J. Phys. Chem. A **2008,** *112,* 117−124

Theoretical Study of the Enthalpies of Formation for $C_{40}H_{56}$ Carotenes

Cheng-Yi Tu, Wen-Hsin Guo, and Ching-Han Hu*

Department of Chemistry, National Changhua University of Education, Changhua 50058, Taiwan

Received: July 19, 2007; In Final Form: August 31, 2007

Theoretical study of the enthalpies of formation (ΔH_f) for polyenes up to nine ethylene units and for several $C_{40}H_{56}$ carotenes including β-carotene, α-carotene, lycopene, and prolycopene is presented. For polyenes and small branched alkenes, we used G2, G3, and G3MP2B3 theories, and the ΔH_f values were evaluated with the atomization, *isodesmic* bond separation, and *homodesmic* schemes. The applicability of six DFT functionals were evaluated by comparing their predictions with those obtained using G3 theory within the atomization scheme. Additivity approaches, including atom equivalents and group equivalents using DFT and semiempirical theories, were explored. We found that group equivalents associated with *isodesmic* reactions are able to provide the most accurate predictions within the test set. The predictions from the six functionals are in good agreement with the G3 results. Among them, B3LYP performs the best, with an average absolute deviation of only 0.30 kcal/mol. The application of DFT in the prediction for the ΔH_f value of $C_{40}H_{56}$ carotenes is promising.

1. Introduction

Carotenoids are an important type of molecules that occur naturally in plants and photosynthetic organisms. They are divided into two categories: carotenes, which consist of only C and H atoms, and xanthophylls, which consist of oxygen atoms at the end groups in addition to C and H. The protective role played by these compounds have long been recognized; they are known to perform free-radical scavenging and singlet oxygen quenching effects in biological tissues.[1−5] Carotenoids are also important plant pigments. They act as light-harvesting antennae in the photosynthetic system and protect the plants from the harmful effects caused by singlet oxygen and triplet chlorophyll.[6]

The $C_{40}H_{56}$ family of carotenes have received probably the most attention from scientists than others. Among them, β-carotene and lycopene are the most abundant in human plasma and have been proposed to perform anticancer and antioxidant activities.[7]

Thermodynamic data, particularly enthalpies of formation (ΔH_f) of molecules, is an important property, from which one can calculate enthalpies of reaction before performing an experiment. Unfortunately, such data for carotenes has been lacking. The extended, delocalized framework of carotenes discourages the determination of their ΔH_f values using experimental or computational means.

In the past decade, accurate prediction of the enthalpies of formation of gas-phase molecules has been achieved via high-level quantum chemistry methods. Several theoretical approaches have been developed for the purpose of accurate predictions of the ΔH_f values of medium-sized molecules, including radicals and ionic species.[8−12] Among them, the GX theories have received much attention from chemists. For neutral, closed-shell molecules consisting of first and second period atoms, chemical accuracy (±1 kcal/mol) from experiment can be expected. For one example, in a set of 38 hydrocarbons, enthalpies of formation predicted using G3 theory have an average absolute deviation (AAD) of 0.69 kcal/mol from the experiment,[11] while its computationally less-costly variant G3-(MP2) has similar accuracy.

In the case of carotenes, molecules of interest are much larger than those that can be studied using G2 and G3 theories. Thus, computational methods that are less CPU demanding (for example, density functional theory, DFT) are necessary. The applicability of such alternative methods relies on error cancellations; thus, *isodesmic* or *homodesmic* reactions should be considered.[13] It has been well demonstrated that DFT has a serious problem with the accumulation of errors as the size of molecule increases.[14,15] The systematic error was shown by Schleyer et al. to be the inadequacy in describing the long-range nonbonded interactions, that is, protobranching stabilization effects using DFT.[16,17] It was also pointed out by Grimme that DFT presents serious difficulty in describing the stereoelectronic alkylation effects.[18] These studies pose warnings for the application of DFT in the computations of ΔH_f for carotenes.

Despite the warning from failed cases of DFT, it remains to be a valuable alternative to ab initio methods for its computational effectiveness. Thus, in this study we will examine the applicability of DFT functionals in the prediction of the enthalpies of formation. The ΔH_f values of carotenes are calculated using different schemes including atomization reaction, isodesmic[13] (bond separation) reaction, and for model polyene compounds a homodesmic reaction (vide infra).

Bond and group additivity approaches based on an empirical experimental data set has been widely used in estimating enthalpies of formation.[19−21] Among the many attempts to predict thermodynamic properties of molecules, the concepts of atomic equivalents (AEs) and bond and group equivalents (BGEs) originated from a similar idea and have been explored by several research groups. For example, Jorgensen et al. developed AE and BGE schemes using semiempirical methods including AM1, PM3, and MNDO. With a large set of training molecules, accurate predictions were obtained.[22] Liu et al. used the AE scheme and converted DFT energies to enthalpies of formation, and for hydrocarbons an AAD of 0.81 kcal/mol from experiment has been achieved.[23] These approaches utilize a set

