J. Phys. Chem. A **2000,** *104,* 4497−4504

Application of the Group Additivity Method to Alkyl Radicals: An ab Initio Study

István Marsi, Béla Viskolcz, and László Seres*

Department of Chemistry, J. Gy. Teachers' Training College, H-6701 Szeged, P.O.B. 396, Hungary

Received: July 14, 1999; In Final Form: February 16, 2000

The geometries and standard enthalpies of formation ($\Delta_fH°$) of 4 paraffins and 31 alkyl radicals at 298 K were obtained by means of complete basis set 4 (CBS-4) ab initio molecular orbital theory. A comparison of the geometry of a molecule and the radicals derivable from it via elimination of a H atom from different positions suggests that only a negligible change occurs in the geometry of the groups immediately adjacent to the radical center. The CBS-4 $\Delta_fH°$s scaled by an empirical correction display an average deviation of 0.4 kcal mol^{-1} from the experimental data. Group values (GVs) were evaluated from the corrected ab initio $\Delta_fH°$s by simultaneous parameter estimation. The group additivity (GA) rule was checked by a statistical analysis of the theoretical data. The derived GVs describe the corrected ab initio $\Delta_fH°$s with an average deviation of less than 0.3 kcal mol^{-1}, which indicates that the GA rule is applicable to alkyl radicals, and the $\Delta_fH°$s of alkyl radicals can be predicted with an accuracy similar to that attained with existing experimental methods. New values of radical, radical-adjacent, and $(C)_2{-}C{-}\dot{C}{-}(C)$ gauche GVs are suggested; together with the accepted alkane GVs, these yield accurate $\Delta_fH°$s of alkyl radicals within estimated errors of 0.5 kcal mol^{-1}. Application of the GVs derived from small species is predicted to perform less satisfactorily in the estimation of $\Delta_fH°$s of $C_4{-}C_6$ species: the average deviation increases from 0.3 to 0.6 kcal mol^{-1}, while the correlation of the individual deviations and carbon numbers of the alkyl radicals changes from nearly zero to −0.36.

1. Introduction

A knowledge of the thermochemistry of short-lived species such as free radicals is important for an understanding of the complex chemical processes of organic chemistry,[1] combustion chemistry,[2] etc.

Application of the group additivity (GA) method (and especially the readily applicable version suggested by Benson et al.[3−5]) has provided a powerful means of studying chemical reactions on the basis of thermochemistry, alleviated planning of experimental studies, evaluation of experimental results, estimation of the equilibrium and rate constants needed by simulation studies in olefin chemistry, oxidation, air pollution, etc. The transferable group values (GVs)[5] relating to molecules are based on well-established experimental data on many compounds.[6] The GVs estimated from the best data on small compounds, and those derived from a greater number of compounds, i.e., including larger molecules in which the contributions may be somewhat different, are within experimental error.[7−9]

The situation is much less favorable in the case of radicals, where the method has been used essentially without any justification, even for the best-known representatives of this group, the alkyl radicals. Here, because of the scarcity of data, only the data on the archetypes of the alkyl radicals, i.e., ethyl, *n*-propyl, isopropyl, *tert*-butyl, isobutyl, and neopentyl, were used in the derivation of the GVs.[5,7−8] Consequently, no comparison can be made with the GVs derived from larger radicals. It is not self-evident, for example, that the relaxation caused by the appearance of a $\dot{C}(H)(C)_2$ or $\dot{C}(C)_3$ group within a chain longer than isopropyl or *tert*-butyl does not result in a change in the enthalpy of formation ($\Delta_fH°$) beyond the usual errors of experimental studies. Consequently, in the application of the method to any of the larger radicals, further GVs are needed. As an example, in the estimation of $\Delta_fH°[(CH_3)_2CH\dot{C}HCH_3]$, the gauche interaction of the methyl groups is expected to be different from that in the alkane.

Since gauche contributions are defined as the increments due to the interactions of non-trans substituents at positions 1 and 2, no gauche contributions to $\Delta_fH°$ can be estimated from the simplest radicals. (This may result in unknown errors when the GA method is applied to branched radicals.) Whenever we wish to make thermochemical calculations on elementary reactions involving branched alkyl radicals $\geq C_5$, we inevitably face the problem of gauche interactions. (The only exception is neopentyl.)

In medium-sized alkyl radicals, two types of such interactions are to be considered: the interactions present in the structures $(C)_2{-}C{-}\dot{C}{-}C$ [or $(C)_2{-}\dot{C}{-}C{-}C$] and $(C)_2{-}C{-}C{-}\dot{C}$ [or $\dot{C}{-}C(C){-}C{-}C$], which we call the radical-gauche 1 (RG1) interaction and the radical-gauche 2 (RG2) interaction, respectively (see Figure 1). In the case of RG1, the gauche interactions of an alkyl group attached to an sp^3 C atom with an alkyl group attached to an sp^2 C atom are taken into account, while in the RG2 interaction, an alkyl group and a group in which the polyvalent atom is a C radical center are attached to sp^3 C atoms.

In this simplifying approach, following Benson's procedure[5] for the treatment of alkane gauche (AG) interactions, no consideration is given to the differences in the interactions of methyl, methylene or methyne groups; i.e., they are regarded as the same.

Since the availability of experimental data on large alkyl radicals is not expected to increase sufficiently to allow estimations as reliable as those on organic compounds in the

---

* Author to whom correspondence should be addressed. E-mail: seresl@jgytf.u-szeged.hu.

10.1021/jp992355e CCC: $19.00

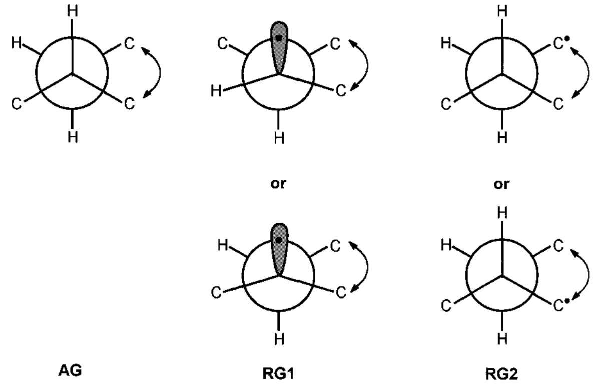


Figure 1. Newman projection of gauche interactions in alkyl radicals.

foreseeable future, we set out to perform theoretical calculations to check whether the GA method is also applicable to larger radicals.[10]

Although ab initio quantum chemical methods are suggested to predict the absolute values of the enthalpies of formation of compounds only within an error of about 2 kcal mol^{-1} [11] in many cases, they perform much better in the estimation of enthalpy differences. Theoretical studies of the transition state structures involved in alkyl[12] and alkoxy[13] radical isomerizations suggest that one element of the GA rules, the ring strain energy, can be predicted reliably[14] for such reactions. Thermodynamic properties of fluorinated carbon groups were recently derived by means of ab initio studies.[15]

In this paper, we report the results of calculations in which empirically corrected ab initio $\Delta_f H^\circ$s of alkyl radicals were obtained. On the basis of these data, a conclusion was drawn as to the applicability of the GA rule to the alkyl radicals. Also, the problem of using GVs derived from the archetype radicals for the estimation of $\Delta_f H^\circ$s of radicals with greater numbers of C atoms is discussed.

