

## In Silico Screening of IL-1 $\beta$ Production Inhibitors Using Chemometric Tools

Amirhossein Sakhteman<sup>a</sup>, Najmeh Edraki<sup>b</sup>, Bahram Hemmateenejad<sup>c</sup>, Ramin Miri<sup>b</sup>, Alireza Foroumadi<sup>d</sup>, Abbas Shafiee<sup>d,e</sup> and Mehdi Khoshneviszadeh<sup>a,b\*</sup>

<sup>a</sup>Department of Medicinal Chemistry, Faculty of Pharmacy, Shiraz University of Medical Sciences, Shiraz, Iran. <sup>b</sup>Medicinal and Natural Products Chemistry Research Center, Shiraz University of Medical Sciences, Shiraz, Iran. <sup>c</sup>Department of Chemistry, Shiraz University, Shiraz, 71454, Iran. <sup>d</sup>Pharmaceutical Sciences Research Center, Tehran University of Medical Sciences, Tehran, Iran. <sup>e</sup>Department of Medicinal Chemistry, Faculty of Pharmacy, Tehran University of Medical Sciences, Tehran, Iran.

### Abstract

The IL-1 $\beta$  plays a major role in inflammatory disorders and IL-1 $\beta$  production inhibitors can be used in the treatment of inflammatory and related diseases. In this study, quantitative relationships between the structures of 46 pyridazine derivatives (inhibitors of IL-1 $\beta$  production) and their activities were investigated by Multiple Linear Regression (MLR) technique Stepwise Regression Method (ES-SWR). The genetic algorithm (GA) has been proposed for improvement of the performance of the MLR modeling by choosing the most relevant descriptors. The results show that eight descriptors are able to describe about 83.70% of the variance in the experimental activity of the molecules in the training set. The physical meaning of the selected descriptors is discussed in detail. Power predictions of the QSAR models developed were evaluated using cross-validation, and validation through an external prediction set. The results showed satisfactory goodness-of-fit, robustness and perfect external predictive performance. The applicability domain was used to define the area of reliable predictions. Furthermore, the *in silico* screening technique was applied in order to predict the structure and potency of new compounds of this type using the proposed QSAR model.

**Keywords:** QSAR; IL-1 $\beta$ ; Pyridazine; *in silico* screening; MLR.

### Introduction

In recent years, many efforts have been done toward the development of new therapeutic agents in the area of inflammatory diseases. Cytokines are multifunctional proteins that responsible for host defense mechanisms such as inflammatory, immune and hematogenic responses (1). Cytokines have been categorized

as being proinflammatory (IL-1 $\beta$ , TNF- $\alpha$ , IL-6...) or anti-inflammatory (TGF- $\beta$ , IL-10, IL-13...) depending on their effects on the immune system. Inflammatory cytokines play an important role in inflammatory disease. Consequently, inhibition the production of IL-1 $\beta$  can reduce levels of these proinflammatory cytokine, and thereby reduce inflammation and prevent destruction effects in diseases such as rheumatoid arthritis (RA), osteoarthritis (OA) and Crohn's disease (2). For the rational design of novel IL-1 $\beta$  production inhibitors, quantitative structure-

\* Corresponding author:

E-mail: m.khoshneviszadeh@gmail.com

activity relationships (QSAR) models and *in silico* screening could be useful (3). QSAR is a powerful tool to study the relationship between activity and structural parameters and could be used for design new biological active compounds and predict their potency, toxicity, ADME, etc. In this study, a series of 5, 6-bis (4-methoxyphenyl)-pyridazin derivatives with IL-1 $\beta$  production inhibitory activity discovered by Matsuda *et al* (4, 5). According to our literatures review, no QSAR study was carried out on the mentioned series. In the present investigation, a quantitative structure–activity relationship was explored using different molecular descriptors. Previous studies demonstrated that linear QSAR models had good predictive ability therefore (6-9), a virtual screening study was then carried out to identify novel biologically active patterns by modification of the original molecules. The study led to the identification of novel structures, which are potent IL-1 $\beta$  production inhibitors based on the QSAR model. The structures were filtered using the domain of applicability of the QSAR model.

#### *Materials and methods*

##### *Data set*

The biological data used in this study are the IL-1 $\beta$  production inhibitory (IC<sub>50</sub>) activity of a series of 5,6-bis (4-methoxyphenyl)-pyridazin derivatives (4, 5). The structural features and biological activity of these compounds are listed in Table 1. The biological data were converted to logarithmic scale (pIC<sub>50</sub>) and then used for subsequent QSAR analysis as dependent variables. The original data set was divided into a training set (n = 35) and a prediction set (n = 11) using the popular Kennard and Stones algorithm (10, 11). This algorithm has been widespread used with great success in many QSAR studies (12-14).

