Simulation of the Radioprotective Action of Mercaptoethylamine Derivatives and its Analogues with their Quantum-Chemical and Information Features

MUKHOMOROV V.K.
Physical Department
State Polytechnic University
St. Petersburg
RUSSIA

Abstract: - Quantum chemistry, condensed matter physics and applied information theory methods are used to reveal the relationship between the molecular structure of radiation injury modifiers and their radioprotective activity in the series of aminothiols and their analogues. Significant electronic and informational parameters of molecules, which are associated with the radioprotective effect of drugs, were determined by statistical analysis methods. Based on the identified significant molecular parameters, possible mechanisms of the biochemical and biophysical radioprotective action of the analyzed chemical compounds are discussed. The detected significant molecular parameters suggest what possible molecular processes these drugs can take part in and what electronic and informational properties radioprotectors molecules should possess.

Key-Words: - Aminothiols, radioprotectors, electronic energy, threshold, pseudopotential, dipole moment, information function, modeling

Received: May 30, 2022. Revised: October 11, 2022. Accepted: November 14, 2022. Published: December 31, 2022.

1 Introduction

The problem of changing the body's radio-sensitivity through the use of various chemical compounds continues to be one of the most topical and intensively developed in modern radiobiology. The study of molecular mechanisms of action of radiation lesion modifiers is of fundamental importance for understanding the triggering effects of radiation and mechanisms of radiation protection. At the same time, deciphering the molecular mechanisms of radiation exposure opens up the prospect of new approaches to the search for effective radio-protective agents. In this connection, of considerable interest are studies concerning the connection between the radioprotective effect of drugs and the electronic structure of molecules and their informational content. Aminothiols are of great interest due to their diverse applications in medicine and organic chemistry. Aminothiols are active fungicides and have an antibacterial effect, exhibit herbicidal activity and antidote properties, as well as antihemolytic and hypotensive effects [1]. In this article, quantitative relationships will be obtained between the features that determine the energy and information properties of the molecules of mercaptoethylamine derivatives and its analogues and their radio-protective effect. The antiradiation properties of these preparations have been studied experimentally in sufficient detail.

2 Problem Formulation

Obviously, modification of the basic molecular structure is one of the ways to influence the molecular factors that determine the radioprotective efficacy of the drug. It is of some interest to reveal the cause-and-effect relationship between the radio-protective activity of molecules of a number of mercaptoethylamine derivatives and their analogues (Table 1) under conditions of varying the molecular structure, which are accompanied by changes in electronic and information properties of molecules. To activate the protective action of the drug, apparently, it is necessary for the molecule to interact with the active centers of the biosystem. This can lead to a restructuring of some physiological processes of the body accompanied by an increase in its radioresistance. The molecule of radioprotector can interact with biologically important macromolecules that are sensitive to the action of radiation. Radiation damage to the biosystem is complex and consists of both the act of excitation or ionization and a number of accompanying rapid processes, such as the migration of charge and energy of secondary

electromagnetic radiation, localization of charge, polarization of the medium, etc. [2]. One of the conditions for the repair of such damage is the contact of the damaged molecule with an exogenous impurity molecule, as well as the presence of charge and energy migration paths between them. That is, there must be long- and short-range (on the molecular scale) interactions leading to the formation of relatively stable intermolecular complexes including biomacromolecules and low-molecular impurities. The excess energy received by biomolecules as a result of irradiation can be used to destroy the intermolecular bonds stabilizing the complexes. It is possible that the weakening of the effect of irradiation on the body is associated with the creation of obstacles to the implementation of radiation damage by hindering the electronic-conformational transformations of macromolecules. It is known [3] that after irradiation in the presence of effective aminothiol radioprotectors, modification of single- and double-stranded DNA breaks occurs.

3 Problem Solution

In this article, various possible manifestations of primary physical and chemical processes at the molecular level, arising under the influence of high-energy radiation, will be compared with the quantum chemical characteristics of molecules. Knowledge of the electronic structure of molecules makes it possible to apply this information in the analysis of various ideas about the mechanisms of the protective action of low-molecular compounds. That is, it is possible to identify which quantum parameters of molecules are the most common and informative for the analyzed series of chemical compounds.

3.1 Electronic properties of molecules

The electronic characteristics of the substituted aminothiols and their analogues were calculated using the semiempirical Hartree-Fock self-consistent field method in the MINDO/3 approximation [4], taking into account the optimization of the spatial geometry of the molecules. The method provides satisfactory results for most standard characteristics of molecules, including molecules containing phosphorus and sulfur atoms. Analysis of the electronic features of the molecules of this series of chemical compounds showed that the most informative molecular parameters are the boundary one-electron molecular orbital (MO) energies: the highest occupied ε_{oc}, the lowest unoccupied ε_{un} spin orbitals, the energy interval $\Delta\varepsilon = \varepsilon_{un} - \varepsilon_{oc}$, as well as the squares of dipole moments of molecules μ^2. Table 1 shows the calculated values of ε_{oc}, ε_{un}, μ and $\Delta\varepsilon$, as well as the radio-protective trait - survival rate (A, %), of irradiated mice at absolutely lethal dose [5,6]. It is well known that the energy of the highest occupied molecular orbital of an isolated molecule determines its ionization potential (in accordance with Koopmans theorem for molecules with closed shells). The energy interval $\Delta\varepsilon$ approximates (the electronic transition is limited by the symmetry of the one-electron levels) the electronic excitation energy of an isolated molecule. Since the results of biological effect assessment of drugs depend, in general, on many different factors that cannot always be taken into account, it is convenient for further statistical analysis to divide all chemical compounds presented in Table 1 into three groups according to the result indicator (A): highly active (A_1 survival rate $\geq$ 60%; relatively low doses), medium active (A_2 = 50%; medium doses) and slightly inactive or inactive drugs ($A_3 \leq$ 30%; high doses). Table 1 shows either the protection range or the maximum possible protection. The protection effect depends significantly on the applied dose of the drug [5] and is limited by various factors, including the toxicity of the drugs. Table 1 shows either the protection range or the maximum possible protection. Using the results of quantum-mechanical calculations (Table 1), we divide all chemical compounds on the basis of $\Delta\varepsilon$ into three groups. The first group includes preparations for which the value of $\Delta\varepsilon < 8.5$ eV. The second group contains chemical compounds for which the energy difference is in the relatively narrow range of 8.5 eV $\leq \Delta\varepsilon \leq$ 9 eV. The third group includes preparations for which the difference $\Delta\varepsilon > 9$ eV. The numerical material can now be presented in the form of a 3×3 contingency table (Table 2). In this case, features A and $\Delta\varepsilon$ can be called interval features. Following the method detailed in [7,8], the empirical values q_{ij} in Table 2 determine the frequencies of occurrence of sign values in admissible areas (1, 2 and 3) defined by interval signs A and $\Delta\varepsilon$. If there were a one-to-one relationship between the attributes, non-zero values would only be on the diagonal of the table. Verification of the statistical reliability of the relationship between the radioprotective effect of substituted aminothiols and the energy value $\Delta\varepsilon$ is performed by comparing the empirical values of frequencies q_{ij} with the expected values. We choose the relative frequencies q_{ij}' as the expected values, so that the distribution over the cells of the table would correspond to the absence of connection between the events.

Table 1
Electronic and information features of substituted aminothiols and their analogues, as well as their bioactivity

No	Chemical compounds	ε_{oc}	ε_{un}	$\Delta\varepsilon$	μ, *D*	*Z*, conv. units	*H*, *bits*	*dH*1, *bits*	A_{exp}, %
		eV							
1	$H_2N(NH=)CNHCH_2CH_2SPO_3H_2$	-9.43	-4.70	4.73	4.48	3.140	2.130	-0.109	100
2	$H_2NCH_2CH_2SC(=NH)NH_2$	-7.63	0.19	7.82	2.53	2.625	1.623	-0.014	100
3	$H_2NC(=NH)CH_2SH$	-8.25	-1.19	7.06	2.48	2.727	1.686	-0.030	100
4	$H_2N(NH=)CNHCH_2CH_2SH$	-7.53	-0.27	7.26	4.38	2.625	1.623	-0.014	100
5	$H_2NCH_2CH_2SPO_3H_2$	-9.57	-4.68	4.89	4.89	3.125	2.078	-0.125	100
6	$H_2NCH_2CH_2SSO_3H$	-9.25	-2.94	6.31	4.82	3.333	2.013	-0.126	100
7	$(CH_3)_2N(NH=)CCH_2SH$	-7.87	-0.02	7.85	2.42	2.471	1.545	0.041	100
8	$H_2NC(=NH)CH_2CH_2SH$	-7.87	-0.03	7.84	3.11	2.572	1.611	0.015	100
9	$H_2NC(=NH)CH_2SSO_3H$	-8.52	-3.00	5.52	3.65	3.600	2.156	-0.141	100
10	$H_2NC(=NH)CH_2SPO_3H$	-8.66	-4.57	4.08	4.86	3.375	2.225	-0.147	100
11	$H_2NCH_2CH_2SC(=S)SH$	-8.63	-1.59	7.04	2,56	3.000	1.725	-0.024	100
12	$H_2NC(=NH)CH_2SSCH_2C(=NH)NH_2$ (T)*)	-7.78	-1.22	6.56	0.94	2.900	1.761	-0.036	100
13	$CH_3C(NH_2)HCH_2SH$	-7.38	-0.21	7.18	0.91	2.286	1.430	0.066	70
14	$H_2NCH_2CH_2SH$	-8.32	-0.16	8.16	2.71	2.364	1.491	0.032	60
15	$(CH_3)_2S=O$	-9.39	-2.12	7.21	3.59	2.600	1.571	0.022	65
16	$H_2NCH_2CH_2SCN$	-8.89	-0.72	8.17	1.98	2.833	1.729	0.000	50
17	$H_2NCHCOOHCH_2SH$	-9.27	-1.43	7.84	2.57	3.000	1.921	-0.024	50
18	$H_2NCH_2CH_2SC_6H_5$	-8.38	-0.08	8.30	2.13	2.571	1.438	0.042	50
19	$H_2NCH_2CH_2CH_2SH$	-8.85	0.50	9.35	2.41	2.286	1.430	0.066	50
20	$H_2NC(=NH)SCH_3$	-8.81	-0.78	8.03	2.85	2.462	1.547	-0.016	50
21	$CH_3CH(SH)CH_2NH_2$	-8.91	0.12	9.03	2.56	2.286	1.430	0.066	50
22	$H_2NCH_2CH_2SC(=O)CH_3$	-8.88	-0.05	8.84	1,30	2.733	1.774	0.025	50
23	$H_2C=CHCH_2NHCH_2CH_2SH$	-8.61	0.14	8.75	2.43	2.333	1.411	0.079	50
24	$CH_3CH_2SC(=NH)NH_2$ (T)	-8.64	0.33	8.97	2.49	2.572	1.611	0.015	50
25	$H_2NCH_2CH_2SCH_3$	-8.89	-0.06	8.83	2.31	2.286	1.430	0.066	30
26	$H_2NCH_2CH_2SCH_2CH_3$	-8.90	-0.16	8.73	2.67	2.235	1.379	0.085	20
27	$H_2NCH_2CH(SH)COOH$ (T)	-9.29	-0.75	8.54	2.56	3.000	1.921	-0.024	10
28	$H_2NCH_2CH_2SCH_2CH=CH_2$	-8.89	-0.07	8.82	2.30	2.333	1.411	0.079	0
29	$H_2NCH_2CH_2CH_2CH_2SH$	-8.85	0.45	9.30	2.59	2.235	1.379	0.085	10
30	$OHCH_2CH_2SH$	-9.80	-0.08	9.72	2.05	2.600	1.571	0.022	0
31	$(CH_3)_2NCH_2CH_2SH$	-8.28	-0.01	8.27	1.09	2.235	1.379	0.085	10
32	$H_2NCH_2C(CH_3)_2SH$	-8.88	0.06	8.94	2.52	2.375	1.424	0.076	10
33	$H_2NC(=O)NHCH_2CH_2SH$**)	-9.40	-0.08	9.32	2.24	2.800	1.857	-0.019	25
34	$H_2NCH_2CH_2OH$	-8.98	1.72	10.7	1.87	2.364	1.491	0.032	0
35	$CH_3CH_2CH_2SH$	-9.44	0.42	9.85	2.49	2.167	1.189	0.110	10
36	$CH_3NHCH_2CH_2SH$ (T)	-8.55	0.12	8.67	2.50	2.286	1.430	0.066	10
37	H_2NCOCH_2SH	-9.85	-0.41	9.43	2.49	3.000	1.961	-0.036	10
38	H_2NCH_2COSH	-9.37	0.09	9.45	1.66	3.000	1.961	-0.036	0
39	$H_2NCOCH_2CH_2SH$	-9.55	0.04	9.59	2.08	2.769	1.823	0.007	0
40	$H_2NOCH_2CH_2SH$	-8.39	0.06	8.45	2.69	2.667	1.781	-0.023	10
41	$OHCH_2CH_2SC(=NH)NH_2$	-8.74	0.17	8.91	2.49	2.800	1.857	-0.019	0
42	$H_2NCH_2CH_2COSH$	-9.02	0.46	9.48	1.79	2.769	1.823	0.007	0
43	$H_2NCOCH_2CH_2SC(=NH)\ NH_2$	-8.89	-0.01	8.88	1.76	2.889	1.877	-0.018	0
44	$H_2NCH_2CH(OH)CH_3$	-8.95	1.78	10.7	1.81	2.625	1.623	-0.014	0
45	$H_2NCH_2CH(Cl)CH_3$	-8.88	0.99	9.86	1.15	2.462	1.489	0.057	0

*) The T index indicates that the drug is toxic. **) With oral administration of the drug at a dose of 1000 mg/kg, there is no protection.

Obviously, in the case of independent events, the joint proportion (relative frequency) is equal to the simple product of the proportions: $p_{ij} = p_i \cdot P_j \equiv q_i \cdot Q_j/N^2$; here $p_i = q_i/N$. The numerical values of q_i and Q_j are given in Table 2. The theoretically expected relative frequencies q_{ij}' are determined as

follows:

$$q_{ij}' = N \cdot p_{ij} \equiv q_i \cdot Q_j / N. \quad (1)$$

Such values of frequencies would occur in the absence of a connection between the events. The *chi-square* test is used to test the *null*-hypothesis that there is no relation between the events.

Table 2
Relationship between the radioprotective effect of substituted aminothiols and their analogues and the value of the electronic sign $\Delta\varepsilon$.

A, %	Characteristic $\Delta\varepsilon$ (in eV)			
	$\Delta\varepsilon < 8.5$	$8.5 \le \Delta\varepsilon \le 9.0$	$\Delta\varepsilon > 9.0$	Total
≥ 60	$q_{11} = 15$ $q_{11}' = 7$	$q_{12} = 0$ $q_{12}' = 3.67$	$q_{13} = 0$ $q_{13}' = 4.33$	$q_1 = 15$ $p_1 = 0.333$ $q_1' = 15$
= 50	$q_{21} = 4$ $q_2' = 4.2$	$q_{22} = 3$ $q_{22}' = 2.2$	$q_{23} = 2$ $q_{23}' = 2.6$	$q_2 = 9$ $p_2 = 0.200$ $q_2' = 9$
≤ 30	$q_{31} = 2$ $q_{31}' = 9.8$	$q_{23} = 8$ $q_{23}' = 5.13$	$q_{33} = 11$ $q_{33}' = 6.07$	$q_3 = 21$ $p_3 = 0.467$ $q_3' = 21$
	$Q_1 = 21$ $P_1 = 0.467$	$Q_2 = 11$ $P_2 = 0.244$	$Q_3 = 13$ $P_2 = 0.289$	$N = 45$ $\sum_{i=1}^{3} P_i =$ $\sum_{j=1}^{3} p_j = 1.$

If the *chi-square* value is less than the table value at a given level of significance and number of degrees of freedom, then the *null*-hypothesis (no relation between the signs) is accepted. Using the data in Table 2, we obtain the following inequality:

$$\chi^2 = \sum (q_{ij} - q_{ij}')^2 / q_{ij}' = 29.4 > \chi_{0.05}^{cr,2}(f = 4) = 9.488. \quad (2)$$

Here the summation is performed over indices i and j from 1 to 3. Number of degrees of freedom $f = (v - 1)\cdot(w - 1) = 4$; here v is the number of rows, w is the number of columns. For a 3×3 contingent table (Table 2) $v = w = d = 3$. Thus, inequality (2) with a probability of 0.95 allows us to reject the *null*-hypothesis and accept that the events A and $\Delta\varepsilon$ are significantly interconnected. The value of the energy interval is associated with the radio-protective effect of drugs. This conclusion is also preserved when choosing the significance level α = 0.001, i.e. 0.1%. Consequently, a decrease in the energy difference $\Delta\varepsilon$ is accompanied by an increase in the radioprotective effectiveness of chemical compounds of a number of aminothiols and their analogues. Obviously, if there were an absolute unambiguous relationship between features, then the table should contain only diagonal elements.

