Digest Journal of Nanomaterials and Biostructures Vol. 4, No. 4, December 2009, p. 843 – 856

# MOLECULAR MODELLING STUDIES OF SOME SUBSTITUED 2-BUTYLBENZIMIDAZOLES ANGIOTENSIN II RECEPTOR A NTAGONISTS AS ANTIHYPERTENSIVE AGENTS

M. C. SHARMA[a*], D. V. KOHLI [a] S. C. CHATURVEDI[b], SMITA.SHARMA[c]
[a]*Pharmaceutical Chemistry Research Laboratory, Department of Pharmaceutical Sciences, Dr. H. S. Gour Vishwavidyalaya, Sagar (M.P) – 470 003, India*
[b]*School of PharmacyD.A.V.V Indore (M.P)-452001 India*
[c]*Chodhary Dilip Singh KanyaMahavidyla Bhind (M.P)-477001*

A series of 33 Compounds antihypertensive drugs 2-butylbenzimidazoles derivatives were subjected to quantitative structure activity relationship analysis with an attempt to derive a correlation between the biological activity as dependent variable and various descriptors as independent variables by using V-Life MDS 3.5. Negative logarithmic value of (–PMIC) was taken as dependent variable, chi3Cluster, T_2_Cl_6, XlogP, T_N_N_slogp, T_O_Cl_5 was taken as independent varable.Docking studies of all compounds using Biopredicta V-Life MDS 3.5,all the molecules docked and scores were calculated and most active compound has high score and successfully docked into the active site This study can help in rational drug design and synthesis of new angiotensin II antagonists as antihypertensive with predetermined affinity.

(Received June 6, 2009; accepted October 7, 2009)

*Keywords:* QSAR, PLS, Angiotensin II receptor, 2-butylbenzimidazole, Docking, Losrtan, antihypertensive agents

## 1. Introduction

Quantitative structure-activity relationship (QSAR) studies represent a powerful tool for relating the biological activities of compounds to their physicochemical properties, which are referred to as descriptors [1] QSAR has been traditionally perceived as a means of establishing correlation between trend in the chemical structure modifications and respective changes of biological activity, Computational chemistry has developed into an important contributor to rational drug design. Quantitative structure activity relationship (QSAR) modelling results in a quantitative correlation between chemical structure and biological activity. [2] In the field of molecular modeling, docking is a method which predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complex.[3] Knowledge of the preferred orientation in turn may be used to predict the strength of association or binding affinity between two molecules using for example scoring functions. Molecular docking can be thought of as a problem of "lock-and-key", where one is interested in finding the correct relative orientation of the "key" which will open up the "lock" (where on the surface of the lock is the key hole, which direction to turn the key after it is inserted, etc.). Here, the protein can be thought of as the "lock" and the ligand can be thought of as a "key". Molecular docking may be defined as an optimization problem, which would describe the "best-fit" orientation of a ligand that binds to a particular protein of interest. However since both the ligand and the protein are flexible, a "hand-in-glove" analogy is more appropriate than "lock-and-key". [4] The renin-angiotensin system (RAS) is of major importance in cardiovascular and renal regulation and has been an attractive target in drug

*Corresponding author: mukesh2206@rediffmail.com

discovery for a long time. The renin-angiotensin system (RAS) plays a key role in regulating cardiovascular homeostasis and electrolyte/ fluid balance in normotensive and hypertensive subjects The main receptors involved in the RAS are the Angiotensin type-1 ($AT_1$ and type-2 ($AT_2$) receptors, both activated by the endogenous octapeptide angiotensin II (AngII) [5] The $AT_1$ receptor is well studied and is mainly known for regulating blood pressure and there are a number of $AT_1$ receptor selective antagonists on the market today used for treating hypertension. The $AT_2$ receptor is less well-understood and interestingly the stimulation of the $AT_1$and $AT_2$ receptors has, in many cases, direct opposing effects [6] The clinical success achieved by angiotensin converting enzyme (ACE) inhibitors in the treatment of hypertension and congestive heart failure has made the RAS a major focus for the discovery of novel antihypertensive agents However, ACE also has kininase activity, and this lack of specifity has been implicated in the occasional side effects of ACE inhibitors such as dry cough and angioedema With the development of ang II receptor antagonists, a more specific attempt to inhibit the activity of the RAS has become the main pharmacological approach. All major pharmaceutical companies embarked on a fast follower program immediately thereafter. Today, irbesartan, candesartan, and valsartan are all established in the market, and others, e.g., tasosartan and telmisartan, are following closely (Figure 1). Some further 20 compounds are in development. Most of these compounds share the biphenyl tetrazole unit or replacements thereof with the original, advanced lead losartan. [7] Almost all of the chemical manipulations within the fundamental skeleton of sarans concerned the substitution of the imidazole ring of losartan with several variously substituted hetero aromatic groups or acyclic structures.

The Renin Angiotensin system plays a central role in the regulation of blood pressure and electrolyte balance. The inhibition of this system has been an important target in antihypertensive therapy.

Angiotensin II receptor Antagonists supersedes the use of ACE inhibitors because of the following reasons:

- They block the conversion of Angiotensin I to Angiotensin II but they do not block the production of Angiotensin II by non- ACE mechanisms.
- Angiotensin II levels are not suppressed completely.
- Angiotensin II levels decrease temporarily but they return to normal levels after few weeks.
- They lead to accumulation of bradykinin levels on the biological fluids.

Adverse effect like angio-edema and dry cough are reported.

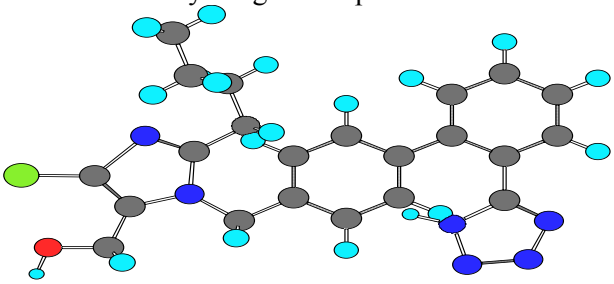


*Fig: 1 Losartan*

## 2. Experimental

We report our attempt to rationalize the physico-chemical and structural features among series of substituted 2-butylbenzimidazoles bearing biphenyl methyl moiety derivatives in relation to their angiotensin II receptor antagonist's activity using QSAR approach. The angiotensin II receptor antagonists activity data of 2-butylbenzimidazoles derivatives were taken from the reported work of Kubo et al. [8] and are presented in Table 1. The biological activity data ($IC_{50}$ in Molar) were converted in to $pIC_{50}$ according to the formula $pIC_{50}$ = (-log ($IC_{50}$). Thus such studies may help for the design and synthesis of better angiotensin II receptor