10.1021/jp075654+ CCC: $40.75

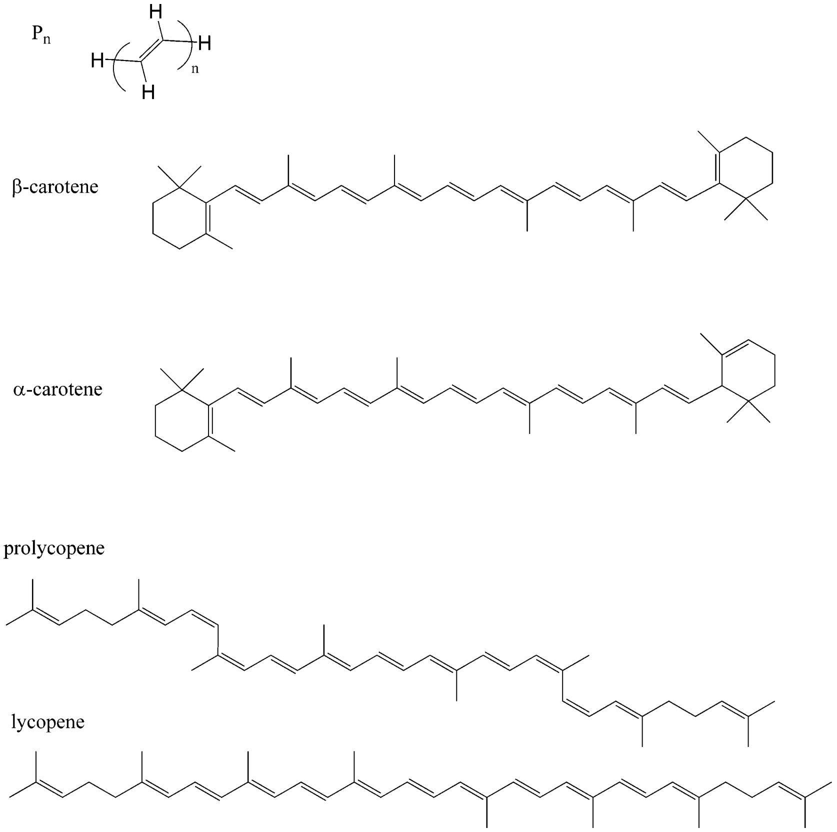

Figure 1. Polyenes (P_n, $n = 1-9$) and $C_{40}H_{56}$ carotenes.

of molecules with known enthalpies of formation, and the AEs of elements and BGEs of bonds and functional groups were obtained via least-square fitting between experimental enthalpies and computed energies. With a few fitted parameters, that is, the equivalents, we are able to predict the enthalpies of formation using the less-accurately computed energies of molecules.

The size of a carotene forbids high-level quantum chemistry computations such as G3. The application of AE and BGE, along with DFT or semiempirical methods, may offer a remedy to this difficult situation. Thus, despite the documented artifacts for the accumulation of errors in DFT, proper usage of DFT seems to still be a reasonable choice for the predictions of ΔH_f for carotenes. Semiempirical theories neglect electron correlation effects; however, these methods are efficient and have been parametrized to reproduce experimental properties for a fairly large set of molecules. One important functional feature of carotene is the long conjugated chain, which has not been the subject in the previous fitting set in obtaining AE or BGE.[22,23] The applicability of additivity approaches deserve to be validated and tested on relevant systems. In this study, we intend to design a method that is suitable for accurate determination of ΔH_f for carotene species.

2. Computational Approach

To account for the extended delocalization of carotenes, we studied a series of all-trans polyene systems P_n : H−(HC═CH)$_n$−H (see Figure 1). The ΔH_f values were calculated using (i) atomization reaction, (ii) isodesmic[13] (bond separation) reaction using eq 1, and (iii) homodesmic reaction using eq 2.

$$P_n + 2(n - 1)CH_4 \rightarrow nC_2H_4 + (n - 1)C_2H_6 \quad (1)$$

$$P_n + (n - 2)C_2H_4 \rightarrow (n - 1)C_4H_6 \quad (2)$$

G2,[10] G3,[11] and G3MP2B3[24] theories were used to compute the ΔH_f values (298 K) of molecules up to P_9. G3MP2B3 is a variation of G3MP2[25] in which B3LYP/6-31G(*d*) geometries and zero-point energies (ZPE) were used instead of MP2/6-31G(*d*) geometries and HF/6-31G(*d*) ZPE. G3MP2B3 has been demonstrated to provide energetics predictions with an accuracy comparable to that of G3MP2 at a lesser computational cost. The DFT predicted ΔH_f will be compared with the available experimental data[26] and with those obtained at the G2 and G3 levels of theories.

Six density functionals will be investigated in our study: B3LYP, BLYP, BPW91, MPWPW91, PBEPBE, and BB95, respectively.[27−33] Among them, B3LYP is a hybrid DFT that consists a three-parameter combination of Becke's and Hartree−Fock exchange energy, and Lee, Yang, and Parr's correlation functional.[27] Other functionals are pure DFT that do not include Hartree−Fock exchange. The DFT theories were applied with the 6-31G(*d*) basis set to obtain optimized geometry and harmonic vibration frequencies. Single-point calculations using the 6-311+G(3df,2p) basis set were followed. The vibrational frequencies were scaled by 0.96 and were used to obtain ZPE and thermal corrections to 298 K. The scale factor is identical to that used in the G3MP2B3 approach.[24]

For the $C_{40}H_{56}$ carotenes (Figure 1), the following bond separation, isodesmic reaction was applied to lycopene:

$$C_{40}H_{56} + 38CH_4 \rightarrow 13C_2H_4 + 26C_2H_6 \quad (3)$$

For the atomization of hydrocarbon C_nH_m (carotenes and branched alkenes) we have

$$\Delta H_f(C_nH_m) = E(C_nH_m) - n[E(C) - \Delta H_f(C)] - m[E(H) - \Delta H_f(H)] \approx E(C_nH_m) - n\epsilon_C - m\epsilon_H \quad (4)$$

TABLE 1: ΔH_f Values of Polyenes (in kcal/mol) Predicted Using Atomization, Isodesmic, and Homodesmic Schemes

		atomization			isodesmic $P_n + 2(n-1)CH_4 \rightarrow n\,C_2H_4 + (n-1)C_2H_6$			homodesmic $P_n + (n-2)C_2H_4 \rightarrow (n-1)C_4H_6$		
	exp[a]	G3	G3MP2B3	G2	G3	G3MP2B3	G2	G3	G3MP2B3	G2
P_1	12.5 ± 0.1	12.33	11.80	12.77						
P_2	26.3 ± 0.2	26.65	25.00	28.03	26.84	27.09	28.22			
P_3	40.2 ± 0.7	40.06	37.97	42.39	40.25	40.74	43.04	39.18	39.26	39.21
P_4		53.24	50.42	56.50	53.44	54.37	57.61	51.83	52.00	51.86
P_5		66.25	62.78	70.50	66.47	67.82	72.08	64.32	64.65	64.41
P_6		79.28	75.06	84.49	79.51	81.20	86.54	76.82	77.23	76.94
P_7		92.38	87.30	98.44	92.62	94.52	100.94	89.40	89.77	89.44
P_8		105.48	99.72		105.74	108.03		101.98	102.48	
P_9		118.45	111.97		118.71	121.35		114.42	115.02	

[a] Experimental values from ref 26.

