Design of Reverse Transcriptase Inhibitors containing Coumarine Nucleus using Molecular Modeling Studies

Suryawanshi Asha* ; Chitre Trupti; Kakad sunil , Asegavkar Kalyani

D.D.Vispute college of pharmacy college and research center, New Panvel Navi Mumbai

Abstract-Benzothiophene derivatives have been shown to have both anti-inflammatory and anti-HIV-1 effects. Benzothiophene fused with the Coumarine derivative is likely to show better Non nucleoside reverse transcriptase inhibitors (NNRTIs) activity. A data set of 26 (N-1, 3-benzo[d] thiazol-2-yl-2-(2-oxo-2H-chromen -4-yl) acetamide derivatives, of series reported as anti -HIV agents was used for Molecular modeling studies. New Chemical Entities were generated based on QSAR studies, incorporating Benzothiophene nucleus with Coumarin. Designed compounds were synthesized and evaluated for anti-HIV activity using In-vitro RT assay.

Keywords: Anti-HIV, Docking study, 2DQSAR, 3DQSAR, Reverse Transcriptase.

I. INTRODUCTION

Acquired immune deficiency syndrome (AIDS) caused by HIV-1 virus [1], is one of the major serious health and has become a major worldwide epidemic.[2] the current status of anti-HIV therapy includes: (1) Reverse transcriptase inhibitors (RTIs): (a) Nucleoside (NRTIs) [3] e.g. Apricitabine (ATC), Elvucitabine (L-d4FC), Amdoxovir (DAPD), Racivir (RCV) (b) Nucleotide (NtRTIs) e.g. Tenofovir (c) Non nucleoside (NNRTIs) e.g. Nevirapine, Delaviridine, Efavirenz, Etravirine and Rilpivirine. (2) Protease inhibitors (PIs)[4,5] e.g. Indinavir (Crixivan), Lopinavir/Ritonavir (Kaletra), Nelfinavir (Viracept), Ritonavir (Norvir), Saquinavir (Invirase), Tipranavir (Aptivus),(3) HIV- integrase inhibitors (INIs) [6] e.g. Raltegravir, Elvitegravir MK-2048, Dolutegravir, S/GSK-1265744 and (4) Fusion inhibitors (FIs) [7] e.g. Niferivoc. Problems associated with the development of drug-resistant viral variants have led to the introduction of highly active antiretroviral therapy (HAART), typically involving concomitant treatment with a mixture of nucleoside, non-nucleoside HIV-1 RT inhibitors and HIV-1 PR (protease) inhibitor along with HIV integrase inhibitors [8].NNRTIs are structurally diverse group of compounds which bind to the viral enzyme Reverse Transcriptase (RT) at the site where it interacts with a specific allosteric non-substrate binding pocket site (Non nucleoside binding pocket-NNBP). NNRTIs non-competitively inhibit RT enzyme, block its mechanism and make it unable to produce a viral DNA [9]. Currently, more than 50 structurally diverse classes of compounds have been identified as NNRTIs. They are reported to suppress HIV-1 replication and are targeted at the NNIBP[9]. Coumarins (2-oxo-2H-chromen) have been found to exhibit a wide range of biological and controlled therapeutic activities, in view of their extensive occurrence in nature and wide range of toxicity [10]. Naturally occurring (+)-Calanolide A [9, 10] is currently undergoing anti-AIDS clinical trials. Structure modifications on (+)-Calanolide A have been carried out with the aim of improving inhibitory potency .Very recently (−)-calanolide B and (+)-calanolide C have been identified as HIV-1 specific RT inhibitors. Currently (−)-Calanolide B (=Costatolide) is under preclinical studies [11]. Suksdorfin which was isolated from fruit lomatium suksdorfii as a lead compound , on further modifications has led to discovery of 3R,4R-di-O-(S) camphanoyl-(-)-cis-khellactone (DCK) and 3R,4R-di-O-(-)-camphanoyl-2, 2 dimethdihydropyrano chromone (DCP) as a potent anti-HIV agents, is still a research focus in the current anti - HIV drug discovery field[9,12] .Benzothiophene derivatives have been shown to have both anti-inflammatory and anti-HIV-1 effects[13]. Originally, these compounds were shown to block expression of cellular adhesion molecules and to exhibit anti-inflammatory properties. Recently, benzothiophene derivatives were shown to block HIV-1 transcription in response to tumor necrosis factor α stimulation of promyelocytes. Additionally, these compounds blocked constitutive HIV-1 transcription in chronically infected cells and induced a latency state in cytokine-activated Cells. Benzothiophene fused with the coumarin derivative is likely to show better Non nucleoside reverse transcriptase inhibitors (NNRTIs) activity. Substituted and fused coumarin derivatives e.g. N-1, 3-benzo[d]thiazol-2-yl-2-(2-oxo-2H-chromen -4-yl) acetamide have been reported recently as anti –HIV- 1 inhibitors[10]. A data set of 26 (N-1, 3-benzo[d]thiazol-2-yl-2-(2-oxo-2H-chromen -4-yl) acetamide derivatives, of series reported as anti -HIV agents [10] was used for Molecular modeling studies.

II. QSAR STUDY

QSAR studies were performed using a training set of 14 molecules. .A test set of six molecules with varied chemical and distributed biological activity was used to

assess the predictive power of generated QSAR models. Six molecules did not fit into either training set and test set hence were dropped from present QSAR studies [14, 15].

The selected series of molecules along with their biological activity data are tabulated in (Table 1).

