Enthalpies of formation from theoretical methods: a new approach based on density functional theory plus semiempirical corrections

2 **AUTHORS**, INCLUDING:

Alessandro Fortunelli
Italian National Research Council

210 PUBLICATIONS   **3,975** CITATIONS

SEE PROFILE

Enthalpies of formation from theoretical methods: a new approach based on density functional theory plus semiempirical corrections

Alessandro Fortunelli[a,*], Massimo Selmi[b]

[a]Instituto di Chimica Quantistica ed Energetica Molecolare del C.N.R., Via Risorgimento 35, I-56126 Pisa, Italy
[b]Scuola Normale Superiore, I-56100 Pisa, Italy

Abstract

A theoretical method to obtain enthalpies of formation of organic compounds is considered. The application of the method requires modest computer facilities, because it is based, for its most computationally demanding part, on density functional theory (DFT) which is now implemented in many efficient programs.

The proposed procedure is based on "isodesmic reactions". Starting from the hypothesis that the main source of error in density functional calculations on isodesmic reactions is the neglect of intra-molecular dispersive contributions to the electronic energy, the DFT results are corrected with a simple semiempirical formula involving a limited number of parameters.

The involved approximations are checked through MP2 calculations with localized molecular orbitals, using the orbital invariant formulation of Møller–Plesset perturbation theory proposed by Pulay and Saebø.

The method has been tested on a set of hydrocarbon molecules: the final results agree well with experimental enthalpies of formation, the standard deviation being lower than $1.0 \, \text{kcal mol}^{-1}$.

1. Introduction

It has been shown in the literature that high level ab initio procedures such as G1 and G2 [1,2] are able to give enthalpies of formation which agree with experimental data within $1-2 \, \text{kcal mol}^{-1}$. Unfortunately these procedures cannot be routinely applied to large size molecules because of their high computational costs.

Ab initio methods plus semiempirical parameters have been used by Ibrahim and Schleyer [3] and by Allinger et al. [4] to estimate enthalpies of formation. These procedures require a large number of empirical parameters which depend on the atoms or groups in the molecule and also on the molecular surroundings around these units.

In the present article an approach for calculating enthalpies of formation of organic compounds (already outlined in Ref. [5]) is developed, which is based on a combination of density functional (DF) and semiempirical techniques, and further justified through second-order Møller–Plesset (MP2) perturbation calculations. The approach is founded on firm theoretical grounds and requires a limited number of semiempirical parameters.

In order to obtain enthalpies of formation we consider "isodesmic reactions", i.e. transformations in which the number of bonds each formal

* Corresponding author.

type is conserved and only the relationships among the bonds are altered [6]. This allows one to concentrate only on the changes of the environment in which the bonds are located, because the correlation energy contributions inherent in individual bonds largely cancel in isodesmic reactions.

In particular, we concentrate attention on hydrocarbons and on a class of isodesmic reactions of particular importance: the bond separation reactions (bsr), in which all formal bond between carbon atoms are separated into the simplest parent (two-carbon-atom) molecules containing these same kinds of bonds. For example, for the cyclopentene bond separation reaction one has

$$C_5H_8 + 5CH_4 \rightarrow 4C_2H_6 + C_2H_4 \qquad (1)$$
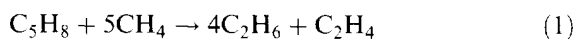
By calculating thermal, rotational, zero-point vibrational and electronic energies for all the molecules involved in Eq. (1) one can obtain the heat of reaction (1) and hence the heat of formation of the cyclopentene molecule in gas-phase

$$\Delta H_f^0(C_5H_8) = 4\Delta H_f^0(C_2H_6) + \Delta H_f^0(C_2H_4)$$
$$- 5\Delta H_f^0(CH_4) - \Delta H_{bsr}^0 \qquad (2)$$

where the ΔH_f^0 values of methane, ethylene and ethane are the experimental ones, and ΔH_{bsr} is the theoretically predicted heat of reaction (1).

2. Method

The hydrocarbon molecules considered in the present work are given in Table 1. The geometry of the molecules was obtained through a molecular mechanics program (using the MM3 force field [7]). The same program was utilized to obtain the zero-point vibrational and rotational contributions to ΔH_{bsr}, whereas the electronic contribution was obtained through single-point DF calculations at the MM3 geometries.

