American-Eurasian Journal of Toxicological Sciences 3 (2): 75-84, 2011
ISSN 2079-2050

3D QSAR *kNNMFA* Approach Studies on Series of Substituted Piperidin-2-One Biphenyl Tetrazoles as Angiotensin II Receptor Antagonists

Mukesh Chandra Sharma and Dharm Veer Kohli

Drug Research Laboratory, Department of Pharmaceutical Sciences,
Dr. H.S. Gour University, Sagar (M.P.) 470003, India

Abstract: k-nearest neighbor molecular field analysis (kNN-MFA) based 3D-QSAR, fragment and knowledge based approach were used to generate and screen a virtual library. k-Nearest Neighbor Molecular Field Analysis (kNN-MFA), a three dimensional quantitative structure activity relationship (3D-QSAR) method has been used in the present to study the correlation between the molecular properties and the angiotensin II receptor activities on a series of piperidin-2-one biphenyl tetrazoles derivatives. kNN-MFA calculations for both electrostatic and steric field were carried out. The master grid maps derived from the best model has been used to display the contribution of electrostatic potential and steric field. The best model with good external and internal predictivity for the training and test set has shown cross validation (q2) and external validation (pred_r^2) values of 0.7493 and 0.7654, respectively. The steric descriptors at the grid points E_891, E_875, S_1118, E_1226, S_306 and E_1108 play an important role in the design of new molecule. The information rendered by 3D QSAR models may lead to a better understanding of structural requirements of angiotensin II receptor antagonists and also aid in designing novel potent antihypertensive molecules.Partial least-squares (PLS) method was applied for QSAR model development considering training and test set approaches.

Key words: Ang II · 3D QSAR · kNN-MFA · Piperidin-2-one · Antihypertensive

INTRODUCTION

During the past several years, investigators interested in the development of new therapies for cardiovascular diseases have focused on the renin-angiotensin system [1]. The renin-angiotensin system (RAS) is known to play a pivotal role in the regulation of fluid, electrolyte balance and blood pressure and is a modulator of cellular growth and proliferation. Now, several efforts have been made for the treatment of hypertension and congestive heart failure. The renin inhibitors, highly specific enzymes responsible for the conversion of angiotensin to angiotensin I, have been indicated experimentally to exert potent antihypertensive effects, but efforts to develop renin inhibitors as drugs have been hampered by their poor oral bioavailability and limited duration of action. The blockade of the Ang II synthesis by converting enzyme inhibitors (CEI) has become a well-established therapeutic application in the treatment of hypertension and heart failure [2]. Renin, a proteolytic enzyme produced mainly in the juxtaglomerular apparatus of the kidney, acts on the circulating alpha globulin angiotensinogen produced by the liver to form the decapeptide Asp-Arg-Val-Tyr-Ileu-His-Pro-Phe-His-Leu,named angiotensin I [3]. Quantitative structure activity relationship (QSAR) searches information relating chemical structure to biological and other activities by developing a QSAR model. Using such an approach one could predict the activities of newly designed compounds before a decision is being made whether these compounds should be really synthesized and tested. The QSAR approach helps to correlate the specific biological activities or physical properties of a series of compounds with the measured or computed molecular properties of the compounds, in terms of descriptors [4-5]. QSAR methodologies save resources and expedite the process of the development of new molecules and drugs. A number of quantitative structure-activity relationship (QSAR) studies related to design of angiotensin II receptor antagonists drugs have also been reported [6-10]. Many different approaches to QSAR have been developed over the years. The rapid increase in three-

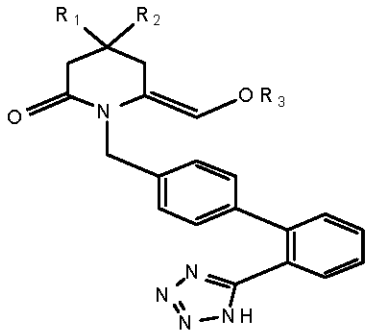
Corresponding Author: M.C. Sharma, Department of Pharmaceutical Sciences, Dr. Hari Singh Gaur University, Sagar (M.P.) 470003, India. E-mail: mukeshcsharma@yahoo.com.

