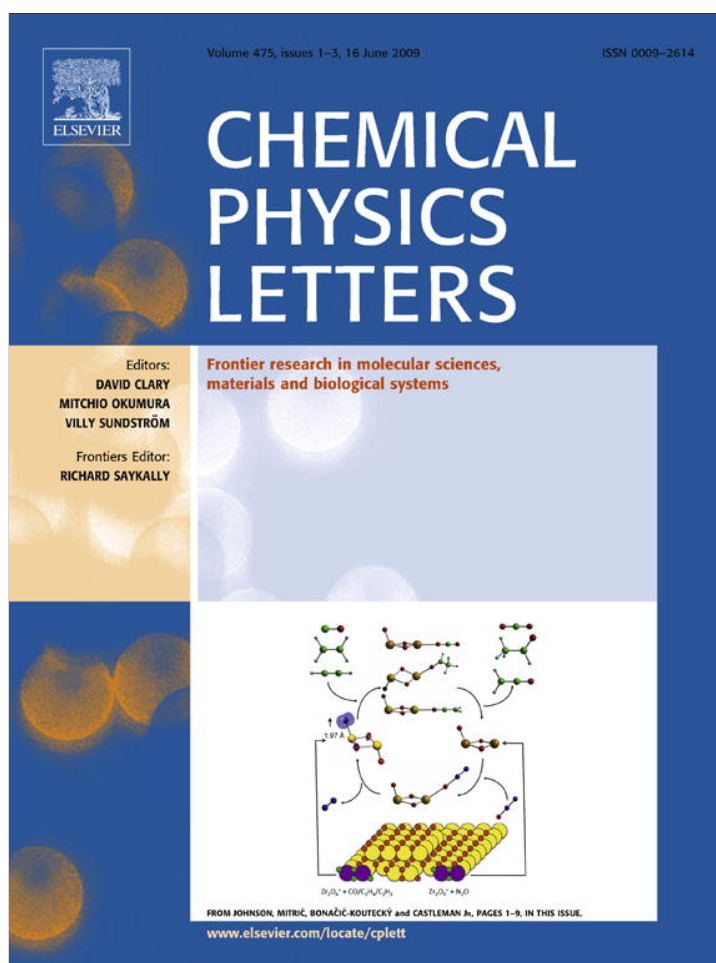


Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



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.

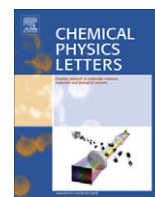
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

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



Contents lists available at ScienceDirect

## Chemical Physics Letters

journal homepage: [www.elsevier.com/locate/cplett](http://www.elsevier.com/locate/cplett)Accurate multi-coefficient electronic structure methods MLSE( $C_n$ )-DFT for thermochemical kinetics

Yi-Lun Sun, Tsung-Hui Li, Jien-Lian Chen, Wei-Ping Hu \*

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

## ARTICLE INFO

## Article history:

Received 26 December 2008

In final form 8 May 2009

Available online 13 May 2009

## ABSTRACT

We have developed a series of new multi-coefficient electronic structure methods, MLSE( $C_n$ )-DFT, that performed equally well on both neutral and charged systems. The lowest average mean unsigned error on 211 thermochemical kinetics data is 0.56 kcal/mol using the MLSE( $C_1$ )-M06-2X method. The simplified MLSE( $C_2$ )-M06-2X method can achieve similar accuracy at 54% of the computational cost. Therefore, it is the most recommended method. The highly simplified, but reasonably accurate, MLSE( $C_3$ )-B3LYP method is an economical alternative for larger systems.

© 2009 Elsevier B.V. All rights reserved.

## 1. Introduction

The exponential improvements in computer processing capabilities in the past two decades has allowed for more accurate quantum chemical methods to be routinely applied to all areas of chemical studies. A traditional goal in developing quantum chemical methods is to obtain relative molecular energies to within a so-called 'chemical accuracy' of  $\sim 1$  kcal/mol. Such accuracy can sometimes be achieved using very high-level theories, such as MP4(SDTQ), QCISD(T), CCSD(T), CCSDTQ with basis sets such as G3Large, aug-cc-pVTZ, aug-cc-pVQZ, or even larger basis sets. However, the computational costs for applying these methods are prohibitively high, except when modeling very small molecules. Alternatively, this goal can also be achieved more economically by using the so-called 'multi-level methods' and 'multi-coefficient methods' [1–25]. In these methods, additive corrections are applied to a base energy in order to account for the incomplete treatment of the correlation energies and the incompleteness of the basis-set sizes (and in some methods, the corrections to the relativistic effects and Born–Oppenheimer approximation are also included). The multi-level methods, for example, the Gaussian- $n$  [6,7], complete basis set (CBS) [8,9], Weizmann- $n$  theories [10–12], correlation-consistent composite approach (ccCA) [13,14], focal-point analysis (FPA) approach [15,16], HEAT protocol [17–19], the recent approach by Klopper et al. [20], and their various variants, have provided very accurate energies for thermochemistry and thermochemical kinetics. A series study by Truhlar et al. [3,4,21,22], Pople and co-workers [24], and Hu and co-workers [5,23] has demonstrated that by using scaled energy components in the multi-level methods, higher accuracy and sometimes higher efficiencies can be achieved. In these

