ISSN 1070-3632, Russian Journal of General Chemistry, 2009, Vol. 79, No. 3, pp. 488–495. © Pleiades Publishing, Ltd., 2009.
Original Russian Text © I.G. Tsygankova, S.M. Zhenodarova, 2009, published in Zhurnal Obshchei Khimii, 2009, Vol. 79, No. 3, pp. 498–505.

Structure–Activity Correlation for a Series of Isatin Derivatives, Inhibitors of Caspase 3

I. G. Tsygankova and S. M. Zhenodarova

Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences, ul. Institutskaya 3, Pushchino, Moscow oblast, 142290 Russia
e-mail: tsygan@iteb.ru

Received August 28, 2008

Abstract—The correlation structure– property was used for the calculation of inhibitor activity of a series of isatin derivatives, inhibitors of caspase 3, based on a fragment description of the molecular structure. A large number of simple models of the similar accuracy was obtained. The frequency analysis of the popularity of certain fragments contribution into the models made it possible to reveal the role of individual elements of the molecular structure. This result is important for prediction of the activity of caspase 3 inhibitors.
DOI: 10.1134/S1070363209030244

Caspases, enzymes from the family of cysteine proteinases, are directly involved into the biochemical mechanisms underlying the apoptosis (programmed cell death) [1, 2]. Any disorder in this process can cause various diseases, like myocardial infarction, stroke, neurodegenerative diseases, cancer, etc. [3], therefore many research teams are actively designing drugs capable of affecting (inducing or blocking) apoptosis. Potential targets in these cases, especially when the goal is the prevention of cells death, commonly are caspases [4], and especial interest is directed at caspase 3 playing the key part in the apoptosis process and therefore being an attractive therapeutic target [5]. It turned out that numerous caspases inhibitors of peptide nature were not suitable to pharmaceutical requirements because of their moderate selectivity, weak resistance to the metabolism, and insufficient penetration into cells [6]. This fact favors the development of nonpeptide inhibitors, and among the latter isatin derivatives seem very promising [7–12]. The studies performed both on pure enzymes and in tissue cultures showed that these compounds selectively inhibited caspase 3 and to a lesser extent caspase 7 [10] and could serve as a basis for drugs in treating diseases characteristic of excessive apoptosis. In the course of design of new pharmaceuticals the access to calculated activity estimation of caspase 3 inhibitors would be very favorable. Therefore the goal of this study was estimation of the activity of inhibitors, the isatin derivatives presented in Table 1, based on the correlation structure–property. The experimental data on the inhibiting activity of isatin derivatives (Table 2) were taken from publications [8, 10]. We studied correlation models designed with the use of fragment molecular descriptors that permitted relatively simple calculations and had a clear structural meaning.

Description of models design. To construct the correlation structure–property we employed the simplest fragments of the inhibitor molecules linked to heavy atoms. This relationship consists of a sum of single-fragment contributions linked to an atom of a definite type, and of pair contributions, determined by a type of a pair of atoms and a number of bonds dividing them [13–15].

$$P \approx \sum_t n_t P_t + \sum_{t,m,r} k_{t,m-r} P_{t,m-r} + \text{const.} \quad (1)$$

Here n_t is the number of fragments linked to atoms of a definite type t, $k_{t,m-r}$ is the number of pairs of fragments of t and m type located at a definite distance r. The distance is determined by the number of chemical bonds between the atoms of the corresponding fragments. Fragment contributions P_t, $P_{t,m-r}$ are the parameters of the correlation equation that are established from the experimentally measured activities by the procedure of linear regression. The numbers n_t and $k_{t,m-r}$ are quantitative descriptors of the molecular structure.

