



Available online through

www.jbsoweb.com

ISSN 2321 - 6328

## Research Article

### AN ALTERNATIVE MODEL OF THE RELATIONSHIP BETWEEN THE CARCINOGENIC ACTIVITY OF POLYCYCLIC AROMATIC HYDROCARBONS AND THEIR MOLECULAR STRUCTURE

Vladimir Mukhomorov \*

University of Naples "Federico II", Naples, Italy

\*Corresponding Author Email: vmukhomorov@mail.ru

Article Received on: 01/07/18 Accepted on: 31/07/18

**DOI: 10.7897/2321-6328.06383**

#### ABSTRACT

A new model is proposed that connects the carcinogenic activity of polycyclic aromatic hydrocarbons with their molecular structure. The model does not use the representation of the molecule local regions existence responsible for carcinogenic activity. Molecular descriptors determine the electronic and information properties of molecules in general. Representations of the physics of condensed states and methods of information theory are used to determine molecular descriptors. To assess the carcinogenicity of a chemical compound, it is sufficient to know its gross formula. A statistically significant relationship between molecular descriptors and carcinogenic activity of polycyclic aromatic hydrocarbons has been established. The properties of molecules that limit the carcinogenic activity of molecules are considered. The model makes it possible to perform an express assessment of the carcinogenic activity of polycyclic aromatic hydrocarbons. To assess the carcinogenicity of a chemical compound, it is sufficient to know its gross formula. It is analyzed the interrelation between molecular descriptors and the physical properties (the first ionization potential of the molecule, the magnitude of the affinity energy, the energy of the most intensive electronic transition, the value of the pair intermolecular interaction, the electron-donor properties) of chemical compounds. Electronic and information descriptors are derived from various initial principles. However, it is established that the descriptors are closely related. We are demonstrating a comparison of the proposed model and the Pullman's model.

**Keywords:** Polycyclic hydrocarbons, Carcinogenicity, Electronic descriptor, Information function, Intermolecular interactions, Donor, Acceptor, Statistical confidence

#### INTRODUCTION

An important and prevalent class of potent carcinogenic compounds present in the environment is polycyclic aromatic hydrocarbons (PAHs), which are found in various petroleum and combustion products derived from heat and power generation and motor vehicle exhausts. Further more, PAHs are generally formed by pyrolysis of organic matters such as tobacco smoking and certain procedure of food preparation, the polycyclic aromatic hydrocarbons exposure to humans is extensive<sup>1</sup>. It is highly likely that this class of chemical compounds is important in human carcinogenesis<sup>2</sup>. PAHs have several features which distinguish them from many of the more discovered carcinogens. They act the site of application. The effective dose is minute, of the order of micrograms, and they have been found to induce tumours in almost every tissue and animal species in which they have been tested<sup>3</sup>. It was realized that PAHs as a group of potential carcinogens have many structural features in common but widely varying cancerogenic activities.

PAHs are involved in intermolecular interactions in the body. This is confirmed by the discovery of hydrocarbon - cancer tissue complexes. Like so many other substances, the carcinogens become firmly bound to proteins, but the importance of this binding for the specific carcinogenic action is not clear<sup>4</sup>. Detailed studies using chemical compounds labeled with radioactive carbon showed that the formation of such complexes plays a major role in the development of tumors. There is a parallelism between the frequency of occurrence of tumors and the size of the complex<sup>1</sup>. It has been established with high accuracy<sup>5</sup> that PAHs are attached to DNA, and

carcinogenic activity is proportional to the number of molecules of hydrocarbons involved in complex formation. For example, it is known that polycyclic hydrocarbons are capable to install into DNA. The use of radioactive indicators made it possible to observe the addition of polycyclic hydrocarbons to proteins and DNA. Moreover, the carcinogenic effect is proportional to the amount of bound hydrocarbon.

The attempts have been made on numerous occasions to correlate carcinogenic potency with the electronic properties of PAHs. For example, Chalvet and Daudel<sup>6</sup>, calculated the localization energy of hydrocarbons, interesting from the viewpoint of carcinogenic effect, with regard to hyperconjugation, using the parameters of Muller<sup>7</sup>. The most successful simulation was performed by Pullman A. and Pullman B.<sup>8,9</sup>. Subsequently, this model was developed in the works of other researchers, for example, in work<sup>10</sup>. In the majority of these models the carcinogenic action is related to the reactivity of the K- and L - regions of these molecules. The finding that osmium tetroxide adds to the 9,10-phenanthrene bond, which become known as the K - region was taken as a model for the attachment of cell constituent to the hydrocarbons. In a series of polycyclic hydrocarbons good correlation has been found between the calculated values for the bond order of the K - region and the rate of addition of osmium tetroxide. However, the correlation with the carcinogens activities of the polycyclic hydrocarbons was not good. The situation was not improved by more elaborate calculations<sup>3</sup>.

**METHOD AND ANALYSIS OF PAHs CARCINOGENIC ACTIVITY**

In this article, we propose an alternative model for interpreting the carcinogenic activity of PAHs. We use the electronic descriptor  $Z$  and the information function  $H$  for a quantitative description of the molecular structure of polycyclic hydrocarbons. These molecular descriptors characterize the molecule in general<sup>11-14</sup>, but not separate local regions of the molecule. The descriptor  $Z$  is often found in condensed matter physics. This descriptor is a multiplier in the effective potential (pseudopotential). The pseudopotential defines a field that acts on electrons. The descriptor  $Z$  determines the average number of valence electrons per atom in the molecule:  $Z = \sum n_i Z_i / N$ , where  $n_i$  is the number of atoms of the  $i$ -th type in the molecule with the number of electrons  $Z_i$  of the outer shell, and  $N$  is the number of atoms in the molecule. The pseudopotential method assumes that only electrons of outer shell are taken into account on the scattering center. It is well known that the chemical properties of molecules are determined by a relatively small group of outer shell electrons. The properties of the remaining electrons of the atom, which are called the electrons of the core, hardly change during the chemical processes the molecule involving. This approximation is called the "frozen-skeleton approximation". External electrons move do not in the real Hartree-Fock field of the molecule, but they move in a much weaker field of the pseudopotential.

The quantification of the descriptor  $H$  is carried out using through the Shannon information function<sup>15</sup>:  $H = -\sum p_i \log_2 p_i$ . The relation  $p_i = n_i / N$  satisfies the following conditions:  $0 \leq p_i \leq 1$ ,  $\sum p_i = 1$ ,  $\sum n_i = N$ . The ratio  $n_i / N$  determines the equity participation of the  $i$ -th type atom in the molecule.