In the equation, E and ΔH_f are the computed enthalpies and enthalpies of formation for C and H atoms, respectively. Using a test set of molecules with known experimental enthalpies of formation, we can replace the terms in the bracket with atomic equivalents (ϵ_C and ϵ_H) that are least-square fitted to reproduce ΔH_f from computed enthalpy. The AE schemes at various DFT theories were obtained separately. Because of the lack of experimental data, we used the G3 predicted ΔH_f of P_n and alkenes from the atomization scheme as our reference.

Similar to the idea of AE, and by using the isodesmic reaction, we can write

$$C_nH_m + lCH_4 \rightarrow (1/2)(3n - m - l)C_2H_4 + (1/2)(m - 2n + 2l)C_2H_6 \quad (5a)$$

$$\Delta H_f(C_nH_m) = E(C_nH_m) + l[E(CH_4) - \Delta H_f(CH_4)] - (1/2)(3n - m - l)[E(C_2H_4) - \Delta H_f(C_2H_4)] - (1/2)(m - 2n + 2l)[E(C_2H_6) - \Delta H_f(C_2H_6)] \approx E(C_nH_m) + l\epsilon_{CH_4} - (1/2)(3n - m - l)\epsilon_{C_2H_4} - (1/2)(m - 2n + 2l)\epsilon_{C_2H_6} \quad (5b)$$

the group equivalents (GEs, ϵ_{CH_4}, $\epsilon_{C_2H_4}$, and $\epsilon_{C_2H_6}$) were least-square fitted to the G3 enthalpies of formation. The GE scheme described here is different from Jorgensen's BGE in the sense that our "group" represents bond separation products (i.e., CH_4, C_2H_4, and C_2H_6). In contrast, Jorgensen's BGE represents contributions to ΔH_f by specific bonds and functional groups.

In addition to DFT, AE and GE schemes were applied with the semiempirical AM1[34] and PM3[35] methods. ΔH_f values evaluated using the empirical parametrizations of Leal were also included for comparison.[36] The computations were performed using the Gaussian03 series of programs.[37]

3. Results and Discussion

Polyenes. The ΔH_f values of polyenes predicted using G2, G3, and G3MP2B3 theories are summarized in Table 1. In general, G3 predicted results with different schemes are in good agreement with available experimental measurements ($P_1 \sim P_3$). The largest deviation arises from the homodesmic reaction for P_3. Results obtained from G3 theory with the atomization and isodesmic schemes are very close. Within the atomization scheme using G2 theory, we see an overestimation of ΔH_f from experimental data and from G3. The G2 deviation from G3 becomes larger as the size of polyene increases. In contrast, the G3MP2B3 predictions underestimate the experimental and G3 results, and the deviation increases for the extended polyenes. For P_9, the G3MP2B3 result is 6.48 kcal/mol smaller than the G3 prediction within the atomization scheme.

We have observed in the atomization scheme that the deviations of G2 results from those of G3 correlate approximately linearly with their differences in "higher-level correction" (HLC), and spin−orbit correction (ΔE(SO)) of G3. In other words, the differences in ΔH_f among G2 and G3 theories arises mainly from their differences in HLC and ΔE(SO) of G3 (the latter is absent in G2). The improved accuracy of the G3 theory is mainly attributed to the HLC and ΔE(SO) adopted in the theory. These corrections, however, cancel out in both the isodesmic and homodesmic schemes because there are equal numbers of α and β electrons at both sides of the reactions.

Within the isodesmic scheme, G3MP2B3 and G2 results are all larger than the experimental and G3 results, particularly G2. For P_7, G2 prediction is 8.32 kcal/mol larger than the G3 result. The overestimation of G3MP2B3 from G3 is less significant. It should be noted, however, that the ΔH_f values obtained from G3MP2B3 are noticeably different within the atomization and isodesmic schemes. G3 theory instead provides very similar predictions with both schemes. For P_9, the G3 predicted ΔH_f value within atomization and isodesmic schemes differs only by 0.26 kcal/mol.

With the homodesmic reaction, G3 theory underestimates the experimental ΔH_f of P_3 by ~1 kcal/mol. In addition, the same reaction using G3MP2B3 and G2 are very close to the G3 values, not only for P_3 but also for all polyenes. For the ΔH_f of P_9 obtained using G3 theory within the homodesmic scheme is ~4 kcal/mol lower than those obtained using the atomization and isodesmic schemes.

Unfortunately, there are only a few experimental enthalpies available for polyenes. In our evaluations of various DFT functionals, we used the ΔH_f obtained with the G3 atomization scheme as a reference. G3 theory with the atomization scheme has been developed via best fit with experimental data and was demonstrated to be superior to the G2 theory. For example, in the G2/97 test set the AAD of G3 predicted ΔH_f at 298 K is only 0.94 kcal/mol, a significant improvement compared with the G2 result (1.56 kcal/mol).[11] Higher accuracy was seen in hydrocarbons in which the AAD is only 0.68 kcal/mol (1.29 kcal/mol for G2).[11] In the study of linear alkanes by Curtis et al., it has been shown that G3 (atomization) results agree very well with the experimental data, and the agreement is further improved when the bond-separation isodesmic scheme was used.[14] In our study, the G3 isodesmic scheme provides very similar predictions to those of the G3 atomization scheme.

