Estimation of activation energies of thermal decomposition of nitro compounds based on structural descriptors

M. I. Kumskov, S. E. Peshkova, L. A. Ponomareva, and K. I. Rezchikova*

N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, 47 Leninsky prosp., 117913 Moscow, Russian Federation. Fax: +7 (095) 135 5328. E-mail: kumskov@cacr.ioc.ac.ru

The structure—property correlation equations for the activation energies of thermal decomposition of nitro compounds in the gaseous phase were obtained based on structural descriptors. The equations were constructed using the BIBIGON computing program system for the data base, which consists of 90 nitro compounds belonging to various chemical classes; the correlation coefficient R is 0.986.

Key words: structural descriptors, molecular fragments, quantitative structure—property relationships, activation energy, gaseous phase, thermal stability, BIBIGON program package.

This work is devoted to the search for the relationships between the structures of *C*-, *N*-, and *O*-nitro compounds and the activation energy (E_a) of thermal decomposition in the gaseous state, which is generally interpreted as the monomolecular cleavage of one of the bonds (C—N, N—N, or O—N). The values of E_a can be obtained experimentally by the manometric method under static conditions at 100—200 °C. The maximum error in determining E_a is 1—2 kcal mol^{-1}.[1–15]

Previously, the values of E_a were also calculated, for example, by quantum-chemical methods[16] or with the use of different topological, informative, and physicochemical indices.[17] In this work, we used a new simpler method for estimating E_a of thermal decomposition of nitro compounds with the use of the quantitative structure—property relationships, which involve only structural descriptors automatically generated using the teaching data base.[18]

Method and Its Applications

In the proposed method, the numbers of repeats of structural fragments (chains of atoms) in the molecule are the parameters used in the construction of the correlation equations. We have implemented this method of specifying structural descriptors of chemical compounds in the BIBIGON program package.[18] This method is convenient because the description of molecules on a computer requires no data other than the structural formulas; these formulas are represented by molecular graphs in the teaching data base.[19]

We have estimated the activation energies of homolytic decomposition of 90 organic nitro compounds, which belong to various chemical classes (aliphatic compounds, aromatic compounds, heterocyclic nitro compounds, nitramines, and nitrates), in the gaseous phase.[1–15,20] The correlation equations were constructed in the following form:

$$E_i = a_0 + \sum_{j=1,k} (a_j \cdot F_j(X_{ij})) + e_i, \quad i = 1, \ldots N, \qquad (1)$$

$$\sum_{i=1,N} (e_i)^2 \rightarrow \min, \qquad (2)$$

where a_0 is the constant; N is the size of the teaching data base; e_i is the error in estimating the property (E_i) of the i-th compound; X_{ij} is the number of repeats of the j-th fragment (a chain of atoms) in the i-th molecular graph; $F_j(X_{ij})$ is the function of the X_{ij} descriptor; a_j are the coefficients determined from condition (2); and k is the number of descriptors involved in the equation.

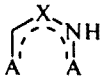
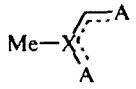
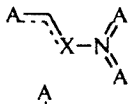
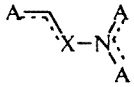
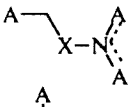
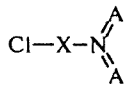
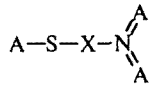
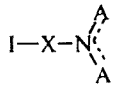
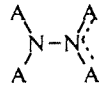
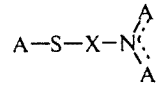
In the BIBIGON program system, prior to generation of descriptors, the structures contained in the data base are preliminary processed: all vertices of molecular graphs are marked in a specific fashion. This procedure makes it possible also to take into account (when the fragments are compared with each other) the particular characteristics such as the number of neighboring atoms, the type of the chemical bond, and the position of the atom in the ring substructure.[18] Hydrogen atoms are not components of chains.

