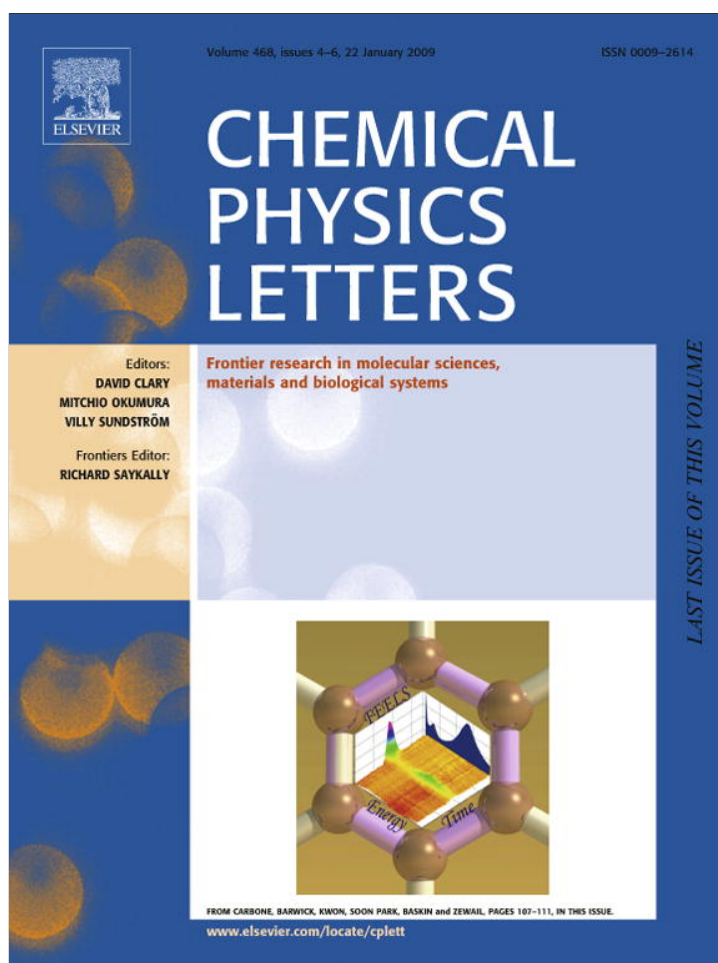


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Chemical Physics Letters

journal homepage: www.elsevier.com/locate/cplett

The MC-DFT approach to the M06-2X, B2K-PLYP, and B2T-PLYP functionals

Jien-Lian Chen, Jie-Ting Hong, Kuo-Jui Wu, Wei-Ping Hu *

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

ARTICLE INFO

Article history:

Received 2 October 2008

In final form 9 December 2008

Available online 24 December 2008

ABSTRACT

We have tested the performance of three recently developed density functionals, including M06-2X, B2K-PLYP, and B2T-PLYP, on thermochemical kinetics using the multi-coefficient density functional theory (MC-DFT). The results indicated that in most cases, the accuracy can be significantly improved by using more than one basis set. Compared with using a single large basis set, the atomization energies can be predicted more accurately at a lesser or equal cost, and the same level of accuracy can be reached using less expensive basis set combinations. The three combinations pdz/MG3S, pdz/ptz/apdz, and pc1/pc2/apc1 are especially attractive in this regard.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Density functional theory (DFT) [1–4] has been the most popular theoretical method for chemists in the past decade because of its high performance-to-cost ratio for many different types of chemical systems. Most of the recently developed density functionals were based on the very successful generalized gradient approximation (GGA) with hybrid exchange functionals. In some new DFT methods (the so-called ‘meta’ DFT [5–8]), the exchange–correlation functionals also include contribution from the kinetic energy density. To achieve higher accuracy, many new types of functionals have also been developed semi-empirically by fitting some of the parameters that define the functionals to best reproduce experimental data. Notably, the recently developed and highly successful M05 [9,10] and M06 series [11–13] of functionals are of this type.

The so-called multi-coefficient methods were first developed by Truhlar and coworkers [14,15]. The basic idea with these methods is to place scaling coefficients before energy correction terms in multi-level wavefunction-based electronic structure methods, such as the Gaussian-2 [16] and Gaussian-3 methods [17]. From there, a series of studies [14,15,18,19] recognized that by using scaled energy components in the multi-level methods, higher accuracy and sometimes higher efficiencies can be achieved. They then proposed to generalize their multi-coefficient methods by combining the wavefunction-based and the density functional methods to create more cost-effective methods. In particular, they developed two so-called ‘doubly hybrid meta DFT’ methods, the MC3BB and MC3MPW [20]. In these methods, the energy is expressed as a linear combination of the energies calculated at the Hartree-Fock, MP2, and hybrid-meta DFT levels. Three coefficients

in these methods were determined empirically against accurate thermochemical kinetics data. The DFT used in the MC3BB is based on the hybrid version of Becke88 exchange functional [21] and the B95 correlation functional [22], and the DFT used in the MC3MPW is based on the hybrid version of the modified Perdew–Wang exchange functionals [23] and the PW91 correlation functional [24]. They showed that the MC3BB was more accurate than any other method of comparable cost. Later, they further include higher-level wavefunction-based correlation energies up to QCISD(T) level into their multi-coefficient method. For example, in their MCG3-MPWB method [25], 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 multi-coefficient hybrid methods called MLSE-DFT [26] using Dunning’s correlation-consistent basis sets [27] for neutral systems. The average errors on a set of 169 molecular energies by MCG3-DFT and MLSE-DFT methods were as low as 0.6–0.7 kcal/mol.

