



SCIREA Journal of Biology

<http://www.scirea.org/journal/Biology>

January 14, 2018

Volume 3, Issue 1, February 2018

INTERRELATION OF CARCINOGENIC PROPERTIES OF METABOLITES WITH THEIR INFORMATION AND ELECTRONIC DESCRIPTORS

VLADIMIR MUKHOMOROV

University of Naples "Federico II" Via Cintia, I-80126, Naples, ITALY

Email: Vmukhomorov@mail.ru

Abstract

The interrelationship between the carcinogenic effects of metabolites and molecular descriptors is modeled. The information function and the average number of valence electrons of molecules are used as molecular descriptors. It is demonstrated that the conversion of non-carcinogenic initial agents to carcinogenic metabolites is accompanied by a systematic increase in molecular descriptors. The factor that limits the carcinogenic activity of chemical compounds is considered.

Keywords: Carcinogen, metabolite, descriptors, oxy compounds, information, valence electrons

Introduction

It was shown earlier [1-3] that the molecular descriptors determining the information content of molecules (H) and the average number of valence electrons of the molecule (Z) make it possible to statistically reliably separate chemical compounds possessing carcinogenic activity from chemical compounds that do not possess carcinogenic activity. The statistical approach was tested [3] on a large number of chemical compounds (more than 600 chemical compounds) of various classes. To do this, samples were being used that contained aromatic amines, nitroso compounds, alkylating agents, halogenated chemical compounds, dyes, sulfur-containing chemical compounds, oxy compounds, mustards, natural chemicals, medicine.

The information function of a molecule is defined as follows [4]

$$H = -\sum_i p_i \log_2 p_i, \quad (1)$$

here $p_i = n_i / N$, n_i is the number of atoms of type i in a molecule, N is the total number of atoms. Summation is performed over all atoms of the molecule. The ratio n_i / N determines equity participation of the i -th kind of atom in the molecule. We actually use the combinatorial Kolmogorov representation [5] in equation (1). Equation (1) is free of any probabilistic assumptions. In this case, the amount of information in the molecule is a function of the number of elements of a finite set of atoms. The information function is an integral characteristic of the molecule. The function H quantifies the diversity of the atomic structure of the chemical compound. The smaller the value of the function H , the more diverse the multicomponent system. The average number of electrons on the outer shell of the atoms of the molecule is determined by the equation

$$Z = -\sum_i n_i Z_i / N. \quad (2)$$

Here Z_i is the number of valence electrons of the i -th atom; n_i is the number of atoms of i -th kind; Z_i is a number of electrons in the outer electron shell. The summation is performed over all the atoms in a molecule N is the total number of atoms. A definition of the mean number of valence electrons close in meaning was used in the work [6]. It is sufficient knowledge only the

gross formula of a chemical compound for the calculation of molecular descriptors (1) and (2). This approach allows express analysis of the carcinogenicity of new chemical compounds.

In accordance with [3] carcinogenic chemical compounds are grouped in the range of mean values $Z_1^{(av)} = 2.92$ and $H_1^{(av)} = 1.56bits$. Non-carcinogenic chemicals are grouped around of mean values: $Z_2^{(av)} = 3.23$ and $H_2^{(av)} = 1.70bits$. These mean values were obtained for a sample that included chemical compounds of different classes. The threshold values of molecular descriptors are equal to $Z^{(th)} = 3.06$, $H^{(th)} = 1.62bits$ [3]. These threshold values statistically reliably separate carcinogens from inactive chemical compounds. The statistical model allows deviations from these values if the chemical compounds belong to only one class of chemical agents. For example, if the sample contains only nitroso compounds, the threshold values of the molecular descriptors are slightly higher than the values given above. However, it should be noted that not all chemical compounds for which $Z \approx Z_1^{(av)}$ and $H \approx H_1^{(av)}$ should be carcinogens.