TABLE 1: Structure of Training And Test Sets Of Compounds Along With Observed And Predicted Activity.

s.no	R_1	R_2	R_3	R_4	R_5	R_6	PEC_{50}	predicted activity	Residue
[h]6a	-H	-H	-H	$-CH_3$	-H	$-NO_2$	-1.07918	-1.61537	0.53619
6b	-H	-H	-H	$-CH_3$	-H	-Cl	-1.32222	-1.47888	0.15666
6c	-H	-H	-H	$-CH_3$	-Cl	-F	-1.76343	-1.6320	-0.13143
6d	-H	-H	$-CH_3$	-H	-H	$-NO_2$	-1.6902	-1.36113	-0.32907
6e	-H	-H	$-CH_3$	-H	-H	-Cl	-1.64345	-1.60025	-0.0432
6f	-H	-H	$-CH_3$	-H	-Cl	-F	-1.23045	-1.62392	0.39347
[h]6g	-H	$-CH_3$	-H	-H	-H	$-NO_2$	-1.50515	-1.35257	-0.15243
6h	-H	$-CH_3$	-H	-H	-H	-Cl	-1.43136	-1.5917	0.16034
[h]6i	-H	$-CH_3$	-H	-H	-H	-F	-1.30103	-1.47033	0.16933
6j	-H	$-CH_3$	-H	-H	-Cl	-F	-1.07918	-1.36907	-0.28989
[h]6k	-H	-H	$-CH_3$	$-CH_3$	-H	$-NO_2$	-0.95424	-1.50506	-0.55106
6l	-H	-H	$-CH_3$	$-CH_3$	-H	-Cl	-1.77815	-1.74441	-0.03374
[h]6m	-H	-H	$-CH_3$	$-CH_3$	-H	-F	-1.38021	-1.62299	0.24299
6n	-H	-H	$-CH_3$	$-CH_3$	-Cl	-F	-1.98677	-1.76774	-0.21903
6o	-H	$-CH_3$	-H	$-CH_3$	-H	-Cl	-1.81291	-1.7745	-0.03841
6p	-H	$-CH_3$	-H	$-CH_3$	-H	-F	-1.57978	-1.65308	0.0733
[h]6q	-H	$-CH_3$	-H	$-CH_3$	-Cl	-F	-1.20412	-1.79798	0.59398
6r	-H	-H	-OH	-H	-H	-NO2	-0.90309	-0.774105	-0.12899
6s	-H	-H	-OH	-H	-H	-Cl	-0.8451	-1.01202	0.16692
6t	-Benzo-		-H	-H	-H	$-OCH_3$	-2	-2.0022	0.0022

Molecules

For QSAR model building, the compounds were divided into training and test set using sphere exclusion method [16].After this selection, a Uni-column statistics[17] of test and training sets (Table 2) showed the accurate selection of test and training sets, as the maximum of the training set was more than that of the test set and the minimum of the training set was less than or equal to that of the test set.

TABLE 2: UNI-COLUMN STATISTICS FOR TRAINING SET AND TEST SET

Set	Average	Max*	Min*	±SD	Sum
Training	-1.4588	-0.8450	-2.0000	0.3729	-20.4230
Test	- 1.3975	-0.9540	-1.7780	0.3400	-8.3850

III. 2D-QSAR

The 2D-QSAR study was carried out using algorithms like partial least squares (PLS), principle component regression (PCR) and multiple linear regressions (MLR) with User Define (UD) as the variable selection method. Various 2D descriptors (290) like physicochemical, alignment independent (AI), topological and atom – type count descriptors were calculated. Removal of invariable column resulted in around (197) descriptors. Representative descriptors were obtained with the help of correlation matrix (Table 2.1).

TABLE 2.1 CORRELATION MATRIX

	T_N_O_6	T_O_O_6	T_C_O_2	Polarizibility AHC	Chi4
T_N_O_6	1				
T_O_O_6	-0.21946	1			
T_C_O_2	0.64167	0.58934	1		
Polarizibility AHC	0.62249	0.52083	0.28023	1	
Chi4	0.42355	0.21078	0.38304	0.62131	1

Various 2D QSAR models were generated using PLS, PCR and MLR combinations. Best model was generated using PLS method. The PLS analysis was used to correlate biological activities with physicochemical properties and in turn chemical composition of the selected series of compounds.2D QSAR equations were selected by optimizing the statistical results generated along with variation of the descriptors in these models. The fitness/pattern plots were also generated for evaluating the dependence of the biological activity on various different types of the descriptors. The frequency of use of a particular descriptor in the population of equations indicated the relevant contributions of the descriptors. The

statistical result of 2D QSAR model along with the contribution of the descriptors is tabulated in (Table 3).

TABLE 3: STATISTICAL RESULTS OF 2D QSAR EQUATION GENERATED BY PLS METHOD.

Statistical parameter	PLS (model A)	PCR (model B)
r2	0.7983	0.7107
r2 SE	0.2031	0.2167
q2	0.5741	0.5697
q2SE	0.2559	0.2581
F test	25.5911	23.0254
Pred _r2	0.5897	0.5646
Pre_r2 SE	0.4594	0.4346
n	20	20
Contributing Descriptors :		
	1. Polarizibility AHC	1.T_C_C_3
	2. Chi4	2.T_N_O_7
	3.T_N_O_6	3.MMFF_29
	4.T_O_O_6	4.T_O_O_6
		5.T_C_O_2

TABLE 3.1 3D QSAR MODEL GENERATED BY SA KNN MFA METHOD FOR COUMARIN DERIVATIVES.

Sr.no.	Parameter	SA KNN MFA
1.	q2	0.7467
2.	Pred.r2	0.7536
3.	q2 SE	0.2388
4.	Pre.r2 se	0.1523
5.	N	14
6.	K nearest neighbor	S_215(-0.4923,0.2387)
7.	Contributing steric parameter	E_444(-0.5036,-0.1730)
8.	Contributing electronic parameter	E_625(-0.9325,-0.4669

A brief idea of the requirement of different physicochemical parameters and their contributions (positive or negative influence on biological activity), required for potential anti-HIV activity was obtained from the 2-D QSAR analysis (Figure 1). Where n = 20 of which training set comprised 14 and test set comprised 6 molecules. r 2 =0.7983, q 2 (cross validated r 2) = 0.5741, F - test = 25.5911, r 2 se = 0.2031, q 2se = 0.2559, pred_ r 2 = 0.5897. The best regression equation obtained is represented in Eq.1 by Model- (PLS) pEC50 = 1.72974 + 0.0681 (T_N_O_6) + 0.3660 (T_O_O_6) + 0.0771 (T_C_O_2) 0.0576 (polarizability AHC) -0.1527 (chi4) .....Eq. 1 In above QSAR models, r 2 is a correlation coefficient which gives explained variance in biological activity. Predictive ability of generated QSAR models was evaluated by q 2 employing LOO method. F value reflects ratio of variance explained by models and variance due to error in regression. High F value indicates that model is statistically significant. Low SE of estimation indicted by r 2 se and q 2 se suggested that models are statistically significant. Predictive ability of QSAR model was also confirmed by external validation of test set compounds denoted by pred_ r 2 and it was found in agreement with accepted criteria of more than 0.3. Among these three models, PLS has come out with very good results as compare to other two models. Results of PLS analysis showed very good predictive ability as indicted by r 2, q 2, F -test, and pred_ r 2 values.

