The Origin of Systematic Error in the Standard Enthalpies of Formation of Hydrocarbons Computed via Atomization Schemes

Gyula Tasi,*[a] Róbert Izsák,[a] Gergely Matisz,[a] Attila G. Császár,[b] Mihály Kállay,[c] Branko Ruscic,[d] and John F. Stanton[e]

The combustion of hydrocarbon fuels involves many chemical species and reactions; for example, saturated hydrocarbons with up to about twelve and eighteen carbon atoms constitute the bulk of gasoline and diesel fuel, respectively. Knowledge of the (temperature-dependent) enthalpies of formation of alkanes is therefore of importance for developing reliable models of fuel combustion. Motivated partly by their industrial importance, several recent theoretical studies have dealt with the structural, electronic, and thermodynamic properties of small- to mid-sized alkanes.[1–8] Extremely accurate thermochemical calculations are now possible for small species with up to about five atoms (for recent examples, see refs. [9–16]). However, the severe associated computational requirements coupled with a steep scaling of cost with system size are such that compromises must be made for larger systems. It is necessary to resort to less demanding electronic structure techniques and to employ methods with relatively small basis sets and additivity approximations; often, somewhat ad hoc empirical corrections must also be used. One of the first successful families of approximate ab initio methods for first-principles thermochemistry were the BAC-MPX ($X = 2$ and 4) techniques.[17] Even more successful have been the Gaussian-n (Gn) approaches of Pople and co-workers, G1,[18] G2,[19] and G3,[20] and their modified versions, including G2(MP2,SVP),[21,22] the CBS methods,[23] as well as the Weizmann-n (Wn) variant W1.[24] Common to all of these model chemistries is a relatively wide range of applic-

[a] Prof. G. Tasi, R. Izsák, G. Matisz
Department of Applied and Environmental Chemistry
University of Szeged, Rerrich B. tér 1, 6720 Szeged (Hungary)
Fax: (+36)62-544-619
E-mail: tasi@chem.u-szeged.hu

[b] Prof. A. G. Császár
Laboratory of Molecular Spectroscopy
Institute of Chemistry, Eötvös University
1518 Budapest 112, P.O. Box 32 (Hungary)

[c] Prof. M. Kállay
Department of Physical Chemistry
Budapest University of Technology and Economics
H-1521 Budapest, P.O. Box 91 (Hungary)

[d] Prof. B. Ruscic
Chemistry Division, Argonne National Laboratory
Argonne, Illinois 60439 (USA)

[e] Prof. J. F. Stanton
Institute for Theoretical Chemistry
Department of Chemistry and Biochemistry
The University of Texas at Austin, Austin, Texas 78712 (USA)

ability which derives from judicious approximations made in the electronic structure treatments. All of these methods should be distinguished from the more accurate model chemistries known as W2,[24] W3,[10] HEAT,[12] and the focal-point approach,[25,26] which can only be applied to small molecules, albeit with considerably higher accuracy.

The G2(MP2,SVP) model[21,22] is one of the simplest versions of the Gn series that can be applied to larger molecules. In this model enthalpies of formation of molecules at 0 K, $\Delta_f H_0^o$, are calculated from theoretical atomization energies and accepted values of enthalpies of formation for gaseous atoms. To adjust these values to 298.15 K, that is, to obtain $\Delta_f H_{298}^o$, as opposed to the more easily calculated $\Delta_f H_0^o$, thermal corrections to the enthalpies of molecules of interest are added via the standard formulas of statistical thermodynamics, using the rigid-rotor, harmonic oscillator (RRHO) approximation.[9,27] According to ref. [28], the G2(MP2,SVP) model supplies better standard enthalpies of formation for small noncyclic hydrocarbons than the seemingly more rigorous G2 model from which it derives, perhaps due to its excellent parameterization. Furthermore, literature studies show that accuracy of the G2(MP2,SVP) model chemistry improves when a) one corrects the G2(MP2,SVP) energy of the carbon atom for the first-order spin–orbit (SPO) interaction (thereby lowering atomization energies, which has the effect of raising the molecular enthalpies of formation);[28] and b) the low-frequency internal rotations are treated as free rotors.[27] Nevertheless, since the relevant higher-level corrections (HLC) of the G2(MP2) model chemistries have been fitted without consideration of SPO and using the CODATA[29] value of $\Delta_f H_0^o(C_{gas})$, one must be careful when assessing such improvements.

As shown here, the remarkable accuracy of G2(MP2,SVP) for standard enthalpies of formation deteriorates for larger alkanes. Moreover, a decidedly systematic feature is noted: apart from methane, standard enthalpies of formation are *underestimated* for all species studied thus far.