antagonists as antihypertensive agents. All the thirty four compounds were built on workspace of molecular modeling software V-Life MDS 3.5, which is a product VLife Sciences Pvt Ltd., India [9].The compounds were then subjected to conformational analysis and energy minimization using montocarlo conformational search with RMS gradient of 0.001 kcal/mol and iteration limit of 10000 using a MMFF94 force field. Montocarlo conformational search method is similar to the RIPS method that generates a new molecular conformation by randomly perturbing the position of each coordinate of each atom in molecule, followed by energy minimization and optimization is necessary process for proper alignment of molecules around template. Most stable structure for each compound was generated after energy minimization and used for calculating various physico-chemical descriptors like thermodynamic, steric and electronic. The various descriptors selected for 2D QSAR were vdWSurfaceArea (van der Waals surface area of the molecule), –vePotential Surface Area (total van der Waals surface area with negative electrostatic potential of the molecule), +vePotentialSurfaceArea (total van der Waals surface area with positive electrostatic potential of the molecule) dipole moment, YcompDipole (y component of the dipole moment), element count, slogP, path count, cluster, distance based topological indices, connectivity index, hydrophobic and hydrophilic areas like SA Most Hydrophilic (Most hydrophilic value on the vdW surface by Audry Method using Slogp), SAMostHydrophobicHydrophilic Distance (distance between most hydrophobic and hydrophilic point on the vdW surface by Audry Method using Slogp), SAHydrophilicArea (vdW surface descriptor showing hydrophilic surface area by Audry Method using SlogP) and SKMostHydrophilic (Most hydrophilic value on the vdW surface by Kellog Method using Slogp), radius of gyration, Wiener's index, moment of inertia, semi-empirical descriptors, HOMO (Highest occupied molecular orbital), LUMO (lowest unoccupied molecular orbital), heat of formation and ionization potential. Besides these all alignment independent descriptors were also calculated. The hydrophobic descriptors govern the movement of a drug molecule across the biological membranes in order to interact with the receptor by van der Waals binding forces whereas both electronic and steric descriptors influence the affinity of a drug molecule necessary for proper drug- receptor interaction. The optimal training and test sets were generated by either random selection method or the sphere exclusion algorithm. A commonly used ratio of training to validation objects (test set), which was also adopted in this work, is 70%: 30% [10]. However, rational splitting was accomplished by applying a sphere-exclusion type algorithm [11-14]. In classical sphere-exclusion algorithm the molecules are selected whose similarities with each of the other selected molecules are not higher than a defined threshold. Each selected molecule generates a hyper-sphere around itself, so that any molecule inside the sphere is excluded from the selection in the train set and driven toward the test set. The number of compounds selected and the diversity among them can be determined by adjusting the radius of the sphere (R).

**Statistical analysis**

Models were generated by using three significant statistical methods, namely, partial least square analysis, multiple regressions, and principle component analysis. The cross-validation analysis was performed using the leave-one-out method. In the selected equations, the cross-correlation limit was set at 0.5, the number of variables at 10, and the term selection criteria at $r^2$. An F value was specified to evaluate the significance of a variable. The higher the F value, the more stringent was the significance level: F test ''in'' as 4 and F test ''out'' as 3.99. The variance cutoff was set at 0, and scaling was auto scaling in which the number of random iterations was set at 100.The following statistical parameters were considered for comparison of the generated QSAR models: correlation coefficient (r), squared correlation coefficient (r2), predictive r2 for external test set (pred r2) for external validation, and Fischer's (F).The predicted r2 (pred_r2) value was calculated using Eq. 1, where yi and yˆi are the actual and predicted activities of the ith molecule in the test set, respectively, and ymean is the average activity of all molecules in the training set. Both summations are over all molecules in the test set. The pred_$r^2$ value indicates the predictive power of the current model for the external test set as follows

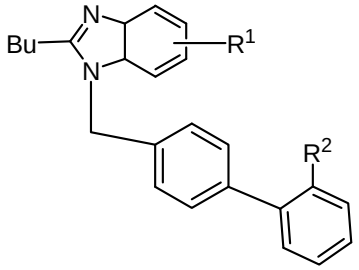
$$\text{pred\_}r^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - y_{mean})^2} \quad (1)$$

Internal validation was carried out using leave-one-out (q2, LOO) method. For calculating q2, each molecule in the training set was eliminated once and the activity of the eliminated molecule was predicted by using the model developed by the remaining molecules. The q2 was calculated using the equation which describes the internal stability of
a model:

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

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.

*Table.1 The substituted 2-butylbenzimidazoles derivatives for angiotensin II antagonists and their $IC_{50}$ values*




| Comp. | $R^1$ | $R^2$ | $IC_{50}$ (M) | -log $IC_{50}$ |
|---|---|---|---|---|
| 13a | H | Tet | 9.0 | 0.954243 |
| 13b | 5-OMe | Tet | 9.1 | 0.959041 |
| 13c | 6-OMe | Tet | 11 | 1.041393 |
| 13d | 5-Cl | Tet | 15 | 1.176091 |
| 13e | 6-Cl | Tet | 31 | 1.491362 |
| 13f | 7-OMe | Tet | 28 | 1.447158 |
| 14a | 4-$C0_2$Me | Tet | 72 | 1.857332 |
| 14b | 5-$C0_2$Me | Tet | 7.4 | 0.869232 |
| 14c | 6-$C0_2$Me | Tet | 4.4 | 0.643453 |
| 14d | 7-$C0_2$Me | Tet | 3.2 | 0.50515 |
| 14e | 5-Me-7-$C0_2$Me | Tet | 8.7 | 0.939519 |
| 14f | 5-Cl-7-$C0_2$Me | Tet | 4.4 | 0.643453 |
| 14g | 6-Me-7-$C0_2$Et | Tet | 9.1 | 0.959041 |
| 14h | 4-$CONH_2$ | Tet | 130 | 2.113943 |
| 141 | 7-$C0_2$Et | Tet | 14 | 1.146128 |
| 14j | 7-$C0_2$Bu | Tet | 12 | 1.079181 |
| 15a | 4-$C0_2$H | Tet | >100 | 2 |
| 15b | 5-$C0_2$H | Tet | 55 | 1.740363 |
| 15c | 6-$CO_2$H | Tet | 90 | 1.954243 |
| 15d | 7-$C0_2$H | Tet | 5.5 | 0.740363 |
| 15e | 5-Me-7-$C0_2$H | Tet | 13 | 1.113943 |
| 15f | 5-Cl-7-$C0_2$H | Tet | 11 | 1.041393 |
| 15g | 6-Me-7-$C0_2$H | Tet | 3.4 | 0.531479 |
| 16a | H | $C0_2$H | 11 | 1.041393 |
| 16b | 7-$C0_2$H | $C0_2$H | 6.6 | 0.819544 |
| 17 | 7-$C0_2$H | 1-Me-Tet | 34 | 1.531479 |
| 18 | 7-CONHi-Pr | Tet | 5.4 | 0.732394 |
| 27 | 7-$CH_2OH$ | Tet | 4.5 | 0.653213 |