Table 1. Compounds included in the set of inhibitors of caspase 3 based on isatin derivatives

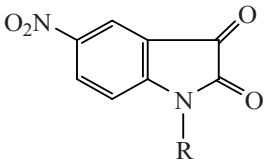
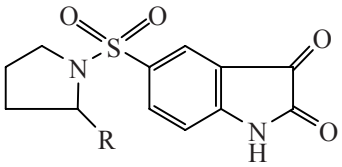
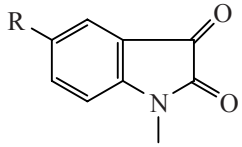
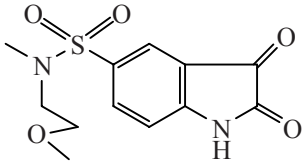
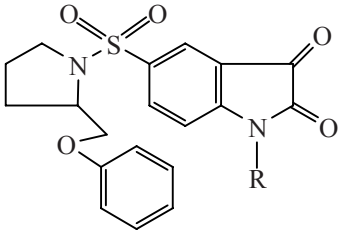
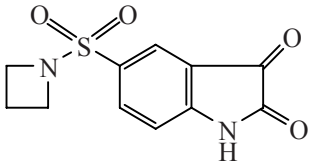
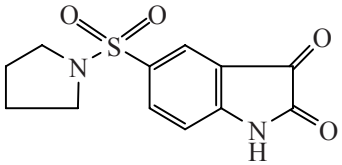
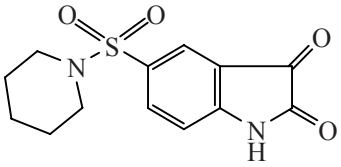
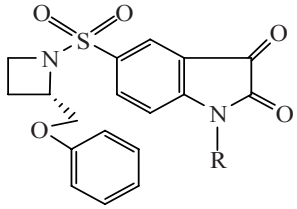
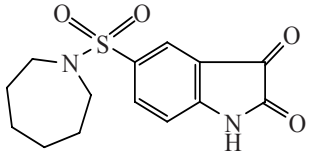
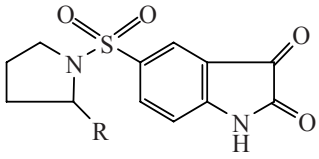
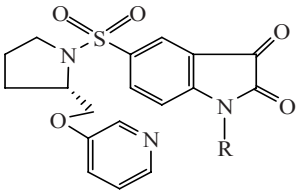
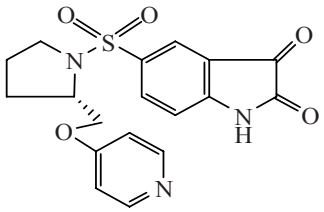
Structure	Comp. no.	R	Structure	Comp. no.	R
	**1**	Me		**14**	(*S*)-MeO_2C
	2	Bz		**15**	(*S*)-*t*-BuO_2C
	3	H		**16**	(*S*)-$Me_2N(O)C$
				17	(*S*)-$PhNHCH_2$
				18	(*R*)-$PhNHCH_2$
	**4**	MeO_2C		**19**	(*S*)-$PhN(O)C$
	5	I		**20**	(*S*)-$PhOCH_2$
	6	CN		**21**	(*S*)-$PhSCH_2$
	**7**			**22**	Me
				23	All
				24	$C_6H_{11}CH_2$
				25	Bz
				26	*t*-BuO_2CCH_2
	**8**			**27**	HO_2CCH_2
				28	4-Pyr-CH_2
				29	4-$MeOPhCH_2$
				30	4-F-$PhCH_2$
				31	4-$MeSPhCH_2$
	**9**			**32**	4-MeO_2CPhCH_2
				33	4-$HOPhCH_2$
				34	6-F-Pyr-3-CH_2
				35	6-F-Pyr-4-CH_2
				36	6-F-Pyr-2-CH_2
	**10**			**37**	H
				38	Me
				39	Bz
				40	4-$MeOPhCH_2$
				41	4-F-$PhCH_2$
	**11**			**42**	4-$MeSPhCH_2$
				43	2-F-$PhCH_2$
				44	6-F-Pyr-3-CH_2
				45	6-F-Pyr-4-CH_2
				46	6-F-Pyr-2-CH_2
	**12**	(*S*)-$MeOCH_2$		**47**	Me
	13	(*R*)-$MeOCH_2$		**48**	Bz
				49	4-$MeOPhCH_2$
				50	4-F-$PhCH_2$
				51	4-$MeSPhCH_2$
				52	H
				**53**	