It was hoped that the calculated data on molecules and radicals larger than the above-mentioned six would permit the derivation of GVs which allow the respective data on larger radicals not included in the present work to be estimated with reasonable errors. It was also hoped that some of the new GVs not available from experimental data would be accurate enough to be used in the thermochemical calculations until more reliable experimental data become available.

The applicability of the GA method and the new GVs derived will be discussed in terms of the computational methods on which the present study is based.

2. Calculations

(**a**) **Quantum Chemical Calculations**. The complete basis set 4 (CBS-4) method of Petersson et al.,[16−19] which uses a basis set extrapolation to estimate residual energy errors, was applied to all species. The geometries and harmonic vibrational frequencies were calculated with HF/3-21G*. The corrections for molecules and radicals were made by using the scaled harmonic frequencies for the vibrational energy, the classical approximations for translation ($^3/_2RT$) and rotation ($^3/_2RT$ for nonlinear molecules and RT for linear molecules) and the PV term. All ab initio calculations were performed with the GAUSSIAN 94 program.[20]

The theoretical enthalpies of formation at 298.15 K were calculated from the enthalpy of the following hypothetical reaction:

$$nC + (n + {}^1/_2)H_2 \rightarrow C_nH_{2n+1} \qquad (A)$$

$$\Delta_f H^\circ(C_nH_{2n+1},g) = H^\circ(C_nH_{2n+1},g) - n[H^\circ(C_{graphite},s) - \Delta_f H^\circ(C,g)] - (n + {}^1/_2)H^\circ(H_2,g) \qquad (1)$$

where the H°s are the total energies with thermal corrections to enthalpies at 298.15 K of the corresponding species, and $\Delta_f H^\circ(C,g) = 171.29 \pm 0.1$ kcal mol^{-1} is the elemental correction for the C atom, taken directly from standard tables.[6,21,22]

(**b**) **Calculation of the Group Contributions**. In alkyl radicals, the 14 chemically realistic GVs listed in Table 1 can reasonably be distinguished.

If $\boldsymbol{h}$ denotes the vector of standard enthalpies of formation ($\boldsymbol{h}(\Delta_f H^\circ, ..., \Delta_f H^\circ)$), and $\boldsymbol{g}$ the vector of the GVs ($\boldsymbol{g}(g_1, g_2, ..., g_{14})$), their dependence expressed in GA rules can be formulated as follows:

$$\boldsymbol{h} = \boldsymbol{X}\boldsymbol{g} \qquad (2)$$

where $(\boldsymbol{X})_{ij}$ denotes the occurrence of the jth group in the ith species.

Multivariate linear regression (MVLR) seemed to be appropriate to derive the individual GVs from the ab initio $\Delta_f H^\circ$s. The scarcity of input $\Delta_f H^\circ$s generally precludes the application of MVLR, although the latter is expected to provide more reliable results[23] than successive evaluation of GVs from eq 2, i.e., treating an appropriate subset of eq 2 as a system of linear equations. Further, MVLR allows a statistical proof of the results of the simultaneous estimation, revealing whether the assumption of the GA rule is valid and, if so, to what extent. [It should be noted here that unweighted MVLR was applied in the calculations, in consequence of the lack of reasonable weights.] Presumably, the inclusion of non-radical-adjacent alkyl groups in the simultaneous estimation does not result in a significant revision of these GVs reliably derived from alkanes. However, it can be informative as to whether the present calculations can reproduce them (as expected), and the errors of the individual GVs estimated in this way can then be compared. Thus, as a first approximation, an attempt will be made to estimate as many GVs as possible on a consistent database.

Unfortunately, the estimation of each of the 14 GVs fails, a problem that unavoidably occurs for any set of alkyl radicals. The explanation of this is connected with the perception of O'Neal and Benson[4] that each of the three alkyl radical GVs and those of four alkyl groups adjacent to the radical center cannot be determined in consequence of their linear dependence. In other words, the rank of an $\boldsymbol{X}'_{m\times 7}$ matrix containing 7 columns referring to the above 7 groups is $<$7. The reasonable number of rows (corresponding to the number of species) in $\boldsymbol{X}'$ is 34, as $\sum_{i=1}^{3}\binom{3+i}{i} = 34$. (All possible combinations of *primary*, *secondary*, and *tertiary* alkyl radical centers, and *primary*, *secondary*, *tertiary*, and *quaternary* radical-adjacent alkyl groups have to be considered.) Any further radical in $\boldsymbol{X}'$ appears as an obviously linearly dependent row. Either elementary calculations, as proposed in ref 4, or systematic computer testing of the possible $\binom{34}{7}$ choices of $\boldsymbol{X}'_{7\times 7}$ minors of $\boldsymbol{X}'_{m\times 7}$ demonstrates that rank $(\boldsymbol{X}'_{7\times 7}) < 7$. When further, e.g., structural features of radicals expressed in additional GVs are considered, new columns have to be entered into $\boldsymbol{X}'$. However, this does not decrease the difference between the number of columns and the rank of the matrix. Thus, rank$(\boldsymbol{X}_{m\times 14}) < 14$ in eq 2, for any m, irrespective of the nature of the species. The root of the problem is obviously the attempt to evaluate radical GVs and GVs of alkyl groups adjacent to the radical center simultaneously.

