Pharmaceutical Sciences, September 2015, 21, 86-93
doi: 10.15171/PS.2015.22
http://journals.tbzmed.ac.ir/PHARM

Research Article

QSAR Analysis of Cyclooxygenase Inhibitors Selectivity Index (COX_1/COX_2): Application of SVM-RBF and MLR Methods

Siavash Tatardar[1], Abolghasem Jouyban[2], Somaieh Soltani[2*], Mostafa Zakariazadeh[3]
[1]*Student Research Committee, Tabriz University of medical sciences, Tabriz, Iran.*
[2]*Drug Applied Research Center and Pharmacy Faculty, Tabriz University of medical sciences, Tabriz, Iran.*
[3]*Biothechnology Research Center, Tabriz University of medical sciences, Tabriz, Iran.*

Article info

Article History:
Received: 4 January 2015
Accepted: 16 April 2015
ePublished: 30 September 2015

Keywords:
COX Inhibitor
QSAR
MLR
SVM-RBF
Selectivity Index

ABSTRACT

Background: Anti-inflammatory inhibitors of cyclooxygenase 2 (COX2) have been shown to increase the risk of adverse cardiovascular events in clinical trials. The studies showed that such adverse events could be due to COX2-induced suppression of prostaglandin I2 (PGI2) synthesis. These adverse effects related to the degree of COX2 selectivity of NSAIDs. Study of the selectivity index of COX1/COX2 is important for development of the new NSAID drugs. Prediction methods of such index have been interested by scientists. ***Methods:*** The selectivity index of a number of 68 molecules from 8 different chemical groups was predicted using MLR and SVM-RBF models. Calculated structural and physicochemical parameters, using the energy optimized molecular structures were applied to develop the desired models. The developed models were validated using LMO, external test set and Y-randomization methods. ***Results:*** Regression coefficient of developed MLR model was 0.825 and 0.752 for training and test sets, while for SVM-RBF model it was 0.628 and 0.863 for training and test sets. The RMSE of the developed models were 0.08, 0.06, 0.29 and 0.16 respectively for train (MLR, SVM-RBF) and test (MLR, SVM-RBF) datasets. ***Conclusion:*** The validation results showed the higher prediction capability for SVM-RBF in comparison with MLR models. The selected descriptors showed the contribution of electronic parameters in conjunction with size and shape parameters in selectivity of studied compounds.

Introduction

Anti-inflammatory role of Non-steroidal anti-inflammatory drugs (NSAIDs) are known for decades, but their rapid drug discovery occurred following the isolation and crystallization of cyclooxygenase (COX), which catalyzes the conversion of Arachidonic acid (AA) to prostaglandins (PG).[1] Two isoforms of COX; i.e COX_1 (or constitutive form) which is responsible for the maintenance of physiologic homeostasis, and COX_2 (or inducible isoform in most of the body tissues) which is responsible for inflammation are studied continuously as anti-inflammatory drug targets.[1,2]

Findings about two isoforms lead to the theories about the association of gastrointestinal and renal adverse effects of classic NSAIDs to their COX_1 inhibitory activities, and attempts to develop COX_2 selective inhibitors, lead to reduced gastrointestinal side effects.

The structural features of selective COX_2 inhibitors studied during the years. According to the differences between COX_1 and COX_2 enzymes active site, which arises mainly from the 523 amino acid residue, it is generally concluded that the selective COX_2 inhibitors have larger molecular size than non-selective inhibitors (Figure1). The structure activity relationship of developed selective COX_2 inhibitors were reviewed and the findings showed the successful diaryl hetero/carbocyclic derivatives (e.g. celecoxib) as selective COX_2 inhibitors.[3,4]

Cardiovascular side effects of some selective COX2 inhibitors resulted in the withdrawal of rofecoxib in 2004 and valdecoxib in 2005 from the market. The safety of this class of drugs was questioned with these findings and it gave a setback to the drug discovery of selective COX-2 inhibitors.4 In addition findings about COX inhibitors role in cancer 5-8 and Alzheimer diseases,9 renewed the medicinal chemists' attention and attempts to study the needed structural features of balanced selectivity.

Many QSAR models are developed to predict and characterize the COX2 and COX1 inhibitory of different inhibitors, while selectivity index of the studied inhibitors have not been studied well until know.10 The complicated nature of this index and its relation to two enzymes makes the model development

***Corresponding Author:** Somaieh Soltani, Tel: (+98) 41 33372250, Fax: (+98) 41 33344798, E-mail: soltanisomaieh@gmail.com

difficult and studies in this issue would help the consequent research. The selectivity index (SI) of a diverse set of COX2 and COX1 inhibitors (Table 1) were studied. The only available model was reported 10 in 2012 for a limited data set using GA-PLS-MLR method in which the SI was not predicted using a single model and the ratio of QSAR models which was developed for COX1 and COX2 selectivity, was used to SI prediction. The present paper reported the QSAR analysis of selectivity index for a various dataset of chemical structures by the application of MLR (multivariate linear regression) and SVM-RBF (support vector machine-radial basis function) methods.