The measure of the strength of the relationship between events can also be quantified using Pearson's contingency coefficient ($0 \le K \le 1$) and Chuprov's contingency coefficient ϕ ($0 \le \phi \le 1$)) [8]:

$$K = \{\chi^2 \cdot d / [(\chi^2 + N)(d - 1)]\}^{0.5} = 0.77,$$
$$\phi = \{\chi^2 / [N \cdot (d - 1)]\}^{0.5} = 0.57. \quad (3)$$

Here $d = 3$, that is, the number of rows and columns of a 3×3 table is the same. Both the coefficients K and ϕ indicate the existence of a strong connection between the events. It is usually assumed that the relationship between signs is close if the inequalities $K \ge 0.5$ and $\phi \ge 0.3$ are satisfied. Pearson's contingency coefficient is comparable to the linear correlation coefficient, which in this case is $|r|$ = 0.82. Let's check whether the average values of $\Delta\varepsilon^{av}$ differ significantly for the three areas of activity: $A_1 \ge 60\%$, $A_2 = 50\%$ and $A_3 \le 30\%$. For regions A_1 and A_2 we obtain the following sample statistics of average values:

$N_1 = 15$, $\Delta\varepsilon_1^{av} = 6.63 \pm 0.33$; 95% confidence interval: (5.93-7.34), $\Delta\varepsilon_1^{min} = 4.08$, $\Delta\varepsilon_1^{max} = 8.16$, $S_1 = 1.27$; Grubbs-Romanovsky homogeneity test for small samples: $\tau^{max} = 1.20 < \tau^{min} = 2.01 < \tau_{0.05}^{cr,2}(N_1) = 2.493 < \tau_{0.05}^{cr,1}(N_1) = 2.617$; the Wilk-Shapiro normality test: $W = 0.894 > W_{0.05}^{cr}(N_1) = 0.881$; the David-Hartley-Pearson normality test [9]: $U1_{0.05}^{cr}(N_1) = 2.97 < U = [(\Delta\varepsilon_1^{max} - \Delta\varepsilon_1^{min})/S_1] = 3.21 < U2_{0.05}^{cr}(N_1) = 4.17$; coefficient of variation: $V_1 = S_1 \cdot 100\% = 19.2\%$; $\delta V_1 = \pm V_1/(2N_1)^{0.5} = \pm 3.56\%$; accuracy of experience: $P_1 = V_1/N_1^{0.5} = 4.96\%$; $N_{1repr} = 12$; according to [10] the representativeness of the sample arithmetic mean is also determined by the inequality: $\Theta = y1 \cdot [(N_1 - 1)/(N_1 \cdot y2 - y1^2)]^{0.5} = 20.23 > \Theta^{cr} = 3$,

here we use the following notations: y1 is the sum of the variants of the series, y2 is the sum of the squares of the variants of the series;

$N_2 = 9$, $\Delta\varepsilon_2^{av} = 8.58 \pm 0.17$; 95% confidence interval: (8.19-8.98), $\Delta\varepsilon_2^{min} = 7.84$, $\Delta\varepsilon_2^{max} = 9.35$, $S_2 = 0.517$, $\tau^{min} = 1.44 < \tau^{max} = 1.478 < \tau_{0.05}^{cr,2}(N_2) = 2.237 < \tau_{0.05}^{cr,1}(N_2) = 2.392$; the Wilk-Shapiro normality test: $W = 0.946 > W_{0.05}^{cr}(N_2) = 0.829$, the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N_2) = 2.59 < U = [(\Delta\varepsilon_2^{max} - \Delta\varepsilon_2^{min})/S_2] = 2.92 < U2_{0.05}^{cr}(N_2) = 3.552$;

V_2 = (6.02 ± 1.42)%. P_2 = 2.01%; N_{2repr} = 8: Θ = 49.8. (4)

It follows from inequalities (4) that the populations $\Delta\varepsilon_1$ and $\Delta\varepsilon_2$ are homogeneous, and the elements of the populations are normally distributed. We first need to check whether the variance of the residuals in the two populations is different. To do this, let's calculate the ratio of the larger variance to the smaller variance. This relation has an F distribution, which should be compared to the table value:

$$F_{1,2} = S_1^2/S_2^2 = 6.03 >$$
$$F_{0.05}^{cr}(f_1 = N_1 - 1; f_2 = N_2 - 1) = 3.23. \quad (5)$$

Since $F > F_{0.05}^{cr}$, the variances are significantly different and the comparison of the mean values should be performed using the following relationship:

$$t = \Delta\varepsilon_2^{av} - \Delta\varepsilon_1^{av} = 1.95 > T^{av} =$$
$$[v_1 \cdot t_{0.05}^{cr}(f_1) + v_2 \cdot t_{0.05}^{cr}(f_2)]/(v_1 + v_2)^{0.5} = 0.66, \quad (6)$$

where $v_1 = S_1^2/N_1$ and $v_2 = S_2^2/N_2$; number of freedom degrees: $f_1 = N_1 - 1$, $f_2 = N_2 - 1$. Here the one-sided Student's test is applied. Since inequality (6) holds, the hypothesis of the equality of the mean values can be rejected. It follows from inequality (6) that at the 95% confidence level, the average values of the energy interval $\Delta\varepsilon^{av}$ for the regions A_1 and A_2 are significantly different and this difference is not random. Using the results (4) we can also calculate the biserial correlation coefficient between the first and second groups [11]:

$$r_{bs} = [(\Delta\varepsilon_2^{av} - \Delta\varepsilon_1^{av})/S] \cdot [N_1 \cdot N_2/(N^2 - N)]^{0.5} = 0.663,$$
$$t = 4.15 > t_{0.05}^{cr}(N - 2) = 1.72,$$
$$S^2 = [(N_1 - 1) \cdot S_1^2 + (N_2 - 1) \cdot S_2^2]/(N - 2), \quad (7)$$

which is significant at the 95% level; here the total $N = N_1 + N_2 = 24$; standard deviation $S = 1.417$.

Now let's compare the areas of bioactivity A_2 and A_3. The population statistics $\Delta\varepsilon_3$ for area A_3 is as follows:

N_3 = 21, $\Delta\varepsilon_3^{av}$ = 9.26 ± 0.15; 95% confidence interval: (8.96-9.56), $\Delta\varepsilon_3^{min}$ = 8.27, $\Delta\varepsilon_3^{max}$ = 10.7, S_3 = 0.664, $\tau^{min} = 1.49 < \tau^{max} = 2.17 << \tau_{0.05}^{cr,2}(N_3) = 2.644 < \tau_{0.05}^{cr,1}(N_3) = 2.750$; the Wilk-Shapiro normality test: $W = 0.933 > W_{0.05}^{cr}(N_3) = 0.918$, the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N_3) = 3.18 < U = [\Delta\varepsilon_3^{max} - \Delta\varepsilon_3^{min})/S_3] = 3.66 < U2_{0.05}^{cr}(N_3) = 4.49$; V_2 = (7.17 ± 1.11)%, P_2 = 1.56%, N_{3repr} = 17; Θ = 7.2. (8)

Since the difference between the variances for regions 2 and 3 of bioactivities is not significant:

$$F = S_3^2/S_2^2 = 1.65 >$$
$$F_{0.05}^{cr}(f_3 = N_3 - 1; f_2 = N_2 - 1) = 3.15, \quad (9)$$

then the comparison of average $\Delta\varepsilon^{av}$ values for regions A_2 and A_3 should be performed using the relation [11,12]:

$$t = \Delta\varepsilon_3^{av} - \Delta\varepsilon_2^{av} = 0.67 > t^{av} =$$
$$t_{0.05}^{cr}(f = N_{23} - 2) \cdot \{N_{23}S_{23}^2/[N_2N_3(N_{23} - 2)]\}^{0.5} = 0.42,$$
$$N_{23} = N_2 + N_3, \quad S_{23}^2 = (N_2 - 1) \cdot S_2^2 + (N_3 - 1) \cdot S_3^2. \quad (10)$$

Inequality (10) also holds when using the two-sided criterion $t_{0.975}^{cr}$ = 2.05. Thus, there is a significant difference between the average values for the energy interval $\Delta\varepsilon$ for the two neighboring regions of bioactivities A_2 and A_3, and the following inequalities are observed: $\Delta\varepsilon_3^{av} > \Delta\varepsilon_2^{av} > \Delta\varepsilon_1^{av}$. Consequently, there is a trend in the relationship between the bioactivity of chemical compounds and the value of the energy interval $\Delta\varepsilon$. The smaller the value of $\Delta\varepsilon$, the lower the energy required to excite an isolated molecule is likely to be. That is, it can be assumed that, in accordance with this sequence, the electron-donor properties of the molecule are enhanced. At the same time, since $\Delta\varepsilon$ is defined as the difference between the MO energies, the decrease in the difference $\Delta\varepsilon$ can be associated with a decrease in the energy scale of the molecular level ε_{un}. This, in turn, leads to an improvement in the acceptor properties of the molecules. According to Szent-Györgyi [13], molecules with a low $\Delta\varepsilon$ value are catalytic electron transmitters and have both good donor and acceptor properties. As is known, the decisive factor in the protective effect of sulfur-containing preparations is the accumulation of radioprotector molecules up to their threshold concentration in the cells of critical organs of the body. The donor-acceptor interaction (a mechanism caused by the exchange of electrons between the filled orbitals of the donor molecule and the vacant orbitals of the acceptor molecule) leads to the binding of the drug in the body. The resulting energy level of the complex lies below the initial states. The resulting binding energy level of the complex lies below the initial states. Delocalization of electron leads to the formation of a molecular complex. For the donor-acceptor mechanism of complex formation, the position on the energy scale of the vacant molecular orbital is important. In the works [3,14] it has been proved that aminothiols as radioprotectors have the ability to form temporary mixed disulfide bonds with the enzymes responsible for the synthesis of DNA precursors. The ability of aminothiols to interact with proteins can lead to

short-term blocking of metabolic processes, including DNA synthesis. The transfer of an electron between molecules is characteristic of many fundamental biological processes, which are accompanied by the formation of complexes with charge transfer.

Let us also perform an additional check for the presence of a systematic shift in the average molecular factor $\Delta\varepsilon$. To do this, we will use the Abbe-Linnick test [15,16] for the sequences of bioactivities in Table 1 ordered by magnitude. For a sample of independent, normally distributed random variables $\Delta\varepsilon$ the trend hypothesis is tested by the following statistics:

$$q = 0.5 \cdot \sum_{i=1}^{N-1} \left(\Delta\varepsilon_{i+1} - \Delta\varepsilon_i\right)^2 / \sum_{i=1}^{N} \left(\Delta\varepsilon_i - \Delta\varepsilon^{av}\right)^2 =$$

$$= 0.306 < q_{0.05}^{cr}(N) = 0.7605,$$

$$Q^* = -(1-q)\cdot[(2N+1)/(2-(1-q)^2)]^{0.5} = -5.37.$$

Here $\Delta\varepsilon^{av} = 8.25$ is the arithmetic mean value; sample volume $N = 45$. The approximate statistical index Q^* is preferably used for sample sizes $N \geq 60$. However, the resulting estimate for smaller sample sizes usually does not contradict the inequality for the statistical test q; Q^* has a standard normal distribution [15]. If $q > q^{cr}$, then we can assume that the observations do not contain a systematic shift of mathematical expectations. Since the reverse inequality $q < q^{cr}$ is satisfied, and the inequality $Q^* = -5.37 < u_{0.05} = -1.645$ is also valid, then the *null*-hypothesis about the equality of the means of the $\Delta\varepsilon_i$ series is rejected (an alternative hypothesis about the presence of a systematic bias is accepted) with a probability of 0.95, in this case; $u_{p/2}$ is the quantile of the normal distribution at $p = 0.10$. Thus, an increase in the energy interval $\Delta\varepsilon$ is associated with a decrease in the value of the effective feature A_{exp}.

Statistical methods are usually used in a complex manner, due to the complexity of the processes under study. One of the most common methods of applied statistics is regression analysis, which is used to determine the functional relationship between the resulting factor and many possible explanatory variables. In this case, the explanatory variables are related to the bioresponse by some regression function. However, as is known [11,17], correlation analysis establishes only the strength of the connection. The analysis showed that the relationship between the radioprotective activity (A, %) of aminothiols from Table 1 and the value of electronic energy $\Delta\varepsilon$ can be approximated by the following empirical non-linear dependence (Fig. 1):

$$A(\Delta\varepsilon) = 1/[1+c\cdot\exp(b_0 + b_1\cdot\Delta\varepsilon)], \quad N = 45. \quad (11)$$

Hereinafter it is assumed that $A \equiv A$(in percent)/100%. Using the Grubbs-Romanovsky τ-test, it can be shown that the initial data for the radioprotective efficacy of drugs satisfy the uniformity condition:

$$\tau = |A^{max/min} - A^{av}| / S = \begin{cases} 1.40^{max} < \tau_{0.05}^{cr}(N) = 3.12, \\ 1.12^{min} < \tau_{0.05}^{cr}(N) = 3.12. \end{cases} \quad (12)$$

Here $A^{av} = 0.444$ is the average value, $S = 0.396$ is the standard deviation. By simple mathematical transformations the approximation (11) can be linearized. The regression becomes linear in the estimated parameters b_i. The resulting features A_{line} of the linearized regression satisfy the homogeneity condition: $\tau^{max} = 1.29 < \tau^{min} = 1.41 < \tau_{0.05}^{cr}(N) = 3.12$; the Kolmogorov-Smirnov normality test: $d = 0.188$, $\lambda = 1.26 < \lambda_{0.95}^{cr} = 1.36$. In this case, statistical tests of linear regression can be used to assess the significance of the regression. The following statistics were obtained for the linearized regression equation:

$N = 45$; $R = 0.82 \pm 0.07$, $R > R_{0.05}^{cr}(N-2) = 0.295$ [11]; for relatively small samples ($N \geq 10$) the statistical significance of the correlation coefficient is determined by the inequality [11]: $t = 0.5\cdot\ln[(1+R)/(1-R)]\cdot(N-3)^{0.5} = 7.50 > t_{0.05}^{cr}(N-2) = 1.68$; the minimum sample size sufficient for the reliability of the correlation coefficient [18]: $N_{0.05}^{min} = 6$: $RMSE = 1.977$, $c = 4.71\cdot10^{-4}$, $b_0 = -7.68 \pm 1.70$, $b_1 = 1.89 \pm 0.20$, $t(b_1) = 9.30 > |t(b_0)| = 4.53 > t_{0.05}^{cr}(N-2) = 2.017$; $F = 86.52 >> F_{0.05}^{cr}(f_1 = 1; f_2 = 43) = 4.06$; sum of squared residuals: $\Sigma_0 = 168.1$. (13)

Since the inequality $t(b_1) > t_{1-\alpha}^{cr}$ takes place for a two-sided critical region, the regression coefficient b_1 is statistically significant at a significance of $1-\alpha$, reliably greater than zero, and reflects a positive relationship between the molecular factor $\Delta\varepsilon$ and the bioresponse. Estimation of significance of the coefficient of determination can be obtained using F - statistics: $F = R^2\cdot(N-m-1)/(1-R^2)$, which is compared to the table value. Here $m = 1$ is the number of explanatory variables. Since $F >> F^{cr}$, we can conclude that the coefficient of determination R^2 is significantly different from zero. The population statistics $\Delta\varepsilon$ for the entire sample ($N = 45$) will be as follows:

$N = 45$; $\Delta\varepsilon^{av} = 8.25 \pm 0.22$; reliability of the average value: $t = 37.5 > t_{0.05}^{cr}(N-2) = 2.017$; 95% confidence interval: (7.81-8.69); $\Delta\varepsilon^{min} = 4.08$, $\Delta\varepsilon^{max} = 10.7$; $S_\varepsilon = 1.417$, $\tau^{max} = 1.67 < \tau^{min} = 2.84 < \tau_{0.05}^{cr}(N) = 3.12$; the Wilk-Shapiro normality test: W

$= 0.923 \approx W_{0.05}^{cr}(N) = 0.926$; the Pearson normality test: $\chi^2 = 8.87 < \chi_{0.05}^{2,cr}(df = 12) = 21.026$; the Romanowsky's normality test: $|\chi^2 - df|/(2df)^{0.5} = 0.64 < 3.0$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N) = 3.83 < U = (\Delta\varepsilon^{max} - \Delta\varepsilon^{min})/S_\varepsilon = 4.5 < U2_{0.05}^{cr}(N) = 5.35$; $V = (17.2 \pm 1.81)\%$; $P = 2.56\%$; $N_{repr} = 36$; $\Theta = 37.6 > 3$. (14)

Here $df = n - l - 1$ is the number of freedom degrees, n is the number of intervals into which the range of variation of the random variable is divided, l is the number of estimated distribution parameters. It follows from inequalities (14) that the population of elements $\Delta\varepsilon$ is homogeneous and has a distribution close to the normal distribution. According to the Chaddock scale [19], the correlation coefficient R (13) is in the range of values, which characterizes the relationship between the explanatory variable and the resultant variable as "close". Thus, at a significance level of 5%, we can recognize the existence of a close relationship between the events. The correlation field that determines the distribution of observations (Fig. 1) indicates the existence of a relationship between the value of the energy interval $\Delta\varepsilon$ and the radio-protective activity of molecules.

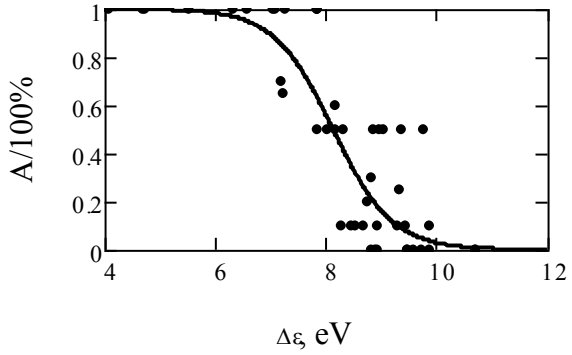

Fig. 1. The scattering diagram of the radio-protective action of a number of substituted aminothiols and their analogues (Table 1) depending on the magnitude of the difference in electronic energies $\Delta\varepsilon$. The solid line is determined by regression (11).

In a fairly narrow range of energies $8\text{eV} \le \Delta\varepsilon \le 9\text{eV}$ (group of chemical compounds Nos.16-24) there is a significant change in the radio-protective properties of low molecular weight compounds. The energy interval $\Delta\varepsilon$ can be compared with the threshold processes of deexcitation of metastable states of biomacromolecules, associated both with the interception of migrating electronic excitation in the biosystem, as well as with the prevention of possible molecular conformational transitions. The threshold action of radioprotectors is important not only when the molecular descriptor $\Delta\varepsilon$ is used as an explanatory feature. In the case when $\Delta\varepsilon < 8.5$ eV, i.e., less than the threshold value, most of the aminothiols analyzed here have a significant prophylactic effect. If the value of $\Delta\varepsilon$ noticeably exceeds the energy range of 8.5 - 9.0 eV, then the chemical compounds of a number of substituted aminothiols, as a rule, are weakly active in the antiradiation action.

From experiments [20,21] aimed at studying the radioprotective effect of cysteamine on mammalian cells, it is known that radio-protector molecules quickly penetrate into cells and reach nuclear DNA without difficulties associated with transport. Low-molecular compounds can interact with DNA [22] and, thus, contribute to the stabilization of the macromolecule structure by participating in the dissipation of the excitation electronic energy into the conformational energy of the impurity nuclear subsystem. It is possible that conformational transitions induced by interaction with low-molecular compounds lead to changes in the electronic state of the active groups of biomacromolecules and their mutual arrangement. This, in turn, can lead to pre-irradiation blocking of DNA replication. It is possible that molecular processes associated with the phenomenon of conformational selection take place [23]. Possibly, the instability of the initial conformation of low-molecular chemical compounds with respect to conformational transitions also has a preventive effect, and the probability of transition to another conformation is higher, the smaller the energy interval $\Delta\varepsilon$ [23].

Analyzing the molecular data from Table 1, one can notice that an increase in the energy interval $\Delta\varepsilon$ is accompanied by a decrease in the dipole moment of the molecule. This qualitative conclusion can be verified using the Abbe-Linnick test (13). After ranking the data by the value of the feature $\Delta\varepsilon$, the following inequalities for dipole moments were obtained $q = 0.531 < q_{0.05}^{cr}(N = 45) = 0.7603$. It follows that for the sample presented in Table 1, there is a significant trend with a statistical significance of 0.95.

Next, let's check whether the relationship between the variables μ and $\Delta\varepsilon$ is linear or non-linear. In accordance with the physical concepts of the participation of polar molecules in intermolecular interactions [24], the dependence of the energy on the value of the dipole moment must be quadratic. Indeed, the value of the energy interval $\Delta\varepsilon$ (in eV) is significantly related to the value of the square of the total dipole moment μ^2 of the

molecule. The regression equation can be linearized by replacing the explanatory variable μ^2 with M:

$$\Delta\varepsilon(M) = a_0 + a_1 \cdot M, \quad (15)$$

$N = 45$; $r_{\mu 2} = -0.74 \pm 0.07$, $|r_{\mu 2}| > r_{0.05}^{cr}(N-2) = 0.310$; statistical significance of the correlation coefficient: $t = |r_{\mu 2}| \cdot (N-2)^{0.5}/(1-r_{\mu 2}^2)^{0.5} = 7.2 > t_{0.05}^{cr}(N-2) = 2.017$ (two-sided critical region); the minimum sample size sufficient for the reliability of the correlation coefficient: $N_{0.05}^{min} = 7$; $RMSE = 0.998$; $a_0 = 9.62 \pm 0.24$, $a_1 = -0.19 \pm 0.03$, $t(a_0) = 40.1 > |t(a_1)| = 7.25 > t_{0.05}^{cr}(N-2) = 2.017$; sum of squares residuals: $\Sigma_0 = 42{,}8$.

The significance of the coefficient of determination is determined using F-statistics: $F = r^2 \cdot (N-2)/(1-r^2) = 52.08 >> F_{0.05}^{cr}(f_1 = 1; f_2 = 43) = 4.06$. For comparison, here is also the correlation coefficient for the regression: $\Delta\varepsilon(\mu) = b_0 + b_1 \cdot \mu$, $r_\mu = -0.66$.