so-called 'multi-coefficient methods', such as the G3S [24], G3S/3 [4], G3SX [25], those in MCCM/3 suite [4], and the MLSE methods [5,23], all of the scaling coefficients for the various energy components were optimized against databases of experimentally derived or high-level theoretical energies. This is in contrast to the above mentioned multi-level methods where no empirical derived coefficients were used for energy components except for only a few highest-level electron correlation terms and for vibrational zero-point and thermal energies in some of the methods. Recently, Truhlar et al. found that incorporating hybrid DFT energies into the multi-coefficient methods [26] can substantially increase the accuracy, with only a modest increase in the computational cost. For example, the mean unsigned error (MUE) to a thermochemical kinetics data set of 169 energies can be improved from 1.00 kcal/mol using the MCG3/3 method [4] to 0.73 kcal/mol using the MCG3-MPWB [26] method, with only 22% increase in computational cost [26,27]. We have followed this example in order to improve our own MLSE methods [5,23] by incorporating DFT energies into the so-called 'MLSE-DFT' methods [27]. By incorporating the additional DFT terms, the MUE was reduced by 0.22 kcal/mol. The basis sets we used in the MLSE methods were Dunning's correlation-consistent basis sets [28–30]. In contrast, a series of fine-tuned Pople-type basis sets were used for the MCG3 and G3S methods. Our choice was based on that Dunning's basis sets provide a very well defined hierarchy of basis sets. Furthermore, they were designed for basis-set extrapolation and for methods that consider high-level electron correlation. However, calculations using Dunning's correlation-consistent basis sets are usually less efficient than those using the Pople-type basis sets in some of the electronic structure programs that are unable to exploit general contractions. Furthermore, there are no intermediate basis sets between the valence triple-zeta and valence double-zeta basis sets. This limits the choices of the theoretical levels. Nevertheless, the MLSE-DFT methods successfully

\* Corresponding author. Fax: +886 5 272 1040.

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

predicted the atomization energies of neutral molecules and reaction energetics of neutral systems with a low mean unsigned error (MUE),  $\sim 0.6$  kcal/mol on 169 energies at modest computational costs. Extending the MLSE-DFT methods to treat the charged systems is conceptually straightforward. This process can be deduced from past experience and from comparison with the MCG3 and G3S methods. That is, an MP2 calculation with an augmented valence triple-zeta quality basis set (e.g., MP2/aug-cc-pVTZ) and an MP4 calculation with valence triple-zeta quality basis set (e.g., MP4SDQ/cc-pVTZ) are necessary. Although the direct addition of these two theoretical levels to the MLSE-DFT method would likely lead to accurate methods for both neutral and charged systems, both calculations are very time-consuming. In this Letter, we present a new set of all-purpose, improved MLSE-DFT methods that include or approximate the two additional theoretical levels with very good performance. As mentioned above, the use of Pople-type basis sets in multi-coefficient methods has been developed and extensively optimized in MCG3, MCG3/3, MCG3-DFT, G3S and G3SX methods. Thus we consider only the Dunning-type basis sets in the current study.

## 2. Methods

In order to simplify the notation, we make the following abbreviations for Dunning's correlation-consistent basis sets:

pdz cc-pV(D + d)Z  
 apdz aug-cc-pV(D + d)Z  
 ptz cc-pV(T + d)Z  
 aptz aug-cc-pV(T + d)Z

The '+d' signifies that an additional set of *d* functions are added to the original correlation-consistent basis sets for the second-row elements [32]. For elements not in the second row, the original cc-pVnZ and aug-cc-pVnZ (*n* = D, T) basis sets were used [28–30].

We used the MLSE-DFT methods we developed previously [27] as the starting point for our improved methods:

$$\begin{aligned}
 E(\text{MLSE-DFT}) = & C_{\text{WF}} \{E(\text{HF}/\text{pdz}) + \\
 & C_{\Delta\text{HF}} [E(\text{HF}/\text{ptz}) - E(\text{HF}/\text{pdz})] + \\
 & C_{\text{E2}} [E2/\text{pdz}] + \\
 & C_{\text{E34}} [E(\text{MP4SDQ}/\text{pdz}) - E(\text{MP2}/\text{pdz})] + \\
 & C_{\text{QC1}} [E(\text{QCISD(T)}/\text{pdz}) - E(\text{MP4SDQ}/\text{pdz})] + \\
 & C_{\text{B}} [E2/\text{ptz} - E2/\text{pdz}] + \\
 & C_{\text{HF+}} [E(\text{HF}/\text{apdz}) - E(\text{HF}/\text{pdz})] + \\
 & C_{\text{E2+}} [E2/\text{apdz} - E2/\text{pdz}] + \\
 & (1 - C_{\text{WF}}) \{E(\text{DFTX}/\text{pdz}) + \\
 & C_{\text{BI}} [E(\text{DFTX}/\text{ptz} - \text{DFTX}/\text{pdz})]\} + E_{\text{SO}},
 \end{aligned} \quad (1)$$

where  $E(\text{theory}/\text{basis set})$  denotes the single-point Born–Oppenheimer energy, which was calculated by combining the particular theory with the basis set. The E2/basis set denotes the second-order energy correction calculated from the MP2 theory using the corresponding basis set. The DFTX denotes a particular hybrid DFT method with *X*% Hartree–Fock exchange energy. Tabulated spin-orbital correction energies ( $E_{\text{SO}}$ ) for atoms and selected open-shell species [33] are also included in Eq. (1). The use of two basis sets in the DFT calculation in Eq. (1) was found to be able to increase the accuracy substantially [27].

