

Pyridones as NNRTIs against HIV-1 mutants: 3D-QSAR and protein informatics

Utsab Debnath · Saroj Verma · Surabhi Jain ·
Setu B. Katti · Yenamandra S. Prabhakar

Received: 9 May 2013 / Accepted: 14 July 2013 / Published online: 25 July 2013
© Springer Science+Business Media Dordrecht 2013

Abstract CoMFA and CoMSIA based 3D-QSAR of HIV-1 RT wild and mutant (K103, Y181C, and Y188L) inhibitory activities of 4-benzyl/benzoyl pyridin-2-ones followed by protein informatics of corresponding non-nucleoside inhibitors' binding pockets from pDBs 2BAN, 3MED, 1JKH, and 2YNF were analysed to discover consensus features of the compounds for broad-spectrum activity. The CoMFA/CoMSIA models indicated that compounds with groups which lend steric-cum-electro-positive fields in the vicinity of C5, hydrophobic field in the vicinity of C3 of pyridone region and steric field in aryl region produce broad-spectrum anti-HIV-1 RT activity. Also, a linker rendering electronegative field between pyridone and aryl moieties is common requirement for the activities. The protein informatics showed considerable alteration in residues 181 and 188 characteristics on mutation. Also, mutants' isoelectric points shifted in acidic direction. The study offered fresh avenues for broad-spectrum anti-HIV-1 agents through designing new molecules seeded with groups satisfying common molecular fields and concerns of mutating residues.

Keywords Pyridones · NNRTIs · HIV-1 RT · CoMFA · CoMSIA · Protein informatics

Abbreviations

NNRTIs	Non-nucleoside reverse transcriptase inhibitors
QSAR	Quantitative structure–activity relationship
CoMFA	Comparative molecular field analysis
CoMSIA	Comparative molecular similarity indices analysis
NNIBP	Non-nucleoside inhibitors binding pocket

Introduction

Acquired immunodeficiency syndrome (AIDS) is one of the most dreaded infections of the present time and needs greater attention to overcome the disease burden on the society [1]. The human immunodeficiency virus-1 (HIV-1) is its causative agent [2]. It is composed of 1 chromosome, 9 genes and 9 proteins [3]. HIV-1 Gag-Pol polyprotein is one among these nine proteins and reverse transcriptase (RT) is a segment of this one. Among different druggable targets of HIV-1, RT is a promising one for the treatment of AIDS [4]. Depending on enzyme's binding domain, the compounds are classified as nucleoside reverse transcriptase inhibitors and non-nucleoside reverse transcriptase inhibitors (NNRTIs). Furthermore, the X-ray crystallographic studies of HIV-1 RT–ligand complexes showed that a variety of non-nucleosides bind to allosteric site of the enzyme [5, 6]. This has prompted researchers to explore diverse non-nucleoside scaffolds as potential RT inhibitors [6, 7]. Additionally, non-nucleoside analogues have gained prominence in the treatment of HIV-1 as they cause less toxicity to the host [8].

While HIV-1 RT is a promising target for drug development, its genome is known to rapidly undergo spontaneous mutations to result in new variants [7]. Besides this,

Electronic supplementary material The online version of this article (doi:10.1007/s10822-013-9667-1) contains supplementary material, which is available to authorized users.

U. Debnath · S. Verma · S. Jain · S. B. Katti ·
Y. S. Prabhakar (✉)

Medicinal and Process Chemistry Division, CSIR-Central Drug
Research Institute, Lucknow 226001, India
e-mail: yenpra@yahoo.com

the pressure of xenobiotics is an additional factor for the manifestation of mutations in the drug targets and thereby in organisms [9]. The fallout of this situation is quick emergence as well as spread of drug-resistant viral strains, failure of treatment followed by clinical complications and surge of more infections. Currently some of the clinically practiced drugs namely Nevirapine, Efavirenz, Delavirdine etc. are reported to have become less effective in the treatment of infection [10–12]. With the resistant strains escaping the action of clinically practiced drugs or rendering them ineffective, there is an urgent need for new molecules which can be efficient against both wild type and resistant strains to reverse the situation. The scenario calls for reassessment of anti-retroviral therapies and strategies aiming at compounds/devises with broad spectrum activity profile against multiple strains.

