International Journal of Pharmacy and Pharmaceutical Sciences

ISSN- 0975-1491 Vol 10, Issue 1, 2018

Original Article

MODELING OF CARBONIC ANHYDRASE (II) INHIBITORY ACTIVITIES OF SULPHONILAMIDE SCHIFF BASES BY ARTIFICIAL NEURAL NETWORK TRAINED WITH DIFFERENT NUMERICAL TECHNIQUES

VANDANA PANDEY

Department of Chemistry, Kurukshetra University, Kurukshetra, 136119, Haryana, India
Email: vandana_p71@rediffmail.com

Received: 25 Sep 2017 Revised and Accepted: 30 Nov 2017

ABSTRACT

Objective: The aim of the present study was to develop robust linear and non-linear Quantitative Structure-Activity Relationship (QSAR) models for exploring the relationship between the structural features of a series of sulphanilamide Schiff bases and their CA (II) inhibition activities.

Methods: QSAR modeling of carbonic anhydrase (II) inhibiting activities of a series of sulphanilamide Schiff bases as a function of theoretically derived molecular descriptors calculated by Dragon software was established linearly by stepwise multiple linear regression (SW-MLR) method and non-linearly by artificial neural network (ANN) method, trained with different numerical techniques namely, Scaled conjugate gradient (SCG), quasi-Newton (BFGS), and Levenberg-Marquardt (LM) algorithm. SW-MLR method was also used to select descriptors from large descriptor pool. After the selection of variables, best selected linear model was validated by Y-randomization test. The applicability domain was assessed by the normalized mean Euclidean distance value for each compound. The prediction quality of proposed non-linear QSAR models was tested externally using validation and test set.

Results: The low value of $R^2_{average}$ = 0.214 from the Y-randomization test and no significant correlation between the selected descriptors indicates that linear model is reliable, and robust. Applicability domain analysis has also revealed that the suggested model has acceptable predictability. To explore non-linear relationship between selected descriptors and the target property, ANN approach trained with three supervised algorithms (BFGS, SCG and LM) was used. Statistical comparison of the quality of models obtained using ANN method trained with above mentioned three algorithms with SW-MLR model shows that ANN with 4-3-1 architecture and trained with LM algorithm has better predictive power as indicated by low $RMSE_{val}$ (0.11), $MAPE_{val}$ (11.95) values and high R^2_{val} (0.96) value.

Conclusion: The results of this work indicated the ANN trained with fastest Levenberg-Marqardt algorithm is a promising tool for establishing non-linear relationship between selected sulphanilamide Schiffbases and their CA (II) inhibition values.

Keywords: Sulphanilamide Schiff bases, Artificial neural network, Scaled conjugate gradient (SCG), Quasi-Newton (BFGS) and Levenberg-Marquardt (LM) algorithm

© 2018 The Authors. Published by Innovare Academic Sciences Pvt Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)
DOI: http://dx.doi.org/10.22159/ijpps.2018v10i1.22775

INTRODUCTION

Carbonic anhydrases (CAs) are a class of metalloenzymes containing Zn^{2+}ion as active site. CAs are involved in catalyzing the inter-conversion of carbonic acid and carbon dioxide to bicarbonate and H_3O+, playing an important role in several physio-pathological processes. Various clinically used drugs have been reported to possess significant CA inhibitory such as derivatives belonging to the sulphonamide, sulphamate or sulphamide families [1-6]. Carbonic anhydrase (II) is one of the fourteen forms of human α carbonic anhydrases. CA (II) is related to many diseases, including glaucoma, tumors, epilepsy and diabetes.

