QSPR MODELING OF GIBBS FREE ENERGY OF CHEMICAL TRANSFORMATIONS OF OIL SHALE COMPONENTS DURING THERMAL TREATMENT

Castro[a)], E.A.; Toropova[b)], A.P.; Toropov[b)], A.A.; Mukhamedjanova[b)], D.V.
[a)] INIFTA, Chemistry Department, Faculty of Exact Sciences, La Plata University, Suc. 4, C.C. 16, 1900 La Plata, Argentina
e-mail: castro@dalton.quimica.unlp.edu.ar / jubert@arnet.com.ar
[b)] Institute of Geology and Geophysics, Republic of Uzbekistan Academy of Sciences, Tashkent, Uzbekistan

Received July 16, 2003. In final form August 28, 2003

Abstract

We examine the encoding of chemical transformations by the systems of the"initial state graphs" and the "final state graphs". By means of the optimization of correlation weights of local invariants in these graph systems, we propose a QSPR model of Gibbs free energy of chemical transformations of oil shale components during thermal treatment. Statistical characteristics of the model are the following: n = 12, r = 0.95, s = 47 kJ/mol, F = 106 (Training Set); n = 8, r = 0.96, s = 40 kJ/mol, F = 78 (Test Set)

Resumen

Se estudia la codificación de transformaciones químicas por el sistema de "gráficos de estado inicial" y "gráficos de estado final".Por medio de la optimización de los pesos de correlación de los invariantes locales, en estos sistemas de gráficos, se propone un modelo QSPR para la energía libre de Gibas de la transformación química de los componentes del aceite de destilado de petróleo. Las características estadísticas del modelo son las siguientes: n = 12, r = 0.95, s = 47 kJ/mol, F = 106 (Training Set); n = 8, r = 0.96, s = 40 kJ/mol, F = 78 (Test Set)

Introduction

Thermal treatment of oil shale components is involving several ecological and technological problems [1]. One of the ways to study ecological and technological aspects of oil shale industry is the QSPR modeling [2] (QSPR/QSAR = Quantitative Structure-Property/Activity Relationships). The aim of QSPR/QSAR studies is the prediction of numerical values of chemical compounds properties and/or activities from their molecular structure. As a general rule, molecular structure is presented by molecular graphs [3]. The modelling of a physical chemistry property is to correlate a descriptor (or a set of descriptors) calculated with molecular graphs and the values of the property for a set of compounds under consideration.

Recently, the Optimization of Correlation Weights of Local Graph Invariants (OCWLI) has been suggested as a important tool of the QSPR/QSAR analysis [3-7]. Labeled Hydrogen- Filled Graphs (LHFGs) have been used as a base of the OCWLI in Refs. [3-7]. The purpose of the present study is the estimation of the OCWLI ability to

model Gibbs free energy of chemical transformations of oil shale components during termal treatment.

Method

The essence of the OCWLI is the following:

1) it is defined a descriptor calculated with values of correlation weights (CWs) of local graph invariants;
2) resorting to the Monte Carlo optimization method these values of the CWs are calculated in such a way that they produce a fitting equation with a correlation coefficient as large as possible between the descriptor (DCW) and the property/activity (PA) data of interest for the given molecular training set;
3) on the base of the training set data it is computed the following equation:

$$PA = C_0 + C_1\ DCW \tag{1}$$

by the Least Squares statistical method. The predictive capability of Eq. (1) with the given test set may then be validated through the calculation via a test set. In this last case, results are true predictions.

The molecular descriptors employed in the present study for QSPR analysis were calculated as a function of molecular architecture of initial substances and products of chemical transformations of oil shale components during termal treatment. Molecular structures of initial substances and molecular structures of products of the mentioned chemical transformations have been presented by special systems of LHFGs. Figure 1 presents such LHFG system on chemical transformation such as:

$$CaCO_3 + SO_3 \longrightarrow CaSO_4 + CO_2$$

The approach under consideration has been tested with 20 chemical transformations of oil shale components during termal treatment from Ref. [1]. The calculation of the molecular descriptors has been obtained by means of the Monte Carlo method correlation weights [3]. The descriptors are computed as follows

$$DCW(state) = \sum_{\text{all vertexes}} [CW(a_k)\ x\ CW(\delta_k)] \tag{2}$$

where a_k denotes the chemical element which is presented by k-th vertex of the graph; δ_k is the degree of the k-th vertex; $CW(a_k)$ and $CW(\delta_k)$ are the correlation weights of the local graph invariants; the "state" is denoting of inital and final states; initial LHFG is denoting of initial substances and final LHFG is denoting of products of the chemical reaction (see Figure 1). The particular algebraic form chosen in Eq. (2) is arbitrary and, in principle, any other one may be equally valid.

