*Russian Journal of General Chemistry, Vol. 71, No. 5, 2001, pp. 712–717. Translated from Zhurnal Obshchei Khimii, Vol. 71, No. 5, 2001, pp. 759–764.*

*Original Russian Text Copyright © 2001 by Golovanov, Tsygankova.*

# Structure–Property Correlation Equation: V.[1] Estimations of the Distribution Factors of Molecules Containing Several Polar Groups in the Octanol–Water System

**I. B. Golovanov and I. G. Tsygankova**

*Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences, Pushchino, Moscow oblast, Russia*

Received August 16, 1999

**Abstract**—A simple structure–property correlation equation was suggested for estimating the distribution factors of molecules in the octanol–water system. The presence of several polar groups in molecules requires inclusion of additional terms taking into account their interaction. Their values are approximately estimated from a limited set of test parameters.

In [2] we suggested an approach allowing fairly accurate estimation of the logarithm of the distribution factor ($\log P$) in the octanol–water system of hydrocarbons and other classes of organic substances containing one functional group ($NH_2$, OH, O, etc.). In terms of this approach, $\log P$ of molecules RX is described by the expression

$$\log P_{RX} = \Sigma \log P_{ii} + \Sigma\Sigma \log P_{ij}, \tag{1}$$

where $\log P_{ii}$ are single-fragment contributions and $\log P_{ij}$ are two-fragment contributions corresponding to interaction of *i*th and *j*th fragments. Although the groups $CH_n$ directly bound to X differ from similar groups in hydrocarbon molecules and their contributions $\log P_{ii}$ should differ from those of $CH_n$ groups in hydrocarbons, these differences can be included in the values of $\log P_{ij}$ corresponding to contributions of groups X; therefore, passing from R–$CH_3$ to R–X requires introduction of a single parameter. In this case, estimation of $\log P$ involves no problems. However, if a molecule includes several groups, it is necessary to take into account both variation of the contributions $\log P_{ii}$ of adjacent groups (they will be different depending on the mutual arrangement of the functional groups) and the contributions $\log P_{XY}$ corresponding to interaction of the functional groups. It is very difficult to estimate these contributions, because the set of experimentally studied examples is limited and does not allow formulation of any reasonable physicochemical criteria for choosing the type of the dependence of $\log P_{XY}$ on the mutual arrangement of X and Y.

Here we consider an approximate method for evaluating the contributions $\log P_{XY}$.

The problems arising in this case can be illustrated by several examples. Consider first the available experimental data for ethers (Table 1). It is seen that the difference between the $\log P$ values for ethers R–O–R' and isoelectronic hydrocarbons R–$CH_2$–R', $\Delta_{CH_2 \to O} = \log P_{ROR'} - \log P_{RCR'}$, in the case of aliphatic R and R' varies from –2.18 to –2.63. Although the scatter of $\Delta_{CH_2 \to O}$ is significant, we can take an average value for ethers with aliphatic substituents: $\Delta_{CH_2 \to O} = -2.42 \pm 0.17$, which allows estimation of $\log P$ of such compounds with a reasonable accuracy. However, this can be done only for aliphatic R and R'. For benzyl methyl ether, $\Delta_{CH_2 \to O}$ is appreciably lower, and if the O atom is bound to two phenyl rings, it becomes practically zero. No general rule can be suggested for choosing $\Delta_{CH_2 \to O}$, but, with a wide set of experimental data available, it is possible to estimate $\Delta_{CH_2 \to O}$ empirically for certain particular cases, as it is done usually.

Solution of specific problems is often based on the assumption that corrections for interaction of polar groups can be estimated from data on their accessible surface areas. It is implicitly assumed in this case that the accessible surface areas of the same groups in different molecules are close. However, as seen from Table 1, the accessible surface areas of ether oxygen atoms in different molecules differ by a factor of more than 2, although the $\Delta_{CH_2 \to O}$ values are similar in all cases. This fact casts doubt on correlation of the contributions $\log P_{XY}$ with accessible surface areas, since there are no clear correlations between

---

[1] For communication IV, see [1].

