 0.017/0.03 Å for  $R_v/R_{pd}$ , though the difference is rather small. The performance of the CP-corrected SCS(MI)-MP2 in the Q basis set is also satisfactory, slightly overestimating the reference  $R_v/R_{pd}$  values by  $-0.02/-0.05$  Å. Both the DFT results deviate significantly from the reference, the errors being somewhat lower for the DFT-D method in this case, i.e., 0.02/0.17 Å for M06-2X and 0.09/0.01 Å for DFT-D.

For the interaction energies, it was found<sup>7</sup> that the higher-order correlation effects not only converge faster with basis set size but also suffer less from BSSE. In the spirit of this idea, we combine the CP-corrected MP2 gradient with the CP-uncorrected  $\Delta$ CCSD(T) correction calculated with a smaller basis set, for instance, 025015; i.e., see the CP-MP2/D+ $\Delta$ /025015 rows in Table 2. The effect of the higher-order correlation on the CP-corrected MP2 gradient is again the elongation of the  $R$ ,  $R_v$ , and  $R_{pd}$  distances. However, in both the T-shaped and the PD

**Table 3. Acetylene Dimer T, MP2 Method<sup>a</sup>**

X	$E_{\text{int}}/X$	$E_{\text{int}}/a\text{TaQ}$	$E_{\text{def}}/X$	$E_{\text{def}}/TQ$	$E_{\text{bind}}/a\text{TaQ}/X$
D	−1.164	−1.748 (−0.094)	0.004	1.598	−1.744 (−0.096)
CP-D	−1.182	−1.718 (−0.064)	0.003	1.585	−1.715 (−0.067)
T	−1.380	−1.669 (−0.015)	0.005	0.065	−1.664 (−0.016)
CP-T	−1.386	−1.665 (−0.011)	0.004	0.063	−1.660 (−0.012)
Q	−1.536	−1.659 (−0.005)	0.005	0.017	−1.654 (−0.006)
CP-Q	−1.537	−1.658 (−0.004)	0.005	0.016	−1.654 (−0.006)
DT	−1.496	−1.627 (0.027)	0.006	0.032	−1.622 (0.026)
TQ	−1.651	−1.654	0.006	0.006	−1.648

<sup>a</sup>  $E_{\text{int}}/X$ : CP-corrected interaction energy in “X” basis set and (CP-)“X” optimized geometry.  $E_{\text{int}}/a\text{TaQ}$ : CP-corrected interaction energy in aTaQ extrapolation and (CP-)“X” optimized geometry.  $E_{\text{def}}/X$  and  $E_{\text{def}}/TQ$ : deformation energy in “X” basis set or TQ extrapolation and (CP-)“X” optimized geometry.  $E_{\text{bind}}/a\text{TaQ}/X$ : CP-corrected interaction energy in aTaQ extrapolation and (CP-)“X” optimized geometry with corrections for the deformation of monomers calculated in the “X” basis set and the corresponding geometry. Values in parentheses represent the error with respect to the “reference” TQ values in the corresponding column.

acetylene dimers, this elongation is in the same direction as the error of the CP-corrected MP2 gradient, thus making such a hybrid method rather inaccurate. Only when the two-point extrapolation of the CP MP2/ $X + \Delta$ /025015 from D to T basis set is carried out are the resulting values for  $R$ ,  $R_v$ , and  $R_{pd}$ , i.e., 2.7134, 2.8538, and 3.2045 Å, finally more accurate than the MP2 CP-corrected extrapolated values. In the case of the PD structure, such a hybrid method converges to the PD structure, unlike the CP-corrected CCSD(T)/D gradient, which converged to the T-shaped structure.

As mentioned in the beginning of this section, CP-corrected CCSD(T)/CBS interaction energies (without relaxation energies) were calculated for the “best” geometries for each method. These energies vary only slightly for the T-shape structure, from  $-1.506$  (SCS(MI)-MP2 geometry) to  $-1.446$  (DFT-D geometry) kcal/mol, i.e., by only 0.06 kcal/mol ( $\sim 4\%$ ). For the T-shaped structure, the values varied from  $-1.382$  to  $-1.323$  kcal/mol, i.e., again only by 0.06 kcal/mol ( $\sim 4\%$ ). This clearly demonstrates the insensitivity of the interaction energy to geometry changes, thus making the method used for geometry optimization far less important than the method used for the single-point interaction energy calculation. There might be other molecular properties that are more strongly affected by deviations in geometry, such as spectroscopic parameters, but this is out of the scope of this article.

As noted at the beginning of this section, we could afford to calculate the CCSD(T)/CBS interaction energies only for the “best” geometries for each method. However, to get a deeper insight into the importance of basis set saturation in the gradient and interaction energy calculations, we carried out CP-corrected MP2 calculations of the interaction energy in several basis sets and close to the CBS limit at the geometries optimized at the MP2 level in the same series of basis sets with and without the CP-correction in the gradient. The deformation energy contributions to the binding energy are also calculated in the series of basis sets as well as close to the CBS limit. Table 3 summarizes these result for the T-shaped acetylene dimer.