Table 2. Inhibitor activity of isatin derivatives

Comp. no.[a]	Experiment		Calculation[d]	Comp. no.[a]	Experiment		Calculation[d]
	IC_{50},[b] nmol l^{-1} [8, 10]	log (IC_{50})[c]	log (IC_{50})		IC_{50},[b] nmol l^{-1} [8, 10]	log (IC_{50})[c]	log (IC_{50})
1	1000.0	3.00	2.91	**29**	14.5	1.16	1.04
2	250.0	2.40	2.41	**30**	12.1	1.08	0.91
3	3000.0	3.48	3.58	**31**	12.4	1.09	1.06
4	15000.0	4.18	3.99	**32**	12.0	1.08	1.04
5	7000.0	3.85	3.85	**33**	13.5	1.13	0.91
6	600.0	2.78	2.91	**34**	10.3	1.01	1.16
7	910.0	2.96	3.08	**35**	21.3	1.33	0.91
8	170.0	2.23	3.08	**36**	9.1	0.96	0.91
9	2800.0	3.45	3.08	**37**	287.0	2.46	2.33
10	2200.0	3.34	3.08	**38**	92.0	1.96	1.41
11	1900.0	3.28	3.08	**39**	9.7	0.99	0.91
12	120.0	2.08	1.91	**40**	8.4	0.92	1.10
13	18000.0	4.26	4.25	**41**	8.8	0.94	0.91
14	170.0	2.23	2.29	**42**	11.3	1.05	1.06
15	70.0	1.85	1.80	**43**	9.4	0.97	0.91
16	410.0	2.61	2.29	**44**	10.9	1.04	1.16
17	31.0	1.49	1.41	**45**	29.2	1.47	0.91
18	5500.0	3.74	3.75	**46**	5.8	0.76	1.16
19	140.0	2.15	2.29	**47**	23.3	1.37	1.41
20	44.0	1.64	2.08	**48**	5.2	0.72	0.91
21	44.0	1.64	1.41	**49**	58.3	1.77	2.08
22	30.0	1.48	1.41	**50**	20.4	1.31	1.41
23	4.6	0.66	1.00	**51**	3.9	0.59	0.47
24	5.2	0.72	0.69	**52**	8.4	0.92	0.91
25	2.5	0.40	0.91	**53**	4.4	0.64	0.74
26	3.1	0.49	0.65				R^2 0.9433
27	170.0	2.23	2.08				*s* 0.27
28	4.2	0.62	0.91				

[a] Here and in Table 5 the numbering of compounds corresponds to that in Table1. [b] IC_{50} is the inhibitor concentration inhibiting caspase 3 by 50%. [c] For convenience the experimental values were transformed into the log scale. [d] In calculation by expression (1) the following values of parameters were used: *Const* =3.58±0.2, P_{CR} = 0.92±0.43, $P_{Car,\ O\text{-}4}$ = 0.67±0.21, $P_{Car,\ F\square 2}$ = 0.47±0.29, $P_{C,\ Car\text{-}19}$ = 0.32±0.18, $P_{CO,\ C\text{-}3}$ = -0.17±0.09, $P_{C,\ Car\text{-}3}$ = -0.25±0.05, $P_{C,\ O\text{-}14}$ = -0.31±0.14, $P_{CS,\ C\text{-}1}$ = -0.71±0.14.

This description of the molecular structure permits a relatively simple calculation of the values of molecular descriptors for various compounds proceeding from the formula and topology of the molecule of the compound. Thus in the fragment approach the molecular descriptors explicitly reflect the specific features of the molecular structure. This feature helps the interpretation of the correlation models obtained and provides a possibility to forecast the variation of activity at the modification of the molecular structure of the inhibitor. However the fragment approach is not free of certain limitations. Firstly, this quantitative description of the molecular structure (by the number of atoms of a definite type, by the number of pairs of

atoms divided by a single bond, by a number of pairs of atoms, divided by two bonds, etc.) is two-dimensional. The spatial features of the molecular structure might be taken into consideration only by special procedures. In particular, in order to distinguish *R*- and *S*-stereoisomers it is possible to introduce fragments of different types. Secondly, the description of versatile structures requires the introduction of a large number of fragment types resulting in numerous molecular descriptors and consequently, in numerous variables in the correlation model. Therefore it is necessary to perform a selection of variables the most significant for revealing the activity and an elimination of insignificant variables. This is an important stage in the designing a correlation model, and a great effort has gone into the search for reliable and efficient algorithms of variables selection [16].