This difficulty cannot be avoided either if MVLR is applied. A typical indication of the failing of MVLR is the (near)

TABLE 1: Overview of Groups and Gauche Interactions in Alkyl Radicals

	alkyl groups	radical adjacent alkyl groups	radical groups
primary	$C-(H)_3(C)$	$C-(H)_3(\dot{C})$	$\dot{C}-(H)_2(C)$
secondary	$C-(H)_2(C)_2$	$C-(H)_2(\dot{C})(C)$	$\dot{C}-(H)(C)_2$
tertiary	$C-(H)(C)_3$	$C-(H)(\dot{C})(C)_2$	$\dot{C}-(C)_3$
quaternary	$C-(C)_4$	$C-(\dot{C})(C)_3$	

gauche interactions	relevant structures
AG	$(C)_2-C-C-C$
RG1	$(C)_2-C-\dot{C}-C$ or $(C)_2-\dot{C}-C-C$
RG2	$(C)_2-C-C-\dot{C}$ or $\dot{C}-C(C)-C-C$

singularity of $\boldsymbol{X}^T \boldsymbol{X}$ in the expression for $\hat{\boldsymbol{g}}$, i.e., the optimum of $\boldsymbol{g}$:

$$\hat{\boldsymbol{g}} = (\boldsymbol{X}^T \boldsymbol{X})^{-1}(\boldsymbol{X}^T \boldsymbol{h})$$

Let us consider that the rank of the product of matrices cannot be greater than the minimum of the ranks of the multiplicands. Thus, all of the 14 GVs cannot be estimated.

Consequently, the approximate identity of a pair formed from an alkyl group and a radical-adjacent alkyl group has to be presumed. $C-(H)_3(\dot{C}) = C-(H)_3(C)$ is generally proposed as the most sustainable choice from a chemical point of view. Unfortunately, the computational limitations involved meant that exclusively $C_{n\leq 6}$ alkyl radicals could be incorporated in the present ab initio calculations. With regard to the sparse occurrence of $C-(\dot{C})(C)_3$ in the investigated $C_{n\leq 6}$ species, it is preferred to apply $C-(\dot{C})(C)_3 = C-(C)_4$. The total number of $C_{2\leq n\leq 6}$ alkyl radicals is 32. From this complete data set, only $\dot{C}H_2C(CH_3)_2CH_2CH_3$ and $CH_3C(CH_3)_2\dot{C}HCH_3$ were disregarded. The latter species would distinguishably improve only the identifiability of the $C-(\dot{C})(C)_3$ group, which was otherwise not estimated.

The identifiability of the GVs was checked in detail via eigenvalue−eigenvector analysis of the matrix $\{g_i g_j(\partial^2 S/\partial g_i\, \partial g_j)\}_{\boldsymbol{g}=\hat{\boldsymbol{g}}}$ composed of the normed second derivatives with respect to g_i and g_j of the sum of squares (S) taken at the (approximate) optimum $\boldsymbol{g} = \hat{\boldsymbol{g}}$.[24] In this task, $\hat{g}_i\hat{g}_j\ \partial^2 S/\partial g_i\, \partial g_j = \hat{g}_i\hat{g}_j\sum_{k=1}^{30} x_{ki}x_{kj}$, $i, j = 1, ..., 13$.

If the MVLR is based on the above 30 radicals, the identifiability of 12 GVs is satisfactory, but that of $C-(C)_4$ is poor. A check was made as to whether the inclusion of further species improved the identifiability. The task becomes treatable through the inclusion of a single alkane, preferably C_2H_6. (It is somewhat surprising that the inclusion of C_2H_6 is recommended to improve the identifiability of $C-(C)_4$ rather than that of neopentane. This is probably due to the manifold co-occurrence of $C-(H)_3(C)$, determined primarily via C_2H_6.) The inclusion of further (either radical or alkane) species is beneficial, but it results in rapidly decreasing further improvements in identifiability. Finally, a database incorporating 30 alkyl radicals and 4 alkanes was selected for estimation (Table 2). Although the numbers of radicals and alkanes involved differ by almost an order of magnitude, this does not depreciate the identifiability even of alkane GVs as $\Delta_f H^\circ$s of radicals can serve as a source of alkane, radical, and radical-adjacent GVs alike. All reasonable combinations were systematically tested with respect to identifiability. In conclusion, the selected set can be considered the optimal one as concerns databases of this size.

3. Results

(**a**) **Geometries**. The conformations of the alkyl chains in the radicals are similar to those in alkanes. In radicals, the most stable conformation of the chain is also the most extended one, i.e., the one in which the C atoms of the backbone of the chain are in a plane, just as in the alkanes. One of the methyl C atoms at the branching of the chain is out of this plane, of course. The regular alkane structure is perturbed at the radical center where the sp^2-like hybrid is located. The α and β bonds are shorter and longer, respectively, than the same types of bonds in the other parts of the radical. Similar observations have been made for different alkyl radicals (Pacansky et al.[25]), branched heptyl radicals (Viskolcz et al.[26]), and other alkyl radicals (Viskolcz et al.[12]).

The equilibrium geometries of the different conformers are shown in Table 3. The smallest distances between two H atoms (2.22−2.29 Å) were found for AG interactions. This relative closeness of H atoms may cause a significant positive contribution to the energy of the radical. For the other two interactions (RG1 and RG2), the distances between the H atoms are larger (ca. 2.6 and 3.0 Å, respectively). The distances of two H atoms in the alkyl chain are ca. 2.5 and 2.6 Å in positions 1,2 and 1,3, respectively. These numbers suggest that the energy contributions of RG1 and RG2 can be expected to be small or negligible. Also, the distance between two C atoms in the 1,4 positions is the shortest for RG2, greater for AG, and the greatest for RG1 interactions. It may be assumed that these are the cause of the different energy contributions of RG1 and RG2, respectively.

(**b**) **Enthalpies of Formation**. The calculated $\Delta_f H^\circ$s of 31 radicals ($C_{n\leq 6}$) and 4 alkanes are listed in Table 2. The enthalpies of formation of simple alkyl radicals have been debated during the past few decades.[27−29] Recent experimental data from different laboratories [7,29−34] are included for comparison, where available.

The relative $\Delta_f H^\circ$s of *primary*, *secondary*, and *tertiary* radicals calculated by CBS-4 are in good agreement with the relative stabilities of the above radicals in all cases. The most stable radicals are the *tertiary* ones, and the *primary* and *secondary* radicals have $\Delta_f H^\circ$s ca. 5 and 3 kcal mol^{-1} higher, independent of the size of the radicals.

The CBS-4 $\Delta_f H^\circ$s agree with the experimental ones within 3 kcal mol^{-1}. (We have considered the higher values available in the literature.[27,31−33]) The agreement is better if the lower suggested values[8] are accepted. The calculated $\Delta_f H^\circ$s are generally significantly lower than the corresponding experimental values, especially if we choose the data of Tsang[27] or Seetula and Slagle[33] for comparison. Dobis and Benson[28,35] suggested that the $\Delta_f H^\circ$s obtained by these laboratories may be too high. The characteristic negative deviations of the experimental $\Delta_f H^\circ$s of alkanes suggest that the CBS-4 method underestimates $\Delta_f H^\circ$.

When making such comparisons, one has to consider the systematic sources of the deviations.

