Revue Roumaine de Chimie, **33**, *8*, 839–845(1988)

TOPOLOGICAL AND EMPIRICAL MODELS. I

THE PREDICTION OF THE GIBBS ENERGIES OF FORMATION FOR ALKANES

OVIDIU IVANCIUC
Polytechnic Institute "Traian Vuia", Faculty of Chemical Technology, Timişoara, Romania

Received May 16, 1987

The prediction ability for the standard Gibbs energies of formation for alkanes by using topological and empirical models is tested. A number of forty topological indices were used in correlations. Best results are obtained with equations using as parameters $^{1}\chi$, $^{2}\chi$, $^{3}\chi_P$, $^{4}\chi_P$, $^{4}\chi_{PC}$, $^{5}\chi_C$, $^{5}\chi_P$, W, I_{orb}, MCI, DRGI and RGI. Through the equations obtained the standard Gibbs energies of formation for alkanes are demonstrated to bear a strong relationship to topological characteristics.

1. INTRODUCTION

Accurate data about the fundamental physical and chemical properties of organic compounds are strongly needed to design and operate a large variety of technological installations. An answer to this inquiry is the existence of large compilations containing information on various physical and chemical properties of chemical compounds, under various conditions,[1] but they are of no great use when one searches for the properties of a new compound or a compound whose properties are not determined. For example, the number of alkane constitutional isomers increases very rapidly with the number of carbon atoms.[2] We cannot expect that the thermodynamic properties of so many chemical species can be determined in the laboratory. So it is practical to obtain relationships between the structure of chemical compounds and their thermodynamic properties in order to compute the unknown thermodynamic properties of isomers with higher number of carbon atoms. To solve this situation, physico-chemical properties were established, beginning with the work done by Kopp.[3] Usually, they are group additivity methods, ranging from those that consider as parameters the type of the chemical bonds to the more complex additivity functions of Smolenskii [4] and Tatevskii. [5]

Another method used in establishing correlations between structure and physico-chemical properties is to express the bonding topology of the molecular graph by a mathematical expression, which may be a matrix,

* Correspondence address : strada N. Grigorescu, 18/2, 2750 Hunedoara, Romania

a polynomial or a numerical index. Numerical indices developed in this way are called topological indices (TI).[6]

The aim of the present paper is to establish equations for the prediction of the Gibbs free energy of formation (gas) of alkane using topological indices, and to make a comparison with the results obtained using various empirical additive schemes.

2. TOPOLOGICAL INDICES, ADDITIVE SCHEMES AND DATA USED

In the following, only the notation will be indicated for the forty topological indices used in correlations:

—The connectivity indices [7-10] ${}^0\chi$, ${}^1\chi$, ${}^3\chi_P$, ${}^3\chi_C$, ${}^4\chi_P$, ${}^4\chi_C$, ${}^4\chi_{PC}$, ${}^5\chi_P$, ${}^5\chi_C$, ${}^5\chi_{PC}$, ${}^6\chi_P$, ${}^6\chi_C$, ${}^6\chi_{PC}$

— The centric and quadratic indices [11] B, C, C′, Q, Q′

B — the centric index

C — the normalized centric index

C′ — the binormalized centric index

Q — the normalized quadratic index

Q′ — the binormalized quadratic index

— The Gordon-Scantlebury index [12] N_2

— The Zagreb group index [13] M_1

— The Wiener index [14,15] W

— The Hosoya index [16] Z

— The Lovasz and Pelikan index [17] X_1

— Information indices on the graph distances [18] I_D^e, I_D^w, $\bar{I}_D^e$, $\bar{I}_D^w$

I_D^e — the information content on the distribution of distances

$\bar{I}_D^e$ — the mean information content on the distribution of distances

I_D^w — the information content on realized distances

$\bar{I}_D^w$ — the mean information content on realized distances

— Informational indices on the polynomial coefficients [18] I_{PC}, $\bar{I}_{PC}$:

I_{PC} — the information content on polynomial coefficients

$\bar{I}_{PC}$ — the mean information content on polynomial coefficients

— The graph orbits information index [19,20] $\bar{I}_{orb}$

— The chromatic information index [21] $\bar{I}_{chr}$

— Information analogous with the Randic index [22] $\bar{I}_\chi$

— The electropy index [23] ε

— Topological indices based on the generalized graph centre concept [24] RCI, DRCI, MCI, RGI, DRGI

— The informational centric index [24] $\bar{I}_C$

Also used in correlations was the number N of carbon atoms.

The source of data for the standard Gibbs energy of formation for alkane at 298 K was the Scott table.[1] Data on 37 isomers of alkanes between butane and octane were used in correlations.