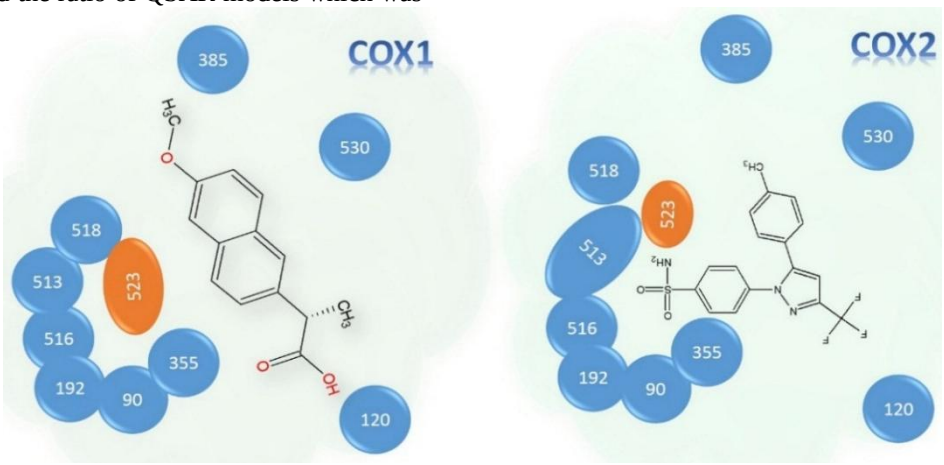


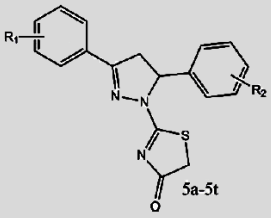
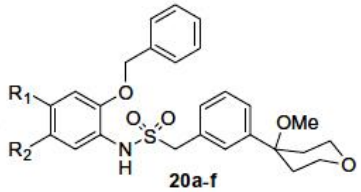
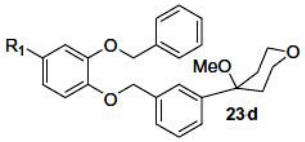
Figure 1. COX1 and COX2 active site differences. The extra hydrophob packet in COX2 because of replacement of IL523 with Val523.

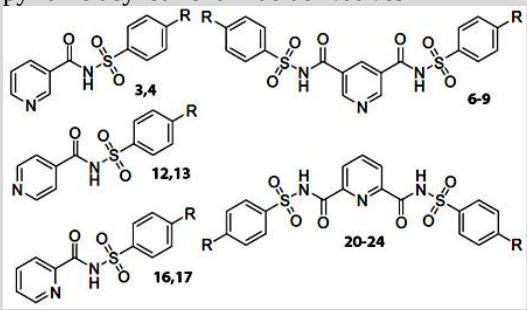
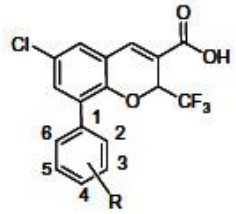
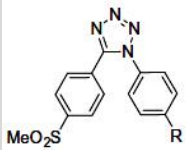
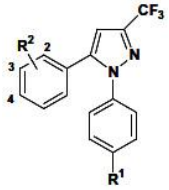
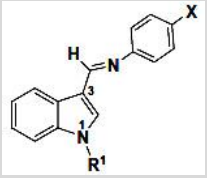
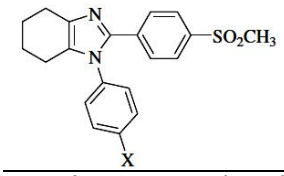
Materials and Methods

Data set

۶۸ different COX inhibitor from 9 different structural categories were obtained from literature.[11-18] The details of COX_2 and COX_1 activity and the corresponding SI (COX_1/COX_2) are presented in Table1.

Table 1. Studied dataset. Molecular structure, observed and predicted logSI values.

