Multicoefficient Density Functional Theory (MC−DFT)

Jien-Lian Chen, Yi-Lun Sun, Kuo-Jui Wu, and Wei-Ping Hu*

Department of Chemistry and Biochemistry, National Chung Cheng University, Chia-Yi 621, Taiwan

Received: July 26, 2007; In Final Form: October 1, 2007

We have systematically tested the performance of several pure and hybrid versions of density functional methods on different types of molecular energies by combining energies calculated using more than one basis sets. Most hybrid functionals show important performance improvement as compared to methods using only a single basis set. The results suggest that, in many cases, scaling the basis set corrections is also important for density functional theory calculation. The best method, the B1B95 functional using the cc-pVDZ/cc-pVTZ/aug-cc-pVDZ basis set combination, achieves an average accuracy of 1.76 kcal/mol on a database of 109 atomization energies, 38 hydrogen-transfer barrier heights, 38 non-hydrogen-transfer barrier heights, 13 ionization potentials, and 13 electron affinities.

Introduction

In the past decade, the density functional theory (DFT)[1−4] has become one of the most popular methods for quantum chemical study. On calculating molecular energies, current sophisticated DFT methods can usually achieve accuracies that can only be obtained by much more expensive, wavefunction-based ab initio methods. The computational costs of the DFT methods scale approximately as the fourth power of the system sizes (N^4). (Strictly speaking, the N^4 scaling is only correct for hybrid functionals. Pure GGAs and meta-GGAs scale as N^3. Both scalings are considerably less steep for practical calculations in finite precision with screening thresholds.) In comparison, the popular ab initio methods such as MP2, MP4(SDQ), and CCSD(T) scale as the fifth (N^5), sixth (N^6), and seventh powers (N^7) of the system sizes, respectively. Thus, the DFT is much more suitable, and sometimes the only applicable, theory to model large systems with reasonable accuracy. Although the DFT is an exact theory, today the form of the exchange-correlation functional can only be approximated with different levels of sophistication. Unlike the wavefunction-based methods which can be improved by including well-defined hierarchical levels of correlation energies, there are usually no known systematic approaches to improve an existing functional. For example, although the performance of some popular hybrid versions of the gradient-corrected functionals, such as B3LYP,[5] MPW1PW91,[6] B1B95,[7] B98,[8] BMK,[9] etc., are very impressive, they still cannot match the accuracy of the wavefunction-based methods that consider very high-level electron correlation, such as QCISD(T) and CCSD(T). The accuracy of the predicted atomization energies by the most accurate DFT methods today is still well above the so-called chemical accuracy of ∼1 kcal/mol, which is desired for some very energy-sensitive studies, such as chemical kinetics and dynamics. To improve the accuracy, there is little one can do other than redesigning a better approximation to the exact exchange-correlation functional whose true form remains elusive. Using a larger basis set with DFT can sometimes improve the results considerably,[10−15] but the improvement is still limited by the quality of the functionals. The traditional wisdom is that the basis-set effects in DFT are somewhat smaller than the wavefunction-based methods that include high-level electron correlation. Recently, the so-called multi-coefficient or multi-level electronic structure methods, such as MCG3/3,[16] G3S,[17] G3SX,[18] and MLSE*n*+*d*[19] have been shown to be able to predict very accurate molecular energies at modest computational costs. In these methods, additive corrections using scaled energy components calculated using a series of basis sets at different levels of theory are applied to a base energy to account for the incomplete treatment of the electron correlation and the incompleteness of the basis set. The scaling coefficients were determined by minimizing the errors to a set of database values derived from experiments or from very high-level calculation. In these methods, lower-level correlated calculations, such as MP2 or MP4, were performed using several different basis sets, and then the energy differences obtained are scaled to account for the corrections due to the basis-set-size effects. These corrections were proved to be essential to the success of these methods. Recently, Truhlar and co-worker[20] have shown that significant performance improvement can be achieved by combining the energies calculated from the multi-coefficient methods and from the hybrid DFT methods. For example, in their MCG3-MPWB method[20] the MCG3/3 energy was combined with energy calculated from the MPW1B95 method, and the performance was significantly improved over the MCG3/3 method with only very modest increase of the computational cost. Recently, we have also developed a set of related multilevel methods (MLSE-DFT)[21] using Dunning's correlation-consistent basis sets[22] for neutral systems. The average errors on a set of 169 molecular energies were as low as ∼0.6 kcal/mol. We found that there was a non-negligible performance gain by using a linear combination of the calculated DFT energies from two different basis sets. This is somewhat surprising because one normally assumes that, although DFT calculation with a larger basis set may give better results, the improvement may not be systematic with respect to the sizes of the basis sets. It would be very interesting to know whether the use of a linear combination of the DFT energies alone (without the wavefunction-based energies) can lead to better performance than the DFT using a single large basis set. Thus, in this paper, we report a systematic study on the effects of using two or three basis sets with several popular DFT methods.

---

* To whom correspondence should be addressed. E-mail: chewph@ccu.edu.tw.

