Pergamon

Bioorganic & Medicinal Chemistry 9 (2001) 3295–3299

BIOORGANIC & MEDICINAL CHEMISTRY

QSAR Study on Inhibition of Brain 3-Hydroxy-anthranilic Acid Dioxygenase (3-HAO): A Molecular Connectivity Approach

Vijay K. Agrawal,[a] Rashmi Sohgaura[a] and Padmakar V. Khadikar[b,*]

[a]*Computer & QSAR Laboratories, A. P. S. University, Rewa-486 003, India*
[b]*Research Division, Laxmi Pest and Fumigation Pvt. Ltd. 3, Khatipura, Indore 452 007, India*

Received 16 April 2001; accepted 30 June 2001

Abstract—The ability of 4,5-, 4,6-disubstituted and 4,5,6-trisubstituted 3-hydroxyanthranilic acid derivatives to reduce the production of the excitotoxin quinolinic acid (QUIN) by inhibition of brain 3-hydroxyanthranilic acid dioxygenase (3-HAO) has been investigated using molecular connectivity indices ($^0\chi^v$, $^1\chi^v$, $^2\chi^v$). The in-vivo inhibition of 3-HAO in rat cortex (pIC_{50}, nM) is used for this purpose. The regression models obtained suggest that the degree of branching of the compounds under study have a dominant role in the observed inhibition potency. The data were used to generate quantitative structure–activity relationship (QSAR) models for estimating the potency of 3-HAO. The information obtained from the correlation should be useful in designing more potent analogues.

Introduction

The cytostatic, non-haem ferrous (Fe^{2+}) enzyme 3-hydroxyanthranilic acid dioxygenase (3-HAO) play an important role in the metabolic transformation of L-tryptophan to nicotinamide in the kynurenine pathway. 3-HAO seems mainly to be localized to the hepatic tissue, but it was also been observed that the enzyme is expressed in the brain as well as in inflammatory cells where the production of quinolinic acid (QUIN) has been shown to be stimulated by certain cytokines. Biological and immunological analysis in the rat suggests that the brain and liver 3-HAO are identical proteins. Increased QUIN levels, or enhanced kynurenine pathway activity, have been implicated in inflammatory diseases of the central nervous system (CNS).

Recently Linderberg et al.[1] have synthesized several analogues of 3-hydroxyanthranilic acid (3-HANA) and reported their ability to reduce the production of QUIN by inhibition of 3-HAO. The potency of the compounds to inhibit 3-HAO was assayed in rat brain homogenate. However, till date, no topological QSAR modeling of the potency of these compounds has been reported in the literature.

*Corresponding author. Tel.: +91-731-531906;
e-mail: vijay-agrawal@lycos.com

Molecular connectivity, which is a method of quantifying the molecular structure of a compound, was introduced by Randic[2] and extensively developed by Kier and Hall.[3] The method describes the structure of a molecule by a set of molecular connectivity indices, $^m\chi^v$, which encode information about size, branching, cyclization, unsaturation, and hetero-atom content of the molecules. This methodology has been widely used in QSAR studies with many classes of biologically active molecules.[4] We have, therefore, used valence connectivity indices ($^0\chi^v$, $^1\chi^v$, $^2\chi^v$) in the present study and observed that excellent results are obtained upon introduction of indicator parameters accounting for the substitutions at R_5 and R_6 in the series of substituted 3-HANA (Fig. 1, Table 1).

Results and Discussion

The descriptors (valence connectivity indexes and indicator parameters) used for proposing QSAR models are presented in Table 2; while the structural details of substituted 3-HANA and their inhibition potency (pIC_{50}, nM) are given in Table 1.

A regression analysis of the potency, using the multiparametric analysis (MPA) resulted into six bi-parametric and three tri-parametric models (Table 4). All these models (except models 7, 8, 11, 13, 14, 15, 17, 18)

0968-0896/01/$ - see front matter
PII: S0968-0896(01)00242-5

were statistically significant according to multi-variate as well as cross-validation methods. The three most significant models obtained being models 4, 10 and 16 (Table 5). The important parameters influencing pIC_{50} (nM) values are $^{0}\chi^{v}$, the size of the R_6 substituent, and the lipophilicity of the R_5 substituent. This means that small electron-withdrawing groups at R_6 and lipophilic electron-withdrawing groups at R_5 would increase the potency of substituted 3-HANA. We now give the details of these findings.