The ΔH_f values predicted using DFT are summarized in Table 2. The results are presented as deviation (DFT − G3) of DFT results from those obtained using the G3 atomization scheme per ethylene unit, that is, deviation/n. For the atomization scheme, we see that all tested functionals are involved with

TABLE 2: ΔH_f Values of Polyenes Predicted Using Atomization, Isodesmic, and Homodesmic Schemes at Various Levels of Density Functional Theories[a]

	atomization						isodesmic					
	B3LYP	BLYP	BPW91	MPWPW91	PBEPBE	BB95	B3LYP	BLYP	BPW91	MPWPW91	PBEPBE	BB95
P_1	−0.21	1.34	−0.74	−5.26	−9.55	−5.96						
P_2	0.68	1.63	−1.92	−7.75	−12.7	−7.27	−0.14	−0.37	0.45	−0.66	−0.78	−0.61
P_3	0.93	1.61	−3.80	−8.72	−13.9	−7.85	−0.23	−0.59	−0.89	−1.02	−1.18	−0.95
P_4	1.02	1.55	−4.29	−9.25	−14.55	−8.22	−0.32	−0.76	−1.12	−1.26	−1.43	−1.19
P_5	1.06	1.49	−4.62	−9.61	−14.97	−8.46	−0.38	−0.89	−1.27	−1.43	−1.61	−1.36
P_6	1.08	1.42	−4.86	−9.87	−15.28	−8.63	−0.43	−1.00	−1.41	−1.57	−1.76	−1.48
P_7	1.07	1.37	−5.07	−10.09	−15.53	−8.79	−0.49	−1.08	−1.55	−1.70	−1.90	−1.61
P_8	1.06	1.29	−5.25	−10.24	−15.75	−8.92	−0.54	−1.18	−1.67	−1.79	−2.04	−1.71
P_9	1.06	1.22	−5.32	−10.36	−15.85	−9.00	−0.56	−1.27	−1.69	−1.86	−2.07	−1.78
lycopene[b]	151.42	190.50	48.39	−62.44	−179.65	−26.87	108.44	106.52	93.73	87.44	79.92	87.02
	homodesmic											
P_3	−0.04	−0.11	−0.14	−0.14	−0.14	−0.14						
P_4	−0.11	−0.22	−0.27	−0.26	−0.27	−0.29						
P_5	−0.15	−0.30	−0.38	−0.37	−0.37	−0.40						
P_6	−0.19	−0.40	−0.48	−0.46	−0.47	−0.48						
P_7	−0.25	−0.46	−0.58	−0.56	−0.57	−0.58						
P_8	−0.29	−0.54	−0.68	−0.63	−0.68	−0.66						
P_9	−0.31	−0.63	−0.69	−0.68	−0.69	−0.71						

[a] Values are in deviations per ethylene unit (deviations/*n*) from G3 with corresponding schemes. [b] Values are enthalpies of formation predicted using various functionals.

accumulation of errors. Not only do the errors increase with molecular sizes but the direction of error accumulation with various DFT is also unpredictable. For example, B3LYP and BLYP overestimate ΔH_f (positive deviations); in contrast, BPW91, MPWPW91, PBEPBE, and BB95 significantly underestimate ΔH_f (negative deviations). Outrageous predictions were produced by the BPW91, MPWPW91, PBEPBE, and BB95 functionals, and the deviation/*n* increases with *n*. The extremely large deviations indicate that the atomization scheme is not an appropriate approach. It should be noted, however, that with B3LYP the deviation reaches a constant as *n* gets large.

Using the isodesmic reaction, we see that all functionals underestimate ΔH_f and the deviations increase with the size of polyenes. In contrast to the atomization scheme, the deviation/*n* of the isodesmic scheme is reduced significantly. B3LYP has the smallest deviation among these functionals. ΔH_f predicted using the homodesmic reaction is the most accurate among the three schemes, and the deviations of B3LYP and BLYP are smaller than those of other functionals.

We computed the ΔH_f of one of the $C_{40}H_{56}$ carotenes, that is, lycopene using the atomization and isodesmic reactions (see Table 2). We see that a very wide range of enthalpies of formation resulted using the atomization scheme. In contrast, the isodesmic scheme predicted that ΔH_f values are less divergent, ranging from 79.92 kcal/mol (PBEPBE) to 108.44 kcal/mol (B3LYP). It is difficult to evaluate the accuracy of the DFT predicted ΔH_f for lycopene, however. When we look into the B3LYP predictions, the trends in the atomization and isodesmic schemes suggest that atomization is likely to overestimate the ΔH_f of P_{11}, which constitute the extended conjugated system of lycopene. Likewise, B3LYP with the isodesmic reaction is likely to underestimate the ΔH_f of P_{11}. The real situation is, however, more complicated because lycopene consists of a methylated conjugated system and side groups. The homodesmic reaction (eq 2) is not available for carotenes.

Atomic Equivalence (AE) and Group Equivalence (GE) Approaches. Enthalpies of formation for lycopene predicted by DFT span a wide range, and the results from different schemes can differ significantly. For this reason, we evaluated the accuracy of AE and GE schemes on the prediction of ΔH_f.

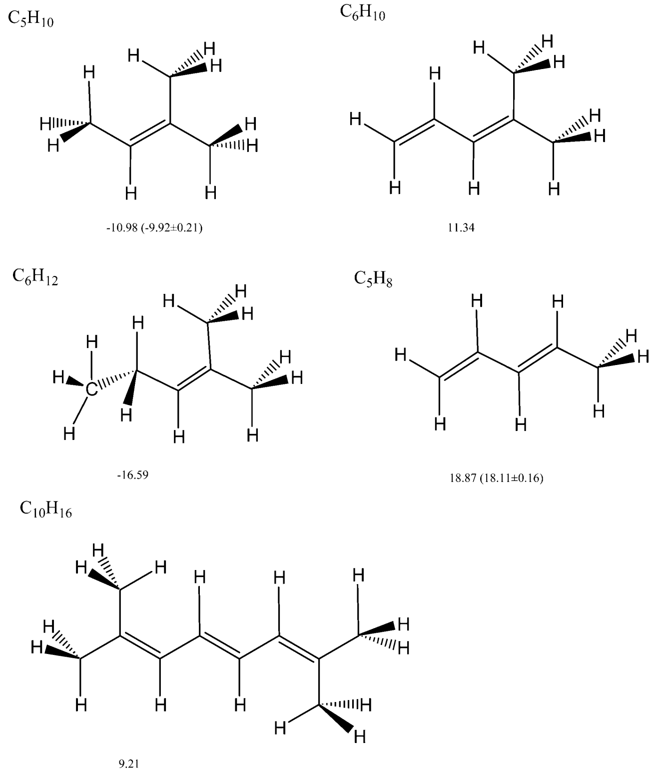

Figure 2. Branched alkenes included in the test set. ΔH_f values calculated using the G3 atomization scheme are shown; available experimental values are included in parentheses.

