JCTC Journal of Chemical Theory and Computation

Assessment of the Performance of DFT and DFT-D Methods for Describing Distance Dependence of Hydrogen-Bonded Interactions

Kanchana S. Thanthiriwatte, Edward G. Hohenstein, Lori A. Burns, and C. David Sherrill*

Center for Computational Molecular Science and Technology, School of Chemistry and Biochemistry and School of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, United States

Received August 18, 2010

Abstract: Noncovalent interactions such as hydrogen bonds, van der Waals forces, and $\pi-\pi$ interactions play important roles influencing the structure, stability, and dynamic properties of biomolecules including DNA and RNA base pairs. In an effort to better understand the fundamental physics of hydrogen bonding (H-bonding), we investigate the distance dependence of interaction energies in the prototype bimolecular complexes of formic acid, formamide, and formamidine. Potential energy curves along the H-bonding dissociation coordinate are examined both by establishing reference CCSD(T) interaction energies extrapolated to the complete basis set limit and by assessing the performance of the density functional methods B3LYP, PBE, PBE0, B970, PB86, M05-2X, and M06-2X and empirical dispersion corrected methods B3LYP-D3, PBE-D3, PBE0-D3, B970-D2, BP86-D3, and ωB97X-D, with basis sets 6-311++G(3df,3pd), aug-cc-pVDZ, and aug-cc-pVTZ. Although H-bonding interactions are dominated by electrostatics, it is necessary to properly account for dispersion interactions to obtain accurate energetics. In order to quantitatively probe the nature of hydrogen bonding interactions as a function of distance, we decompose the interaction energy curves into physically meaningful components with symmetry-adapted perturbation theory (SAPT). The SAPT results confirm that the contribution of dispersion and induction are significant at and near equilibrium, although electrostatics dominate. Among the DFT/DFT-D techniques, the best overall results are obtained utilizing counterpoise-corrected ωB97X-D with the aug-cc-pVDZ basis set.

1. Introduction

Hydrogen-bonding (H-bonding) interactions, in addition to $\pi-\pi$ stacking interactions, provide a significant contribution to the stability and conformational arrangement of nucleic acids.[1−5] Though it is well understood that H-bonding arises mainly from electrostatic interactions,[6] the accurate determination of attendant interaction energies has been the subject of many theoretical studies.[1−23] While H-bonding between DNA/RNA base pairs is limited to O−H···N and N−H···N bonding patterns, the O−H···O motif is also accessible with nonstandard base pairs such as the uracil dimer.[22] As a step to better understanding hydrogen bonding in biomolecules, bimolecular complexes of formic acid (FaOO), formamide (FaON), and formamidine (FaNN) have been studied as prototypes for H-bonding in DNA/RNA base pairs.

Dissociation potential energy curves of the H-bonded complexes are governed not only by the dominant electrostatic component, but also by dispersion interactions near equilibrium and at medium distances. While it is well established that Hartree−Fock (HF) and density functional theory (DFT) generally recover most of the interaction energies in the H-bonded systems, these methods fail to

* To whom correspondence should be addressed: E-mail: sherrill@gatech.edu.

10.1021/ct100469b
Published on Web 12/02/2010

describe dispersion interactions accurately.[24−29] Reliable computation of noncovalent interactions requires high-level treatment of electron correlation by methods such as coupled-cluster theory with singles and doubles including perturbative triples, CCSD(T).[30] This last approach has emerged as the most accurate computationally affordable method that can be applied to small systems and is considered the "gold standard" for chemical accuracy,[31] though it is unfortunately limited by its $O(N^7)$ complexity in its standard form. To describe noncovalent interactions in large systems, less computationally expensive methods must be employed.

Density functional theory[32] is used extensively to investigate a variety of chemical systems, as it generally requires fewer (for nonhybrid functionals) or equivalent (for hybrid functionals) computer resources compared to Hartree−Fock theory and provides reasonably accurate results.[32] The application of DFT to noncovalent interactions, however, has been hampered by the failure of most density functionals to describe "long-range" electron correlation.[24−29] Several approaches exist for improving existing density functionals. The addition of empirical terms to model dispersion interactions (i.e., the DFT-D method) is the most popular approach.[33−39] The hybrid meta exchange correlation functionals developed by Truhlar and co-workers are also promising candidates.[40−43] Truhlar's M05-2X and M06-2X functionals are said to account for "medium-range" electron correlation, which is sufficient for describing the dispersion interactions within many smaller complexes near their equilibrium geometries.[43]

We have examined potential energy curves of the FaOO−FaOO, FaON−FaON, FaNN−FaNN, FaOO−FaON, FaON−FaNN, and FaOO−FaNN bimolecular complexes, which we will refer to as the HBC6 test set. We present benchmark-quality complete-basis-set (CBS) extrapolated CCSD(T) potential energy curves for these molecules and also assess the performance of popular DFT and DFT-D methods. Previously, most investigations characterizing H-bonded complexes have focused on equilibrium configurations (from either geometry optimizations or crystal structures).[3,21] Recently, Hobza and co-workers also studied H-bonded complexes (formamide dimer, methylamine dimer, and methanol dimer) along the dissociation pathway.[5] The present work significantly extends these computations by examining five new H-bonded complexes, larger basis sets for the coupled-cluster correction, a much more extensive array of DFT methods, and wavefunction-based symmetry-adapted perturbation theory (SAPT) analysis to gain additional insight into the physical basis for attractive interactions in our target complexes. The systematic study performed here should be very helpful in elucidating the fundamental nature of H-bonding interactions (including their distance dependence) and in providing additional benchmark CCSD(T)/CBS data for H-bonded systems, which we have found necessary in independent work on parametrization of dispersion-including DFT methods.

2. Theoretical Methods

Geometry optimizations were performed for the homogeneous and heterogeneous bimolecular complexes of FaOO, FaON, and FaNN in planar hydrogen-bonded configurations. The structures of these dimers along the dissociation coordinate were fully optimized at the CCSD(T)/aug-cc-pVDZ[44,45] level of theory under the constraint of fixed intermonomer distances. We defined the dissociation coordinate for our systems as the distance between the central carbon atoms of each monomer in the bimolecular complexes. It is noteworthy that most of the complexes in HBC6, particularly FaOO−FaOO and FaOO−FaNN, show double proton transfer with "short, strong" or "low-barrier" hydrogen bonds.[46−52] Because our interest here is primarily in noncovalent interactions and not in proton transfer reactions, we have neglected very short intermolecular distances where the proton transfer occurs. Single-point energy computations along the dissociation coordinate were performed with second order Møller−Plesset perturbation theory (MP2) using Dunning's aug-cc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ basis sets,[44,45] as well as CCSD(T) with the aug-cc-pVDZ and aug-cc-pVTZ basis sets.