It is worth mentioning that no degeneracy is observed in the activity (pIC_{50}, nM), however, degeneracy is observed in valence connectivity indices used (Table 3). This is obvious as these indices belong to second generation topological indices. According to Balaban,[5] such indices, inspite of their observed degeneracy, are found to give excellent QSAR models. The same is found to be true in the present case also.

A perusal of Tables 4 and 5 shows that out of the six bi-parametric models, the model-4 is found to be an excellent model for modeling pIC_{50} (nM). This model-4 is found as under:

$$pIC_{50}\ (nM) = -1.0617\ (\pm 0.0903)\ {}^{0}\chi^{v} + 1.3269\ (\pm 0.1779)Ip_5 + 5.6907$$

$$n = 16, \quad Se = 0.3472, \quad R_A^2 = 0.9150, \quad R^2 = 0.9264,$$

$$R = 0.9625, \quad F = 81.758, \quad p = 4.330 \times 10^{-8}$$

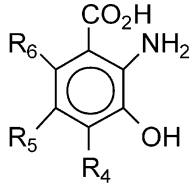

Figure 1. Substituted 3-HANA used in the present study (see Table 1 for details).

Table 1. Substituted 3-hydroxy-anthranilic acid derivatives and their inhibition potency against 3-HAO

Compd no.	R_4	R_5	R_6	pIC_{50} (nM)
1	Cl	H	H	−0.778
2	Br	H	H	−0.301
3	F	H	H	−1.380
4	Br	OEt	H	−1.748
5	Br	Bt	H	0.532
6	Br	Me	H	−0.363
7	Cl	Br	H	0.585
8	Cl	Cl	H	0.583
9	Br	H	Br	− 0.763
10	Br	H	MeO	−2.079
11	Cl	H	Cl	−1.004
12	Cl	H	Ph	−4.041
13	Cl	$(CH_2)_2$	$(CH_2)_2$	−2.301
14	Cl	Me	Me	−0.914
15	Cl	Me	Cl	−0.643
16	Cl	H	Me	−0.892

Here, and thereafter, n is the number of compounds; Se, standard error of estimation; R_A^2, adjustable R^2; R^2, coefficient of determination; R, correlation coefficient; F, F-ratio, and p, probability.

This model for the set of 16 data points containing only two correlating parameters ($^{0}\chi^{v}$ and Ip_5) is an excellent situation.[6]

The R_A^2 value for the aforementioned correlation is found to be the highest among all the six bi-parametric models (Table 5). R_A^2 is a measure of% explained variation in the dependent variable, $pIC_{50,}$ nM in the present case, that takes into account the relationship between the number of cases (compounds used) and number of independent variables, topological indices and indicator parameters used in the regression model; whereas R^2 will increase when an independent variable is added. Decrease of R_A^2 in remaining bi-parametric models in our case indicates that the corresponding added variable does not reduce the unexplained variation in enough to offset the degree of freedom.

It is important to record that the parameters used in models 5 to 16 are some combination of $^{1}\chi^{v}$, $^{2}\chi^{v}$, Ip_5, and Ip_6. The valence connectivity index $^{1}\chi^{v}$ distinguishes the degree of unsaturation and the presence of hetero-atoms, while $^{2}\chi^{v}$ gives more information about branching and degree of unsaturation. Hence, the lower qualities of these models (except for 10 and 16) compared to model-4 may be attributed to these factors.

Table 2. Molecular valence connectivity indices and indicator parameters for substituted 3-hydroxy-anthranilic acid derivatives

Compd no.	Ip_5	Ip_6	$^{o}\chi^{v}$	$^{1}\chi^{v}$	$^{2}\chi^{v}$
1	0	0	5.913	3.040	2.162
2	0	0	5.913	3.040	2.162
3	0	0	5.913	3.040	2.162
4	1	0	7.951	4.156	2.638
5	1	0	6.213	3.145	2.638
6	1	0	6.835	3.456	2.478
7	1	0	6.213	3.145	2.268
8	1	0	6.213	3.145	2.268
9	0	1	6.213	3.145	2.283
10	0	1	7.244	3.569	2.503
11	0	1	6.213	3.145	2.283
12	0	1	9.222	5.247	3.570
13	1	1	8.586	5.086	3.611
14	1	1	7.758	3.879	3.006
15	1	1	7.136	3.568	2.605
16	0	1	6.835	3.456	2.526

