



Received on 04 August, 2014; received in revised form, 10 October, 2014; accepted, 15 December, 2014; published 01 April, 2015

THREE-DIMENSIONAL QUANTITATIVE STRUCTURE ACTIVITY RELATIONSHIP STUDIES ON DIVERSE STRUCTURAL CLASSES OF NATURAL FLAVONOIDS AS AMV AND HIV REVERSE TRANSCRIPTASE INHIBITORS USING CoMFA AND CoMSIA

Shravan Kumar Gunda*, Sri Swathi Mutya, Sharada Durgam, Vijaya Lakshmi Lokirevu and Mahmood Shaik

Bioinformatics Division, Osmania University, Hyderabad, India

Keywords:

3D-QSAR; CoMFA; CoMSIA;
FlexX Docking; Natural flavonoids;
AMV-RT; HIV-RT

Correspondence to Author:

Shravan Kumar Gunda

Research Associate,
Bioinformatics Division,
Osmania University,
Tarnaka, Hyderabad – 500007
Telangana State, India.


E-mail: gunda14@gmail.com

ABSTRACT: In this study, comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) were performed on a series of 73 experimentally reported natural flavonoid derivatives as AMV and HIV Reverse transcriptase (RT) inhibitors with pIC_{50} values ranging from 4.00 to 5.92. Based on alignment, highly predictive CoMFA and CoMSIA were obtained with cross-validated q^2 value of 0.760 and 0.696 respectively, and non-cross-validated partial least-squares (PLS) analysis with the optimum components of five showed a conventional r^2 of 0.964 and 0.977 respectively. The CoMFA and CoMSIA models have been further validated for their stability and robustness using group validation and boot-strapping techniques and for their predictive abilities. The analysis of CoMFA contour maps provided insight into the possible modification of the molecules for better activity.

INTRODUCTION: Reverse transcriptase (RT) is a multifunctional enzyme. It possesses RNA and DNA directed DNA polymerases and ribonuclease H activities¹. HIVRT has 2 enzymatic activities, a DNA polymerase that copies either RNA or DNA template, and an RNaseH that can cleave RNA if the RNA is part of a DNA/RNA duplex. These two enzymatic activities of Reverse transcriptase cooperate to convert the RNA into a double-stranded linear DNA. HIV infects vital cells in the human immune system such as helper T cells (specifically CD4+T cells), dendritic cells and macrophages². HIV RT is a heterodimer composed of p51 and p66 subunits³. Both the ends of the provirus are flanked by a repeated sequence known as the long terminal repeats (LTRs).

The genes of HIV are located in the central region of the proviral DNA and they encode at least nine proteins. Avian myeloblastosis virus is a type C oncogenic RNA virus of the sub family orthoretrovirinae. It is responsible for acute myeloblastic leukemia (AML) when injected in newly hatched chickens. It also causes myelocytomatosis, osteopetrosis, lymphoid leukosis and nephroblastoma. AMV is an acute leukemia virus which causes a myeloblastic leukemia in birds and transforms myeloid hematopoietic cells *in vitro*⁴.

Avian Myeloblastosis Virus contains a continuous sequence of approximately 1,000 nucleotides which may represent a gene responsible for acute myeloblastic leukemia in chickens⁵. Avian Myeloblastosis Virus Reverse Transcriptase is an RNA-directed DNA polymerase⁶. Both AMV-RT and HIV-1-RT are heterodimers composed of two non-identical monomer subunits. RT of HIV-1 and AMV are of tremendous medical interest as they

QUICK RESPONSE CODE 	DOI: 10.13040/IJPSR.0975-8232.6(4).1517-27
Article can be accessed online on: www.ijpsr.com	
DOI link: http://dx.doi.org/10.13040/IJPSR.0975-8232.6(4).1517-27	

the target enzymes best known for developing HIV and AMV drugs.

Flavonoids (Bio Flavonoids) are a class of plant secondary metabolites. Flavonoids (specifically Flavonoids such as the Catechins) are the most common group of polyphenolic compounds in the human diet and are found ubiquitously in plants⁷. Flavonoids are the most important plant pigments for flower coloration producing red/blue or yellow pigmentation in petals which attract pollinator animals.

The widespread distribution of flavonoids, their variety and their relatively low toxicity compared to other active plant compounds (for instance alkaloids) mean that many animals, including humans, ingest significant quantities in their diet⁸. The flavones are characterized by a planar structure because of a double bond in the central aromatic ring. An important effect of flavonoids is the scavenging of oxygen-derived free radicals.

Flavonoids also possess anti HIV, antiviral, anti inflammatory, Anti allergic and anti carcinogenic etc., properties^{9, 10}. Because of the worldwide spread of HIV since the last few decades, investigations of the antiviral activity of flavonoids have mainly focused on HIV. Many natural flavonoids and their products are capable to inhibit various stages of the replication cycle of the virus.

Methodology:

Molecular structures and optimization:

A series of seventy three natural flavonoid molecules selected for the present study were taken from an earlier report¹¹. The compound names and their biological activities are given in **Table-1(a)** and **1(b)**. *In vitro* reverse transcriptase activities were converted into the corresponding pIC₅₀ (-log IC₅₀) values. These experimental activities are used as dependent variables in comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) parameters as independent variables. The total set of RT inhibitors was randomly divided into training set (55 compounds) and test set (18 molecules).

