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: vandanap71@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\textsuperscript{2}_{\text{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\textsuperscript{2}_{\text{Val}}(0.96) value.

Conclusion: The results of this work indicated the ANN trained with fastest Levenberg-Marquardt 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

INTRODUCTION

Carbonic anhydrases (CAs) are a class of metalloenzymes containing Zn\textsuperscript{2+} as active site. CAs are involved in catalyzing the interconversion of carbonic acid and carbon dioxide to bicarbonate and H\textsubscript{2}O\textsuperscript{+}, 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 physicochemical 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 sulphanilamids'e.

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 \text{K}_{\text{CA}(II)}. 

DOI: http://dx.doi.org/10.22159/ijpps.2018v10i1.22775
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 Clure [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. 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].

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, 3D Morse, 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 (R²), 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 they fall in the range [0,1-0,9]. ANN with standard numerical optimization techniques including Scale 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 (R²), root mean of squared errors (RMSE) and mean absolute percent 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

<table>
<thead>
<tr>
<th>S. No.</th>
<th>RI</th>
<th>R2</th>
<th>LogKCA(II)</th>
<th>SW-MLR&lt;sup&gt;a&lt;/sup&gt;</th>
<th>4-3-1&lt;sup&gt;b&lt;/sup&gt; (LM)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>4-5-1&lt;sup&gt;c&lt;/sup&gt; (SCG)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>4-6-1&lt;sup&gt;d&lt;/sup&gt; (BFG)&lt;sup&gt;c&lt;/sup&gt;</th>
<th>N. M. D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Phenyl</td>
<td>H</td>
<td>1.4314</td>
<td>0.949</td>
<td>1.154</td>
<td>1.122</td>
<td>1.15</td>
<td>0.001</td>
</tr>
<tr>
<td>2</td>
<td>2-Nitrophenyl</td>
<td>H</td>
<td>1.3222</td>
<td>0.882</td>
<td>0.961</td>
<td>1.004</td>
<td>1.118</td>
<td>0.264</td>
</tr>
<tr>
<td>3</td>
<td>4-Hydroxysphenyl</td>
<td>H</td>
<td>1.2788</td>
<td>1.200</td>
<td>1.023</td>
<td>1.013</td>
<td>1.056</td>
<td>0.206</td>
</tr>
<tr>
<td>4</td>
<td>4-Octachlorophenyl</td>
<td>H</td>
<td>1.2788</td>
<td>1.160</td>
<td>0.873</td>
<td>0.767</td>
<td>1.035</td>
<td>0.000</td>
</tr>
<tr>
<td>5</td>
<td>4-Dimethylanilinophenyl</td>
<td>H</td>
<td>0.9031</td>
<td>0.653</td>
<td>0.963</td>
<td>0.957</td>
<td>0.731</td>
<td>0.072</td>
</tr>
<tr>
<td>6</td>
<td>4-Cyanophenyl</td>
<td>H</td>
<td>1.0414</td>
<td>1.078</td>
<td>1.106</td>
<td>1.178</td>
<td>1.202</td>
<td>0.088</td>
</tr>
<tr>
<td>7</td>
<td>2,4-Dimethoxy-6-formylphenyl</td>
<td>H</td>
<td>0.301</td>
<td>0.622</td>
<td>0.592</td>
<td>0.515</td>
<td>0.604</td>
<td>0.565</td>
</tr>
<tr>
<td>8</td>
<td>3,5-Methoxy-4-acetoxophenyl</td>
<td>H</td>
<td>1</td>
<td>0.680</td>
<td>0.466</td>
<td>0.486</td>
<td>0.44</td>
<td>0.000</td>
</tr>
<tr>
<td>9</td>
<td>2,3-Dihydroxy-5-formylphenyl</td>
<td>H</td>
<td>0.301</td>
<td>0.736</td>
<td>1.2</td>
<td>1.387</td>
<td>1.051</td>
<td>0.001</td>
</tr>
<tr>
<td>10</td>
<td>2-Hydroxy-3-methoxy-5-formylphenyl</td>
<td>H</td>
<td>0.4771</td>
<td>0.757</td>
<td>0.486</td>
<td>0.408</td>
<td>0.496</td>
<td>0.039</td>
</tr>
<tr>
<td>11</td>
<td>3-Methoxy-4-hydroxy-5-bromophenyl</td>
<td>H</td>
<td>0.6021</td>
<td>0.730</td>
<td>0.738</td>
<td>0.728</td>
<td>0.744</td>
<td>0.110</td>
</tr>
<tr>
<td>12</td>
<td>5-Methyl-2-furyl</td>
<td>H</td>
<td>0.6021</td>
<td>0.544</td>
<td>1.022</td>
<td>1.105</td>
<td>0.935</td>
<td>0.176</td>
</tr>
<tr>
<td>13</td>
<td>Pyrrol-2-y1</td>
<td>H</td>
<td>0.301</td>
<td>0.855</td>
<td>0.869</td>
<td>0.779</td>
<td>0.904</td>
<td>0.046</td>
</tr>
<tr>
<td>14</td>
<td>4-Methoxy-1(5)-yl</td>
<td>H</td>
<td>1.0792</td>
<td>0.921</td>
<td>0.824</td>
<td>0.693</td>
<td>0.863</td>
<td>0.059</td>
</tr>
<tr>
<td>15</td>
<td>3,5-Dimethyl-4(5)-yl</td>
<td>H</td>
<td>0.9542</td>
<td>0.736</td>
<td>1.2</td>
<td>1.387</td>
<td>1.051</td>
<td>0.001</td>
</tr>
<tr>
<td>16</td>
<td>4-Pyridyl</td>
<td>H</td>
<td>0.699</td>
<td>1.099</td>
<td>1.132</td>
<td>1.126</td>
<td>1.139</td>
<td>0.044</td>
</tr>
<tr>
<td>17</td>
<td>4-Methoxystyryl</td>
<td>Me</td>
<td>-0.9208</td>
<td>-0.864</td>
<td>-0.325</td>
<td>-0.634</td>
<td>-0.256</td>
<td>0.278</td>
</tr>
<tr>
<td>18</td>
<td>4-Dimethylamino styryl</td>
<td>Me</td>
<td>-1.0000</td>
<td>-1.132</td>
<td>-0.332</td>
<td>-0.623</td>
<td>-0.266</td>
<td>0.701</td>
</tr>
<tr>
<td>19</td>
<td>3,4,5-Trimethoxy styryl</td>
<td>Me</td>
<td>-0.6198</td>
<td>-0.235</td>
<td>-0.067</td>
<td>-0.156</td>
<td>-0.073</td>
<td>1.000</td>
</tr>
<tr>
<td>20</td>
<td>4-Methoxy styryl</td>
<td>Ph</td>
<td>0.1761</td>
<td>0.144</td>
<td>-0.091</td>
<td>-0.112</td>
<td>-0.085</td>
<td>0.052</td>
</tr>
<tr>
<td>21</td>
<td>3,4,5-Trimethoxy styryl</td>
<td>Ph</td>
<td>0.3711</td>
<td>0.239</td>
<td>0.097</td>
<td>0.418</td>
<td>-0.056</td>
<td>0.304</td>
</tr>
</tbody>
</table>
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 of the 25 compounds used in training set were fitted. Here JGI8 = Mean topological charge index of order 8 (2D descriptors), Mor20u = signal 20/unweight (3D-MoRSE descriptors), R7u+ = R maximal autocorrelation of lag 7/unweight (GETAWAY descriptors) and G1s = 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