Chemical group	code	Normalized COX_1	Normalized COX_2	logSI (obs)	logSI (MLR pred)	logSI (SVM-RBF pred)
dihydro-pyrazolyl-thiazolinone derivatives [11]  5a-5t	A01	42.4	0.5	1.93	2.11	1.77
	A02	40.8	6.4	0.81	0.95	1.08
	A03	48.6	4.5	1.03	0.98	1.07
	A05*	**45.3**	**1.2**	**1.58**	**0.80**	**1.60**
	A09	44.1	7.2	0.79	0.85	0.92
	A10	47.2	6.3	0.88	0.72	1.70
	A11	**33.2**	**2.7**	**1.09**	**0.56**	**0.74**
	A12	34.3	19.4	0.26	0.71	0.57
	A13	37.7	11.7	0.51	0.61	0.47
	A14	48.5	7.6	0.81	0.67	0.65
	A15	44.3	4.7	0.97	0.75	0.82
	A19	40.2	18.4	0.34	0.51	0.62
	A20	45.6	16.3	0.45	0.61	0.65
Nimesulide derivatives [12]  20a-f	B01	14.8	2.3	0.81	0.87	0.97
	B02	24.4	3.7	0.82	0.92	0.84
	B03	15.6	2.9	0.73	0.87	0.87
	B04	16.5	4.9	0.53	0.81	0.76
	B05	7.9	1.5	0.72	0.80	0.88
	B06	11.6	0.97	1.08	0.94	0.69
 23d	B07	44.2	38.8	0.04	-0.06	0.43

pyridine acyl sulfonamide derivatives[13] 	C01	41.2	12.6	0.51	0.86	0.72
	C02	38	7.8	0.69	1.14	0.70
	C03	43.8	6.3	0.84	0.76	0.51
	C04	**39.7**	**6**	**0.82**	**0.81**	**1.02**
	C05	35.8	3.7	0.99	0.70	1.08
	C06	40.3	2.6	1.19	1.46	1.05
	C07	41.5	1.9	1.34	1.43	0.80
	C08	35.6	0.8	1.65	1.53	0.85
	C09	39.8	5.6	0.85	1.50	0.68
	C10	**35.8**	**6.6**	**0.73**	**0.20**	**0.80**
	C11	**40.5**	**4.9**	**0.92**	**0.50**	**0.81**
	C12	36.9	7.9	0.67	0.26	1.17
	C13	40.2	5.6	0.86	1.01	1.31
	C14	29.8	3.4	0.94	0.95	1.49
	C15	42.5	2.8	1.18	1.24	1.43
Benzopyran derivatives [14] 	D01	91.4	0.76	2.08	1.38	1.71
	D02	3.87	21.3	-0.74	-0.45	-0.29
	D03	1.39	1.82	-0.12	0.24	-0.28
	D04	2.44	6.38	-0.42	0.23	-0.26
	D05	2.91	3.48	-0.08	0.10	-0.13
	D06	2.91	0.19	0.91	0.76	0.96
	D07	0.24	0.02	1.06	0.77	1.18
	D08	4.34	0.1	1.62	1.78	1.47
	D09	17.7	0.08	2.35	1.83	2.05
	D10	**32**	**0.36**	**1.95**	**2.06**	**1.94**
	D11	75.9	0.24	2.5	2.12	2.34
	D12	26.5	13.4	0.3	0.74	0.23
1,5-diaryl-substituted tetrazoles [15] 	E01	3700	200	1.27	1.49	1.40
	E02	811	45	2.26	2.24	2.04
	E03	420	2	2.32	2.34	2.13
	E04	640	2.7	2.37	1.62	1.56
	E05	500	110	0.65	1.50	1.12
Isomeric acetoxy analogs of celecoxib [16] 	F03	80.3	4.1	1.29	1.24	1.58
	F04	0.7	0.02	1.67	1.37	1.57
N-1 and C-3 substituted indole Schiff bases[17] 	**G01**	**62.3**	**30.4**	**0.3**	**-0.05**	**0.76**
	G02	78.2	11.6	0.83	0.96	1.11
	G03	91.1	0.71	2.11	1.72	0.68
	G04	81.5	17.5	0.67	1.12	1.81
	G05	95.5	0.84	2.06	2.14	0.83
	G06	80.2	12.4	0.81	0.54	2.05
	G07	86.1	6.53	1.12	0.96	0.97
	G08	**68.4**	**17.2**	**0.6**	**0.97**	**0.93**
1,2-diaryl-4,5,6,7-tetrahydro-1Hbenzo[d] imidazoles[18] 	H01	56.5	0.55	2.01	1.99	2.08
	H02	**55.7**	**0.34**	**2.21**	**1.99**	**2.15**
	H03	61.6	0.67	1.96	1.91	1.82
	H04	40.1	0.62	1.81	1.93	1.89
	H05	42.2	0.37	2.06	1.96	1.82
	H06	36.1	0.69	1.72	2.01	1.88

*Data for test set are shown in bold.