All the 3D structures of flavonoid derivatives were sketched by using Sybyl program package version

6.7¹² on a silicon graphic workstation. Present study, each structure of 73 compounds was first minimized using Tripo's force field with a 0.005 kcal/mol Å energy gradient convergence criterion. Charges were calculated by the Tripo's force field method at the beginning and Gasteiger-Huckel^{13, 14} was considered for the further calculations.

Molecular alignment:

Molecular alignment is the most sensitive parameter in three dimensional quantitative structure activity relationship studies. The quality and predictive power of the model were directly dependent on the alignment rule. CoMFA results are not sensitive to a number of factors such as alignment, lattice shifting step size and probe atom type.

Structural alignment play important role in prediction of CoMFA models and the reliability of the contour models depend strongly on the structural alignment of the molecules. The molecular alignment was achieved by SYBYL routine align database module. Because of the highest potency, compound NOR-ICARIIN was used as a template to align the other 73 compounds from the series by common substructure alignment, using the align database command in SYBYL 6.7^{15, 16}.

The template molecule is typically most active, lead molecule and the compound contains more number of functional groups. Most common substructure used for alignment, and the superimposed structure after alignment is presented in **Fig.1**.

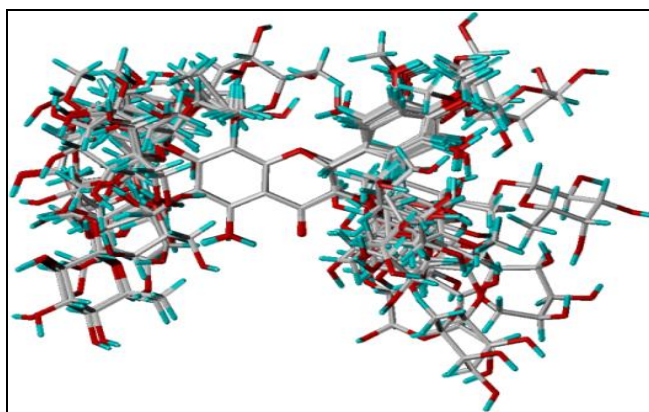


FIG.1: STRUCTURAL ALIGNMENT OF THE COMPOUNDS CONSISTING OF THE TRAINING SET AND TEST SET FOR CONSTRUCTING THE 3D QSAR/CoMFA AND CoMSIA MODELS

CoMFA studies:

After alignment, the molecules are placed one by one into a 3D cubic lattice with 2 Å grids. In CoMFA. Steric (Lennard-Jones) and electrostatic (Coulomb potentials) fields are calculated at each grid point using a sp³-hybridized carbon probe with a Vander Waals radius of 1.52 Å and +1.0 charge (default probe atom in SYBYL program). A 30 kcal/mol energy cutoff was applied, which means the steric and electrostatic energies greater than 30 kcal/mol are truncated to that value, thus, can avoid infinity of energy values inside molecule¹⁷. With standard options for scaling of variables, regression analysis was carried out using the fully cross-validated partial least squares (PLS) method (leave one out). Optimum number of components (N=5) used in the model derivation.

The column filtering was set 2.0 kcal/mol to get better signal to noise ratio by omitting those lattice points whose energy variation was below this threshold. The cross-validated coefficient q² was calculated according to the following equation:

$$q^2 = 1 - \frac{\sum (Y_{pred} - Y_{actu})^2}{\sum (Y_{actu} - Y_{mean})^2}$$

Where Y_{pred}, Y_{actu} and Y_{mean} are predicted, actual and mean values of the target property (pIC₅₀), respectively; and

$$\sum (Y_{actu} - Y_{mean})^2 = PRESS$$

PRESS is the prediction error sum of the squares, derived from the LOO method. The ONC (Optimum number of components) corresponding to the lowest PRESS value was used for deriving the final Partial least square regression models. By using the same number of components performed the Non-cross-validation to calculate conventional r².

The q² quantifies the predictive ability of the model. It was determined by LOO (Leave-One-Out) procedure of cross validation in which each molecule is successively removed from the model derivation and its pIC₅₀ value can be predicted using the model build from the remaining molecules. The CoMFA model is calculated by two statistical parameters: q² and r². q² indicates predictive capability of the model, should be >0.5.

The value r² indicates self-consistency of the model, it should be >0.9.

CoMSIA studies:

In the present CoMSIA¹⁸ investigations, five different similarity fields including steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor interactions were calculated using the sp³ carbon probe atom with a+1 charge atom and a radius of 1.0 Å. CoMSIA approach is a substitution to perform 3D-QSAR by CoMFA. Molecular similarity is compared in terms of similarity indices.

