Title

Beyond Energies: Geometries of Nonbonded Molecular Complexes as Metrics for Assessing Electronic Structure Approaches

Journal

Journal of Chemical Theory and Computation, 11(4)

ISSN

1549-9618

Authors

Witte, Jonathon
Goldey, Matthew
Neaton, Jeffrey B
et al.

Publication Date

2015-04-14

DOI

10.1021/ct501050s

Peer reviewed

Beyond Energies: Geometries of Non-bonded Molecular Complexes as Metrics for Assessing Electronic Structure Approaches

Journal:	*Journal of Chemical Theory and Computation*
Manuscript ID:	ct-2014-01050s.R1
Manuscript Type:	Article
Date Submitted by the Author:	03-Mar-2015
Complete List of Authors:	Witte, Jonathon; University of California, Berkeley, Chemistry Goldey, Matthew; University of Chicago, Institute for Molecular Engineering Neaton, Jeffrey; Lawrence Berkeley National Laboratory, Molecular Foundry Head-Gordon, Martin; University of California, Berkeley, Chemistry

Beyond Energies: Geometries of Non-bonded Molecular Complexes as Metrics for Assessing Electronic Structure Approaches

Jonathon Witte,[†,‡] Matthew Goldey,[¶] Jeffrey B. Neaton,[§,‡,‖] and Martin Head-Gordon[*,†,⊥]

Department of Chemistry, University of California, Berkeley, California 94720, United States, Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States, Institute for Molecular Engineering, The University of Chicago, Chicago, Illinois 60637, United States, Department of Physics, University of California, Berkeley, California 94720, Kavli Energy Nanosciences Institute at Berkeley, Berkeley, California 94720, United States, and Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

E-mail: mhg@cchem.berkeley.edu

Abstract

Electronic structure approaches for calculating intermolecular interactions have traditionally been benchmarked almost exclusively on the basis of energy-centric metrics.

[*]To whom correspondence should be addressed
[†]Department of Chemistry, University of California, Berkeley, California 94720, United States
[‡]Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States
[¶]Institute for Molecular Engineering, The University of Chicago, Chicago, Illinois 60637, United States
[§]Department of Physics, University of California, Berkeley, California 94720, United States
[‖]Kavli Energy Nanosciences Institute at Berkeley, Berkeley, California 94720, United States
[⊥]Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

Herein, we explore the idea of utilizing a metric related to geometry. On a diverse series of non-covalently-interacting systems of different sizes, from the water dimer to the coronene dimer, we evaluate a variety of electronic structure approximations with respect to their abilities to reproduce coupled-cluster-level geometries. Specifically, we examine Hartree-Fock, second-order Møller-Plesset perturbation theory (MP2), attenuated MP2, scaled MP2, and a number of density functionals, many of which include empirical or nonempirical van der Waals dispersion corrections. We find a number of trends that transcend system size and interaction type. For instance, functionals incorporating VV10 nonlocal correlation tend to yield highly accurate geometries; ωB97X-V and B97M-V in particular stand out. We establish that intermolecular distance, as measured by, e.g., the center-of-mass separation of two molecules, is the geometric parameter that deviates most profoundly among the various methods. This property of the equilibrium intermolecular separation, coupled with its accessibility via a small series of well-defined single-point calculations, makes it an ideal metric for the development and evaluation of electronic structure methods.

1 Introduction

There exist a tremendous number of approaches to approximately solving the non-relativistic, time-independent Schrödinger equation for a many-electron system, and it is often not clear *a priori* what the optimal choice for a particular application is. The well-defined hierarchy of wavefunction-based (WF) methods offers a clear path to obtaining highly accurate energies, though at great expense. The simplest method for adding electronic correlation to the Hartree-Fock (HF) mean field ansatz,[1,2] second-order Møller-Plesset perturbation theory (MP2),[3] scales as $O(N^5)$, and more highly-correlated methods, such as coupled-cluster theory with single, double, and perturbative triple excitations – CCSD(T)[4] – exhibit even worse scaling ($\mathrm{O}(N^7)$ in the case of CCSD(T)). The resolution-of-the-identity (RI) approximation[5–8] partly addresses this cost by reducing the computational prefactor, but the under-

lying scaling of the method to which it is applied is left unchanged. Additionally, correlated WF methods exhibit slow convergence with respect to basis set size, a consequence of their inclusion of excited-state determinants and the correspondingly large number of basis functions required to accurately describe the virtual space associated therewith.[9,10] Though still in its infancy, attenuation of the Coulomb operator is one promising means of addressing not only these issues, but basis set superposition error (BSSE) as well.[11,12]

In stark contrast to the clear-cut hierarchy of WF methods, the well-known Jacob's ladder[13] of density functional theory (DFT)[14,15] offers no clear way to systematically improve results. Moreover, the inherently local description provided by DFT within standard approximations renders it incapable of recovering long-range dispersion.[16] In the absence of strong permanent electrostatic interactions, these second-order effects are crucial for the correct description of non-covalent interactions, which will be the focus of this paper. In the past decade, a variety of means of accounting for long-range van der Waals dispersion forces have been proposed, from simple pairwise C_6 corrections to the exchange-correlation energy[17–21] to the inclusion of fully nonlocal correlation kernels.[22–25]

When evaluating the performance of any of these various electronic structure approaches, the recent literature has focused almost exclusively on the ability of the method to reproduce "exact" electronic energies. Whenever a promising new method is developed, a flurry of studies arise wherein this method is applied to a variety of systems of physical interest. The common thread in such studies is their focus on energies; for instance, in a non-covalently interacting system, the metric of choice is typically the binding energy, defined – at least for size-consistent methods – as the difference between the total energy of the system and the energies of its constituent molecules. Although there may be subtle differences in the precise definition of the binding energy among various studies (e.g., the fragments may or may not be allowed to relax) the fact remains that it is an objective metric: it provides a means of methodically comparing electronic structure approaches. Unfortunately, this sort of energy-centric approach to method evaluation is far from perfect. The ideal method would

recover the entire "exact" potential energy surface, not just a single point; it would reproduce "exact" geometries as well as energies.

This energy-centric focus has been justified repeatedly over the years by studies in which intramolecular geometric parameters were considered. It has been well established that – in any reasonable basis – even HF yields accurate bond lengths and angles for many systems, and differences between various methods, as measured on the basis of these metrics, tend to be minimal.[26–31] Although these sorts of intramolecular metrics are sufficient for describing geometries of small single molecules, they are inadequate in the context of large molecules and systems composed of multiple molecules, i.e. systems involving significant non-covalent interactions. In such systems, quantitative comparison of geometries is difficult; distilling $3N-6$ geometric degrees of freedom into a single objective metric is a distinctly non-trivial endeavor. Despite being difficult, this is an important avenue of research. Intermolecular interactions are orders of magnitude weaker than covalent bonds. They involve relatively shallow potential surfaces, and as such, some of these softer degrees of freedom may be useful for the evaluation of electronic structure approaches.

In the context of non-covalently-interacting systems, the systematic evaluation of electronic structure methods with regard to their description of geometries is a largely undeveloped idea. There have been a number of studies in which binding energy curves were generated and studied, though the principle metric of evaluation has in every case been the equilibrium binding energy, not the location of the minimum nor the shape of the curve.[32–38] There have also been a handful of studies in recent years in which hydrogen bond lengths predicted by various methods were compared;[39–41] additionally, there has been a study by Vydrov and Van Voorhis in which the performances of various van der Waals density functionals were evaluated on the basis of their ability to predict intermolecular separation in small CO_2-containing complexes,[42] and a study by Remya and Suresh in which a tremendous number of density functionals were screened on the basis of their abilities to minimize the overall root-mean-square deviation with regard to CCSD geometries of ten small complexes.[43] In the

same time frame, there have been countless studies in which electronic structure approaches have been evaluated solely on the basis of their abilities to reproduce single-point energies. Moreover, the conclusions that can be drawn from these few geometry-based studies are limited, in some cases by confinement to a single interaction motif, in others by the questionable quality of the reference structures, and in all cases by a focus on only small systems.

In this work, we evaluate a variety of electronic structure approximations with regard to their abilities to reproduce complete-basis CCSD(T)-level geometric parameters on a diverse set of systems for which high-quality reference data is readily available. We explore the impacts of interaction type and system size on the performances of the various methods. Moreover, we establish a procedure for obtaining geometric parameters for larger systems, for which multidimensional optimizations with CCSD(T) are prohibitively expensive. We find that although a number of deficiencies of various methods – such as the characteristic overbinding of MP2 – are simply amplified during the transition to larger systems, some are not.

2 Computational Methods

We have examined a wide variety of electronic structure methods with respect to their abilities to reproduce coupled-cluster geometries of non-covalently-interacting molecules. Methods examined include HF,[1,2] MP2,[3,44] attenuated MP2,[11,12] simple-scaled MP2 (sMP2, with same-spin and opposite-spin coefficients set to 0.60 for aug-cc-pVDZ, 0.75 for aug-cc-pVTZ),[11,12,45] spin-component scaled MP2 (SCS-MP2),[46] scaled opposite-spin MP2 (SOS-MP2),[47] B3LYP,[48–51] PBE,[52] M06,[53] M06-L,[54] M06-2X,[53] M11,[55] ωB97X,[56] ωB97X-D,[57] ωB97X-V,[58] B97M-V,[59] vdW-DF2,[25,60,61] VV10,[23,52,60] LC-VV10,[23] and Grimme -D2 and -D3 corrections to PBE and B3LYP.[18,62,63] For DFT-D3, two different damping functions were utilized: the original zero-damping scheme of Grimme,[62] which we refer to simply as DFT-D3, and the damping scheme of Becke and Johnson,[19,64] which we denote as DFT-D3

(BJ).

