Journal of the Korean Chemical Society
2015, *Vol. 59, No. 6*
Printed in the Republic of Korea
http://dx.doi.org/10.5012/jkcs.2015.59.6.483

Quantum Chemical Studies of Some Sulphanilamide Schiff Bases Inhibitor Activity Using QSAR Methods

Elham Baher[†], Naser Darzi[‡,*], Ali Morsali[‡], and Safar Ali Beyramabadi[‡]

[†]*Department of Chemistry, faculty of science, Golestan university, Gorgan, Iran, P.O. Box 49138-15739, Iran*
[‡]*Department of Chemistry, Mashhad Branch, Islamic Azad University, Mashhad, Iran, P.O. Box 91735-413, Iran.*
[*]*E-mail: n_darzi138@yahoo.com*
(Received March 27, 2015; Accepted September 4, 2015)

ABSTRACT. The different calculated quantum chemical descriptors by DFT method were used for prediction of some sulphanilamide Schiff bases inhibitor activity as a binding constant (log K). Multiple linear regression (MLR) and artificial neural network (ANN) were employed for developing the useful quantitative structure activity relationship (QSAR) model. The obtained results presented superiority of ANN model over the MLR one. The offering QSAR model is very easy to computation and Physico-Chemically interpretable. Sensitivity analysis was used to determine the relative importance of each descriptor in ANN model. The order of importance of each descriptor according to this analysis is: molecular volume, molecular weight and dipole moment, respectively. These descriptors appear good information related to different structure of sulphanilamide Schiff bases can participate in their inhibitor activity.

Key words: Sulphanilamide Schiff bases, QSAR, DFT

INTRODUCTION

The carbonic anhydrases are part of a group of enzymes that catalyze the conversion of carbon dioxide in water to bicarbonate and protons.[1] The zinc ion in carbonic anhydrases enzyme structure is its active site and for this reason they are classified as metalloenzymes. One of operations of the enzyme is to interconvert carbon dioxide and bicarbonate in relation to maintain acid-base balance in blood, stomach and other tissues.[2] Also, CO_2 transmission of tissues to red blood cells and of red blood cells to the lungs is other operation of this enzyme. Changing the operation of this enzyme disrupts the normal functions of body. Inhibitor drugs of carbonic anhydrase enzymes can have different effects. For example, inhibitors of this enzyme in the kidney acts as a diuretic and type two enzyme inhibitors in the treatment of glaucoma (with mechanism to reduce the synthesis of aqueous humor) function. Also, this enzyme in the stomach by the stomach HCL control is to reduce of heartburn.[1,3] Among these, sulphanilamide Schiff bases are important group of drugs that are capable to inhibiting carbonic anhydrase. Carbonic anhydrase inhibitors are down through their connection with anion SO_2NH with ZN^{-2} of enzymes. They also provide drugs such as antibacterial, antimicrobial, anticonvulsant, anti-inflammatory, anti-cancer and anti-tumor is also used.[3]The attempt to improve the quality of carbonic anhydrase inhibitor drugs needs to examine a large number of experiments. But these experiments are generally expensive and time consuming. Among these, quantitative structure activity relationship (QSAR) method offers an encouraging method for prediction of inhibitor activity based on the extracted descriptors from the molecular structures. The advantage of this technique is dependent to only the knowledge of chemical structures and is not need to any experimental data. Up to now, some of researchers have tried to carry out these studies for new drug design.[4-8] One of these researcher; E. Eroglu presented a QSAR model based on a group of sulfonamide Schiffbase inhibitors of carbonic anhydrase enzyme by using Codessa Pro methodology.[4] In another study by E. Eroglu et al., they developed QSAR study on some bioactive sulfonamide compounds For all the compounds, initial geometry optimizations were down using the molecular mechanics force fields.[5] The study conducted by V.K. Agrawal *et al.* presented QSAR model just only using distance-based topological indices.[6] S. Singh *et al.* offered 3D-QSAR CoMFA analyses on inhibitors of the enzyme Rv3588c (including some diazenyl benzene sulfonamides). In their model, electrostatic agents shown important role.[7]

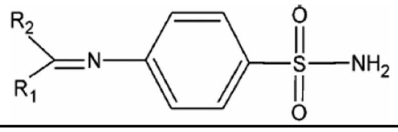
As well known, two factors are very important for the prediction ability of QSAR models; descriptors with sufficient information about structures and the modeling method (simplicity and reasonability). Despite of the various researches has been done on carbonic anhydrase inhibi-

tors, due to the importance of this topic, we tried to create a simple and reasonable QSAR model that can provide more efficient than previous models in qualitative and quantitative prediction up to help in the development and production of new carbonic anhydrase inhibitor drugs. Two modeling methods; multiple linear regression (MLR) and artificial neural network (ANN) were used. Meanwhile, high accuracy in calculating of quantum chemistry descriptors with method DFT can help effectively to achieve this goal.