10.1021/jp0758871 CCC: $40.75

Methods

The density functionals we tested in the current study are BB95,[23,24] MPWB95,[6,24] MPWPW91,[6] BLYP,[23,25] and TPSSKCIS[26,27] and the hybrid versions (i.e., mixed with Hartree−Fock or exact exchange energies) B1B95,[24] MPW1B95,[12] MPW1PW91,[6] TPSS1KCIS,[20] B3LYP, and B98. The "1"s in the first four functionals signify that they are one-parameter (X) hybrid DFT with the names of exchange and correlation functionals before and after "1", respectively. The parameter X determines the mixing percentage of the exact exchange energy to the total exchange energy. The last two functionals include the exact exchange energies by design. We first tested the performance of these functional on a set of 109 atomization energies (AEs) from the MGAE109/05 database[28] compiled by Truhlar and co-workers. The basis sets used are 6-31+G(d,p), MG3S,[29] cc-pVDZ, and cc-pVTZ. The MG3S basis set is the G3LargeMP2 basis set[30] without the diffuse functions on hydrogens. The energies calculated by using two different basis sets (B1 and B2) are

$$E_{2B} = E(\mathrm{DFT/B1}) + c_1\,[E(\mathrm{DFT/B2}) - E(\mathrm{DFT/B1})\,] + E_{SO} \quad (1)$$

where "DFT" is any density functional method listed above and c_1 is a coefficient determined by minimizing the mean unsigned error (MUE) to the database values. The E_{SO} is the tabulated spin-orbital coupling energy for selected open-shell species.[31] The B1/B2 combinations were restricted to 6-31+G(d,p)/MG3S and cc-pVDZ/cc-pVTZ. For the hybrid functionals, the mixing percentages X were also optimized to minimize the MUE except for the B3LYP and B98 functionals where the original parameters were used. (The X value in the B98 functional cannot be changed from the standard input file in the Gaussian 03 program.[32]) When two basis sets are used, the X values were restricted to be same.

We then tested the performance of these functionals on a set of 38 hydrogen transfer barrier heights (HTBHs) and a set of 38 non-hydrogen transfer barrier heights (NHTBHs) from the HTBH38/04[20] and NHTBH38/04[33] databases which were also compiled by Truhlar and co-workers. The coefficient c_1 in eq 1 and the X value were similarly determined as in the tests for AEs.

Since the cc-pVDZ/cc-pVTZ combination lacks the diffuse basis functions and is thus inappropriate for anionic systems, we have also included the aug-cc-pVDZ basis set in the current study, and the new energies using three basis sets (B1/B2/B3) were then calculated as

$$E_{3B} = E(\mathrm{DFT/B1}) + c_1\,[E(\mathrm{DFT/B2}) - E(\mathrm{DFT/B1})] + c_2\,[E(\mathrm{DFT/B3}) - E(\mathrm{DFT/B1})] + E_{SO} \quad (2)$$

The performance of the hybrid DFT methods on the atomization energies and barrier heights mentioned above and a set of 13 ionization potentials and 13 electron affinities from the IP13/3 and EA13/3 databases,[28] respectively, was tested using the aug-cc-pVTZ, MG3S basis sets and the combinations of cc-pVDZ/cc-pVTZ/aug-cc-pVDZ, 6-31+G(d,p)/MG3S, and cc-pVDZ/MG3S basis sets. In addition, for the B1B95, MPW1B95, B98, and B3LYP functionals we also used Jensen's pc-1, pc-2, aug-pc-1, and aug-pc-2 basis sets[34] which were developed for DFT calculations. The aug-pc-2 basis set and the combination of pc-1/pc-2/aug-pc-1 basis sets was also tested on all of the data mentioned above. The coefficients c_1 and c_2 in eqs 1 and 2 and the X values (except for B3LYP and B98 functionals) were determined to minimized the MUEs of all 211 data to the database values. It is noted that the current methods and the basis-set extrapolation parts of other multicoefficient methods, such as MCG3-DFT, G3S, MLSE-DFT, etc., all shared the same formalism with the basis-set extrapolation scheme by Schwenke[35] that the coefficients of the basis-set extrapolation terms are not derived from theory but are determined empirically based on experimental data or energies from accurate calculation.

TABLE 1: Performance on 109 Atomization Energies and the Parameters of the Pure MC−DFT Methods

	MUE (kcal/mol)	c_1
BB95		
cc-pVDZ	9.54	
cc-pVTZ	11.12	
cc-pVDZ/cc-pVTZ	9.34	0.301
6-31+G(d,p)	8.97	
MG3S	11.01	
6-31+G(d,p)/MG3S	8.39	−0.824
MPWB95		
cc-pVDZ	11.68	
cc-pVTZ	16.06	
cc-pVDZ/cc-pVTZ	11.41	−0.431
6-31+G(d,p)	13.63	
MG3S	15.82	
6-31+G(d,p)/MG3S	13.08	−0.986
MPWPW91		
cc-pVDZ	8.37	
cc-pVTZ	9.44	
cc-pVDZ/cc-pVTZ	7.97	0.406
6-31+G(d,p)	7.95	
MG3S	9.48	
6-31+G(d,p)/MG3S	7.86	−0.267
TPSSKCIS		
cc-pVDZ	7.10	
cc-pVTZ	5.60	
cc-pVDZ/cc-pVTZ	4.72	0.634
6-31+G(d,p)	5.58	
MG3S	5.35	
6-31+G(d,p)/MG3S	5.19	0.733
BLYP		
cc-pVDZ	13.65	
cc-pVTZ	7.28	
cc-pVDZ/cc-pVTZ	6.31	1.400
6-31+G(d,p)	8.37	
MG3S	4.03	
6-31+G(d,p)/MG3S	4.03	0.990

In the following sections, we will also call the methods using eqs 1 and 2 the multi-coefficient density functional theory or the "MC−DFT" approach. All of the electronic structure calculation was performed using the Gaussian 03 program.