**Table 1.** Values of log*P* for some ethers and isoelectronic hydrocarbons

| Ether | $\log P_{eth}$ [3] | Isoelectronic hydrocarbon | $\log P_{HC}$ [3] | $\Delta_{CH_2 \to O}$ | $S_O$, Å$^2$ |
|---|---|---|---|---|---|
| $CH_3$–O–$CH_3$ | 0.10 | $CH_3$–$CH_2$–$CH_3$ | 2.36 | –2.26 | 14.91 |
| $C_3H_7$–O–$CH_3$ | 1.21 | *n*-$C_5H_{12}$ | 3.39 | –2.18 | 10.97 |
| $C_2H_5$–O–$C_2H_5$ | 0.89 | *n*-$C_5H_{12}$ | 3.39 | –2.50 | 6.95 |
| $C_4H_9$–O–$CH_3$ | 1.66 | *n*-$C_6H_{14}$ | 3.90 | –2.24 | 10.97 |
| $C_4H_9$–O–$C_2H_5$ | 2.03 | *n*-$C_7H_{16}$ | 4.66 | –2.63 | 6.95 |
| $C_3H_7$–O–$C_3H_7$ | 2.03 | *n*-$C_7H_{16}$ | 4.66 | –2.63 | 6.95 |
| Tetrahydrofuran | 0.46 | $C_5H_{10}$ | 3.00 | –2.54 | 11.20 |
| $C_6H_5$–$CH_2$–O–$CH_3$ | 1.92 | $C_6H_5$–$C_3H_7$ | 3.63 | –1.71 | 8.53 |
| $C_6H_5$–O–$CH_3$ | 2.08 | $C_6H_5$–$C_2H_5$ | 3.15 | –1.07 | 13.67 |
| $C_6H_5$–O–$C_6H_5$ | 4.21 | $C_6H_5$–$CH_2$–$C_5H_5$ | 4.14 | +0.07 | 7.34 (8.65) |
| Dibenzofuran | 4.12 | Fluorene | 4.18 | –0.06 | 14.76 |

**Table 2.** Values of log*P* for some di- and triethers and isoelectronic hydrocarbons

| Ether | $\log P_{eth}$ [3] | Isoelectronic hydrocarbon | $\log P_{HC}$ [3] | Δ' | $\Delta_{OO}$ |
|---|---|---|---|---|---|
| $CH_3$–O–$CH_2$–O–$CH_3$ | 0.18 | *n*-$C_5H_{12}$ | 3.39 | –3.21 | 1.63 |
| $C_2H_5$–O–$CH_2$–O–$C_2H_5$ | 0.84 | *n*-$C_7H_{16}$ | 4.66 | –3.82 | 1.02 |
| 1,3-Dioxolane | –0.37 | Cyclopentane | 3.00 | –3.37 | 1.47 |
| Trioxane | –0.43 | Cyclohexane | 3.44 | –3.87 | 1.13 |
| 2,4,6-Trimethyl-1,3,5-trioxane | 0.67 | 2,4,6-Trimethylcyclohexane | 4.20 | –3.53 | 1.24 |
| $CH_3$–O–$(CH_2)_2$–O–$CH_3$ | –0.21 | *n*-$C_6H_{14}$ | 3.90 | –4.11 | 0.73 |
| $C_2H_5$–O–$(CH_2)_2$–O–$C_2H_5$ | 0.66 | *n*-$C_8H_{18}$ | 5.18 | –4.52 | 0.34 |
| 1,4-Dioxane | –0.27 | Cyclohexane | 3.44 | –3.71 | 1.13 |

the accessible surface areas of polar groups and their contributions $P_X$.

The empirical corrections for interaction of polar groups can be found from experimental data. Consider, e.g., data for diether and triether molecules (Table 2). If the ether O atoms did not interact with each other, the difference between the log*P* values of an ether containing *n* O atoms and the isoelectronic hydrocarbon would be equal to –2.42*n*. However, this is not the case. The difference between the experimental value and that estimated from this relationship defines the correction for the interaction of O atoms, $\Delta_{OO}$:

$$\Delta_{OO} = (\log P_{eth} - \log P_{HC}) - (-2.42n),$$

where $\log P_{eth}$ is log*P* for a molecule R–O–R'–O–R" containing *n* O atoms, $\log P_{HC}$ is log*P* for the isoelectronic hydrocarbon molecule, and –2.42 is the correction for introduction of one oxygen atom. At $n > 2$ this value should be divided by the number of interacting O–O pairs.