EXPERIMENTAL

Experimental values of some sulphanilamide Schiff bases inhibitor activity as a binding constant (log K) was used.[9] This data set consists of 35 derivatives of sulphanilamide Schiff bases. The structure and values of log K for these compounds were shown in *Table* 1. Quantum chemical descriptors have enough and useful information about feature of molecules. Several quantum chemical descriptors

***Table* 1.** The structure, experimental values and Predicted values of sulphanilamide schiff bases inhibitor activity (log K)



No	R1	R2	Exp.	MLR	ANN	Residual-ANN
1	Phenyl	H	1.431	1.280	1.1762	0.255
2	2-Hydroxyphenyl	H	0.1697	1.006	0.586	-0.416
3	2-Nitrophenyl	H	1.322	1.012	1.177	0.145
4	4-Chlorophenyl	H	1.447	1.172	1.167	0.280
5	4-Hydroxyphenyl	H	0.2022	1.047	0.177	0.025
6	4-Methoxyphenyl	H	1.279	0.808	1.177	0.102
7	4-Dimethylaminophenyl	H	0.903	0.796	0.898	0.005
8	4-Nitrophenyl	H	0.5857	1.228	0.697	-0.111
9	4-Cyanophenyl	H	1.041	1.049	1.038	0.0026
10	3-Methoxy-4-hydroxyphenyl	H	0.903	0.880	0.901	0.0016
11	3,4-Dimethoxyphenyl	H	-0.302	0.697	-0.589	0.286
12	3-Methoxy-4-acetoxyphenyl	H	1.000	0.762	0.989	0.011
13	2,3-Dihydroxy-5-formylphenyl	H	0.396	1.061	0.545	-0.149
14	2-Hydroxy-3-methoxy-5-formylphenyl	H	0.477	0.957	0.590	-0.113
15	3,4,5-Trimethoxyphenyl	H	0.477	0.590	0.590	-0.113
16	3-Methoxy-4-hydroxy-5-bromophenyl	H	0.602	0.776	0.589	0.013
17	2-Furyl	H	0.699	1.077	0.698	0.001
18	5-Methyl-2-furyl	H	0.602	0.721	0.602	0.000
19	Pyrol-2-yl	H	0.301	0.951	0.589	-0.288
20	Imidazol-4(5)-yl	H	1.079	1.066	1.078	0.001
21	2-Pyridyl	H	0.954	0.947	0.953	0.001
22	3-Pyridyl	H	0.599	1.339	0.903	-0.303
23	4-Pyridyl	H	0.699	1.087	1.177	-0.478
24	Styryl	Me	0.409	0.857	0.589	-0.180
25	4-Methoxystyryl	Me	-0.921	0.483	-0.689	-0.232
26	4-Dimethylamino styryl	Me	-1.000	0.568	0.089	-1.089
27	3,4,5-Trimethoxy styryl	Me	-0.620	0.435	-0.589	-0.031
28	Styryl	Ph	-0.252	0.688	0.589	-0.841
29	4-Methoxy styryl	Ph	0.176	0.077	0.589	-0.413
30	4-Dimethylaminostyryl	Ph	0.228	0.228	0.549	-0.321
31	3,4,5-Trimethoxy styryl	Ph	0.371	0.258	0.589	-0.218
32	3,4,5-Trimethoxy styryl	4-$MeOC_6H_4$	0.104	0.676	0.687	-0.583
33	3-Nitrostyryl	4-$MeOC_6H_4$	-1.163	0.089	0.043	-1.206
34	3,4,5-Trimethoxy styryl	4-$NH_2C_6H_4$	-0.071	0.026	0.589	-0.660
35	3,4,5-Trimethoxy styryl	4-PhC_6H_4	0.395	0.104	0.189	0.206

***Table* 2.** The correlation matrix for the selected descriptors

	dipole	volume	MW
dipole	1	0.605	0.616
volume		1	0.636
MW			1

can be generated by Gaussian (09).[10] Gaussian program can be obtained these descriptors based on the molecular structure with high accuracy. Thus, the structure of components is drawn with the Chemoffice software and was prepared in a suitable file for the Gaussian program. DFT method was used with B3lyp function[11,12] and 6-311++G (d, p) basis set for geometrical optimization. After descriptor generation, screening and selection procedure should be down. Then, by final descriptors, MLR and ANN models were developed for investigation of relation between log K and characteristics of each sulphanilamide Schiff bases. For this order, the data set was divided to two groups; training (28) and prediction sets (7). Y-ranking method was employed for division of these groups.