This Communication reveals that the systematic degradation of G2(MP2,SVP) enthalpies of formation with molecular size derives, to a large extent, from the value used for the enthalpy of formation of the carbon atom. This is not completely surprising, since the CODATA[29] value of $\Delta_f H_0^o(C_{gas})$ has a relatively large uncertainty, ± 0.45 kJ mol^{-1}, and associated problems have been noted in previous publications.[12–14,30]

Table 1 lists experimental as well as G2(MP2,SVP) standard enthalpies of formation for a number of C_1—C_{13} alkanes. In the calculations, only the principal isotopic species was considered for each molecule, and the RRHO approximation was applied to obtain enthalpy corrections. For molecules with conformational flexibility, the number of unique conformers are listed in parentheses in Table 1; an ideal mixture of freely interconverting conformers was assumed and Boltzmann statistics were used to determine the population of each conformer. To the best of our knowledge, there are no other alkanes with a small number of conformers for which dependable experimental enthalpies of formation are available. Note here that the experimental enthalpies of formation of the alkanes C_8 and C_9 have the largest deviation from the appropriate values of the com-

**Table 1.** Experimental and G2(MP2,SVP) standard enthalpies of formation of alkanes [kJ mol^{-1}].[a]

C_i	Molecules[b]	Exptl.[c]	Model A	Model B	Model C	Model D	Model E	Model F
C_1	methane (1)	-74.549 ± 0.060	-74.50	-74.04	-74.16	-73.70	-73.68	-73.45
C_2	ethane (1)	-83.800 ± 0.300	-83.89	-82.97	-83.21	-82.29	-82.25	-81.79
C_3	propane (1)	-104.700 ± 0.500	-105.42	-104.04	-104.40	-103.02	-102.96	-102.46
C_4	butane (2)	-125.600 ± 0.670	-126.53	-124.69	-125.17	-123.33	-123.24	-123.21
C_4	isobutane (1)	-135.007 ± 0.410	-136.28	-134.44	-134.92	-133.08	-133.00	-132.37
C_5	pentane (4)	-146.400 ± 0.670	-147.47	-145.17	-145.76	-143.46	-143.36	-143.97
C_5	isopentane (2)	-154.500 ± 0.840	-156.05	-153.75	-154.34	-152.04	-151.94	-151.96
C_5	neopentane (1)	-168.500 ± 1.000	-172.37	-170.07	-170.66	-168.36	-168.26	-167.25
C_6	hexane (12)	-167.200 ± 0.790	-168.88	-166.12	-166.83	-164.07	-163.94	-165.23
C_6	2,2-dimethylbutane (1)	-186.100 ± 1.000	-190.07	-187.31	-188.03	-185.26	-185.14	-184.84
C_6	2,3-dimethylbutane (2)	-178.300 ± 1.000	-181.73	-178.97	-179.68	-176.92	-176.80	-176.97
C_7	2,2,3-trimethylbutane (1)	-204.800 ± 1.100	-212.22	-209.00	-209.83	-206.61	-206.46	-206.54
C_8	2,2,3,3-tetramethylbutane (1)	-225.900 ± 1.900	-238.35	-234.67	-235.62	-231.94	-231.77	-230.83
C_9	di-*tert*-butylmethane (1)	-241.569 ± 0.832	-255.58	-251.44	-252.51	-248.37	-248.18	-248.20
C_{13}	tri-*tert*-butylmethane (1)	-235.738 ± 2.460	-244.25	-238.27	-239.81	-233.83	-233.56	-234.56
c			0.99928	0.99913	0.99918	0.99896	0.99895	0.99923
r			0.97099	0.98695	0.98278	0.99910	0.99986	0.99998
rmsd			5.95	3.73	4.22	2.96	2.96	2.79
max			14.01	9.87	10.94	6.80	6.61	6.63

[a] Statistical data included in the last four rows of the table: c, linear correlation coefficient; r, linear regression coefficient referring to computed versus experimental enthalpies of formation; max, absolute value of the maximum deviation between experimental and computed enthalpies of formation (kJ mol^{-1}); and rmsd, root-mean-square deviation between experimental and computed enthalpies of formation (kJ mol^{-1}). [b] Numbers in parentheses refer to the number of non-isomorphic conformers considered during the computation of the enthalpies of formation. [c] The experimental value listed for methane comes from the ATcT database of Ruscic et al.;[32] the remaining experimental values are from the NIST WebBook.[34]

putational models investigated. Leaving out these data from the statistical analyses would not change the conclusions of this Communication, the proposed empirical correction of $\Delta_f H_0^o(C_{gas})$ would be lower by 0.2 kJ mol^{-1}, well within the error bar given below.

