

JCTC

Journal of Chemical Theory and Computation

Efficient Diffuse Basis Sets: cc-pVxZ+ and maug-cc-pVxZ

Ewa Papajak, Hannah R. Leverentz, Jingjing Zheng, and Donald G. Truhlar*

Department of Chemistry and Supercomputing Institute,
University of Minnesota, Minneapolis,
Minnesota 55455-0431

Received December 28, 2008

Abstract: We combine the diffuse basis functions from the 6-31+G basis set of Pople and co-workers with the correlation-consistent basis sets of Dunning and co-workers. In both wave function and density functional calculations, the resulting basis sets reduce the basis set superposition error almost as much as the augmented correlation-consistent basis sets, although they are much smaller. In addition, in density functional calculations the new basis sets, called cc-pVxZ+ where $x = D, T, Q, \dots$, or $x = D+d, T+d, Q+d, \dots$, give very similar energetic predictions to the much larger aug-cc-pVxZ basis sets. However, energetics calculated from correlated wave function calculations are more slowly convergent with respect to the addition of diffuse functions. We also examined basis sets with the same number and type of functions as the cc-pVxZ+ sets but using the diffuse exponents of the aug-cc-pVxZ basis sets and found very similar performance to cc-pVxZ+; these basis sets are called minimally augmented cc-pVxZ, which we abbreviate as maug-cc-pVxZ.

1. Introduction

The choice of basis set for electronic structure calculations is a key factor in determining both the reliability of a calculation and its cost.¹ In recent years, the correlation consistent (cc) basis sets of Dunning and co-workers^{2–6} have become very popular. These are arranged in two series, in particular cc-pVxZ and aug-cc-pVxZ with $xZ = DZ$ (“double- ζ ”), TZ (“triple- ζ ”), QZ (“quadruple- ζ ”), etc. The prefix aug- (for “augmented”) denotes the addition of diffuse basis functions, i.e., Gaussian functions with small exponential parameters times spherical harmonics. In the aug-cc-pVxZ basis sets, one adds quite a few diffuse

Table 1. Number of Contracted Basis Functions for $H_2SO_4(NH_3)$

	$x = D$	$x = D+d$	$x = T$	$x = T+d$	$x = Q$	$x = Q+d$
cc-pVxZ	113	118	254	259	484	489
cc-pVxZ+	137	142	278	283	508	513
maug-cc-pVxZ	137	142	278	283	508	513
aug-cc-pVxZ	187	192	395	400	714	719

functions, e.g., in aug-cc-pVDZ one adds 4 diffuse functions on each H and 9 on each heavy (i.e., non-hydrogenic) atom, whereas in aug-cc-pVTZ one adds 9 diffuse functions on each H and 16 on each heavy atom. Another popular set of basis functions is 6-31+G(d)⁷ and 6-311+G(2df,p) of Pople and co-workers⁸ in which the single plus sign denotes the addition of no diffuse functions on hydrogens and 4 on each heavy atom. (The same Gaussian exponential parameters are used for the diffuse basis functions in these two bases.) We have recently had some success combining the plus strategy with the cc basis sets, yielding cc-pVDZ+, cc-pVTZ+, and cc-pVQZ+, and so we decided to make some systematic comparisons of the performance of these strategies when applied to the same test problems. Those systematic comparisons are the primary subject of this article. We also show that for density functional calculations very similar results can be obtained by drastically pruning the aug-cc-pVxZ basis sets; this latter strategy is called minimal aug-cc-pVxZ, abbreviated maug-cc-pVxZ.

The above considerations also extend to a sequence⁶ of basis sets with $xZ = (D+d)Z, (T+d)Z$, etc. in which five extra d functions are added to each element from Al to Ar. Table 1 compares the number of basis functions in the various series of basis sets for one of our test cases, namely the complex of H_2SO_4 with NH_3 (similar trends would be observed for other sample molecules). The table shows that tremendous cost savings are attainable with the plus basis sets, if they are accurate enough. Are they?