One of the importance stages to determine the effective factors on inhibitor activity of sulphanilamide Schiff bases is screening of generated quantum chemical descriptors because of removing inappropriate descriptors. For this reason, descriptors with high correlation and constant values were eliminated. Finally, 11 descriptors were reminded for selecting step. For descriptor selection, SPSS (V.19)[13] software in backward mode was employed and then 3 descriptors were selected for model generation. The correlation matrix for the selected descriptors is in *Table* 2. According to this table, there is no correlation between these descriptors.

RESULTS AND DISCUSSION

Linear Model

At first linear relationship between log K and descriptors was investigated. For this reason, reminded descriptors in pre-selected analysis were used for developing MLR model. For model developing, training set was used and then prediction set was employed for evaluation the mentioned model. Thus, by using backward-MLR, 7 models were created. The best model with high R-value, fewer standard errors (SE) and greater F-value was chosen of all obtained models. Break-point procedure can help to find these conditions. This analysis was performed by plotting the squared correlation coefficient and SE values of the obtained models versus the number of descriptors involved in each model. The break-point plot was displayed in *Fig.* 1.

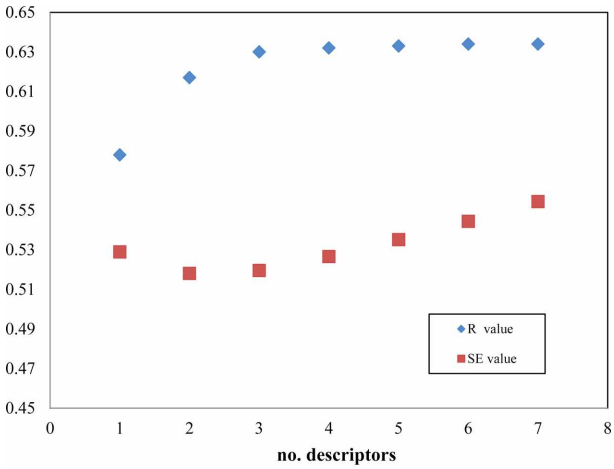


***Figure* 1.** Break-point plot.

As can be seen this figure, the model with 3 descriptors has the best condition. The descriptors that entered in this model are: molecular weight (MW), dipole moment (dipol) and molecular volume (V). Equivalent (1) presents obtained linear model.

Log K=2.034 (±0.457) – 0.057 (±0.053) dipol −0.003 (±0.002) MW −0.008 (±0.004) V;

$$R = 0.630 \quad SE = 0.519 \qquad (1)$$

Also, Q^2 and SPRESS were calculated after the leave-7-out cross-validation to evaluate credibility and robustness of this model.[14] The obtained results were presented in *Table* 3. The prediction values of log K based on MLR-cross-validation are in *Table* 1.

Artificial Neural Network Model

Artificial neural network (ANN) is mathematical system that was simulated from biological neural networks.[15] ANN consist nodes or neurons as a processing element in some layers. At present work, the back-propagation network was employed to investigation of nonlinear relationship in the case of sulphanilamide Schiff bases inhibitor activity and selected descriptors. To training the network, the biases and weights between the layers in a direction to minimize the output errors were changed.[16–18] The ANN programs

***Table* 3.** Statistical parameters for MLR and ANN models

	MLR		ANN	
Sets	Training	Prediction	Training	Prediction
R	0.663	0.523	0.829	0.543
SE	0.469	0.699	0.380	0.751
Q^2	0.583		0.67	
SPRESS	0.671		0.689	