Dipole moment statistics of the molecules:

$N = 45$, $\mu^{av} = 2.54 \pm 0.14$; reliability of the average value: $t = 18.1 > t_{0.05}^{cr}(N) = 2.014$; 95% confidence interval: (2.25-2.83); $\mu^{min} = 0.91$, $\mu^{max} = 23.9$; $S_\mu = 4.89$; $\tau^{min} = 1.69 < \tau^{max} = 2.45 < \tau_{0.05}^{cr}(N) = 3.12$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N) = 3.75 < U = [(\mu^{max} - \mu^{min})/S_\mu] = 4.13 < U2_{0.05}^{cr}(N) = 5.26$; $N_{repr} = 36$, $P = 6.7\%$; $\Theta = 17.6 > 3.0$. (16)

The populations μ (16) and $\Delta\varepsilon$ (14) are homogeneous, and the distribution of their elements is close to the normal distribution.

The presence of a large dipole moment of the molecule contributes to the long-range (on the molecular scale) electrostatic interaction, which can lead to the emergence of a microgradient and concentration of the drug in the local area of the biophase. In addition, the presence of a dipole moment in a molecule suggests that such molecules should have the property of hydrophilicity. However, the simultaneous presence of groups of CH_2 atoms in molecules gives molecules the opposite property - hydrophobicity. The hydrophobicity of the molecule increases with an increase in the number of CH_2 groups. In this case, the molecules from Table 1 refer to amphiphilic molecules containing both hydrophobic and hydrophilic sites from groups of atoms. As is known, the existence of polar and nonpolar parts of a molecule promotes the aggregation of low-molecular chemical compounds with the formation of molecular clusters, including those with biological molecules.

It was shown [24] that the contribution to the interaction energy, which is proportional to the square of the dipole moment of the molecule, is due to the polarization properties of the target biosubstrate and the dipole-dipole interaction. The higher the induction electron polarizability of an object, the stronger its interaction with a low-molecular compound. It is well known that if some molecule has a constant dipole moment, then this dipole moment causes a shift of charges in the neighboring molecule, i.e. there is an induced dipole moment. This results in an attraction between the molecules due to constant and induced dipole moments. Such interactions were called orientation-induction interactions. In the dipole approximation, the interaction energy is proportional to μ^2 and R^{-6}; here R is the distance between the centers of gravity of the molecules. It should also be noted that the polarization properties of electronically excited molecular systems are significantly higher compared to the polarization of molecules in their ground electronic state. The importance of orientation-induction interactions for the activation of bioactivity of molecules suggests that the region with which the molecule interacts should have high polarization properties. The presence of a large dipole moment of the molecule in general indicates the hydrophilicity of the molecule. Indeed, for the majority of active ($A \geq 60\%$) in radio-protective chemical compounds (Table 1) dipole moment on average exceeds the dipole moment of ineffective in radio-protective chemical compounds. Let us check whether the average values of the dipole moment of chemical compounds differ significantly for the activity regions A_1 ($N_1 = 15$), A_2 ($N2 = 9$), and A_3 ($N_3 = 21$). The following statistics have been obtained:

$N_1 = 15$, $\mu_1^{av} = 3.22 \pm 0.34$; reliability of the average value: $t = 9.5 > t_{0.05}^{cr}(N_1) = 2.131$; 95% confidence interval: (2.49-3.95); $\mu_1^{min} = 0.91$, $\mu_1^{max} = 4.89$, $S_{\mu 1} = 1.315$, $\tau^{max} = 1.26 < \tau^{min} = 1.76 < \tau_{0.05}^{cr,2}(N_1) = 2.497 < \tau_{0.05}^{cr,1}(N_1) = 2.617$; the Wilk-Shapiro normality test: $W = 0.915 > W_{0.05}^{cr}(N_1) = 0.881$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N_1) = 2.97 \approx U = [(\mu_1^{max} - \mu_1^{min})/S_{\mu 1}] = 3.03 < U2_{0.05}^{cr}(N_1) = 4.170$; $P = 10.5\%$; $N_{repr} = 12$, $\Theta = 9.5$. (17)

$N_2 = 9$, $\mu_2^{av} = 2.30 \pm 0.15$; reliability of the average value: $t = 15.3 > t_{0.05}^{cr}(N_2) = 2.262$; 95% confidence interval: (1.95-2.65); $\mu_2^{min} = 1.18$, $\mu_2^{max} = 2.16$, $S_{\mu 2} = 0.463$, $\tau^{min} = 1.18 < \tau^{max} = 2.16 < \tau_{0.05}^{cr,2}(N_2) = 2.237 < \tau_{0.05}^{cr,1}(N_2) = 2.392$; the Wilk-Shapiro normality test: $W = 0.873 > W_{0.05}^{cr}(N_2) = 0.829$, the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N_2) =$

$2.590 < U = [(\mu_2^{max} - \mu_2^{min})/S_{\mu 2}] = 3.35 < U2_{0.05}^{cr}(N_2) = 3.552$; $P = 6.7\%$; $N_{repr} = 7$, $\Theta = 15.3$. (18)

$N_3 = 21$, $\mu_3^{av} = 2.15 \pm 0.10$; reliability of the average value: $t = 21.5 > t_{0.05}^{cr}(N_3) = 2.086$; 95% confidence interval: (1.93-2.36); $\mu_3^{min} = 1.09$, $\mu_3^{max} = 2.69$, $S_{\mu 3} = 0.471$, $\tau^{min} = 1.15 < \tau^{max} = 2.246 < \tau_{0.05}^{cr,2}(N_3) = 2.644 < \tau_{0.05}^{cr,1}(N_3) = 2.75$; the Wilk-Shapiro normality test: $W = 0.888 \approx W_{0.05}^{cr}(N_3) = 0.908$, the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N_3) = 3.180 < U = [(\mu_3^{max} - \mu_3^{min})/S_{\mu 3}] = 3.27 < U2_{0.05}^{cr}(N_3) = 4.490$; $P = 4.8\%$; $N_{repr} = 17$, $\Theta = 20.9$. (19)

It follows that the populations are homogeneous and have a distribution close to the normal distribution. Let us check the significance of the difference between the average values of μ_1^{av} and μ_2^{av}. Let us first find the ratio of the larger sample variance to the smaller sample variance:

$$F_{1,2} = S_{\mu 1}^2/S_{\mu 2}^2 = 8.1 >$$

$$F_{0.05}^{cr}(f_1 = N_1 - 1;\ f_2 = N_2 - 1) = 3.23. \quad (20)$$

The $F_{1,2}$ value exceeds the tabulated value, so the variances should be considered different at a significance level of $\alpha = 0.05$. Since inequality (20) is satisfied, to determine the significance of the difference in the average values, we can use the approximate following relation (Cochran-Cox test [12,25]):

$$tS_\mu^2 = t_{0.05}^{cr}(f_1 = N_1 - 1)S_{\mu 1}^2 + t_{0.05}^{cr}(f_2 = N_2 - 1)S_{\mu 2}^2,$$

$$S_\mu = [S\mu_1^2/N_1 + S_{\mu 2}^2/N_2]^{0.5}, \quad (21)$$

$$t = |\mu_1^{av} - \mu_2^{av}| = 0.92 > T^{av} = tS_\mu^2/S_\mu = 0.66.$$

A one-sided significance test is applied here. It follows from inequality (21) that at the significance level $\alpha = 0.05$, the average values of the dipole moment μ differ significantly for regions A_1 and A_2.

Then, using relations [11,12] compare the average values of μ_2^{av} and μ_3^{av} for regions A_2 and A_3:

$$F_{2,3} = S_{\mu 3}^2/S_{\mu 2}^2 = 1.03 >$$

$$F_{0.05}^{cr}(f_2 = N_2 - 1;\ f_3 = N_3 - 1) = 2.45,$$

$$t = |\mu_2^{av} - \mu_3^{av}| = 0.15 < t^{av} = t_{0.05}^{cr}(f = N_{2,3} - 2)\times$$

$$\{N_{2,3}\cdot S_{2,3}^2/[N_2\cdot N_3\cdot(N_{2,3} - 2)]\}^{0.5} = 0.94,$$

$$N_{2,3} = N_2 + N_3, \quad S_{2,3}^2 = (N_2 - 1)\cdot S_{\mu 2}^2 + (N_3 - 1)\cdot S_{\mu 3}^2. \quad (22)$$

Consequently, for weakly active and inactive chemical compounds, the average values of the populations μ_2 and μ_3 at the 5% significance level do not differ significantly and we can accept the null hypothesis, that is, they belong to the same set. In this case, the samples μ_2 and μ_3 can be combined. The following statistics for the combined population were obtained:

$N_{2+3} = 30$; $\mu_{2+3}^{av} = 2.19 \pm 0.09$; reliability of the average value: $t = 24.3 > t_{0.05}^{cr}(N_{2+3}) = 2.042$; 95% confidence interval: (2.02-2.37); $\mu_{2+3}^{min} = 1.09$, $\mu_{2+3}^{max} = 2.85$, $S_{2+3} = 0.463$, $\tau^{min} = 1.42 < \tau^{max} = 2.38 < \tau_{0.05}^{cr}(N_{2+3}) = 2.96$; the Wilk-Shapiro normality test: $W = 0.895 \approx W_{0.05}^{cr}(N_{2+3}) = 0.927$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N_{2+3}) = 3.470 < U = [(\mu_{2+3}^{max} - \mu_{2+3}^{min})/S_{2+3}] = 3.80 < U2_{0.05}^{cr}(N_{2+3}) = 4.890$; $P = 3.9\%$; $N_{2+3repr} = 24$, $\Theta = 25.6$. (23)

Statistics (23) demonstrates that the sample $N_{2+3} = 30$ is homogeneous, and the population elements have a distribution close to normal. It is now possible to compare the average dipole moment values for molecules belonging to the group of bioactive drugs ($A_1 \geq 60\%$) and for molecules belonging to the pooled population ($A_2 = 50\%$ and $A_3 \leq 30\%$). Using relations (5) and (6) we obtain the following inequalities:

$$F = S_1^2/S_{2+3}^2 = 8.07 >$$

$$F_{0.05}^{cr}(f_1 = N_1 - 1;\ f_{2+3} = N_{2+3} - 1) = 2.05,$$

$$t = |\mu_1^{av} - \mu_{2+3}^{av}| = 1.03 > T^{av} = 0.75. \quad (24)$$

Therefore, the average values of the dipole moments of molecules for bioactive and inactive drugs at the 95% confidence level differ significantly. Thus, an increase in the sample size (up to $N_{2+3} = 30$) containing weakly active drugs and inactive chemical compounds does not change the significance of the difference between the average values (21) and (24). Consequently, we can recognize that the distinction for the average values is not random.

Let us also check the hypothesis of the existence of a relationship between the value of the anti-radiation activity of chemical compounds (Table 1) and the value of the square of the dipole moment of molecules. For this purpose we will use the method of contingency of features. Indeed, for the majority of radioprotective active chemical compounds ($A_1 \geq 60\%$), the following inequality holds for the square of the dipole moment: $\mu^2 > \mu^{2,av} = 7.35D^2$. At the same time, for weakly active and inactive drugs, the following inequality is more likely: $\mu^2 < \mu^{2,av} = 7.35D^2$. The average value of the dipole moment square $\mu^{2,av}$ will be taken as its threshold value. Using the data of Table 1, we will compile a 3×2 contingency table (Table 3), which presents the relative frequencies q_{ij} of the appearance of features in the i-th row and j-th column, as well as theoretically expected frequencies q_{ij}', determined

by the formula (1). To test the hypothesis of the significance of the relationship between the feature μ^2 and the radioprotective efficacy of substituted aminothiols and their analogues, we use the *chi*-square test (2): $\chi^2 = 20.69 > \chi_{0.05}^{2,cr}(f = 2) = 5.99$.

It follows from this inequality that the *null*-hypothesis should be rejected and the existence with a probability of 0.95 of a significant relationship between the molecular feature μ^2 and the radioprotective activity of chemical compounds should be recognized.

Table 3

Interrelation of the radioprotective action of substituted aminothiols and their analogs with the electronic parameter μ^2

A, %	Sign μ^2		
	$\geq 7.35D^2$	$< 7.35D^2$	Total
≥ 60	$q_{11} = 9$ $q_{11}' = 3$	$q_{12} = 6$ $q_{12}' = 12$	$q_1 = 15$ $p_1 = 0.333$ $q_1' = 15$
$= 50$	$q_{21} = 1$ $q_{21}' = 1.8$	$q_{22} = 8$ $q_{22}' = 7.2$	$q_2 = 9$ $p_2 = 0.200$ $q_2' = 9$
< 50	$q_{31} = 0$ $q_{31}' = 4.20$	$q_{23} = 21$ $q_{23}' = 16.8$	$q_3 = 21$ $p_3 = 0.467$ $q_3' = 21$
	$Q_1 = 10$ $P_1 = 0.222$	$Q_2 = 35$ $P_2 = 0.778$	$N = 45$ $\sum_{i=1}^{3} P_i =$ $\sum_{j=1}^{3} p_j = 1.00$

Apparently, the value of the square of the dipole moment of the molecule influences the radioprotective effect of the drug. The strength of the bond is characterized by the contingency coefficients (2) and (3):

$$K = [\chi^2/(\chi^2 + N)]^{0.5}/K_{max} = 0.74,$$
$$\phi = [\chi^2/(d-1)/N)]^{0.5} = 0.68. \quad (25)$$

The correction K_{max} depends on the number of rows and the number of columns of the 3×2 contingency table and can be quantified from the following ratio [8]:

$$K_{max} = 0.5\cdot[((v-1)/v)^{0.5} + ((w-1)/w)^{0.5}] = 0.762. \quad (26)$$

Here $w = 3$ is the number of rows and $v = 2$ is the number of columns, respectively; d is the smaller of the two numbers v and w. Both contingency coefficients (26) indicate a fairly strong relationship between the value of the molecular trait μ^2 and the resulting trait. This conclusion does not allow rejecting the initial assumption about the importance of the accumulation of low-molecular compounds in the body and their participation in intermolecular bonds through dipole-dipole and orientation-induction interactions in the radioprotective effect.

Let us check the significance of the relationship between the radioprotective properties of drugs and the square of the dipole moment of the molecule using the following regression equation:

$$A(\mu^2)/100 = 1/[1+c\cdot\exp(b_0 + b_1\cdot\mu^2)], \quad (27)$$

linearized regression statistics:

$N = 45$; $R = -0.60 \pm 0.12$, $|R| > R_{0.05}^{cr}(N-2) = 0.310$; the minimum sample size sufficient for the reliability of the correlation coefficient: $N_{0.05}^{min} = 11$; $RMSE = 2.76$, $c = 4.71\cdot10^{-4}$, $b_0 = 10.41 \pm 0.66$, $b_1 = -0.34 \pm 0.07$, $t(b_0) = 15.68 > |t(b_1)| = 4.85 > t_{0.05}^{cr}(N-2) = 2.017$; $F = 23.52 > F_{0.05}^{cr}(f_1 = 1; f_2 = 43) = 4.06$; $\Sigma_0 = 327.2$ (28)

In applied statistics, an approximate rule has been established: if the absolute value of the correlation coefficient R exceeds the average error of the coefficient by at least three times, then we can reliably assume that the relationship between the signs is not random. Used in statistics (28), the tabular values of the Student and Fisher tests indicate the significance of the relationship between the value of the attribute μ^2 and the radioprotective activity of chemical compounds in Table 1. However, as further analysis showed, the joint consideration of the explanatory variables $\Delta\varepsilon$ and μ^2 in the regression did not lead to a significant improvement in the regression. That is, the inclusion of an additional explanatory variable μ^2 in the regression (11) does not have a significant effect on the resulting variable. At the same time, an additional check of the relationship (15) of molecular features $\Delta\varepsilon$ and μ^2 showed that for bioactive drugs ($A_1 \geq 60\%$) there is a significant linear relationship:

$\Delta\varepsilon(\mu^2) = a_0 + a_1\cdot\mu^2$, $N_1 = 15$, $r_1 = -0.69 \pm 0.15$, adjusted correlation coefficient [11]: $|r_1^*| = 0.71 > r_{0.05}^{cr}(N_1 - 2) = 0.514$; estimation of the significance of the correlation coefficient, taking into account the Hotelling corrections [11]: $u_H = 0.794 > u_{0.05}(N_1) = 0.523$; the minimum sample size sufficient for the reliability of the correlation coefficient: $N_{0.05}^{min} = 8$; $RMSE = 0.955$; $a_0 = 7.92 \pm 0.45$, $a_1 = -0.11 \pm 0.03$, $t(a_0) = 17.63 > |t(a_1)| = 3.43 > t_{0.05}^{cr}(N_1 - 2)$; $F =$

$11.75 > F_{0.05}^{cr}(f_1 = 1; f_2 = 13) = 4.67$; straightness sign [18]: $K = [N\cdot(1 - R^2)]^{0.5} = 2.80 < K^{thr} = 3.00$. (29)

To further test, the *null*-hypothesis of correlation coefficient insignificance we will use the following inequality, which applies to samples with a volume ≈ 10 [11]:

$$t = 0.5\cdot\ln[(1 + r_1)/(1 - r_1)]\cdot(N_1 - 3)^{0.5} =$$
$$2.94 > t_{0.05}^{cr}(N_1 - 2) = 2.16. \quad (30)$$

This inequality rejects the *null*-hypothesis at the significance level $\alpha = 0.05$. Consequently, we can agree that the correlation coefficient (29) is significant. The relationship of explanatory features $\Delta\varepsilon$ and μ^2 can lead to their collinarity. To check for collinearity between variables, we use the Farrar-Glauber test [26]:

$$\chi^2 = -[N_1 - 1 - (2m + 5)/6]\cdot\ln(1 - r_1^2) =$$
$$8.30 > \chi_{0.05}^{2,cr}(f = 1) = 3.841. \quad (31)$$

The number of explanatory variables m =1. Since $\chi^2 > \chi^{2,cr}$, the hypothesis about the presence of collinearity does not contradict the original data. According to the Cheddock scale [19], for the adjusted correlation coefficient r_1^* (29), the relationship is characterized as "close connection". An increase in the feature μ^2 is associated with a decrease ($a_1 < 0$) in the value of the energy interval $\Delta\varepsilon$. However, for the areas of bioactivity A_2 and A_3, there is no significant relationship between the molecular features $\Delta\varepsilon$ and μ^2. Indeed, the correlation coefficients at a significance level of 5% are noticeably lower than the admissible critical values $|r_2| = 0.17 < r_{0.05}^{cr}(f = 7) = 0.666$ [11] and $|r_3| = 0.36 < r_{0.05}^{cr}(f = 19) = 0.433$, respectively. It is important to note that the relationships of molecular features for bioactive chemical compounds and inactive (or weakly active) drugs differ significantly. That is, there is a structural shift in the relationships of molecular features. This structural shift can be taken into account as an additional, qualitative property of chemical compounds that separates molecules that are bioactive in terms of radioprotection from inactive or weakly active drugs.

Dipole electrostatic forces can lead to a significant change in the distribution of positive ions inside the cell [27,28]. This, in turn, is reflected in the mechanism of DNA replication exposed to intense radiation. In addition, the electrostatic field component of the dipole (vector value), directed along the double helix chain, strongly polarizes the electrons of nucleotide bases, especially in electronically excited states, which also affects DNA replication. Intense external irradiation can sensitize this process, while the action of the electrostatic dipole field stabilizes it.

Dipole-dipole or dipole-induction interactions (both interactions are proportional to μ^2) can also determine the direction of movement of a low-molecular compound to the activation center in the biosystem (for example, to DNA, RNA), as well as the binding to it. The dynamic equilibrium between the bound state of the impurity with the target biosubstrate and the disconnected state of the molecules is determined by anisotropic short-range interaction forces, of which the forces responsible for the formation of charge transfer complexes are the most effective. For homologous series of compounds, the ability to complex formation depends significantly on the energy parameter ε_{un} of the acceptor [29], and the electron-acceptor properties of molecules are the stronger, the lower the molecular level ε_{un} lies on the energy scale.