Actually, to calculate the specific number  $p_i$ , we use Kolmogorov's combinatorial approach<sup>16</sup>. The information function makes it possible provides insight into the diversity of a multicomponent system. Information function  $H$  is used for a quantitative determination of the measure of organization or diversity of multi-component subsystems. Information function defines the diversity of a set of  $n$  elements. The smaller the information function value, the more diverse (in the relative content of atoms in molecules) a multicomponent system. The choice of the base of logarithm in equation was not a matter of principle importance.

Descriptors  $Z$  and  $H$  are set of indexes for the study of cancerogenic activity which have previously<sup>11</sup> been adopted for study of the relation between molecular structure of chemical compounds and their bioactivity. We can get some additional knowledge about the physical interpretation of the molecular descriptor  $Z$ . For this purpose, we will analyze the additive components of the total energy of the pair intermolecular interaction of polycyclic hydrocarbons with model molecular systems.

The fact that osmium tetroxide is attached at position 9,10 of the phenanthrene molecule (K-bond<sup>8</sup>) served as the basis for allowing the possibility of binding cyclic hydrocarbons to the active center of the organism precisely by this region of the molecule. Caillet and Pullman<sup>17</sup> carried out model studies of the additive components of the intermolecular interactions of polycyclic aromatic hydrocarbons with tetramethyl-uric acid. The table 1 demonstrates the following additive contributions to the pair interaction energy:  $E_{elect}$  is the electrostatic energy,  $E_{pol}$  is the polarization interaction energy,  $E_{disp}$  is the dispersion energy, and  $E_{rep}$  is the energy of the short-range exchange repulsion.

**Table 1. Additive components of the pair interaction energy of tetramethyl-uric acid with polycyclic hydrocarbons.**

N	Chemical compound	Brutto formula	Contributions to the energy of pair interaction, kcal/mol [17]					Z <sup>*</sup>	H <sup>*</sup> , bits
			E <sub>elect</sub>	E <sub>pol</sub>	E <sub>disp</sub>	E <sub>rep</sub>	E <sub>total</sub>		
1	Dibenzo-1,2,3,4-pyrene	C <sub>24</sub> H <sub>14</sub>	-1.93	-1.11	-10.37	4.95	-8.46	2.90	0.95
2	Anthanthren	C <sub>22</sub> H <sub>12</sub>	-1.34	-0.93	-8.72	4.79	-6.20	2.94	0.94
3	Perylene	C <sub>20</sub> H <sub>12</sub>	-1.35	-0.86	-7.80	4.25	-5.76	2.88	0.96
4	Benzo-1,2-pyrene	C <sub>20</sub> H <sub>12</sub>	-1.77	-0.96	-9.23	4.51	-7.45	2.88	0.96
5	Benzo-3,4-pyrene	C <sub>20</sub> H <sub>12</sub>	-1.38	-0.86	-8.64	4.48	-6.40	2.88	0.96
6	Pyrene	C <sub>16</sub> H <sub>10</sub>	-1.46	-0.78	-7.76	4.12	-5.88	2.85	0.96
7	Dibenzo-1,2,5,6-anthracene	C <sub>22</sub> H <sub>14</sub>	-1.09	-0.79	-8.23	3.88	-6.23	2.83	0.96
8	Chrysene	C <sub>18</sub> H <sub>12</sub>	-1.18	-0.80	-7.92	4.15	-5.75	2.80	0.97
9	Benzo-1,2-anthracene	C <sub>18</sub> H <sub>12</sub>	-1.18	-0.80	-7.92	4.15	-5.75	2.80	0.97
10	Phenantrene	C <sub>14</sub> H <sub>10</sub>	-1.57	-0.65	-7.13	3.77	-5.58	2.75	0.98
11	Anthracene	C <sub>14</sub> H <sub>10</sub>	-1.27	-0.79	-6.84	3.67	-5.22	2.75	0.98

<sup>\*)</sup> The elements of the sets  $Z$  and  $H$  satisfy the normal distribution. Statistics of Wilk-Shapiro is:  $W_Z = 0.94 > W_H = 0.92 > W_{0.95}^{(cr)}(f = 11) = 0.85$ .

Statistical analysis has shown that there is a reliable relationship between the descriptor  $Z$  with the contribution to the energy of short-range repulsion, the polarization contribution, and the dispersion interaction.

We obtained the following correlation equation for the relationship of the descriptor  $Z$  with the energy  $E_{rep}$ :

$$E_{rep}(Z) = a_0 + a_1 Z, \quad N = 11, \quad R = 0.88, \quad a_0 = -12.28 \pm 3.00, \quad a_1 = 5.82 \pm 1.06, \\ F = 30.3 > F_{0.05}^{(cr)}(f_1 = 1; f_2 = 9) = 5.1; \quad W_{E_{rep}} = 0.96 > W_{0.95}^{(cr)}(f = 11) = 0.85. \quad (1)$$

Here  $f$  is the number of freedom degrees. The statistical reliability of the correlation coefficient is confirmed by the inequality:  $t = 0.5 \cdot \ln[(1+|R|)/(1-|R|)] \cdot (N-3)^{1/2} = 3.89 > t_{0.95}^{(cr)}(f = N-2) = 2.26$ . The relationship between the contributions  $E_{pol}$  and  $E_{disp}$  with the molecular descriptor  $Z$  has the following statistics:

$$E_{pol}(Z) = a_0 + a_1 Z, \quad N = 11, \quad R = -0.77, \quad a_0 = 3.48 \pm 1.21, \quad a_1 = -1.52 \pm 0.43, \\ F = 12.7 > F_{0.05}^{(cr)}(f_1 = 1; f_2 = 9) = 5.1; \quad W_{E_{pol}} = 0.93 > W_{0.95}^{(cr)}(f = 11) = 0.85; \\ \text{statistical reliability of the correlation coefficient: } t = 2.85 > t_{0.95}^{(cr)}(f = 9) = 2.26. \quad (2)$$

$$E_{disp}(Z) = a_0 + a_1 Z, \quad N = 11, \quad R = -0.74, \quad a_0 = 25.3 \pm 10.3, \quad a_1 = -11.8 \pm 3.62, \\ F = 10.7 > F_{0.05}^{(cr)}(f_1 = 1; f_2 = 9) = 5.1; \quad W_{E_{disp}} = 0.95 > W_{0.95}^{(cr)}(f = 11) = 0.85; \\ \text{statistical reliability of the correlation coefficient: } t = 2.66 > t_{0.95}^{(cr)}(f = 9) = 2.26. \quad (3)$$