The use of geometrical parameters and vibrational frequencies taken from a molecular mechanics program is justified by the fact that these are very similar to the experimental ones for hydrocarbons, whereas only with very sophisticated and computationally expensive ab initio methods is it possible to improve them: see in

Ref. [5] the fact that the enthalpies of formation change by only 0.3 to $1.0\,\text{kcal mol}^{-1}$ if using DF geometries and frequencies.

The electronic contribution to ΔH_{bsr} was calculated within the density functional theory (DFT), using the GAUSSIAN92/DFT [8] set of codes with the standard D95** basis set [9]. The DF calculations employed the correlation potential of Perdew [10], summed to standard Hartree–Fock Hamiltonian (hybrid functional), the resulting Kohn–Sham equations [11,12] being solved self-consistently. The results of these calculations are also given in Table 1, together with the experimental data (taken from Ref. [13]).

It is to be noted that when calculating enthalpies of formation it is essential to achieve a high level of accuracy, because of the tiny quantities involved: in the present calculations this is assured by the fact that only the correlation energy functional is evaluated numerically, whereas the much larger exchange contribution is evaluated in the standard way.

3. Density functional results

From Table 1, one gets an average absolute error of $6\,\text{kcal mol}^{-1}$ with respect to the experimental values. We guess that this error is mainly due to the electronic contribution and, in particular, may be attributed [5] to the Coulomb correlation of electrons belonging to different bonds or atomic domains. In fact, for bond separation reactions one may suppose that in passing from reactants to products the correlation energy arising from electrons belonging to the same portion of atoms or bonds is conserved, whereas the non-local contributions are not. One sees, for example, that the number of formal bonds and first-neighbor pairs of formal bonds is conserved in bond separation reactions, whereas the number of second-neighbor and more distant pairs of formal bonds is always greater in reactants. The correlation potential used in the DF calculations of Table 1 is a functional of the electronic density and its gradients; hence it is not able to describe non-local effects. In fact, one would require a truly non-local potential in order to be able to reproduce them: a potential in which

Table 1
Experimental[a] and DFT values (through bond separation reactions) of heats of formation in the gas phase for some hydrocarbons, and the difference (diff) between the two set of data[b]: diff = ΔH_f^{298K}(exp) − ΔH_f^{298K}(calc)

Compound	ΔH_f^{298K}(exp)	ΔH_f^{298K}(calc)	Diff
Cyclopentane	−18.74	−13.99	−4.75
Cyclohexane	−29.43	−24.02	−5.41
Norbornane	−12.42	−3.90	−8.52
[2.2.2]Bicyclooctane	−22.58	−12.41	−10.17
[3.3.0]Bicyclooctane	−22.07	−10.63	−11.44
Cyclopentene	7.87	12.83	−4.96
Cyclohexene	−1.28	3.43	−4.71
Cyclopentadiene	32.00	37.35	−5.35
Dicyclopentadiene-endo	42.20	60.79	−18.59
Norbornadiene	58.00	71.69	−13.69
Norbornene	21.40	30.80	−9.40
1,3-Butadiene	26.33	28.33	−2.00
2-Ethyl-1-pentene	−17.90	−12.23	−5.67
Propane	−24.82	−24.35	−0.47
n-Butane	−30.15	−28.42	−1.73
n-Pentane	−35.00	−32.49	−2.51
Isobutene	−4.04	−2.99	−1.05
1-Butene	−0.03	1.46	−1.49
2-Methyl-1-butene	−8.68	−6.07	−2.62
2-Ethyl-1-butene	−12.32	−9.16	−3.16
cis-3-Hexene	−11.38	−7.31	−4.07
cis-2-Butene	−1.67	−0.15	−1.52
Methylencyclohexane	−6.02	−2.35	−3.67
Methylencyclopentane	2.87	7.62	−4.75
Cyclobutene	31.00	45.57	−14.57
2-Methyl-1,3-butadiene	18.10	24.27	−6.17

[a] Experimental values from Ref. [13]. [b] All data are in kcal mol^{-1}.

the correlation functional at a point r_1 not only depends on the value of ρ (and, possibly, of its derivatives: gradient-corrected functionals) at the same point r_1, but also on the values of ρ over all space (in particular, at those points r_2 in which there is an accumulation of charge due to another bond) [5]. The error in DF calculations on these reactions may thus be attributed to the neglect of the correlation of electrons belonging to different bonds: the currently used density functionals do not properly take into account the stabilization of reactants due to inter-bond dispersion effects, and this explains the negative sign in error values.