Experimental $\Delta_f H^\circ$s of paraffins are derived from heats of combustion and thus involve $\Delta_f H^\circ(CO_2,g)$ and $\Delta_f H^\circ(H_2O,l)$, which are not known to better than 0.05 kcal mol^{-1}. Also, $\Delta_{comb} H^\circ$(paraffin C,g) is probably not known to better than 0.1 kcal mol^{-1}. As a consequence, the absolute accuracy of experimental $\Delta_f H^\circ$s for molecules is restricted and the uncertainty increases with increasing number of C atoms.

Further, since ab initio calculations yield enthalpies of atomization at 0 K, $\Delta_f H^\circ_{298}(C,g)$ and $\Delta_f H^\circ_{298}(H,g)$ are not

TABLE 2: Ab Initio Calculated, Empirical Corrected, and Experimental Enthalpies of Formation of Alkanes and Alkyl Radicals (kcal mol^{-1}; 1 cal = 4.184 J)

	theory			experiment			
species	N	CBS-4	CBS-4 corrected	Seetula et al.[33]	Cohen, Benson[7]	Cohen, Benson[8]	Berkowitz et al.[29]
CH_3	0.5	35.24	34.69	34.9	35.1	35.1	34.8
C_2H_5	1.5	28.63	28.79	28.4	28.4	28.4	28.9
n-C_3H_7	2.5	23.27	24.14	24.09	23.4	23.4	
i-C_3H_7	2.5	20.12	20.99	20.70	20.0	21.0	21.5
1-C_4H_9	3.5	17.29	18.87	19.34	18.4		
2-C_4H_9	3.5	14.58	16.16	15.94	15.0	16.0	16.1
i-C_4H_9	3.5	16.02	17.60	17.38	16.0		
t-C_4H_9	3.5	10.70	12.28	12.38	9.4	11.0	12.3
1-C_5H_{11}	4.5	11.39	13.68	14.12[a]			
2-C_5H_{11}	4.5	8.62	10.91	10.7[a]			
3-C_5H_{11}	4.5	9.06	11.35	11.2[a]			
2-M-1-butyl	4.5	10.45	12.74	12.6[a]			
2-M-2-butyl	4.5	5.18	7.47	7.7[a]	5.0		
3-M-2-butyl	4.5	7.07	9.36	9.2[a]			
3-M-1-butyl	4.5	9.40	11.69				
neopentyl	4.5	7.43	9.72		8.9	9.2	
1-C_6H_{13}	5.5	5.21	8.22				
2-C_6H_{13}	5.5	2.51	5.52				
3-C_6H_{13}	5.5	2.74	5.75				
2-M-1-pentyl	5.5	4.20	7.21				
2-M-2-pentyl	5.5	−1.37	1.64	2.47[b]	0.0		
2-M-3-pentyl	5.5	1.21	4.22				
4-M-2-pentyl	5.5	0.25	3.26				
4-M-1-pentyl	5.5	3.73	6.74				
3-M-1-pentyl	5.5	3.86	6.87				
3-M-2-pentyl	5.5	1.51	4.52				
3-M-3-pentyl	5.5	−0.09	2.92		0.0		
2-ethyl-1-butyl	5.5	4.20	7.21				
2,3-DM-1-butyl	5.5	2.97	5.98				
2,3-DM-2-butyl	5.5	−1.85	1.16				
3,3-DM-1-butyl	5.5	0.88	3.89				
ethane	2	−20.75	−20.23			−20.03	
propane	3	−26.51	−25.28			−25.02	
i-butane	4	−34.51	−32.57			−32.07	
neopentane	5	−43.29	−40.64			−40.18	

[a] Estimated values. [b] Reference 34.

known to better than 0.1 kcal mol^{-1}. The largest correction needed involves the frequencies and their related zero-point energies.

These uncertainties add up rapidly and are very likely sources of systematic deviations with size. As a consequence, the results of ab initio calculations must be corrected to allow for the errors that are either independent of or dependent on the number of C atoms in the species.

The major objective of applied quantum chemistry is the prediction of molecular structure and energy. The smallest difference between experimental and calculated $\Delta_f H°$s was found for the methyl radical, and the largest one for neopentane. Evidently, there is a correlation between the size of a species and the deviation from the experimental $\Delta_f H°$ (see Figure 2). We suggest an empirical correction to reproduce the known experimental data, and by using such a correction predict theoretical GVs for the estimation of $\Delta_f H°$ for other radicals and molecules. This correction, similarly to the higher level of correction (HLC) method, is based on a very simple assumption: the combined number of spin-paired electrons differs on the two sides of eq A. An alkyl radical has only one unpaired electron, whereas the C atom in the triplet state has two unpaired electrons. If the numbers of spin-paired (β) and spin-unpaired (α) electrons are summed on the two sides of eq A, there is a difference. Since the energy of α electrons is somewhat higher than that of β electrons, this energy contribution due to pairing of spin-unpaired electrons may lead to theoretical $\Delta_f H°$s significantly lower than the experimental ones.

We assume a linear empirical correction depending on the difference in α and β electron pairs:

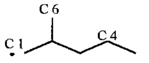
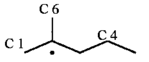
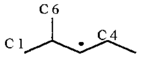
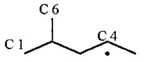
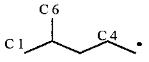
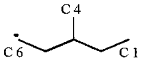
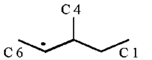
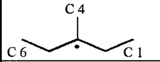
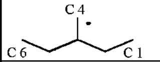
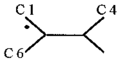
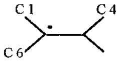
$$\Delta_f H^\circ_{\text{corr}} = \Delta_f H^\circ_{\text{CBS}-4} + A \times N + B$$

where A and B are adjusting parameters ($A = 0.7116$ and $B = -0.9075$ kcal mol^{-1}), and N is the negative difference between the numbers of spin-unpaired and spin-paired electron pairs on the two sides of eq A, listed in Table 2.

The corrected enthalpies of formation ($\Delta_f H^\circ_{\text{corr}}$) of radicals and molecules are likewise listed in Table 2. The relative stabilities of structural isomers did not change the correction. The corrected values are in good agreement with most experimental $\Delta_f H°$s. The deviations are <0.4 kcal mol^{-1} in all but four cases, as can be seen in Figure 2. It may be assumed that these experimental values are too low and need further experimental study. The deviations of the experimental and corrected $\Delta_f H°$s do not depend on the size of the species, also expressed by the change in the correlation coefficient of deviations and carbon number from −0.8774 to 0.2059, due to the applied correction. These facts suggest that the CBS-4 method with the proposed empirical correction may predict reasonable $\Delta_f H°$s for other species, too.