In the first column, i.e., “ $E_{\text{int}}/X$ ”, the CP-corrected MP2 interaction energies calculated in the “X” basis set at the CP-corrected and CP-uncorrected MP2 geometries obtained in the same basis set are shown. The interaction energies range from  $-1.164$  (D)

Table 4. Acetylene Dimer PD, MP2 Method<sup>a</sup>

X	$E_{\text{int}}/X$	$E_{\text{int}}/\text{aTaQ}$	$E_{\text{def}}/X$	$E_{\text{def}}/\text{TQ}$	$E_{\text{bind}}/\text{aTaQ}/X$
D	−0.928	−1.629 (−0.066)	0.003	1.576	−1.626 (−0.068)
CP-D	−0.951	−1.591 (−0.028)	0.003	1.557	−1.588 (−0.030)
T	−1.243	−1.566 (−0.003)	0.004	0.061	−1.562 (−0.004)
CP-T	−1.246	−1.556 (0.007)	0.004	0.059	−1.552 (0.006)
Q	−1.404	−1.566 (−0.003)	0.005	0.016	−1.561 (−0.003)
CP-Q	−1.405	−1.565 (−0.002)	0.005	0.016	−1.560 (−0.002)
DT	−1.381	−1.542 (0.021)	0.005	0.041	−1.538 (0.020)
TQ	−1.540	−1.563	0.006	0.006	−1.558

<sup>a</sup>  $E_{\text{int}}/X$ : CP-corrected interaction energy in “X” basis set and (CP-)“X” optimized geometry.  $E_{\text{int}}/\text{aTaQ}$ : CP-corrected interaction energy in aTaQ extrapolation and (CP-)“X” optimized geometry.  $E_{\text{def}}/X$  and  $E_{\text{def}}/\text{TQ}$ : deformation energy in “X” basis set or TQ extrapolation and (CP-)“X” optimized geometry.  $E_{\text{bind}}/\text{aTaQ}/X$ : CP-corrected interaction energy in aTaQ extrapolation and (CP-)“X” optimized geometry with corrections for deformation of monomers calculated in “X” basis set and the corresponding geometry. Values in parentheses represent the error with respect to the “reference” TQ values in the corresponding column.

to −1.715 kcal/mol (TQ), i.e., the difference being 0.551 kcal/mol (~32%). For all three basis sets, the interaction energies on the CP-corrected geometries are slightly more attractive, but at most only by 0.018 kcal/mol (in the D basis set). If we calculate MP2/CBS interaction energies at these geometries, i.e., the second column “ $E_{\text{int}}/\text{aTaQ}$ ”, these interaction energies only range from −1.627 (DT) to −1.748 (D) kcal/mol; i.e., the span is reduced to only 0.121 kcal/mol (~7%). If we set the TQ geometry to be the reference one, the span of errors in the interaction energies (in parentheses in the second column) reduces even further below 0.1 kcal/mol (0.094 kcal/mol). This clearly demonstrates the importance of the quality of the underlying basis set used for the geometry and single-point calculations. The third and the fourth columns, i.e., “ $E_{\text{def}}/X$ ” and “ $E_{\text{def}}/\text{TQ}$ ”, show the deformation energies as obtained in the “X” basis set, or in the close-to-CBS limit, “TQ”. For this particular complex, the deformation energies are very small, about 0.006 kcal/mol, and are practically invariant to the basis set size (changes from 0.003 to 0.006 kcal/mol from CP-D to TQ). Much larger values are obtained for the CBS limit of the relaxation energy as calculated from geometries optimized in one particular basis set. Values above 1.5 kcal/mol obtained for D and CP-D structures indicate that the monomer geometries (both in the complex and in the isolated systems) are quite different from those obtained in larger basis sets. Indeed, the distance between the carbon atoms within the monomers changes from 1.230 Å in D to 1.209 Å in the Q basis set. In a similar fashion, the length of the C–H bonds shortens from 1.075/1.077/1.076 Å in D to 1.061/1.064/1.062 Å, where the first number corresponds to the C–H bond of the monomer “on top”, according to Figure 2 the most distant from the horizontally placed monomer, the second number refers to the C–H of the same monomer pointing toward the monomer on the bottom, and the third value corresponds to the symmetric C–H bonds of the monomer on the bottom.

Similar effects can be observed for the PD structure, shown in Table 4.

The range of the interaction energies in different basis sets and geometries is now 0.659 kcal/mol (~42%), i.e., from −0.928 (D) to −1.587 kcal/mol (TQ). The CBS interaction energies calculated at different geometries span, again, a much narrower

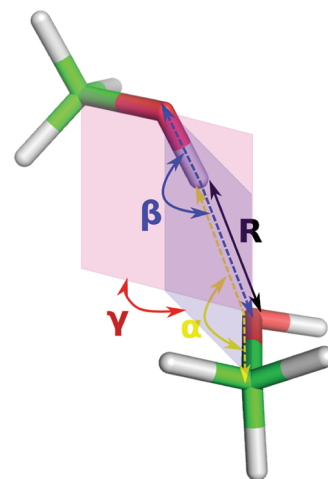


Figure 3. Methanol dimer: structure and the analyzed geometrical parameters.

interval of only 0.088 kcal/mol (~6%), i.e., from −1.538 (DT) to −1.626 (D) kcal/mol. The deformation energies calculated in the same basis set as the gradient are practically identical to those of the T-shaped structure and so is the CBS estimate of the deformation energy for each geometry. For both the T-shaped and the PD structures, the largest error in the MP2/CBS interaction energy (both with and without deformation energy) is obtained for the geometry obtained in the D basis set and is below 0.1 kcal/mol. The errors at the geometries obtained in the T basis set and better are an order of magnitude lower. For such small errors, it is almost meaningless to analyze the deviations for each basis set or extrapolation; nevertheless, the CP-correction to the gradient noticeably increases the accuracy and, at least for the D basis set and the DT extrapolation, does not increase the accuracy compared to the T basis set itself.

**4.2. Methanol Dimer.** In the H-bonded methanol dimer, three intermolecular geometry parameters will be discussed (see Figure 3).