Using the descriptors Z and H , we analyzed the carcinogenic activity of metabolites.. The starting chemical compounds are not carcinogens or weakly active. For example, cholesterol is a weak carcinogen ($Z = 2.16$, $H = 1.04bits$). The reference [7] denies the possible carcinogenicity of cholesterol. Whereas according to [8] the product of its conversion of 6 β -peroxy- Δ^4 -cholesten-3-on is a strong carcinogen (descriptors increase to the values: $Z = 2.27$, $H = 1.15bits$). That is, the increase in the carcinogenic effect is accompanied by an increase in the values of the descriptors. A similar situation is observed for tryptophan ($Z = 2.89$, $H = 1.60bits$) which does not have carcinogenic activity. As a result of the tryptophan metabolism three chemical compounds are formed [8 (Appendix II. Russian edition)]: 3-hydroxykinurenine ($Z = 2.96$, $H = 1.68bits$), 2-amino-3-hydroxyacetophenol ($Z = 2.90$, $H = 1.60bits$), 3-hydroxyanthranilic acid ($Z = 3.22$, $H = 1.72bits$). These chemicals are carcinogens. The molecular descriptors of all metabolites exceed the descriptors of the tryptophan. Apparently, such an increase in the values of metabolites descriptors is often seen. For example, the conversion of aflatoxin B1 ($Z = 3.31$, $H = 1.47bits$) [9] to carcinogenic aflatoxin M1 ($Z = 3.39$, $H = 1.50bits$) also leads to an increase in the molecular descriptors. The conversion of 4-aminobiphenyl ($Z = 2.67$, $H = 1.21bits$) to *N*-oxy-4-aminobiphenyl ($Z = 2.92$, $H = 1.53bits$) is also accompanied by an increase in descriptor values. The carcinogenic activity of *N*-oxy-4-aminobiphenyl is higher than the initial agent. It is suggested that the "final" carcinogen of aflatoxin B₁ is its metabolite 2,3-epoxyaflatoxin B₁ ($Z = 3.33$, $H = 1.52bits$). The disappearance of the carcinogenic activity of aflatoxin B1 in hypophysectomy testifies to the indirect carcinogenic effect of the initial compound, as the metabolism significantly changes. The descriptors of all metabolites (Table 1) are closer to the threshold values than for the initial chemical compounds. However, it is important to note that

the descriptors Z and H remain below the threshold value. All metabolites are carcinogenic substances, while their values are less than the threshold values.

Metabolites (Table 1) have not only a carcinogenic effect, but they also lead to intoxication of an organism [9]. That is, an increase in the molecular descriptors of metabolites is accompanied by an increase in the toxic properties of molecular compounds. The increase in toxicity with increasing Z descriptor is noted for other biologically active chemical compounds, for example, for antifungal agents [1].

| N | The original chemical compound | Carcinogenic activity | Metabolite | Activity |
|---|--|--------------------------|--|-------------|
| 1 | Cholesterol $Z = 2.16, H = 1.04bits$ | Data is not adequate [7] | 6- β -Peroxy- Δ^4 -cholesten-3-on $Z = 2.27, H = 1.1bits$ | Very active |
| 2 | Tryptophan $Z = 2.89, H = 1.60bits$ | Inactive | 3-Oxykinurenine $Z = 2.96, H = 1.68bits$ | Active |
| | | | 2-Amino-3-hydroxyacetophenol $Z = 2.90, H = 1.60bits$ | Active |
| | | | 3-Oxyanthranilic acid $Z = 3.22, H = 1.72bits$ | Active |
| 3 | Aflatoxin B ₁ $Z = 3.31, H = 1.47bits$ | Slightly active | Aflatoxin M ₁ $Z = 3.39, H = 1.50bits$ | Active |
| | | | 2,3- Epoxiflatoxin B ₁ $Z = 3.33, H = 1.52bits$ | Active |
| 4 | Safrole $Z = 2.82, H = 1.35bits$ | Slightly active | 1'-Hydroxisafrole $Z = 2.96, H = 1.43bits$ | Active |
| | | | C ₁₂ H ₁₂ O ₄ *) [9] $Z = 3.00, H = 1.45bits$ | Active |
| 5 | 4-Aminobiphenyl $Z = 2.67, H =$ | Slightly active | N-Oxy-4-aminobiphenyl $Z = 2.92, H = 1.53bits$ | Active |