Recently, Grimme and coworkers have developed a related hybrid method, which they called ‘double-hybrid exchange–correlation functionals’ [28,29]. In these functionals, the correlation energy calculated at the MP2 level based on the Kohn–Sham orbitals was added to the correlation part of a hybrid exchange–correlation functional. For example, the B2-PLYP and MPW2-PLYP functionals were developed using the Lee–Yang–Parr correlation functional [30] with the hybrid version of the Becke88 [21] and the modified Perdew–Wang exchange functionals [23], respectively. The ‘2’ signifies that there are two empirical parameters, which are the mixing ratios (or coefficients) of the exact exchange energy in the exchange functional and the MP2 correlation energy (E2) in the correlation functional. Very encouraging results were obtained using these methods on the G3/05 test set [31]. More recently, Martin and coworker re-parametrized these methods based

* Corresponding author. Fax: +886 5 2721040.

E-mail address: chewph@ccu.edu.tw (W.-P. Hu).

on the AE6 [32] and W3 training set [33] for thermochemistry and the BH6 set [32] for representative barrier heights. They developed the B2T-PLYP, B2K-PLYP, and MPW2K-PLYP double-hybrid functionals [34]. The first of these was designed for thermochemical study and the other two for thermochemical kinetics. Their test results showed that these functionals give very good performance for a large variety of molecular energies. A related double-hybrid approach was recently used by Head-Gordon and coworkers [35] to improve the DFT calculation on long-range interactions.

Very recently, our group developed a set of multi-coefficient DFT (MC-DFT) [36] methods. In these methods, the DFT calculation is carried out using two or three basis sets of different sizes. The electronic or Born–Oppenheimer energy is calculated as

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

or

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

where B2 and B3 are basis sets larger in sizes than that of B1. The coefficients c_1 , c_2 , and the mixing coefficient of the exact exchange energy in the hybrid exchange functional were determined to minimize the mean unsigned errors (MUEs) with respect to a set of 211 accurate database values. For functionals such as B1B95, B98 [37], and MPW1PW91, the MC-DFT approach can improve the accuracy significantly (as compared to the results using a single basis set) with very modest increase in computational cost [36]. In many cases, the MC-DFT results are even better than the results obtained using a much larger basis set. Our goal in the current study is to test the MC-DFT approach on the three recently developed functionals, M06-2X, B2K-PLYP, and B2T-PLYP. We would like to determine whether the MC-DFT approach can be applied to these new functionals to obtain even more accurate or more efficient multi-coefficient versions of these functionals.

It is noted that the current MC-DFT approach and the basis-set extrapolation parts of other multi-coefficient methods, such as MCG3-DFT [25], G3S [38], MLSE-DFT [26], etc., all shared the same formalism with the basis-set extrapolation scheme by Schwenke [39], that the coefficients of the basis-set extrapolation terms are not derived from theory, but are rather determined empirically based on experimental data or energies from accurate calculation.

2. Method

The basis sets used in the tests included the MG3S [40], aug-cc-pVTZ (aptz) [27], aug-pc-2 (apc2), aug-pc-3 (apc3) [41], and the combinations of cc-pVDZ/MG3S (pdz/MG3S), cc-pVDZ/cc-pVTZ/aug-cc-pVDZ (pdz/ptz/apdz), cc-pVDZ/aug-cc-pVDZ/aug-cc-pVTZ (pdz/apdz/aptz), pc-1/pc-2/aug-pc-1 (pc1/pc2/apc1), pc-1/aug-pc-1/aug-pc-2 (pc1/apc1/apc2). The abbreviations in parentheses will be used in the rest of the Letter. All the empirical parameters in the functionals are in their original values without modification. The MC-DFT coefficients c_1 , c_2 , in Eqs. (1) and (2) were determined by minimizing the MUEs with respect to a set of 211 accurate thermochemical kinetics data (the ‘training set’) including 109 atomization energies (AEs) from the MGAE109/05 database [10], 38 hydrogen transfer barrier heights (HTBHs) and 38 non-hydrogen transfer barrier heights (NHTBHs) from the HTBH38/04 [10] and NHTBH38/04 [42] databases, and 13 ionization potentials and 13 electron affinities from the IP13/3 [10] and EA13/3 databases [10]. These databases were compiled by Truhlar and coworkers and were also used to determine the MC-DFT coefficients in a previous work [36]. Spin-orbit corrections for the selected species [14] were included in the calculation for the training set. The resulting methods were also tested on a set of 8 proton affinities from the

PA8 database [10,13], a set of 4 alkyl bond dissociation energies from the ABDE4 database [10,13], a set 10 difficult energies from the DC10 databases [13], and a set of 31 AEs from the training set of the W3 thermochemistry method reported by Martin and coworkers [33]. The PA8, ABDE4, and DC10 databases were also developed by Truhlar and coworkers, and they were used as more sensitive test sets in the current study. All the electronic structure calculations were performed using a locally modified version of the GAUSSIAN 03 program [43].