Apparently, the molecular descriptor  $Z$  is allowed intermolecular interaction to be identified and described in accordance with the specification short-range components of intermolecular interactions that determine the formation of complexes of organic substances with molecular systems that are simulating biosystems. It is important to emphasize that the relationship exists for the dispersion interaction, which gives the dominant contribution to the total interaction energy. As the

analysis showed, there is not a linear relationship for long-range electrostatic interactions. However, it is necessary to note the following important fact. The relationships (1) - (3) are not unexpected. Similar relationships were found for other classes of chemical compounds. For example, there is a close interrelationship between  $Z$  descriptor (and also  $H$  descriptor) and the absolute value of the dispersion interaction energy for halogenated compounds<sup>11</sup>:

$$E_{disp}(Z) = a_0 + a_1 Z, \quad N = 12, \quad R = 0.96, \quad a_0 = -1.10 \pm 0.95, \quad a_1 = 2.52 \pm 0.24, \\ F = 108.50 \gg F_{0.05}^{(cr)}(f_1 = 1; f_2 = 10) = 5.00; \quad W_{E_{disp}} = 0.95 > W_{0.95}^{(cr)}(f = 12) = 0.86; \text{ statistical reliability of the correlation} \\ \text{coefficient: } t = 5.83 > t_{0.95}^{(cr)}(f = 10) = 2.23. \quad (4)$$

Statistical analysis showed a close relationship between the molecular descriptors  $Z$  and  $H$ . However, polycyclic hydrocarbons have peculiarity in comparison with other chemical compounds of different classes:  $R$  is positive<sup>11</sup>. For PAHs, the relationship has a statistically significant opposite direction ( $R$  is negative):

$$H(Z) = a_0 + a_1 \cdot Z, \quad N = 11, \quad R = -0.96, \quad a_0 = 1.49 \pm 0.05, \quad a_1 = -0.186 \pm 0.02, \\ RMSE = 0.0035, \quad t(a_0) = 29.6 > |t(a_1)| = 10.48 > t_{0.05}^{(cr)}(f = 9) = 2.26, \\ F = 109.98 > F_{0.05}^{(cr)}(f_1 = 1, f_2 = 9) = 5.1; \text{ statistical reliability of the correlation coefficient: } t = 5.83 > t_{0.95}^{(cr)}(f = 10) = 2.23. \\ (5)$$

Statistical analysis confirms the opposite trend in the interrelationship between the descriptors, even if we use a wider range of PAHs ( $N = 41^8$ ,  $R = -0.996$ ,  $RMSE = 0.0012$ ). The set of elements  $Z$  and  $H$  satisfy the normal distribution. An increase the sample size leads to the appearance of a weak nonlinearity (Fig. 1A). In this case, the RMSE decreases to 0.0005. That is,

the interrelationship is approaching to a functional dependence. We emphasize that the descriptors had to be retrieved through different principles. The nonlinearity of the interrelationship occurs for most classes of chemical compounds<sup>11</sup>. Apparently, the low value of RMSE (5) is due to the homogeneity of the electron distribution for classical polycyclic hydrocarbons.

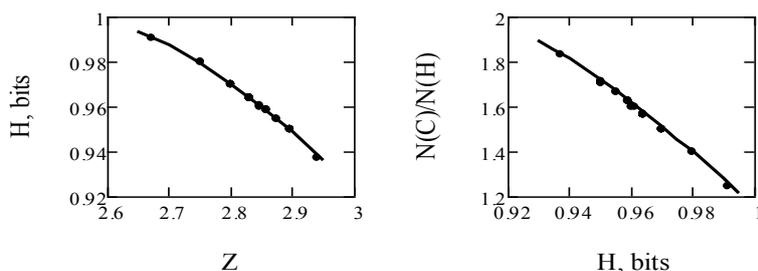


Fig.1. (A) The relationship between the descriptors  $H$  and  $Z$ . (B) The relationship between the descriptors  $\gamma = N(C)/N(H)$  and  $H$ .

It follows from the statistics (1) - (5) that the molecular descriptors  $Z$  and  $H$  are statistically significantly informative in order to characterize the pair intermolecular interaction. We will use these molecular descriptors to evaluate the carcinogenicity of polycyclic hydrocarbons.

The most active carcinogens are molecules that have the greatest value of the molecular descriptor  $Z$ . For example, molecules of dibenzo-3,4,9,10-pyrene and dibenzo-3,4,8,9-pyrene, for which descriptor  $Z$  is equal to 2.895. Taking into account the table 1 and the relationships (1) - (3), it can be assumed that molecules with a high value of  $Z$  are preferred in paired dispersion interactions that determine the attraction. Pair intermolecular Three rings: anthracene (– ;  $Z = 2.750$ ), phenanthrene (– ;  $Z = 2.750$ ); four rings: tetraphene (+/- ;  $Z = 2.800$ ), benzo(c)phenanthrene (+/- ;  $Z = 2.800$ ), pyrene (+/- ;  $Z = 2.846$ ); five rings: dibenz-1,2,7,8-anthracene (+ ;  $Z = 2.833$ ), dibenz-1,2,3,4-phenanthrene (+ ;  $Z = 2.833$ ), dibenz-1,2,5,6-anthracene (++) ;  $Z = 2.833$ ), benz-3,4-pyrene (++++ ;  $Z = 2.875$ ). (6)

We note that an increase in the descriptor  $Z$  is accompanied by an increase in the carcinogenic activity of polycyclic hydrocarbons for a number of compounds used in the model of the Pullmans. Pullmans proposed<sup>8,9</sup> to distinguish three groups of carcinogenic substances according to their relative activity from over 50 polycyclic hydrocarbons. The first group (inactive or weakly active) is benz-3,4-phenanthrene ( $Z = 2.80$ ,  $H = 0.97$ bits), dibenz-1,2,5,6-phenanthrene ( $Z = 2.83$ ,  $H = 0.96$ bits), dibenz-1,2,3,4-phenanthrene ( $Z = 2.83$ ,  $H = 0.96$ bits); the second group (active) is dibenz-1,2,5,6-anthracene ( $Z = 2.83$ ,  $H = 0.96$ bits), dibenz-1,2,7,8-anthracene ( $Z = 2.83$ ,  $H = 0.96$ bits); the third group (the most active carcinogens) is benz-3,4-pyrene ( $Z = 2.875$ ,  $H = 0.955$ bits), dibenz-1,2,3,4-pyrene ( $Z = 2.895$ ,  $H = 0.95$ bits), dibenz-3,4,8,9-pyrene ( $Z = 2.895$ ,  $H = 0.95$ bits), dibenz-3,4,9,10-pyrene ( $Z = 2.895$ ,  $H = 0.95$ bits). Thus, in this case we can also make a comparative assessment of the carcinogenicity for each group of polycyclic hydrocarbons using the molecular descriptor of  $Z$  (as well as the descriptor  $H$ ; see Fig. 1A). According to the level of carcinogenicity, three groups of compounds maintain their consistency. The greater the value of the descriptor  $Z$ , the more likely the intermolecular interaction is stronger.