Polarizibility AHC: This descriptor evaluates molecular polarizability using sum of atomic polarizabilities using the atomic hybrid component (AHC). This is Negative contributing descriptor and its contribution is 27% in the 2D QSAR equation. Chi4:-This descriptor signifies a retention index (fourth order) derived directly from gradient retention times. This is Negative contributing descriptor and its contribution is 13%.

T_C_O_2: This is Alignment Independent (AI) descriptors which means the count of number of carbon atoms (single double or triple bonded) separated from oxygen atom by 2 bond distances in a molecule. This descriptor showed positive contribution toward anti-HIV activity in selected QSAR equation (Eq. 1) and its contribution is approx 21% (Figure. 1). Positive contribution of this descriptor revealed the increase of anti-HIV activity of coumarin analogues with presence of oxygen atom at R3 position of coumarin ring.

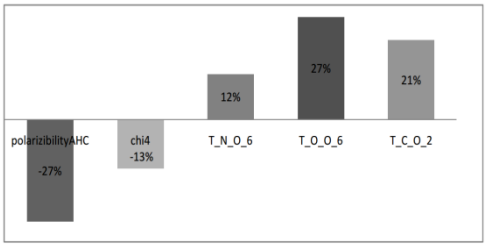
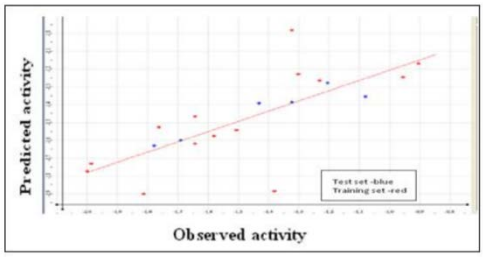


Figure1. Contributions of descriptors for biological activity developed using PLS equation.

T_O_O_6: This is also Alignment Independent (AI) descriptors. This is the count of number of Oxygen atoms (single double or triple bonded) separated from any other Oxygen atom (single double or triple bonded) by 6 bonds in a molecule. This descriptor is positively contributing and it contributes 27%. Positive contribution of this descriptor shows increase the anti-HIV activity of coumarin analogues with the presence of oxygen atom.T_N_O_6: This is Alignment Independent (AI) descriptors. This is the count of number of Oxygen atoms (single double or triple bonded) separated from Nitrogen atom by 6 bond distance in a molecule. This is positive contributing descriptor and it contributed 12%. Positive contribution of this descriptor reveals that presence of Nitrogen atom at R5 position of Coumarin ring may increase the anti-HIV activity.

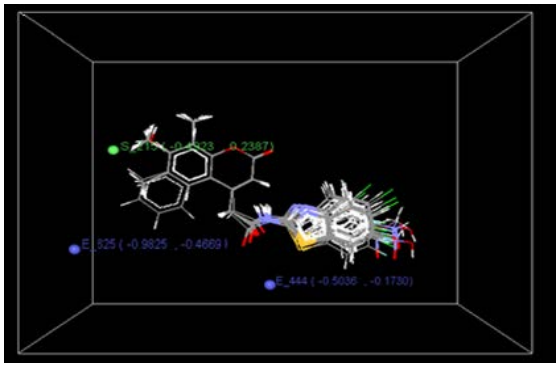
The 3D QSAR models were generated using k-Nearest Neighbor Molecular Field Analysis (SA kNN MFA) using Simulated Annealing (SA) variable selection method [20, 21]. 3D QSAR models were selected based on value of statistical parameters and the best SA kNN-MFA. 3D QSAR model having a q2 (r2cv) of 0.7467 and pred. r2 of

0.7536 (Table 3.1). The plots of observed versus predicted activity of both training and test sets molecules helped in cross-validation of kNN-QSAR model are depicted in (Figure 2).



Model –Simulated Annealing kNN MFA Model.pEC50=- E_625 (-0.9825,-0.4669) - E_444 (-0.5036,-0.1730) + S_215 (-0.4923, 0.2387)…..Eq.2 In 3D QSAR studies, 3D data points generated around the Benzothiazole-2-yl (2-oxo-2H-chromen-4-yl) acetamide pharmacophore were used to optimize the electrostatic and steric requirements of the coumarin nucleus for anti-HIV activity (Figure 4).

Figure 4. Stereoview of the molecular rectangular field grid around the superposed molecular units of benzo [d]thiazol-2-yl-2-(2-oxo-2H-chromen-4-yl) acetamides series of compounds using SA kNN MFA method.

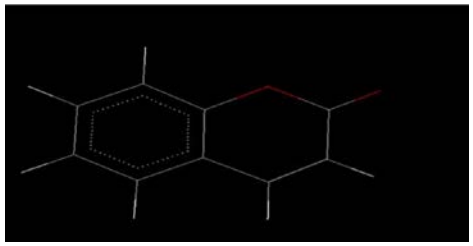


The range of property values in the generated data points helped for the design of NCEs. These ranges were based on the variation of the field values at the chosen points using the most active molecule and its nearest neighbor set. The points generated in SA KNN MFA 3D QSAR model are S_215, E_444 and E_625 respectively. These points suggested the significance and requirement of steric and electrostatic (electronegative as well as electropositive) properties as mentioned in the ranges in parenthesis for structure–activity relationship and maximum biological activity of coumarin analogues. Negative values in steric field descriptors indicated the requirement of less bulky group, for enhancing the biological activity of coumarin analogues.

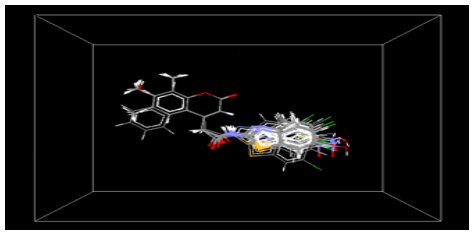
Therefore less steric substituent's as well as bulky steric substituent's were preferred at the position of generated data points S_215 (-0.4923,0.2387),around coumarin pharmacophore. Similarly the negative values of electrostatic descriptors suggested the requirement of electronegative groups at the position of generated data point E_444 (-0.5036,-0.1730) and E_625 (-0.9825,-0.4669) respectively around pharmacophore for maximum activity.

