ICC
Iranian Chemical Communication
Original Research Article
Payame Noor University http://icc.journals.pnu.ac.ir

QSAR studies and application of genetic algorithm - multiple linear regressions in prediction of novel p2x7 receptor antagonists' activity

Alireza Banaei, Eslam Pourbasheer*, Fatemeh Haggi

Department of Chemistry, Payame Noor University, P.O. BOX 19395-3697, Tehran, Iran

Received: 25 July 2015, Accepted: 6 January 2016, Published: 6 January 2016

Abstract

Quantitative structure-activity relationship (QSAR) models were employed to predict the activity of P2X7 receptor antagonists. A data set consisted of 50 purine derivatives was utilized in the model construction where 40 and 10 of these compounds were in the training and test sets respectively. A suitable group of calculated molecular descriptors was selected by employing stepwise multiple linear regressions (SW-MLR) and genetic algorithm-multiple linear regressions (GA-MLR) as variable selection tools. The proposed MLR models were fully confirmed applying internal and external validation techniques. The obtained results of this QSAR study showed the superiority of the GA-MLR model over the SW-MLR model. As a result, the obtained GA–MLR model could be applied as a valuable model for designing similar groups of P2X7 receptor antagonists.

Keywords: QSAR; genetic algorithms; P2x7 receptor antagonists; purine derivatives.

Introduction

P2X7 receptor, a plasma membrane receptor for extracellular adenosine-5-triphosphate (ATP) predominantly explained in inflammation relevant cells, has been known as an important regularizer of both IL-1 maturation and externalization [1-3]. Activation of $P2X_7$ receptors results in two distinct responses, depending on the exposure

*Corresponding author: Eslam Pourbasheer
Tel: +98 (45) 33519448, Fax: +98 (45) 33519448
E-mail: pourbasheer@ut.ac.ir

time to agonist. A brief excitation of the receptor by extracellular ATP leads to opening membrane channel permeable to small cations (Na^+ ,Ca^{2+}, K^+) [4-6]. $P2X_7$ receptor is implicated in the regulation, expression and secretion of cytokines and inflammatory mediators including interleukins (IL-1) [2], IL-2, [7], IL-18 [8] and tumor necrosis factor-α [9]. Notably, the $P2X_7$ receptor plays an essential role in the processing and release of pro-inflammatory cytokine IL-1β in the immune system by a complex sequence of events. Initially, the activated $P2X_7$ receptor causes the reduction of K^+ that leads to the excitation of IL-1β converting enzyme (caspase-1), and further convert LPS-activated pro-IL-1β to mature IL-1β [2]. Due to the presence of $P2X_7$ receptors on cells of immune system (macrophages, microglia, etc.) and the relationship between $P2X_7$ activation and cytokine or glutamate release, this receptor may play an important role in the development and progression of various disease states or conditions such as chronic inflammation [10-11], neurodegeneration [12-15], and chronic pain [16-17]. Therefore, the development of antagonists of the $P2X_7$ receptor could be a therapeutic strategy to treat inflammatory diseases.

Several $P2X_7$ receptor antagonists with diverse scaffolds have been discovered so far. Till date no $P2X_7$ receptor antagonist reached into the market ultimately as a drug due to lack of potency and efficacy. The analogous molecules have now been discontinued because of lack of efficacy [8]. Therefore, there is still a need to identify and develop clinically effective and well tolerated $P2X_7$ receptor antagonists.

Nowadays, there is a growing interest to use the computational approaches in predicting the activities of new chemical structures before synthesis. Among the computational methods, quantitative structure-activity relationship (QSAR) is a well known method to describe the chemical structures-biological activities interactions [18-24]. The first step, also the most important step in construction the QSAR models is the choice of a series of molecular descriptors with the higher effect on the biological activity [25].

Recently, some QSAR studies have been reported about the p2x7 receptor antagonists activities [26-27]. In the

present study, multiple linear regression (MLR) technique along with stepwise (SW) and genetic algorithm (GA) variable selection methods were used for generating the QSAR models [28]. The impetus of the present study is to propose and develop a novel QSAR model to predict the antagonist potency of purine derivatives as $P2X_7$ receptors.

Materials and Methods

Data set

In this QSAR study, a dataset consisted of 50 compounds of purine derivatives as $P2X_7$ receptors antagonists, was collected from the literature [29-30]. The chemical structures and activity data for whole compounds are given in Table 1. The activity data as IC_{50} (nM) values for the total compounds were converted to logarithmic scale pIC_{50} (M) and then applied for QSAR analysis as the dependent variables. The whole dataset were accidentally separated into two training and test sets in which 40 and 10 compounds were obtained for each set respectively. The test set was applied to determine the predictive ability of the created model based on training set [31].