Table 3. Correlation matrix for inter-correlation of molecular parameters and their correlation with the activity

	pIC_{50}(nM)	$^{o}\chi^{v}$	$^{1}\chi^{v}$	$^{2}\chi^{v}$	Ip_5	Ip_6
pIC_{50} (nM)	1.0000					
$^{0}\chi^{v}$	−0.7818	1.0000				
$^{1}\chi^{v}$	−0.7171	0.8772	1.0000			
$^{2}\chi^{v}$	−0.7122	0.9222	0.7414	1.0000		
Ip_5	0.3773	0.2182	0.0408	0.2652	1.0000	
Ip_6	−0.5289	0.5101	0.5331	0.4687	−0.2500	1.0000

Furthermore, models 7, 8, 11, 13, 15, 17, and 18 suffer from the statistical defect in that the coefficient of Ip_6 term is smaller than its standard deviation. Such models are not allowed statistically.

Furthermore, bi-parametric models involving Ip_6 as one of the correlating parameter are found to be less significant. Thus, the substituents at R_6 have negative role in exhibiting the activity.

The step-wise regression analysis resulted into nine tri-parametric models (Table 4). The quality of these regression models as presented in Table 5 indicated that the models 10 and 16 are the most significant models for modeling the activity. These models are found as under:

Model 10;

$$pIC_{50}(nM) = -1.3468\ (\pm 0.1846)^{0}\chi^{v} + -0.4144(\pm 0.2389)^{1}\chi^{v} + 1.4243\ (\pm 0.1748)Ip_5 + 6.0887$$

$$n = 16, \quad Se = 0.3231, \quad R_A^2 = 0.9264, \quad R^2 = 0.9411,$$

$$R = 0.9701, \quad F = 63.928$$

Model 16;

$$pIC_{50}(nM) = -1.1512\ (\pm 0.1084)^{0}\chi^{v} + -1.4397(\pm 0.2389)Ip_5 + 0.2974\ (\pm 0.2163)Ip_6 + 6.1027$$

$$n = 16,\ Se = 0.3359,\ R2A = 0.9205,\ R2 = 0.9363,$$
$$R = 0.9677,\ F = 58.867$$

The remaining tri-parametric models were discarded on the grounds that in their cases the coefficient of Ip_6 term is considerably smaller than its corresponding standard deviation, making them statistically insignificant.

The aforementioned results and discussion indicated that models 4, 10 and 16 are the only appropriate models for modeling the inhibition potency. Which out of these three models is the best, is deduced by calculating quality factor[7,8] ($Q = R / Se$). The Q-values for models 4 10 and 16 are found as 2.7722, 3.0025, and 2.8809 respectively suggesting that the model-10 is the best model for modeling the activity.

In order to confirm our findings we have opted cross-validation method and calculated cross-validation parameters for models 4, 6, 10, 12, 16 and 18 and are presented in Table 6. The highest r^2_{CV} value and the lowest values of S_{PRESS} and PSE obtained for model-10 again indicate it to be the most appropriate model for modeling the inhibition potency.

It is worth recording that S_{PRESS} is found to be the same as that of respective values of standard error of estimation (Se). This means that both Se and S_{PRESS} are not good parameters for deciding quality of regression. Hence, it is PSE which establishes quality as well as predictive ability of the proposed models.

Final conclusion is obtained by calculating pIC_{50} values from models-4, 10 and 16 and comparing them with the observed values of pIC_{50}. Such a comparison is made in Table 7 and demonstrated in Figure 2. The predictive correlation potential ($R^2 = 0.941$) is in favour of model-10.

Conclusion

From the aforementioned results and discussion we conclude that both geometrical and electronic structural features of the substituted 3-HANA play an important role in the inhibition potency and that currently available topological methods such as valence connectivity are capable of probing these features in detail.