In Comparative Molecular Similarity Indices Analysis, a distance-dependent Gaussian-type physicochemical function has been adopted to avoid uniqueness at the atomic positions and dramatic changes of potential energy for those grids in the proximity of the surface. In CoMFA steric and electrostatic fields were calculated. Primarily, the intention is to division the different properties into various placements where they play a decisive role in determining the biological activity¹⁹. In general, molecular similarity indices, i.e., A_{F,K} between the compounds of interest were computed by placing a probe atom at the intersections of the lattice points using below equation:

$$A_{F,K}^q(j) = \left| \sum_{i=1}^n W_{probe,k} W_{ik} e^{-a r_{iq}^2} \right|$$

Where q is a grid point, i is a summation index over all atoms of the molecule j under computation, W_{ik} is actual value of the physicochemical property k of atom i, and W_{probe,k} is value of the probe atom.

In the present study, similarity indices were computed using a probe atom (W_{probe,k}) with charge +1, radius 1Å, hydrophobicity +1, and attenuation factor a of 0.3 for the Gaussian type distance. The statistical valuation for the CoMSIA analyses was performed in the same manner as described for CoMFA.

Molecular Docking:

Molecular docking studies were performed using flexX software²⁰ installed on Silicon Graphics Inc octane2 workstation using the package SYBYL 6.7,

to investigate the binding mode between the inhibitors and RT. FlexX is a fragment based method. FlexX handles the flexibility of the ligand by decomposing the ligand into fragments and performs the incremental construction algorithm directly inside the protein active site. This method allows conformational flexibility of the ligand while keeping the protein rigid. The base fragment is selected such that it has most potential interaction groups and the fewest alternative conformations.

All the 73 molecules which were used in QSAR studies are taken for molecular docking studies. The crystal structure of RT (PDB ID: 3HVT) in complex with NVP (Nevirapine) was used in the study. While creating RDF file, active site was defined within a radius 6.5Å of the ligand. Formal charges were assigned to all the molecules and the molecules were docked. FlexX generated 30 different conformations in the active site. All these conformations are ranked according to the FlexX score.

RESULTS AND DISCUSSION:

3D QSAR Studies:

CoMFA and CoMSIA 3D-QSAR models were derived using AMV and HIV RT inhibitors. Molecule name, their experimental pIC_{50} , predicted, residual and dock score values are given in **Table 1(a)** and **1(b)**.

CoMFA analysis:

Fifty five compounds out of the total seventy three RT inhibitors were used as training set and eighteen compounds were used as test set. The test set compounds were selected randomly so that the structural diversity and wide range of activity in the dataset were considered included. Partial least

square analysis was carried out for the training set and a cross-validated q^2 of 0.760 for five components.

The non cross-validated PLS analysis with the optimum components revealed a conventional r^2 value of 0.964, F value = 264.110 and an estimated standard error of estimate (SEE) 0.083. The steric field descriptors contribution is 59.4 % of the variance, while the electrostatic field contribution is 40.6 % of the variance. 100 runs were carried out for Bootstrap analysis for further validation of the model by statistical sampling of the original dataset to create new datasets. Correlation between CoMFA and CoMSIA experimental vs predicted were shown in **Fig 3 (a)** and **3 (b)** respectively.

CoMSIA analysis:

The CoMSIA analyses were performed using five descriptor fields: steric, electrostatic, hydrophobic, hydrogen bond donor and acceptor. The CoMSIA study disclosed a cross validated q^2 of 0.696 with optimum number of component 6, a conventional r^2 of 0.977 with a standard error of estimate 0.067 and F value 339.477.

The steric field contribution 13.2 % of the variance and, the electrostatic descriptor explains 21.4 %, the hydrophobic field explains 18.1% while the hydrogen bond donor explains 20.3 % of the variance and hydrogen bond acceptor field contribution is 27.0%. For Bootstrap 100 runs was then carried out for model validation by statistical sampling of the original dataset to create new datasets. This yielded higher r^2 bootstrap value 0.962 for CoMSIA with standard error of estimate 0.082 affirming the statistical validity of the developed models. The statistical analysis is given in **Table 2**.

TABLE 1 (a): COMPOUNDS USED IN TRAINING SET

C.No	Compound Name	pIC_{50}	CoMFA		CoMSIA		Dock score
			Predicted	Residual	Predicted	Residual	
1	Acacetin	4.310	4.360	-0.050	4.386	-0.076	-17.3
2	Fortunellin	4.710	4.669	0.041	4.663	0.047	-16.0
3	Rhoifolin	4.630	4.669	-0.039	4.632	-0.002	-15.5
4	Baicalein-7-o-glc	4.000	4.553	-0.553	4.551	-0.551	-19.4
5	Baicalein	4.020	4.224	-0.204	4.196	-0.176	-18.8
6	Baicalin	4.580	4.281	0.299	4.117	0.463	-16.2
7	Oroxin b	4.260	4.147	0.113	4.056	0.204	-14.0
8	Baicalin methyl ester	4.010	4.344	-0.334	4.332	-0.322	-20.0
9	Baicalein-6-o-glc	4.910	4.569	0.341	4.553	0.357	-15.9