This is confirmed by the fact that MP2 calculations [14] on the same hydrocarbons, in which the inter-bond dispersion effects are approximately taken into account, show a notable decrease of the errors in the calculated enthalpies of formation.

In Ref. [5] this suggested a way to improve the DFT results by adding a semiempirical correction which takes into account dispersion effects. In the following section such a suggestion is checked.

4. Semiempirical corrections

In the localized picture of molecular orbitals one may associate valence orbitals to bonds or lone pairs. Using the orbital invariant formulation of the Møller–Plesset perturbation theory proposed by Pulay and Saebø [15] it is possible to obtain the various intra- and inter-bond contributions to the molecular correlation energy at all orders in perturbation theory. As an application of this methodology, the results of MP2 calculations with localized orbitals are reported in Table 2 for a subset of the compounds of Table 1: the

Table 2
Reaction differences of MP2 correlation energies for isodesmic reactions[a]; first column, total MP2; other columns, contributions from vicinal pairs, non-vicinal pairs, and intrabond contributions, respectively

Compound	$\Delta E_{bsr}^{(2)}$(tot)	$\Delta E_{bsr}^{(2)}$(vic)	$\Delta E_{bsr}^{(2)}$(dist)	$\Delta E_{bsr}^{(2)}$(intra)
Cyclopentane	0.009680	0.000684	0.008657	−0.000061
Cyclopentene	0.010652	0.001188	0.010064	−0.001037
Cyclobutene	0.008842	0.002190	0.005641	0.000207
n-Pentane	0.006979	−0.000103	0.006874	−0.000156
n-Butane	0.004504	−0.000042	0.004434	−0.000117
Propane	0.002119	−0.000003	0.002092	−0.000070
cis-2-Butene	0.005205	0.000289	0.005828	−0.001007

[a] All data are in atomic units.

contributions from intra-bond ($\Delta E_{bsr}^{(2)}$(intra)), first-neighbor ($\Delta E_{bsr}^{(2)}$(vic)) or more distant ($\Delta E_{bsr}^{(2)}$(dist)) pairs of orbitals to MP2 total reaction energy are explicitly distinguished. The MP2 calculations employed the standard 6-31G* basis set [16]. The occupied and virtual orbitals were separately localized through the Boys procedure [17].

From Table 2 it appears that the contributions arising from distant pairs of bonds ($\Delta E_{bsr}^{(2)}$(dist)) give the main contribution to the total $\Delta E_{bsr}^{(2)}$ of hydrocarbon isodesmic reactions (about 90% on average), the contributions from vicinal pairs being very small and tending to compensate those from intra-bond pairs in some cases.

Based on these results one can write a simple expression which approximates the non-vicinal inter-bond terms as a sum of (intra-molecular) dispersive interactions arising from instantaneous dipoles directed along the bonds [5].

Using the common approximation of dipolar expansion, one thus gets

$$E_d \approx \sum_{i}^{n_B} \sum_{j>i}^{n_B}{}' E_{d_{ij}} = -\sum_{i}^{n_B} \sum_{j>i}^{n_B}{}' \frac{a_i^2 a_j^2}{\epsilon_i + \epsilon_j} f(\mathbf{b}_i, \mathbf{b}_j, \mathbf{R}_{ij})$$

(3)

where E_d is the intra-molecular dispersion energy, $E_{d_{ij}}$ are the contributions of each pair of non-vicinal formal bonds to E_d, n_B is the total number of formal bonds in the molecule, a_i, ϵ_i are parameters typical of each formal bond (to be derived from theoretical calculations or from fitting the experimental data), $f(\mathbf{b}_i, \mathbf{b}_j, \mathbf{R}_{ij})$ is the standard

dipole–dipole interaction

$$f(\mathbf{b}_i, \mathbf{b}_j, \mathbf{R}_{ij}) = \frac{1}{R_{ij}^6}\left[\frac{3}{2}\frac{\mathbf{b}_i \cdot \mathbf{R}_{ij}}{R_{ij}}\frac{\mathbf{b}_j \cdot \mathbf{R}_{ij}}{R_{ij}} - \frac{1}{2}\mathbf{b}_i \cdot \mathbf{b}_j\right]^2$$

(4)

(squared because virtual, i.e. dispersive, interactions are considered), $\mathbf{R}_{ij}$ is the vector joining the middle points of the ith and jth bonds, $\mathbf{b}_i$ is a unit vector lying along the ith bond. We stress again that the nearest-neighbor pairs of formal bonds have not been included in the sums in Eq. (3), at variance with what is performed in Ref. [5] where two constant terms (ΔE_1 and ΔE_2) had been introduced to describe (together with the nearest-neighbor $E_{d_{ij}}$ terms) the correlation energy associated with intra-bond and nearest-neighbor pairs of bonds. Here it is supposed that, as suggested by the results of Table 2, these contributions should altogether play a minor role in hydrocarbon isodesmic reactions, and neither nearest-neighbor pairs nor constant terms were introduced into Eq. (3).