The first improved method, called the MLSE(C1)-DFT, was developed by adding the energy components obtained from the MP2/aptz, and MP4D/ptz levels to (1):

$$\begin{aligned}
 E(\text{MLSE(C1)-DFT}) = & C_{\text{WF}} \{E(\text{HF}/\text{pdz}) + \\
 & C_{\text{E2}} [E2/\text{pdz}] + \\
 & C_{\text{E34SDQ}} [E(\text{MP4SDQ}/\text{pdz}) - E(\text{MP2}/\text{pdz})] + \\
 & C_{\text{QC1D}} [E(\text{QCISD}/\text{pdz}) - E(\text{MP4SDQ}/\text{pdz})] + \\
 & C_{\text{QC1}} [E(\text{QCISD(T)}/\text{pdz}) - E(\text{QCISD}/\text{pdz})] + \\
 & C_{\text{B1E2}} [E2/\text{ptz} - E2/\text{pdz}] + \\
 & C_{\text{HF+}} [E(\text{HF}/\text{apdz}) - E(\text{HF}/\text{pdz})] + \\
 & C_{\text{E2+}} [E2/\text{apdz} - E2/\text{pdz}] + \\
 & C_{\text{B2E2}} [E2/\text{aptz} - E2/\text{apdz}] + \\
 & C_{\text{B1E34}} [E(\text{MP4D}/\text{ptz}) - E(\text{MP4D}/\text{pdz})] + \\
 & (1 - C_{\text{WF}}) \{E(\text{DFTX}/\text{pdz}) + \\
 & C_{\text{DFT+}} [E(\text{DFTX}/\text{apdz} - \text{DFTX}/\text{pdz})]\}.
 \end{aligned} \quad (2)$$

As in the MLSE-DFT method, the  $C_{\text{WF}}$  term consists of energies obtained from wavefunction-based methods, while the  $(1 - C_{\text{WF}})$  term consists of energies from DFT calculation. The MP4D theory was used instead of the MP4SDQ in the  $C_{\text{B1E34}}$  term because our tests showed that the MP4D theory provided better overall performance at a lower computational cost. The higher-level electron correlation contribution above MP4D was handled by the QCISD(T) and DFT calculation. The inclusion of the spin-orbital correction energies does not improve the accuracy, and thus the  $E_{\text{SO}}$  term was not used in this method. For the DFTX terms, several hybrid DFT methods, including MPW1B95 [34,35], TPSS1KCIS [36,37], MPW1PW91 [34], B1B95 [35,38], M06-2X [39] and B3LYP [40,41] were tested in the method. Although Eq. (2) contains many energy terms, the wavefunction part only requires four single-point *ab initio* calculations: MP2/aptz, MP4D/ptz, MP2/apdz, and QCISD(T)/pdz. The relative importance of these calculations is discussed briefly in the supplementary material. For the DFT calculation in Eq. (2), we now use the apdz and pdz basis sets since this combination was found to provide the best results.

The computational cost of MLSE(C1)-DFT is significantly higher than that of MLSE-DFT because of the expensive MP2/aug-cc-pVTZ calculation. One way to lower the cost is to reduce the size of the aug-cc-pVTZ basis set. Usually, the diffuse functions with higher angular momentum are less important than those with lower angular momentum. Consequently, the second method simplified the aug-cc-pVTZ basis sets by omitting the *f* diffuse functions for the second-row elements, omitting the *d* and *f* diffuse functions for the first-row elements, and omitting all diffuse functions for hydrogen. The resulting basis set is abbreviated as aptzs (simplified aug-cc-pVTZ). By replacing the MP2/aptz with the MP2/aptzs calculation, we developed the following MLSE(C2)-DFT method:

$$\begin{aligned}
 E(\text{MLSE(C2)-DFT}) = & C_{\text{WF}} \{E(\text{HF}/\text{pdz}) + \\
 & C_{\text{E2}} [E2/\text{pdz}] + \\
 & C_{\text{E34SDQ}} [E(\text{MP4SDQ}/\text{pdz}) - E(\text{MP2}/\text{pdz})] + \\
 & C_{\text{QC1D}} [E(\text{QCISD}/\text{pdz}) - E(\text{MP4SDQ}/\text{pdz})] + \\
 & C_{\text{QC1}} [E(\text{QCISD(T)}/\text{pdz}) - E(\text{QCISD}/\text{pdz})] + \\
 & C_{\text{B1E2}} [E2/\text{ptz} - E2/\text{pdz}] + \\
 & C_{\text{HF+}} [E(\text{HF}/\text{apdz}) - E(\text{HF}/\text{pdz})] + \\
 & C_{\text{E2+}} [E2/\text{apdz} - E2/\text{pdz}] + \\
 & C_{\text{B2E2}} [E2/\text{aptzs} - E2/\text{apdz}] + \\
 & C_{\text{B1E34}} [E(\text{MP4D}/\text{ptz}) - E(\text{MP4D}/\text{pdz})] + \\
 & (1 - C_{\text{WF}}) \{E(\text{DFTX}/\text{pdz}) + \\
 & C_{\text{DFT+}} [E(\text{DFTX}/\text{apdz} - \text{DFTX}/\text{pdz})]\}.
 \end{aligned} \quad (3)$$

The above simplifications significantly reduce the computation cost of the MLSE(C2)-DFT method compared with the MLSE(C1)-DFT method. However, two large basis sets, ptz and aptzs, are still

required for the MP2 calculation. Moreover, the MP4D/ptz calculation is also very expensive. To make the method even more computationally affordable, we completely eliminate the calculation using the ptz basis set in the following MLSE(C3)-DFT method:

$$\begin{aligned}
 E(\text{MLSE(C3)-DFT}) = & C_{\text{WF}}\{E(\text{HF}/\text{pdz})+ \\
 & C_{\text{E2}} [E2/\text{pdz}] + \\
 & C_{\text{E34D}} [E(\text{MP4D}/\text{pdz}) - E(\text{MP2}/\text{pdz})] + \\
 & C_{\text{E34SDQ}} [E(\text{MP4SDQ}/\text{pdz}) - E(\text{MP4D}/\text{pdz})] + \\
 & C_{\text{QCI}} [E(\text{QCISD(T)}/\text{pdz}) - E(\text{MP4SDQ}/\text{pdz})] + \\
 & C_{\text{HF+}} [E(\text{HF}/\text{apdz}) - E(\text{HF}/\text{pdz})] + \\
 & C_{\text{E2+}} [E2/\text{apdz} - E2/\text{pdz}] + \\
 & C_{\text{B2HF}} [E(\text{HF}/\text{aptzs}) - E(\text{HF}/\text{apdz})] + \\
 & C_{\text{B2E2}} [E2/\text{aptzs} - E2/\text{apdz}] + \\
 & C_{\text{BMP4+}} [E(\text{MP4SDQ}/\text{apdz}) - E(\text{MP4SDQ}/\text{pdz})] \\
 & (1 - C_{\text{WF}}) \{E(\text{DFTX}/\text{pdz})\} + E_{\text{SO}}.
 \end{aligned} \quad (4)$$

Compared to the C2 method, all the calculation involving the ptz basis set is omitted. An additional MP4SDQ/apdz calculation is added (the  $C_{\text{BMP4+}}$  term) to partially compensate for the reduced accuracy caused by omitting the MP4D/ptz calculation. The spin-orbital correction energies are also included. In the current method, only one basis set (pdz) is used for the DFT calculation because we found almost no improvement using two basis sets in this method. One may question why only the double-zeta basis sets were used for the methods developed in the current study. Our tests showed that since the main contribution of energies was from the wavefunction-based energy terms, the role of DFT terms was providing minor additional exchange-correlation corrections. Only very minor, if at all, improvement would be obtained if larger basis sets or basis set combinations were used in the DFT terms. More detailed data on this matter is available in the supplementary material.

The coefficients (including  $X$ ) in Eqs. ((2)–(4)) were determined by minimizing the MUEs with respect to a set of 211 accurate thermochemical kinetics data (the ‘training set’). (For M06-2X and B3LYP functionals, the default values of  $X$  were used and not re-optimized.) These data include 109 main-group atomization energies (AEs) from the MGAE109/05 database [42], 38 hydrogen-transfer barrier heights (HTBHs), 38 non-hydrogen-transfer barrier heights (NHTBHs) from the HTBH38/04 [26] and NHTBH38/04 [43] databases, 13 ionization potential (IPs) energies and 13 electron affinity (EAs) energies from the IP13/3 and EA13/3 databases [44], respectively. The NHTBH38/04 database values were determined using the W1 theory. All other database values were derived from reliable experimental measurement [45]. These databases were compiled by Truhlar and co-workers and were previously used to determine the MCG3-DFT and MLSE-DFT coefficients [26,29]. The electronic structure calculation was performed using the GAUSSIAN 03 program [46]. It is noted that the basis-set extrapolation parts

of the current and other previous multi-coefficient methods all shared the same formalism with the basis-set extrapolation scheme by Schwenke [47], 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.

