American-Eurasian Journal of Toxicological Sciences 3 (2): 85-91, 2011
ISSN 2079-2050

3D QSAR Approach on Substituted Isoxazolidines Derivatives as Angiotensin II Receptor Antagonist

Mukesh Chandra Sharma and Dharm Veer Kohli

Department of Pharmaceutical Sciences, Drug Research Laboratory,
Dr. H.S. Gour University, Sagar (M.P.) 470 003, India

Abstract: 3D-QSAR studies were performed using the stepwise variable selection k-nearestneighbor molecular field analysis approach; a leave-one-out cross-validated correlation coefficient (q2) of 0.7143 and a non-cross-validated correlation coefficient (r2) of 0.7481 were obtained. The structures of all compounds were built on a workspace of Vlife MDS 3.5 molecular modeling software and 3D QSAR models were generated by applying a partial least square (PLS) linear regression analysis coupled with a stepwise variable selection method and simulated annealing. Both derived models were found to be statistically significant in terms of regression and internal and external predictive ability ($r2$ = 0.72342 and 0.6473, pred_$r2$ = and 0.6718). Contour maps using this approach showed that steric, electrostatic and hydrophobic effects determine binding affinities. The results obtained from QSAR studies could be used in designing better Ang II activity among the congeners in future.

Key words: Angiotensin II · Isoxazolidines · QSAR · kNN-MFA

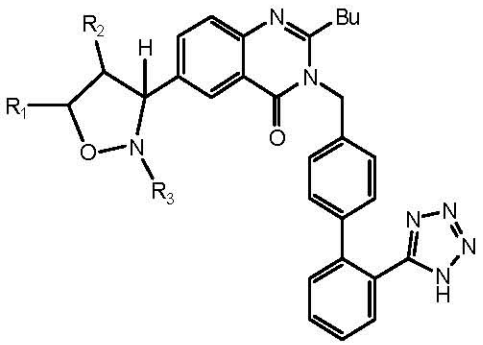
INTRODUCTION

The renin-angiotensin system (RAS) plays an important role in blood pressure regulation and electrolyte homeostasis [1]. Angiotensin II (AII) is the biologically active component of the RAS and is responsible for most of the peripheral effects of this system. The octapeptide angiotensin II (Ang II) is produced by the renin angiotensin system (RAS) and is a potent vasoconstrictor and thus plays an important role in the pathophysiology of hypertension [2]. This directed many researchers toward the designing of drugs to block the effect of Ang II either by inhibiting the angiotensin converting enzyme (ACE) or renin or by blocking the Ang II receptor [3]. Renin, an enzyme produced primarily by the juxtaglomerular cells of the kidney, catalyzes the conversion of angiotensinogen into an inactive substance, angiotensin I (A-I). Angiotensin-converting enzyme (ACE) then converts Ang -I to the physiologically active angiotensin II (Ang-II), which causes potent vasoconstriction, aldosterone secretion and sympathetic activation. All of these actions contribute to the development of hypertension [4-5]. The discovery of potent and orally active non-peptide Ang II antagonists such as losartan and eprosartan has encouraged the development of a large number of similar compounds.[6] Quantitative structure activity relationship (QSAR) is one of the major tools in drug discovery to explore ligand– receptor/enzyme interactions, especially when either the structural details of the target are not known or protein binding data of ligand is unavailable. [7]. A number of quantitative structure-activity relationship (QSAR) studies related to design of angiotensin II receptor antagonists drugs have also been reported [8-13]. The present study is aimed to elucidate the structural features of the development of a 3D QSAR kNNMFA quantitative structure activity relationship with the aid of various physicochemical parameters has been an important task in lead optimization. The relevance of the model for the design of novel derivatives should be assessed not only in terms of predictivity, but also in terms of their ability to provide a chemical and structural explanation of their binding interaction. Here we propose a general model for the antagonist and present minimal structural requirement for an Angiotensin II antagonist.