2. Tests

Several issues were considered in choosing and designing our tests. First, diffuse functions are more important, in general, for energetics than for geometries and are more important for electron affinities, barrier heights, and noncovalent interactions than for ionization potentials and bond energies. Second, when simply examining the absolute accuracy of some computed properties, such as geometries or strengths of noncovalent interactions, the improvement in the basis set due to adding diffuse functions may be masked by other aspects of the calculations that lead to larger errors than the basis set's incompleteness. Third, a special aspect of basis set incompleteness that demands attention, especially in the computation of

* Corresponding author e-mail: truhlar@umn.edu.

noncovalent interactions, is basis set superposition error (BSSE),^{9,10} and if this is not small it is not even clear *how* to extract an estimate of the interaction energy from the calculations. For example, should one include a counterpoise correction (CpC) and how should that correction be defined? The CpC becomes not only ambiguous but also costly when one considers clusters containing more than two molecules,¹¹ and we therefore wish to keep it small and avoid it.

With these considerations in mind, we decided to carry out six sets of test calculations. In the first four we compare calculated ionization potentials, electron affinities, bond energies, and barrier heights to databases of accurate (or at least best available) estimates. In the fifth and sixth we examine the size of the basis set superposition error for a set of noncovalent complexes.

The databases used for the first four tests are as follows:

- ionization potentials: database IP13/3,^{12,13} containing C, S, O, Si, P, Cl, OH, PH, SH, PH₂, O₂, S₂, and Cl₂;
- electron affinities: database EA13/3,^{12,13} containing the same 13 species;
- atomization energies: database AE6,¹⁴ containing SiH₄, SiO, S₂, propyne, glyoxal, and cyclobutane;
- barrier heights: database DBH24/08,¹⁵ consisting of 4 subdatabases: HATBH6, NSBH6, UABH6, and HTBH6 containing 6 barrier heights each (forward and reverse barrier heights of three reactions) for heavy-atom transfer, nucleophilic substitution, unimolecular and association reactions, and hydrogen transfer, respectively.

As explained in previous work, all energetic data in these databases are zero point exclusive, and all calculations for comparison to these databases use standard (QCISD/MG3) geometries.^{12–15} All the results presented here include spin–orbit contributions¹⁶ for the C, O, F, Si, S, Cl, OH, and HS species.

For the fifth and sixth test, we studied ten complexes, which are shown in Figure 1. For each complex we calculated the BSSE by the functional counterpoise method⁹ of Boys and Bernardi and expressed it as a percentage

$$\%CpC = \frac{|CpC|}{BE} \times 100\% \quad (1)$$

where BE is the CpC-corrected binding energy. To avoid comparing calculations at different geometries (geometry can have a large effect on BSSE), all CpC and BE calculations were carried out at standard geometries. In particular, the geometries for Cl₂(NH₃), HCl(H₂S), (C₂H₄)₂, and (CH₄)₂ were taken from the NCCE31/05 database,^{17–19} the geometry of oa-H₂SO₄(NH₃) complex is a nonstationary point selected to show a particular interaction motif, and the others were optimized by the M06-2X¹⁹ density functional with the MG3S¹² basis set. The geometries that are not published elsewhere are given in the Supporting Information.

Because basis set requirements are different for density functional theory²⁰ (DFT) and wave function theory²¹ (WFT), we made calculations with both. For density functionals we consider B3LYP,²² M05-2X,¹⁸ M06,¹⁹ M06-2X,¹⁹ M06-L,²³ and M06-HF.²⁴ Note that M06-L is a local density functional, and the others are all nonlocal (also called hybrid). For WFT, we consider Hartree–Fock²⁵ (HF), second-order perturbation theory²⁶ (MP2), coupled cluster theory with single and double