To obtain the AEs and GEs that are suitable for predicting ΔH_f for carotenes, we have included several branched alkenes in our test set (see Figure 2). These molecules were chosen because they consist of residual methyl groups that are relevant to carotenes. The equivalents are obtained via least-square fit to the G3 data. The ΔH_f values of polyenes and branched alkenes predicted using AE and GE at various levels of density functional theories are summarized in Tables 3 and 4, respectively. The optimized equivalents are also included in the tables.

For the predictions using AE (see Table 3), the accuracy of DFT predicted ΔH_f is improved significantly as shown by the

TABLE 3: ΔH_f Values of Polyenes Predicted Using Atomic Equivalent (AE) at Various Levels of Density Functional Theories[a]

	B3LYP	BLYP	BPW91	MPW1PW91	PBEPBE	BB95
P_2	−2.36 (−1.07)	−3.99 (−2.74)	0.55 (−0.14)	−0.57 (0.69)	0.62 (1.89)	−2.02 (−0.77)
P_3	−1.64 (−0.64)	−2.78 (−1.83)	−0.77 (0.27)	−0.07 (0.93)	0.87 (1.88)	−1.20 (−0.20)
P_4	−1.09 (−0.36)	−1.76 (−1.10)	−0.28 (0.48)	0.25 (0.97)	0.94 (1.67)	−0.67 (0.05)
P_5	−0.57 (−0.12)	−0.87 (−0.50)	0.07 (0.55)	0.38 (0.83)	0.84 (1.28)	−0.23 (0.19)
P_6	−0.12 (0.10)	−0.17 (−0.04)	0.23 (0.48)	0.38 (0.57)	0.58 (0.78)	0.12 (0.31)
P_7	0.17 (0.13)	0.55 (0.41)	0.15 (0.14)	0.12 (0.05)	0.07 (0.00)	0.22 (0.16)
P_8	0.44 (0.08)	0.94 (0.46)	−0.08 (−0.42)	−0.02 (−0.36)	−0.66 (−1.02)	0.25 (−0.14)
P_9	0.82 (0.20)	1.19 (0.51)	0.29 (−0.24)	−0.17 (−0.77)	−0.72 (−1.31)	0.39 (−0.19)
C_5H_8	−2.33 (−1.55)	−3.36 (−2.61)	−2.17 (−1.32)	−1.49 (−0.72)	−0.48 (0.29)	−2.08 (−1.36)
C_5H_{10}	−0.41 (−0.39)	−0.54 (−0.56)	1.47 (1.61)	1.75 (1.75)	−0.58 (−0.58)	−0.57 (−0.62)
C_6H_{12}	1.37 (0.72)	1.76 (1.09)	0.07 (−0.41)	−0.06 (−0.73)	0.43 (−0.23)	1.34 (0.66)
C_6H_{10}	−0.75 (−0.41)	−0.88 (−0.58)	−1.08 (−0.63)	−0.83 (−0.48)	−0.21 (0.14)	−0.71 (−0.35)
$C_{10}H_{16}$	2.07 (1.20)	3.60 (2.69)	0.45 (−0.22)	−0.20 (−1.02)	−0.41 (−1.24)	1.70 (0.97)
AAD	1.09 (0.54)	1.72 (1.16)	0.59 (0.53)	0.48 (0.76)	0.57 (0.95)	0.88 (0.46)
ϵ_C[b]	−38.12912 (−38.13538)	−38.12007 (−38.12603)	−38.12599 (−38.13192)	−38.12594 (−38.13193)	−38.08326 (−38.08924)	−38.12166 (−38.12764)
ϵ_H[b]	−0.58100 (−0.59186)	−0.57302 (−0.58363)	−0.57812 (−0.58881)	−0.57764 (−0.58829)	−0.57464 (−0.58527)	−0.57251 (−0.58308)
lycopene[c]	118.79 (109.61)	131.15 (121.95)	108.50 (99.97)	101.59 (92.70)	94.72 (85.74)	112.66 (103.73)

[a] Values for the test set molecules are in deviations from G3 with atomization scheme. Results obtained using electronic energies are in parentheses. [b] In hartrees. [c] Values are enthalpies of formation predicted using AEs of various functionals.

TABLE 4: ΔH_f Values of Polyenes Predicted Using Group Equivalent (GE) at Various Levels of Density Functional Theories[a]