3. Results and discussion

Table 1 shows the MUEs obtained using the three functionals on the training set of 211 data. For the M06-2X functional, the MG3S and apc2 basis sets give impressive MUEs of 1.68 and 1.60 kcal/mol, respectively. The pdz/MG3S and pdz/ptz/apdz combination give even more impressive results of 1.47 and 1.46 kcal/mol, respectively. The pc1/pc2/apc1 combination also gives a similarly small MUE of 1.51 kcal/mol. It should be mentioned that the M06-2X functionals have already been highly optimized using the MG3S basis set. Thus, the room for improvement using the MC-DFT approach is expected to be smaller than some of the more traditional DFTs such as B1B95 and B98, as shown in a previous study [36]. For the B2K-PLYP functional, the apc2 performs (MUE = 1.85 kcal/mol) significantly better than the MG3S basis set (MUE = 4.13 kcal/mol). For some reason, the B2K-PLYP/MG3S calculation performs poorly on the 109 AEs. Interestingly, using the pdz/MG3S combination reduces the MUE obtained using the MG3S alone by 50%. The pc1/apc1/apc2 and pc1/pc2/apc1 combinations can further lower the MUE to 1.60 and 1.65 kcal/mol, respectively. The pdz/ptz/apdz combination also performs very well with an MUE of 1.62 kcal/mol. The B2T-PLYP functional performs similarly to the B2K-PLYP. The pc1/apc1/apc2 combination significantly reduces the MUE obtained using the apc2 basis set by 0.4 kcal/mol. Interestingly, using the pc1/apc1/apc2 combination, the B2T-PLYP performs even better than the B2K-PLYP, even though the former was based primarily on thermochemistry. In addition to potentially having higher accuracy, the MC-DFT approach in most cases can reach a particular level of accuracy with a much smaller computational cost. As shown in Table 1, using the pc1/pc2/apc1 combination usually gives smaller MUEs than using the apc2 basis set. The quantitative comparison of computational cost will be presented at the end of this section.

Tables 2–4 show the performance of the M06-2X, B2K-PLYP, and B2T-PLYP functionals, respectively, on various types of energies. For the M06-2X, the improvement using the MC-DFT approach is mainly due to the reduction of MUE on the atomization energies. For example, the MG3S basis set gives an MUE of 1.88 kcal/mol on the 109 AEs while the pdz/MG3S combination gives an MUE of 1.50 kcal/mol. Similarly, the ptz and pc2 basis sets give MUEs of 2.49 and 2.07 kcal/mol (data not shown), respectively, while the pdz/ptz/apdz and pc1/pc2/apc1 combinations give

Table 1
Mean unsigned errors (kcal/mol) for the training set.

	M06-2X	B2K-PLYP	B2T-PLYP
MG3S	1.68 (2.21) ^a	4.13 (5.31)	3.69 (4.88)
aptz	1.89 (2.62)	2.05 (3.04)	2.34 (3.36)
apc2	1.60 (2.08)	1.85 (2.52)	1.94 (2.51)
pdz/MG3S	1.47 (1.97)	2.05 (2.80)	1.88 (2.50)
pdz/ptz/apdz	1.46 (2.06)	1.62 (2.33)	1.74 (2.46)
pdz/apdz/aptz	1.55 (2.20)	1.76 (2.60)	1.82 (2.67)
pc1/pc2/apc1	1.51 (2.05)	1.65 (2.38)	1.67 (2.27)
pc1/apc1/apc2	1.54 (2.06)	1.60 (2.22)	1.54 (2.03)

^a Numbers in parentheses are the RMS errors.

Table 2
Mean unsigned errors (kcal/mol) on various types of energies by the M06-2X functional.

	AE	IP	EA	HTBH	NHTBH	MUE ^a	PA8	ABDE	DC10	W3
MG3S	1.88	2.54	2.07	1.13	1.22	1.68 (2.21) ^b	1.64	0.71	10.78	2.43 (4.10) ^b
aptz	2.28	2.70	1.55	1.14	1.37	1.89 (2.62)	2.06	0.79	13.45	2.90 (4.97)
apc2	1.80	2.39	1.40	1.09	1.34	1.60 (2.08)	1.75	0.79	9.54	2.23 (3.68)
apc3	1.71	2.37	1.52	1.00	1.33	1.54 (1.98)	1.67	0.82	NA ^c	2.21 (3.57)
pdz/MG3S	1.50	2.62	1.65	1.05	1.32	1.47 (1.97)	1.76	0.69	9.57	2.06 (3.75)
pdz/ptz/apdz	1.47	3.05	1.95	0.97	1.21	1.46 (2.06)	2.42	0.98	9.62	2.18 (3.99)
pdz/apdz/aptz	1.66	2.69	1.29	1.10	1.37	1.55 (2.20)	2.03	0.95	12.02	2.25 (4.33)
pc1/pc2/apc1	1.73	2.23	1.53	0.95	1.19	1.51 (2.05)	1.68	0.79	9.17	1.96 (3.47)
pc1/apc1/apc2	1.72	2.25	1.34	1.03	1.34	1.54 (2.06)	1.65	0.82	9.62	2.00 (3.38)

^a Mean unsigned errors for the training set of 211 data.