##### *Descriptor generation*

The chemical structure of molecules was constructed using Hyperchem package (Version 7, Hypercube Inc., Florida 32601 USA). The Z-matrices of the structures were provided by the software and transferred to the Gaussian 98 program. Complete geometry optimization was performed taking the most extended

conformations as starting geometries. Semi empirical molecular orbital calculations (AM1) of the structures were performed using Gaussian 98 program. A large numbers of molecular descriptors were calculated using HyperChem, Gaussian 98 and Dragon Packages. Gaussian 98 was employed for calculation of different quantum chemical descriptors including dipole moment (DM), local charges, HOMO and LUMO energies, hardness (g); softness (S); electronegativity (v); and electrophilicity (x). Dragon software calculated different functional groups, topological, geometrical and constitutional descriptors for each molecule, and some chemical parameters including molecular volume (V), molecular surface area (SA), hydrophobicity (log P), hydration energy (HE) were calculated using Hyperchem software.

##### *Data processing and modeling*

The calculated descriptors were collected in a data matrix, D. First the descriptors were checked for constant or near-constant values and those detected were removed from the original data matrix. Then, the correlation of descriptors with each other's and with the activity data was determined. Among the collinear descriptors detected ( $r > 0.9$ ), one of them that had the highest correlation with activity was retained, and the rest were omitted (15, 16). For the development of QSAR equation stepwise multiple linear regression (MLR) was used. In stepwise regression, a multiple-term linear equation was built step-by-step. The basic procedures involve (i) identifying an initial model, (ii) iteratively <stepping>, that is repeatedly altering the model at the previous step by adding or removing a predictor variable in accordance with the <stepping criteria> (in our case, probability of F = 0.05 for inclusion; probability of F = 0.1 for exclusion for the forward selection method), and (iii) terminating the search when stepping is no longer possible given the stepping criteria, or when a specified maximum number of steps have been reached. Specifically, at each step, all variables are reviewed and evaluated to determine which one will contribute most to the equation. That variable will then be included in the model, and the process starts again. A limitation of the stepwise regression search approach is that

it presumes there is a single 'best' subset of X variables and seeks to identify it. There is often no unique 'best' subset, and all possible regression models with a similar number of X variables as in the stepwise regression solution should be fitted subsequently to study whether some other subsets of X variables might be better (17, 18). We used GA for variable selection in MLR regression. The applied GA used a binary representation as the coding technique for the given problem: The presence or absence of a descriptor in a chromosome was coded by 1 or 0. The GA performed its optimization by variation and selection via the evaluation of the fitness function. The operators used here were crossover and mutation. The probability for the application of these operators was varied linearly with the generation renewal ( $0 \pm 1\%$  for mutation and  $60 \pm 90\%$  for crossover). The population size was 100. For a typical run, the evolution of the generation was stopped when 90% of the generations took the same fitness. The fitness function was the root mean square error of cross-validation (19).

#### Model validation

The goodness-of-fit of the resulted QSAR models were judged using statistical parameters such as correlation coefficient ( $R^2$ ), standard error of regression (SE), and variance ratio (F) at specified degrees of freedom. The generated QSAR equations were also validated by leave-one-out cross-validation correlation coefficient ( $Q^2$ ) and root mean square error of cross-validation (RMSECV)(20). In addition, an external test set composed selected 11 molecules was used to judge the overall prediction ability of the resulted models (see Table 1). According to Tropsha *et al.* (21). The predictive ability of a QSAR model should be tested on an external set of data that has not been taken into account during the process of developing the model. In particular, to assess the predictive power of QSAR models the correlation coefficient between the predicted and observed activities of compounds from an external test ( $R^2$ ), the correlation coefficients for regressions through the origin (predicted versus observed activities, or observed versus predicted activities, i.e.,  $R_0^2$  or  $R_0'^2$ , respectively), and the slope of the

regression lines through the origin ( $K$  and  $K'$ , respectively) were calculated. Tropsha *et al.* (21, 22). considered a QSAR model to be predictive, if all of the following conditions are satisfied: (i)  $Q^2 > 0.5$ , (ii)  $R^2 > 0.6$ , (iii)  $R_0^2$  or  $R_0'^2$  is close to  $R^2$ , such that  $[(R^2 - R_0^2)/r^2]$  or  $[(R^2 - R_0'^2)/R^2] < 0.1$  and (iv)  $0.85 \leq K \leq 1.15$  or  $0.85 \leq K' \leq 1.15$ . In addition, according to the recommendation of Roy and Roy [18], an additional statistic for external validation ( $rm^2$ ) was calculated as  $rm^2 = R^2 * [1 - (R^2 - R_0^2)^{1/2}]$ . For a model with good external predictability,  $rm^2$  value should be  $> 0.5$ . (23)

#### Y-randomization test

This technique ensures the robustness of a QSAR model. The dependent variable vector  $pIC_{50}$  is randomly shuffled and a new QSAR model is developed using the original independent variable matrix. The new QSAR models (after several repetitions) are expected to have low  $R^2$  and  $Q^2$  values. If the opposite happens then an acceptable QSAR model cannot be obtained for the specific modeling method and data (12).