We group the preparations in Table 1 according to the value of the sign ε_{un} into two groups, prone to complex formation ($\varepsilon_{un} < 0$) and inactive in this respect ($\varepsilon_{un} > 0$). Using the method of conjugation of qualitative features, we determine the statistical characteristics of the relationship between the radioprotective activity of preparations and their ability to form complexes with charge transfer. In this case, the feature ε_{un} is a dichotomous feature that can take only two qualitative values - either negative or positive. The significance of the relationship between features is established using the *chi-square* test. Using the numerical values q_{ij} and q_{ij}', presented in the 3×2 contingency table (Table 4), using relations (2) and (25), we determine the statistics of the relationship between the molecular trait ε_{un} and the biological response (A_{exp},%): $\chi^2 = 9.71 > \chi_{0.05}^{2,cr}(f = (v - 1)\cdot(w - 1)) = 5.99$, $\phi = 0.465$, $K = 0.553$. These results indicate a significant relationship between the radioprotective activity of drugs and the position on the energy scale of the electronic level ε_{un} in isolated molecules.

Let us compare the average values of ε_{un}^{av} (in eV) for the activity regions A_1, A_2, and A_3. The following statistics were obtained:

$N_1 = 15$, $\varepsilon_{un1}^{av} = -1.77 \pm 0.47$; 95% confidence interval: (-2.77, -0.77); $\varepsilon_{un1}^{min} = -4.70$, $\varepsilon_{un1}^{max} = 0.19$, $S_{un1} = 1.81$, $\tau^{max} = 1.08 < \tau^{min} = 1.62 < \tau_{0.05}^{cr,2}(N_1) = 2.493 < \tau_{0.05}^{cr,1}(N_1) = 2.617$; the Wilk-Shapiro normality test: $W = 0.860 \approx W_{0.05}^{cr}(N_1) = 0.881$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N_1) = 2.970 \approx U = [(\varepsilon_{un1}^{max} - \varepsilon_{un1}^{min})/S_{un1}] = 2.71 = U2_{0.05}^{cr}(N_1) = 4.170$, (32)

Table 4
Relationship between the radioprotective activity of substituted aminothiols and their analogues and the electronic attribute ε_{un}

A, %	Sign ε_{un} (in eV)		
	Negative	Positive	Total
≥ 60	$q_{11} = 14$ $q_{11}' = 9.33$	$q_{12} = 1$ $q_{12}' = 5.67$	$q_1 = 15$ $p_1 = 0.33$ $q_1' = 13$
= 50	$q_{21} = 7$ $q_{21}' = 5.60$	$q_{22} = 4$ $q_{22}' = 3.40$	$q_2 = 11$ $p_2 = 0.20$ $q_2' = 11$
< 50	$q_{31} = 9$ $q_{31}' = 13.07$	$q_{23} = 12$ $q_{23}' = 7.93$	$q_3 = 21$ $p_3 = 0.467$ $q_3' = 21$
	$Q_1 = 28$ $P_1 = 0.62$	$Q_2 = 17$ $P_2 = 0.38$	$N = 45$ $\sum_{i=1}^{3} P_i = \sum_{j=1}^{3} p_j$ $=1.00$

$N_2 = 9$, $\varepsilon_{un2}^{av} = -0.22 \pm 0.03$; 95% confidence interval: (-0.70, 0.26); $\varepsilon_{un2}^{min} = -1.43$, $\varepsilon_{un2}^{max} = 0.50$, $S_{un2} = 0.627$, $\tau^{max} = 1.15 < \tau^{min} = 1.93 < \tau_{0.05}^{cr,2}(N_2) = 2.237 < \tau_{0.05}^{cr,1}(N_1) = 2.392$; the Wilk-Shapiro normality test: $W = 0.919 > W_{0.05}^{кр}(N_2) = 0.829$, the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N_2) = 2.590 < U = [(\varepsilon_{un2}^{max} - \varepsilon_{un2}^{min})/S_{un2}] = 3.07 = U2_{0.05}^{cr}(N_2) = 3.552$, (33)

$N_3 = 21$, $\varepsilon_{un3}^{av} = 0.23 \pm 0.03$; 95% confidence interval: (-0.05, 0.50); $\varepsilon_{un3}^{min} = -0.75$, $\varepsilon_{un3}^{max} = 1.78$, $S_{un3} = 0.611$, $\tau^{min} = 1.60 < \tau^{max} = 2.55 < \tau_{0.05}^{cr}(N_3) = 2.80$; the Wilk-Shapiro normality test: $W = 0.819 < W_{0.05}^{cr}(N_3) = 0.908$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N_3) = 3.18 < U = [(\varepsilon_{un3}^{max} - \varepsilon_{un3}^{min})/S_{un3}] = 4.14 < U2_{0.05}^{cr}(N_3) = 4.490$, (34)

These populations are homogeneous and have a distribution of elements close to the normal distribution. Using (5), (6), (8), (9) and the results of statistics (32) - (34), we check the significance of the difference between the average values of ε_{un1}^{av}, ε_{un2}^{av}, and ε_{un3}^{av}. The following inequalities were obtained:

$$F_{1,2} = S_{un1}^2/S_{un2}^2 = 8.32 > F_{0.05}^{cr}(f_1 = N_1 - 1; f_2 = N_2 - 1) = 3.23,$$
$$t = |\varepsilon_{un1}^{av} - \varepsilon_{un2}^{av}| = 1.55 > T^{av} = 0.907, \quad (35)$$

$$F_{2,3} = S_{un2}^2/S_{un3}^2 = 1.05 < F_{0.05}^{cr}(f_2 = N_2 - 1; f_3 = N_3 - 1) = 3.52,$$
$$t = |\varepsilon_{un2}^{av} - \varepsilon_{un3}^{av}| = 0.44 > t^{av} = 0.417. \quad (36)$$

Obviously, the average values of ε_{un1}^{av} and ε_{un2}^{av} differ significantly from each other. Considering inequality (36) we can admit that average values of ε_{un1}^{av} and ε_{un3}^{av} are also significantly different. Thus, the average values of ε_{un}^{av} for the three bioactivity regions A_1, A_2, and A_3 differ significantly from each other at the 95% confidence level, and the following sequence of inequalities is observed: $\varepsilon_{un3}^{av} > \varepsilon_{un2}^{av} > \varepsilon_{un1}^{av}$. For bioactive radioprotectors ($A_1 \geq 60\%$), the molecular features ε_{un1} are grouped around $\varepsilon_{un1}^{av} = -1.77$ eV, while for weakly active or inactive chemical compounds, the energy values ε_{un} are most likely localized in the region (-0.22, 0.23) eV.

At one time, Bacq and Alexander [30] propose a hypothesis according to which one of the possible mechanisms of radioprotection is due to the fact that the radioprotector molecule can neutralize radicals. As is known, the primary products of radiolysis are electrons, free radicals, and excited molecules. For the first selected area of the compounds in Table 1, the energy level ε_{un1} is negative (with the exception of drug No.2). That is, the transfer of an electron from a radical or other reaction center to this orbital can become energetically favorable. For example, it is known that under the action of radiation in the aquatic environment, chemically highly active and very mobile (diffusion coefficient $4.96 \cdot 10^{-5}$ cm^2/s) hydrated electrons (e_{aq}) are formed [31,32]. For compounds of the first group (Nos. 1, 5, 6, 9, 10), the energy of the unoccupied one-electron level ε_{un1} lies on the energy scale below the main energy level of a hydrated electron, which is equal to -2.82 eV [33]. The electron hydration time is ≈ 0.24 ps. The hydrated electron is a very powerful reducing agent, and the addition reaction proceeds at a high rate [31,32]. Consequently, it is energetically favorable for the electron e_{aq} to move from the hydrated state to the lower molecular level of the radioprotector. The high polarizability of –SH and –SS– groups can become a site of attack by a hydrated electron [32]. It is noted in the literature that the rate constant of the reaction of a hydrated electron with effective radioprotectors (for example, cysteine, cysteamine, or cystamine) is even higher than with oxygen. In [34] it is noted that self-trapped electrons can cause a break in the carbon-sulfur bond in organic compounds. In this case, the radioprotector molecule can act as a neutralizer of the active radical. At the same time, the participation of the radioprotector molecule in other possible mechanisms of radioprotection is also not denied here. For most of the compounds from Table 1 that

do not have effective radioprotection, the values of ε_{un} are either positive or small (in absolute value) negative values. In this case, energy is required to attach a hydrated electron.

It is not excluded that molecules of radio-protectors can participate both in the processes of "healing" of damages in the DNA structure, caused by radiation exposure [35] and participate in the interaction with the target molecule prior to irradiation. In the latter case, such binding (including disulfide bonds) leads to a change in the organism's response to the subsequent action of radiation [3]. It is possible that the adsorption of low molecular compounds in the biosystem before irradiation increases the activity of biological processes - DNA replication, synthesis of RNA, proteins, etc., which in turn increases the body's resistance to radiation. Moreover, the radioprotector molecule must have such a spatial configuration that allows the molecule to adapt to local areas of the biophase. For example, cysteine ($HSCH_2CH(NH_2)COOH$) is known to have radioprotective properties, whereas its optical stereoisomer *iso*-cysteine has no protective properties, although the one-electron molecular-orbital energies of these molecules are virtually the same. That is, the spatial correspondence (complementarity) between the functional groups of the protector molecules and the biological object is important. The emerging steric hindrances can limit the donor-acceptor properties of molecules, since effective binding of molecules occurs with a strong overlap of electron shells - the maximum overlap of the electron-donor orbital with the electron-acceptor orbital.

Two enantiomeric forms of the same molecule often have different biological activities. This is because receptors, enzymes, antibodies, and other elements of the body also have chirality, and the structural mismatch between these elements and chiral molecules prevents their interaction. A similar situation occurs, for example, with regard to the effect on blood pressure of the enantiomers, the left-handed isomer of adrenaline compared to the right-handed isomer. These two molecules differ only in the spatial arrangement of the structural elements of the molecule, which is reflected in the interaction of the molecule with the adrenoreceptor. Structural correspondence or mismatch between drug molecules and chiral molecules of the biophase turned out to be a decisive factor for the different manifestations of the biological activity of drugs [36]. In this case, a significant role is played by the vector quantity - the dipole moment of the molecule, as well as the hydrophobic - hydrophilic regions of the radioprotector molecules.

It can be assumed that in the process of implementation in the body of the protective properties of a radioprotector, an important role is played by the ability of the drug to participate in the formation of complexes with charge transfer. Thereby, perhaps, a temporary inhibition of biochemical processes is carried out. It is known [37] that the intermolecular forces that determine the formation of complexes with charge transfer are small (compared to covalent bonds), but, nevertheless, they have a significant effect on the conformational transitions of macromolecules in a polar dielectric medium. It is known [37] that the intermolecular forces that determine the formation of complexes with charge transfer are small (compared to covalent bonds), but, nevertheless, they have a significant effect on the conformational transitions of macromolecules in a polar dielectric medium. From the point of view of the manifestation of radioprotective action by drugs, the donor properties of radioprotectors have been repeatedly discussed in the literature [38–40]. Damage repair in this case is related both to the actual process of intermolecular electron transfer to the ionized bioobject, which leads to "healing" of the damage, and to the possible formation of complexes with charge transfer.

The donor properties of chemical compounds generally depend on many electronic and steric properties of the interacting molecules. However, for the homologous series of chemical compounds, the electronic processes of electron transfer are significantly related to the position on the energy scale of the highest filled ε_{oc} molecular orbital. It is well known that the higher on the energy scale is the MO level of energy ε_{oc}, the stronger are the donor properties of the molecule. Let's check whether there is a statistically significant relationship between the anti-radiation protection of the preparations and the position of the one-electron MO of the energy level relative to the threshold value ε_{oc}^{thr}. For the sample presented in Table 1 (N = 45), as a threshold (boundary) value ε_{oc}^{thr}, we take the average value ε_{oc}^{av} = -8.78 eV (95% confidence interval is: (-8.60, -8.96) eV); the Wilk-Shapiro normality test: $W = 0.958 > W_{0.05}^{cr}(N) = 0.945$. Further, we again use the statistical method of contingencies. Let's make a 3×2 contingency table (Table 5). Using the results presented in Table 5, as well as formulas (2), (23) and (24), one can obtain statistics on the relationship between the radio-protective effectiveness of substituted aminothiols and their analogues and the position on the energy

scale of the one-electron MO level ε_{oc}^{av}: $\chi^2 = 10.95 > \chi_{0.05}^{2,cr}(f = 2) = 5.99$, $\phi = 0.49$, $K = 0.581$. Sufficiently high statistical characteristics indicate the existence of a significant relationship between the radioprotective activity of molecules and their donor properties. It is important to emphasize the physicochemical meaning of the energy parameters ε_{oc} and ε_{un}, which determine the level of the redox potential of the molecule

Table 5
Relationship between the radioprotective action of molecules and the electronic parameter ε_{oc}

A, %	Sign ε_{oc} (in eV)		
	$\lvert\varepsilon_{oc}\rvert \le 8.78$	$\lvert\varepsilon_{oc}\rvert > 8.78$	Total
≥ 60	$q_{11} = 11$ $q_{11}' = 6.0$	$q_{12} = 4$ $q_{12}' = 9.0$	$q_1 = 15$ $p_1 = 0.33$ $q_1' = 15$
$= 50$	$q_{21} = 3$ $q_{21}' = 3.6$	$q_{22} = 6$ $q_{22}' = 5.4$	$q_2 = 9$ $p_2 = 0.244$ $q_2' = 9$
< 50	$q_{31} = 4$ $q_{31}' = 8.4$	$q_{23} = 17$ $q_{23}' = 12.6$	$q_3 = 21$ $p_3 = 0.47$ $q_3' = 21$
	$Q_1 = 18$ $P_1 = 0.400$	$Q_2 = 27$ $P_2 = 0.600$	$N = 45$ $\sum_{i=1}^{3} P_i =$ $\sum_{j=1}^{3} p_j = 1.00$

Let's check the significance of the difference between the average values ε_{oc1}^{av}, ε_{oc2}^{av} и ε_{oc3}^{av}. Using relations (5), (6), (8), (9) the following inequalities were obtained:

$$F_{1,2} = (S_{oc1}/S_{oc2})^2 = 8.93 > F_{0.05}^{cr}(f_1 = N_1 - 1; f_2 = N_2 - 1) = 3.23,$$
$$t = |\varepsilon_{oc1}^{av} - \varepsilon_{oc2}^{av}| = 0.400 > T^{av} = 0.365, \quad (37)$$

$$F_{3,2} = (S_{oc3}/S_{oc2})^2 = 2.85 < F_{0.05}^{cr}(f_3 = N_3 - 1; f_2 = N_2 - 1) = 3.15,$$
$$t = |\varepsilon_{oc2}^{av} - \varepsilon_{oc3}^{av}| = 0.24 < t^{av} = 0.26. \quad (38)$$

The average values of ε_{oc1}^{av} for highly active chemical compounds differ significantly, at a confidence level of 0.95, from the values of ε_{oc2}^{av} and ε_{oc3}^{av} for weakly active or inactive drugs (37). At the same time, for chemical compounds for which the bioactivity $A_2 = 50\%$ or $A_3 < 50\%$, the energies of the highest occupied orbital do not differ statistically significantly (38).

Population statistics ε_{oc1} (in eV):

$A_1 \ge 60\%$, $N_1 = 15$, $\varepsilon_{oc1}^{av} = -8.41 \pm 0.19$; 95% confidence interval: (-8.81, -8.00), $\varepsilon_{oc1}^{min} = -9.57$, $\varepsilon_{oc1}^{max} = -7.38$; $S_{oc1} = 0.738$; $\tau^{max} = 1.40 < \tau^{min} = 1.57 < \tau_{0.05}^{cr,2}(N_1) = 2.493 < \tau_{0.05}^{cr,1}(N_1) = 2.617$; the Wilk-Shapiro normality test: $W = 0.925 > W_{0.05}^{cr}(N_1) = 0.881$; $|V| = 8.8\%$; $P = 2.3\%$; $N_{1repr} = 12$, $|\Theta| = 44.1$, (39)

$A_2 = 50\%$, $N_2 = 9$, $\varepsilon_{oc2}^{av} = -8.80 \pm 0.08$; 95% confidence interval: (-8.99, -9.27), $\varepsilon_{oc2}^{min} = -9.27$, $\varepsilon_{oc2}^{max} = -8.38$; $S_{oc2} = 0.247$; $\tau^{max} = 1.72 < \tau^{min} = 1.89 < \tau_{0.05}^{cr,2}(N_2) = 2.237 < \tau_{0.05}^{cr,1}(N_2) = 2.392$; the Wilk-Shapiro normality test: $W = 0.940 > W_{0.05}^{cr}(N_2) = 0.829$; $|V| = 2.8\%$; $P = 0.9\%$; $N_{2repr} = 8$, $|\Theta| = 5.28$, (40)

$A_3 < 50\%$, $N_3 = 21$, $\varepsilon_{oc3}^{av} = -9.04 \pm 0.09$; 95% confidence interval: (-9.23, -8.85), $\varepsilon_{oc3}^{min} = -9.85$, $\varepsilon_{oc3}^{max} = -8.85$; $S_{oc3} = 0.417$; $\tau^{max} = 0.46 < \tau^{min} = 1.95 < \tau_{0.05}^{cr,2}(N_3) = 2.644 < \tau_{0.05}^{cr,1}(N_3) = 2.750$; the Wilk-Shapiro normality test: $W = 0.943 > W_{0.05}^{cr}(N_3) = 0.908$; $|V| = 4.6\%$; $P = 1.0\%$; $N_{3repr} = 19$, $|\Theta| = 19.6$. (41)

Sets ε_{oc} (39) - (41) are homogeneous and have a distribution close to the normal distribution.