### 3. Results and discussion

Table 1 lists the MUEs obtained by the current MLSE( $C_n$ )-M06-2X ( $n=1, 2$ ) and MLSE(C3)-B3LYP methods. The results from several other recent multi-coefficient methods are also listed for comparison. Some researchers prefer using the root-mean-square error (RMSE) as measurement for accuracy since they may provide better assessment of the accuracy for smaller sample sizes and for systems with irregular distribution of errors. We thus also listed the RMSEs in Table 1 if they are available. Among the choices of the density functionals, for  $n=1$  and 2, the methods using the M06-2X functional performed best, while for  $n=3$ , the method using the B3LYP functional performed best. The MLSE(C1)-M06-2X method provided the lowest overall MUE of 0.56 kcal/mol on the training set. Compared to the MCG3-MPWB method (overall MUE = 0.73 kcal/mol), the MLSE(C1)-M06-2X method predicted more accurate atomization energies and significantly more accurate non-hydrogen-transfer barrier heights. The MLSE(C2)-M06-2X method, which is a simplified version of the MLSE(C1)-M06-2X method described in the previous section, also gave a very satisfactory overall MUE of 0.59 kcal/mol. Compared to the C1 method, the C2 method calculated similar MUEs based on atomization energies and barrier heights, but they gave higher MUEs based on ionization potentials and electron affinities. This suggests that the simplification of the aptz basis set did not significantly affect the calculation accuracy for the neutral systems. The MLSE(C3)-B3LYP method calculated an overall MUE of 0.62 kcal/mol. The errors in the IPs, EAs and barrier heights are 0.1–0.2 kcal/mol higher than the C1 method. The overall MUEs of the MLSE(C3)-B3LYP methods are approximately 0.06 kcal/mol higher than the C1 methods. This is still quite satisfactory, considering the significant simplification over the C1 methods. As seen in Table 1, the trend in the overall MUE and RMSE is very similar. The current methods also compare favorably with the G3S/3 and G3SX methods. Compared with the MLSE-TS method [27] we developed previously for neutral systems, the MLSE(C3)-B3LYP method performed similarly, and the C1 and C2 methods predicted much better non-hydrogen-transfer barrier heights.

Table 2 compares the overall MUEs obtained by using different DFT functionals in the MLSE( $C_n$ )-DFT methods. For the C1 and C2 methods, the M06-2X functional performed best, while the best functional for the C3 method is B3LYP. The MPW1B95 functional, which has been the choice in the MCG3-DFT methods, performed

**Table 1**  
Mean unsigned errors (kcal/mol) obtained using the MLSE( $C_n$ )-DFT methods.

	AE	IP	EA	HTBH	NHTBH	Overall MUE	Overall RMSE
MLSE(C1)-M06-2X	0.62	0.55	0.63	0.47	0.43	0.56	0.79
MLSE(C2)-M06-2X	0.65	0.60	0.69	0.50	0.44	0.59	0.82
MLSE(C3)-B3LYP	0.62	0.68	0.82	0.45	0.68	0.62	0.85
MLSE-TS <sup>a</sup>	0.62	–	–	0.55	0.69	0.61	0.90
MCG3/3 <sup>b</sup>	1.04	0.95	0.92	0.84	1.00	0.98	1.27 <sup>c</sup>
MCG3-MPWB <sup>b</sup>	0.75	0.67	0.86	0.54	0.84	0.73	0.95 <sup>c</sup>
G3S/3 <sup>b</sup>	0.94	1.02	1.21	0.76 <sup>d</sup>	–	0.92 <sup>c</sup>	1.26 <sup>c</sup>
G3SX <sup>b</sup>	0.85	1.07	1.06	0.67	0.60	0.80	1.19 <sup>c</sup>

<sup>a</sup> For neutral systems only.

<sup>b</sup> Obtained from Refs. [4,26,31].

<sup>c</sup> Not including NHTBH.

<sup>d</sup> MUE of 44 hydrogen transfer barrier heights from Ref. [4].

**Table 2**  
Comparison of MUEs (kcal/mol) obtained by different DFT functionals.

	MLSE(C1)-DFT	MLSE(C2)-DFT	MLSE(C3)-DFT
M06-2X	0.556	0.585	0.637
B3LYP	0.605	0.632	0.617
MPW1B95	0.586	0.592	0.651
TPSS1KCIS	0.616	0.599	0.643
MPW1PW91	0.612	0.603	0.650
B1B95	0.593	0.596	0.653

second best in the C1 and C2 methods. The M06-2X functional is the second best for the C3 method. Apparently, the very accurate M06-2X functional is also very suitable to be used in the current multi-coefficient methods. For the C1 and C3 methods, the M06-2X and B3LYP functionals, respectively, are better than other functionals by a relatively wide margin. For the C2 method, the accuracy is less sensitive to the choice of functionals. The TPSS1KCIS, MPW1PW91, and B1B95 functional performed similarly, and they are significantly less accurate than the best functionals in the C1 and C3 methods. More detailed performance dependence on the functionals and basis sets is included in the supplementary material.

The optimized coefficients of the three new methods in Table 1 are shown in Table 3. The coefficients of other new methods are included in the supplementary material. We found that the DFT energy only contributed to approximately 10–20% of the total energy in Eqs. (2)–(4), and the primary contribution to the MLSE(C $n$ )-DFT energy was, as expected, from the wavefunction-based terms. However, our study also showed that if the DFT terms were omitted, the calculated MUEs would be 0.1–0.2 kcal/mol higher. Thus, the DFT terms are important to reach higher accuracy with only very small increase in computational cost.