When significant descriptors are chosen for Eq. (1), the BIBIGON system can generate new features by calculating particular functions F (of one variable) of the parameters of descriptors. The number of descriptors analyzed is generally many times greater than the number of compounds involved in the teaching data base, and therefore, Eqs. (1) are constructed by the method of the self-organization of linear models.[18,21] Two ways of representing the fragments are available in the BIBIGON

Translated from *Izvestiya Akademii Nauk. Seriya Khimicheskaya*, No. 8, pp. 1942—1945, August, 1996.

1066-5285/96/4508-1840 $15.00

Table 1. Parameters of Eqs. (1a) and (1b)

Coefficient	Values of the coefficients		Functions		Fragment	
	Equation (1a)	Equation (1b)	Equation (1a)	Equation (1b)	Equation (1a)	Equation (1b)
a_0	19.86	23.22	1	1	Const	Const
a_1	11.87	2.93	Z_{av}	Z	A—Me	
a_2	2.35	36.62	max − Z	Z_{av}		
a_3	0.13	−10.44	Z	Z		
a_4	3.20	217.70	Z	Z_{av}		
a_5	41.87	11.89	Z_{av}	Z_{avn}		
a_6	−0.98	1.17	max − Z	max − Z		
a_7	−11.94	6.21	Z	max − Z		
a_8	248.86	−3.14	Z_{av}	Z		
a_9	14.73	−22.64	Z_{avn}	Z_{av}		
a_{10}	20.06		Z_{av}		A—F	
a_{11}	−1.33		Z			
a_{12}	6.74		max − Z			
a_{13}	−2.54		Z			
a_{14}	18.82		Z_{av}			
a_{15}	−21.35		Z_{av}			

Note. X is an arbitrary atom corresponding to the "empty" code of the descriptor; A is an arbitrary atom involved in the formation of the d marker; --- is an aromatic bond; Z is the value of the Z descriptor without its functional transformation; Z_{av} is the average value of the Z descriptor in the graph with respect to all chains with the given length; max − Z is the difference between the maximum value of the descriptor in the treaching data base and its current value; Z_{avn} is the average value of the max − Z difference in the graph for all chains with the given length (the length of the chain is equal to the number of edges).

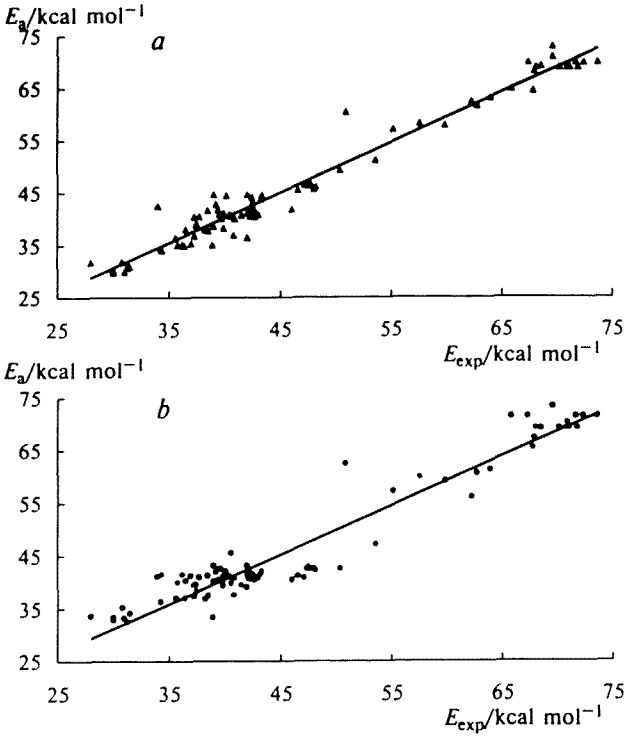


Fig. 1. Diagrams of deviations of the estimated activation energies (E_a/kcal mol^{-1}) of thermal homolytic decomposition of nitro compounds from the experimental values for the compounds involved in the teaching data base for Eqs. (1a) (*a*) and (1b) (*b*) (see Table 1).

