QSAR Study for a Series of 31 Peptide-Mimetic Analogues with the Ability to Inhibit HIV-1 Protease Using Descriptors Based Analysis

T. Anthoney Swamy[*] and Gummadi Samuel Moses[+]

[*]Department of Chemistry, University of Eastern Africa, Baraton, P O Box 2500-30100, Eldoret, Kenya

[+]P.G Dept. of Chemistry, Government College (NAAC 'A' Grade), Rajahmundry, Andhra Pradesh, India

ABSTRACT: Human Immunodeficiency Virus (HIV) is the causative agent of the pandemic disease Acquired Immune Deficiency Syndrome (AIDS). HIV acts to disrupt the immune system which makes the body susceptible to opportunistic infections. Untreated, AIDS is generally fatal. Twenty five years of research by countless scientists around the world has led to the discovery and exploitation of several targets in the replication cycle of HIV. Many lives have been saved, prolonged and improved as a result of this massive effort. One particularly successful target has been the inhibition of HIV protease. In combination with the inhibition of HIV reverse transcriptase, protease inhibitors have helped to reduce viral loads and partially restore the immune system. The knowledge says that proteins (peptides) or peptide mimetic structural features are playing predominant role in the anti- HIV agents apart from the flavanoid moiety. Hence , we have considered 31 peptide-mimetic inhibitors with biological activities to proceed with quantitative structure-activity relationship using Descriptors Based Analysis

Key words: Descriptors Based Analysis, QSAR, *HIV-1 Protease Inhibitors*

[*]***Corresponding Author: Email: drtanthony2011@yahoo.com***

INTRODUCTION:

The quantitative structure-activity relationship (QSAR) research field provides medicinal chemists with the ability to predict drug activity by mathematical equations which construct a relationship between the chemical structure and the biological activity **[1,2]**. These mathematical equations are in the form of Y = bX+e that describe a set of predictor variables (X) with a predicted variable (Y) by means of a regression vector (b) **[3]**. The QSAR study for a series of 31 peptide-mimetic analogues with the ability to inhibit HIV-1 Protease has been considered **[4]**. The selected molecules in the current study belong to class of peptide-mimetic molecules.

Molecules for QSAR:

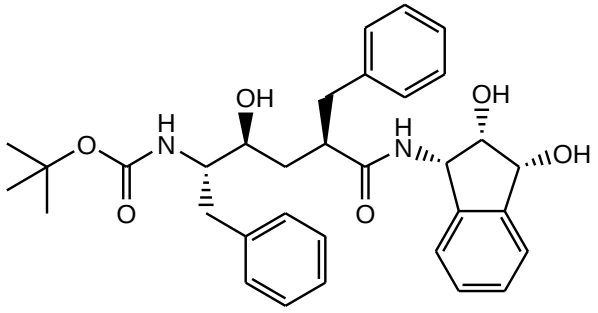
Molecules with different substituents are generated electronically with their biological activities. All the molecules were coded with a prefix "Exa". Code, Molecular Formula, Molecular Weight and the –log IC50 (can be represented as pIC50) values are given in the Table 1.1. The high is the value of pIC50, the more is the activity. Hence, molecule Exa33

was chosen as reference molecule among all these molecules because of its highest activity. This molecule will be used understand the results of the present study.

Table 1.1:

Code given	Experimental pIC50	Formula	Mol Weight
Exa1	9.6	C33H40N2O5	544.681
Exa3	8.11	C34H42N2O5	558.708
Exa4	9.72	C35H44N2O6	588.734
Exa5	9.59	C34H39F3N2O5	612.679
Exa6	9.64	C35H42N2O5	570.718
Exa7	9.22	C33H35F5N2O5	634.633
Exa8	9.54	C34H42N2O5	558.708
Exa9	9.51	C33H41N3O5	559.696
Exa10	9.57	C33H41N3O7	591.695
Exa11	5.53	C26H34N2O5	454.559
Exa12	9.8	C33H40N2O6	560.68
Exa13	7.56	C29H38N2O5	494.622
Exa14	9.14	C33H39IN2O5	670.578
Exa15	8.27	C34H40N2O6	572.691
Exa16	9.28	C32H39N3O5	545.669
Exa17	9.6	C33H40N2O5S	576.746
Exa18	9.77	C37H48N2O5	600.787
Exa19	6.94	C31H38N2O4	502.644
Exa20	8.02	C29H38N2O5	494.622
Exa21	7.47	C33H40N2O4	528.682
Exa22	6.16	C33H42N2O5	546.697
Exa23	6.79	C29H40N2O6	512.638
Exa24	7.18	C35H42N2O6	586.718
Exa25	6.67	C34H42N2O5	558.708
Exa30	4.52	C31H44N2O4	508.692
Exa31	6.89	C33H40N2O4	528.682
Exa32	6.84	C32H46N2O4	522.719
Exa33	10	C33H40N2O6	560.68
Exa34	7.41	C32H40N2O5	532.67
Exa49	5.33	C28H40N2O5	484.628
Exa50	5.86	C28H40N2O6	500.627