#### Applicability domain

Finally the QSAR model was used to identify novel active compounds via an *in silico* screening procedure, and thus the definition of its domain of applicability is of particular importance. The utility of a QSAR model is based on its accurate prediction ability for new compounds. A model is valid only within its training domain, and new compounds must be assessed as belonging to the domain before the model is applied. The applicability domain is assessed by the leverage values for each compound. The threshold  $h^*$  is generally fixed at  $3(k + 1)/n$  ( $n$  is the number of training set compounds, and  $k$  is the number of model parameters), On the other hand, when the leverage value of a compound is lower than the threshold value, the probability of accordance between predicted and observed values is as high as that for the training set compounds (24).

#### In silico screening

Finally *in silico* screening procedure was carried out to identify a variety of potential novel lead compounds by presenting structural

modifications on the original dataset. Throughout the screening procedure, only the predictions that fall into the domain of applicability were considered reliable.

## Results and Discussion

On the basis of Kennard-Stones algorithm, the dataset of the 46 pyridazine derivatives was divided into a training set (35 compounds) and a prediction or test set (11 compounds, see Table 1). Stepwise regression was used on the training data set to develop MLR QSAR model.

$$\text{pIC}_{50} = -102.168 (\pm 15.264) + 119.255 (\pm 17.542) \text{MATS4m} + 0.106 (\pm 0.026) \text{RDF105u} - 0.168 (\pm 0.024) \text{RDF100u} - 5.458 (\pm 1.131) \text{GATS3v} + 0.269 (\pm 0.053) \text{RDF075v} + 1.222 (\pm 0.293) \text{C-005} + 0.073 (\pm 0.022) \text{RDF095u} - 0.006 (\pm 0.002) \text{Surface area}$$

$$N = 35 \quad R^2 = 0.837 \quad R^2_{\text{adj}} = 0.780 \quad F = 16.079 \\ \text{S.E.} = 0.330 \quad Q^2 = 0.733 \\ \text{RMScv} = 0.374 \quad R^2_{\text{pred}} = 0.754$$

The regression coefficients in this model are significant at the 95% level. This equation can predict and explain 75.4% and 83.7%, respectively, of the variance of the inhibitory activity.

The possibility of having included outliers in our dataset was investigated by calculating the standard residuals. Standardized residuals greater than 2.5 or less than -2.5 are considered large and indicate the exclusion of the respective data from the data set. The calculated standardized residuals were within the above upper and lower limits for all the compounds, and thus, none of them were excluded from the data set as outlier (3). The Figure1 showed that the good model was obtained with eight descriptors.

The above equation showed that the most significant descriptors are Moran autocorrelation of lag 4 weighted by mass (MATS4m), Radial Distribution Function - 105 / unweighted (RDF105u), Radial Distribution Function - 100 / unweighted (RDF100u), Geary autocorrelation of lag 3 weighted by van der Waals volume (GATS3v), Radial Distribution Function - 075 / weighted by van der Waals volume (RDF075v),

CH3X (C-005), Radial Distribution Function - 095 / unweighted (RDF095u) and surface area. The correlation matrix (Table 2) indicated that the eight selected descriptors are not highly correlated. Variance Inflation Factor (VIF) values for the eight descriptors (25-27), also shown in Table 2, demonstrate that the model contains no multicollinearity.

### *Interpretation of the Selected Descriptors*

The variety of factors such as molecular electrostatic potential, polarizability, hydrophobicity, and lipophobicity influence the binding of ligand to its target. MATS4m and GATS3v are Autocorrelation of Topological Structure. The 2D-autocorrelation descriptors explain how the values of certain functions, at intervals equal to the *lag*, are correlated. The 2D autocorrelation descriptors represent the topological structure of the compounds, but are more complex in nature when compared to the classical topological descriptors. The computation of these descriptors involves the summations of different autocorrelation functions corresponding to different structural lags and leads to different autocorrelation vectors corresponding to the lengths of the sub-structural fragments. Basically, the pool of 2D autocorrelation descriptors defines a wide 2D space. On behalf of a greater applicability, physicochemical properties (atomic masses, atomic van der Waals volumes, atomic Sanderson electronegativities, and atomic polarizabilities) were inserted as weighting components. As a result, these descriptors address the topology of the structure or parts thereof in association with a specific physicochemical property. Bearing in mind this aspect, the interpretation of 2D autocorrelation descriptors was uneasy. Based on derived model the positive regression coefficient of MATS4m shows that small path lengths and branching in the molecule (lag 4 weighted by atomic mass) contribute to higher activity (28). RDF105u, RDF100u, RDF075v and RDF095u are Radial Distribution Function descriptors; the 3D coordinates of the atoms of molecules can be transformed into a structure code that has a fixed number of descriptors irrespective of the size of a molecule. This task is performed by a structure coding technique referred to as

Radial Distribution Function code (RDF code). In general, there are some prerequisites for a structure code:

- Independent from the number of atoms, i.e., the size of a molecule,
- Unambiguity regarding the three-dimensional arrangement of the atoms, and
- Invariance against translation and rotation of the entire molecule.