Results and Discussion

(1) Performance on Atomization Energies. Table 1 shows that performance of five pure DFT methods on atomization energies. Since the performances of pure DFTs are significantly inferior to those of the hybrid DFTs, the results here basically serve only as comparison to the results that will be presented later. There is a somewhat unpleasant feature of the pure DFTs that the performance in many cases deteriorates from double-ζ basis set to triple-ζ basis sets, as shown in the table for BB95, MPWB95, and MPW1PW91 functionals. In these cases, the c_1 values are small or even negative. That is, although increasing the basis-set size does not improve the performance, the MC−DFT method is robust enough to find suitable c_1 values to improve the performances slightly over the single-basis-set results. For the TPSSKCIS and BLYP functionals, the triple-ζ basis sets perform better, and the cc-pVDZ/cc-pVTZ combina-

TABLE 2: Performance and the Parameters of the Hybrid MC−DFT Methods on 109 Atomization Energies

	MUE (kcal/mol)	*X*	c_1
	B1B95		
cc-pVTZ	3.26	28	
cc-pVDZ/cc-pVTZ	2.34	28	1.322
MG3S	2.58	28	
6-31+G(d,p)/MG3S	2.58	28	1.065
	B1B95-AE		
cc-pVTZ	3.13	26	
cc-pVDZ/cc-pVTZ	2.08	33	1.514
MG3S	2.58	28	
6-31+G(d,p)/MG3S	2.58	28	1.065
	MPW1B95		
cc-pVTZ	3.25	31	
cc-pVDZ/cc-pVTZ	3.24	31	0.950
MG3S	2.92	31	
6-31+G(d,p)/MG3S	2.88	31	0.944
	MPW1B95-AE		
cc-pVTZ	3.10	33	
cc-pVDZ/cc-pVTZ	2.97	36	1.228
MG3S	2.56	34	
6-31+G(d,p)/MG3S	2.54	34	1.057
	MPW1PW91		
cc-pVTZ	5.04	25	
cc-pVDZ/cc-pVTZ	3.24	25	1.548
MG3S	4.16	25	
6-31+G(d,p)/MG3S	3.47	25	1.449
	MPW1PW91-AE		
cc-pVTZ	4.02	19	
cc-pVDZ/cc-pVTZ	3.24	26	1.595
MG3S	3.45	20	
6-31+G(d,p)/MG3S	3.35	22	1.306
	TPSS1KCIS-AE		
cc-pVTZ	3.57	13	
cc-pVDZ/cc-pVTZ	3.55	13	1.046
MG3S	3.17	13	
6-31+G(d,p)/MG3S	2.79	18	1.545
	B3LYP		
cc-pVTZ	4.83	20	
cc-pVDZ/cc-pVTZ	3.39	20	1.333
MG3S	4.28	20	
6-31+G(d,p)/MG3S	3.34	20	1.574
	B98		
cc-pVTZ	4.12	21.98	
cc-pVDZ/cc-pVTZ	2.05	21.98	1.399
MG3S	3.02	21.98	
6-31+G(d,p)/MG3S	2.68	21.98	1.278

tion reduces the MUE by 16% and 13%, respectively, over the cc-pVTZ results. The performance gain using the 6-31+G(d,p)/MG3S combination is negligible.

Table 2 shows the performance of the hybrid DFT methods on the same set of AEs and the optimized parameters. In all cases, the triple-ζ basis sets perform significantly better than the double-ζ basis sets, and thus the results by a single double-ζ basis set are not listed. (The c_1 values are now all close or greater than one.) Except for the B3LYP and B98 functionals, the mixing percentages (X) of the exact exchange energy were reoptimized against the MGAE109/05 database values, and we append a suffix of "-AE" to the name of the functional for which the X value is reoptimized. For the B1B95, MPW1B95, and MPW1PW91 functionals, the original X values were also used for comparison. The optimized X values for the B1B95-AE functional using the MG3S basis set and the 6-31+G(d,p)/MG3S combination are the same ($X = 28$) as in the original B1B95 functional. Very significant improvement using dual basis sets (as compared to using a single basis set) is observed in many cases. The most impressive cases are the B98, B1B95-AE, B3LYP, and B1B95 functionals using the cc-pVDZ/cc-pVTZ combination with improvement in accuracy by 50%, 34%, 30%, and 28% respectively. The 6-31+G(d,p)/MG3S combination with the B3LYP, MPW1PW91, TPSS1KCIS-AE, and B98 functionals also have the improvement in accuracy by 22%, 17%, 12%, and 11%, respectively. The improvement for the 6-31+G(d,p)/MG3S combination is less impressive but is nonetheless important in the cases mentioned above. For a single-basis-set calculation, the one with MG3S basis set is in most cases better than the one with cc-pVTZ basis set. However, since the cc-pVDZ/cc-pVTZ combination gains more important performance improvement, the resulting methods with the B98, B1B95, B1B95-AE, MPW1PW91, and MPW1PW91-AE functionals are more accurate in predicting AEs than the same functionals using the MG3S basis set or the 6-31+G(d,p)/MG3S combination. The most accurate methods in Table 2 are the B98 and B1B95-AE functionals using the cc-pVDZ/cc-pVTZ combination with MUEs of 2.05 and 2.08 kcal/mol, respectively. The most popular B3LYP functional can be improved considerably with MUEs of 4.28 kcal/mol using the MG3S basis set and of 3.34 kcal/mol using the 6-31+G(d,p)/MG3S combination. Table 2 also shows the performance of the B1B95 functional with cc-pVDZ/cc-pVTZ combination gains another 11% improvement by changing X from 28 to 33. The MPW1PW91 functional gains significant improvement by reoptimizing X using single basis set, but very small improvement using dual basis sets. The MPW1B95 and MPW1B95-AE functionals show very small improvement using dual basis sets. All the hybrid functionals perform much better than their pure DFT counterparts in Table 1.

