Empirical Stability Rules and Conjugation in the Estimation of Enthalpies of Formation

MICHAEL L. MAVROVOUNIOTIS

Department of Chemical Engineering and Systems Research Center, A. V. Williams Building (115), Room 2149, University of Maryland, College Park, Maryland 20742

Some empirical rules frequently used in the comparison of stabilities of similar compounds can be explained by conjugation. A method is proposed for the quantification of these empirical arguments and their application to the estimation of enthalpies of formation. A simple version of the method is demonstrated for alkanes. By using four adjustable parameters, the method fits the enthalpies of formation of alkanes with a standard deviation of 2.8 kJ/mol, while a group contribution method with four parameters has a standard deviation of 4.7 kJ/mol.

INTRODUCTION

The accurate estimation of physical and thermodynamic properties from the molecular structure of a compound is an important but difficult problem, while experimental procedures for determining properties are costly and time consuming.[1,2]

In additive group contribution methods,[1,3-7] a property is estimated as a summation of contributions of the groups comprising the structure of the molecule. The advantage of these methods is that they can give quick estimates without requiring substantial computational resources. The basic group contribution approach, however, represents a very simplified view of chemical molecular structures, largely ignoring concepts such as delocalized bonds, resonance, and other electronic interactions among groups. In order to take such complications into account, some group contribution methods introduce special corrections. However, this leads to additional adjustable parameters which may or may not be physically meaningful.

The use of additive contributions for the ideal-gas enthalpy of formation, which is the property of interest here, assumes that groups do not interact with each other, i.e., the particular way in which groups are connected does not affect their contributions to the enthalpy of formation. For this reason, the estimated enthalpies of formation are often unable to distinguish among isomers. For example, the 2.7 kJ/mol difference in the enthalpies of formation of 2-methylpentane and 3-methylpentane[1] cannot be predicted by group contribution methods without extra corrections. The difference may at first appear to be insignificant in this case, but it becomes very important if one considers some of the applications of enthalpies of formation. In chemical reaction processes, the selectivity is often determined by the relative stability of very similar isomeric compounds. A small difference in the stabilities of two compounds or reaction intermediates may have a substantial impact on the selectivity of a process. These differences are more pronounced in compounds that are affected significantly by resonance, i.e., hybrids of a number of conjugate forms. Group contribution methods do not handle resonance in any systematic way (except for the common discrimination between aromatic and nonaromatic carbons).[4]

The relative stabilities of similar compounds are often compared, in organic chemistry, by using empirical rules. This paper focuses on the quantification and utilization of such empirical rules in the estimation of enthalpies of formation. While the approach discussed here may ultimately lead to methods for the estimation of many physical and chemical properties of organic compounds (including unstable reaction intermediates), the scope of this brief paper will be confined to enthalpies of formation of alkanes.
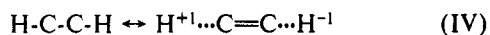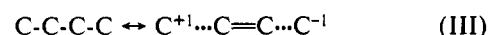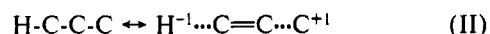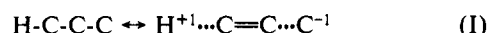
EMPIRICAL RULES

Many empirical rules used in organic chemistry are based on the idea of conjugation or resonance. Whenever a com-

Table I. Enthalpies of Formation of C_5H_{12} Isomers (Ref 1)

compound	enthalpy of formation (kJ/mol)
n-pentane	−146.5
2-methylbutane	−154.6
2,2-dimethylpropane	−166.1

pound can be described by several electronic arrangements (which are compatible with the basic topology of the molecule), its stability is enhanced. The stabilization is more pronounced when the alternative structural formulas are energetically similar. The real compound is viewed as a *hybrid* of all its *conjugates* (structural formulas). Conjugation can explain interactions among atoms that are separated by several bonds and is especially pronounced in the delocalization of π-electrons. Group contribution methods do not take into account the conjugates and do not view the compound as a hybrid (except in the special case of aromatic compounds). Instead, they work with a single form of the molecular structure—essentially the most dominant of the conjugates. Conjugation can extend over several bonds and involve single and double bonds; in this paper, no distinction will be made between hyperconjugation (involving σ bonds) and conjugation involving π bonds.