In drug research, discovering remedies to tackle mutant organisms and overcoming drug-resistance are big challenges. In case of the HIV-1, it is desirable to have RT inhibitors which can bind to the allosteric site of wild type as well as mutant strains and elicit equal or comparable therapeutic response against all of them. In the recent past Benjahad, Nguyan, Guillemont and co-workers have investigated different 4-benzyl/benzoyl pyridin-2-ones and related compounds as inhibitors of the HIV-1 RT of wild type and its clinically important mutants [13–15]. This has resulted in the identification of different pyridin-2-one derivatives as potent inhibitors of wild type and mutant strains which concomitantly showed moderate to reasonable activity against the other strains.

The centre stage of medicinal chemistry is biological response of probing molecules which give first insight of the activity (or receptor) space in terms of the chemical landscape. The quantitative structure–activity relationship (QSAR) and molecular modelling protocols and bioinformatics studies are selected rational drug design approaches which provide direction to the design of new or modified therapeutics. They fine-tune the notional insight of biological response in terms of the structure space and its properties, thereby give an opportunity to understand and explore the possibilities around the prospective chemical scaffold. Different researchers have studied the QSAR and modelling aspects anti-HIV agents stemming from varying scaffolds [16–18]. For the NNRTI scaffolds, these investigations have identified/ratified the importance hydrophobicity/steric and butterfly-like disposition of the structural moieties [16–18]. Our earlier studies on 2,3-diarylthiazolidinones and related compounds as NNRTIs has highlighted the significance of aryl, heteroaryl moieties and their topology for the activity [19–25]. A vast majority of these studies have deliberated on single viral strains.

Drug target mutation and drug resistance are chained events. Even though the former event paves the way for the

later, the tipping is registered only due to the later event. Moreover, as the former event is part of a natural course, we have little handle to contain it except comprehending the phenomenon. This prompted us to undertake a comparative molecular field analysis (CoMFA) [26] and comparative molecular similarity indices analysis (CoMSIA) [27] directed 3D-QSAR study of wild type and mutant HIV-1 RT inhibitory activities of 4-benzyl/benzoyl pyridin-2-ones [13–15] and simultaneously explore the protein informatics of the non-nucleoside inhibitors' binding pocket (NNIBP) to discover the consensus structural features of the compounds for activity against multiple strains. The putative bioactive conformations of the compounds for the CoMFA/CoMSIA study were adopted from the X-ray co-crystal structures and docking experiments. Furthermore, to shed some light on the changing characteristics of the NNIBP of HIV-1 RT on mutation, we analyzed the FASTA sequences and X-ray crystal structures of the wild type and mutant proteins with different bioinformatics tools. The findings of the CoMFA/CoMSIA together with the NNIBP characteristics were used to design new molecules as potential broad spectrum inhibitors of HIV-1 strains. The results are presented hereunder.

Materials and methods

Data set

One hundred and seventy-eight 4-benzyl/benzoyl pyridin-2-one derivatives (Fig. 1; hereafter referred as pyridone derivatives) and their HIV-1 RT inhibitory activities against wild type and mutant strains K103N, Y181C and Y188L (IC₅₀, molar dose of compound to produce 50 % inhibition activity) reported in the literature were considered for this study [13–15]. Even though the total number of compounds is one hundred and seventy-eight, for some compounds activities against mutant strains were not reported. Apart from this, for a limited number of analogues the anti HIV-1 activities were reported using additional mutant strains. Since the number of such analogues is small, these additional mutants are not made part of the current study. The datasets considered for the study showed good variation in terms of structures and their activity. To maintain brevity, only a limited number of compounds referred in the discussion are shown in Table 1. However, complete data has been provided as part of supplementary material. For the purpose of computation, the activities of the compounds were transformed into logarithm of inverse of inhibitory concentration and expressed as pIC₅₀. The protein informatics of HIV-1 RT wild type and mutants were investigated using their FASTA sequences and X-ray crystal structures retrieved from the protein databank. In

this study pdb crystal structures 2BAN [28], 3MED [29], 1JKH [30], and 2YNF [31] were deployed as representatives of wild type, K103N, Y181C and Y188L strains, respectively.

Before commencing CoMFA/CoMSIA study, the compounds with reported activities against wild type and mutant strains were arbitrarily partitioned into training and test sets with the later one holding one-third of the compounds. This has resulted in following composition of training and test sets for the wild type and mutant strains: wild type, 119 compounds in training set and 59 compounds in test set; K103N, 97 compounds in training set and 48 compounds in test set; Y181C, 98 compounds in training set and 48 compounds in test set; Y188L, 91 compounds in training set and 45 compounds in test set. Table 2 shows the spread of compounds' activity in the training and test sets. Only the training set compounds are used for the model generation and test set is used for the external validation of developed models.