Throughout this study, all MP2 calculations employ the RI approximation in conjunction with the auxiliary basis sets of Weigend et al.,[65] but the RI prefix has been omitted. All calculations were performed with a development version of Q-Chem 4.2,[66] with the exception of the calculations on the A21x12 dataset reported in Table 1, which were performed with PSI4.[67] Molecular structures were generated with Avogadro.[68] In this work, we have utilized three distinct datasets which can broadly be characterized by the sizes of their constituent systems; due to computational constraints, the manner in which these three classes of systems were treated differs, and will be detailed forthwith.

2.1 Small Systems

The first dataset, henceforth referred to as A21, is comprised of the first 21 systems in the A24 dataset.[69] These systems were optimized by Řezáč and Hobza[69] at the ΔCCSD(T)/CBS level, with a two-point (aug-cc-pVTZ, aug-cc-pVQZ) Helgaker extrapolation of the counterpoise-corrected MP2 correlation energy and a counterpoise-corrected coupled-cluster correction in the aug-cc-pVDZ basis.[10,70–72] The systems contained in this dataset are small enough that high-quality geometries are readily available, and so we have performed unconstrained relaxations using all of the methods detailed above, in order to compare the resultant structures to the benchmark-level structures of Řezáč and Hobza.[69] All geometry optimizations were initialized with the relevant ΔCCSD(T)/CBS structure and a Hartree-Fock Hessian generated in the 6-31+G* basis. Tight convergence criteria were employed: the DIIS error was converged to 10^{-8}, the maximum component of the gradient was converged to 1.5×10^{-4}, the maximum atomic displacement was converged to 6×10^{-5}, the energy change of successive optimization cycles was converged to 10^{-9}, and integral threshholds of 10^{-14} were used. All calculations were performed in the aug-cc-pVTZ basis.[73,74] No symmetry was exploited, though the point group symmetry of each optimized structure matched in every instance the symmetry of the reference structure. Optimiza-

tions were performed in Cartesian coordinates. A fine Lebedev integration grid consisting of 99 radial points and 590 angular points was utilized in the computation of the semilocal exchange-correlation components of all of the density functionals; the coarser SG-1 grid was used for nonlocal correlation in the relevant methods, i.e. those involving vdW-DF2 or VV10 nonlocal correlation.[75]

A variety of metrics were employed to compare the geometries associated with each method with the benchmark ΔCCSD(T)/CBS geometries. One such metric, the overall root-mean-square deviation (RMSD) is given by

$$RMSD = \sqrt{\frac{\sum_{i=j}^{N} d(x_i, y_i, z_i, x_j, y_j, z_j)^2}{N}} \quad (1)$$

where N is the number of atoms in the structure and $d(x_i, y_i, z_i, x_j, y_j, z_j)$ is the Euclidean distance between the points (x_i, y_i, z_i) and (x_j, y_j, z_j). Specifically, we define the overall RMSD as the minimum such number, allowing for rigid transformations of the coordinate systems i and j associated with the reference structure and the optimized structure. The overall RMSD thus encompasses both inter- and intramolecular errors in geometry. We also utilized an intermolecular metric, namely the closest point of contact between the two molecules in each system, and various intramolecular metrics, specifically bond length and bong angle root-mean-square errors.

2.2 Medium Systems

The second dataset, henceforth referred to as M12, consists of a well-balanced subset of the S66x8 dataset of Řezáč et al.,[33] as well as the CO_2-benzene complex.[76] Unfortunately, computational constraints have precluded explicit multi-dimensional optimizations of structures of this size at suitably high levels of theory (i.e. CCSD(T)); as a result, we have utilized cubic interpolation of various single point energies corresponding to rigid displacement along

a single intermolecular coordinate. The particular coordinates and displacements used are defined in the original works of Řezáč et al.[33] and Witte et al.,[76] and are depicted in Figure 1.

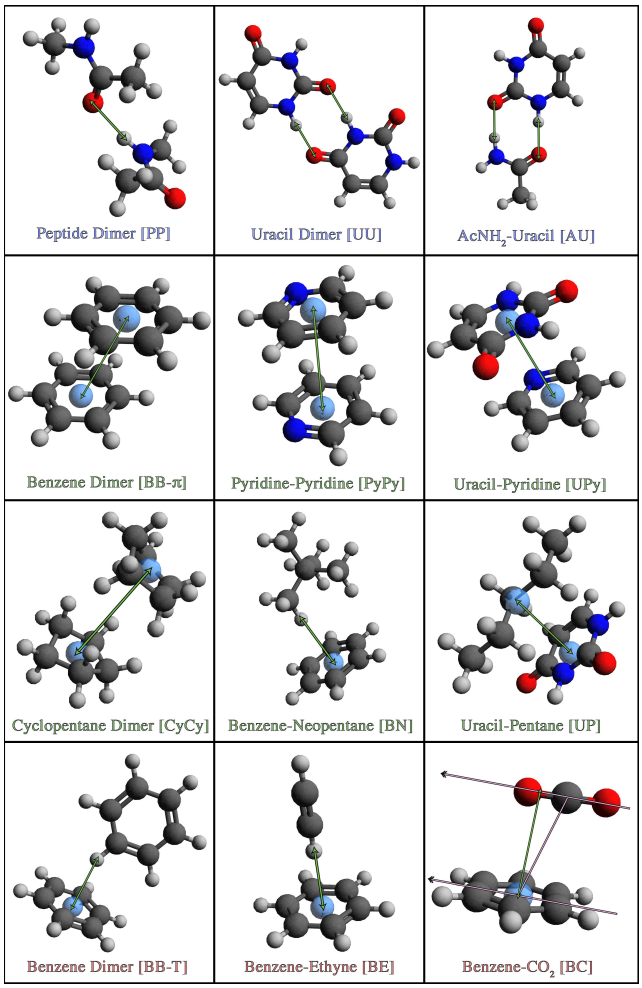


Figure 1: Structures of the systems in the M12 dataset. A light blue sphere corresponds to the center of mass of a particular molecule. The type of interaction is indicated by the color of the text: hydrogen-bonded systems are blue, dispersion-bound systems are green, and systems with mixed interactions are red. Green double-headed arrows indicate the relevant intermolecular axis for each system. In brackets, abbreviations for the complexes are introduced.

We have numerically examined the validity of this sort of approach. A representative summary of our findings is provided in Table 1. A variety of means of interpolating along the potential energy surface have been examined: a cubic spline, a quartic spline, and a least-squares-optimized function consisting of a decaying exponential and a power series in r^{-1},

i.e. a simplification of the model of Tang and Toennies.[77] For the M12 set, the differences in interpolated equilibrium distances obtained with these methods are on the order of 0.001 Å.

Table 1: Justification of methods. We have considered the quality of the reference data with regard to basis set size on both small (top left) and moderately-sized (bottom left) systems, the impact of the specific form of function used for interpolation (top right), and the difference between interpolation and constrained optimization (bottom right). UAD, SAD, and MAX denote respectively the unsigned, signed, and maximum average difference in interpolated equilibrium separation across the indicated dataset, in units of Å. r and BE denote the equilibrium distance (Å) and binding energy (kcal mol^{-1}), respectively.

Benchmark Quality: Basis Size[a]				Interpolation Type[b]			
	UAD	SAD	MAX		UAD	SAD	MAX
MP2: (aQ,a5) vs (aT,aQ)[c]	0.001	0.000	-0.002	Quartic vs Cubic	0.000	0.000	0.001
ΔCC: aT vs aD[d]	0.004	-0.004	-0.009	LSQ vs Cubic	0.001	0.000	-0.004
Benchmark Quality: CO_2-Benzene[e]				**Validity of Interpolation[f]**			
	r	BE			r	BE	
MP2/CBS (aT,aQ) + ΔCC (aD)	3.248	-2.60		Optimized	3.255	-2.65	
MP2/CBS (Q,5) + ΔCC (T)[g]	3.251	-2.66		Interpolated	3.253	-2.64	
Difference	0.003	0.06		Difference	0.001	-0.01	

[a]Results pertain to the A21x12 set.
[b]Results pertain to the M12 set.
[c]Extrapolation to the CBS limit was performed according to the scheme of Helgaker et al.[10,70] aT, aQ, and a5 denote aug-cc-pVTZ, aug-cc-pVQZ, and aug-cc-pV5Z, respectively.
[d]RI-CCSD(T) correction to MP2 correlation energy.[71] aD and aT denote aug-cc-pVDZ and aug-cc-pVTZ, respectively.
[e]Equilibrium distances r and binding energies BE were interpolated with a cubic spline.
[f]Case study on the CO_2-benzene complex using VV10.
[g]T, Q, and 5 denote cc-pVTZ, cc-pVQZ, and cc-pV5Z, respectively.

We have also investigated whether interpolation is a suitable substitute for explicitly performing a constrained optimization along the relevant axis. There appears to be no difference between the two approaches; the results of a case study of VV10 on the CO_2-benzene complex are provided in Table 1. Further proof is provided by a recent study on the parallel-displaced benzene dimer in which both the in-plane shift and interplane spacing were optimized at the CCSD(T)/aug-cc-pVTZ level of theory.[78] The optimized parameters correspond to a center-of-mass separation of 3.84 Å, which compares quite favorably with the value of 3.86 Å obtained by interpolating along the ΔCCSD(T)/CBS binding curve of

the M12 set.