dimensional structural information (3D) of bioorganic molecules, coupled with the development of fast methods for 3D structure alignment (e.g. active analogue approach), has led to the development of 3D structural descriptors and associated 3D QSAR methods. The most popular 3D QSAR methods are comparative molecular field analysis and comparative molecular similarity analysis [11-12]. In the present study; we have applied k-nearest neighbour molecular field analysis (kNNMFA) [13]. The present study is aimed to elucidate the structural features of piperidin-2-one biphenyl tetrazoles derivatives required for angiotensin II receptor antagonists and to obtain predictive three-dimensional QSAR models to guide the rational synthesis of novel antihypertensive activity. With the above facts and in continuation of our research for newer anti-hypertensive agent in the present study, we report here the development of a new method (kNNMFA) that adopts a k-nearest neighbor principle for generating relationships of molecular fields with the experimentally reported activity to provide further insight into the key structural features required to design potential drug candidates of this class. This method utilizes the active analogue principle that lies at the foundation of medicinal chemistry.

MATERIALS AND METHODS

Data Set for Analysis: QSAR studies were performed on a series of piperidin-2-one biphenyl tetrazoles reported work were taken [14]. The biological activities of these sixteen compounds were expressed in terms of IC_{50} values for angiotensin II receptor antagonists. For correlation purposes, the values were converted to negative logarithmic scale-log IC_{50}. These compounds along with their biological data are presented in Table 1. The molecular structure of all the sixteen molecules were sketched using VLife MDS 3.5 [15] software in the 2D builder module and then the structures were converted to 3D space for further analysis. The structures were sketched using the 2D draw application and converted to 3D structures. Three-dimensional structures were drawn for each molecule and the molecular geometries optimized using Monte Carlo conformational search fields and charges. All the compounds were drawn in Chem DBS using fragment database and then subjected to energy minimization using batch energy minimization method. Conformational search were carried out by systemic conformational search method and all the compounds were aligned by template based method.

Table 1: Structural and biological data piperidin-2-one



Com	R_1	R_2	R_3	IC_{50} (nM)	-logIC_{50}
1	Me	Me	OEt	5800	3.763
2	Me	Et	OEt	250	2.397
3*	Et	Et	OEt	20	1.301
4	Me	H	OEt	200	2.301
5	Me	n-butyl	OEt	470	2.672
6	Et	n-butyl	OEt	180	2.255
7	n-propyl	n-propyl	OEt	330	2.518
8	Me	n-propyl	OEt	120	2.079
9*	Me	i-propyl	OEt	120	2.079
10	Et	n-propyl	OEt	90	1.954
11	Me	n-pentyl	OEt	680	2.832
12*	-cyclopentyl	-	OEt	540	2.732
13	-cyclohexyl	-	OEt	40	1.602
14	Et	Et	Me	3100	3.491
15	Me	Et	OMe	320	2.505
16*	Et	Et	OMe	210	2.322

*test compound

Table 2: Description of descriptor used in the 3D QSAR model

Com	S_162,	S_306	S_723	S_1124	E_229	E_874	E_1108
1	-0.00674	-0.0797	-0.00840	-0.0771	0.018569	0.017519	0.015124
2	-0.00644	-0.0736	-0.00751	-0.0673	0.031145	0.030585	0.028563
3	-0.00801	-0.0914	-0.00920	-0.0811	0.015318	0.014744	0.012758
4	-0.00675	-0.0692	-0.00611	-0.0479	0.017891	0.021523	0.023665
5	-0.00805	-0.0901	-0.00871	-0.0716	-0.001890	-0.005280	-0.008330
6	-0.00709	-0.0772	-0.00733	-0.0607	0.035918	0.031953	0.026197
7	-0.00775	-0.0919	-0.01001	-0.0945	0.032008	0.028563	0.023748
8	-0.00925	-0.1037	-0.00935	-0.0699	0.020339	0.023172	0.023802
9	-0.01020	-0.1215	-0.01219	-0.1019	0.014393	0.009809	0.046750
10	-0.01030	-0.1712	-0.02284	-0.0208	0.006638	0.009763	0.010712
11	-0.01033	-0.0214	-0.01223	-0.0099	0.031822	0.031668	0.029998
12	-0.00574	-0.0726	-0.00808	-0.0762	0.015106	0.015309	0.014159
13	-0.00574	-0.0726	-0.00808	-0.0542	0.015106	0.015309	0.014159
14	-0.00325	-0.0349	-0.00328	-0.0273	-0.013020	-0.015580	-0.018610
15	-0.00683	-0.0751	-0.00722	-0.0609	0.007135	0.002869	-0.002540
16	-0.00739	-0.0807	-0.00769	-0.0644	0.003906	0.002790	0.000723

