Journal of the Korean Chemical Society
1999, *Vol. 43, No. 6*
Printed in the Republic of Korea

향상된 결합 파라미터를 도입한 순이론적 탄화수소 생성엔탈피 결정

P. Duchowicz · E. A. Castro*
CEQUINOR, Departamento de Química, Facultad de Ciencias Exactas, Universidad Nacional de La Plata
(1999. 8. 30 접수)

Hydrocarbon Enthalpies of Formation from *ab initio* Calculations Improved Through Bond Parameters

P. Duchowicz and E. A. Castro*
CEQUINOR, Departamento de Química, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, C.C. 962, La Plata 1900, Argentina
(Received August 30, 1999)

요 약. 순이론적인 계산으로부터 얻은 전체 에너지로부터 탄화수소의 생성 엔탈피를 계산하는 새로운 방법을 제안하였다. 이 방법은 결합 파라미터를 도입하여 원자의 전체 갯수에 근거를 두는 과거의 방법론을 개선한 것이다. 현 방법론의 결과는 우수하며 생성 엔탈피 계산의 평균오차는 실험 오차보다 작은 것으로 나타났다. 제안하는 방법론을 확실히 하기 위한 몇 가지 가능한 확장법도 제시하였다.

ABSTRACT. A new predictive scheme to calculate hydrocarbon enthalpies of formation from *ab initio* total energy calculations is described. The method improves a previous computation procedure (based on the total number of atoms) through the inclusion of bond parameters. Present results are good enough and the average absolute errors in the computed values of enthalpies of formation are lower than the experimental uncertainties. Some possible extensions are pointed out in order to reach definitive conclusions about the proposed methodology.

INTRODUCTION

Accurate prediction of thermodynamic properties of molecules makes up the primary aim of the great majority of electronic structure calculations. There has been considerable advance in recent years towards the development of high-accuracy first-principles models for computational thermochemistry.[1,2] The critical evaluation of these theoretical models of electronic structure is a prerequisite to their proper application.[3] The standards for assessment of quantum mechanical methods against experimental data were established by the Pople group nearly 25 years ago.[4] At the same time, the impressively large and rapidly increasing set of *ab initio* data for a wide variety of molecules gives the chemist an unprecedent good chance to evaluate the energies of many molecular species for which no experimental values are available.

Among the current thermodynamic properties, the standard enthalpy of formation $\Delta H^{o}_{f}(g)$, which measures thermodynamic stability, is most usually employed in chemical research and thermochemical applications. The accurate estimation of the heats of formation for chemical compounds is very important in different fields of chemistry.[5-7] $\Delta H^{o}_{f}(g)$ is the standard reaction enthalpy for the formation of the compound in gas phase from its elements in their reference standard state. The reference state of an element is its most stable state, which is at 298K and 1 atm.[8]

It has already shown that high level *ab initio* methods like G2 and G3 of Pople *et al.*[9,10] are capable to give enthalpies of formation which agree with experiment within about 1-2 kcal/mol. However, these procedures are not routinely applied to large seize molecules due to their high computational costs. Thus, *ab initio* methods have been supplemented by semiempirical parameters to estimate $\Delta H^{\circ}_{f}(g)$.[11-14] These procedures require a variable number of empirical parameters, which take into account the atoms or groups in the molecule as well as the molecular environments around these constitutional units. The main advantage of this sort of mixed methods relies upon the fact they are quite inexpensive, rather simple to apply and furthermore their accuracy degree is similar to the high-level *ab initio* procedures, *i.e.* 1-2 kcal/mol.[15]

In a relatively recent paper, Cioslowski *et al.*, have presented an alternative approach to converting Hartree Fock (HF) and Density Functional Theory energies of molecules, ions, and radicals to standard enthalpies of formation.[16] They employed a combination of atomic equivalents bond density functions and corrections for molecular charge and spin multiplicity to produce accurate enthalpy estimates for most organic and inorganic compounds of the first- and second-row elements.

However, when analysing the error statistics for $\Delta H^{\circ}_{f}(g)$ derived from the computed total electronic energies there are several maximum absolute errors quite large that makes this approach rather dubious (see Table 2 in Ref. 16).