In chemometric research, quantitative structure-activity relationships (QSAR) studies offer the advantage of being more environment friendly than experimental approaches in molecular design and sustainable pharmacy. QSAR models are potentially important in making it possible to evaluate large number of chemicals without using conventional laboratory procedure as well as reducing number of tests on animals during drug development. Role of QSAR in accessing and reducing risks for sustainable development is well documented [7, 8]. These models are mathematical equations constructing a relationship between theoretical descriptors [9] obtained from chemical structures and biological activities. There are several approaches in QSAR modeling. Linear modeling approaches such as multiple linear regression (MLR), Partial least square (PLS) are developed to extract the maximum information from complex data matrices based on their linear behavior. In contrast, artificial neural networks (ANNs) have been used for exploring non-linear modeling and optimization when underlying mechanisms are very complex [10-12]. Generalization, convergence and complexity are some of the important factors in training of a multilayer feed-forward artificial neural network which influence its performance. These factors are highly dependent upon the type of numerical technique or algorithm used for training.

For training of multilayer feed-forward artificial neural networks, the backpropagation (BP) algorithm is generally preferred due to its simplicity. Among various methods of training, second order methods include Scaled conjugate gradient (SCG), quasi-newton (BFGS), and Levenberg-Marquardt (LM) algorithm [13, 14].

Several QSAR studies of CA (II) inhibition using theoretical and physiochemical descriptors of various groups of molecules have been reported [15-17]. In this context, QSAR studies on, CA (II) inhibition activities of Schiff bases sulfanilamides based primarily on topological descriptors were reported by various researchers [18-20].

There is no report on the use of ANN trained with different algorithms in the QSAR modeling of CA(II) inhibition activities of Schiff bases sulphanilamide's.

Therefore, the purpose of present study was to examine the accuracy of ANN trained with different numerical techniques and to statistically compare the previously reported results with the results obtained from linear and non-linear modeling for better prediction of CA(II) inhibition activity in terms of log KCA(II).

MATERIALS AND METHODS

Dataset

The series of 35 Schiff base sulphanilamide compounds (fig. 1) and their CA (II) inhibition activities were taken from the work published by Supernan and Clare [17, 19]. The 3D structures of the compounds in the form of SDF files were generated from the Pubchem database using its various utilities. The activity data were first converted into logarithmic scale and then values of logKCA (II) were used for subsequent QSAR modeling as the response variables. The molecular formula of the compounds along with their logKCA (II) inhibitory activities are presented in the table 1. Twenty-four molecules were used to build the QSAR model and the rest eleven were used as external validation and test set.

Instrumentation

E-Dragon software [21] was used to calculate theoretical descriptors. All calculations present in this work were carried out on a personal computer with a Window XP operating system. SPSS software [22] was used for SW-MLR analysis. ANN calculations were performed with Matlab [23].

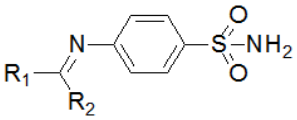


Fig. 1: Structure of representative sulphanilamide compounds

Generation and selection of molecular descriptors

E-Dragon software [21] was used to calculate a total of 979 1-, 2-and 3-D descriptors including constitutional, molecular properties, topological descriptors, connectivity indices, information indices, topological charge indices, geometrical descriptors, WHIM, 3DMorse, Getaway and RDF descriptors for each molecule. Since large number of descriptors are calculated for each molecule. The calculated descriptors were first analyzed to check the existence of constant and near constant variables, which were removed. Furthermore, the correlation of the descriptors with each other and with target property (logKCAII) was examined in order to decrease the redundancy.

SW-MLR method

Stepwise multiple linear regression (SW-MLR) method was applied for each category of 1-, 2-and 3-D descriptors to get reduced pool of descriptors. In stepwise technique one parameter at a time is added to a model and always in the order of most significant to least significant in terms of F-test values [24, 25]. Statistical parameters were calculated subsequently for each step in the process, so the significance of the added parameter could be verified. The goodness of the correlation is tested by the regression coefficient (R^2), the standard error of the estimate (SEE) and the F-test [26]. Finally, twenty-five best selected descriptors from various categories were further subjected to stepwise multiple linear regression to get most significant descriptors.

ANN method

In the present work Matlab software package was used for implementation of three layered fully connected, feed forward computational neural network. For further improvement of performance in comparison with that of SW-MLR method, ANN approach was used for mapping non-linear relationship between theoretical descriptors selected from SW-MLR method and logKCA (II) inhibitory activities, In ANN approach, each neuron in any layer is fully connected with the neurons of adjacent layers.

