International Pharmacy Acta

Original article

Journal Homepage: http://intpharmacta.com

DOI: http://dx.doi.org/10.22037/ipa.v1i2.21052

ISSN: 2645-3266

QSAR Modeling of COX-2 Inhibitory Activity of Thiazinan, Benzthiazinan, and Benzdiazinan Derivatives

Sara Akbarpour Avini[a], Tannaz Zebardast[b], Zahra Hajimahdi[c*] and Afshin Zarghi [c*]

Received: 11/10/2017
Accepted: 05/02/2018
Received and accepted date of paper)

Authors' Affiliations:
[a] Active Pharmaceutical Ingredients Research Center, Pharmaceutical Sciences Branch, Islamic Azad University, Tehran, Iran (APIRC);
[b]Department of Pharmaceutical Chemistry, Faculty of Pharmaceutical Chemistry, Pharmaceutical Sciences Branch, Islamic Azad University, Tehran, Iran (IAUPS);
[c]Department of Medicinal Chemistry, School of Pharmacy, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
****Correspondance:***
Zahra Hajimahdic and Afshin Zarghi; PhD; Department of Medicinal Chemistry, School of Pharmacy, Shahid Beheshti University of Medical Sciences, Tehran, Iran;
Tel +1-734-764-2209
Fax +1-734-647-2110
E-mail: z.hajimahdi@sbmu.ac.ir, zarghi@sbmu.ac.ir

How to reference this article:
Int Pharm Acta. 2018, June1:2:190-197

Abstract

Introduction:

COX-2 inhibitory activities of some thiazinan, benzdiazinan, and benzthiazinan derivatives were modeled by quantitative structure–activity relationship (QSAR) with stepwise-multiple linear regression (SW-MLR) method.

Methods and Results:

The built model was robust and predictive with correlation coefficient (R^2) of 0.840 for training and 0.522 for test groups. The quality of the model was evaluated by leave-one out (LOO) cross validation and LOO correlation coefficient (Q^2) was 0.639. We also investigated a leverage approach for definition of applicability domain of model. According to QSAR model results, COX-2 inhibitory activity of selected data set had correlation with *VE3_Dzm* (Logarithmic coefficient sum of the last eigenvector from Barysz matrix weighted by mass), *GATS6c* (geometrical structure of the considered molecules in a complex way), and *GATS5i* (Geary autocorrelation of lag 5 weighted by ionization potential) descriptors derived from their structures.

Conclusion:

This study established a linear QSAR model for prediction of COX-2 inhibitory of thiazinan, benzthiazinan and benzdiazinan derivatives.

Keywords: COX-2 inhibitors, Thiazinan, Benzdiazinan, Benzthiazinan, Multiple linear regression, QSAR.

1. Introduction

Non-steroidal anti-inflammatory drugs (NSAIDs) are known for anti-inflammatory effects and widely used for decades. Their function is due to inhibition of cyclooxygenase (COX) enzyme, which catalyzes the conversion of arachidonic acid (AA) to prostaglandins (PG) [1]. In 1989, two isoforms of cyclooxygenase were identified. Cyclooxygenase-1 (COX-1) or prostaglandin H_1 synthase is present in many tissues and involved in many physiological processes, such as gastric cytoprotection, kidney function, and platelet aggregation. COX-2 or prostaglandin H_2 synthase is expressed in tissues in response to inflammatory stimuli and involved in pathologic processes [2].

All classical NSAIDs, such as aspirin, ibuprofen, and indomethacin, are non-selective and inhibit both COX-1 and COX-2. Therefore, these drugs caused some adverse effects such as gastric ulceration and renal damage [3, 4]. It is believed that selective COX-2 inhibitors provide anti-inflammatory effect, devoid of gastric and renal side effects, leading to an improved safety profile [5]. Subsequently, large efforts have been made to introduce selective COX-2 inhibitors [6-9]. As documented, COX-2 is found up-regulated in cancer cells and neurological disorders, such as Parkinson and Alzheimer's diseases [10, 11]. Therefore, attempts to design and discovery of novel selective COX-2 inhibitors are essential and attractive in medicinal chemistry.