Descriptors calculation

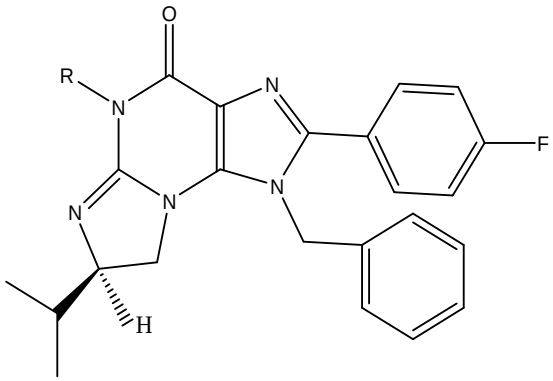
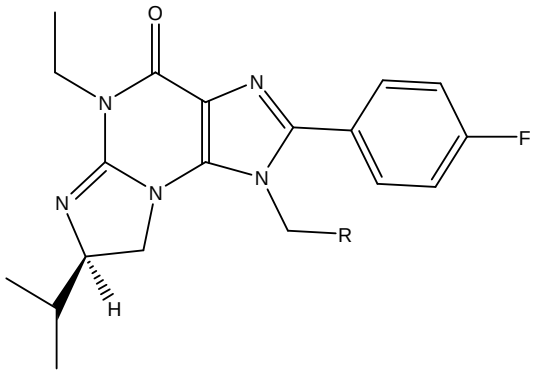
The second dimensional structures were constructed in Hyperchem 7.0 software and these structures were optimized using molecular mechanics force field (MM+) and semi-empirical (AM1) methods with the root mean square gradient of 0.01 kcal mol^{-1}. Calculation of molecular descriptors for each compound of data set has been followed by using the Dragon 2.1 software. Totally, 1497 different molecular descriptors were generated for each compound. In order to decrease the redundancy existing in the descriptors data matrix, the correlations of descriptors with each other and with pIC50 of the molecules are examined, and collinear descriptors ($R > 0.9$) are detected. Those of the descriptors which have the pair wise correlation coefficient above 0.9 and having the lower correlation with pIC_{50} values are removed from the data matrix. They were excluded in the pre-reduction step and 387 descriptors were left for variable selection.

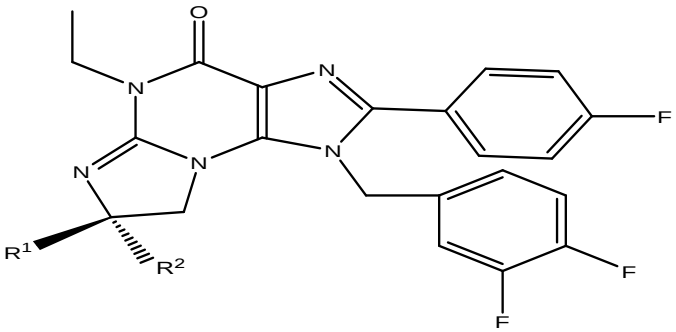
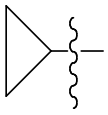
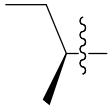
Genetic algorithm (GA)

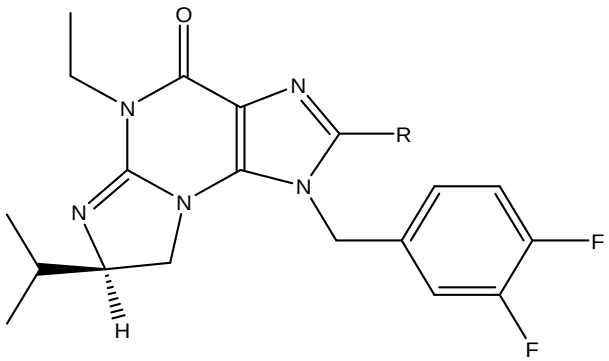
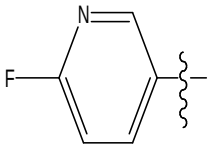
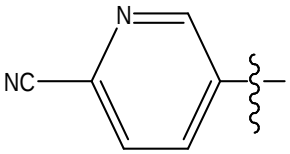
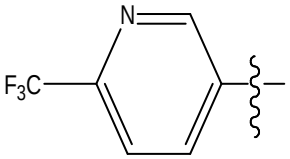
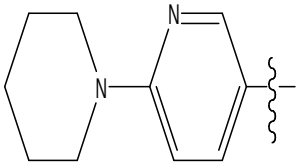
Choosing appropriate descriptors for QSAR studies is difficult, because there are no absolute rules that govern this choice. In other words, the problem is to choose the best possible group of descriptors from all present descriptors that can predict with minimum error in