Table 4. Proposed models for modeling the activity

Model No.	Parameters used	Regression expression
1	$^{0}\chi^{v}$	$pIC_{50} = -0.9148\ (\pm 0.1950)\ ^{0}\chi^{v} + 5.3409$
2	$^{1}\chi^{v}$	$pIC_{50} = -1.1121\ (\pm 0.2889)\ ^{1}\chi^{v} + 3.1086$
3	$^{2}\chi^{v}$	$pIC_{50} = -1.7638\ (\pm 0.4646)\ ^{2}\chi^{v} + 3.4735$
4	$^{0}\chi^{v}$, Ip_5	$pIC_{50} = -1.0617\ (\pm 0.0903)^{0}\chi^{v} + 1.3269\ (\pm 0.1779)Ip_5 + 5.6907$
5	$^{1}\chi^{v}$, Ip_5	$pIC_{50} = -1.1379\ (\pm 0.2436)^{1}\chi^{v} + 0.9393\ (\pm 0.3623)Ip_5 + 2.7335$
6	$^{2}\chi^{v}$, Ip_5	$pIC_{50} = -2.1639\ (\pm 0.2740)^{2}\chi^{v} + 1.4048\ (\pm 0.2552)Ip_5 + 3.7790$
7	$^{0}\chi^{v}$, Ip_6	$pIC_{50} = -0.8098\ (\pm 0.2283)^{0}\chi^{v} - 0.4057\ (\pm 0.4499)Ip_6 + 4.8194$
8	$^{1}\chi^{v}$, Ip_6	$pIC_{50} = -0.9428\ (\pm 0.3432)^{1}\chi^{v} - 0.4725\ (\pm 0.5109)Ip_6 + 2.7238$
9	$^{2}\chi^{v}$, Ip_6	$pIC_{50} = -1.4736\ (\pm 0.5181)^{2}\chi^{v} - 0.5767\ (\pm 0.4826)Ip_6 + 3.0307$
10	$^{0}\chi^{v}$, $^{1}\chi^{v}$, Ip_5	$pIC_{50} = -1.3468\ (\pm 0.18461)^{0}\chi^{v} + 0.4144\ (\pm 0.2389)^{1}\chi^{v} + 1.4243\ (\pm 0.1748)Ip_5 + 6.0887$
11	$^{0}\chi^{v}$, $^{1}\chi^{v}$, Ip_6	$pIC_{50} = -0.7390\ (\pm 0.4272)^{0}\chi^{v} - 0.1147\ (\pm 0.5755)^{1}\chi^{v} - 0.3860\ (\pm 0.4779)Ip_6 + 4.7418$
12	$^{1}\chi^{v}$, $^{2}\chi^{v}$, Ip_5	$pIC_{50} = -0.3444\ (\pm 0.2449)^{1}\chi^{v} - 1.7317\ (\pm 0.4054)^{2}\chi^{v} + 1.3189\ (\pm 0.2536)Ip_5 + 3.9962$
13	$^{1}\chi^{v}$, $^{2}\chi^{v}$, Ip_6	$pIC_{50} = -0.5449\ (\pm 0.4408)^{1}\chi^{v} - 0.9255\ (\pm 0.6742)^{2}\chi^{v} - 0.3840\ (\pm 0.4981)Ip_6 + 3.5518$
14	$^{0}\chi^{v}$, $^{2}\chi^{v}$, Ip_5	$pIC_{50} = -0.8565\ (\pm 0.2287)^{0}\chi^{v} - 0.4780\ (\pm 0.4900)^{2}\chi^{v} + 1.3568\ (\pm 0.1808)Ip_5 + 5.4637$
15	$^{0}\chi^{v}$, $^{2}\chi^{v}$, Ip_6	$pIC_{50} = -0.8710\ (\pm 0.5423)^{0}\chi^{v} + 0.1401\ (\pm 1.1170)^{2}\chi^{v} - 0.4054\ (\pm 0.4680)Ip_6 + 4.8880$
16	$^{0}\chi^{v}$, Ip_5, Ip_6	$pIC_{50} = -1.1512(\pm 0.1084)^{0}\chi^{v} + 1.4397\ (\pm 0.1907)Ip_5 + 0.2974(\pm 0.2163)Ip_6 + 6.1027$
17	$^{1}\chi^{v}$, Ip_5, Ip_6	$pIC_{50} = -1.0906(\pm 0.3057)^{1}\chi^{v} + 0.9041\ (\pm 0.3969)Ip_5 - 0.1293(\pm 0.4688)Ip_6 - 2.6422$
18	$^{2}\chi^{v}$, Ip_5, Ip_6	$pIC_{50} = -2.2384\ (\pm 0.3460)^{2}\chi^{v} + 1.4533\ (\pm 0.2940)Ip_5 + 0.1205(\pm 0.3209)Ip_6 + 3.8829$