To generate three dimensional structures of these molecules, a reference structure is always needed from the literature. To find exact or similar structural features of this class of molecules, an attempt was made to search and explore the Cambridge Structure Database (CSD) **[5]**. But relevant results were satisfactorily not generated. Hence, it was searched in Protein Data Bank (PDB) **[6]**. Out of the several hits obtained, the coordinates of the small molecule was utilized from the PDB Code 1BDR. Exa33 was generated based on imported coordinates.



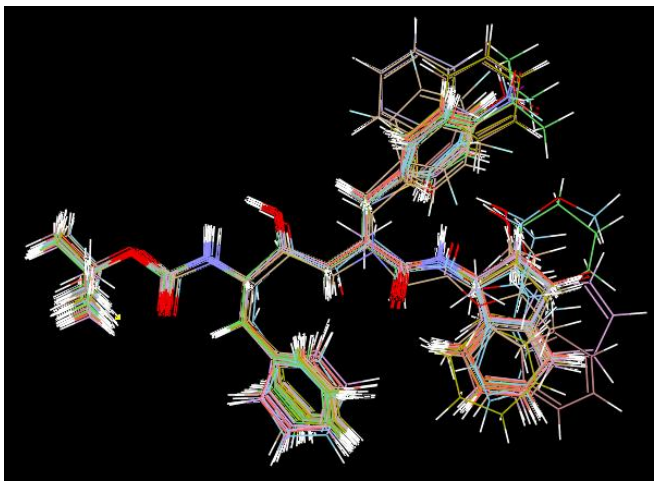
(Exa33)

The conformer with lowest energy was considered for final coordinates of Exa33. It is claimed to have global minimum. This conformer was used to generate remaining all the molecules. Molecular orbital minimization was done directly for these molecules.

Alignment:

To generate 3D QSAR, alignment is the most critical step. It has to be done to see maximum alignment occurs when all the molecules overlay on each other. Due to the criticality, alignment of molecules was done manually. The skeleton that is common to all the molecules was considered to achieve the best alignment. The scaffold thus preferred for these molecules is properly aligned and subjected to alignment. The overlay of all the molecules is given in Figure 1.1.

Figure 1.1: Manual alignment of molecules



RESULT AND DISCUSSION:

The names of the descriptors available in individual set are given in their respective headings. In the first step, separate stepwise selection-based GFA analyses were performed using different types of descriptors, and then, a GFA equation was obtained utilizing the pool of all calculated descriptors. Validation parameters thus obtained for different GFA equations are shown along with each equation. Finally, descriptors with positive effect on the biological activity were selected and perform GFA analysis.

QSAR equations were generated using Generic Function Approximation (GFA) technique in this study.

- The population size for GFA was kept at 1000 and Generations at 50,000. Number of models to be generated kept as 10, children per generation factor is 1, mutation probability 0.1 and equation length allowed is between 3 and 10.
- Friedman LOF as scoring function with smoothness parameter 0.5, linear splines and quadratic splines as additional terms are denied.
- All descriptor values are taken as independent variables on x-axis and the biological Activity is taken as dependant variable on y-axis. Initially test molecules are removed from observations and the predicted activities for test molecules are compared with experimental Biological Activities

- The molecules chosen for test set are Exa15, Exa21, Exa24, Exa25 and Exa34 throughout these studies

GFA calculates only linear combinations of the input variables. However, there are times when it is desirable to attempt to generate regression equations including square or cubic terms, or cross terms including more than one variable. To accomplish this, the original input columns have been used to generate extra columns and added in the study. The regression algorithm then acts upon this expanded set of input columns. For each set of descriptors, simple and quadratic extra terms have been tried as two independent studies to understand whether the same set of descriptors are responsible for the prediction of biological activities.