Calculated descriptors

The 2D structures of all molecules were drawn and converted to 3D structures using HyperChem 8.0 software. The energy of molecules were optimized

using molecular mechanics (MM+) followed by semi-empirical (AM_1) methods. 1664 descriptor were calculated using Dragon 5.4 software which reduced to 1200 descriptors after discarding constant and near constant values (correlation coefficient >0.99). All calculations were performed on an Intel inside Corei7 personal computer using SPSS 19 and Statistica 7 software.

Dataset preparation, outlier detection and descriptor selection

The data set was randomly divided to training (59 data points) and prediction (9 data points) set after sorting according to their SI. SI was calculated using equation 1. The IC_{50} values of data points were normalized using celecoxib as standard compound before logSI calculation.

$$\text{logSI}=\log\left(\frac{IC_{50}COX_1}{IC_{50}COX_2}\right) \quad \text{Eq. (1)}$$

The data set were checked to define outlier compounds. To do this we used data scoring method based on correlation of normalized logSI with standardized logSI. The data points with scores >3 or <-3 were considered as outlier and removed from further investigation. The simple stepwise regression were used to select the descriptors.

QSAR model development

MLR and SVM-RBF methods were used to develop the desired models based on the selected descriptors. MLR model were developed using the selected features based on stepwise regression method and were judged through R (correlation coefficient) and RMSE (root mean square error) values for training and prediction sets. The validity of the developed models was further evaluated using Leave many out cross validation and Y- randomization methods.

Following the introduction of SVM in 1990s as a useful tool for solving complicated problems (e.g. interactions between the ligand and its biological target),[19] most of the researchers developed QSAR models[20-25] based on SVM using linear (i.e. regression) or nonlinear (i.e. ANN) kernels.

Like ANN, SVMs are able to describe nonlinear relationships better than linear methods and most of the developed QSARs were able to approximate the structure–relationship issues as one of the most studied nonlinear problems well.[19]

SVM was originally developed for linear two-class classification with margin (i.e. minimal distance from the separating hyper plane to the closest data points). SVM learning machine searches for an optimal separating hyper-plane. The solution is only based on the data points placed at the margins (i.e. support vectors). Nonlinear SVM can be developed by transforming of problem into a feature space using a set of nonlinear basis functions. A kernel representation is applied to write the solution as a weighted sum of the values of certain kernel function evaluated at the support vectors.[25] Gaussian radial basis function (RBF) was used as a nonlinear kernel function in the present paper. The RBF is the most popular choice of kernel types used in SVMs[26] because of its flexibility in fitting data.

The SVM-RBF model was optimized according to the gamma parameter jointly with the C parameter [27] based on the 10 fold leave many out cross validation optimization procedure. Gamma is a kernel function parameter which handles the nonlinear pattern and the smaller gamma will lead to lower bias, where C is the penalty parameter which handles the hyper plane and larger C will lead to lower bias while the over fitting could occur. Consequently a joint optimization within a correct range of C and Gamma should be done to develop an acceptable SVM. The developed models were externally validated using a test data set.

Results and Discussion

Prepared data set

Table 1 shows the prepared data set. Test data are shown as bold data points. The outlier analysis results are shown in Figure 2. According to the results there is not any data point out of border values.

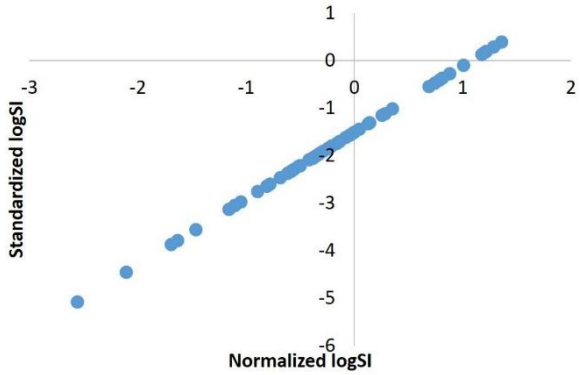


Figure 2. Standard scoring of the studied data set.

Selected descriptors

All variables were passed the tolerance criterion of 0.05 to be entered in a single equation. Also, a variable is not entered if it would cause the tolerance of another variable already in the model to drop below the tolerance criterion. The final selected parameters were cross correlated and the all descriptors have correlation coefficient less than 0.6). Table 2 shows the details of selected descriptors and their cross correlation results. According to the standardized coefficients of the selected descriptors the contribution of each descriptor in the model correlation capability is shown by score value in table 2. The most contributed descriptor gained the least score. The selected descriptors showed the contribution of electronic parameters in conjunction with size and shape parameters in selectivity of studied compounds.

