Iranian Journal of Analytical Chemistry
Volume 4, Issue 1, March 2017 (40-50)

مجله ایرانی شیمی تجزیه
دوره ۴، شماره ۱، فروردین ماه ۱۳۹۶ (۵۰-۴۰)

مطالعات ارتباط کمی ساختار- فعالیت برروی بازدارنده‌های گیرنده هیستامین با استفاده از ژنتیک الگوریتم - رگرسیون خطی چندگانه

اسلام پوربشیر*، مریم کارگر مقدم، علیرضا بنائی
بخش شیمی، دانشگاه پیام نور، صندوق پستی ۳۶۹۷-۱۹۳۹۵، تهران، ایران
تاریخ دریافت: ۹ آبان ۱۳۹۵ تاریخ پذیرش: ۲۱ اسفند ۱۳۹۵

Quantitative Structure-Activity Relationship Studies on the Histamin H3 Receptor Inhibitors Using the Genetic Algorithm-Multiple Linear Regressions

Eslam Pourbasheer*, Maryam Kargar Moghadam, Alireza Banaei
Department of Chemistry, Payame Noor University (PNU), P.O. BOX, 19395-3697, Tehran, Iran
Received: 30 October 2016 **Accepted:** 11 March 2017

چکیده
یک مدل ارتباط کمی ساختار- فعالیت برای پیش‌بینی فعالیت آنتاگونیستی مشتقات بنزیل تترازول به‌عنوان گیرنده هیستامین ایجاد شده‌است. انواع مختلف توصیف‌کننده‌های مولکولی برای نشان‌دادن جنبه‌های مختلف ساختارهای مولکولی استفاده شده است. در این روش، تمامی مجموعه داده‌های ترکیبات به دو بخش آموزشی و آزمایشی تقسیم شده‌اند. مدل ارتباط بین توصیف‌کننده‌های مولکولی و فعالیت بیولوژیکی مولکول‌ها با استفاده از رگرسیون خطی چندگانه، و با استفاده از روش گام به گام و ژنتیک الگوریتم ایجاد شده است. مقایسه نتایج حاصل شده، نشان‌دهنده برتری روش ژنتیک الگوریتم – رگرسیون خطی چندگانه نسبت به روش گام به گام رگرسیون خطی چندگانه می‌باشد. مدل ارتباط کمی ساختار فعالیت با استفاده از روش ارزیابی متقاطع یک نمونه خارج از رده و آماره فیشر، دسته آزمون خارجی (N=64, R2=0.808, F= 30.806, Q2adj=0.782, Q2LOO=0.751, Q2LGO=0.669) و روش بهم ریختگی تصادفی تأیید شد. در نتیجه، مدل ارتباط کمی ساختار- فعالیت ایجاد شده می‌تواند به عنوان ابزار ارزشمند برای طراحی ساختارهای مشابه آنتاگونیست‌های جدید گیرنده هیستامین مورد استفاده قرار گیرد.

واژه‌های کلیدی
ارتباط کمی ساختار- فعالیت؛ ژنتیک الگوریتم؛ رگرسیون خطی چندگانه؛ گیرنده هیستامین.

Abstract
A quantitative structure-activity relationship model has been created for forecasting the antagonist potency of benzyl tetrazole derivatives as human histamine receptors. Various kinds of molecular descriptors were used to represent different aspects of the molecular structures. In this method, the whole data set for the compounds were divided into the training and test sets. The model of relationships between molecular descriptors and biological activity of molecules were created by using stepwise multiple linear regressions and a genetic algorithm. Comparison of the results obtained indicated the superiority of the genetic algorithm based multiple linear regression over the stepwise based multiple linear regression. The ultimate quantitative structure-activity relationship model (N =64, R2=0.808, F= 30.806, Q2adj= 0.782, Q2LOO = 0.751, Q2LGO=0.669) was fully approved using the leave-one-out cross-validation method, Fischer statistics (F), external test set and the Y-randomization test. As a result, the produced quantitative structure-activity relationship model could be applied as a valorous instrumentation for sketching analogous groups of new antagonists of histamine receptors.

Keywords
Quantitative Structure-Activity Relationship; Genetic Algorithms; Multiple Linear Regressions; Histamine Receptor.

1. INTRODUCTION

Alzheimer's disease (AD) is the most joint form of neurodegenerative insanity [1-3]. It forms for nearly 50–60% of the general occasions of dementia between persons over the age of 65 years [4]. Unfortunately, the remedial options for AD are limited. Alzheimer's disease (dementia) is a very active area of histamine H3 receptor research [5-6]. One of the mostly prescribed anti-Alzheimer's drugs are H3R antagonists [7-8]. Histamine is a biogenic amine that is a wide range of biological functions, including neurotrans-

*Corresponding Author: pourbasheer@ut.ac.ir

mission, inflammation and smooth muscle contraction; these effects are through the activation of different G protein-coupled receptors (GPCRs) [9].

