INTERNATIONAL JOURNAL OF RESEARCH IN PHARMACY AND CHEMISTRY

Available online at www.ijrpc.com

Research Article

STUDIES ON STRUCTURAL INSIGHT OF 2-AMINO-6-ARYLSULFONYLBENZONITRILE DERIVATIVES AS ANTI HIV AGENTS

Utpal Chandra De[1*], Sudhan Debnath[2] and Debanjan Sen[3]

[1]Department of Chemistry, Tripura University, Suryamaninagar-799022, Tripura, India.
[2]Department of Chemistry, M.B.B. College, Agartala-799004, Tripura, India.
[3]Bengal institute of Pharmaceutical Sciences, Kalyani, West Bengal, India.

ABSTRACT

Comparative Molecular Field Analysis (CoMFA), Comparative Molecular Similarity Indices Analysis (CoMSIA) and docking studies were performed on a series of 2-amino-6-arylsulfonylbenzonitriles (AASBs) and congeners as selective as anti HIV agents. The statistically significant model was developed for the training set of 42 molecules and was validated by a test set of 18 compounds. The PLS analysis yielded the best predictive CoMFA model having R_{cv}^2 = 0.603, R_{nv}^2 (non-cross-validated) = 0.998, F value = 1346.132, R_{bs}^2 =0.999 with standard error of estimate (SEE) 0.05 and R_{pred}^2 = 0.796 while the CoMSIA model resulted R_{cv}^2 = 0.506, R_{nv}^2 (non-cross-validated) = 0.989, F value= 288.023, R_{bs}^2 = 0.995 with SEE 0.119 and R_{pred}^2 = 0.540. Results analysis indicated that steric, electrostatic, hydrophobic and hydrogen bonding feature plays a significant role in selectivity of the compounds to act as anti HIV agent. The contour maps obtained from 3D QSAR studies and docking analysis also supported the activity trend of the selected molecules. The development of ligand based 3D QSAR model, docking studies and subsequent structural insight analysis of selected compounds as anti HIV agent is discussed which could be helpful in anti HIV drug design.

Keywords: anti HIV agents, 3D QSAR, docking, CoMFA, CoMSIA.

INTRODUCTION

Human immunodeficiency virus (HIV) specifically damages the immune system and causes the disease called AIDS. It has developed a worldwide disastrous scenario as far as human health is concerned. World Health Organization (WHO) estimated that about 33.4 million [31.1 million–35.8 million] people are living with HIV worldwide in which about 3.7 million children's are less than 15 years old and they are inborn HIV infected mainly due to mother-child transmission. It is also estimated that about 2.7 million people were newly infected in 2008 and 2 million [1.7– 2.4 million] people died of AIDS related illness in 2008[1-3]. There are two species of HIV *viz*. HIV-1 and HIV-2 which generally attack humans' immune system and HIV-1 is the most disastrous as it easily get transmitted and causes the majority of HIV infections. During the last two decades, various anti-HIV-1 drugs *viz*. nucleoside and nucleotide reverse transcriptase inhibitors (NRTIs & NtRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs) and protease inhibitors (PIs) have been developed[4-10]. On the other hand Reverse Transcriptase (RT) inhibitors prevent completion of synthesis of the double-stranded viral DNA via binding to an allosteric hydrophobic pocket (non-

nucleoside RT inhibitor binding site) and blocks essential enzymatic function of protease. This key role of RT in the HIV-1 life cycle makes it as a major target for the development of anti HIV agents. The efficacy of RT inhibitors is severely limited by the emergence of HIV-1 drug-resistant mutants[11]. Therefore, the search for new, selective and potent drugs which will be able to inhibit mutant forms too remains a challenge. In this context we have studied CoMFA and CoMSIA analyses followed by docking for a series of 2-amino-6-arylsulfonylbenzonitriles (AASBs) and congeners to have structural insight for new anti HIV drug design. CoMFA was introduced by Cramer[12] assuming that the interaction between an inhibitor and its molecular target is primarily non-covalent in nature and shape dependent. CoMSIA is similar to CoMFA in terms of fields around the molecule which assumed that changes in binding affinities of ligands are related to changes in molecular properties represented by different fields. In addition to steric and electrostatic fields alike in CoMFA, CoMSIA also consider the hydrophobic, hydrogen bond donor and hydrogen bond acceptor fields to cover the major contributions to ligand for effective binding interaction with active site of the receptor[13].

MATERIAL AND METHOD

Data Set

A series of 60 compounds of 2-amino-6-arylsulfonylbenzonitriles and congeners with precise IC_{50} values as anti HIV agents were selected from the literature[14] where some descriptors (topological, geometrical and quantum-chemical) were generated from CODESSA to describe the molecules. Chemical structures and corresponding activities (experimental, predicted and residual activity) for the complete set of compounds (divided into training and test sets based on principal component analysis) are presented in Table I. The robustness and predictive ability of the models were evaluated by external validation using the test set of 18 compounds. The 3D QSAR studies were carried out in SYBYL[15] installed on an Intel core i3 processor, 3.06 GHz PC with Windows 7 Home Premium operating system. The 2D structure of the molecules were built in Chemdraw Ultra 11 software and subsequently converted into sybylmol2 format in chemDraw.

Data Set Alignment

A good alignment is the most essential element for CoMFA and CoMSIA analysis although a number of other factors of the aligned compounds may have bearing on results[16]. The quality and the predictive ability of a model is directly depends on the alignment rules. In the present study, superimposition of the molecules were carried out on the basis of 'common substructure' alignment procedure[17] using compound 52 (Table I) as template molecule. Two aromatic rings with sulphur linker were used as the common scaffold for alignment of the entire compound in the series. The optimization of the data set were performed using the energy minimize option (Tripos Force Field) with a 0.05 kcal/ (Å mol) energy gradient criterion and Gasteiger-Hückel charge command of SYBYL package. Each analog was aligned (Fig.1) to the template by rotation and translation to minimize the root mean square deviation between atoms in the template and the corresponding atoms in the analog using align database option in SYBYL package of Tripos, Inc.