As a final justification of our methodology, we have addressed the quality of our benchmarks. For the A21 and M12 sets, the reference data is ΔCCSD(T)/CBS, with a two-point (aug-cc-pVTZ, aug-cc-pVQZ) extrapolation of the MP2 correlation energy and a correction for higher-order correlation effects in the aug-cc-pVDZ basis.[10,70,71] Specifically, we have investigated the impact of utilizing larger basis sets for each of these components on interpolated equilibrium distances in the A21x12 dataset – a set constructed in a manner analogous to the S66x8 set – wherein we rigidly have scaled the center-of-mass separation of each system in the A21 set by a factor of 0.9 to 2.0, in increments of 0.1. The A21 set, along with the relevant intermolecular axis utilized for the generation of the A21x12 set, is depicted in Figure 2. As is evident from Table 1, the reference data is indeed sufficiently high-quality: using larger bases changes the interpolated equilibrium intermolecular distances by less than 0.01 Å. As a case study of whether such a choice of basis sets is sufficient for larger systems, we have examined the CO_2-benzene complex at two different levels of theory, and found that our reference data is indeed sufficiently converged with respect to basis set size, as regards both equilibrium geometry and binding energy.

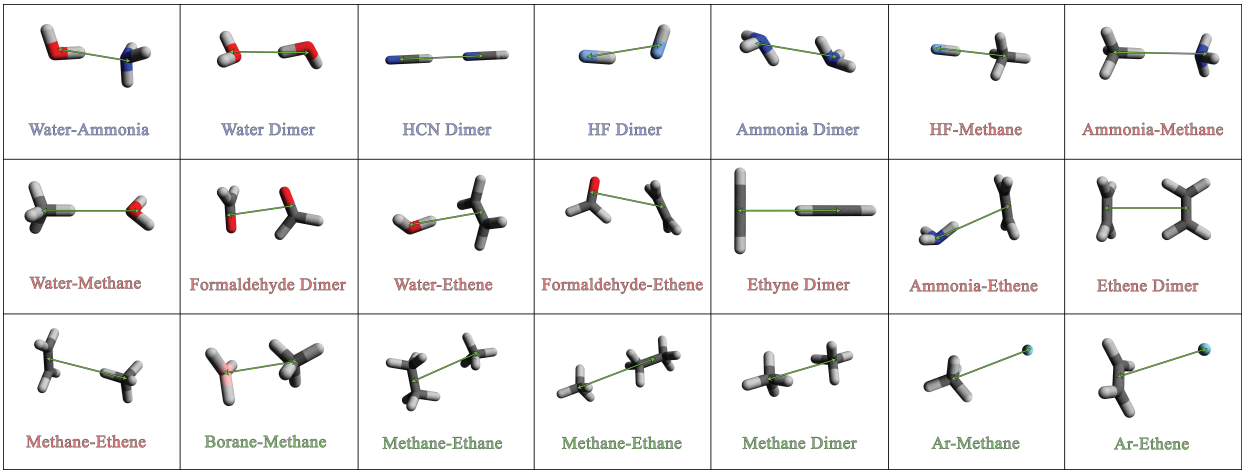


Figure 2: Structures of the systems in the A21x12 dataset. The type of interaction is indicated by the color of the text: hydrogen-bonded systems are blue, dispersion-bound systems are green, and systems with mixed interactions are red. Green double-headed arrows connect the centers of masses of the two molecules in each system, and hence indicate the relevant intermolecular axis for each system.

As in the case of the A21 set, our calculations on the M12 set utilize a (99,590) Lebedev integration grid for semilocal components of density functionals, with the SG-1 grid being used for nonlocal correlation. Furthermore, integral threshholds of 10^{-14} were used, the DIIS error was converged to 10^{-8}, and no symmetry was exploited. All calculations were performed in the aug-cc-pVTZ basis. For the MP2 calculations, the frozen core approximation was employed.[79] MP2 results were corrected for basis set superposition error (BSSE) a la Boys and Bernardi.[72]

2.3 Large System

In an attempt to probe the transferability of observed trends to even larger systems, we have treated the coronene dimer in an analogous manner to how the M12 set was treated. The reference geometry for the coronene dimer, determined at the QCISD(T)/h-aug-cc-pVDZ level by Janowski et al.,[80] is depicted in Figure 3. This geometry corresponds to an interplane spacing of 3.458 Å, with an in-plane shift of 1.553 Å. We constructed a potential energy surface for each method with a series of seven of single-point energy calculations; in so doing, we sampled interplane separations from 3.008 Å to 3.908 Å, in increments of 0.158 Å, holding the in-plane shift and all intramolecular parameters constant. To aid convergence, densities determined at the LDA level were used as a starting point for all jobs, and the criterion for determining wavefunction convergence was lowered to 10^{-6}. All calculations were performed in the aug-cc-pVDZ basis, and all methods were corrected for BSSE in the usual manner.[72] A (99,590) Lebedev grid was used for the evaluation of semilocal components of density functionals, the SG-1 grid was used for the evaluation of nonlocal correlation, integral threshholds were set to 10^{-14}, and no symmetry was exploited.

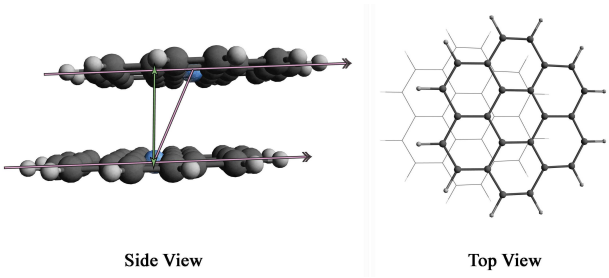


Figure 3: Structure of the coronene dimer. Light blue spheres correspond to the centers of masses of the coronene monomers. A green double-headed arrow indicates the relevant intermolecular axis.

3 Results and Discussion

3.1 Small Systems

A summary of results pertaining to the A21 dataset is provided in Table 2 and Figure 4. The optimized structures themselves can be found in the Supporting Information. It is evident from Table 2 that of the methods examined, ωB97X-V is the top performer, treating the various classes of interactions equally well. Moreover, the similarity between the overall root-mean-square deviation (RMSD) and overall weighted root-mean-square deviation (wRMSD) indicates that ωB97X-V performs equally well for both weak and strong interactions, as opposed to a method such as B3LYP, which performs disproportionately well on stronger interactions, i.e. hydrogen bonds. Within the A21 dataset, the various modifications of MP2 perform reasonably well for geometry optimizations, with attenuated MP2 (attMP2) slightly outperforming other versions of MP2. MP2 systematically underestimates intermolecular distances, as evidenced in Figure 4, which can be attributed somewhat to BSSE, (see Supporting Information), but primarily to the overbinding endemic to the method. This overbinding and concomitant underestimation is alleviated somewhat by the attenuation of the Coulomb operator; some systems are actually underbound by attMP2, which suggests the addition of a long-range dispersion correction could be profitable. This has been explored by Huang et al., who found that attenuation of dispersion-corrected MP2 can indeed improve

the description of intermolecular interactions, though at the possible cost of a poorer picture of intramolecular interactions.[81] Simple-scaling of the MP2 correlation energy, on the other hand, leads to systematic overestimation of intermolecular separation. This is an artifact of the fact that the scaling coefficent was optimized with respect to errors in interaction energies in the S66 dataset, and is hence too small for the purposes of this dataset. Scaling of individual components of the MP2 correlation energy (SCS-MP2 and SOS-MP2) is similarly underwhelming here.

Table 2: Average overall root-mean-square deviations (RMSD) and weighted root-mean-square deviations (wRMSD) in geometries of complexes in various subsets of the A21 Dataset. Weights were determined from the relative binding energies of each complex. Within each subset, the methods are listed in order of ascending RMSD or wRMSD. All calculations were performed in the aug-cc-pVTZ basis with tight convergence criteria. Calculations involving density functionals utilized a (99,590) grid.