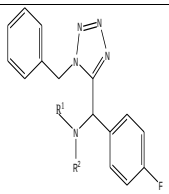
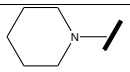
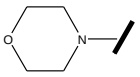
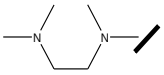
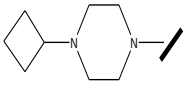
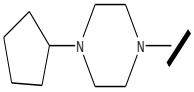
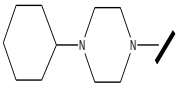
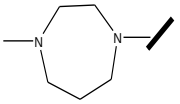
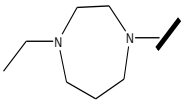
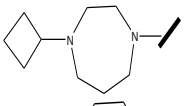
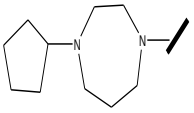
The histamine H3 receptor is predominantly explained in the central nervous system (CNS) and is recognized as a presynaptic receptor controlling the release of histamine and the release of several neurotransmitters [9-10]. The experimental determination of the inhibition activity of histamine is difficult, expensive and time-consuming, So an effective way to gain a complete data set without the necessity of performing expensive laboratory experiments is the application of quantitative structure activity relationship (QSAR) techniques[11-12]. Among computational methods, QSAR studies are one of the most important areas in chemometrics[13-15]. QSAR techniques are rapidly developing and have been widely used by chemists for predicting compounds' properties, including biological activity, physical properties, and toxicity [16-19]. The generated QSAR model could be used as a valuable tool for predicting the antagonist potency of benzyl tetrazoles derivatives as human histamine (H3) receptors. In the present work, Multiple linear regression (MLR) was used to derive the QSAR equations and stepwise (SW) and genetic algorithm (GA) methods were used for the selection of the most relevant descriptors from all of the descriptors[20-22]. The main aim of this report is to establish a new QSAR model for predicting the antagonist potency of benzyl tetrazoles derivatives as histamine (H3) receptors.

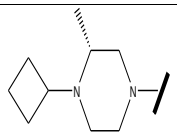
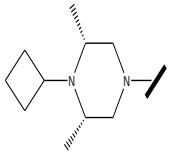
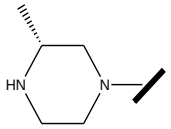
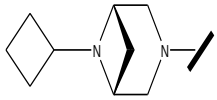
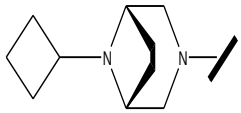
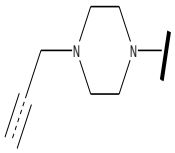
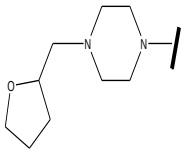
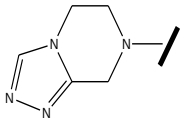
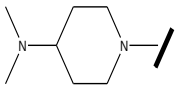
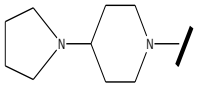
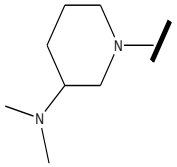
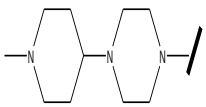
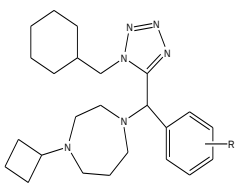
2. EXPERIMENTAL

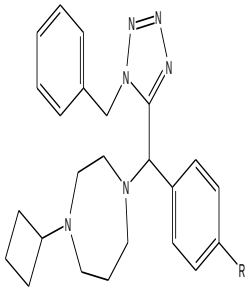
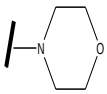
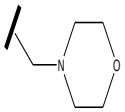
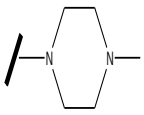
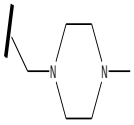
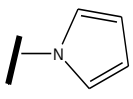
2.1 Data set

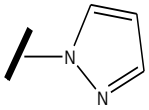
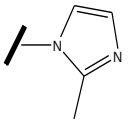
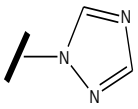
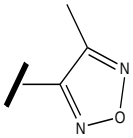
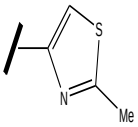
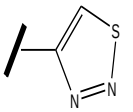
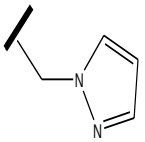
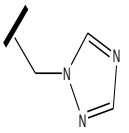
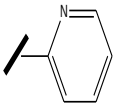
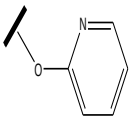
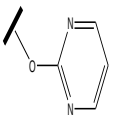
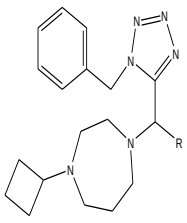
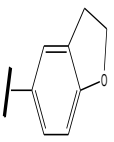
Data used in this QSAR study, a series of 64 benzyl tetrazoles derivatives as histamine H3 receptor antagonists were collected from the literature [23]. The chemical structures and activity data for the complete set of compounds can be seen in Table 1.