Formally, the Radial Distribution Function of an ensemble of N atoms can be interpreted as the probability distribution to find an atom in a spherical volume of radius r. The equation represents the Radial Distribution Function code as it is used in this investigation:

$$g(r) = f \times \sum_i^{N-1} \sum_j^N A_i A_j e^{-B(r-r_{ij})^2}$$

Where f is a scaling factor and N is the number of atoms. By including characteristic atomic properties A of the atoms i and j, the RDF codes can be used in different tasks to fit the requirements of the information to be represented.

The exponential term contains the distance  $r_{ij}$  between the atoms i and j and the smoothing parameter B, which defines the probability distribution of the individual distances.  $g(r)$  was calculated at a number of discrete points with defined intervals. The radial distribution function in this form meets all the requirements for 3D structure descriptors: it is independent of the number of atoms, i. e., the size of a molecule, it is unique regarding the three-dimensional arrangement of the atoms, and it is invariant against translation and rotation of the entire molecule. Additionally, the RDF descriptors can be restricted to specific atom types or distance ranges to represent specific information in a certain three-dimensional structure space, e.g. to describe steric hindrance or structure/activity properties of a molecule. Finally, the RDF descriptors are interpretable by using simple rules sets, and thus it provides a possibility for conversion of the code back into the corresponding 3D structure. Besides information about interatomic distances in the entire molecule, the RDF descriptors provide

further valuable information, e.g. about bond distances, ring types, planar and non-planar systems and atom types. By using different weighting schemes, which include atom types, electronegativity, atom mass or van der Waals radii, RDF can be adjusted to select among those atoms of molecule, which give rise to an important descriptor in deriving an appropriate QSAR (29-31).

Final descriptor C005 is one of the Ghos-Crippen atom-centred fragments related to the methyl group attached to any electronegative atom (O, N, S, P, Se, halogens) fragment. It gives information about the number of predefined structural features in the molecule. Based on the produced QSAR equation a high value of MATS4m, RDF105u, RDF075v, C-005 and RDF095u give a positive contribution to the IL-1 production inhibition. On the other hand, a high value of GATS 3v and Surface area give a negative contribution to the inhibition.

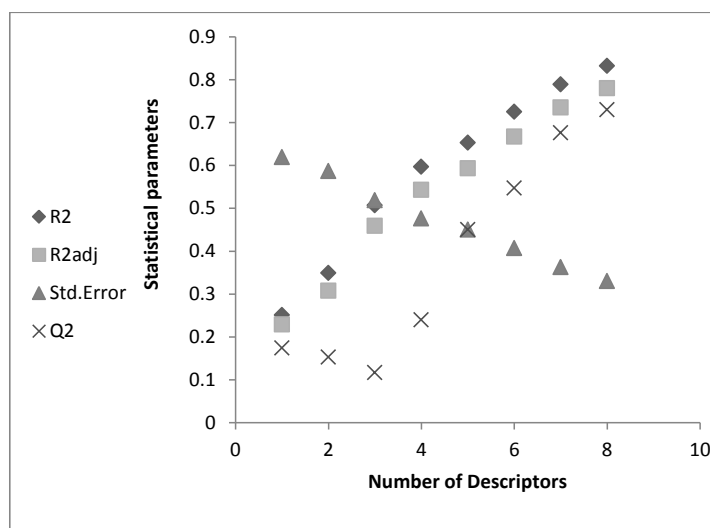
#### Model Validation

The IL-1 $\beta$  production inhibition predictability of the proposed model was evaluated by using the external set of 11 compounds (Table 1). The proposed QSAR model has all conditions to be considered as predictive models.

$$\begin{aligned} R^2_{\text{pred}} &= 0.754 > 0.6 \\ [(R^2_{\text{pred}} - R_0^2) / R^2_{\text{pred}}] &= -0.078 < 0.1 \\ [(R^2_{\text{pred}} - R_0^2) / r^2_{\text{pred}}] &= -0.061 < 0.1 \\ \text{rm}^2 &= 0.595 > 0.5 \\ K &= 0.83, K1.00 = \end{aligned}$$

This model was further validated by applying the Y-randomization test. Several random shuffles of the Y vector were performed. The low  $R^2$  ( $0.0 < R^2 < 0.34$ ) and  $Q^2$  ( $0.0 < Q^2 < 0.25$ ) values indicate that the good results in our original model are not due to a chance correlation or structural dependency of the training set. The extrapolation method was applied to the compounds that constitute the test set. The results are presented in Table 1. None of the 11 compounds fell outside from the domain of the model (warning leverage limit = 0.77).