The model A data in Table 1 were obtained via the original G2(MP2,SVP) method[21, 22] using the canonical CODATA[29] experimental value of $\Delta_f H_0^o(C_{gas})$, 711.194 kJ mol^{-1}. Model B is the same as model A except that an improved estimate obtained from high-level electronic structure calculations,[31] $\Delta_f H_0^o(C_{gas}) = $ 711.65 kJ mol^{-1}, is used instead of the CODATA value. In Models C and D the experimental SPO correction of -29.6 cm^{-1} to the atomic G2(MP2,SVP) energy of carbon is added (C$=$A$+$SPO and D$=$B$+$SPO), as suggested in ref. [28]. Model E data were obtained via relaxing the value of $\Delta_f H_0^o(C_{gas})$ within model C, that is, minimizing the root-mean-square deviation (rmsd) function with respect to the value of $\Delta_f H_0^o(C_{gas})$ (see also Figure 1). Finally, the suggestion of Nicolaides et al.[27] was considered: Model F is the same as model E except that the low-frequency internal rotations, that is, those with frequencies below 260 cm^{-1}, were treated as free rotors. The last rows of Table 1 show the results of linear regression analyses between the experimental and the computed enthalpies of formation and the rmsd.

The data in Table 1 reveal that the linear correlation between the experimental and the computed enthalpies of formation is excellent, independent of the choice of the model. However, Model A, that is, the original G2(MP2,SVP) method, systematically underestimates the standard enthalpies of formation of alkanes, except for methane. The more carbon atoms the molecule contains, the larger the discrepancy generally becomes
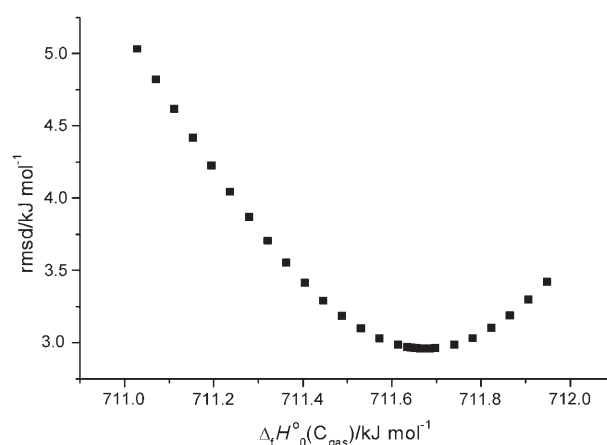
Figure 1. Dependence of the rmsd error (Table 1) on $\Delta_f H_0^o(C_{gas})$ within Model C.

between the computed and the experimental values. The linear regression coefficient for Model A is $r = 0.971$.

Model B data reveal a significant improvement in the computed G2(MP2,SVP) enthalpies of formation. The 6.0 kJ mol^{-1} rmsd of Model A decreases to 3.7 kJ mol^{-1}, and r increases significantly to 0.987.

The inclusion of the SPO stabilization for the carbon atom (from the averaged state that is addressed in non-relativistic calculations) improves both the rmsd and r values of Models A and B. The linear regression coefficient of Model D is an impressive 0.999. Furthermore, the rmsd between Models A and C, 1.7 kJ mol^{-1}, decreases substantially, to 0.7 kJ mol^{-1}, when Model B is improved to Model D. Thus, it is considerably more

important to employ a seemingly improved enthalpy of formation of C_{gas} in the computation than the inclusion of the SPO correction. This is not unexpected due to the fact that the latter was effectively absorbed into the parameterization of the method.

Figure 1 reflects an important result of this Communication, as it shows the rmsd as a function of $\Delta_f H_0^o(C_{gas})$, that is, relaxing the value of the enthalpy of formation of carbon within Model C. The unusually steep dependence of the rmsd on $\Delta_f H_0^o(C_{gas})$ results in an estimate of 711.68 kJ mol^{-1} for $\Delta_f H_0^o(C_{gas})$. The Model E data in Table 1 were obtained with this estimate.

An analysis similar to that of Model E was performed for Model F. This resulted in another, slightly higher estimate for $\Delta_f H_0^o(C_{gas})$, 711.91 kJ mol^{-1}. The Model F data in Table 1 were obtained with this estimate. Relative to Model E, the rmsd decreases only slightly and the change in the regression coefficient is marginal.