Such estimates are rather rough, since they do not take into account the hydrophobic interaction of the radicals R, R', and R" and the differences between the conformations of the ether and isoelectronic hydrocarbon. Table 2 shows that $\Delta_{OO}$ depends on the mutual arrangement of the O atoms. For example, in the case of the 1,3 arrangement of oxygen atoms (separation by one $CH_2$ group) $\Delta_{OO}$ varies from 1.02 to 1.63, whereas in the case of the 1,4 arrangement (separation by two $CH_2$ groups) $\Delta_{OO}$ varies from 0.34 to 1.13. The ranges of variation of these corrections overlap, and it is difficult to make definite conclusions, but for rough estimations of log*P* of ethers we can take $\Delta_{OO}^{13} = 1.23 \pm 0.16$ and $\Delta_{OO}^{14} = 0.73 \pm 0.26$.

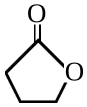
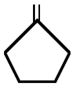
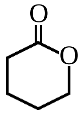
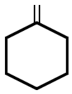
This approach is consistent with reasonable physicochemical assumptions. Indeed, $\Delta_{OO}$ should be equal to zero at large distances when the O atoms can be considered as noninteracting (the distance should be larger than ~6 Å—the distance allowing accommodation of a water molecule) and should increase with decreasing O–O distance. The analytical form of the

**Table 3.** Corrections for interaction of O atoms in ether molecules

| Molecule | $\Delta_{OO}^{exp}$ [3] | $\Delta_{OO}^{calc\,a}$ | $\Delta_{OO}^{calc\,b}$ | Molecule | $\Delta_{OO}^{exp}$ [3] | $\Delta_{OO}^{calc\,a}$ | $\Delta_{OO}^{calc\,b}$ |
|---|---|---|---|---|---|---|---|
| $R–O–CH_2–O–R'$ | 1.02 – 1.63 | 1.31 | 1.26 | 2,4,6-Tetramethyl-1,3,5-trioxane | 1.24 | 1.31 | 1.26 |
| 1,3-Dioxolane | 1.47 | 1.31 | 1.26 | $R–O–(CH_2)_2–O–R'$ | 0.34 – 0.73 | 0.56 | 0.41 |
| Trioxane | 1.13 | 1.31 | 1.26 | 1,4-Dioxane | 1.13 | 1.13 | 1.03 |

[a] From relation (2), $k_1 = 5.864 \pm 1.56$, $k_2 = 0.643 \pm 0.1$. [b] From relation (2), $k_1 = 8.960 \pm 0.80$, $k_2 = 0.844 \pm 0.035$ (as in diols).

**Table 4.** Value of $\log P$ for some esters, lactones, and isoelectronic hydrocarbons

| Ester (lactone) | $\log P_{est}$ [3] | Hydrocarbon | $\log P_{HC}$ [3] | Δ |
|---|---|---|---|---|
| $CH_3–COO–CH_3$ | 0.18 | $CH_3–C(CH_2)–CH_2–CH_3$ | 2.82 | –2.64 |
| $C_2H_5–COO–CH_3$ | 0.89 | $C_2H_5–C(CH_2)–CH_2–CH_3$ | 3.30 | –2.41 |
| $CH_3–COO–C_2H_5$ | 0.73 | $CH_3–C(CH_2)–CH_2–C_2H_5$ | 3.33 | –2.60 |
| $C_2H_5–COO–C_2H_5$ | 1.21 | $C_2H_5–C(CH_2)–C_3H_7$ | 3.84 | –2.63 |
| $CH_3–COO–C_3H_7$ | 1.23 | $CH_3–C(CH_2)–C_4H_9$ | 3.84 | –2.61 |
| $CH_3–COO–C_4H_9$ | 1.78 | $CH_3–C(CH_2)–C_5H_{11}$ | 4.35 | –2.57 |
|  | –0.64 |  | 2.71 | –3.35 |
|  | –0.35 |  | 3.22 | –3.57 |

dependence of $\Delta_{OO}$ on $R_{OO}$ is unknown; based on available data, we can suggest the following relation:

$$\Delta_{OO} \approx k_1 \exp(-k_2 R_{OO}), \qquad (2)$$

where $k_1$ and $k_2$ are parameters determined from data for test examples, and $R_{OO}$ is the distance between the oxygen atoms. The values of $k_1$ and $k_2$ can be determined from the experimental values of $R_{OO}$ avaibable for $R–O–CH_2–O–R'$, $R–O–(CH_2)_2–O–R'$, and 1,4-dioxane, assuming that at $R_{OO}$ = 7 Å $\Delta_{OO} \sim 0$. From these data we determined the parameters $k_1$ and $k_2$: $k_1 = 5.864 \pm 1.56$ and $k_2 = 0.643 \pm 0.1$; using these values, we obtained the values of $\Delta_{OO}$ nicely consistent with those obtained from the experimental data (Table 3).