Estimates of the CBS limit of the CCSD(T) interaction energies were obtained first by extrapolating to the CBS MP2 limit using the two-point extrapolation scheme of Halkier et al.[53] with aug-cc-pVTZ and aug-cc-pVQZ basis sets. Next, we added a "coupled-cluster correction", ΔCCSD(T), estimated as the difference between CCSD(T) and MP2 in a smaller basis set. Janowski and Pulay[54] have emphasized that basis sets beyond aug-cc-pVDZ can be important for high-quality estimates of ΔCCSD(T), and we have recently confirmed this for several other van der Waals dimers.[16,21] In this work, the estimates of ΔCCSD(T) were evaluated using aug-cc-pVDZ/aug-cc-pVTZ extrapolations of the MP2 and CCSD(T) correlation energies. The final CCSD(T)/CBS estimation was obtained by summing the energies of Hartree−Fock/aug-cc-pVQZ, the extrapolated MP2/CBS correlation, and the estimated ΔCCSD(T) correction. All CCSD(T) and MP2 computations were performed with the core electrons frozen.

All geometry optimizations were performed with the ACES II program suite,[55] and single-point energy computations utilized the Molpro2009 package of ab initio programs.[56] DFT computations along each curve were performed using the Q-Chem 3.2 suite of programs[57] utilizing a Lebedev grid with 302 angular points for each of the 100 radial shells. This grid is larger than the default of most electronic structure program packages, but it is necessary to avoid artifacts due to numerical integration for noncovalent interactions, particularly when using meta-GGA functionals.[24,58] Energy evaluations were performed for the B3LYP,[59] PBE,[60] PBE0,[61] B970,[62] BP86,[63,64] M05-2X,[40] M06-2X,[41,42] and ωB97X-D[39,65] functionals along with the 6-311++G(3df,3pd), aug-cc-pVDZ, and aug-cc-pVTZ basis sets. Here, we denote the original B97 functional as B970 to distinguish it from the reparameterized B97 functional developed as part of the B97-D method of Grimme,[37] which has previously been shown to perform well for potential curves of more weakly bound van der Waals dimers.[17] The empirical dispersion correction (denoted -D2 and -D3) introduced by Grimme[37,38] was added to B3LYP, PBE, PBE0, B970, and BP86 functionals. The dispersion interaction energy (E_{disp}) is given by

$$E_{\mathrm{disp}} = -\sum_{n}\sum_{\mathrm{AB}} s_n \frac{C_n^{\mathrm{AB}}}{R_{\mathrm{AB}}^n} f_{\mathrm{d},n}(R_{\mathrm{AB}}), n = 6, 8 \quad (1)$$

where s_n is a global scaling factor, C_n^{AB} denotes the averaged nth-order dispersion coefficient for atom pair AB, R_{AB} is interatomic distance of atom pair AB, and $f_{\mathrm{d},n}(R_{\mathrm{AB}})$ is the damping function. The use of a damping function minimizes double counting of "short-range" correlation effects captured by the density functional. In the -D2 approach, the damping function is given by

$$f_{\mathrm{d},n}(R_{\mathrm{AB}}) = \frac{1}{1 + \mathrm{e}^{-d(R_{\mathrm{AB}}/R_{\mathrm{r}}-1)}}, n = 6 \quad (2)$$

where R_{r} represents the sum of atomic radii and d is a global damping parameter that controls the "sharpness" of the damping function. The R^{-6} term is the leading term in the expansion of the dispersion energy of which higher order (R^{-8}, R^{-10}, ...) terms are omitted. The empirical dispersion energies were scaled by a global factor, s_6, which is 1.05 for B3LYP, 0.75 for PBE, 0.6 for PBE0, and 1.05 for BP86, as recommended by Grimme[37] and 0.75 for B970.[66,67] The damping function of the -D3 approach is given by

$$f_{\mathrm{d},n}(R_{\mathrm{AB}}) = \frac{1}{1 + 6(R_{\mathrm{AB}}/s_{\mathrm{r},n}R_{\mathrm{AB}}^0)}, n = 6, 8 \quad (3)$$

where $s_{\mathrm{r},n}$ is the order-dependent scaling factor of the cutoff radii R_{AB}^0. The -D3 correction was added to the B3LYP, PBE, PBE0, and BP86 functionals. Chai and Head-Gordon's[39] ωB97X-D functional incorporates a similar R^{-6} dispersion term to improve its performance for treating noncovalent complexes. The dispersion corrected functionals will be referred to as either DFT-D2 or DFT-D3 appropriately. Grimme's DFT-D2 empirical dispersion correction was implemented in Q-Chem 3.2 by two of the authors (C.D.S. and E.G.H.).

To reduce basis set superposition error, we applied the Boys−Bernardi scheme[68] to yield counterpoise-corrected interaction energies, $E_{\mathrm{cp}}^{\mathrm{int}}$. We also obtained "relaxed" interaction energies, $E_{\mathrm{rlx}}^{\mathrm{int}}$, by adding to $E_{\mathrm{cp}}^{\mathrm{int}}$ the deformation energies:

$$E_{\mathrm{rlx}}^{\mathrm{int}} = E_{\mathrm{cp}}^{\mathrm{int}} + (E_{\mathrm{A}}^{\mathrm{A}} - E_{\mathrm{A}'}^{\mathrm{A}}) + (E_{\mathrm{B}}^{\mathrm{B}} - E_{\mathrm{B}'}^{\mathrm{B}}) \quad (4)$$

where $E_{\mathrm{A}}^{\mathrm{A}} - E_{\mathrm{A}'}^{\mathrm{A}}$ represents the deformation energy required to take monomer A from its equilibrium geometry to that in the complex (and, as denoted by the superscript, this energy difference is evaluated in the basis of monomer A). Thus, $E_{\mathrm{rlx}}^{\mathrm{int}}$ denotes the counterpoise-corrected interaction energy of the dimer relative to infinitely separated monomers in their equilibrium geometries, whereas $E_{\mathrm{cp}}^{\mathrm{int}}$ denotes the counterpoise-corrected interaction energy of the dimer relative to infinitely separated monomers frozen at the monomer geometries they adopt in the dimer.