Table 2. Selected descriptors definitions and their Pearson inter correlations.

Descriptor Category	Description	Model number	Score	Descriptor Name	JGI8	Infective80	BEHp1	HOMT	L1s	X4A	X3A	Hypertens50	R4ett	GATS5p
2D autocorrelations	mean topological charge index of order 8	1	.5	JGI8	1									
Drug-like indices	Ghose-Viswanadhan-Wendoloski antiinfective-like index at 80%	2	3	Infective-80	.564	1								
Burden eigenvalues	largest eigenvalue n. 1 of Burden matrix weighted by polarizability	3	2	BEHp1	.517	.486	1							
Geometrical	HOMA total	4	1	HOMT	.165	-.182	.516	1						
WHIM	1st component size directional WHIM index / weighted by I-state	5	6	L1s	-.141	-.434	-.047	.516	1					
Connectivity indices	average connectivity index of order 4	9	4	X4A	-.100	.277	.316	.173	-.247	1				
Connectivity indices	average connectivity index of order 3	8	7	X3A	-.332	.062	.019	-.076	-.290	.831	1			
Drug-like indices	Ghose-Viswanadhan-Wendoloski antihypertensive-like index at 50%	6	10	Hypertens-50	.358	.202	.173	.078	-.110	.102	-.014	1		
GETAWAY	R maximal autocorrelation of lag 4 / weighted by Sanderson electronegativity	10	8	R4e+	.025	.480	.203	-.372	-.336	.396	.421	.050	1	
2D autocorrelations	Geary autocorrelation of lag 5 weighted by polarizability	7	9	GATS5p	-.522	-.275	-.080	.142	-.111	.383	.494	-.091	.063	1

Model construction

MLR model

10 different MLR model was developed using the selected descriptors (the descriptors were added to the model according to the model number in Table 2). The best model (Eq. 2) with R^2=0.825, F=23.00 and SEE=0.33 possessed RMSE value of 0.08 and 0.29 for training and test set respectively. The regression coefficient of predicted versus observed values for test set was 0.752.

logSI=-63.9+84.7 JGI8-1.2 Infective80+16.8 BEHp1-0.3 HOMT+0.1 L1s+67.6 X4A-35.6 X3A+0.8 Hypertens50-12.3 R4ett-0.7 GATS5p Eq.(2)

The significant contributing of the descriptors addition to the model correlation and prediction capability were checked by the variation of adjusted R2 and standard error of the estimate (S.E.E) of the developed models following each descriptor addition to the model. Figure 3 showed the results in which the adjusted R2 increase and S.E.E. decrease after each descriptor addition. Removing of the least scored descriptor (Hypertens50) led to no significant variation in developed model parameters while the other drug like property (i.e. Infective80 with score 3) was necessary for a significant correlation. The predicted versus experimental selectivity index values correlation are shown in Figure 4.