The aim of this paper is to present a very simple approach to compute hydrocarbon enthalpies of formation from *ab i nitio* calculations improved through bond parameters which possess a satisfactory chemical accuracy to be applicable for chemical predictive purposes and whose largest deviations do not present any sort of pathological behaviour.

The paper is organized as follows: Next section deals with the method employed to calculate hydrocarbon enthalpies of formation from *ab initio* calculation improved via bond parameters. We present the methodological antecedents as well as other closely related procedures. Then, we discuss the whys and wherefores of the chosen trial set to apply the method. Numerical results are discussed and comparisons are made with other similar procedures, in order to highlight the relative merits of the present method. Finally, we analyse the possibility and suitability to extend the calculations to other molecular sets in order to reach more general conclusions about the proposed methodology.

METHOD

For a given chemical system S, $\Delta H^{\circ}_{f}(g)$ is given by

$$\Delta H^{\circ}_{f}(g)=H(S)+\sum_i n_i(S)\,(\Delta H^{\circ}_{f}(i)-H(i)) \qquad (1)$$

where $n_i(S)$ is the number of atoms of the element i constituting the molecule S. In his turn, H(i) and H(S) at 298.15K are equal to

$$H(R)=E(R)+E_{ZP}(R)+E_{thermo}(R);\ R=S,\ i \qquad (2)$$

where E and E_{ZP} are the total and zero point energy, respectively and E_{thermo} is the difference between the enthalpy at T=298.15K and the energy at T= 0K.

Combining the atomic quantities appearing in Eq. (1) yields the expression

$$\Delta H^{\circ}_{f}(S)=H(S)+\sum_i n_i(S)\ h_i \qquad (3)$$

where h_i is an atomic enthalpy equivalent. A more expedient calculation scheme is gotten via the absoprtion of zero-point energies and thermal corrections into atomic equivalents e_i

$$\Delta H^{\circ}(S)=E(S)+\sum_i n_i(S)\ e_i \qquad (4)$$

which allows one to get rid of the expensive calculations of vibrational frequencies. h_i and e_i are fitted equivalents and their use eliminates large errors associated with the neglect of atomic correlations. The employment of relationships based on Eq. (4) to predict molecular $\Delta H^{\circ}_{f}(g)$ has given reasonably accurate results.[11,14,17-23]

Following a similar foundation, Herndon presented a simple enough protocol to convert HF *ab initio* total electronic energies for hydrocarbons to accurate heats of formation.[24] The optimum procedures use the numbers of carbon and hydrogen atoms and 6-31 G* energies as independent variables. Later on, the numerical relationships were improved through higher-order equations.[25] This approximative scheme is better than the previous one based on atomic equivalents since a significative lower number of parameters are required.

However, on spite of its satisfactory degree of accuracy to predict molecular enthalpies of formation, we deem there is room for further improvements. In fact, in order to have a quite general approach to compute $\Delta H^{\circ}_{f}(g)$, the number and nature of the chemical bonds should be included within the calculation scheme. Thus, it should take into account the existing differences among isomers, which under the present approach are considered equivalents.

According to Cioslowski,[16] a correction term must account for the electron correlation effects associated with bond formation. He proposed a relationship such as

$$\Delta H^{\circ}_{f}(S)=H(S)+\sum_{i} h_1(Z_i)+\sum_{i,j} h_2(Z_i,Z_j,\alpha_{ij}) \qquad (5)$$

where Z_p (p=i,j) is the atomic number corresponding to nucleous p and α_{ij} represent a set of atomic parameters. In Eq. (5) the first sum runs over all the nuclei present in S, whereas the second summation encompass all the atomic interaction lines between nuclear attractor. Since basic quantities involved in Eq. (5) are derived from the theory of atoms in molecules[26] and parameters are finally fitted through a regular regression procedure, we have considered more appropriated to resort to a simpler equation like

$$\Delta H^{\circ}_{f}(S)=A\ E(S)+\sum_{i} a_i n_i+\sum_{i,j} b_{i,j} n_{i,j} \qquad (6)$$

where n_i is the number of i-atoms and $n_{i,j}$ is the number of i-j bonds, and where A, a_i, and $b_{i,j}$ are empirical parameters which are obtained by multilinear regression analyses. This equations generalizes Herndon relationships (1-3).[24]