The findings of 2D and 3D QSAR studies provided the overall substitution pattern (electrostatic and steric pattern) required around the Benzothiazol-2-yl-2-(2-oxo-2H-chromen-4-yl)acetamides (Figure 3b).

Figure 3.(a) Common template used for alignment of coumarin derivatives and (b) overlain of benzo[d]thiazol-2-yl-2-(2-oxo-2H-chromen-4-yl)acetamides series of compounds using V-Life MDS 3.5.



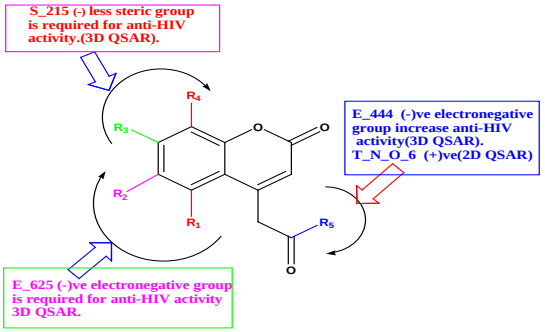
(a)



(b)

Descriptors generated in 2D QSAR equation signified the importance of coumarin nucleus for anti-HIV activity of compounds. Substitution pattern around (2-oxo-2H-chromen-4-yl) acetaldehyde pharmacophore showed in (Figure 5) was used for the design of NCEs using CombiLib tool of vLife MDS software.

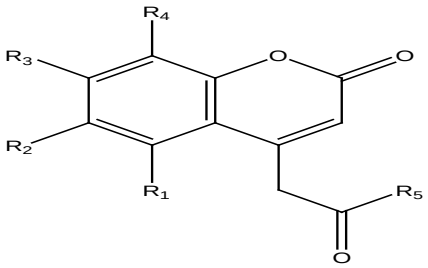
Figure 5. The outcomes of QSAR studies show requirements around the (2-oxo-2H- chromen-4-yl) acetaldehyde pharmacophore for anti-HIV activity



Designed compounds were subjected to Lipinski's screen (the rule of 5)[22] to ensure drug like pharmacokinetic profile of the designed compounds in order to improve their oral bioavailability. The parameters used as Lipinski's filters are: A= Number of hydrogen bond acceptor (<10), D= Number of hydrogen bond donor (<5),

R=Number of rotatable bond (<10), X= Xlog P (X) (<5), W= Molecular weight (<500 g/ mol), S= Polar surface area (<140 Å). More than one hundreds of molecules were generated using CombiLib tool which follows the Lipinski's rule, but we have selected most active molecules on the basis of predicted activity. The most potent compounds have negative values of pEC50, similarly the predicted activity values for NCE's were found to be lie towards negative side. The results are shown in (Table 4).

TABLE 4: Structure of Designed Nces Along With Predicted Activity Obtained By Pls Equation Generated By 2d Qsar



Comp	R1	R2	R3	R4	R5	Screen Result	Screen score	Predicted activity
A1	-H	-H	-H	-H	Ethyl 2- amino 4,5,6,7 tetrahydrobenzothiophene- 3- carboxylate	ADRWS	5	-0.68683
A2	-H	-H	-H	-CH3	Ethyl 2- amino 4,5,6,7 tetrahydrobenzothiophene- 3- carboxylate	ADRWS	5	-0.84531
A3	-H	-H	-H	-OH	Ethyl 2- amino 4,5,6,7 tetrahydrobenzothiophene- 3- carboxylate	ADRXWS	6	- 0.81137
A4	-H	-H	-CH3	-H	Ethyl 2- amino 4,5,6,7tetrahydrobenzothiophene- 3- carboxylate	ADRXWS	6	-0.83127
A5	-H	-H	-CH3	-CH3	Ethyl 2- amino 4,5,6,7tetrahydrobenzothiophene- 3- carboxylate	ADRWS	5	-0.97127
A6	-H	-H	-CH3	-OH	Ethyl 2- amino 4,5,6,7tetrahydrobenzothiophene- 3- carboxylate	ADRWS	5	-0.75653
A7	-H	-H	-OH	-H	Ethyl 2- amino 4,5,6,7tetrahydrobenzothiophene- 3- carboxylate	ADRXWS	6	- 0.74745
A8	-H	-OH	-OH	-H	Ethyl 2- amino 4,5,6,7 tetrahydrobenzothiophene- 3- carboxylate	ADRXWS	6	- 0.94745
A9	-H	-CH3	-CH3	-H	5- Chloro- 6-fluorobenzo -thiozol-2- amine	ADRXWS	6	- 0.91943
A10	-Cl	-CH3	-CH3	-H	Ethyl 2- amino 4,5,6,7tetrahydrobenzothiophene- 3- carboxylate	ADRXWS	6	- 0.96949
A11	-OH	-CH3	-OH	-H	5- Chloro- 6-fluorobenzo -thiozol-2- amine	ADRXWS	6	- 0.96949
A12	-H	-OH	-OH	-CH3	Ethyl 2- amino 4,5,6,7tetrahydrobenzothiophene- 3- carboxylate	ADRXWS	6	- 0.34745
A13	-H	CH3	-CH3	-H	Ethyl 2- amino 4,5,6,7tetrahydrobenzothiophene- 3- carboxylate	ADRXWS	6	- 0.91943
A14	**- H**	-OH	-H	-Imidazole	Ethyl 2- amino 4,5,6,7tetrahydrobenzothiophene- 3- carboxylate	ADWS	4	-0.58605
A15	-H	-H	- OH	-	5- Chloro- 6-fluorobenzo	ADRWS	5	- 0.89149

