J. Phys. Chem. A **2007,** *111,* 5851−5854

Accurate Theoretical Determination of the Structure of Aromatic Complexes Is Complicated: The Phenol Dimer and Phenol···Methanol Cases

Michal Kolář and Pavel Hobza*

Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic and Center for Biomolecules and Complex Molecular Systems, Flemingovo nám. 2, 166 10 Prague 6, Czech Republic

Received: February 22, 2007; In Final Form: April 16, 2007

The structure of the phenol dimer and phenol···methanol complexes was determined by gradient optimization using the Hartree−Fock (HF), MP2, DFT, and RI-DFT-D methods with various basis sets. Theoretical rotational constants were compared with experimental values and the following conclusions were made: (1) HF and DFT methods fail to predict cluster geometries; (2) MP2 with a medium basis set yields reliable cluster geometries but only because of a compensation for errors; (3) when the AO basis set is enlarged, the geometry becomes incorrect, and the theoretical geometry becomes reliable only when the higher correlation energy contributions (CCSD(T)) are included; and (4) the RI-DFT-D procedure covering the dispersion energy provides excellent geometries.

Introduction

Recent years have brought significant progress to the calculation of interaction energies of extended complexes containing aromatic subsystems. It is now well-known that MP2 stabilization energy when determined with extended basis sets (or even at the complete basis set (CBS) limit) is overestimated, which is especially true for stacking structures.[1,2] To reduce the stabilization energy of these structures, it is necessary to include the CCSD(T) term. The CCSD(T) correction term is definitely not negligible, and in our database,[3] consisting of more than 160 complexes, it reaches up to 4 kcal/mol (∼25% of interaction energy). It is necessary to include this correction term not only in the case of stacked structures of aromatic hydrocarbons and DNA bases but also in that of complexes of amino acids.[4] An important question thus arises, namely, how to determine the structure of these systems. When using the MP2/extended basis set treatment (which has nowadays become feasible even for extended complexes), overestimated stabilization energies are expected to yield incorrect geometries. The only straightforward way to prove it would be to compare the structures determined at various theoretical levels with experimental data. The problem lies in the fact that experimental gas-phase structures of extended (aromatic) complexes are extremely rare. Phenol···methanol[5] and phenol···phenol[6,7] complexes would provide unique experimental data.

The latter complex is highly interesting as the structure of the dimer is characterized by two different types of noncovalent interactions: the O−H···O hydrogen bonding and the stacking of two aromatic rings. The hydrogen bonding is mainly of an electrostatic origin while the aromatic stacking is mainly due to London dispersion interactions, both of which participate in determining the equilibrium structure.

In our previous study,[8] we calculated the structure of the phenol dimer using the MP2 method combined with 6-31G-(d,f) and TZVPP basis sets, and the latter calculations agreed relatively well with the experimental rotational constants (with the average error being about 1.5%). A recently published paper[7] on the phenol dimer which contained in addition to the ground-state geometries also the geometries in the first excited-state prompted us to reinvestigate the structure of the dimer. We used a wider set of computational techniques (HF, MP2, DFT, CC2), and unlike in our previous study, we systematically used the Gaussian 03[9] code, which uses stricter convergence criteria (maximum force = 4.5×10^{-4}; root-mean-square force = 3.0×10^{-4}; maximum displacement = 1.8×10^{-3}; root-mean-square displacement = 1.2×10^{-3} which approximately corresponds to an energy change of about 10^{-8}) than the previously used Turbomole[10] code (energy change 10^{-6}; gradient change 10^{-4}). In the case of a flat potential energy surface, tighter convergence criteria can yield a different structure.

The presented study shows that geometries determined with the MP2 procedure employing an extended AO basis set can be less reliable than those calculated with the same method but using only medium AO basis sets. An improvement in the former procedure can only be obtained by considering the CCSD(T) corrections. Two aromatic complexes will be investigated, the phenol dimer and the phenol···methanol complex.