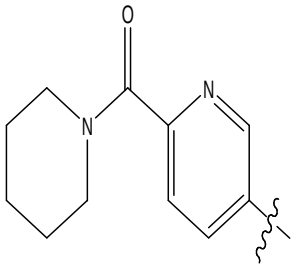
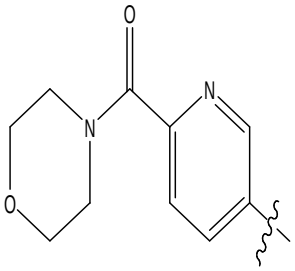
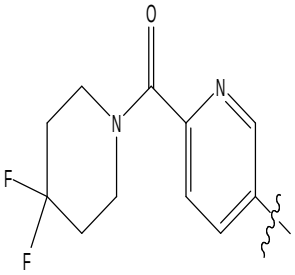
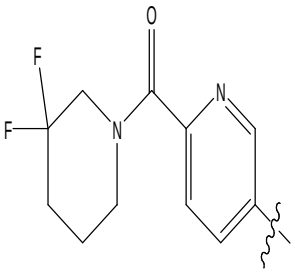
comparison to the experimental data. A generally accepted method for this problem is genetic algorithm-based multiple linear regression (GA-MLR). Genetic algorithm is a random optimization procedure that imitates selection in nature and has been demonstrated to be a very useful tool in QSAR studies with many merits [32]. In this research, GA-MLR was employed to establish a QSAR model. The fitness function applied in this research was the leave-one-out cross-validated correlation coefficient (Q^2_{LOO}). The GA-MLR program is implemented in Matlab 6.5 software [33].

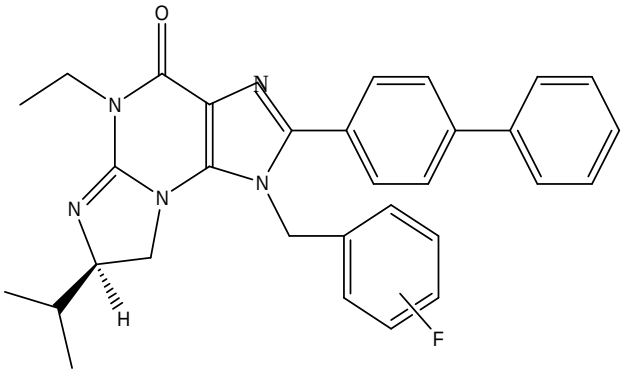
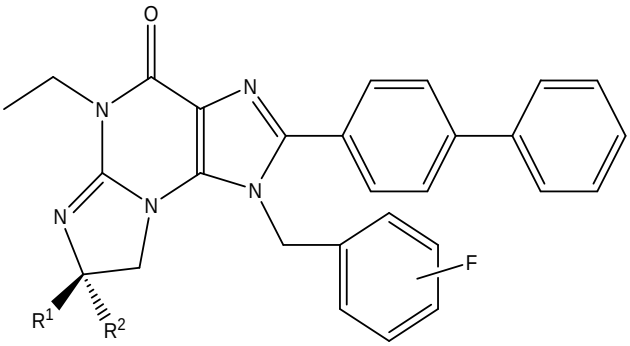
Table 1. Chemical structures and the corresponding experimental and predicted pIC_{50} values by GA-MLR method

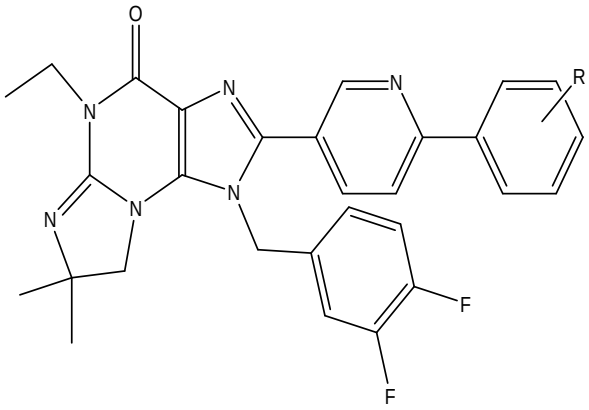
NO.	R	R^1	R^2	Fluorine substitution pattern on benzyl ring	Exp	Pred
						
1	Et	-	-	-	5.5	5.53
						

NO.	R	R^1	R^2	Fluorine substitution pattern on benzyl ring	Exp	Pred
2	3F,4FPh	-	-	-	6.5	6.70
3	2F,4FPh	-	-	-	6.0	6.04
4[a]	4FPh	-	-	-	6.2	6.29
5	3F,5FPh	-	-	-	6.3	6.26
6	3,4,5-TriF-Ph	-	-	-	6.5	6.33
						
7	-	Butyl	H	-	6.2	6.50
8	-		H	-	6.6	6.69
9	-		H	-	6.4	6.28
10[a]	-	Me	Me	-	6.4	6.87

NO.	R	R^1	R^2	Fluorine substitution pattern on benzyl ring	Exp	Pred
						
11[a]	4-CF_3OPh	-	-	-	6.6	6.59
12	3-CF_3OPh	-	-	-	6.5	6.67
13	4CN-Ph	-	-	-	6.65	6.60
14	4CN,3F-Ph	-	-	-	6.7	6.70
15		-	-	-	6.6	6.48
16		-	-	-	6.8	7.01
17		-	-	-	7.0	7.23
18		-	-	-	6.9	6.84