The ether O atoms are the simplest examples of polar groups; their "conformational state" does not vary, and we can assume that in all cases they are characterized by the same value of $\log P_{ii}$. The electronic states (and also such characteristics as conformation, $\log P_{ij}$ contributions, etc.) of polar groups consisting of several atoms can be different, and therefore the use of a single value of $\log P_{ij}$ will not give adequate results. Let us illustrate this problem by the example of esters, lactones, and isoelectronic hydrocarbons. Some data on $\log P$ of these molecules are given in Table 4.

It is seen that in going from a hydrocarbon to the isoelectronic ester the group ${}_{-CH_2}{>}C{=}CH_2$ is replaced by ${}_{-O}{>}C{=}O$; the corresponding correction is –2.57 ± 0.07, whereas in going to lactones the correction is –3.46 ± 0.11. Formally, the ester groups in these compounds are similar (–COO–), but actually they differ significantly: In esters the lone electron pairs of the O atoms are in the *trans* conformation relative to the C–O bond, and in lactones, in the *cis* conformation.

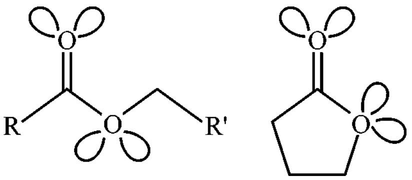


**Table 5.** Values of log*P* for some alcohols and isoelectronic hydrocarbons

| Alcohol | $\log P_{ROH}$ [3] | Hydrocarbon | $\log P_{HC}$ [3] | $\Delta_{OH}$ | Alcohol | $\log P_{ROH}$ [3] | Hydrocarbon | $\log P_{HC}$ [3] | $\Delta_{OH}$ |
|---|---|---|---|---|---|---|---|---|---|
| $CH_3$–OH | –0.77 | $C_2H_6$ | 1.81 | –2.58 | $C_5H_{11}$–OH | 1.56 | *n*-$C_6H_{14}$ | 3.90 | –2.34 |
| $C_2H_5$–OH | –0.31 | $C_3H_8$ | 2.36 | –2.67 | $C_6H_{13}$–OH | 2.03 | *n*-$C_7H_{16}$ | 4.66 | –2.63 |
| $C_3H_7$–OH | 0.25 | *n*-$C_4H_{10}$ | 2.89 | –2.64 | $C_7H_{15}$–OH | 2.72 | *n*-$C_8H_{18}$ | 5.18 | –2.46 |
| $C_4H_9$–OH | 0.88 | *n*-$C_5H_{12}$ | 3.39 | –2.51 | | | | | |

**Table 6.** Values of log*P* for some polyhydric alcohols

| Polyhydric alcohol | $\log P_{R(OH)_2}$ [3] | Monohydric alcohol | $\log P_{ROH}$ [3] | $\Delta_{OH}$ | $\Delta_{OH\cdots OH}$ |
|---|---|---|---|---|---|
| OH–$CH_2$–$CH_2$–OH | –1.36 | $CH_3$–$CH_2$–$CH_2$–OH | 0.25 | –1.61 | 0.95 |
| OH–$(CH_2)_3$–OH | –1.04 | *n*-$C_4H_9$–OH | 0.88 | –1.92 | 0.64 |
| OH–$(CH_2)_4$–OH | –0.83 | *n*-$C_5H_{11}$–OH | 1.56 | –2.39 | 0.17 |

They have different dipole moments (1.67 and 4.12 D [4]) and, apparently, interact with solvent molecules quite differently. Hence, the correction for introduction of the COO group should depend on its conformation, and generally it can vary in the range from –2.57 to –3.46.

Consider then the molecules of alcohols and isoelectronic hydrocarbons. Some data on log*P* of these molecules are given in Table 5. It is seen that replacement of the $CH_3$ group by the OH group requires a correction $\Delta_{CH_3 \to OH} = -2.56 \pm 0.08$. However, if we compare dihydric and monohydric alcohols (Table 6), we see that introduction of the second OH group is described by different corrections depending on the mutual arrangement of these groups. This fact is due to interaction of OH groups with each other. As in the case of ethers, the correction for interaction of the OH groups can be described as

$$\Delta_{OH\cdots OH} \simeq \log P_{OH} - \log P_{HC} - (-2.56 \times 2),$$

where $\log P_{OH}$ is log*P* of diol, $\log P_{HC}$ is log*P* of the isoelectronic hydrocarbon, and –2.56 is the correction for introduction of one OH group.