| | | | | |
|----|---|------------------------------|--|--|
| | 1.21bits | | | |
| 6 | 2-Naphtylamine $Z = 2.70, H = 1.23bits$ | Slightly active | 2-Amino-1-naphthol $Z = 2.77, H = 1.65bits$ | Very active |
| | | | 2-Naphthylhydroxyamine $Z = 2.86, H = 1.45bits$ | Very active |
| | | | Bis (2-hydroxylamine)-1-naphthyl phosphate $Z = 3.17, H = 1.75bits$ | Very active |
| | | | Bis (2-amino-1-naphthyl) phosphate $Z = 3.17, H = 1.75bits$ | Very active |
| 7 | 1,2-Benzo(a)pyrene $Z = 2.13, H = 0.96bits$ | Carcinogenic Nonmutagenic | 7,8-Dihydroxy-9,10-epoxy-7,8,9,10-tetrahydro-1,2-benzopyrene $Z = 2.46, H = 1.30bits$ | Carcinogenic Very active Mutagenic |
| 8 | Herbicide AAP $Z = 2.88, H = 0.95bits$ | Carcinogenic | 7,8-Dihydroxy-9,10-epoxy-7,8,9,10-tetrahydro-1,2-benzopyrene $Z = 2.82, H = 1.42bits$ | Active |
| 9 | 4-Acetylaminostilbene $Z = 2.73, H = 1.33bits$ | Slightly active | <i>N</i> -Oxy-4-acetylaminostilbene $Z = 2.82, H = 1.42bits$ | Very active |
| 10 | 4-Aminobiphenyl $Z = 2.67, H = 1.40bits$ | Carcinogenic | 3-Oxy-4-aminobiphenyl $Z = 2.80, H = 1.59bits$ | Very active |
| | | | <i>N</i> -Oxy-2-acetylaminofluorene $Z = 2.80, H = 1.40bits$ | Very active |

| | | | | |
|----|--|------------------------------|--|--|
| | =1.21bits | | <i>N</i> -Oxy-4-aminobiphenyl <i>Z</i> = 2.80, <i>H</i> =1.40bits | Very active |
| | | | 4-Amino-3-hydroxy-diphenylsulfate <i>Z</i> = 3.24, <i>H</i> =1.79bits | Very active |
| 11 | Methylurea <i>Z</i> = 2.73, <i>H</i> = 1.69bits | Inactive | <i>N</i> -Nitroso- <i>N</i> -methylurea <i>Z</i> = 3.33, <i>H</i> =1.89bits | Very active |
| 12 | <i>N</i> -Nitrozodimethylamine <i>Z</i> = 2.73, <i>H</i> = 1.69bits | Carcinogenic Nonmutagenic | Methyl diazohydroxide <i>Z</i> = 3.00, <i>H</i> = 1.79bits | Carcinogenic Very active Mutagenic |
| 13 | Cycasin <i>Z</i> = 3.03, <i>H</i> =1.72bits | Active | Diazomethane <i>Z</i> = 3.21, <i>H</i> =1.52bits | Very active |
| 14 | 4-Amino-4-acetylaminobiphenyl <i>Z</i> = 2.76, <i>H</i> =1.36bits | Active | <i>N</i> -Oxy-4-acetylaminobiphenyl <i>Z</i> = 2.87, <i>H</i> =1.46bits | Active |
| 15 | 2-Acetylaminofluorene <i>Z</i> = 2.80, <i>H</i> =1.35bits | Active | <i>N</i> -Oxy-2-acetylaminofluorene <i>Z</i> = 2.90, <i>H</i> =1.45bits | Very active |
| | | | 2-Oxy-2-acetylaminofluorene <i>Z</i> = 2.90, <i>H</i> =1.45bits | Active |
| 16 | 3-Acetylaminofluorene | Inactive | 2-Oxy-2-acetylaminofluorene <i>Z</i> = 2.90, <i>H</i> =1.45bits | Active |