Docking protocol

The structures of the entire compound were drawn in ChemBioDraw Ultra 11.0, converted to corresponding mole file and then imported to maestro project table. Imported ligands were prepared using LigPrep option of maestro 9.5 version[18] with OPLS_2005 force field for subsequent docking studies. While performing this step, chiralities were determined form 3D structure and original states of ionization were retained. Epik module was used for generating possible physicochemical states at target pH 7.2 ± 2. Complexes between RT and any of AASBs were not available and hence we used the 3D structure of RT complexes with 5-bromo-3-(pyrrolidin-1-ylsulfonyl)-1H-indole-2-carboxamide (PDB entry 2RF2)[19] as it shares some (less but comparatively more than others) similarities to the ligand dataset (Table I). Receptor optimization, grid generation followed by extra precession (XP) flexible docking of 1-60 (Table I) were performed successively using the 'protein preparation wizard', 'receptor grid generation' and 'glide' option respectively implemented in Maestro 9.2 software package. Protein Preparation Wizard module, is designed to ensure chemical correctness, hydrogen optimization and minimization of crude protein structure using OPLS_2005 force field and RMSD of 0.30 Å for distance tolerance. As the receptor was co-crystallized with a ligand, it was

excluded during receptor grid generation[20]. 400 best poses for each ligand from 10,000 poses which passes through the initial glide screen were kept for energy minimization. The energy minimization was performed in maximum of 100 steps with a distant dependent dielectric constant of 2.0. To soften the potential for non polar parts of the receptor, a scaling factor and partial charge cutoff was set as default value at 1.0 and 0.25 whereas for ligand these values were at 0.8 and 0.15. No constraint was set for ligands during docking.

CoMFA and CoMSIA Procedures

Standard CoMFA and CoMSIA procedures[21,22] were performed for the dataset (**Table I**) to correlate the biological activity (pIC_{50}) to their steric, electrostatic, hydrophobic and hydrogen bonding feature. Descriptor fields of both the CoMFA and CoMSIA were calculated by placing the aligned molecules in a 3D cubic lattice with grid spacing of 1 Å and extending to 4 Å units in all three dimensions within defined region. Steric and electrostatic fields were deduced using Tripos force field method. In CoMFA method a sp^3 hybridized carbon atom with +1 charge was used as a probe atom and 30 kcal mol^{-1} energy cut off was applied. CoMSIA fields such as steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor were deduced using same lattice grid employed for CoMFA prediction. The attenuation factor was set to the default value of 0.3 of SYBYL package.

Partial least square (PLS) analysis

PLS analysis[23,24] used to linearly correlate the CoMFA and CoMSIA fields to biological activity and was carried out by the leave-one-out (LOO) and leave-group-out (group of 10 compounds) cross-validation methods, respectively. The optimal number of components, usually corresponding to the highest cross-validated squared correlation coefficient (Q^2) (Table II), was selected on the basis of the lowest standard error of estimate (SEE). In order to avoid over-fitting, the model with higher component was accepted and used only when the q^2 differences between two components were larger than 10%. For further assessment of statistical confidence and robustness of the model, a 100-cycle bootstrap validation was performed wherein each run some ligands are excluded and/or included more than once and the mean correlation coefficient (R_{bs}^2) was calculated (Table II).

Model validation

In general high cross validated correlation coefficient (R_{cv}^2) for a 3D-QSAR model suggests the fitness of the model to have a high predictive power; however, in some QSAR studies urged that it is not always a sufficient condition[25-27] for high predictive abilities. Thus to confirm the predictive power of a model, it has to be tested with various external validation method. In the present study both CoMFA and CoMSIA predicted models were subjected to different statistical validation[28] such as calculation of a *b, k,* R_0^2, R^2 and R_{pred}^2. It is reported[33] that 3D-QSAR models generally considered acceptable if they satisfy all the conditions *viz.* $a \sim 1$, $b \sim 0$, $0.85 < k < 1.15$, $R_{cv}^2 > 0.5$, $R_{nv}^2 > 0.6$, $R_{bs}^2 > 0.8$, $R^2 > 0.6$ and $R_{pred}^2 > 0.6$ where R_{cv}^2 (= LOO cross-validated correlation coefficient), R_{nv}^2(=non cross-validated correlation coefficient), R_{bs}^2(= bootstrap validated correlation coefficient) for training set (internal validation); R^2(=correlation coefficient for regression between experimental and predicted activity) a (= regression coefficient), *b* (= intercept of regression line), *k* (= coefficient of regression through origin), R_0^2 (= correlation coefficient for regression through origin). If the regression equation for experimental activity ($Y_{exp(test)}$) and predicted activity ($Y_{pred(test)}$) of test set compounds is represented by $Y_{pred(test)} = a\, Y_{exp(test)} + b$, then R, 'a' and 'b' will be given by the following equations:

$$a = \frac{\sum (Y_{\exp(test)} - \bar{Y}_{\exp(test)})\ (Y_{pred(test)} - \bar{Y}_{pred(test)})}{\sum (Y_{pred(test)} - \bar{Y}_{pred(test)})^2}$$

$$b = (\bar{Y}_{\exp(test)} - a\bar{Y}_{pred(test)})$$

$$R^2 = 1 - \frac{\sum (Y_{\exp(test)} - Y_{pred(test)})^2}{\sum (Y_{\exp(test)} - \bar{Y}_{\exp(test)})^2}$$

Here $\bar{Y}_{\exp(test)}$ and $\bar{Y}_{pred(test)}$ are the mean experimental and mean predicted pIC_{50} values respectively, of test set compounds.