10	Scutellarein	4.000	4.218	-0.218	4.087	-0.087	-18.6
11	Scutellarin	4.030	4.000	0.030	4.190	-0.160	-16.0
12	Cirsimaritin	4.290	4.273	0.017	4.238	0.052	-14.9
13	Cirsimaritin	4.060	4.265	-0.205	4.326	-0.266	-12.5
14	Isovitexin	4.610	4.326	0.284	4.411	0.199	-23.9
15	Luteolin	4.100	4.046	0.054	4.101	-0.001	-16.0
16	6-hydroxyluteolin	4.030	4.075	-0.045	3.986	0.044	-16.0
17	Pedalitin	4.030	4.117	-0.087	4.017	0.013	-16.3
18	Homoorientin	4.490	4.309	0.181	4.411	0.079	-20.0
19	Swertiajaponin	4.660	4.561	0.099	4.639	0.021	-17.4
20	Norwogonin	4.430	4.206	0.224	4.356	0.074	-17.7
21	Wogonin	4.430	4.322	0.108	4.398	0.032	-14.1
22	Isoswertisin	4.580	4.685	-0.105	4.630	-0.050	-12.6
23	Skullcapflavine II	4.740	4.562	0.178	4.490	0.250	-15.6
24	Galangin	4.370	4.195	0.175	4.306	0.064	-14.6
25	Geraldol	4.270	4.136	-4.136	4.200	-4.200	-15.9
26	6-hydroxy-kaempferol	4.010	4.151	-0.141	4.055	-0.045	-16.8
27	Kaempferol	4.130	4.323	-0.193	4.241	-0.111	-15.1
28	Kaempferide	4.180	4.182	-0.002	4.242	-0.062	-14.1
29	Quercetin	4.380	4.104	0.276	4.066	0.314	-16.1
30	Tamarexitin	4.220	4.124	0.096	4.149	0.071	-14.9
31	Rhamnetin	4.290	4.127	0.163	4.148	0.142	-16.6
32	Ombuin	4.160	4.190	-0.030	4.266	-0.106	-20.2
33	Quercimeritrin	4.160	4.252	-0.092	4.072	0.088	-13.8
34	Isoquercitrin	4.600	4.502	0.098	4.356	0.244	-13.7
35	Quercitrin	4.660	4.672	-0.012	4.955	-0.295	-14.1
36	Rutin	4.630	4.685	-0.055	4.800	-0.170	-18.9
37	Robinetin	4.050	4.151	-0.101	4.248	-0.198	-18.8
38	Myricetin	4.080	4.200	-0.120	4.137	-0.057	-14.8
39	Quercetagenin	4.000	4.102	-0.102	3.924	0.076	-15.4
40	Gossypetin	4.070	4.354	-0.284	4.375	-0.305	-15.9
41	Gossypin	4.420	4.174	0.246	4.176	0.244	-14.4
42	Trifolin	4.630	4.547	0.083	4.487	0.143	-13.7
43	Panasenoside	5.360	5.142	0.218	5.249	0.111	-17.8
44	Rhamnocitrin	4.170	4.256	-0.086	4.479	-0.309	-15.7
45	Robinin	5.690	5.300	0.390	5.377	0.313	-16.2
46	Hyperin	4.630	4.725	-0.095	4.622	0.008	-14.3
47	Datisctin	4.550	4.331	0.219	4.408	0.142	-21.7
48	Morin	4.390	4.304	0.086	4.323	0.067	-17.8
49	Anhydroicaritin	4.480	4.463	0.017	4.371	0.109	-11.4
50	Nor-icariin	5.920	5.429	0.491	5.254	0.666	-8.6
51	Dihydrobaicalein	4.280	4.343	-0.063	4.238	0.042	-16.0
52	Astibilin	5.440	5.070	0.370	4.856	0.584	-11.2
53	Daidzin	4.490	4.897	-0.407	5.032	-0.542	-19.1
54	Biochanin-a	4.750	4.740	0.010	4.692	0.058	-14.5
55	Formononetin	4.800	4.669	0.131	4.699	0.101	-15.1

TABLE 1 (b): COMPOUNDS USED IN TEST SET

C.No	Compound Name	pIC ₅₀	CoMFA		CoMSIA		Dock Score
			Predicted	Residual	Predicted	Residual	
1	Beta-naphthoflavone	5.380	4.810	0.570	4.890	0.490	-16.4
2	Pratol	4.860	4.210	0.650	4.420	0.440	-18.0
3	Oroxylin a	5.060	4.930	0.130	4.730	0.330	-16.5
4	Hispidulin	4.630	4.470	0.160	4.210	0.420	-14.6
5	Vitexin	4.600	4.140	0.460	4.290	0.310	-18.0
6	Swertisin	5.180	4.940	0.240	4.790	0.390	-15.4
7	Flavocommelin	5.220	4.720	0.500	4.810	0.410	-16.4
8	Eupafolin	4.760	4.500	0.260	4.290	0.470	-18.2
9	Cirsiliol	5.070	4.910	0.160	4.900	0.170	-18.1
10	Isoscutellarein	4.650	4.190	0.460	4.270	0.380	-16.5