The architecture of ANN is such that (i) number of neurons in the input layer is equal to number of descriptors selected from SW-MLR method (ii) the number of hidden neurons is optimized and (iii) one neuron is placed in the output layer whose output is the target activity for each molecule. The input vectors and output values were preprocessed so that that they fall in the range [0.1-0.9]. ANN with standard numerical optimization techniques including Scaled conjugate gradient (SCG), quasi-Newton (BFGS), and Levenberg-Marquardt (LM) were applied for training of the network

Method validation

The predictive power of QSAR methods is evaluated internally as well as externally using validation and test set as recommended by Golbraikh and Tropsha [27]. For internal validation, Y-randomization technique was performed to check robustness of the model. External validation was performed by dividing the data set into training, validation and test set randomly in such a way that ratio of vectors for training, validation and testing were 0.7, 0.15 and 0.15 respectively. As a result, 24, 6 and 5 Schiff base sulphanilamide compounds respectively chosen from the data set of 35 molecules for training, validation and test set. Finally, the performance of the prediction system was evaluated using the following common statistics: coefficient of determination (R^2), root mean of squared errors (RMSE) and mean absolute per cent error (MAPE). The Applicability Domain (AD) is assessed by the normalized mean distance values for each compound.

Table 1: Experimental, calculated logKCA (II) values and normalised mean distance values of sulphanilamide compounds

S. No.	R_1	R_2	LogKCa(II)[a]	SW-MLR[b]	4-3-1 (LM)[b]	4-5-1 (SCG)[b]	4-6-1 (BFG)[b]	N. M. D.
1	Phenyl	H	1.4314	0.949	1.154	1.212	1.15	0.001
2	2-Nitrophenyl	H	1.3222	0.882	0.961	1.004	1.118	0.264
3	4-Hydroxyphenyl	H	1.2788	1.200	1.023	1.013	1.056	0.206
4	4-Methoxyphenyl	H	1.2788	1.160	0.873	0.767	1.035	0.000
5	4-Dimethylaminophenyl	H	0.9031	0.653	0.963	0.957	0.731	0.072
6	4-Cyanophenyl	H	1.0414	1.078	1.106	1.178	1.202	0.088
7	3,4-Dimethoxyphenyl	H	0.4771	0.871	0.798	0.752	0.831	0.018
8	3-Methoxy-4-acetoxyphenyl	H	1	0.680	0.466	0.486	0.44	0.000
9	2,3-Dihydroxy-5-formylphenyl	H	0.301	0.622	0.592	0.515	0.604	0.565
10	2-Hydroxy-3-methoxy-5-formylphenyl	H	0.4771	0.757	0.486	0.408	0.496	0.039
11	3-Methoxy-4-hydroxy-5-bromophenyl	H	0.6021	0.730	0.738	0.728	0.744	0.110
12	5-Methyl-2-furyl	H	0.6021	0.544	1.022	1.105	0.935	0.176
13	Pyrol-2-yl	H	0.301	0.855	0.869	0.779	0.904	0.046
14	Imidazol-4(5)-yl	H	1.0792	0.921	0.824	0.693	0.863	0.059
15	2-Pyridyl	H	0.9542	0.736	1.2	1.387	1.051	0.001
16	4-Pyridyl	H	0.699	1.099	1.132	1.126	1.139	0.044
17	4-Methoxystyryl	Me	−0.9208	-0.864	-0.325	-0.634	-0.256	0.278
18	4-Dimethylamino styryl	Me	−1.0000	-1.132	-0.332	-0.623	-0.266	0.701
19	3,4,5-Trimethoxy styryl	Me	−0.6198	-0.235	-0.067	-0.156	-0.073	1.000
20	4-Methoxy styryl	Ph	0.1761	0.144	-0.091	-0.112	-0.085	0.052
21	3,4,5-Trimethoxy styryl	Ph	0.3711	0.239	0.097	0.418	-0.056	0.304