Again, if the equation of the regression line drawn through the origin (*ro*) is given by

$$Y^{ro}_{pred\,(test)} = KY_{\exp(test)}$$

then, the co-efficient of regression (*K*) and correlation co-efficient of regression through origin would be represented by the following equation:

$$K = \frac{\sum Y_{\exp(test)} Y_{pred\,(test)}}{\sum Y^2_{pred\,(test)}}$$

$$R^2_o = 1 - \frac{\sum (Y_{\exp(test)} - Y^{ro}_{pred\,(test)})^2}{\sum (Y_{\exp(test)} - \bar{Y}_{\exp(test)})^2}$$

The predictive correlation coefficient (R_{pred}^2) based on the test set molecules, was calculated using the following equation:

$$R^2_{pred} = \frac{\sum (Y_{\exp(test)} - \bar{Y}_{pred(trn)})^2 - \sum (Y_{\exp(test)} - Y_{pred(trn)})^2}{\sum (Y_{\exp(test)} - \bar{Y}_{pred(trn)})^2}$$

Where $Y_{pred\,(trn)}$ and $\bar{Y}_{pred\,(trn)}$ are the predicted and mean of predicted pIC_{50} values respectively, of training set compounds.

RESULTS AND DISCUSSION

CoMFA and CoMSIA

Various 3D QSAR models were generated using different molecules in the training set and five statistically significant models were recorded with significant CoMFA/CoMSIA parameters for 42 molecules included in the training set (Table I). The best model was selected based on the values of statistical parameters *viz* R_{cv}^2, R^2, R_{pred}^2 etc. (Table II, III).

Analysis of COMFA/CoMSIA model

The statistical parameters associated with CoMFA and CoMSIA models are listed in Table II & III. The best CoMFA model of dataset gave a cross-validated correlation coefficient (R^2_{cv}) of 0.603 (>0.5) with an optimum number of components (ONC) of 8 with column filtering value of 2.0 k.cal/mol, which is an indication of model's reliability and predictive ability. A high non-cross validated correlation coefficient (R_{nv}^2) of 0.998 with a standard error estimate (SEE) of 0.05, excellent F value of 1346.132 and predictive correlation coefficient (R^2_{pred}) of 0.796 also confirms the robustness and predictive ability of the model. The best CoMSIA model of same dataset gave a cross-validated correlation coefficient (R^2_{cv}) of 0.506 (>0.5) with an ONC of 8 and column filtering value of 2.0 k.cal/mol, which further indicated the model's reliability to predict the pIC_{50} values. A reasonably high value of non-cross validated correlation coefficient (R_{nv}^2) of 0.989 with a standard error estimate (SEE) of 0.119, excellent F value of 288.023 and predictive correlation coefficient (R^2_{pred}) of 0.540 also indicated model's reliability. The experimental and predicted pIC_{50} values of the training set and test sets are given in Table I, whereas the graph of experimental versus predicted pIC_{50} values of the training and test sets are illustrated in Fig. 2.

Discussion of the Contour Plots

CoMFA and CoMSIA results were graphically interpreted by field contribution maps. Coefficient contour maps using the field type *"StDev*Coeff"* were generated. To select the appropriate contour levels for each feature, the respective histograms of actual field values were analyzed. Contour levels that produced chemically meaningful contour maps were chosen. The contour maps of CoMFA model highlight those regions in space where the aligned molecules would favorably or unfavorably interact with a possible environment whereas the contribution maps of the CoMSIA approach denote those region occupied by the ligands that would 'favor' or 'disfavor' the presence of a group with a particular physico-chemical property. Contour diagrams, exposed with more active compound 52, of the CoMFA and CoMSIA models considering steric and electrostatic features are shown in figure 3 and figure 4. As the contour diagrams derived from CoMFA and CoMSIA are almost alike they are discussed together. Analyzing the CoMFA steric contour plot (Fig. 3a) reveals that the green tracing around the aromatic ring A indicate highly favorable steric effect and yellow tracing indicate disfavored steric factor for activity enhancement. The CoMSIA steric contour plot

(Fig. 3b) also shows almost a similar green and yellow region to that of CoMFA plot. In electrostatic contours of CoMFA (**Fig. 4a**) and CoMSIA (**Fig. 4b**), blue contour representing the regions where electro-positive substituents are advantageous while the red contour represents favored region for partial negative charge to have better binding interaction with active site of reverse transcriptase. The corresponding steric and electrostatic field contributions in best model were 0.628 and 0.372 (CoMFA), 0.158 and 0.149 (CoMSIA) which implied the dominance of steric field over electrostatic field for interactions to the active site of HIV-1 RT.

Selectivity Analysis

For selectivity analysis, the total database has been subdivided into three different groups (1-18; 19-32; 33-60) (Table I) depending on their skeletal structure. The green region, stretching out from the vicinity of 3 and 5-positions of A ring as observed in the CoMFA plot (Fig. 3a) for the dataset suggested that substitution at these area by bulky groups are favorable for better inhibition of RT. The experimental pIC_{50} values of highest active compounds 15, 29 and 52 (for three groups in Table I) demonstrated this feature in which these positions are occupied by methyl groups. The reasonably high experimental pIC_{50} values of compounds 16, 17 (16>17), 31, 53, 54, 55, 57, 58, 59, 60 (53>54>55>58=59>57>60) may be due to methyl substitutions at 5-position of ring A which is an indication of favorable steric effect involvement at this position for ligand-receptor binding interaction. Moreover, the experimental pIC_{50} values of compounds 8, 10, 11, 39, 42, 45 also indicated the activity enhancement due to steric reason at 3-position of ring A and activity follows (10>8>11) the order of substituents size (Br>Cl>F). Yellow contours near to 2 and 4-positions of ring A suggested that bulky substituents at these positions are highly unfavorable for activity. This was supported by the experimental pIC_{50} values of 2-substituted compounds 7, 9 (7>9 as size of Br>Cl), 38, 41, 44 (44>38>41 as size of Br>Cl>F), and 4-substituted compounds 40, 43 (40>43 as size of Br>Cl) in the selected dataset (Table I).