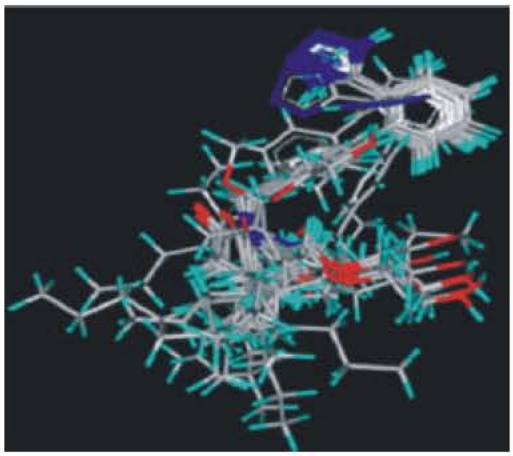


Fig. 1: Superposition of compounds in the training and test sets using the template-based alignment method

Optimized molecules were aligned Fig. 1 by template based method using the most active molecule as a template. The optimized batch of molecules was selected for calculation of the physiochemical descriptors. The descriptor pool was reduced by eliminating out the descriptors with constant and near constant values. Further reduction in the descriptor pool was done by ousting the descriptors that are highly degenerate and difficult to interpret. A correlation analysis was performed between biological data and remaining descriptors, most of which were molecular and electro topological descriptors and the descriptors those were showing very low correlation with inhibitory activity were also removed. The descriptors selected for modeling activity of the piperidin-2-one biphenyl tetrazoles are summarized in Table 2.

Methodology: We hereby report the models, as generated by kNN-MFA in conjunction with stepwise (SW) forward-backward variable selection methods. In the kNN-MFA method, several models were generated for the selected members of training and test sets and 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. 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.

Molecular Modeling and Alignment: Conformational search were carried out by systemic conformational search method and lowest energy conformers were selected. All the compounds were aligned by template based method. The selection of template molecule for

alignment was done by considering the following facts: a) the most active compound; b) the lead or commercial compound; c) the compound containing the greatest number of functional group[16,17] Generally, the low energy conformer of the most active compound is selected as a reference.

Selection of Training and Test Set: The dataset of sixteen molecules was divided into training and test set by Sphere Exclusion (SE) [18] method for model 1, 2, 3 and 4 having dissimilarities values of 8.9, respectively with -$\log IC_{50}$ activity field as dependent variable and various 3D descriptors calculated for the compounds as independent variables. In the present kNN-MFA study, (-16.31654 to 16.6743) × (-11.3176 to 14.32176) × (-10.6542 to 12.4398) A° grid at the interval of 2.00 was generated around the aligned compounds. This algorithm allows constructing training sets covering all descriptor space areas occupied by representative points. The higher the dissimilarity level is, the smaller the training set is and the larger the test set is. It is expected that the predictive ability of QSAR models generally decreases when the dissimilarity level increases. The method resulted in selection and four compounds as test set and remaining others as training set.The steric and electrostatic interaction energies are computed at the lattice points of the grid using a methyl probe of charge +1 of Gasteiger-Marsili type. These interaction energy values are considered for relationship generation and utilized as descriptors to decide nearness between molecules. The method described above has been implemented in software, Vlife Molecular Design Suite (Vlife MDS), which 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. For calculation of field descriptor values, using Tripos force field both electrostatic and steric field type with cut offs 10.0 and 30.0 Kcal/mol respectively were selected and charge type was selected as Gasteiger-Marsili. Dielectric constant was set to 1.0 considering the distance dependent dielectric function. This resulted in calculation of 4818 field descriptors (1606 for each steric, electrostatic) for all the compounds in separate columns. QSAR analysis was performed after removal of all the invariable columns, as they do not contribute to QSAR. The optimal test and training data set were generated using manual selection method.

The kNN methodology relies on a simple distance learning approach whereby an unknown member is classified according to the majority of its k-nearest neighbors in the training set. The nearness is measured by an appropriate distance metric (e.g., a molecular similarity measure calculated using field interactions of molecular structures). The standard kNN method is implemented simply as follows: Calculate distances between an unknown object (u) and all the objects in the training set; select k objects from the training set most similar to object u, according to the calculated distances; and classify object u with the group to which the majority of the k objects belongs. An optimal k value is selected by optimization through the classification of a test set of samples or by leave-one out cross-validation.