	B3LYP	BLYP	BPW91	MPW1PW91	PBEPBE	BB95
P_2	0.08 (0.10)	−0.09 (−0.06)	1.15 (−0.22)	−0.28 (−0.26)	−0.10 (−0.08)	−0.09 (−0.05)
P_3	0.29 (0.28)	0.28 (0.29)	−0.29 (0.21)	0.16 (0.17)	0.30 (0.32)	0.33 (0.36)
P_4	0.33 (0.32)	0.46 (0.47)	0.07 (0.43)	0.41 (0.42)	0.52 (0.52)	0.46 (0.47)
P_5	0.33 (0.31)	0.51 (0.49)	0.30 (0.52)	0.48 (0.47)	0.58 (0.55)	0.49 (0.46)
P_6	0.28 (0.29)	0.37 (0.39)	0.33 (0.47)	0.42 (0.42)	0.46 (0.46)	0.44 (0.42)
P_7	0.05 (0.08)	0.25 (0.28)	0.12 (0.14)	0.10 (0.09)	0.10 (0.10)	0.13 (0.12)
P_8	−0.19 (−0.22)	−0.20 (−0.23)	−0.24 (−0.40)	−0.10 (−0.12)	−0.47 (−0.51)	−0.24 (−0.32)
P_9	−0.32 (−0.34)	−0.79 (−0.75)	0.01 (−0.21)	−0.31 (−0.32)	−0.38 (−0.39)	−0.51 (−0.53)
C_5H_8	−0.86 (−0.84)	−1.02 (−0.99)	−1.80 (−1.37)	−1.31 (−1.29)	−0.92 (−0.90)	−0.91 (−0.93)
C_5H_{10}	−0.35 (−0.36)	−0.49 (−0.50)	1.48 (1.61)	1.75 (1.73)	−0.60 (−0.63)	−0.53 (−0.61)
C_6H_{12}	0.46 (0.29)	0.24 (0.08)	−0.15 (−0.38)	−0.17 (−0.37)	0.70 (0.51)	0.62 (0.39)
C_6H_{10}	−0.24 (−0.16)	−0.11 (−0.02)	−0.96 (−0.64)	−0.77 (−0.68)	−0.36 (−0.27)	−0.31 (−0.20)
$C_{10}H_{16}$	0.14 (0.28)	0.41 (0.55)	−0.03 (−0.16)	−0.44 (−0.27)	0.16 (0.32)	0.17 (0.40)
AAD	0.30 (0.30)	0.40 (0.39)	0.53 (0.52)	0.52 (0.51)	0.43 (0.43)	0.40 (0.40)
ϵ_{CH_4}[b]	−40.50471 (−40.45718)	−40.41857 (−40.46493)	−40.43947 (−40.48705)	−40.43699 (−40.48352)	−40.38065 (−40.42709)	−40.41486 (−40.46114)
$\epsilon_{C_2H_4}$[b]	−78.64043 (−78.58711)	−78.53979 (−78.59175)	−78.56564 (−78.61894)	−78.56302 (−78.61516)	−78.46369 (−78.51574)	−78.53709 (−78.58899)
$\epsilon_{C_2H_6}$[b]	−79.82306 (−79.74685)	−79.68220 (−79.75652)	−79.72133 (−79.79665)	−79.71803 (−79.79263)	−79.61365 (−79.68810)	−79.68032 (−79.75884)
lycopene[c]	102.09 (101.64)	103.92 (103.54)	104.37 (100.52)	99.57 (99.24)	99.65 (99.24)	99.43 (98.80)

[a] Values for the test set molecules are in deviations from G3 with atomization scheme. Results obtained using electronic energies are in parentheses. [b] In hartrees. [c] Values are enthalpies of formation predicted using GEs of various functionals.

TABLE 5: ΔH_f Values of $C_{40}H_{56}$ Carotenes Predicted Using the GE Scheme with Density Functional Theories[a]

	B3LYP	BLYP	BPW91	MPW1PW91	PBEPBE	BB95
lycopene[b]	102.09(101.64)	103.92(103.54)	104.37(100.52)	99.57(99.24)	99.65(99.24)	99.43(98.80)
prolycopene[b]	114.28(113.22)	116.01(115.09)	117.24(112.95)	111.74(110.91)	111.85(110.98)	111.89(110.47)
β-carotene[b]	89.81(87.62)	103.32(101.35)	91.56(86.68)	84.61(82.59)	81.89(79.83)	87.20(84.59)
α-carotene[b]	91.33(89.38)	105.50(103.79)	94.55(89.29)	86.96(85.15)	84.18(82.37)	89.33(87.23)
β-carotene[c]	80.05(82.95)	87.59(90.59)	89.15(87.00)	83.44(86.42)	84.78(87.72)	79.46(81.72)
α-carotene[c]	81.57(84.71)	89.77(93.03)	92.13(89.61)	85.76(88.98)	87.06(90.26)	81.59(84.35)

[a] Results obtained using electronic energies are in parentheses. [b] Isodesmic reaction using eq 3. [c] Isodesmic reaction using eq 6.

average absolute deviations. With respect to AE, most functionals (except BLYP) have AAD of smaller than 1 kcal/mol. Among them, AAD of MPW1PW91 is the smallest (0.48 kcal/mol). Not only are the AADs are very small butthe deviations obtained from the AE scheme are also scattered. Error accumulation is not seen. Within the test set, the largest errors occur in P_2, C_5H_8, and $C_{10}H_{16}$.

In addition to the usage of computed enthalpies, the fitting process was repeated excluding zero-point energy and thermal corrections, where the results are shown in parentheses. We see that estimations via AE using electronic energies provide predictions that are not at all inferior to those obtained using computed enthalpies. In several cases (B3LYP, BLYP, BPW91, and BB95), the AADs of results using electronic energies are smaller than those using computed enthalpies. This observation is perhaps not surprising because the approach of using atomic equivalents involves translating computed energies or enthalpies into enthalpies of formation with two parameters. From this point of view, ZPE and thermal corrections may be considered as approximately proportional to molecular size and their total energies. Interestingly, our test calculations show that quite accurate fit can be achieved with HF, MP2, and B3LYP electronic energies, even with the minimum STO-3G basis set (see the Supporting Information).

The ΔH_f values of lycopene predicted using AE are still quite scattered, however, ranging from 94.72 (PBEPBE) to 131.15 kcal/mol (BLYP). Thus, despite the excellent agreement within the test set, we are not able to justify the application of such an approach in the evaluation of the enthalpies of formation for carotenes. Unlike in the training set, the predicted ΔH_f values using electronic energies are ~9 kcal/mol smaller than those obtained using computed enthalpy for all functionals.

Group equivalent predicted enthalpies of formation (see Table 4) that are even more accurate than those from atomic equivalents. The most accurate results are obtained from B3LYP, with an AAD of only 0.30 kcal/mol. BLYP and BB95 results are quite accurate: the AADs are 0.40 kcal/mol. The AADs using electronic energies instead of computed enthalpies are nearly the same. Similar to the case in the AE approach, the deviations from G3 data are scattered and accumulation of errors is not involved. The largest error occurs in C_5H_8. Overall, we see that predictions within the GE scheme agree better with the G3 results than the AE approach.

In contrast, the ΔH_f values of lycopene predicted using the GE approach are much less scattered than those predicted using AE. The smallest ΔH_f predicted for lycopene is 99.43 kcal/mol (BB95), and the largest ΔH_f is 104.37 kcal/mol. Not only does the DFT predicted ΔH_f lie within a small range but the predictions using computed electronic energies are also very close to those using computed enthalpies. Three of the functionals (MPW1PW91, PBEPBE, BB95) provide nearly identical results. Considering the fact that the functional forms for the exchange and correlation energies of these density functionals were developed independently, we would propose that DFT, along with the GE approach, is able to provide accurate estimates for the ΔH_f values of lycopene and related $C_{40}H_{56}$ carotenes. We estimate the ΔH_f of lycopene to be 99−104 kcal/mol.