NO.	R	R^1	R^2	Fluorine substitution pattern on benzyl ring	Exp	Pred
19[a]		-	-	-	7.0	7.06
20		-	-	-	7.2	7.0
21		-	-	-	7.2	7.41
22		-	-	-	7.4	7.47

NO.	R	R^1	R^2	Fluorine substitution pattern on benzyl ring	Exp	Pred
						
23	-	-	-	3,4	6.7	6.73
24[a]	-	-	-	2	6.0	6.05
25	-	-	-	3	6.5	6.62
26	-	-	-	4	6.6	6.90
27	-	-	-	2,3	6.0	6.15
28	-	-	-	2,4	6.0	6.23
29	-	-	-	2,5	5.9	5.92
30	-	-	-	3,5	6.9	6.52
						
31	-	Me	Me	3,4	7.0	7.25

NO.	R	R^1	R^2	Fluorine substitution pattern on benzyl ring	Exp	Pred
32	-	Cyclohexyl	-	3,4	6.4	6.56
33[a]	-	Me	Me	3,4	6.9	7.23
34[a]	-	Me	Me	4	6.5	6.70
35	-	Me	Me	2,4	6.1	6.11
36	-	Me	Me	3,5	7.0	7.19
						
37	-	H	-	-	7.6	7.29
38[a]	-	2-Fluoro	-	-	7.2	6.91
39	-	2-Methoxy	-	-	7.0	7.16
40[a]	-	3-Methyl	-	-	7.6	7.41
41	-	3-Trifluoromethyl	-	-	6.9	7.08
42	-	3-Fluoro	-	-	7.5	7.28
43	-	3-Trifluoro	-	-	6.9	6.67

NO.	R	R^1	R^2	Fluorine substitution pattern on benzyl ring	Exp	Pred
		methoxy				
44	-	4-Methyl	-	-	7.4	7.15
45[a]	-	4-Trifluoro methyl	-	-	6.8	6.62
46	-	4-Fluoro	-	-	7.3	7.28
47	-	4-Methoxy	-	-	7.3	7.32
48	-	4-Trifluoro methoxy	-	-	5.7	6.41
49	-	2,4-Difluoro	-	-	7.0	6.86
50	-	3,5-Difluoro	-	-	7.2	7.22

[a]Test set

Results and discussion

For choosing training and test sets, the compounds are sorted from low to high experimental activity values. Then the whole data set was haphazardly partitioned into a training set (containing 40 compounds) and a test set (containing 10 compounds) with proportion 80% and 20%, respectively. Both sets are demonstrated in Table 1. For the selection of the most relevant descriptors, stepwise variable selection method was applied. Finally, SW–MLR model based on training set and using the chosen descriptors was created, and the following linear equation was obtained:

pIC_{50}= -3.077(2.40) -46.511(12.350) JGI6 +133.295(25.231) JGI10 +0.834(0.085) Mor09m -6.481(0.828) R8m +138.424(32.706) JGI9 +83.597(15.915) G1m +0.052(0.010) RDF070m (1)

N_{train}= 40, R^2_{train}= 0.906, R^2_{test}= 0.239, R^2_{adj}= 0.885, F_{train}= 43.859, F_{test}= 0.330, $RMSE_{train}$= 0.154, $RMSE_{test}$= 0.597, Q^2_{LOO}= 0.855, Q^2_{LGO}=0.843, Q^2_{BOOT}=0.844

In the above equation, N is the number of chemical compounds available in training set, R^2 is the squared correlation coefficient, R^2_{adj} is adjusted R^2, *RMSE* is the root mean square error, F is the Fisher F statistic and Q^2_{LOO}, Q^2_{LGO} and Q^2_{BOOT} show the squared cross-validation coefficients for leave one out, leave group out and bootstrapping, respectively. The obtained statistical parameters indicated that the SW-MLR model created acceptable results for the training set, but it did not create satisfactory results for the test set. Therefore, the genetic algorithm method was applied to select the best subset of descriptors. The GA-MLR model and its statistical parameters are presented as:

pIC_{50}= 8.383(±0.718) -73.224(±16.384) JGI6 +252.521(±28.016) JGI9 -0.068(±0.018) RDF085p -0.492(±0.072) Mor04m -5.978(±1.063) R8m -35.464(±8.806) R5e+ (2)

N_{train}= 40, R^2_{train}= 0.833, R^2_{test}= 0.755, R^2_{adj}= 0.802, F_{train}= 27.385, F_{test}= 1.499, $RMSE_{train}$= 0.206, $RMSE_{test}$ = 0.228, Q^2_{LOO}= 0.762, Q^2_{LGO}= 0.676, Q^2_{BOOT}= 0.734

As can be seen, the statistical parameters obtained by GA–MLR model demonstrate the acceptable results for both training and test sets. The GA-MLR model with six selected descriptors was applied to predict the activity values. The experimental and predicted values for pIC_{50}, were provided in Table 1. The plot of the predicted activities against the experimental activities was indicated in Figure 1. As can be seen from Table 1 and Figure 1, the calculated values of pIC_{50} are in good agreement with the experimental values.