The interaction energies of the six H-bonded dimers were computed with symmetry-adapted perturbation theory.[69,70] SAPT computes the interaction in terms of individual energy components. For the purpose of our analysis, we will group the individual terms as

$$E_{\mathrm{electrostatic}} = E_{\mathrm{elst}}^{(10)} + E_{\mathrm{elst,resp}}^{(12)} + E_{\mathrm{elst,resp}}^{(13)} \quad (5)$$

$$E_{\mathrm{exchange}} = E_{\mathrm{exch}}^{(10)} + E_{\mathrm{exch}}^{(11)} + E_{\mathrm{exch}}^{(12)} \quad (6)$$

$$E_{\mathrm{induction}} = E_{\mathrm{ind,resp}}^{(20)} + E_{\mathrm{exch-ind,resp}}^{(20)} + {}^{t}E_{\mathrm{ind}}^{(22)} + {}^{t}E_{\mathrm{exch-ind}}^{(22)} + \delta E_{\mathrm{ind,resp}}^{(\mathrm{HF})} \quad (7)$$

$$E_{\mathrm{dispersion}} = E_{\mathrm{disp}}^{(20)} + E_{\mathrm{disp}}^{(30)} + E_{\mathrm{disp}}^{(21)} + E_{\mathrm{disp}}^{(22)} + E_{\mathrm{exch-disp}}^{(20)} \quad (8)$$

The SAPT computations have been performed with our recently developed SAPT program.[71] This program has been implemented within the framework of PSI 3.4.[72] The level of SAPT shown above, which we will denote SAPT2+(3), when paired with the aug-cc-pVDZ basis, can be expected to provide energies within 0.4 kcal mol^{-1} of benchmark results for systems of this size.[73] All SAPT computations have been performed under the density fitting (DF) approximation using the auxiliary aug-cc-pVDZ-RI basis.[74]

3. Results and Discussion

The performance of density functional methods for H-bonded interaction energies at different distances is illustrated in Figure 1. Relaxed, CP-corrected interaction energies vs dissociation coordinate distance have been plotted for each of the six complexes of HBC6 using an aug-cc-pVTZ basis set in combination with B3LYP-D3, PBE-D3, PBE0-D3, BP86-D3, B970-D2, M05-2X, M06-2X, and ωB97X-D and their underlying uncorrected DFT functionals where applicable. The interaction energy curves for B3LYP-D2, PBE-D2, PBE0-D2, and BP86-D2 can be found in the Supporting Information. The CCSD(T)/CBS curves are also presented for reference. Although standard DFT methods are often thought to be reliable for describing H-bonding interactions, it can be clearly observed that DFT methods B970, B3LYP, and BP86 are significantly underbound, with average absolute errors (MAEs) at the potential minima of 1.2−2.0 kcal mol^{-1}. In contrast, PBE-D3, BP86-D3, PBE0-D3, and B3LYP-D3 approaches are overbound with respect to the reference energies with MAE values of 1.6−1.8 kcal mol^{-1}. Surprisingly, the uncorrected PBE and PBE0 functionals exhibit far better performance than their DFT-D counterparts, with MAEs of 0.43 and 0.35 kcal mol^{-1} at their equilibrium geometries. These two are the most accurate DFT treatments, along with B970-D2, M05-2X, ωB97X-D, and M06-2X, all of which yield potential minimum MAE values less than 0.5 kcal mol^{-1}. All methods display the correct asymptotic behavior at the dissociation limit.

Mean absolute error statistics for $E_{\mathrm{rlx}}^{\mathrm{int}}$ with respect to CCSD(T)/CBS values across the entire potential energy curves are compiled in Table 1 for the functionals examined. All three basis sets show similar performance. The DFT methods B970, B3LYP, and BP86 as well as the dispersion corrected DFT methods PBE-D3, PBE0-D3, BP86-D3, and their -D2 counterparts consistently give overall MAE values (weighted average over six test cases, rightmost columns) greater than 0.9 kcal mol^{-1} for all basis sets. M05-2X, B970-D, ωB97X-D, and M06-2X have errors reliably less than 0.4 kcal mol^{-1}. The best-performing functional (for relaxed interaction energies) is M06-2X, with ωB97X-D the second-best. The -D3 MAE for both the PBE0 and BP86 functionals