We have used the following mathematical relationship for the descriptor of the chemical transformation

$$DCW = DCW(initial) - DCW(final) \tag{3}$$

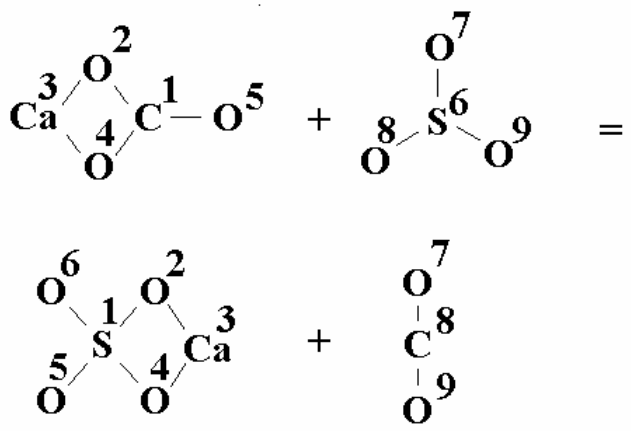

Figure 1*: LHFG system for modeling the chemical transformation: $CaCO_3 + SO_3 = CaSO_4 + CO_2$*

Table 1: *Correlation weights of local graph invariants obtained by Monte Carlo method optimization procedure.*

Invariant	Correlation weights
Atoms (a_k)	CW(a_k)
H	2.25000
C	2.00050
O	1.30953
Mg	4.17192
Si	5.88036
S	0.32528
Ca	8.04469
Fe	0.01978
Vertex degree (δ_k)	CW(δ_k)
1	1.00000
2	1.00781
3	1.00000
4	1.36384

Results and discussion

The model of the Gibbs free energy values were obtained via the standard Least Squares Method and final result is the following:

$$\Delta G = 1519.00\ DCW - 35.00 \quad (4)$$

n = 12, r = 0.95, s = 47, F = 106 (Training Set)

n = 8, r = 0.96, s = 40, F = 78 (Test Set)

Table 2: *Adjacency matrix of the initial LHFG: $CaCO_3 + SO_3$*

Atom		C	O	Ca	O	O	S	O	O	O
C	1	0	1	0	1	1	0	0	0	0
O	2	1	0	1	0	0	0	0	0	0
Ca	3	0	1	0	1	0	0	0	0	0
O	4	1	0	1	0	0	0	0	0	0
O	5	1	0	0	0	0	0	0	0	0
S	6	0	0	0	0	0	0	1	1	1
O	7	0	0	0	0	0	1	0	0	0
O	8	0	0	0	0	0	1	0	0	0
O	9	0	0	0	0	0	1	0	0	0

where n is the number of chemical reactions, r is the correlation coefficient, s is the standard error estimation, and F is the Fischer statistical F-ratio.

Table 1 lists the correlation weights for the computation of the DCW with Eq. (3). Table 2 contains the adjacency matrix corresponding to the initial state from Figure 1. The calculation of the DCW(initial) is shown in Table 3. Table 4 presents the adjacency matrix corresponding to the final state from Figure 1. The computation of the DCW(final) is exemplified in Table 5. From the figures given in these Tables one can verify that

$$DCW = DCW(\text{ initial }) - DCW(\text{final}) = 18.3109 - 18.4449 = -0.1340$$

Table 3*: Calculation of the DCW(initial) for $CaCO_3 + SO_3$*
DCW(initial) = 18.3109

a_k	δ_k	$CW(a_k)$	$CW(\delta_k)$	$CW(a_k) + CW(\delta_k)$
C 1	0003	2.0005	1.0000	2.0005
O 2	0002	1.3095	1.0078	1.3198
Ca3	0002	8.0447	1.0078	8.1075
O4	0002	1.3095	1.0078	1.3198
O5	0001	1.3095	1.0000	1.3095
S 6	0003	0.3253	1.0000	0.3253
O7	0001	1.3095	1.0000	1.3095
O8	0001	1.3095	1.0000	1.3095
O9	0001	1.3095	1.0000	1.3095

Table 6 presents the training set, the test set, the observed (*i.e.* experimental) Gibbs free energy values taken from Ref. [1], and the Gibbs free energies calculated with Eq. (4). We have tried several partitions of the complete molecular set for the training and test sets, but final results do not depend significantly of them.

The LHFG corresponding to the illustrative system (Figure 1) of the initial (Table 2 and Table 3) and final LHFG (Table 4 and Table 5) allows one to extend the

multilinear regression analysis by QSPR modeling of the Gibbs free energies of chemical tranformation under consideration.