^b Numbers in parentheses are the RMS errors.

^c Not calculated due to the resource limitation.

Table 3
Mean unsigned errors (kcal/mol) on various types of energies by the B2K-PLYP functional.

	AE	IP	EA	HTBH	NHTBH	MUE ^a	PA8	ABDE	DC10	W3
MG3S	6.58	1.68	3.30	0.93	1.44	4.13 (5.41) ^b	0.52	2.14	13.77	4.59 (5.13) ^b
aptz	2.82	1.60	2.26	0.69	1.31	2.05 (3.04)	0.62	1.38	12.52	3.08 (4.27)
apc2	2.35	1.85	2.66	0.70	1.27	1.85 (2.52)	0.85	1.01	6.80	2.73 (3.49)
apc3	1.59	1.57	1.94	0.85	1.21	1.41 (1.97)	0.84	1.73	NA ^c	1.64 (2.44)
pdz/MG3S	2.19	1.92	3.82	1.19	1.94	2.05 (2.80)	1.03	2.01	8.51	2.14 (2.90)
pdz/ptz/apdz	1.80	1.83	1.62	0.98	1.70	1.62 (2.33)	1.12	1.68	6.36	2.04 (2.96)
pdz/apdz/aptz	2.30	1.68	1.16	0.75	1.49	1.76 (2.60)	0.75	1.34	10.04	2.33 (3.49)
pc1/pc2/apc1	1.73	2.25	2.16	0.89	1.84	1.65 (2.38)	1.62	1.06	7.53	1.74 (2.46)
pc1/apc1/apc2	1.92	1.85	1.34	0.74	1.56	1.60 (2.22)	1.00	0.92	6.57	2.08 (2.78)

^a Mean unsigned errors for the training set of 211 data.

^b Numbers in parentheses are the RMS errors.

^c Not calculated due to the resource limitation.

Table 4
Mean unsigned errors (kcal/mol) on various types of energies by the B2T-PLYP functional.

	AE	IP	EA	HTBH	NHTBH	MUE ^a	PA8	ABDE	DC10	W3
MG3S	5.71	2.05	2.87	0.95	1.47	3.69 (4.88) ^b	0.53	3.72	14.65	3.07 (3.64) ^b
aptz	3.08	1.93	1.91	1.25	1.62	2.34 (3.36)	0.65	3.22	14.63	2.38 (3.59)
apc2	2.30	2.14	2.15	1.22	1.50	1.94 (2.51)	0.83	2.93	8.49	1.94 (2.54)
apc3	1.60	1.92	1.63	1.10	1.38	1.49 (1.90)	0.92	5.29	NA ^c	1.38 (1.84)
pdz/MG3S	1.98	2.30	4.05	0.93	1.66	1.88 (2.50)	0.99	3.82	4.86	1.83 (2.40)
pdz/ptz/apdz	1.81	2.21	1.49	1.19	2.01	1.74 (2.46)	1.01	3.53	7.70	1.55 (2.26)
pdz/apdz/aptz	2.15	1.99	1.33	1.24	1.54	1.82 (2.67)	0.76	3.18	10.66	1.84 (2.80)
pc1/pc2/apc1	1.67	2.57	2.31	1.19	1.66	1.67 (2.27)	1.48	2.99	5.70	1.51 (1.99)
pc1/apc1/apc2	1.56	2.14	1.45	1.26	1.59	1.54 (2.03)	0.96	2.80	5.44	1.52 (1.94)

^a Mean unsigned errors for the training set of 211 data.

^b Numbers in parentheses are the RMS errors.

^c Not calculated due to the resource limitation.

MUEs of 1.47 and 1.73 kcal/mol, respectively. For the W3 test set, which contains 31 AEs, the MG3S and pc2 give MUEs of 2.43 and 2.60 kcal/mol, respectively, while the pdz/MG3S and pc1/pc2/apc1 combinations give MUEs of 2.06 and 1.96 kcal/mol, respectively. The prediction of electron affinity also benefits from the MC-DFT approach. For example, the MG3S gives an MUE of 2.07 kcal/mol for the 13 EAs while the MG3S/pdz combination gives an MUE of 1.65 kcal/mol. A similar improvement is observed for the pc1/pc2/apc1 combination. Some of the combinations also show significant improvement on the barrier heights. For example, the ptz basis set gives a MUE of 1.33 kcal/mol for the HTBH38 while the pdz/ptz/apdz combination gives a MUE of 0.97 kcal/mol. Also, the apdz and pc2 basis sets give MUEs of 1.62 and 1.55 kcal/mol, respectively for the NHTBH38 while the pdz/ptz/apdz and pc1/pc2/apc1 combinations give MUEs of 1.21 and

1.19 kcal/mol, respectively. The MC-DFT results for the IP13, PA8, ABDE4, and DC10 are similar to those obtained using a single basis set. Using the large apc3 basis set with the M06-2X functional improves the apc2 results only slightly and is not more accurate than using some of the basis set combinations.