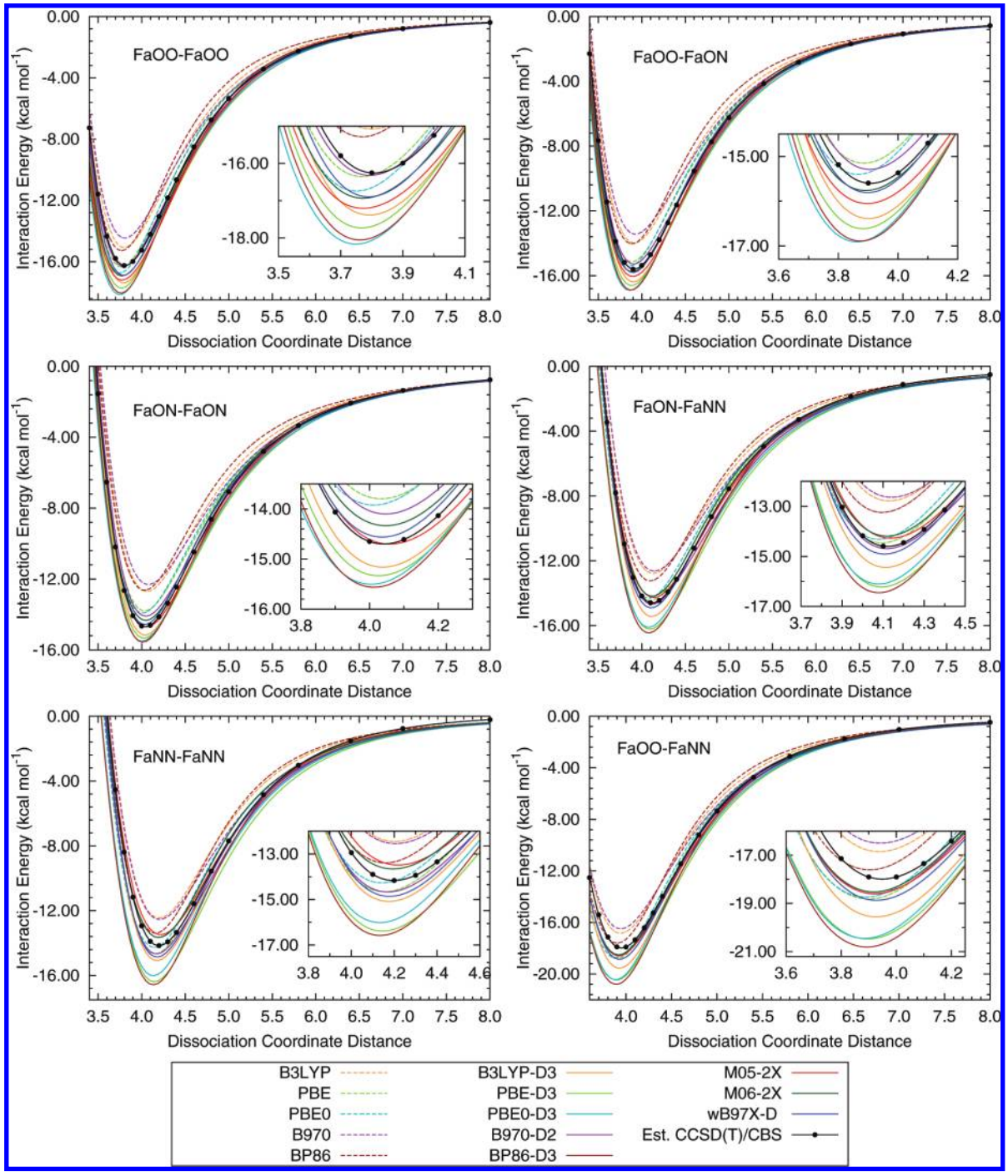

Figure 1. Relaxed counterpoise-corrected interaction energies at the DFT(-D)/aug-cc-pVTZ and CCSD(T)/CBS levels of theory.

is about 0.06−0.1 kcal mol^{-1} higher than the -D2 error. PBE-D3 and B3LYP-D3 show a MAE value of 0.1−0.2 kcal mol^{-1} improvement in reliability compared with their -D2 counterparts, although PBE-D3 has an error higher than 1 kcal mol^{-1}. We note that the -D3 correction to the popular B3LYP functional leads to overall MAEs of around 0.6 kcal mol^{-1}, which is not as good as the meta-GGA functionals, ωB97X-D, or B970-D2, but better than the others.

It is also worthwhile to examine the counterpoise-corrected interaction energies, E_{cp}^{int} (without the addition of deformation energies), because these are somewhat more convenient for use as benchmarks for assessing new theoretical methods. The MAE statistics are tabulated in Table 2, and errors with respect to the CCSD(T)/CBS reference are plotted in Figure 2 (negative values indicate overbinding). Our results show a considerable increase in the error at intermolecular distances shorter than the equilibrium distance. Most methods are overbound at closer intermonomer distances except B3LYP, B970, and BP86, which are highly underbound. However, the dispersion corrected B970-D functional is the second best performing functional with the 6-311++G(3df,3pd) and aug-cc-pVTZ basis sets, and the third best with the aug-cc-pVDZ basis set. The DFT/DFT-D methods B3LYP,

Table 1. Errors in DFT(-D) Relaxed Interaction Energies with Respect to CCSD(T)/CBS[a]

	mean absolute error (MAE) in kcal mol^{-1}							
Method	FaOO−FaOO	FaON−FaON	FaNN−FaNN	FaOO−FaON	FaON−FaNN	FaOO−FaNN	HBC6	HBC6(D2)
				6-311++G(3df,3pd)				
B970	0.95	1.59	1.02	1.27	1.31	0.83	1.17	
B3LYP	0.77	1.41	1.22	1.07	1.31	0.76	1.10	
BP86	0.86	1.44	0.71	1.13	0.87	0.68	0.95	
PBE0	0.50	0.44	0.78	0.44	0.51	0.50	0.53	
PBE	0.36	0.55	0.80	0.32	0.34	0.44	0.47	
PBE0-D3	1.07	0.69	1.75	0.88	1.23	1.31	1.15	1.09
PBE-D3	0.90	0.51	2.01	0.71	1.27	1.42	1.13	1.35
BP86-D3	0.91	0.53	1.96	0.71	1.24	0.26	1.12	1.03
B3LYP-D3	0.68	0.27	0.71	0.48	0.50	0.87	0.58	0.71
M05-2X	0.54	0.20	0.42	0.26	0.29	2.33	0.33	
B970-D2	0.16	0.43	0.67	0.24	0.19	0.40	0.35	
ωB97X-D	0.33	0.15	0.60	0.14	0.26	0.51	0.33	
M06-2X	0.36	0.33	0.41	0.24	0.35	0.27	0.33	
				aug-cc-pVDZ				
B970	1.00	1.53	1.01	1.27	1.28	0.81	1.16	
B3LYP	0.80	1.34	1.17	1.06	1.25	0.72	1.06	
BP86	0.84	1.39	0.62	1.12	0.82	0.62	0.91	
PBE0	0.40	0.30	0.69	0.31	0.39	0.45	0.42	
PBE	0.27	0.47	0.84	0.24	0.26	0.46	0.42	
PBE0-D3	1.05	0.79	1.78	0.92	1.30	1.36	1.20	1.14
PBE-D3	0.90	0.60	2.06	0.74	1.34	1.49	1.18	1.40
BP86-D3	0.89	0.57	1.98	0.72	1.35	1.41	1.15	1.03
B3LYP-D3	0.65	0.39	0.76	0.50	0.56	0.91	0.62	0.75
M05-2X	0.38	0.38	0.50	0.14	0.41	0.24	0.34	
B970-D2	0.13	0.44	0.69	0.28	0.19	0.44	0.36	
ωB97X-D	0.34	0.16	0.66	0.18	0.35	0.56	0.37	
M06-2X	0.26	0.29	0.32	0.10	0.22	0.25	0.24	
				aug-cc-pVTZ				
B970	0.93	1.55	0.97	1.23	1.26	0.79	1.13	
B3LYP	0.78	1.38	1.19	1.06	1.28	0.74	1.08	
BP86	0.86	1.41	0.70	1.10	0.83	0.68	0.93	
PBE0	0.54	0.43	0.80	0.46	0.52	0.53	0.54	
PBE	0.39	0.51	0.84	0.33	0.35	0.47	0.48	
PBE0-D3	1.08	0.73	1.80	0.91	1.27	1.34	1.19	1.13
PBE-D3	0.91	0.55	2.05	0.73	1.31	1.45	1.16	1.38
BP86-D3	0.92	0.57	1.99	0.75	1.27	1.41	1.15	1.07
B3LYP-D3	0.67	0.29	0.74	0.50	0.53	0.89	0.60	0.73
M05-2X	0.73	0.13	0.30	0.47	0.17	0.45	0.37	
B970-D2	0.16	0.38	0.73	0.18	0.21	0.44	0.35	
ωB97X-D	0.32	0.16	0.61	0.15	0.25	0.50	0.33	
M06-2X	0.48	0.20	0.37	0.25	0.27	0.34	0.32	