Molecular conformations and alignment

The wild type HIV-1 RT pdb 2BAN [28] has compound **44** as ligand in its co-crystal structure. It is a compound with high order of activity against wild type. Since the biological response is an outcome of interaction between the ligated molecule(s) and the receptor (macromolecule), the pose of the former in the latter's binding domain offers clue to the molecular space for the activity. Here, the accessibility of suitable protein–ligand co-crystal structures or such information frees the 3D-QSAR from bias towards an arbitrarily chosen 'best' conformation, which may be 'global minimum' or any other reference conformation of high active compound. Moreover, the co-crystal structure of the ligand while defining the desirable chemical structure space of the receptor also gives some

credence of 'direct drug design' concepts. Hence the pose of compound **44** in 2BAN is adopted as bioactive conformation of wild type. Furthermore, compounds **126**, **61** and **60** are high activity analogues against K103N, Y181C and Y188L strains, respectively. The putative bioactive conformations of these compounds were determined by respectively using the pdb 3MED [29], 1JKH [30], and 2YNF [31]. These co-crystal coordinates were used in AutoDock [32] to identify the best docked conformations of compounds **126**, **61** and **60** against respective viral strain (Fig. 2). All these conformers were used in the SKETCH module of SYBYL [33] as templates to build the molecular structures of remaining compounds. The generated 3-D structures were energy minimized using the Tripos force field [34] with a distance-dependent dielectric function and the Powell conjugate gradient algorithm with a convergence criterion of 0.01 kcal/mol Å and iterations up to one thousand. Partial atomic charges were calculated using the Gastieger-Huckel method. Since the pyridone moiety is the common core unit for all the compounds, the conformers obtained against the respective viral strains were aligned using this common moiety of the template molecules (Fig. 1). The structures were aligned on the template molecules with the atom based root-mean-square fitting method [35]. It may be reemphasized here that in this study the putative bioactive conformations of all compounds against each viral strain were built from the appropriate co-crystal and docked poses of high active compounds which may be referred as "pose-then-build" approach. This has provided comparable core conformations for all the analogues. Besides this, it is computationally faster and less tiresome. On the contrary, opting for the docked poses of all compounds in the alignment may lead to highly proliferated/diffused molecular aggregates. As molecular alignment is a key step of CoMFA/CoMSIA, this may affect the outcome of the analysis.

Fig. 1 General structure of 4-benzyl/benzoyl pyridin-2-ones and related compounds present in the database. The diversity in the substitution is briefly shown in R_1 , R_2 , R_3 , R_4 and X . The R/R' in the substituents denote alkyl and functionalized alkyl moieties; Y/Z denote hetero atoms; n (0–2) indicates variable ring/bridge size. The asterisk-marked points/atoms are common to all structures and are used for the CoMFA/CoMSIA alignment of the molecules

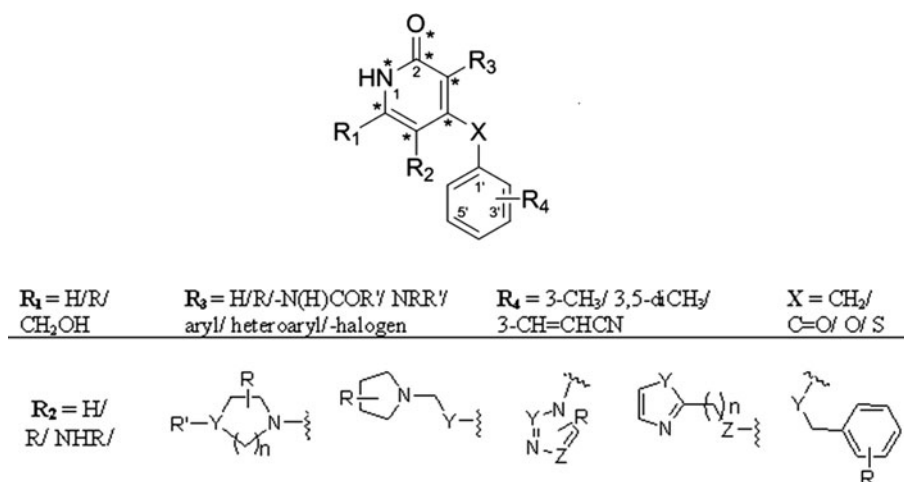


Table 1 Selected pyridone derivatives (Fig. 1) and their reported and CoMFA/CoMSIA predicted HIV-1 RT inhibitory activities against Wild type and mutant strains