(**c**) **Group Additivity Rule and Group Values**. The GVs derived via MVLR are presented in Table 4. The GVs in columns A−F were obtained through different approaches, to be discussed below. A common feature of columns A−F is that the estimation was successful: the GVs obtained were capable

TABLE 3: Main Geometrical Properties of Gauche Interactions

Species	Gauche int.	$R_{C1\text{-}C4}$	$r_{C6\text{-}C4}$	$r_{H61\text{-}H41}$
	AG	3.901	3.159	2.226
	RG1 (2)	3.306	3.305	2.659
	RG1	3.903	3.356	2.634
	RG2	3.898	3.087	2.665
	AG	3.926	3.164	2.225
	AG	3.131		2.290
	RG2		3.070	2.897
	AG	3.138		2.281
	RG1		3.333	2.595
	RG1 (2)	3.252	3.252	2.643
	RG2 (2)	3.068	3.069	3.038
	AG		3.159	2.274
	RG2	3.102		2.828
	AG(?)		3.161	2.239
	RG1 (2)	3.363		2.690

of describing all the 34 ab initio data within an average error of $\leq$0.35 kcal mol^{-1}, and with a highest deviation of $<$0.7 kcal mol^{-1}. The worst agreement was found for the neopentyl radical. (A comparison of the $\Delta_fH°$s obtained via ab initio calculations and from GVs is depicted in Figure 3, while that of the GV-based and experimental data is to be seen in Figure 4.) The agreement of these data indicates that the GA method is applicable for the calculation of standard enthalpies of formation, and suggests that it also yields reliable data for radicals not included in this study. The relevance of simple linear GA rules for (alkyl) radicals is confirmed by the multiple correlation coefficient $\rho_m > 0.999$.

The GVs in columns A–F can be commented on as follows:

(A) In column A, GVs from uncorrected ab initio data are shown. Although the fitting is good, several GVs seem to differ from those recently suggested by Cohen and Benson.[8] This may stem from the systematic error caused in the ab initio data by the input errors and by the difference in the numbers of spin-paired electrons.

(B) After application of the suggested correction, the data in column B were obtained. The overall increases in $\Delta_fH°$s in consequence of the correction applied are reflected in all the GVs calculated. The values of the groups in radical-adjacent positions are significantly lower than those in the molecules. The decrease in this difference with the order of the affected C atoms and the radical groups themselves does not seem to meet expectations. The correlation of the GVs is probably responsible for the distortion of these parameters. Primarily the GV data for radical groups and radical-adjacent alkyl groups display a strong correlation. For both $C-(H)_3(\dot{C})$ vs $\dot{C}-(H)_2(C)$ and $C-(H)_3(\dot{C})$ vs $\dot{C}-(H)(C)_2$, the strength of the correlation is > 0.97, which is readily explainable by the frequent co-occurrence of these groups. The structural similarity of $C-(H)_3(C)$ and $C-(H)_3(\dot{C})$ enables their identity to be introduced, which is a simplification widely used in GA calculations. The GVs in Column C were obtained in this way.

(C) The estimation of the 12 GVs selected on the assumption $C-(H)_3(\dot{C}) = C-(H)_3(C)$ resulted in RG2 = 0.01 kcal mol^{-1}, a negligible contribution. RG2 was set to zero and the fit was repeated. The parameters did not change beyond 0.02 kcal mol^{-1}. The GVs obtained can be seen in column C.

In this task, both the ab initio data to be fitted and the fitting itself meet the criteria that must be satisfied by reliable thermochemical data. The standard deviation, the variances and the correlation of the parameters have only moderate values. The latter are not given in Table 4; their maximum is 0.83 for $C-(H)_3(C)$ and $C-(C)_4$. The remainder are below 0.8, with a typical value of <0.5. The derivatives $\partial g_i/\partial h_j$ $_{\text{input}}$ were also calculated to establish whether the input of the $\Delta_fH°$s of the individual compounds had an extreme influence on the GVs. Ethane (or, if this was not present in the input, another simple alkane molecule, see (D)) had a primary effect on 4 of the 12 GVs. This influence is positive, but fortunately not predominant.

Comparison of the “best” estimated parameters in column C with the “low” literature values[8] reveals an average deviation of 0.5 kcal mol^{-1}; the three largest deviations are for $\dot{C}-(C)_3$, $C-(H)(\dot{C})(C)_2$, and $\dot{C}-(H)_2(C)$: 1.83, 1.03, and 0.85, respectively. This discrepancy may be caused by the accepted $\Delta_fH°$s of the alkyl radicals.

The GVs for alkyl groups can be obtained with different accuracies in the range ±0.1–0.7 kcal mol^{-1} in the 12-parameter task. While those of $C-(H)_3(C)$ and $C-(H)_2(C)_2$ are essentially insensitive to the various changes in the models, the conditions are less favorable for *tertiary* and *quaternary* groups. This is due to the differences in their frequencies in the species.

The RG2 interaction was estimated to be nearly zero, permitting its complete neglect. In contrast, it is advisable to consider RG1 and AG, although the latter seems to be lower than the accepted 0.8 kcal mol^{-1}.[5] The slightly negative value of RG1 confirms the simple image that gauche interactions on an alkyl radical backbone represent a more relaxed structure.

The present calculations predict in all cases that alkyl groups are more stable than the relevant radical-adjacent alkyl groups, as can be expected. Assumption of the identity of each alkyl and radical-adjacent alkyl group is an admissible approximation that results in a maximum error of 0.96 kcal mol^{-1} for *tertiary* radicals.

Bond dissociation enthalpies (BDHs) can be calculated to check the GVs for radical groups, by applying

$$\mathrm{BDH} = \Delta_fH°(\dot{R}) + \Delta_fH°(H) - \Delta_fH°(RH)$$

The good agreement convincingly demonstrates that the errors in the individual parameters do not accumulate (see Table 5), although the BDHs seem to be slightly overestimated in comparison with the experimental ones.[36]

(D) The first members of a homologous series exhibit different behavior from that of the further members in many cases. In order to check the influence of this, the calculations were repeated with the omission of the smallest radicals and alkanes, $C_{n<4}$ (column D). The resulting changes in GVs were neither significant nor unidirectional. The standard deviation for the regression remained the same as in column C. The $\Delta_fH°$s of $C_{n<4}$ radicals and alkanes calculated from the GVs in columns C and D were compared with one another and with the corrected CBS-4 data. Negligible increases in the *primary* and *secondary* radical GVs and AG were observed, while most of the other groups displayed a similar degree of decrease. The deviations for the small species not present in the input slightly increased:

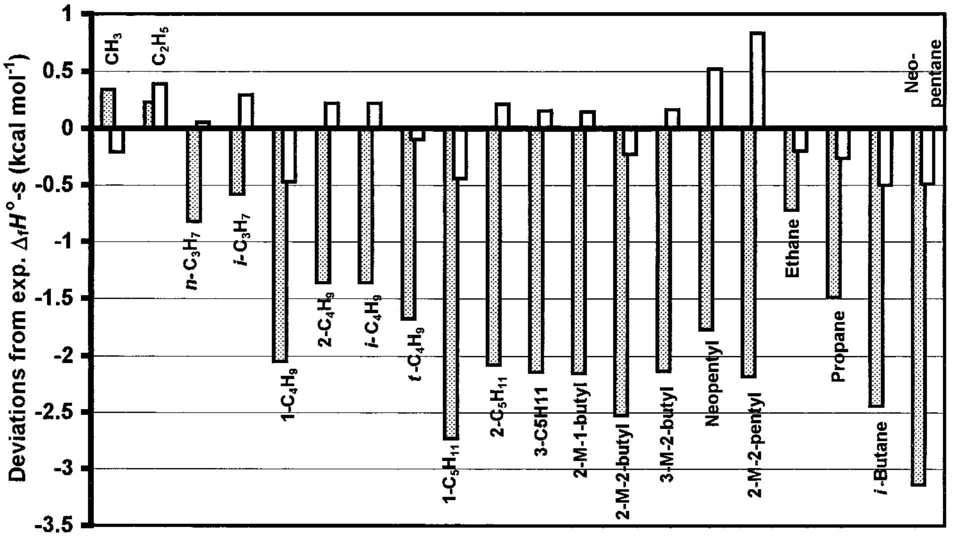


Figure 2. Deviations of CBS-4, empirically corrected CBS-4 and experimental $\Delta_f H°$s: (■) CBS-4, experimental; (□) corrected CBS-4, experimental.

TABLE 4: Estimated GVs (kcal mol^{-1}) with Variances for 95% Confidence[a]

group	A	B	C	D	E	F	G lit.
$C-(H)_3(C)$	−10.29 (±0.12)	−10.03 (±0.12)	−10.11 (±0.15)	−10.09 (±0.37)	−10.12	−10.0*	−10.15
$C-(H)_2(C)_2$	−6.04 (±0.09)	−5.33 (±0.09)	−5.32 (±0.13)	−5.41 (±0.20)	−5.05	−5.0*	−4.99
$C-(H)(C)_3$	−3.70 (±0.33)	−2.54 (±0.33)	−2.34 (±0.39)	−2.53 (±0.92)	−2.23	−2.4*	−2.03
$C-(C)_4$	−2.12 (±0.53)	−0.51 (±0.53)	0.18 (±0.59)	−0.08 (±1.41)	−0.18	−0.1*	−0.12
$C-(H)_3(\dot{C})$	−11.60 (±0.50)	−11.34 (±0.50)	≡$C-(H)_3(C)$	≡$C-(H)_3(C)$	−10.0 (assumed)	−**10.0** (assumed)	−10.0 (assumed)
$C-(H)_2(\dot{C})(C)$	−6.85 (±0.51)	−6.14 (±0.50)	−4.93 (±0.21)	−5.06 (±0.32)	−4.65	−**4.93** (±0.31)	−5.0
$C-(H)(\dot{C})(C)_2$	−3.87 (±0.62)	−2.70 (±0.61)	−1.38 (±0.41)	−1.66 (±0.94)	−1.08	−**1.55** (±0.45)	−2.4
$\dot{C}-(H)_2(C)$	40.42 (±0.48)	40.33 (±0.47)	39.24 (±0.30)	39.57 (±0.46)	38.91	**38.68** (±0.37)	38.4
$\dot{C}-(H)(C)_2$	43.25 (±0.98)	43.61 (±0.97)	41.24 (±0.40)	41.43 (±0.59)	41.22	**40.84** (±0.42)	41.0
$\dot{C}-(C)_3$	45.67 (±1.52)	46.47 (±1.51)	42.82 (±0.66)	42.76 (±1.23)	42.63	**42.34** (±0.62)	41.0
AG	0.54 (±0.21)	0.54 (±0.21)	0.48 (±0.29)	0.52 (±0.36)	0.20	0.8*	0.8
RG1	−0.21 (±0.18)	−0.22 (±0.18)	−0.17 (±0.26)	−0.09 (±0.30)		−**0.20** (±0.36)	
RG2	0.15 (±0.16)	0.15 (±0.16)					
s^*	0.17	0.17	0.25	0.26	0.62	0.35	
ρ_m	0.9999	0.9999	0.9999	0.9999	0.9978	0.9937	

[a] A: Uncorrected CBS-4 $\Delta_f H°$s of 30 alkyl radicals and 4 alkane molecules; 13 GVs estimated. B: Corrected CBS-4 $\Delta_f H°$s of 30 alkyl radicals and 4 alkane molecules; 13 GVs estimated. C: Corrected CBS-4 $\Delta_f H°$s of 30 alkyl radicals and 4 alkane molecules; 12 GVs estimated. D: Corrected CBS-4 $\Delta_f H°$s of 27 alkyl radicals and 2 alkane molecules with $C_{n\geq 4}$; 12 GVs estimated. E: Corrected CBS-4 $\Delta_f H°$s of 30 alkyl radicals; 10 GVs derived in a step-by-step procedure (essentially the same set as treated by Cohen and Benson[8]). F: Corrected CBS-4 $\Delta_f H°$s of 29 alkyl radicals; 6 GVs estimated; those marked * fixed as given in Cohen's review.[9] G: Literature: empirical GVs of Cohen and Benson.[8]

the average of the changes was 0.18 kcal mol^{-1}. Thus, it is essentially indifferent whether the GVs are derived by considering all 30 radicals or by omitting the smallest ones. These facts indicate that the Benson procedure for the estimation of GVs from the first members yields essentially correct results. Since the proposed GVs show no expressed dependence on the number of C atoms, they are expected to describe the $\Delta_f H°$s of $C_{n>6}$ isomers, too. This expectation appears to be confirmed by the fact that the deviations of $\Delta_f H°$s based on GVs vs the occurrence of groups yielded a correlation that was moderate ($\max(r_{ij}) = 0.638$ for $C-(C)_4$ in column C) or negligible ($r_{ij} < 0.3$ for any other groups in models B−F).