Computations

Methods. Geometry and energy characteristics of all the systems investigated were determined by the Hartree−Fock, MP2, DFT/B3LYP, and CC2 procedures using the cc-pVDZ and cc-pVTZ AO basis sets and TZVP auxilary basis set for all RI methods.[11] Geometry optimizations were performed using the standard procedure as well as the counterpoise-corrected gradient optimization.

Besides the widely used standard methods mentioned, we also used the recently introduced RI-DFT-D method,[12] which covers the London dispersion energy. Because of the parametrization of the method to known, accurate CCSD(T) energies (via the damping function associated with dispersion energy), the method yields accurate interaction energies comparable to the CCSD-(T) values. In this study, we used the TPSS functional[14] and three different AO basis sets: (i) the TZVP basis set, (ii) the Pople 6-311++G(3df,3pd), for which we used the abbreviation LP, and (iii) the modified aug-cc-pVQZ basis set, where both g-functions and the most diffuse f-functions were removed from

10.1021/jp071486+ CCC: $37.00

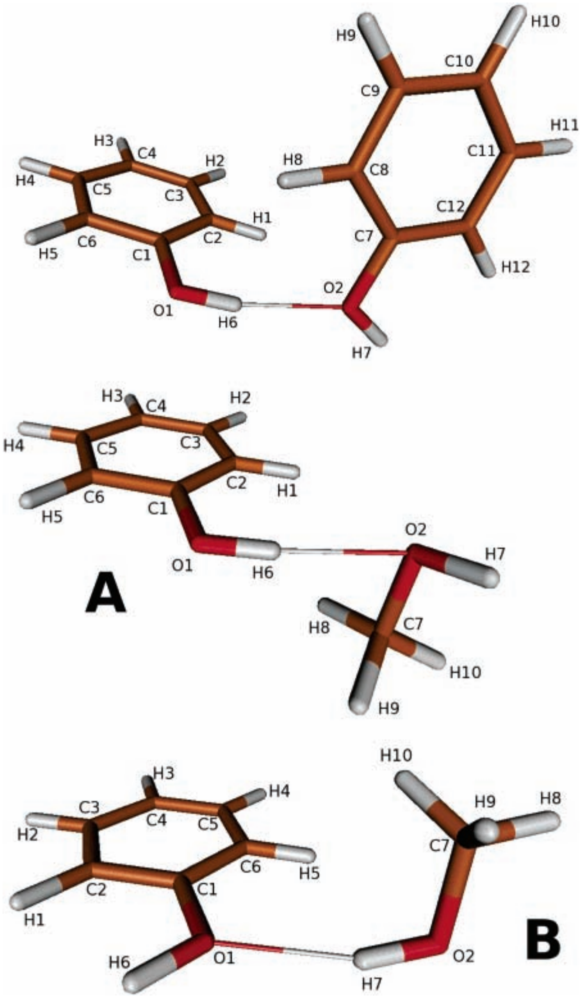

Figure 1. Optimized structures and numbering of phenol dimer and phenol···methanol complexes; the latter complex possesses two structures: A (phenolic hydrogen participates in H-bonding) and B (methanolic hydrogen in H-bonding).

heavy atoms and analogical modifications were made for hydrogens (the abbreviation aQZ′). The RI-DFT-D energies are practically basis set superposition error (BSSE)-free thanks to both the large basis set and the method employed. This means that with this technique, it is not necessary to perform the time-consuming counterpoise-corrected gradient optimization.

Interaction Energies. Interaction energy was systematically determined, including the basis set superposition error (BSSE). In the case of counterpoise-corrected optimization, the final interaction energies covered the deformation energies.

Rotational Constants, Intersystem Angle, and Intersystem Distance. Rotational constants were calculated from the optimized geometries and compared with the experimental values. The final error (FE; in %) was defined as an arithmetic mean of errors of three rotational constants (the difference between the theoretical and experimental values divided by the experimental one) multiplied by 100. The meaning of the intersystem dihedral angle α(C1−O1−O2−C7) is apparent from Figure 1. The intersystem distance *R* is the distance between the centers of the mass of the two subsystems.

Throughout the study we used Gaussian 03[9], Turbomole,[10] and Molpro[13] suites.