11	Scutevulin	5.070	4.420	0.650	4.540	0.530	-16.6
12	Hyperoside	4.410	4.720	-0.310	4.740	-0.330	-15.2
13	Peltatoside	4.850	5.110	-0.260	5.110	-0.260	-12.5
14	Myricitrin	4.200	4.650	-0.450	4.610	-0.410	-12.8
15	Liquiritin	4.880	4.260	0.620	4.420	0.460	-11.3
16	Dihydrooroxilin a	5.420	4.990	0.430	4.790	0.630	-14.5
17	Daidzein	5.360	4.680	0.680	4.770	0.590	-15.6
18	B-anhydroicaritin	5.360	4.810	0.550	4.880	0.480	-15.5

TABLE 2: STATISTICAL ANALYSIS OF CoMFA AND CoMSIA

Field Name	CoMFA		CoMSIA	
q^2	0.760		0.696	
r^2	0.964		0.977	
Standard Error of Estimate	0.083		0.067	
F value	264.110		339.477	
Cross Validation	0.744		0.702	
Bootstrap	Mean	Std.dev	Mean	Std.dev
SEE	0.067	0.036	0.053	0.032
r^2	0.972	0.013	0.984	0.009
Field Contributions (%)				
Steric	59.4		13.2	
Electrostatic	40.6		21.4	
Hydrophobic	-		18.1	
Donor	-		20.3	
Acceptor	-		27.0	

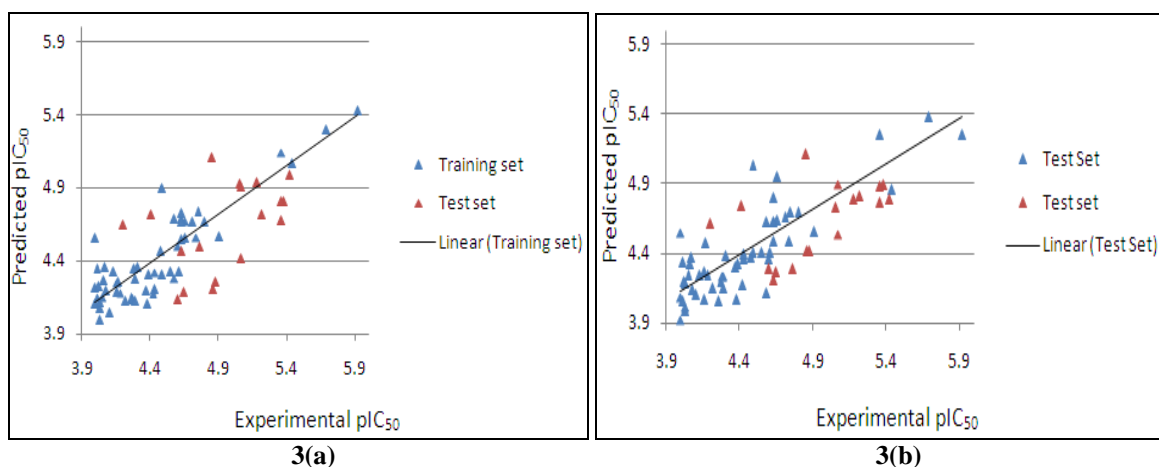


FIG 3: PREDICTED AND OBSERVED ACTIVITIES OF TRAINING AND TEST SETS USING (a) CoMFA AND (b) CoMSIA MODELS

Contour Analysis:

From the final non-cross validated analysis the steric and electrostatic fields for CoMFA were plotted as contour maps. Scalar coefficient results and standard deviation of a specific column of the data table ($SD \times \text{coeff}$) were calculated from the field energies of each lattice point and they are always plotted as the CoMFA percentage contribution equation. These contour maps illustrate regions that differ in molecular fields that are linked with variations in biological activity.

The steric and electrostatic CoMFA maps are represented in **Figure 4**; the steric, electrostatic,

hydrophobic, hydrogen bond donor and hydrogen bond acceptor CoMSIA fields are shown in **Figure 5**. These contour maps reflect definite properties through specifically colored sections depicted in the maps such as steric tolerances [green – high tolerance (80% contribution) and yellow – low tolerance (20% contribution)], electrostatic tolerance [red – decreased positive charge tolerance (20% contribution) and blue – decreased tolerance for negative charge (80% contribution)], hydrophobic regions [yellow – favored (80% contribution) and white – disfavored (20% contribution)], hydrogen bond donor [cyan – favorable (80% contribution) and purple –

unfavorable (20% contribution)] and acceptor regions [magenta – favorable (80% contribution) and red – unfavorable (20% contribution)], respectively.

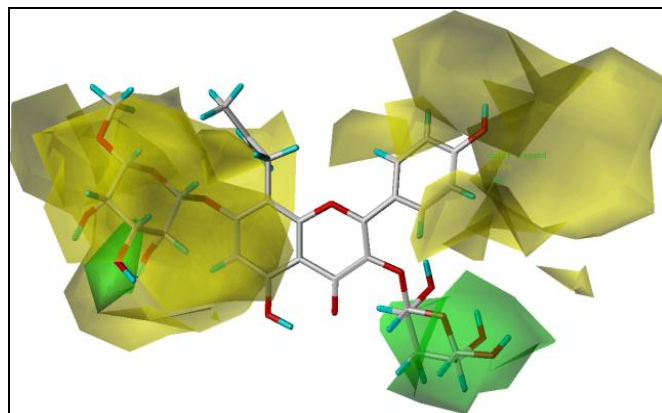
Steric and Electrostatic contour maps are shown in **Fig 4**. CoMFA models explain the variations between molecules having differences in steric and electrostatic interactions. **Fig 4(a)** shows the steric contour map for the CoMFA models with the highly active NOR-ICARIIN ($pIC_{50}=5.92$) as a reference. The steric interaction is represented by green and yellow contours. Green contours indicate the favorable, where as yellow contours indicate unfavorable for biological activity.