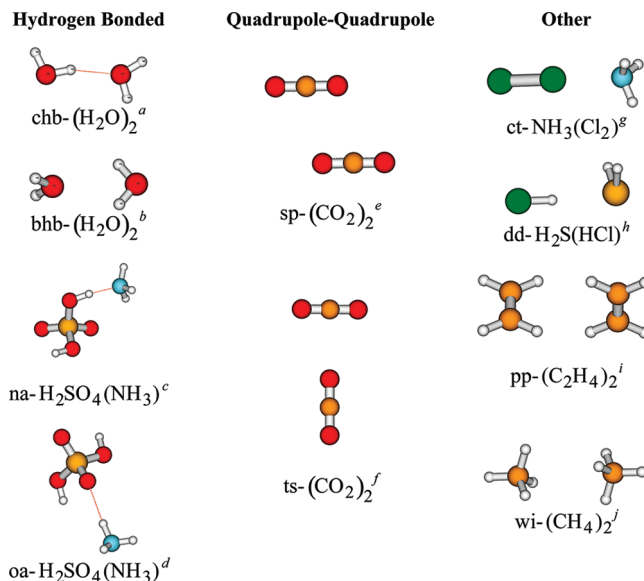


Figure 1. Ten complexes used for the BSSE tests: (a) conventionally hydrogen-bonded (chb) water dimer, (b) bifurcated hydrogen-bonded (bhb) water dimer, (c) hydrogen-bonded dimer of sulfuric acid and ammonia where nitrogen is a hydrogen acceptor (na), (d) hydrogen-bonded dimer of sulfuric acid and ammonia where oxygen is a hydrogen acceptor (oa), (e) slipped parallel (sp) carbon dioxide dimer, (f) T-shaped (ts) carbon dioxide dimer, (g) charge transfer (ct) dimer of chlorine and ammonia, (h) dipole-dipole (dd) hydrogen chloride and sulfur hydride dimer, (i) π - π stacking (pp) ethene dimer, and (j) weak interaction (wi) methane dimer.

Table 2. Most Diffuse Exponents of Various Symmetries

<i>l</i>	basis set	H	C	N	O	S	Cl
<i>s</i>	cc-pVTZ	0.1027	0.1285	0.1787	0.2384	0.1322	0.1625
	cc-pVTZ+	0.1027	0.0438	0.0639	0.0845	0.0405	0.0483
	aug-cc-pVTZ	0.02526	0.04402	0.0576	0.07376	0.0497	0.0591
	maug-cc-pVTZ	0.1027	0.04402	0.0576	0.07376	0.0497	0.0591
<i>p</i>	cc-pVTZ	0.388	0.1209	0.1725	0.214	0.1098	0.1301
	cc-pVTZ+	0.388	0.0438	0.0638	0.0845	0.0405	0.0483
	aug-cc-pVTZ	0.102	0.03569	0.0491	0.05974	0.0351	0.0419
	maug-cc-pVTZ	0.388	0.03569	0.0491	0.05974	0.0351	0.0419
<i>d</i>	cc-pVTZ	1.057	0.318	0.469	0.645	0.269	0.344
	cc-pVTZ+	1.057	0.318	0.469	0.645	0.269	0.344
	aug-cc-pVTZ	0.247	0.1	0.151	0.214	0.101	0.135
	maug-cc-pVTZ	1.057	0.318	0.469	0.645	0.269	0.344

excitations²⁷ (CCSD), and CCSD with quasiperturbative connected triple excitations²⁸ (CCSD(T)).

All density functional calculations were performed using the Gaussian 03²⁹ program and the MN-GFM³⁰ functional module. CCSD(T) calculations were carried out by Molpro-2006.1.³¹

3. Diffuse Functions

Since the exponential diffuse functions originally developed⁷ for use with the 3-21G, 6-21G, and 6-31G basis sets were later found⁸ to also be appropriate for use with the 6-311G basis set, we form the new “plus” basis sets by simply adding a diffuse *sp* shell to each non-hydrogenic atom, using the standard exponents that are used in both the 6-31+G and 6-311+G basis sets. To give the reader some idea of how these compare to the most diffuse exponents in the augmented correlation-consistent basis sets, Table 2 compares the exponents of the most diffuse *s*, *p*, and *d* functions in the triple- ζ basis set series under

Table 3. Mean Unsigned Errors (MUEs) (in kcal/mol) in Ionization Potentials

	B3LYP	M06-2X	CCSD(T)
cc-pVDZ	3.14	2.98	10.23
cc-pVDZ+	4.63	2.87	8.85
aug-cc-pVDZ	4.77	2.96	5.68
cc-pVTZ	4.15	2.39	3.44
cc-pVTZ+	4.59	2.63	3.15
aug-cc-pVTZ	4.67	3.45	1.93

Table 4. Mean Unsigned Errors (MUEs) (in kcal/mol) in Electron Affinities

	B3LYP	M06-2X	CCSD(T)
cc-pVDZ	20.84	18.77	30.98
cc-pVDZ+	3.05	4.05	10.49
aug-cc-pVDZ	3.03	3.69	5.27
cc-pVTZ	10.14	8.35	15.05
cc-pVTZ+	2.34	3.06	5.34
aug-cc-pVTZ	2.51	2.98	2.01

consideration here. We see that the exponential parameters of s and p diffuse functions of the plus and aug basis sets are similar.