1.1.1. AlogP descriptors

Descriptors of this group are: ALogP, ALogP_AtomClass, ALogP_AtomClassName, ALogP_AtomMRScore, ALogP_AtomScore, ALogP_Count, ALogP_MR

a) Simple Terms:

$$pIC50 = -8.47494 - 0.804867 * ALogP + 0.13412 * ALogP\_MR$$

Validation parameters of the QSAR equation are:

R−squared	0.596417 ;
;Adjusted R−squared	0.56759 ;
Cross validated R−squared	0.524501
F−value	20.6893
Friedman LOF	4.47689

b) **Quadratic terms :** pIC50 = −98.0852− 14.8524 * ALogP + 1.7772 * ALogP_MR + 0.147535 * ALogP * ALogP − 0.00668428 * ALogP_MR * ALogP_MR + 0.0804638 * ALogP * ALogP_MR

Validation parameters of the QSAR equation are:

R−squared	0.66444
Adjusted R−squared	0.597328
Cross validated R−squared	0.50323
F−value	9.90046
Friedman LOF	7.11663

1.1.2. Shadow descriptors

Descriptors of this group are:
Shadow_nu,Shadow_Xlength,Shadow_XY,Shadow_XYfrac,Shadow_XZ,Shadow_XZfrac,Shadow_Ylength,Shadow_YZ,Shadow_YZfrac,Shadow_Zlength,Molecular_Volume

a) Simple terms: pIC50 = 11.4785 + 21.8535 * Shadow_nu − 7.50035 * Shadow_Xlength + 0.513785 * Shadow_XY − 121.177 * Shadow_XYfrac − 36.4785 * Shadow_XZfrac −

1.78899 * Shadow_Ylength − 0.896842 * Shadow_YZ + 85.5375 * Shadow_YZfrac + 14.6775 * Shadow_Zlength + 0.0517894 * Molecular_Volume

Validation parameters of the QSAR equation are:

R−squared	0.737005
Adjusted R−squared	0.605507
Cross validated R−squared	0.564086
F−value	5.6047
Friedman LOF	5.47498

b) Quadratic terms : pIC50= 279.225 − 9.55942 * Shadow_Xlength − 665.482 * Shadow_XYfrac − 0.0318853 * Shadow_XZ * Shadow_XZ − 442.928 * Shadow_XZfrac * Shadow_XZfrac + 0.000203754 * Shadow_XY * Molecular_Volume + 9.20537 * Shadow_XYfrac * Shadow_XZ − 3.05175 * Shadow_XYfrac * Shadow_YZ + 0.220964 * Shadow_XZ * Shadow_Ylength + 429.105 * Shadow_XZfrac * Shadow_YZfrac − 0.101626 * Shadow_Ylength * Shadow_YZ

Validation parameters of the QSAR equation are:

Friedman LOF	9.79469
R−squared	0.918039
Adjusted R−squared	0.877058
Cross validated R−squared	0.663727
F−value	22.4017

1.1.3. Molecular descriptors

Descriptors of this group are:

RadOfGyration,Apol,CoordDimension,FormalCharge,Gasteiger_Charges,IsChiral,logD,Molecular_Formula,Molecular_Mass,Molecular_Solubility,Molecular_Weight,OmegaAngle_Atoms,PhiAngle_Atoms,pKa,pKa_Atom,PsiAngle_Atoms,VSA_AlogP,VSA_AtomicAreas,VSA_MR,VSA_PartialCharge,HBA_Count,HBD_Count,NPlusO_Count,Num_AromaticBonds,Num_AromaticRings,Num_AtomClasses,Num_Atoms,Num_Bonds,Num_BridgeBonds,Num_BridgeHeadAtoms,Num_ChainAssemblies,Num_Chains,Num_ExplicitHydrogens,Num_Fragments,Num_H_Acceptors,Num_H_Acceptors_Lipinski,Num_H_Donors,Num_H_Donors_Lipinski,Num_Hydrogens,Num_Macro_Chains,Num_Macro_Residues,Num_MetalAtoms,Num_NegativeAtoms,Num_PositiveAtoms,Num_RingAssemblies,Num_RingBonds,Num_Rings,Num_Rings3,Num_Rings4,Num_Rings5,Num_Rings6,Num_Rings7,Num_Rings8,Num_Rings9Plus,Num_RotatableBonds,Num_SpiroAtoms,Num_StereoAtoms,Num_StereoBonds,Num_TerminalRotomers,Num_TrueStereoAtoms,Num_UnknownStereoAtoms,Num_UnknownStereoBonds,Num_UnknownTrueStereoAtoms,Organic_Count

a) Simple terms: pIC50 = 3.05591− 0.000325452 * Apol − 1.30703 * HBD_Count + 0.525605 * Num_AromaticBonds + 1.62567 * Num_H_Acceptors + 0.867844 * Num_RingBonds − 4.21771 * Num_Rings − 0.464445 * Num_RotatableBonds + 1.49638 * Num_StereoAtoms + 1.6739 * Num_StereoBonds

Validation parameters of the QSAR equation are:

R−squared	0.881827
Adjusted R−squared	0.822741
Cross validated R−squared	0.790237
F−value	14.9244
Friedman LOF	2.46009

b) Quadratic terms: pIC50 = 6.48628 + 11.1752 * Num_RotatableBonds + 2.13722 * Num_StereoBonds * Num _StereoBonds + 7.04055 * RadOfGyration * HBD_Count − 1.98162 * RadOfGyration * Num_RotatableBonds − 0.00027881 * Apol * HBD_Count + 0.256401 * logD * Num_RingAssemblies − 3.7477 * HBA_Count * Num_Fragments + 0.382605 * HBA_Count * Num_StereoAtoms − 36.0761 * HBD_Count * Num_Fragments + 0.802899 * Num_AromaticRings * Num_H_Acceptors

Validation parameters of the QSAR equation are:

Friedman LOF	27.0423
R−squared	0.935305
Adjusted R−squared	0.902958
Cross validated R−squared	0.874448
F−value	28.9145

1.1.4. Surface Area descriptors

Descriptors of this group are:

Molecular_FractionalPolarSASA,Molecular_FPSA,Molecular_PolarSASA,Molecular_Polar SurfaceArea,Molecular_SASA,Molecular_SAVol,Molecular_SurfaceArea

a) Simple terms: pIC50 = −34.0596 − 89.9176 * Molecular_FractionalPolarSASA + 209.592 * Molecular_FPSA + 0.111179 * Molecular_PolarSASA − 0.347966 * Molecular_PolarSurfaceArea − 0.0389914 * Molecular_SASA + 0.0429279 * Molecular_SAVol + 0.0713482 * Molecular_SurfaceArea

R−squared	0.582562
Adjusted R−squared	0.455515
Cross validated R−squared	0.465931
F−value	4.58542
Friedman LOF	6.70182

b) Quadratic terms: pIC50 = 247.714 − 3.71903 * Molecular_SASA + 3.22997 * Molecular_SAVol − 5026.9 * Molecular_FractionalPolarSASA * Molecular_FractionalPolarSASA − 0.00249837 * Molecular_SAVol * Molecular_SAVol − 0.00520636 * Molecular_SurfaceArea * Molecular_SurfaceArea + 19.7328 * Molecular_FractionalPolarSASA * Molecular_PolarSurfaceArea − 9.36478 * Molecular_FractionalPolarSurfaceArea * Molecular_PolarSurfaceArea + 3.29919 * Molecular_FractionalPolarSurfaceArea * Molecular_SAVol − 0.0058211 * Molecular_PolarSurfaceArea * Molecular_SASA + 0.00749071 * Molecular_SASA * Molecular_SurfaceArea

Validation parameters of the QSAR equation are:

Friedman LOF	14.2083
R−squared	0.819348
Adjusted R−squared	0.729022
Cross validated R−squared	0.594844
F−value	9.07099

1.1.5. Topological descriptors

Descriptors of this group are:

BIC,CHI_0,CHI_1,CHI_2,CHI_3_C,CHI_3_CH,CHI_3_P,CHI_V_0,CHI_V_1,CHI_V_2,CHI_V_3_C,CHI_V_3_CH,CHI_V_3_P,CIC,E_ADJ_equ,E_ADJ_mag,E_DIST_equ,E_DIST_mag,IAC_Mean,IAC_Total,IC,JX,JY,Kappa_1,Kappa_1_AM,Kappa_2,Kappa_2_AM,Kappa_3,Kappa_3_AM,PHI,SC_0,SC_1,SC_2,SC_3_C,SC_3_CH,SC_3_P,SIC,V_ADJ_equ,V_ADJ_mag,V_DIST_equ,V_DIST_mag,Wiener,Zagreb

a) Simple terms : pIC50 = −1891.28 − 7.30755 * CHI_2 − 2.1642 * CHI_3_P + 452.586 * CIC − 0.0422683 * E_DIST_equ + 307.884 * IC − 19.5992 * JX − 9.09679 * Kappa_2 − 1.31157 * SC_3_C + 726.97 * SIC + 0.0310257 * Wiener

Validation parameters of the QSAR equation are:

Friedman LOF	1.65875
R−squared	0.92032
Adjusted R−squared	0.880481
Cross validated R−squared	0.757863
F−value	23.1005

b) Quadratic terms: pIC50 = −244.197 − 0.0488316 * E_DIST_equ + 0.0399067 * Wiener + 1.69616 * Zagreb − 0.170634 * CHI_0 * CHI_0 + 18.7913 * CIC * CIC − 20.6398 * JY * JY + 87.3427 * BIC * JY + 1.72075 * BIC * SC_3_P − 0.209625 * CHI_3_P * SC_3_C + 0.0336156 * E_ADJ_mag * JY

Validation parameters of the QSAR equation are:

Friedman LOF	2.51706
R−squared	0.954776
Adjusted R−squared	0.932164
Cross validated R−squared	0.897392
F−value	42.2242

1.1.6. Jurs and Dipole descriptors

Descriptors of this group are:

Dipole_mag,Dipole_X,Dipole_Y,Dipole_Z,PMI_mag,PMI_X,PMI_Y,PMI_Z,Jurs_DPSA_1, Jurs_DPSA_2,Jurs_DPSA_3,Jurs_FNSA_1,Jurs_FNSA_2,Jurs_FNSA_3,Jurs_FPSA_1,Jurs_ FPSA_2,Jurs_FPSA_3,Jurs_PNSA_1,Jurs_PNSA_2,Jurs_PNSA_3,Jurs_PPSA_1,Jurs_PPSA _2,Jurs_PPSA_3,Jurs_RASA,Jurs_RNCG,Jurs_RNCS,Jurs_RPCG,Jurs_RPCS,Jurs_RPSA,J urs_SASA,Jurs_TASA,Jurs_TPSA,Jurs_WNSA_1,Jurs_WNSA_2,Jurs_WNSA_3,Jurs_WPS A_1,Jurs_WPSA_2,Jurs_WPSA_3

a) Simple terms: pIC50 = 94.0459 + 0.969675 * Dipole_X + 0.781794 * Dipole_Z − 0.177928 * PMI_mag + 0.0637395 * PMI_Y + 0.181672 * PMI_Z − 0.0285471 * Jurs_DPSA_2 − 460.536 * Jurs_FNSA_3 − 622.982 * Jurs_RNCG+ 0.0283648 * Jurs_TASA + 0.0567449 * Jurs_WPSA_1

Validation parameters of the QSAR equation are:

Friedman LOF	2.07106
R−squared	0.900515
Adjusted R−squared	0.850772
Cross validated R−squared	0.69321

b) Quadratic terms : pIC50 = −16.5658 + 0.030601 * PMI_X − 0.52526 * Jurs_PNSA_3 + 0.384683 * Dipole_mag * Dipole_mag − 1.85986e−005 * PMI_X * PMI_X − 2.2202 * Dipole_mag * Jurs_RNCS − 0.00482866 * Dipole_X * Jurs_WNSA_1 − 0.00737263 * Dipole_Y * Jurs_TPSA + 0.562743 * PMI_X * Jurs_FNSA_3 + 0.142526 * PMI_X * Jurs_RNCG + 26.0499 * Jurs_FPSA_3 * Jurs_RNCS

Validation parameters of the QSAR equation are:

Friedman LOF	2.31339
R−squared	0.980642
Adjusted R−squared	0.970962
Cross validated R−squared	0.721373
Significance−of−regression F−value	101.315

Finally selected set of descriptors from previous analyses:

a) Simple terms

Descriptors of this group are:

Dipole_X,Dipole_Z,PMI_Y,PMI_Z,Jurs_TASA,Jurs_WPSA_1,CIC,IC,SIC,Wiener,Molecular_FractionalPolarSurfaceArea,Molecular_PolarSASA,Molecular_SAVol,Molecular_SurfaceArea,Num_AromaticBonds,Num_H_Acceptors,Num_RingBonds,Num_StereoAtoms,Num_StereoBonds,Shadow_nu,Shadow_XY,Shadow_YZfrac,Shadow_Zlength,Molecular_Volume,ALogP_MR

pIC50 = 118.41 + 0.00118573 * PMI_Y − 0.0106889 * Jurs_TASA + 0.0123104 * Jurs_WPSA_1 − 15.7012 * CIC − 87.9084 * SIC + 0.358241 * Num_AromaticBonds + 1.13254 * Num_StereoAtoms + 1.42365 * Num_StereoBonds − 3.92053 * Shadow_nu − 3.82288 * Shadow_Zlength

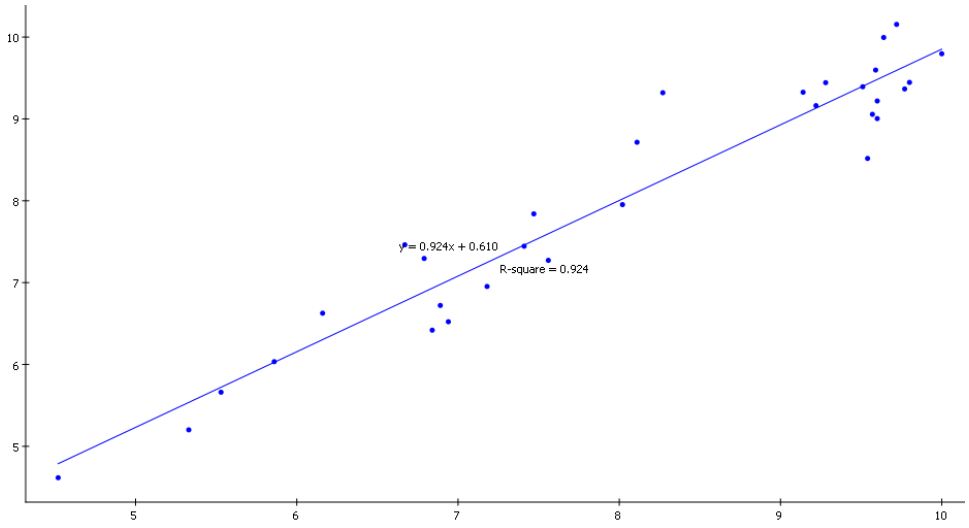
Validation parameters of the QSAR equation are:

R−squared	0.924211
Adjusted R−squared	0.886317
Cross validated R−squared	0.845504
F−value	24.3891
Friedman LOF	1.57775

Above equation could explain 92.4 % of the variance and predict 84.5 % of the variance. The positive coefficients of the PMI_Y, Jurs_WPSA_1, number of aromatic bonds, stereo atoms and bonds are conducive to the biological activity. Among these, the most important impact is due to only stereo atoms and bonds. It is obvious that the compounds having 6 and 5 stereo bonds are with high activity in the **Table 1.2**. Increasing the number of stereo bonds and hence the stereo atoms conduce for the biological activity.

The relevant plot of experimental vs predicted biological activities are shown in Figure 1.2.

Figure 1.2: 2D Plot generated for selected final descriptors



Experimental pIC50

b) Quadratic terms

Descriptors of this group are:

ALogP_MR,ALogP,Shadow_XY,Molecular_Volume,Shadow_XYfrac,Shadow_XZ,Shadow_Ylength,Shadow_XZfrac,Shadow_YZfrac,Num_RotatableBonds,Num_StereoBonds,RadOfGyration,HBD_Count,logD,Num_RingAssemblies,HBA_Count,Num_StereoAtoms,Num_AromaticRings,Num_H_Acceptors,Molecular_FractionalPolarSASA,Molecular_PolarSurfaceArea,Molecular_FractionalPolarSurfaceArea,Molecular_SAVol,Molecular_SASA,Molecular_SurfaceArea,Wiener,Zagreb,CIC,BIC,SC_3_P,E_ADJ_mag,JY,PMI_X,Dipole_mag,PMI_X,Jurs_FNSA_3,Jurs_RNCG,Jurs_FPSA_3,Jurs_RNCS

pIC50 = 49.0746 − 40.1009 * Jurs_RNCS − 0.00260708 * Shadow_XZ * Shadow_XZ
+ 0.000284825 * ALogP_MR * Molecular_Volume − 0.000447779 * ALogP * E_ADJ_mag
− 0.00919213 * Shadow_XY * Shadow_Ylength − 7.4555 * Shadow_XYfrac * BIC −
0.0328001 * Shadow_YZfrac * Molecular_SAVol
+ 6.60067 * RadOfGyration * Jurs_RNCS + 0.0025081 * Num_RingAssemblies *
Molecular_SurfaceArea + 0.375537 * Num_StereoAtoms * Num_AromaticRings

Validation parameters of the QSAR equation are:

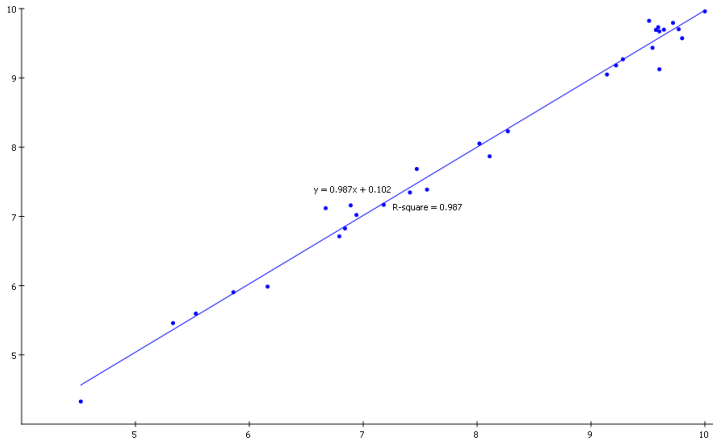
Friedman LOF	5.28471
R−squared	0.987357
Adjusted R−squared	0.981036
Cross validated R−squared	0.968734
F−value	156.192

Above equation could explain 98.7 % of the variance and predict 96.8 % of the variance. The positive coefficients of the ALogP_MR, Molecular_Volume, RadOfGyration, Jurs_RNCS, Molecular_SurfaceArea, Num_StereoAtoms, Num_AromaticRings are conducive to the biological activity. Among these, the high impact is due to RadOfGyration, Jurs_RNCS and

moderate impact is due to stereo atoms and Num_AromaticRings. Increasing the number of stereo bonds and hence the stereo atoms conduce for the biological activity.