Parameter “R” represents the length of the H-bond, “α” is the angle between the C–O bond of the hydrogen acceptor and the hydrogen involved in bonding, “β” is the angle between the O–H bond and hydrogen acceptor oxygen and “γ” is the dihedral angle between the C–O and H covalently bound and C–O and H weakly bound defined planes of the two monomers. The results of the all-coordinate optimization are summarized in Table 5.

The analysis of the geometrical parameters is more complicated than in the case of the acetylene dimer because of the large number of intermolecular degrees of freedom. Furthermore, an analysis of such small errors obtained using all of the methods considered here is more or less irrelevant.

Let us start the analysis by comparing the CCSD(T)/CBS interaction energies (calculated as MP2/(aTaQ) + ΔCCSD(T)/025015) at geometries optimized using various methods and basis sets. The best agreement with the reference interaction energy, −5.898 kcal/mol, i.e. the one calculated at the complex geometry obtained using the MP2/(TQ)+ΔCCSD(T)/025015 gradient, is obtained for the MP2/(TQ) geometry, with an error of only 0.01 kcal/mol. This is in good agreement with the well-known capability of the MP2 method to properly describe (small) H-bonded complexes. The ordering of the rest of the

**Table 5. Methanol Dimer—Selected Geometrical Parameters Obtained from the Unconstrained Geometry Optimization on Various Theoretical Levels and Basis Sets<sup>a</sup>**

method	hybrid/basis	R [Å]	$\alpha$ [deg]	$\beta$ [deg]	$\gamma$ [deg]	RMSD [Å]	$E_{\text{int}}$
CC	MP2(DT <sup>b</sup> )+ $\Delta$ (025015)	1.8888	102.85	163.57	116.93	0.1220	−5.898
	MP2(TQ)+ $\Delta$ (025015)	1.8876	105.09	167.20	124.79	—	
	MP2(DT)+ $\Delta$ (D)	1.9008	103.84	163.95	118.55	0.0979	
	MP2(TQ)+ $\Delta$ (D)	1.9003	106.99	168.10	125.79	0.0362	
	CP D	1.9711	105.02	156.70	76.40	0.6185	
	CP T	1.9233	104.57	160.59	99.08	0.3152	
MP2.5	D	1.8812	102.90	153.06	108.34	0.3825	−5.861
	MP2(DT)+ $\Delta$ (025015)	1.8890	104.01	164.23	117.66	0.1049	
	MP2(TQ)+ $\Delta$ (025015)	1.8887	107.11	167.87	124.21	0.0357	
	MP2(DT)+ $\Delta$ (D)	1.8972	104.84	164.16	119.13	0.1023	
	MP2(TQ)+ $\Delta$ (D)	1.8972	108.23	168.12	125.39	0.0608	
	CP D	1.9640	105.42	157.04	77.30	0.6046	
MP3	D	1.8955	103.34	153.48	107.91	0.3745	−5.828
	MP2(DT)+ $\Delta$ (025015)	1.8979	104.00	164.73	116.51	0.1025	
	MP2(TQ)+ $\Delta$ (025015)	1.8977	107.31	169.19	122.70	0.0451	
	MP2(DT)+ $\Delta$ (D)	1.9144	105.65	164.88	119.46	0.0894	
	MP2(TQ)+ $\Delta$ (D)	1.9148	109.58	169.26	125.43	0.0808	
	CP D	1.9744	106.01	157.22	78.59	0.5825	
SCS(MI)-MP2	Q	1.8901	107.90	165.97	119.03	0.1111	−5.814
	CP T	1.9033	106.93	159.82	95.74	0.3756	
	CP Q	1.9019	108.39	165.14	111.52	0.1725	
MP2	D	1.8690	102.41	152.62	108.95	0.3888	−5.888
	T	1.8754	103.63	160.74	115.80	0.1971	
	Q	1.8773	105.38	165.23	121.06	0.0758	
	DT	1.8798	104.21	163.15	119.23	0.1285	
	TQ	1.8794	106.91	166.91	124.79	0.0506	
	CP D	1.9550	104.88	156.59	77.60	0.6115	
	CP T	1.9183	104.94	159.05	97.37	0.3569	
	CP Q	1.8958	105.49	163.99	112.60	0.1492	
M06-2X	6-311++G(3df,3pd)	1.9566	103.57	146.46	86.94	0.6136	−5.654
DFT-D	TPSS/6-311++G(3df,3pd)	1.8834	111.72	177.37	126.70	0.1964	−5.943

<sup>a</sup> “ $E_{\text{int}}$ ” corresponds to the CCSD(T)/CBS interaction energy calculated at the particular complex geometry. RMSD stands for the root mean square deviation of the optimized geometry with regard to the reference geometry (i.e., RMSD of “—”). <sup>b</sup> D, T, ... stand for Dunning's cc-pVXZ basis sets. DT and TQ stand for the (X→Y) Helgaker type of basis set extrapolation. “025015” stands for the 6-31G\*\* (0.25, 0.15) modified Pople's basis set, where the exponent of the most diffuse d functions on heavy (C, N, O, ...) atoms is changed from 0.8 to 0.25, and the exponent of the p function of hydrogens is changed from 1.1 to 0.15.

<sup>c</sup> For a particular method “X”, the  $\Delta$  term is defined as  $\Delta = X - \text{MP2}$ .

methods, in terms of accuracy, is as follows: MP2.5, DFT-D, MP3, SCS(MI)-MP2, and M06-2X, with errors of −0.037, 0.045, −0.070, −0.084, and −0.244 kcal/mol.