22	3,4,5-Trimethoxy styryl	4-$MeOC_6H_4$	0.1038	-0.047	0.216	0.196	-0.078	0.651
23	3-Nitrostyryl	4-$MeOC_6H_4$	−0.1871	-0.051	-0.022	0.087	0.106	0.050
24	3,4,5-Trimethoxy styryl	4-$NH_2C_6H_4$	−0.0706	-0.191	-0.174	-0.276	-0.092	0.234
	External set							
1	2-Hydroxyphenyl*	H	1.6128	-0.106	1.504	1.819	1.755	0.218
2	4-Nitrophenyl*	H	0.699	1.381	0.873	0.767	1.035	0.087
3	3,4,5-Trimethoxyphenyl*	H	0.4771	0.818	0.496	0.526	0.549	0.326
4	3-Pyridyl*	H	0.9031	1.266	1.067	0.899	1.018	0.072
5	Styryl*	Ph	−0.2518	-0.013	-0.246	-0.528	-0.156	0.219
6	3,4,5-Trimethoxy styryl*	4-PhC_6H_4	0.3945	0.106	0.456	0.223	-0.077	0.033
7	4-Chlorophenyl**	H	1.4472	1.219	0.992	0.927	1.051	0.013
8	3-Methoxy-4-hydroxyphenyl**	H	0.9031	0.781	0.915	0.929	0.896	0.203
9	2-Furyl**	H	0.699	0.930	0.847	0.735	0.885	0.150
10	Styryl**	Me	−0.4089	-1.013	-0.345	-0.0624	-0.362	0.012
11	4-Dimethylaminostyryl**	Ph	0.2279	-0.359	-0.007	-0.044	0.026	0.222

a = observed logKCA(II) activity; b= calculated logKCA(II) acitivity by different methods; * = compounds in validation set; ** = compounds in test set, N. M. D.= normalized Mean Distance

RESULTS AND DISCUSSION

Linear approach

The reduced data set containing twenty-five descriptors was further subjected to stepwise regression analysis in order to select a limited number of descriptors significantly contributing to the prediction of logKCA (II) inhibitory activity of Schiff bases of sulfanilamides. As the aim was to select only 4 or 5 descriptors, considering the number of compounds in the data set was 35. Finally, four descriptors namely JGI8, Mor20u, R7u$^+$ and G1s showing high accordance with inhibitory activity logKCA (II) were selected out and activities of the 24 compounds used in training set were fitted.

Here JGI8 = Mean topological charge index of order 8 (2D autocorrelations), Mor20u = signal 20/unweight (3D-MoRSE descriptors), R7u+= R maximal autocorrelation of lag 7/unweight (GETAWAY descriptors) andG1s =1st component symmetry directional WHIM index/weighted by atomic electro topological states (WHIM descriptor). The different DRAGON classes, to which these descriptors belong, are briefly described as follows. 2D autocorrelations descriptors are spatial autocorrelations, calculated from molecular graph. 3D MoRSE descriptors are very flexible 3D structure encoding framework for cheminformatics. GETAWAY descriptors are calculated from the leverage/geometry matrix obtained by centered atomic coordinates and Weighted Holistic Invariant molecular (WHIM) descriptors are geometrical descriptors based on statistical indices built to capture relevant 3D information regarding molecule. A correlation matrix was obtained among all the descriptors used, in final selection of the model because regression equation is useless if descriptors are highly correlated. It can be seen from the correlation matrix (table 2), there is no significant correlation between the selected descriptors.