| | | | | |
|---|---|---|---|---|
| 28 | 7-$CH_2OMe$ | Tet | 6.0 | 0.778151 |
| 29 | 7-$CH_2NMe_2$ | Tet | 24 | 1.380211 |
| 30 | 7-Me | Tet | 3.3 | 0.518514 |
| 31 | 7-$CH_2COiiEt$ | Tet | 2.5 | 0.39794 |
| 32 | 7-OH | Tet | 11 | 1.041393 |
| 33 | 7-$CH_2C0_2H$ | Tet | 26 | 1.39794 |

**Multiple Linear Regression Analysis**

The stepwise multiple regression analyses were carried out using the statistical software openstat2, version 6.5.1, designed and standardized by Bill Miller and Stat Val. Correlation matrix was obtained to justify the use of more than one variable in the study. The variables used were with maximum correlation to activity and minimum inter-correlation with each other. From the statistical viewpoint, the ratio of the number of samples (N) to the number of variables used (M) should not be very low; usually it is recommended that N/M≥5.

*Table 2. Selected physico-chemical parameters of 2-butylbenzimidazole.*

| H-Donor Count | XlogP | slogp | H cou. | kappa3 | T_C_N_5 | T_N_N_5 | T_O_Cl_5 | T_N_O_4 | T_C_Cl_5 | T_2_Cl_6 | T_2_O_4 | T_N_O_4 | T_C_N_2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 6.348 | 5.2743 | 0 | 5.2394 | 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 11 |
| 2 | 6.221 | 5.0032 | 1 | 5.9568 | 19 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 11 |
| 2 | 6.221 | 5.0032 | 1 | 5.9568 | 19 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 11 |
| 1 | 6.967 | 5.9277 | 0 | 5.4820 | 17 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 11 |
| 1 | 6.967 | 5.9277 | 0 | 5.4820 | 17 | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 11 |
| 2 | 6.221 | 5.0032 | 1 | 5.7969 | 18 | 0 | 0 | 1 | 0 | 0 | 2 | 1 | 11 |
| 2 | 4.139 | 4.7276 | 2 | 6.0277 | 18 | 0 | 0 | 2 | 0 | 0 | 6 | 2 | 11 |
| 2 | 4.139 | 4.7276 | 2 | 6.1895 | 19 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 11 |
| 2 | 4.139 | 4.7276 | 2 | 6.1895 | 19 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 11 |
| 2 | 4.139 | 4.7276 | 2 | 6.0277 | 18 | 0 | 0 | 2 | 0 | 0 | 4 | 2 | 11 |
| 2 | 4.42 | 5.0360 | 2 | 6.2702 | 19 | 0 | 0 | 2 | 0 | 0 | 4 | 2 | 11 |
| 2 | 4.758 | 5.381 | 2 | 6.2702 | 18 | 0 | 2 | 2 | 2 | 0 | 4 | 2 | 11 |
| 2 | 4.42 | 5.0360 | 2 | 6.1124 | 19 | 0 | 0 | 2 | 0 | 0 | 4 | 2 | 11 |
| 3 | 5.072 | 3.8995 | 2 | 5.7969 | 19 | 1 | 0 | 1 | 0 | 0 | 3 | 1 | 12 |
| 2 | 4.589 | 5.1177 | 2 | 6.1124 | 19 | 0 | 0 | 2 | 0 | 0 | 4 | 2 | 11 |
| 2 | 5.489 | 5.8979 | 2 | 6.9135 | 19 | 0 | 0 | 2 | 0 | 0 | 4 | 2 | 11 |
| 2 | 5.92 | 4.9725 | 2 | 5.7969 | 17 | 0 | 0 | 2 | 0 | 0 | 6 | 2 | 11 |
| 2 | 5.92 | 4.9725 | 2 | 5.9568 | 18 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 11 |
| 2 | 5.92 | 4.9725 | 2 | 5.9568 | 18 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 11 |
| 2 | 5.92 | 4.9725 | 2 | 5.7969 | 17 | 0 | 0 | 2 | 0 | 0 | 4 | 2 | 11 |
| 2 | 6.201 | 5.2809 | 2 | 6.0277 | 18 | 0 | 0 | 2 | 0 | 0 | 4 | 2 | 11 |
| 2 | 6.539 | 5.6259 | 2 | 6.0277 | 17 | 0 | 2 | 2 | 1 | 0 | 4 | 2 | 11 |
| 2 | 6.201 | 5.2809 | 2 | 5.8721 | 18 | 0 | 0 | 2 | 0 | 0 | 4 | 2 | 11 |
| 1 | 6.144 | 5.7924 | 2 | 5.2577 | 5 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 7 |
| 1 | 7.01 | 6.8255 | 2 | 5.8142 | 7 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 7 |
| 2 | 0 | 6.0107 | 2 | 5.7225 | 17 | 0 | 0 | 2 | 0 | 0 | 4 | 2 | 13 |
| 3 | 7.818 | 6.1107 | 2 | 8.1432 | 22 | 1 | 0 | 1 | 0 | 0 | 2 | 1 | 13 |
| 1 | 6.038 | 5.0868 | 1 | 5.5555 | 17 | 0 | 0 | 1 | 0 | 0 | 2 | 1 | 11 |
| 1 | 4.912 | 5.7977 | 1 | 6.0277 | 19 | 0 | 0 | 1 | 0 | 0 | 2 | 1 | 11 |
| 2 | 6.527 | 5.2319 | 1 | 6.3579 | 21 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 13 |
| 1 | 6.629 | 5.5827 | 0 | 5.3265 | 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 11 |
| 2 | 6.9 | 5.8525 | 2 | 6.5025 | 20 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 11 |
| 2 | 5.974 | 4.9799 | 1 | 5.3265 | 17 | 0 | 0 | 1 | 0 | 0 | 3 | 1 | 11 |
| 2 | 6.021 | 4.9014 | 2 | 6.3579 | 18 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 11 |