[a] MAEs are calculated across each potential energy curve. The column HBC6 denotes the MAEs over all six potential energy curves.

B970, BP86, PBE-D2/-D3, and PBE0-D2/-D3 show poor performance for H-bonded E_{cp}^{int}, similar to the results for E_{rlx}^{int}. M05-2X, M06-2X, B970-D2, and ωB97X-D demonstrate MAE values of less than 0.4 kcal mol^{-1}, with ωB97X-D and B970-D as the best performing. The MAE of -D3 in the unrelaxed interaction energies for BP86 and PBE0 is up to 0.08 kcal mol^{-1} higher than -D2 with respect to CCSD(T)/CBS values. B3LYP-D3 and PBE-D3 have improved the reliability by 0.1−0.2 kcal mol^{-1} from B3LYP-D2 and PBE-D2 functionals.

Our computations show that the DFT methods B3LYP, BP86, and B970 cannot describe the H-bonding interaction quantitatively. All three offer MAE values greater than 1.25 kcal mol^{-1} and across the entire test set are significantly underbound with respect to the CCSD(T)/CBS reference. Although these methods should be able to describe the dominant electrostatic contributions, in order to obtain accurate interaction energies, the long- and medium-range dispersion contributions also need to be added. After careful observation of the B3LYP, B970, and BP86 error curves in Figure 2, one might suggest that they behave as $1/R^6$. B3LYP-D3 and BP86-D3 compensate for this deficiency (although they often overcompensate at short intermolecular distances), and B970-D2 shows excellent performance with MAE values less than 0.30 kcal mol^{-1}. Hence, in order to accurately represent interaction energies in H-bonded systems, the dispersion energy contribution needs to be included. Unlike B3LYP, BP86, and B970, our results show that standard PBE and PBE0 are fairly reliable across the potential energy curves for all three basis sets. The MAEs in interaction energies for PBE and PBE0 are 0.52 and 0.40 kcal mol^{-1}, respectively, with the aug-cc-pVTZ basis. Since pure PBE and PBE0 already behave well for the interaction energies of these systems, the addition of an empirical dispersion energy term makes PBE-D2/-D3 and PBE0-D2/-D3 worse for interaction energies in H-bonded systems. Interaction energy error curves for M06-2X are better than for M05-

Table 2. Errors in DFT(-D) Unrelaxed Interaction Energies with Respect to CCSD(T)/CBS[a]

	mean absolute error (MAE) in kcal mol^{-1}							
method	FaOO−FaOO	FaON−FaON	FaNN−FaNN	FaOO−FaON	FaON−FaNN	FaOO−FaNN	HBC6	HBC6(D2)
				6-311++G(3df,3pd)				
BP86	1.24	1.80	1.02	1.55	1.40	1.13	1.36	
B970	1.06	1.59	1.36	1.32	1.47	1.02	1.31	
B3LYP	0.90	1.44	1.56	1.16	1.49	1.01	1.26	
PBE	0.44	0.93	0.30	0.71	0.46	0.25	0.52	
PBE0	0.40	0.43	0.45	0.34	0.33	0.37	0.39	
PBE0-D3	0.97	0.67	1.31	0.82	1.01	1.09	0.98	0.90
BP86-D3	0.51	0.25	1.27	0.35	0.76	0.79	0.65	0.61
PBE-D3	0.47	0.17	1.24	0.28	0.70	0.78	0.60	0.80
B3LYP-D3	0.56	0.28	0.35	0.41	0.32	0.60	0.42	0.61
M05-2X	0.63	0.14	0.55	0.36	0.29	0.22	0.37	
M06-2X	0.35	0.20	0.34	0.21	0.27	0.20	0.26	
B970-D2	0.14	0.46	0.23	0.30	0.18	0.17	0.25	
ωB97X-D	0.29	0.08	0.32	0.15	0.17	0.39	0.23	
				aug-cc-pVDZ				
BP86	1.41	1.91	1.06	1.69	1.48	1.23	1.47	
B970	1.23	1.69	1.41	1.47	1.55	1.11	1.42	
B3LYP	1.07	1.54	1.60	1.31	1.56	1.11	1.37	
PBE	0.59	1.01	0.29	0.83	0.51	0.32	0.59	
PBE0	0.33	0.51	0.42	0.33	0.30	0.34	0.37	
PBE0-D3	0.82	0.59	1.28	0.70	0.95	1.02	0.89	0.81
BP86-D3	0.36	0.18	1.24	0.23	0.70	0.70	0.57	0.57
PBE-D3	0.32	0.15	1.23	0.19	0.65	0.72	0.54	0.57
B3LYP-D3	0.38	0.25	0.32	0.29	0.27	0.50	0.33	0.41
M05-2X	0.37	0.34	0.69	0.14	0.48	0.12	0.36	
M06-2X	0.19	0.31	0.35	0.16	0.32	0.16	0.25	
B970-D2	0.28	0.54	0.19	0.42	0.25	0.10	0.30	
ωB97X-D	0.18	0.12	0.29	0.10	0.14	0.32	0.19	
				aug-cc-pVTZ				
BP86	1.25	1.81	0.98	1.56	1.39	1.12	1.35	
B970	1.06	1.59	1.30	1.32	1.44	1.01	1.29	
B3LYP	0.92	1.45	1.53	1.18	1.48	1.02	1.27	
PBE	0.45	0.93	0.31	0.72	0.44	0.24	0.52	
PBE0	0.43	0.44	0.46	0.36	0.34	0.39	0.40	
PBE0-D3	0.96	0.66	1.36	0.81	1.03	1.10	0.98	0.91
BP86-D3	0.51	0.27	1.32	0.36	0.78	0.81	0.67	0.64
PBE-D3	0.46	0.16	1.29	0.27	0.72	0.79	0.61	0.81
B3LYP-D3	0.53	0.27	0.39	0.39	0.33	0.59	0.41	0.51
M05-2X	0.75	0.13	0.44	0.48	0.20	0.34	0.39	
M06-2X	0.43	0.13	0.30	0.26	0.20	0.21	0.26	
B970-D2	0.11	0.42	0.29	0.27	0.14	0.18	0.24	
ωB97X-D	0.27	0.09	0.33	0.13	0.15	0.36	0.22	