Table 2: Correlation matrix for the inter-correlation of selected descriptors

	JGI8	Mor20u	R7u+	G1s
JGI8	**1**			
Mor20u	0.47604	1		
R7u+	0.63326	0.1867	1	
G1s	0.73423	0.44847	0.42863	1

In the present work, these descriptors were used for construction of both linear and nonlinear models. The best selected model obtained by SW-MLR method contained four descriptors resulted in a strong correlation to experimental pIC_{50} values (R^2 =0.83, S=0.303 R^2_{adj}= 0.80). As results suggest, 83% of variance in the training data matrix could be explained by selected four descriptors. The F ratio in the Annova table shows that independent variables statistically significantly predict the dependent variable F(4,19)=24.348, p<.005 suggest the regression model is good fit of data. As for as collinearity statistics concern, the value of tolerance ranges from 0.20-0.76 which is>0.1 and VIF ranges from 1.3-4.9 which is<10. Selecting the four descriptors as independent variables, parameters and unstandardized coefficient values of the stepwise regression multi parametric model are depicted in the table 3.

Table 3: The values of coefficients and collinearity statistics for the SW-MLR model

Model	Unstandardized coefficients		Standardized coefficients	Collinearity statistics	
	B	Std. error	Beta	Tolerance	VIF
(Constant)	3.903	2.053			
JGI8	102.928	26.343	.806	.202	4.957
Mor20u	.626	.163	.408	.764	1.309
R7u+	42.645	13.227	.362	.683	1.464
G1s	-35.241	12.829	-.514	.246	4.070

The results of the QSAR modeling by stepwise multiple linear regression method hinted the predominance of 2D topological (JG18) and 3D GETAWAY (R7u+) descriptor over other descriptors in the model influencing the logKCA (II) inhibitory activity of the studied compounds due to their relatively high numerical coefficient. In order to ensure the robustness of the proposed model,

Y-randomization test was performed by generating fifty random models, resulted quite low average R^2 = 0.214, which confirmed that the internal validation of the proposed QSAR model is quite robust. But external validation parameters of the proposed SW-MLR model were not satisfactory. In the previous QSAR study of the set of 35 sulfanilamide Schiff bases [19], numerous models with molecular descriptors and indicator variables were tested (32 models with up to seven parameters). In model 32 with seven parameters, a value for R^2= 0.879 was obtained. In the present study, which has only four parameters, it is evident that the results for this set of compounds are quite satisfactory.

Domain of applicability

One of the OECD principles for model validation requires defining the applicability domain (AD) for the QSAR model for reliable prediction. Several AD approaches have been already proposed and classified into four major categories i,e, range based method, geometric method, distance based method and probability density method [28]. Distance based approaches calculate the distance of the query compound from defined points within the descriptor space of training data. Some commonly used distance measures in the QSAR studies include Mahalanobis, Euclidean and city block distance [29, 30]. In the present paper, AD is verified by Euclidean based approach. It is based on mean distance scores calculated by distance norms. At first, normalized mean distance scores for training set compound were calculated with values (0=least diverse to1= for most diverse). Then normalized distance for test set were calculated, and those test compounds with score outside 0 to 1 ranges and said to be outside AD. The normalized mean distance scores for both training and test compounds are presented in the table 1. The results show that all compounds fall within the applicability domain of model as their normalized mean distance score fall within the range of 0 and 1.

Although, the linear model is quite satisfactory, as the results suggest, in order to improve predictive performance and to explore non-linear relationship between selected descriptors and logKCA (II) activities, ANN approach trained with different algorithms was used for mapping.