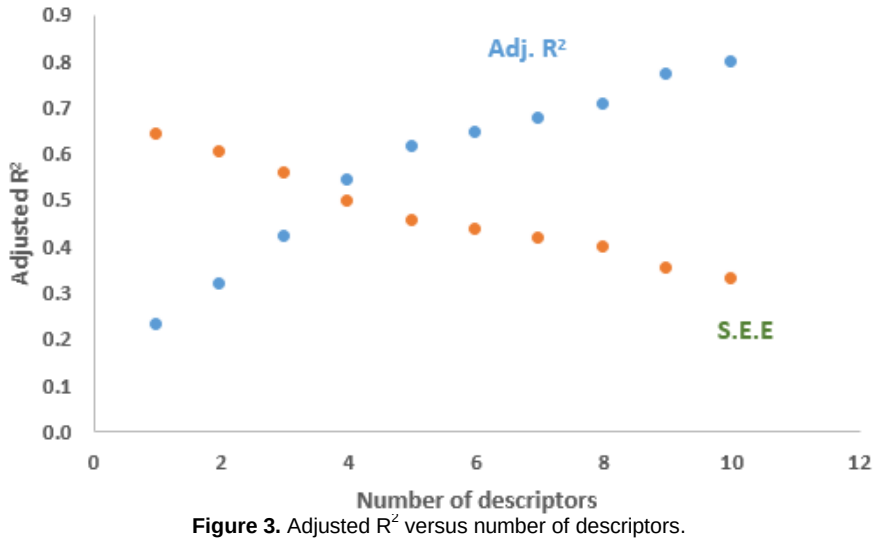
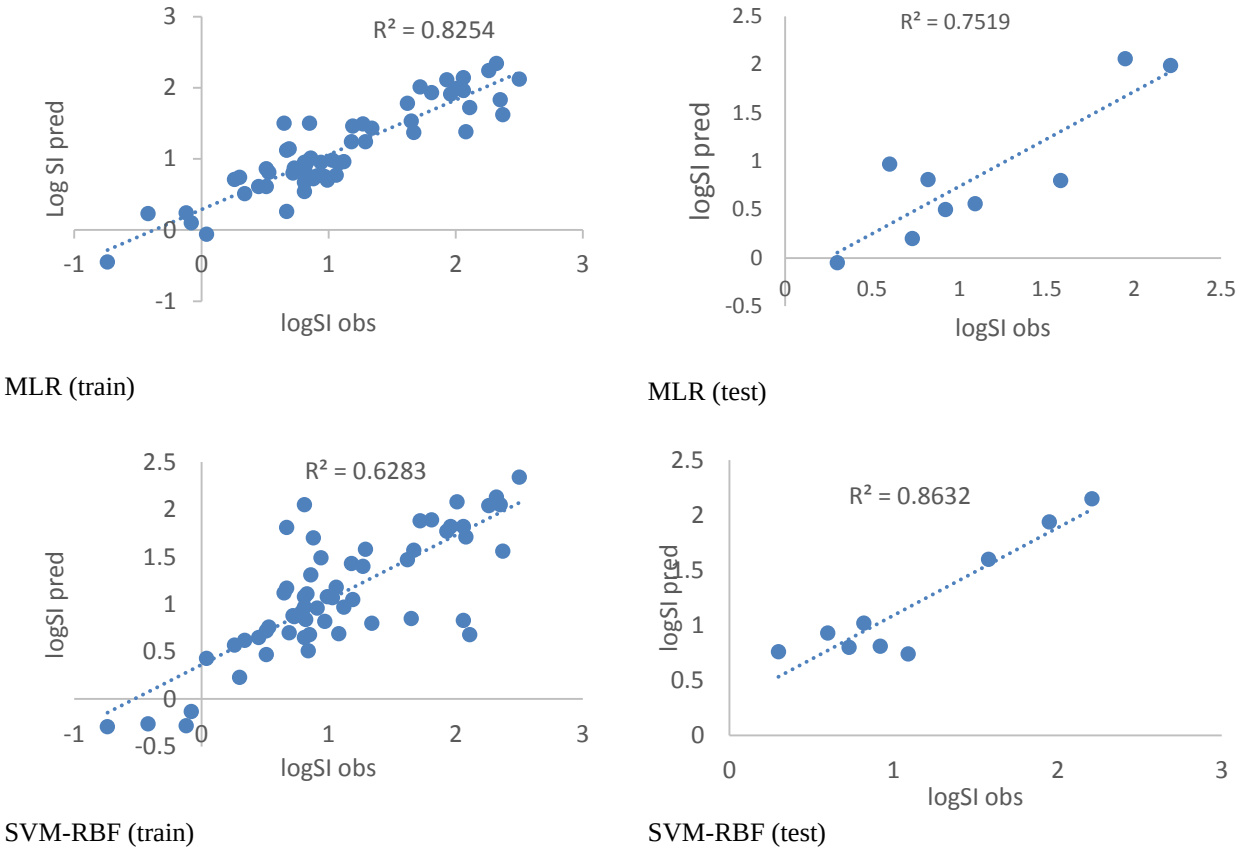


Figure 3. Adjusted R^2 versus number of descriptors.



MLR (train) MLR (test)

SVM-RBF (train) SVM-RBF (test)

Figure 4. Predicted versus observed logSI.

SVM-RBF model

The selected descriptors were applied to the developing of SVM-RBF model. The model was optimized according to the Gamma value in the range of 0.01-0.2, and the Gamma value of 0.15 resulted to the best SVM-RBF model. The resulted model parameters were: capacity=10.000, epsilon=0.100, gamma=0.15. Number of support vectors for the best model was 34, in which 18 of them were bounded. Train and test sets regression coefficient (RMSE) was 0.628 (0.06) and 0.863 (0.16) respectively. Cross-validation error was 0.068 for 10 fold cross validation.

The differences between correlation coefficient and RMSE of training set between SVM-RBF and MLR models reveals the better correlation of back calculated logSI values using MLR method, while the SVM-RBF model was able to predict the external dataset SI significantly better than MLR model. The predicted

versus experimental selectivity index values correlation are shown in Figure 4.