The suggested method, according to the high predictive ability, could be a useful tool to the costly and time consuming experiments for



**Figure 1.** The statistical parameters according to number of descriptors entered in model.

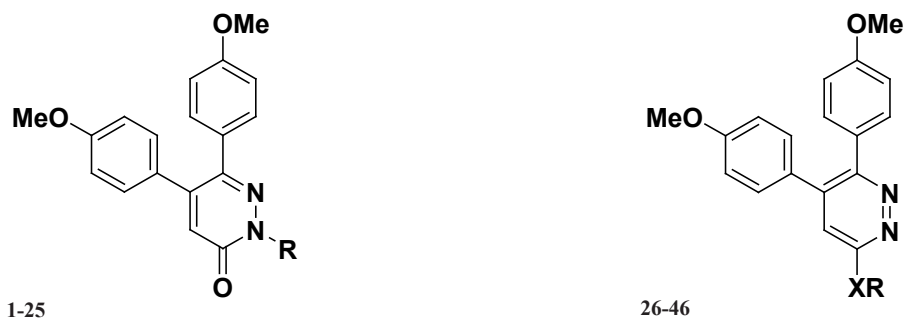
determining the IL-1  $\beta$  production inhibitory activity of pyridazine derivatives. The method can also be used to the screen virtual compounds in order to identify derivatives with desired activity.

#### *In silico screening*

The *in silico* screening was applied to the design of new structures with potential IL-1  $\beta$  production inhibitors according to the developed QSAR model. The role of the *in silico* screen was as a guide to the identification of the most promising new synthetic targets. For this purpose, first we selected potent inhibitors from data set (3, 4, 17, 22, 23, 24, 28, 31, 32, 33, 34, 35, 36, 37) and retain the main scaffold (4-methoxy phenyl moiety) and change the heterocyclic core, aromatic moiety and connection between aromatic moiety and central heterocyclic core. The modifications incorporated in the virtual screening study were chosen based on their synthetic feasibility. The suggested structures were not involving the use of unusual ring fragments or functional groups that cannot be prepared using established protocols. The chemistry of 1,2,4-triazine was well understood and introducing practical modifications here was considered synthetically viable, the *in silico* screen began with the replacement of

the pyridazine core by 1,2,4-triazine (Table 3). The model tolerated the introduction of 1,2,4-triazine since all those studied were within the domain of applicability. The compounds 49 and 51 showed the best activity, 6.83 and 6.56, respectively. In some cases the activities lower and just one compound has higher activity than original compound. The next step was incorporated 1,2,4-triazine instead of pyridazine core in second series of compounds that the different substituted aromatic moiety connect to central heterocyclic by O or S atoms. This model tolerated and the activity was retained (Table 4.).

The other modification was replacement connective atoms (S or O) to N, this modification gave structures that showed good activity and were within the models domain of applicability (Table 5) but introducing the second nitrogen to linker slightly lead to decrease of predicted activity and was clearly within the domain of applicability (Table 6). Interestingly, replacement of pyridazine to pyridine ring led to enhanced biological activity and comfortably within the domain of applicability (Table 7). The *In silico* study suggested that presence of heterocyclic ring containing nitrogen group was necessary for inhibition IL-1  $\beta$  production and increasing number of nitrogen ring diminished



**Table 1.** IL-1  $\beta$  production inhibitory activity—observed and predicted using developed model.

Compound	R or XR	pIC <sub>50</sub> observed	pIC <sub>50</sub> predicted	Leverage (limit= 0.77)
1	H	4.50	4.72	
2	Bn	6.58	6.71*	0.25
3	4-FBn	6.82	6.80*	0.16
4	2,4-F <sub>2</sub> Bn	6.74	6.19	
5	3,4-Cl <sub>2</sub> Bn	6.67	6.46	
6	4-ClBn	6.27	5.49*	0.13
7	3,4-F <sub>2</sub> Bn	6.26	6.29	
8	2,4-Cl <sub>2</sub> Bn	5.96	5.84	
9	2,6-Cl <sub>2</sub> Bn	5.5	5.04*	0.23
10	2,4,6-Cl <sub>3</sub> Bn	5.02	5.35	
11	Me	5.94	5.87	
12	Et	6.21	5.70	
13	iPr	5.88	5.93	
14	cPr	5.09	5.44	
15	cPrCH <sub>2</sub>	5.74	5.95	
16	cPn	6.29	7.04*	0.28
17	cPnCH <sub>2</sub>	6.95	7.08	
18	3-PyCH <sub>2</sub>	6.36	6.44*	0.20
19	4-PyCH <sub>2</sub>	6.28	6.21	
20	2-PyCH <sub>2</sub>	5.83	6.54	
21	Cinnamyl	6.21	6.59	
22	2,4-F <sub>2</sub> Cinnamyl	7.00	6.70	
23	4-ClCinnamyl	7.00	7.05	
24	4-FCinnamyl	6.77	6.87*	0.11
25	2,4-ClCinnamyl	6.56	6.54	
26	2,4-F <sub>2</sub> PhNH	5.71	5.96	
27	2,4-F <sub>2</sub> PhNPr	5.50	5.37	
28	3,4,5-Cl <sub>3</sub> PhS	6.99	7.09	
29	2,4-F <sub>2</sub> PhS	5.73	4.84*	0.247