The multi-collinearity, among descriptors available in model, was inspected by calculating the variation inflation factors (VIF) as below:

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

In above formula, r is correlation coefficient of multiple regressions among one descriptor with other descriptors in the QSAR model. A VIF equal to 1 displays that there is not any relationship among descriptors; if VIF value states among 1 to 5, it shows that proposed model is acceptable, and VIF greater than 10 means that the obtained model is not proper [22]. The multi-collinearity results are shown in Table 2. As is obvious in Table 2, the obtained VIF values for the most

descriptors are less than 5, which means, that the descriptors used in GA-MLR model are fairly independent of each other.

Table 2. The linear model based on the six descriptors selected by GA-MLR method

Descriptors	coefficients	Std.Error	MF[a]	VIF[b]	Chemical meanings
Constant	8.383	0.718	0	0	-
JGI6	-73.224	16.384	1.742	1.938	Mean topological charge index of order6
JGI9	252.521	28.016	-2.475	1.359	Mean topological charge index of order9
RDF085p	-0.068	0.018	0.503	1.846	Radial Distribution Function - 085 / weighted by atomic polarizabilities
Mor04m	-0.492	0.072	-0.854	1.908	3D-MoRSE - signal 04 / weighted by atomic masses
R8m	-5.978	1.063	1.462	3.058	R autocorrelation of lag 8 / weighted by atomic masses
R5e+	-35.464	8.806	0.622	1.122	R maximal autocorrelation of lag 5 / weighted by atomic Sanderson electronegativities

[a]Mean effect

[b]Variation inflation factors

Also, the correlation coefficient matrix of pair selected descriptors is shown in Table 3. A check of Table 3 indicates that the highest correlation coefficient value between pairwise descriptors is smaller than 0.57, which means that these descriptors are not highly dependent on each other. The mean effect (MF) value is applied to estimate the relative importance and also the contribution of each descriptor in model. This parameter demonstrates the relative importance of a descriptor in comparison to the other selected descriptors in the model. The MF sign indicates the variation direction in the values of the dependent variable as a result of the increase or decrease in the independent variables values [24]. The MF values are presented in Table 2. The leave one out and leave group out cross-validation techniques used to assessment the predictive power of the created models. In order to appraise the robustness and predictive ability of the models, Q^2_{BOOT} is also calculated based on bootstrapping repeated 5000 times [34]. The cross-validation results

indicate that the created GA-MLR model is a valid model so; it can be applied to calculate the activity values of the $P2X_7$ receptor antagonists.

Table 3. Correlation coefficient matrix of the selected descriptors by GA-MLR

	JGI6	**JGI9**	**RDF085p**	**Mor04m**	**R8m**	**R5e+**
JGI6	1	0	0	0	0	0
JGI9	-0.08	1	0	0	0	0
RDF085p	-0.17	-0.09	1	0	0	0
Mor04m	-0.55	0.16	-0.03	1	0	0
R8m	0.57	0.19	-0.54	-0.48	1	0
R5e+	0.29	0.19	-0.07	-0.16	0.17	1

The Williams plot, the plot of the standardized residuals (δ) against hat values (h), is applied to determining the applicability domain (AD) of the obtained GA-MLR model [35]. The leverage of a compound in the original variable space is defined as:

$$h_i = x_i^T \left(X^T \right)^{-1} x_i \qquad (4)$$

Where *xi* is the descriptor vector of the considered compound and X is the descriptor matrix derived from the training set descriptor values. The warning leverage (h*) is defined as:

$$h^* = \frac{3p}{n} \qquad (5)$$

Where n is the number of calibration compounds, p is the number of model variables plus one. The leverage (h) greater than the warning leverage (h^*) suggested that the compound was very influential on the model. Moreover, a value of 3 for standardized residual is commonly used as a cut-off value for accepting predictions, because points that lie ±3 standardized residual from the mean cover 99% of normally distributed data.

The Williams plot (Figure 2) indicates that only one compound (No. 29 in the training set) has the leverage (h) more than the warning h^* value of 0.53, thus it can be considered as structural outlier (*X* outlier). Also in this Figure, there is one compound (No. 48 in the training set) with standard residuals >3δ. Thus compound number 48 can be considered as response outlier (*Y* outlier).

In order to further appraise the robustness of the model generated by the GA-MLR procedure, Y-randomization test was utilized [31]. In this test, the activity values for set of molecules was shuffled randomly and after several replications the constructed model indicated to have less value for R^2 and Q^2_{LOO} values (Table 4).