Steps Involved in kNN-MFA Method: Molecules are optimized before alignment optimization is done by MOPAC energy minimization and optimization is necessary process for proper alignment of molecules around template. kNN-MFA method requires suitable alignment of given set of molecules, alignment are template based. This is followed by generation of 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.The optimal training and test set were generated using sphere exclusion method. Model was generated by various kNN methods and models validated internally and externally by leave one out, external validation. Predict the activity of test set of compounds.

Model Quality and Validation

Internal and External Validation: Internal validation is carried out using 'leave-one-out' (LOO) method. The cross-validated coefficient, q^2, is calculated using the following equation:

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

Where y_i and $\hat{y}_i$ 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 four 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.

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

Where y_i and $\hat{y}_i$ 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 [19] A Z score value is calculated by the following formula:

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

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. The probability (α) of significance of randomization test is derived by comparing Z score value with Z score critical value, if Z score value is less than 5.0; otherwise it is calculated by the formula as given in the literature. For example, a Z score value greater than 3.95 indicates that there is a probability (α) of less than 0.001 that the QSAR model constructed for the real dataset is random. The randomization test suggests that all the developed models have a probability of less than 1% that the model is generated by chance.

RESULTS AND DISCUSSION

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 PLS components in the model); 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).. This method utilizes the active analogue principle that lies at the foundation of medicinal chemistry [20-22]. 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. 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. This method utilizes the active analogue principle that lies at the foundation of medicinal chemistry adopts a k-Nearest Neighbor principle for generating relationship of molecular fields with the experimentally reported activity (Ajmani *et al*., 2006). The kNN-MFA steric and electrostatic fields thus generated were scaled by the standard method in the software. 3D QSAR models were generated by kNN-MFA in conjunction with Stepwise (SW) Forward Backward selection methods. From these models, the one with q^2 = 0.5 and pred_r^2 = 0.5 was selected to predict the activity of training and test set of compounds. The steric (S) and electrostatic (E) descriptors specify the regions, where variation in the structural features of different compounds in the training set leads to increase or decrease in activities. The number accompanied by the descriptors represents its position in the 3D MFA grid.

Model-1:

pIC50 = 0.62346(S_1124)-0.0043(S_162)-0.2703(S_723) - 0.2208 (E_229)

$N_{traning}$ = 12, N_{test} = 4, Optimum Components = 3, DF = 33, r^2 = 0.8134, q^2 = 0.7621, F test = 65.871, r^2_se = 0.3681, q^2_se = 0.5301, pred_r^2 = 0.7138, pred_r^2se = 0.2514, ZScore $Q^{\wedge 2}$ = 0.1543, Best Rand $Q^{\wedge 2}$ = 0.5417.

Model 1 atom based alignment shows a q^2 (cross validated r^2) of 0.7621. The descriptors S_1124, S_162, S_723 and E_229 are the steric and electrostatic field energy of interactions between probe (CH_3) and compounds at their corresponding spatial grid points of 1124, 162, 723 and 229. The external predictability of the above 3D-QSAR model using the test set was determined by Pred_r^2, which is 0.7138. So the above results indicate that 3D-QSARmodel for angiotensin II receptor antagonists generate 76% and 70 % internal and external model prediction, respectively and *F* value of 65.871 and

Table 3: 3D-QSAR-derived observed and predicted activities of compounds

Comp	Observed Activity	Predicted activity-1	Predicted activity-2	Predicted activity-3	Predicted activity-3
1	3.763	3.68700	3.832	3.6980	3.8840
2	2.397	2.28073	2.512	2.5080	2.3130
3	1.301	1.23100	1.374	1.4310	1.1970
4	2.301	2.19400	2.241	2.0560	2.1101
5	2.672	2.49100	2.441	2.7430	2.8320
6	2.255	2.03200	2.135	2.0920	2.3140
7	2.518	2.58200	2.643	2.3840	2.3930
8	2.079	1.95400	2.115	2.2140	1.9320
9	2.079	2.14100	2.368	2.1970	2.2340
10	1.954	1.85900	1.767	2.0980	1.7840
11	2.832	2.69300	2.772	2.8940	2.6720
12	2.732	2.66000	2.857	2.5176	2.4870
13	1.602	1.46400	1.674	1.5280	1.7240
14	3.491	3.32600	3.564	3.6090	3.3850
15	2.505	2.32000	2.612	2.3350	2.4380
16	2.322	2.21100	2.387	2.1570	2.1140