**Table 1.** (Continue).

30	PhS	4.69	4.53	
31	2,3,4,5,6-F <sub>5</sub> PhO	7.39	7.07	
32	PhO	7	6.86	
33	2,3-F <sub>2</sub> PhO	6.92	7.04	
34	2-CNPhO	6.82	6.40	
35	3-CNPhO	6.79	6.94	
36	2,5-F <sub>2</sub> PhO	6.74	6.16	
37	2,3,5,6-F <sub>4</sub> PhO	6.72	6.77	
38	2,6-F <sub>2</sub> PhO	6.45	6.20	
39	4-MeOPhO	6.35	6.41	
40	3,4-F <sub>2</sub> PhO	6.33	6.49*	0.07
41	4-NO <sub>2</sub> PhO	6.20	6.43	
42	4-PyO	5.90	5.62	
43	3,5-F <sub>2</sub> PhO	5.86	5.74*	0.11
44	2,4-F <sub>2</sub> PhO	5.63	5.85	
45	3-NO <sub>2</sub> PhO	5.92	5.27*	0.31
46	2,4-Cl <sub>2</sub> PhO	5.59	5.91	

\*Selected as test set.

biological activity.

### Conclusion

In this study we have identified eight descriptors that successfully model the IL-1  $\beta$  production inhibitory activity. The validation

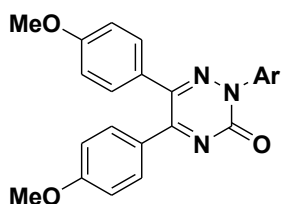
procedures utilized in this work (separation of data into independent training and validation sets, Y-randomization) illustrated the accuracy of produced QSAR model. Based on the developed QSAR model, we have designed novel structures that could be further investigated as novel effective IL-1 beta

**Table 2.** Correlation matrix for the eight selected descriptors.

	MATS4m	GATS3v	RDF095u	RDF100u	RDF105u	RDF075v	C-005	Surface area	VIF*
MATS4m	1	0.598	0.185	0.244	0.129	0.307	-0.273	0.480	3.785
GATS3v		1	0.064	0.143	0.242	0.379	-0.005	0.214	3.143
RDF095u			1	0.612	0.511	0.235	-0.154	0.548	2.830
RDF100u				1	0.608	0.555	-0.209	0.614	2.719
RDF105u					1	0.639	-0.099	0.585	2.744
RDF075v						1	-0.305	0.485	1.478
C-005							1	-0.132	2.389
Surface area								1	3.563

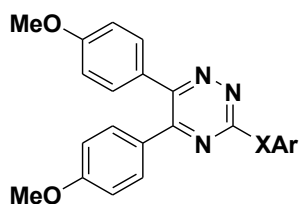
\*VIF less than 10 demonstrates that the model contains no multicollinearity.





**Table 3.** Predicted activities of selected 1,2,4-triazin-3(2H)-one derivatives.

ID	Ar	pIC <sub>50</sub> predicted	Leverage
47	4-FBn	6.42	0.182
48	2,4-F <sub>2</sub> Bn	5.65	0.167
49	cPnCH <sub>2</sub>	6.83	0.206
50	2,4-F <sub>2</sub> Cinnamyl	6.51	0.061
51	4-ClCinnamyl	6.56	0.142
52	4-FCinnamyl	6.41	0.103

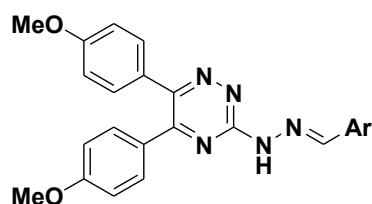


**Table 4.** Predicted activities of selected 3-aryl 1,2,4-triazin derivatives.

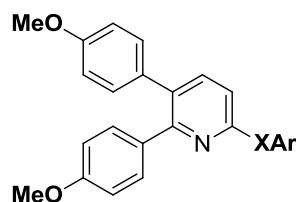
ID	XAr	pIC <sub>50</sub> predicted	Leverage
53	3,4,5-Cl <sub>3</sub> PhS	6.96	0.307
54	2,3,4,5,6-F <sub>5</sub> PhO	6.80	0.150
55	PhO	6.93	0.182
56	2,3-F <sub>2</sub> PhO	6.75	0.055
57	2-CNPhO	6.10	0.037
58	3-CNPhO	6.71	0.100
59	2,5-F <sub>2</sub> PhO	5.86	0.094
60	2,3,5,6-F <sub>4</sub> PhO	6.48	0.107