[a] MAEs are calculated across each potential energy curve. The column HBC6 denotes the MAEs over all six potential energy curves.

2X, but both MAE values are under 0.40 kcal mol^{-1} for all three basis sets. The best performing DFT/DFT-D functional for unrelaxed interaction energies of the HBC6 set is ωB97X-D, which gives MAE values of 0.23, 0.19, and 0.22 kcal mol^{-1} with 6-311++G(3df,3pd), aug-cc-pVDZ, and aug-cc-pVTZ, respectively.

The HBC6 test set shows that appropriately chosen DFT-D methods and meta-GGA methods can accurately describe noncovalent interactions in H-bonded systems. To further understand the nature of these H-bonding interactions, we decompose the interaction energy to its components (electrostatics, exchange, induction, and dispersion) as described in eqs 5−8. We use a ternary diagram to interpret the results of our SAPT computations.[75] In a ternary diagram, each of three quantities is represented by a vertex which denotes 100% of the considered quantity, and the opposite side of the triangle is 0% in that quantity. Any other value between 0% and 100% can be placed between the vertex for that quantity and the opposite side. Figure 3 shows the ternary diagram for the attractive components of the total interaction energies: electrostatics, induction, and dispersion. For each system in the HBC6 test set, we plot a dot for each intermolecular distance R. The dots that are closer to the electrostatics vertex belong to the large intermolecular distances, and dots that are closer to the middle of the diagram correspond to the shorter intermolecular distances for each system. The colored lines correspond to the SAPT attractive components at the minima of the FaOO−FaOO, FaON−FaON, and FaNN−FaNN systems. It is clear that electrostatics dominate at long range, taking over from induction and dispersion contributions, which account for 10−40% and 10−20%, respectively, of attraction at short to intermediate distances. This decomposition analysis agrees with our results, in which DFT methods generally fail to describe the interaction energies in H-bonded systems accurately, while DFT-D methods and meta-GGA

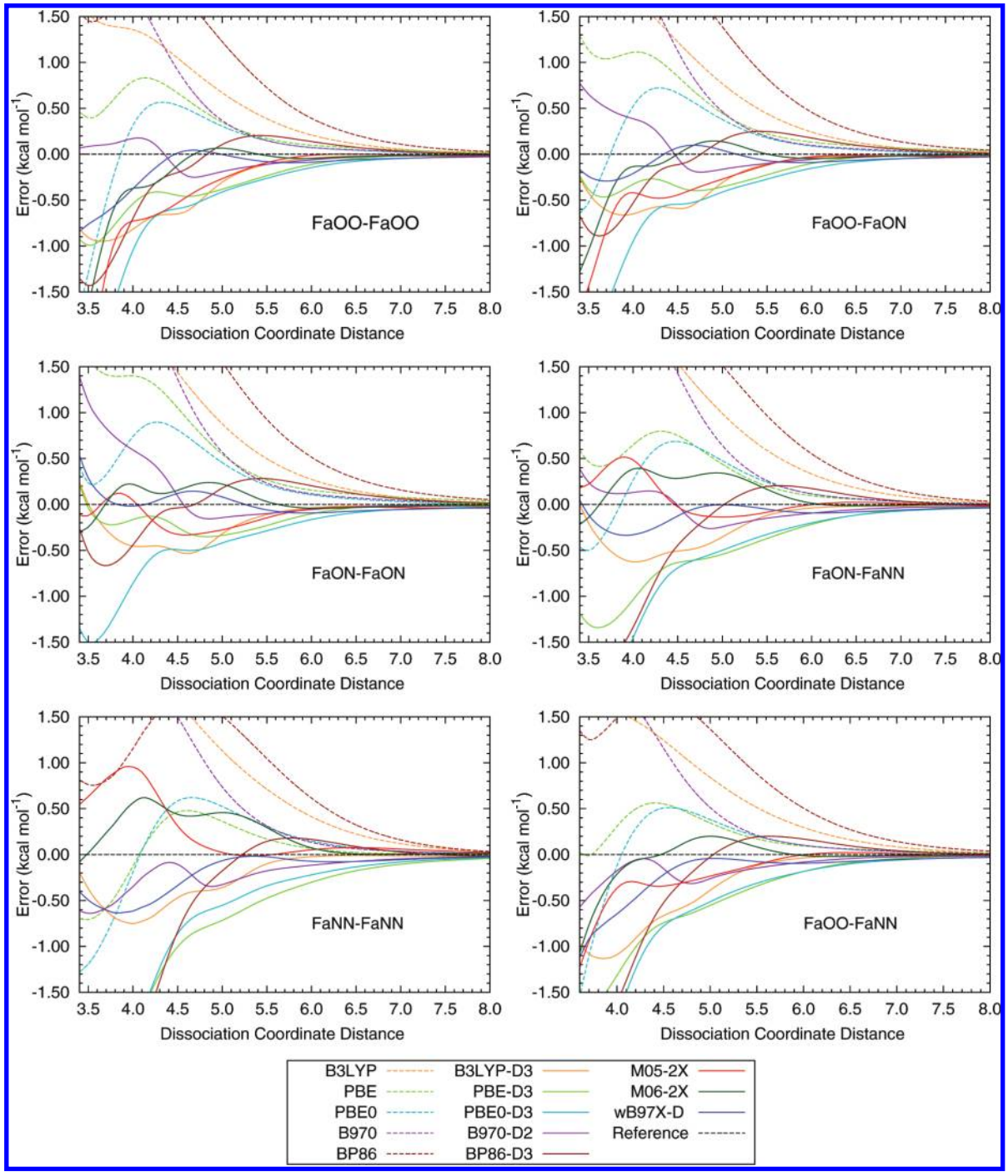

Figure 2. Errors in counterpoise-corrected DFT(-D)/aug-cc-pVTZ unrelaxed interaction energies with respect to CCSD(T)/CBS.

methods that claim to recover "medium-range" electron correlation show success.