The performed statistical analysis showed that the radioprotective effectiveness of a number of mercaptoethylamine derivatives and their analogs depends on the different electronic properties of the molecules. As it turned out, there are some threshold values for all discussed electronic characteristics of samples from Table 1, and going beyond these values leads to a significant change in the preventive properties of drugs. This result allows us to use the methods of multivariate regression analysis and by analogy with equation (11) we can write the following nonlinear regression equation:

$$A/100 = 1/[1 + c \cdot \exp(b_0 + b_1 \cdot \varepsilon_{oc} + b_2 \cdot \varepsilon_{un} + b_3 \cdot \Delta\varepsilon + b_4 \cdot \mu^2)]. \quad (42)$$

For the regression parameter c the value obtained for the regression (11) was taken. Regressions of this type are usually called combined forms of regression. It is important to note that in regression (42) the explanatory variables $\Delta\varepsilon$ and μ^2 as well as ε_{un} and $\Delta\varepsilon$ are closely related. For example, the relationship between the molecular features ε_{un} and $\Delta\varepsilon$ is as follows: $N = 45$, $r = 0.92$. In applied statistical analysis, it is roughly accepted [8] that if the value of the pairwise correlation coefficient $|r| > 0.8$, then the explanatory variables are collinear. Below are the detailed statistics of the relationship. At the same time, there is practically no relationship between the explanatory variables ε_{un} and ε_{oc}: $r =$

0.16. In general, the presence of a paired linear relationship between several explanatory variables is defined as multicollinearity. Multicollinearity between variables can lead to a decrease in the accuracy of regression estimation and even to the impossibility of assessing the influence of explanatory variables on the resulting attribute [8]. It is known, that if one of the explanatory variables can be represented as a linear combination of other explanatory variables, then the system of normal equations may not have a unique solution. Therefore, the variable $\Delta\varepsilon$ can be excluded from the regression equation. As a result, we obtain the following three-factor regression equation:

$$A/100 = 1/[1+c\cdot\exp(b_0 + b_1\cdot\varepsilon_{oc} + b_2\cdot\varepsilon_{un} + b_3\cdot\mu^2)]. \quad (43)$$

After linearizing equation (43), the following multiple regression statistics were obtained:

$N = 45$, multiple correlation coefficient: $R_1 = 0.870 > R_{0.05}^{cr}(f_1 = m; f_2 = \nu) = 0.415$ [41], $R_1^2 = 0.757$, adjusted coefficient of determination [11]: $R_1^{*2} = 0.74$, $RMSE = 1.745$; $c = 4.71\cdot10^{-4}$, $b_0 = -19.664 \pm 4.022$, $b_1 = -3.306 \pm 0.455$, $b_2 = 1.406 \pm 0.305$, $b_3 = -0.104 \pm 0.075$; $|t(b_1)| = 7.26 > |t(b_0)| = 4.88 > t(b_2) > 4.61 > t_{0.05}^{cr}(f = N - m - 1) = 2.021 > |t(b_3)| = 1.39$; the significance of the coefficient of multiple determination: $F = 41.86 > F_{0.05}^{cr}(f_1 = m; f_2 = N - m - 1) = 2.83$; $\Sigma = 124.85$; AIC = 1.1537, SC = 1.3588, SS = 0.2660. (44)

Here $m = 3$ is the number of explanatory variables; $\nu = N - m - 1$; $R_{1-\alpha}(m;\nu)$ is the multiple correlation coefficient. The residuals of regression (43) are normally distributed. Kolmogorov - Smirnov normality test for regression residuals: $d_{max} = 0.1022$, $\lambda = d_{max}\cdot N^{0.5} = 0.686 < \lambda_{0.8}^{cr} = 1.07$. The Wilk-Shapiro normality test is also performed: $W = 0.951 > W_{0.05}^{cr}(N = 45) = 0.945$. Σ is the sum of squares of the regression residuals. Standardized (normalized) regression coefficients b_i^* are defined as follows [42]:

$$b_1^* = b_1\cdot S_{\varepsilon oc}/S_{Act} = -0.570 \pm 0.079,$$
$$b_2^* = b_2\cdot S_{\varepsilon un}/S_{Act} = -0.599 \pm 0.130,$$
$$b_3^* = b_3\cdot S_{\mu 2}/S_{Act} = -0.179 \pm 0.129. \quad (45)$$

Here, the index Act $\equiv A/100\%$ (after linearization of the regression equation); $S_{\varepsilon oc} = 0.586$, $S_{\varepsilon un} = 1.447$, $S_{\mu 2} = 5.87$, $S_{Act} = 2.234$. Standardized coefficients make it possible to compare quantitatively the influence of each explanatory variable on the variability of the resulting attribute. Using the standardized coefficients (45), one can determine the approximate coefficient of determination, which in an additive form allows one to make estimates of the relative contributions of each explanatory variable to the variability of the resulting attribute:

$$R_{appr}^2 = b_1^*\cdot r_{\varepsilon oc,Act} + b_2^*\cdot r_{\varepsilon un,Act} + b_3^*\cdot r_{\mu 2,Act}$$
$$= 0.263 + 0.378 + 0.116 = 0.757. \quad (46)$$

The approximate coefficient of determination (46) coincides with the coefficient of determination of the regression (44). Here $r_{\varepsilon oc,Act} = -0.456$, $r_{\varepsilon un,Act} = 0.648$, $r_{\mu 2,Act} = -0.594$ are paired correlation coefficients between the explanatory variables and the resulting feature (after transformation to a linear form). All correlation coefficients in absolute value are greater than the permissible table value $r_{0.05}^{cr}(f = N - 2) = 0.300$ [11]. From relation (46) it follows that the maximum contribution to the explanation of the variability of bioactivity comes from the electronic energies ε_{un} (26.3%) and ε_{oc} (37.8%). The contribution from the dipole moment of the molecule is much lower and amounts to only 11.6%.

Statistics (44) also provides information criterion Akaike [43] relative quality of a linear statistical model for a given data set. The information criterion is defined as follows:

$$AIC = 2m/N + \ln(\Sigma/N). \quad (47)$$

Here m is the number of explanatory variables; Σ is the sum of the squares of the regression residuals; N is the number of observations. The AIC test establishes a trade-off between the magnitude of the residual sum of squares and the number of explanatory variables. The first term is the penalty for using additional variables, the second term is the penalty for large variance. As the number of variables in a linear model increases, the first term in (47) increases and the second term decreases, because usually increasing the number of variables in a regression reduces the residual sum of squares. Regression residuals should be normally distributed. The equation (47) usually also includes a constant value of $1 + \ln(2\pi)$, which is omitted here because it is not essential for comparison tests. The Akaike test quantifies the relative amount of information that is lost when building a statistical model. The less information is lost (that is, the smaller the AIC value (47)), the higher the quality of the model. When comparing statistical models, preference is given to the model for which the AIC test is the smallest, that is, the model that minimizes information loss. The test is useful only when comparing linear statistical models, and the size of the compared samples N must be the same. Recently, the Schwarz criterion has also been frequently used [44]:

$$SC = (m + 1)\ln(N)/N + \ln(\Sigma/N). \quad (48)$$

The Schwartz criterion is similar to the Akaike criterion, but uses more stringent penalty functions. Practical applications of the (47) and (48) tests have shown that the Schwartz test is somewhat more reliable than the Akaike test in relative comparison of statistical models. The estimate obtained using this indicator is considered consistent. For a comparative assessment of the quality of models, one can also use the ratio: $SS = \Sigma^{0.5}/(N - m)$. This ratio is usually closely related [45] to the test values of Akaike and Schwartz.

Thus, approximately 24% of the variance remains unexplained, which can be attributed to unaccounted for factors or due to random variation in the raw data. The adjusted coefficient of determination is determined as follows [8,11]:

$$R^{*2} = 1 - (1 - R^2)(N - 1)/(N - m - 1). \quad (49)$$

Obviously, the adjusted coefficient of determination (49) depends on the number of explanatory variables in the regression. The adjusted coefficient of determination is used for the purpose of comparing models with different numbers of factors, so that the number of explanatory variables does not affect the R^2 statistics. The multiple correlation coefficient is determined from the following relationship:

$$R_{1-\alpha}(f_1 = m; f_2 = v) = [m \cdot F_{1-\alpha}(m,v)/(v + m \cdot F_{1-\alpha}(m,v)]^{0.05}. \quad (50)$$

Here $v = N - m - 1$; $F_{1-\alpha}$ is $100 \cdot (1 - \alpha)\%$ quantile of distribution $F(m;v)$. The *null*-hypothesis H_0: $R = 0$ is rejected at the α significance level, since $R > R_{1-\alpha}^{cr}$. At a given significance level $\alpha = 0.05$, the multiple correlation coefficient is much larger than the critical value and, therefore, its difference from zero is not accidental. In applied statistics, it is accepted that if the coefficient of determination $R^2 > 0.75$, then the relationship between the effective feature and the explanatory variables can be characterized as strong. The significance of the correlation coefficient can also be checked using the *t*-criterion (15):

$$t = R_1 \cdot (N - m - 1)^{0.5}/(1 - R_1^2)^{0.5} = 11.3 > t_{0.05}^{cr}(f = 41) = 2.021. \quad (51)$$

According to inequality (51) we can admit that the model adequately describes the relationship between the resultant variable and the explanatory variables. Signs ε_{oc} and ε_{un} can be defined as intensive indicators (to a lesser extent this applies to μ^2), which are directly related to cause-and-effect relationships between bioactivity and the structure of molecules. The statistics of the sampling sets ε_{oc} and ε_{un}, (the statistics of the set μ^2 are given in (16)) and the resulting feature A^*_{act} (values after linearization are used) will be as follows:

ε_{oc}^{av} = -8.79 ± 0.09; N = 45; 95% confidence interval: (-8.96, -8.60), ε_{oc}^{min} = -9.85, ε_{oc}^{max} = -7.38; S_{oc} = 0.59; $\tau^{min} = 1.81 < \tau^{max} = 2.41 < \tau_{0.05}^{cr}(N) = 3.12$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N) = 3.75 < U = [(\varepsilon_{oc}^{max} - \varepsilon_{oc}^{min})/S_{oc}] = 4.19 < U2_{0.05}^{cr}(N) = 5.26$;

ε_{un}^{av} = -0.53 ± 0.22; N = 45; 95% confidence interval: (-0.96, -0.09), ε_{un}^{min} = -4.70, ε_{un}^{max} = 1.78; S_{un} = 1.45; $\tau^{max} = 1.59 < \tau^{min} = 2.88 < \tau_{0.05}^{cr}(N) = 3.12$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N) = 3.75 < U = [(\varepsilon_{un}^{max} - \varepsilon_{un}^{min})/S_{un}] = 4.47 < U2_{0.05}^{cr}(N) = 5.26$;

A_{act}^{*av} = 7.51 ± 0.51; N = 45; 95% confidence interval: (6.50 - 8.53), A_{act}^{*min} = 2.70, A_{act}^{*max} = 11.9; S_A = 3.40; $\tau^{max} = 1.29 < \tau^{min} = 1.42 < \tau_{0.05}^{cr}(N) = 3.12$; the Kolmogorov-Smirnov normality test: d_{max} = 1.26, $\lambda = d_{max} \cdot N^{0.5} = 1.26 < \lambda_{0.95}^{cr} = 1.36$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N) = 3.75 > U = [(A_{act}^{*max} - A_{act}^{*min})/S_A] = 2.71 < U2_{0.05}^{cr}(N) = 5.26$. (52)

The negative value of the coefficient b_1 means that with increasing energy ε_{oc} (negative value), the radioprotective properties of the preparations increase. At the same time, a decrease in the ε_{un} level on the MO energy scale is accompanied by a decrease in the antiradiation activity of drugs ($b_2 > 0$). An additional check showed that the explanatory factors ε_{un} and μ^2 are closely related. The correlation coefficient is $|r_{23}| = 0.786 > R_{0.05}^{cr}(f = N - 2) = 0.300$. For the remaining explanatory variables the following pairwise correlations were obtained: r_{12} = 0.162 и r_{13} = 0.118. The regression residuals (43) are normally distributed (the Wilk-Shapiro test: W = 0.957 > $W_{0.05}^{cr}(N = 45) = 0.925$). The presence of multicollinearity of the variables is tested with the Farrar-Glauber test:

$$\chi^2 = -(N - 1 - (2m + 5)/6) \cdot \ln\left(\det \begin{vmatrix} r_{1,1} & r_{1,2} & r_{1,3} \\ r_{2,1} & r_{2,2} & r_{2,3} \\ r_{3,1} & r_{3,2} & r_{3,3} \end{vmatrix}\right)$$

$$= 45.8 > \chi_{0.05}^{cr}(f = m(m - 1)) = 12.592. \quad (53)$$

The collinearity of the explanatory variables is also indicated by the value:

$$t_{23} = |r_{23}| \cdot (N - m)^{0.5}/(1 - r_{23}^2)^{0.5} = 8.24 > t_{0.05}^{cr}(f = N - m) = 2.02. \quad (54)$$

Since inequalities (53) and (54) are satisfied, the hypothesis of multicollinearity does not contradict the original data. One of the highly correlated

explanatory variables must be eliminated from the regression equation. Which of the variables should be removed is determined as follows. The least significant of all regression coefficients is the coefficient b_3 (44). Next, the values $t_{ij} = r_{ij} \cdot (N - m)^{0.5}/(1 - r_{ij})^{0.5}$ (i = 1,2,3 and j = 1,2,3), are calculated. The index for t_{23} has the maximum value: $t_{23} = 8.24 > t_{0.05}^{cr}(f = N - 3) = 2.02 > t_{12} = 1.08 > t_{13} = 0.7$. Therefore, the third explanatory variable μ^2 can be excluded from the regression (44). Thus, the regression equation (44) can be replaced by a two-factor equation:

$$A/100 = 1/[1+c \cdot \exp(b_0 + b_1 \cdot \varepsilon_{oc} + b_2 \cdot \varepsilon_{un})] . \quad (55)$$

After linearizing the regression equation, the following statistics were obtained:

$N = 45$, the multiple correlation coefficient is equal: $R_2 = 0.86 > R_{0.05}^{cr}(f_1 = 2; f_2 = 42) = 0.365$ [41], $R_2^2 = 0.740$, $R_2^{*2} = 0.740$, $RMSE = 1.764$; the significance of the coefficient of multiple determination: $F = 60.48 > F_{0.05}^{cr}(f_1 = m; f_2 = N - m - 1) = 3.22$; $c = 6.79 \cdot 10^{-4}$, $b_0 = 20.362 \pm 4.034$, $b_1 = -3.319 \pm 0.46$, $b_2 = 1.744 \pm 0.18$; $t(b_2) > 9.36 > |t(b_1)| = 7.29 > |t(b_0)| = 5.05 > t_{0.05}^{cr}(f = 42) = 2.02$; $\Sigma_1 = 130.73$ is the sum of squares of residuals; the test of normality of the population of residuals: $W = 0.951 > W_{0.05}^{cr}(N = 45) = 0.945$; the Kolmogorov-Smirnov normality test for the residuals: $d_{max} = 1.102$, $\lambda = 0.6855 < \lambda_{0.2}^{cr} = 1.07$; the regression quality tests: AIC = 1.1554, SC = 1.3203, SS = 0.2659; $b_1^* = b_1 \cdot S_{oc}/S_A = -0.573 \pm 0.079$, $b_2^* = b_2 \cdot S_{un}/S_A = 0.749 \pm 0.079$. (56)

Reducing the number of variables in regression (55) compared to regression (44) preserves the quality of the regression. Moreover, the Schwarz test (SC) indicates an improvement in the quality of the regression. The following estimates of the contribution of each explanatory variable to the variability of the resultant variable were obtained:

$$R_{appr}^2 = b_1^* \cdot r_{\varepsilon oc,Act} + b_2^* \cdot r_{\varepsilon un,Act} = 0.263 + 0.481 = 0.744. \quad (57)$$

The approximate coefficient of determination (57) is very close to the coefficient of determination (56). Thus, the molecular parameters ε_{oc} and ε_{un} actually determine the pharmacodynamic stage of drug action. Regression (55) does not contradict regression (43). As mentioned above, the pair correlation coefficient between the explanatory variables εv and ε_{un} is insignificant: $r_{1,2} = 0.16$. Therefore, it can be recognized that there is no collinearity between the explanatory variables. Since the regression residuals (55) are normally distributed ($W = 0.951 > W_{0.05}^{cr}(N) = 0.945$), the collinearity of the explanatory variables can be quantified using the Farrar-Glauber relation:

$$\chi^2 = -[N - 1 - (2m + 5)/6] \cdot \ln\left(\det\begin{vmatrix} r_{1,1} & r_{2,1} \\ r_{1,2} & r_{2,2} \end{vmatrix}\right) =$$

$$1.10 < \chi_{0.05}^{2,cr}(f = 1) = 3.841. \quad (58)$$

Since inequality (58) is satisfied, we can agree that there is no significant collinearity between the variables at the 95% confidence level.

It is known [5] that elongation of the hydrocarbon chain in the $NH_2(CH_2)_kSH$ molecule for k = 2, 3, 4 leads to a decrease in the radioprotective effect of the chemical compound. Indeed, using the data of Table 1, it can be seen that for preparations Nos. 19 and 29 (k = 3 and 4), the electron-acceptor ability of these compounds noticeably weakens compared to preparation No. 14 (k = 2), the hydrophobic contribution is increased and at the same time the energy $\Delta\varepsilon$ increases. In accordance with equations (42) and (43), these changes can lead to a decrease in radiation protection. In addition, for the chain of carbon-hydrogen atoms $(CH_2)_k$ there is a change in the effective charges of carbon atoms (positive values). The effective charge of an atom characterizes the shift of the electron density along the chemical bond and this is a quantitative measure of the polarization of the chemical bond. The greater the change, the farther the carbon atom is located from the acceptor. Such electron density distribution leads to the appearance of centers with different reactivity. It is possible that the different biological activity of α-homocysteine and β-homocysteine is related to this.

The replacement of the amine group in compound No.14 by a methyl group (No. 35) or by an isoelectronic (in terms of the number of electrons on the outer shell) hydroxyl group (No. 30) changes the electronic properties of the molecules so that it leads to a decrease in their antiradiation action and simultaneously reduces the donor-acceptor properties of the molecules as a whole. The electron affinity (A, eV) values are known for some substituents [46]. It is well known, that the measure of the electron affinity of an atom, molecule, or group of atoms is the amount of energy released when an electron is attached to it. A comparative analysis of the observed electron affinity values for the substituents in the R_1 position of NH_2 (A = 0.74 eV), CH_3 (A = 1.05-1.08 eV), $N(CH_3)_2$ (A = 1.08 eV), $NHCH_3$ (A = 1.56 eV) and OH (A = 1.83 ± 0.04 eV) demonstrates that this sequence is associated with a decrease in the radio-protective effect of the drugs (Nos. 14, 35, 31, 30, 36): 60 (70),

10, 10(40), 10(50), and 0%(0%), respectively. The brackets indicate the radioprotection given in [47]. Therefore, the $R_1CH_2CH_2R_2$ molecule in this case is asymmetric in terms of the energy parameter, that is, in terms of the electron affinity (A) of the substituents. The molecule is, as it were, "polarized" (R_2 = SH, A = 2.32±0.01 eV [46]) by its ability to accept or donate an electron. The greater this energy "asymmetry", the higher the radioprotective effect of the chemical compound.

In the process of irradiation in a living organism, a hydroxyl radical ($OH^{\cdot}$) arises, which is extremely chemically active and destroys almost any molecule it encounters. Acting on SH-groups, histidine and other amino acid residues of proteins, hydroxyl $OH^{\cdot}$ causes denaturation of the latter and inactivates enzymes. In this case, the radioprotector molecule containing the SH group can intercept the hydroxyl molecule. Since the SH group has a high electron affinity, electron transfer from the radical to the radioprotector molecule is possible. In nucleic acids, the $OH^{\cdot}$ radical destroys carbohydrate bridges between nucleotides and, thus, breaks DNA and RNA chains, resulting in mutations and cell death. In addition, the decay of the negative molecular ion produces the $H^{\cdot}$ ion, which has a very high kinetic energy [48]. Apparently, the presence of non-protein SH groups in the molecule is a necessary condition for the effectiveness of low molecular weight aminothiols, but not sufficient. It was established [4] that there is a connection between the protective effect of radioprotectors and the concentration of SH-groups in body tissues.