| | | | | |
|----|---|-------------|--|-----------------|
| | $Z = 2.80, H = 1.35bits$ | | | |
| 17 | Benz-(1,2)-anthracene $Z = 2.80, H = 0.97bits$ | Active | 4'-Oxybenz-(1,2)-anthracene $Z = 2.90, H = 1.15bits$ | Active |
| 18 | 9,10-Dimethylbenz-(1,2)-anthracene $Z = 2.67, H = 0.99bits$ | Active | 4'-Oxy-9,10-dimethylbenz-(1,2)-anthracene $Z = 2.76, H = 1.14bits$ | Active |
| 19 | Chrysene $Z = 2.80, H = 0.98bits$ | Active | 1'-Oxychrysene $Z = 2.90, H = 1.15bits$ | Active |
| 20 | <i>N,N'</i> -Dimethyl-4-aminoazobenzene (DAB) $Z = 2.59, H = 1.34bits$ | Active | <i>N</i> -Hydroxy-4-monomethylaminoazo-benzene $Z = 2.79, H = 1.47bits$ | Active |
| 21 | 7,12-Dimethylbenz(a) anthracene $Z = 2.67, H = 0.99bits$ | Very active | 7-Oxymethyl-12-methylbenz(a) anthracene $Z = 2.76, H = 1.14bits$ | Inactive, toxic |
| 22 | 4-Aminostilbene $Z = 2.64, H = 1.19bits$ | Active | <i>N</i> -Oxy-4-acetylaminostilbene $Z = 2.82, H = 1.42bits$ | Active |
| 23 | Bensidine $Z = 2.69, H$ | Active | 3,3'- Oxybensidine $Z = 2.93, H = 1.59bits$ | Active |
| | | | 4'-_____Acetylamino-4-aminobiphenyl | Active |

| | | | | |
|--|-----------|--|--|--------|
| | =1.32bits | | $Z = 2.77, H = 1.45bits$ | Active |
| | | | 4'-____Acetylamino-4-amino-3-oxybiphenyl $Z = 2.88, H = 1.54bits$ | |
| | | | 3-Oxybensidine $Z = 2.82, H = 1.49bits$ | |

Table 1. Carcinogenic activity, electronic descriptor and information functions for the initial chemical compounds and their metabolites.

The reference [10] indicates that the molecules of nitroso compounds are not "direct" carcinogens. Most nitroso compounds acquire the ability to form tumors as a result of metabolic transformations in the body. The method of quantitative evaluation of the carcinogenic activity of chemical compounds requires knowledge of only the initial molecular gross formula. For example, Table 2 shows the homologous series of nitrosomethylalkylamines. This table shows their comparative carcinogenic activity. The experiment was performed in rats [10].