An efficient way of equilibrium calculations for the ideal gases can be carried using standard thermodynamic properties of isomer groups, determined for alkanes in [25]. We have used their results concerning standard Gibbs energy of formation for alkane isomer groups in order to compare with the predictive ability of Ti's. With the same intention, we have used the following empirical additive schemes: the method of Tatevskii,[5]

with nine parameters, the scheme of Seyfer and Smolenskii,[26] with twenty parameters, the method of van Krevelen,[27] with eight parameters and the method of Thisch [28,29] with ten parameters.

The statistical parameters for estimating the accuracy of correlations are the correlation coefficient R, the standard deviation s and the Fisher statistics F.

3. CORRELATIONAL ABILITY OF THE TOPOLOGICAL INDICES

Table 1 lists the regression coefficients as well as the statistical parameters obtained with equation

$$\Delta G^0_{f,\,298} = A_0 + A_1 P \qquad (1)$$

where the property P represents the standard Gibbs energy of formation calculated with an additive scheme. However, we have to note that the additive schemes are intended to give a direct estimation of the physicochemical property under consideration, not through the agency of a correlation equation. We used this way in order to get a set of statistical parameters to compare with those obtained through equations (3), (5), (7) and (9). A good additive scheme should have A_0 close to 0 and A_1

Table 1

Regression coefficients and statistical parameters for equations of type (1)

Method	A_0	A_1	R	s	F
Tatevskii	0.478	1.01	0.987	1.87	1370.1
van Krevelen	2.19	1.05	0.968	2.97	523.8
isomer groups	7.82	1.24	0.953	3.60	345.9
Seyfer	−0.339	0.774	0.946	3.86	295.6
Thisch	3.98	1.17	0.880	5.63	120.5

close to 1. The method of Tatevskii satisfies these requirements, being the best one among the additive schemes used in this study:

$$\Delta G^0_{f,\,298} = 0.478\ (0.711) + 1.01(0.06)G \qquad (2)$$

where G represents the standard Gibbs energy of formation computed with the method of Tatevskii and in parenthesis there are the 95% confidence intervals of the coefficients.

Table 2 gives the best seven equations of type

$$\Delta G^0_{f,\,298} = A_0 + A_1 TI \qquad (3)$$

Table 2

Regression coefficients and statistical parameters for the best equations of type (3)

TI	A_0	A_1	R	s	F
N	−58.0	9.186	0.952	3.63	339.7
$\bar{I}_D^W$	−68.4	17.87	0.941	4.01	271.1
$^0\chi$	−60.9	11.54	0.937	4.15	251.2
$^1\chi$	−53.7	18.82	0.932	4.29	232.0
ε	−35.6	0.235	0.924	4.53	204.6
I_D^W	−15.9	0.098	0.908	4.98	163.7
W	−19.5	0.503	0.903	5.09	154.7

The best set of statistical parameters is obtained by the use of the carbon number N, but it is not a satisfactory equation :

$$\Delta G^0_{f,\,298} = -58.0(7.25) + 9.19(1.01)N \qquad (4)$$

In order to obtain better estimations, the set of 820 biparametric equations of type (5) was computed and the best ten equations are presented in table 3.

$$\Delta G^0_{f,\,298} = A_0 + A_1TI1 + A_2TI2 \qquad (5)$$

Table 3

Regression coefficients and statistical parameters for the best ten equations of type (5)

TI1	TI2	A_0	A_1	A_2	R	s	F
$^1\chi$	$^4\chi_{PC}$	−52.27	7.27	3.79	0.981	2.35	431.3
N	$^3\chi_P$	−49.03	6.55	6.61	0.976	2.64	336.3
$^3\chi_P$	$\bar{I}_D^W$	−21.09	9.36	0.062	0.975	2.66	332.4
$^2\chi$	M_1	−33.79	15.96	3.23	0.975	2.66	331.4
$^3\chi_P$	W	−23.47	9.54	0.315	0.975	2.67	329.1
$^1\chi$	$^5\chi_C$	−53.93	18.49	7.94	0.973	2.77	304.6
$^3\chi_C$	M_1	−40.91	−9.41	2.10	0.973	2.77	303.3
$^1\chi$	$\bar{I}_C$	−51.04	24.46	−13.97	0.973	2.78	301.0
$^1\chi$	RGI	−54.40	15.40	0.595	0.971	2.86	284.5
$^1\chi$	$^3\chi_P$	−44.88	12.51	7.89	0.971	2.90	277.2