Certain convergence and method-specific patterns in the geometry changes are worth noting here. The convergence of the H-bond length at the CP-uncorrected and CP-corrected MP2 levels is monotonous, but from the opposite direction. CP-uncorrected MP2 H-bond lengths converge from below (1.8690 Å, in D, to 1.8773 Å, in Q basis set), while the CP-corrected bond length converges from above (1.9550 Å, in D, to 1.8958 Å, in Q). Even more distinctive is the change of the dihedral angle “ $\gamma$ ”, increasing by about 12° at the CP-uncorrected MP2 level, but by about 35° at the CP-corrected level. However, judging by the insensitivity of MP2/CBS interaction energies to these angles in Table 6, the potential energy surface of the dimer is either quite shallow in this direction or the change of the dihedral angle is compensated by the change of other geometrical parameters.

The MP2 interaction energies calculated and optimized with the same basis set vary by 2.39 kcal/mol (~38%), which is comparable to acetylene dimer results. The convergence of the MP2/CBS interaction energies calculated at the (CP-corrected) MP2 geometries in a series of basis sets and basis set extrapolations varies by only 0.19 kcal/mol (3.13%). Inclusion of the deformation energy changes the span of errors only marginally, now being −0.22 kcal/mol (3.84%), smearing out the systematic, though numerically insignificant, higher accuracy of the CP-uncorrected MP2 geometries.

**4.3. Benzene···Water Complex.** The benzene···water complex is a representative of the O—H··· $\pi$  interactions and is thus somewhat more challenging to describe than the “standard” H bond in the methanol dimer (Figure 4).

Among all of the possible degrees of freedom in this complex, we chose the distance between the center of mass of the benzene ring and the hydrogen atom involved in the interaction as well as three separate angles. The intermolecular coordinate “ $\alpha$ ” is the



Table 6. Methanol Dimer, MP2 Method<sup>a</sup>

X	$E_{\text{int}}/X$	$E_{\text{int}}/\text{aTaQ}$	$E_{\text{def}}/X$	$E_{\text{def}}/\text{TQ}$	$E_{\text{bind}}/\text{aTaQ}/X$
D	−3.845	−5.836 (0.069)	0.280	1.157	−5.556 (0.222)
CP-D	−4.098	−5.725 (0.180)	0.138	1.124	−5.587 (0.191)
T	−5.114	−5.909 (−0.004)	0.183	0.225	−5.726 (0.052)
CP-T	−5.098	−5.855 (0.050)	0.141	0.200	−5.713 (0.065)
Q	−5.597	−5.910 (−0.005)	0.140	0.143	−5.770 (0.008)
CP-Q	−5.587	−5.892 (0.013)	0.127	0.134	−5.765 (0.013)
DT	−5.507	−5.900 (0.005)	0.151	0.186	−5.748 (0.030)
TQ	−5.886	−5.905	0.127	0.127	−5.778

<sup>a</sup>  $E_{\text{int}}/X$ : CP-corrected interaction energy in “X” basis set and (CP-)“X” optimized geometry.  $E_{\text{int}}/\text{aTaQ}$ : CP-corrected interaction energy in aTaQ extrapolation and (CP-)“X” optimized geometry.  $E_{\text{def}}/X$  and  $E_{\text{def}}/\text{TQ}$ : deformation energy in “X” basis set or TQ extrapolation and (CP-)“X” optimized geometry.  $E_{\text{bind}}/\text{aTaQ}/X$ : CP-corrected interaction energy in aTaQ extrapolation and (CP-)“X” optimized geometry with corrections for deformation of monomers calculated in “X” basis set and the corresponding geometry. Values in parentheses represent the error with respect to the “reference” TQ values in the corresponding column.

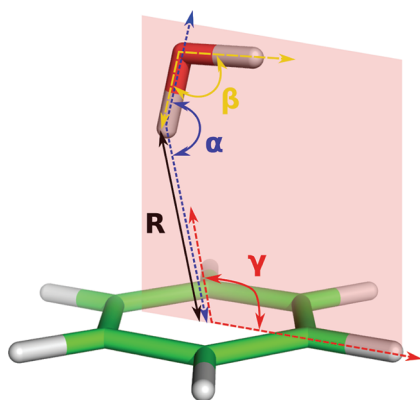


Figure 4. Benzene...water: structure and the analyzed geometrical parameters.

angle between the oxygen, the hydrogen atom involved in the interaction, and the center of mass of the benzene ring, while  $\gamma$  is the angle between the hydrogen atom involved in the interaction, the center of mass of the benzene ring, and the hydrogen atom (of the C–H bond), lying in the intersection of the plane of the water molecule and the benzene ring. Unlike the previous case studies, here we also investigate the intramolecular angle “ $\beta$ ” formed between the hydrogen atoms and the oxygen in water, which was observed to vary with the method/basis set used in the optimization.

The obtained optimized geometrical parameters along with the CCSD(T)/CBS interaction energies are shown in Table 7.

The basis set dependence of the MP2 gradients and interaction energies is analyzed in Table 8, analogously to the previous test cases.

The CCSD(T)/CBS interaction energies calculated at the geometries obtained with the tested methods again show only a narrow spread of about 0.14 kcal/mol ( $\sim 4\%$ ). Comparing the CCSD(T)/CBS interaction energies calculated at the MP2/(TQ) geometry (Table 7) with the one obtained at the MP2/(TQ) level (Table 8), i.e.,  $-3.35$  vs  $-3.55$  kcal/mol, we see that the effect of higher-order correlation is rather small. This can also

be clearly observed in the geometrical parameters, where MP2 delivers essentially the same results as the CCSD(T) method, with the RMSD for DT extrapolation being 0.08 vs 0.09 Å.