Similar tests and parametrization for the functionals listed in Tables 1 and 2 have been performed in a series of studies by Truhlar and co-workers using the MG3S and 6-31+G(d,p) basis set.[12,13,20] Our current results using a single basis set are entirely consistent with their studies. Our MPW1B95-AE functional using the MG3S basis set performed slightly better than their MPW1B95 because they intentionally used a very small training set (AE6) for the functional.[12] They concluded that both B1B95 and MPW1B95 outperformed B98 functionals in predicting AEs. Although this is true using a single MG3S basis set, however, Table 2 shows that, using the cc-pVDZ/cc-pVTZ combination, the B98 functional (MUE = 2.05 kcal/mol) outperformed B1B95/MG3S (MUE = 2.58 kcal/mol) and MPW1B95/MG3S (MUE = 2.92 kcal/mol) by a wide margin.

(2) Performance on Barrier Heights. Table 3 shows the performance of the hybrid functionals on the barrier heights. Here the parameters were determined to minimize the MUEs to the 76 barrier height values in the HTBH38/04 and NHTBH38/04 databases, and the resulting methods are named by appending "-BH" to the name of the functional for which the X values have been reoptimized. The performance of three previously developed DFT methods for thermochemical kinetics by Truhlar and co-workers, the MPW1K,[36] BB1K,[37] and MPWB1K[12] functionals, are also shown in the table. These methods are based on the MPW1PW91, B1B95, and MPW1B95 functionals, respectively, with the X optimized for thermochemical kinetics. Interestingly, in the current study, we obtained exactly the same X values for these functionals using the MG3S basis set even though our training set is different from theirs which did not contain any NHTBHs. As shown in Table 3, the performance on BHs was dramatically improved using the cc-pVDZ/cc-pVTZ combination in most cases. The B1B95-BH methods with the cc-pVDZ/cc-pVTZ combination performed very impressively with a MUE of 0.98 kcal/mol for the 76

TABLE 3: Performance on the Barrier Heights and the Parameters of the Hybrid MC−DFT Methods

	MUE (kcal/mol)					
	HTBH	NHTBH	All 76	AEs	X	c_1
		B1B95-BH				
cc-pVTZ	1.35	2.65	2.00	8.53	45	
cc-pVDZ/cc-pVTZ	0.91	1.04	0.98	3.46	39	1.981
MG3S (BB1K)	1.16	1.42	1.29	6.31	42	
6-31+G(d,p)/MG3S	1.03	1.44	1.23	4.89	42	1.706
		MPW1B95-BH				
cc-pVTZ	1.42	2.74	2.08	6.99	48	
cc-pVDZ/cc-pVTZ	0.91	1.12	1.02	5.49	42	2.026
MG3S (MPWB1K)	1.29	1.44	1.37	4.62	44	
6-31+G(d,p)/MG3S	1.31	1.39	1.35	3.79	43	1.291
		MPW1PW91-BH				
cc-pVTZ	1.40	2.98	2.19	13.36	46	
cc-pVDZ/cc-pVTZ	1.33	1.65	1.49	4.84	41	1.924
MG3S (MPW1K)	1.32	1.77	1.55	11.02	42.8	
6-31+G(d,p)/MG3S	1.32	1.77	1.55	11.02	42.8	1.000
		TPSS1KCIS-BH				
cc-pVTZ	1.90	3.16	2.53	11.30	43	
cc-pVDZ/cc-pVTZ	1.75	1.60	1.68	6.17	36	2.102
MG3S	1.72	1.59	1.66	9.43	40	
6-31+G(d,p)/MG3S	1.73	1.47	1.60	6.00	38	1.905
		B3LYP-BH				
cc-pVTZ	4.75	6.65	5.70	4.83	20	
cc-pVDZ/cc-pVTZ	3.04	3.66	3.35	11.57	20	2.540
MG3S	4.23	4.58	4.41	4.28	20	
6-31+G(d,p)/MG3S	3.50	4.24	3.87	11.30	20	3.961
		B98-BH				
cc-pVTZ	4.49	5.37	4.93	4.12	22	
cc-pVDZ/cc-pVTZ	2.32	2.52	2.42	12.99	22	2.737
MG3S	4.16	3.41	3.78	3.02	22	
6-31+G(d,p)/MG3S	3.54	2.97	3.26	10.04	22	3.183