In electrostatic contour maps of CoMFA (Fig. 4a) and CoMSIA (Fig. 4b), the blue contour area (80% contribution) stretching out from 2-H and 4-H in ring A and from 2-position of ring B suggested that the partial electropositive charge in these region would increase the inhibitory activity of the selected compounds while red contours (20% contribution) in the vicinity of 3-H of ring A represent the favorable area for partial negative charge. These features clearly demonstrated the activity of compounds (2-position) 2, 4, 7, 9 (7>2>4>9), 19, 26 (26>19), 34, 36, 38, 41, 44, 46 (34>36>44>38>41>46) and (4-position) 6, 13 (13>6), 22, 25 (25>22), 37, 40, 43, 48 (40>37>43>48) where weak electron donating substituents (like $-OCH_3$) and electron withdrawing substituents (like F, Cl, Br) affected the activity. Moreover, the activity of 12, 14, 20, 27, 28, 35, 47, 49 where weak electron withdrawing substituents at 3-position of ring A resulted high inhibitory activity where it followed 12 >14, 20 ~ 27 > 28 as electron withdrawing effect follows the order of CF_3>CN~OCH_3. The CoMSIA steric contour (Fig. 3b) was in good agreement with the CoMFA steric contours and hence needles to discuss separately. The red contour in the vicinity of sulphonyl group suggested that partial negative charge in this region will be favorable for activity of the compounds. This feature was clearly demonstrated by the activity order of 15(S), 29(SO) and 52(SO_2) which followed the order of 52>29>15 as partial negative charge follows SO_2>SO>S.

CoMSIA Hydrophobic and H-bond Donor/Acceptor Regions

The hydrophobic, hydrogen bond donor and hydrogen bond acceptor featured contour maps of CoMSIA are displayed in **Fig. 5**. The corresponding hydrophobic, hydrogen bond donor and hydrogen bond acceptor fields' contribution were 0.369, 0.204, and 0.120 respectively, which implies that hydrophobic contribution is important for interactions with the active site of HIV-1 RT although hydrogen bond (donor and acceptor) contribution also played an important role in ligand receptor binding interactions.

In hydrophobic contour (**Fig. 5a)** the orange colour (80% contribution) near to 3 and 5 position of A ring indicated favorable hydrophobic interaction with active site of the receptor while white region near 2 and 4-position of ring A is unfavorable for this kind of interaction. This result was corroborated by the experimental activity (pIC_{50} values) of more active compounds in each group *viz.* 15, 16, 17, 18 (15>16>17>18 as hydrophobicity of $-CH_3$ group is greater than either of Cl, $-OCH_3$ and CF_3) 29, 31, 32 (29>31>32), 52, 53, 54, 55, 57, 58, 59, 60 (Table I) in which either of the positions between 3 and 5 of ring A is occupied

by $–CH_3$ group. In ring B, hydrophobic interaction at 3-position (orange contour) is likely to favor activity and at 2 and 4-positions (white contour) this interaction would render detrimental effect on activity.

In hydrogen bond donor contour (Fig. 5b), the cyan color (80% contribution) in the vicinity of 3-H and 4-H of A ring and in the lower region of the $-NH_2$ group in B ring is favored for hydrogen bond donor like substitution while purple colour near to 3-H of both ring A and ring B is not suitable for such substitution. The red contours (Fig. 5b) covering the area of 2, 4-positions of ring A and 2, 5-position of ring B should considered to be disfavor regions for hydrogen bond acceptor like substitution during design of new inhibitors of this class. Magenta colour around $–SO_2$ group indicates that oxygen atom may be involved in favorable hydrogen bond acceptor interaction with the receptor.

Docking Analysis

Analysis of docking poses of all AASBs, 1-60 (Table I) showed that all inhibitors displayed an H-bond donor interaction between the backbone carbonyl moiety of Lys101 amino group (positioned at the RT inhibitor binding site) and one hydrogen atom of NH_2 group of B ring (Fig. 6a). Docking poses of all compounds also suggested that hydrogen bonding feature of these compounds contributed, to a smaller extent, to the binding interaction with active site of the receptor which is in good agreement with the CoMFA and CoMSIA findings.

Several lipophilic interactions were also detected with different hydrophobic pockets. Docking poses showed that two phenyl rings are in good interaction with hydrophobic region while nitrile group (-CN) oriented towards the hydrophilic region of the receptor site (fig. 6b). Careful inspection of the docking pose (Fig. 6b) revealed that the region tracing out from the position of 3H and 5H of B ring may be altered for more hydrophobic interaction with the receptor site which also corroborated with the CoMFA/CoMSIA results.

RESIDUE INTERACTION

The docking poses of ligands with the receptor and their pIC_{50} values were used for interaction fingerprinting analysis. The whole chem-informatics study was carried out using Maestro 9.2 script to generate interaction finger printing matrices (fig 7).

Back bone interaction

The target protein 2RF2 contain two chains namely chain A and chain B. All most all types of receptor-ligand interactions were observed with the active site residues of chain A. Molecule 50 and 60 showed highest back bone interaction where as compound 27, 30, 38, 41, 42, 53 and compound 58 showed moderate back bone interaction with His 235, Lys 101, Lys 103, Pro 236 residues of active site of chain A. The –CN group and $–NH_2$ group attached with the B ring of compounds undergoes a strong interaction with Lys 101 of active site of enzyme. In molecule 13, 40, 48, 57, the hydrogen atom of $–NH_2$ group was shown to be responsible for this type of back bone interaction. But in case of molecule 57, the $–NH_2$ group and –CN group showed back bone interaction with Pro 236 residue. Same type of interaction was found in molecule 5, 6, 51, 56. The $–SO_2$ group was found to interact with Gly 190, Val 189, Tyr 188 and this type of interactions were also found in molecule 33, 36, 38, 39, 42, 45, 46, 49, 50, 52, 60. However, in case of molecule 60, the $–OCH_2CH_2CH_2CH_3$ (at 3-position) substituent of ring A also showed back bone interaction with Ile 180. Similar type of interaction was also found in molecule 56 where the CF_3 group of ring A interacted with Ile 180.