One particularly interesting feature of the CP-corrected gradients of the tested methods can be clearly identified. The angle  $\alpha$ , describing the bending of the water molecule toward the benzene ring, converges much faster for the CP-corrected gradients compared to the plain ones. For instance, the values of  $\alpha$  for the plain MP2 gradients in the D to Q basis sets are 115, 134, and 148°, while the value for CP-corrected gradient in the D basis set is 151°, very close to the CBS limit of about 155°. Concerning the MP2/CBS interaction energies calculated at these geometries (see Table 8), the geometry obtained using the plain gradient in the D basis set appears to be almost converged. Geometries obtained using the CP-corrected gradients, along with all tested basis sets (and extrapolations), deliver geometries for which the MP2/CBS interaction energy differs in hundredths of a kilocalorie per mole.

The SCS(MI)-MP2 method does not perform better than the MP2 method itself, judging from the RMSD values. However, the small elongation of the benzene...water distance caused by the higher-order correlation, e.g., 2.441 Å on the best CCSD(T)/CBS estimate in Table 7 vs 2.428 Å on the MP2/(TQ) level, is reproduced at the spin-component scaled MP2 level, 2.504 vs 2.438 Å for SCS(MI)-MP2/Q vs MP2/Q. The performance of the DFT methods is noticeably worse, but still satisfactory. Unlike M06-2X, the TPSS-D method predicts a significantly larger benzene...water separation, 2.739 Å (+ 0.3 Å from the reference), but both methods clearly underestimate the  $\alpha$  angle by as much as approximately 30°.

**4.4. Pyrrol Dimer.** The largest investigated complex, for which the analytic CCSD(T) geometry gradient is still computationally feasible, is the pyrrol dimer in a T-shape tilted conformation, see Figure 5.

According to the known character of  $\pi$ – $\pi$  system interaction, the pyrrol dimer clearly has the largest dispersion contribution among the tested complexes. Due to the fact that the  $\pi$ – $\pi$  dispersion interaction is the most challenging one to describe, the most diverse results for the tested methods are expected. Because of the size of the system, only a limited set of results, in terms of the variety and size of the basis sets, could be obtained, thus offering only a limited space for demonstration of the general trends. The mutual distance between the pyrrol rings is monitored by the parameter “ $R$ ”, which is the distance between the hydrogen, bound to nitrogen, pointing toward the second pyrrol ring and the center of mass of that second pyrrol ring.

The “ $\alpha$ ” and “ $\beta$ ” angles then reflect the tendency of the rings to fold against each other, converging to the so-called “sandwich” structure.

To our surprise, the range of the CCSD(T)/CBS interaction energies obtained for “the best” geometries for each test method is again very narrow, less than 0.1 kcal/mol ( $\sim 1\%$ ; Table 9).

Concerning the similarity of the reference MP2/(DT)+ $\Delta$ CCSD(T)/025015 geometry with the ones obtained using lower level methods, the best match is obtained using the MP2/(DT)+ $\Delta$ MP2.5/025015 hybrid gradient. MP2 and MP3 results are in full agreement with the general trend of MP2 to overestimate the magnitude of the  $\pi$ – $\pi$  dispersion interaction, thus shortening the intermolecular distances, and MP3 having the opposite effect. SCS(MI)-MP2, known for having a much better description of this type of noncovalent interaction, also delivers results in excellent agreement with CCSD(T), being the



**Table 7. Benzene...Water—Selected Geometrical Parameters Obtained from the Unconstrained Geometry Optimization on Various Theoretical Levels and Basis Sets<sup>a</sup>**

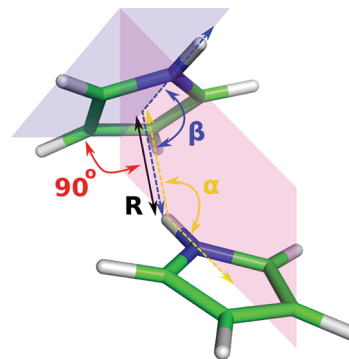
method	hybrid/basis	R [Å]	$\alpha$ [deg]	$\beta$ [deg]	$\gamma$ [deg]	RMSD [Å]	$E_{\text{int}}$
CC	MP2(DT <sup>b</sup> )+ $\Delta^c$ (025015)	2.4623	138.52	102.97	102.70	0.0927	−3.374
	MP2(TQ)+ $\Delta$ (025015)	2.4414	152.31	103.67	102.94	—	
MP2.5	D	2.6960	114.48	100.57	105.95	0.1997	−3.346
	MP2(DT)+ $\Delta$ (025015)	2.4671	139.47	103.15	103.56	0.0783	
	MP2(TQ)+ $\Delta$ (025015)	2.4582	154.12	103.83	103.17	0.0198	
	MP2(DT)+ $\Delta$ (D)	2.4751	140.40	103.19	103.52	0.0706	
	MP2(TQ)+ $\Delta$ (D)	2.4688	155.49	103.86	103.15	0.0311	
	CP D	2.6158	151.03	101.26	100.94	0.0703	
MP3	D	2.7363	114.49	100.88	105.72	0.1945	−3.336
	MP2(DT)+ $\Delta$ (025015)	2.5010	137.98	103.37	103.46	0.0819	
	MP2(TQ)+ $\Delta$ (025015)	2.4887	153.33	104.04	103.12	0.0253	
	MP2(DT)+ $\Delta$ (D)	2.5150	140.02	103.44	103.25	0.0681	
	MP2(TQ)+ $\Delta$ (D)	2.5110	155.66	104.09	103.03	0.0455	
	CP D	2.6539	149.95	101.51	101.20	0.0800	
SCS(MI)-MP2	Q	2.5042	149.37	104.09	103.28	0.0225	−3.291
	CP T	2.5021	153.65	103.76	102.39	0.0301	
	CP Q	2.5152	155.92	104.23	102.93	0.0484	
MP2	D	2.6540	115.12	100.21	105.96	0.2031	−3.346
	T	2.4852	134.28	102.16	103.84	0.1102	
	Q	2.4379	148.08	103.10	103.35	0.0284	
	DT	2.4373	140.47	102.91	103.78	0.0777	
	TQ	2.4277	154.88	103.61	103.20	0.0184	
	CP D	2.5860	151.18	100.94	100.93	0.0622	
	CP T	2.4613	154.49	102.67	101.77	0.0187	
	CP Q	2.4471	155.18	103.26	102.65	0.0174	
M06−2X	6-311++G(3df,3pd)	2.5254	126.88	103.31	103.25	0.1538	−3.234
DFT-D	TPSS/6-311++G(3df,3pd)	2.7387	124.75	102.73	104.84	0.1351	−3.361