				Imidazole	-thiozol-2- amine			
A16	-OH	-OH	- OH	-H	Ethyl 2- amino 4,5,6,7tetrahydrobenzothiophene 3-carboxylate	ADRXWS	6	- 1.14745
A17	-OH	-H	- OH	-pyrrole	Ethyl 2- amino 4,5,6,7tetrahydrobenzothiophene-3- carboxylate	ADRXWS	6	- 1.14745
A18	-Cl	-OH	–OH	-pyrrole	Ethyl 2- amino 4,5,6,7tetrahydrobenzothiophene-3- carboxylate	ADRXWS	6	- 1.20745
A19	**-F**	-H	-OH	-H	Ethyl 2- amino 4,5,6,7tetrahydrobenzothiophene-3- carboxylate	ADRWS	5	- 0.28745
A20	**-OH**	-H	$-CH_3$	-H	Ethyl 2- amino 4,5,6,7 tetrahydrobenzothiophene-3- carboxylate	ADRXWS	6	-0.78605
A21	-H	**-OH**	-H	-OH	Ethyl 2- amino 4,5,6,7 tetrahydrobenzothiophene- 3-carboxylate	ADRXWS	6	-0.76872
A22	-OH	-H	- H	-OH	5-Chloro-6-chlorobenzo -thiozol-2- amine	ADRXWS	6	- 0.66949
A23	-OH	-H	Pyrrole	-OH	Ethyl 2- amino 4,5,6,7 tetrahydrobenzothiophene- 3-carboxylate	ADRWS	5	- 0.56449
A24	-H	-OH	-H	-H	Ethyl 2- amino 4,5,6,7tetrahydrobenzothiophene-3- carboxylate	ADRXWS	6	-0.76872
A25	- H	-F	-OH	- H	5- Chloro- 6-fluorobenzo -thiozol-2- amine	ADRXWS	6	- 0.99149
A26	-OH	-H	Imidazole	-OH	Ethyl 2- amino 4,5,6,7 tetrahydrobenzothiophene-3- carboxylate	ADRXWS	6	-1.26872

Compounds qualifying all required parameters set for Lipinski's Screen/filter is indicated by ADRXWS strings.Substitution at R1, R2 and R5: 3D QSAR studies helped to find out the electrostatic requirements at these positions. The electrostatic data point generated was E_444, E_625 It was found that the electronegative groups like: -OH,-NH2, -Cl,-SH etc were essential for potent anti-HIV activity and accordingly the substitutions were carried out for designing of NCEs. Electronegative - NH group along with heterocyclic ring (benzothiozol) at R5 position in the reported series[10] was found to have less predicted activity as compared to our substitution of 2-Aminobenzothiophene by Lipinski's rule . Hence 2 - Amino benzothiophene was substituted at the R5 position for the anti-HIV- activity According to literature survey 13.Substituent - H and -CH3 group at R1 and R2 position have less predicted activity as compared to- OH group so it is replaced by - OH group for enhancing the anti-HIV activity[10]. Substitution at R3 and R4:3D QSAR studies also showed requirement of less steric groups at R3 and R4 position. The steric data point generated was S_215 (-0.4923, 0.2387) indicates less sterically group to bulky steric group were required at R3 and R4 position for anti-HIV activity. Steric grid point generated on the R4 and R5 which is both less steric to bulky steric so at R4 position - H is replaced by -OH group which have better predicted activity as compared to -H . At R3 position there is -OH group (electronegative) and- CH3 group (electropositive) present in the reported series which have nearly the same predicted activity as compared to NCEs – OH, and - CH3 group was not replaced by other substituent.

IV. DOCKING STUDY

The molecular docking tool, GLIDE (Schrodinger Inc., USA) was used for ligand docking studies in to the enzyme reverse transcriptase binding pocket. The crystal structures of reverse transcriptase were obtained from protein data bank. (PDB Cod: 1Ikw8). All structures were

prepared for docking using 'protein preparation wizard' in Maestro wizard 8.5[13]. Water molecules in the crystal structures were deleted. The protein preparation was carried out in two steps, preparation and refinement. After ensuring chemical correctness, the hydrogen's were added where hydrogen atoms were missing. Side chains that are not close to the binding cavity and do not participate in salt bridges were neutralized. In the refinement component, a restrained impact minimization of the co-crystallized complex was carried out. This helps in reorientation of side chain hydroxyl group. It uses the OPLS.AA force field [24] for this purpose. Grids were defined by centering them on the ligand in the crystal structure using the default box size. The ligand were built using maestro build panel and prepared by Ligprep 2.2 module which produces the low energy conformer of ligands using MMFF94 force field[25]. The lower energy conformations of the ligands were selected and were docked into the grid generated from protein structures using standard precision (SP) docking mode[26]. In this docking method the ligands are flexible and receptor is rigid except the protein active site which has slight flexibility. The final evaluation is done with glide score (docking score) and single best pose is generated as the output for particular ligand. G-score = vdw + Coul + Lipo + H bond + Metal + Rot B +Si …….Eq.3

Where, vdW= Van der Waal energy; Coul= Coulomb energy; Lipo= lipophilic contact term; H Bond = hydrogen-bonding term; Metal= metal-binding term; RotB = penalty for freezing rotatable bonds; Site = polar interactions at the active site. The G-score indicates the binding affinity of the designed compound to the receptor/enzyme. The G-score of the standard compound Efavirenz (Efz) was found to be -9.62692. The G-scores of the designed NCEs A20, A4, A7, A3, A21 were found to -9.95871,-9.94714,-9.79032,-9.99503,-9.52184
respectively. The close analysis of these results suggested that the designed NCEs have significantly good G-score with the standard compound. H-bond is one of the most widely used parameter for the evaluation of the docking results, as it is an influential parameter in the activity of the any compound. The number of H-bond interactions in the standard compounds was compared with that of designed NCEs. The number of H-bond in the standard compound Efavirenz was found to be 1 (Table 5).

Table 5: Results of Molecular Docking studies.