The QSAR equations were constructed for efficacy data of both species of malarial parasite with the physcio-chemical descriptors and indicator variables. The statistical quality of the equations[15] was judged by the parameters like correlation coefficient (r), explained variance ($r^2$), standard error of estimate(s) and the variance ratio or overall significance value (F).The accepted equations are validated for stability and predictive ability using "leave –one-out" and cross validation technique. The statistical parameters used to access the quality of the models are the predictive sum of squares (PRESS) of validation. Finally, the standard cross-validation correlation coefficient $r^2$ and $q^2$ are also calculated.

$$PRESS = \Sigma (Ypred - Y\ obs)^2$$

$$S_{press} = \sqrt{PRESS/ (n-k-1)}$$

n= no. of compounds used for cross-validation

$y_i$= experimental value of the physic-chemical property for the ith sample

y= value predicted by the model built without the sample i

*Table 3. Observed, and predicted antihypertensive activity.*

| Com.No | Actual Activity | Observed Activity | Predict.1 | Predict.2 | Predict.3 | Predict.4 |
|---|---|---|---|---|---|---|
| 13a | 9.0 | 0.954243 | 1.0532 | 1.1101 | 1.18719 | 1.03812 |
| 13b | 9.1 | 0.959041 | 0.94721 | 1.08213 | 1.151233 | 0.98321 |
| 13c | 11 | 1.041393 | 1.101504 | 1.15882 | 1.222731 | 1.14378 |
| 13d | 15 | 1.176091 | 1.182841 | 1.22126 | 1.196317 | 1.23717 |
| 13e | 31 | 1.491362 | 1.3922 | 1.4174 | 1.512571 | 1.47925 |
| 13f | 28 | 1.447158 | 1.471948 | 1.49528 | 1.390786 | 1.46315 |
| 14a | 72 | 1.857332 | 1.826863 | 1.92157 | 1.932962 | 1.80435 |
| 14b | 7.4 | 0.869232 | 0.964136 | 0.99934 | 1.059279 | 0.927331 |
| 14c | 4.4 | 0.643453 | 0.774013 | 0.726352 | 0.76321 | 0.71631 |
| 14d | 3.2 | 0.50515 | 0.605624 | 0.73877 | 0.728897 | 0.53812 |
| 14e | 8.7 | 0.939519 | 0.886349 | 0.9789 | 0.895442 | 0.96147 |
| 14f | 4.4 | 0.643453 | 0.908205 | 0.754828 | 0.714895 | 0.67352 |
| 14g | 9.1 | 0.959041 | 1.0123 | 1.016205 | 1.10525 | 1.02156 |
| 14h | 130 | 2.113943 | 2.144284 | 2.1472 | 2.06418 | 1.99562 |
| 141 | 14 | 1.146128 | 0.99816 | 1.083347 | 1.0981 | 1.2138 |
| 14j | 12 | 1.079181 | 1.171158 | 1.15491 | 1.434366 | 1.12812 |
| 15a | >100 | 2 | 1.1995 | 2.11423 | 1.890381 | 2.13814 |
| 15b | 55 | 1.740363 | 1.670321 | 1.71104 | 1.809287 | 1.921074 |
| 15c | 90 | 1.954243 | 1.81162 | 1.863081 | 2.012542 | 1.892743 |
| 15d | 5.5 | 0.740363 | 0.79204 | 1.02137 | 0.770821 | 0.92814 |
| 15e | 13 | 1.113943 | 1.06932 | 0.924116 | 1.163231 | 1.249701 |
| 15f | 11 | 1.041393 | 1.169151 | 1.148119 | 1.13196 | 1.09381 |
| 15g | 3.4 | 0.531479 | 0.700019 | 0.68137 | 0.815886 | 0.497264 |
| 16a | 11 | 1.041393 | 0.967553 | 1.182836 | 1.12429 | 1.168932 |
| 16b | 6.6 | 0.819544 | 0.841858 | 0.741724 | 0.909735 | 0.936145 |
| 17 | 34 | 1.531479 | 1.616548 | 1.468433 | 1.507579 | 1.612678 |
| 18 | 5.4 | 0.732394 | 1.09627 | 0.81376 | 0.93355 | 0.81482 |
| 27 | 4.5 | 0.653213 | 0.723771 | 0.976807 | 0.69405 | 0.746148 |
| 28 | 6.0 | 0.778151 | 0.73353 | 0.885394 | 0.846768 | 0.7179 |
| 29 | 24 | 1.380211 | 1.291442 | 1.41972 | 1.416597 | 1.447813 |
| 30 | 3.3 | 0.518514 | 0.64276 | 0.57215 | 0.66381 | 0.59025 |
| 31 | 2.5 | 0.39794 | 0.769545 | 0.41765 | 0.679956 | 0.552789 |
| 32 | 11 | 1.041393 | 1.367292 | 1.134277 | 1.191203 | 1.10386 |
| 33 | 26 | 1.39794 | 1.40951 | 1.34281 | 1.446217 | 1.371367 |

## 3. Results and discussion

Rationalization of physicochemical characters for antihypertensive activity was performed with the help of 33 compounds of 2-butylbenzimidazoles derivatives analogs using regression analysis technique (partial least square analysis, multiple regressions, and principle component analysis) were applied to generate models. The sphere exclusion method was adopted for division of the training and test sets.

Uni-Column Statistics: Training set

| Column Name | Average | Max | Min | StdDev | Sum |
|---|---|---|---|---|---|
| IC_50 | 5.9656 | 6.6021 | 5.0000 | 0.4037 | 149.1410 |

Uni-Column Statistics: Test set

| Column Name | Average | Max | Min | StdDev | Sum |
|---|---|---|---|---|---|
| IC_50 | 5.6378 | 6.3565 | 4.8861 | 0.5032 | 50.7406 |

The unicolumn statistical analysis can be interpreted to mean that the standard deviation of the training set is higher than that of the test set, indicating that there is a wide spread of activity in the training set with respect to the means compared with the test set. The mean and standard deviation for the training and test sets provide insight into the relative difference in the mean and point density distribution of the two sets. The minimum and maximum values in both the training and test sets are compared such that the maximum of the test set should be less than that of the training set. The minimum of the test set should be greater than that of the training set, suggesting the interpolative behavior of the test set (i.e., derived within the minimum–maximum range of the training set).