CoMFA Contour Analysis:

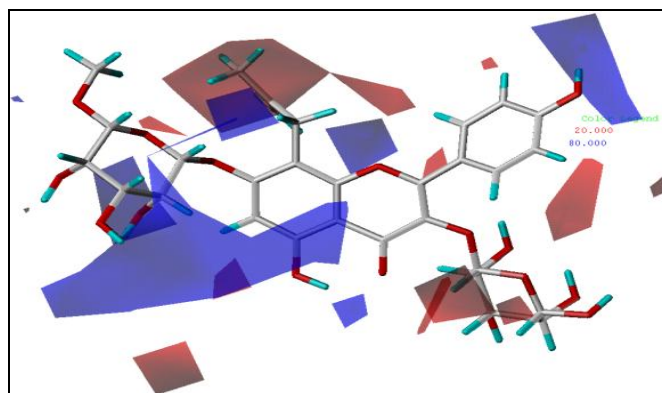
In the CoMFA contour map, two green contours and two yellow contours are present. A large green contour present at 3rd position of ORha group and a small contour was found at R7 position i.e., at OGlC group. This indicates that a bulky substituent's were preferred in this region. Isopentenyl group is present at side chain of R8 position of a most active compound NOR-ICARIIN shown in **Fig 4(a)**. A green contour appears which means that sterically favored substituent's will improve the biological activities of RT inhibitors. Green contour regions indicate favorable for steric, while yellow contours indicate unfavorable for biological activity. So addition of a bulky group at this position is favorable to the inhibitory concentration. A large yellow contour appears at ORha and OGlC groups of the R3 and R7 positions respectively. It suggests that these two positions are unfavorable.

The electrostatic contour map of CoMFA displayed in **Fig 4(b)**. Blue color contours indicates positively charged groups which increase biological activity and also electron deficient for high binding affinity. Most active compound NOR-ICARIIN shows large blue contour at side chain of R7 position of OGlC group and two small sized contours present at R8 and R4' positions. It indicating that these areas contains more positively charged substituent's at these areas. Areas where negatively charged groups enhance biological activity are contoured by red contours. Four small red color contours were present around the molecule. A medium contour

present at R8 position of Isopentenyl group which indicates more negatively charged substituent's present in these areas. A red contour present at R3 position of ORha group and also a red contour located at near R5' position, which indicates more negative charge substituent's present at these positions.



4(a) STERIC FIELDS: GREEN CONTOURS REPRESENT STERICALLY FAVORED REGIONS; YELLOW CONTOURS INDICATE STERICALLY DISFAVORED REGIONS.



4(b) ELECTROSTATIC FIELDS: RED REGIONS REPRESENT NEGATIVE POTENTIAL FAVORED; BLUE REGIONS INDICATE POSITIVE CHARGE FAVORED.

FIG. 4: COMFA CONTOUR MAPS

CoMSIA Contour Analysis:

The CoMSIA steric contours are very similar to those of the CoMFA steric contours shown in **Fig 5 (a)**. In CoMSIA, a large green contour present at R3 position of Isopentenyl group and a small green contour located nearby R7 position, which indicates that a bulky substituent's were preferred in this region. Green contours appear which means that sterically favorable substituent's will improve the biological activities of RT inhibitors. When compare to CoMFA steric green contours CoMSIA

green contours are reduced in size. Two large yellow contour present at R7 and R4' positions.

The electrostatic contour map of CoMSIA displayed in **Fig 5(b)**. Two medium sized blue contour present at R3 and R7 positions of ORha and OGlc groups indicating more positive charge constituents present at these positions. Most active compound NOR-ICARIN shows a medium sized red contour at side chain of R3 position, three small sized red contours are present at ORha and isopentenyl groups of R7 and R8 positions, which

indicating that more negative charge constituents present at these positions.

Fig 5(c) displays hydrophobic contour maps. Four white contours, a large one at the R3 position and three small contours at the R3, R7 and R3' positions, indicating that enhanced hydrophobic interactions disfavor the activity. A large yellow contour present at main chain and a small yellow contour at R7 position of OGlc group, which favors the increase the biological activity. Substituting hydrophilic groups in the molecule can radically increase the activity of RT.