Significance tests for Models A through F suggest that Model A is significantly different from all the other models, Models B and C are somewhat similar, while differences in Models D, E, and F are statistically insignificant. Results from Models D, E, and F, based on the G2(MP2,SVP) technique without refitting the HLC, suggest a total correction of the reference dissociation asymptote of 0.80–1.05 kJ mol^{-1} per carbon atom. Out of this, about 0.35 kJ mol^{-1} can be ascribed to SPO, the rest of it is suggested to be due to the use of the (imprecise) CODATA value of $\Delta_f H_0^o(C_{gas})$. Averaging Model-D, E, and F $\Delta_f H_0^o(C_{gas})$ values gives the best empirical estimate of this study, $\Delta_f H_0^o(C_{gas}) = (711.75 \pm 0.45)$ kJ mol^{-1}.

In summary, the original G2(MP2,SVP) model can be improved slightly for alkanes by including the spin–orbit effect in carbon, as already noted in ref. [28]. It is also clear that treating the low-frequency internal rotations as free rotors—as suggested in ref. [27]—results in only a marginal (statistically insignificant) improvement in the predicted enthalpies of formation of larger alkanes. The most significant source of systematic error in enthalpies of formation for larger organic compounds based on atomization-energy schemes appears to be that associated with the enthalpy of formation of atomic carbon. The best empirical estimate of this study for the enthalpy of formation of C_{gas} [(711.75 $\pm$ 0.45) kJ mol^{-1}] suggests that the CODATA[29] value—(711.194 $\pm$ 0.45) kJ mol^{-1}—is too low by roughly 0.5 kJ mol^{-1}. This estimate for $\Delta_f H_0^o(C_{gas})$ is in good agreement with an ab initio value,[31] (711.65 $\pm$ 0.32) kJ mol^{-1}, and with an interim Active Thermochemical Tables (ATcT)[32] value of Ruscic[33] quoted in ref. [12], (711.79 $\pm$ 0.21) kJ mol^{-1}. The suggested increase in $\Delta_f H_0^o(C_{gas})$ will increase the enthalpies of formation of the radicals CH[13] and CH$_2$[14] computed from accurate ab initio total atomization energies. These new values, in turn, considerably improve the agreement between enthalpies of formation of C_2H_2 computed via different isodesmic reactions within the HEAT protocol (see footnote 101 of ref. [12]). This Communication has focused on the G2(MP2,SVP) model chemistry because it is sufficiently simple to apply, so that calculations on relatively large molecules are possible, thereby allowing the problems pointed out here to be quantified numer-

ically. The main conclusion of this paper supports a suggestion based on the ATcT approach,[30] namely, that establishing a new value for $\Delta_f H_0^o(C_{gas})$ is needed. This will affect *all* theoretical-model chemistries based upon atomization energies and *all* modeling studies and chemical-engineering calculations employing, directly or indirectly, $\Delta_f H^o(C_{gas})$.

[1] J. M. Goodman, *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 876.
[2] a) G. Tasi, F. Mizukami, *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 632; b) G. Tasi, F. Mizukami, I. Pálinkó, J. Csontos, W. Györffy, P. Nair, K. Maeda, M. Toba, S. Niwa, Y. Kiyozumi, I. Kiricsi, *J. Phys. Chem. A* **1998**, *102*, 7698; c) G. Tasi, F. Mizukami, *J. Math. Chem.* **1999**, *25*, 55; d) G. Tasi, F. Mizukami, J. Csontos, W. Györffy, I. Pálinkó, *J. Math. Chem.* **2000**, *27*, 191.
[3] D. F. DeTar, *J. Phys. Chem. A* **1998**, *102*, 5128.
[4] E. Koglin, R. J. Meier, *Chem. Phys. Lett.* **1999**, *312*, 284.
[5] P. C. Redfern, P. Zapol, L. A. Curtiss, K. Raghavachari, *J. Phys. Chem. A* **2000**, *104*, 5850.
[6] M. S. Deleuze, W. N. Pang, A. Salam, R. C. Shang, *J. Am. Chem. Soc.* **2001**, *123*, 4049.
[7] P. Vansteenkiste, V. Van Speybroeck, E. Pauwels, M. Waroquier, *Chem. Phys.* **2005**, *314*, 109.
[8] L. Pollack, T. L. Windus, W. A. De Jong, D. A. Dixon, *J. Phys. Chem. A* **2005**, *109*, 6934.
[9] B. Ruscic, J. E. Boggs, A. Burcat, A. G. Császár, J. Demaison, R. Janoschek, J. M. L. Martin, M. Morton, M. J. Rossi, J. F. Stanton, P. G. Szalay, P. R. Westmoreland, F. Zabel, T. Bérces, *J. Phys. Chem. Ref. Data* **2005**, *34*, 573.
[10] A. D. Boese, M. Oren, O. Atasoylu, J. M. L. Martin, M. Kállay, J. Gauss, *J. Chem. Phys.* **2004**, *120*, 4129.
[11] M. S. Schuurman, S. R. Muir, W. D. Allen, H. F. Schaefer III, *J. Chem. Phys.* **2004**, *120*, 11586.
[12] A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez, J. F. Stanton, *J. Chem. Phys.* **2004**, *121*, 11599.
[13] A. G. Császár, P. G. Szalay, M. L. Leininger, *Mol. Phys.* **2002**, *100*, 3879.
[14] A. G. Császár, M. L. Leininger, V. Szalay, *J. Chem. Phys.* **2003**, *118*, 10631.
[15] B. A. Flowers, P. G. Szalay, J. Vazquéz, J. F. Stanton, M. Kállay, J. Gauss, A. G. Császár, *J. Phys. Chem. A* **2004**, *108*, 3195.
[16] D. Feller, D. A. Dixon, J. S. Francisco, *J. Phys. Chem. A* **2003**, *107*, 1604.
[17] M. D. Allendorf, C. F. Melius, P. Ho, M. R. Zachariah, *J. Phys. Chem.* **1995**, *99*, 15285.
[18] J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari, L. A. Curtiss, *J. Chem. Phys.* **1989**, *90*, 5622.
[19] L. A. Curtiss, K. Raghavachari, G. W. Trucks, J. A. Pople, *J. Chem. Phys.* **1991**, *94*, 7221.