Table 5. Quality of regression models presented in Table 4

Model no.	Se	R_A^2	R^2	R	F	P
1	0.7688	—	0.6112	−0.7818	22.006	3.470×10^{-4}
2	0.8597	—	0.5143	−0.7171	14.822	1.768×10^{-3}
3	0.8655	—	0.5072	−0.7122	14.410	1.966×10^{-3}
4	0.3472	0.9150	0.9264	0.9625	81.758	4.330×10^{-8}
5	0.7239	0.6306	0.6798	0.8245	13.802	6.094×10^{-4}
6	0.4921	0.8293	0.8521	0.9231	37.436	4.032×10^{-6}
7	0.7740	0.5778	0.6341	0.7693	11.263	1.453×10^{-3}
8	0.8637	0.4742	0.5443	0.7378	7.764	6.044×10^{-3}
9	0.8525	0.4877	0.5560	0.7457	8.139	5.105×10^{-3}
10	0.3231	0.9264	0.9411	0.9701	63.928	1.191×10^{-7}
11	0.8042	0.5491	0.6353	0.7970	6.967	5.721×10^{-3}
12	0.4746	0.8412	0.8730	0.9343	27.493	1.162×10^{-5}
13	0.8357	0.5077	0.6061	0.7786	6.156	8.903×10^{-3}
14	0.3478	0.9147	0.9318	0.9653	54.631	2.869×10^{-7}
15	0.8050	0.5432	0.6345	0.7966	6.945	5.787×10^{-3}
16	0.3359	0.9205	0.9363	0.9677	58.667	1.892×10^{-7}
17	0.7511	0.6023	0.6819	0.8257	8.573	2.592×10^{-3}
18	0.5092	0.8172	0.8538	0.9240	23.350	2.681×10^{-5}

Table 6. Cross-validation parameters

	Cross-validation parameters					
Model no.	PRESS	SSY	$\frac{PRESS}{SSY}$	r_{CV}^2	S_{PRESS}	PSE
4	1.5672	19.7129	0.0795	0.9205	0.3472	0.3130
6	3.1482	18.1319	0.1736	0.8264	0.4921	0.4436
10	1.2531	20.2801	0.0618	0.9382	0.3231	0.2798
12	2.7028	18.5773	0.1455	0.8545	0.4746	0.4110
16	1.3540	19.9261	0.0680	0.9321	0.3359	0.2909
18	3.116	18.1685	0.1713	0.8287	0.5092	0.4410

Table 7. Comparison of experimental and calculated inhibition potency pIC_{50} (nM) using models 4, 10 and 16

Compd. No.	pIC_{50}(nM) Exp.	Calculated pIC_{50} (nM) from		
		Model-4	Model-10	Model-16
1	−0.778	−0.587	−0.615	−0.704
2	−0.301	−0.587	−0.615	−0.704
3	−1.380	−0.587	−0.615	−0.704
4	−1.748	−1.425	−1.473	−1.610
5	0.532	0.423	0.448	0.390
6	−0.363	−0.242	−0.260	−0.326
7	0.585	0.423	0.449	0.390
8	0.583	0.423	0.449	0.390
9	−0.763	−0.906	−0.976	−0.752
10	−2.079	−1.999	−2.188	−1.939
11	−1.004	−0.906	−0.976	−0.752
12	−4.041	−4.099	−4.157	−4.216
13	−2.301	−2.099	−1.943	−2.045
14	−0.914	−1.221	−1.328	−1.091
15	−0.643	−0.501	−0.619	−0.375
16	−0.892	−1.566	−1.094	−1.468

Experimental

Pharmacology

As stated earlier the inhibition potency (IC_{50}, nM) were adopted from the literature[1] and were converted into their log (pIC_{50}) units. According to ref 1 the compounds were tested for their ability to inhibit 3-HAO homogenates of rat brain tissue according to the method of Faster et al.[9] The details can be found in ref 1.

The molecular connectivity indices

The connectivity index $\chi = \chi(G)$ of a graph G is defined by Randic[2] as under

$$\chi = \chi(G) = \sum_{ij} [d_i \; d_j]^{-0.5} \tag{1}$$

where d_i is the valence of a vertex i, equal to the number of bonds connected to the atom i, in G, representing the graph of a compound. Meaning of d_j is analogous.