In search for discovery of selective COX-2 inhibitors, several QSAR (quantitative structure–activity relationship) works have been carried out [12-14]. QSAR has been perceived as an approach to relate molecule structures to a wide variety of physical, chemical, biological, and technological properties. QSAR predicts that how chemical structure modifications can change proposed property, without experimental measurement. So development of a QSAR model reduces the cost of drug design and synthesis [15-17]. Major steps in constructing QSAR models include finding a data set of molecules, molecular descriptors calculation, variable selection, model building, and model validation.

The purpose of the present study was to explore the significant structural features responsible for COX-2 inhibitory activity of some thiazinan, benzdiazinan, and benzthiazinan derivatives which have been synthesized in our laboratory [18-20]. The results of *in-vitro* cyclooxygenase inhibition assays and molecular docking studies confirmed that thiazinan, benzdiazinan, and benzthiazinan scaffolds were capable of COX-2 inhibition (Figure 1). Consequently, QSAR analysis was carried out by multiple linear regressions (MLR) to model COX-2 inhibitory activity of compounds with their structural features.

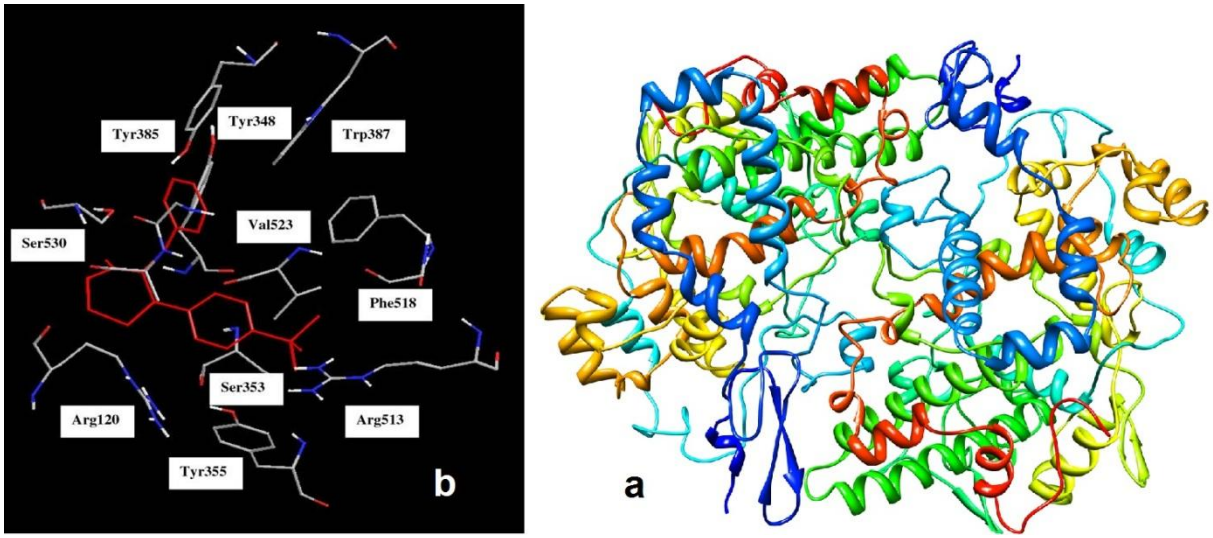


Figure 1. (a) structure of COX-2 and (b) Compound 17 docked in the active site of murine COX-2 isozyme [18]

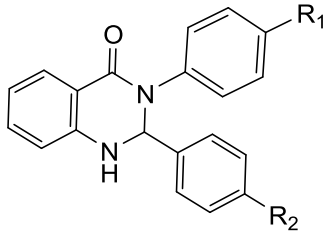
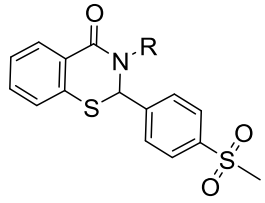
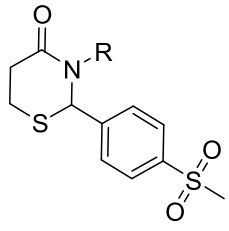
2. Materials and Methods