It can also be noted that the groups of R_1 atoms have a relatively low electron affinity, but a high ionization potential, which is noticeably higher than that of the SH substituent (I = 10.5 eV). For example, the ionization potentials of the NH_2 and OH groups are known [46] to be 11.4 and 13.18 eV, respectively. It is possible that the function of the R_1 substituents is to orient the radioprotector molecule in space (for example, due to intermolecular hydrogen bonding) in such a way that the SH substituent is available for interaction with radicals, and the NH_2 group is oriented in such a way as to participate in the formation of the $NH^{+}\cdots N$ hydrogen bond. There are experimental confirmations [49] that in real biological systems there is a hydrogen bond $NH^{+}\cdots N$. A radioprotector molecule can participate in the formation of such a specific bond, given its specific spatial arrangement. This assumption is supported by the fact that the symmetric $SHCH_2CH_2SH$ molecule (meaning symmetry in terms of the electron affinity energy A) exhibits no radio-protective effect [50], although the electron affinity values of both acceptor substituents R_1 = SH and R_2 = SH are the highest of the substituents presented here. A similar situation exists for S,S-ethyldiisothiuronium and S,S-propyldiisothiuronium molecules. Both of these compounds do not have effective radioprotection [51]. The same characteristic changes are revealed upon passing from compound No. 2 to compound No. 41, when the substituent NH_2 (in position R_1) changes to the isoelectronic group of OH atoms. Replacement of the hydrogen atom (electron affinity of the hydrogen atom A = 0.77 eV [46]) at the amine group NH_2 in chemical compound No.14 with the group of atoms $H_2C{=}CHCH_2$ (electron affinity A ≈ 0.1 - 2.1 eV [46] (comparison of various data; a semi-empirical quantum-chemical calculation gives a value of A = 1.0 eV.)) or per group of CH_3 atoms (electron affinity A = 1.05–1.08 eV [46]) is also accompanied by a decrease in survival (preparations No. 23 and No.36). Molecule No.34 can also be added to this scheme. Replacement of the SH substituent (No. 10) with the isoelectronic OH substituent (No.34), with a lower electron affinity energy, also leads to a decrease in radioprotective activity. In this series of compounds, it is *β*-mercaptoethylamine (cystamine, becaptan, mercamin) that has the best radioprotective properties. Obviously, substituents through chemical bonds affect the electronic distribution in the entire molecule. Addition of the acceptor substituent to the hydrocarbon chain shifts the electron density along the chain of σ-bonds of carbon atoms toward the acceptor, and the greater the shift, the further away from the acceptor the carbon atom is. This, in turn, is accompanied not only by shifts in the electron density in covalent chemical bonds, but also by the energy of molecules, which inevitably affects the physical and chemical properties of the drug.

The lack of radioprotective effect of *iso*-cysteine (an isomer of cysteine) compared to cysteine may be due to the conformational properties of the molecule. The distance between the groups of SH and NH_2 atoms for *iso*-cysteine varies so much in three-dimensional space that this does not allow their donor-acceptor properties to manifest themselves. *Iso*-cysteine does not form mixed disulfides with proteins, but, like radiosensitizers, binds to them by other intermolecular bonds [3]. The substitution of the thiol group SH (electron affinity is 2.32 eV) for the *iso*-electronic (according to the number of valence electrons) OH group (No.34; electron affinity is 1.83 eV) in chemical compound No.14 also noticeably reduces the complexation activity of this compound. Comparing preparations No.30, No.35 with preparations No.14

and No.34, it can be noted that the highest radioprotective activity is achieved if the R_1 substituent in the $R_1CH_2CH_2R_2$ molecule has a relatively low electron affinity (for example, NH_2: A = 0.74 eV; I = 11.4 eV), while the R_2 substituent (for example, SH: A = 2.32 eV; I = 10.4 eV) has a noticeably higher value than the R_1 substituent. A decrease in the affinity energy for the R_2 substituent is accompanied by a decrease in the protection effect. For example, for molecules No.14 (SH substituent, A = 2.32 eV), No.16 (SCN substituent, A = 2.17 eV) and No.25 (SCH_3 substituent, A ≈ (2.0 – 2.5) eV) there is the following radioprotection sequence: 60, 50 and 30%. Moreover, the change in the value of the ionization potential I of substituents has the opposite direction to the change in the value of electron affinity.

It is important to note that among the chemical compounds of Table 1, it is the SH substituent that has the highest (experimentally observed) value of electron affinity among the substituents used here. However, this is only one side of the properties of the molecules associated with the manifestation of the radioprotective effect of the drugs. A large positive value of ε_{un} (1.72 eV) for molecule No.34 indicates that this chemical compound is practically incapable of forming intermolecular complexes due to donor-acceptor interactions. In addition, the large value of the molecular parameter $\Delta\varepsilon$ (10.7 eV) apparently also completely excludes the possibility of the molecule's participation in electron-conformation transitions. The blocking of the SH and NH_2 functional groups (Nos. 25, 26, 28, 31) violates the threshold conditions established above for the energy parameters of the molecules, which correlates with a decrease in the radioprotective effect. The substitution of a hydrogen atom in the SH group is also accompanied by a decrease in the electron affinity of the substituent, which is associated with a change in the radioprotective activity of the drugs: No.14 (R_2 = SH, A = 2.32 eV, A = 60%); No.16 (R_2 = SCN, A = 2.17 eV, A = 50%); No.26 (R_2 = SCH_2CH_3, A = 1.18 eV, A = 20%). For comparison, the electron affinity of the sulfur atom is 2.077 eV. If the addition of other atoms to the sulfur atom, for example, drugs No.14 and No.16, lead to an increase in the electron affinity of the substituent compared to the affinity of the sulfur atom, then, as follows from Table 1, the radioprotective activity of the drug is noticeable. At the same time, the addition of a group of CH_2CH_3 atoms to the sulfur atom (No.26) reduces the substituent's electron affinity to such an extent that it becomes less than the atomic value for sulfur. The hydrophobic properties of the molecule also increase. For such a substituent, this is accompanied by a noticeable decrease in the antiradiation activity of the chemical compound. In addition, the substitution of the hydrogen atom at the amine and thiol groups creates steric hindrances that hinder the participation of these compounds in the processes of electron transfer, intermolecular approach, and conformational selection. In particular, for example, the formation of complexes with charge transfer is most effective at such distances between the reagents when there is a significant overlap of the interacting molecular orbitals.

Comparison of the influence of changes in the factors of equation (55) on the variability of radioprotective activity of sulfur-containing amino acids: cysteine (No.17) and *iso*-cysteine (No.27), the first of which has pronounced antiradiation protection, is of some interest. Without denying the possibility of the participation of cysteine in the defense of the body through other possible mechanisms of protection against intense radiation [6], which are not discussed here, the following circumstance should be noted. Moving the carboxyl group from the α-position to the β-position with respect to the mercapto group unfavorably changes the important energy parameter of the molecule ε_{un}, and the estimates of bioactivities in this case differ by more than a factor of two using regression (55).

The SH and CH groups can participate in the formation of an intermolecular hydrogen bond, and the bond strength is characterized by the following sequence: OH > NH > SH > CH. The SH substituent belongs to the classical proton donors with participation in the formation of a hydrogen bond. As the hydrogen bonding energy increases, the redistribution of electron density affects all of the atoms of the molecules that make up the molecular complex, which can ultimately lead to profound changes in the physical and chemical properties of substances.

3.2 The relationship of information and electronic features of molecules

It was shown [52] that the information molecular character $dH1 = p_H \cdot \log_2 p_H - p_C \cdot \log_2 p_C$ is related to the biological activity of a chemical compound. Here, the p_H and p_C probabilities determine the proportion of hydrogen and carbon atoms in the molecule. The total molecular information function for a discrete set of atoms is quantified as follows [45]: $H = -\Sigma_i p_i \log_2 p_i$, $p_i = n_i/N$, n_i is the number of atoms of sort i; N is the total number of atoms in the molecule. The summation is performed over all sorts of atoms in the molecule.An analysis of the interrelations of molecular factors showed that the

information function $dH1$ is related to the value of the electronic energy ε_{un}. The following regression was obtained for bioactive chemical compounds (Nos.1-15) (radioprotective activity is equal to $A_1 \geq 60\%$):

$\varepsilon_{un}(dH1)_1 = a_{01} + a_{11} \cdot dH1_1$, $N_1 = 15$, $R_1 = 0.86 \pm 0.07$, $R_1^* = 0.87 > R_{0.05}^{cr}(N_1 - 2) = 0.514$; estimation of the significance of the correlation coefficient, taking into account Hotelling's corrections [11]: $u_H = 1.214 > u_{0.05}(N_1) = 0.523$; the minimum sample size sufficient for the reliability of the correlation coefficient: $N_{0.05}^{min} = 5$; $RMSE(S_1) = 0.951$, $a_{01} = -0.92 \pm 0.28$, $a_{11} = 21.65 \pm 3.53$, $t(a_{11}) = 6.14 > |t(a_{01})| = 3.25 > t_{0.05}^{cr}(N_1 - 2) = 2.16$; $F = 37.67 > F_{0.05}^{cr}(f_1 = 1; f_2 = 13) = 4.67$; sum of the residuals squares: $\Sigma_1 = 11.75$; the Wilk-Shapiro normality test for the residuals: $W = 0.962 > W_{0.05}^{cr}(N_1) = 0.881$; straightness sign: $K = 1.97 < K^{thr} = 3.00$ [18]. (59)

Similarly, we write a linear regression for a sample containing chemical compounds Nos.16-25 (the bioactivity is equal to $A_2 = 50\%$):

$\varepsilon_{un}(dH1)_2 = a_{02} + a_{12} \cdot dH1_2$, $N_2 = 9$, $R_2 = 0.82 \pm 0.12$, $R_2^* = 0.84 > R_{0.05}^{cr}(N_2 - 2) = 0.666$; assessment of the significance of the correlation coefficient, taking into account the Hotelling corrections: $u_H = 1.037 > u_{0.05}(N_2) = 0.693$; the minimum sample size sufficient for the reliability of the correlation coefficient: $N_{0.05}^{min} = 6$; $RMSE(S_2) = 0.383$, $a_{02} = -0.60 \pm 0.16$, $a_{12} = 13.69 \pm 3.61$, $t(a_{12}) = 3.80 > |t(a_{02})| = 3.70 > t_{0.05}^{cr}(N_2 - 2) = 2.365$; $F = 14.4 > F_{0.05}^{cr}(f_1 = 1; f_2 = 7) = 5.59$; sum of the residuals squares: $\Sigma_2 = 1.027$; the Wilk-Shapiro normality test for the residuals: $W = 0.954 > W_{0.05}^{cr}(N_2) = 0.829$; straightness sign: $K = 1.72 < K^{thr} = 3.00$. (60)

For small sample sizes $N \leq 15$ the best estimate [11] of the correlation coefficient is $R^* = R \cdot [1 + 0.5(1 - R^2)/(N - 3)]$. Let's check whether the two regressions (59) and (60) can be combined into one regression, i.e. the same relationship of signs for these samples or different. To do this, we use the Chow test [53]. We first obtain the regression for the combined sample, i.e. including the populations A_1 and A_2:

$\varepsilon_{un}(dH1_{1+2}) = a_0 + a_1 \cdot dH1_{1+2}$, $N = 24$, $R = 0.89 \pm 0.04$, $R^* = 0.90 > R_{0.05}^{cr}(N - 2) = 0.404$; the minimum sample size sufficient for the reliability of the correlation coefficient: $N_{0.05}^{min} = 5$; $RMSE(S) = 0.783$, $a_0 = -0.89 \pm 0.16$, $a_1 = 21.69 \pm 2.37$, $t(a_1) = 8.93 > |t(a_0)| = 5.45 > t_{0.05}^{cr}(N - 2) = 2.074$; $F = 79.7 > F_{0.05}^{cr}(f_1 = 1; f_2 = 22) = 4.30$; sum of the residuals squares: $\Sigma = 13.50$, normality of residual distribution (Wilk-Shapiro test): $W = 0.966 > W_{0.05}^{cr}(N) = 0.918$; straightness sign: $K = 2.14 < K^{thr} = 3.00$. (61)

Population statistics of ε_{un}:

$N = 24$; $\varepsilon_{un}^{av} = -1.19 \pm 0.34$; 95% confidence interval: (-1.88, -0.49), $\varepsilon_{un}^{min} = -4.70$, $\varepsilon_{un}^{max} = 0.50$; $S_{un} = 1.647$; $\tau^{max} = 1.03 < \tau^{min} = 2.13 < \tau_{0.05}^{cr,2}(N) = 2.701 < \tau_{0.05}^{cr,1}(N) = 2.800$; the Pearson normality test: $\chi^2 = 3.81 < \chi_{0.05}^{2,cr}(df = 12) = 21.026$, the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N) = 3.25 \approx U = [(\varepsilon_{un}^{max} - \varepsilon_{un}^{min})/S_{нс}] = 3.16 < U2_{0.05}^{cr}(N) = 4.60$; $N_{repr} = 19$; (62)

population statistics of $dH1_{1+2}$:

$N = 24$, $dH1_{1+2}^{av} = -0.014 \pm 0.014$; 95% confidence interval (-0.043, 0.015), $dH1_{1+2}^{min} = -0.147$, $dH1_{1+2}^{max} = 0.079$; $S_{dH11+2} = 0.069$; $\tau^{max} = 1.35 < \tau^{min} = 1.93 < \tau_{0.05}^{cr,2}(N) = 2.701 < \tau_{0.05}^{cr,1}(N) = 2.800$; the Pearson normality test: $\chi^2 = 0.83 < \chi_{0.05}^{2,cr}(df = 13) = 22.362$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N) = 3.25 < U = [(dH1_{1+2}^{max} - dH1_{1+2}^{min})/S_{dH11+2}] = 3.28 < U2_{0.05}^{cr}(N) = 4.60$; $N_{repr} = 19$. (63)

The Chow test is defined by the following inequality:

$$F = [(\Sigma - \Sigma_1 - \Sigma_2) \cdot (N - 2m - 2)] \times [(\Sigma_1 + \Sigma_2) \cdot (m + 1)]^{-1} = 0.566 < F_{0.05}^{cr}(f_1 = m + 1; f_2 = N - 2m - 2) = 3.49. \quad (64)$$

Inequality (64) indicates, first, that the two regressions can be combined into a single regression, and, second, that there is no structural shift in the relationship between the energy ε_{un} and the attribute $dH1$. Both regressions (59) and (60) are statistically significant. The combined regression is of higher quality than the separate regressions (59) and (60). According to the Cheddock scale [19], the linear relationship between the attributes ε_{un} and $dH1$ (59) and (60) is characterized as "very close". At the same time, for the region of weak bioactivity $A_3 \leq 30\%$, sample volume $N_3 = 21$ (Nos. 25-45) there is no relationship between the signs. The linear correlation coefficient is insignificant: $R_3 = 0.09 < R_{0.05}^{cr}(f = 19) = 0.433$; $F = 0.16 << F_{0.05}^{cr}(f_1 = 1; f_2 = 19) = 4.38$. In this case, the events are mutually independent for any pair of random values ε_{un} and $dH1$. Thus, there is a structural shift in the relationship between the signs of ε_{un} and $dH1$ when moving from bioactive to inactive or weakly active drugs. It can be assumed that such a shift in the relationships is associated with a change in the anti-radiation activity of chemical compounds.

The statistics of populations $dH1_1$, $dH1_2$, $dH1_3$ will be as follows:

$A_1 \geq 60\%$: $N_1 = 15$, $dH1_1^{av} = -0.039 \pm 0.019$; 95% confidence interval: (- 0.079, 0.0005), $dH1_1^{min} = -0.147$, $dH1_1^{max} = 0.066$; $S_{dH11} = 0.072$, $\tau^{max} = 1.46 < \tau^{min} = 1.50 < \tau_{0.05}^{cr,2}(N_1) = 2.493 < \tau_{0.05}^{cr,1}(N_1) = 2.617$; the Wilk-Shapiro normality test: $W = 0.905 > W_{0.05}^{cr}(N_1) = 0.881$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N_1) = 2.970 = U = [(dH1_1^{max} - dH1_1^{min})/S_{dH11}] = 2.96 < U2_{0.05}^{cr}(N_1) = 4.17$; $N_{repr} = 12$;

$A_2 = 50\%$: $N_2 = 9$, $dH1_2^{av} = 0.028 \pm 0.013$; 95% confidence interval: (-0.001, 0.057), $dH1_2^{min} = -0.024$, $dH1_2^{max} = 0.079$; $S_{dH12} = 0.038$, $\tau^{max} = 1.34 < \tau^{min} = 1.37 < \tau_{0.05}^{cr,2}(N_2) = 2.493 < \tau_{0.05}^{cr,1}(N_2) = 2.617$; the Wilk-Shapiro normality test: $W = 0.937 > W_{0.05}^{cr}(N_2) = 0.829$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N_2) = 2.59 < U = [(dH1_2^{max} - dH1_2^{min})/S_{dH12}] = 2.71 < U2_{0.05}^{cr}(N_2) = 3.552$; $N_{repr} = 7$;

$A_3 \leq 30\%$: $N_3 = 21$, $dH1_3^{av} = 0.028 \pm 0.011$; 95% confidence interval: (0.006 - 0.050), $dH1_3^{min} = -0.038$, $dH1_3^{max} = 0.110$; $S_{dH13} = 0.049$, $\tau^{min} = 1.35 < \tau^{min} = 1.34 < \tau_{0.05}^{cr}(N_3) = 2.64$; the Wilk-Shapiro normality test: $W = 0.890 \approx W_{0.05}^{cr}(N_3) = 0.908$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N_3) = 3.18 \approx U = [(dH1_3^{max} - dH1_3^{min})/S_{dH13}] = 3.02 < U2_{0.05}^{cr}(N_3) = 4.49$; $N_{repr} = 17$. (65)

Populations $dH1_{i=1,2,3}$ are homogeneous and have a distribution of elements close to normal. Let us check the significance of the difference between the average values of the information function $dH1^{av}$ for the bioactivity areas A_1, A_2, and A_3. For the compared regions A_1 and A_2, A_2 and A_3, in accordance with relations (5), (6), (8) and (9), the following inequalities were obtained:

$$F_{1,2} = S_{dH11}^2/S_{dH12}^2 = 3.67 > F_{0.05}^{cr}(f_1 = N_1 - 1; f_2 = N_2 - 1) = 3.52,$$
$$t = |dH1_1^{av} - dH1_2^{av}| = 0.065 > T^{av} = 0.041, \quad (66)$$

$$F_{3,2} = S_{dH13}^2/S_{dH12}^2 = 1.69 < F_{0.05}^{cr}(f_3 = N_3 - 1; f_2 = N_2 - 1) = 3.44,$$
$$t = |dH1_2^{av} - dH1_3^{av}| = 0.057 > t^{av} = 0.024. \quad (67)$$

Inequalities (66) and (67) indicate that the mean values of the information function $dH1$ are significantly different for regions A_1 and A_2 and for regions A_3 and A_2. Thus, the information function $dH1$, as well as the quantum molecular signatures ε_{un} and ε_{oc}, allows us to separate bioactive drugs from weakly active or inactive chemical compounds. Consequently, the electronic sign ε and the information function $dH1$, derived from different representations of the molecular structure, do not contradict each other.

It can also be shown that the regression coefficients a_{11} (59) and a_{12} (60) differ statistically insignificantly. Let us preliminarily check whether the variances of the residuals differ significantly [8]. The verification is carried out using a relation that has an F-distribution:

$$F = (S_1/S_2)^2 = 2.73 < F_{0.05}^{cr}(f_1 = N_1 - 2; f_2 = N_2 - 2) = 3.55. \quad (68)$$

The numerator (68) has a large dispersion. This result also does not contradict the Romanovsky test [54]: $Q = S_1^2 \cdot (N_1 - 3)/[S_2^2 \cdot (N_1 - 1)] = 2.05$, $S_\Xi = \{2 \cdot (N_1 + N_2 - 4)/[(N_2 - 1) \cdot (N_1 - 5)]\}^{0.5} = 0.845$, $\Xi = |Q - 1|/S_\Xi = 1.24 < 3.0$. There is a large despersion in the numerator for Q. In this case, you can use the following relation to estimate the difference between the regression coefficients a_{11} and a_{12} [8]:

$$S^2 = [(N_1 - 2) \cdot S_1^2 + (N_2 - 2) \cdot S_2^2]/(N_1 + N_2 - 4),$$
$$\Omega_{12} = 1/[(N_1 - 1) \cdot S_{dH11}^2] + 1/[(N_2 - 1) \cdot S_{dH12}^2],$$
$$t = |a_{11} - a_{12}|/(S^2 \cdot \Omega_{12})^{0.5} = 0.75 < t_{0.05}^{cr}(f = N_1 + N_2 - 4) = 2.08. \quad (69)$$

Inequality (69) allows us to agree with the null hypothesis that the regression coefficients a_{11} and a_{12}, which determine the slope of the lines, differ insignificantly from each other.