The predicted ΔH_f values of lycopene, prolycopene, β-carotene, and α-carotene are summarized in Table 5. For prolycopene, the predicted values of various DFT functionals are reminiscent of the results in lycopene, whereas the ΔH_f values of lycopene are ~12 kcal/mol higher. Again, in prolycopene three of the functionals (MPW1PW91, PBEPBE, BB95) provide nearly identical results. Similarly, we see that the ΔH_f values predicted for prolycopene using electronic energy and computed enthalpy are very similar with most DFT functionals (except BPW91).

Much more widely distributed predictions are seen in β-carotene and α-carotene. For β-carotene, the deviation is more than 20 kcal/mol, ranging from the smallest PBEPBE value (81.89 kcal/mol) to the largest BLYP value (103.32 kcal/mol). The deviation can be attributed to the ability of describing relative energies between the α and β ring structures and their open-chain counterparts. Our GE scheme for carotenes was designed with eq 3, with which the formal linkage between all non-hydrogen atoms of the lycopene (and prolycopene) are retained. In the case of carotenes with ring end groups (α-carotene and β-carotene), 11 C═C and 30 C−C bonds are involved. The bond separation scheme of eq 6 should be considered:

$$C_{40}H_{56} + 42CH_4 \rightarrow 11C_2H_4 + 30C_2H_6 \qquad (6)$$

Although both eqs 3 and 6 are uniformly defined bond separation reactions, the ΔH_f values of α-carotene and β-carotene computed using eq 6 differ from those obtained by using their relative enthalpies (or simply, by using eq 3) with respect to that of lycopene. The difference of B3LYP for example, is ~10 kcal/mol with two equations. This poses a dilemma in deciding which equation to use. At this stage we suggest to use eq 3 for all carotenes, from which the ΔH_f values for α-carotene and β-carotene are the relative energies with respect to that of lycopene. The accuracy of DFT thus relies on their ability to account for the relative energies of compounds with open or ring end groups.

Semiempirical and Parametrized Methods. Carotenes are a large system for high-level quantum chemistry theories; thus, it is intriguing to examine the reliability of semiempirical and parametrized methods. In Table 6 we summarized AM1 and PM3 computed ΔH_f values. In addition to our fitted AEs and GEs, Jorgensen's parametrizations for AE and BGE were computed.[22] The enthalpies of formations predicted using Leal's parametrizations are also included.[36]

The AADs summarized in Table 6 revel that for all of the schemes applied AM1 predictions compare better with the G3 results than PM3. The AE and GE schemes in this study can provide fair predictions within the test set. The GE scheme performs better than the AE scheme, wherein the AADs of the GE scheme with AM1 and PM3 theories are 0.59 and 0.61 kcal/

TABLE 6: ΔH_f Values of Polyenes Predicted Using Semiempirical Methods[a]

		AM1					PM3				
	Leal[b]	standard	AE[d]	BGE[d]	AE[e]	GE[e]	standard	AE[d]	BGE[d]	AE[e]	GE[e]
P_2	−0.64	3.22	2.80	−3.23	1.90	0.09	4.34	3.61	−3.38	2.75	0.17
P_3	−0.47	2.78	1.78	−3.44	0.61	−0.82	4.93	3.92	−3.02	1.35	−0.69
P_4	−0.08	4.52	2.93	−1.47	1.50	0.45	7.81	6.52	−0.38	2.23	0.74
P_5	0.48	4.86	2.68	−0.90	0.99	0.32	9.04	7.47	0.62	1.48	0.52
P_6	1.03	5.33	2.56	−0.20	0.62	0.32	10.44	8.59	1.78	0.88	0.46
P_7	1.50	5.63	2.28	0.33	0.07	0.16	11.64	9.51	2.74	0.09	0.21
P_8	1.97	6.88	1.94	0.81	−0.61	−0.14	12.85	10.44	3.71	−0.77	−0.11
P_9	2.58	6.15	1.61	1.29	−1.12	−0.27	14.00	11.31	4.63	−1.53	−0.33
C_5H_8	−0.55	0.87	0.54	−3.49	−0.69	−1.78	1.74	0.78	−3.38	0.35	−1.21
C_5H_{10}	0.76	1.65	2.07	1.23	0.46	0.42	−1.26	−2.40	−2.09	−0.26	−0.31
C_6H_{12}	0.61	0.54	1.05	1.46	−0.89	−0.22	−0.30	−1.65	−1.06	0.91	1.88
C_6H_{10}	−1.19	4.25	4.01	−1.22	2.45	2.08	1.57	0.38	−2.54	0.38	−0.16
$C_{10}H_{16}$	−1.34	1.00	0.34	−3.27	−2.12	−0.69	−0.53	−2.45	−1.28	−3.31	−1.26
AAD	0.97	3.70	2.21	1.58	1.07	0.59	6.05	5.20	2.42	1.25	0.61
lycopene[c]	87.26	82.13	76.44	81.54	74.24	86.62	78.06	79.66	96.53	69.95	83.61
prolycopene[c]	na	86.36	86.55	91.65	84.36	96.74	79.63	84.04	100.91	74.34	88.00
β-carotene[c]	57.25	58.38	57.38	63.18	56.38	68.77	52.96	57.38	75.24	47.67	61.33
α-carotene[c]	61.36	62.18	62.38	67.48	60.18	72.56	56.77	61.19	78.06	51.48	65.14

[a] Values are in deviations from G3 [b] Empirical parameters from ref 36. [c] Values are enthalpies of formation. [d] Equivalents from ref 22. [e] Equivalents from this study. Refer to the Supporting Information for the equivalents.

mol, respectively. Compared with the DFT predictions, the agreement of the semiempirical method is inferior, indicating that the errors of DFT theories are likely to be more systematic (and to be corrected by parametrization) than those of semiempirical theories.[38] The GE scheme predicted ΔH_f values of lycopene using AM1 and PM3 are 86.62 and 83.61 kcal/mol, respectively. These values are significantly smaller than those predicted using DFT (99−104 kcal/mol). A surprisingly good result is seen with the empirical parameters of Leal,[36] where an AAD of 0.97 kcal/mol is achieved. The ΔH_f of lycopene predicted with Leal's parameters is 87.26 kcal/mol.