Hydrogen-Bonded		Mixed Interactions		Dispersion-Bound		All			
Method	RMSD (Å)	Method	RMSD (Å)	Method	RMSD (Å)	Method	RMSD (Å)	Method	wRMSD (Å)
MP2	0.010	ωB97X-V	0.016	attMP2	0.009	ωB97X-V	0.014	ωB97X-V	0.014
ωB97X-V	0.011	LC-VV10	0.024	MP2	0.012	LC-VV10	0.028	attMP2	0.022
attMP2	0.012	vdW-DF2	0.035	ωB97X-V	0.013	attMP2	0.035	MP2	0.022
SCS-MP2	0.013	ωB97X-D	0.036	B97M-V	0.020	vdW-DF2	0.036	LC-VV10	0.027
sMP2	0.013	sMP2	0.039	B3LYP-D3	0.021	B97M-V	0.037	B97M-V	0.027
B3LYP	0.014	M11	0.050	B3LYP-D3 (BJ)	0.024	ωB97X-D	0.037	sMP2	0.029
B97M-V	0.017	SOS-MP2	0.053	VV10	0.042	MP2	0.038	SCS-MP2	0.029
M11	0.018	B97M-V	0.057	vdW-DF2	0.042	sMP2	0.042	ωB97X-D	0.031
M06-2X	0.019	attMP2	0.062	LC-VV10	0.043	M11	0.045	B3LYP-D3	0.032
B3LYP-D3	0.020	SCS-MP2	0.065	SCS-MP2	0.048	SCS-MP2	0.048	B3LYP-D3 (BJ)	0.034
B3LYP-D3 (BJ)	0.020	PBE	0.065	ωB97X-D	0.050	B3LYP-D3	0.053	vdW-DF2	0.037
LC-VV10	0.021	MP2	0.068	M06-L	0.054	B3LYP-D3 (BJ)	0.054	M11	0.038
SOS-MP2	0.023	M06	0.071	ωB97X	0.057	SOS-MP2	0.055	ωB97X	0.041
ωB97X-D	0.025	B3LYP-D3	0.089	M06	0.059	M06	0.058	SOS-MP2	0.042
PBE	0.027	B3LYP-D3 (BJ)	0.089	M11	0.061	VV10	0.067	M06	0.045
M06	0.029	M06-L	0.092	PBE-D3	0.066	M06-L	0.067	M06-2X	0.049
ωB97X	0.031	PBE-D3 (BJ)	0.093	PBE-D3 (BJ)	0.067	ωB97X	0.069	M06-L	0.051
PBE-D3	0.031	ωB97X	0.096	B3LYP-D	0.070	PBE-D3 (BJ)	0.072	B3LYP-D	0.052
vdW-DF2	0.032	VV10	0.097	M06-2X	0.070	PBE-D3	0.074	PBE-D3	0.056
M06-L	0.033	PBE-D3	0.101	sMP2	0.072	M06-2X	0.084	VV10	0.060
PBE-D3 (BJ)	0.035	PBE-D	0.126	PBE-D	0.083	B3LYP-D	0.090	PBE-D3 (BJ)	0.061
B3LYP-D	0.035	M06-2X	0.126	SOS-MP2	0.084	PBE-D	0.095	PBE	0.062
VV10	0.038	B3LYP-D	0.130	PBE	0.205	PBE	0.096	PBE-D	0.074
PBE-D	0.047	B3LYP	0.152	HF	0.639	HF	0.330	B3LYP	0.138
HF	0.079	HF	0.270	B3LYP	1.128	B3LYP	0.398	HF	0.184

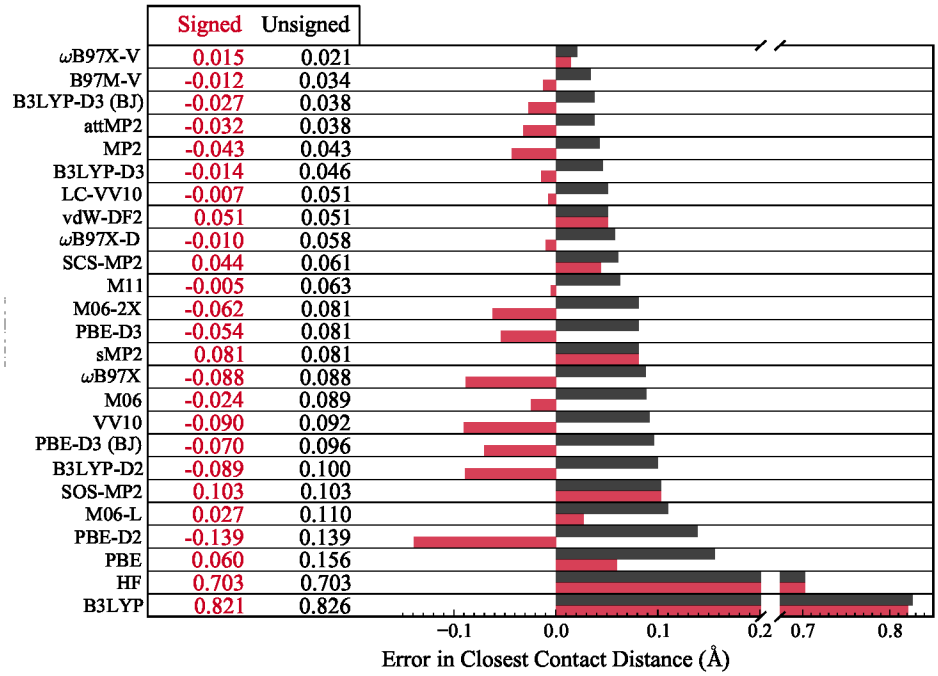


Figure 4: Average signed and unsigned errors in closest-contact distances in geometries of complexes from the A21 Dataset. The methods are listed in order of ascending unsigned error. All calculations were performed in the aug-cc-pVTZ basis with tight convergence criteria. Calculations involving density functionals utilized a (99,590) grid.

Most of the density functionals examined provide rather mediocre geometries for this dataset; the descriptions of mixed and dispersion-dominated interactions in particular leave much to be desired. There do appear to be trends in the performances of the various types of functionals, however. Functionals incorporating some amount of exact exchange tend to outperform those with approximate exchange kernels, particularly on hydrogen-bonded systems. Going a step further, long-range corrected hybrids appear to offer a superior description to global hybrids or functionals with no exact exchange. The best functionals incorporate range-separated exchange in conjunction with some sort of long-range dispersion correction. The nature of the dispersion tail is particularly important: the standard DFT-D2 treatment of Grimme yields complexes in which the closest-contact distance is vastly underestimated; in fact, PBE-D2 actually performs worse than PBE, a consequence of the fact that PBE exchange alone often overbinds in the context of non-covalent interactions.[82] The

more repulsive DFT-D3 and DFT-D3 (BJ) variants are thus better suited for these smaller systems: B3LYP-D3 and PBE-D3 exhibit less severe underestimation of intermolecular separation than their -D2 counterparts, regardless of whether zero-damping or BJ-damping is employed, as evidenced in Figure 4. The differences between -D2 and -D3 can be attributed to both differences in their respective C_6 coefficients as well as differences in the damping functions of the two methods. Among the -D3 variants studied, both the zero-damping and BJ-damping schemes yield similar results. A variant employing Wu-Yang damping[83] – the function of choice in the original -D2 prescription – in conjunction with -D3 C_6 coefficients is given in the Supporting Information, and suggests the difference between -D2 and -D3 is in this case primarily due to the starkly different Fermi-type damping function.

The performances of the various nonlocal functionals are somewhat less intuitive. The standard exchange-matching issues of vdW-DF2 are manifest in the systematically too-large closest-contact distance, as illustrated in Figure 4. The behaviors of the various functionals with a VV10 tail, however, are less predictable: variants with range-separated exchange, namely ωB97X-V and LC-VV10, yield structures very similar to the ΔCCSD(T) benchmarks, whereas VV10 in its original iteration is somewhat lackluster in its description of the A21 set. This cannot be understood simply to be a result of a deficiency in any single component of the VV10 functional. As is apparent in Figure 4, VV10 underestimates intermolecular separations, whereas LC-VV10 on average overestimates them, despite the fact that the rPW86 exchange incorporated in VV10 is generally more repulsive than the (short-range) PBE exchange of LC-VV10. There is clearly a subtle interplay between the exchange components and the nonlocal tail at work. The fact that the parameters of the nonlocal tail were optimized on different datasets further complicates this issue.

One thing that bears mentioning here is that although we have distilled the comparison of geometries into two numbers – the overall RMSD and the error in closest-contact distance – these two metrics alone only tell part of the story. The overall RMSD encompasses all of the discrepancies between a structure associated with a particular method and the reference

structure; what is lacking, however, is a breakdown of whence these disparities arise. The error in closest-contact distance, on the other hand, is a much more focused metric; it is primarily a measure of intermolecular separation, though it can be obfuscated by symmetry-preserving rotations, provided such transformations exist. A variety of other possible metrics exist. For instance, two simple intramolecular metrics might be bond length or bond angle RMS errors; these have been tabulated for the methods examined in the A21 dataset, and can be found in the Supporting Information. For the A21 set and the methods examined, these errors are an order of magnitude smaller than the overall RMSD; moreover, the distributions associated with these errors are much narrower than the distribution associated with either closest-contact error or overall RMSD, and hence the utility of such metrics is limited. Perhaps the most interesting story told by this supplementary data is the lack of a difference between the base functionals and those incorporating a -D2 or -D3 dispersion correction: addition of a simple empirical dispersion correction does not significantly impact intramolecular parameters in small molecules. This may not be the case, however, for large or extended molecules, particularly those involving significant non-covalent intramolecular interactions.

3.2 Medium Systems

It is evident from Table 2 and Figure 4 that the primary source of overall deviation in geometries in the chosen methods in the A21 set is the intermolecular separation. Thus, in systems where performing unconstrained optimizations are infeasible, it is possible to probe geometric differences solely on the basis of a single intermolecular coordinate. Moreover, we established earlier (see Table 1) that interpolation based on a series of intelligently-chosen points is, to a small degree of uncertainty (on the order of 0.01 Å), equivalent to explicitly performing a constrained optimization along the relevant coordinate. Examination of Figure 5 further supports this notion; with few exceptions, the methods investigated in this study yield well-behaved binding energy curves, and so it is no surprise that any sort of reasonable

choice of interpolation scheme would identify the same minimum. Such is the basis for our treatment of the M12 set of medium-sized molecular systems.