<table>
<thead>
<tr>
<th>JGI8</th>
<th>Mor20u</th>
<th>R7u+</th>
<th>G1s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mor20u</td>
<td>0.47604</td>
<td>1</td>
<td>0.1867</td>
</tr>
<tr>
<td>R7u+</td>
<td>0.63326</td>
<td>0.1867</td>
<td>1</td>
</tr>
<tr>
<td>G1s</td>
<td>0.73423</td>
<td>0.44847</td>
<td>0.42863</td>
</tr>
</tbody>
</table>

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 pIC50 values ($R^2 = 0.83, S = 0.303 R_{adj}^2 = 0.80$). As results suggest, 83% of variance in the training data matrix could be explained by selected four descriptors. The F ratio in the Anova table shows that independent variables statistically significantly predict the dependent variable F(4,19) = 24.348, p < 0.005 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

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized coefficients</th>
<th>Standardized coefficients</th>
<th>Collinearity statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. error</td>
<td>Beta</td>
</tr>
<tr>
<td>(Constant)</td>
<td>3.903</td>
<td>2.053</td>
<td>0.868</td>
</tr>
<tr>
<td>JGI8</td>
<td>102.928</td>
<td>26.343</td>
<td>0.408</td>
</tr>
<tr>
<td>Mor20u</td>
<td>0.626</td>
<td>0.163</td>
<td>0.362</td>
</tr>
<tr>
<td>R7u+</td>
<td>42.645</td>
<td>13.227</td>
<td>0.362</td>
</tr>
<tr>
<td>G1s</td>
<td>-35.241</td>
<td>12.829</td>
<td>-0.514</td>
</tr>
</tbody>
</table>

The results of the Q SAR modeling by stepwise multiple linear regression method hinted the predominance of 2D topological (JGI8) 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 compounds were calculated with values $0$=least diverse to 1= 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 $\rho$ proposed by Andrea and Kalayeh [35], was used. The number of neurons were defined from 3-6, as $\rho$ 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 0.01-0.1 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](image)

![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)](image)
Table 4 shows the superiority of ANN trained with Levenberg-Marquardt algorithm over conjugate gradient algorithm (SCG), quasi-Newton (BFGS) algorithm and SW-MLR algorithm. Table 1 shows 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 Golbanih and Tropsha [25] for external predictability (validation and test set). The R² (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' R² and r² 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 capability of 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 bases 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

3. Supuran CT, Clare BW. Carbonic anhydrase inhibitors-part 57: Quantum chemical QSAR of a group of 1,3,4-thiadiazolin-5-yl 1,3,4-thiadiazolinyl sulfonamides with carbonic anhydrase inhibitory properties. Eur J Med Chem 1999;34:41-50.
22. SPSS for windows Statistical package for IBM PC, SPSS Inc.
23. MATLAB, Math works Inc; 2017.