Table 2. Characteristics of the best equations with different numbers of descriptors for 81 compounds of the teaching data base

Equation	Number of descriptors	R	S	F
(1a)	13	0.984	2.56	159.55
(1b)	12	0.982	2.66	159.20
(1c)	11	0.980	2.76	161.43
(1d)	10	0.979	2.86	165.10

system, which lead to generation of "empty" and "full" chains. In "full" chains, each atom is coded, whereas in "empty" chains, only the first and the last atoms are coded, *i.e.*, the code of every "inner" atom is "empty" ("xxxxx").

The atom code consists of five symbols and has the following form:

The atom code = element (two symbols), the d marker, the b marker, and the r marker.

The d marker denotes the number of neighboring atoms (except for H). The b marker describes the types of chemical bonds for a given atom using the designation "s" for the case when all bonds are single, "d" for the case when there is one double bond, "t" for the occurrence of a triple bond, "a" for the existence of an aromatic bond, and "w" for two double bonds. The r marker identifies the participation of the atom in cyclic fragments as "c" for acyclic atom, "r" for an intracyclic atom, and "s" for an intracyclic atom with an attached substituent.

We studied the models consisting of "empty" as well as of "full" chains. The sizes of the chains generated varied from 3 to 5 atoms with the participation functions F of generation of complex features as well as without these functions. The quality of the models was estimated from the multiple correlation coefficient R, standard deviation S, and the Fisher criterion F. The prognostic efficiency was established from the prediction error vector, *i.e.*, we calculated the values of R_v, S_v, and F_v, which are the analogs of the R, S, and F parameters, respectively.

The best model was obtained with the use of "empty" chains with the participation of F functions. This model consists of 60 correlation equations with different numbers of descriptors (R = 0.968—0.986; F = 134.64—156.89). The BIBIGON system makes it possible to select equations intended for the prediction/estimation of the properties of new compounds, from the model. Of all equations constructed, the parameters of two correlation equations are given in Table 1. Equation (1a) is based on 15 descriptors and is characterized by R = 0.986, F = 175.19, and S = 2.42; the parameters of the "cross validation" are as follows: R_v = 0.980, F_v = 162.63, and S_v = 2.83. Equation (1b) is based on nine descriptors and has the following characteristics: R = 0.968, F = 136.28, and S = 3.48; the parameters of the "cross validation" are as follows: R_v = 0.934, F_v = 116.63, and S_v = 3.93. Figure 1 exhibits a comparison of the experimental values of E_a of thermal decomposition of a number of nitro compounds and the values calculated from the above-mentioned equations.

With the aim of testing the prognostic efficiency of the equations of the best model, the teaching data base was arranged in order of increasing E_a, and each tenth compound was excluded, *i.e.*, the test list, which consisted of nine structures, was obtained. Within the framework of the best model ("empty" chains, included functions), equations of type (1) with different numbers of descriptors were constructed for the remaining 81 compounds. The characteristics of the four best equations [(1a)—(1d)] found are given in Table 2; the values of E_a for nine test compounds, which were calculated using these equations, are listed in Table 3.

Therefore, we obtained equations for the calculation of the activation energy of thermolysis, which give more accurate values than those obtained previously[17] (R = 0.974 and S = 3.24; in our calculations, R = 0.986 and S = 2.42). Besides, the parameters introduced were calculated[17] with the use of different scales and normalizations, and the correlation equation reported contained no data on the characteristic features of chemical structures. The equations obtained in this work clearly demonstrate contributions of particular fragments of the molecular structure to the change in the activation energy. The weighting coefficients of the descriptors involved in this equation account for this contribu-

Table 3. Activation energies of thermal decomposition (E/kcal mol^{-1}) in the gaseous phase estimated for nine test structures using Eqs. (1a)—(1d), whose characteristics are given in Table 2