Methodology: The Ang II receptor antagonistic activity data of Substituted quinazolinone were taken from the reported work [14]. The activity data given as IC50 values, whereas IC_{50} refers to the experimentally determined nanomolar concentration of the Isoxazolidines derivatives.

Corresponding Author: Mukesh C. Sharma Drug Research Laboratory, Department of Pharmaceutical Sciences, Dr. H.S. Gour University, Sagar (M.P.) 470 003, India. E-mail: mukeshcsharma@yahoo.com.

Table 1: Biological activity data and structures of Series



Comp.	R_1	R_2	R_3	IC_{50}(nM)	pIC_{50}
1	CH_2OH	H	CH_3	61	1.78533
2	CH_2OH	H	CH_3	43	1.633468
*3	-$(CH_2)_3$-	H	CH_3	98	1.991226
4	-CH_2 CH_2CO-	H	CH_3	42	1.623249
5	-CH_2 CH_2CO-	H	CH_3	98	1.991226
6	-CH_2 CH_2CO-	H	CH_2Ph	79	1.897627
*7	-CH_2OCO-	H	CH_3	150	2.176091
8	-CH_2OCO-	H	CH_3	68	1.832509
9	-CH_2OCO-	H	CH_2Ph	94	1.973128
*10	-CH_2OCO-	H	CH_2Ph	177	2.247973
11	-$CH_2CH_2CH_2CO$-	H	CH_3	900	2.954243
*12	-1,3-cyclopentyl-	H	CH_3	154	2.187521
13	-$(CH_2)_6$-	H	CH_3	180	2.255273

*test compound

The biological activity values reported in nanomolar units were converted to their molar units and then further to negative logarithmic scale and subsequently used as the dependent variable for the QSAR analysis. The -log values of pIC50 (-logIC_{50}) along with the structure of the compounds in the series are listed in (Table 1).

Optimization of Molecules Structure and Alignment of Molecules: Three-dimensional structures of all compounds have been constructed using MDS 3.5 [15] and their geometries were subsequently optimized to make the conformations having least potential energy. Energy minimizations were performed using Merck molecular force field (MMFF) and MMFF charge followed by considering distance-dependent dielectric constant of 1.0 and convergence criterion of 0.01 kcal/mol. The 2D structures were converted to 3D structures by sending them to MDS. Energy minimization and geometry optimization was conducted using Merck Molecular Force Field (MMFF) method with Root Mean Square (RMS) gradient set to 0.01 Kcal/mol A^0 and iteration limit to 10000. Molecular alignment is a crucial step in 3D-QSAR study to obtain meaningful results. This method is based on moving of molecules in 3D space, which is related to the conformational flexibility of molecules. The goal is to obtain optimal alignment between the molecular structures necessary for ligand–receptor interactions.

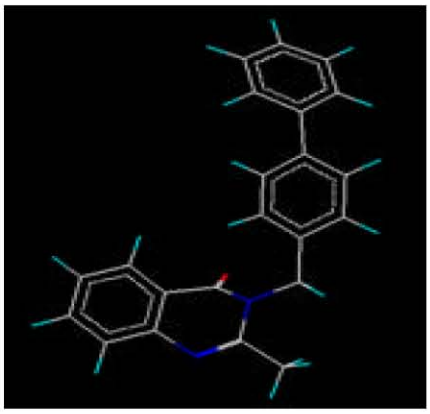


Fig. 1(a): Template structure for alignment all ligands

All molecules in the data set were aligned by template-based method where a template is built by considering common substructures in the series. The template structure, i.e. Isoxazolidines ring was used for alignment by considering the common elements of the series as shown in Figure 1(a). The reference molecule is chosen in such a way that it is the most active among the series of molecules considered. The superimposition of all molecules based on minimizing root mean square deviation (RMSD) is shown in Figure 1(b).In the kNN-MFA method, several models were generated for the selected members of training and test sets and