Side chain interaction

Compounds 2, 8, 28, 33, 42, 47, 49 showed side chain interaction with Leu 234, Trp 229, Tyr 188, Tyr 181, Val 179, Gly 190, and Val 106. Sulphur atom in molecule 9, oxygen of $-SO_2$ group in molecule 44, $–SO_2$ group as a whole in case of molecule 60 shows this type of side chain interaction.

Hydrophobic Interaction

Highest hydrophobic interaction was found in compound 1, 8, 10, 15, 21, 28, 29, 37, 47, 49, 52, 53, 55, 59, and moderate hydrophobic interaction was found in 2, 4, 5, 8, 14, 25, 26, 34, 42, 47, 48 and 49. Leu 234, Trp 229, Tyr 188, Tyr 181, Val 106, Gly 190, Val 179 residues were shown to be responsible for hydrophobic interactions with the ligands. The ring A and its substituents were shown to undergo a strong interaction with Tyr 181. The aromatic ring A was also involved in a strong hydrophobic interaction with Phe 227. These features were found in molecule 14, 35, 55, 57, 59. In case of molecule 6, 13, 20, 21, 29, 30, 50, 51, 55, 59 substituents either at 3 position of ring A were involved in hydrophobic interaction with Trp 229.

Besides this, in molecule 25, 37, 40, 48, 51, 56, the –CN and NH_2 group of ring B was found to involved in hydrophilic interaction with Leu 234.

Hydrogen Bonding Interaction

All most all the compound showed hydrogen bond donor (HBD) interaction with Lys 101 and hydrogen bond acceptor (HBA) interaction with Hie 235. Among them compound 2, 8, 14, 23, 28, 35, 37, 39, 42, 47, 48, 51, 55 and 59 showed prominent HBA interaction. In all cases the -NH_2 group of ring B was involved in HBA with Hie 235. Molecule 48 and molecule 57 also showed strong hydrogen bond donor (HBD) interaction with Val 179, an active site residue of receptor molecule.

Aromatic Residue

Interaction between active site of the receptor and aromatic residue was found highest in molecule 14, 22, 35, 37, 49, 51, 55 and 59. Active site residues Tyr 181, Tyr 188, Trp 229, Phe 227 of chain A were participated to form aromatic residue mediated interaction. The B ring of molecule 5, 8, 10, 22, 25, 38, 42, 43, 57, 49 also showed aromatic residue mediated interaction with Tyr 318. The aromatic residue mediated interaction of ring A with Tyr 181 was found in case of compound 10, 39, 40, 48 and 57. The –SO_2 group also participated in aromatic residue mediated interaction with Tyr 188. It was found in compound 4, 6, 22, 25, 39, 60.

Charge residue

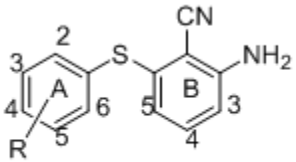
Some interaction due to charge residue was found in compound 25, 40, 43 and 57 with Tyr 181 of active site of chain A of the receptor molecule.

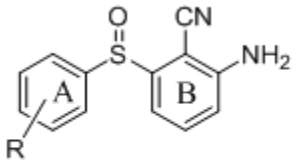
CONCLUSIONS

3D QSAR has been established on a series of 2-amino-6-arylsulfonylbenzonitriles and congeners as selective anti HIV-1 reverse transcriptase inhibitors employing the most widely used techniques *viz.* CoMFA/CoMSIA and docking study. A good correlation between docking pose orientation and CoMFA/CoMSIA contour maps on more active molecule confirmed the reliability and robustness of the derived QSAR model. The structural requirements identified in the present study for the selected compounds could be utilize to design novel, potent and selective HIV-1 RT inhibitors.

ACKNOWLEDGEMENT

One of the authors Dr. U.C. De is thankful to Schrödinger, LLC, NY for providing software support, CSIR, New Delhi, Govt. of India for providing financial assistance (EMR-II/2389, 07.06.2010) to procure a computer and OS. The authors also thankful to Dr. A. Saha, Raja Bazaar Science College, Kolkata, for providing some other software support.





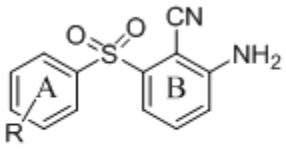


Table I: 2-Amino-6-arylsulfonylbenzonitrile derivatives

Compound No.	R	Experimental Activity	Predicted Activity		Residual Activity	
			COMFA	COMSIA	COMFA	COMSIA
1	H	1.836	1.803	1.769	0.033	0.067
2	2-OCH_3	2.367	2.341	2.224	0.026	0.143
3	3-OCH_3	2.222	2.274	2.47	-0.052	-0.248
4*	2-CH_3	1.796	2.126	2.898	-0.33	-1.102
5*	3-CH_3	2.215	2.340	2.612	-0.125	-0.397
6	4-CH_3	0.939	0.963	1.127	-0.024	-0.188
7	2-Cl	2.387	2.408	2.240	-0.021	0.147
8*	3-Cl	2.131	2.421	2.453	-0.29	-0.322
9	2-Br	1.523	1.488	1.269	0.035	0.254
10	3-Br	2.292	2.292	2.333	0	-0.041