The following question arises: what is a “diffuse” basis function? In practice the exponent of the most diffuse valence basis function in the cc-pV x Z series decreases as x increases. A similar trend is found in even-tempered basis sets.³² Eventually, one might wonder, does not the most diffuse basis function become indistinguishable from a “diffuse” basis function? In principle, yes, but in practice, no. For example consider the most diffuse s on carbon in the cc-pV x Z bases. For $x = 2-6$, these are 0.1596, 0.1285, 0.1111, 0.1019, and 0.08635,³³ respectively, whereas the added diffuse s function in the plus set has exponent 0.0438, and the added diffuse s functions in the aug sets have exponents of 0.0469, 0.04402, 0.04145, 0.0394, and 0.0354. Thus there is a wide gap (0.086 to 0.044) between the most diffuse valence basis function and the added “diffuse” basis function, and one would have to go to unrealistically high x before there is any ambiguity; so the definition of “diffuse” is operationally clear.

One could imagine optimizing diffuse functions for various properties such as atomic or molecular Rydberg states, atomic or molecular electron affinities, atomic or molecular polarizabilities, atomic or molecular electron densities, and so forth. One would obtain different results by optimizing against different property databases, but the optimum diffuse functions by any of these approaches would be expected, based on experience, to be a factor of $\sim 2-4$ smaller than the exponent of the least diffuse valence basis function.

4. Results and Discussion

Tables 3–6 give errors relative to the databases^{12–15} explained in Section 2.

Tables 3 and 4 show results for ionization potentials and electron affinities. The DFT errors are almost the same for the plus and aug basis sets, even for electron affinities. The coupled cluster calculations, however, are more slowly convergent with respect to the addition of diffuse functions. In order to illustrate this even more clearly, we added the heavy-atom diffuse d functions of the aug-cc-pVTZ basis to cc-pVDZ+. The

Table 5. Mean Unsigned Errors per Bond (MUEPBs) (in kcal/mol) in Atomization Energies^a

	B3LYP	M06-2X	CCSD(T)
cc-pVDZ	2.80	2.07	8.87
cc-pVDZ+	3.16	2.38	8.87
aug-cc-pVDZ	2.81	2.13	8.07
cc-pVTZ	0.85	0.40	3.12
cc-pVTZ+	0.94	0.44	3.11
aug-cc-pVTZ	0.94	0.42	2.62

^a The MUE in atomization energies is divided by the average number (4.83) of bonds per molecule in the AE6 database.

Table 6. Mean Unsigned Errors (MUEs) (in kcal/mol) in the Barrier Heights of the DBH24/08 Database

	HATBH6	NSBH6	UABH6	HTBH6	DBH24
B3LYP/cc-pVDZ	8.63	12.89	2.21	6.48	7.55
B3LYP/cc-pVDZ+	7.45	3.91	2.02	5.98	4.84
B3LYP/aug-cc-pVDZ	7.53	4.15	2.17	5.71	4.89
B3LYP/cc-pVTZ	6.86	9.59	1.65	4.88	5.75
B3LYP/cc-pVTZ+	6.69	4.23	1.71	4.68	4.33
B3LYP/aug-cc-pVTZ	6.81	3.32	1.70	4.71	4.13
M06-2X/cc-pVDZ	2.78	9.64	1.29	1.94	3.91
M06-2X/cc-pVDZ+	1.90	1.34	1.24	1.53	1.50
M06-2X/aug-cc-pVDZ	1.46	1.25	1.35	1.27	1.33
M06-2X/cc-pVTZ	1.08	5.21	1.02	1.43	2.18
M06-2X/cc-pVTZ+	0.77	0.76	1.13	1.30	0.99
M06-2X/aug-cc-pVTZ	0.67	0.66	1.10	1.30	0.93
CCSD(T)/cc-pVDZ	3.94	7.94	1.40	3.03	4.08
CCSD(T)/cc-pVDZ+	4.35	1.08	1.44	1.76	2.16
CCSD(T)/aug-cc-pVDZ	2.27	1.99	0.68	0.76	1.42
CCSD(T)/cc-pVTZ	1.90	5.27	0.47	1.41	2.26
CCSD(T)/cc-pVTZ+	1.65	0.81	0.55	0.99	1.00
CCSD(T)/aug-cc-pVTZ	0.95	0.78	0.33	0.70	0.69