Results and Discussion

TABLE 1: Experimental and Theoretical (HF, MP2, B3LYP) Characteristics of the Phenol Dimer[a]

	exptl[b]	basis set	HF	MP2	B3LYP
α	63	cc-pVDZ	99/109	54/71	115/112
		cc-pVTZ	*c*/108	47/40	*c*/119
R	5.251	cc-pVDZ	6.239/5.990	5.328/5.035	6.242/5.856
		cc-pVTZ	*c*/6.174	4.894/3.601	*c*/6.146
E		cc-pVDZ	−3.734	−4.959	−4.133
		cc-pVTZ	*c*	−6.518	*c*
A	1.41699	cc-pVDZ	1.9166/1.8334	1.3893/1.3321	1.9838/1.7914
	cc-pVTZ	*c*/1.9246	1.2723/1.0007	*c*/1.9933	
B	0.31351	cc-pVDZ	0.2268/0.2422	0.3039/0.3307	0.2230/0.2517
		cc-PVTZ	*c*/0.2291	0.3603/0.6230	*c*/0.2313
C	0.28811	cc-pVDZ	0.2212/0.2374	0.2752/0.3118	0.2217/0.2442
		cc-pVTZ	*c*/0.2272	0.3203/0.4905	*c*/0.2258
FE		cc-pVDZ	29.8/27.8	3.1/6.6	30.6/20.4
		cc-pVTZ	*c*/27.9	12.0/66.2	*c*/29.5

[a] Theoretical calculations were performed using the cc-PVDZ and cc-pVTZ basis sets. The values in the nominator and denominator correspond to counterpoise-corrected and counterpoise-uncorrected characteristics. The dihedral angle α(C1−O1−O2−C7) is given in deg, the distance *R* of the centers of mass in given in Å, interaction energies are given in kcal/mol, rotational constants are given in GHz, and error (FE) is given in %. [b] Reference 7. [c] Not calculated

Phenol Dimer. Table 1 shows the experimental and theoretical characteristics of the dimer in question determined utilizing various theoretical procedures in conjunction with the cc-pVDZ and cc-pVTZ basis sets, of which the rotational constants will be investigated first. The final error, FE, is prohibitively large for the HF and DFT calculations, and no improvement is obtained when passing from cc-pVDZ to cc-pVTZ basis set. The same is true for passing from standard to counterpoise-corrected gradient optimization. The explanation is straightforward: neither method covers the London dispersion energy, which is responsible for attraction between the aromatic rings. Consequently, the distance *R* between the aromatic rings is overestimated by almost 1 Å (cf. Table 1), and the intersystem dihedral angle α is too large (by more than 40°).

The correlated MP2 method covers the London dispersion energy, hence, closer agreement with the experiment is expected. When investigating the cc-pVDZ results, we found that the FE error was considerably smaller. This error was further reduced when the counterpoise-corrected constants were taken into consideration, and these values basically agreed with our previous results (ref 8). Good agreement with experimentally proved values was also demonstrated when the angle α was considered where the standard and counterpoise-corrected cc-pVDZ values bracketed the experimental value. Similarly, the uncorrected and counterpoise-corrected distances R also bracketed the experimental value, but the error was still not negligible (about 0.1 Å). We had expected both errors to be reduced significantly when passing to the considerably larger cc-pVTZ basis set. To our surprise, the opposite was true, and the FE error dramatically increased: in the case of the standard optimization to 66% and in the case of the counterpoise-corrected one to a still considerably large value of 12%. (The situation is, however, more complicated, and it is a priori not clear that the counterpoise-corrected gradient optimization yields geometry which is closer to accurate geometry.) The intersystem angle α told us that the dihedral angle was too small, which was confirmed by the significantly underestimated distance *R*. The incorrect results from the MP2/cc-pVTZ calculations can only be explained by the overestimated stabilization energy, which was confirmed by the fact that the FE error was smaller when using the counterpoise-corrected optimization. In this case, the geometry was optimized on the basis of the BSSE-corrected total energy while in the other case it was optimized on the basis of uncorrected total energy. It was mentioned above that the overestimation of the MP2 stabilization energy was removed