To give a better idea of the empirical rules of interest here, and their relation to conjugation, two specific rules will be considered in this section. Consider first the widely used rule stating that branched alkanes are more stable than their straight-chain isomers.[8] For example, the alkanes with the molecular formula C_5H_{12} have enthalpies of formation shown in Table I. The differences observed here cannot be due to steric hindrance among side chains because steric interactions correspond to the reverse from the observed order of stability. The basis of the rule lies with an electronic conjugation effect. One would expect conjugates that involve as few bonds as possible to exert the strongest influence on the stability of a compound. However, the number of carbon–carbon bonds and the number of carbon–hydrogen bonds do not differ among these isomers, and one-bond conjugates cannot be used to compare the stabilities. The next set of candidates would involve three bonds:

$$\text{H-C-C-C} \leftrightarrow \text{H}^{+1}\cdots\text{C}{=}\text{C}\cdots\text{C}^{-1} \qquad (I)$$

$$\text{H-C-C-C} \leftrightarrow \text{H}^{-1}\cdots\text{C}{=}\text{C}\cdots\text{C}^{+1} \qquad (II)$$

$$\text{C-C-C-C} \leftrightarrow \text{C}^{+1}\cdots\text{C}{=}\text{C}\cdots\text{C}^{-1} \qquad (III)$$

$$\text{H-C-C-H} \leftrightarrow \text{H}^{+1}\cdots\text{C}{=}\text{C}\cdots\text{H}^{-1} \qquad (IV)$$

Note that by omitting diradical conjugates we are indirectly neglecting the effects of spin and spin–spin interactions, focusing our attention on the behavior of electron pairs. If carbon and hydrogen have different electronegativities, the conjugates of type (I) or type (II) should be the most relevant. Specifically, if carbon is less electronegative than hydrogen, which is the currently accepted view,[9] then conjugate (II) is

Table II. Analysis of Effect of HCCC Conjugates on Enthalpies of Formation of Three Isomeric Methylheptanes

compound	enthalpy of formation (ref 1) (kJ/mol)	no. of HCCC conjugates
2-methylheptane	−215.6	32
3-methylheptane	−212.8	30
4-methylheptane	−212.2	30
n-octane[a]	−208.6	26

[a] For comparison.

Table III. Analysis of Effect of Number of Side Chains and Number of Conjugates on Enthalpies of Formation of Isomeric Pentenes

ΔH_f (kJ/mol)	compound	side chains	HCC=C conjugates	HCCC conjugates
−20.93	1-pentene	1	2	8
−28.09	cis-2-pentene	2	5	3
−31.78	trans-2-pentene	2	5	3
−36.34	2-methyl-1-butene	2	5	7
−42.58	2-methyl-2-butene	3	9	6

more important. If, on the other hand, carbon is more electronegative than hydrogen, as accepted within some approaches,[10] then conjugate (I) is more important. This issue is not important here, because, for neutral compounds, the number of conjugates of the two types is the same. These will be collectively called HCCC conjugates. Each distinct sub-chain of one hydrogen followed by three carbons, H-C-C-C, gives a distinct conjugate, even if the hydrogen or some of the carbons are also participating in other such subchains. There are 14 HCCC conjugates for *n*-pentane, 20 for 2-methylbutane, and 36 for 2,2-dimethylpropane. In agreement with the provided enthalpies of formation, the higher number of conjugates has a stabilizing effect. As an example in which group contributions are not able to distinguish among the isomers, consider the three isomeric methylheptanes, for which the enthalpies of formation and numbers of HCCC conjugates are shown in Table II. While group contribution methods would predict the same enthalpy of formation for the three methylheptanes, inspection of the number of HCCC conjugates (Table II) shows that, in agreement with the data, 2-methylheptane is more stable and the other two are roughly equivalent.