**Table 5.** Predicted activities of selected 3-amino 1,2,4-triazin derivatives.

ID	XAr	pIC <sub>50</sub> predicted	Leverage
61	3,4,5-Cl <sub>3</sub> PhNH	7.51	0.350
62	2,3,4,5,6-F <sub>5</sub> PhNH	6.67	0.120
63	PhNH	6.89	0.182
64	2,3-F <sub>2</sub> PhNH	6.72	0.055
65	2-CNPhNH	6.29	0.034
66	3-CNPhNH	6.90	0.101
67	2,5-F <sub>2</sub> PhNH	5.84	0.099
68	2,3,5,6-F <sub>4</sub> PhNH	6.36	0.103

**Table 6.** Predicted activities of selected 3-benzylidenehydrazine 1,2,4-triazin derivatives.

ID	XAr	pIC <sub>50</sub> predicted	Leverage
69	3,4,5-Cl <sub>3</sub> Ph	7.24	0.384
70	2,3,4,5,6-F <sub>5</sub> Ph	6.06	0.212
71	Ph	6.32	0.146
72	2,3-F <sub>2</sub> Ph	5.44	0.242
73	2-CNPh	6.08	0.135
74	3-CNPh	6.09	0.194
75	2,5-F <sub>2</sub> Ph	5.66	0.197
76	2,3,5,6-F <sub>4</sub> Ph	5.18	0.442

**Table 7.** Predicted activities of selected three aryl pyridine derivatives.

ID	XAr	pIC <sub>50</sub> predicted	Leverage
77	3,4,5-Cl <sub>3</sub> PhS	7.54	0.472
78	2,3,4,5,6-F <sub>5</sub> PhO	7.32	0.293
79	PhO	7.19	0.194
80	2,3-F <sub>2</sub> PhO	7.26	0.143
81	2-CNPhO	6.59	0.040
82	3-CNPhO	7.18	0.135
83	2,5-F <sub>2</sub> PhO	6.39	0.105
84	2,3,5,6-F <sub>4</sub> PhO	7.01	0.178

production inhibitors. The proposed method, due to the high predictive ability, tends a useful alternative to the costly and time consuming experiments for cytokine production inhibitory activity.

### Acknowledgment

The authors also acknowledge the financial support of Shiraz and Tehran University of Medical Sciences, Vice Chancellor of

Research. This study was part of the PhD thesis project of Mehdi Khoshneviszadeh.