<sup>a</sup> “ $E_{\text{int}}$ ” corresponds to the CCSD(T)/CBS interaction energy calculated at the particular complex geometry. RMSD stands for the root mean square deviation of the optimized geometry with regard to the reference geometry (i.e., RMSD of “−”). <sup>b</sup> D, T, ... stand for Dunning's cc-pVXZ basis sets. DT and TQ stand for the (X→Y) Helgaker type of basis set extrapolation. “025015” stands for the 6-31G\*\* (0.25, 0.15) modified Pople's basis set, where the exponent of the most diffuse d functions on heavy (C, N, O, ...) atoms is changed from 0.8 to 0.25 and the exponent of the p function of hydrogens is changed from 1.1 to 0.15. <sup>c</sup> For a particular method “X”, the  $\Delta$  term is defined as  $\Delta = X - \text{MP2}$ .

**Table 8. Benzene...Water, MP2 Method<sup>a</sup>**

X	$E_{\text{int}}/X$	$E_{\text{int}}/\text{aTaQ}$	$E_{\text{def}}/X$	$E_{\text{def}}/\text{TQ}$	$E_{\text{bind}}/\text{aTaQ}/X$
D	−1.509	−3.656 (−0.107)	0.051	1.738	−3.605 (−0.078)
CP-D	−2.203	−3.587 (−0.038)	0.019	1.642	−3.568 (−0.041)
T	−2.735	−3.554 (−0.005)	0.036	0.150	−3.518 (0.009)
CP-T	−2.906	−3.583 (−0.034)	0.019	0.118	−3.564 (−0.037)
Q	−3.250	−3.564 (−0.015)	0.024	0.048	−3.539 (−0.012)
CP-Q	−3.274	−3.562 (−0.013)	0.019	0.042	−3.542 (−0.015)
DT	−3.220	−3.516 (0.033)	0.032	0.086	−3.484 (0.043)
TQ	−3.552	−3.549	0.022	0.022	−3.527

<sup>a</sup>  $E_{\text{int}}/X$ : CP-corrected interaction energy in “X” basis set and (CP-)“X” optimized geometry.  $E_{\text{int}}/\text{aTaQ}$ : CP-corrected interaction energy in aTaQ extrapolation and (CP-)“X” optimized geometry.  $E_{\text{def}}/X$  and  $E_{\text{def}}/\text{TQ}$ : deformation energy in “X” basis set or TQ extrapolation and (CP-)“X” optimized geometry.  $E_{\text{bind}}/\text{aTaQ}/X$ : CP-corrected interaction energy in aTaQ extrapolation and (CP-)“X” optimized geometry with corrections for deformation of monomers calculated in “X” basis set and the corresponding geometry. Values in parentheses represent the error with respect to the “reference” TQ values in the corresponding column.

**Figure 5. Pyrol dimer: Structure and the Analyzed Geometrical Parameters.**

second most accurate one here. The basis set convergence of the MP2 geometries is not monotonous, leading to quite different results for the DT and TQ extrapolations. This is further manifested in noticeably distinct MP2.5 and MP3 results, when used with these extrapolated MP2 gradients. However, the

**Table 9. Pyrrol Dimer—Selected Geometrical Parameters Obtained from the Unconstrained Geometry Optimization on Various Theoretical Levels and Basis Sets<sup>a</sup>**

method	hybrid/basis	R [Å]	$\alpha$ [deg]	$\beta$ [deg]	RMSD [Å]	$E_{\text{int}}$
CC	MP2(DT <sup>b</sup> )+ $\Delta^c$ (025015)	2.2554	136.30	95.07	—	−6.664
MP2.5	D	2.2863	133.10	95.85	0.0329	
	MP2(DT)+ $\Delta$ (025015)	2.2518	136.00	96.08	0.0131	
	MP2(TQ)+ $\Delta$ (025015)	2.2823	132.64	95.93	0.0395	−6.713
	CP D	2.3175	141.60	95.92	0.0985	
MP3	D	2.3254	136.86	95.45	0.0476	
	MP2(DT)+ $\Delta$ (025015)	2.2864	140.20	95.66	0.0624	
	MP2(TQ)+ $\Delta$ (025015)	2.3080	139.38	95.56	0.0581	−6.635
	CP D	2.3644	145.26	95.54	0.1552	
SCS(MI)-MP2	Q	2.2800	136.32	96.69	0.0242	
	CP T	2.2596	137.04	96.67	0.0279	
	CP Q	2.2836	137.64	96.82	0.0402	−6.656
MP2	D	2.2578	129.18	96.22	0.0842	
	T	2.2350	130.66	96.57	0.0713	
	Q	2.2509	129.32	96.72	0.0858	
	DT	2.2270	130.63	96.55	0.0768	
	TQ	2.2747	125.09	96.40	0.1431	−6.672
	CP D	2.2782	137.53	96.37	0.0421	
	CP T	2.2444	134.30	96.64	0.0237	
	CP Q	2.2496	131.86	96.84	0.0516	
M06−2X	6-311++G(3df,3pd)	2.6979	97.64	87.94	0.5564	−6.722
DFT-D	TPSS/6-311++G(3df,3pd)	2.3420	132.84	97.71	0.0391	−6.678