In Table 3 we see that for the B2K-PLYP functional, the improvement using the MC-DFT approach on AEs is also very significant. For example, using the MG3S basis set gives a very disappointed MUE of 6.58 kcal/mol on the 109 AEs while the pdz/MG3S combination dramatically improves the results to 2.19 kcal/mol. Similarly for the 109 AEs using the ptz and pc2 basis sets give MUEs of 5.25 and 4.18 kcal/mol, respectively, while the pdz/ptz/apdz and pc1/pc2/apc1 combinations give MUEs of 1.80 and 1.73 kcal/mol, respectively. To predict reasonable AEs using the B2K-PLYP functionals with a single basis set, a triple-zeta quality basis set with diffuse

functions (such as aptz, apc2) is clearly required. However, using combinations such as pdz/ptz/apdz and pc1/pc2/apc1 can provide significantly better results. For the W3 test set, the MG3S and pc2 give MUEs of 4.59 and 4.12 kcal/mol, respectively, while the pdz/MG3S and pc1/pc2/apc1 combinations give MUEs of 2.14 and 2.04 kcal/mol, respectively. The prediction of EA again can benefit from the MC-DFT approach for some of the basis set combinations. For example, the apdz basis set gives an MUE of 3.74 kcal/mol for the 13 EAs while the pdz/ptz/apdz combination gives an MUE of 1.62 kcal/mol. A similar improvement is observed for the pc1/pc2/apc1 combination. For the B2K-PLYP functional, the MC-DFT results for the IP, PA, HTBH, NHTBH, and ABDE sets are similar to those obtained using a single basis set. For the DC10 set, the apc2 basis set gives a MUE of 6.80 kcal/mol, which is significantly better than using any basis sets or basis set combinations with the M06-2X functional (MUEs = ~10 kcal/mol). The combinations of pdz/ptz/apdz and pc1/pc2/apc1 also give similar results, which are significantly better than using a single basis set. This suggests that including the second-order correction makes the B2K-PLYP a more robust method towards some difficult cases. Using the large apc3 basis set with the B2K-PLYP functional improves the results quite significantly. The MUE of the training set is now 1.41 kcal/mol which is compared to the MUE of 1.85 kcal/mol using the apc2 basis set and the MUEs of ~1.6 using the best basis set combinations. The MUE to the W3 test set is now 1.64 kcal/mol which is compared to the MUE of 2.73 kcal/mol using the apc2 basis set and the MUE of 1.74 using the best basis set combination. This is in contrast to the M06-2X functional where only very limited improvement is observed using the apc3 basis set. Using the pc1/pc2/apc3 or the pc1/apc1/apc3 basis set combinations can only lower the MUE by an additional ~0.1 kcal/mol.

Table 4 shows that the B2T-PLYP functional performs slightly better than the B2K-PLYP on the 109 AEs. This is expected since it is parameterized for thermochemistry. Similar to B2K-PLYP, the improvement using the MC-DFT approach on AEs is also very significant. For the W3 test set, the apc2 basis set gives an MUE of 1.94 kcal/mol, which is significantly better than the MUEs of 2.23 and 2.73 kcal/mol obtained using the M06-2X and B2K-PLYP functionals, respectively. The pdz/ptz/apdz, pc1/pc2/apc1, and pc1/apc1/apc2 combinations give very low MUEs of 1.55, 1.51, and 1.52 kcal/mol, respectively. For the HTBH38 and NHTBH38 sets, the B2T-PLYP is less accurate than the B2K-PLYP functional which was designed to achieve good accuracy on barrier heights. In most cases, however, the differences are not very significant. For the DC10 set, the MC-DFT performance is even better than the B2K-PLYP functionals. In particular, the pdz/MG3S, pc1/pc2/apc1, and pc1/apc1/apc2 combinations give very impressive MUEs of 4.86, 5.70, and 5.44 kcal/mol, respectively. For the ADBE set, the performance of M06-2X is much better than B2K-PLYP, and the performance of B2T-PLYP is less satisfactory.

It is well-known that second-order correction energies converge relatively slowly with respect to the basis-set size [34] as compared to the Hartree-Fock or DFT energies. Thus, it is desirable to investigate whether using separate scaling coefficients for DFT and E2 energies would improve the performance. That is, Eqs. (1) and (2) can be rewritten as:

$$E'_{2B} = E(\text{DHDFT}/\text{B1}) + c_1 \{ [E(\text{HDFT}/\text{B2}) - E(\text{HDFT}/\text{B1})] + c'_1 [E(\text{E2}/\text{B2}) - E(\text{E2}/\text{B1})] \} \quad (3)$$

$$E'_{3B} = E(\text{DHDFT}/\text{B1}) + c_1 \{ [E(\text{HDFT}/\text{B2}) - E(\text{HDFT}/\text{B1})] + c'_1 [E(\text{E2}/\text{B2}) - E(\text{E2}/\text{B1})] \} + c_2 \{ [E(\text{HDFT}/\text{B3}) - E(\text{HDFT}/\text{B1})] + c'_2 [E(\text{E2}/\text{B3}) - E(\text{E2}/\text{B1})] \} \quad (4)$$