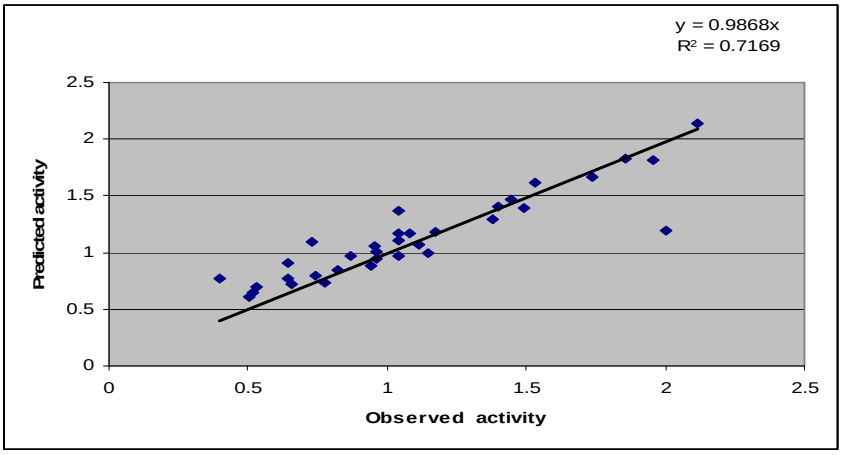
**Model 1**

**$Log_{10}$(IC_50) = +** 0.4250 Hydrogen Count -0.1844 T_2_O_4 + 0.6073 T_N_O_4 +0.2364 - 0.5364 T_C_N_2 + 1.2774 kappa3 + 7.0026 1

Optimum Components = 2, Degrees of Freedom = 15, n = 25, $r^2$= 0.7152, q2= 0.6753, F test = 22.3215 r2 se = 0.4321, q2 se = 0.4690, pred_$r^2$ = 0.6875, SEE = 0.112, SECV= 0.310, SEP=0.021, best_ran_$r^2$ = 0.325, best_ran_$q^2$ = 0.116, Zscore_ran_r2 =1.012, Zscore_ran_q2= 1.214, α_ran_$r^2$ = <0.0001, α _ran_$q^2$ = <0.001

Model 1 explains 71 % of the variance. The model shows an internal predictive (q2= 0.6753) of 67% and a predictivity for the external test (pred_$r^2$= 0.6875) of 69%. The overall statistical significance level was found to exceed 99.9%. In model 1, the influential descriptor kappa3, Hydrogen Count, T_C_N_2, T_N_O_4, T_2_O_4. The presence of a phenyl ring in the compound is favoured for activity. Thus, derivatives with an $5^{th}$ position Benzimidazoles analogs tend to be more potent than the compact substituents in that position. Antihypertensive activity of substituted Benzimidazoles analogs inferred that 5th position of benzimidazoles is more susceptible compared with 4th, 6th, and $7^{th}$ position for change in the activity and also suggested that substitution at 5 th position of benzimidazole, such as kappa3, was necessary for AII antagonistic activity, unsubstantiated 4th, 5th, and 6th were better than any substitution at this position. The contribution of groups also helps in understanding the binding of AII antagonist with $AT_1$ receptor by means of possible hydrogen bond interaction in between T_C_N_2 of benzimidazole.

Model.1



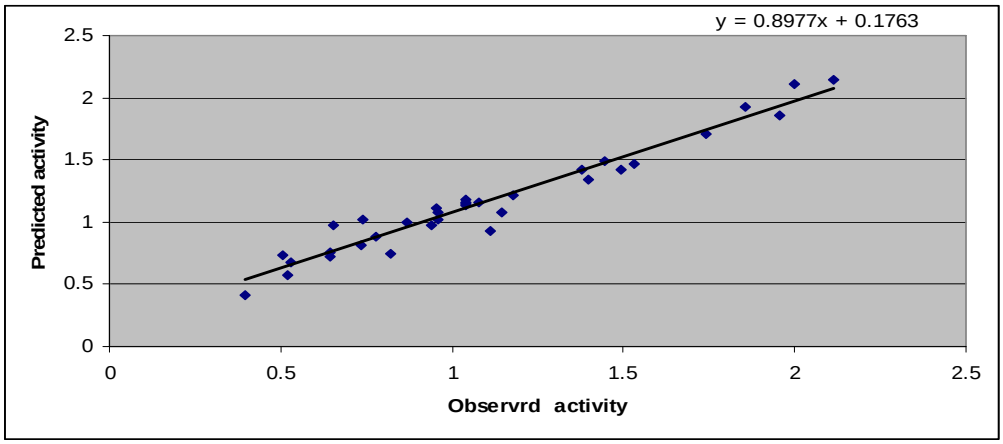

**Model.2**

**$Log_{10}$(IC_50) = +** 0.3838 slogp + 0.8714 T_2_Cl_6 + 0.6073 chi3Cluster +0.2364 +0.394 T_N_N_5 - 0.3801 T_C_N_5

Optimum Components = 2, Degrees of Freedom = 15, n = 25, $r^2$= 0.8152, q2= 0.7016, F test = 31.321, $r^2$ se = 0.3351, q2 se = 0.3690, pred_$r^2$ = 0.7295, SEE = 0.120, SECV= 0.211, SEP=0.190, best_ran_$r^2$= 0.165, best_ran_$q^2$ =0.216, Zscore_ran_r2 =0.231, Zscore_ran_q2= 0.102, α_ran_$r^2$ =<0.0001, α_ran_$q^2$ = <0.001

Model 2 generated using the multiple regression analysis method with 0.8152 as the coefficient of determination (r2) was considered using the same molecules in the test and training sets. The model can explain 76.9 % of the variance in the observed activity values. The model shows an internal predictive power (q2= 0.7016) of 70% and predictivity for the external test set (pred_r2 = 0.7295) of about 72 %. The F test value of 31.321 shows the overall statistical significance level for 99.99% of the model. Model 2 also shows a positive correlation with slogp, T_2_Cl_6, T_N_N_5, and a negative correlation with T_C_N_5.

Model.2

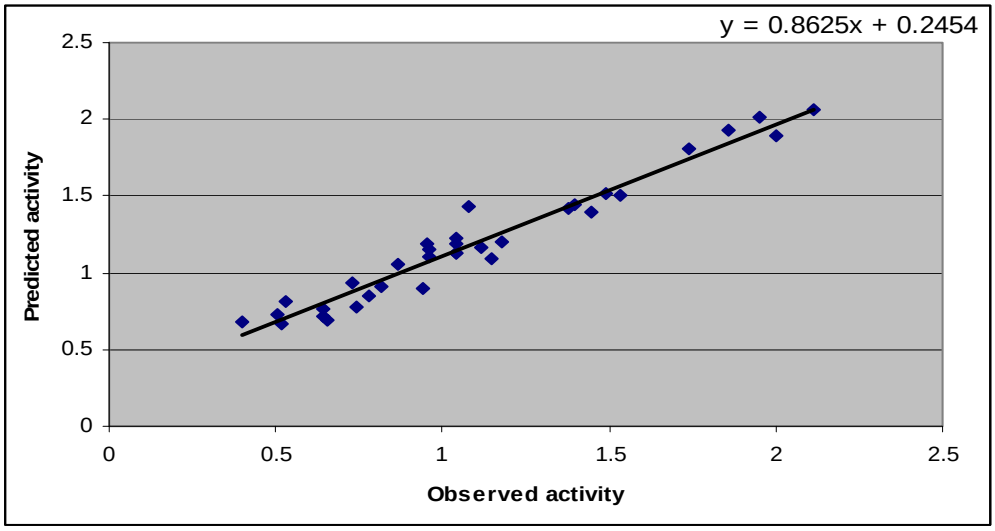



**$Log_{10}$ (IC_50) = +** 0.7838 H-Donor Count + 0.8714 T_2_Cl_6 + 0.6073 chi5chain+0.2364 +0.394 T_N_N_5 - 0.3801 T_2_O_4 + 4.2238