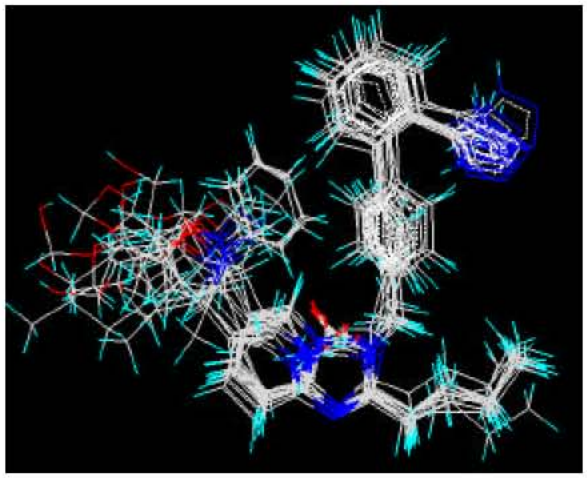


Fig. 1(b) 3D view of aligned molecules

the corresponding best models are reported herein. VLife Molecular Design Suite (VLifeMDS) allows user to choose probe, grid size and grid interval for the generation of descriptors. The variable selection methods along with the corresponding parameters are allowed to be chosen and optimum models are generated by maximizing q^2. k-nearest neighbor molecular field analysis (kNN-MFA) requires suitable alignment of given set of molecules. This is followed by generation of a common rectangular grid around the molecules. The steric and electrostatic interaction energies are computed at the lattice points of the grid using a methyl probe of charge +1. These interaction energy values are considered for relationship generation and utilized as descriptors to decide nearness between molecules. The term descriptor is utilized in the following discussion to indicate field values at the lattice points. The optimal training and test sets were generated using the sphere exclusion algorithm [16]. This algorithm allows the construction of training sets covering descriptor space occupied by representative points. Once the training and test sets were generated, kNN methodology was applied to the descriptors generated over the grid. The dissimilarity level was set to 5.9, as the higher the dissimilarity level, the lesser the predictive ability of QSAR model. The method resulted in selection and four compounds as test set and remaining others as training set.This is done to test the internal stability and predictive ability of the QSAR models. Developed QSAR models were validated by the following procedure:

Internal and External Validation: The regression coefficient r^2 is a relative measure of fit by the regression equation. It represents the part of the variation in the observed data that is explained by the regression. Internal validation is carried out using 'leave-one-out' (LOO) method [17]. The cross-validated coefficient, q^2, is calculated using the following equation:

$$q^2 = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \hat{y}_{\text{mean}})^2}$$

Where y_i and y_i are the actual and predicted activity of the i^{th} molecule in the training set, respectively and y_{mean} is the average activity of all molecules in the training set. However, a high q^2 value does not necessarily give a suitable representation of the real predictive power of the model for antimalarial ligands. So, an external validation is also carried out in the present study. The external predictive power of the model is assessed by predicting pIC_{50} value of the three test set molecules, which are not included in the QSAR model development. The predictive ability of the selected model is also confirmed by pred_r^2 or rCVext^2.

$$\text{pred\_}r^2 = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \hat{y}_{\text{mean}})^2}$$

Where y_i and y_i are the actual and predicted activity of the i^{th} molecule in the test set, respectively and y_{mean} is the average activity of all molecules in the training set. The robustness of the selected model is checked by Y - randomization test. The robustness of the models for training sets is examined by comparing these models to those derived for random datasets. Random sets are generated by rearranging the activities of the molecules in the training set. The significance of the models hence obtained is derived based on a calculated Z score[18]. A Z score value is calculated by the following formula:

$$\text{Zscore} = \frac{(h - \mu)}{\sigma}$$

Where h is the q^2 value calculated for the actual dataset, μ the average q^2 and σ is its standard deviation calculated for various iterations using models build by different random datasets.