For the isatin derivatives under study we chose 16 types of fragments linked to heavy atoms. In this selection we introduced the following simplifying assumptions: (1) the carbon atoms in the six-membered isatin ring and the atoms of phenyl substituents were regarded as one type fragments, aromatic; (2) all carbon atoms in the allyl substituent (compound **23**, Table 1) were regarded as fragments of one type, aliphatic; (3) nitrile, nitro and sulfo groups were regarded as unique fragments whose types were designated as CN, NO, and SO respectively; (4) to the nitrogen atom in amido and amino groups and heterocycles corresponded a fragment of NH type disregarding the degree of substitution; (5) asymmetrical carbon atoms in the *S* and *R* is conformers were regarded as different type fragments. The designations of all types of the descriptors are presented in Table 3.

Table 3. Fragment types and their designations for the studied set of isatin derivatives

Designation of fragment type	Corresponding atoms or groups of atoms
Car	Aromatic carbon
C	Aliphatic carbon
C_R	Asymmetrical carbon atom of pyrrolidine or azetidine ring in the *R*-configuration
C_S	Asymmetrical carbon atom of pyrrolidine or azetidine ring in the *S*-configuration
CO	Carbonyl group
Niz	Nitrogen in the isatin framework
Np	Nitrogen in the pyridine ring
NH	Nitrogen in amino group or amide
O	Ether group
OH	Hydroxy group
CN	Cyano group
NO	Nitro group
SO	Sulfo group
SH	Thiol group
F	Fluorine
I	Iodine

Discussion of results. In compounds considered the maximum distance between the fragments in a pair (the chain of dividing bonds) equaled 21. Therefore accounting for probable combinations of fragments types and distances in a pair the initial number of molecular descriptors reached 2872. Subsequent to removal from this set of descriptors with zero value and descriptors appearing in a single compound, and at representing the group of mutually correlated descriptors by a single descriptor only 187 significant descriptors remained. Their values for the total set of compounds formed a predictors matrix for the multiple linear regression.

Further the algorithm was employed formerly described in [17] consisting in a successive iteration using a stepped regression for selecting the most efficient variables. The following restrictions served as criterion of the selection: No more than 8 variables in the model; the standard deviation of the calculated estimates from the experimental findings no more than 0.3 log units corresponding to less than 10% of the total range of the changes in the experimental activity in the set of compounds under study. As a result after 6 iterations we obtained only 11 models fitting to the chosen criterions. Every one of them included 8 variables. In all eleven models only 36 descriptors were utilized from the 187 descriptors selected for the regression.

Descriptors appearing in three or more models are given in Table 4 and Table 5, and their values are presented for all compounds from the set of isatin derivatives. The estimates obtained in the model best by the statistical indices are given in the last column of Table 2. The results presented show that the calculated

Table 4. The most often selected descriptors

Designation[a]	Popularity (% to the number of obtained models)	Sign of the contribution
Car,Car-8	27.3	–
CO,Car-13	27.3	+
CO	45.5	+
C_S,C-1	45.5	–
CO,Car-2	54.5	+
C_S,C-2	54.5	–
C_R	72.7	+
C,CO-3	81.8	–
Car,I-2	100.0	+

[a] Here and in Table 5 the following designations of two-fragment descriptors are used: Car, Car-8 is a pair of aromatic carbon atoms divided by eight bonds; CO, Car-13, and CO, Car-2 is an aromatic carbon removed from a carbonyl group at thirteen and two bonds respectively; C_S, C-1 и C_S, C-2 is an aliphatic carbon at a distance of one and two bonds respectively from an asymmetrical carbon atom in the *S*-configuration; C, CO-3 is an aliphatic carbon removed from a carbonyl group at three bonds; Car, I-2 are atoms of aromatic carbon and iodine divided by two bonds.