2.1. Data set

A set of 21 compounds bearing thiazinan, benzthiazinan, and benzdiazinan scaffolds which have been synthesized and evaluated as selective COX-2 inhibitors in our laboratory was used as a data set searching for the QSAR model (18-20). All the biological data expressed as IC_{50} were converted into pIC_{50} (-log IC_{50}) values. The total set of molecules was randomly separated into a training set (16 compounds) for generating QSAR model and a test set (5 compounds) for validating the quality of the model. Selection of a training and test set molecules was done carefully such the that test set contains a range of low, moderate, and high activity compounds similar to the training set. The general chemical structures and

biological activity values of all of the compounds are shown in Table 1.

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

Compounds	No	R		pIC_{50}		
		R_1	R_2	Exp.	Pred.	Res.
	1*	Cl	SO_2Me	6.76	6.37	0.158
	2*	F	SO_2Me	7.15	7.05	0.010
	3	H	SO_2Me	6.45	6.54	0.008
	4	Me	SO_2Me	7.00	6.94	0.002
	5	OMe	SO_2Me	7.15	7.02	0.017
	6*	SO_2Me	Cl	6.65	6.85	0.037
	7*	SO_2Me	F	7.04	7.06	0.000
	8	SO_2Me	H	6.92	6.90	0.000
	9	SO_2Me	Me	6.95	7.00	0.001
	10	SO_2Me	OMe	6.95	6.91	0.046
	11	Benzyl		7.15	7.14	0.000
	12	Phenyl		6.92	6.88	0.001
	13	Phenethyl		7.096	7.16	0.004
	14	4-F-phenyl		7.30	7.32	0.000
	15	4-Me-phenyl		7.22	7.11	0.012
	16	4-OMe-phenyl		6.95	7.12	0.026
	17*	Benzyl		7.22	7.13	0.008
	18	Butyl		7.096	7.12	0.000
	19	Pentyl		7.096	7.15	0.002
	20	Phenethyl		7.15	7.08	0.004
	21	Propyl		6.95	6.94	0.000
*Test Set						

2.2. Geometry optimization and molecular descriptor calculation

The 3D chemical structures of the molecules were obtained using the HyperChem 7.0 software [21]. The initial optimization was conducted using the molecular mechanics force field (MM+) procedure included in HyperChem, and then semi-empirical method AM1 was applied to optimize the molecules' geometry using the Polak-Ribiere algorithm until the root mean square gradient was 0.001 kcal mol^{-1}. PaDEL [22] and Dragon [23] softwares were used to calculate the descriptors, and finally 2942 molecular descriptors were calculated. The calculated descriptors were first analyzed for the existence of constant or near constant (>90%) variables. The detected ones were then removed. Secondly, the descriptors' correlation with each other and with the activity (pIC_{50}) of the molecules was examined, and the collinear descriptors (i.e. correlation coefficient between descriptors is greater than 0.9) were detected. Among the collinear descriptors, the one exhibiting the highest correlation with the activity was retained, and the others were removed from the data matrix, and finally 320 descriptors were remained.

3. Results

Stepwise multiple linear regression (SW-MLR) method was applied to select the most important set of descriptors. Stepwise method is a combination of two strategies; forward selection and backward elimination. Stepwise regression performs multiple regression by adding or removing descriptors to improve the fitness of the model [24]. Figure 2 shows the results of SW-MLR process. To investigate the optimum number of descriptors to be used in the equation, the numbers of descriptors were plotted against regression coefficient, R^2, and Standard Error of Estimate (SEE). As can be seen, R^2 increases with the increasing number of descriptors, while the values of SEE decrease with the increasing number of descriptors. According to the rule of thumb, at least five compounds should be included in the equation for every descriptor. So, regarding the number of compounds in train set (16 compounds), R^2, and SEE values, the most suitable model is three parametric one.