Optimum Components = 2, Degrees of Freedom = 12, n = 25, $r^2$= 0.7452, q2= 0.6273, F test 27.321 r2 se = 0.4251, q2 se = 0.3381, pred_$r^2$ = 0.7019, SEE = 0.191, SECV= 0.377, SEP=0.319, best_ran_$r^2$= 0.212, best_ran_$q^2$ =0.431, Zscore_ran_r2 =0.431
Zscore_ran_q2= 0.279, α_ran_$r^2$ =<0.00001, α_ran_$q^2$ = <0.01

Model 3 generated using PLS regression analysis method with 0.7421 as the Coefficient of determination ($r^2$) was considered using the same molecules in the test and training sets. The model can explain 70 % of the variance in the observed activity values. The model shows an internal predictive power (q2= 0.6273) of 60% and predictivity for the external test set (pred_r2 = 0.7019) of about 70 %. The F test value of 27.321 shows the overall statistical significance level for 99.99% of the model. Model 3 also shows a positive correlation with H-Donor Count, T_2_Cl_6, T_N_N_5, chi5chain and a negative correlation with T_2_O_4 and constant.
Model.3

Model.3




**Model.4**
**$Log_{10}$ (IC_50) = + 1**.2462 chi3Cluster + 0.4815 T_2_Cl_6 + 0.4313 XlogP +0.5186 T_N_N_5 - 0.4412 slogp + 1.221

Optimum Components =3, Degrees of Freedom = 17, n = 25, $r^2$= 0.8425, q2= 0.7151, F test 47.312 r2 se = 0.4511, q2 se = 0.2511, pred_$r^2$ = 0.8271, SEE = 0.191, SECV= 0.42576 , SEP=0.01027 , best_ran_$r^2$= 0.524, best_ran_$q^2$ =0.621, Zscore_ran_r2 =0.662, Zscore_ran_q2= 0.421, α_ran_$r^2$ =<0.00001, α _ran_$q^2$ = <0.01

Model 4 generated using PLS regression analysis method with 0.8425as the Coefficient of determination ($r^2$) was considered using the same molecules in the test and training sets. The model can explain 84 % of the variance in the observed activity values. The model shows an internal predictive power (q2= 0.7151) of 72 % and predictivity for the external test set (pred_r2 = 0.8271) of about 82.71 %. The F test value of 47.312 shows the overall statistical significance level for 99.99% of the model.

*Table: Correlation matrix of model-4*

| Parameters | chi3Cluster | T_2_Cl_6 | XlogP | T_N_N_5 | slogp |
|---|---|---|---|---|---|
| chi3Cluster | 1.0000 | | | | |
| T_2_Cl_6 | 0.22164 | 1.0000 | | | |
| XlogP | 0.6517 | 0.729153 | 1.0000 | | |
| T_N_N_5 | 0.4403 | 0.422649 | 0.306134 | 1.0000 | |
| slogp | 0.3751 | 0.1632 | 0.42854 | 0.4358 | 1.000000 |

**Docking Studies**

The structure of the ligand was prepared in MOL2 format using the 2D sketcher module of V-Life MDS 3.5 and MMFF 94 charges were assigned to the ligand atoms. The ligand was subjected to minimization until converged to a maximum derivative of 0.001 kcal_A ˚ A docking study was carried out in order to explain the experimentally observed low $AT_1$ affinity. For this purpose a previously developed model of the $AT_1$ receptor was used.[16] For the definition of the active site, several site-directed mutagenesis studies have been performed in order to examine the residues involved either in ligand binding (agonists or antagonists) or in signal transduction.[17-27] these studies are some times controversial and do not establish unequivocally the active site of the AT1 receptor [17,19,23,24,27] Other similar studies [28-31] a binding profile was proposed in the previous report where n-butyl chain interacts with L1 lined by Phe182, Phe171, Ala163 and Ser103 and the biphenyl system interacts with L2 lined by Val108 through Van der Waal interactions while N-3 of benzimidazole nucleus and terminal carboxyl group interact through H-bonding with Tyr113 and Tyr184 H-bonding interactions The key amino acids for binding of nonpeptide $AT_1$ antagonists are Val108,Ser 109, Asn111,Leu 112,Tyr 113, Ala163, Arg167,Ala 181, His 183,Lys199, Ser252, Trp 253,His256,and Asn295,. These amino acids have been reported for specific interactions with non-peptide antagonists of the class of SARTANs. Other amino acids mostly in the third and seventh helices are important for the maintenance of the integrity of the pocket. These reported amino acids cover a wide area of the receptor thus cannot determine a unique binding pocket. The pocket selected for further docking studies was the one which included three of the most important amino acids (Val108, Lys199, Trp 253 and His256) for antagonist binding. This pocket is located near the polar head groups and exceeds down to the hydrophobic core of the phospholipid bilayers of the membrane in which the receptor is embedded. Thus the specific pocket may be easily accessible by the ligands after their diffusion which confirms the proposed two-step mechanism of action for $AT_1$ antagonists. All the others residues were comprised in a limited region compatible for the interaction with ligand .Furthermore aligning the $AT_1$ receptor model with the bovine rhodopsin crystal structure came out that the residues listed above (principally Valine (108), Lys (199), and Leu (265)) delimited a region that corresponded to the binding site of retinal. The best docked structure of 14F minimum energy (-21.382) ligand highlighted an interaction of the methyl,carboxyl,chlorine substituent with Lys (199), while Valine (108) interacted with the biphenyl system, and Ala 181 with the *n*-butyl substituent; Leu (265) to interact with the ligand. As regards the role of the tetrazole ring in AT1 receptor binding, all the recent models suggest an ionic interaction with Lys (199), The results obtained from the V-Life MDS 3.5[9] Biopredicta Tools docking studies showed that the binding disposition of the ligand was similar to that of losartan with the biphenyl ring located in the lipophilic pocket delimited by Valine (108), Valine 179, Trp (253), His (256),Ile(288), 2-butyl substituent interacting in the secondary lipophilic pocket created by Ser (109), Leu (112), Tyr (113), Asn (168). However as shown in (**Figure3,4,5,6**) with respect Ligand 14F was interaction with Hydrogen Bond, Hydrophobic interaction, PI Stacking interaction, vdW interaction.