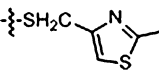
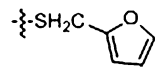



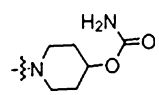
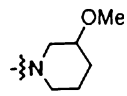
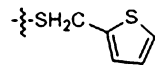


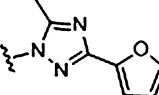

S.No ^a	CNo ^b	R ₂	R ₃	pIC ₅₀ (Wild)			pIC ₅₀ (K103N)			pIC ₅₀ (Y181C)			pIC ₅₀ (Y188L)		
				Obs ^c	Pr1 ^d	Pr2 ^e	obs	Pr1	Pr2	Obs	Pr1	Pr2	Obs	Pr1	Pr2
1	2	Et	NMe ₂	8.40	8.06	8.15	8.00	7.63	6.94	7.20	6.74	6.79	6.80	6.99	6.31
2	16	(CH ₂) ₂ OMe	NMe ₂	8.40	8.42	8.59	6.80	7.13	7.15	5.90	6.46	6.43	5.00	5.18	5.47
3	19	Et	NMe ₂	9.00	8.37	8.56	8.52	8.16	7.53	7.14	7.00	7.13	6.20	6.33	6.17
4	29	Et	NHCOEt	5.40	5.10	4.92	4.00	3.92	4.59	4.00	3.94	4.27	NR ^f	3.11	2.37
5	30	Et	NHCOPr	4.00	3.76	3.90	4.00	3.68	4.12	4.00	4.16	4.45	NR	3.29	2.14
6	32	Et	NHPr	7.70	7.96	8.06	5.80	5.55	5.38	4.80	5.70	4.97	4.00	4.00	4.27
7	35	Et	NBu ₂	4.30	4.41	4.81	4.00	4.63	4.25	4.00	3.53	3.72	NR	3.09	4.73
8	37	Et	N(CH ₂ Ph) ₂	4.00	3.72	4.17	4.30	4.34	4.33	4.10	4.13	4.17	NR	3.22	2.73
9	44	Et	N(Me)C ₂ H ₄ OMe	9.00	8.75	8.34	7.10	6.88	6.57	6.10	5.93	5.29	6.40	6.39	6.61
10	46	Et	N(Me)CH ₂ CN	8.40	8.19	8.00	6.50	6.92	6.82	5.80	5.99	5.85	4.90	5.31	5.27
11	54	Et	N(Me)Et	9.00	8.93	8.27	8.22	8.38	7.38	7.49	8.09	6.60	6.60	6.54	6.36
12	55	Et	N(Me)C ₂ H ₄ OMe	9.00	9.30	8.86	7.70	8.17	7.93	7.40	7.24	6.74	6.50	6.73	6.23
13	56	Et	I	8.90	8.54	8.14	8.50	8.31	8.51	7.70	7.18	7.83	7.30	6.63	6.43
14	59	OEt	I	9.20	8.69	8.68	8.20	8.03	7.86	8.20	7.83	7.45	6.90	6.93	7.04
15	60		I	9.00	8.56	8.72	9.00	8.84	8.93	8.20	7.61	7.57	8.10	7.94	7.98
16	61		I	8.90	8.97	8.82	9.00	8.63	8.40	8.30	7.60	7.69	7.50	7.43	6.78
17	84		I	8.70	8.64	8.57	8.60	7.64	7.83	8.00	7.50	7.08	7.50	7.99	7.67
18	92		I	8.20	8.45	8.36	7.90	8.05	7.52	7.40	7.66	8.04	6.90	7.11	7.59
19	93		I	8.20	8.41	7.94	8.10	8.48	7.98	7.50	7.57	8.00	7.50	7.41	7.37
20	101		I	8.20	8.34	8.40	8.70	8.61	8.71	7.60	7.37	7.39	7.30	7.45	6.89
21	102		I	8.20	8.11	8.14	8.20	8.31	8.40	7.80	7.23	7.55	7.40	7.19	7.21
22	119		I	8.00	8.01	8.29	8.00	8.32	8.36	7.40	7.30	7.22	6.10	5.72	5.46
23	120		I	9.20	9.27	8.75	8.90	8.21	8.21	8.00	8.26	7.85	7.50	7.05	7.22
24	126		I	8.90	8.95	8.92	9.51	9.49	9.38	8.10	8.09	7.90	7.80	7.59	7.57
25	130		I	8.70	8.64	8.71	7.90	7.54	7.93	7.50	7.59	7.74	7.50	7.15	6.99