Molecule No.	G-score	No. of Hydrogen bond interaction	E model	Good VDW	Bad VDW	Ugly VDW
A20	-9.95871	1	-48.3428	372	6	2
A4	-9.94714	2	-41.9358	331	12	3
A7	-9.79032	2	-62.7803	416	13	0
A3	-9.99503	1	-57.3422	406	10	0
A21	-9.52184	3	-49.4344	340	16	0
Efz(std)	-9.62692	1	-56.5465	238	9	0

The no. of H-bond interactions for the designed compounds A20, A4, A7, A3 and A21 were found to be 1, 2,2,1and 3 respectively, indicating significantly better binding interactions with the enzyme. The following contacts are represented in the form of Van der Waals (vdw) Interactions: Good vdw interactions, Bad vdw interactions, Ugly vdw interactions. It was found that designed compound A20, A4, A7, A3 and A21 has more number of good vdw interactions, less number of bad vdw and ugly contacts when compared with the standard Efavirenz (Table 5). In conclusion G score and E model in addition to number of H-bond interactions, number of good, bad and ugly vdw contact decide the possible binding affinity and in turn potency of the designed NCEs. A4, A7, A21 showed the H-bond interaction with Lys 101 residue (Figure 7, 8, 9). The H atom of amine of Lys101 showed the H-bond interaction with C=O, H-bond distance of Efz is 1.990 Å (Figure 6) and compounds A4, A7, A21 is 1.933 Å, 2.387 Å, 1.992 Å.

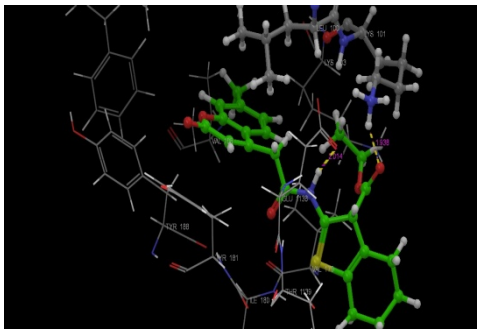


Figure 6. Binding mode of standard Efavirenz into binding pocket

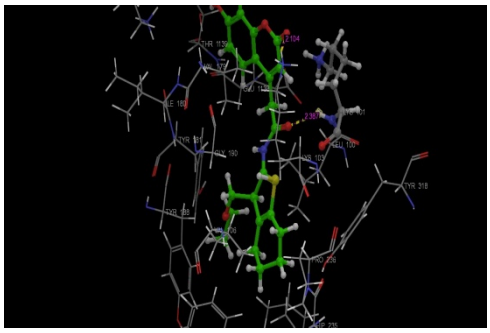


Figure7: Binding mode of A4 into binding pocket of RTenzyme

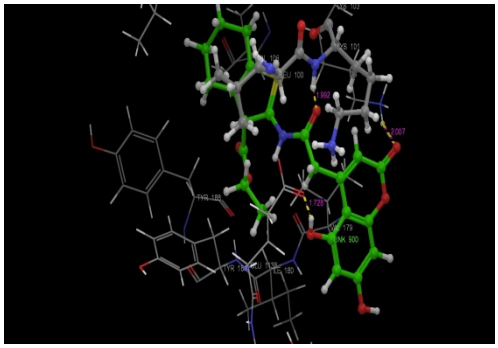


Figure 8. Binding mode of A7, into binding pocket of RT enzyme

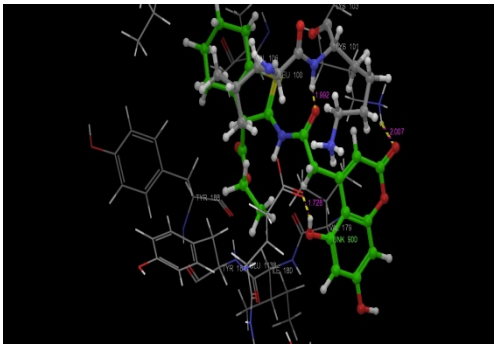


Figure 9. Binding mode of standard A21 into binding pocket of RT enzyme.
of RT enzyme

The ADME profile of the designed NCEs was studied using the Qik-Prop [21] tool of Schrodinger software. In addition to predict molecular properties, QikProp provides the ranges for comparing the molecules properties with those of 95% of known drugs. QikProp also flags 30 types of reactive functional groups that may cause false positives in High Through put Screen assays (HTS). The range of values that cause a molecule to be flagged can be similar or dissimilar to other known drugs. Forty four physical descriptors and pharmaceutically relevant properties of coumarin analogues were analyzed using QikProp, amongst which significantly contributing descriptors are reported here which are required for predicting the drug like properties of molecules (Table 6).

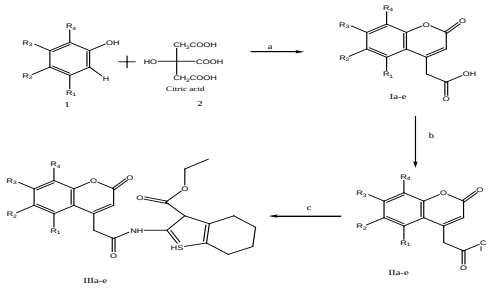
TABLE 6: Results of ADME Properties Prediction Studies

Molecule	Mol.MW	logPo/w	logBB	% Human Oral	Rule	CNS
A20	429.487	4.483	-0.210	100	0	-2
A4	427.514	4.601	-1.103	98.23	0	1
A7	415.503	3.644	-1.549	100	0	1
A3	443.514	3.835	-1.844	100	0	0
A21	445.486	2.944	-2.205	100	0	-2
Efz	315.679	3.497	0.033	100	0	1

These properties are as follows: Molecular weight (mol _MW) (150–650), Octanol/water partition coefficient (Log Po/w) (-2 to 6.5), Aqueous solubility (QPlog S) (-6.5 to 0.5), CNS (-2 to 2), Percent human oral absorption (≥80% is high, ≤25% is poor). The first three properties are based on Lipinski rule of five ,molecular weight (mol_ MW) less than 650, partition coefficient between Octanol and water (log Po/w) between -2 and 6.5 and solubility (QPlog S) greater than -7 . CNS parameter indicated about the ability of the drug to pass through the central nervous system which is mandatory for inhibition of CNS disorder caused due to HIV infection. All designed compounds showed ADME properties in acceptable range

V. GENERAL METHOD

Synthesis of coumarin- 4- acetic acid(step I) [10,18] (scheme1) :A mixture of citric acid (1 mol) and conc. sulphuric acid (32 ml) was stirred for half an hour, then the temperature was slowly raised up to 40-500C during an interval of 10 to 15 min and as soon as the evolution of gas slackened, the flask was removed from the heating mantle, allowed to stand for 15 min till the reaction mixture became clear (yellow colored) and free from carbon monoxide bubbles; this was then cooled to 0-100C. To this solution, substituted phenol (1 mol) was added at chilled condition. After the addition of substituted phenols, the reaction mixture was stirred at room temperature for 48 hrs. The reaction mixture was then poured onto crushed ice; the separated solid was filtered and dissolved in saturated sodium bicarbonate solution which on acidification gave the compounds. (Yield 50 to 75%).Melting point (uncorrected) 1300c.