The MLR analysis with a stepwise selection was carried out to relate the pIC_{50} to a three set of descriptors. Data analysis was performed by SPSS software (statistic version 17.0). It is described by the following equation 1:

pIC_{50} = 8.07462(+/-0.42888) + 0.05657(+/-0.00786) *VE3_Dzm* - 1.3153(+/-0.33665) *GATS5i* +0.5073(+/-0.22693) *GATS6c*

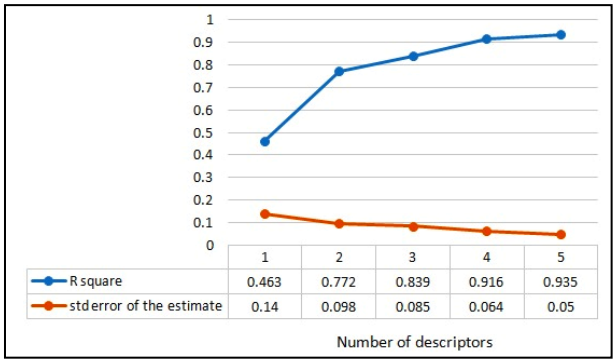


Figure 2. Influences of the number of descriptors on the R^2 and SEE of the regression model

The built SW-MLR model produced good results and could predict the activity of train and test sets with R^2 = 0.840 and R_{test}^2 = 0.522, respectively (Table 2). The predicted pIC_{50} values of the compounds in the training and test sets by SW-MLR model and residuals values (experimental pIC_{50} - predicted pIC_{50}) are shown in Table 1. The plot of the predicted pIC_{50} versus the experimental pIC_{50} obtained by the SW-MLR modeling, is demonstrated in Figure 3.

Table 2. Statistical parameters of SW-MLR model

Training Set		Test Set	F	Q^2_{LOO}
SEE	R^2	R^2	20.89	0.639
0.08541	0.840	0.522		

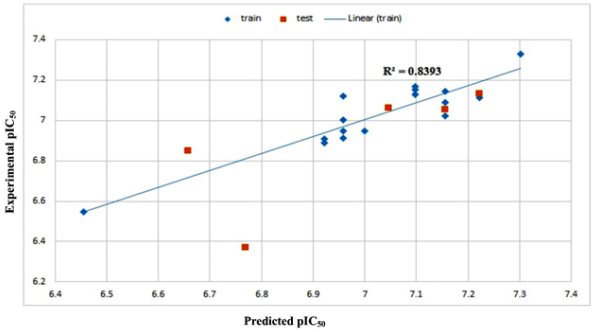


Figure 3. The predicted pIC_{50} values by the SW-MLR modeling versus the experimental pIC_{50} values

To validate the obtained model, the cross-validation analysis was performed using leave-one-out (LOO) method. For LOO cross-validation, one compound is removed from the train set, and the model is recalculated using the rest of the train set. The predicted activity for that point is then compared to its actual value. This is repeated until each data point is left off once. The cross-validated regression coefficient, Q^2, was found as 0.639, and hence this model can be termed as statistically significant.

The robustness of the resulting model was further validated by applying Y-randomization test. Several random shuffles were performed on dependent variable (pIC_{50}), and new QSAR models were built. The low R^2 and Q^2_{LOO} values show that the good results in the obtained model are regarded as reasonable and not because of a chance correlation (Table 3).

Table 3. R^2 and Q^2_{LOO} values of SW-MLR after several Y-randomization test

Iteration	R^2	Q^2_{LOO}
1	0.104	-0.313
2	0.075	-0.332
3	0.091	-0.418
4	0.148	-0.526
5	0.198	-0.373
6	0.080	-0.412
7	0.280	-0.003
8	0.181	-0.369
9	0.117	-0.587
10	0.124	-0.515

The values of the selected descriptors of SW-MLR model are shown in Table 4. Collinearity is a major disadvantage in MLR models [25]. Therefore, the inter-correlation between the three selected descriptors in SW-MLR model was calculated (Table 5). The results from Table 5 showed that the correlation coefficient value of each pair descriptors was less than 0.57; thus, the selected descriptors by stepwise method were completely independent.

Table 4. The descriptors values were used in model construction.