The relevant plot of experimental vs predicted biological activities are shown in Figure 1.3.

Figure 1.3: 2D Plot generated for selected final descriptors



Experimental pIC50

Table 1.2: The values of the descriptors used for the study

Name	BiolAct	ALogP_MR	Num_StereoAtoms	Num_StereoBonds	Num_H_Acceptors	Molecular_FPSA	CIC	IC	SIC	Dipole_X	Dipole_Z	Shadow_nu	Shadow_YZfrac	Shadow_Zlength
Exa1	9.6	154.872	6	0	5	0.2	1.597	3.725	0.7	-0.46	-0.727	3.153	0.732	6.448
Exa3	8.11	159.373	6	0	5	0.192	1.686	3.672	0.685	-0.531	-0.941	3.134	0.665	6.48
Exa4	9.72	165.282	7	0	6	0.219	1.39	4.036	0.744	-1.411	-0.497	3.173	0.696	6.526
Exa5	9.59	160.846	7	0	5	0.186	1.607	3.852	0.706	-2.275	-0.646	3.092	0.74	6.47
Exa6	9.64	165.19	6	1	5	0.19	1.435	3.957	0.734	-0.538	-0.642	3.045	0.68	6.633
Exa7	9.22	155.954	6	0	5	0.185	1.618	3.873	0.705	-2.092	-0.631	2.971	0.658	6.719
Exa8	9.54	159.913	6	0	5	0.193	1.542	3.816	0.712	-0.387	-0.578	2.967	0.738	6.825
Exa9	9.51	159.572	6	0	6	0.24	1.542	3.816	0.712	-0.695	-0.749	3.152	0.77	6.407
Exa10	9.57	163.329	6	0	7	0.264	1.534	3.892	0.717	7.513	-1.519	3.087	0.726	6.479
Exa11	5.53	125.601	5	0	5	0.237	1.223	3.821	0.758	-0.473	-0.534	3.098	0.727	6.668
Exa12	9.8	156.566	6	0	6	0.232	1.542	3.816	0.712	-1.043	-1.348	3.102	0.765	6.51
Exa13	7.56	139.422	6	1	5	0.216	1.121	4.049	0.783	-0.553	-0.381	3.033	0.678	6.82
Exa14	9.14	167.28	6	0	5	0.186	1.542	3.816	0.712	-1.222	-0.677	3.092	0.709	6.526
Exa15	8.27	160.259	6	0	6	0.221	1.539	3.854	0.715	-1.544	0.345	3.225	0.69	6.256
Exa16	9.28	152.715	6	0	6	0.223	1.597	3.725	0.7	-1.538	-0.627	3.189	0.724	6.362

Exa17	9.6	162.654	6	0	6	0.236	1.47	3.888	0.726	-1.249	-0.232	3.212	0.647	6.259
Exa18	9.77	173.538	7	0	5	0.175	1.607	3.852	0.706	-0.347	-0.534	2.994	0.716	6.961
Exa19	6.94	145.93	4	0	4	0.171	1.684	3.526	0.677	-0.746	-2.226	3.155	0.663	6.399
Exa20	8.02	140.054	6	1	5	0.217	1.285	3.885	0.751	-0.399	0.15	2.897	0.736	6.829
Exa21	7.47	153.51	5	0	4	0.166	1.494	3.791	0.717	-0.609	-2.181	3.253	0.713	6.25
Exa22	6.16	156.647	5	0	5	0.194	1.679	3.643	0.685	-1.251	-1.755	3.2	0.743	6.865
Exa23	6.79	140.891	6	0	6	0.235	1.379	3.831	0.735	-1.713	-1.353	2.831	0.737	6.755
Exa24	7.18	164.366	6	0	6	0.195	1.55	3.877	0.714	1.697	-3.035	3.063	0.725	7.091
Exa25	6.67	159.51	6	0	5	0.193	1.549	3.809	0.711	0.57	-0.138	2.856	0.717	7.176
Exa30	4.52	147.062	5	0	4	0.168	1.258	3.952	0.759	-0.473	-1.117	3.134	0.738	6.902
Exa31	6.89	153.664	5	0	4	0.166	1.687	3.599	0.681	-0.641	-0.972	3.387	0.735	6.62
Exa32	6.84	151.48	6	0	4	0.163	1.395	3.853	0.734	-0.617	-2.108	2.916	0.711	6.796
Exa33	10	156.08	7	0	6	0.232	1.537	3.82	0.713	-1.182	-0.306	3.315	0.68	6.403
Exa34	7.41	151.892	5	0	5	0.2	1.61	3.675	0.695	-0.662	-1.086	3.015	0.728	6.675
Exa49	5.33	137.079	4	0	6	0.17	1.629	3.5	0.682	-0.909	0.051	2.951	0.72	6.607
Exa50	5.86	137.61	6	0	6	0.223	1.751	3.419	0.661	-1.085	-2.263	2.609	0.668	7.39