CCSD(T) MUE dropped to only 2.73 kcal/mol for ionization potentials and 3.57 kcal/mol for electron affinities. Although diffuse functions are important for DFT,¹² it is apparently not necessary to exhaust the diffuse space as seems to be required in WFT.

Table 5 shows results for atomization energies. Here the diffuse functions have only a small effect. Table 6, however, shows a large effect for barrier heights, especially for nucleophilic substitution reactions (which is expected since these reactions involve anions). In all DFT cases (all four subdatabases, two density functionals, and both double and triple- ζ tests), the plus basis sets account quite well for the effect of diffuse functions seen with the larger aug basis sets. They are less successful for the WFT tests, but they often pick up a good fraction of the diffuse function increment at much less cost than is required for aug calculations.

A specific example of barrier height calculations is presented in Table 7. The forward and reverse barrier heights of one of the reactions in the DBH24/08 database were calculated using M06-2X density functional with cc-pV x Z, cc-pV x Z+, and aug-cc-pV x Z ($x = T$ and Q) basis sets. Adding diffuse s and p functions to the heavy atoms changes the calculated barrier height by 4.64–9.46 kcal/mol. The additional effect of adding diffuse functions on hydrogens and higher angular momentum diffuse functions (d and f when $x = T$ and d, f , and g when $x = Q$) on heavy atoms provides a further change in barrier heights of only 0.70–0.85 kcal/mol. The numbers of contracted Gaussian basis functions for this transition state are 146, 158, and 230 for the cc-pVTZ, cc-pVTZ+, and aug-cc-pVTZ basis sets, respectively; these numbers are 285, 297, and 424 for the

Table 7. Barrier Heights (kcal/mol) of the Reaction $\text{OH}^- + \text{CH}_3\text{F} \leftrightarrow \text{HOCH}_3 + \text{F}^-$ Calculated Using cc-pVxZ, cc-pVxZ+, and aug-cc-pVxZ ($x = \text{T}$ and Q) Basis Sets with the M06-2X Density Functional^a

basis set	V_f	V_r
cc-pVTZ	-12.71	8.67
cc-pVTZ+	-3.25	17.52
aug-cc-pVTZ	-2.40	18.28
cc-pVQZ	-8.38	12.91
cc-pVQZ+	-2.37	17.85
aug-cc-pVQZ	-1.60	18.55

^a V_f is the forward barrier height, and V_r is the reverse barrier height.

Table 8. %CpC Values Averaged Over Results from Five Levels of DFT, Four Levels of WFT, and Two Geometries of $\text{H}_2\text{SO}_4(\text{NH}_3)$

basis set	average %CpC
cc-pVTZ	17.0
cc-pV(T+d)Z	17.0
cc-pVTZ+	8.1
cc-pV(T+d)Z+	8.2
aug-cc-pVTZ	8.9
aug-cc-pV(T+d)Z	8.7

cc-pVQZ, cc-pVQZ+, and aug-cc-pVQZ basis sets, respectively. The relative computational costs of the cc-pVxZ+ basis sets are 1.2 for $x = \text{T}$ and 1.1 for $x = \text{Q}$. In the case of the aug-cc-pVxZ basis sets they are 4.4 for $x = \text{T}$ and 5.5 for $x = \text{Q}$. These costs are calculated as the ratio of the cost of a single-point energy calculation for the plus or aug basis set relative to the cost for the corresponding plain basis sets (cc-pVxZ); calculations were performed on a Linux cluster with two dual-core 2.6 GHz AMD Opteron processors. Therefore the plus basis sets take account of the main effect of diffuse functions for these barrier height calculations without significantly increasing computational costs.