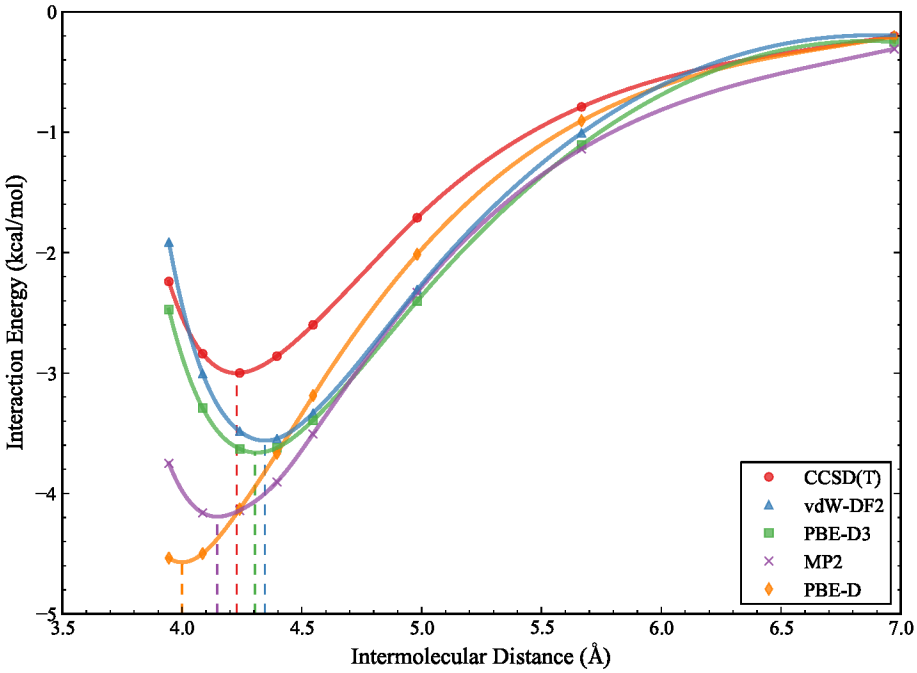


Figure 5: Selection of binding energy curves for the cyclopentane dimer (CyCy in the M12 dataset). Equilibrium separations associated with each method are denoted by vertical dashed lines, and were determined via interpolation with a cubic spline. All calculations were performed in the aug-cc-pVTZ basis. Calculations involving density functionals utilized a (99,590) grid.

A summary of results pertaining to the M12 dataset are provided in Figures 6 and 7. In light of the small uncertainty associated with the interpolation, any difference in intermolecular separation larger than 0.03 Å can be safely deemed significant. For certain cases where the minimum associated with a particular method lies outside of the range explored, the second largest (or smallest, as appropriate) separation examined is reported, thereby providing a lower bound for the error. As a consequence of this and the general problem of interpolating a minimum from a flat surface, those errors associated with particularly underbinding methods, namely PBE, HF, and B3LYP on systems involving dispersive interactions, can only be interpreted in a qualitative sense.

Hydrogen-Bonded			Dispersion (π-π)			Dispersion (Other)			Mixed			All			
Method	ASE	AUE	Method	ASE	AUE	Method	ASE	AUE	Method	ASE	AUE	Method	ASE	AUE	MAX
ωB97X	0.00	0.00	sMP2	-0.01	0.02	B3LYP-D3 (BJ)	0.01	0.01	B3LYP-D3 (BJ)	0.01	0.01	B3LYP-D3 (BJ)	0.00	0.02	0.03
M06-2X	0.00	0.00	SOS-MP2	0.02	0.02	B97M-V	-0.01	0.01	VV10	0.01	0.01	B97M-V	0.01	0.02	0.04
MP2 (CP)	0.00	0.00	LC-VV10	-0.02	0.02	B3LYP-D3	-0.02	0.02	B97M-V	0.02	0.02	VV10	0.00	0.02	-0.05
ωB97X-V	0.01	0.01	ωB97X-D	0.01	0.02	VV10	-0.01	0.02	sMP2	0.00	0.02	B3LYP-D3	0.01	0.03	0.05
M06-L	-0.01	0.01	B3LYP-D3 (BJ)	0.00	0.02	ωB97X	0.01	0.02	ωB97X	0.01	0.02	LC-VV10	-0.03	0.03	-0.06
M11	0.01	0.01	M06-L	-0.03	0.03	LC-VV10	-0.03	0.03	B3LYP-D3	0.03	0.03	sMP2	0.02	0.03	0.13
B97M-V	0.01	0.01	B97M-V	0.03	0.03	MP2 (CP)	-0.03	0.03	SOS-MP2	0.03	0.03	ωB97X-D	-0.01	0.03	-0.11
M06	0.01	0.01	VV10	0.03	0.03	SCS-MP2	0.02	0.05	ωB97X-D	0.03	0.03	ωB97X	0.03	0.03	0.10
B3LYP-D3	-0.01	0.01	M06	0.04	0.04	ωB97X-V	0.05	0.05	SCS-MP2	-0.04	0.04	M06	0.01	0.03	-0.10
sMP2	0.01	0.01	PBE-D2	-0.04	0.04	M06	-0.05	0.05	LC-VV10	-0.04	0.04	M06-L	-0.03	0.03	-0.08
ωB97X-D	-0.01	0.01	B3LYP-D3	0.05	0.05	attMP2	-0.06	0.06	M06-L	0.00	0.04	ωB97X-V	0.04	0.04	0.08
PBE-D3	-0.01	0.01	SCS-MP2	-0.07	0.07	PBE-D3 (BJ)	0.06	0.06	ωB97X-V	0.04	0.04	SCS-MP2	-0.01	0.04	0.10
VV10	-0.01	0.01	PBE-D3 (BJ)	0.07	0.07	sMP2	0.07	0.07	M06	0.04	0.04	SOS-MP2	0.04	0.04	0.16
attMP2	-0.02	0.02	B3LYP-D2	-0.07	0.07	M06-L	-0.07	0.07	PBE-D3 (BJ)	0.05	0.05	PBE-D3 (BJ)	0.04	0.05	0.10
B3LYP-D3 (BJ)	-0.02	0.02	ωB97X-V	0.07	0.07	ωB97X-D	-0.07	0.07	MP2 (CP)	-0.06	0.06	MP2 (CP)	-0.06	0.06	-0.17
SCS-MP2	0.02	0.02	M11	-0.07	0.07	PBE-D3	0.09	0.09	PBE-D2	-0.08	0.08	M11	-0.06	0.07	-0.13
PBE-D3 (BJ)	-0.02	0.02	ωB97X	0.09	0.09	SOS-MP2	0.09	0.09	M11	-0.09	0.09	attMP2	-0.07	0.07	-0.14
PBE	0.01	0.02	M06-2X	-0.09	0.09	M11	-0.11	0.11	PBE-D3	0.09	0.09	M06-2X	-0.08	0.08	-0.13
LC-VV10	-0.02	0.02	attMP2	-0.12	0.12	M06-2X	-0.11	0.11	M06-2X	-0.09	0.09	PBE-D2	-0.08	0.08	-0.23
MP2	-0.03	0.03	PBE-D3	0.12	0.12	MP2	-0.12	0.12	attMP2	-0.10	0.10	PBE-D3	0.07	0.08	0.13
B3LYP-D2	-0.03	0.03	MP2 (CP)	-0.14	0.14	vdW-DF2	0.12	0.12	B3LYP-D2	-0.10	0.10	B3LYP-D2	-0.10	0.10	-0.28
PBE-D2	-0.04	0.04	vdW-DF2	0.14	0.14	PBE-D2	-0.16	0.16	vdW-DF2	0.14	0.14	vdW-DF2	0.12	0.12	0.16
B3LYP	0.04	0.04	MP2	-0.20	0.20	B3LYP-D2	-0.21	0.21	MP2	-0.14	0.14	MP2	-0.12	0.12	-0.25
SOS-MP2	0.04	0.04	PBE	0.63	0.63	PBE	0.51	0.51	PBE	0.34	0.34	PBE	0.38	0.38	0.86
vdW-DF2	0.07	0.07	HF	1.30	1.30	HF	1.33	1.33	B3LYP	0.62	0.62	HF	0.89	0.89	1.78
HF	0.14	0.14	B3LYP	1.35	1.35	B3LYP	1.57	1.57	HF	0.77	0.77	B3LYP	0.89	0.89	1.78

Figure 6: Average signed (ASE) and unsigned (AUE) errors (in units of Å) in interpolated equilibrium intermolecular separations in geometries of complexes in various subsets of the M12 dataset. Maximum error for each method (MAX) is given in last column (in units of Å). Within each subset, methods are sorted in order of ascending AUE. Signed values are colored as such: positive errors are blue, negative errors are red, and the tint of the color correlates with the magnitude of the error. Equilibria correspond to the interpolated (cubic spline) minima of the binding energy curves. All calculations were performed in the aug-cc-pVTZ basis. Calculations involving density functionals utilized a (99,590) grid.

Due to the uncertainty associated with each interpolated equilibrium separation, it is difficult to draw the same sorts of conclusions for the M12 set as we did for the A21 set. Nevertheless, a few trends are apparent. Among the MP2 methods, standard MP2 is the worst performer, underestimating the intermolecular separation in every system and providing geometries worse than any of the standard density functional approaches (with the exception of PBE and B3LYP). The treatment of systems involving $\pi - \pi$ interactions is particularly bad – this is a manifestation of BSSE and the general unsuitability of uncoupled HF polarizabilities (and hence C_6 coefficients) for describing such systems.[84] In a larger basis, where BSSE is reduced, this underestimation of intermolecular distance is somewhat less drastic, though still substantial, as evidenced by the counterpoise-corrected (CP) MP2 results in Figures 6 and 7. Attenuation of the Coulomb operator (attMP2) addresses the underestimation to a

similar extent as CP-correction, though at a fraction of the cost. Simple-scaling of the MP2 correlation energy (sMP2) yields better agreement with the coupled-cluster benchmarks; this can be at least in part attributed to the optimization of the scaling coefficient on the S66 dataset, a set which contains eleven of the twelve systems in M12. However, the dangers of using a single scaling coefficient for a variety of systems are hinted at by the large error for the method on the cyclopentane dimer, as well as the lackluster performance of sMP2 with regard to the A21 set. Separate scaling of the different components of the MP2 correlation energy (SOS-MP2 and SCS-MP2) is generally inferior to simple-scaling.