barrier heights. This is even better than the QCISD(T)/MG3S results (MUE = 1.10 kcal/mol).[33] The MPW1B95-BH methods with the same combination also performed very well with a MUE of 1.02 kcal/mol. This is compared to the MUEs of 1.54, 1.29, and 1.37 kcal/mol by the MPW1K, BB1K, and MPWB1K methods, respectively. Additionally, the B1B95-BH method with the cc-pVDZ/cc-pVTZ combination gives a MUE of 3.46 kcal/mol on 109 AEs as compared to MUEs of 11.02, 6.31, and 4.62 kcal/mol by the MPW1K, BB1K, and MPWB1K methods. That is, the B1B95-BH method predicts very accurate BHs and at the same time gives significantly better AEs than other kinetic models. The MPW1PW91-BH method with the cc-pVDZ/cc-pVTZ combination performs similarly on the BHs as the original MPW1K method, however, the performance on AEs (MUE = 4.84 kcal/mol) is much better. The TPSS1KCIS-BH method performs well on the BHs but not very impressively for AEs. The B3LYP and B98 methods do not perform well on the BHs using any basis sets or basis set combinations.

(3) Performance on All Data. As mentioned in the Methods section, in order to make more accurate prediction on charged systems, we have added aug-cc-pVDZ, aug-cc-pVTZ, aug-pc-1, and aug-pc-2 basis sets that contain diffuse functions to our tests. The 6-31+G(d,p) and MG3S basis sets already contain diffuse functions for heavy atoms. In this part of the test, the parameters c_1 and c_2 in eqs 1 and 2 as well as the X values (except for the B3LYP and B98 functionals) are determined to minimize the MUEs to all of the 211 values in the MGAE109/05, IP13/3, EA13/3, HTBH38/04, and NHTBH38/04 databases. The resulting methods with the X values reoptimized are named by appending "-All" to the functional names. The test results and the parameters are shown in Table 4 where the performances of some recently developed highly optimized DFT methods including the BMK, PW6B95,[13] M05-2X,[28] and M06-2X[38] functionals are also included for comparison. It is noted that the X value of BMK functional cannot be modified in the Gaussian 03 program, and the last three functionals are not yet available in the distributed version of Gaussian 03, and thus, only the single-basis-set results from literature are listed. As expected, the performance using the single MG3S, and aug-pc-2 basis set is very similar with the latter slightly better. The average errors using a single aug-cc-pVTZ basis set is 0.1−0.7 kcal/mol larger than using the MG3S basis set. This is consistent with earlier studies[10,14,15] that, when using a single basis set, the performance of Dunning's basis sets for DFT calculation is less satisfactory than the highly polarized Pople-type basis sets. Although the purpose of the combination cc-pVDZ/cc-pVTZ/aug-cc-pVDZ may seem to approximate the larger aug-cc-pVTZ basis set, Table 4 shows clearly that in several cases the methods using the combination outperformed the ones using a single aug-cc-pVTZ or MG3S basis set by a wide margin. For example, the B1B95-All functional using the cc-pVDZ/cc-pVTZ/aug-cc-pVDZ combination, the overall best method, gives a MUE of 1.76 kcal/mol while the B1B95-All/aug-cc-pVTZ, B1B95-All/MG3S, B1B95-All/aug-pc-2 methods give MUEs of 2.84, 2.57, and 2.48 kcal/mol, respectively. The B1B95-All functional with the pc-1/pc-2/aug-pc-1 combination also performs reasonably well with a MUE of 2.02 kcal/mol. Once again, the performance can be dramatically improved using an appropriate basis set combination. If the aug-cc-pVDZ or the aug-pc-1 basis set in the above combinations is replaced by the aug-cc-pVTZ or the aug-pc-2 basis set, respectively, the performance is not improved in most cases, but the computational cost increases significantly. Thus it seems unnecessary to use the aug-cc-pVTZ or the aug-pc-2 basis set in the current MC−DFT approach. The improvement using the 6-31+G(d,p)/MG3S combination is usually less impressive. Interestingly, for the B1B95-All, MPW1PW91-All, and B98 functionals, the cc-pVDZ/MG3S combinations do result in substantial improvement over the single-basis-set MG3S calculation. Except for the MPW1B95-All functional, using the pc-1/pc-2/aug-pc-1 combination also results in significant improvement over the single aug-pc-2 basis set. For the B1B95 and MPW1B95 functionals, the cc-pVDZ/cc-pVTZ/aug-cc-pVDZ combination performs better than the pc-1/pc-2/aug-pc-1 combination, while for the B98 and B3LYP functionals, the pc-1/pc-2/aug-pc-1 combination performs slightly better. Table 4 also shows that for the pc-1/pc-2/aug-pc-1 combination, the contribution from the aug-pc-1 basis set is less significant with relatively smaller c_2 coefficients.