<sup>a</sup> “ $E_{\text{int}}$ ” corresponds to the CCSD(T)/CBS interaction energy calculated at the particular complex geometry. RMSD stands for the root mean square deviation of the optimized geometry with regard to the reference geometry (i.e., RMSD of “—”). <sup>b</sup> D, T, ... stand for Dunning's cc-pVXZ basis sets. DT and TQ stand for the (X→Y) Helgaker type of basis set extrapolation. “025015” stands for the 6-31G\*\* (0.25, 0.15) modified Pople's basis set, where the exponent of the most diffuse d functions on heavy (C, N, O, ...) atoms is changed from 0.8 to 0.25, and the exponent of the p function of hydrogens is changed from 1.1 to 0.15. <sup>c</sup> For a particular method “X”, the  $\Delta$  term is defined as  $\Delta = X - \text{MP2}$ .

**Table 10. Pyrrol Dimer, MP2 Method<sup>a</sup>**

X	$E_{\text{int}}/X$	$E_{\text{int}}/\text{aTaQ}$	$E_{\text{def}}/X$	$E_{\text{def}}/\text{TQ}$	$E_{\text{bind}}/\text{aTaQ}/X$
D	−5.585	−8.174 (−0.160)	0.156	2.331	−8.018 (−0.228)
CP-D	−5.854	−8.025 (−0.011)	0.114	2.237	−7.911 (−0.121)
T	−7.097	−7.988 (0.026)	0.165	0.283	−7.823 (−0.033)
CP-T	−7.144	−7.956 (0.058)	0.149	0.260	−7.807 (−0.017)
Q	−7.591	−7.983 (0.031)	0.177	0.200	−7.807 (−0.017)
CP-Q	−7.589	−7.963 (0.051)	0.161	0.183	−7.803 (−0.013)
DT	−7.754	−7.919 (0.095)	0.175	0.203	−7.744 (0.046)
TQ	−8.035	−8.014	0.224	0.224	−7.790

<sup>a</sup>  $E_{\text{int}}/X$ : CP-corrected interaction energy in “X” basis set and (CP-)“X” optimized geometry.  $E_{\text{int}}/\text{aTaQ}$ : CP-corrected interaction energy in aTaQ extrapolation and (CP-)“X” optimized geometry.  $E_{\text{def}}/X$  and  $E_{\text{def}}/\text{TQ}$ : deformation energy in “X” basis set or TQ extrapolation and (CP-)“X” optimized geometry. Values in parentheses represent the error with respect to the “reference” TQ values in the corresponding column.

differences are slightly buffered by the presence of higher-order correlation.

CP-corrected MP2 gradients are superior to the plain ones in all of the tested basis sets; however, MP2.5 and MP3 CP-corrected results in the D basis sets are inferior to those obtained on the computationally cheaper, hybrid MP2/(DT)+ $\Delta$ /025015 level.

There are large discrepancies in the quality of the results produced by the two DFT methods tested in this work. While the TPSS+D delivers one of the most accurate results, with a RMSD

value comparable with that of MP2.5/CBS, the M06-2X method converges to the most distinct geometry among all of the tested methods. This method produces the largest displacement of the pyrrol rings, with deviations of more than 0.2 Å on average. The  $\alpha$  angle is, on the other hand, 30% smaller, resembling the “sandwich” rather than the “T-shaped” structure.

Concerning the detailed basis set convergence of the plain- and CP-corrected MP2 gradients and interaction energies, shown in Table 10, it is clear that for such a large molecular system, basis set saturation is reached much faster.

The MP2/(aTaQ) interaction energy calculated at the plain MP2 gradient geometry in the D basis set is only 0.16 kcal/mol different than the one obtained for the most accurate MP2/(TQ) geometry. When the interaction energies are calculated in the same basis sets that were used for the geometry optimization, CP-corrected gradients are naturally more accurate. Nevertheless, when the CP-corrected T and extrapolated DT gradients are compared, the latter one leads to a more accurate interaction energy, −7.14 vs −7.75 kcal/mol, for a significantly lower computational cost.

## 5. CONCLUSIONS

We compared the efficiency of the CP-corrected gradients in terms of rate of convergence with respect to basis set size, with extrapolated and hybrid gradients. Essentially, what we have tried to do is transfer the ideas successfully used for noncovalent interaction energy calculations to geometry optimizations.

Furthermore, we have also approached the problem from a pragmatical point of view, seeking to find the level of theory for geometry description that is needed for the subsequent single-point interaction energy calculations (performed at the reliable and accurate CCSD(T) level of theory).