number nearest neighbors k of 4 were observed with this model. Statistical significance of the model indicated by *Z score* value of 0.1543 and *α* of >0.001. The plot of contributions of steric and electrostatic field interactions model in Fig. 2(a) indicates relative regions of the local fields (steric and electrostatic) around the aligned molecules. The above model is validated by predicting the biological activities of the test molecules, as indicated in Table 3. The plot of observed versus predicted activities for the test compounds is represented in Fig. 3(a).

Model-2:

pIC50 = -0.0412 (E_891) - 10.0000(E_875) +3.0467(S_1118) -1.0130 (E_1226) + 0.2309 (S_306) -0.2261 (E_1108)

$N_{traning}$ = 12, N_{test} = 4, Optimum Components = 4, DF = 31, r^2 = 0.7843, q^2 = 0.7493, F test = 54.16, r^2_se = 0.5397, q^2_se = 2.5973, pred_r^2 = 0.7954, pred_r^2se = 0.3188, ZScore $Q^{\wedge 2}$ = 1.3257, Best Rand $Q^{\wedge 2}$ = 0. 3862.

Model 2 is indicated Steric and electrostatic interaction field at lattice points Corresponding spatial grid points of 891, 875, 1118, 1226, 306 and 1108. These points suggested the significance and requirement of electrostatic and steric properties as mentioned in the ranges in parenthesis for structure–activity relationship and maximum biological activity of Ang II antagonists. The external predictability of the above 3D-QSAR model using the test set was determined by Pred_r^2, which is 0.7954. So the above results indicate that 3D-QSARmodel for angiotensin II receptor antagonists generate 61% and 39 % internal and external model prediction, respectively and *F* value of 54.16 and number nearest neighbors k of 4 were observed with this model. Statistical significance of the model indicated by *Z score* value of 1.325 and *á* of >0.001. The plot of contributions of steric and electrostatic field interactions model Fig; 2(b) indicates relative regions of the local fields (steric and electrostatic) around the aligned molecules. The above model is validated by predicting the biological activities of the test molecules, as indicated in Table 3. The plot of observed versus predicted activities for the test compounds is represented in Fig. 3(b).

Model-3:

pIC50 = +25.8951(S_1124) -0.3112 (S_954) +4.9401 (E_874)

$N_{traning}$ = 12, N_{test} = 4, Optimum Components = 4, DF = 35, r^2 = 0.7486, q^2 = 0.7013, F test = 29.88, r^2_se = 0.9662, q^2_se = 2.6583, pred_r^2 = 0.6746, pred_r^2se = 0.3188, ZScore $Q^{\wedge 2}$ = 3.166, Best Rand $Q^{\wedge 2}$ = 0. 3862.

Model-4:

pIC50 = -0.0615 (E_1232) -0.0675 (S_544) +0.1265 (E_993) -0.0579 (S_1159) 0.0858(E_831) 0.0492

$N_{traning}$ = 12, N_{test} = 4, Optimum Components = 4, DF = 26, r^2 = 0.7281, q^2 = 0.6845, F test = 37.854, r^2_se = 0.2448, q^2_se = 0.3798, pred_r^2 = 0.6583, pred_r^2se = 0.5437, ZScore $Q^{\wedge 2}$ = 1.1875, Best Rand $Q^{\wedge 2}$ = 0. 3299.