4. Conclusions

We have studied counterpoise corrected interaction energies (with and without deformation energy corrections) for the HBC6 test set of H-bonded complexes. The benchmark CCSD(T)/CBS potential curves reported here should be helpful in testing and parametrizing new approximate methods for noncovalent interactions and supplement existing benchmark data which is primarily available for potential energy minima only. DFT-D and meta-GGA DFT methods provide significant improvements over traditional density functionals for these systems. H-bonded complexes are dominated by electrostatic influences, and thus traditional density functionals are generally capable of computing reasonable binding energies. However, our results have shown that the DFT methods B3LYP, B970, and BP86 perform poorly for our test set, while PBE and PBE0 provide accurate results. The meta-GGA functionals and several of the DFT-D methods perform even better, and this improvement is

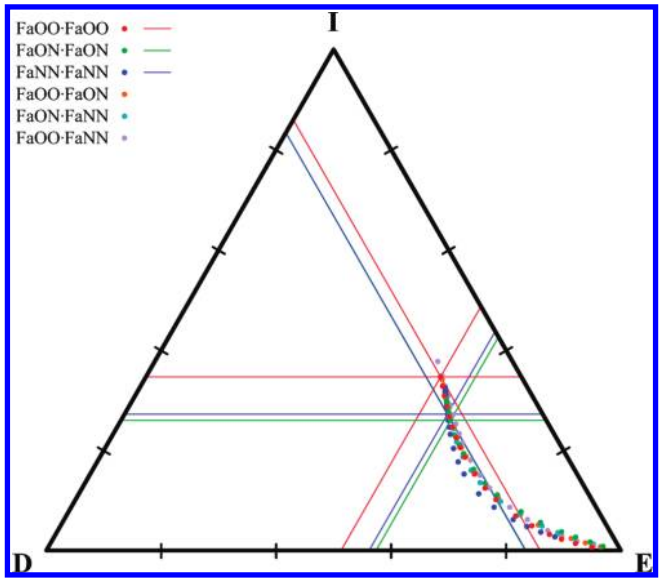

Figure 3. Relationship between Electrostatics (E)−Dispersion (D)−Induction (I) in SAPT analysis for HBC6 test set. Proximity to a vertex indicates an increasing fraction of the attraction coming from that component. The dots that are closer to the electrostatics vertex belong to the large intermolecular distances, and dots that are closer to the middle of the diagram correspond to the shorter intermolecular distances for each system (see the text).

ascribed to the importance of dispersion interactions at medium range, as revealed by SAPT analysis. Thus, although H-bonded interactions are dominated by electrostatic interactions, the contribution of London dispersion forces is not negligible in computing accurate interaction energies.

Acknowledgment. This material is based upon work supported by the National Science Foundation through grant CHE-1011360.

Supporting Information Available: Errors in DFT(-D2/-D3) relaxed interaction energies, errors in DFT(-D2/-D3) unrelaxed interaction energies, Cartesian coordinates of optimized geometries, and counterpoise-corrected interaction energies for all considered DFT/DFT-D methods of the HBC6 test set. This information is available free of charge via the Internet at http://pubs.acs.org/.

References

(1) Pitoňák, M.; Riley, K. E.; Neogrady, P.; Hobza, P. *ChemPhysChem* **2008**, *9*, 1636–1644.

(2) Jurečka, P.; Hobza, P. *J. Am. Chem. Soc.* **2003**, *125*, 15608–15613.

(3) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.

(4) Sinnokrot, M. O.; Sherrill, C. D. *J. Am. Chem. Soc.* **2004**, *126*, 7690–7697.

(5) Riley, K. E.; Pitonak, M.; Cerny, J.; Hobza, P. *J. Chem. Theory Comput.* **2010**, *6*, 66–80.

(6) Scheiner, S. *Hydrogen Bonding: A Theoretical Perspective*; Oxford Univ Press: New York, 1997.

(7) Frisch, M. J.; Del Bene, J. E.; Binkley, J. S.; Schaefer, H. F. *J. Chem. Phys.* **1986**, *84*, 2279–2289.

(8) Feller, D.; Boyle, C. M.; Davidson, E. R. *J. Chem. Phys.* **1987**, *86*, 3424–3440.

(9) Szalewicz, K.; Cole, S. J.; Kolos, W.; Bartlett, R. J. *J. Chem. Phys.* **1988**, *89*, 3662–3673.

(10) Feller, D. *J. Chem. Phys.* **1992**, *96*, 6104–6114.

(11) Manalo, M. N.; Prez, L. M.; Li Wang, A. *J. Am. Chem. Soc.* **2009**, *129*, 11298–11299.

(12) Tapavicza, E.; Lin, I.-C.; von Lilenfeld, A.; Tavernelli, I.; Coutinho-Neto, M. D.; Rothlisberger, U. *J. Chem. Theory Comput.* **2007**, *3*, 1673.

(13) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 415–432.

(14) Mignon, P.; Loverix, S.; De Proft, F.; Geerlings, P. *J. Phys. Chem. A* **2004**, *108*, 6038–6044.

(15) Šponer, J.; Jurečka, P.; Hobza, P. *J. Am. Chem. Soc.* **2004**, *126*, 10142–10151.

(16) Sherrill, C. D.; Takatani, T.; Hohenstein, E. G. *J. Phys. Chem. A* **2009**, *113*, 10146–10159.

(17) Vazquez-Mayagoitia, A.; Sherrill, C. D.; Apra, E.; Sumpter, B. G. *J. Chem. Theory Comput.* **2010**, *6*, 727–734.

(18) Tsuzuki, S.; Lüthi, H. P. *J. Chem. Phys.* **2001**, *114*, 3949–3957.

(19) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149–5155.

(20) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 5590–5596.