Thus, even at this level of biparametric equation the topological method appears to give a reasonable basis for describing the effects of

molecular structure on the standard Gibbs energy of formation. The best biparametric equation is obtained with two connectivity indices :

$$\Delta G^0_{f,298} = -52.34(13.48) + 17.27(2.39)^1\chi + 3.79(0.53)^4\chi_{PC} \quad (6)$$

To test further the usefulness and validity of the topological model, the set of 10,660 three parameters correlation equations of type (7) was computed and the best ten of them were selected and presented in table 4.

$$\Delta G^0_{f,298} = A_0 + A_1TI1 + A_2TI2 + A_3TI3 \quad (7)$$

Table 4

Regression coefficients and statistical parameters for the best ten equations of type (7)

TI1	TI2	TI3	A_0	A_1	A_2	A_3	R	s	F
$^2\chi$	M_1	MCI	−28.97	−14.00	2.497	0.036	0.989	1.83	479.9
$^1\chi$	$^3\chi_P$	MCI	−38.61	8.69	6.21	0.033	0.988	1.91	441.0
$^1\chi$	$^4\chi_{PC}$	$\bar{I}_{orb}$	−54.24	19.73	3.45	−2.88	0.988	1.92	435.3
$^3\chi_P$	$^4\chi_P$	MCI	−22.36	9.46	8.39	0.042	0.988	1.92	433.7
$^1\chi$	$^5\chi_C$	$^5\chi_{PC}$	−50.08	16.57	5.95	3.89	0.987	1.99	400.4
$^1\chi$	$^5\chi_P$	$^5\chi_C$	−59.36	20.82	−11.72	6.26	0.986	2.03	389.7
$^3\chi_P$	W	MCI	−23.49	7.94	0.223	0.028	0.985	2.08	369.4
$^0\chi$	$^2\chi$	RGI	−61.83	14.33	−11.15	0.784	0.985	2.08	368.9
$^1\chi$	$^4\chi_{PC}$	MCI	−47.71	14.49	2.94	0.020	0.985	2.09	366.0
$^1\chi$	$^4\chi_P$	$^4\chi_{PC}$	−45.39	13.72	6.06	4.76	0.985	2.09	363.7

Comparing the two important statistical parameters, correlation coefficient and standard deviation, of equation (2) with those of equation (8), we conclude that the topological model has a greater predictive ability.

$$\Delta G^0_{f,298} = -28.97(7.28) - 14.2(2.22)^2\chi + 2.497(0.396)M_1 + 0.036(0.006)MCI \quad (8)$$

Taking into account the great number of equations with four parameters which can be formed by combinations of 41 parameters, we computed only a limited set of equations of type (9) from which a selection is presented in table 5. Of course, due to the limited search, they are not the best equations of type (9).

$$\Delta G^0_{f,298} = A_0 + A_1TI1 + A_2TI2 + A_3TI3 + A_4TI4 \quad (9)$$

Use of two connectivity indices, the Zagreb group index M_1, and a topological index based on the generalized graph centre concept, MCI, provides the basis for an excellent correlation :

$$\Delta G^0_{f,298} = -28.90(7.64) - 13.11(2.58)^2\chi + 1.76(0.35)^4\chi_P +$$

$$+2.37(0.47)M_1 + 0.034(0.007)MCI \quad (10$$

Table 5

Regression coefficients and statistical parameters for a selection of equations of type (9)

TI1	TI2	TI3	TI4	A_0	A_1	A_2	A_3	A_4	R	s	F
$^2\chi$	$^4\chi_P$	M_1	MCI	−28.90	−13.11	1.76	2.37	0.034	0.990	1.78	380.5
$^1\chi$	$^3\chi_P$	$^5\chi_C$	$^5\chi_{PC}$	−47.48	14.65	2.99	4.68	3.10	0.990	1.79	376.4
$^1\chi$	$^3\chi_P$	$^4\chi_{PC}$	I_{orb}	−51.41	17.97	3.04	2.36	−3.16	0.989	1.81	369.5
$^1\chi$	$^4\chi_P$	$^4\chi_{PC}$	MCI	−42.18	11.71	5.32	3.89	0.018	0.989	1.87	344.2
$^1\chi$	$^3\chi_P$	$^4\chi_{PC}$	MCI	−41.64	10.58	4.14	1.09	0.028	0.989	1.88	342.6
$^1\chi$	$^3\chi_P$	I_D^e	MCI	−41.89	10.59	5.84	−0.054	0.034	0.988	1.93	323.2
$^1\chi$	$^3\chi_P$	W	MCI	−37.46	8.02	6.33	0.019	0.033	0.988	1.94	321.2
$^3\chi_P$	W	I_D^w	MCI	−25.24	7.98	0.447	−0.045	0.030	0.986	2.09	274.9

Statistical evaluation of the computed equations concerning the estimation of standard Gibbs energies of formation for alkanes indicates that the topological models give best correlations and are using fewer parameters than the additive schemes.