The final corrections to the calculated enthalpies of formation of Table 1 are obtained by calculating the E_d values from Eq. (3) for all the compounds involved in the isodesmic reaction, and then making the difference between products and reactants (weighted by the stoichiometric coefficients).

Equation (3) is based on a picture of the correlation energy of two (distant) charge clouds, which depends on their relative orientation, shape and distance. This "dispersion-like" energy due to the attraction of instantaneous dipoles directed

along the bonds has stabilizing effects and is thus negative in sign for both products and reactants. However, the difference between products and reactants is the most important quantity in isodesmic reactions: it properly turned out that this is always negative for the compounds listed, if choosing the parameters as explained below.

The parameters in Eq. (3) (two for each formal bond) might have been obtained from the ab initio MP2 calculations with localized orbitals, but we preferred to get them from direct fit of the experimental enthalpies of formation. In Table 3 the enthalpies of formation calculated in this way and the corresponding errors for the hydrocarbons of Table 1 (with the exception of the conjugated ones) are reported. The standard deviation is 0.84 kcal mol^{-1} and the parameters of the fit are $a_{C-C} = 9.16$, $a_{C=C} = 10.11$, $a_{C-H} = 10.65$, $\epsilon_{C-C} = 9.16$, $\epsilon_{C=C} = 2.65$ and $\epsilon_{C-H} = 16.30$, if measuring the R_i in Å and the energies in kcal mol^{-1}.

The fitting procedure is non-linear, because of the non-linear form of Eq. (3). However, if one considers the factors of $f(\mathbf{b}_i, \mathbf{b}_j, \mathbf{R}_{ij})$ in Eq. (3) as the parameters of the fit, the expression may be fitted with a linear procedure. The corresponding parameters came out to be very close to those obtained from the non-linear fit (the differences being $\leq 4\%$, with the exception of the C=C/C=C interaction): this constitutes further support to the internal consistency of the approach. Also, the introduction of ΔE_1 and ΔE_2 as two additional fitting parameters (as described in Ref. [5]) does not substantially improve the quality of the fit: the standard deviation is slightly decreased to 0.77 kcal mol^{-1}.

It is finally to be noted that, if one utilizes the standard Hartree–Fock results for calculating the electronic contributions to ΔH_{bsr} (i.e. if one does not correct the Fock operator and the total energy through the Perdew's correlation functional) the corresponding standard deviation comes out to be larger (2.0 kcal mol^{-1}) than the one obtained using the hybrid functional results. This can be understood by thinking that, through

Table 3
Experimental[a] and calculated (through bond separation reactions: DFT + semiempirical corrections) values of heats of formation in the gas phase for some hydrocarbons, and the difference (diff) between the two set of data[b]: diff $= \Delta H_f^{298K}(\text{exp}) - \Delta H_f^{298K}(\text{calc})$

Compound	$\Delta H_f^{298K}(\text{exp})$	$\Delta H_f^{298K}(\text{calc})$	Diff
Cyclopentane	−18.74	−18.81	0.07
Cyclohexane	−29.43	−28.47	−0.96
Norbornane	−12.42	−13.82	1.40
[2.2.2]Bicycloctane	−22.58	−22.08	−0.50
[3.3.0]Bicycloctane	−22.07	−21.31	−0.76
Cyclopentene	7.87	7.43	0.44
Cyclohexene	−1.28	−0.78	−0.50
Dicyclopentadiene-endo	42.20	43.15	−0.95
Norbornadiene	58.00	57.74	0.26
Norbornene	21.40	20.01	1.39
2-Ethyl-1-pentene	−17.90	−16.66	−1.24
Propane	−24.82	−25.07	0.25
n-Butane	−30.15	−29.98	−0.17
n-Pentane	−35.00	−34.90	−0.10
Isobutene	−4.04	−4.60	0.56
1-Butene	−0.03	0.20	−0.23
2-Methyl-1-butene	−8.68	−8.87	0.19
2-Ethyl-1-butene	−12.32	−13.33	1.01
cis-3-Hexene	−11.38	−10.11	−1.27
cis-2-Butene	−1.67	−1.43	−0.25
Methylencyclopentane	2.87	3.65	−0.78
Cyclobutene	31.00	31.21	−0.21