Table 4 compares the computational cost and MUEs from various multi-level and multi-coefficient methods and from the recently developed M06-2X density functional method [39]. For medium-sized molecules, the MLSE(C1)-M06-2X method was 25% faster than the accurate G3SX method. It also performed significantly better on the current thermochemical kinetics data set. The MLSE(C2)-M06-2X method achieved similar accuracy but required only 54% of the computational cost of the C1 method. Thus, it is the most recommended method from this study. The most economical MLSE(C3)-B3LYP method required 67% of the computational cost of the MLSE(C2)-M06-2X method, and provides an MUE only 5% higher. This method is comparable in cost to the MCG3-MPWB method but achieves higher accuracy. Consequently, the MLSE(C3)-B3LYP method is another economical alternative to the C2 and C1 methods, especially for larger systems. As also seen in Table 4, the M06-2X/MG3S method provides an MUE of 1.68 kcal/mol at a very modest cost. While the cost of the

**Table 3**  
Optimized coefficients of the MLSE(C $n$ )-DFT methods.

	MLSE(C1)-M06-2X	MLSE(C2)-M06-2X	MLSE(C3)-B3LYP
C <sub>WF</sub>	0.80087	0.81024	0.88226
C <sub>E2</sub>	1.04438	1.04515	0.94430
C <sub>E34D</sub>			0.67265
C <sub>E34SDQ</sub>	0.89730	0.91059	1.04817
C <sub>QCID</sub>	1.24570	1.25425	
C <sub>QC1</sub>	0.99005	1.07955	1.17690
C <sub>B1E2</sub>	-1.22019	-1.04670	
C <sub>HF+</sub>	0.19534	0.28851	
C <sub>E2+</sub>	0.63885	0.71751	-0.37629
C <sub>B2HF</sub>			1.23505
C <sub>B2E2</sub>	0.85051	0.71956	2.09452
C <sub>BMP4+</sub>			1.63976
C <sub>B1E34</sub>	1.79806	1.77047	
C <sub>DFT+</sub>	2.93878	2.48901	

**Table 4**  
Computational cost.<sup>a</sup>

	Cost <sup>b</sup>	MUE
MLSE(C1)-M06-2X	107.2	0.56
MLSE(C2)-M06-2X	57.7	0.59
MLSE(C3)-B3LYP	38.7	0.62
MLSE-TS	34.4	0.61 <sup>c</sup>
MCG3/3	27.6	0.98 <sup>d</sup>
MCG3-MPWB	33.8	0.73 <sup>d</sup>
CBS-Q	46.8	1.23 <sup>d,e</sup>
G3S/3	124.5	0.92 <sup>d,e</sup>
G3SX	143.1	0.80 <sup>d</sup>
M06-2X/MG3S	11.9	1.68
M06-2X/aug-cc-pVTZ	58.9	1.89
M06-2X/aug-pc-2	63.0	1.60

<sup>a</sup> Determined using computers with an Intel E6600 CPU and a version of GAUSSIAN 03 program optimized for the hardware architecture.

<sup>b</sup> Total CPU time relative to an MP2/6-31+G(d,p) calculation for C<sub>5</sub>H<sub>5</sub>N, C<sub>2</sub>Cl<sub>4</sub>, C<sub>4</sub>H<sub>4</sub>O, C<sub>4</sub>H<sub>4</sub>S, C<sub>4</sub>H<sub>5</sub>N, CF<sub>3</sub>CN, and SiCl<sub>4</sub>. The CPU time for M06-2X and B3LYP functionals is almost identical.

<sup>c</sup> MUE for neutral systems only.

<sup>d</sup> Obtained from Refs. [4,26,31].

<sup>e</sup> MUE without NHTBH.

**Table 5**  
Unsigned errors (kcal/mol) for the TAE6 test set.

	MLSE(C1) M06-2X	MLSE(C2) M06-2X	MLSE(C3) B3LYP	MLSE TS	MCG3 <sup>a</sup> TS	G3SX <sup>a</sup>
C <sub>10</sub> H <sub>8</sub>	1.22	1.27	0.61	0.08	0.96	0.14
C <sub>8</sub> H <sub>18</sub>	1.03	0.74	1.08	1.78	1.07	0.69
C <sub>10</sub> H <sub>16</sub>	0.67	0.73	0.54	0.10	0.21	0.16
O <sub>3</sub>	0.66	0.18	0.47	1.49	1.25	0.87
OCS	0.99	1.17	1.15	1.09	1.79	2.06
P <sub>4</sub>	1.74	2.41	4.49	1.99	0.90	8.96
Average	1.06	1.08	1.39	1.09	1.03	2.14

<sup>a</sup> Obtained from Ref. [26].