Name	*VE3_Dzm*	*GATS5i*	*GATS6c*
1	-19.738	0.813	0.952
2	-12.370	0.617	0.963
3	-16.214	0.797	0.867
4	-9.897	0.814	0.990
5	-7.657	0.823	0.912
6	-7.804	0.859	0.684
7	-7.258	0.708	0.649
8	-7.799	0.853	0.780
9	-6.666	0.829	0.780
10	-5.747	0.873	0.612
11	-4.074	0.894	0.934
12	-9.132	0.845	0.871
13	-2.635	0.940	0.940
14	-6.961	0.644	0.977
15	-6.248	0.856	1.017
16	-5.111	0.861	0.922
17	-3.728	0.899	0.890
18	-2.455	0.932	0.821
19	-1.931	0.976	0.924
20	-2.936	0.951	0.847
21	-2.958	1.031	0.783

Table 5. Correlation coefficient matrix of the selected descriptors by SW-MLR

	VE3_Dzm	*GATS5i*	*GATS6c*
VE3_Dzm	1	0.573	-0.158
GATS5i		1	-0.136
GATS6c			1

After internal and external validation, it cannot be claimed that this QSAR model is reliable for unknown sample unless its domain of application is defined. If the predictive value of the sample falls into this applicability domain, the value may be considered reliable. The leverage along with the Williams plot is usually used to define applicability domain of a model. The Williams plot is definedas the plot of the standardized residuals versus the leverage (*h*). In this plot, two horizontal lines and one vertical line mark a safety area. Compounds with standard residuals >3 standard deviation units and leverage higher than the warning h^* are regarded as outliers. The leverage (h_i) of every compound is calculated by the following equation:

$h_i = x_i (X^T X)^{-1} x_i^T$

In this equation, x_i is the descriptor-row vector of the query molecule and X is the $k \times n$ matrix containing the k descriptor values for each one of the n training molecules. The critical leverage h^* (the vertical line) is fixed at $3(k + 1)/n$ [26, 27]. From the Williams plot (Figure 4), it is obvious that all data points fall within the safety area in the model. All of the compounds have the leverage lower than the warning h^* value of 0.75. As a result, it can be said that the model is acceptable for prediction purpose.