Next we turn to BSSE. The first BSSE test is restricted to the two $\text{H}_2\text{SO}_4(\text{NH}_3)$ complexes. We averaged the %CpC over the two geometries, with five levels of DFT (M05-2X and all four members of the M06 functional family) and four levels of WFT (HF, MP2, CCSD, and CCSD(T)) for a total of 18 cases for each basis set. The results are in Table 8, and they are dramatic; either the plus or the aug basis sets reduce %CpC by a factor of 2. With DFT, the computer time for the aug calculations in Table 8 is about 4 times larger than the cc-pVTZ calculations, whereas the time for the cc-pVTZ+ calculations is only 1.5 times larger than that for cc-pVTZ with the four hybrid density functionals and almost the same as cc-pVTZ with M06-L.

The cc-pV(T+d)Z+ basis is very similar in quality to the MG3S basis set¹² that our group has used successfully in many applications. For example, MG3S has 269 contracted functions vs 259 (Table 1) for cc-pV(T+d)Z+, and they have identical diffuse s and p exponents (Table 2). However MG3S gives 13.5% for the test in Table 8, so cc-pVTZ+ should be considered as a possible improvement for calculating noncovalent interactions.

The final test involves averaging %CpC over all ten complexes with the M06-2X density functional and the HF, MP2, CCSD, and CCSD(T) levels of WFT, for a total of 50 cases for

Table 9. Average %CpC Values Over Four Levels of WFT and One Level of DFT for All Ten Complexes

basis set	average
cc-pV(T+d)Z	39.4
cc-pV(T+d)Z+	25.3
aug-cc-pV(T+d)Z	20.2

Table 10. Mean Unsigned Errors (MUEs) (in kcal/mol) in the Barrier Heights of the DBH24/08 Database Using maug-cc-pVDZ and maug-cc-pVTZ Basis Sets^a

	HATBH6	NSBH6	UABH6	HTBH6	DBH24
B3LYP/maug-cc-pVDZ	7.66	3.31	2.02	5.83	4.71
M06-2X/maug-cc-pVDZ	1.95	1.25	1.23	1.39	1.46
CCSD(T)/maug-cc-pVDZ	4.15	0.99	1.44	1.85	2.11
B3LYP/maug-cc-pVTZ	6.71	3.28	1.71	4.68	4.09
M06-2X/maug-cc-pVTZ	0.73	0.73	1.14	1.29	0.97
CCSD(T)/maug-cc-pVTZ	1.70	0.30	0.55	1.00	0.89

^a The maug-cc-pVxZ basis set is the aug-cc-pVxZ with all diffuse functions removed except for the s and p diffuse functions on non-hydrogenic atoms.

Table 11. Counterpoise Correction (CpC, kcal/mol) from M06-2X Calculations for Various Dimers Using the cc-pV(T+d)Z+ and maug-cc-pV(T+d)Z Basis Sets

dimer	cc-pV(T+d)Z+	maug-cc-pV(T+d)Z ^a	ΔCpC^b
chb-(H ₂ O) ₂	0.16	0.12	-0.03
bhb-(H ₂ O) ₂	0.25	0.24	-0.02
na-H ₂ SO ₄ (NH ₃)	0.38	0.35	-0.02
oa-H ₂ SO ₄ (NH ₃)	0.11	0.11	0.00
sp-(CO ₂) ₂	0.09	0.06	-0.03
ts-(CO ₂) ₂	0.08	0.06	-0.02
ct-NH ₃ (Cl ₂)	0.21	0.23	0.02
dd-H ₂ S(HCl)	0.13	0.10	-0.02
pp-(C ₂ H ₄) ₂	0.06	0.06	0.00
wi-(CH ₄) ₂	0.03	0.03	0.00
MSD ^c			-0.01

^a The maug-cc-pV(T+d)Z basis set is the aug-cc-pV(T+d)Z with all diffuse functions removed except for the s and p diffuse functions on non-hydrogenic atoms. ^b This column gives the energy difference (in kcal/mol) between the two previous columns; the cc-pV(T+d)Z+ CpC result is subtracted from the maug-cc-pV(T+d)Z CpC one. ^c MSD stands for "mean signed difference" and is the average of the ΔCpC values (in kcal/mol).

each basis set. The results are in Table 9. Here the cc-pV(T+d)Z+ basis set reduces %CpC by 36%, whereas aug-cc-pV(T+d)Z reduces it by 49%, averaged over the 50 cases.