Model Validation

Leave many out cross validation

10 fold Leave many out (LMO) cross correlation was done by removing of 9 molecule in each step (the data was sorted according to SI beforehand), and the MLR and SVM-RBF models were developed based on remained compounds. The developed MLR model was used to predict the excluded compounds activity and the correlation coefficients are summarized in Table 3. No significant variation occurred in correlation coefficient and the models were reliable according to the LMO cross validation. The 10 fold cross validation of SVM-RBF model were done during model development procedure and the resulted cross validation error was 0.068.

Table 3. LMO cross validation results.

Model number	Correlation coefficient
1	0.872
2	0.886
3	0.911
4	0.893
5	0.896
6	0.901
7	0.900
8	0.879
9	0.880
10	0.866
Mean (±SD)	0.888 (0.014)

Y- Randomization

The logSI values were shuffled 10 times and the correlation coefficient of the developed models were calculated for each step. Results are shown in Table 4. The regression coefficients below 0.2 indicates that the random correlation was not happened during model development.

Table 4. Y randomization results.

Shuffled data	Randomized Regression coefficient
1	0.150
2	0.094
3	0.121
4	0.193
5	0.220
6	0.139
7	0.198
8	0.193
9	0.211
10	0.117
Mean (±SD)	0.164 (0.04467)

Conclusion

The results suggest that both two and three dimensional descriptors are participated in selectivity index prediction, while development of COX_2 QSAR model using the selected descriptors was not possible. Application of the selected descriptors for COX1 QSAR model development resulted in a 4 descriptors significant MLR model. Both MLR and SVM-RBF models were able to predict the selectivity with acceptable errors (in comparison with experimental celecoxib selectivity RSD% for different studies which was has been used as reference SI for data normalization), while SVM-RBF model was produced more accurate results and can be used as a validated model for similar data sets. This study was a step forward to enable researchers to estimate the selectivity index of studied compounds, but the availability of a reliable dataset which contains both COX_1 and COX_2 selective compounds inhibition activities with acceptable experimental RSD% would help to improve the developed models accuracy and applicability domain.

Acknowledgment

This paper is contributed to a part of Mr Siavash Tatardar PharmD thesis and partially funded by Tabriz university of medical sciences through grant number **52/3547.**

Conflict of Interest

The authors report no conflicts of interest.

References

1. Simmons DL, Botting RM, Hla T. Cyclooxygenase Isozymes: The Biology of Prostaglandin Synthesis and Inhibition. *Pharmacol Rev* 2004;56:387-437. doi: 10.1124/pr.56.3.3
2. Flower RJ. The development of COX2 inhibitors. *Nat Rev Drug Discov* 2003;2:179-191. doi:10.1038/nrd1034
3. Zarghi A, Arfaei S. Selective COX-2 Inhibitors: A Review of Their Structure-Activity Relationships. *Iran J Pharm Res* 2011;10:655-683.
4. Asit K. Chakraborti SKG, Raj Kumar, Hashim F. Motiwala and Pradeep S. Jadhavar. Progress in COX-2 Inhibitors: A Journey So Far. *Curr Med Chem* 2010;17:1563-1593. doi: 10.2174/092986710790979980
5. Balkwill F, Mantovani A. Inflammation and cancer: Back to Virchow? *Lancet* 2001;357:539-545. doi:10.1016/S0140-6736(00)04046-0
6. Harris R, Beebe, Alshafie. Reduction in cancer risk by selective and nonselective cyclooxygenase-2 (COX-2) inhibitors. *J Exp Pharmacol* 2012;4:91-96. doi:10.2147/JEP.S23826
7. Khan Z, Khan N, Tiwari RP, Sah NK, Prasad GB, Bisen PS. Biology of Cox-2: an application in cancer therapeutics. *Current drug targets*