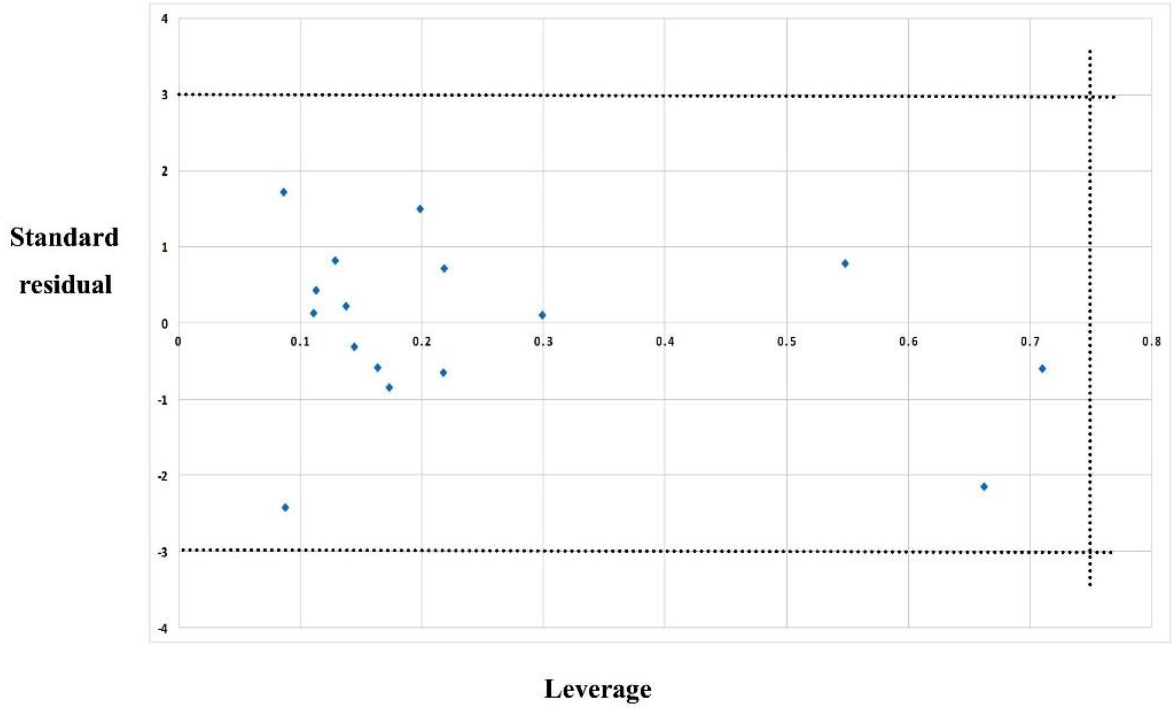


Figure 4. The William plot for the SW-MLR model

4. Discussion and Conclusion

The QSAR models provide some information about structural features influencing the biological activity of the studied compounds. Thus, interpretation of the descriptors appeared in the QSAR model could help to obtain some insight into the factors that are likely to have effects on the pIC_{50}. According to the internal and external validation parameters, SW-MLR model possessed good fitting ability, good predictive ability, and high stability. The relative importance of the descriptors in the QSAR model was identified based

on their standardized regression coefficients. The calculated MLR coefficients cannot be used because the descriptors in the final MLR models have not the same units. Standardized regression coefficients of the selected descriptors in SW-MLR model are shown graphically in Figure 5. As can be seen, *VE3-Dzm* is more significant than the other descriptors with a positive sign. An explanation of the selected descriptors using handbook of molecular descriptors follows next [28, 29].

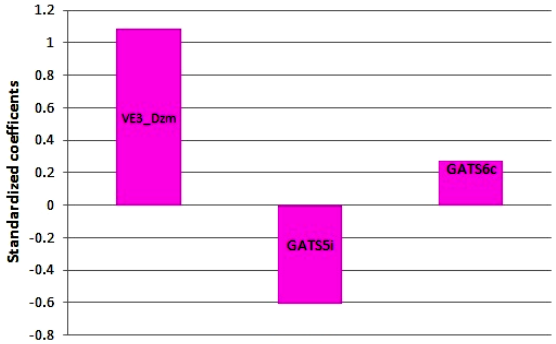


Figure 5. Standardized coefficients versus descriptor values in MLR

VE3-Dzm refers to logarithmic coefficient sum of the last eigenvector from Barysz matrix weighted by mass. Another group of topological descriptors, used in DT models are those calculated from a Barysz matrix. The Barysz matrix is a weighted distance matrix that counts for the presence of heteroatoms and multiple bonds in a molecule [30]. This descriptor has a positive effect on the inhibitory activity. Hence, it is concluded that increasing the value of this descriptor causes an increase of COX-2 inhibitory activity.

GATS6c is related to the geometrical structure of the considered molecules in a complex way.

The Geary autocorrelation (GATS) is defined as [31]:

$$GATS = \frac{(A-1)\sum_{I=1}^{A}\sum_{J=1}^{A}(wi-wj)2}{2\Delta k\sum_{i=1}^{A}(wi-\varpi)2}$$

Where $\boldsymbol{\varpi}$ is the average value of atomic property $\mathbf{w_i}$ over the molecule. The descriptors *GATS6c* is the Geary autocorrelations with lag k = 6 and the property **w** being the atomic charges. Since it presented a positive sign in MLR equation, increasein the value of this descriptor will lead to increase in the activity (pIC_{50}).

GATS5i displays Geary autocorrelation of lag 5 weighted by ionization potential. In this descriptor, the Geary coefficient is a distance-type function, that function is any physicochemical property calculated for each atom of the molecule, such as atomic mass, polarizability, etc. The physicochemical property in this case is ionization potential [32]. As can be seen it has negative sign; therefore, increase in the *GATS5i* descriptor leads to decrease in pIC_{50} value.