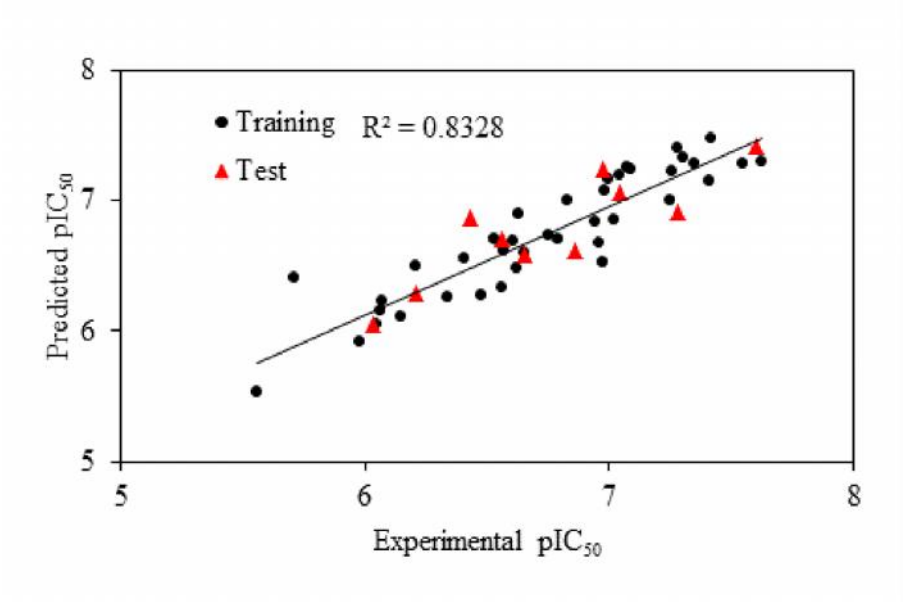


Figure1. The predicted pIC50 values by the GA-MLR modeling against the experimental pIC50 values

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

No	Q^2	R^2
1	0.002	0.153
2	0.003	0.223
3	0.117	0.339
4	0.022	0.114
5	0.002	0.139
6	0.084	0.073
7	0.307	0.038
8	0.016	0.127
9	0.017	0.194
10	0.131	0.360

Interpretation of descriptors

Analysis of the descriptors contained in the GA–MLR model can provide better insights to design some new compounds with higher activities. The brief descriptions of the selected descriptors are presented in Table 2.

The first two descriptors are JGI6 (Mean topological charge index of order 6) and JGI9 (Mean topological charge index of order 9) which belong to the Galvez topological charge indices. These descriptors explain

charge transfer among pairs of atoms and therefore global charge transfer in a molecule [36]. In Table 2, JGI 6 coefficient has negative sign which suggests that the pIC_{50} value is contrariwise related to this descriptor. The positive coefficient sign and the highest mean effect value for JGI9 illustrates that this descriptor has a direct and significant effect on the pIC_{50} value of the studied compounds.

The third appearing descriptors in the model is RDF085p (Radial Distribution Function - 085 / weighted by atomic polarizabilities) which belong to the radial distribution function (RDF) descriptors. RDF in this form meets all the requirements for the 3D structure descriptors. It is independent of the atom number (i. e. the size of a molecule), and is unique regarding the three-dimensional arrangement of the atoms and also it is constant versus the translation and rotation of the whole molecule. In addition, these descriptors can be limited to specific atom types or distance ranges to demonstrate specific information in a certain three-dimensional structure space (e.g. to describe the steric hindrance or the structure / activity properties of a molecule) [37]. In this descriptor, weighting scheme is the atomic polarizabilities which illustrate that the polarizabilities of the molecule atoms play a significant role in RDF085p descriptor. The RDF085p coeeficient has a negative sign, which illustrates that this descriptor has negative effect on the pIC_{50} value, which means that increasing the value of above descriptor with increasing the atomic polarizabilities for each compound atoms, the pIC_{50} value is reducing.

The Mor04m (3D-MoRSE - signal 04 / weighted by atomic masses) is a type of the 3D-MoRSE descriptors. 3D-MoRSE descriptors (3D Molecule Representation of Structures based on Electron diffraction) are obtained from Infrared spectra simulation applying a generalized scattering function [35]. The negative sign of Mor04m descriptor in Table 2 illustrates that with increasing atomic masses will decrease the pIC_{50} value.

The next two descriptors are R8m (R autocorrelation of lag 8 / weighted by atomic masses) and R5e+ (R maximal autocorrelation of lag 5 / weighted by atomic Sanderson electronegativities) that are belong to the Getaway descriptors. GETAWAY

descriptors are based on a leverage matrix similar to that determined in statistics and the most commonly applied for regression diagnostics. These molecular descriptors try to match 3D-molecular geometry provided by the molecular influence matrix and the atom relatedness by molecular topology, with chemical information by using various atomic weights (atomic mass, polarizability, van der Waals volume and electronegativity, etc.) [38]. The negative signs of R8m and R5e+ descriptors indicate that the pIC_{50} value is contrariwise related to these descriptors.

As a summary of the above discussion, we can conclude that the mean topological charge index, the atomic polarizabilities, the atomic masses and the atomic Sanderson electronegativities of molecules play an important role in antagonist potency of purine derivatives as P2X7 receptor.