Available data on log*P* for linear alcohols and the corrections $\Delta_{OH\cdots OH}$ are listed in Table 6. It is seen that as the distance between the interacting OH groups (determined by the number of separating CC bonds) increases, the corrections for interaction decrease, as in the case of ethers.

In aqueous and ethanolic solutions, diols exist as mixtures of various conformers whose relative content is unknown. Therefore, it is difficult to make definite conclusions without additional assumptions. We assumed that the structures of diols are similar to those observed in the gas phase [5–7]. This allowed us to choose the $R_{OH\cdots OH}$ values and estimate the corrections for the $\Delta_{OH\cdots OH}$ interaction by relation (2). The parameters $k_1$ and $k_2$ were chosen from data for 1,2-ethanediol, 1,3-propanediol, and 1,4-butanediol. In the latter two cases we assumed the O···O distance to be close to 3.0 Å. In this case the obtained parameters $k_1$ and $k_2$ were 23.933 ± 1.28 and 1.256 ± 0.02, respectively. With these parameters we calculated the values of $\Delta_{OH\cdots OH}$ and log*P* for diols (Table 7).

The agreement between the calculated and experimental values of $\Delta_{OH\cdots OH}$ is satisfactory taking into account the roughness of the assumptions made.

The parameteres $k_1$ and $k_2$ determining the corrections for interaction of the ether O atoms and hydroxy groups differ considerably. However, if we take the common parameters $k_1$ and $k_2$ for determination

**Table 7.** Corrections for interaction of OH groups in diol molecules

| Molecule | $\Delta^{exp}_{OH\cdots OH}$ [3] | $\Delta^{calc}_{OH\cdots OH}$ [a] | $\Delta^{calc}_{OH\cdots OH}$ [b] |
|---|---|---|---|
| OH–$CH_2$–$CH_2$–OH | 0.95 | 0.837 | (0.94) |
| OH–$(CH_2)_3$–OH | 0.64 | 0.553 | (0.712) |
| OH–$(CH_2)_4$–OH | 0.17 | 0.553[c] | (0.712) |

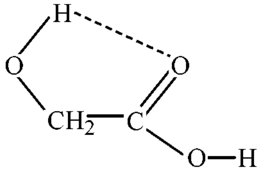
[a] $k_1 = 23.933 \pm 1.28$, $k_2 = 1.256 \pm 0.02$. [b] $k_1 = 8.960 \pm 0.80$, $k_2 = 0.844 \pm 0.035$ (as in diethers). [c] Disagreement with the experiment is probably due to the fact that from the whole set of conformers in equilibrium we took into account only the conformer in which the H bond is the most preferable in energy.

**Table 8.** Values of log$P$ for some F-containing alcohols

| F-Alcohol | log$P_{R_FOH}$ [3] | Alcohol | log$P_{ROH}$ [3] | $\Delta_{CH_3\to F}$ |
|---|---|---|---|---|
| $CF_3$–$CH_2$–OH | 0.41 | $CH_3$–$CH_2$–OH | –0.31 | 0.72 |
| $C_2F_5$–$CH_2$–OH | 1.23 | *n*-$C_3H_7$–OH | 0.25 | 0.98 |
| $C_3F_7$–$CH_2$–OH | 1.94 | *n*-$C_4H_9$–OH | 0.88 | 1.06 |

of both $\Delta_{O\cdots O}$ and $\Delta_{OH\cdots OH}$ ($k_1$ = 8.960 ± 0.80, $k_2$ = 0.844 ± 0.035), the agreement of the log$P$ values calculated using such corrections (Tables 3, 7, fourth columns) with the experiment is also satisfactory.

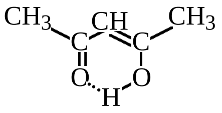
Along with the above data, there are also data on interaction of the OH group with the carboxy and carbonyl groups. In the molecule of hydroxyacetic acid, an intramolecular H bond is formed:

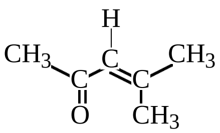


Correction for interaction of the OH group with the carboxy group can be estimated from the comparison of data for hydroxyacetic and propionic acids (log$P$ –1.11 and 0.33, respectively).