For calculation using a single basis set, the energy can be written as:

$$E'_{1B} = E(\text{HDFT}/\text{B1}) + c'_0 E(\text{E2}/\text{B1}) \quad (5)$$

where the $E(\text{DHDFT}/\text{B1})$ refers to the energy calculated using one of the double-hybrid DFT methods using the B1 basis set while the $E(\text{HDFT}/\text{B1})$, $E(\text{HDFT}/\text{B2})$, and $E(\text{HDFT}/\text{B3})$ refer to the energies calculated by the corresponding hybrid DFT (DHDFT without E2) using the B1, B2, and B3 basis sets, respectively. The $E(\text{E2}/\text{B1})$, $E(\text{E2}/\text{B2})$, and $E(\text{E2}/\text{B3})$ represent the second-order correction energies calculated using the B1, B2, and B3 basis sets, respectively. In the multi-coefficient double-hybrid methods mentioned in Tables 1–4, the coefficients c'_0 , c'_1 and c'_2 were set to the original E2 mixing coefficients of the double-hybrid DFT method. In Eqs. (3)–(5), these values were also allowed to be optimized against the data in the training set. These new methods are similar to the MP2-SAC methods described by Truhlar and coworkers [14,15]. Table 5 shows the MUEs of the training set of the two double-hybrid DFT methods using various basis sets and basis set combinations. The reduction of the MUEs relative to the values in Table 1 is also presented. Using separate scaling coefficients for E2 results in a relatively wide range of improvement. The most dramatic improvement is found in calculation using a single MG3S basis set. For example, the MUEs obtained using the B2K-PLYP/MG3S and B2T-PLYP/MG3S methods were reduced from 4.13 and 3.69 kcal/mol to 1.73 and 1.78 kcal/mol, respectively. These values are now similar to those obtained using the apc2 basis set. Furthermore, the two methods just mentioned give significantly better results on the DC10 set. The pdz/MG3S combination also shows significant improvement of 0.2–0.5 kcal/mol in MUEs. For B2K-PLYP with pdz/MG3S combination, the MUE on DC10 is now 5.63 kcal/mol, compared to 8.51 kcal/mol in Table 3. It is apparent that the disappointing performance of the MG3S basis set with the double-hybrid DFT in Tables 1, 3 and 4 is due to the insufficient E2 energy. Adding an additional ~10% of E2 energy improves the results significantly. In Table 5, the pdz/apdz/aptz combination also shows important improvement of ~0.3 kcal/mol in MUEs, and this makes it the most accurate combination for B2K-PLYP. On the other hand, the improvement with the pc1/pc2/apc1 and pc1/apc1/apc2 combinations is relatively small (~0.1 kcal/mol).

Some authors may prefer using the root-mean-square errors (RMSEs) in reporting the accuracy. Thus we have also included

Table 5

Performance on the training set using separate scaling coefficients for hybrid DFT and E2 energy.

	MUE (kcal/mol)	Improvement ^a
B2K-PLYP		
MG3S	1.73 (2.34) ^b	2.40 (3.07)
aptz	1.88 (2.81)	0.17 (0.23)
apc2	1.72 (2.41)	0.13 (0.11)
pdz/MG3S	1.59 (2.23)	0.46 (0.57)
pdz/ptz/apdz	1.56 (2.18)	0.06 (0.15)
pdz/apdz/aptz	1.50 (2.07)	0.26 (0.53)
pc1/pc2/apc1	1.58 (2.33)	0.07 (0.05)
pc1/apc1/apc2	1.59 (2.20)	0.01 (0.02)
B2T-PLYP		
MG3S	1.78 (2.28)	1.91 (2.60)
aptz	1.95 (2.90)	0.39 (0.46)
apc2	1.68 (2.22)	0.26 (0.29)
pdz/MG3S	1.62 (2.09)	0.26 (0.31)
pdz/ptz/apdz	1.63 (2.33)	0.11 (0.13)
pdz/apdz/aptz	1.61 (2.21)	0.21 (0.46)
pc1/pc2/apc1	1.64 (2.21)	0.03 (0.06)
pc1/apc1/apc2	1.51 (1.97)	0.03 (0.06)

^a The improvement is relative to the values in Table 1.

^b Numbers in parentheses refer to the RMS errors.

Table 6
Computational cost^{a,b,c}.

	M06-2X	B2K-PLYP
MG3S	11.6	18.3
pdz/MG3S	12.6	19.5
ptz	11.0	14.0
aptz	58.9	82.0
pdz/ptz/apdz	16.1	19.5
pc2	11.3	13.9
apc2	63.0	87.4
apc3	1670	1970
pc1/pc2/apc1	16.0	19.4

^a Tested using computers with Intel E6600 CPUs and a version of GAUSSIAN 03 program optimized for the hardware architecture.

^b Total CPU time relative to MP2/6-31+G(d,p) calculation for C₅H₅N, C₂Cl₄, C₄H₄O, C₄H₄S, C₄H₅N, CF₃CN, and SiCl₄.