The next significant point to analyse is the molecular set to be chosen for the present analysis. The first option is to resort to a wide spectrum of possibilities, including different sort of organic and inorganic molecules as well as ions and radicals. We deem this choice is not a very sensible one due to the empirical character of the approach. In fact, in order to compute $\Delta H^{\circ}_{f}(g)$ for a given molecule this method takes into consideration just a few molecular paramerters (*i.e.* just three molecular parameters for alkanes plus total electronic energy) which evidently constitutes an oversimplification when comparing with a strict complete theoretical calculation (see Eqs. (1,2)). A clear illustration of this state of affairs is given by Ciolowsky *et al*'s results (see Tables 2-7 in Ref. 16) where maximum absolute errors are extremely large (*i.e.* ~50 kcal/mol !!!). Then, the second alternative is to concentrate in a rather specialized set, which is the current approach for this sort of analysis. However, this option does not necessarily implies a lack of molecular variations within such restricted choice. For example, Herndon's choice of 65 hydrocarbons comprises examples of planar, nonplanar, alternant and no-alternant aromatic hydrocarbons, alkyl- and alkenyl-substituted benzene derivatives, acyclic and polycyclic alkanes, strained and unstrained olefines and alkynes. Disparate structures of highly strained compounds such as cyclopropene and cubane, combined with polycyclic aromatics like anthracene and perylene, do not require separate parametrizations for different types of C and H atoms. Thus, in this study we have selected this set of molecules as a first step to test the proposed method.

RESULTS

Experimental $\Delta H^{\circ}_{f}(g)$ and the negatives of calculated HF 6-31G* electronic energies, optimized at 6-31G* level are listed in columns 2 and 3, respectively, of *Table* 1. The experimental $\Delta H^{\circ}_{f}(g)$ are modeled by the following equations, obtained by multilinear regression analysis

$$\Delta H^{\circ}_{f}(g)=593.858373\ E+22498.95\ n_C+338.850811\ n_H-0.620987\ n_{C-C}-1.181333\ n_{C=C}-0.397153\ n_{C\equiv C}-0.701003\ n_{C\text{-}Cyclopropane} \qquad (7)$$

Standard Error=1.4437, R^2=0.9987,
Average Deviation= 1.08

$$\Delta H^{\circ}_{f}(g)=589.899403\ E+22349.17\ n_C+336.410813\ n_H-0.626674\ n_{C-C}-1.33081 ln_{C=C}-0.768674n_{C\equiv C}-0.787758n_{C\text{-}Cyclopropane}+0.787197 \qquad (8)$$

Standard Error=1.4504, R^2=0.9987,
Average Deviation=1.08

where E is the total electronic energy (in atomic units) calculated at the 6-31G* basis set level, n_X is the number of X-atoms, and $n_{X\text{-}Y}$ stands for the number of X-Y bonds.

It is interesting to note that regression equations can be interpreted as being mainly concerned with total electronic energy and the number of constituents atoms with correction terms related to the different chemical bonds.

Table 1. Hydrocarbon 631G* total electronic energies (a.u.) and enthalpies of formation (kcal/mol)