(E) The previous simplifications were essentially unavoidable, and the nature of the task forced us to introduce them. The following ones were aimed at testing the applicability of the GVs derived from the smallest species in a step-by-step procedure. In column E, four alkyl, three radical-adjacent alkyl, three radical, and one gauche GVs were distinguished. RG1 and RG2 interactions were neglected, and the $C-(H)_3(\dot{C})$ was assumed to be −10.00 kcal mol^{-1}; in other words, an attempt was essentially made to reproduce the GVs of the model of Cohen and Benson[8] on an ab initio database. (As compared with the latter, the equivalence of $C-(\dot{C})(C)_3$ and $C-(C)_4$ is applied in our treatment.) The GVs obtained in this way differ from those in column C by less than 0.37 kcal mol^{-1}. However, these GVs reproduce the corrected ab initio $\Delta_f H°$s less satisfactorily: an increase can be found both in the average deviation (from 0.26 to 0.62 kcal mol^{-1}) and in the correlation of carbon number and deviation (from -1.5×10^{-6}, practically zero to −0.357). This can be also an argument to prefer MVLR to step-by-step calculations whenever possible.

(F) A comprehensive revision of numerous GVs relating to molecules was made by Cohen,[9] making use of wide, pure experimental based data. It can be expected that the GVs relating to alkane groups and alkane gauche interactions derived in the present paper will agree with the set published in ref 9. The GVs in column F of Table 4 meet these expectations: if the GVs transferable from alkanes are fixed, the change in the estimated parameters is 0.32 kcal mol^{-1} on average. The parameter variances show a more expressed increase, which is readily explainable by the fact that Cohen's GVs were built in as error-free data, and consequently, the errors of the estimation are concentrated in the parameters left. In conclusion, this comparison corroborates the validity of the present GVs, partly directly, and partly indirectly. The GV-based $\Delta_f H°$s are in satisfactory agreement with the experimental data.[33] The average deviation is somewhat lower than that obtained via the use of Cohen and Benson's GVs,[8] 0.33 and 0.77 kcal mol^{-1}, respec-

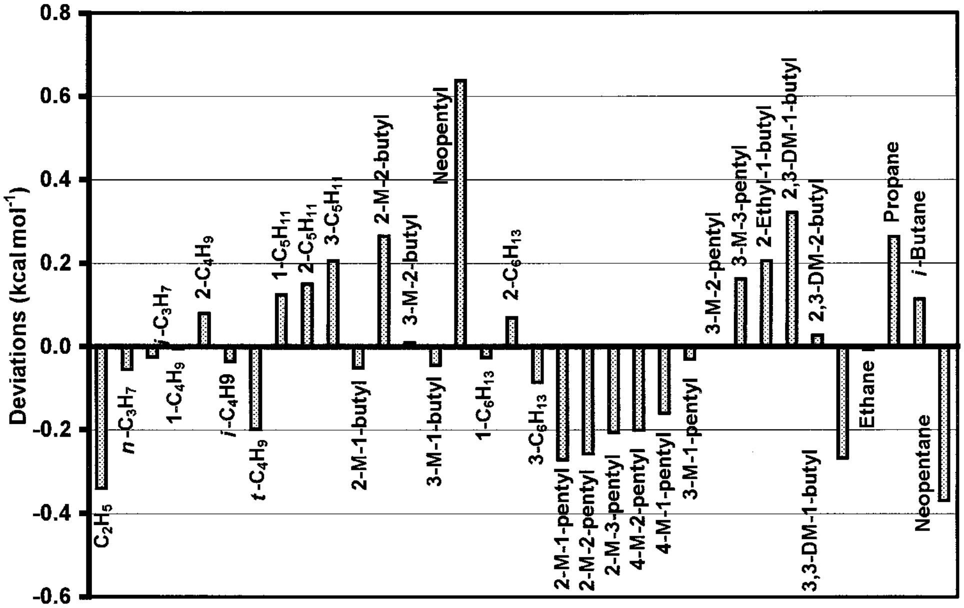


Figure 3. Deviations of CBS-4 corrected and GV-based (Table 4, column C) $\Delta_f H°$s.

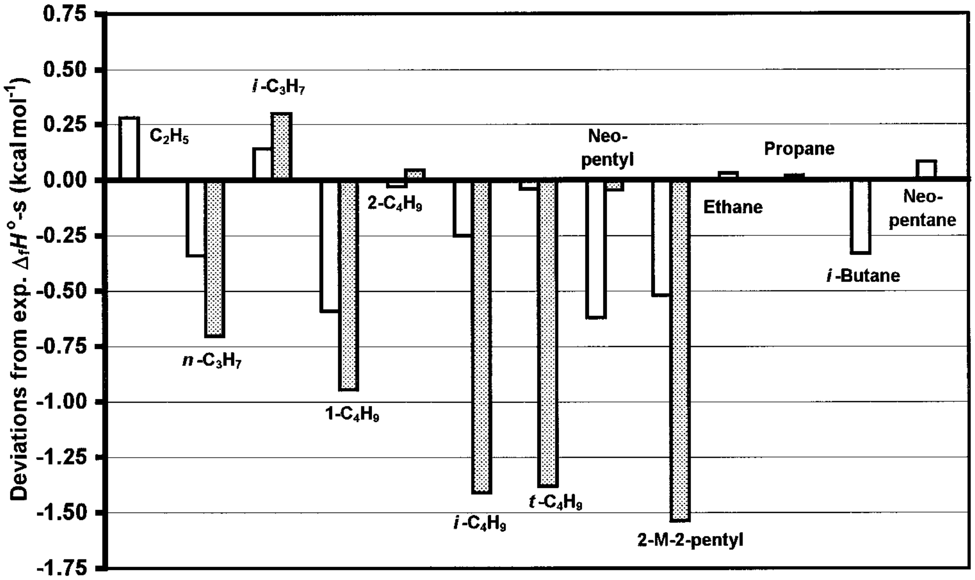


Figure 4. Deviations of GV-based and experimental $\Delta_f H°$s: (■) on Cohen and Benson's GVs;[8] (□) on GVs derived in this paper (Table 4, column F).

TABLE 5: Calculated and Experimental Bond Dissociation Enthalpies (BDHs, kcal mol^{-1})

	av BDH calcd	exptl BDH[36]
RCH_2-H	101.83 (±0.40)	101.10
R_1R_2CH-H	99.09 (±0.33)	98.45
$R_1R_2R_3C-H$	97.21 (±0.33)	96.5

tively, although in the latter the deviation is by definition zero in several cases, due to the applied step-by-step evaluation of GVs (see Figure 4).

4. Conclusions

The present ab initio method of computing the thermochemical properties of a homologous series is successful and the computational time needed is acceptable. The main advantage of this procedure is that the systematic errors can be decreased by empirical corrections, and the accuracy of the data can be estimated. The suggested empirical correction on CBS-4 enthalpies reproduces recent experimental data reasonably well. Further studies to reveal the nature of this correction are needed.