Scheme1.Synthesis of various Ethyl 2-(2-(2-oxo-2H-2 chromen-4-yl)acetamido)-4,5,6,7 tetrahydro -3H benzo[d]thiophene-3- caboxylate .Reagent and condition :a)Conc. H2SO4, stirring for 48hrs at room temperature;b) POCl3 ; c) Ethanol solvent , Ethyl 2- amino 4,5,6,7-tetrahydrobenzothiophene- 3- carboxylate ,microwave 15min at 600C.

Synthesis of coumarin- 4-acetyl chloride (step II)[19] : A mixture of step I (2g,0.185mol) and 4ml POCl3 were refluxed for 1hr at 50 to 60oc and slowly poured the reaction mixture into crushed ice (caution) .Filtered and finally recrystallized from ethanol water.Yield 80 to 90 %. Melting point (uncorrected) 106 0c.

Synthesis of Ethyl 2-(2-(2-oxo-2H-2 chromen-4-yl) acetamido)-4, 5, 6, 7-tetrahydro--3H benzo[d]thiophene-3-caboxylate (step III)[20]: A mixture of stepII 1g (4.19mmol) and Ethyl 2- amino 4,5,6,7-tetrahydrobenzothiophene- 3- carboxylate 0.944g (4.19mmol) were dissolved in ethanol and this reaction is carried out in the microwave at 850C for 15min. After cooling at room temperature, crystalline product is obtained in about 3-4 hrs. It was recrystallised using ethanol. Yield 80% .Melting point (uncorrected) 2450c.

Ethyl 2- (2-(8-hydroxy-2-oxo-2H-chromen-4-yl)acetamido)-4, 5, 6, 7-tetrahydro-3H benzo [d]thiophene -3- caboxylate (III a) (A3). Percent yield: 76 % (solid). Melting point (uncorrected) 105-1180C. FTIR (K Br): cm-13667.99 (OH stretch); 3301.54(C-H stretch); 2900 (S-H stretch); 1646.91(C=O stretch of coumarin); 3573.3(N-H

stretch (20amide)); 1280.5(C-O-C (ester) stretch); 1145.51 (C=S stretch).

Ethyl 2-(2-(7-methyl-2-oxo-2H- chromen-4-yl)acetamido)-4, 5, 6, 7-tetrahydro-3H-benzo [d]thiophene -3-caboxylate (III b) (A4). Percent yield: 70%. Melting point (uncorrected) 120-1350C. FTIR (KBr):cm-12935.13(C-H stretch); 340920 (N-H stretch(20amide)); 1693.19(C=O stretch of coumarin);1646.91 (C=C stretch); 1411.64(C-O stretch (carboxylic acid)); 1276.65(C-O-C (ester) stretch); 1157.08 (C=S stretch).

Ethyl 2 - (2-(7-hydroxy-2-oxo-2H- chromen-4-yl) acetamido)-4, 5, 6, 7-tetrahydro-3H-benzo [d]thiophene-3-caboxylate (III c) (A7). Percent yield: 78%. Melting point (uncorrected) 130-1350C. FTIR (KBr): cm-1 3567.91 (OH stretch); 3021.54(C-H stretch); 2900(S-H stretch); 1646.91(C=O stretch of coumarin); 3403.3(N-H stretch (20amide)); 1280.5(C-O-C (ester) stretch); 1145.51 (C=S stretch).

Ethyl 2 - (2-(6-hydroxy-8-methyl-2-oxo-2H- chromen-4-yl) acetamido)-4, 5, 6, 7-tetrahydro-3H-benzo[d]thiophene-3- caboxylate (III d) (A20). Percent yield 75%. Melting point (uncorrected) 115-1220 C. FTIR (KBr):cm-1 3565.67 (OH stretch) ;3497.68 (N-H(20amide) stretch); 2938.98(C-H stretch(methyl)) ; 1675.05(C=O stretch of coumarin)1596.77(C=C stretch) ;1488.78 (C-O stretch); 1280.5 (C-O-C stretch).

Ethyl 2 - (2-(6, 8-dihydroxy-2-oxo-2H-chromen-4-yl) acetamido)-4, 5, 6, 7-tetrahydro-3H-benzo[d]thiophene-3-caboxylate (III e) (A21). Percent yield78%. Melting point (uncorrected) 108-125 0C. FTIR (KBr):cm-1 3405.67(OH stretch) ; 3407.68(N-H stretch (20 amide)) ; 2938.99(C-H stretch); 1646.91(C=O stretch of coumarin); 1589.06 (C=C stretch) ; 1488.78(C=S stretch);1280.5 (C-O-C (ester) stretch).

VI. BIOLOGICAL ACTIVITY

RT activity was carried out at Nation Centre for Cell Science Pune using Roche Reverse Transcriptase Assay, colorimetric. In this procedure 1ng recombinant HIV-1-RT is diluted in Lysis buffer (20ul/well). Lysis buffer without HIV-1-RT addition was used as a negative control. RT inhibitors (20ul) were diluted in Lysis buffer and 20 ul of reaction mixture (solution 3a or 3b) per reaction tube was added and incubated for 1 h at 37 0C. For the number of micro plate (MP) modules to be used, enough foil bags were opened and MP modules were put into the frame in the correct orientation. MP modules were ready to use and need not be rehydrated prior to addition of the samples. The samples (60 ul) were transferred into the wells of the MP modules. MP modules were covered with a cover foil and incubated for 2 h at 37 0C. The solution was removed completely and rinsed five times with 250 ul of washing buffer per well (solution6) for 30sec each and washing buffer was removed carefully. Later on 200 ul of anti-DIG-POD working dilution (5 ul/ml, solution 5a) were added per well, MP modules were covered with a cover foil and incubated for 1 h at 37 0C. The solution was completely removed. MP modules were rinse five times with 250 ul of washing buffer per well (solution 6) for 30 seconds each and washing buffer was carefully removed. Then 200 ul of ABTS substrate solution (1 tablet in 5ml of substrate diluents, bottle 9) was added per well and incubated at room temperature until color development (green color) was sufficient for photometric detection (10–30 min). Using micro plate (ELISA) reader, absorbance of the samples at 405 nm was measured (reference wavelength: approx. 490 nm). (Table: 7)