Evaluation of the QSAR Models: The developed QSAR models are evaluated using the following statistical measures: n, (the number of compounds in regression); k, (number of variables); DF, (degree of freedom); optimum component, (number of optimum); r^2 (the squared

correlation coefficient), F test (Fischer's value) for statistical significance, q^2 (cross-validated correlation coefficient); pred_r^2, (r^2 for external test set); Z score, (Z score calculated by the randomization test); best_ran_q^2, (highest q^2 value in the randomization test); best_ran_r^2, (highest r^2 value in the randomization test). The regression coefficient r^2 is a relative measure of fit by the regression equation. It represents the part of the variation in the observed data that is explained by the regression. However, a QSAR model is considered to be predictive, if the following conditions are satisfied: $r^2 > 0.6$, $q^2 > 0.6$ and pred_$r^2 > 0.5$. The F-test reflects the ratio of the variance explained by the model and the variance due to the error in the regression. High values of the F-test indicate that the model is statistically significant. The low standard error of Pred_r^2se, q^2_se and r^2_se shows absolute quality of fitness of the model.

RESULTS AND DISCUSSION

The QSAR study of Substituted Isoxazolidines derivatives through PLS methodology, based on various feature selection methods Using VLife MDS 3.5 software resulted in the following statistically significant models.Molecules with better biological activity compared with the existing Substituted Isoxazolidines derivatives were searched. The k-NN MFA models were developed using step-wise forward-backward method with cross correlation limit set to 0.5 and term selection criteria as q^2. F-test 'in' was set to 4.0 and F-test 'out' to 3.99. As some additional parameters, variance cut-off was set as 2 Kcal/mol ? and scaling and auto scaling, additionally the k-nearest neighbor parameter setting was done within the range of 2-5 and prediction method was selected as distance based weight average. This method utilizes the active analogue principle that lies at the foundation of medicinal chemistry [19-21].

Model-A Simulated Annealing: E_146 (-0.4673 0.2746) + H_386 (0.1801 0.5561)- S_136 (-0.0493 -0.0134)-S_496 (-0.2903 -0.1696)-E_547 (-0.0710 1.6214).

k Nearest Neighbour= 3, n = 13,Degree of freedom = 17,q2 = 0.7143,q^2_se = 0.6748,Predr^2 = 0.6473, pred_r^2se = 0.4352

Model-B Step-Wise Forward-Backward: S_147 (-0.0389 -0.0135) + H_240 (0.1151 0.4416) + E_573 (0.2350 10.0000)-S_761 (-0.0537 -0.0388)-S_63 (-0.0214, -0.0043)

k Nearest Neighbour= 3,n = 13,Degree of freedom = 17,q^2 = 0.7481,q2_se = 0.3765 Pred_r^2 = 0.6718, pred_r^2se = 0.6753.

Model-A,[Figure-1(c)] the atom based alignment shows a *q2* (cross validated r^2) of 0.7143 with five descriptors namely E_146, H_386, S_136, S_496 and E_547. E_146, H_386, S_136, S_496 and E_547 are the steric, electrostatic and hydrophobic field energy of interactions between probe (CH_3) and compounds at their corresponding spatial grid points of 146, 386, 136, 496 and 547.A non-cross-validated *r2* of 0.7724, *F* value of 30.435 and number nearest neighbors *k* of 3 were observed with this model. I.e. all the values are proved statistically significant. The steric, electrostatic and hydrophobic contributions were 54 and 19%, respectively and exhibited good external prediction with *r2 pred* of 0.6473. Statistical significance of the model indicated by *Z score* value of 0.548 and α of >0.0001. The above model is validated by predicting the biological activities of the test molecules, as indicated in Table 3. Model-B [Figure-1(d)] the kNN-MFA model generated from template based alignment showed *q2* (cross validated *r2*) of 0.6680 with five descriptors namely S_147, H_240, E_573, S_761 and S_63. A non-cross validated *r2* of 0.7481, *F* value of 45.576 and number nearest neighbors *k* of 3 were observed with this model. The steric, electrostatic and hydrophobic contributions were 62,18 and 20%, respectively and