A similar empirical rule states that the stability of alkenes increases with their degree of substitution, i.e., the number of side chains attached to the double-bonded carbons.[8] Consider the set of isomeric pentenes with their enthalpies of formation and the number of side chains on double-bonded carbons, shown in Table III. The empirical rule cannot distinguish between 2-methyl-1-butene and 2-pentene. The HCCC conjugates discussed for alkanes are relevant for this example, but they take a secondary role because of the important class of HCC=C conjugates which involve fewer bond disruptions:

$$H\text{-}C\text{-}C\text{=}C \leftrightarrow H^{-1}\cdots C\text{=}C\text{-}C^{+1}$$

The numbers of HCC=C conjugates (Table III) capture the same ordering as the original empirical rule; with the additional evidence from HCCC conjugates it is also possible to distinguish between 2-methyl-1-butene and 2-pentene. However, the cis and trans versions of 2-pentene remain indistinguishable; this limitation is inherent in the fact that the analysis performed here covers only electronic effects within the existing skeleton of the molecule but ignores steric effects.

ESTIMATION OF THE ENTHALPY OF FORMATION

Note that 2-methylpentane has two conjugates more than 3-methylheptane and is roughly 3 kJ/mol more stable; 3-methylheptane has four conjugates more than *n*-octane and

is, almost proportionately, 7 kJ/mol more stable. This suggests that utilization of the empirical rules in a quantitative framework may be feasible. In the estimation of the ideal-gas enthalpy of formation, intermolecular interactions are not taken into account. The energy of a compound is influenced primarily by the electron clouds of the different bonds and their interactions. The ABC approach proposed here is based on using contributions of Atoms and Bonds within Conjugate forms of a compound.[11] In other words, the principle of additivity is applied at the level of conjugate forms (rather than the actual compound); furthermore, instead of using contributions of whole groups, the proposed approach entails contributions only from individual atoms and bonds. The goal of the approach is to capture interactions among bonds and atoms in the numbers and kinds of conjugate forms of the compound, rather than in the kinds of functional groups comprising the compound.

Let E be the enthalpy of formation of the (hybrid) compound and E_i be the energy (enthalpy of formation) of conjugate i, obtained by summing contributions from atoms and bonds. The preceding comments suggest that E is a well-defined function of the E_i's, i.e., $E = E(E_1, E_2, ..., E_{n-1}, E_n)$. There are several desirable qualitative properties for the function E. For example, the hybrid must be more stable than each individual conjugate; if a conjugate is extremely unstable, then it should not affect the hybrid at all; if a conjugate is extremely stable, it alone should determine the energy of the hybrid; and the function should be invariant with respect to the choice of reference states. For the demonstrations in this paper, a qualitative plausible function that satisfies these requirements is

$$E = -A \ln \left[\sum_i \exp(-E_i/A)\right] \tag{1}$$

where A is any *positive* constant. A simple scheme utilizing this combination function in the estimation of the enthalpies of formation of alkanes will be demonstrated here. For a class as restricted as alkanes, it is not possible to determine contributions directly because the contributions are not independent. Thus, the enthalpy of formation of the dominant conjugate is simply taken to be a linear function in n, the number of carbons:

$$E_i(\text{dominant conjugate}) = bn + d \tag{2}$$

The enthalpy of formation of an HCCC conjugate differs from the dominant conjugate by a constant term independent of the number of carbons:

$$E_i(\text{HCCC conjugate}) = bn + d + h \tag{3}$$

The energy of each conjugate depends only on the numbers of bonds and not on more detailed structural features; the different types of conjugates should nonetheless permit accurate estimation of the enthalpy of formation of the compounds. If the number of HCCC conjugates is N, the enthalpy of formation of an alkane is given by