values of the inhibitor activity of isatin derivatives are well consistent with the experimental values for all compounds of the set. Compound no. **8** with an azetidine ring is an exclusion where the deviation of the calculated estimate from the experimental value exceeded threefold the standard deviation value. This fact may be due to the strain in the four-membered ring resulting in the difference in behavior of its atoms from the common ones. We made a strong simplification in regarding them as fragments of the same type as in acyclic substituents that led to a large error.

It is important that among the frequently selected descriptors both those with negative and positive contribution are present (Table 4). The structural features corresponding to the former descriptors are favorable for the activity; the details of the molecular structure corresponding to the latter reduce the inhibitor activity. The numerical values of the selected descriptors are given in Table 5; in the last row of Table 5 the dispersions of descriptors values over the total set of compounds under consideration are presented. Confronting the dispersion value with the sign of the contribution into the relation 1 of the corresponding descriptor we note that the descriptors with the positive contributions CO; CO, Car-2; CO, Car-13; Car, I-2 have a small dispersion; it means that the values of these descriptors weakly change along the total set of isatin derivatives. Therefore their relative role in the value of the inhibitor activity is small, and they may be regarded as a certain addition to the fixed term of expression 1. For instance, in all the models of the collection a descriptor Car, I-2 is present corresponding to a pair iodine atom–atom of aromatic carbon at a distance of two bonds. The positive contribution of this descriptor made possible a good description of compound no. 5 endowed with low activity. However it is not possible to make a

Table 5. Description of structures of studied isatin derivatives with the most frequently chosen descriptors

Comp. no.	Car,Car-8	CO,Car-13	CO	C_S,C-1	CO,Car-2	C_S,C-2	C_R	C,CO-3	C,I-2
1	0	0	2	0	4	0	0	1	0
2	4	0	2	0	4	0	0	1	0
3	0	0	2	0	4	0	0	0	0
4	0	0	3	0	6	0	0	1	0
5	0	0	2	0	4	0	0	1	2
6	0	0	2	0	4	0	0	1	0
7	0	0	2	0	4	0	0	0	0
8	0	0	2	0	4	0	0	0	0
9	0	0	2	0	4	0	0	0	0
10	0	0	2	0	4	0	0	0	0
11	0	0	2	0	4	0	0	0	0
12	0	0	2	2	4	2	0	0	0
13	0	0	2	0	4	0	1	0	0
14	0	0	3	1	4	2	0	2	0
15	0	0	3	1	4	2	0	5	0
16	0	0	3	1	4	2	0	2	0

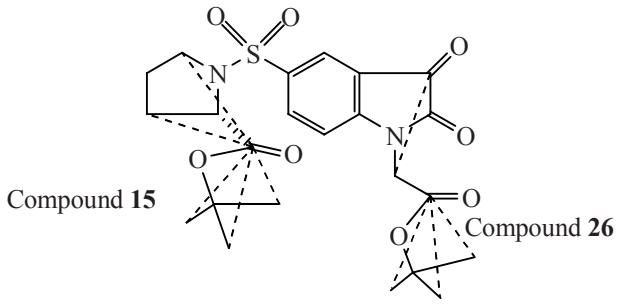
Table 5. (Contd.)