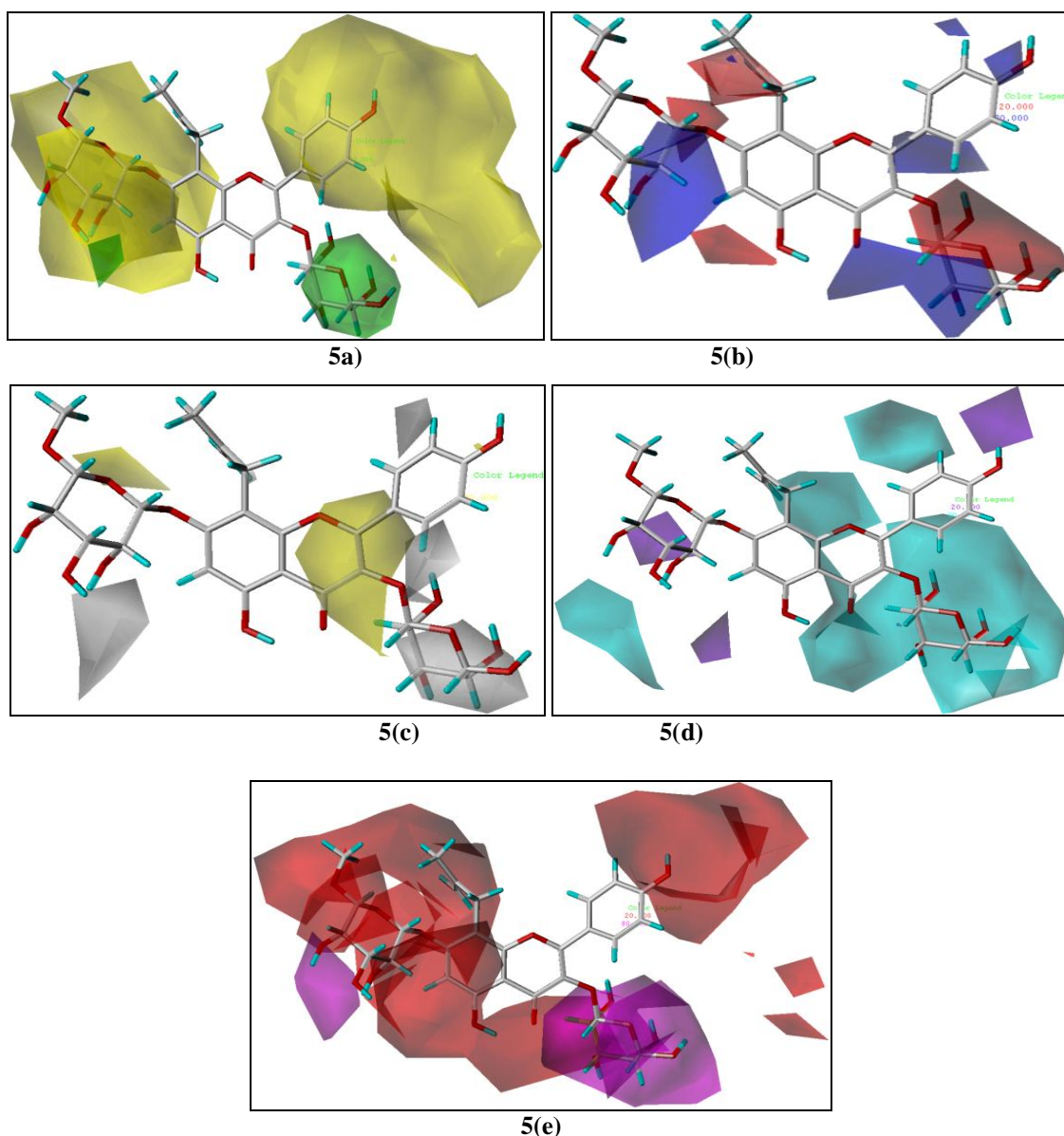


FIG. 5: CoMSIA STDEV* COEFF CONTOUR MAPS BASED ON COMPOUND 72. (5a) STERIC FIELDS: GREEN AND YELLOW CONTOURS REPRESENT STERIC FAVORABLE AND UNFAVORABLE REGIONS, RESPECTIVELY. (5b) ELECTROSTATIC FIELDS: BLUE AND RED CONTOURS REPRESENT REGIONS THAT FAVOR ELECTROPOSITIVE AND ELECTRONEGATIVE GROUPS, RESPECTIVELY. (5c) HYDROPHOBIC FIELDS: YELLOW FIELDS INDICATE REGIONS WHERE HYDROPHOBIC GROUPS COULD ENHANCE THE ACTIVITY; WHILE WHITE FIELDS INDICATE REGIONS WHERE HYDROPHOBIC GROUPS DECREASE ACTIVITY. (5d) HYDROGEN-BOND DONOR FIELDS: CYAN AND PURPLE CONTOURS REPRESENT FAVORABLE AND UNFAVORABLE HYDROGEN-BOND DONOR REGIONS, RESPECTIVELY. (5e) HYDROGEN-BOND ACCEPTOR FIELDS: MAGENTA AND RED CONTOURS REPRESENT FAVORABLE AND UNFAVORABLE HYDROGEN-BOND ACCEPTOR REGIONS, RESPECTIVELY.