	Hydrogen-Bonded			Dispersion (π-π)			Dispersion (Other)			Mixed			All
Method	PP	UU	AU	BB-π	PyPy	UPy	CyCy	BN	UP	BB-T	BE	BC	AUE
B3LYP-D3 (BJ)	-0.02	-0.02	-0.02	-0.03	-0.01	0.03	-0.01	0.00	0.03	0.01	0.00	0.02	0.02
B97M-V	0.01	0.01	0.01	0.03	0.04	0.02	-0.02	0.00	-0.02	0.03	0.01	0.01	0.02
VV10	-0.02	-0.01	-0.01	0.03	0.04	0.04	-0.05	0.00	0.01	0.02	0.00	0.01	0.02
B3LYP-D3	-0.01	-0.01	-0.01	0.04	0.04	0.05	0.00	-0.03	-0.02	0.01	0.02	0.04	0.03
LC-VV10	-0.02	-0.03	-0.03	-0.01	-0.02	-0.03	-0.01	-0.04	-0.03	-0.03	-0.06	-0.02	0.03
sMP2	0.02	0.01	0.01	-0.02	-0.01	0.01	0.13	0.03	0.05	-0.01	-0.02	0.04	0.03
ωB97X-D	-0.02	-0.01	-0.01	-0.02	0.00	0.04	-0.11	-0.06	-0.03	0.00	0.02	0.07	0.03
ωB97X	0.00	0.00	0.00	0.10	0.09	0.07	-0.02	0.03	0.02	0.04	-0.01	0.02	0.03
M06	0.00	0.01	0.01	0.03	0.05	0.04	-0.10	-0.02	-0.03	0.04	0.05	0.04	0.03
M06-L	0.00	-0.01	0.00	-0.02	-0.03	-0.03	-0.06	-0.08	-0.07	0.00	-0.06	0.05	0.03
ωB97X-V	0.01	0.00	0.00	0.07	0.08	0.06	0.05	0.04	0.05	0.05	0.02	0.05	0.04
SCS-MP2	0.02	0.02	0.02	-0.09	-0.08	-0.03	0.10	-0.04	0.01	-0.06	-0.05	0.00	0.04
SOS-MP2	0.05	0.04	0.04	0.00	0.01	0.04	0.16	0.04	0.08	0.01	0.01	0.07	0.04
PBE-D3 (BJ)	-0.01	-0.03	-0.03	0.05	0.06	0.09	0.05	0.04	0.10	0.05	0.02	0.09	0.05
MP2 (CP)	0.01	0.00	0.00	-0.17	-0.16	-0.08	0.00	-0.06	-0.02	-0.09	-0.06	-0.05	0.06
M11	0.00	0.01	0.01	-0.06	-0.07	-0.08	-0.11	-0.08	-0.13	-0.06	-0.10	-0.10	0.07
attMP2	-0.02	-0.02	-0.02	-0.14	-0.13	-0.09	-0.01	-0.09	-0.07	-0.10	-0.11	-0.07	0.07
M06-2X	-0.01	0.00	-0.01	-0.10	-0.09	-0.09	-0.12	-0.09	-0.13	-0.07	-0.08	-0.13	0.08
PBE-D2	-0.04	-0.03	-0.03	-0.04	-0.05	-0.04	-0.23	-0.13	-0.12	-0.10	-0.11	-0.02	0.08
PBE-D3	0.00	-0.02	-0.02	0.12	0.12	0.13	0.08	0.06	0.12	0.09	0.07	0.12	0.08
B3LYP-D2	-0.05	-0.02	-0.02	-0.07	-0.08	-0.06	-0.28	-0.17	-0.18	-0.13	-0.12	-0.05	0.10
vdW-DF2	0.07	0.07	0.07	0.14	0.13	0.16	0.12	0.12	0.13	0.16	0.14	0.11	0.12
MP2	-0.03	-0.02	-0.02	-0.25	-0.22	-0.14	-0.08	-0.16	-0.11	-0.18	-0.14	-0.11	0.12
PBE	0.06	-0.01	-0.01	0.86	0.59	0.45	0.48	0.51	0.54	0.43	0.21	0.38	0.38
HF	0.22	0.10	0.09	1.78	1.37	0.75	1.44	1.26	1.30	1.00	0.51	0.81	0.89
B3LYP	0.09	0.01	0.01	1.78	1.67	0.62	1.44	1.58	1.68	0.85	0.37	0.64	0.89

Figure 7: Errors in interpolated equilibrium intermolecular separations (in units of Å) in geometries of complexes of the M12 dataset. The last column, the average unsigned error (AUE) across all systems (in units of Å), represents the metric by which the methods are sorted. Signed values are colored as such: positive errors are blue, negative errors are red, and the tint of the color correlates with the magnitude of the error. Equilibria correspond to the interpolated (cubic spline) minima of the binding energy curves. All calculations were performed in the aug-cc-pVTZ basis. Calculations involving density functionals utilized a (99,590) grid. For the abbreviations, see Figure 1.

Among the density functionals examined, most of the top performers incorporate some form of long-range dispersion correction. Specifically, those functionals with VV10 nonlocal correlation reproduce ΔCCSD(T)/CBS equilibrium separations well: B97M-V in particular is consistently quite accurate. Among the functionals lacking any form of long-range van der Waals correction, ωB97X, M06, and M06-L stand out. Their impressive performances can be at least somewhat attributed to the significant overlap between the systems examined

here and the systems in their respective training sets. Interestingly enough, these methods actually perform better on these larger systems than on the small systems in the A21 set (cf. Figures 4 and 6). Thus, in going from small systems (A21) to medium systems (M12), we do not see an unambiguous amplification of deficiencies. The systematic underestimation of intermolecular separation associated with MP2 and its attenuated variant, the overestimation of vdW-DF2 and ωB97X-V, and the overestimation of the DFT-D3 methods relative to their DFT-D2 counterparts are significantly magnified by the growth in system size, but the relative performances of certain other methods reverse completely. For instance, VV10 offers an excellent description of every system in the M12 set despite being one of the worst methods with respect to the A21 set.

There is one more interesting point to be made that is illustrated clearly by Figure 5: overbinding is not synonymous with underestimation of intermolecular separation, i.e. at least for some methods, horizontal and vertical motion of the binding energy curve associated with a given system are often decoupled. Throughout this article, we have made a point of maintaining this distinction by referring to methods as "overestimating intermolecular separations," rather than by simply calling them "underbinding." In general, we might expect that a method that yields a too-long intermolecular separation would be underbinding, but it is clear from Figure 5 that this is emphatically not the case. For instance, in the case of the cyclopentane dimer, PBE-D3 and vdW-DF2 both overestimate equilibrium intermolecular separations despite being overbinding with respect to energy. In a similar manner, M06 underbinds every system in M12 with respect to energy yet predicts a compressed geometry for half of the dispersion-bound systems. This highlights a serious deficiency in the standard approach of comparing methods solely on the basis of binding energies. This shortcoming is further amplified by the practice of comparing energies calculated with different methods on the same geometry: the binding energies of the cyclopentane dimer as predicted by PBE-D3 and MP2 at the ΔCCSD(T) minimum are indistinguishable, though at their respective minima they differ by nearly 0.2 kcal/mol.

3.3 Large System

Interpolated equilibrium interplane separations and binding energies predicted by each of the methods for the parallel-displaced coronene dimer are provided in Figure 8. Note that the values of 3.91 Å and 0 kcal mol^{-1} reported for PBE, B3LYP, and HF simply indicate that these methods were all repulsive at the maximum separation examined, 3.908 Å. Errors are expressed relative to current best guesses of 3.458 Å and 23.45 kcal mol^{-1}.[59,80] It is worth mentioning, however, that these reference values are not nearly as ironclad as those used for the A21 and M12 sets: the interplane separation corresponds to QCISD(T)/h-aug-cc-pVDZ, i.e. cc-pVDZ on hydrogens and alternating carbons, and aug-cc-pVDZ on other carbons,[80] and the binding energy is given by a counterpoise-corrected MP2/CBS (aTZ,aQZ Helgaker extrapolation)[10,70] energy corrected for higher-order correlation in this same small basis.[59] Other notable reports for binding energy include -19.98 kcal mol^{-1} and -24.36 kcal mol^{-1}, values which were obtained using different prescriptions for incorporating an MP2 energy correction and a different small basis for the ΔQCISD(T) correction.[80,85] As a result of this general uncertainty in the "true" equilibrium interplane separation and binding energy, any discussion of our results for the coronene dimer can only be semi-quantitative.