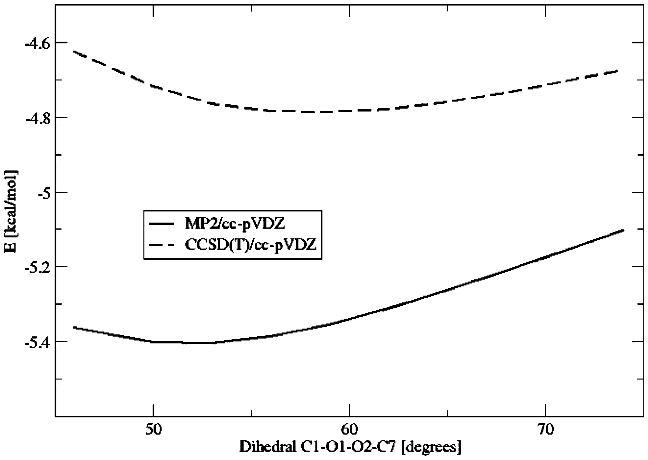

Figure 2. Dependence of CCSD(T) and MP2 interaction energies of phenol dimer on the dihedral angle.

by considering the higher correlation energy terms. Passing from the MP2 to the CC2 procedure did not provide any improvement (not shown), which leads us to the conclusion that an explicit consideration of the CCSD(T) term is inevitable.

The CCSD(T) calculations are even with medium basis sets time-consuming, and for a cluster of the present size gradient optimization is clearly impractical. The only possible way of estimating the role of higher correlation contributions is thus to perform a series of single-point calculations. Figure 2 shows a one-dimensional scan of the dihedral angle α(C1−O1−O2−C7) in dependence on the MP2 and CCSD(T) uncorrected interaction energies (cc-pVDZ basis set). Evidently, both energy curves are quite flat, and the rather large change in the dihedral angle is connected with only a small energy change. Figure 2 shows that the CCSD(T) energy curve is shifted to lower stabilization energies and the higher dihedral angle α, which better agrees with the experimental value (in comparison with the MP2 values). The minima found at the CCSD(T) and MP2 energy curves correspond to 59° and −4.789 kcal/mol, and 53° and −5.404 kcal/mol, respectively. These results clearly indicate the need to determine the geometry of the phenol dimer (and also of any complexes with aromatic electrons) at higher levels than the MP2 level. It is true that when using a medium basis set, the MP2 data agree reasonably well with the experimental values, but this is evidently caused by the compensation of error (the increase in MP2 stabilization energy when the extended basis set is used and its decrease when a method beyond MP2 is applied). The problem remains that this error compensation cannot be relied upon to provide valid results.

Among the various techniques enabling the achievement of CCSD(T) performance at a much lower cost, the modified DFT methods covering the London dispersion energy are in first place. These computationally very favorable methods[12,15,16] make it possible to perform the gradient optimization even for extended complexes. Table 2 summarizes the RI-DFT-D/TPSS results for the phenol dimer when the TZVP, aQZ′, and LP basis sets were used. The calculated FE errors were small, and when extended LP and aQZ′ basis sets were used, the error was close to 1%. The intersystem dihedral angle α determined using all three basis sets agreed well with the experiment, which was also true about the intersystem distance *R*. It should be mentioned here that the performance of the rather small TZVP basis set was very good, which is especially promising for calculations of extended complexes.

Phenol⋯Methanol Complex. Table 3 summarizes both the experimental and the theoretical characteristics of the complex in question. Unlike the previous complex, which possesses only one structure, this one has two possible structures where either phenolic or methanolic hydrogen can participate in H-bonding. On the basis of different acidities of the two hydrogens, it is expected that the former will be more probable. Stabilization energies for this structure are systematically larger than stabilization energies for the other structure, but the latter ones are definitively not negligible; the CH_3OH⋯phenol structure will thus coexist with the C_6H_5OH⋯methanol one (population of both structures at room temperature will be about 1:200). When investigating the FE values, we found again smaller values for the C_6H_5OH⋯methanol structure, and the best agreement with the experiment was achieved using the LP and aQZ′ basis sets; the TZVP values were slightly larger. It must be, however, concluded that the assignment of theoretical constants to the experimental ones is not unambiguous. The RI-DFT-D/aQZ′ values for structures A and B agree fairly well with the experimental numbers, which prevents a clear decision from being made.