***Table 4**: Adjacency matrix of the final LHFG: $CaSO_4 + CO_2$*

Atom		S	O	Ca	O	O	O	O	C	O
S	1	0	1	0	1	1	1	0	0	0
O	2	1	0	1	0	0	0	0	0	0
Ca	3	0	1	0	1	0	0	0	0	0
O	4	1	0	1	0	0	0	0	0	0
O	5	1	0	0	0	0	0	0	0	0
O	6	1	0	0	0	0	0	0	0	0
O	7	0	0	0	0	0	0	0	1	0
C	8	0	0	0	0	0	0	1	0	1
O	9	0	0	0	0	0	0	0	1	0

***Table 5**: Calculation of the DCW(final) for $CaSO_4 + CO_2$*
DCW(final) = 18.4449

a_k	δ_k	$CW(a_k)$	$CW(\delta_k)$	$CW(a_k)xCW(\delta_k)$
S1	0.0004	0.3253	1.3638	0.4436
O2	0.0002	1.3095	1.0078	1.3198
Ca3	0.0002	8.0447	1.0078	8.1075
O4	0.0002	1.3095	1.0078	1.3198
O5	0.0001	1.3095	1.0000	1.3095
O6	0.0001	1.3095	1.0000	1.3095
O7	0.0001	1.3095	1.0000	1.3095
C8	0.0002	2.0005	1.0078	2.0161
O9	0.0001	1.3095	1.0000	1.3095

Conclusions

Statistical characteristics of Gibbs free energy model of the training set are practically the same as those corresponding to the test set. Results for this last set are true predictions so that we deem that the optimization of correlation weights of local graph invariants of the "initial state" and "final state" graph systems may be used as a suitable tool for rather satisfactory predictions of Gibbs free energies.

The great majority of thermodynamic models studied within the realm of QSPR theory comprise different hydrocarbons sets and substituted hydrocarbons, where there is not a large variety of components atoms (i.e. the main components are C and H atoms). In this study the molecular set includes 8 different atoms, which demands an accurate topological description of each one of them.

***Table 6**: Training and test sets of observed and calculated with Eq. (4) Gibbs free energies of chemical transformations of oil shale components during thermal treatment.*

			Gibbs free Energy [kJ/mol]		
ID	Chemical Transformation	DCW	exp.	calc.	exp-calc
	Training set				
1	$CaCO_3 + SO_3 = CaSO_4 + CO_2$	-0.134	-218.9	-238.5	19.6
2	$FeCO_3 + SO_3 = FeSO_4 + CO_2$	-0.134	-177.9	-238.5	60.7
3	$FeCO_3 = FeO + CO_2$	0.005	-137.7	-27.4	-110.3
4	$MgSO_4 = MgO + SO_3$	0.171	232.8	224.7	8.1
5	$FeSO_4 = FeO + SO_3$	0.139	204.3	176.1	28.2
6	$CaSO_4 + SiO_2 = CaSiO_3 + SO_3$	0.164	257.8	214.1	43.7
7	$CaO + SiO_2 = CaSiO_3$	-0.037	-91.6	-91.2	-0.4
8	$FeO + SiO_2 = FeSiO_3$	0.025	-62.1	3.0	-65.1
9	$MgO + SiO_2 = MgSiO_3$	-0.007	-37.9	-45.6	7.7
10	$FeCO_3 + SiO_2 = FeSiO_3 + CO_2$	0.030	25.4	10.6	14.8
11	$CaO + H_2S = CaS + H_2O$	-0.008	-69.3	-47.2	-22.2
12	$FeCO_3 + H_2S = FeS + H_2O + CO_2$	-0.003	-25.1	-39.6	14.4
	Test set				
1	$CaCO_3 = CaO + CO_2$	0.068	130.5	68.3	62.3
2	$CaSO_4 = CaO + SO_3$	0.202	349.3	271.8	77.5
3	$MgSO_4 + SiO_2 = MgSiO_3 + SO_3$	0.164	195.3	214.1	-18.8
4	$FeSO_4 + SiO_2 = FeSiO_3 + SO_3$	0.164	203.4	214.1	-10.7
5	$CaCO_3 + SiO_2 = CaSiO_3 + CO_2$	0.030	39.3	10.6	28.7
6	$MgCO_3 + SiO_2 = MgSiO_3 + CO_2$	0.030	11.0	10.6	0.4
7	$FeO + H_2S = FeS + H_2O$	-0.008	-48.4	-47.2	-1.2
8	$CaCO_3 + H_2S = CaS + H_2O + CO_2$	0.060	66.5	56.1	10.4

Acknowledgment

Authors thank some valuable comments made by the referee which has been useful to improve the final version of this article.

References

[1] Stanotina, S.B.; Reshetov, V.A.; Morkovin, V.V. *Russ. J. Phys. Chem.*, **1999**, *73*, 806
[2] Bonchev, D.; Kier, L.B. *J. Math. Chem.* **1992**, *9*, 75
[3] Toropova, A.P.; Toropov, A.A. *Russ. J. Struct. Chem.* **2001**, *42*, 1230
[4] Toropov, A.A.; Toropova, A.P. *J. Mol. Struct. (Theochem)* **2001**, *538*, 197
[5] Toropov, A.A.; Toropova, A.P. *J. Mol. Struct. (Theochem)* **2002**, *581*, 11
[6] Mercader, A.; Castro, E.A.; Toropov, A.A. *Chem. Phys. Lett*. **2000**, *330*, 612
[7] Duchowicz, P.R.; Castro, E.A.; Toropov, A.A. Comp. Chem. **2002**, *26*, 327