| N | Compound ON-N-CH ₃ (CH ₂) _{m} CH ₃ | Gross formula | Activity (Act) | Z | $H, bits$ | π |
|-----|---|--|-------------------|------|-----------|-------|
| 1 | $m = 2$ | C ₄ H ₁₀ N ₂ O | 4 | 2.47 | 1.55 | 1.04 |
| 2 | $m = 3$ | C ₅ H ₁₂ N ₂ O | 4 | 2.40 | 1.49 | 1.56 |
| 3 | $m = 4$ | C ₆ H ₁₄ N ₂ O | 4 | 2.35 | 1.45 | 2.08 |
| 4 | $m = 5$ | C ₇ H ₁₆ N ₂ O | 4 | 2.31 | 1.41 | 2.60 |
| 5 | $m = 1$ | C ₃ H ₈ N ₂ O | 3 | 2.57 | 1.61 | 0.52 |
| 6 | $m = 6$ | C ₈ H ₁₈ N ₂ O | 3 | 2.28 | 1.37 | 3.12 |
| 7 | $m = 7$ | C ₉ H ₂₀ N ₂ O | 3 | 2.25 | 1.34 | 3.64 |
| 8 | $m = 8$ | C ₁₀ H ₂₂ N ₂ O | 3 | 2.23 | 1.32 | 4.16 |
| 9 | $m = 9$ | C ₁₁ H ₂₄ N ₂ O | 2 | 2.21 | 1.30 | 4.68 |
| 10 | $m = 10$ | C ₁₂ H ₂₆ N ₂ O | 2 | 2.20 | 1.28 | 5.20 |

| | | | | | | |
|----|----------|--------------------|---|------|------|------|
| 11 | $m = 11$ | $C_{13}H_{28}N_2O$ | 2 | 2.18 | 1.26 | 5.72 |
| 12 | $m = 12$ | $C_{14}H_{30}N_2O$ | 2 | 2.17 | 1.25 | 6.24 |

Table 2. Comparative carcinogenic activity [10] of homologous series of nitrosomethylalkylamines.

The parameter $\pi = \sum_m \pi_m$ is the sum of additional additive contributions to the hydrophobicity of the molecule; $\pi_m(\text{CH}_2) = \log(P_{\text{CH}_2}) - \log(P_H)$ (Table 2), where P_{CH_2} and P_H are, respectively, the partition coefficient for a substituted member and the parent member of a congeneric series. The index P_H is the distribution coefficient in the system of octanol-water for the original molecule. The magnitude of this index characterizes the ability of a molecule to penetrate into the body to the places of action. An additional contribution to the hydrophobicity of the molecule can be determined by the method of additive increments [11]. The contribution of each CH_2 fragment to the hydrophobicity of the molecule is assumed to be $\pi(\text{CH}_2) = 0.52$. As shown by the analysis, the molecular descriptors Z , H and π are interrelated (Fig. 1). Although the descriptors Z , H and π were derived from different principles, the relationship between the characteristics is very close. This relationship is close to functional dependence ($RMSE \ll 1$).

The increase in the length of the hydrocarbon side-chain is accompanied by an increase in the comparative carcinogenic activity (a four-point scale) of the chemical compound. A further increase in the side-chain leads to a decrease in carcinogenic activity (Fig. 2). That is, in this case, the hydrophobicity of molecules is the limiting factor of the carcinogenic activity of the agent. A similar limiting effect of hydrophobicity has been observed for radioprotectors (*N*-substituted *S*-2-aminoethylthiosulfates) [1]. It is important to note that all molecular descriptors are smaller than threshold values.