If the OH group in hydroxyacetic acid did not interact with the carboxy group, its log$P$ would be 0.33 – 2.56 = –2.23. The difference of this value from the experimental data is due to the interaction OH···O=C; its contribution $\Delta_{OH\cdots O=C}$ is +1.12. Estimation of $\Delta_{OH\cdots O=C}$ by relation (2) with the latter set of $k_1$ and $k_2$ gives the value of +0.94.

The effect of intermolecular interaction of polar groups is considerably stronger in acetylacetone, as judged from comparison of its log$P$ with that of 4-methylpent-3-en-2-one:

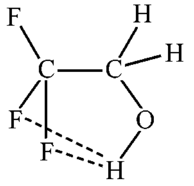


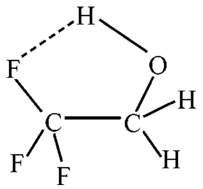


Their log$P$ is 0.40 and 0.61, respectively. If the OH group introduced into the ketone molecule did not form the intramolecular H bond, log$P$ of acetylacetone would be 0.61 – 2.56 = –1.95. Hence follows $\Delta_{OH\cdots O=C}$ = +2.35; estimation of $\Delta_{OH\cdots O=C}$ by relation (2) gives the value of +1.20.

It should be noted that in the acetylacetone molecule the intramolecular H bond is extremely strong, affecting the characteristics of the bonded fragments. This fact is apparently responsible for the large deviation of the calculated $\Delta_{OH\cdots O=C}$ values from the experiment and for considerable increase in this contribution in acetylacetone as compared to hydroxyacetic acid.

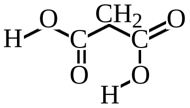
The effect of interaction of polar groups is considerably smaller in alcohols containing such groups as $CF_3$, $C_2F_5$, etc. In such molecules, intramolecular H bonds OH···F are formed:



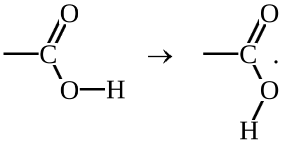


Since the log$P$ values of $C_nF_m$ and $C_nH_m$ molecules differ insignificantly ($CF_4$ ~1.18, $CH_4$ ~1.09; $C_2F_6$ ~2.0, $C_2H_6$ ~1.81, etc.), the difference between the values of log$P$ for alcohols containing the $CF_3$ group and "common" alcohols gives a correction for formation of the intramolecular H bond OH···F (Table 8). It is seen that these corrections are close to those for the isoelectronic diols and can be fairly accurately estimated by relation (2).

In conclusion, let us consider data for malonic acid.



Its log$P$ is –0.81, and, to fit this value, we should assume $\Delta_{OH\cdots O=C}$ = 0.93. In this case, the pattern is similar to that in acetylacetone (closure of a six-membered ring with formation of a strong intramolecular H bond), but the correction for interaction is lower by a factor of about 2.5. The point is that in malonic acid formation of the H bond is accompanied by conformational changes within a polar group. Indeed, for the ring closure the carboxy group should rearrange into the “perturbed” conformational state, similarly to passing from esters to lactones:



We will give no further examples, as they do not refine the pattern. The above data show that interaction of polar groups makes the molecule more hydrophobic ($\Delta_{XY} > 0$). The correction for interaction of polar groups $\Delta_{XY}$ decreases with increasing distance between the fragments X and Y and can be approximately estimated by relation (2). In this case, possible conformational changes of the polar group should be taken into account.

## REFERENCES

1. Golovanov, I.B. and Tsygankova, I.G., *Zh. Obshch. Khim.*, 2001, vol. 71, no. 4, pp. 538–543.
2. Golovanov, I.B. and Tsygankova, I.G., *Zh. Obshch. Khim.*, 2000, vol. 70, no. 2, pp. 223–230.
3. Hansch, C. and Hoekman, D., *Exploring QSAR*, Washington: Am. Chem. Soc., 1995.
4. Osipov, O.A., Minkin, V.I., and Garnovskii, A.D., *Spravochnik po dipol’nym momentam* (Handbook on Dipole Moments), Moscow: Vysshaya Shkola, 1971.
5. Nagy, P.J., Dunn, W.J., Alagona, G., and Ghio, C., *J. Am. Chem. Soc.*, 1991, vol. 113, no. 18, pp. 6719–6729.
6. Die, T., Topol, J.A., and Burt, S.K., *J. Phys. Chem.*, 1994, vol. 98, no. 4, pp. 1121–1128.
7. Traettenberg, M. and Hedberg, K., *J. Am. Chem. Soc.*, 1994, vol. 116, no. 4, pp. 1382–1387.