So far these tests have been applied to the plus basis sets. Next we consider whether similar performance would be obtained with the maug-cc sets. (One can anticipate, especially from the discussion in Section 3, that the performance *will* be similar, but it may be useful to quantitatively validate this.) Table 10 shows some results, similar to Table 6, for maug-cc basis sets; the results for maug-cc-pVDZ are similar to those in Table 6 for the plus basis sets. The results for maug-cc-pVDZ and maug-cc-pVTZ are very similar for reactions in the DBH24/08 database to those for cc-pVDZ+ and cc-pVTZ+ basis sets. For %CpC we tested the maug-cc-pV(T+d)Z basis set; a comparison of the cc-pV(T+d)Z+ basis set with the maug-cc-pV(T+d)Z basis set for the binding energies and basis set superposition errors of the dimers shown in Figure 1 is given in Table 11. Again, we see a very similar performance for cc+ and maug-cc basis sets, as expected.

A strategy similar to the plus and maug strategies is to delete all diffuse functions of the aug-cc-pVxZ basis sets on hydrogenic atoms and keep all the diffuse functions on non-hydrogenic atoms.^{34–36} Another strategy is to delete the highest-angular-momentum diffuse functions on both hydrogenic atoms and non-hydrogenic atoms. Both these strategies have been employed more widely than the plus or maug strategy, although they both lead to much larger basis sets (for a given x) than plus or maug. Note that reducing the number of basis functions for a given x either decreases computation time for that x or allows the use of higher x and therefore increases accuracy.

The effect of diffuse functions on CpC has also been examined previously.^{35–37}

It is beyond the scope of the present study to review the interesting theoretical issues underlying the applicability (or not) of DFT to electron affinities, but the interested reader can consult a review⁴⁰ for this. Jensen,⁴¹ in a more recent study, examined the importance of diffuse s , p , and d functions for computing electron affinities, multipole moments, and polarizabilities by DFT and found that diffuse s and p functions are more important than diffuse d functions, but diffuse d functions are sometimes important. He also pointed out the difficulty or impossibility of optimizing the diffuse functions in certain cases.

Selective pruning of unaugmented cc-pVxZ basis sets has also been examined. Wilson and co-workers examined such truncations^{38,39} and showed, for example, that h and g functions can often be deleted from the cc-pV5Z basis set with only a small loss of accuracy for energies or geometries.

5. Concluding Remarks

While the addition of diffuse functions is necessary, in many cases only diffuse functions of low angular momentum (i.e., diffuse s and p functions) are needed, and furthermore they are often needed only on heavy atoms. The cc-pVxZ+ basis sets are designed to take advantage of this, and as a consequence their basis set size and cost are much smaller than those of the aug-cc-pVxZ basis sets, which include diffuse functions of both low and high angular momentum on all atoms. However, the quality of DFT results obtained by using cc-pVTZ+ is much closer to that of aug-cc-pVTZ than to that of cc-pVTZ. Similarly, results for cc-pVDZ+ are closer to those of aug-cc-pVDZ than to those of cc-pVDZ. In fact, in many cases the plus basis sets perform equally well as the aug ones. The benefits of using cc-pVxZ+ basis sets with WFT are less dramatic than in the case of DFT. However, they can serve as a good intermediate step between cc-pVTZ and aug-cc-pVTZ, and in fact they are often affordable where the full aug basis is not.

Acknowledgment. The authors are grateful to Steven Mielke for discussions. This work was supported by the United States Department of Energy, Office of Basic Sciences under grant no. DE-FG02-86ER13579.

Supporting Information Available: Five geometries of complexes optimized by M06-2X/MG3S and one [oa-H₂SO₄(NH₃)] contrived geometry for the dimers shown in Figure 1 and geometries of the corresponding individual gas-phase monomers with which the binding energies of those dimers were

calculated. This material is available free of charge via the Internet at <http://pubs.acs.org/>.