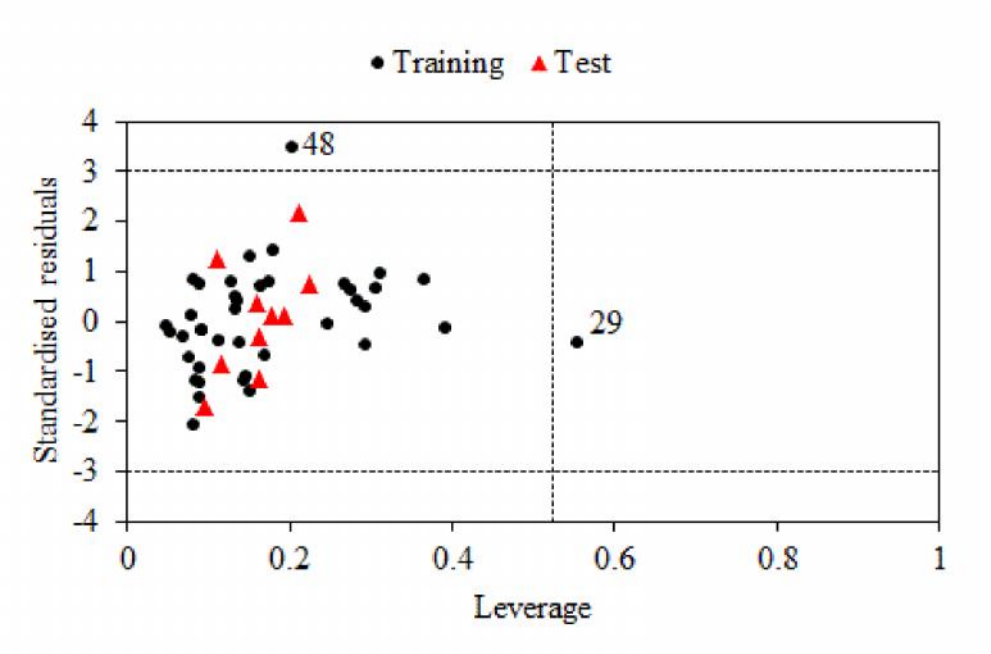


Figure 2. The Williams plot of GA-MLR model for the training and test sets

Conclusion

In this study, multiple linear regressions as a simple and very fast technique were applied to build a quantitative relation between the molecular structures and biological activity of purine derivatives. Stepwise and genetic algorithm, were used as powerful methods to select the best descriptors. The stability, robustness and predictive strength of the created linear models were confirmed utilizing

cross-validation (leave-one-out and leave-group-out), external test set and Y-randomization procedures. Comparison among the obtained data for models demonstrated that GA-MLR is a premier model with acceptable statistical quality and low relative errors. Therefore, the developed GA-MLR model can be useful to calculate the activity values of novel derivatives, and design of new P2x7 receptors.

Acknowledgements

We are grateful to the Payame Noor University for the financial support.

References

[1] M. Tsukimoto, H. Harada, M. Degawa, *Drug. Discov.Today.*, **2007**,*4*, 33-37.

[2] D. Ferrari, C. Pizzirani, E. Adinolfi, R.M. Lemoli, A. Curti, M. Idzko, E. Panther, F. Di Virgilio, *J. Immunol.*, **2006**,*176*, 3877-3883.

[3] C.A. Gabel, *Purinerg. Signal.*, **2007**, *3*, 27-38.

[4] C. Virginio, A. MacKenzie, R.A. North, A. Surprenant, *J. Physiol.*,**1999**, *519*, 335-346.

[5] A. Michel, R. Kaur, I. Chessell, P. Humphrey, *Brit. J. pharmacol.*, **2000**, *130* 513-520.

[6] R.A. North, *Physiol. Rev.*, **2002**, *82*, 1013-1067.

[7] W.H. Loomis, S. Namiki, R.S. Ostrom, P.A. Insel, W.G. Junger, *J. Biol. Chem.*, **2003**, *278*,4590-4596.

[8] N. Mehta, S. Chand, M. Singh Bahia, O. Silakari, *Lett. Drug. Des. Discov.*, **2012**, *9*, 185-198.

[9] T. Suzuki, I. Hide, K. Ido, S. Kohsaka, K. Inoue, Y. Nakata, *J. Neurosci.*, **2004**, *24*,1-7.

[10] F. Di Virgilio, V. Vishwanath, D. Ferrari, On the role of the P2X7 receptor in the immune system, Purinergic and Pyrimidinergic Signalling II, Springer **2001**, pp. 355-374.

[11] F. Di Virgilio, S. Falzoni, C. Mutini, J.M. Sanz, P. Chiozzi, *Drug. Develop. Res.*, **1998**, *45*, 207-213.

[12] R.A. Le Feuvre, D. Brough, O. Touzani, N.J. Rothwell, *J. Cerebr. Blood F. Met.*, **2003**, *23*,381-384.