Structure	E_{exp}	$E_{(1a)}$	$E_{(1b)}$	$E_{(1c)}$	$E_{(1d)}$	E_{av}	$E_{exp} - E_{av}$
4-Nitropyridine	69.65	71.58	70.39	64.83	64.65	67.86	2.09
n-Nitroaniline	73.66	70.95	71.35	71.98	71.55	71.46	2.02
4-Nitrobiphenyl	62.80	56.45	54.70	54.99	54.63	55.19	7.60
1,1-Dinitroethyl 2,4-dinitrophenyl sulfide	37.50	37.35	38.67	38.18	38.15	38.09	−0.59
Tetranitromethane	40.90	39.04	39.21	39.31	39.30	39.22	−1.69
Fluorotrinitromethane	42.50	41.63	41.72	41.80	41.79	41.74	−0.77
Bromofluorodinitromethane	39.50	43.61	43.67	43.76	43.76	43.07	4.20
Iodotrinitromethane	34.40	40.19	40.34	40.47	40.46	40.37	−5.97
Fluorodinitromethane	47.50	44.97	45.01	45.12	45.12	45.06	2.45

Note. $E_{av} = (E_{(1a)} + E_{(1b)} + E_{(1c)} + E_{(1d)})/4$.

tion. The relationships deduced allow the prediction of activation energies of thermal decomposition of nitro compounds from the structural formulas on a personal computer using the BIBIGON program package.

We thank Professor V. A. Shlyapochnikov for helpful discussion and valuable advice. The work was financially supported in part by the Russian Foundation for Basic Research (Project Nos. 94-03-09323, 94-01-00041, and 96-01-01518).