References

- (1) Davidson, E. R.; Feller, D. *Chem. Rev.* **1986**, *86*, 681.
- (2) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.
- (3) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796. Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
- (4) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98*, 1358.
- (5) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1994**, *100*, 2975.
- (6) Dunning, T. H., Jr.; Peterson, K. A.; Wilson, A. K. *J. Chem. Phys.* **2001**, *114*, 9244.
- (7) Clark, T.; Chandrasekhar, J.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294.
- (8) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.
- (9) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
- (10) Schwenke, D. W.; Truhlar, D. G. *J. Chem. Phys.* **1985**, *82*, 2418. Simon, S.; Duran, M.; Dannenberg, J. J. *J. Chem. Phys.* **1996**, *105*, 11024. Errata: 1987, *86*, 3760.
- (11) Valiron, P.; Mayer, I. *Chem. Phys. Lett.* **1997**, *275*, 46.
- (12) Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 1384.
- (13) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 3898.
- (14) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *107*, 8996. Erratum: *J. Phys. Chem. A* **2004**, *108*, 1460.
- (15) Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 569. Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2009**, *5*, 808–821.
- (16) Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 1643.
- (17) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 5656.
- (18) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364.
- (19) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.
- (20) Kohn, W.; Becke, A. D.; Parr, R. G. *J. Phys. Chem.* **1996**, *100*, 12974.
- (21) Raghavachari, K.; Anderson, J. B. *J. Phys. Chem.* **1996**, *100*, 12960.
- (22) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.
- (23) Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 194101/1.
- (24) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 13126.
- (25) Roothaan, C. C. *J. Rev. Mod. Phys.* **1951**, *23*, 69.
- (26) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.
- (27) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910.
- (28) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.
- (29) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski,

- wski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian03, Revision C.02*; Gaussian, Inc.: Wallingford, CT, 2004.
- (30) Zhao, Y.; Truhlar, D. G. *MN-GFM: Minnesota Gaussian Functional Module—version 3.0*; University of Minnesota: Minneapolis, MN, 2007.
- (31) Werner, H.; Knowles, P. J.; Scheutz, M.; Lindh, R.; Celani, P.; Korona, T.; Rauhut, G.; Manby, F. R.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Ecker, F. H. C.; Hertzner, A.; Lloyd, W.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pitzer, R.; Schumann, U.; Stoll, H.; Stone, A. J.; Torroni, R.; Thorsteinsson, T. *1997, MOLPRO 2006.1*; University College Consultants Ltd.: Cardiff, Wales, UK, 2006.
- (32) Bardo, R. D.; Ruedenberg, K. *J. Chem. Phys.* **1973**, *59*, 5956.
- (33) Wilson, A. K.; van Mourik, T.; Dunning, T. H., Jr. *THEOCHEM* **1997**, *388*, 339.
- (34) Sinnokrot, M. O.; Sherill, C. D. *J. Phys. Chem. A* **2003**, *107*, 8377.
- (35) Hwang, R.; Muh, S. B.; Lee, J. S. *Mol. Phys.* **2003**, *101*, 1429.
- (36) Elsohly, A. M.; Tschumpter, G. S. *Int. J. Quantum Chem.* **2009**, *109*, 91.
- (37) Halkier, A.; Koch, H.; Jorgensen, P.; Christiansen, O.; Nielsen, I. M. B.; Helgaker, T. *Theor. Chem. Acc.* **1997**, *97*, 150.
- (38) Prascher, B. P.; Wilson, B. R.; Wilson, A. K. *J. Chem. Phys.* **2007**, *127*, 124110.
- (39) Mintz, B.; Friskell, S.; Shah, A.; Wilson, A. K. *Int. J. Quantum Chem.* **2007**, *107*, 3077.
- (40) Rienstra-Kiracofe, J. C.; Tsuchumper, G. S.; Schaefer, H. F., III.; Nanadi, S.; Ellison, G. B. *Chem. Rev.* **2002**, *102*, 1231.
- (41) Jensen, F. *J. Chem. Phys.* **2002**, *117*, 9234.

CT800575Z