***Table* 4.** Architecture of ANN

Transfer function	Sigmoidal
No. of hidden layer nods	7
Weight learning rate	0.4
Bias learning rate	0.5
Momentum	0.6
No. of input layer nods	3
No. of output layer nods	1

were written in FORTRAN 77. A transfer function of three-layer network is a sigmoid. The selected descriptors by backward-MLR and log K were used as inputs and outputs layer, respectively. The inputs and outputs values of ANN were normalized between 0.1 and 0.9. The initial weights were selected randomly between −0.3 and 0.3. The number of nodes in the hidden layers, learning rate, and momentum would be optimized before training the network. In order to evaluate the performance of the ANN, the standard error of training (SET) and the standard error of prediction (SEP) were calculated. The training iteration was stopped at overtraining point, where SEP is started to increase. Then the trained network was used to calculation the log K values of prediction set. The optimized values of the number of nodes in the hidden layers, learning rate, and momentum and ANN characteristics are given in *Table* 4.

In order to further investigate the credibility of obtained ANN model leave-7-out cross validation method was performed and obtained Q^2 and SPRESS was evaluated. The results of cross-validation test were shown in the *Table* 3. The prediction values of log K based on ANN-cross-validation are in *Table* 1.

Models Review

Predicted residual sum of square error (SPRESS) appears to good estimate of the real predictive error of the models. Also, parameter Q^2 is important as the predictive power. With respect to *Table* 3, SPRESS value of ANN and MLR model is relatively equal. Therefore, this parameter cannot compare real predictive error of the obtained models. But referring to Q^2, Q^2 of ANN model is greater than MLR model that indicates the obtained ANN model has good predictive power. As well, other obtained statistical parameters of these models in *Table* 3 such as R-value and SE indicate ANN model is successful than MLR model to prediction of log K as the inhibitor activity of sulphanilamide Schiff bases. For further investigation on ANN model, the plot of ANN predicted values of log K against experimental values of them was shown in *Fig.* 2. This figure indicates the agreement between the predicted and

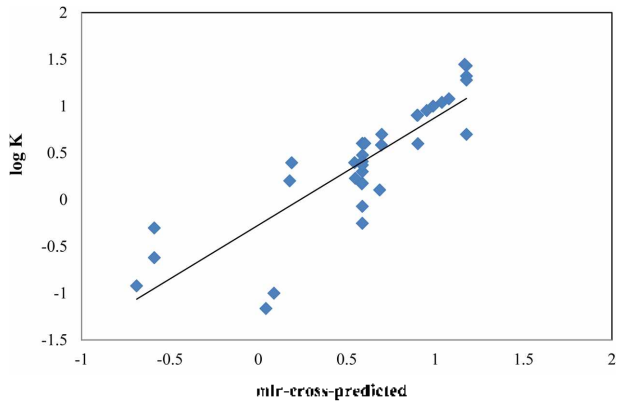


***Figure* 2.** ANN predicted values of log K versus Experimental values of log K plot.

experimental values of log K. Also, the residual values between ANN predicted and experimental values of log K were in the *Table* 1.

Sensitivity Analysis And Descriptors Interpretation

Generally drug molecules operate on particular targets at the cellular level and affect therapeutic operation based upon binding to receptors that modify the cellular activity. Whereof log K is the logarithm of sulphanilamide Schiff base binding constant in to target cell, furthermore, by interpreting descriptors in the developed model, it is possible to gain some vision into the factors that are likely to govern the log K of inhibitor activity of sulphanilamide Schiff bases. Here, a brief interpretation of these factors is performed based on the results of sensitivity analysis. Sensitivity analysis is the method to determine the relative importance of each descriptor in ANN model. The procedure of this approach is based on the sequential removal of descriptor by zeroing the related connection weights in the first layer of the ANN.[19] For each sequentially zeroing, the root mean square error of prediction set (RMSEP) as the prediction error of this network was calculated. Normally, RMSEP value increases in this way. Then, differences between RMSEP and root mean square error of established ANN (RMSE) was calculated as DRMSE. Each descriptor with greater value of DRMSE is more important. The DRMSE values are shown for each descriptor in *Fig.* 3. According to the *Fig.* 3, the order of importance of each descriptor in ANN model is fallowing:

V > MW > dipol

V: The molecular volume is specified as the volume of area within which a molecule is constrained by its neighbors. It can be computed both of experimental observation and theoretical calculated. The calculate value of this