The hydrogen bond donor effect could be explained by the presence of cyan and purple colored plots; purple plots explain the favorable hydrogen bond donor fields while cyan contour plot explains the unfavorable donor fields. A large contour located at R3 position of ORha group is covered by cyan color indicates that hydrogen bond donor groups exhibit negative effect on the biological activity. Three small sized purple colored contours, two at R7 position, and one at R4' position of OH group that explains hydrogen bond donor groups exhibit positive effect on the biological activity. The hydrogen bond donor contour maps are shown in **Fig 5 (d)**.

The CoMSIA hydrogen bond acceptor contours are represented in **Fig. 5 (e)** indicates areas where hydrogen-bond acceptors in the ligand promote or decrease binding affinities of the molecule. A magenta contour in the acceptor field surrounds at R3 position of ORha group and a small contour at R7 position, indicates that substitution of hydrogen bond acceptor groups at these regions increases the inhibitory activity of the molecules against RT. This correlation can be explained by analyzing the trends in biological data for some of the compounds. R7, R8 and R4' groups are covered by

red contour which significantly decrease the molecule's RT inhibitory activity.

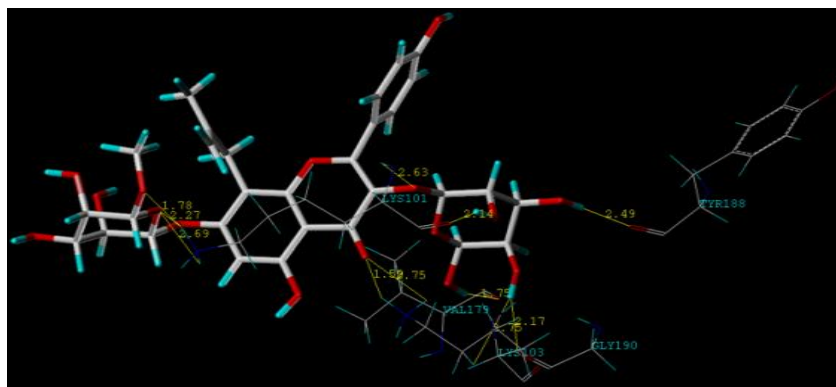
Docking results:

Docking results showed that all the molecules are forming at least two hydrogen bonds with important amino acids Lys101 and Lys103 of protein. Molecules also showed hydrogen bond interactions with other active site amino acids like Val179, Tyr188 and Gly190 of protein. Highly active molecule is participating in greater number of interactions with important amino acids Lys101, Lys103, Val179, Tyr188 and Gly190 than least active molecules.

The interactions between highly active molecule Nor-icariin and amino acids of RT binding pocket are shown in **Fig 6**. As depicted from figure, receptor and ligand are tightly bound to each other by forming a network of hydrogen bonding interactions. The binding between compound and RT amino acids are depicted as 11 hydrogen bonding interactions: Five hydrogen bonds between terminal hydroxyl group of Nor-icariin and 'O' of ionized Lys101. Three hydrogen bonds for Lys103 and a single hydrogen bond for Val179, Tyr188 and Gly190. All the distances of most and least active compounds are shown in **Table 3**.

TABLE 3: ALL THE DISTANCES OF MOST AND LEAST ACTIVE COMPOUNDS

Interacting Amnio acid	No. of Interactions	Distance
Most active compound NOR-ICARIIN		
LYS101	5	1.78, 2.14, 2.27, 2.63, 2.69
LYS103	3	1.59, 2.75, 2.75
VAL179	1	1.75
TYR188	1	2.49
GLY190	1	2.17
Least active compound-BAICALEIN-7-O-Glc		
LYS101	2	2.06, 2.07
LYS102	2	1.97, 2.21
LYS103	4	1.76, 1.79, 1.91, 2.28



(a) INTERACTIONS OF MOST ACTIVE COMPOUND NOR-ICARIIN WITH RT.

18. Anugolu RK, Gunda SK and Mahmood S, 3D QSAR CoMFA/CoMSIA and docking studies on azole dione derivatives, as anti-cancer inhibitors, *Int J Comput Biol Drug Des.* 2012; 5:111-136.
19. Khanfar MA, Youssef DT and El Sayed KA, 3D-QSAR studies of latrunculin-based actin polymerization inhibitors using CoMFA and CoMSIA approaches. *Eur J Med Chem.* 2010; 45: 3662-3668.
20. Gunda SK, Narasimha SK and Shaik M. P56(lck) kinase inhibitor studies: a 3D QSAR approach towards designing new drugs from flavonoid derivatives, *Int J Comput Biol Drug Des.* 2014;7: 278-294.

How to cite this article:

Gunda SK, Sri Mutya S, Durgam S, Lokirevu VL and Shaik M: Three-Dimensional Quantitative Structure Activity Relationship Studies on Diverse Structural Classes of Natural Flavonoids as AMV and HIV Reverse Transcriptase Inhibitors Using CoMFA AND CoMSIA. *Int J Pharm Sci Res* 2015; 6(4): 1517-27. doi: 10.13040/IJPSR.0975-8232.6(4).1517-27.

All © 2013 are reserved by International Journal of Pharmaceutical Sciences and Research. This Journal licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.

This article can be downloaded to **ANDROID OS** based mobile. Scan QR Code using Code/Bar Scanner from your mobile. (Scanners are available on Google Playstore)