[13] R. Le Feuvre, D. Brough, N. Rothwell, *Eur. J. Pharmacolo.*, **2002**, *447*, 261-269.

[14] D. Rampe, L. Wang, G.E. Ringheim, *J. neuroimmunol*, **2004**, *147*,56-61.

[15] X. Wang, G. Arcuino, T. Takano, J. Lin, W.G. Peng, P. Wan, P. Li, Q. Xu, Q.S. Liu, S.A. Goldman, *Nature. Med.*, **2004**, *10*, 821-827.

[16] V. Raghavendra, J.A. DeLeo, *Adv. Mol. Cell Biol.*, **2003**, *31*, 951-966.

[17] E.D. Milligan, S.F. Maier, L.R. Watkins, Seminars in Pain Medicine, Elsevier, **2003**, 171-183.

[18] E. Pourbasheer, S. Riahi, M.R. Ganjali, P. Norouzi, *Fuller. Nanotub. Car. N.*, **2011**, *19*, 585-598.

[19] E. Pourbasheer, S. Riahi, M.R. Ganjali, P. Norouzi, *J. Enzyme. Inhib. Med. Chem.*, **2010**, *25*, 844-853.

[20] E. Pourbasheer, A. Beheshti, H. Khajehsharifi, M.R. Ganjali, P. Norouzi, *Med. Chem. Res.*, **2013**, *22*, 4047-4058.

[21] E. Pourbasheer, R. Aalizadeh, M.R. Ganjali, P. Norouzi, J. Shadmanesh, C. Methenitis, *Med. Chem. Res.*, **2014**, *23*, 2264-2276.

[22] E. Pourbasheer, R. Aalizadeh, M.R. Ganjali, P. Norouzi, A. Banaei, *Med. Chem. Res.*, 2014, *23*, 3082-3091.

[23] A. Habibi-Yangjeh, E. Pourbasheer, M. Danandeh-Jenagharad, *Bull. -Korean Chem. Soc.*, **2008**, *29*, 833.

[24] E. Pourbasheer, S. Ahmadpour, R. Zare-Dorabei, M. Nekoei, *Arab. J. Chem.*, **2013**,.

[25] R. Sabet, A. Fassihi, Int. J. Mol. sci., 9, **2008**, 2407-2423.

[26] S. Janardhan, S. Seth, V. Viswanadhan, Mol. Divers., 18, **2014**, 161-181.

[27] M. Ahmadi, M. Shahlaei, *Res. Pharmaceut. Sci.*, 10, **2015**.307-325.

[28] S. Riahi, E. Pourbasheer, R. Dinarvand, M.R. Ganjali, P. Norouzi, *Chem. Biol. Drug. Des.*, **2009**, *74*, 165-172.

[29] J.J. Matasi, S. Brumfield, D. Tulshian, M. Czarnecki, W. Greenlee, C.G. Garlisi, H. Qiu, K. Devito, S.-C. Chen, Y. Sun, R. Bertorelli, W. Geiss, V.-D. Le, G.S. Martin, S.A. Vellekoop, J. Haber, M.L. Allard, *Bioorg. Med. Chem. Lett.*, 2011, *21*, 3805-3808.

[30] S. Brumfield, J.J. Matasi, D. Tulshian, M. Czarniecki, W. Greenlee, C. Garlisi, H. Qiu, K. Devito, S.-C. Chen, Y. Sun, R. Bertorelli, J. Ansell, W. Geiss, V.-D. Le, G.S. Martin, S.A. Vellekoop, J. Haber, M.L. Allard, *Bioorg. Med. Chem. Lett.*, **2011**, *21*, 7287-7290.

[31] E. Pourbasheer, R. Aalizadeh, M.R. Ganjali, P. Norouzi, J. Shadmanesh, *J. Saudi. Chem. Soc.*, **2014**, *18*, 681-688.

[32] E. Pourbasheer, R. Aalizadeh, M.R. Ganjali, P. Norouzi, *Med. Chem. Res.*, **2014**, *23*, 57-66.

[33] T. MathWorks, Genetic algorithm and direct search toolbox user's guide, Math. Inc. USA, **2005**.

[34] R. Wehrens, H. Putter, L. Buydens, *Chemom. Intell. Lab. Sys.*, **2000**, *54*, 35-52.

[35] E. Pourbasheer, R. Aalizadeh, M.R. Ganjali, P. Norouzi, *Struct. Chem.*, **2014**, *25*, 355-370.

[36] R. Bazl, M.R. Ganjali, H. Derakhshankhah, A.A. Saboury, M. Amanlou, P. Norouzi, *Med. Chem. Res.*, **2013**, *22*, 5453-5465.

[37] M. Nekoei, M. Mohammadhosseini, A.A. Gharahbagh, *Anal. Bioanal. Electro*, **2009**, *1*, 159-168.

[38] B. Tiperciuc, V. Zaharia, R. Campean, M. Curticapean, A. Costescu, M. Diudea, *MATCH-Commun. Math. Co.*, **2008**, *60*, 985-996.