^c The cost of the MCG3-MPWB and MLSE-TPSS1KCIS methods is 32.8 and 34.4, respectively.

the RMSEs in Tables 1–5 for comparison. As one can see, in almost all cases, the MUEs and RMSEs follow the same trends.

Table 6 shows the relative computational cost of the M06-2X and B2K-PLYP functionals using various basis sets and basis set combinations for seven medium-sized molecules. The cost of the very accurate multi-level methods MCG3-MPWB [25] and MLSE-TPSS1KCIS [26] is also included for comparison. For valence triple-zeta basis sets, the cost of B2K-PLYP is 30–50% higher than that of M06-2X. The cost of using the MG3S, ptz, and pc2 is similar, as is the cost of the basis set combinations using these basis sets as the largest components. The cost of using the aptz or apc2 basis sets is 4–5 times higher. The cost of using the apc3 basis set is approximately 20 times higher than that of apc2 and is approximately 100 times higher than that of the pdz/ptz/apdz (or pc1/pc2/apc1) combination. Although the apc3 basis set gives very impressive results with the B2K-PLYP and B2T-PLYP functionals, its use on large molecules is impractical. As can be seen in Tables 1–5, some of the results using the apc2 or aptz basis sets in the combinations are reasonably good, but they are only marginally better than much less expensive combinations such as pdz/MG3S, pdz/ptz/apdz, and pc1/pc2/apc1. As seen in Table 5, the cost of using aptz or apc2 basis sets is considerably higher than that of using the multi-level methods MCG3-DFT or MLSE-DFT which include very high-level correlation energies and give much better MUEs (0.6–0.7 kcal/mol). Thus, for small to medium-sized molecules, the use of these basis sets is also not well-justified. For larger molecules in energy-sensitive applications, the three cost-effective combinations mentioned above seem most plausible.

4. Concluding remarks

We have demonstrated that the multi-coefficient DFT (MC-DFT) approach can be applied to three recently developed DFT functionals with success. Some of the cost-effective basis set combinations such as pdz/MG3S, pdz/ptz/apdz, and pc1/pc2/apc1 are especially attractive. With the MC-DFT approach, all three functionals can reach very satisfactory 1.5–1.7 kcal/mol MUEs for the training set. By scaling the hybrid-DFT and the E2 parts of the energies separately for the double-hybrid functionals, very dramatic improvement can be obtained for the MG3S basis set in all cases, and significant improvement can also be obtained for the pdz/MG3S, pdz/ptz/apdz, and pdz/apdz/ptz combinations. In comparison to the results of using a single large basis set, the most important advantages of the MC-DFT approach are twofold. First, the atomization energies (the AE, DC10, and W3 sets), which are the most important data for thermochemistry, can be predicted more accurately at less than or equal cost. Second, the same level of accuracy

can be reached using much less expensive basis set combinations. These two properties make the MC-DFT approach especially attractive for practical application. The MC-DFT approach can also be applied to the ‘doubly hybrid meta DFT’ by Truhlar and coworkers, and this will be the subject for a future study. Preliminary results showed that the performance is similar to that of the B2K-PLYP in the current study. Very recently, Martin and coworkers have developed a new, more robust version of double-hybrid functional B2GP-PLYP [44], and Truhlar and coworkers have also developed two new hybrid meta functionals M08-HX and M08-SO [45]. It will be interesting to test whether the MC-DFT approach can improve the performance of these new functionals. The encouraging results obtained in the current and in a previous study [36] suggest that some of the MC-DFT methods might soon become methods of choice for the application on thermochemistry and thermochemical kinetics study. Furthermore, the development of new accurate semi-empirical functionals in the future may also take advantage of the MC-DFT approach for better performance.

Acknowledgments

This work is supported by the National Science Council of Taiwan, Grant number NSC-95-2113-M-194-007. The authors are grateful to Prof. Donald G. Truhlar for the suggestion and discussion on the M06-2X functional. We are also grateful to the National Center for High-Performance Computing (NCHC) for providing part of computational resources.