We can generally conclude the following:

1. When the CCSD(T)/CBS (extrapolated MP2 augmented with the higher-order correlation corrections from a CCSD(T)) interaction energy calculated at geometries optimized at various theoretical WFT/DFT levels is used as a measure, only small differences are obtained (less than 5% for the tested complexes).
2. Comparing the RMSD of a complex geometry optimized using a particular WFT/DFT method and basis set with respect to the reference geometry, obtained at the highest computationally feasible level (extrapolated MP2 gradient augmented with the higher-order correlation corrections from CCSD(T)), more pronounced differences are noticeable. CCSD(T) geometries are most faithfully reproduced by the MP2.5 and MP3 methods, followed by the comparably well performing SCS(MI)-MP2 and MP2 methods, finally by the worst performing DFT-D and M06-2X methods.
3. A detailed basis set investigation of the MP2 gradient and interaction energies confirmed the known fact that geometries converge much faster with basis set size than do interaction energies. Geometry optimization using the cc-pVTZ basis set delivers practically converged geometries with respect to the CBS ones.
4. Basis set extrapolation of gradients was shown to improve the results and can be considered as a low-cost alternative to the use of CP-corrected gradients.
5. A hybrid scheme consisting of an MP2 gradient calculated at a larger, "X+1", basis set (or basis set extrapolation) and the post-MP2 (higher-order correlation) gradient in a smaller, "X", basis set was shown to deliver geometries close to those obtained using the plain post-MP2 "X+1" gradient. Analogously to a similar hybrid scheme, which nowadays is routinely used for the calculation of interaction energies, such a hybrid gradient scheme can save a huge amount of computer time, when high accuracy is desired.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: jiri.cerny@img.cas.cz.

## ACKNOWLEDGMENT

This work was a part of the research project no. Z40550506 of the Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, and it was supported by Grant Nos. LC512 and MSM6198959216 from the Ministry of Education, Youth and Sports of the Czech Republic. The support of Praemium Academiae, Academy of Sciences of the Czech Republic, awarded to P.H. in 2007 is also acknowledged. This work was also supported by the Korea Science and Engineering Foundation (World Class Univ. program: R32-2008-000-10180-0) and by the Slovak Grant Agency VEGA under the contract No. 1/0520/10 and the Slovak Research and Development Agency under the contract No. APVV-0059-10. J.C. gratefully acknowledges support by the Institutional Research concept No. AV0Z50520701 of the Academy of Sciences of the Czech

Republic, and an important part of the calculations was performed using the "Chinook" supercomputer at the Environmental Molecular Sciences Laboratory of the Pacific Northwest National Laboratory.

## REFERENCES

- (1) Lee, E. C.; Kim, D.; Jurečka, P.; Tarakeshwar, P.; Hobza, P.; Kim, K. S. *J. Phys. Chem. A* **2007**, *111*, 3446.
- (2) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2004**, *108*, 10200.
- (3) Šponer, J.; Riley, K. E.; Hobza, P. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2595.
- (4) Riley, K. E.; Pitoňák, M.; Jurečka, P.; Hobza, P. *Chem. Rev.* **2010**, *110*, 5023.
- (5) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 289.
- (6) Janesko, B. G.; Henderson, T. M.; Scuseria, G. E. *J. Chem. Phys.* **2009**, *131*, 034110.
- (7) Dąbkowska, I.; Jurečka, P.; Hobza, P. *J. Comput. Chem.* **2005**, *122*, 204322.
- (8) Riley, K. E.; Pitoňák, M.; Černý, J.; Hobza, P. *J. Chem. Theory Comput.* **2010**, *6*, 66.
- (9) Černý, J.; Hobza, P. *Chem. Commun.* **2010**, 46, 383–385.
- (10) Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2005**, *7*, 1624–1626.
- (11) Min, S. K.; Lee, E. C.; Lee, H. M.; Kim, D. Y.; Kim, D.; Kim, K. S. *J. Comput. Chem.* **2008**, *29*, 1208.
- (12) Černý, J.; Tong, X.; Hobza, P.; Müller-Dethlefs, K. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2780.
- (13) Hobza, P.; Zahradník, R. *Weak Intermolecular Interactions In Chemistry and Biology*; Elsevier Scientific Pub. Co.: Amsterdam, 1980.
- (14) Simon, S.; Duran, M.; Dannenberg, J. J. *J. Chem. Phys.* **1996**, *105*, 11024.
- (15) Salvador, P.; Duran, M. *J. Chem. Phys.* **1999**, *111*, 4460.
- (16) Hobza, P.; Havlas, Z. *Theor. Chem. Acc.* **1998**, *99*, 372.
- (17) Hobza, P.; Bludský, O.; Suhai, S. *Phys. Chem. Chem. Phys.* **1999**, *1*, 3073.
- (18) Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2010**, *6*, 66.
- (19) Boese, A. D.; Jansen, G.; Torheyden, M.; Höfener, S.; Klopper, W. *Phys. Chem. Chem. Phys.* **2011**, *13*, 1230.
- (20) Janowski, T.; Pulay, P. *Chem. Phys. Lett.* **2007**, *447*, 27.
- (21) Paizs, B.; Salvador, P.; Császár, A. G.; Duran, M.; Suhai, S. *J. Comput. Chem.* **2001**, *22*, 196.
- (22) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243.
- (23) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Olsen, J. *Chem. Phys. Lett.* **1999**, *302*, 437.
- (24) Pitoňák, M.; Neogrády, P.; Černý, J.; Grimme, S.; Hobza, P. *Chem. Phys. Chem.* **2009**, *10*, 282.
- (25) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–169.
- (26) Lotrich, V.; Flocke, N.; Ponton, M.; Yau, A. D.; Perera, A.; Deumens, E.; Bartlett, R. J. *J. Chem. Phys.* **2008**, *128*, 194104.
- (27) DiStasio, R. A.; Head-Gordon, M. *Mol. Phys.* **2007**, *105*, 1073.
- (28) Frisch, M. J. et al. *Gaussian 03*; Gaussian Inc.: Wallingford CT, 2003.
- (29) Jurečka, P.; Černý, J.; Hobza, P.; Salahub, D. R. *J. Comput. Chem.* **2007**, *28*, 555–569.
- (30) Pitoňák, M.; Neogrády, P.; Řezáč, J.; Jurečka, P.; Urban, M.; Hobza, P. *J. Chem. Theory Comput.* **2008**, *4*, 1829.