Best results are obtained with equations which contain as parameters combinations of the following topological indices : the connectivity indices $^1\chi$, $^2\chi$, $^3\chi_P$, $^4\chi_P$, $^4\chi_{PC}$, $^6\chi_C$, $^6\chi_{PC}$, the Wiener index W, the graph orbits information index I_{orb}, and the topological indices based on the generalized graph centre concept MCI, DRGI and RGI.

Through the equations obtained, the standard Gibbs energies of formation of alkanes is demonstrated to bear a strong relationship to topological characteristics.

REFERENCES

[1] D. W. Scott, "Chemical Thermodynamic Properties of Hydrocarbons and Related Substances : Properties of the Alkane Hydrocarbons, C_1 through C_{10} in the Ideal Gas State from 0 to 1500 K", U. S. Department of the Interior, Bureau of Mines Bulletin 666, 1974.

[2] J. V. Knop, W. R. Muller, K. Szymanski and N. Trinajstić, "Computer Generation of Certain Classes of Molecules", Kemija i Industrii, Zagreb 1985, p. 38.

[3] H. Kopp, *Ann. Chem.*, **41**, 79, 169 (1842).

[4] E. A. Smolenskii, *Zhur. Fiz. Khim.*, **38**, 1288 (1964).

[5] V. M. Tatevskii, "Metody rascheta fiziko-khimicheskikh svoystv uglevodorodov", Toptekhizd., Moscou, 1960.

[6] A. T. Balaban, I. Motoc, D. Bonchev and O. Mekenyan, *Top. Curr. Chem.*, **114**, 21 (1983), Eds. M. Charton and I. Motoc, Springer, Berlin.

[7] M. Randic, *J. Amer. Chem. Soc.*, **97**, 6609 (1975).

[8] L. B. Kier, L. H. Hall, W. J. Murray and M. Randić, *J. Pharm. Sci.*, **64**, 1971 (1975).

[9] L. B. Kier, L. H. Hall, W. J. Murray and M. Randić, *J. Pharm. Sci.*, **65**, 1226 (1976).

[10] L. B. Kier and L. H. Hall, "Molecular Connectivity in Chemistry and Drug Research", Academic, New York, 1976.

[11] A. T. Balaban, *Theor. Chim. Acta*, **53**, 355 (1979).

[12] M. Gordon and G. R. Scantlebury, *Trans. Faraday Soc.*, **60**, 605 (1964).

[13] I. Gutman and N. Trinajstić, *Chem. Phys. Lett.*, **17**, 535 (1972).

[14] H. Wiener, *J. Amer. Chem. Soc.*, **69**, 17, 2636 (1947).

[15] H. Wiener, *J. Chem. Phys.*, **15**, 766 (1947).

[16] H. Hosoya, *Bull. Chem. Soc. Japan*, **44**, 2332 (1971).

[17] L. Lovasz and J. Pelikan, *Period. Math. Hung.*, **3**, 175 (1973).
[18] D. Bonchev and N. Trinajstić, *J. Chem. Phys.*, **67**, 4517 (1977).
[19] N. Rashevsky, *Bull. Math. Biophys.*, **30**, 229 (1955).
[20] E. Trucco, *Bull. Math. Biophys.*, **18**, 129 (1956).
[21] A. Mowshowits, *Bull. Math. Biophys.*, **30**, 175, 225, 387, 533 (1968).
[22] S. H. Bertz, *J. Amer. Chem. Soc.*, **103**, 3599 (1981).
[23] W. T. Yee, K. Sakamoto and Y. J. L'Haya, *Rept. Univ. Electro-commun.*, **27**, 53 (1976).
[24] D. Bonchev, A. T. Balaban and U. Mekenyan, *J. Chem. Inf. Comput. Sci.*, **20**, 106 (1980).
[25] R. A. Alberty and C. A. Gehrig, *J. Phys. Chem. Ref. Data*, **13**, 1173 (1984).
[27] A. L. Seyfer and E. A. Smolenskii, *Zhur. Fiz. Khim.*, **38**, 2230 (1964).
[26] D. van Krevelen and H. Chermin, *Chem. Eng. Sci.*, **1**, 238 (1952).
[28] T. P. Thisch and T. K. Trong, *Canad. J. Chem. Eng.*, **54**, 344 (1976).
[29] T. P. Thisch, I. L. Duran and R. S. Remalho, *Ind. Eng. Chem.-Proc., Res. Dev.*, **10**, 576 (1971).