(21) Takatani, T.; Hohenstein, E. G.; Malagoli, M.; Marshall, M. S.; Sherrill, C. D. *J. Chem. Phys.* **2010**, *132*, 144104.

(22) Jurečka, P.; Hobza, P. *Chem. Phys. Lett.* **2002**, *365*, 89–94.

(23) Mo, Y. *J. Mol. Model.* **2006**, *12*, 665–672.

(24) Johnson, E. R.; Wolkow, R. A.; DiLabio, G. A. *Chem. Phys. Lett.* **2004**, *394*, 334–338.

(25) Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2005**, *7*, 1624–1626.

(26) Allen, M. J.; Tozer, D. J. *J. Chem. Phys.* **2002**, *117*, 11113–11120.

(27) Hobza, P.; Šponer, J.; Reschel, T. *J. Comput. Chem.* **1995**, *16*, 1315–1325.

(28) Kristyán, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175–180.

(29) Kurita, N.; Sekino, H. *Int. J. Quantum Chem.* **2003**, *91*, 355–362.

(30) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

(31) Lee, T. J.; Scuseria, G. E. Achieving Chemical Accuracy with Coupled-Cluster Theory. In *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*; Langhoff, S. R., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1995.

(32) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford: New York, 1989; Volume *16* International Series of Monographs on Chemistry.

(33) Zimmerli, U.; Parrinello, M.; Koumoutsakos, P. *J. Chem. Phys.* **2004**, *120*, 2693–2699.

(34) Jurecka, P.; Cerny, J.; Hobza, P.; Salahub, D. R. *J. Comput. Chem.* **2007**, *28*, 555–569.

(35) Wu, Q.; Yang, W. *J. Chem. Phys.* **2002**, *116*, 515–524.

(36) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463–1473.

(37) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

(38) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.

(39) Chai, J.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615.

(40) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Phys.* **2005**, *123*, 161103.

(41) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364–382.

(42) Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 194101.

(43) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

(44) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(45) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

(46) Ushiyama, H.; Takatsuka, K. *J. Chem. Phys.* **2001**, *115*, 5903–5912.

(47) Shetty, S.; Pal, S.; Kanhere, D. G.; Goursot, A. *Los Alamos Natl. Lab., Prepr. Arch., Condens. Matter* **2004**, 1–30.

(48) Podolyan, Y.; Gorb, L.; Leszczynski, J. *J. Phys. Chem. A* **2002**, *106*, 12103–12109.

(49) Miura, S.; Tuckerman, M. E.; Klein, M. L. *J. Chem. Phys.* **1998**, *109*, 5290–5299.

(50) Kim, Y.; Lim, S.; Kim, Y. *J. Phys. Chem. A* **1999**, *103*, 6632–6637.

(51) Dedikova, P.; Pitonak, M.; Neogrady, P.; Cernusak, I.; Urban, M. *J. Phys. Chem. A* **2008**, *112*, 7115–7123.

(52) Aplincourt, P.; Bureau, C.; Anthoine, J.-L.; Chong, D. P. *J. Phys. Chem. A* **2001**, *105*, 7364–7370.

(53) Halkier, A.; Klopper, W.; Helgaker, T.; Jørgensen, P.; Taylor, P. R. *J. Chem. Phys.* **1999**, *111*, 9157–9167.

(54) Janowski, T.; Pulay, P. *Chem. Phys. Lett.* **2007**, *447*, 27–32.

(55) ACES II program is a product of the Quantum Theory Project, University of Florida. Authors: Stanton, J. F. Gauss, J. Watts, J. D. Nooijen, M. Oliphant, N. Perera, S. A. Szalay, P. G. Lauderdale, W. J. Gwaltney, S. R. Beck, S. Balková, A. Bernholdt, D. E. Baeck, K.-K. Sekino, H. Rozyczko, P. Huber, C. Bartlett. R. J. Integral packages included are VMOL (Almöf, J. Taylor, P. R.), VPROPS (P. R. Taylor), and a modified version of the ABACUS integral derivative package(Helgaker, T. U. Aa. Jensen, H. J. Olsen, J. Jørgensen, P. Taylor P. R.).

(56) Werner, H.-J. MOLPRO, version 2009.1. http://www.molpro.net (accessed Nov 2010).

(57) Shao, Y.; et al. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.

(58) Johnson, E. R.; Becke, A. D.; Sherrill, C. D.; Di Labio, G. A. *J. Chem. Phys.* **2009**, *131*, 034111.

(59) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

(60) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(61) Adamo, C.; Scuseria, G. E.; Barone, V. *J. Chem. Phys.* **1999**, *111*, 2889–2899.

(62) Becke, A. D. *J. Chem. Phys.* **1997**, *107*, 8554–8560.

(63) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(64) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822–8824.

(65) Chai, J. D.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 084106.

(66) The DFT-D2 s_6 parameter for the B970 functional was obtained by minimizing the mean absolute percent deviation for the S22 test set calculated with un-counterpoise-corrected and equally weighted aug-cc-pVDZ and aug-cc-pVTZ basis sets. A more thorough procedure, taking into account BSSE-corrected interaction energies, shifts the recommended parameter only slightly, to 0.80.

(67) Burns, L. A.; Vázquez-Mayagoitia, Á.; Sumpter, B. G.; Sherrill, C. D. Manuscript in preparation.

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

(69) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887–1930.

(70) Williams, H. L.; Szalewicz, K.; Jeziorski, B.; Moszynski, R.; Rybak, S. *J. Chem. Phys.* **1993**, *98*, 1279–1292.

(71) Hohenstein, E. G.; Sherrill, C. D. *J. Chem. Phys.* **2010**, *132*, 184111.

(72) Crawford, T. D.; Sherrill, C. D.; Valeev, E. F.; Fermann, J. T.; King, R. A.; Leininger, M. L.; Brown, S. T.; Janssen, C. L.; Seidl, E. T.; Kenny, J. P.; Allen, W. D. *J. Comput. Chem.* **2007**, *28*, 1610–1616.

(73) Hohenstein, E. G.; Sherrill, C. D. *J. Chem. Phys.* **2010**, *133*, 014101.

(74) Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175–3183.

(75) Singh, N. J.; Min, S. K.; Kim, D. Y.; Kim, K. S. *J. Chem. Theory Comput.* **2009**, *5*, 515–529.

CT100469B