11	3-F	2.009	2.009	1.917	0	0.092
12	3-CN	2.762	2.762	2.376	0	0.386
13*	4-CN	1.359	1.226	1.458	0.133	-0.099
14	3-CF_3	1.893	1.887	2.037	0.006	-0.144
15	3,5-$(CH_3)_2$	3.367	3.397	3.294	-0.03	0.073
16	3-Cl, 5-CH_3	2.754	2.734	2.685	0.02	0.069
17	3-OCH_3, 5-CH_3	2.699	2.676	3.058	0.023	-0.359
18*	3-OCH_3, 5-CF_3	2.292	2.226	2.583	0.066	-0.291
19	2-OCH_3	2.319	2.347	2.355	-0.028	-0.036
20	3-OCH_3	1.796	1.765	1.856	0.031	-0.06
21*	3-CH_3	1.534	2.452	2.146	-0.918	-0.612
22	4-CH_3	1.310	1.311	1.176	-0.001	0.134
23	2-Br	1.407	1.423	1.406	-0.016	0.001
24	3-Br	4.097	4.193	4.119	-0.096	-0.022
25	4-Br	1.694	1.682	1.730	0.012	-0.036
26	2-CN	2.409	2.384	2.055	0.025	0.354
27	3-CN	1.848	1.867	1.689	-0.019	0.159
28	3-CF_3	1.398	1.372	1.673	0.026	-0.275
29	3,5-$(CH_3)_2$	3.469	3.430	3.318	0.039	0.151
30*	2-Cl,5-Cl	2.007	2.127	2.741	-0.12	-0.734
31	3-Cl, 5-CH_3	3.495	3.493	3.599	0.002	-0.104
32	3-OCH_3, 5-CF_3	2.684	2.684	2.854	0	-0.17
33	H	2.699	2.522	2.280	0.177	0.419
34*	2-OCH_3	3.222	2.711	2.108	0.511	1.114
35*	3-OCH_3	3.046	3.186	2.814	-0.14	0.232
36*	2-CH_3	2.638	2.622	2.249	0.016	0.389
37	4-CH_3	2.022	1.999	1.727	0.023	0.295
38	2-Cl	2.387	2.581	2.444	-0.194	-0.057
39*	3-Cl	3.229	3.105	2.891	0.124	0.338
40*	4-Cl	2.523	2.680	1.900	-0.157	0.623
41	2-Br	2.301	2.325	2.571	-0.024	-0.27
42*	3-Br	3.268	2.626	2.542	0.642	0.726
43	4-Br	1.699	1.873	1.756	-0.174	-0.057
44	2-F	2.523	2.455	2.321	0.068	0.202
45*	3-F	2.523	2.338	2.235	0.185	0.288
46	2-CN	2.268	2.219	2.349	0.049	-0.081
47	3-CN	2.62	2.621	2.479	-0.001	0.141
48	4-CN	1.097	1.062	1.954	0.035	-0.857
49	3-CF_3	2.456	2.517	2.511	-0.061	-0.055
50	2-Cl, 5-Cl	3.523	3.441	3.616	0.082	-0.093
51*	3-Cl, 5-Cl	4.155	3.750	3.719	0.405	0.436
52	3-CH_3, 5-CH_3	5.000	4.964	4.558	0.036	0.442
53	3-Br, 5-CH_3	4.699	4.688	4.878	0.011	-0.179
54	3-Cl, 5-CH_3	4.523	4.497	4.464	0.026	0.059
55	3-OCH_3, 5-CH_3	4.301	4.311	4.489	-0.01	-0.188
56*	3-OCH_3, 5-CF_3	4.046	3.488	3.816	0.558	0.23
57*	3-OH, 5-CH_3	3.367	3.990	4.294	-0.623	-0.927
58*	3-OCH_2CH_3, 5-CH_3	4.222	4.097	4.353	0.125	-0.131
59	3-$O(CH_2)_2CH_3$, 5-CH_3	4.222	4.259	4.078	-0.037	0.144
60	3-$O(CH_2)_3CH_3$, 5-CH_3	3.222	3.218	3.297	0.004	-0.075

**compounds are of test set*

Table II: Internal validation data of CoMFA/CoMSIA

Models		Q^2 (R_{cv}^2)	R_{nv}^2	F value	SEE	R_{bs}^2	Std. Deviation
1	CoMFA	0.603	0.998	1346.132	0.050	0.999	0.001
	CoMSIA	0.506	0.989	288.023	0.119	0.995	0.061
2	CoMFA	0.521	0.998	418.093	0.051	0.999	0.001
	CoMSIA	0.364	0.986	223.410	0.140	0.995	0.003
3	CoMFA	0.562	0.998	1330.603	0.052	0.998	0.001
	CoMSIA	0.682	0.986	221.756	0.138	0.998	0.001
4	CoMFA	0.603	0.998	349.294	0.048	0.999	0.001
	CoMSIA	0.371	0.993	342.199	0.096	0.995	0.004
5	CoMFA	0.326	0.999	600.523	0.037	0.999	0.001

	CoMSIA	0.291	0.980	149.337	0.166	0.995	0.003

Table III: External validation data of CoMFA/CoMSIA

Models		*a*	*b*	R^2	*k*	R_0^2	R_{pred}^2
1	CoMFA	1.054	-0.144	0.889	1.005	0.999	0.796
	CoMSIA	0.807	0.519	0.744	0.981	0.995	0.540
2	CoMFA	0.802	0.566	0.745	0.993	0.998	0.591
	CoMSIA	0.518	1.373	0.647	0.963	0.984	0.192
3	CoMFA	0.865	0.361	0.794	0.987	0.997	0.639
	CoMSIA	0.652	0.907	0.667	0.946	0.958	0.349
4	CoMFA	0.764	0.787	0.709	1.040	0.984	0.433
	CoMSIA	0.538	1.362	0.586	0.995	0.999	0.082
5	CoMFA	1.022	-0.351	0.865	0.904	0.841	0.641
	CoMSIA	0.628	0.905	0.765	0.936	0.968	0.407