Comp. no.	Car,Car-8	CO,Car-13	CO	C_S,C-1	CO,Car-2	C_S,C-2	C_R	C,CO-3	C,I-2
17	8	1	2	2	4	2	0	0	0
18	8	1	2	0	4	0	1	0	0
19	8	1	3	1	5	2	0	2	0
20	8	1	2	2	4	2	0	0	0
21	8	1	2	2	4	2	0	0	0
22	8	1	2	2	4	2	0	1	0
23	8	1	2	2	4	2	0	2	0
24	8	1	2	2	4	2	0	2	0
25	12	1	2	2	4	2	0	1	0
26	8	3	3	2	4	2	0	4	0
27	8	3	3	2	4	2	0	1	0
28	10	1	2	2	4	2	0	1	0
29	12	1	2	2	4	1	0	1	0
30	12	1	2	2	4	2	0	1	0
31	12	1	2	2	4	2	0	1	0
32	12	1	3	2	5	2	0	1	0
33	12	1	2	2	4	2	0	1	0
34	11	1	2	2	4	2	0	1	0
35	10	1	2	2	4	2	0	1	0
36	12	1	2	2	4	2	0	1	0
37	8	1	2	2	4	0	0	0	0
38	8	1	2	2	4	1	0	1	0
39	12	1	2	2	4	1	0	1	0
40	12	1	2	2	4	1	0	1	0
41	12	1	2	2	4	1	0	1	0
42	12	1	2	2	4	1	0	1	0
43	12	1	2	2	4	1	0	1	0
44	11	1	2	2	4	1	0	1	0
45	10	1	2	2	4	1	0	1	0
46	12	1	2	2	4	0	0	1	0
47	7	1	2	2	4	2	0	1	0
48	11	1	2	2	4	2	0	1	0
49	7	1	2	2	4	2	0	0	0
50	8	0	2	2	4	2	0	0	0
51	11	1	2	2	4	2	0	1	0
52	11	1	2	2	4	2	0	1	0
53	11	1	2	2	4	2	0	1	0
Dispertion[a]	23.17	0.42	0.13	0.75	0.11	0.78	0.04	0.86	0.08

[a] Dispersion of descriptors values over the total set of compounds under consideration.

conclusion on the role of iodine for the activity in general since the studied set of isatin derivative contains only one such compound.

The most often chosen descriptor with a negative contribution, i.e., favorable for the activity, is the descriptor C, CO-3 corresponding to the position of an aliphatic carbon three bonds away from a carbonyl group. In compounds where these pairs are formed by the isatin carbonyl groups and carbon atoms of a substituent at a nitrogen (see figure) a high activity was observed (for instance, nos. **23, 24**, and **26**). In



A tentative structure of isatin derivative showing the substituents present in compounds **15** and **26**. The pairs of fragments included in the descriptor CO, C-3 are connected by a dash line.

compounds where these contributions originated from the carbonyl groups of substituents at a pyrrolidine ring (see figure) the activity was lower (for instance, no. **15**). It is presumable that the hydrophobic substituents at the isatin nitrogen are essential for the high inhibitor activity.

A frequently selected descriptor with a positive contribution is a single-fragment descriptor equal to the number of fragments C_R, namely for the studied set of compounds this descriptor indicates that the asymmetrical carbon atom has the *R*-configuration. Two descriptors with negative contributions, C_S, C-2 and C_S, C-1 corresponding to the location of an aliphatic carbon at a distance of 1 and 2 bonds from an asymmetrical carbon atom, have nonzero values in all compounds with the *S*-configuration (Table 1 and Table 5). Thus these three descriptors, C_R, C_S, C-2, and C_S, C-1 reflect the fact that the configuration of the asymmetrical carbon atom essentially affects the activity. Actually, the activity of *R*-derivatives is by 3-4 orders of magnitude lower (it is the most pronounced in the pairs of compounds nos. **12, 13**, **17,** and **18** in Table 1).

A pair descriptor with a positive contribution CO, Car-13 has the largest value in compounds nos. **26** (active) and **27** (less active), therefore its role cannot be unambiguously understood.

Hence using fragment descriptors we succeeded in obtaining correlation models having clear structural sense and permitting the calculation of inhibitor activity of isatin derivatives with high precision. It should also be stressed that the role of individual descriptors and the corresponding structural features was revealed in the analysis of the collection of models and not of a single best equation. In the analysis the sign of the parameter of the correlation expression (of descriptor contribution) and the quantitative values of descriptors along the total set of isatin derivatives were taken into the consideration.