Let us also compare the correlation coefficients [8]:

$$\Lambda = |z_1 - z_2| \cdot [(N_1 - 3)^{-1} + (N_2 - 3)^{-1}]^{-0.5} = 0.272 < \Lambda_{0.05}^{cr} = 1.96, \quad (70)$$

here $z = 0.5 \cdot \ln[(1 + R)/(1 - R)] = 1.1513 \cdot \lg[(1 + R)/(1 - R)]$ is the normalizing Fisher transform [11] for the correlation coefficient R. It follows from inequality (70) that the correlation coefficients of the regressions also do not differ significantly. Using the values of z_1 and z_2, let us test the hypothesis that the composite estimate (com) of the correlation coefficient is different from zero:

$$z^{com} = [z_1 \cdot (N_1 - 3) + z_2 \cdot (N_2 - 3)] \times (N_1 + N_2 - 6)^{-1} = 0.941. \quad (71)$$

The test is carried out with the help of the following ratio, which has a normal distribution:

$$\Lambda = z^{com} \cdot [N_1 + N_2 - 6]^{0.5} = 3.99 > \Lambda_{0.05}^{cr} = 1.96. \quad (72)$$

Inequality (72) suggests that there is a significant relationship between the molecular features ε_{ns} and $dH1$ for the bioactivity regions A_1 and A_2 at the 5% level of significance. This result does not contradict

the conclusion that follows from (69). Thus, it is advisable to split the total sample into two parts only if the decrease in variance is significantly greater than the remaining unexplained variance when using two regressions.

Further analysis showed that the electronic energies ε_{un} and $\Delta\varepsilon$ for a number of chemical compounds from Table 1 are very closely related to each other:

$\Delta\varepsilon(\varepsilon_{un}) = a_0 + a_1 \cdot \varepsilon_{un}$, $N = 45$, $R = 0.92 \pm 0.02$, $R > R_{0.05}^{cr}(N - 2)$; the minimum sample size sufficient for the reliability of the correlation coefficient: $N_{0.05}^{min} < 5$; $RMSE = 0.581$, $a_0 = 8.74 \pm 0.09$, $a_1 = 0.94 \pm 0.06$, $t(a_0) = 94.6 > t(a_1) = 15.43 > t_{0.05}^{cr}(N - 2) = 2.014$; $F = 238.1 > F_{0.05}^{cr}(f_1 = 1; f_2 = 43) = 4.08$; regression residuals are normally distributed (the Wilk-Shapiro test): $W = 0.960 > W_{0.05}^{cr}(N) = 0.945$; sum of residuals squares: $\Sigma = 14.55$; straightness sign: $K = 2.63 < K^{thr} = 3.00$. (73)

Now let's check whether the variational series has a structural shift when moving from bioactive chemical compounds (region A_1) to relatively weakly bioactive drugs ($A_2 = 50\%$). The following two regressions were obtained for bioactive chemical compounds (Nos. 1-15):

$\Delta\varepsilon(\varepsilon_{un})_1 = a_{01} + a_{11} \cdot \varepsilon_{un}$, $N_1 = 15$, $R_1 = 0.95 \pm 0.03$, $R_1^* = 0.96 > R_{0.05}^{cr}(N_1 - 2) = 0.514$; the minimum sample size sufficient for the reliability of the correlation coefficient: $N_{0.05}^{min} < 5$; $RMSE_1 = 0.424$; $a_{01} = 7.81 \pm 0.16$, $a_{11} = 0.67 \pm 0.06$, $t(a_{01}) = 60.0 > t(a_{11}) = 10.6 > t_{0.05}^{cr}(N_1 - 2) = 2.160$; $F = 112.0 > F_{0.05}^{cr}(f_1 = 1; f_2 = 13) = 4.67$; regression residuals are normally distributed (the Wilk-Shapiro test): $W = 0.976 > W_{0.05}^{cr}(N_1) = 0.881$; sum of squares of residuals: $\Sigma_1 = 2.348$; straightness sign: $K = 1.21 < K^{thr} = 3.00$, (74)

and for weak drugs (Nos. 16-24):

$\Delta\varepsilon(\varepsilon_{un})_2 = a_{02} + a_{12} \cdot \varepsilon_{un}$, $N_2 = 9$, $R_2 = 0.92 \pm 0.06$, $R^* = 0.93 > R_{0.05}^{cr}(N_2 - 2) = 0.666$; estimation of the significance of the correlation coefficient, taking into account the Hotelling corrections: $u_H = 1.431 > u_{0.05}(N_2) = 0.693$; the minimum sample size sufficient for the reliability of the correlation coefficient: $N_{0.05}^{min} < 5$; $RMSE_2 = 0.211$; $a_{02} = 8.75 \pm 0.08$, $a_{12} = 0.76 \pm 0.12$, $t(a_{02}) = 116.5 > t_{0.05}^{cr}(a_{12}) = 6.40 > t_{0.05}^{cr}(N_2 - 2) = 2.365$; $F = 40.9 > F_{0.05}^{cr}(f_1 = 1; f_2 = 7) = 5.59$; the Wilk-Shapiro test for residuals: $W = 0.904 > W_{0.05}^{cr}(N_2) = 0.829$; sum of squares of residuals: $\Sigma_2 = 0.313$; straightness sign: $K = 1.15 < K^{thr} = 3.00$. (75)

Linear regression for the merged population $N = N_1 + N_2$ has the following statistics:

$\Delta\varepsilon(\varepsilon_{un}) = a_0 + a_1 \cdot \varepsilon_{un}$, $N = 24$, $R = 0.93 \pm 0.03$, $R > R_{0.05}^{cr}(N - 2) = 0.404$; the minimum sample size sufficient for the reliability of the correlation coefficient: $N_{0.05}^{min} < 5$; $RMSE = 0.538$, $a_0 = 8.31 \pm 0.14$, $a_1 = 0.80 \pm 0.07$, $t(a_0) = 61.0 > t(a_1) = 11.73 > t_{0.05}^{cr}(N - 2) = 2.074$; $F = 137.5 > F_{0.05}^{cr}(f_1 = 1; f_2 = 22) = 4.30$; the Wilk-Shapiro test for residuals: $W = 0.915 = W_{0.05}^{cr}(N) = 0.916$; sum of squares of residuals: $\Sigma = 6.368$; the sign of straightness: $K = 1.80 < K^{thr} = 3.00$. (76)

Population statistics $\Delta\varepsilon$ and ε_{un} for pooled samples:

$N = 24$, $\Delta\varepsilon^{av} = 7.37 \pm 0.29$; 95% confidence interval: (6.77-7.96), $\Delta\varepsilon^{min} = 4.08$, $\Delta\varepsilon^{max} = 9.35$; $S_{\Delta 1} = 1.417$; $\tau^{max} = 1.40 < \tau^{min} = 2.32 < \tau_{0.05}^{cr,2}(N) = 2.701 < \tau_{0.05}^{cr,1}(N) = 2.800$; the Wilk-Shapiro test: $W = 0.923 > W_{0.05}^{cr}(N) = 0.918$, the Pearson normality test: $\chi^2 = 1.03 < \chi_{0.05}^{2,cr}(df = 11) = 19.675$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N) = 3.34 < U = [(\Delta\varepsilon^{max} - \Delta\varepsilon^{min})/S_{\Delta 1}] = 3.72 < U2_{0.05}^{cr}(N) = 4.71$; $P = 3.9\%$; $N_{prepr} = 19$,

$N = 24$, $\varepsilon_{un}^{av} = -1.19 \pm 0.34$; 95% confidence interval: (-1.88,-0.49), $\varepsilon_{un}^{min} = -4.70$, $\varepsilon_{un}^{max} = 0.50$; $S_{un} = 1.647$; $\tau^{max} = 1.03 < \tau^{min} = 2.13 < \tau_{0.05}^{cr,2}(N) = 2.701 < \tau_{0.05}^{cr,1}(N) = 2.800$; the Wilk-Shapiro test: $W = 0.819 < W_{0.05}^{cr}(N) = 0.918$, the Pearson normality test: $\chi^2 = 3.81 < \chi_{0.05}^{2,cr}(df = 12) = 21.026$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N) = 3.34 \approx U = [(\varepsilon_{un}^{max} - \varepsilon_{un}^{min})/S_{un}] = 3.20 < U2_{0.05}^{cr}(N) = 4.71$; $N_{prepr} = 19$. (77)

Then the ratio (64) for the Chow test is calculated:

$F = 13.94 > F_{0.05}^{cr}(f_1 = m + 1; f_2 = N - 2m - 2) = 3.49$. (78)

Thus, in accordance with inequality (78), it can be assumed that the relationship between molecular features $\Delta\varepsilon$ and ε_{un} undergoes a statistically significant structural shift in the transition from the A_1 bioactivity region to the A_2 bioactivity region. Therefore, it is not recommended to use regression (76) built on pooled samples to interpret the relationship of features. The difference $\Sigma - \Sigma_1 - \Sigma_2$ is an indicator of the improvement in the quality of the model when the sample size is divided into two parts. Thus, the *null*-hypothesis about the absence of a structural shift in the sample data is rejected. Therefore, for statistical analysis, two samples should not be combined into one, and the transition from region A_1 to region A_2 has a qualitative jump in the relationship of molecular features ε_{un} and $\Delta\varepsilon$. Similarly, we check for the presence of a structural shift for the relationship of explanatory variables for samples from areas A_2 and A_3. For inactive or

weakly bioactive chemical compounds (Nos. 25-45), linear regression has the following statistics:

$\Delta\varepsilon(\varepsilon_{un})_3 = a_{03} + a_{13}\cdot\varepsilon_{un}$, $N_3 = 21$, $R_3 = 0.79 \pm 0.09$, $R_3^* = 0.80 > R_{0.05}^{cr}(N_3 - 2) = 0.433$; estimation of the significance of the correlation coefficient, taking into account the Hotelling corrections: $u_H = 1.023 > u_{0.05}(N_3) = 0.438$; the minimum sample size sufficient for the reliability of the correlation coefficient: $N_{0.05}^{min} = 6$; $RMSE = 0.417$; $a_{03} = 9.07 \pm 0.10$, $a_{13} = 0.86 \pm 0.15$, $t(a_{03}) = 93.18 > t(a_{12}) = 5.63 > t_{0.05}^{cr}(N_3 - 2) = 2.093$; $F = 31.7 > F_{0.05}^{cr}(f_1 = 1; f_2 = 19) = 4.38$; the Wilk-Shapiro test for the residuals: $W = 0.970 > W_{0.05}^{cr}(N_3) = 0.908$; the sum of the squares residuals: $\Sigma_3 = 3.305$; straightness sign: $K = 2.75 < K^{thr} = 3.00$. (79)

The statistics of the population of elements ε_{un} for area A_3 will be as follows:

$N_3 = 21$, $\varepsilon_{un}^{av} = 0.23 \pm 0.13$; 95% confidence interval: (-0.05, 0.50), $\varepsilon_{un}^{min} = -0.75$, $\varepsilon_{un}^{max} = 1.78$; $S_{un} = 0.611$, $\tau^{min} = 1.60 < \tau^{max} = 2.54 < \tau_{0.05}^{cr,2}(N_3) = 2.644 < \tau_{0.05}^{cr,1}(N_3) = 2.750$; the Wilk-Shapiro normality test: $W = 0.816 < W_{0.05}^{cr}(N_3) = 0.908$; the Pearson normality test: $\chi^2 = 16.6 < \chi_{0.05}^{2,cr}(df = 15) = 24.996$; the David-Hartley-Pearson normality test: $U1_{0.05}^{cr}(N_3) = 3.18 < U = [(\varepsilon_{un}^{max} - \varepsilon_{un}^{min})/S_{un3}] = 4.14 < U2_{0.05}^{cr}(N_3) = 4.49$; $N_{repr} = 17$. (80)

For the combined sample (areas A_2 and A_3) the linear regression statistics will be as follows:

$\Delta\varepsilon(\varepsilon_{un}) = a_0 + a_1\cdot\varepsilon_{un}$, $N = 30$, $R = 0.84 \pm 0.06$, $R^* = 0.85 > R_{0.05}^{cr}(N - 2) = 0.361$; the minimum sample size sufficient for the reliability of the correlation coefficient: $N_{0.05}^{min} = 5$; $RMSE = 0.385$, $a_0 = 8.97 \pm 0.07$, $a_1 = 0.90 \pm 0.11$, $t(a_0) = 126.2 > t(a_1) = 8.06 > t_{0.05}^{cr}(N - 2) = 2.048$; $F = 65.0 > F_{0.05}^{cr}(f_1 = 1; f_2 = 28) = 4.20$; regression residuals are normally distributed (the Wilk-Shapiro test): $W = 0.957 > W_{0.05}^{cr}(N_3) = 0.927$; the sum of the squares residuals: $\Sigma = 4.158$; straightness sign: $K = 2.88 < K^{thr} = 3.00$. (81)

Using relation (69), as well as the results (73), (76) and (81), the following inequality for the Chow test was obtained:

$$F = 1.94 < F_{0.05}^{cr}(f_1 = m + 1; f_2 = N - 2m - 2) = 3.34. \quad (82)$$

Since $F < F^{cr}$, the *null*-hypothesis of structural stability of the variation series at the 95% confidence level should be accepted. Therefore, combining samples A_2 and A_3 into one sample is allowed. That is, for the relationship of molecular features $\Delta\varepsilon$ and ε_{un}, there is a qualitative and quantitative structural shift in the transition from bioactive drugs (A_1 region) to inactive or relatively weakly active drugs (A_2 and A_3 regions). Thus, the relationships between molecular features $\Delta\varepsilon$ and ε_{un} for bioactive and inactive chemical compounds differ significantly. Figure 2 clearly shows the structural shift for the relationship of signs $\Delta\varepsilon$ and ε_{un}, which separates bioactive chemical compounds from inactive or relatively weakly active drugs. There are also two lines of regression equations (74) and (81). Thus, taking into account the results (64), (74)-(76) and (79), we can assume the existence of homogeneous or heterogeneous information arrays. Taking into account the significant relationship between the features $\Delta\varepsilon$ and ε_{un} (73), as well as significant statistics (59) and (60), structural changes can also be found in the relationships of the feature $\Delta\varepsilon$ with the molecular features Z [52] and $dH1$ (59). The molecular feature Z is associated with the pseudopotential of the molecule [45,55].

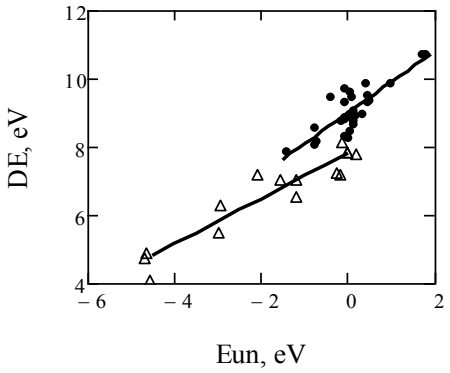

Fig.2. Scatterplots for bioactive drugs (Δ) and for inactive or weakly active chemical compounds (•).1 - linear regression (74). 2 - linear regression (81). The following designations are used here: DE ≡ $\Delta\varepsilon$, Eun ≡ ε_{un}.

It is important to check the presence of such a relationship, since the sign $\Delta\varepsilon$ (or sign ε_{un}) and the molecular signs Z and $dH1$ were obtained for samples based on different physical concepts of the molecular structure. As shown in [45], the $dH1$ factor correlates with the hydrophobic properties of molecules, that is, it is associated with the pharmacodynamic stage of drug action. Let's check the relationship between the signs $\Delta\varepsilon$ and Z. For bioactive chemical compounds ($A_1 \geq 60\%$), the following straight-line regression was obtained:

$\Delta\varepsilon(Z)_1 = a_{01} + a_{11}\cdot Z_1$, $N_1 = 15$, $R_1 = -0.82 \pm 0.09$, $|R_1^*| = 0.83 > R_{0.05}^{cr}(N_1 - 2) = 0.514$; estimation of the significance of the correlation coefficient, taking into account Hotelling's corrections: $u_H = 1.086 > u_{0.05}(N_1) = 0.523$; the minimum sample size sufficient for the reliability of the correlation

coefficient: $N_{0.05}^{min} = 6$; $RMSE = 0.763$; $a_{01} = 14.1 \pm 1.48$, $a_{11} = -2.60 \pm 0.51$, $t(a_0) = 9.52 > |t(a_1)| = 5.07 > t_{0.05}^{cr}(N_1 - 2) = 2.160$; $F = 25.7 > F_{0.05}^{cr}(f_1 = 1; f_2 = 13) = 4.67$; the Wilk-Shapiro test for regression residuals: $W = 0.887 > W_{0.05}^{cr}(N_1) = 0.881$; the sum of the residuals squares: $\Sigma_1 = 7.581$; the sign of straightness: $K = 2.16 < K^{thr} = 3.00$. (83)

The statistics of the Z_1 population will be as follows:

$N_1 = 15$, $Z_1^{av} = 2.85 \pm 0.10$; 95% confidence interval: (2.63-3.07), $Z_1^{min} = 2.286$, $Z_1^{max} = 3.60$, $S_{Z1} = 0.397$, $\tau^{min} = 1.42 < \tau^{max} = 1.89 < \tau_{0.05}^{cr,2}(N_1) = 2.493 < \tau_{0.05}^{cr,1}(N_1) = 2.617$; the Wilk-Shapiro normality test: $W = 0.950 > W_{0.05}^{cr}(N_1) = 0.881$; $V = 13.9\%$; $P = 3.6\%$; $N_{repr} = 12$; $\Theta = 27.8$. (84)

Linear regression for the region $A_2 = 50\%$:

$N_2 = 9$, $\Delta\varepsilon(Z)_2 = a_{02} + a_{12} \cdot Z_2$, $R_2 = - 0.68 \pm 0.20$, $|R_2^*| = 0.72 > R_{0.05}^{cr}(N_2 - 2) = 0.666$; estimation of the significance of the correlation coefficient, taking into account Hotelling's corrections: $u_H = 0.741 > u_{0.05}(N_2) = 0.692$; the minimum sample size sufficient for the reliability of the correlation coefficient: $N_{0.05}^{min} = 8$; $RMSE = 0.405$, $a_{02} = 12.2 \pm 1.46$, $a_{12} = -1.39 \pm 0.57$, $t(a_0) = 8.32 > |t(a_1)| = 2.45 > t_{0.05}^{cr}(N_2 - 2) = 2.365$; $F = 6.02 > F_{0.05}^{cr}(f_1 = 1; f_2 = 7) = 5.59$; the Wilk-Shapiro test for regression residuals: $W = 0.943 = W_{0.05}^{cr}(N_2) = 0.829$; the sum of the residuals squares: $\Sigma_1 = 1.150$; the sign of straightness: $K = 2.08 < K^{thr} = 3.00$. (85)

For the population N_2, the Z_2 statistics will be as follows:

$N_2 = 9$, $Z_2^{av} = 2.56 \pm 0.08$; 95% confidence interval (2.37-2.76), $Z_2^{min} = 2.286$, $Z_2^{max} = 3.00$; $S_{Z2} = 0.252$; $\tau^{min} = 1.10 < \tau^{max} = 1.73 < \tau_{0.05}^{cr,2}(N_2) = 2.237 < \tau_{0.05}^{cr,1}(N_2) = 2.392$; the Wilk-Shapiro normality test: $W = 0.927 > W_{0.05}^{cr}(N_2) = 0.829$; $V = 9.8\%$; $P = 3.3\%$; $N_{repr} = 8$; $\Theta = 30.5$. (86)