A

B

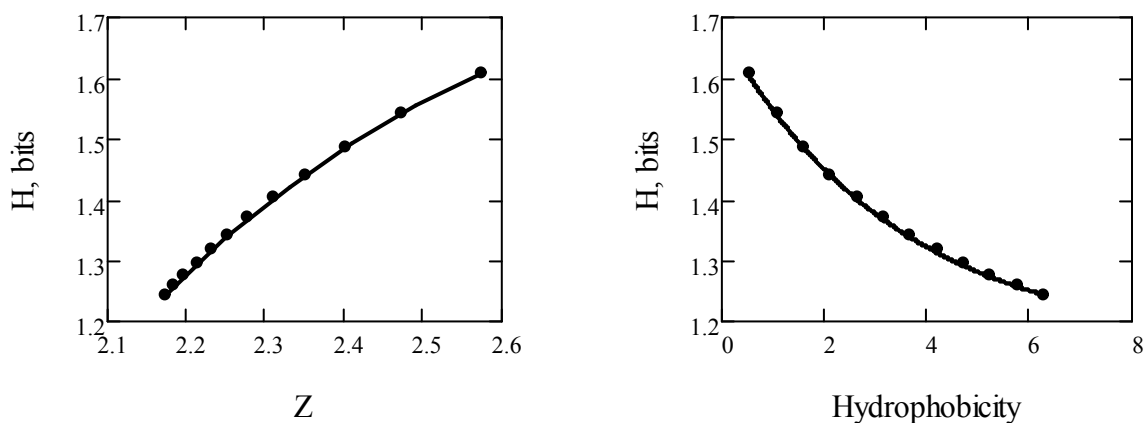


Fig.1. (A). Interrelation of molecular descriptors Z and H for a number of nitrosomethylalkylamines. (•) – the data in table 2. The regression line is defined by equation: $H(Z) = B + A \cdot \exp(-C \cdot Z)$, $A = -58.2 \pm 0.96$, $B = 1.89 \pm 0.01$, $C = 2.07 \pm 5.59$, $RMSE = 0.0002$.

(B) Interrelation of information function H of nitrosomethylalkylamines with their hydrophobicity. (•) – the data in table 2. The regression line is defined by equation: $H(\pi) = B + A \cdot \exp(-C \cdot \pi)$, $A = 0.55 \pm 0.01$, $B = 1.16 \pm 0.01$, $C = 0.29 \pm 0.01$, $RMSE = 0.002$.

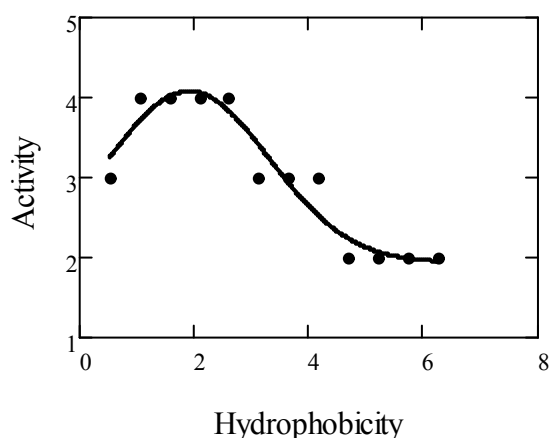


Fig.2. Interrelation of the carcinogenic activity of nitrosomethylalkylamines with the hydrophobicity parameter. (•) – comparative carcinogenic activity (Table 2). The envelope regression line is defined by equation: $Act(\pi) = D + A \cdot \exp(-(\pi - B)^2 / C^2)$, $A = 2.14 \pm 0.23$, $B = 1.91 \pm 0.15$, $C = -1.98 \pm 0.33$, $D = 1.94 \pm 0.21$, $RMSE = 0.28$.

It were examined [8] the agents similar in molecular structure to 2-acetylaminofluorene: 3-acetylaminodibenzothiophene, 3-acetylaminodibenzothiophene-5-oxide, 3-acetylaminodibenzofuran. All of them ended up being carcinogens, and 3-acetylaminodibenzothiophene is even more active carcinogen than 2-acetylaminofluorene.

We introduce an additional descriptor, that is, the information function of redundancy. The dimensionless redundancy information function is defined as follows:

$$D = 1 - H / H_{\max} . \quad (3)$$

Here $H_{\max} = \log_2(n)$, n is the number of different atoms in the molecule. Table 3 shows the row of chemical compounds close related structure 2-acetylaminofluorene.