Non-linear approach

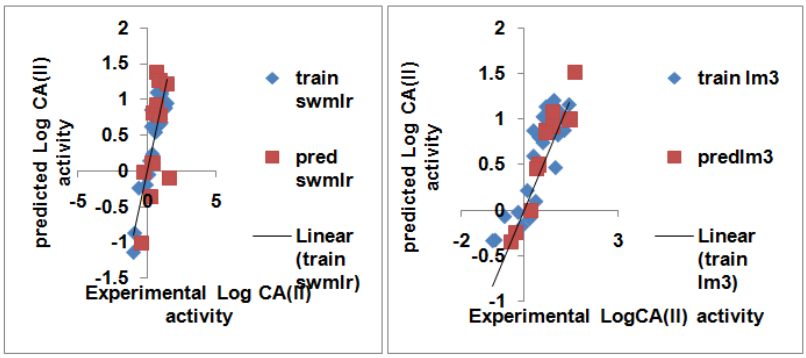
For successful training of the back propagation neural network, various factors should be considered including the number of hidden layers, the number of neurons in input and hidden layers, type of training algorithm, choice of activation function, number of epochs and learning rate. The SW-MLR selected four descriptors were used as inputs to the network, whereas, logK (CAII) inhibitory activity was used as the output value. As in most of the applications of ANN to chemistry, one hidden layer seems to be sufficient [31], a fully connected 3-layered feed forward network with back propagation pattern with mean squared error (MSE) as the performance function was used in the present study. The back propagation (BP) algorithm is a well-known method for supervised training of a multilayer feed-forward artificial neural network that adopts the gradient descent principle. However, the neural networks trained with back propagation algorithm exhibit slow learning rate. Many faster numerical techniques were proposed to speed up the convergence of the BPNN [32, 33]. Among these, scaled conjugate gradient algorithm (SCG), quasi-Newton (BFGS) algorithm, and Levenberg-Marquardt (LM) algorithm are three back propagation second order fast training algorithms that use standard numerical optimization techniques. These are well suited to neural network training where the performance function is MSE. The scaled conjugate gradient algorithm (SCG), is gradient based training algorithm. It is a very good general purpose training algorithm. Quasi-Newton (BFGS) method converges faster since it does not require calculation of second derivatives. The Levenberg-Marquardt algorithm is a variation of Newton's method [34]. It provides a balance between convergence of steepest descent and the speed of Newton's method.

In this study, above mentioned three training algorithms were evaluated for the dataset divided in three parts namely training, validation and test sets. The transfer function in the first layer was tan-sigmoid, and the output layer transfer function was linear. To select the number of nodes, the concept of ratio ρ proposed by Andrea and Kalayeh [35], was used. The number of neurons were defined from 3-6, as ρ ranges from 2-1.04. MSE value for the prediction sets were calculated by changing number of neuron in the hidden layer. Change in learning rate in the range of.001-0.1 has no considerable effect on the MSE of the prediction set in the ANN with various numbers of hidden neurons. Predicted logKCA (II) values for the external set using above mentioned three algorithms along with linear SW-MLR method are presented in the table 1.

Finally, the performance of the prediction system was evaluated using the following common statistics: Coefficient of determination (R^2), root mean of squared errors (RMSE) and mean absolute percent error (MAPE). These statistical parameters for SW-MLR and ANN trained with different algorithms are listed in the table 4.

Table 4: Statistical parameters obtained by applying SW-MLR and ANN trained with different algorithms to the validation and test set

	SW-MLR	4-3-1 (lm)	4-5-1 (scg)	4-6-1 (bfg)
R2 (val)	0.006	0.96	0.98	0.86
R2 (test)	0.9	0.86	0.88	0.88
RMSE(val)	0.796	0.11	0.161	0.253
RMSE(test)	0.4	0.23	0.28	0.216
MAPE(val)	80.59	11.95	31.09	40.41
MAPE(Test)	34.46	34.52	43.26	30.87



(a) (b)

Fig. 2: Plot of experimental vs predicted activity for the QSAR model obtained by SW-MLR method (a) and ANN (trained with L-M algorithm)(b)

Table 4 shows the superiority of ANN trained with Levenberg-Marquardt algorithm over conjugate gradient algorithm (SCG), quasi-Newton(BFGS) algorithm and SW-MLR as RMSE and MAPE values for validation and test set improved from SW-MLR to ANN (trained with LM). Finally ANN model trained with LM algorithm satisfied parameters proposed by Golbraikh and Tropsha [25] for external predictability (validation and test set). The R^2 (pred) value of the QSAR model trained with Levenberg-Marquardt algorithm is 0.907, indicating a good goodness-of-fit of the model. The calculated values of other parameters k, k', R^2 and$R'_0{}^2$ are found to be 1.01, 0.93, 0.9987 and 0.9993 respectively these values are within the range, ascertaining the fitting ability, stability, reliability and predictive ability of the proposed model.