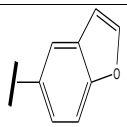
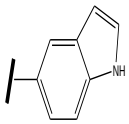
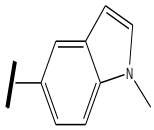
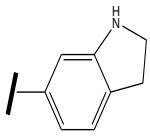
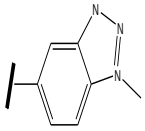
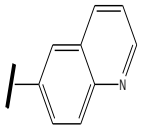
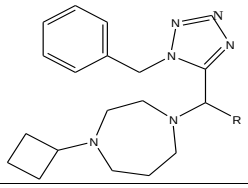
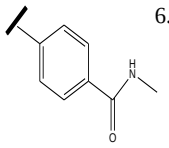
Table 1. Chemical structures and the corresponding observed and predicted pIC_{50} values by GA-MLR method.

NO	General structure	R^1 R^2	R	EExp.(pIC_{50})	Pred.(pIC_{50})
1			_	6.45	6.24
2			_	8.66	7.82
3[a]			_	8.47	8.82
4[a]			_	4.7	4.94
5			_	5.75	5.67
6[a]			_	6.62	6.24
7[a]			_	6.29	6.37
8			_	6.74	5.99
9			_	5.31	5.88
10			_	5.44	5.99

NO	General structure	R^1 R^2	R	EExp.(pIC_{50})	Pred.(pIC_{50})
11			_	5.61	6.03
12			-	5.61	5.31
13			-	4.85	5.62
14			-	4.7	4.72
15			-	6.32	7.64
16			-	5.72	5.89
17			-	7.19	6.75
18[a]			-	7.77	7.22
19			-	7.55	8.01
20			-	7.38	6.99
21			-	7.85	7.71
22			-	7.55	7.09
23		_	2-C1	7.29	7.54
24		_	2-OCHF	7.47	7.64
25		_	3-C1	6.94	6.84

NO	General structure	R^1 R^2	R	EExp.(pIC_{50})	Pred.(pIC_{50})
26[a]		_	3-OCHF	7.72	7.82
27		_	4-H	7.68	8.19
28		_	4-^{i}Pr	8.52	8.50
29		_	4-CF_2	8.17	8.87
30		_	4-OEt	6.26	7.79
31		_	4-OCHF	9.22	7.89
32		_	4-OCF_3	9.3	8.72
33		_	4-C1	9.4	9.25
34		_	4-Br	8.19	8.88
35		_	4-CONH	8.92	8.31
36		_	4-CONH	9.7	8.50
37[a]		_	4-CONH	8.8	8.49
38		_	4-Pyrrolidin-2-One	8.04	8.05
39[a]		_	4-SO_2Me	9.4	8.63
40			4-SO_2NM	8.44	8.41
41		_		7.51	6.71
42		_		8.23	8.41
43		_		9.05	9.52
44		_		7.23	7.61
45		_		7.39	7.21

NO	General structure	R^1 R^2	R	EExp.(pIC_{50})	Pred.(pIC_{50})
46		_		6.92	7.36
47		_		7.01	7.36
48		_		7.49	7.33
49		_		8.8	8.69
50		_		7.03	6.61
51		_		7.33	7.93
52[a]		_		4.7	5.58
53[a]		_		7.24	7.65
54		_		5.56	5.79
55		_		9.52	9.36
56[a]		_		8.8	8.69
57		_		7.6	7.22

NO	General structure	R^1 R^2	R	EExp.(pIC_{50})	Pred.(pIC_{50})
58		_		9.15	8.54
59		_		8.34	8.16
60		_		7.92	8.51
61		_		7.48	7.09
62[a]		_		6.55	6.83
63[a]		_		6.99	7.69
64		_		6.13	5.99

The inhibitory activity values are expressed as the half maximal inhibitory concentration (IC_{50}). The activity data [IC_{50} (nM)] was converted to the logarithmic scale pIC_{50} [-log IC_{50} (nM)] and then used for the subsequent QSAR analyses as the response variables. The data set was randomly divided into two subsets: the training set containing 51 compounds (80%) and the test set containing 13 compounds (20%). The training set was applied to build a regression model, and the test set was applied to evaluate the predictive ability of the model obtained.