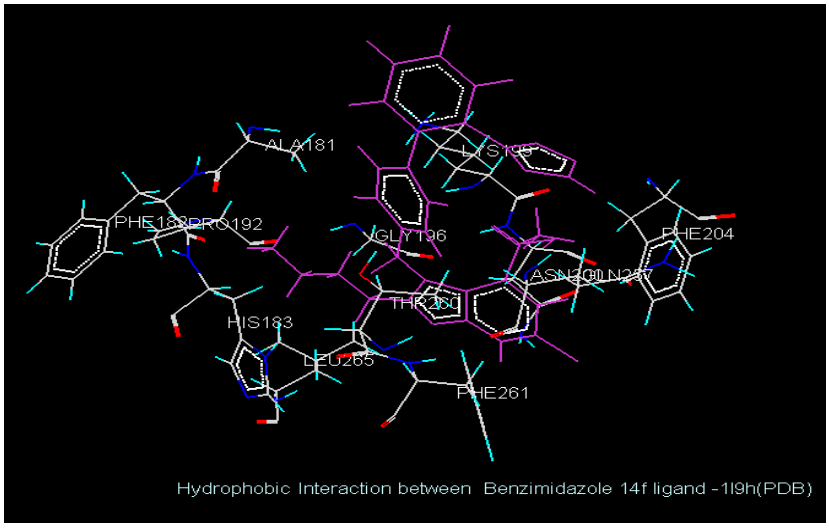



*Fig: 3 ligand – receptor Interaction Hydrophobic Interactions*

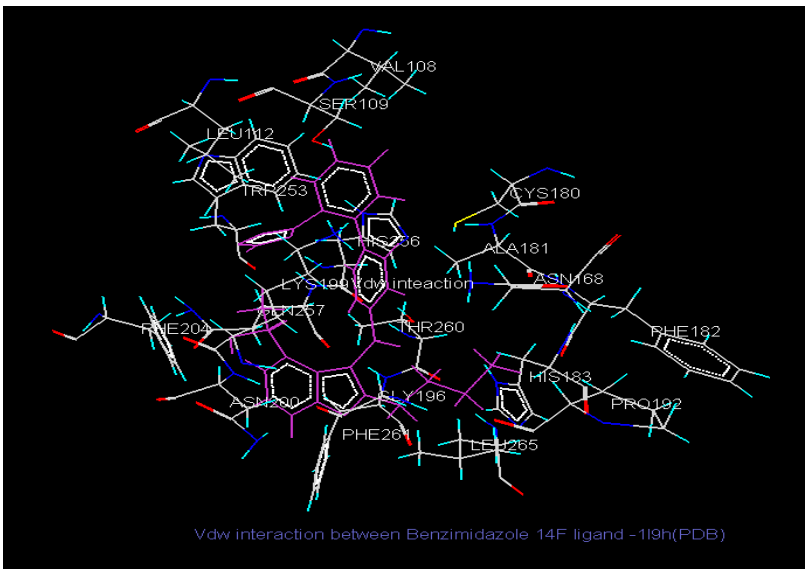



*Fig 4: ligand – receptor Interaction Vdw Interaction*

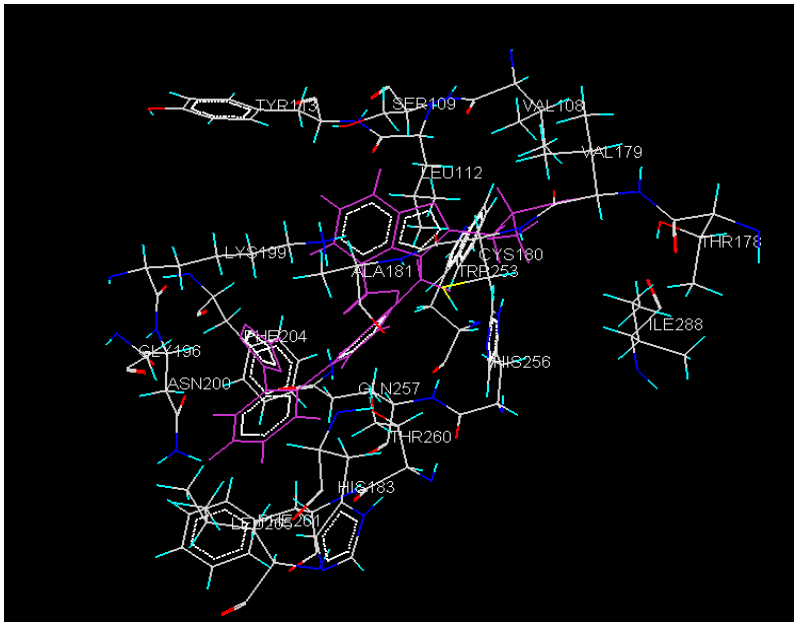



*Fig: 5 Hydrogen Bond, Vdw interaction, Hydrophobic, ligand – receptor Interaction*