| N | Chemical compounds | Gross formula | Activity | D | Z | H , bits |
|-----|---------------------------------------|---------------------|----------|------|------|------------|
| 1 | 3-Acetylaminodibenzothiophene | $C_{14}H_{13}NOS$ | +++ | 0.35 | 2.87 | 1.53 |
| 2 | 2-Acetylaminofluorene | $C_{15}H_{13}NO$ | ++ | 0.33 | 2.80 | 1.35 |
| 3 | 3-Acetylaminophenanthrene | $C_{16}H_{13}NO$ | + | 0.33 | 2.84 | 1.34 |
| 4 | 2-Acetylaminophenanthrene | $C_{14}H_{13}NO$ | + | 0.32 | 2.76 | 1.36 |
| 5 | 3-Acetylaminodibenzothiophene-5-oxide | $C_{14}H_{13}NO_2S$ | + | 0.30 | 2.97 | 1.62 |
| 6 | 3-Acetylaminodibenzofuran | $C_{14}H_{13}NO_2$ | + | 0.27 | 2.87 | 1.46 |
| 7 | 2-Aminofluorene | $C_{13}H_{10}N$ | + | 0.25 | 2.79 | 1.20 |
| 8 | 2-Aminoanthracene | $C_{14}H_{11}N$ | + | 0.25 | 2.77 | 1.19 |

Table 3. A number of aromatic amines close to 2-acetylaminofluorene.

The table 3 demonstrates correlation of the descriptors D , Z and H with carcinogenic activity value for related chemical compounds. The higher the values of the descriptors D , Z and H , the higher the carcinogenic activity. Descriptor D has a monotonous relationship with the carcinogenic activity of chemical compounds.

The reference [12] provides information on the carcinogenic activity (on a five-point scale) 1-vinyl-1-hydroperoxide cyclohexane-3 (activity is equal to 5 units) and 1-vinylcyclohexane-3 (activity is equal to 1 unit). For these chemical compounds, the following descriptor values have been obtained: $Z = 2.62$, $H = 1.34\text{bits}$, $D = 0.16$ and $Z = 2.20$, $H = 0.97\text{bits}$, respectively. Table 4 shows the activity of chemical compounds on a five-point scale.

| N | Chemical compounds | Gross formula | Activity | D | Z | H , bits |
|-----|---------------------------------------|-------------------|----------|-------------|------|------------|
| 1 | 1-Vinyl-1-hydroperoxide cyclohexane-3 | $C_8H_{11}O_2$ | 5 | 0.16 | 2.62 | 1.34 |
| 2 | 1-Ethyl-oxy-3,4-epoxycyclohexane | $C_8H_{12}O_2$ | 3 | 0.17 | 2.55 | 1.32 |
| 3 | 1,2- Epoxybutane | C_4H_6O | 1 | 0.17 | 2.55 | 1.32 |
| 4 | d1- Diepoxybutane | $C_4H_6O_2$ | 0 | 0.08 | 2.83 | 1.46 |
| 5 | Meso-diepoxybutane | $C_4H_6O_2$ | 0 | 0.08 | 2.83 | 1.46 |
| 6 | Benzene peroxide | $C_{14}H_{10}O_4$ | 0 | 0.10 | 3.21 | 1.43 |
| 7 | Styrene oxide | C_8H_8O | 0 | 0.20 | 2.70 | 1.26 |
| 8 | Lauroylperoxide | $C_{12}H_{23}O_3$ | 0 | 0.21 | 2.34 | 1.25 |
| 9 | 9,10- Epoxystearic acid | $C_{18}H_{33}O_3$ | 0 | 0.25 | 2.27 | 1.19 |
| 10 | 6,7,9,10- Epoxystearic acid | $C_{18}H_{32}O_4$ | 0 | 0.21 | 2.37 | 1.25 |
| 11 | Hexa-epoxysvalol | $C_{30}H_{54}O_4$ | 0 | 0.20 | 2.43 | 1.26 |

*) The bold type indicates the values of descriptor D that are higher than the values for active carcinogens.

Table 4. Carcinogenic activity of oxy compounds [12].

Chemical compounds numbers 8-10 have high values of descriptor D . The values of the descriptors Z and H for these agents are below the threshold values. However, these agents do not show carcinogenic activity. The absence of carcinogenic activity may be due to non-optimal hydrophobic properties of the molecules [13]. This is due to the presence of long hydrocarbon chains with a large number of atomic groups CH, CH₂ and CH₃. This range of values of the index m prevents the manifestation of biological activity if the number of groups $m > 10$ [1-3]. That is, the hydrophobic properties of molecules can limit their biological activity.