In summary, this study established a linear QSAR model for prediction of COX-2 inhibitory of thiazinan, benzthiazinan, and benzdiazinan derivatives. The model was validated using LOO cross-validation, Y-randomization, and external test set. The built model had a good self- and external-predictive power. The calculated applicability domain of the model showed that the obtained model was acceptable for prediction purpose. Based on QSAR models results, *VE3-Dzm*, *GATS6c*, and *GATS5i* were found to be important factors, controlling the COX-2 inhibitory activity.

References

1. Simmons DL, Botting RM, Hla T. Cyclooxygenase isozymes: the biology of prostaglandin synthesis and inhibition. Pharmacol Rev. 2004;5:387-437.

2. Smith W, DeWitt D, Garavito R. Cyclooxygenases: structural, cellular, and molecular biology. Annu Rev Biochem. 2000;69:145–82.

3. Henry D. Side-effects of non-steroidal anti-inflammatory drugs. Bailliere Clin Rheum. 1998;2:425–54.

4. Raskin J. Gastrointestinal effects of nonsteroidal anti-inflammatory therapy. Am J Med. 1999;106:4S–12S.

5. Zarghi A, Arfaei S. Selective COX-2 inhibitors: A review of their structure-activity relationships. Iran J Pharm Res. 2011;10:655-83.

6. Arfaie S, Zarghi A. Design, synthesis and biological evaluation of new (E)- and (Z)-1,2,3-triaryl-2-propen-1-ones as selective COX-2 inhibitors. Eur J Med Chem. 2010;45:4013-7.

7. Ghodsi R, Azizi E, Zarghi A. Design, Synthesis and Biological Evaluation of 4-(Imidazolylmethyl)-2-(4-methylsulfonyl phenyl)-Quinoline Derivatives as Selective COX-2 Inhibitors and *In-vitro* Anti-breast Cancer Agents. Iran J Pharm Res. 2016;15:169-77.

8. Sabakhi I, Topuzyan V, Hajimahdi Z, Daraei B, Arefi H, Zarghi A. Design, Synthesis and Biological Evaluation of New 1, 4-Dihydropyridine (DHP)

Derivatives as Selective Cyclooxygenase-2 Inhibitors. Iran J Pharm Res. 2015;14:1087-93.

9. Zarghi A, Sabakhi I, Topuzyan V, Hajimahdi Z, Daraei B. Design, Synthesis and Biological Evaluation of 5-Oxo-1,4,5,6,7,8-Hexahydroquinoline Derivatives as Selective Cyclooxygenase-2 Inhibitors. Iran J Pharm Res. 2014;13 (supplement):61-9.

10. Shen FQ, Wang ZC, Wu SY, Ren SZ, Man RJ, Wang BZ, et al. Synthesis of novel hybrids of pyrazole and coumarin as dual inhibitors of COX-2 and 5-LOX. Bioorganic & medicinal chemistry letters. 2017;27:3653-60.

11. Xiang Z, Ho L, Valdellon J, Borchelt D, Kelley K, Spielman L, et al. Cyclooxygenase (COX)-2 and cell cycle activity in a transgenicmouse model of Alzheimer's disease neuropathology. Neurobiol Aging. 2002;23:327-34.

12. Soltani S, Abolhasani H, Zarghi A, Jouyban A. QSAR analysis of diaryl COX-2 inhibitors: comparison of feature selection and train-test data selection methods. Eur J Med Chem. 2010;45(7):2753-60.

13. Akbari S, Zebardast T, Zarghi A, Hajimahdi Z. QSAR Modeling of COX -2 Inhibitory Activity of Some Dihydropyridine and Hydroquinoline Derivatives Using Multiple Linear Regression (MLR) Method. Iran J Pharm Res. 2017;16:525-32.

14. Irannejad H, Kebriaieezadeh A, Zarghi A, Montazer-Sadegh F, Shafiee A, Assadieskandar A, et al. Synthesis, docking simulation, biological evaluations and 3D-QSAR study of 5-Aryl-6-(4-methylsulfonyl)-3-(metylthio)-1,2,4-triazine as selective cyclooxygenase-2 inhibitors. Bioorg Med Chem. 2014;22:865-73.