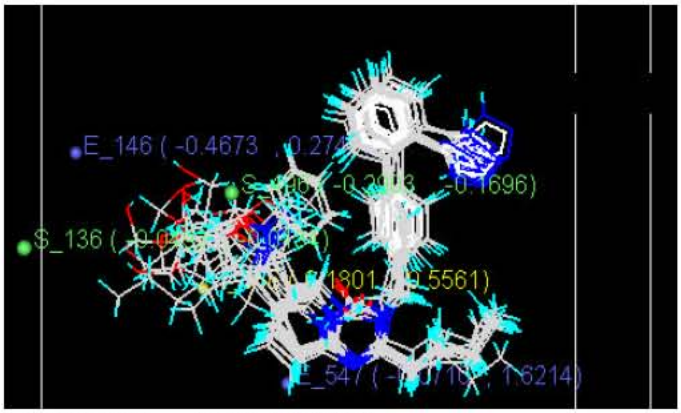


Fig. 1(c): Contribution plot of 3D QSAR model SA-kNNMFA

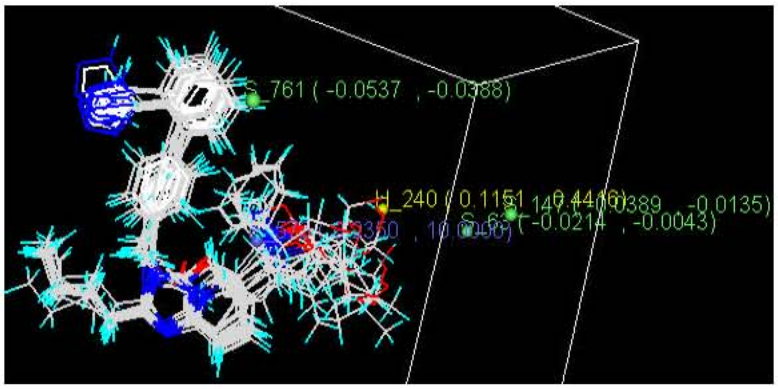


Fig. 1(d): Contribution plot of 3D QSAR model SW-kNNMFA

Table 2: 3D descriptors required for Angiotensin II affinity

S_136	S_761	E_146	E_547	E_573	H_240	H_386
-0.00752	-0.01214	0.630398	0.341021	0.179779	0.328875	0.324026
-0.00833	-0.01328	0.639316	0.501047	0.322779	0.312311	0.309952
-0.00846	-0.01343	0.590388	0.320405	0.178303	0.322065	0.315834
-0.00799	-0.01271	0.622685	0.358871	0.210634	0.366332	0.360443
-0.00793	-0.01272	0.569901	0.301318	0.158773	0.37287	0.367998
-0.00931	-0.01516	0.700978	0.313883	0.123483	0.310951	0.30707
-0.01313	-0.0215	0.448049	0.237053	0.133566	0.305833	0.300585
-0.00871	-0.01382	0.53667	0.277949	0.152232	0.302035	0.297274
-0.00909	-0.01465	0.773677	0.418354	0.223218	0.305762	0.301571
-0.00877	-0.01386	0.401862	0.122661	0.009063	0.301841	0.297192
-0.00958	-0.01535	0.575294	0.259747	0.088694	0.314695	0.308753
-0.00988	-0.01608	0.628009	0.31024	0.165483	0.330749	0.327037
-0.00915	-0.01462	0.493148	0.201454	0.066009	0.317216	0.312489