In the case of hetero-systems the connectivity is given in terms of valence delta values δ_i^v and δ_j^v of atoms i and j and is denoted by χ^v. This version of the connectivity index is called the valence connectivity index and defined as[3,4] under:

$$\chi^v = \chi^v(G) = \sum_{ij} [\delta_i^v \; \delta_j^v]^{-0.5} \tag{2}$$

where the sum is taken over all bonds i–j of the molecule. Valence delta values are given by

$$\delta_i^v = \frac{Z_i^v - H_i}{Z_i - Z_j - 1} \tag{3}$$

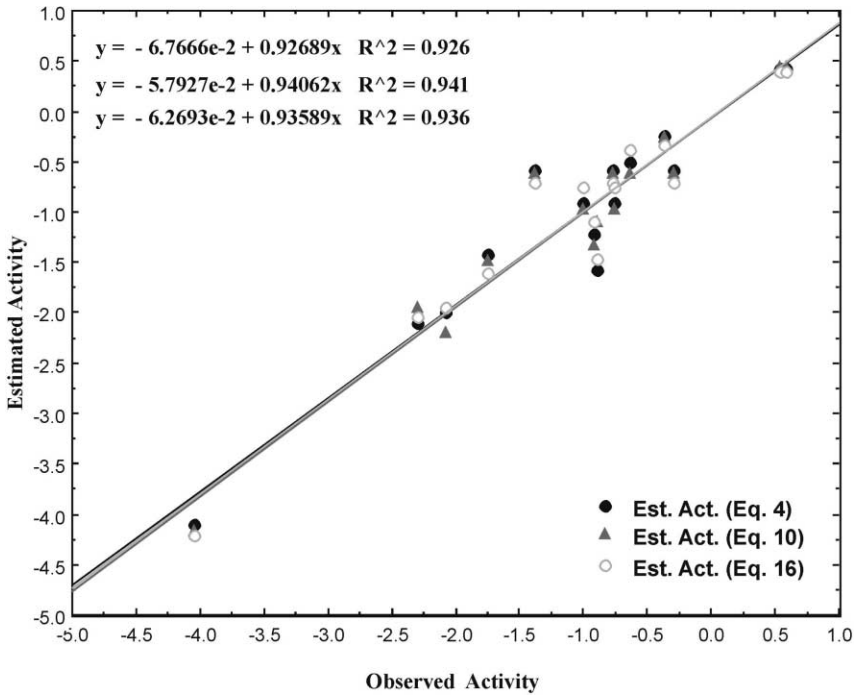

Figure 2. Correlation of observed activity: inhibition potency (pIC_{50}, nM) with their estimated values.

where Z_i is the atomic number of atom i, Z_i^v is the number of valence electron of the atom i and H_i is the number of hydrogen atoms attached to atom i. the δ_i^v values are available in the books of Kier and Hall.[3,4]

Recall that nowadays the connectivity and the valence connectivity indices expressed by eqs (1) and (2) are termed as first-order connectivity and first-order valence connectivity index respectively. Lower or higher order indices are also possible which are defined analogously.

Indicator parameters

Two indicator parameters Ip_5 and Ip_6 are used for substitutions at R_5 and R_6 respectively and their values are taken as unity. In absence of substituition at these places both Ip_5 and Ip_6 take the value of zero.

Statistical analysis

All statistical analysis[6,10,11] were performed by least squares linear regression program namely Regress-1 software supplied by Professor I. Lukovits, Hungarian Academy of Sciences, Budapest, Hungary.

Acknowledgements

Authors are thankful to Prof. I. Lukovits, Hungarian Academy of Sciences, Budapest, Hungary for providing Softwares for computation of the topological indices and for making regression analysis.

References and Notes

1. Linderberg, H.; Hillberg, S.; Bjork, S.; Gotthamar, B.; Hogberg, T.; Persson, K.; Schwarcz, R.; Luthman, J.; Jonasson, R. *Eu. J. Med. Chem.* **1999**, *34*, 729.
2. Randic, M. *J. Am. Chem. Soc.* **1975**, *97*, 6609.
3. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York, 1976.
4. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure–Activity Relationship*; Wiley: New York, 1986.
5. Balaban, A. T. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 23.
6. Chatterjee, S.; Hadi, A. S.; Price, B. *Regression Analysis by Examples*. 3rd ed.; Wiley: New York, 2000.
7. Pogliani, L. *Amino Acids* **1994**, *6*, 141.
8. Pogliani, L. *J. Phys. Chem.* **1996**, *100*, 18065.
9. Foster, A. C.; White, B. J.; Schwarez, R. *J. Neuro. Chem.* **1986**, *47*, 22.
10. Lucic, B.; Trinajstic, N. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 121.
11. Lucic, B.; Trinajstic, N.; Slid, S.; Karelson, M.; Katritzky, A. R. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 610.