Molecule	−Energy	ΔH°_{f} (g) (exp)	ΔH°_{f} (g) Eq. (7)	ΔH°_{f} (g) Eq. (8)
Methane	40.19517	−17.79	−15.88	−15.51
Acethylene	76.81783	54.55	56.30	56.38
Ethylene	78.03172	12.56	12.33	12.578
Ethane	79.22876	−20.04	−20.28	−20.04
Propyne	115.86432	44.41	44.24	44.25
Allene	115.86110	45.31	44.81	44.88
Propene	117.07147	4.79	4.28	4.41
Propane	118.26365	−24.93	−26.06	−25.96
1,3-Butadiene	154.91965	26.01	25.60	25.62
2-Butyne	154.90926	34.69	33.11	33.02
1-Butene	156.10608	0.07	−1.27	−1.43
(Z)-2-Butene	156.10786	−1.77	−1.77	−1.77
(E)-2-Butene	156.11041	−2.84	−3.28	−3.28
Isobutene	156.11067	−4.27	−3.44	−3.43
Cyclobutane	156.09720	6.78	4.50	4.59
n-Butane	157.29840	−30.33	−30.52	−30.55
Isobutane	157.29897	−32.24	−30.86	−30.88
Cyclopentadiene	192.79172	32.12	32.66	32.83
1,3-Pentadiene	193.95916	18.29	17.69	17.60
1,4-Pentadiene	193.94721	25.27	24.78	24.65
Cyclopentene	193.97719	8.44	6.92	7.05
Cyclopentane	195.16358	−18.26	−19.37	−19.28
n-Pentane	196.33302	−35.60	−35.52	−35.68
Cyclohexane	234.20800	−29.49	−30.19	−30.20
Cyclopropene	115.82305	66.22	67.35	67.40
Cyclopropane	117.05887	12.73	11.71	11.92
Cyclobutane	154.89961	37.45	37.44	37.52
Neopentane	196.33383	−40.14	−36.01	−36.16
Cubane	307.39391	148.69	146.52	146.41
Bicyclo(1.1.0)butane	154.87177	51.90	53.91	54.03
Bicyclo(2.1.0)pentane	193.92697	37.70	36.68	36.75
Bicyclo(2.2.0)hexane	232.96556	29.90	29.32	29.27
Bicyclo(2.2.1)heptane	272.06116	−12.40	−11.90	−11.84
Bicyclo(2.2.2)octane	311.10358	−23.67	−21.53	−21.58
Spiropentane	193.91753	44.25	42.29	42.32
Bicyclo(2.1.0)pentene	192.71022	79.70	81.00	80.98
Bicyclo(2.2.0)hexene	231.76849	62.50	61.95	61.90
n-Hexane	235.36779	−39.94	−40.62	−40.91
Cycloheptatriene	269.68233	43.56	44.33	44.30
Norbornadiene	269.65251	59.18	61.98	61.97
Quadricyclane	269.61822	81.04	82.22	82.35
Cyclooctatetraene	307.52422	70.30	70.98	70.76
Benzene	230.70310	19.80	17.64	17.91
Naphthalene	383.35500	36.00	34.03	34.20
Anthracene	535.99880	55.20	55.24	55.27
Phenantrene	536.00980	49.70	48.71	48.78
Benz(a)anthracene	688.65688	68.10	67.96	67.91
Chrysene	688.66090	66.00	65.57	65.54

Table 1. (Continued)

Molecule	−Energy	ΔH°_f (g) (exp)	ΔH°_f (g) Eq. (7)	ΔH°_f (g) Eq. (8)
Triphenylene	688.66030	66.50	65.93	65.89
Benzo(c)phenanthrene	688.64950	69.60	72.34	72.27
Pyrene	611.76800	54.00	54.86	55.03
Perylene	764.40650	78.40	79.22	79.23
Acenaphthene	460.26060	37.23	36.50	36.43
Biphenylene	459.01460	99.80	98.75	98.62
Acenaphthylene	459.07380	62.20	63.59	63.70
Azulene	383.28260	73.53	77.03	76.91
Fluoranthene	611.74562	69.20	68.16	68.24
Cis-Stilbene	537.13326	60.30	59.61	59.44
Trans-Stilbene	537.13943	56.40	55.95	56.00
Biphenyl	460.25394	43.30	41.24	41.30
o-Xylene	308.77622	4.56	5.32	5.35
m-Xylene	308.77724	4.14	4.72	4.75
p-Xylene	308.77704	4.31	4.84	4.87
Styrene	307.58540	35.40	34.25	34.29
Toluene	269.74016	11.95	11.19	11.34

Table 2. Alkenes 6-31G* total electronic energies (a.u.) and enthalpies of formation (kcal/mol)