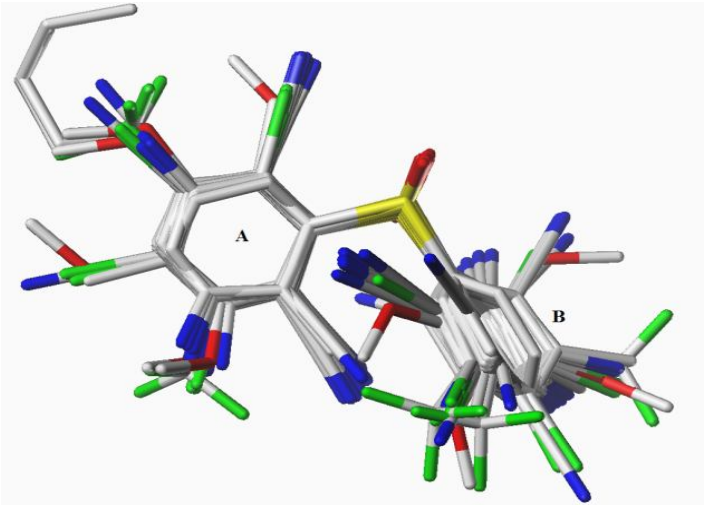


Fig. 1: Substructure based (SYBYL standard) aligned Dataset

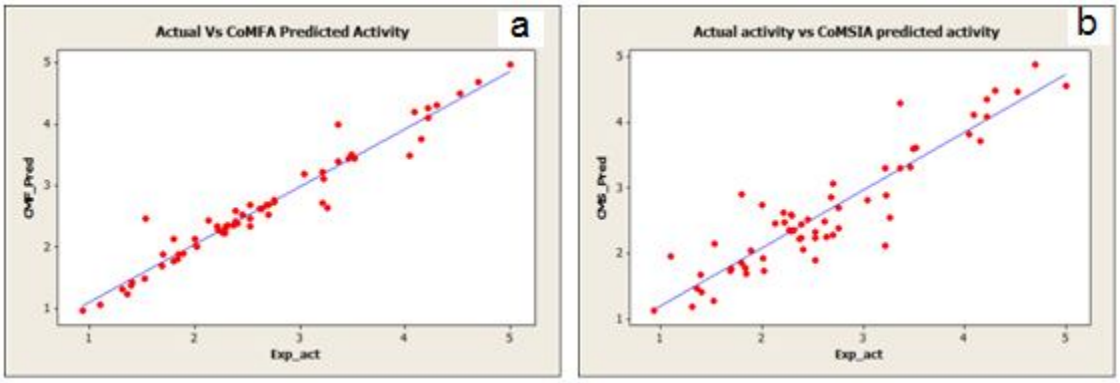


Fig. 2: Graph of experimental vs. predicted activity (a) CoMFA (b) CoMSIA

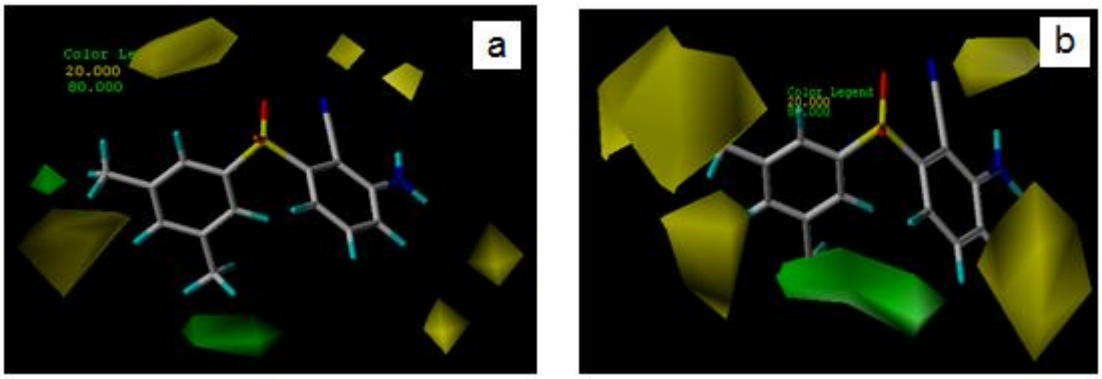


Fig. 3: Steric contour maps around most active compound 52 (a) CoMFA (b) CoMSIA Green: favorable; yellow: unfavorable

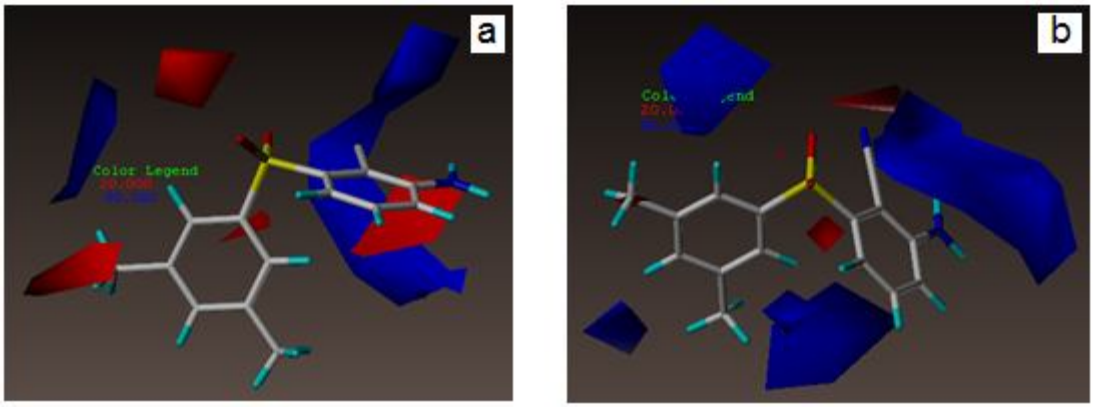


Fig. 4: Electrostatic contour maps around most active compound 52 (a) CoMFA (b) CoMSIA Blue: favorable; red: unfavorable