REFERENCES

1. Thornberry, N.A. and Lazebnik, Y., *Science*, 1998, vol. 281, no. 5381, p. 1312.
2. Fuentes-Prior, P. and Salvesen, G.S., *Biochem. J.,* 2004, vol. 384, no. 2, p.201.
3. *Programmirovannaya kletochnaya gibel'* (Programmed Cell Death), Novikov, V.S., Ed., St. Petersburg: Nauka, 1996.
4. Talanian, R.V., Brady, K.D., and Cryns, V.L., *J. Med. Chem*., 2000, vol.43, no. 18, p.3371.
5. Scott, C.W., Sobotka-Briner, C., Wilkins, D.E., Jacobs, R.T., Folmer, J.J., Frazee, W.J., Bhat, R.V., Ghanekar, S.V., and Aharony, D. , *J. Pharm. Exp. Therap*., 2003, vol.304, no. 1, p. 433.
6. Becker, J.W., Rotonda, J., Soisson, S.M., Aspiotis, R., Bayly, C., Francoeur, S., Gallarot, M., Gardia-Calvo, M., Giroux, A., Grimm, E., Han, Y., McKay, D., Nicholson, D.W., Peterson, E., Renaud, J., Roy, S., Thornberry, N., and Zamboni, R., *J. Med. Chem*., 2004, vol. 47, no. 10, p. 2466.
7. Lee, D., Long, S.A., Adams, J.L., Chan, G., Vardye, K.S., Francis, T.A., Kikly, K., Winkler, J.D., Sung, Ch.M., Debouck, Chr., Richardson, S., Levy, M.A., DeWolf, W.E., Jr, Keller, P.M., Tomaszek, Th., Head, M.S., Ryan, M.D., Haltiwanger, R.C., Liang, P.-H., Janson, Ch.A., McDevitt, P.J., Jochanson, K., Concha, N.O., Chan, W., Suva, L.J., Gowen, M., and Nutall, M.E., *J. Biol. Chem*., 2000, vol.275, no. 21, p. 16007.
8. Lee, D., Long, S.A., Murray, J.H., Adams, J.L., Nutall, M.E., Nadeau, D.P., Kikly, K., Winkler, J.D., Sung, Ch.-M., Ryan, M.D., Levy, M.A., Keller, P.M., and DeWolf, W.E., Jr., *J. Med. Chem*., 2001, vol. 44, no. 12, p. 2015.
9. Chapman, J.G., Magee, W.P., Stukenbrok, H.A., Beckius, G.E., Milia, A.J., and Tracey, W.R., *Eur. J. Pharmacol*., 2002, vol. 456, no. 1, p. 59.
10. Chu, W., Zhang, J., Zeng, Ch., Rothfuss, J., Tu, Zh., Chu, Y., Reichert, D.E., Welch, M.J., and Mach, R.H., *J. Med. Chem*., 2005, vol. 48, no. 24, p. 7637.
11. Lee, D.Y., Hyun, K.H., Park, H.Y., Lee, K.A., Lee, B.-S., and Kim, C.K., *Bull. Korean Chem. Soc*., 2006, vol. 27, no. 2, p. 273.
12. Chu, W., Rothfuss, J., d'Avignon, A., Zeng, Ch., Zhou, D., Hotchkiss, R.S., and Mach, R.H., *J. Med. Chem*., 2007, vol. 50, no. 15, p. 3751
13. Golovanov, I.B. and Zhenodarova, S.M., *Zh. Obshch. Khim*., 2003, vol. 73, no. 1, p. 90.

14. Golovanov, I.B. and Tsygankova, I.G., *Zh. Obshch. Khim*., 1999, vol. 69, no. 8, p. 1275.
15. Golovanov, I.B. and Tsygankova, I.G., *Quantative Structure Activity Relationship (QSAR),* 2000, vol. 19, p. 554.
16. Tsygankovam I.G., *Curr.Comp.-Aided Drug Des*., 2008, vol. 4, no. 2, p. 132.
17. Tsygankova, I.G. and Zhenodarova, S.M., *Zh. Obshch. Khim*., 2008, vol. 78, no. 9. p. 1529.