2.2 Descriptors calculation

The first step to obtain a QSAR model is to encode the structural features of the molecules, which are often called molecular descriptors[24]. A molecular descriptor is the most important factor affecting the quality of a QSAR model[25]. The chemical structures of the 64 studied molecules were drawn into the Hyperchem (Version8.0) and the pre-optimized was conducted using MM+ molecular mechanics force field. Then a more precise optimization of these molecules was performed by the semi-empirical AM1 method. For these the molecular structures we carried out geometry optimization calculations using the Polak–Ribiere algorithm[26] until the root mean square gradient was 0.01. After optimization of the chemical structures of the studied molecules, the molecular descriptors were entered on the DRAGON 2.1 software for the calculation of the different types of theoretical descriptors for each molecule. Then, a total of 1497 molecular descriptors were calculated, they included a) 0D-constitutional descriptors, b) 1D-functional groups, atom centered fragments, c) 2D-topological descriptors, walk and path counts, ..., d) 3D-Randic molecular profiles, geometrical descriptors, ..., e) charge descriptors and f) molecular properties. The calculated descriptors with a constant or near constant value were first

analyzed and those detected were then eliminated from the data matrix. In addition, to decrease the redundancy existing in the descriptors, the correlations of the descriptors with each other and with the activity (pIC_{50}) values of the molecules were examined. Then, the collinear descriptors (i.e. $r > 0.9$) were detected. Among the collinear descriptors, the one presenting the highest correlation with the activity (pIC_{50}) was retained and the other descriptors were removed from the data matrix. Finally, 441 molecular descriptors remained.

2.3 Genetic algorithm (GA)

Genetic algorithm is a stochastic optimization technique that has been inspired by evolutionary principles[27]. The distinctive aspect of GA for researchers is finding the global optimal solution in a complex multidimensional search space. The first step in GA is a creation of an initial population of solutions (chromosomes). Each chromosome consists of a sequence of independent structures called genes which represent the descriptors. In the next step, each subset of chromosome is evaluating by its fitness to predict pIC_{50} values. Then, a fraction of the children of the next generation is produced by crossover and the rest by mutation from the parents based on their scaled fitness scores. This process is continuous until the evaluation of generation took 90% of same fitness [28]. The fitness function of our study was leave-one-out cross-validation (Q^2_{LOO}). The diagrammatic presentation of the genetic algorithm is shown in Fig. 1.

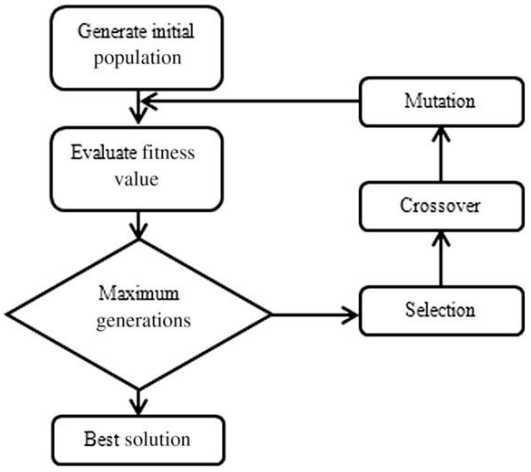


Fig. 1. The diagrammatic presentation of the genetic algorithm.

3. RESULT AND DISCUSSION

The pIC_{50} activity of 64 compounds (a training set of 51 compounds and a test set of 13 compounds from the whole available compounds with known activity), benzyl tetrazoles as histamine H3 receptor antagonists were taken from published results and used for the present QSAR study. The two sets are listed in Table 1 that the test set compounds are marked. After dividing the data set into a training set and a test set, the next stage was to choose the main factors that were the most important for the antagonist potency of the benzyl tetrazoles derivatives. The test data were not involved by any means in the process of selecting the most suitable descriptors or in the extension of the QSAR model. They were discussed as a completely unknown external set of data, which was applied only to test the accuracy of the produced model. In this QSAR study, we applied a stepwise regression (SW) and genetic algorithm (GA) as two common and powerful methods for selecting the most important descriptors, and then constructed multiple linear regression (MLR) models using the selected six descriptors. By SW-MLR modeling, the six most significant descriptors are nR04, IC4, MATS7v, RDF040v, Mor05v and G1u. The SW-MLR model is described by the following equation:

pIC_{50} = - 20.886 (±6.364) + 1.402 (±0.357) nR04 + 2.741 (±0.964) IC4 + 12.683 (±3.503) MATS7v - 0.139 (±0.059) RDF040v - 0.861 (±0.220) Mor05v + 61.631 (±22.677) G1u (1)