Table 1 continued

S.No ^a	CNo ^b	R ₂	R ₃	pIC ₅₀ (Wild)			pIC ₅₀ (K103N)			pIC ₅₀ (Y181C)			pIC ₅₀ (Y188L)		
				Obs ^c	Pr1 ^d	Pr2 ^e	obs	Pr1	Pr2	Obs	Pr1	Pr2	Obs	Pr1	Pr2
26	143		I	8.60	8.49	8.57	9.21	8.97	8.95	7.60	7.29	7.52	7.30	7.04	7.18

^a S.No is serial number of the entry in the table. For the compounds shown in the table the variations at R₁, R₄ and X (Fig. 1) are as follows: R₁ is CH₂OH in S.No.3 and Me in all others; R₄ is 3-Me in S.Nos. 2, 9 and 10, 3-C₂H₂CN in S.Nos. 11 and 12, and 3, 5-di-Me in all others; X is CO in S.Nos. 1–3, CH₂ in S.Nos. 4–12, O in S.Nos. 13–21 and S in S.No. 22–26

^b CNo is compound number as used in the discussion

^c Observed activity

^d CoMFA predicted activity

^e CoMSIA predicted activity

^f Not reported

Table 2 Distribution pattern of pyridones' HIV-1 RT inhibitory activities in training and test sets

Dataset	Compounds	Activity (pIC ₅₀) distribution			
		Min	Max	Avg	SD
<i>Wild type</i>					
Training	119	3.80	9.20	7.84	1.22
Test set	59	4.20	9.10	7.77	1.01
<i>K103N</i>					
Training	97	4.00	9.51	7.50	1.36
Test set	48	4.30	9.00	7.28	1.08
<i>Y181C</i>					
Training	98	4.00	8.30	6.85	1.14
Test set	48	4.00	8.20	6.76	0.94
<i>Y188L</i>					
Training	91	4.00	8.10	6.32	0.93
Test set	45	4.00	7.50	6.38	0.82

CoMFA and CoMSIA indices

For CoMFA indices, the molecular fields were sampled at default lattice spacing of 2 Å with the lattice extending to 4 Å beyond the aligned molecules in all directions. A sp³ carbon probe atom with default van der Waals radius and unit positive charge was used to generate the steric (Lennard-Jones 6–12 potentials) and electrostatic (Coulombic potentials) field energies with a distance-dependent dielectric constant at each lattice point. In order to minimize the influence of noise, the generated fields were scaled using the CoMFA-Standard method with default cut-off energy (30 kcal/mol) and column filter (2.0 kcal/mol) [33].

For CoMSIA indices of the molecules, the same lattice box as used for the CoMFA was adopted. The indices were derived according to Klebe et al. [36]. Following the

default procedure as implemented in SYBYL, carbon with unit positive charge was used for the computation of similarity indices (A_F). The indices (A_F) for molecule j with atoms i at a grid point q were calculated using Eq. 1.

$$A_{F,K}^q(j) = - \sum_{i=1}^n \omega_{\text{probe},k} \omega_{ik} e^{-\alpha r_{iq}^2} \quad (1)$$

The CoMSIA method involves computation of five different physicochemical properties (k) namely steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor as molecular similarity indices (Eq. 1). The computation uses a Gaussian type distance-dependent function between the grid point q and each atom i in the molecule. A default value of 0.3 was used as the attenuation factor (α). In CoMSIA, the steric indices are related to the third power of the atomic radii; the electrostatic descriptors are derived from partial atomic charges; the hydrophobic fields are derived from atom based parameters [37] and the hydrogen-bond donor and hydrogen-bond acceptor atoms within a putative protein environment are derived from experimental values [38].

Model derivation

The CoMFA and CoMSIA 3D-QSAR models were derived using the PLS regression procedure of SYBYL [39]. The predictive ability of the model was measured in terms of cross-validated ($r^2(r_{cv}^2$ or Q^2)) as shown in Eq. 2.

$$r_{cv}^2 = 1 - \frac{\sum (Y_{\text{predicted}} - Y_{\text{observed}})^2}{\sum (Y_{\text{observed}} - Y_{\text{mean}})^2} \quad (2)$$

Where $Y_{\text{predicted}}$, Y_{observed} , and Y_{mean} are predicted, actual, and mean values of the target property (pIC₅₀), respectively. The number of components leading to the lowest standard error of prediction (SEP) were used as optimum