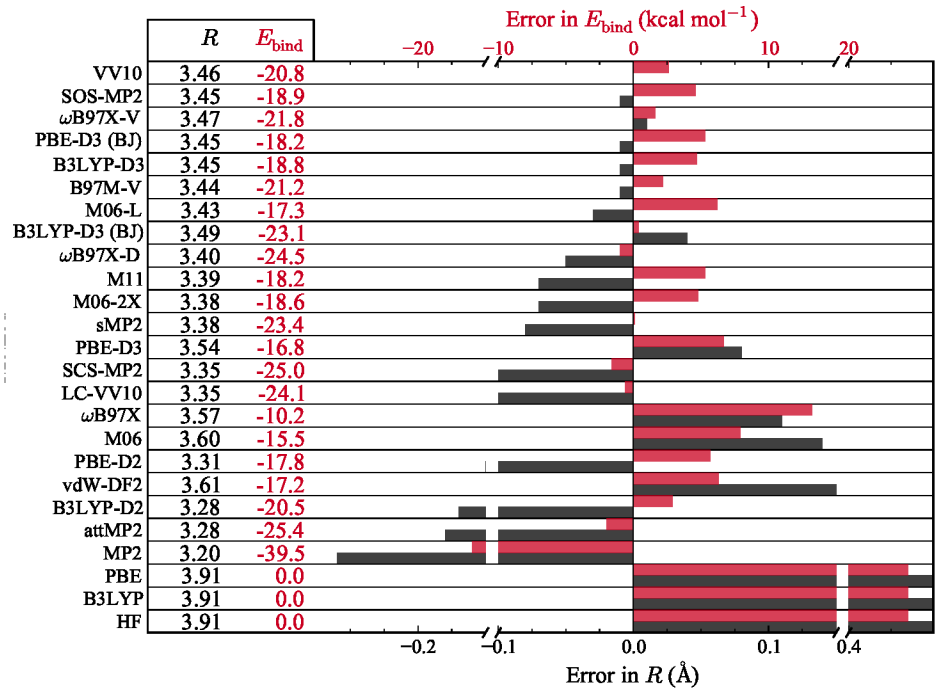


Figure 8: Errors in interpolated equilibrium intermolecular separations (in units of Å) and binding energies (in units of kcal mol^{-1}) for the parallel-displaced coronene dimer. The methods are listed in order of ascending error in intermolecular separation. Equilibria correspond to the interpolated (cubic spline) minima of the binding energy curves. All calculations were performed in the aug-cc-pVDZ basis. Calculations involving density functionals utilized a (99,590) grid. With the exception of attenuated MP2, all methods incorporate a correction for BSSE.[72] Errors in separation are relative to $R = 3.458$ Å reported by Janowski et al.,[80] corresponding to QCISD(T)/h-aug-cc-pVDZ, and errors in binding energy are relative to $E_{\text{bind}} = -23.45$ kcal mol^{-1}, which corresponds to QCISD(T)/CBS as reported by Mardirossian and Head-Gordon.[59]

It is apparent from Figure 8 that some of the general trends observed on the M12 set transfer reasonably well to the case of the coronene dimer: for instance, most methods involving VV10 nonlocal correlation perform well; DFT-D2 underestimates intermolecular separation relative to both variants of DFT-D3; some form of correction to MP2 is important, standard meta-GGA functionals perform surprisingly well; etc. This does not seem surprising, since the coronene dimer is often thought, to a first approximation, to be largely just a bigger version of the benzene dimer. This picture is very limited, however: comparison of our data for the coronene dimer with those for the parallel-displaced benzene-benzene dimer demonstrate that there exists only a very weak correlation between percent errors in

geometries of the benzene dimer and those in the coronene dimer; this correlation is weaker still – if not entirely absent – when comparing errors in binding energies between the two systems. Thus, we advocate the use of care when extrapolating to larger systems. Moreover, as was previously illustrated in Figure 5 for the cyclopentane dimer, neither the signs nor relative magnitudes of the errors in geometry and energy for the coronene dimer are in any way correlated. This point is illustrated still further by Table 3, in which the Pearson's correlation coefficient for percent errors in interpolated equilibrium energy and intermolecular separation across the M12 set is listed for each method.

Table 3: Pearson's correlation coefficient, R, for percent errors in interpolated equilibrium energy and intermolecular separation for the M12 dataset. Methods are divided into two sets – wavefunction-based (WFT) and density functional theory (DFT) – and listed in order of descending correlation coefficient within each set. For details about the calculations, see Figure 7.

WFT		DFT	
Method	R	Method	R
SOS-MP2	0.95	PBE	0.96
MP2 (CP)	0.94	B3LYP	0.86
HF	0.93	PBE-D	0.84
SCS-MP2	0.93	M06	0.84
attMP2	0.88	ωB97X	0.83
MP2	0.88	ωB97X-D	0.82
sMP2	0.87	LCVV10	0.74
		B97M-V	0.71
		B3LYP-D3	0.63
		vdW-DF2	0.61
		PBE-D3 (BJ)	0.40
		M06-L	0.39
		PBE-D3	0.22
		M11	0.22
		M06-2X	0.19
		VV10	0.17
		B3LYP-D3 (BJ)	0.13
		B3LYP-D	-0.28
		ωB97X-V	-0.29

It is clear that equilibrium binding energies and intermolecular separations are only weakly correlated for a number of methods; for some approaches, they are even somewhat anticorrelated, i.e. the methods overbind while overestimating intermolecular separation. It is also notable that wavefunction-based approaches, on average, seem to exhibit a stronger correlation between equilibrium binding energy and intermolecular separation than density

functionals. This being said, this sort of orthogonality between energy and separation observed for a number of methods is not a flaw; rather, it is simply an interesting phenomenon that highlights yet again the fact that merely comparing energies is an insufficient means of assessing the performance of a given method for intermolecular interactions.

4 Conclusion

In this work, we have systematically assessed the abilities of a variety of electronic structure approximations to replicate coupled-cluster-level geometries of non-covalent complexes. Methods examined include HF, MP2, and several common DFT exchange-correlation functionals with and without various dispersion corrections. A variety of systems were studied: the A21 set of small (2-4 heavy atoms) systems, the M12 set of moderately-sized (8-14 heavy atoms) systems, and the parallel-displaced coronene dimer (48 heavy atoms). For the A21 set, ΔCCSD(T)/CBS geometries are readily available for comparison.[69] However, for the larger systems, multidimensional optimizations at such a level of theory are prohibitively expensive. Thus, we have established the validity of a protocol for utilizing binding energy curves along a single intermolecular coordinate to probe the performance of a given method with regard to geometries: interpolation with a cubic spline yields a minimum consistent with explicit constrained optimization, even with a relatively large distance between sampled geometries. Although the overall root-mean-square deviation is the most comprehensive metric for differentiating among methods, this sort of measure of error in intermolecular separation is a reasonable substitute.

We find that the relative performances of the various electronic structure methods for reproducing CCSD(T) geometries is dependent on not only the predominant interaction type, but also the size of the molecular system. Nevertheless, a number of general trends that transcend system size are evident. Those methods incorporating the VV10 brand of nonlocal correlation tend to yield quite accurate geometries; the recently-developed func-

tionals ωB97X-V and B97M-V in particular are remarkably consistent, although the former tends to overestimate equilibrium separations, especially in the context of small systems. The vdW-DF2 method, on the other hand, systematically predicts too-large intermolecular separations, with the associated error dramatically increasing during the transition from small systems to moderately-sized systems. Conventional GGA functionals (e.g. PBE) yield wildly inaccurate geometries for systems in which dispersion interactions are dominant. The addition of some form of correction for long-range dispersion generally improves their performances; furthermore, for the systems examined, DFT-D3 and DFT-D3 (BJ) are unambiguously superior to DFT-D2 with respect to geometries, though not necessarily with respect to equilibrium binding energies. Conventional semilocal functionals, such as ωB97X and the Minnesota functionals, yield decidedly mediocre geometries, particularly for dispersion-dominated interactions. Furthermore, some of these functionals – most notably M06 and M06-L – suffer from nonphysical grid-dependent oscillations, particularly for the coronene dimer.[86] Such oscillations introduce a large degree of uncertainty into the equilibrium intermolecular separations and binding energies; after all, an ill-behaved potential energy surface has an ill-defined minimum.

Among the wavefunction-based approaches examined, Hartree-Fock theory yields grossly inadequate geometries – even for small hydrogen-bonded complexes, on which it might be expected to perform reasonably well. A perturbative correction for electronic correlation vastly improves upon this HF picture: MP2 yields highly accurate geometries for small molecules. However, the shortcomings of the method are manifest in the deterioration of the quality of its predicted geometries of larger dispersion-dominated systems. In systems with upwards of 8 heavy atoms, MP2 yields geometries that are at best on par with standard density functionals, even when corrected for BSSE. Attenuation of the Coulomb operator achieves a similar effect to counterpoise correction; both approaches systematically underestimate intermolecular separation in dispersion-bound systems. It is likely that increasing the size of the parameter space – by, e.g., combining attenuation with scaling – would improve the

description of geometries further, as it has been shown to do with energies.[87]

The evaluation of electronic structure methods with regard to their description of geometries adds an important additional dimension to benchmarking binding energies. In the past, relative energies have served as the primary means of comparison of various methods. This procedure has several shortcomings. There is no information regarding the shape of the potential energy surface associated with each method. Moreover, it is not a particularly fair comparison: typically, the same geometry is used for all methods, such that the reported energy is identically the equilibrium energy for only one method. A self-consistent treatment with each method, wherein the structure is relaxed prior to the energy calculation, is a more balanced approach. Incorporating some form of geometric metric – e.g., some measure of intermolecular separation for non-covalent complexes – into the training and selection of new density functionals may lead to the development of more robust methods, and can be achieved with relative ease. This has been done, in a fashion, in the development of the HCTH, τ-HCTH, and BMK functionals.[88–90] These functionals were parameterized on experimental geometries of a number of small molecules by incorporating the computed gradient at these reference geometries into the penalty function for each method. What we are proposing here is extending this sort of idea to non-covalent interactions, which have a handful of highly-variable intermolecular degrees of freedom. The inclusion of a single additional metric is by no means an end-all solution, as information relating to the shape of the potential energy surface is still absent, but it is a step in the right direction, and it is something for which sufficiently high quality reference data already exists and further data can be relatively straightforwardly generated.