As seen in Table 4, the B1B95-All functional using the cc-pVDZ/cc-pVTZ/aug-cc-pVDZ combination gives very balanced performance on all types of energies. In fact, for the average MUE this method outperformes all other methods in Table 4 except for the most recent M06-2X functional which was extensively optimized for many types of molecular energies using 33 empirically determined parameters embedded in the functional. In comparison, our best method in this study uses an existing functional with three empirically determined parameters, and it achieves very similar performance. This makes the M06-2X and our B1B95-All with the cc-pVDZ/cc-pVTZ/aug-cc-pVDZ combination methods of choice for thermochemistry and thermochemical kinetics. The PW6B95 functional gives similarly impressive performance for thermochemistry. The BMK functional and our MPW1B95-All functional using the cc-pVDZ/cc-pVTZ/aug-cc-pVDZ combination also give very good and balanced performance with average MUEs of ∼2.0

TABLE 4: Mean Unsigned Errors (MUEs) on Various Types of Energies and Parameters of the Hybrid MC−DFT Methods

	AE	HTBH	NHTBHs	IP	EA	average[a]	X	c_1	c_2
B1B95-All									
aug-cc-pVTZ	3.12	2.84	2.38	2.01	2.70	2.84	28		
cc-pVDZ/cc-pVTZ/aug-cc-pVDZ	1.95	1.43	1.41	2.29	1.64	1.76	33	1.485	0.375
MG3S	2.58	2.80	2.27	2.18	3.16	2.57	28		
6-31+G(d,p)/MG3S	2.58	2.77	2.24	2.16	3.13	2.56	28	1.103	
cc-pVDZ/MG3S	2.11	1.62	2.15	2.35	1.90	2.03	33	1.247	
aug-pc-2	2.62	2.52	2.14	1.97	2.70	2.48	30		
pc-1/pc-2/aug-pc-1	2.28	1.37	2.07	1.82	1.90	2.02	36	1.472	0.180
MPW1B95-All									
aug-cc-pVTZ	2.79	2.72	2.10	1.95	2.63	2.59	33		
cc-pVDZ/cc-pVTZ/aug-cc-pVDZ	2.44	1.48	1.36	2.31	1.56	2.01	38	1.203	0.534
MG3S	2.56	2.50	1.90	2.58	3.19	2.47	34		
6-31+G(d,p)/MG3S	2.55	2.47	1.86	2.64	3.17	2.45	34	1.103	
cc-pVDZ/MG3S	2.59	1.94	1.76	2.19	1.33	2.22	36	1.145	
aug-pc-2	2.54	2.36	1.89	1.97	2.65	2.36	35		
pc-1/pc-2/aug-pc-1	2.67	1.87	1.82	1.91	1.59	2.26	39	1.089	0.381
MPW1PW91-All									
aug-cc-pVTZ	3.98	4.48	3.94	3.74	2.08	3.93	20		
cc-pVDZ/cc-pVTZ/aug-cc-pVDZ	3.35	1.79	1.82	4.14	1.82	2.75	26	1.782	0.196
MG3S	3.54	4.08	3.71	3.75	2.44	3.61	22		
6-31+G(d,p)/MG3S	3.67	3.12	2.71	3.67	2.82	3.35	27	1.585	
cc-pVDZ/MG3S	3.01	2.89	3.14	4.15	4.40	3.17	26	1.306	
TPSS1KCIS-All									
aug-cc-pVTZ	3.67	4.38	4.99	2.61	2.45	3.90	15		
cc-pVDZ/cc-pVTZ/aug-cc-pVDZ	4.12	2.57	3.23	3.06	1.98	3.48	23	1.346	0.516
MG3S	3.41	4.02	4.77	2.58	3.03	3.69	17		
6-31+G(d,p)/MG3S	3.16	2.95	3.46	2.48	3.82	3.17	23	1.823	
cc-pVDZ/MG3S	3.55	3.23	3.62	2.80	1.86	3.35	20	1.194	
B98									
aug-cc-pVTZ	4.34	4.21	3.35	3.14	1.63	3.90	21.98		
cc-pVDZ/cc-pVTZ/aug-cc-pVDZ	2.10	3.50	3.25	3.75	2.48	2.68	21.98	1.379	0.396
MG3S	3.02	4.16	3.41	3.21	1.84	3.23	21.98		
6-31+G(d,p)/MG3S	2.68	4.06	3.32	3.21	1.85	3.03	21.98	1.283	
cc-pVDZ/MG3S	2.00	3.82	2.93	3.47	3.30	2.67	21.98	1.176	
aug-pc-2	3.12	4.18	3.24	3.22	1.58	3.25	21.98		
pc-1/pc-2/aug-pc-1	2.05	3.63	3.06	2.97	1.74	2.55	21.98	1.531	−0.082
B3LYP									
aug-cc-pVTZ	5.21	4.28	4.58	4.66	2.51	4.73	20		
cc-pVDZ/cc-pVTZ/aug-cc-pVDZ	3.57	3.84	5.01	5.16	2.39	3.90	20	1.374	0.234
MG3S	4.28	4.23	4.58	4.72	2.29	4.23	20		
6-31+G(d,p)/MG3S	3.34	4.01	4.51	4.73	2.07	3.68	20	1.606	
cc-pVDZ/MG3S	3.36	3.90	4.11	5.11	5.68	3.84	20	1.191	
aug-pc-2	4.16	4.29	4.47	4.69	2.56	4.18	20		
pc-1/pc-2/aug-pc-1	3.23	3.85	4.45	4.47	2.20	3.58	20	1.476	−0.058
BMK									
aug-cc-pVTZ	2.81	1.48	1.53	3.63	1.51	2.31	42		
cc-pVDZ/cc-pVTZ/aug-cc-pVDZ	2.38	1.27	1.17	3.98	2.17	2.05	42	1.123	0.450
MG3S	2.22	1.32	1.27	4.21	1.56	1.97	42		
6-31+G(d,p)/MG3S	2.20	1.31	1.27	4.17	1.49	1.95	42	1.060	
cc-pVDZ/MG3S	2.15	1.29	1.31	4.22	1.44	1.93	42	1.035	
PW6B95[b]	1.88	3.14	3.24	1.78	2.93	2.37	28		
M05-2X[c]	2.26	1.34	1.77	3.54	2.03	2.07	56		
M06-2X[d]	1.88	1.13	1.22	2.54	2.07	1.68	54		