In this regard, we note the following relationship. Anticarcinogenic properties and selectively action agents on malignant cells have been studied [14]. The following agents have been analyzed: furfuryl-6-aminopurine ($Z = 3.20$), N^{NH_2} -puryl-6-tryptamine ($Z = 3.20$), N^{NH_2} -puryl-6-tyramine ($Z = 3.00$), N^{NH_2} -puryl-6-histamine ($Z = 3.15$), N^ϵ -puryl-6-lysine ($Z = 2.91$), N, N' -dipuryl-6-ethylenediamine ($Z = 3.24$). It is important to note that all these agents are characterized by a rather high Z descriptor value, which is higher than the mean values of the descriptors for active carcinogens.

It is suggested [15] that carcinogens induce DNA single-strand breaks. In this case, purine bases (especially guanine) are the target for them. In this regard, it should be noted that the molecular descriptor Z for all purine bases is larger than the threshold values. The descriptors maximum values are achieved for guanine ($Z = 3.50$, $H = 1.82\text{bits}$). For other purine bases the value of the molecular descriptor Z is also higher than the threshold values: adenine ($Z = 3.33$), guanine ($Z = 3.50$), thymine ($Z = 3.36$), cytosine ($Z = 3.23$), uracil ($Z = 3.5$).

It should be noted that the statistical approach is not sensitive in the analysis of isoelectronic molecular systems, as well as for isomer molecules. The statistical method gives the most satisfactory results in the analysis of homologous series of related compounds. Such series of chemical compounds are characterized by data sets that satisfy the statistical requirements for homogeneity and compatibility of the set elements.

References

- [1] Mukhomorov V.K. *Modeling of Chemical Compounds Bioactivity. Relationships of Structure - Bioactivity*. Lambert Academic Publisher. Saarbrücken. Germany. 2012. (in Russian).
- [2] Mukhomorov V.K. *Biomedical Statistics and Information*. 1, 24 (2016).
- [3] Mukhomorov V.K. *International Journal of Chemical and Biomedical Science*. 3, 34 (2017).

- [4] Shannon C. A mathematical theory of communication. *Bell Techn. Journal*. 27, 379 (1948).
- [5] Kolmogorov A.N. *Theory of information and theory of algorithms*. Nauka. Moscow. 1987. (in Russian).
- [6] Veljković V., Lalović D. *Experientia*. 33, 1228 (1977).
- [7] *Carcinogenic substances*. Handbook. Ed. Turusov V.S. (IARC Monographs on the Evaluation the Carcinogenic Risk of Chemicals to Humans). Moscow. 1987.
- [8] Badger G.M. *The Chemical Basis of Carcinogenic Activity*. Charles C. Thomas – Publisher: Springfield-Illinois-USA. 1962.
- [9] Rubenchik B.L. *Biochemistry of Carcinogenesis*. Kiev. 1977.
- [10] Rubenchik B.L. *Formation of carcinogens from nitrogen compounds*. Kiev.1990.
- [11] Leo A., Hansch C., Elkins D. *Chem. Reviews*. 71, 525 (1971).
- [12] Orris K., Yan Duuren B.L., Nelsen N. Carcinogenic activity of oxy compounds. In: *Proceedings of the VIII International of Anti-Cancer Congress*. 1963. Vol. 2. Moscow. pp.305-311.
- [13] Mukhomorov V.K. *Advances in Biological Chemistry*. 1, 1 (2011).
- [14] Hydvedy T.Y., Arky I., Antoni F., Kőteles G. In: *Proceedings of the VIII International of Anti-Cancer Congress*. 1963. vol.2. pp. 225.
- [15] Vilenchik M.M. *Regularities of the Molecular-Genetic effect of Chemical Carcinogens*. Moscow. 1977.