Thus, the use of molecular descriptors  $Z$  and  $H$  does not contradict the criteria of carcinogenicity suggested by the Pullmans, which were called the K and L regions of the molecules<sup>9</sup>. Quantum chemical calculation on a variety of compounds led to the formulation of the K and L regions model by the Pullmans. The K region of polycyclic hydrocarbons carcinogenesis states that a potent carcinogen must possess a K region with high electron density. For example, dibenz-1,2,3,4-pyrene, dibenz-3,4,8,9-pyrene and dibenz-3,4,9,10-pyrene do not have the L region, but they have a suitable index of K region and therefore they are carcinogens. The molecular descriptors  $Z = 2.895$  and  $H = 0.950$ bits also indicate the carcinogenicity of these compounds. Descriptors not just indicate, but the

Anthracene ( $Z = 2.75$ ) > benz-1,2-anthracene ( $Z = 2.80$ ) > dibenz-1,2,5,6-anthracene ( $Z = 2.83$ ) > benz-3,4-pyrene ( $Z = 2.875$ ). (7)

Obviously, this sequence of fixations correlates with the value of descriptor  $Z$ . This gradation of the easiness of fixation is confirmed by experience. Photochemical oxidation very easily occurs in non-carcinogenic anthracene molecules ( $Z = 2.75$ ) and naphthacene ( $Z = 2.80$ ). Photooxidation becomes the more difficult with an increase of the number of side rings. Dibenz-1,2,5,6-anthracene molecule ( $Z = 2.83$ ) weakly react, and benz-3,4-pyrene molecule ( $Z = 2.875$ ) is completely inactive in the

interactions can only be a prerequisite for the formation of complexes. Strong covalent chemical interactions or interaction with transfer of charge might arise at the final stage of complex formation. This situation was indeed observed<sup>8</sup> in the complexation of dibenzo-1,2,5,6-anthracene with the cellular receptor.

Here we will analyze the experimental data that were used by Pullman A. and Pullman B.<sup>8,9</sup> for interpretation of carcinogenic activity of PAHs. Using the data of the Pullmans<sup>8,9</sup>, we make the following sequence of polycyclic hydrocarbons containing from 3 to 5 condensed rings:

descriptors correctly determine the sequence in the relative activity of this series of compounds. At the same time, compounds of small dimensions composed of less than four condensed benzene rings: benzene ( $Z = 2.50$ ,  $H = 1.00$ bits), naphthalene ( $Z = 2.67$ ,  $H = 0.99$ bits), anthracene ( $Z = 2.75$ ,  $H = 0.98$ bits), and phenanthrene ( $Z = 2.75$ ,  $H = 0.98$ bits) are inactive.

Pullmans also pointed out<sup>9</sup> that their criteria have exceptions for some polycyclic hydrocarbons. First, benz-3,4-phenanthrene, dibenz-1,2,5,6-phenanthrene and dibenz-1,2,3,4-phenanthrene do not subject to the criteria of the Pullmans<sup>9</sup>. Perhaps this is due to the inaccuracy of the established threshold values for the criteria of the L and K regions. As shown above, the use as a quantitative indicator of the molecular descriptor  $Z = 2.80$ - $2.83$  does not deny the weak carcinogenicity of these compounds. This is also confirmed by the test for the promotive activity of mice for benz-3,4-phenanthrene. The test was positive<sup>18</sup>. Secondly, according to the Pullmans<sup>9</sup>, anthanthrene should have carcinogenic activity. The value of descriptor  $Z = 2.94$  point to this. According to modern data<sup>18</sup>, anthanthrene is indeed of carcinogenic activity. At the same time, the strong carcinogen the dibenz-3,4,6,7-pyrene does not have K region. Therefore, this compound should not have carcinogenic activity. At the same time, the molecular descriptor for this polycyclic hydrocarbon is very high ( $Z = 2.895$ ) and hence this substance must be a carcinogen. This result does not contradict observations<sup>18</sup>.

We note further that the most studied addition reactions (according to the Pullmans<sup>9</sup>) occurring in the L region of the molecule are: a) fixation of maleic anhydride and b) photochemical oxidation. The descriptor  $Z$  (or the descriptor  $H$ ) enable tracking the facility of fixing maleic anhydride. Pullmans<sup>9</sup> indicated that the facility of fixation is consistently decreasing in the following series of compounds:

photooxidation reaction. Methyl substitution of the hydrocarbon increases the reactivity of the chemical compound. Thus, for example, 9,10-dimethylbenzanthracene ( $Z = 2.67$ ,  $H = 0.99$ bits) is much more active than the initial hydrocarbon benzanthracene ( $Z = 2.80$ ,  $H = 0.97$ bits), in reactions with maleic anhydride and reaction of photooxidation.

These experimental facts also correlate with the variation of the molecular descriptor  $Z$  (as well as descriptor  $H$ ). Indeed, the molecule of benz-1,2-pyrene ( $Z = 2.875$ ) possesses a high carcinogenic activity<sup>19</sup>, as well as the following polycyclic hydrocarbons (in order of increasing activity): dibenz-1,2,3,4-pyrene ( $Z = 2.895$ ), dibenz-3,4,8,9-pyrene ( $Z = 2.895$ ), dibenz-3,4,9,10-pyrene ( $Z = 2.895$ ), and also ovalene ( $Z = 3.087$ ). The ovalene has pronounced carcinogenic, mutagenic and teratogenic properties. As demonstrated in article<sup>20</sup>, the dibenz-3,4,6,7-pyrene is very active (no less active than the benz-3,4-pyrene or the dibenz-3,4,8,9-pyrene). The dibenz-3,4,6,7-pyrene

has the descriptor  $Z = 2.895$  (activity: ++/++++). At the same time, it is important to note that dibenz-3,4,6,7-pyrene does not have a sufficiently active  $K$  region. Consequently, this polycyclic compound is not active from the position of the Pullmans model.