$$E = -A \ln \{\exp[-(bn + d)/A) + N \exp[-(bn + d + h)/A)]\} \tag{4}$$

which can be transformed into

$$E = bn + d - A \ln [1 + N \exp(-h/A)] \tag{5}$$

where n and N are known for each compound, and b, d, h, and A are unknown constants, independent of the compound. Using the enthalpies of formation of 51 alkanes with 2–10 carbons from ref 1, a regression was performed (Table IV) and the adjustable parameters were estimated[11] (in kJ/mol) as $b = -16.6$, $d = -47.8$, $h = 146.2$, and $A = 43.9$. The standard deviation of the regression was ±2.8 kJ/mol. For comparison, a group contribution fit was performed by using four groups whose contributions (in kJ/mol) were estimated

Table IV. Estimation of Enthalpies of Formation of Saturated Hydrocarbons[a]

compound	no. of carbons	HCCC conjugates	ethahlpy of formation	estimation error	
				proposed method	group contributions
enthane	2	0	−84.7	3.7	−5.7
propane	3	6	−103.9	−2.2	−6.4
n-butane	4	10	−126.2	−1.4	−3.9
isobutane	4	18	−134.6	−1.4	1.1
n-pentane	5	14	−146.5	−2.1	−3.5
2-methylbutane	5	20	−154.6	0.2	1.3
2,2-dimethylpropane	5	36	−166.1	−1.0	7.5
n-hexane	6	18	−167.3	−1.9	−2.6
2-methylpentane	6	24	−174.4	−0.1	1.5
3-methylpentane	6	22	−171.7	−1.1	−1.5
2,2-dimethylbutane	6	36	−185.7	2.1	7.2
2,3-dimethylbutane	6	28	−177.9	0.1	1.4
n-heptane	7	22	−187.9	−1.5	−1.8
2-methylhexane	7	28	−195.1	0.7	2.0
3-methylhexane	7	26	−192.4	−0.4	−0.7
2,2-dimethylpentane	7	40	−206.3	3.4	8.0
2,3-dimethylpentane	7	30	−199.4	3.5	3.0
2,4-dimethylpentane	7	34	−202.1	3.2	5.7
3,3-dimethylpentane	7	36	−201.7	1.5	5.3
3-ethylpentane	7	24	−189.8	−1.3	−3.3
2,2,3-trimethylbutane	7	42	−204.9	0.7	3.2
n-octane	8	26	−208.6	−0.8	−1.0
2-methylheptane	8	32	−215.6	1.6	2.7
3-methylheptane	8	30	−212.8	0.3	−0.1
4-methylheptane	8	30	−212.2	−0.3	−0.7
2,2-dimethylhexane	8	44	−224.9	2.9	6.7
2,3-dimethylhexane	8	34	−214.1	−1.4	−2.1
2,4-dimethylhexane	8	36	−219.6	2.8	3.4
2,5-dimethylhexane	8	38	−222.8	4.6	6.6
3,3-dimethylhexane	8	40	−220.3	0.8	2.1
3,4-dimethylhexane	8	32	−213.1	−0.9	−3.1
3-ethylhexane	8	28	−211	0.0	−1.9
2,2,3-trimethylpentane	8	44	−220.3	−1.7	−1.2
2,24-trimethylpentane	8	50	−224.3	−1.2	2.8
2,3,3-trimethylpentane	8	42	−216.6	−4.2	−4.9
2,3,4-trimethylpentane	8	38	−217.6	−0.6	−2.0
2-methyl-3-ethylpentane	8	32	−211.3	−2.7	−4.9
3-methyl-3-ethylpentane	8	36	−215.1	−1.7	−3.1
n-nonane	9	30	−229.2	0.1	−0.3
2-methyloctane	9	30	−229.2	0.1	−3.6
2,2-dimethylheptane	9	48	−247	6.0	8.9
2,2,3-trimethylhexane	9	48	−241.4	0.4	0.0
2,2,4-trimethylhexane	9	50	−243.4	1.3	2.0
2,2,5-trimethylhexane	9	54	−254.2	9.9	12.8
3,3-diethylpentane	9	36	−232.1	−1.3	−6.0
2,2,3,3-tetramethylpentane	9	54	−237.4	−6.9	−9.3
2,2,3,4-tetramethylpentane	9	47	−237.1	−3.3	−7.6
2,2,4,4-tetramethylpentane	9	66	−242.1	−8.2	−4.6
2,3,3,4-tetrmethylpentane	9	42	−236.4	−1.0	−8.3
n-decane	10	34	−249.8	1.1	0.5
3,3,5-trimethylheptane	10	50	−258.7	0.0	−2.5