Molecule	−Energy	ΔH°_f(g) (exp.)	ΔH°_f(g) (Eq. 7)	ΔH°_f(g) (Eq. 8)
Cis-2-Pentene	195.14229	−6.60	−6.66	−6.80
Trans-2-Pentene	195.14504	−7.62	−8.29	−8.42
2-Methyl-2-butene	195.14577	−9.99	−8.73	−8.85
2-Methyl-2-pentene	234.18020	−15.99	−13.62	−13.87
2-Methyl-1-butene	195.14374	−8.44	−7.52	−7.65
3,3-Dimethyl-1-butene	234.17395	−14.46	−9.91	−10.19
3-Methyl-1-butene	195.14116	−6.60	−5.99	−6.13
2,3-Dimethyl-2-butene	234.17701	−16.30	−11.72	−11.99
2,3-dimethyl-1-butene	234.17700	−14.96	−11.72	−11.99
1-Methylcyclopentene	233.01793	−0.91	−1.71	−1.70
3-Methylcyclopentene	233.01300	1.77	1.21	1.21
Cycloheptene	272.04603	−2.20	−2.85	−2.99
1-Methylcyclohexene	272.05863	−10.35	−10.33	−10.42
Norbornene	270.86184	21.51	22.06	22.13
Cis-4,4-Dimethyl-2-Pentene	273.20350	−17.35	−17.27	−17.66
Trans-4,4-Dimethyl-2-Pentene	273.21254	−21.22	−17.27	−17.66
2-Bicyclo(2.2.2)octene	309.91244	4.90	7.57	7.55
Cyclohexene	233.01965	−1.20	−2.74	−2.71

Average absolute error Eq. (7)=1.61 kcal/mol , Average absolute error Eq. (8)=1.55 kcal/mol

All the data listed in *Table* 1 were used to obtain the multilinear regression equations. The statistical parameters seem to show that Eq. (7) and Eq. (8) furnish very good correlations of the ΔH°_f(g) data. The calculated values of enthalpies of formation are listed in Table 1, columns 3 and 4, respectively. Only nine molecules exhibit $\Delta\Delta H^{\circ}_f$(g) (exp.calc.) values larger than ±2 kcal/mol by either regression analysis. Only two molecules are badly predicted, with a $\Delta\Delta H^{\circ}_f$(g) around 4 kcal/mol. The average deviations are quite good (*i.e.* 1.08 kcal/mol) since the usual experimental uncertainties lie about ±2 kcal/mol.

When comparing with previous results obtained in a similar, but less generalized context (see Refs. 24 and 25, for example), present predictions are somewhat better than those precedent ones. In fact, standard deviations and mean deviations of equations (7) and (8) are lower than corresponding parameters in previous correlations.

As a further test of the present approximation, we have computed the heats of formation of 18 alkenes not included in the testing set. They are those reported reported by Schmitz and Chen in their study of heats of formation from *ab initio* theory and a group equivalent scheme for a set of alkenes.[27] From their list of 26 alkenes (see *Table* 1 in Ref. 27) we have excluded those molecules included in our test set (*i.e.* ethylene, propylene, 1-butene, *cis*-2-butene, *trans*-2-butene, isobutene, and cyclopentene) plus *trans*-1,2-*di*-*t*-butylethylene. This last molecule presented an anomalous numerical behaviour, probably due to some sort of error in the total electronic energy or/and the experimental heat of formation. Results are presented in *Table* 2 and theoretical predictions show again a satisfactory agreement with experimental values since the average absolute errors are less than 2 kcal/mol and there are only two rather large deviations between theoretical and experimental data (~4 kcal/mol) exceeding the usual experimental uncertainties (~2 kcal/mol).

DISCUSSION

We have shown the need to take into account the bond contribution when modeling suitable equations to convert HF total electronic energies of molecules to standard enthalpies of formation. Although previous approximations using just the number of constituent atoms gave satisfactory enough predictions of $\Delta H^{\circ}_{f}(g)$, the inclusion of bond contributions improve enthalpy estimates for hydrocarbon molecules. The average absolute errors in the computed values of heat of formation is 1.08 kcal/mol for the set of 65 hydrocarbon that include quite different sort of molecules, from systems as small as CH_4 up to a large species like perylene. Furthermore, another genuine predictive calculation for a set of 18 alkenes yielded once again a rather satisfactory agreement with experimental data.