[a] MUE for all 211 data. [b] From ref 13 using the MG3S basis set. [c] From ref 28 using the MG3S basis set. [d] From ref 38 using the MG3S basis set.

kcal/mol. The B98 functional does not perform very well on all data because it underestimates the barrier heights significantly.

In principle, the MC−DFT approaches can also be applied to the new PW6B95, M05-2X, and M06-2X functionals. Unfortunately, these functionals are not yet widely available and we currently do not have access to these functionals. It would be very interesting to test these functionals using the current approaches in the future.

In fact, the cc-pVDZ/cc-pVTZ/aug-cc-pVDZ basis set combination can also be used to optimize the performance on AEs or BHs alone. The detailed results are included in the Supporting Information. The improvement over the cc-pVDZ/cc-pVTZ combination is in most cases insignificant, except for MPW1B95-AE where the improvement is 23%. For B1B95-BH, the improvement on the barrier height is only 2% (from 0.98 to 0.96 kcal/mol), but there is 25% improvement on AEs (from 3.46 to 2.59 kcal/mol). This makes the B1B95-BH with the cc-pVDZ/cc-pVTZ/aug-cc-pVDZ combination another method of choice for thermochemical kinetics.

The current MC−DFT approaches can also be viewed as a convenient and economical way of extrapolating the basis sets. The current study suggests that one probably does not need to use very large basis sets such as 6-311+G(3d2f, 3d2f, 2pd),

TABLE 5: Computational Cost[a]

basis set	time (s)[b]
6-31+G(d,p)	71
MG3S	588
cc-pVDZ	57
cc-pVTZ	584
aug-cc-pVDZ	223
aug-cc-pVTZ	3043
pc-1	57
pc-2	624
aug-pc-1	447
aug-pc-2	3661

[a] Tested using the B1B95 functional by computers with Intel E6600 CPU and a version of Gaussian 03 program optimized for the hardware architecture. [b] Total CPU time in seconds to calculate C_5H_5N, C_2Cl_4, C_4H_4O, C_4H_4S, C_4H_5N, CF_3CN, and $SiCl_4$.

aug-cc-pVQZ, or aug-pc-3 to reach the accuracy at the basis-set limit. Curtiss et al.[14] showed that for the 147 enthalpies of formation in the G2/97 set, the MUEs calculated with the aug-cc-pVQZ basis set using the B98 and B3LYP functionals are 2.4 and 3.3 kcal/mol, respectively. These values are similar to the cc-pVDZ/cc-pVTZ results of 2.1 and 3.4 kcal/mol in Table 2. Effectively, the coefficients in the current MC−DFT approaches were determined to extrapolate the basis set to an optimal size so as to make the best agreement with the database results.

(4) Computational Cost. Table 5 shows the computational cost for the B1B95 functional using various basis sets. For medium-sized molecules, the computational cost for MG3S, cc-pVTZ, and pc-2 basis sets is similar. The computational cost of 6-31+G(d,p), cc-pVDZ, and pc-1 calculation is almost negligible. Thus the cost for the cc-pVDZ/cc-pVTZ or cc-pVDZ/MG3S combinations is only slightly higher than that using the MG3S basis set. The sum of cost using cc-pVDZ and aug-cc-pVDZ basis sets is about 50% of that using cc-pVTZ basis set; thus, the cost for cc-pVDZ/cc-pVTZ/aug-cc-pVDZ combination is approximately 50% higher than using the single MG3S or cc-pVTZ basis set. The cost for pc-1/pc-2/aug-pc-1 combination is approximately 80% higher than using the single pc-2 basis set. As shown in Tables 2−4, the improvement using the MC−DFT approach is, in many cases, very significant, and thus the extra cost is well justified. The cost of using aug-cc-pVTZ or aug-pc-2 basis set is 5−6 times higher that that of cc-pVTZ or MG3S and with only very limited further performance improvement in a few cases. Thus, the use of the aug-cc-pVTZ or aug-pc-2 basis sets is not recommended in the current MC−DFT approach.