TABLE: 7 ANTI-HIV ACTIVITIES OF SYNTHESIZED COMPOUNDS AGAINST RT ENZYME

Compound	Optical Density	% RT Inhibition
Blank	0.000	0.000
$Control_1$	0.313	0.000
$Control_2$	0.435	0.000
A20	0.103	72.45
A4	0.159	57.14
A7	0.194	48.12
A3	0.309	17.37
A21	0.170	54.54
Nevirapine	0.035	90.12

VII. CONCLUSION

Molecular modeling studies were performed to design NCE's to inhibit RT enzyme for anti-HIV activity. The results of QSAR studies revealed that presence of electronegative group at R1, R2 and R5 position and less steric group containing electronegative and electropositive substituent at R3and R4 position increase the RT inhibiting activity. Also it was proved from the designed library of (2-oxo-2H- chromen-4-yl) acetaldehyde derivatives, that substitution of less steric(Electropositive and electronegative) groups at R3 position and also at R1 and R2 position electronegative(-OH) groups is subjected to good binding affinity for RT enzyme and have same binding pocket as standards containing amino acids Lys 101 of RT enzyme, Prediction of ADME properties of designed compound helped in selecting the compounds having drug-like properties. Finally, this compounds which shows good predicted activity in QSAR with good docking score and passes ADME properties can be subjected to wet lab study i.e. synthesis and evaluation of RT inhibitory assay studies shows the percentage of inhibition . Amongst all the synthesized compounds, A20 and A4 are

showing better % of inhibition. This indicates that pharmacophoric requirement obtained from 3D QSAR is essential for anti-HIV activity. But compound A3 which shows the poor % of inhibition may be due to the absence of electronegative group at R2 position. For new users this can be used as ideal pharmacophore to design new chemical entities to inhibit RT enzyme.

ACKNOWLEDGEMENTS

Authors are thankful to Dr. Debashis Mitra (scientist –F), NCCS (Nation Centre for Cell Sciences) Pune, for providing facilities to carry out In-vitro Reverse Transcriptase (RT) assay.

REFERENCES

[1]. Cheng, Z.; Zhang, Y.; W, Fu. Eur. J. Med. Chem. 2010, 45, 3970.
[2]. Pawar ,V.; Lokwani, D.; Bhandari, S.; Mitra, D.; Sabde,S.; Bothara, K.; Madgulkar, A. Bioorg. Med. Chem. 2010, 18, 3198.
[3]. Olomola, T. O.; Klein, R..;. Lobb, K. A.; Sayed, Y.; Kaye, P. T. Tetrahedron Lett .2010, 51, 6325.
[4]. Zhan ,P.;X, Chen.; Li ,D.; Fang ,Z.; Clercq, E. D.; Liu, X.. Medicinal Research Review.2011.Wiley Periodicals.Inc. DOI 10.1002/med.20241.
[5]. Bhavsar, D.; Trivedi, J.; Parekh, S.; Savant, M.; Thakrar, S.; Bavishi, A.; Radadiya, A.; Vala, H.; Lunagariya, J.; Parmar, M.; Paresh, L.; Loddo, R .; Shah, A. Bioorg. Med. Chem. Lett .2011, 21, 3443.
[6]. César, G. Z. J.; Alfonso, M.G. G.; Marius, M. M.; Elizabeth, E.M.; Ángel, C. B.M.; Maira, H.R. ; Guadalupe, C. L. M.; Manuel, J. E.; Ricardo, R. C. Fitoterapia. 2011, 82, 1027.
[7]. Zhou, T.; Shi, Q.; Chen, C.H.; Zhu, H.; Huan, L.; Ho, P.; Lee K. H. Bioorg.Med. Chem.2010, 18,667.
[8]. Maggiora, G. M. J. Chem. Inf. Model.; 2006, 46, 1535.
[9]. Verma, R P.; Hansch, C .Bioorg .Med.Chem.2005, 13, 4597.
[10]. Gobbi, A.; Man-Ling, L. J. Chem .Inf. Comput. Sci. 2003, 43,317.
[11]. Zheng, W.; Tropsha, A. J. Chem. Inf. Comput. Sci. 2000, 40, 185.
[12]. Verma, J.; Khedkar, V.M.; Coutinho, E.C. 2010, 95.
[13]. GLIDE, version 5.0, Molecular Docking Tool of Schrodinger Inc. New York, USA.
[14]. Manvar, A.; Malde, A.; Verma, J.; Virsodia, V.; Mishra , A .; Upadhyay, K.; Acharya, H.; Coutinho, E.; Shah ,A. Eur. J. Med. Chem. 2008 ,43, 2395 .
[15]. Patel, D.; Patel, N .; Kumari, P.; Patel, N. International J.Chem. 2011 , 3,118 .
[16]. Singh, I.; Kaur, H.; Kumar, S.; Kumar, A.; Lata, S.; Kumar, A. International J.Chem.Tech. Research. 2010, 2 , 1745.
[17]. Klos, A. M.; D, Tymoshenko. Z, Yang. J. Med. Chem., 2008, 51, 5843.
[18]. Gouda, M. A.; Berghot, M. A.; EL-GHANI, G. E. ABD.; Elattar, K. M.; Khalil ,G. M. Turk. J. Chem. 2011, 35, 815.
[19]. Kathiravan, M.; Shishoo, C.; Chitre, T.; Mahadik, K.; Jain, K. Synth.Commun. 2007 ,37,4273.
[20]. L. Porre`s, O. Mongin, C. Katan, M. Charlot, T. Pons, J. Mertz, M. Blanchard-Desce, Org. Lett. 1 (2004) 47-50
[21]. VLifeMDS; Molecular Design Suite version 3.5, V-life Sciences Technologies Pvt. Ltd., Pune, India. (www.vlifesciences.com)