2011;12:1082-1093. doi:10.2174/138945011795677764
8. Wang D, Dubois RN. The role of COX-2 in intestinal inflammation and colorectal cancer. *Oncogene* 2010;29:781-788. doi:10.1038/onc.2009.421
9. Akiyama H, Barger S, Barnum S, Bradt B, Bauer J, Cole GM, et al. Inflammation and Alzheimer's disease *Neurobiol Aging* 2000;21:383-421. doi:10.1016/S0197-4580(00)00124-X
10. 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;45:2753-2760. doi:10.1016/j.ejmech.2010.02.055
11. Qiu KM, Yan R, Xing M, Wang H-H, Cui H-E, Gong H-B, et al. Synthesis, biological evaluation and molecular modeling of dihydro-pyrazolyl-thiazolinone derivatives as potential COX-2 inhibitors. *Bioorg Med Chem* 2012;20:6648-6654. doi:10.1016/j.bmc.2012.09.021
12. Li Y, Chen SH, Ou TM, Tan JH, Li D, Gu LQ, et al. Syntheses and characterization of nimesulide derivatives for dual enzyme inhibitors of both cyclooxygenase-1/2 and 5-lipoxygenase. *Bioorg Med Chem* 2011;19:2074-2083. doi:10.1016/j.bmc.2011.01.043
13. Lu X, Zhang H, Li X, Chen G, Li QS, Luo Y, et al. Design, synthesis and biological evaluation of pyridine acyl sulfonamide derivatives as novel COX-2 inhibitors. *Bioorg Med Chem* 2011;19:6827-6832. doi:10.1016/j.bmc.2011.09.034
14. Xing L, Hamper BC, Fletcher TR, Wendling JM, Carter J, Gierse JK, et al. Structure-based parallel medicinal chemistry approach to improve metabolic stability of benzopyran COX-2 inhibitors. *Bioorg Med Chem Lett* 2011;21:993-996. doi:10.1016/j.bmcl.2010.12.023
15. Al-Hourani BJ, Sharma SK, Mane JY, Tuszynski J, Baracos V, Kniess T, et al. Synthesis and evaluation of 1,5-diaryl-substituted tetrazoles as novel selective cyclooxygenase-2 (COX-2) inhibitors. *Bioorg Med Chem Lett* 2011;21:1823-1826. doi:10.1016/j.bmcl.2011.01.057
16. Abdur Rahim M, Praveen Rao PN, Bhardwaj A, Kaur J, Huang Z, Knaus EE. Isomeric acetoxy analogs of celecoxib and their evaluation as cyclooxygenase inhibitors. *Bioorg Med Chem Lett* 2011;21:6074-6080. doi:10.1016/j.bmcl.2011.08.053
17. Kaur J, Bhardwaj A, Huang Z, Knaus EE. N-1 and C-3 substituted indole Schiff bases as selective COX-2 inhibitors: Synthesis and biological evaluation. *Bioorg Med Chem Lett* 2012;22:2154-2159. doi:10.1016/j.bmcl.2012.01.130
18. Zarghi A, Reihanfard H, Arfaei S, Daraei B, Hedayati M. Design and synthesis of new 1,2-diaryl-4,5,6,7-tetrahydro-1H-benzo[d] imidazoles as selective cyclooxygenase (COX-2) inhibitors. *Med Chem Res* 2012;21:1869-1875. doi:10.1007/s00044-011-9709-y
19. Liang Y, Xu Q, Li H, Cao D. *Support Vector Machines and Their Application in Chemistry and Biotechnology*, CRC press, FL, USA;2011.
20. Soltani S, Haghaei H, Shayanfar A, Vallipour J, Asadpour Seynali K, Jouyban A. QSBR Study of Bitter Taste of Peptides: Application of GA-PLS in Combination with MLR, SVM, and ANN Approaches. *BioMed Res Int* 2013;501310. doi:10.1155/2013/501310
21. Hernández N , Kiralj R, Ferreira MMC, Talavera I. Critical comparative analysis, validation and interpretation of SVM and PLS regression models in a QSAR study on HIV-1 protease inhibitors. *Chemom Intell Lab Syst* 2009;98:65-77. doi:10.1016/j.chemolab.2009.04.012
22. Fernandez M, Caballero J, Fernandez L, Sarai A. Genetic algorithm optimization in drug design QSAR: Bayesian-regularized genetic neural networks (BRGNN) and genetic algorithm-optimized support vectors machines (GA-SVM). *Mol Diver* 2011;15:269-289. doi:10.1007/s11030-010-9234-9
23. Tresadern G, Cid JM, Trabanco AA. QSAR design of triazolopyridine mGlu2 receptor positive allosteric modulators. *J Mol Graph Model* 2014;53:82-91. doi:10.1016/j.jmgm.2014.07.006
24. Jahangiri R, Soltani S, Barzegar A. A Review of QSAR Studies to Predict Activity of ACE Peptide Inhibitors.*Pharm Sci* 2014;20:122-129.
25. Darnag R, Minaoui B, Fakir M. Application of support vector machines for prediction of anti-HIV activity of TIBO Derivatives. *Chem Mater Res* 2013;3:119-124. doi.org/10.1016/j.ejmech.2010.01.002
26. Gaspar P, Carbonel J, Oliveira JL. On the parameter optimization of Support Vector Machines for binary classification. *J Integr Bioinform* 2012;9:201. doi: 10.2390/biecoll-jib-2012-201.
27. Huang CL, Wang CJ. A GA-based feature selection and parameters optimizationfor support vector machines. *Expert Syst appl* 2006;31:231-240. doi:10.1016/j.eswa.2005.09.024