N_{train} = 51, R^2_{train}= 0.775, R^2_{test} = 0.512, R^2_{adj}=0.744, F_{train}= 25.281, F_{test}=0.655, $RMSE_{train}$= 0.611, $RMSE_{test}$ = 0.967, Q^2_{LOO}= 0.709, Q^2_{LGO}=0.673

In this equation, N is the number of training set compounds, Q^2_{LOO} and Q^2_{LGO}, are the squared cross-validation coefficients for leave one out and leave group out respectively. The R^2 is the squared correlation coefficient, R^2adj is adjusted R^2 and F is the Fisher F-statistic. The values in parentheses are the standard deviations. The good values for the training set and the poor values for the prediction set indicate that the SW-MLR procedure did not produce good results. For more investigation, the GA-MLR was used to select the best set of variables. To investigate the optimum number of descriptors, the influence of the number of the descriptors were selected until one to ten descriptors.

Finally, the GA–MLR analysis led to the derivation of one model, with six variables, which is described by equation 2:

pIC_{50} = 39.823 (±13.439) + 1.738 (±0.280) nR04 - 16.152 (±5.247) LP1 - 1.288 (±0.368) RDF020m + 0.116 (±0.019) Mor02m + 76.186 (±20.365) G1u - 26.512 (±5.506) HATS2e (2)

N_{train} = 51, R^2_{train}= 0.808, R^2_{test} = 0.903, R^2_{adj}=0.782, F_{train}= 30.806, F_{test}=6.391, $RMSE_{train}$= 0.565, $RMSE_{test}$ = 0.454, Q^2_{LOO}= 0.751, Q^2_{LGO}=0.669

Then, the built model was used to predict the test set data. The prediction results are given in Table 1. The predicted values of pIC_{50} for both the training and test sets using equation 2 were plotted against the experimental pIC_{50} values in Fig. 2. As can be seen from Table 1, the prediction results for the pIC_{50} are in good agreement with the experimental values in Fig. 2.

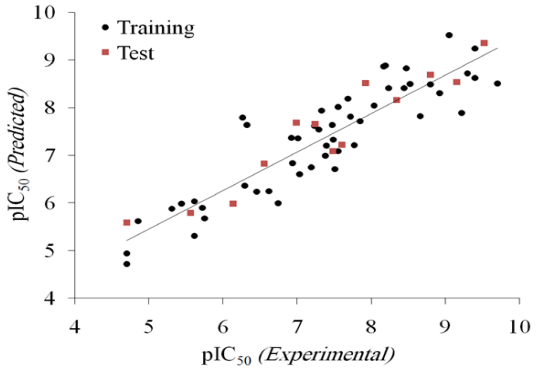


Fig. 2. The predicted pIC_{50} values by the GA-MLR modeling vs. the experimental pIC_{50} values.

Multi-collinearity between the above six descriptors were checked by calculating their variation inflation factors (VIF), which can be presented as follows:

$$VIF = \frac{1}{1 - r^2} \qquad (3)$$

where r is the correlation coefficient of the multiple regression between one variable and the others in the model. If VIF equals 1.0, then no intercorrelation exists for each variable; if VIF falls into the range 1.0–5.0, the related model is acceptable; and if VIF is larger than 10.0, the related model is unstable and recheck is necessary [29].

The corresponding VIF values of the six descriptors are shown in Table 2. As can be seen from Table 2, most of the variables had VIF values of less than 5, indicating that the obtained model has statistic significance.

The selected variables of GA-MLR model are shown in Table 3, and the correlation matrix obtained in this case is shown in Table 3. From Table 3, it could be seen that the correlation coefficient value of each pair descriptors was less than 0.43, which meant that the selected descriptors were independent. Testing the stability, predictive power and generalization ability of the models is a very important step in expression their quality [30]. In the paper, the ability of the GA-MLR model was evaluated by leave-one-out (LOO) and leave-group-out (LGO) cross-validation correlation coefficients[31, 32].