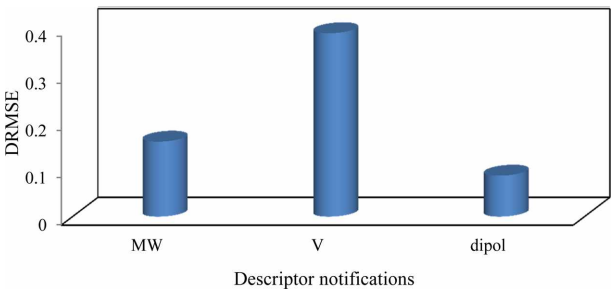


***Figure* 3.** Importance of descriptors by Sensitivity analysis.

parameter by DFT method has very good veracity. The molecular volume characterized the bulk of compound. Furthermore, it seems that the molecular size of sulphanilamide Schiff bases can be important factor on their inhibitor activity.[20]

MW: Molecular Weight with middle mean effect is related to molecular size. This descriptor is atom-type sensitive. The fitting of this descriptor in achieved QSAR model appears correlation between biological activity and the size of sulphanilamide Schiff bases, over again.[21]

Dipol: Dipole Moment indicates information about the charge distribution in molecule. A molecule has a dipole when separation appends in molecular charge. The athurity of the dipole moment depends on the difference in the electronegativity of the atoms in the molecule. This parameter can be important whereas it represents the dependency of interaction between sulphanilamide Schiff base drugs and target issue.[20]

CONCLUSION

In view of the above obtained results, the ANN model has superiority over the MLR one. Therefore, the achieved ANN model can be able to estimate the log K of sulphanilamide Schiff bases as the inhibitor activity, properly. Also, advantages of the obtained ANN model are simplicity and reasonability. The selected different quantum chemical descriptors are very easy to computation and Physico-Chemically interpretable in the developed ANN model. These descriptors appear good information related to different structure of sulphanilamide Schiff bases can participate in their inhibitor activity.

Acknowledgments. Publication cost of this paper was supported by the Korean Chemical Society.

REFERENCES

1. Supuran, C. T.; Winum, J. Y. (Eds.). *Drug Design of Zinc-Enzyme Inhibitors: Functional, Structural, and Disease Applications*; John Wiley & Sons: U.S.A., 2009.
2. Badger, M.; Price, G. *Annu. Rev. Plant Physiol. Plant Mol. Bio.* **1994**, *45*, 369.
3. Supuran, T. C.; Scozzafava, A.; Casini, A. *Medicinal Research Reviews*. **2003**, *23*, 146.
4. Eroglu, E. *Int. J. Mol. Sci.* **2008**, *9*, 181.
5. Eroğlu, E.; Türkmen, H.; Güler, S.; Palaz, S.; Oltulu, O. *Int. J. Mol. Sci.* **2007**, *8*, 145.
6. Agrawal, V. K.; Srivastava, S.; Khadikar, P. V.; Supuran, C. T. *Bioorganic Medicinal Chemistry*. **2003**, *11*, 5353.
7. Singh, S.; Supuran, C. T. *J. Enzyme. Inhib. Med. Chem.* **2014**, *29*, 449.
8. Qiu, H. Y.; Wang, Z. C.; Wang, P. F.; Yan, X. Q.; Wang, X. M.; Yang, Y. H.; Zhu, H. L. *RSC Advances*. **2014**, *4*, 39214.
9. Supuran, C. T.; Clare, B. W. *Eur. J. Med. Chem.* **1998**, *33*, 489.
10. Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.; Wallingford, CT. 2009.
11. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
12. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
13. SPSS/PC, Statistical Package for IBMPC, Quiad software, Ontario, 1986.
14. Fatemi, M. H.; Baher, E. *SAR QSAR Environ. Res.* **2009**, *20*, 77.
15. Patterson, D. W. *Artificial Neural Networks: Theory and Applications*; Prentice Hall: U.S.A., 1996.
16. Bose, N. K; Liang, P. *Neural Network-Fundamentals*; McGraw-Hill: U.S.A., 1996.
17. Beal, M. T.; Hagan, H. B.; Demuth, M. *Neural Network Design*; PWS: U.S.A., 1996.
18. Jalali-Heravi, M.; Fatemi, M. H. *Anal. Chim. Acta*. **2000**, *415*, 95.
19. Nord, L. I.; Jacobsson, S. P. *Chemom. Intell. Lab. Syst.* **1998**, *44*, 153.
20. Tadeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*, eds; Wiley-VCH: Weinheim, 2000.
21. Wang, J.; Krudy, G.; Hou, T.; Zhang, W.; Holland, G.; Xu, X. *J. Chem. Inf. Model.* **2007**, *47*, 1395.