The GA rules were checked by statistical analysis of the simultaneous parameter estimation. It was found that the GA rules are generally useful to check the reality of the thermochemical properties calculated for a homologous series by ab initio methods. New values of group contributions to $\Delta_f H°$ are suggested to evaluate more accurate enthalpies of formation of branched and large alkyl radicals. Among the different gauche contributions defined in the present work, besides the "classical" AG, RG1 has a nonnegligible contribution. Since the suggested GVs reflect the thermochemical features of several members of homologous series, they may have more expressed chemical reality than those from the archetypes of the individual species. In other words, their application allows the GA principle approach to move from soft modeling to hard modeling.

The alkyl radical GVs derived in earlier studies were obtained by step-by-step evaluation from data on the first members of the homologous series of alkyl radicals,[5] or the alkane GVs used in the estimation were less reliable than those recently obtained by Cohen,[9] who neglected all the nonexperimental data. Since the data derived in the present work are also free from such shortcomings, we suggest that the GVs listed in column F are the best currently available.

Acknowledgment. B.V. is grateful for the award of a Magyary Zoltán Postdoctoral Fellowship (AMFK 535/2), for the Hungarian Scientific Research Fund (OTKA F030436) and for PFP 4164 funds.

Supporting Information Available: Listed energies from ab initio calculations. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Pilling, M. J.; Proudler, V. K. *Int. Rev. Phys. Chem.* **1990**, *9*, 329.

(2) Baulch, D. L.; Cobos, C. J.; Cox, R. A.; Esser, C.; Frank, P.; Just, Th.; Kerr, J. A.; Pilling, M. J.; Troe, J.; Walker, R. W.; Warnatz, J. *J. Phys. Chem. Ref. Data* **1992**, *21*, 411.

(3) Benson, S. W.; Buss, J. H. *J. Chem. Phys.* **1958**, *29*, 546.

(4) O'Neal, H.; E.; Benson, S. W. *Int. J. Chem. Kinet.* **1969**, *1*, 221.

(5) Benson, S. W. *Thermochemical Kinetics,* 2nd ed.; John Wiley & Sons: New York, 1976.

(6) Afeefy, H. Y.; Liebman, J. F.; Stein, S. E. *Neutral Thermochemical Data* in *NIST Chemistry WebBook;* NIST Standard Reference Database Number 69; Mallard, W. G., Linstrom, P. J., Eds.; National Institute of Standards and Technology: Gaithersburg MD 20899, Nov 1998 (http://webbook.nist.gov).

(7) Cohen, N.; Benson, S. The thermochemistry of alkanes and cycloalkanes. In *The chemistry of alkanes and cycloalkanes*; Patai, S., Rappoport, Z., Eds.; John Wiley & Sons: New York, 1992; Chapter 6, p 259.

(8) Cohen, N.; Benson, S. W. *Chem. Rev.* **1993**, *93*, 2419.

(9) Cohen, N. *J. Phys. Chem. Ref. Data,* **1996**, *25*, 1411.

(10) Lay, T. H.; Bozzelli, J. W. *J. Phys. Chem. A* **1997**, *101*, 9505.

(11) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **1997**, *106*, 1063.

(12) Viskolcz, B.; Lendvay, G.; Körtvélyesi, T.; Seres, L. *J. Am. Chem. Soc.* **1996,** *118*, 3006.

(13) Lendvay, G.; Viskolcz, B. *J. Phys. Chem. A.* **1998**, *102*, 10777.

(14) Yamauchi, N.; Miyoshi, A.; Kosaka, K.; Koshi, M.; Matsui, H. *J. Phys. Chem. A* **1999**, *103,* 2723.

(15) Yamada, T.; Lay, T. H.; Bozzelli, J. W. *J. Phys. Chem.* **1998**, *102*, 7286.

(16) Nyden, M. R.; Petersson, G. A. *J. Chem. Phys.* **1981**, *75*, 1843.

(17) Petersson, G. A.; Al-Laham, M. A. *J. Chem. Phys.* **1991**, *94*, 6081.

(18) Petersson, G. A.; Tensfeldt, T.; Montgomery, J. A. *J. Chem. Phys.* **1991**, *94*, 6091.

(19) Montgomery, J. A.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1994**, *101*, 5900.

(20) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Pittsburgh, PA, 1995.

(21) Chase, Jr., M. W.; Davis, C. A.; Downey, Jr., J. R.; Frurip, D. J.; McDonald, R. A.; Syverud, A. N. *J. Phys. Chem. Ref. Data* **1985**, *14*, Suppl. No. 1.

(22) *Handbook of Chemistry and Physics*, 77th ed.; Lide, D. R., Ed.; CRC Press: Boca Raton, FL, 1997.

(23) Schmitz, L. R.; Chen, Y. R. *J. Comput. Chem.* **1994**, *15*, No. 12. 1437.

(24) Beck, J. V.; Arnold, K. J. *Parameter Estimation in Engineering and Science*; John Wiley and Sons: New York, 1977; pp 481−487.

(25) Pacansky, J.; Waltman, R. J.; Barnes, L. A. *J. Phys. Chem.* **1993**, *97*, 10694.

(26) Viskolcz, B.; Lendvay, G.; Seres, L. *J. Phys. Chem. A.* **1997**, *101*, 7119.

(27) Tsang, W. *J. Am. Chem. Soc.* **1985**, *107*, 2872, and references therein.

(28) Dobis, O.; Benson, S. W. *J. Am. Chem. Soc.* **1995**, *117*, 8171, and references therein.

(29) Berkovitz, J.; Ellison, G. B.; Gutman, D. *J. Phys. Chem.* **1994**, *98*, 2744, and references therein.

(30) Seekins, P. W.; Pilling, M. J.; Niiranen, J. T.; Gutman, D.; Krasnoperov, L. N. *J. Phys. Chem.* **1992**, *96*, 9847.

(31) Seetula, J. A.; Gutman, D. *J. Phys. Chem.* **1990**, *94*, 7529.

(32) Seetula, J. A.; Russell, J. J.; Gutman, D. *J. Am. Chem. Soc.* **1990**, *112*, 1347.

(33) Seetula, J. A.; Slagle, I. R. *J. Chem. Soc., Faraday Trans.* **1997**, *93*, 1709.

(34) Seres, L.; Dinse, M.; Scherzer, K.; Görgényi, M., to be published.

(35) Dobis, O.; Benson, S. W. *J. Phys. Chem. A.* **1997**, *101*, 6030.

(36) Seakins, P. W.; Pilling, M. J.; Niiranen, J. T.; Gutman, D.; Krasnoperov, L. N. *J. Phys. Chem.* **1992**, *96*, 9847.