The addition reaction of osmium tetroxide was investigated in detail to study the reactivity of the  $K$  region. The following sequence of polycyclic compounds has been drawn up<sup>8</sup> in order of increasing their reactivity with respect to this agent:

Phenanthrene ( $Z = 2.75$ ;  $I = 7.75$  eV;  $-$ ) < benz-1,2-anthracene ( $Z = 2.80$ ;  $I = 7.50$  eV;  $+/-$ ) < dibenz-1,2,5,6-anthracene ( $Z = 2.83$ ;  $I = 7.09-7.80$  eV;  $++$ ) < benz-3,4-pyren ( $Z = 2.875$ ;  $I = 7.19$  eV;  $++++$ ). (8)

It is not difficult to see that in this case there is also a parallelism of the reactivity with the magnitude of the descriptor  $Z$ . The relative carcinogenic potencies of molecules increase in the same sequence. Here we give the values of the first ionization potentials ( $I$ ) of the molecules<sup>21</sup>. The increase in descriptor  $Z$  has been accompanied by a decrease in the ionization potential of the molecule for a given series (8) of polycyclic hydrocarbons. There is the proportionality (8) between the magnitude of descriptor  $Z$  and the magnitude of the first ionization potential of PAHs. It can be assumed that the lower the ionization energy of the molecule, the better the donor properties of the molecule. The increase in the descriptor  $Z$  often coincides with the dynamics of the increase in the donor ability of molecules. Obviously, for the sequence of observations (63), the descriptor  $Z$  and the associated descriptors  $H$  and  $I$  are reliably informative. The variation of a number of chemical compounds (8) indicates the existence of a relationship between the reactivity of molecules and the magnitude of molecular descriptors.

We can be taken the descriptor value  $Z \approx 2.83$  as the boundary value in accordance with the sequence (8). This value approximately separates carcinogens from inactive polycyclic hydrocarbons. In this connection, we note that two dibenzophenanthrenes XIII and XIV (the numbering of the Pullmans<sup>9</sup>) does not contain the reaction  $K$  region. Consequently, they cannot be attributed to the carcinogens. However, the molecular descriptor ( $Z = 2.83$ ) indicates that these chemical compounds may have a weak carcinogenic activity. This is confirmed by observations.

Molecular descriptors are also useful for quantitatively assessing the reactivity of the  $L$  region of molecules. Observation have shown<sup>22</sup> that the most potently active carcinogenic hydrocarbons are remarkably susceptible to oxidation by lead tetraacetate, that is there is a parallelism between carcinogenic activity and this form of chemical reactivity. The sequence order<sup>9</sup> of decreasing reactivity towards  $Pb(OAc)_4$  is:

Anthracene ( $Z = 2.75$ ,  $H = 0.98$ bits;  $-$ ) > benz-1,2-anthracene ( $Z = 2.80$ ,  $H = 0.97$ bits ( $-/+$ )) > dibenz-1,2,5,6-anthracene ( $Z = 2.83$ ,  $H = 0.96$ bits ( $+$ )) = 0. (9)

Molecular descriptors  $Z$  and  $H$  indicate their parallelism with the reactivity of the  $L$  region.

Another molecular descriptor can be proposed for the purpose of rapid assessment of the carcinogenic activity of PAHs. This descriptor is determined by the ratio of the number of carbon atoms  $N(C)$  to the number of hydrogen atoms  $N(H)$ :  $\gamma =$

$N(C)/N(H)$ . As shown by statistical analysis, the descriptor  $\gamma$  is closely related to the molecular descriptors  $Z$  and  $H$  (Fig. 1B). There is a statistically significant close relationship between the descriptors for the 41 polycyclic hydrocarbons represented by the Pullmans<sup>9</sup>. For example, the relationship between the information function  $H$  and the descriptor  $\gamma$  can be approximated by a linear form:

$$\gamma(H) = a_0 + a_1 \cdot H, \quad R = -0.997, \quad N = 41, \quad RMSE = 0.0079, \quad a_0 = 11.86 \pm 0.13, \\ a_1 = -10.67 \pm 0.14, \quad |t(a_1)| = 78.6 \gg t_{0.05}^{(cr)}(f=34) = 2.03. \quad (10)$$

The increase in the sample size leads to the appearance of a statistically significant weak nonlinearity for the relationship between the descriptors of  $\gamma$  and  $H$  (Fig. 1B). A similar close relationship exists between the descriptors  $Z$  and  $\gamma$ . It is important to note that the molecular descriptors that were obtained on the basis of different baseline principles are closely interrelated. The application of these molecular descriptors leads to identical results.

following sequence: phenanthrene ( $\gamma = 1.40$ ) < benz-1,2-anthracene ( $\gamma = 1.50$ ) < dibenz-1,2,5,6-anthracene ( $\gamma = 1.57$ ) < benz-3,4-pyrene ( $\gamma = 1.67$ ) < dibenz-3,4,6,7-pyrene ( $\gamma = 1.71$ ) < valen ( $\gamma = 2.29$ ). At the same time, the molecular descriptor  $\gamma$  is obviously related to the amount of carcinogenic activity: an increase in the carcinogenicity of the substance is accompanied by an increase in the value of  $\gamma$  (as well as the descriptor  $Z$ ).

The above experimental facts can also be interpreted using the molecular descriptor  $\gamma$ , which is in no way connected with the  $K$  and  $L$  regions of the polycyclic molecule. For example, the carcinogenicity of chemical compounds increases in the

It is of fundamental importance that we do not use the idea of special local regions of molecules (that is, the  $K$  and  $L$  regions). On the contrary, we use molecular descriptors that characterize the properties of molecules in general. These descriptors make it possible to successfully interpret the observations that underlie

the model of the Pullmans. Obviously, within the framework of our model, it is possible to remove the following important remark made by Ladik: "... why only the hydrocarbons that have joined the *K* region are carcinogenic, while the hydrocarbons that have joined in position *L* are inactive"<sup>23</sup>.