[20] L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Rassolov, J. A. Pople, *J. Chem. Phys.* **1998**, *109*, 7764.

[21] B. J. Smith, L. Radom, *J. Phys. Chem.* **1995**, *99*, 6468.

[22] L. A. Curtiss, P. C. Redfern, B. J. Smith, L. Radom, *J. Chem. Phys.* **1996**, *104*, 5148.

[23] a) J. W. Ochterski, G. A. Petersson, J. J. A. Montgomery, *J. Chem. Phys.* **1996**, *104*, 2598; b) J. A. Montgomery, Jr., M. J. Frisch, J. W. Ochterski, G. A. Petersson, *J. Chem. Phys.* **1999**, *110*, 2822.

[24] J. M. L. Martin, G. de Oliveira, *J. Chem. Phys.* **1999**, *111*, 1843.

[25] W. D. Allen, A. L. L. East, A. G. Császár, *Structures and Conformations of Non-Rigid Molecules* (Eds.: J. Laane, M. Dakkouri, B. van der Veken, H. Oberhammer), Kluwer, Dordrecht, **1993**, p. 393.

[26] A. G. Császár, W. D. Allen, H. F. Schaefer III, *J. Chem. Phys.* **1998**, *108*, 9751.

[27] A. Nicolaides, A. Rauk, M. N. Glukhovtsev, L. Radom, *J. Phys. Chem.* **1996**, *100*, 17 460.

[28] L. A. Curtiss, K. Raghavachari, P. C. Redfern, J. A. Pople, *J. Chem. Phys.* **1997**, *106*, 1063.

[29] J. D. Cox, D. D. Wagman, V. A. Medvedev, *CODATA Key Values for Thermodynamics*, Hemisphere, New York, **1989** (see also http://www.codata.org/codata/databases/key1.html).

[30] B. Ruscic, R. E. Pinzon, G. von Laszewski, D. Kodeboyina, A. Burcat, D. Leahy, D. Montoya, A. F. Wagner, *J. Phys. Conf. Ser.* **2005**, *16*, 561.

[31] J. F. Stanton, M. Kállay, A. G. Császár, B. Ruscic, unpublished results.

[32] B. Ruscic, R. E. Pinzon, M. L. Morton, G. von Laszevski, S. J. Bittner, S. G. Nijsure, K. A. Amin, M. Minkoff, A. F. Wagner, *J. Phys. Chem. A* **2004**, *108*, 9979.

[33] B. Ruscic, unpublished interim result from Active Thermochemical Tables version 1.25 using the Core (Argonne) Thermochemical Network version 1.048.

[34] H. Y. Afeefy, J. F. Liebman, S. E. Stein, *Neutral Thermochemical Data in NIST Chemistry WebBook*, NIST Standard Reference Database No. 69 (Eds.: P. J. Linstrom, W. G. Mallard), June 2005, National Institute of Standards and Technology, Gaithersburg, MD (http://webbook.nist.gov).