Comparing linear pairwise regressions (83) and (85) we can note the decrease in the quality of regressions when the bioactivity of chemical compounds decreases. Let's check if the two regressions (83) and (85) are significantly different. Again, we will use the Chou test (64). Let's previously the regression for the combined sample:

$\Delta\varepsilon(Z) = a_0 + a_1 \cdot Z$, $N = 24$, $R = - 0.79 \pm 0.08$, $|R^*| = 0.80 > R_{0.05}^{cr}(N - 2) = 0.404$; the minimum sample size sufficient for the reliability of the correlation coefficient $N_{0.05}^{min} = 6$; $RMSE = 0.882$; $a_{01} = 15.7 \pm 1.37$, $a_1 = - 3.02 \pm 0.50$, $t(a_0) = 11.43 > |t(a_1)| = 6.1 > t_{0.05}^{cr}(N - 2) = 2.074$; $F = 37.2 > F_{0.05}^{cr}(f_1 = 1; f_2 = 22) = 4.30$; the Wilk-Shapiro test for regression residuals: $W = 0.957 > W_{0.05}^{cr}(N) = 0.916$; the sum of the residuals squares: $\Sigma = 17.15$; the sign of straightness: $K = 2.93 < K^{thr} = 3.00$. (87)

The statistics of the population Z will be as follows:

$N = 24$, $Z^{av} = 2.74 \pm 0.08$; 95% confidence interval: (2.59-2.90), $Z^{min} = 2.286$, $Z^{max} = 3.60$; $S_Z = 0.372$; $\tau^{min} = 1.22 < \tau^{max} = 2.31 < \tau_{0.05}^{cr,2}(N) = 2.701 < \tau_{0.05}^{cr,1}(N) = 2.800$; the Wilk-Shapiro normality test: $W = 0.929 > W_{0.05}^{cr}(N) = 0.916$; $V = 13.6\%$; $P = 2.8\%$; $N_{repr} = 19$; $\Theta = 36.1$. (88)

Using the statistics (83)-(87) we check the Chow test (64):

$$F = 9.64 > F_{0.05}^{cr}(f_1 = 2; f_2 = 20) = 3.49. \quad (89)$$

Inequality (89) allows us to reject the *null*-hypothesis and admit that regressions (83) and (85) are significantly different. This is also indicated by the sign of straightness of the combined sample K (87), which practically coincides with the threshold value. The tendency to reduce the radioprotective activity of chemical compounds (Table 1) is accompanied by a tendency to reduce the quality of regression equations. Moreover, for a sample containing inactive or weakly active drugs ($N_3 = 21$), the relationship between molecular features $\Delta\varepsilon$ and Z decreases almost to zero (correlation coefficient $R = 0.03$). Thus, in this case, when the electronic attribute Z is used as an explanatory variable, there is a structural shift for the relationship between the factors $\Delta\varepsilon$ and Z during the transition from active to inactive chemical compounds in terms of radioprotection. The result obtained does not contradict the statistical conclusions (73), (76) and (80). It is important to note that the resulting variables in (73) - (80) and the explanatory variables ε_{un} and Z in (84) - (86) were obtained based on completely different ideas about the structure of the molecule. The sign ε_{un} is determined using quantum mechanical calculations of the electronic structure of molecules, the sign Z is associated with the pseudopotential of the molecule, and the signs H and $dH1$ are informational functions of the molecule.

The list of chemical compounds included compounds for which a noticeable antiradiation protective effect could be expected, but, nevertheless, these drugs in practice are not effective radioprotectors. One of the possible reasons for limiting biological activity is the processes associated with the hydrophobic properties of molecules. One of the possible reasons

for limiting biological activity is the processes associated with the hydrophobic properties of molecules. For this reason, such chemical compounds as, for example, Nos. 28, 29, 36, 44 may be ineffective in terms of radioprotection. The results presented in Table 6 indicate the existence of a relationship between the radioprotective effect of drugs and their molecular informational sign *dH*1. There are two qualitative assessments for *dH*1. For active drugs ($A_1 \geq 60\%$), the *dH*1 value is predominantly negative, while for inactive or weakly active chemical compounds, the *dH*1 value is positive. The *chi-square* test at a significance level of $\alpha = 0.05$, as well as the contingency coefficients *K* and *C*, make it possible to draw a statistically justified conclusion about the presence of a statistical relationship between the radioprotective effect and the value of the molecular information sign *dH*1 [52]. Indeed, since $\chi^2 > \chi^{2,cr}$ (Table 6), then with a probability of 0.95 we can accept the hypothesis of the existence of a relationship between the resulting feature (bioactivity) and the explanatory sign *dH*1.

Table 6
Relationship between the radioprotective effect of substituted aminothiols and their analogues and the information factor *dH*1.

A, %	Sign *dH*1, *bits*		
	The negative	The positive	Total
≥ 60	$q_{11} = 11$ $q_{11}' = 6.67$	$q_{12} = 4$ $q_{12}' = 8.33$	$q_1 = 15$ $p_1 = 0.33$ $q_1' = 15$
= 50	$q_{21} = 2$ $q_{21}' = 4.00$	$q_{22} = 7$ $q_{22}' = 5.00$	$q_2 = 9$ $p_2 = 0.20$ $q_2' = 9$
≤ 30	$q_{31} = 7$ $q_{31}' = 9.33$	$q_{23} = 14$ $q_{23}' = 11.67$	$q_3 = 21$ $p_3 = 0.47$ $q_3' = 21$
	$Q_1 = 20$ $P_1 = 0.444$	$Q_2 = 25$ $P_2 = 0.556$	$N = 45$ $\sum_{i=1}^{3} P_i = \sum_{j=1}^{3} p_j$ $=1.00$
Statistics of the contingent signs $\chi^2 = 7.91 > \chi_{0.05}^{2,cr}(f = 2) = 5.99$, $\phi = 0.419$, $K = 0.507$			

Similarly, it can be shown that the independently determined molecular sign *Z* (Table 1) is also significantly related to the energy interval $\Delta\varepsilon$ for bioactive chemical compounds (region A_1). A linear regression equation was obtained, for which the correlation coefficient turned out to be significant and, accordingly, has the value $|R^*| = 0.83 > R_{0.05}^{cr}(N_1 - 2) = 0.514$; $F = 14.4 > F_{0.05}^{cr}(f_1 = 1; f_2 = 13) = 4.67$. Evidence of linearity of regression: $K = 2.16 < K^{cr} = 3.0$. At the same time, for area A_3 (sample size N_3), there is no relationship between features $\Delta\varepsilon$ and *Z*. The correlation coefficient is insignificant and equals $|R| = 0.04 < R_{0.05}^{cr}(N_3 - 2) = 0.433$. Thus, in this case, when using the sign *Z* as an explanatory variable, there is a structural shift in the relationship between $\Delta\varepsilon$ and *Z* during the transition from bioactive chemical compounds to inactive ones. Figures 3A and 3B show significant relationships between informational molecular features (*H*, *dH*1) and pseudopotential feature *Z*, which are evaluated using different ideas about the molecular structure.

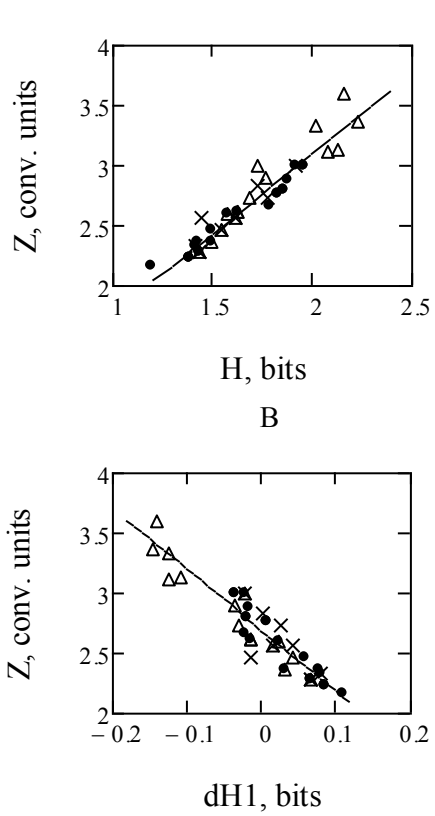

Fig. 3. Δ – area $A_1 \geq 60\%$; × – area $A_2 = 50\%$; • – area $A_3 \leq 30\%$. A. The linear regression: $Z(H) = a + b \cdot H$, $N = 45$, $a = 0.46 \pm 0.11$, $b = 13.22 \pm 0.06$, *RMSE* = 0.101, $R = 0.96 \pm 0.04$. B. The linear regression: $Z(dH1) = a + b \cdot dH1$, $N = 45$, $a = 2.69 \pm 0.02$, $b = -5.09 \pm 0.26$, $R = -0.95 \pm 0.05$, *RMSE* = 0.119.

Obviously, in this case, the relationship is characterized by a homogeneous variance of the random error of the regression model. Taking into account the close relationship (Figures.3A and 3B)

of molecular features: the factor *Z*, the total information function *H* (associated with the diversity of the molecular structure) and the partial information function *dH*1, as well as the relationship (83) - (86), it is possible to establish relationships $\Delta\varepsilon(H)$ and $\Delta\varepsilon(dH1)$ for areas of bioactivity A_1, A_2 and A_3.

The initial sample did not include, for example, such chemical compounds: mercamine ascorbate (dose 0.59 mM/kg; protection 70%), mercamin nicotinate (dose 0.76 mM/kg; protection 0%) or 1-amino-3-mercaptopropane (dose 2.26 mM/kg; protection 10%) [56]. For these compounds, quantum chemical calculations of the electronic structure of molecules have not been performed. However, an approximate theoretical estimate of their radioprotective activity can be obtained. The information function of *dH*1 was calculated for these drugs: -0.025, 0.023 and 0.066 *bits*. These estimates do not contradict the results given in Table 6.

4 Conclusion

The discovered interrelations of molecular factors with the radioprotective action of low molecular weight aminothiols and their analogs demonstrated that the bioactivity of drugs is complex and depends on a combination of various molecular factors. These factors determine the possible participation of radioprotectors in the primary radiation physicochemical processes occurring in the body, increasing its radioresistance. It is important to note that the energy factors of molecules are characterized by some threshold values that separate highly active radiation injury modifiers from weakly active drugs.

The relationships established make it possible to assess the radioprotective effectiveness of a chemical compound of a number of substituted aminothiols and their analogues without performing complex and cumbersome quantum mechanical calculations of the electronic structure of molecules. It is essential that the information molecular features, as well as the electronic factor *Z*, make it possible to make an approximate assessment of the bioactivity of the drug, having information only about the gross formula of the chemical compound. An important result is also the fact that for the analyzed aminothiols, the quantum mechanical parameter of the molecule $\Delta\varepsilon$ is statistically significantly associated with both the *Z* factor and molecular information functions, which were obtained based on different ideas about the molecular structure, not directly based on quantum mechanical calculations of the electron molecular structures. Statistically significant molecular parameters ε_{oc} and ε_{un} characterize electronic processes in which exogenous molecules can participate as radioprotectors, and a significant molecular factor μ^2 determines the ability of these molecules to accumulate in local areas of the biophase prior to irradiation. It should also be noted that explanatory variables, the evaluation of which is based on the use of different and independent representations, namely, on detailed quantum chemical calculations, the use of partial information functions of molecules, or the pseudopotential method, which plays an important role in the quantum theory of solids, are given in this case, to comparable results.

References:

[1] Rachinsky Yu.F., Slavachevskaya N.M., *Chemistry of Aminothiols and some of Their Derivatives*, Moscow, Leningrad, Khimia, 1965, 295 p. (in Russian).

[2] Eidus L.Kh., *Physical and Chemical Fundaments of Radiobiological Processes and Radiation Protection*, Moscow, 1969 (in Russian).

[3] Romantsev E.F., Blokhina V.D., et al., *Biochemical Fundaments of Radioprotectors Action*, Moscow, Atomizdat, 1980, 168 p. (in Russian).

[4] Shchembelov G.A., Ustinyuk V.M., Mamaev V.M., Ischenko V.M., Gloriozov I.P., Luzhkov V.B., Orlov V.V., Simkin V.Y., Pupyshev V.I., Burmistrov V.N., *Quantum Chemical Methods of Molecule Calculation,* Ed. by V.M. Ustynyuk, Moscow, Khimia, 1980, 256 p. (in Russian).

[5] Sweeney T.R., *A Survey of Compounds from the Antiradiation Drug Development Program*, Washington, 1979.

[6] Romantsev E.F., *Radiation and Chemical Protection*, Moscow, 1968 (in Russian).

[7] Fleiss J.L., *Statistical Methods for Rates and Proportios*, New York – Chichester, John Wiley & Sons, 1981.

[8] Förster E., Rönz B., *Metohden der Korrelations- und Regressionsanalyse*, Verlag Die Wirtschaft, Berlin, 1979.

[9] Jonson N., Leone F., *Statistics and Experimental Design. In Engineering and the Physical Sciences*, Vol. 1, John Wiley & Sons, 1977.

[10] D'Agostino R.B., *The Amer. Statistion*, Vol.24, No.6, 1970, pp.14-15.

[11] Sachs L., *Statistische Auswertungsmethoden*, Springer-Verlag, Berlin, New York, 1972.
[12] Pustylnik E.I., *Statistical Methods of Analysis and Processing of Observations*, Moscow, Nauka, 1968, 288 p. (in Russian).
[13] Szent-Györgyi A., *Introduction to a Submolecular Biology*, Academic Press Inc., New-York, London, 1960.
[14] Kalistratov G.V., Romantsev E.F., *Voprosy of Med. Chem.*, Vol. 11(4), 1965, pp. 38-41. (in Russian).
[15] Kobzar A.I., *Applied Mathematical Statistics. For Engineers and Scientists*, Moscow, Fizmatlit, 2016, 816 p. (in Russian).
[16] Linnik Yu.V., *The Method of Least Squares and Bases of Mathematical-Statistical Theory of Observation Processing*, Moscow, 1962, 350 p. (in Russian).
[17] Kimble G.A., *How to Use (and Misuse) Statistics*, Prentice-Hall, Inc., Englewood, N.J., 1978.
[18] Zaitsev G.N., *Mathematics in Experimental Botany*, Moscow, Nauka, 1990, 295 p. (in Russian).
[19] Chaddock R.E., *Principles and Methods of Statistics*, Boston, NewYork, 1925, 471 p.
[20] Pitra C., Hartwig M., Korner I.J., Malz W., *Stud. Biophys*., Vol.62, 1977, pp. 31-42.
[21] Purdic J.W., *Radiat. Res*., Vol.7, 1979, pp.303 -311.
[22] Broch H., Cabropl D., Vasilescu D., *Int. Quant. Chem*., Vol.7, 1980, pp. 283-291.
[23] Berger A.J.K., *Fed. Proc*., Vol.50, 1981, pp.2723-2728.
[24] Mukhomorov V.K., *WSEAS Transactions on Systems*, Vol.21, 2022, pp. 116-133.
[25] Cochran W.G., Cox C.M., *Experimental Design*, New York, John Willey & Sons, 1957.
[26] Farrar D.E., Glauber R.R., *The Review of Economics and Statistics*, Vol. 49(1), 2018, pp. 92-107.
[27] Hoffman T.A., Ladic J., *Cancer Res*., Vol.21, 1961, pp.474-482.
[28] Hoffman T.A., Ladic J., *Adv. Chem. Phys*., Vol.7, 1964, pp.84-87.
[29] Mukhomorov V.K., *Chem.-Pharm. Journal*, Vol.18 (1), 1984, pp.17-25. (in Russian).
[30] Bacq Z., Alexander P., *Fundamentals of Radiobiology*, Oxford, London, New York, 1966, 562 p.
[31] Hart E., Anbar M., *Hydrated electron*, Moscow, Atomizdat, 1973 (in Russian).
[32] Pikaev A.K., *Solvated Electron in Radiation Chemistry*, Moscow, Nauka, 1969. (in Russian).
[33] Mukhomorov V.K., *High Energy Chemistry*, Vol.18(4), 1984, pp.297-300.
[34] Alpatova N.M., Zabusova S.E., Tomilov A.P., *Russian Chemical Reviews*, Vol. LV(2), 1986, pp. 261-276.
[35] *Modern problems of radiation research*, Ed. by L.H. Eidus, Moscow, Nauka, 1972, 343 p.
[36] *Biochemical pharmacology*, Ed. by P.V. Sergeev, N.L. Shimanovsky, Medical Information Agency, Moscow, 2010 (in Russian).
[37] Slifkin M., *Molecular Interactions*, Vol. 2, Ed. by Ratajczak, W.J.Orville-Thomas, A Wiley-Interscience Publ., Chichester-New York, John Wiley and Sons, 1981.
[38] Vladimirov V.G., Mukhomorov V.K., Strel'nikov Yu.E. and others, *Radiobiology*, Vol. 23(5), 1983, pp.615-618.
[39] Izmozherov N.A., Aybinder N.E., Afonina T.D., *Some theoretical aspects of radiation protection*, Moscow, 1980, pp.17-33. (in Russian).
[40] Nagata Ch., Yamaguchi T., *Radiat. Res*., Vol.73, 1978, pp.430-438.
[41] Likeš J., Laga J., *Zakladni Statisticke Tabulky*, Praha, SNLT, 1978.
[42] Dmitriev E.A., *Mathematical Statistics in Soil Science*, Moscow, Ed. Moscow State University, 1995, 319 p. (in Russian).
[43] Akaike Hirotogu, *IEEE Transactions on Automatic Control*, Vol.19(6), 1974, pp. 716-723.
[44] Schwarz G., *Annals of Statistics*, Vol.6, 1978, pp.461-464.
[45] *Mukhomorov V.K.*, Statistical Modeling of the Bioactivity of Chemical Compounds. Part 1, 350p.; Part 2, 301p., LAP Lambert Academic Publishing RU, 2021 (in Russian).
[46] Gurvich L.V., Karachevtsev G.V., Kondratyev V.N., Lebedev Yu.A., Medvedev V.A., Potapov V.K., Khodeev Yu.S., *Chemical bond breaking energies. Ionization potentials and electron affinity*, Moscow, Nauka, 1974, 351p. (in Russian).
[47] Suvorov N.N., Shashkov V.S., *Chemistry and pharmacology of means for prevention of radiation injuries*, Moscow, Atomizdat, 1975 (in Russian).
[48] Massey H., *Negative Ions*, Cambridge University Press, London - New York - Melbourne, 1976.

[49] Zundel G., *Hydration and Intermolecular Interaction*, Academic Press, New York, London, 1969.
[50] Langendorff H., Stapelton G., *Physiol. Rev*., Vol.33, 1053, p.77.
[51] Husinaga S., Klobukowski M., Sakai Y., *J. Phys. Chem.*, Vol.88(21),1984, pp. 4880-4886.
[52] Mukhomorov V.K., *J. Chem. Eng. Chem. Res*., Vol. 1(1), 2014, pp. 54-65.
[53] Chow Gregory C., *Econometrica*, Vol. 28, 1960, pp. 591-605.
[54] Romanovsky V.I., *Elementary Course of Mathematical Statistics*, Moscow, Gosplanizdat, 1939, 359 p. (in Russian).
[55] Veljkovič V., Lalovič D.I., *Phys. Lett*., Vol.45A(1), 1973, pp.59-60.
[56] Tank L.I., *Proceedings of the Scientific Conference. Thiol Compounds in Medicine*, Kyiv, 1957 (in Russian).