[a] Experimental values from Ref. [13]. [b] All data are in kcal mol^{-1}.

the correlation functional, one manages to take into account a part of the local (non-dispersive) correlation contributions to $dE_{bsr}^{(2)}$(intra) and $\Delta E_{bsr}^{(2)}$(vic), thus further reducing their importance in the final value of $\Delta E_{bsr}^{(2)}$.

For comparison, the corresponding standard deviation of the MM3 ΔH_{bsr} values for the compounds of Table 1 is about $1.6\,kcal\,mol^{-1}$.

5. Conclusions

The analysis of the results shows that the proposed procedure is suitable for the calculations of enthalpies of formation for non-conjugated hydrocarbons, including also rather strained molecules.

This procedure is based primarily on DFT. The semiempirical correction is intended to describe only that part of the correlation energy which is not accessible to the currently used density functionals. This allows one to limit the number of semiempirical parameters, to make them general and easily derivable, and to provide them with a sound theoretical basis, still getting sufficiently accurate final values of predicted reaction enthalpies.

Conjugated molecules have not been included in the semiempirical procedure, because this is based on interactions between localized charge distributions, whereas in conjugated π systems it is difficult to define localized bonds for which Eq. (3) may describe correctly the attractive forces arising from dispersion effects.

Finally, we remark that the present approach is easily generalized to molecules containing heteroatoms, the number of parameters being always limited to two for each type of formal bond or lone pair.

Acknowledgements

References

[1] J.A. Pople, M.A. Head-Gordon and D.J. Fox, J. Chem. Phys., 90 (1989) 5622.
[2] L.A. Curtiss, K. Raghavachari and G.W. Trucks, J. Chem. Phys., 94 (1991) 7221.
[3] M.R. Ibrahim and P. von Ragué Schleyer, J. Comp. Chem., 6 (1985) 157, and references cited therein.
[4] N.L. Allinger, L.R. Schmitz, I. Motoc, C. Bender and J.K. Labanowsky, J. Phys. Org. Chem., 3 (1990) 732.
[5] A. Fortunelli and M. Selmi, Chem. Phys. Lett., 223 (1994) 390.
[6] W.J. Hehre, L. Radom, P. von Ragué Schleyer and J.A. Pople, Ab Initio Molecular Orbital Theory, J. Wiley and Sons, New York, 1986, pp. 300–308.
[7] N.L. Allinger, Y.H. Yuh and J.-H. Lii, J. Am. Chem. Soc., 111 (1989) 8551.
[8] M.J. Frish, G.W. Trucks, M. Head-Gordon, P.M.W. Gill, M.W. Wong, J.B. Foresman, B.G. Johnson, H.B. Schlegel, M.A. Robb, E.S. Reploge, R. Gomperts, J.L. Andrés, K. Raghavachari, J.S. Binkley, C. Gonzalez, R.L. Martin, D.J. Fox, D.J. DeFrees, J. Baker, J.J.P. Stewart and J.A. Pople, GAUSSIAN92/DFT, Gaussian Inc., Pittsburg, PA, 1993.
[9] T.H. Dunning and P.J. Hay, Modern Theoretical Chemistry, Plenum, New York, 1976, pp. 1–28.
[10] J.P. Perdew, Phys. Rev. B, 33 (1986) 8822.
[11] W. Kohn and L. Sham, Phys. Rev. A, 140 (1965) 1133.
[12] R.G. Parr and W. Yang, Density Functional Theory of Atoms and Molecules, Oxford University Press, New York, 1989.
[13] D.R. Stull, E.F. Westrum and G.C. Sinke, The Chemical Thermodynamic of Organic Compounds, John Wiley and Sons, New York, 1969.
[14] M. Selmi and J. Tomasi, J. Phys. Chem., in press.
[15] P. Pulay and S. Saebø, Theor. Chim. Acta, 69 (1986) 357.
[16] W.J. Hehre, R. Ditchfield and J.A. Pople, J. Chem. Phys., 56 (1972) 2257.
[17] S.F. Boys, Rev. Mod. Phys., 32 (1960) 296.