These results show that the combination of 2d-and 3d-descriptors can be used successfully for QSAR modeling of sulfanilamide Schiff's bases. The plots of the predicted logK (CAII) inhibitory activities versus the experimental values, obtained by SW-MLR (a) and ANN trained with Levenberg-Marquardt (LM) algorithm (b), are demonstrated in the fig. 2. The values of statistical parameters as well as graphical representation demonstrate superior non-linear mapping capabilities of the ANN model which is important from the point of view of the drug design of such therapeutical agents.

CONCLUSION

Linear (SW-MLR) and non-linear (ANN trained with three different numerical techniques) QSAR modeling of sulfanilamide Schiff's base inhibitors of the physiologically relevant isozyme CAII have been carried out using various important theoretical descriptors. Numerical techniques employed in this paper include Scaled conjugate gradient (SCG), quasi-Newton (BFGS), and Levenberg-Marquardt (LM) algorithm. The results of this work indicate that use of ANN trained with second order algorithms has a great potential for determining non-linear relationship between structural features and logK(CAII) inhibitory activity sulfanilamide Schiff's bases. In particular, BFGS conjugate algorithm and Levenberg-Marquardt are the best in terms of accuracy. The predictive accuracy of linear and non-linear models, together offers the possibility of designing potent selective inhibitors.

CONFLICT OF INTERESTS

Declared none

AUTHOR CONTRIBUTION

All the work have been carried out by me.