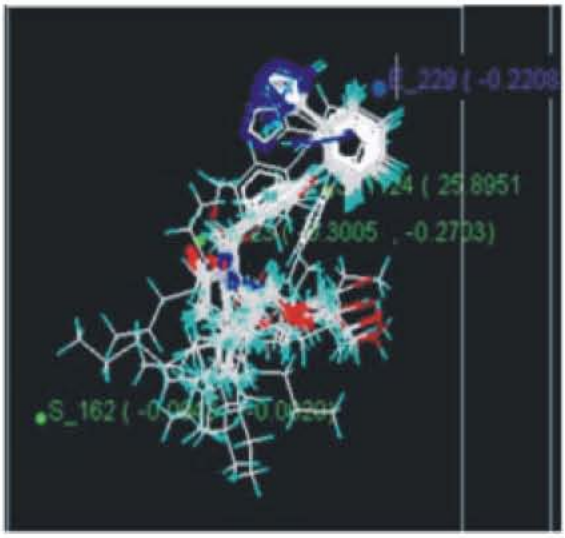


Fig. 2(a): Contribution plot for steric and electrostatic interactions model-1

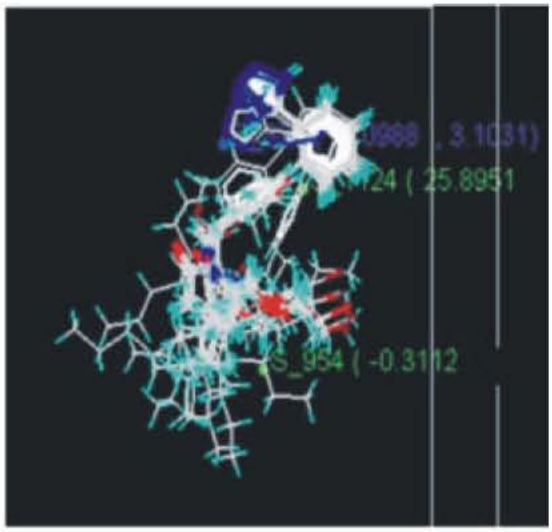


Fig. 2(c): Contribution plot for steric and electrostatic interactions model-3

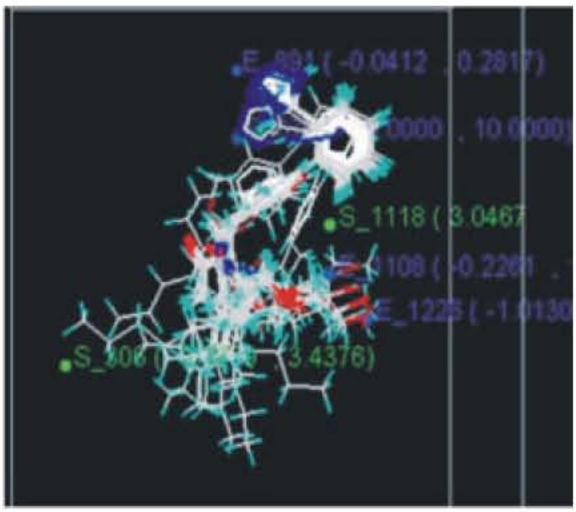


Fig. 2(b): Contribution plot for steric and electrostatic interactions model-2

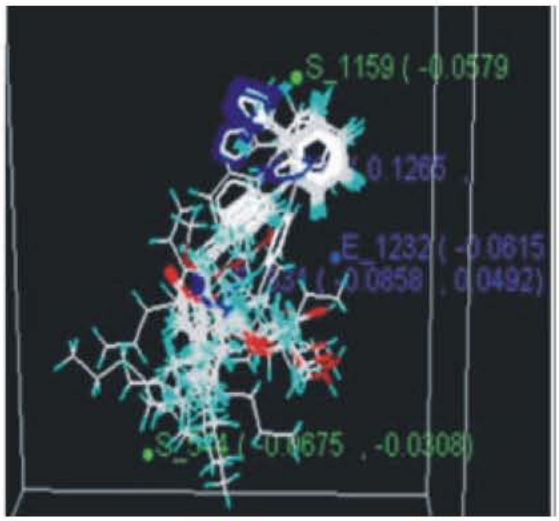


Fig. 2(d): Contribution plot for steric and electrostatic interactions model-4

Model 3 indicated steric and electrostatic contributions were 26 and 74 %, respectively and exhibited good external prediction with r^2 *pred* of 0.6746, F value 29.88. Statistical significance of the model indicated by *Z score* value of 3.166 and α of >0.001. The randomization test suggests that the developed model have a probability of less than 1% that the model is generated by chance. The plot of contributions of steric and electrostatic field interactions model Fig. 2(c) indicates relative regions of the local fields (steric and electrostatic) around the aligned molecules. Green and blue balls represent steric and electrostatic field effects, respectively. The above model is validated by predicting the biological activities of the test molecules, as indicated in Table 3. The plot of observed versus predicted activities for the test compounds is represented in Fig. 3c). From 3D-QSAR model 1 and 2 it is observed that electrostatic field with negative coefficient (E_229, E_111, E_1298, E_1316) is far from the moiety, indicating that electronegative groups are unfavourable on this site and presence of electronegative groups decrease the activity of piperidin-2-one biphenyl tetrazoles compounds. Presence of electrostatic field in model 3 with positive coefficient (E_874) suggests that electropositive (electron-withdrawing) substituent may be favorable on the position of biphenyl template.