15. Hajimahdi Z, Ranjbar A, Abolfazl Suratgar A, Zarghi A. QSAR Study on Anti-HIV-1 Activity of 4-Oxo-1,4-dihydroquinoline and 4-Oxo-4H-pyrido[1,2-a]pyrimidine Derivatives Using SW-MLR, Artificial Neural Network and Filtering Methods. Iran J Pharm Res. 2015;14(Suppl):69-75.

16. Hajimahdi Z, Safizadeh F, Zarghi A. QSAR Analysis for Some 1, 2-Benzisothiazol-3-one Derivatives as Caspase-3 Inhibitors by Stepwise MLR Method. Iran J Pharm Res. 2016;15:439-48.

17. Faraji N, Zebardast T, Zarghi A, Hajimahdi Z. QSAR Modeling of Aminopeptidase N/CD13 (APN) Inhibitory Activity of some Leucine Ureido Derivatives by GA-MLR and SW-MLR Methods. Lett Drug Des Discov. 2017;14:1348 - 57.

18. Zebardast T, Zarghi A, Daraie B, Hedayati M, Dadrass OG. Design and synthesis of 3-alkyl-2-aryl-1,3-thiazinan-4-one derivatives as selective cyclooxygenase (COX-2) inhibitors. Bioorg Med Chem Lett. 2009;19:3162-5.

19. Zarghi A, Zebardast T, Hajighasemali F, Alipoor E, Daraie B, Hedayati M. Design and Synthesis of New 1,3-Benzdiazinan-4-one Derivatives as Selective Cyclooxygenase (COX-2) Inhibitors. Arch Pharm Chem Life Sci. 2012;345:257-64.

20. Zarghi A, Zebardast T, Daraie B, Hedayati M. Design and synthesis of new 1,3-benzthiazinan-4-one derivatives as selective cyclooxygenase (COX-2) inhibitors. Bioorg & Med Chem. 2009;17:5369-73.

21. Froimowitz M. HyperChem: a software package for computational chemistry and molecular modeling. Biotechniques. 1993;14:1010-3.

22. Yap C. PaDel-descriptor: An open source software to calculate molecular descriptors and fingerprints. J Comput Chem. 2011;32:1466-74.

23. Todeschini R, Consonni V, Pavana M. Milano chemometrics and QSPR Group. 2008.

24. Nathans LL, Oswald FL, Nimon K. Interpreting multiple linear regression: A guidebook of variable importance. Pract Assess, Res Eval. 2012;17:1-19.

25. Mukaka M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi Med J. 2012;24:69-71.

26. Wu W, Zhang C, Lin W, Chen Q, Guo X, Qian Y, et al. Quantitative structure-property relationship (QSPR) modeling of drug-loaded polymeric micelles via genetic function approximation. PLOS ONE 2015;10:e0119575.

27. Gramatica P. Principles of QSAR models validation: internal and external. QSAR Comb Sci 2007;26: 694-701.

28. Todeschini R, Consonni V. Handbook of molecular descriptors: John Wiley & Sons; 2008.

29. Todeschini R, Consonni V. Methods and Principles in Medicinal Chemistry: Handbook of Molecular Descriptors. Volume 11, Editor: Mannhold R., Kubinyi H. and Timmerman H. Wiley-VCH: Weinheim; 2000.

30. Antanasijević J, Pocajt V, Antanasijević D, Trišović N, Fodor-Csorba K. Prediction of clearing temperatures of bent-core liquid crystals using decision trees and multivariate adaptive regression splines. Liquid Crystals. 2016;43:1028-37.

31. Weinebeck A, Kaminski S, Murrenhoff H, Leonhard K. A new QSPR-based prediction model for biofuel lubricity. Tribol Int. 2017;115(Supplement C):274-84.

32. Ghasemi JB, Zohrabi P, Khajehsharifi H. Quantitative structure–activity relationship study of nonpeptide antagonists of CXCR2 using stepwise multiple linear regression analysis. Monatshefte für Chemie-Chemical Monthly. 2010;141:111-8.