## References

- (1) Enomoto H, Sawa A, Suhara H, Yamamoto N, Inoue H, Setoguchi C, Tsuji F, Okamoto M, Sasabuchi Y, Horiuchi M and Ban M. Synthesis and pharmacological evaluation of 1,1,3-substituted urea derivatives as potent TNF-alpha production inhibitors. *Bioorg. Med. Chem. Lett.* (2010) 20: 4479-82.
- (2) Tullis JS, Vanrens JC, Natchus MG, Clark MP, De B, Hsieh LC and Janusz MJ. The development of new triazole based inhibitors of tumor necrosis factor-alpha (TNF-alpha) production. *Bioorg. Med. Chem. Lett.* (2003) 13: 1665-8.
- (3) Melagraki G, Afantitis A, Sarimveis H, Koutentis PA, Kollias G and Igglessi-Markopoulou O. Predictive QSAR workflow for the in silico identification and screening of novel HDAC inhibitors. *Mol. Divers.* (2009) 13: 301-11.
- (4) Matsuda T, Aoki T, Ohgiya T, Koshi T, Ohkuchi M and Shigyo H. Synthesis and bioactivities of novel pyridazine derivatives: inhibitors of interleukin-1 beta (IL-1beta) production. *Bioorg. Med. Chem. Lett.* (2001) 11: 2369-72.
- (5) Matsuda T, Aoki T, Koshi T, Ohkuchi M and Shigyo H. Synthesis and bioactivities of novel 5,6-bis(4-methoxyphenyl)-2H-pyridazin-3-one derivatives: inhibitors of interleukin-1 beta (IL-1beta) production. *Bioorg. Med. Chem. Lett.* (2001) 11: 2373-5.
- (6) Iman MDA. QSAR and QSTR study of pyrimidine derivatives to improve their therapeutic index as antileishmanial agents. *Med. Chem. Res.* (2013) 22: 5029-35.
- (7) Pereira F, Latino DA and Gaudencio SP. QSAR-assisted virtual screening of lead-like molecules from marine and microbial natural sources for antitumor and antibiotic drug discovery. *Molecules* (2015) 20: 4848-73.
- (8) Huang SH, Tung CW, Fulop F and Li JH. Developing a QSAR model for hepatotoxicity screening of the active compounds in traditional Chinese medicines. *Food Chem. Toxicol.* (2015) 78: 71-7.
- (9) Georgey H. Synthesis, cytotoxic activity and 2D-QSAR study of some imidazoquinazoline derivatives. *Molecules* (2014) 19: 3777-92.
- (10) Rodionova OY and Pomerantsev AL. Subset selection strategy. *J. Chemometrics* (2008) 22: 674-85.
- (11) Daszykowski M, Walczak B and Massart D. Representative subset selection. *Anal. Chim. Acta* (2002) 468: 91-103.
- (12) Afantitis A, Melagraki G, Sarimveis H, Koutentis PA, Markopoulos J and Igglessi-Markopoulou O. A novel QSAR model for predicting induction of apoptosis by 4-aryl-4H-chromenes. *Bioorg. Med. Chem.* (2006) 14: 6686-94.
- (13) Shahlaei M and Pourhossein A. A 2D image-based method for modeling some c-Src tyrosine kinase inhibitors. *Med. Chem. Res.* (2013) 22: 1-14.
- (14) Xu J, Zhu L, Fang D, Wang L, Xiao S, Liu L and Xu W. QSPR studies of impact sensitivity of nitro energetic compounds using three-dimensional descriptors. *J. Mol. Graph. Model* (2012) 36: 10-9.
- (15) Khoshneviszadeh M, Edraki N, Miri R, Foroumadi A and Hemmateenejad B. QSAR study of 4-aryl-4H-chromenes as a new series of apoptosis inducers using different chemometric tools. *Chem. Biol. Drug Des.* (2012) 79: 442-58.
- (16) Khoshneviszadeh M, Edraki N, Miri R and Hemmateenejad B. Exploring QSAR for substituted 2-sulfonyl-phenyl-indol derivatives as potent and selective COX-2 inhibitors using different chemometrics tools. *Chem. Biol. Drug Des.* (2008) 72: 564-74.
- (17) Leonard JT and Roy K. QSAR by LFER model of HIV protease inhibitor mannitol derivatives using FA-MLR, PCRA, and PLS techniques. *Bioorg. Med. Chem.* (2006) 14: 1039-46.
- (18) Darlington RB. *Regression and linear models* McGraw-Hill New York (1990).
- (19) Cho SJ and Hermsmeier MA. Genetic algorithm guided selection: variable selection and subset selection. *J. Chem. Inf. Comput. Sci.* (2002) 42: 927-36.
- (20) Kumar H, Siddiqui AA and Gupta SP. A Quantitative structure-activity relationship study on some imidazoles and 2-aminopyridines acting as nitric oxide synthase (NOS) inhibitors. *Lett. Drug Des. Discov.* (2013) 10: 666-74.
- (21) Tropsha A, Gramatica P and Gombar VK. The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. *QSAR Comb. Sci.* (2003) 22: 69-77.
- (22) Golbraikh A and Tropsha A. Beware of q<sup>2</sup>! *J. Mol. Graph. Model* (2002) 20: 269-76.
- (23) Roy PP and Roy K. On some aspects of variable selection for partial least squares regression models. *QSAR Comb. Sci.* (2008) 27: 302-13.
- (24) Gramatica P. Principles of QSAR models validation: internal and external. *QSAR Comb. Sci.* (2007) 26: 694-701.
- (25) Hari Narayana Moorthy NS, Ramos MJ and Fernandes PA. QSAR analysis of isosteviol derivatives as glucosidase inhibitors with element count and other descriptors. *Lett. Drug Des. Discov.* (2011) 8: 14-25.
- (26) Soni LK, Gupta AK and Kaskhedikar SG. Exploration of QSAR modelling techniques and their combination to rationalize the physicochemical characters of nitrophenyl derivatives towards aldose reductase inhibition. *J. Enzyme Inhib. Med. Chem.* (2009) 24: 1002-7.
- (27) Holder AJ, Ye L, Yourtee DM, Agarwal A, Eick JD and Chappelou CC. An application of the QM-QSAR method to predict and rationalize lipophilicity of simple monomers. *Dental Materials* (2005) 21: 591-8.
- (28) Asadollahi T, Dadfarnia S, Shabani AMH, Ghasemi JB and Sarkhosh M. QSAR models for CXCR2 receptor

- antagonists based on the genetic algorithm for data preprocessing prior to application of the PLS linear regression method and design of the new compounds using *in silico* virtual screening. *Molecules* (2011) 16: 1928-55.
- (29) Gonzalez MP, Teran C, Teijeira M and Helguera AM. Radial distribution function descriptors: an alternative for predicting A2 A adenosine receptors agonists. *Eur. J. Med. Chem.* (2006) 41: 56-62.
- (30) Gonzalez MP, Gandara Z, Fall Y and Gomez G. Radial Distribution Function descriptors for predicting affinity for vitamin D receptor. *Eur. J. Med. Chem.* (2008) 43: 1360-5.
- (31) Fedorowicz A, Zheng L, Singh H, and Demchuk E. QSAR study of skin sensitization using local lymph node assay data. *Int. J. Mol. Sci.* (2004) 5: 56-66.

---

This article is available online at <http://www.ijpr.ir>

---