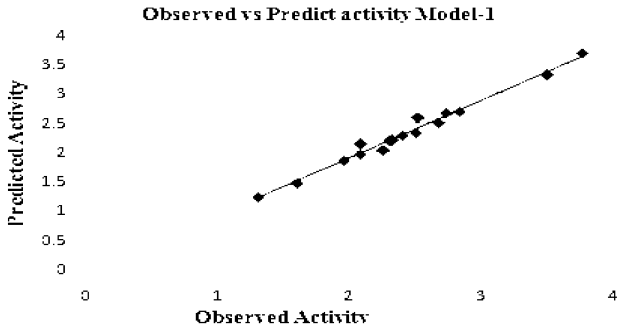


Fig. 3(a): Graph between Observed and Predicted activity of 3D QSAR model 1

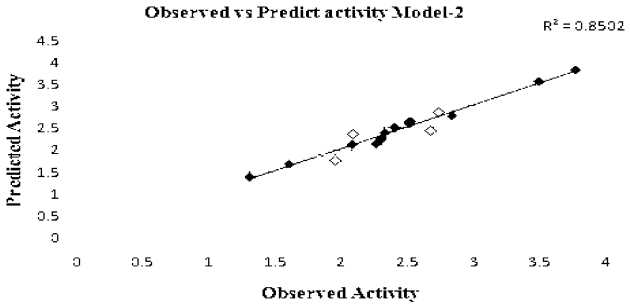


Fig. 3(b): Graph between Observed and Predicted activity of 3D QSAR model 2

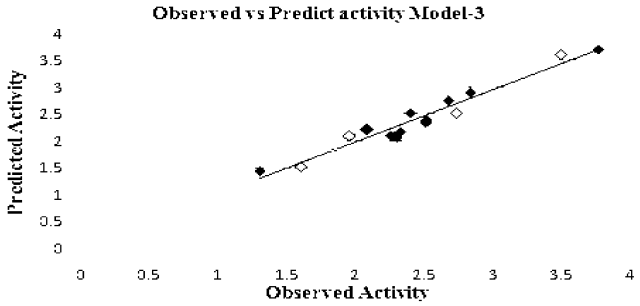


Fig. 3(c): Graph between Observed and Predicted activity of 3D QSAR model 3

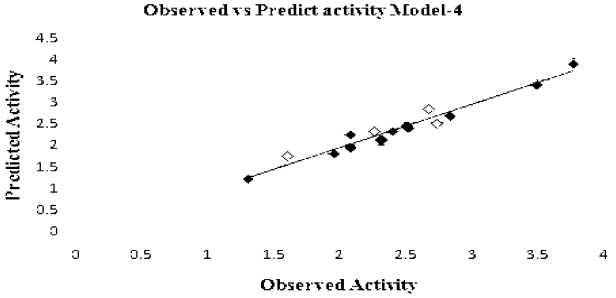


Fig. 3(d): Graph between Observed and Predicted activity of 3D QSAR model 4

Presence of steric descriptors with positive coefficients simultaneously at R^1 and R^2 positions of the piperidin-2-one ring, such as S_1124 and S_954, suggests the favourability of bulky groups in these regions for producing potent Ang II receptor activity. Model 4 indicate S_544, S_1159 As far as steric field is concerned, a negative range indicated that a negative steric potential was favourable for increased activity and hence a less bulky substituent group was preferred in that region. E_831,E_993,E_1232 negative values of electrostatic field descriptors (blue) indicates that negative electronic potential is required to increase activity and more electronegative substituents group is preferred in that position (Fig. 2(d). The plot of observed versus predicted activities for the test compounds is represented in Fig. 3(d). To ascertain the true predictivity of the model, applying leave-one-out method of cross validation using weighted k-nearest neighbor was performed for all the analysis. A Cross-validation analysis was also performed by applying leave-one-out technique using weighted *k*-nearest neighbor method. It is known that the CoMFA method provides significant value in terms of a new molecule design, when contours of the PLS coefficients are visualized for the set of molecules. Similarly, the kNN-MFA models provide direction for the design of new molecules in a rather convenient way. The points which contribute to the kNN-MFA model in data set are termed as the distribution point map. The regression analysis of kNN field was performed using distance-based weighted average and the predictive value of the model was evaluated by standard leave-one out (LOO) cross validation. The cross-validated correlation coefficient served as a measure of the quality of the model. To test the utility of the model as a predictive tool, an external set of compounds with known activities (the test set) were used. The predictive r^2 calculation was based on molecules in the test set and was used to evaluate the predictive power of the kNN-MFA model.