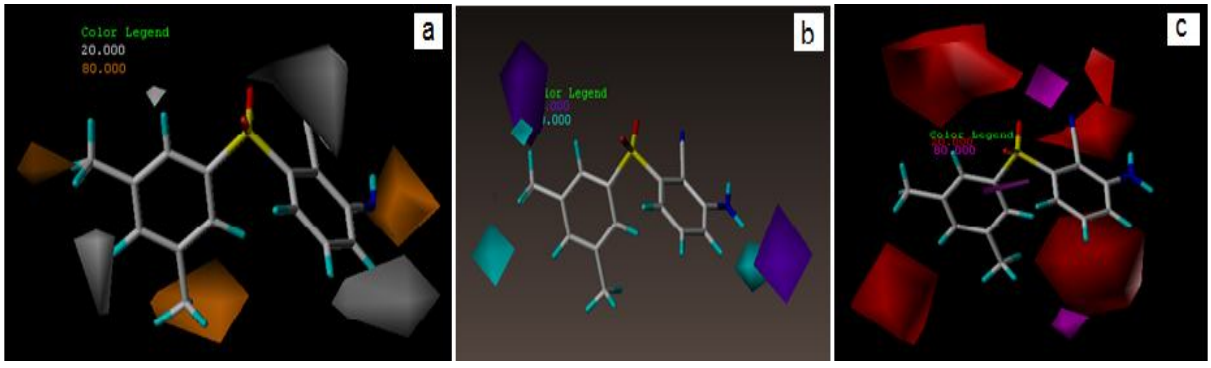


Fig. 5: CoMSIA contour maps around most active compound 52 (a) Hydrophobic features (b) H-bond donor (c) H-bond acceptor. Orange, cyan and magenta: favorable; white, purple and red: unfavorable

REFERENCES

1. WHO Library Cataloguing-in-Publication Data, Global report: UNAIDS report on the global AIDS epidemic 2010. Joint United Nations programme on HIV/AIDS (UNAIDS).
2. http://www.unaids.org/en/KnowledgeCentre/HIVData/Epidemiology/ep
3. http://www.who.int/hiv/mediacentre/news
4. Balzarini J. Current status of the non-nucleoside reverse transcriptase inhibitors of human immunodeficiency virus type 1. Curr Top Med Chem. 2004;4:921–944.
5. Pedersen OS and Pedersen EB. The flourishing syntheses of non-nucleoside reverse transcriptase inhibitors. Synthesis. 2000;4:479–495.
6. Campiani G, Ramunno A, Maga G, Nacci V, Fattorusso C, Catalanotti B, Morelli E and Novellino E. Non-nucleoside HIV-1 reverse transcriptase (RT) inhibitors: past, present, and future perspectives. Curr Pharm Des. 2002;8:615–657.
7. Clercq ED. Non-nucleoside reverse transcriptase inhibitors (NNRTIs): past, present, and future. Chem Biodiversity. 2004;1:44–64.
8. Pauwels R. New non-nucleoside reverse transcriptase inhibitors (NNRTIs) in development for the treatment of HIV infections. Curr Opin Pharmacol. 2004;4:437–446.
9. Barbaro G, Scozzafava A, Mastrolorenzo A and Supuran CT. Highly active antiretroviral therapy: current state of the art, new agents and their pharmacological interactions useful for improving therapeutic outcome. Curr Pharm Des. 2005;11:1805–1843.
10. Jonckheere H, Anne J and Clercq ED. The HIV-1 reverse transcription (RT) process as target for RT inhibitors. Med Res Rev.2000;20:129–154.
11. Pauwels R. New non-nucleoside reverse transcriptase inhibitors (NNRTIs) in development for the treatment of HIV infections. Curr Opin Pharmacol. 2004;4:437–446.
12. Cramer RD, Patterson DE and Bunce JD. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. J Am Chem Soc. 1988;110:5959–5967.
13. Klebe G and Abraham U. Comparative molecular similarity index analysis (CoMSIA) to study hydrogen-bonding properties and to score combinatorial libraries. J Comput Aided Mol Des. 1999;13:1–10.
14. Hu R, Doucet JP, Delamar M and Zhang R. QSAR models for 2-amino-6-arylsulfonylbenzonitriles and congeners HIV-1 reverse transcriptase inhibitors based on linear and nonlinear regression methods. Eur.J Med Chem. 2009;44:2158–2171.
15. SYBYL, version 1.7; Tripos Associates Inc: 1699S Hanley Rd, St. Louis, MO 631444, USA, 2005.
16. Cho SJ and Tropsha A. Structure-based alignment and comparative molecular field analysis of acetylcholinesterase inhibitors. J Med Chem.1995;38:1060–1066.
17. Cramer III RD, Clark M, Simeroth P and Patterson DE. Recent developments in comparative molecular field analysis (CoMFA). Pharm. Chem Libr.1991;16:239–242.
18. Maestro, Version 9.2, SchrÖdinger, LLC, NY2008.
19. http://www.rcsb.org, DOI. 10.2210/pdb2rf2/pdb
20. Cichero E, Cesarini S, Spallarossa A, Mosti L and Fossa P. Acylthiocarbamates as non-nucleoside HIV-1 reverse transcriptase inhibitors: docking studies and ligand-based CoMFA and CoMSIA analyses. J Mol Model. 2009;15: 871–884.
21. Zhang L, Tsai KC, Du L, Fang H, Li M and Xu W. How to Generate Reliable and Predictive CoMFA Models. Curr Med Chem. 2011;18:923-930.
22. Klebe G, Abraham U and Mietzner T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. J Med Chem.1994;37:4130–4146.
23. Richard D, Cramer III RD, Bunce JD, Patterson DE, and Frank IE. Crossvalidation, bootstrapping, and partial least squares compared with multiple regression in conventional QSAR studies. Quant Struct Act Relat. 1988;7:18–25.

24. Bush BL and Nachbar RB. Sample-distance partial least squares: PLS optimized for many variables, with application to CoMFA. J Comput Aided Mol Des. 1993;7:587–619.
25. Roy PP, Paul S, Mitra I and Roy K. On Two Novel Parameters for Validation of Predictive QSAR Model. Molecules. 2009;14:1660-1701.
26. Roy K. On some aspects of validation of predictive QSAR models. Expert Opin Drug Discov. 2007;2:1567–1577.
27. Roy PP and Roy K. On some aspects of variable selection for partial least squares regression models. QSAR Comb Sci. 2008;27:302–313.
28. Mitra I, Roy PP, Kar S, Ojha P and Roy K. On further application of r^2m as a metric for validation of QSAR models. J Chemometrics. 2010; 24:22–33.