Table 2. Details of the constructed GA–MLR model

Descriptors	coefficients	Std.Error	MF[a]	VIF[b]	Chemical meanings
Constant	39.823	13.439	0	0	-
nR04	1.738	0.280	-0.041	1.997	number of 4-membered rings
LP1	-16.152	5.247	1.227	1.344	Lovasz-Pelikan index (leading eigenvalue)
RDF020m	-1.288	0.368	0.051	1.356	Radial Distribution Function - 2.0 / weighted by atomic masses
Mor02m	0.116	0.019	-0.089	1.429	3D-MoRSE - signal 02 / weighted by atomic masses
G1u	76.186	20.365	-0.342	1.823	1st component symmetry directional WHIM index / unweighted
HATS2e	-26.512	5.506	0.194	1.676	leverage-weighted autocorrelation of lag 2 / weighted by atomic Sanderson electronegativities

[a] Mean effect
[b] Variation inflation factor

Table 3. Correlation coefficient matrix of the selected descriptors

	nR04	LP1	RDF020m	Mor02m	G1u	HATS2e
nR04	1	0	0	0	0	0
LP1	0.274	1	0	0	0	0
RDF020m	0.248	-0.331	1	0	0	0
Mor02m	0.361	-0.076	0.231	1	0	0
G1u	-0.483	-0.065	-0.183	-0.492	1	0
HATS2e	-0.532	-0.033	-0.192	-0.253	0.437	1

The Q^2_{LOO} and Q^2_{LGO}, parameters for the MLR model are shown in equation 2. The results show that GA-MLR model has a good internal and external predictive power**.**

There are several methods for defining the applicability domain (AD) of QSAR models, but the most common one is determining the leverage values for each compound. To visualize the AD of a QSAR model, the Williams plot – the plot of the standardized residuals versus the leverage values (h) was exploited in this study [33]. By analyzing the applicability domain of the model (2) from the Williams plot (Fig. 3) it can be seen that three compounds (No.6, No.14 and No.25) from the training set with h > h* and the standardized residuals <3δ are identified as structural outliers and high leverage chemicals.

The QSAR model was confirmed by applying Y-randomization. Several random shuffles of the Y vector (pIC_{50}) were performed and the low R^2 and Q^2_{LOO} values that were obtained showed that the good results in the original model use were not due to a chance correlation or structural dependency of the training set [34]. The results of the Y-randomization test are presented in Table 4.

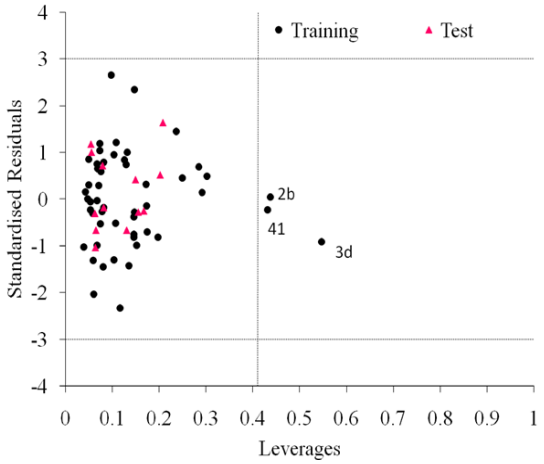


Fig. 3. The William plot of the GA-MLR model.

Table 4. R^2_{train} and Q^2_{LOO} values after several Y-randomization tests

No	Q^2	R^2
1	0.0350	0.1029
2	0.1425	0.0448
3	0.0256	0.2392
4	0.0067	0.1184
5	0.0130	0.0816
6	0.2287	0.0281
7	0.1516	0.0546
8	0.0145	0.1471
9	0.0031	0.1417
10	0.0359	0.0594

4. CONCLUSIONS

In this article, a QSAR study of 64 histamine H3 receptor antagonists was performed based on the theoretical molecular descriptors calculated by the DRAGON software and selected by SW-MLR and GA-MLR methods. The GA-MLR model clearly demonstrates good correlations between the structure and activity of the studied compounds. This model has good stability, robustness and predictive ability, which were verified by internal validation (cross-validation by LOO, LGO, and Y-randomization) and external validation**.**

REFERENCES

[1] Y.w. Zhang, S. Liu, X. Zhang, W.B. Li, Y. Chen, X. Huang, L. Sun, W. Luo, W.J. Netzer, R. Threadgill, G. Wiegand, R. Wang, S.N. Cohen, P. Greengard, F.F. Liao, L. Li and H. Xu, A Functional Mouse Retroposed Gene Rps23r1 Reduces Alzheimer's β-Amyloid Levels and Tau Phosphorylation, *Neuron*. 64 (2009) 328-340.

[2] H.O. Tayeb, H.D. Yang, B.H. Price and F.I. Tarazi, Pharmacotherapies for Alzheimer's disease: beyond cholinesterase inhibitors, *Pharmacol. Ther.* 134 (2012) 8-25.

[3] D. Cvetkovic-Dozic, M. Skender-Gazibara and S. Dozic, Neuropathological hallmarks of Alzheimers disease, *Arch. Oncol.* 9 (2001) 195-199.