Table 3 Observed and predicted activities of statistically significant models

Observed activity	3D QSAR-SA-kNNMFA Predicted activity	3D QSAR-SW-kNNMFA Predicted activity
1.78	1.61	1.52
1.63	1.43	1.37
1.99	1.74	1.8
1.62	1.52	1.41
1.99	2.18	1.67
1.89	1.9	1.65
2.17	2.02	1.89
1.83	1.97	2.11
1.97	1.63	2.26
2.24	1.97	1.85
2.95	2.69	3.17
2.18	2.00	1.91
2.25	2.44	2.51

exhibited good external prediction with *r2 pred* of 0.6718.Statistical significance of the model indicated by *Z score* value of 1.254 and α of >0.0001. The above model is validated by predicting the biological activities of the test molecules, as indicated in Table 3. E_146 (-0.4673 0.2746), E_547 (-0.0710 1.6214) Negative value in electrostatic field descriptors indicates that negative electronic potential is required to increase activity and more electronegative substituents group is preferred in that position, E_573 (0.2350 10.0000) positive range indicates that group that imparting positive electrostatic potential is favorable for activity so less electronegative group is preferred in that region. S_136 (-0.0493 -0.0134), S_496 (-0.2903 -0.1696), S_147 (-0.0389 -0.0135), S_761 (-0.0537 -0.0388) and S_63 (-0.0214, -0.0043) negative coefficient in steric descriptors indicates that negative steric potential is favorable for

activity and less bulky substituents group is preferred in that region, Positive contribution of H_386 (0.1801 0.5561) and H_240 (0.1151 0.4416) to Hydrogen group nearer to R respectively indicates that positive hydrophobic field is favorable for increasing the activity. Hence less hydrophobic or more hydrophilic substituent groups R_1 and R_3 positions are preferred. The proposed models, due to the high internal and external predictive ability, can therefore act as a useful aid to the costly and time consuming experiments for determining the molar concentration of a compound required to achieve better angiotensin II receptor activity.

CONCLUSION

This model employs sphere exclusion method for generation of optimal training and test set and kNN-MFA with stepwise variable selection. The model was validated by cross-validation techniques, randomization and external test set prediction. These models could be usefully employed to prioritize chemicals for synthesis or in search of novel scaffolds from screening of chemical databases. The contours were found to be more or less qualitatively similar in accord with the experimentally observed activity. This is of great aid to design dual inhibitors activity. The analysis of contours has provided a clue about the structural requirement for the observed biological activity. This analysis could be of help in the rational design of potential drug antihypertensive with an enhanced inhibitory potency.

ACKNOWLEDGEMENT

The author wishes to express gratitude to V-life Science Technologies Pvt. Ltd for providing the software for the study and Head, School of Pharmacy, Devi Ahilya Vishwavidyalaya for providing facilities to carry out the work.