REFERENCES

1. Supuran CT, Clare BW. Carbonic anhydrase inhibitors, Part24. A quantitative structure-activity relationship study of positively charged sulfonamide inhibitors. Eur J Med Chem 1995; 30:687–96.
2. Nishimori I, Minakuchi T, Morimoto K, Sano S, Onishi S, Takeuchi H, *et al.* Carbonic anhydrase inhibitors: DNA cloning and inhibition studies of the alpha-carbonic anhydrase from Helicobacter pylori, a new target for developing sulfonamide and sulfamate gastric drugs. J Med Chem 2006;49:2117-26.
3. Supuran CT, Clare BW. Carbonic anhydrase inhibitors-part 57: Qunatum chemical QSAR of a group of 1,3,4-thiadiazole-and 1,3,4-thiadiazoline disulfonamide with carbonic anhydrase inhibitory properties. Eur J Med Chem 1999;34:41-50.
4. Clare BW, Supuran CT. Carbonic anhydrase inhibitors, part 41. Quantitative structure-activity correlations involving kinetic rate constantsof 20 sulfonamide inhibitors from a non-congeneric series. Eur J Med Chem 1997;32:311–9.
5. Clare BW, Supuran CT. Carbonic anhydrase inhibitors. Part 61. Quantum chemical QSAR of a group of benzenedisulfonamides. Eur J Med Chem 1999;34:463–74.
6. Khadikar PV, Clare BW, Balaban AT, Supuran CT, Agrawal VK, Singh J, *et al.* QSAR prediction of CAI, CAII, CAIV inhibitory activities: relative potentialof balaban and balaban type indices. Rev Roum Chem 2006;51:703–17.
7. Raj K, Muthukumar V. Quantitative structure-activity relationship analysis of the anticonvulsant activity of erythrinine. Asian J Pharm Clin Res 2016;9:125-9.
8. Katritzky AR, Lobanov VS, Karelson M. QSPR: thecorrelation and quantitative prediction of chemical and physicalproperties from structure. Chem Soc Rev 1995;24:279–87.
9. Roy SA. QSAR studyon the schiff bases 2, 4, 6-trichlororphenylhydrazine using freely available online 2D. Asian J Pharm Clin Res 2013;6:67-70.
10. Zupan J, Gasteiger J. Neural networks for chemists: An Introduction. VCH, New York; 1993.
11. Hassoun MH. Fundamentals of artificial neural works. MIT Press: Cambridge, MA; 1995.
12. Basheer A, Hajmeer M. Artificial neural networks: fundamentals, computing, design, and application. J Microbiol Methods 2000;43:3-31.
13. Zupan J, Gasteiger J. Neural networks in chemistry and drug design, Wiley-VCH, Weinheim, Germany; 1999.
14. Bose NK, Liang P. Neural Networks, Fundamentals, McGraw-Hill, New York, NY, USA; 1996.
15. Melagraki G, Afantitis A, Sarimveis H, Igglessi-Marapouln O, Supuran CT. QSAR study on para-substituted aromatic-sulphonamides as carbonic anhydrase II inhibition using topological information indices. Bioorg Med Chem 2006; 14:1108-14.
16. Balaban AT, Basak SC, Beteringhe A, Mills D, Supuran CT. QSAR study using topological indices for inhibition of anhydrase II by sulphonamides and schiff's base. Moldivers 2004;8:401-12.
17. Supuran CT, Clare BW. Carbonic anhydrase inhibitors, Part 47. Quantum chemical quantitativestructure-activity relationships for a group of sulfanilamide Schiff base inhibitors of carbonic anhydrase. Eur J Med Chem 1998; 33:489-500.
18. Saxena A, Khadikar PV. QSAR studies on sulfanilamide schiff 's base inhibitors of carbonic anhydrase. Acta Pharm 1999;49:171–9.
19. Agrawal VK, Shrivastava S, Khadikar PV, Supuran CT. Quantitative structure-activity relationshipstudies on sulfanilamide schiff bases: CA inhibitors. Bioorg Med Chem 2003;11:5353-62.
20. Krenkel G, Castro EA. Synthesis and structure–activity relationship of novel, highly potent metharyl and methcycloalkyl cyclooxygenase-2 (COX-2) selective inhibitors. Mole Med Chem 2003;1:13-22.
21. Todeschini R. Milano chemometrics and QSPR group. Available from: http://michem.disat.unimib.it/chm/staff/staff.html, http:// www.vcclab.org/lab/edragon/. [Last accessed on 20 Aug 2017].
22. SPSS for windows Statistical package for IBM PC, SPSS Inc.
23. *MATLAB*, Math works Inc; 2017.
24. Franke R. Theoratical drug design methods, Elsevier. Amsterdam; 1984. p. 184-95.
25. Franke R, Gruska A. Chemometric methods in molecular design. Waterbeemd H Vande. Ed. VCH, Weinheim; 1995. p. 13-163.
26. Snedecor GW, Cochran WG. Statistical methods, Oxford and IBH publishing Co. Pvt. Ltd., New Delhi; 1967. p. 381-418.
27. Golbraikh A, Tropsha A. Beware of q2!. J Mol Graph Model 2002;20:269-76.
28. Sahigara F, Mansouri K, Ballabio D, Mauri A, Consonni V, Todeschini R. Comparison of different approaches to define the applicability domain of QSAR models. Molecules 2012;17:4791-810.
29. Netzeva TI, Worth A, Aldenberg T, Benigni R, Cronin MTD, Gramatica P, *et al.* Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships. The report andrecommendations of ECVAM Workshop 52. Altern Lab Anim 2005;33:155–73.
30. Jaworska J, Nikolova Jeliazkova N, Aldenberg T. QSAR applicabilty domain estimation byprojection of the training set descriptor space: a review. Altern Lab Anim 2005;33:445–59.
31. Devillrs J. Neural networks in QSAR and drug design, Academic Press: London, UK; 1996.
32. Schneider G. Neural networks are useful tools for drug design. Neural Netw 2000;13:15-60.
33. ZupanJ J, GasteigerJ. Neural networks for chemists. An Introduction, VCH Publishers, Weinheim (Germany); 1993.

34. Afantitis A, Malagraki G, Sarimveis H, Igglessi-Markopoulou O, Kollias G. A novel QSAR model for predicting the inhibition of CXCR3 receptor by 4-N-aryl-[1, 4] diazepane ureas. Eur J Med Chem 2009;44:877-84.
35. Andrea TA, Kalayeh H. Applications of neural networks in quantitative structure-activity relationships of dihydrofolate reductase inhibitors. J Med Chem 1991; 34:2824-36.