[4] S.D. Bembenek, J.M. Keith, M.A. Letavic, R. Apodaca, A.J. Barbier, L. Dvorak, L. Aluisio, K.L. Miller, T.W. Lovenberg and N.I. Carruthers, Lead identification of acetylcholinesterase inhibitors–histamine H3 receptor antagonists from molecular modeling, *Bioorg. Med. Chem.Lett.* 16 (2008) 2968-2973.

[5] R. Nirogi, V. Kandikere, G. Bhyrapuneni, N. Muddana, R. Saralaya, R.K. Ponnamaneni and A.K. Manoharan, In vivo receptor occupancy assay of histamine H< sub> 3</sub> receptor antagonist in rats using non-radiolabeled tracer, *J. Pharmacol. Toxicol. Methods.* 65 (2012) 115-121.

[6] R.R. Dandu, J.A. Gruner, J.R. Mathiasen, L.D. Aimone, G. Hostetler, C. Benfield, R.J. Bendesky, V.R. Marcy, R. Raddatz and R.L. Hudkins, Synthesis and evaluation of pyridazinone-phenethylamine derivatives as selective and orally bioavailable histamine H3 receptor antagonists with robust wake-promoting activity, *Bioorg. Med. Chem. Lett.* 21 (2011) 6362-6365.

[7] O. Roche, M. Nettekoven, W. Vifian and R.M.R. Sarmiento, Refinement of histamine H3 ligands pharmacophore model leads to a

new class of potent and selective naphthalene inverse agonists, *Bioorg. Med. Chem. Lett.* 18 (2008) 4377-4379.

[8] R.M. Rodríguez Sarmiento, M.H. Nettekoven, S. Taylor, J.M. Plancher and H. Richter, O. Roche, Selective naphthalene H3 receptor inverse agonists with reduced potential to induce phospholipidosis and their quinoline analogs, *Bioorg. Med. Chem. Lett.* 19 (2009) 4495-4500.

[9] R.O. Millán-Guerrero, L.M. Baltazar-Rodríguez, M.I. Cárdenas-Rojas, M. Ramírez-Flores, S. Isais-Millán, I. Delgado-Enciso, R. Caballero-Hoyos and B. Trujillo-Hernández, A280V polymorphism in the histamine H3 receptor as a risk factor for migraine, *Arch. Med. Res.* 42 (2011) 44-47.

[10] Y. Mitobe, S. Ito, T. Mizutani, T. Nagase, N. Sato and S. Tokita, Development of a selective and potent radioactive ligand for histamine H3 receptors: A compound potentially useful for receptor occupancy studies, *Bioorg. Med. Chem. Lett,* 19 (2009) 4075-4078.

[11] E.P. Reza Aalizadeh and M.R. Ganjali, Analysis of B-RafV600E inhibitors using 2D and 3D-QSAR, molecular docking and pharmacophore studies, *Mol. Divers* 19 (2015) 915-930.

[12] E. Pourbasheer, S. Riahi, M.R. Ganjali and P. Norouzi, QSAR study of C allosteric binding site of HCV NS5B polymerase inhibitors by support vector machine, *Mol. Divers.* 15 (2011) 645-653.

[13] E. Pourbasheer, R. Aalizadeh, H. Shiri, A. Banaei and M. Ganjali, 2D and 3D-QSAR analysis of pyrazole-thiazolinone derivatives as EGFR kinase inhibitors by CoMFA and CoMSIA, *Curr. Comput Aided. Drug. Des.* (2015).

[14] E. Pourbasheer, R. Bazl and M. Amanlou, Molecular docking and 3D-QSAR studies on the MAPKAP-K2 inhibitors, *Med. Chem. Res.* 23 (2014) 2252-2263.

[15] E. Pourbasheer and M. Amanlou, 3D-QSAR analysis of anti-cancer agents by CoMFA and CoMSIA, *Med. Chem. Res.* 23 (2014) 800-809.

[16] A. Beheshti, S. Riahi and M.R. Ganjali, Quantitative structure–property relationship study on first reduction and oxidation potentials of donor-substituted phenylquinolinylethynes and phenylisoquinolinylethynes: Quantum chemical investigation, *Electrochim. Acta*. 54 (2009) 5368-5375.

[17] Y.M. Alvarez-Ginarte, R. Crespo-Otero, Y. Marrero-Ponce, P. Noheda-Marin, J.M. Garcia de la Vega, L.A. Montero-Cabrera, J.A. Ruiz García, J.A. Caldera-Luzardo and Y.J. Alvarado, Chemometric and chemoinformatic analyses of anabolic and androgenic activities of testosterone and dihydrotestosterone analogues, *Bioorg. Med. Chem*.16 (2008) 6448-6459.

[18] J. Caballero, M. Fernández, M. Saavedra and F.D. González-Nilo, 2D Autocorrelation, CoMFA, and CoMSIA modeling of protein tyrosine kinases' inhibition by substituted pyrido [2, 3- d] pyrimidine derivatives, *Bioorg. Med. Chem*. 16 (2008) 810-821.