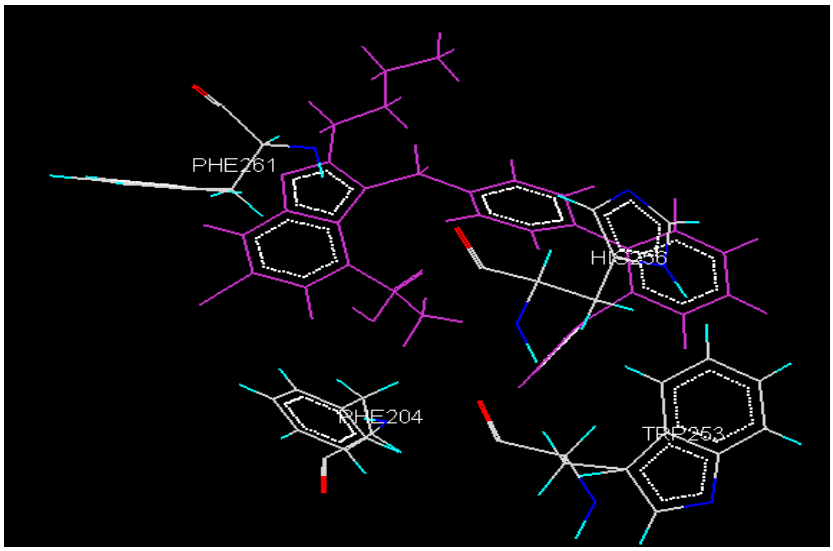



*Fig: 6 PI Staking ligand – receptor Interaction*

## 4. Conclusions

The 2D QSAR studies were conducted with a series of angiotensin II antagonists, and some useful molecular models were obtained. The physicochemical and Alignment-independent descriptors were found to have an important role in governing the change in activity. Hence, these models are very useful for further optimization of antihypertensive activities. The current study provides better insight into the designing of more potent antihypertensive agents in the future before their synthesis. All the n-butyl chains carboxylate, phenyl and tetrazole ring substituents were treated as fully ionized groups. For each ligand, the best docked structure was chosen, and this receptor-based alignment was used for further studies. The docking of best molecule into the AT1 receptor confirmed that Tyr 113, ser 109, Lys (199), val (108), and Asn (168), His 256, Trp

253 interacted with the receptor. this study illustrates a new hypothesis about the binding interaction of non-peptide antagonists inside the AT1 receptor, encouraging further investigations on new residues that might be fundamental for the ligand-receptor interaction. Furthermore, because AT1 antagonists are an interesting therapeutic target.

**Acknowledgement**

The author thanks V-life Science Technologies Pvt. Ltd Pune [India] for providing the software for the study,

## Reference

[1] Hansch, C.; Fujita, J. Am. Chem. Soc. **86**, 1616-1630. (1964).
[2] X.J. Yao, J.P. Panaye, J.P. Doucet, R.S. Zhang, H.F. Chen, M.C. Liu, Z.D. Hu, B.T. Fan, J. Chem. Inf. Comput. Sci. **44** 1257, 1266 (2004)
[3] Jorgensen WL, Science **254** 5034 954,(1991).
[4] M. de Gasparo, K.J. Catt, T. Inagami, J.W. Wright, T. Unger, Pharmacol. Rev. **52**, 415 (2000).
[5] M. Stoll, T. Unger, Angiotensin and its AT2 receptor: new insights into an old system, Regul. Pept. **99** 175 (2001).
[6] Israili, Z. H., J. Hum. Hypertens. **14** (Suppl. 1), S73-S86 [2000]
[7] Keiji Kubo:, Yoshiyuki Inada, Yasuhisa Kohara, Yoshihiro Sugiura. Mami Ojima, Katsuhiko Itoh, Yeshivas Furukawa, Kohei Nishikawa, and Takehiko Naka J. Med. Chem. **36**,1772-1784 (1993).
[8] V-Life MDS 3.5, Molecular Design Suite, VLife Sciences Technologies, Pvt. Ltd. Pune, India, (2003).
[9] A.K. Chakraborti, B. Gopalakrishnan, M.E. Sobhia, A. Malde, Eur. J. Med. Chem. **38**, 975 (2003).
[10] A. Coi, F.L. Fiamingo, O. Livi et al., Bioorg. Med. Chem. **17**, 319 (2009).
[11] A. Golbraikh, A. Tropsha, J. Comp. Aid. Molec. Des. **16**, 357 (2002).
[12] A. Golbraikh, J. Chem. Inf. Comput.Sci. **40**, 414 (2000).
[13] R. Guha, J.R. Serra, P.C. Jurs, J. Mol. Graph. Model **23**, 1-14 (2004)
[14] H.Kubinyi.Quant.Struct.Act.Relat., **13**,285 (1994).
[15] A.Patny, Prashant V. Desai, and Mitchell A. Avery, PROTEINS: Structure, Function, and Bioinformatics **65**, 824–842 (2006).
[16] Fierens, F. L. P.; Vanderbeyden, P. M. L.; Gaborik, Z.; LeMinb, T.; DeBacker, J. P.; Hunyady, L.; Ijzerman, A.; Vauquelin, G. J. RAAS, **1**, 283 (2001).
[17] Hjorth, S. A.; Schambye, H. T.; Greenlee, W. J.; Schwartz, T. W. J. Biol. Chem. **269**, 30953, (1994).
[18] Hunyady, L.; Balla, T.; Catt, K. J. Trends Pharmacol. Sci. **17**, 135, (1996)
[19] Miura, S. I.; Feng, Y. H.; Husain, A.; Karnik, S. S. J. Biol. Chem. **274**, 7103 (1999).
[20] Monnot, C.; Bihoreau, C.; Conchon, S.; Curnow, K. M.; Corvol, P.; Clauser, E. J. Biol. Chem. **271**, 1507 (1996)
[21] Nirula, V.; Zheng, W.; Sothinathan, R.; Sandberg, K. Br. J. Pharmacol.,**119**, 1505 (1996)
[22] Noda, K.; Saad, Y.; Kinoshita, A.; Boyle, T. P.; Graham, R. M.; Husain, A.; Karnik, S. S. J. Biol. Chem **270**, 2284 (1995).
[23].Prendergast, K.; Adams, K.; Greenlee, W. J.; Nachbar, R.B.; Patchett, A. A.; Underwood, D. J. J. Comput. Aided Mol. Des., **8**, 491 (1994).
[24] Schambye, H. T.; Hjorth, S. A.; Weinstock, J.; Schwartzn, W. Mol. Pharmacol. **47**, 425 (1995).
[25] Wang, C.; Jayadev, S.; Escobedo, J. A. J. Biol. Chem. **270**, 16677 (1995).
[26] Zoumpoulakis, P.; Daliani, I.; Zervou, M.; Kyrikou, I.; Siapi, E.; Lamprinidis, G.; Mikros, E.; Mavromoustakos, T. Chem. Phys. Lip., 125, 13 [2003]

[27] Ashton, W. T.; Cantone, C. L.; Chang, L. L.; Hutchins, S. M.; Strelitz, R. A.; Mcross, M.;Chang, R. S. L.; Lotti, V. J.; Faust, K. A.; Chen, T.; Bunting, P.; Schorn, T. N.; Kivlighn, S. D.; Siegl, P. K. S. J. Med. Chem., **36**, 591 (1993).
[28] Datar, P. A.; Desai, P. V.; Coutinho, E. C. J. Chem. Inf. Comput. Sci, **44**, 210 (2004)
[29] Berrllini, G.; Cruciani, G.; Mannhold, R. J. Med. Chem., **48**, 4389 (2005).
[30] Tuccinardi,T.; Calderone, V.; Rapposelli, S.; Martinelli, A. J. Med. Chem. **49**, 4305 (2006).