TABLE 2: Experimental and Theoretical (RI-DFT-D) Characteristics of the Phenol Dimer[a]

		RI-DFT-D/TPSS		
	exptl[b]	basis set TZVP	basis set LP	basis set aQZ′
α	63	66	67	61
R	5.251	5.288	5.298	5.201
Δ*E*		−4.804	−4.062	−3.360
A	1.41699	1.4422	1.4262	1.3997
B	0.31351	0.2997	0.3055	0.3185
C	0.28811	0.2798	0.2865	0.2925
FE		3.0	1.3	1.4

[a] The dihedral angle α(C1−O1−O2−C7) is given in degrees, the distance *R* of the centers of mass are in Å, interaction energies are in kcal/mol, rotational constants are in GHz, and error (FE) is in %; theoretical calculations were performed using the TZVP, LP, and aQZ′ basis sets. [b] Reference 7.

TABLE 3: Experimental and Theoretical (MP2, RI-DFT-D) Characteristics of the Phenol⋯Methanol Complex[a]

		MP2	
	exp[b]	basis set cc-pVDZ	basis set cc-pVTZ
Δ*E*		4.959	−6.581
A	3290.8	3233.7/3172.8	3226.4/2971.9
B	792.2	841.1/832.7	849.1/867.0
C	685.6	726.2/706.3	734.2/742.7
FE		4.6/3.9	5.4/9.2

		RI-DFT-D/TPSS		
	exptl[b]	basis set TZVP	basis set LP	basis set aQZ′
Δ*E*		−7.484/-3.995	−6.366/-3.156	−5.774/-2.560
A	3290.8	3601.8/3362.7	3332.6/3384.0	3312.6/3311.7
B	792.2	802.4/751.9	793.0/751.3	795.9/779.7
C	685.6	690.6/648.2	680.3/645.4	683.6/665.3
FE		3.8/4.2	0.7/4.6	0.5/1.7

[a] The interaction energies are in kcal/mol, rotational constants are in MHz, and error (FE) is in %. Theoretical calculations were performed using the cc-pVDZ and cc-pVTZ and TZVP, LP, and aQZ′ basis sets, respectively. Values in the nominator and denominator refer to structures A and B. [b] Reference 5.

Like in the case of the phenol dimer, passing to a larger basis set is in the case of MP2 calculations not connected with improving the accuracy of the FE error estimate but with its worsening, which indicates the need to include the higher correlation energy terms. The RI-DFT-D values, which effectively cover these terms, are systematically smaller than MP2/

cc-pVTZ ones. Further, the FE error for the $C_6H_5OH\cdots$methanol structure is systematically smaller than this value for the other structure.

Effect of Basis Set on the Geometry. Accurate geometry (and also stabilization energy) is obtained when using a highly correlated method with extended basis set, and the CCSD(T)/CBS level represents one possibility. For gradient geometry optimization, this level is impractical, and some compromise should be selected. The use of basis sets with diffuse functions (like aug-cc-pVDZ) in combination with MP2 method is generally not recommended because the respective stabilization energy can be overestimated. Hence, an extended basis set without diffuse functions is to be used. In this case we rely on compensation of errors, and the question remains which basis set should be used. In other words we try to find a level where the compensation of errors is the most efficient. As shown in this paper the MP2/cc-pVDZ level provides better geometries than the MP2/cc-pVTZ level. We are certainly aware of the fact that it is impossible to rely on the compensation of errors mentioned and a computational procedure yielding accurate geometries as well as stabilization energies should be selected. If a method provides accurate stabilization energies, than it is possible to expect that accurate geometries will also result. The RI-DFT-D procedure represents such a level, and it was shown repeatedly that RI-DFT-D stabilization energies agree very well with the benchmark data. In this paper we have shown that the method also yields excellent geometries. Contrary to wave function theories (WFT), the DFT methods are less sensitive to the presence of diffuse functions, and in the present paper we have shown that the TZVP basis set yields interaction energies and geometries similar to those of LP and aQZ′ basis sets containing diffuse functions. Let us close this paragraph by stating that the RI-DFT-D method provides excellent geometries also for isolated extended systems with folded structures (e.g., helicenes) where standard WFT calculations fail.[17]