From point of view of the Pullmans model it is extremely puzzling<sup>9</sup> to observe that, whereas benzo-3,4-pyrene is cancerogen, the 2' and 3' methyl derivatives are total inactive. However, the descriptors  $\gamma$  and  $Z$  for benz-3,4-pyrene are 1.67 and 2.88, and for methyl derivatives are 1.50 and 2.80, respectively. That is, the magnitude of the descriptors decreases. A similar decrease in descriptors occurs for inactive 6,7-dimethyl-3,4-benzphenanthrene ( $\gamma = 1.25$ ,  $Z = 2.67$ ) and 4,5-dimethylchrysene ( $\gamma = 1.13$ ,  $Z = 2.59$ ). This does not contradict the condition established above, which separates carcinogens from non-carcinogenic compounds. However, defining the link between molecular descriptors and the activity of substituted polycyclic hydrocarbons with various substituents requires additional studies.

## DISCUSSION

Some polycyclic hydrocarbons may have a high  $Z$  descriptor value, but they do not have carcinogenic activity. This may be due to the limiting factors of the manifestation of biological activity<sup>11</sup>. For example, the molecular descriptors  $Z$ ,  $H$  and  $\gamma$  do not impossible to distinguish isomer molecules. The benz-3,4-pyrene ( $Z = 2.875$ ) is very active, and benzo(e)pyrene ( $Z = 2.875$ ) is weakly active. Both molecules are alternant hydrocarbons. However, the affinity energy  $A$  for these

Benz-3,4-pyrene ( $Z = 2.875$ ;  $A = 0.77$  eV; ++++); dibenz-1,2,5,6-anthracene ( $Z = 2.833$ ;  $A = 0.68$  eV; ++); dibenz-1,2,5,6-phenanthrene ( $Z = 2.833$ ;  $A = 0.31$  eV; +); dibenz-1,2,3,4-phenanthrene ( $Z = 2.833$ ;  $A = 0.40-0.43$  eV; +); dibenz-1,2,7,8-anthracene ( $Z = 2.833$ ;  $A = 0.23-0.69$  eV; +/-); benz-1,2-pyrene ( $Z = 2.875$ ;  $A = 0.49-0.61$  eV; -); dibenz-1,2,3,4-anthracene ( $Z = 2.833$ ;  $A = 0.22-0.54$  eV; -).  
(11)

The experimental values of the affinity energy have been taken from the reference book<sup>21</sup>. To this series of compounds, the following polycyclic hydrocarbon can be added: dibenz-3,4,6,7-pyrene ( $Z = 2.895$ ,  $A = 0.82$  eV; ++++). The lower bounds of the affinity energy (11) indicate their parallelism with the carcinogenic activity of the molecules. It is possible that activity pyrene ( $Z = 2.846$ ; +/-) is limited by the value of the affinity energy of the molecule ( $A \geq 0.09$  eV).

Pullman pointed out a number of very active polycyclic hydrocarbons, which contain six condensed rings: dibenz-3,4,8,9-pyrene (++++), dibenz-3,4,9,10-pyrene (++++) and dibenz-1,2,3,4-benzpyrene (+++). For these compounds, the descriptor  $Z$  is equal to 2.895, and the energies of the affinity are 1.1 eV, 0.80-0.84 eV, and 0.80-0.87 eV, respectively<sup>21</sup>. That is, these molecular descriptors do not contradict our model. According to limited data<sup>18</sup> anthanthrene is also the carcinogen:  $Z = 2.941$ ,  $A = 1.0$  eV. In this respect, dibenz-3,4,8,9-pyrene is very indicative, which has a very high carcinogenic activity. According to the classification given by the Dyachkov<sup>25</sup>, the activity of this compound is defined as (++++). This molecule has a high molecular descriptor  $Z = 2.895$  and a high affinity energy  $A = 1.1$  eV<sup>21</sup>. At the same time, coronene, whose molecule has a very high descriptor  $Z = 3.00$ , does not have carcinogenic activity. For this chemical compound, the affinity energy is comparatively low:  $A = 0.30$  eV<sup>21</sup>. This may reduce the carcinogenic activity of the chemical compound. Apparently, the requirement  $Z \geq 2.83$  is a necessary, but not sufficient

condition for the carcinogenicity of classical PAHs. The molecule must also have the high affinity energy. At the same time, only high affinity energy does not guarantee the presence of carcinogenic properties in the molecule, if the descriptor  $Z$  is near the boundary value. For example, benz(a)tetracene ( $Z = 2.83$ ,  $A = 1.1$  eV), pentaphene ( $Z = 2.83$ ,  $A = 1.2$  eV) and naphthacene ( $Z = 2.80$ ,  $A = 0.88$  eV) have high affinity energy but do not have a carcinogenic activity. A similar situation is observed for inactive dibenz-1,2,7,8-naphthacene, which has six condensed rings. However, in this case the descriptor  $Z$  has a very low value of 2.814, although the affinity energy is quite high 0.9 eV<sup>21</sup>. Unfortunately, there is a very limited number of classical polycyclic hydrocarbons for which experimental the affinity energy data are known.

We can more clearly identify the role of the molecule in the transfer of an electron if the object with which the molecule interacts is accurately known. The transfer of an electron is usually accompanied by endoenergetic or exoenergetic processes that can affect the state of the objects with which the polycyclic hydrocarbon interacts.

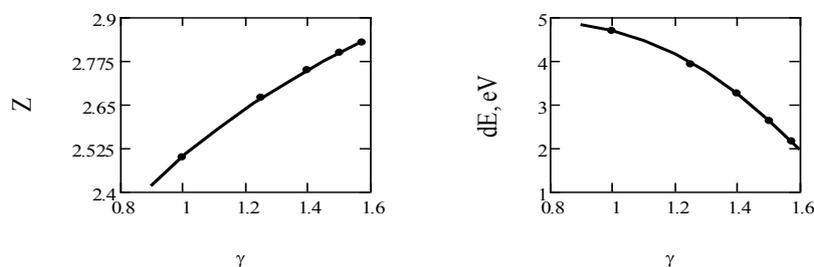
Since the carcinogen molecule interacts with molecular structures, in particular RNA and DNA<sup>26</sup>, the locally released energy during electron transfer can be directed to the destruction of these structured molecular systems. The concentration of hydrocarbons<sup>5</sup> in the interaction region significantly enhances the effect of the released energy.

The addition of an alkyl group to a polycyclic hydrocarbon may limit the carcinogenicity of the substance<sup>22</sup>. First, it reduces the molecular descriptors of  $Z$  and  $\gamma$ . Secondly, an increase in the length of the alkyl group is accompanied by a change in the hydrophobic properties of the substance. This may impede the development of the biological to be activity of the chemical compound<sup>11</sup>. According to the Pullmans, steric hindrances can also be limiting factors. Shape, size, steric and solubility factors of the compounds are of importance for cancerogenic activity, but in the last resort this depends on the recipient tissue. It depends on the species, the particular strain of the test animal, its age, sex, its nutritional and hormonal state, and the phase of mitotic activity of a particular cell<sup>3</sup>.