Acknowledgement

The research was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences under Award DE-FG02-12ER16362. Work at the Molecular Foundry was supported by the U.S. Department of

Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering under Contract No. DE-AC02-05CH11231.

Supporting Information Available

Optimized structures for all molecules in the A21 set for each method examined here may be found online, as may a few additional tables. Additionally, raw binding energies are provided for each method at each separation in every system in the M12 set and the corenene dimer. This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Roothaan, C. *Rev. Mod. Phys.* **1951**, *23*, 69–89.

(2) Hall, G. G. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **1951**, *205*, 541–552.

(3) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.

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

(5) Eichkorn, K.; Treutler, O.; Ohm, H.; Marco, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283–290.

(6) Weigend, F.; Haser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1998**, *294*, 143–152.

(7) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. *Chem. Phys. Lett.* **1993**, *208*, 359–363.

(8) Jung, Y.; Sodt, A.; Gill, P. M. W.; Head-Gordon, M. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6692–6697.

(9) Hobza, P.; Selzle, H.; Schlag, E. *J. Phys. Chem.* **1996**, *100*, 18790–18794.

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

(11) Goldey, M.; Head-Gordon, M. *J. Phys. Chem. Lett.* **2012**, *3*, 3592–3598.

(12) Goldey, M.; Dutoi, A.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2013**, *15*, 15869–15875.

(13) Perdew, J. P.; Ruzsinszky, A.; Tao, J.; Staroverov, V. N.; Scuseria, G. E.; Csonka, G. I. *J. Chem. Phys.* **2005**, *123*, 062201.

(14) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.

(15) Kohn, W.; Sham, L. *Phys. Rev.* **1965**, *140*, A1133–A1138.

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

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

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

(19) Grimme, S. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 211–228.

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

(21) Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009**, *102*, 073005.

(22) Vydrov, O.; Van Voorhis, T. *Phys. Rev. Lett.* **2009**, *103*, 063004.

(23) Vydrov, O. A.; Van Voorhis, T. *J. Chem. Phys.* **2010**, *133*, 244103.

(24) Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* **2004**, *92*, 246401.

(25) Lee, K.; Murray, D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. *Phys. Rev. B* **2010**, *82*, 081101.

(26) Hehre, W. J.; Radom, L.; v.R. Schleyer, P.; Pople, J. A. *Ab Initio Molecular Orbital Theory*, 1st ed.; John Wiley and Sons: New York, 1986; pp 145–202.

(27) Helgaker, T.; Jo, P. *J. Chem. Phys.* **1997**, *106*, 6430–6440.

(28) Sánchez Márquez, J.; Fernández Núñez, M. *J. Mol. Struct. THEOCHEM* **2003**, *624*, 239–249.

(29) Riley, K.; Holt, B. O.; Merz, K. *J. Chem. Theory Comput.* **2007**, *3*, 407–433.

(30) Guido, C. a.; Knecht, S.; Kongsted, J.; Mennucci, B. *J. Chem. Theory Comput.* **2013**, *9*, 2209–2220.

(31) Tamblyn, I.; Refaely-Abramson, S.; Neaton, J. B.; Kronik, L. *J. Phys. Chem. Lett.* **2014**, *5*, 2734–2741.

(32) Gráfová, L.; Pitonák, M.; ÅŸezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2010**, *6*, 2365–2376.

(33) Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2011**, *7*, 2427–2438.

(34) Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2012**, *8*, 4285–4292.

(35) Goerigk, L.; Kruse, H.; Grimme, S. *ChemPhysChem* **2011**, *12*, 3421–3433.

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

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

(38) Takatani, T.; David Sherrill, C. *Phys. Chem. Chem. Phys.* **2007**, *9*, 6106–6114.

(39) Rao, L.; Ke, H.; Fu, G.; Xu, X.; Yan, Y. *J. Chem. Theory Comput.* **2009**, *5*, 86–96.

(40) Thanthiriwatte, K. S.; Hohenstein, E. G.; Burns, L. a.; Sherrill, C. D. *J. Chem. Theory Comput.* **2011**, *7*, 88–96.

(41) Hujo, W.; Grimme, S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 13942–13950.

(42) Vydrov, O. A.; Van Voorhis, T. *J. Chem. Theory Comput.* **2012**, *8*, 1929–1934.

(43) Remya, K.; Suresh, C. H. *J. Comput. Chem.* **2013**, *34*, 1341–1353.

(44) Head-Gordon, M.; Pople, J.; Frisch, M. *Chem. Phys. Lett.* **1988**, *153*, 503–506.

(45) Gordon, M. S.; Truhlar, D. G. *J. Am. Chem. Soc.* **1986**, *108*, 5412–5419.

(46) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095–910.

(47) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. *J. Chem. Phys.* **2004**, *121*, 9793–802.

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

(49) Lee, C.; Yang, W.; Parr, R. *Phys. Rev. B* **1988**, *37*, 785–789.

(50) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

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

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

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

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

(55) Peverati, R.; Truhlar, D. *J. Phys. Chem. Lett.* **2011**, *2*, 2810–2817.

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

(57) Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

(58) Mardirossian, N.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924.

(59) Mardirossian, N.; Head-Gordon, M. *J. Chem. Phys.* **2015**, *142*, 074111.

(60) Murray, E. D.; Lee, K.; Langreth, D. C. *J. Chem. Theory Comput.* **2009**, *5*, 2754–2762.

(61) Perdew, J.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 244–249.

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

(63) Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456–1465.

(64) Johnson, E. R.; Becke, A. D. *J. Chem. Phys.* **2005**, *123*, 24101.

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

(66) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; DiStasio, R. a.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. a.; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. a.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. a.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock, H. L.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer, H. F.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.

(67) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Evangelista, E. G. H. F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 556–565.

(68) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. *J. Cheminform.* **2012**, *4*, 1–17.

(69) Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2013**, *9*, 2151–2155.

(70) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243–252.

(71) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Phys. Chem. A* **1999**, *103*, 8265–8271.

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

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

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

(75) Gill, P. M.; Johnson, B. G.; Pople, J. A. *Chem. Phys. Lett.* **1993**, *209*, 506–512.

(76) Witte, J.; Neaton, J. B.; Head-Gordon, M. *J. Chem. Phys.* **2014**, *140*, 104707.

(77) Tang, K.; Toennies, J. *Zeitschrift für Phys. D Atoms, Mol. Clust.* **1986**, *1*, 91–101.

(78) Miliordos, E.; Aprà, E.; Xantheas, S. S. *J. Phys. Chem. A* **2014**, *118*, 7568–7578.

(79) Cohen, M.; Kelly, P. *Can. J. Phys.* **1966**, *44*, 3227–3240.

(80) Janowski, T.; Ford, A. R.; Pulay, P. *Mol. Phys.* **2010**, *108*, 249–257.

(81) Huang, Y.; Goldey, M.; Head-Gordon, M.; Beran, G. J. O. *J. Chem. Theory Comput.* **2014**, *10*, 2054–2063.

(82) Ruzsinszky, A.; Perdew, J. P.; Csonka, G. I. *J. Phys. Chem. A* **2005**, *109*, 11015–11021.

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

(84) Szabo, A.; Ostlund, N. S. *J. Chem. Phys.* **1977**, *67*, 4351–4360.

(85) Sedlak, R.; Janowski, T.; Pitoňák, M.; Řezáč, J.; Pulay, P.; Hobza, P. *J. Chem. Theory Comput.* **2013**, *9*, 3364–3374.

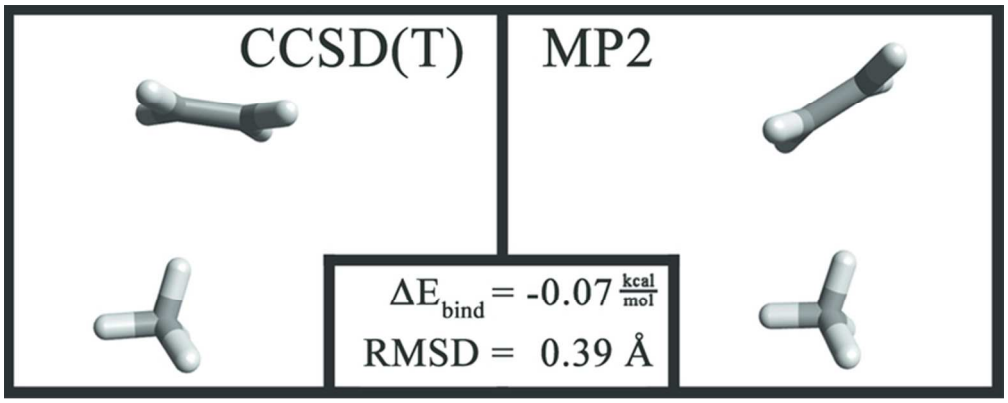
(86) Mardirossian, N.; Head-Gordon, M. *J. Chem. Theory Comput.* **2013**, *9*, 4453–4461.

(87) Goldey, M.; Head-Gordon, M. *J. Phys. Chem. B* **2014**, *118*, 6519–6525.

(88) Boese, A. D.; Handy, N. C. *J. Chem. Phys.* **2001**, *114*, 5497.

(89) Boese, A. D.; Handy, N. C. *J. Chem. Phys.* **2002**, *116*, 9559.

(90) Boese, A. D.; Martin, J. M. L. *J. Chem. Phys.* **2004**, *121*, 3405–16.



For Table of Contents Only
32x12mm (600 x 600 DPI)