[a] The data were taken from ref 1.

as −45.21 for CH_3-, −19.86 for -CH_2-, 2.15 for >CH-, and 22.22 for >C<. In the group contribution fit, whose residuals are given in Table IV, the standard deviation was ±4.7 kJ/mol.

With the same number of parameters (four), the conjugation approach achieved a standard deviation of 2.8 kJ/mol, compared to 4.7 kJ/mol for the group contribution approach. The method's performance approaches the accuracy of experimental measurements, which ranges from 1 to 2 kJ/mol.[2]

DISCUSSION

The objective of the work presented here is the development of a new computer-based approach for the estimation of properties of organic compounds from their molecular structure. The ABC approach[11] under development is based on the contributions of atoms and bonds in the properties of conjugate forms of a molecular structure. Conjugate forms, which are alternative formal arrangements of valence electrons in a molecule, give rise to empirical rules explaining the relative stability of similar compounds. The proposed method must be computer-based because the generation and analysis of a fairly large number of conjugates cannot be done manually.

One nontrivial computational aspect of the method is the generation of the conjugates. It can be accomplished through use of a set of operators, each of which constructs at most one bond or dissociates at most one bond. The basic bond-forming and bond-dissociating operators are denoted as Q_i and Q_{-i}. The operator Q_i constructs a bond of order i from a bond of order $i-1$, between adjacent atoms; the operator Q_{-i} dissociates (heterolytically) a bond of order i into a bond of order $i-1$. Naturally, each operator is only applicable under certain conditions. As an example, Figure 1 shows the formation of conjugates through operators.
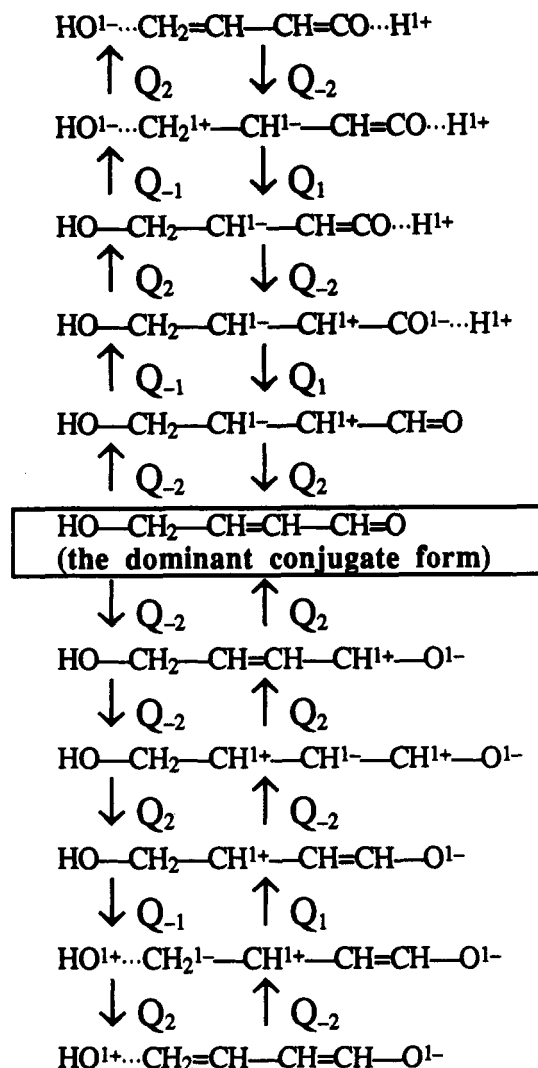
ACKNOWLEDGMENT