Concluding Remarks

In the current study, we have demonstrated that for density functional theory calculation, significant performance improvement on molecular energies can be achieved by using two or three different basis sets in the so-called MC−DFT approach. Although the current study does not mean to make exhaustive tests using all known functionals and on all types of molecular energies, our results clearly show that the atomization energies and the barrier heights, which are the most important properties for thermochemistry and thermochemical kinetics, respectively, can be predicted much more accurately than the conventional way of using a single basis set. We found that the best functional for the MC−DFT approach is the B1B95 functional with adjustable mixing percentage of the exact exchange energy. From the current study the recommended MC−DFT methods for predicting the atomization energies are the B1B95-AE and B98 functionals with the cc-pVDZ/cc-pVTZ basis set combination. The recommended MC−DFT methods for predicting the barrier heights are the B1B95-BH with the cc-pVDZ/cc-pVTZ or cc-pVDZ/cc-pVTZ/aug-cc-pVDZ combinations. For overall performance (109 AEs, 76 BHs, 13 IPs, 13 EAs), the recommended MC−DFT method is the B1B95-All functionals with the cc-pVDZ/cc-pVTZ/aug-cc-pVTZ combination which performs almost as well as the extensively optimized M06-2X functional with the MG3S basis set. It remains to be seen how the MC−DFT approach performs on the various types of weak interactions[13,28,38,39] and on molecules containing transition metals.[38−44] The most important finding in the current study is that, in a way, the DFT calculation can be improved systematically by using more than one basis sets. It also remains to be seen whether the design of new functionals can benefit from the MC−DFT approach. That is, if empirical parameters are included in the functionals, one may also attempt to optimize the coefficients in eq 2 and 3 simultaneously with the empirical parameters. Since the MC−DFT energy is a linear combination of two or three DFT energies, the implementations of energy gradients for geometry optimization and hessians for frequency calculation are relatively straightforward.[45]

Acknowledgment. This work is supported by the National Science Council of Taiwan, Grant No. NSC-95-2113-M-194-007. We are grateful to the National Center for High-Performance Computing (NCHC) for providing part of computing resources.

Supporting Information Available: Tables of performance on AEs and BHs using the cc-pVDZ/cc-pVTZ/aug-cc-pVDZ basis set combination. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864.
(2) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.
(3) Laird, B. B.; Ross, R. B.; Ziegler, T. In *Chemical Application of Density-Functional Theory*; Laird, B. B., Ross, R. B., Ziegler, T., Eds.; ACS Symposium Series 629; American Chemical Society: Washington, DC, 1996; pp 1−11.
(4) Scuseria, G. E.; Staroverov, V. N. In *Theory and Applications of Computational Chemistry: The First Forty Years*; Dykstra, C. E., Frenking, G., Kim, K. S., Scuseria, G. E., Eds.; Elsevier B. V.: Amsterdam, The Netherlands, 2005; pp 669−724.
(5) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (b) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.
(6) Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664.
(7) Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040.
(8) Schmider, H. L.; Becke, A. D. *J. Chem. Phys.* **1998**, *108*, 9624.
(9) Boese, A. D.; Martin, J. M. L. *J. Chem. Phys.* **2004**, *121*, 3405.
(10) Bauschlicher, C. W., Jr.; Partridge, H. *Chem. Phys. Lett.* **1995**, *240*, 533.
(11) Bauschlicher, C. W., Jr. *Chem. Phys. Lett.* **1995**, *246*, 40.
(12) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 6908.
(13) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 5656.
(14) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **2005**, *123*, 124107.
(15) Boese, A. D.; Martin, J. M. L.; Handy, N. C. *J. Chem. Phys.* **2003**, *119*, 3005.
(16) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 3898.
(17) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 1125.
(18) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2001**, *114*, 108.
(19) Lee, T.-H.; Chen, H.-R.; Hu, W.-P. *Chem. Phys. Lett.* **2005**, *412*, 430.
(20) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2005**, *7*, 43.
(21) Sun, Y.-L.; Lee, T.-H.; Chen, J.-L.; Wu, K.-J.; Hu, W.-P. *Chem. Phys. Lett.* **2007**, *442*, 220.

(22) (a) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007. (b) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796. (c) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 1358.

(23) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

(24) Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040.

(25) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(26) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(27) Toulouse, T.; Savin, A.; Adamo, C. *J. Chem. Phys.* **2002**, *117*, 10465.

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

(29) Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 1384.

(30) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1999**, *110*, 4703.

(31) Fast, P. L.; Corchado, J.; Sanchez, M. L.; Truhlar, D. G. *J. Phys. Chem. A* **1999**, *103*, 3139.

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

(33) Zhao, Y.; González-García, N.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 2012.

(34) (a) Jensen, F. *J. Chem. Phys.* **2001**, *115*, 9113. (b) Jensen, F. *J. Chem. Phys.* **2002**, *116*, 7372. (c) Jensen, F. *J. Chem. Phys.* **2002**, *117*, 9234.

(35) Schwenke, D. W. *J. Chem. Phys.* **2005**, *122*, 014107.

(36) Lynch, B. J.; Fast, P. L.; Harris, M.; Truhlar, D. G. *J. Phys. Chem. A* **2000**, *104*, 4811.

(37) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 2715.

(38) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, in press.

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

(40) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 11127.

(41) (a) Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *124*, 224105. (b) Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 194101.

(42) Furche, F.; Perdew, J. P. *J. Chem. Phys.* **2006**, *124*, 044103.

(43) Jensen, K. P.; Roos, B. O.; Ryde, U. *J. Chem. Phys.* **2007**, *126*, 014103.

(44) Quintal, M. M.; Karton, A.; Iron, M. A.; Boese, A. D.; Martin, J. M. L. *J. Phys. Chem. A* **2006**, *110*, 709.

(45) Chuang, Y.-Y.; Truhlar, D. G. *J. Phys. Chem. A* **1999**, *103*, 651.