MLSE(C2)-M06-2X method is approximately seven times higher than the cost of the M06-2X/MG3S, the accuracy is approximately three times higher. Using larger basis sets, such as aug-pc-2 [48] or aug-cc-pVTZ, with the M06-2X functional is as expensive as the MLSE(C2)-M06-2X method, and it offers no significant improvement in accuracy over the M06-2X/MG3S method. Therefore, for calculations that require high accuracy, it is worth incurring the computational cost from using the MLSE(C2)-M06-2X method or other accurate multi-coefficient methods listed in Table 4.

In order to further evaluate the new methods developed in the current study, we tested their performance on the TAE6 data set [26] in Table 5. The TAE6 set, which was compiled by Truhlar and coworkers, consists of atomization energies from six 'tough' molecules as shown in Table 5. These include two large molecules, *n*-octane and naphthalene, and four small molecules, C<sub>10</sub>H<sub>8</sub>, OCS, O<sub>3</sub>, and P<sub>4</sub>, which are notoriously difficult [26] to obtain accurate AEs using quantum chemical methods. These molecules are *not* included in our training set. As shown in the table, the MLSE(C $n$ )-M06-2X ( $n = 1, 2$ ) and MLSE(C3)-B3LYP methods performed satisfactorily on the TAE6 set, with MUEs of 1.06, 1.08, and 1.39 kcal/mol, respectively. The results were also compared with those obtained by MLSE-TS, MCG3-TS [26], as well as the G3SX methods in Table 5. The MLSE(C $n$ )-M06-2X ( $n = 1, 2$ ), MLSE-TS, and MCG3-TS methods yielded similar values for MUE,  $\sim 1$  kcal/mol. Thus, the new methods developed in this study were robust enough to treat these difficult systems reasonably accurately.

#### 4. Concluding remarks

The 'hybrid' approach that combines the wavefunction-based energies and the DFT energies with adjustable scaling coefficients have been proved to be the most cost-effective way to obtain accurate molecular energies for larger (>5 heavy atoms) chemical systems. This approach first appeared in the 'doubly-hybrid' methods by Truhlar and coworker in 2004 which combined the HF, MP2, and hybrid DFT energies. The best method MC3BB outperformed either the MP2 or B1B95 methods by a wide margin. Then in 2005 the MCCM suite of methods were developed by the same authors. In particular, the various MCG3-DFT and MCQCISD-DFT methods gave surprisingly good performance/cost ratio. In 2006 Grimme proposed a related DFT method B2-PLYP [49] that includes the perturbative second-order correlation calculated using the Kohn–Sham orbitals. This approach was later further developed by Martin and co-workers [50]. These so called 'double-hybrid' DFTs performed significantly better than the corresponding hybrid DFTs. Much more exciting development in DFT, or electronic structure methods in general, using the hybrid approach is anticipated in the near future.

The aim of the current study was to develop a set of accurate and economical hybrid multi-coefficient methods for the study of thermochemical kinetics. The purpose of these new MLSE(*Cn*)-DFT methods is thus different from the recently developed highly accurate composite methods such as Wn, ccCA, FPA, and HEAT which can achieve amazingly high accuracy (<1 kJ/mol) in predicting the atomization energies and heats of formation for small molecules. In these composite methods, the correlated calculation (MP2, MP4, coupled cluster) with a triple-zeta basis set was usually just a starting point, and much larger basis sets are required in the calculation. This is in contrast to our current methods in which the most time-consuming calculation for medium-sized molecules is MP2/aptz or MP2/aptzs. We expect that the new MLSE(*Cn*)-DFT methods can easily be applied to many types of interesting chemical systems with 10–15 heavy atoms, and they will be invaluable for accurate study of thermochemistry and kinetics.

#### Acknowledgements

This work is supported by the National Science Council of Taiwan, Grant No. NSC 97-2113-M-194-004. We are grateful to the National Center for High-Performance Computing (NCHC) for providing part of computation resources.

#### Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.cplett.2009.05.015.