Figure 1. Formation of conjugates through bond-forming operators, Q_i, and bond-dissociating operators, Q_{-i}.

neering Research Center of the University of Maryland; by the Systems Research Center funded by the National Science Foundation (NSFD CDR 8803012); and by a Research Initiation Award from the National Science Foundation (CTS 9010549).

REFERENCES AND NOTES

(1) Reid, R. C.; Prausnitz, J. M.; Poling, B. E. *The Properties of Gases and Liquids*, 4th ed.; McGraw-Hill: New York, 1987.

(2) Cox, J. D.; Pilcher, G. *Thermochemistry of Organic and Organometallic Compounds*; Academic Press: New York, 1970.

(3) Joback, K. G.; Reid, R. C. Estimation of Pure-Component Properties from Group Contributions. *Chem. Eng. Commun.* **1987**, *57*, 233–243.

(4) Benson, S. W. *Thermochemical Kinetics*; Wiley: New York, 1968.

(5) Domalski, E. S.; Hearing, E. D. Estimation of the Thermodynamic Properties of Hydrocarbons at 298.15 K. *J. Phys. Chem. Ref. Data* **1988**, *14*, 1637–1678.

(6) Mavrovouniotis, M. L.; Bayol, P.; Lam, T.-K. M.; Stephanopoulos, G.; Stephanopoulos, G. A Group-Containing Method for the Estimation of Equilibrium Constants of Biochemical Reactions. *Biotechnol. Tech.* **1988**, *2*, 23–28.

(7) Mavrovouniotis, M. L. Group Contributions to the Gibbs Energy of Formation of Biochemical Compounds in Aqueous Solution. *Biotechnol. Bioeng.* **1990**, *36*, 1070–1082.

(8) Morrison, R. T.; Boyd, R. N. *Organic Chemistry*, 3rd ed.; Allyn and Bacon: Boston, 1973.

(9) Fliszar, S. *Charge Distributions and Chemical Effects*; Springer-Verlag: New York, 1983.

(10) Mullay, J. A Simple Method for Calculating Atomic Charges in Molecules. *J. Am. Chem. Soc.* **1986**, *108*, 1770–1775.

(11) Mavrovouniotis, M. L. Estimation of Properties from Conjugate Forms of Molecular Structures: The ABC Approach. *Ind. Eng. Chem. Res.* **1990**, *29*, 1943–1953.

On Using the Adjacency Matrix Power Method for Perception of Symmetry and for Isomorphism Testing of Highly Intricate Graphs

GERTA RÜCKER and CHRISTOPH RÜCKER*

Institut für Organische Chemie und Biochemie, Universität Freiburg, Albertstrasse 21, D-7800 Freiburg, FRG

A modification of the adjacency matrix power method described recently for the perception of symmetry in graphs is introduced, which expands the limits of the method far beyond the realm of chemically interesting graphs. The procedure finds the automorphism partition even for intricate graphs without performing a tree search. The calculation effort increases with the problem size polynomially for all tested cases, including strongly regular graphs, two-level regular graphs, and graphs corresponding to balanced incomplete block designs (BIBD). An equally powerful computer program for testing isomorphism of graphs based on the adjacency matrix power method is introduced.

Recently we described a computer program (TOPSYM) for the perception of constitutional (topological) symmetry in

molecular graphs which is based on a very elementary principle, raising the adjacency matrix to its higher powers,