[19] M. Fatemi and E. Baher, Quantitative structure–property relationship modelling of the degradability rate constant of alkenes by OH radicals in atmosphere, *SAR. QSAR. Environmen. Res.* 20 (2009) 77-90.

[20] E. Pourbasheer, R. Aalizadeh, M.R. Ganjali, P. Norouzi, J. Shadmanesh and C. Methenitis, QSAR study of Nav1.7 antagonists by multiple linear regression method based on genetic algorithm (GA-MLR), *Med. Chem. Res.* 23 (2014) 2264-2276.

[21] E. Pourbasheer, R. Aalizadeh, M.R. Ganjali, P. Norouzi and A. Banaei, QSAR study of mGlu5 inhibitors by genetic algorithm-multiple linear regressions, *Med. Chem. Res.* 23 (2014) 3082-3091.

[22] M. Nekoei, M. Salimi, M. Dolatabadi and M. Mohammadhosseini, Prediction of antileukemia activity of berbamine derivatives by genetic algorithm–multiple linear regression, *Monatsh. Chem.* 142 (2011) 943-948.

[23] A.J. Davenport, C.C. Stimson, M. Corsi, D. Vaidya, E. Glenn, T.D. Jones, S. Bailey, M.J. Gemkow, U. Fritz and D.J. Hallett, Discovery of substituted benzyl tetrazoles as histamine H3 receptor antagonists, *Bioorg. Med. Chem.Lett.* 20 (2010) 5165-5169.

[24] M. Nekoei, M. Mohammadhosseini and E. Pourbasheer, QSAR study of VEGFR-2 inhibitors by using genetic algorithm-multiple linear regressions (GA-MLR) and genetic algorithm-support vector machine (GA-SVM): a comparative approach, *Med. Chem. Res.* (2015) 1-10.

[25] L. Firoozpour, K. Sadatnezhad, S. Dehghani, E. Pourbasheer, A. Foroumadi, A. Shafiee and M. Amanlou, An efficient piecewise linear model for predicting activity of caspase-3 inhibitors, DARU, *J. Pharmaceut. Sci.* 20 (2012) 31-36.

[26] E. Polak and G. Ribiere, Note sur la convergence de méthodes de directions conjuguées." *ESAIM: Mathematical Modelling and Numerical Analysis* -

Modélisation Mathématique et Analyse Numérique 3 (1969) 35-43.

[27] S. Riahi, M.R. Ganjali, E. Pourbasheer, F. Divsar, P. Norouzi and M. Chaloosi, Development and validation of a rapid chemometrics-assisted spectro-photometry and liquid chromatography methods for the simultaneous determination of the phenylalanine, tryptophan and tyrosine in the pharmaceutical products, *Curr. Pharm. Anal.* 4 (2008) 231-237.

[28] Z. Bayat and S. Vahdani, A simple and readily integratable approach to toxicity prediction for some of the anti cancer drugs, *J. Chem. Pharm. Res*. 3 (2011) 93-102.

[29] S. Vahdani and Z. Bayat, A Quantitative Structure-Activity Relationship (QSAR) Study of Anti-cancer Drugs, *Der. Chemica. Sinica.* 2 (2011) 235-243.

[30] A.S. Om, J.C. Ryu and J.H. Kim, Quantitative structure-activity relationships for radical scavenging activities of flavonoid compounds by GA-MLR technique, *Mol. Cell. Toxicol.* 4 (2008) 170-176.

[31] M. Adimi, M. Salimi, M. Nekoei, E. Pourbasheer and A. Beheshti, A quantitative structure activity relationship study on histamine receptor antagonists using the genetic algorithm multi parameter linear regression method, *J. Serb. Chem. Soc.* 77 (2012) 639–650.

[32] J. Li, B. Lei, H. Liu, S. Li, X. Yao, M. Liu and P. Gramatica, QSAR study of malonyl-CoA decarboxylase inhibitors using GA-MLR and a new strategy of consensus modeling, *J. comput. chem.* 29 (2008) 2636-2647.

[33] H. Kušić, B. Rasulev, D. Leszczynska, J. Leszczynski and N. Koprivanac, Prediction of rate constants for radical degradation of aromatic pollutants in water matrix: A QSAR study, *Chemosphere.* 75 (2009) 1128-1134.

[34] M. Jalali-Heravi and M. Asadollahi-Baboli, QSAR Analysis of Platelet-derived Growth Inhibitors Using GA-ANN and Shuffling Crossvalidation, *QSAR. Comb. Sci.* 27 (2008) 750-757.