#### References

- [1] F.B. Brown, D.G. Truhlar, Chem. Phys. Lett. 117 (1985) 307.
- [2] M.S. Gordon, D.G. Truhlar, J. Am. Chem. Soc. 108 (1986) 5412.
- [3] P.L. Fast, J.C. Corchado, M.L. Sánchez, D.G. Truhlar, J. Phys. Chem. A 103 (1999) 5129.
- [4] B.J. Lynch, D.G. Truhlar, J. Phys. Chem. A 107 (2003) 3898.
- [5] T.-H. Li, C.-H. Mou, W.-P. Hu, Chem. Phys. Lett. 397 (2004) 364.
- [6] L.A. Curtiss, K. Raghavachari, P.C. Redfern, V. Rasslov, J.A. Pople, J. Chem. Phys. 109 (1998) 7764.
- [7] L.A. Curtiss, P.C. Redfern, K. Raghavachari, J. Chem. Phys. 126 (2007) 84108.
- [8] M.J. Frisch et al., J. Chem. Phys. 110 (1999) 2822.
- [9] M.J. Frisch et al., J. Chem. Phys. 112 (2000) 6532.
- [10] J.M.L. Martin, G. de Oliveira, J. Chem. Phys. 111 (1999) 1843.
- [11] A. Karton, E. Rabinovich, J.M.L. Martin, B. Ruscic, J. Chem. Phys. 125 (2006) 144108.
- [12] A. Karton, P.R. Taylor, J.M.L. Martin, J. Chem. Phys. 127 (2007) 064104.
- [13] N.J. DeYonker, T.R. Cundari, A.K. Wilson, J. Chem. Phys. 124 (2006) 114104.
- [14] N.J. DeYonker, T. Grimes, S. Yockel, A. Dinescu, B. Mintz, T.R. Cundari, A.K. Wilson, J. Chem. Phys. 125 (2006) 104111.
- [15] A.L.L. East, W.D. Allen, J. Chem. Phys. 99 (1993) 4638.
- [16] A.G. Csaszar, W.D. Allen, H.F. Schaefer III, J. Chem. Phys. 108 (1998) 9751.
- [17] A. Tajti et al., J. Chem. Phys. 121 (2004) 11599.
- [18] Y.J. Bomble et al., J. Chem. Phys. 125 (2006) 064108.
- [19] M.E. Harding, J. Vazquez, B. Ruscic, A.K. Wilson, J. Gauss, J.F. Stanton, J. Chem. Phys. 128 (2008) 114111.
- [20] W. Klopper, B. Ruscic, D.P. Tew, F.A. Bischoff, S. Wolfsegger, Chem. Phys. 356 (2009) 14.
- [21] C.M. Tratz, P.L. Fast, D.G. Truhlar, Phys. Chem. Commun. 2 (1999), article 14.
- [22] P.L. Fast, D.G. Truhlar, J. Phys. Chem. A 104 (2000) 6111.
- [23] T.-H. Li, H.-R. Chen, W.-P. Hu, Chem. Phys. Lett. 412 (2005) 430.
- [24] L.A. Curtiss, K. Raghavachari, P.C. Redfern, J.A. Pople, J. Chem. Phys. 112 (2000) 1125.
- [25] L.A. Curtiss, P.C. Redfern, K. Raghavachari, J.A. Pople, J. Chem. Phys. 114 (2001) 108.
- [26] Y. Zhao, B.J. Lynch, D.G. Truhlar, Phys. Chem. Chem. Phys. 7 (2005) 43.
- [27] Y.-L. Sun, T.-H. Li, J.-L. Chen, K.-J. Wu, W.-P. Hu, Chem. Phys. Lett. 442 (2007) 220.
- [28] T.H. Dunning Jr., J. Chem. Phys. 90 (1989) 1007.
- [29] R.A. Kendall, T.H. Dunning Jr., R.J. Harrison, J. Chem. Phys. 96 (1992) 6796.
- [30] D.E. Woon, T.H. Dunning Jr., J. Chem. Phys. 98 (1993) 1358.
- [31] J. Zheng, Y. Zhao, D.G. Truhlar, J. Chem. Theory Comput. 3 (2007) 569.
- [32] T.H. Dunning Jr., K.A. Peterson, A.K. Wilson, J. Chem. Phys. 114 (2001) 9244.
- [33] P.L. Fast, J. Corchado, M.L. Sánchez, D.G. Truhlar, J. Phys. Chem. A 103 (1999) 3139.
- [34] C. Adamo, V. Barone, J. Chem. Phys. 108 (1998) 664.
- [35] A.D. Becke, J. Chem. Phys. 104 (1996) 1040.
- [36] J. Tao, J.P. Perdew, V.N. Staroverov, G.E. Scuseria, Phys. Rev. Lett. 91 (2003) 146401.
- [37] T. Toulouse, A. Savin, C. Adamo, J. Chem. Phys. 117 (2002) 10465.
- [38] A.D. Becke, Phys. Rev. A 38 (1988) 3098.
- [39] Y. Zhao, D.G. Truhlar, Theor. Chem. Acc. 120 (2008) 215.
- [40] A.D. Becke, J. Chem. Phys. 98 (1993) 1372.
- [41] A.D. Becke, J. Chem. Phys. 98 (1993) 5648.
- [42] Y. Zhao, N.E. Schultz, D.G. Truhlar, J. Chem. Theory Comput. 2 (2006) 364.
- [43] Y. Zhao, N. González-García, D.G. Truhlar, J. Phys. Chem. A 109 (2005) 2012.
- [44] B.J. Lynch, Y. Zhao, D.G. Truhlar, J. Phys. Chem. A 107 (2003) 1384.
- [45] P.L. Fast, M.L. Sánchez, J.C. Corchado, D.G. Truhlar, J. Chem. Phys. 110 (1999) 11679.
- [46] M.J. Frisch et al., GAUSSIAN 03, Revision D.02, Gaussian, Inc., Wallingford CT, 2004.
- [47] D.W. Schwenke, J. Chem. Phys. 122 (2005) 014107.
- [48] F.J. Jensen, Chem. Phys. 116 (2002) 7372.
- [49] S. Grimme, J. Chem. Phys. 124 (2006) 034108.
- [50] A. Karton, A. Tarnopolsky, J.-F. Lamere, G.C. Schatz, J.M.L. Martin, J. Phys. Chem. A 112 (2008) 12868.