CONCLUSIONS

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 Ang II antagonist activity. Model 2 is giving very significant results. The master grid obtained for the various kNN-MFA models show that negative value in electrostatic field descriptors indicates the negative electronic potential is required to increase activity and more electronegative substituents group is preferred in that position, positive range indicates that the group which imparts positive electrostatic potential is favorable for activity so less electronegative group is preferred in that region. Positive value of steric descriptors reveals that positive steric potential is favorable for increase in activity and more bulky group is preferred in that region. On the basis of the spatial arrangement of the various shapes, electrostatic and steric potential contributions model proposed in this work. The location, range of function values at the field points selected by model provide the clues for the design of new molecules thus giving insight on structural requirement for designing more potent analogues.

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. Vallotton, M.B., 1987. The Renin-Angiotensin System. Trends. Pharmacol. Sci., 8: 69.
2. Ferrario, C.M., 1990. The Renin-Angiotensin System: Importance in Physiology and Pathology. J. Cardiovasc. Pharmacol., 15(3): 51-55.
3. Fouad, F.M. and R.C. Tarazi, 1986. Restoration of cardiac function and structure by converting enzyme inhibition possibilities and limitations of long-term treatment in hypertension and heart failure. J. Cordiovasc. Pharmcol., 8: S53-S57.
4. Doweyko, A.M., 2008. QSAR: Dead or alive? J. Comput. Aided Mol. Des., 22: 81-89.
5. Scior, T., J.L. Medina-Franco, Q.T. Do, K. Martinez-Mayorga, J.A.Y. Rojas and P. Bernard, 2009. How to recognize and workaround pitfalls in QSAR studies: A critical review. Curr. Med. Chem., 16: 4297-4313.
6. Belvisi, L., G. Bravi, G. Catalano, M. Mabiliab, 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.
7. 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).

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. Saudi Chemical Society (In press).
10. Yoo, S.E., S.K. Kim, S.H. Lee, K.Y. Yi and D.W. Lee, 1999. A comparative molecular field analysis and molecular modeling studies on pyridylimidazole type of angiotensin II antagonists. Bioorg. Med. Chem., 7: 2971-2976.
11. 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.
12. Klebe, G., U. Abraham and T. Mietzner, 1994. Molecular similarity indexes in a comparative-analysis (CoMSIA) of drug molecules to correlate and predict their biological-activity. J. Med. Chem., 37: 4130-4146.
13. Ajmani, S., K. Jhadav and S.A. Kulkarni, 2006. Three-dimensional QSAR using k-nearest neighbor method and its interpretation. J. Chem. Inf. Model., 46: 24-31.
14. Murray, W.V., P. Lalan, A. Gill, M.F. Addo, J.M. Lewis, D.K.H. Lee, R. Rampulla, M.P. Wachter, J.D. Hsi and D.C. Underwood, 1992. Substituted piperidin-2.one biphenyl tetrazoles Angiotensin II antagonists. Biorg. Med. Chem. Lett., 2(12): 1775-1779.
15. Vlife MDS software package, version 3.5, supplied by Vlife science technologies Pvt. Ltd, 1, Akshay 50, Anand park, Aundh, Pune, India 411007.
16. Agarwal, A. and E.W. Taylor, 1993. 3-D QSAR for intrinsic activity of 5-HT_{1a} receptor ligands by the method of comparative molecular field analysis. J. Computat. Chem., 14: 237-245.
17. Baurin, N., E. Vangrevelinghe and L.M. Allory, 2000. 3D-QSAR CoMFA study on imidazolinergic I_2 ligands: a significant model through a combined exploration of structural diversity and methodology. J. Med. Chem., 43: 1109-1122.
18. 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.
19. 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.
20. 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.
21. 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.
22. 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. Lettre., 2(1): 150-161.