REFERENCES

1. Ferrario, C.M., 1990. The Renin-Angiotensin System: Importance in Physiology and Pathology. J. Cardiovasc. Pharmacol., 15(3): 51-55.
2. Burnier, M. and H.R. Brunner, 1997. Angiotensin II receptor antagonists antihypertensive agents. Expert. Opin. Investig. Drugs., 6: 489-500.
3. Goodfriend, T.L., M.E. Elliott and K.J.Catt, 1996. Angiotensin receptors and their antagonists. *N. Engl. J. Med.,* 334:1649-54.
4. Messerli, F.H., M.A. Weber and H.R. Brunner, 1996. Angiotensin II receptor inhibition. Arch. Intern. Med., 156: 1957-65.
5. Bauer, J.H. and G.P. Reams, 1996. The angiotensin II type 1 receptor antagonists: a new class of antihypertensive drugs. Arch. Intern. Med., 155: 1361-8.
6. Wexler, R.R., W.J. Greenlee, J.D. Irvin, M.R.Goldberg, K. Prendergast, R.D. Smith and P.B.M.W.M. Timmermans, 1996. Nonpeptide Angiotensin II Receptor Antagonists: The Next Generation in Antihypertensive Therapy. J. Med. Chem., 39: 625-656.
7. Kubinyi, H., R. Mannhold, L.R. Krogsgaard and H.E. Timmerman, 1993. Methods and principles in medicinal chemistry, vol 1. VCH, Weinheim, pp: 91-95.
8. Sharma, M.C., D.V. Kohli, S.C. Chaturvedi and S. Sharma, 2009. Molecular Modeling Studies of Some substituted 2-butylbenzimidazoles angiotensin II receptor antagonists as antihypertensive agents. Digest. J. Nanomat. Biostruct., 4(4): 843-856.
9. Sharma, M.C., S. Sharma, N.K. Sahu and D.V. Kohli, 2011. 3D QSAR kNNMFA studies on 6-Substituted Benzimidazoles Derivatives As Nonpeptide Angiotensin II Receptor Antagonists: A Rational Approach to antihypertensive agents. J. Saud. Chem. Soc. (In press).
10. Sharma, M.C., S. Sharma, N.K. Sahu and D.V. Kohli, 2011. QSAR Studies of some Substituted imidazolinones Derivatives angiotensin II receptor antagonists using Partial Least Squares Regression (PLSR) Based Feature Selection. J. Saud. Chem. Soc., (In press).
11. Datar, P.A., P.V. Desai and E.C. Coutinho, 2004. A 3D-QSAR of angiotensin II (AT1) receptor antagonists based on receptor surface analysis. J. Chem. Inf. Model., 44: 210-220.
12. Belvisi, L., G. Bravi, G. Catalano, M. Mabilia, A. Salimbeni and C. Scolastico, 1996. A 3D QSAR CoMFA study of nonpeptide angiotensin II receptor antagonists. J. Comput. Aided. Mol. Des., 10: 567-582.
13. Yan, P.Z., B.C. Chen, H. Shi, H.J. Jian, L.W. Hai, L.S. Guo and Q.Y. Ru, 2007. QSAR study of angiotensin II antagonists using robust boosting partial least squares regression. Anal. Chim. Acta., 593: 68-74.
14. Levin, J.I., A.M. Venkatesan, P.S. Ghan, T.K. Bailey, G. Vice and J. Coupet, 1994. 6-Substituted quinazolinone Angiotensin II Receptor Antagonists Bioorg. Med. Chem. Lett., 4(15): 1819-1824.

15. Vlife, M.D.S, 2004. VlifeMDS molecular design suite. VLife Sciences Technologies Pvt. Ltd., Pune, India.
16. Golbraikh, A. and A. Tropsha, 2002. Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. J. Comp. Aided. Mol. Design., 16: 357-369.
17. Cramer, R.D., D.E. Patterson and J.D. Bunce, 1988. Comparative molecular field analysis (CoMFA) 1. Effect of shape on binding of steroids to carrier proteins. J. Am. Chem. Soc., 110: 5959-67.
18. Zheng, W. and A. Tropsha, 2000. Novel variable selection quantitative structure–property relationship approach based on the knearest-neighbor principle. J. Chem. Inf. Comput. Sci., 40: 185-194.
19. Sharma, M.C. and S. Sharma, 2010. 3D- Quantitative Structure-Activity Relationship Analysis of Some 2-Substituted Halogen benzimidazoles Analogues with Antimycobacterial activity. Int. J. Chem. Tech. Res., 2(1): 606-614.
20. Sharma, M.C., S. Sharma, D.V. Kohli and S.C. Chaturvedi, 2010. Three Dimensional Quantitative Structural–Activity Relationship (3D-QSAR) Studies some 3-{4-[3-(2-aryl-phenoxy) butoxy]-phenyl} Propionic acids as novel PPAR γ/δ agonists. Der. Pharma. Chemica., 2(1): 82-90.
21. Sharma, M.C., S. Sharma, D.V. Kohli and S.C. Chaturvedi, 2010. QSAR and k-Nearest Neighbour Molecular Field Analysis (k-NN MFA) Classification Analysis of Studies of Some Bemzimidazoles Derivatives Antibacterial activity Against Escherichia coli. Der. Pharmacia. Lett., 2(1): 150-161.