References

1. G. P. Sharnin, V. V. Zverev, I. Sh. Saifillin, I. S. Gaizin, I. F. Falyakhov, and V. V. Dubikhin, *Izv. Akad. Nauk SSSR, Ser. Khim.*, 1981, 521 [*Bull. Acad. Sci. USSR, Div. Chem. Sci.*, 1981, **30**, 368 (Engl. Transl.)].
2. V. G. Matveev, *Kinetika i mekhanizm termicheskogo razlozheniya aromaticheskikh soedinenii v gazovoi faze* [*Kinetics and Mechanism of Thermal Decomposition of Organic Compounds in the Gaseous Phase*], Ph. D. Thesis, Institute of Chemical Physics in Chernogolovka, Russian Academy of Sciences, Chernogolovka, 1978 (in Russian).
3. V. G. Matveev, V. V. Dubikhin, and G. M. Nazin, *Izv. Akad. Nauk SSSR, Ser. Khim.*, 1978, 783 [*Bull. Acad. Sci. USSR, Div. Chem. Sci.*, 1978, **27**, 675 (Engl. Transl.)].
4. V. G. Matveev, V. V. Dubikhin, and G. M. Nazin, *Kinet. Katal.*, 1976, **17**, 280 [*Kinet. Catal.*, 1976, **17** (Engl. Transl.)].
5. V. N. Grebennikov, V. N. Erashko, G. M. Nazin, A. V. Sultanov, and S. A. Shevelev, *Izv. Akad. Nauk SSSR, Ser. Khim.*, 1977, 310 [*Bull. Acad. Sci. USSR, Div. Chem. Sci.*, 1977, **26**, 273 (Engl. Transl.)].
6. F. I. Dubovitskii, G. B. Manelis, and L. P. Smirnov, *Zh. Fiz. Khim*, 1961, **35**, 521 [*Russ. J. Phys. Chem.*, 1961, **35** (Engl. Transl.)].
7. G. M. Nazin and G. B. Manelis, *Izv. Akad. Nauk SSSR, Ser. Khim.*, 1972, 811 [*Bull. Acad. Sci. USSR, Div. Chem. Sci.*, 1972, **21** (Engl. Transl.)].
8. G. M. Nazin, G. B. Manelis, and F. I. Dubovitskii, *Usp. Khim.*, 1968, **37**, 1443 [*Russ. Chem. Rev.*, 1968, **37** (Engl. Transl.)].
9. G. M. Nazin, G. B. Manelis, and F. I. Dubovitskii, *Izv. Akad. Nauk SSSR, Ser. Khim.*, 1971, 1239 [*Bull. Acad. Sci. USSR, Div. Chem. Sci.*, 1971, **20**, 1147 (Engl. Transl.)].
10. A. Appin, O. Todes, and Yu. Khariton, *Zh. Fiz. Khim.*, 1936, **8**, 866 (in Russian).
11. V. V. Zverev, I. Sh. Saifullin, and G. P. Sharnin, *Izv. Akad. Nauk SSSR, Ser. Khim.*, 1978, 313 [*Bull. Acad. Sci. USSR, Div. Chem. Sci.*, 1978, **27**, 269 (Engl. Transl.)].
12. K. K. Andreev, *Termicheskoe razlozhenie i gorenie vzryvchatykh veshchestv* [*Thermal Decomposition and Combustion of Explosives*], Nauka, Moscow, 1966, 312 pp. (in Russian).
13. V. I. Vedeneev and A. A. Kibkalo, *Konstanty skorosti gazofaznykh monomolekulyarnykh reaktsii* [*Rate Constants for Gas-Phase Monomolecular Reactions*], Nauka, Moscow, 1972, 164 pp. (in Russian).
14. K. K. Andreev, A. P. Glazukhova, N. D. Maurina, and B. S. Svetlov, *Zh. Fiz. Khim.*, 1958, **32**, 1726 [*Russ. J. Phys. Chem.*, 1958, **32** (Engl. Transl.)].
15. F. I. Dubovitskii and B. L. Korsunskii, *Usp. Khim.*, 1981, **50**, 1828 [*Russ. Chem. Rev.*, 1981, **50** (Engl. Transl.)].
16. V. I. Minkin, B. Ya. Simkin, and R. M. Minyaev, *Kvantovaya khimiya organicheskikh soedinenii. Mekhanizmy reaktsii* [*Quantum Chemistry of Organic Compounds: Reaction Mechanisms*], Khimiya, Moscow, 1986, 248 pp. (in Russian).
17. T. S. Pivina, D. V. Sukhachev, L. K. Maslova, V. A. Shlyapochnikov, and N. S. Zefirov, *Dokl. Akad. Nauk*, 1993, **330**, 468 [*Dokl. Chem.*, 1993, **330** (Engl. Transl.)].
18. M. I. Kumskov, L. A. Ponomareva, E. A. Smolenskii, D. F. Mityushev, and N. S. Zefirov, *Izv. Akad. Nauk, Ser. Khim.*, 1994, 1391 [*Russ. Chem. Bull.*, 1994, **43**, 1317 (Engl. Transl.)].
19. I. V. Stankevich, *Grafy v strukturnoi khimii* [*Graphs in Structural Chemistry*], in *Primenenie teorii grafov v khimii* [*Application of the Theory of Graphs in Chemistry*], Eds. N. S. Zefirov and S. N. Kuchanov, Nauka, Novosibirsk, 1988, 7 (in Russian).
20. B. L. Korsunskii, G. V. Sitonin, B. S. Fedorov, and F. I. Dubovitskii, *Izv. Akad. Nauk SSSR, Ser. Khim.*, 1989, 790 [*Bull. Acad. Sci. USSR, Div. Chem. Sci.*, 1989, **38**, 710 (Engl. Transl.)].
21. N. S. Zefirov, M. I. Kumskov, L. A. Ponomareva, D. F. Mityushev, and E. A. Smolenskii, *Tez. dokl. IX Vsesoyuz. konfer. po khimicheskoi informatike* [*Abstract of Papers, IX All-Union Conf. on Chemical Informatics*], Chernogolovka, 1992, 159 (in Russian).

Received March 21, 1995;
in revised form February 16, 1995