Present approach gives computationally inexpensive theoretical predictions of molecular thermochemistry with a very good accuracy which is lower than that of calorimetric measurements. Furthermore, the chosen descriptors set (number of atoms and bonds) are the simplest among the huge amount of current parameters employed in this sort of calculation. Thus, we have at hand a simple enough and sufficiently accurate algorithm to predict enthalpies of formation which may be of interest to a broad audience of experimental chemists.

Since present approach have ben tested for one representative set of molecules (*i.e.* hydrocarbons), the final conclusions are not totally definitive, albeit they are quite encouraging. Now, it should be necessary to extend this procedure to quite different sets of molecules (organic molecules including heteroatoms, inorganic compounds, ions, radical, etc.). At present, research along this line is being carried out in our laboratory and results will be published elsewhere in the near future.

REFERENCES

1. *Computational Thermochemistry*; K. K. Irikura and D. J. Frurip, Ed.; ACS Symposium Series No. 677, Washington, D.C., 1998.
2. S. Hammerun *Chem. Phys. Lett.* **1999**, *300*, 529.
3. G. A. Petersson; D. K. Malick; W. G. Wilson; J. W. Ochterski; J. A. Montgomery, Jr.; M. J. Frisch *J. Chem. Phys.* **1998**, *109*, 10570.
4. J. A. Pople; S. Binkley; R. Seeger *Int. J. Quantum Chem.: Quantum Chem. Symp.* **1976**, *10*, 1.
5. D. R. Stull; E. F. Westrum, Jr.; G. C. Sinke *The Chemical Thermodynamics of Organic Compounds*; Wiley, New York, 1969.
6. E. S. Donalski *J. Phys. Chem. Ref. Data* **1872**, *1*, 211.
7. R. F. Gould, Ed.; Advanced Propellant Chemistry, Advances in Chemistry; Series No. 54, American Chemcial Society, Washington, D.C., 1966.
8. P. A. Atkins *Physical Chemistry*, 5th Ed.; Freeman, New York, 1994.
9. L. A. Curtiss; K. Raghavachari; J. A. Pople *J. Chem. Phys.* **1993**, *98*, 1293.
10. L. A. Curtiss; K. Raghavachari; P. C. Redfern; V. Rassolov; J. A. Pople *J. Chem. Phys.* **1998**, *109*, 7764.
11. M. R. Ibrahim; P. R. Schleyer *J. Comput. Chem.* **1985**, *6*, 157.
12. N. L. Allinger; L. R. Schmitz; I. Motoc; C. Bender; J. K. Labanowsky *J. Phys. Org. Chem.* **1990**, *3*, 732.
13. C. Vericat; E. A. Castro *Commun. Math. Comput.*

Chem. MATCH **1996**, *34*, 167.
14. E. A. Castro *J. Mol. Struct. THEOCHEM* **1994**, *304*, 93.
15. R. A. Friesner; R. B. Murphy; M. D. Beachy; M. N. Ringnalda; W. T. Pollard; B. D. Dunietz; Y. Cao *J. Phys. Chem.* **1999**, *103*, 1913.
16. J. Cioslowski; G. Liu; P. Piskorz *J. Phys. Chem. A* **1998**, *102*, 9890.
17. Z. Yala *J. Mol. Struct. THEOCHEM* **1990**, *207*, 217.
18. C. Vericat; E. A. Castro *Egyp. J. Chem.* **1998**, *41*, 109.
19. E. A. Castro; L. Gavernet; M. Firpo *Acta Chim. Slov.* **1997**, *44*, 327 (1997).
20. E. A. Castro *J. Mol. Struct. THEOCHEM* **1995**, *339*, 239.
21. E. A. Castro *J. Chem. Soc. Pak.* **1995**, *17*, 156.
22. E. A. Castro *Ing. Cs. Quim.* **1996**, *16*, 75.
23. E. A. Castro *Bol. Soc. Quim. Perú* XLII, **1996**, *17*, 239.
24. W. C. Herndon *Chem. Phys. Lett.* **1995**, *234*, 82.
25. E. A. Castro *Comp. Chem.* **1997**, *21*, 305.
26. R. F. W. Bader *Atoms in Molecules: A Quantum Theory*; Clarendon Press: Oxford, UK, 1990.
27. L. R. Schmitz; Y. R. Chen *J. Comput. Chem.* **1994**, *15*, 1437.