In accordance with the model of the Pullmans, the chemical compounds LVIII and LIX (the numbering of the Pullmans<sup>9</sup>) are large spatial size and should be carcinogens. The same conclusion follows from calculations of molecular descriptors, which are  $Z = 2.96$  and  $3.00$ ,  $\gamma = 1.88$  and  $2.00$ , respectively. However, it was not possible to find the results of the pilot testing of these substances.

## CONCLUSION

Thus, two different models based on the different principles produced close results (considered comparable), or that any differences found could be explained. At the same time, it is not necessary to perform complex and labor-intensive quantum chemical calculations. It is enough to know only the gross formula of hydrocarbon. At the same time, to determine the quantitative characteristics of the K and L regions, knowledge of the various quantum chemical parameters of the molecules is



**Fig. 2. (A) The relationship between descriptors  $Z$  and  $\gamma$ . Statistics of the interrelation:  $Z(\gamma) = B + A \cdot \exp(-C \cdot \gamma)$ ,  $B = 3.23 \pm 0.03$ ,  $A = -2.09 \pm 0.04$ ,  $C = 1.05 \pm 0.06$ ,  $R^2 = 0.99$ ,  $RMSE = 0.0012$ . (B) The relationship between descriptors  $\Delta E$  and  $\gamma$ . Statistics of the interrelation:  $\Delta E(\gamma) = A \cdot \gamma^2 + B \cdot \gamma + C$ ,  $B = 7.85 \pm 0.99$ ,  $A = -4.78 \pm 0.39$ ,  $C = 1.63 \pm 0.63$ ,  $R^2 = 0.99$ ,  $RMSE = 0.028$ . Here we use the following notation:  $dE \equiv \Delta E$ .**

Qualitative observation have shown that the most potently active carcinogenic hydrocarbons are remarkably susceptible to oxidation by lead tetraacetate is a rough correlation between carcinogenic activity and this form of chemical reactivity. It is important to emphasize that all the molecular descriptors used in this work ( $Z$ ,  $H$ ,  $\gamma$ ,  $A$ ,  $I$ ,  $\Delta E$ ,  $\lambda$ ) characterize the molecule as a whole. These descriptors have been useful for interpretation of the cancerogenic activity of polycyclic molecules, so that their application to the study of the relationship between electronic structure and activity was quite justified. The principles of this model may be easily understood, and model will necessary in

$\text{CH}_3\text{SOCH}_3$  ( $Z = 2.60$ ) >  $\text{HCON}(\text{CH}_3)_2$  ( $Z = 2.50$ ) >  $[(\text{CH}_3)_2\text{N}]_3\text{PO}$  ( $Z = 2.35$ ) >  $\text{CH}_3\text{OCH}_2\text{CH}_2\text{OCH}_3$  ( $Z = 2.57$ ) > tetrahydrofuran ( $Z = 2.30$ ) >  $\text{H}_3\text{C}_2\text{OCH}_2\text{CH}_2\text{OC}_2\text{H}_5$  ( $Z = 2.27$ ) >  $(\text{C}_2\text{H}_5)_2\text{O}$  ( $Z = 2.14$ ).

The donor ability reaches a minimum when the descriptor value  $Z$  for  $(\text{C}_2\text{H}_5)_2\text{O}$  practically coincides with the value  $Z = 2.17$  for the organometallic compound  $(\text{C}_4\text{H}_9)_2\text{SnCl}$ .

required. The molecular descriptors proposed here may give new information for the model of Pullmans, but they can also be used as independent indicators of the carcinogenicity of classical polycyclic hydrocarbons. Molecular descriptors, on the one hand, can be used for rapid assessment of the carcinogenicity of polycyclic aromatic hydrocarbons, and on the other hand, the physical meaning contained in these descriptors allows us to point out the physicochemical processes in which bioactive molecules may participate. It may be useful to stress the simple way in which the model explains the activity and inactivity of a closely related compounds.

It is also useful to emphasize the high information content of the  $Z$  descriptor. For example, table 2 shows the experimental values of the energy  $\Delta E$  of the most intense electronic transition<sup>27</sup> for several aromatic hydrocarbons.

**Table 2. The energy of the most intense electronic transition  $\Delta E$  and the molecular descriptors of polyacenes**

Chemical compound	$\Delta E$ , eV [27]	$Z$	$H$ , bits	$\gamma$
Benzene	4.70	2.50	1.00	1.00
Naphthalene	3.94	2.67	0.99	1.25
Anthracene	3.27	2.75	0.98	1.40
Naphthacene	2.63	2.80	0.97	1.50
Pentacene	2.16	2.83	0.96	1.57

Figure 2 shows, that energy  $\Delta E$  correlates with the molecular descriptor  $\gamma$ . Since the descriptor  $\gamma$  is closely related to the descriptor  $Z$  (Fig. 2A), the energy of the electronic transition also obviously correlates with the value of the molecular descriptor  $Z$ . The smallness of the RMSE value indicates that this relationship is approaching a functional relationship.

understanding the mechanism of carcinogenic action of polycyclic aromatic hydrocarbons.

We can also note the high information content of the  $Z$  descriptor. This descriptor correlates with electron donor activity not only for classical polycyclic hydrocarbons, but also for other classes of compounds. The quadrupole splitting of the gamma-resonance line of tin in the Mössbauer spectrum of tin dibutyl chloride was studied<sup>28</sup>. For seven solvents had been obtained the following order of a decrease in the donor properties of solvents towards to the organometallic compound  $(\text{C}_4\text{H}_9)_2\text{SnCl}$ :

The change in the carcinogenic properties of molecules with a change in the molecular potential does not contradict the known notions of the mechanism of chemical carcinogenesis. The known data<sup>29</sup>, as well as quantum-chemical calculations<sup>30</sup>, allow

us to conclude that, at least in a series of close congener chemical compounds, their carcinogenicity is directly dependent on the ability to electrophilic attack.

It is suggested<sup>29</sup> 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 = 1.82bits). For other purine bases the values of the molecular descriptors Z and H are also higher than the threshold values: adenine (Z = 3.33; H = 1.59bits), guanine (Z = 3.50; H = 1.82bits), thymine (Z = 3.20; H = 1.83bits), cytosine (Z = 3.23; H = 1.83bits), uracil (Z = 3.50; H = 1.92bits)<sup>14</sup>.