References

- [1] P. Hohenberg, W. Kohn, Phys. Rev. 136 (1964) B864.
- [2] W. Kohn, L. Sham, J. Phys. Rev. 140 (1965) A1133.
- [3] B.B. Laird, R.B. Ross, T. Ziegler, in: B.B. Laird, R.B. Ross, T. Ziegler, (Eds.), Chemical Application of Density-Functional Theory, ACS Symposium Series 629, American Chemical Society, Washington, DC, 1996, p. 1.
- [4] G.E. Scuseria, V.N. Staroverov, in: C.E. Dykstra, G. Frenking, K.S. Kim, G.E. Scuseria, (Eds.), Computational Chemistry: The First Forty Years, B.V. Elsevier, 2005, p. 669.
- [5] A.D. Boese, N.C. Handy, J. Chem. Phys. 116 (2002) 9559.
- [6] J. Tao, J.P. Perdew, V.N. Staroverov, G.E. Scuseria, Phys. Rev. Lett. 91 (2003) 146401.
- [7] V.N. Staroverov, G.E. Scuseria, J.P. Perdew, J. Chem. Phys. 119 (2003) 12129.
- [8] Y. Zhao, D.G. Truhlar, J. Phys. Chem. A 109 (2003) 5656.
- [9] Y. Zhao, N.E. Schultz, D.G. Truhlar, J. Chem. Phys. 123 (2005) 161103.
- [10] Y. Zhao, N.E. Schultz, D.G. Truhlar, J. Chem. Theory Comput. 2 (2006) 364.
- [11] Y. Zhao, D.G. Truhlar, J. Chem. Phys. 125 (2006) 194101.
- [12] Y. Zhao, D.G. Truhlar, J. Phys. Chem. A 110 (2006) 13126.
- [13] Y. Zhao, D.G. Truhlar, Theor. Chem. Acc. 120 (2008) 215.
- [14] P.L. Fast, J.C. Corchado, M.L. Sánchez, D.G. Truhlar, J. Phys. Chem. A 103 (1999) 3139.
- [15] P.L. Fast, J.C. Corchado, M.L. Sánchez, D.G. Truhlar, J. Phys. Chem. A 103 (1999) 5129.
- [16] L.A. Curtiss, K. Raghavachari, G.W. Trucks, J.A. Pople, J. Chem. Phys. 94 (1991) 7221.
- [17] L.A. Curtiss, K. Raghavachari, P.C. Redfern, V. Rassolov, J.A. Pople, J. Chem. Phys. 109 (1998) 7764.
- [18] P.L. Fast, D.G. Truhlar, J. Phys. Chem. A 104 (2000) 6111.
- [19] B.J. Lynch, D.G. Truhlar, J. Phys. Chem. A 107 (2003) 3898.
- [20] Y. Zhao, B.J. Lynch, D.G. Truhlar, J. Phys. Chem. A 108 (2004) 4786.
- [21] A.D. Becke, Phys. Rev. A 38 (1988) 3098.
- [22] A.D. Becke, J. Chem. Phys. 104 (1996) 1040.
- [23] C. Adamo, V.J. Barone, Chem. Phys. 108 (1998) 664.
- [24] J.P. Perdew, in: P. Ziesche, H. Eschig (Eds.), Electronic Structure of Solids '91, Akademie Verlag, Berlin, 1991, p. 11.
- [25] Y. Zhao, B.J. Lynch, D.G. Truhlar, Phys. Chem. Chem. Phys. 7 (2005) 43.
- [26] Y.-L. Sun, T.-H. Lee, J.-L. Chen, K.-J. Wu, W.-P. Hu, Chem. Phys. Lett. 442 (2007) 220.
- [27] T.H.J. Dunning Jr., Chem. Phys. 90 (1989) 1007.
- [28] S.J. Grimme, Chem. Phys. 124 (2006) 034108.
- [29] T. Schwabe, S. Grimme, Phys. Chem. Chem. Phys. 8 (2006) 4398.
- [30] C. Lee, W. Yang, R.G. Parr, Phys. Rev. B 37 (1988) 785.
- [31] L.A. Curtiss, P.C. Redfern, K.J. Raghavachari, Chem. Phys. 123 (2005) 124107.
- [32] B. Lynch, D.G. Truhlar, J. Phys. Chem. A 108 (2004) 1460.
- [33] A.D. Boese, M. Oren, O. Atasoylu, J.M.L. Martin, M. Kállay, J. Gauss, J. Chem. Phys. 120 (2004) 4129.
- [34] A. Tarnopolsky, A. Kartom, R. Sertchook, D. Vuzman, J.M.L. Martin, J. Phys. Chem. A 112 (2008) 3.

- [35] T. Benighaus, R.A. DiStasio, R.C. Lochan, J.-D. Chai, M.J. Head-Gordon, *Phys. Chem. A* 112 (2008) 2702.
- [36] J.-L. Chen, Y.-L. Sun, K.-J. Wu, W.-P. Hu, *J. Phys. Chem. A* 112 (2008) 1064.
- [37] H.L. Schmider, A.D. Becke, *J. Chem. Phys.* 108 (1998) 9624.
- [38] L.A. Curtiss, K. Raghavachari, P.C. Redfern, J.A. Pople, *J. Chem. Phys.* 112 (2005) 1125.
- [39] D.W. Schwenke, *J. Chem. Phys.* 122 (2005) 014107.
- [40] B.J. Lynch, Y. Zhao, D.G. Truhlar, *J. Phys. Chem. A* 107 (2003) 1384.
- [41] F.J. Jensen, *Chem. Phys.* 116 (2002) 7372.
- [42] Y. Zhao, N. González-García, D.G. Truhlar, *J. Phys. Chem. A* 109 (2005) 2012.
- [43] M.J. Frisch et al., *GAUSSIAN 03*, Revision E.01, Gaussian, Inc., Wallingford CT, 2007.
- [44] A. Karton, A. Tarnopolsky, J.-F. Lamère, G.C. Schatz, J.M.L. Martin, *J. Phys. Chem. A* 112 (2008) 12868.
- [45] Y. Zhao, D.G. Truhlar, *J. Chem. Theory Comput.* 4 (2008) 1849.