Conclusions

(i) MP2 calculations with medium AO basis sets yield reliable structures. Enlarging the AO basis set provides overestimated MP2 stabilization energies and inaccurate MP2 geometries. The geometrical characteristics are improved by the inclusion of the higher correlation energy contributions. The reliable results from the MP2/medium basis set calculations are thus a result of the compensation of errors.

(ii) The CCSD(T) calculations are impractical for the geometry optimization of the complexes of the present size. The RI-DFT-D/TPSS calculations, which are also feasible in the case of extended complexes, yield reliable geometrical constants even with a rather small TZVP basis set. When larger LP and aQZ′ basis sets are adopted, excellent geometrical characteristics results are achieved.

(iii) Conclusions made for the phenol dimer are also valid for the phenol$\cdots$methanol complex, and also here the best agreement with the experimental rotational constants is obtained employing the RI-DFT-D/TPSS/aQZ′ procedure. The best structure geometries are attached in Supporting Information.

Acknowledgment. This work is part of the Research Project No. Z4055905 and was supported by Grants EC512 (from the Ministry of Education of the Czech Republic) and A400550510 (from the Grant Agency of the Academy of Sciences of the Czech Republic).

Supporting Information Available: Phenol$\cdots$phenol DFT-D/TPSS/LP, phenol$\cdots$methanol A DFT-D/TPSS/aQZ′, and phenol$\cdots$methanol B DFT-D/TPSS/aQZ′ structure geometries. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Sinnokrot, M. O.; Sherill, C. D. *J. Phys. Chem. A* **2004**, *108*, 10200.

(2) Jurečka, P.; Hobza, P. *J. Am. Chem. Soc.* **2003**, *125,* 15608.

(3) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985.

(4) Vondrášek, J.; Bendová, L.; Klusák, V.; Hobza, P. *J. Am. Chem. Soc.* **2005**, *127*, 2615.

(5) Schmitt, M.; Kuepper, J.; Spangenberg, D.; Westpahl, A. *Chem. Phys.* **2000**, *254*, 349.

(6) Weichert, A.; Riehn, C.; Weichert, A.; Brutschy, B. *J. Phys. Chem. A* **2001**, *105*, 5679.

(7) Schmitt, M.; Boehm, M.; Ratzer, C.; Kruegler, D.; Kleinermanns, K.; Kalman, I.; Berden, G.; Meerts, W. L. *ChemPhysChem* **2006**, *7*, 1241.

(8) Hobza, P.; Riehn, C.; Weichert, A.; Brutschy, B. *Chem. Phys.* **2002**, *283*, 331.

(9) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

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

(11) Feyereisen, M.; Fitzgerald, G. Komornicki, A. *Chem. Phys. Lett.* **1993**, *208*, 359.

(12) Jurečka, P.; Černý, J.; Hobza, P.; Salahub, D. R. *J. Comput. Chem.* **2007**, *28*, 555.

(13) Werner, H.-J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schütz, M.; Celani, P.; Korona, T.; Rauhut, G.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Hampel C.; Hetzer, G.; Lloyd, A. W.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pitzer, R.; Schumann, U.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T. *MOLPRO, a package of ab initio programs*, version 2006.1; see http://www.molpro.net.

(14) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146.

(15) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463; **2006**, *27*, 1787.

(16) Elstner, M; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149.

(17) Rulíšek, L.; Valdés, H.; Klusák, V.; Exner, O.; Starý, I.; Hobza, P. *J. Comp. Chem.*, submitted.