## REFERENCES

- Gäslund A., Jernström B. "DNA-carcinogen interaction: covalent DNA – adducts of benzo(a)pyrene 7,8-dihydrodiol 9,10 – epoxides studied by biochemical and biophysical techniques". *Quarterly Reviews of Biophysics*, 1989; 22: 1-37.
- Evaluation of Carcinogenic Risk. 1972, 1983. International Agency for Research on Cancer Monograph.
- Schoental R. "Carcinogenic by Polycyclic Aromatic Hydrocarbons and by Certain Other Carcinogens". In: Clar E. *Polycyclic Hydrocarbons*. Academic Press. London and New York. Springer Verlag. Berlin-Göttingen-Heidelberg. vol.1, 1964.
- Heidelberger C. In: Ciba Foundation Symposium on "Carcinogenesis: Mechanism of Action", p.179, Churchill, London. 1959.
- Brookes P., Lawley P.D. "Evidence for the Binding of Polynuclear Aromatic Hydrocarbons to the Nucleic Acids of Mouse Skin: Relation between Carcinogenic Power of Hydrocarbons and their Binding to Deoxyribonucleic Acid" // *Nature*, 1964; 202: 781-784 (1964). <https://www.nature.com/articles/202781a0>.
- Chalvet O., Daudel R. "Contribution a l'étude du mecanisme d'action des hydrocarbures carcinogens" // *Compt. Rend.*, 1956; 242: 416-418.
- Muller N., Pickett L.C., Mulleken R.C. "Hyperconjugation and Spectrum of the Benzenium Ion Phototype of Aromatic Carbonium Ions" // *J. Am. Chem. Soc.* 1956; 76: 4770-4778.
- Pullman B. *La Biochimie Electronique*. Press Universitaires de France. 1963.
- Pullman A., Pullman B. "Electronic Structure and Carcinogenic Activity of Molecules of Aromatic Substances" In: *Advances in Cancer Research*. Eds. by Greenstein J., Haddow A. Vol. 3. 1955. Acad. Press. Inc. Publisher. New-York.
- Koutecky J., Zahradnik R. "On the Problem of the Connection between the Electronic Structure of Polynuclear Aromatic Hydrocarbons and their Carcinogenic Effect". // *Cancer Research*, 1961; 21: 457-462.
- Mukhomorov V.K. *Modeling of Chemical Compounds Bioactivity. Relationships of Structure - Bioactivity*. Lambert Academic Publisher. Saarbrücken. Germany. 2012. (in Russian).
- Mukhomorov V.K. "Statistical Aspects of the Interrelation Between the Biological Activity of Chemical Compounds and their Molecular Structure". // *Biomedical Statistics and Informatics*, 2016; 1: 24-34. <http://www.sciencepublishinggroup.com/j/bsi>
- Mukhomorov V.K. "Statistical Modeling and Prediction of Carcinogenic Activity of Chemical Compounds". // *Intern. J. of Chem. and Biomedical Science*, 2017; 3: 39-63. <http://www.aascit.org/journal/ijcbs>
- Mukhomorov V.K. "Interrelation of Carcinogenic Properties of Metabolites with their Information and Electronic Descriptors", *SCIREA Journal of Biology*, 2018; 3: 1-14. <http://www.scirea.org/journal/Biology>
- Quastler G. "The ABC of Information Theory" In: *Theory of Information in Biology*. Ed. Blumenfeld L.A. Moscow.1960.
- Kolmogorov A.N. *Information Theory and Theory of Algorithms*. Nauka. Moscow.1987 (in Russian).
- Caillet J., Pullman B. *Molecular Association in Biology* (Pullman B. ed.), Academic Press, 1968, 217 p.
- Carcinogenic substances. Handbook. Ed. Turusov V.S. (IARC Monographs on the Evaluation the Carcinogenic Risk of Chemicals to Humans). Moscow. 1987.
- Badger G.M. *The Chemical Basis of Carcinogenic Activity*. Charles C. Thomas – Publisher: Springfield-Illinois-USA. 1962.
- Vylegzhanin N.I. "The experience of studying the blastomogeneous properties of 3, 4, 6, 7-dibenzopyrene" // *Oncology Affairs*. 1952; Issue 5:135-141 (in Russian).
- Gurvich L.V., Karachevtsev G.V., Kondratiev Yu. A. *The Energy of Rupture of Chemical Bonds. Potentials of Ionization and Electron Affinity* M. Nauka, 1974 (in Russian).
- Fieser L.F., Putnam S.T "Rate of Oxidation of Aromatic Hydrocarbons by Lead Tetraacetate". // *J. Am. Chem. Soc.*, 1947; 69: 1041-1046.
- Ladik J. *Quantenbiochemie für Chemiker und Biologen*. Akadimiai Kiado. Budapest. 1972.
- <http://chem21.info/info/404503/>
- Dyachkov P.N. "Quantum-chemical Calculations in the Study of the Mechanism of Action and Toxicity of Alien Substances". *The Outcomes of Science and Technology. Series of Toxicology*. Vol. 16. Moscow. 1990. 280 pp. (in Russian).
- Heidelberger C., Moldenhauer M.G. "The Interaction of Carcinogenic Hydrocarbons with Tissue Constituents". // *Cancer Research*, 1956; 16: 442-449.
- Peacock T.E. *Electronic Properties of Aromatic and Heterocyclic Molecules*. Acad. Press. London and New York. 1965.
- Goldanskii, V.I. "Zur Gamma-Resonanzspektroskopie (Mösbauerspektroskopie) in der Chemie". // *Angew. Chem.*, 1967; 79: 844-858.
- Vilenchik M.M. *Regularities of the Molecular-Genetic effect of Chemical Carcinogens*. Moscow. 1977 (in Russian).
- Pullman B., Pullman A. *Quantum Biochemistry*. 1965.

## Cite this article as:

Vladimir Mukhomorov. An alternative model of the relationship between the carcinogenic activity of polycyclic aromatic hydrocarbons and their molecular structure. *J Biol Sci Opin* 2018;6(3): 58-65.  
<http://dx.doi.org/10.7897/2321-6328.06383>

Source of support: Nil; Conflict of interest: None Declared

Disclaimer: JBSO is solely owned by Moksha Publishing House - A non-profit publishing house, dedicated to publish quality research, while every effort has been taken to verify the accuracy of the contents published in our Journal. JBSO cannot accept any responsibility or liability for the site content and articles published. The views expressed in articles by our contributing authors are not necessarily those of JBSO editor or editorial board members.