Table 1.3: Predicted and experimental pIC50 values

Code given	**Experimental pIC50**	**Predicted pIC50**
Exa1	9.6	9.183
Exa3	8.11	8.052
Exa4	9.72	9.795
Exa5	9.59	9.599
Exa6	9.64	9.534
Exa7	9.22	9.224
Exa8	9.54	9.279
Exa9	9.51	9.535
Exa10	9.57	9.424
Exa11	5.53	5.505
Exa12	9.8	9.845
Exa13	7.56	7.298
Exa14	9.14	9.248
Exa15	8.27	8.367
Exa16	9.28	9.327
Exa17	9.6	9.507
Exa18	9.77	9.808
Exa19	6.94	6.803
Exa20	8.02	8.212

Exa21	7.47	8.096
Exa22	6.16	6.236
Exa23	6.79	6.89
Exa24	7.18	7.171
Exa25	6.67	6.711
Exa30	4.52	4.519
Exa31	6.89	6.896
Exa32	6.84	6.879
Exa33	10	10.198
Exa34	7.41	7.156
Exa49	5.33	5.286
Exa50	5.86	5.945

CONCLUSION

According to simple term the positive coefficients of the PMI_Y, Jurs_WPSA_1, number of aromatic bonds, stereo atoms and bonds are conducive to the biological activity. Among these, the most important impact is due to only stereo atoms and bonds. It is obvious that the compounds having 6 and 5 stereo bonds are with high activity in the Table 1.2. Increasing the number of stereo bonds and hence the stereo atoms conduce for the biological activity.

According to quadratic term the positive coefficients of the ALogP_MR, Molecular_Volume, RadOfGyration, Jurs_RNCS, Molecular_SurfaceArea, Num_StereoAtoms, Num_AromaticRings are conducive to the biological activity. Among these, the high impact is due to RadOfGyration, Jurs_RNCS and moderate impact is due to stereo atoms and Num_AromaticRings. Increasing the number of stereo bonds and hence the stereo atoms conduce for the biological activity.

ACKNOWLEDGEMENT:

I would sincerely thank **Dr. M. Kishore Kumar**, **bio-Campus**, **Hyderabad** for his guidance and invaluable technical assistance. My grateful acknowledgments are due to the management of **bio-Campus, Hyderabad**, for providing excellent working atmosphere with modern equipment in drug discovery and design during my stay at their campus.

REFERENCES:

[1] Hansch, C.; Fujita, T. (1964) ρ σ π Analysis: A method for the correlation of biological activity and chemical structure. J. Am. Chem. Soc., 86, 1616-1626.

[2] Kubinyi, H. (1993) Methods and Principles in Medicinal Chemistry; Wolff, M. E., Ed.; VCH: Weinheim, Vol. 1, pp 497-571.

[3] Song, M.; Breneman, C. M.; Sukumar, N. (2004) Three-dimensional quantitative structure-activity relationship analyses of piperidine-based CCR5 receptor antagonists. Bioorg. Med. Chem. 12, 489-499.

[4] Holloway, M. K., J. M. Wai, et al. (1995). "A priori prediction of activity for HIV-1 protease inhibitors employing energy minimization in the active site." J Med Chem 38(2): 305-17.

[5] http://www.ccdc.cam.ac.uk/products/csd/

[6] http://www.rcsb.org/pdb/home/home.do

[7] Hansch, C.; Leo, A.; Hoekman, D. (1995) Exploring QSAR. Hydrophobic, Electronic and Steric Constants; American Chemical Society: Washington, DC.

[8] Hahn, M. (1995) Receptor Surface Models. 1. Definition and Construction. J. Med. Chem., 38, 2080-2090.

[9] Franke, R. Theoretical Drug Design Methods; Elsevier: Amsterdam, 1984.

[10] Franke, R.; Gruska, A. (1995) In Chemometric Methods in Molecular Design; van de Waterbeemd, H., Ed.; VCH: Weinheim, Vol. 2, pp 113-163.

[11] Rohrbaugh, R.H., Jurs, P.C., (1987) "Descriptions of Molecular Shape Applied in Studies of Structure/Activity and Structure/Property Relationships", Analytica Chimica Acta, 199, 99-109.

12] Rogers, D.; Hopfinger, A. (1994) J. Application of genetic function approximation to quantitative structure-activity relationships and quantitative structure-property relationships. J. Chem. Inf. Comput. Sci., 34, 854-866.