In general, the enthalpies of formation predicted using semiempirical and empirical methods agree fairly with the G3 atomization results. The predicted ΔH_f values for carotenes using empirical and semiempirical methods are, however, significantly smaller than those predicted using DFT. The order of enthalpies of formation of the $C_{40}H_{56}$ carotenes are in the same order as those predicted using DFT, that is,: β-carotene < α-carotene < lycopene < prolycopene. The GE and AE approaches described in this study were based on comparisons to the G3 predicted enthalpies of formation. However, the methodology described here is applicable to data sets within which experimental data are available. The strategy of a methodology to better predict heats of formation for carotenes involving different side groups is currently under investigation.

4. Conclusions

Studies for the enthalpies of formation for polyenes and $C_{40}H_{56}$ carotenes were studied. Because of the lack of experimental data, we used the G3 results via the atomization scheme to evaluate the accuracy of DFT and semiempirical methods. We have demonstrated that the application of DFT, along with group equivalents (GEs) and atomic equivalents (AEs) is able to provide accurate predictions for ΔH_f. The predictions of DFT with GE is the most reliable, particularly with the B3LYP functional. For carotenes that do not possess ring end groups, the predicted ΔH_f values for carotenes from various DFT functionals are very close. Semiempirical AM1 and PM3 approaches with the GE scheme agree reasonably with G3 atomization. However, when they are applied on carotenes, the ΔH_f values are significantly smaller than the DFT results. The empirical additivity parameters used by Leal were shown to be quite accurate within the test set; however, the predicted ΔH_f value is also much smaller than that from DFT. In all test cases, the GE scheme provides predictions that are superior to the AE scheme and is recommended to be used with DFT for the computations of the ΔH_f values of carotenes.

Acknowledgment. We acknowledge that the National Science Council of Taiwan, Republic of China, supported this work. We also thank the National Center for High-Performance Computing for computer time and facilities.

Supporting Information Available: Selected optimized geometrical parameters of compounds and data for the AE and GE fits using minimum basis sets. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Burton, G. W.; Ingold, K. U. *Science* **1984**, *224*, 569.
(2) Britton, G. *FASEB J.* **1995**, *9*, 1551.
(3) Sies, H.; Stahl, W. *Am. J. Clin. Nutr.* **1995**, *62* (suppl), 1315S.
(4) Krinsky, N. I. *Annu. Rev. Nutr.* **1993**, *13*, 561.
(5) Edge, R.; McGarvey, D. J.; Truscott, T. G. *J. Photochem. Photobiol., B* **1997**, *41*, 189.
(6) Bohm, F.; Haley, J.; Truscott, T. G.; Schalch, W. *J. Photochem. Photobiol., B* **1993**, *21*, 219.
(7) Peto, R.; Doll, R.; Buckley, J. D.; Sporn, M. B. *Nature* **1981**, *290*, 201.
(8) Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A., Jr. *J. Chem. Phys.* **1996**, *104*, 2598.
(9) Mayer, P. M.; Parkinson, C. J.; Smith, D. M.; Radom, L. *J. Chem. Phys.* **1998**, *108*, 604.
(10) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221.
(11) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.
(12) Temelso, B.; Sherrill, C. D.; Merkle, R. C.; Freitas, R. A., Jr. *J. Phys. Chem. A* **2006**, *110*, 11160.
(13) Hehre, W. J.; Ditchfield, R.; Radom, L.; Pople, J. A. *J. Am. Chem. Soc.* **1970**, *92*, 4796.
(14) Redfern, P. C.; Zapol, P.; Curtiss, L. A.; Raghavachari, K. *J. Phys. Chem. A* **2000**, *104*, 5850.
(15) Check, C. E.; Gilbert, T. M. *J. Org. Chem.* **2005**, *70*, 9828.
(16) Wodrich, M. D.; Corminboeuf, C.; Schleyer, P. v. R. *Org. Lett.* **2006**, *8*, 3631.
(17) Wodrich, M. D.; Wannere, C. S.; Mo, Y.; Jarowski, P. D.; Houk, K. N.; Schleyer, P. v. R. *Chem.—Eur. J.* **2007**, in press.
(18) Grimme, S. *Angew. Chem., Int. Ed.* **2006**, *45*, 4460.
(19) Cohen, N.; Benson, S. W. *Chem. Rev.* **1993**, *93*, 2419.
(20) Wodrich, M. D.; Schleyer, P. v. R. *Org. Lett.* **2006**, *8*, 2135.

(21) Gronert, S. *J. Org. Chem.* **2006**, *71*, 1209.

(22) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L. *J. Comput. Chem.* **2002**, *23*, 498.

(23) Mole, S. J.; Zhou, X.; Liu, R. *J. Phys. Chem.* **1996**, *100*, 14665.

(24) Baboul, A. G.; Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **1999**, *110*, 7650.

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

(26) Afeefy, H. Y.; Liebman, J. F.; Stein, S. E. "*Neutral Thermochemical Data*" in NIST Chemistry WebBook, NIST Standard Reference Database Number 69; Linstrom, P. J., Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg, MD; June 2005 (http://webbook.nist.gov).

(27) Becke, A. D., *J. Chem. Phys.* **1993**, *98*, 5648.

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

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

(30) Perdew, J. P.; Burke, K.; Wang, Y. *Phys. Rev. B* **1996**, *54*, 16533.

(31) Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664.

(32) Perdew, J. P.; K. Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

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

(34) Dewar, M.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899.

(35) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221.

(36) Leal, J. P. *J. Phys. Chem. Ref. Data* **2006**, *35*, 55.

(37) 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 C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(38) Thiel, W. *Quantum-Mechanical Prediction of Thermochemical Data*; Kluwer Academic Publishers: Boston, MA, 2001; pp 235−245.