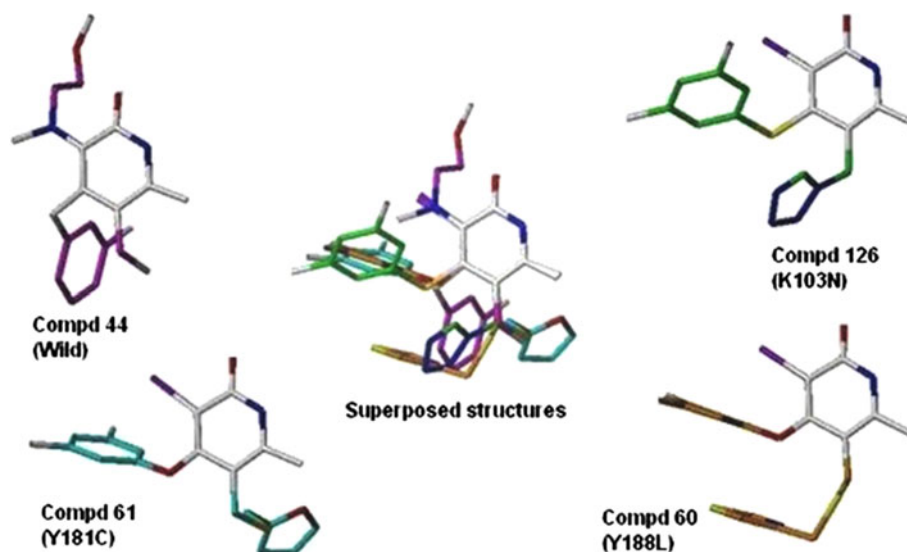


Fig. 2 The putative bioactive conformations of compounds **44**, **126**, **61** and **60** identified from the docking experiment for binding to the HIV-1 RT wild, K103N, Y181C and Y188L strains, respectively and their pyridone moiety based superposed structures. In all structures the carbon skeleton of pyridone moiety is shown in gray. The default colour codes are used for the other heavy atoms (oxygen in red,

nitrogen in blue and sulfur in yellow). In order to facilitate the structure space differences, the remaining part of the carbon skeleton in each compound is shown in different colour as follows: compound **44** in magenta; compound **126** in green; compound **61** in cyan; compound **60** in orange; hydrogens are suppressed for clarity

number of components to generate the final PLS regression models. The models were validated through the bootstrapping analysis for 100 runs and the cross-validation analysis (leave-half-out and leave 20 % out; each 50 runs) [39]. The CoMFA and CoMSIA equations were plotted as contour maps to express the percentage contribution of respective fields to the activity. This was done by considering the field energies at each lattice point as a multiple of the regression coefficient and corresponding standard deviation. Here, the contour maps help in identifying the important regions where changes may affect the binding preference and thereby facilitate the recognition of key features contributing to the interactions between the ligand and the active site of a receptor. Furthermore, the developed CoMFA and CoMSIA models were validated by predicting the activity of the external test set compounds. A model with predictive r^2 value (r^2_{pred}) more than 0.5 may be considered as statistically significant.

Protein informatics

FASTA sequences of identified proteins were aligned with CLUSTALW [40]. The sequence variations with propensity between wild and mutant strains were analyzed using BioEdit software [41]. The residues of ligand binding pocket were identified in MOE (Molecular Operating Environment) software [42]. The bioinformatics tools namely SIFT (Sorting Intolerant from Tolerant) [43], PANTHER (Protein ANalysis THrough Evolutionary

Relationships) [44], I-Mutant 2.0 [45], and CUPSAT (Cologne University Protein Stability Analysis Tool) [46] were used to study the influence of point mutations on the protein structure. The changes in packing volumes, accessible surface areas and free energy of folding of wild type and mutant proteins were studied in VADAR (Volume, Area, Dihedral Angle Reporter) [47].

Results and discussion

CoMFA and CoMSIA statistics

The goodness-of-fit of derived CoMFA and CoMSIA models for pyridone derivatives as HIV-1 RT inhibitors against wild type and mutants K103N, Y181C and Y188L are shown in Table 3. The optimum value of Q^2 was used as criteria to select the significant models. For the wild type HIV-1 RT, the optimum CoMFA model has been found to involve five PLS components with Q^2 value of 0.786. It showed non-cross validated r^2 value of 0.947 and predicted 74.1 % variance ($r^2_{pred} = 0.741$) in the activity of external test set compounds (Table 3). Taking it as reference, five PLS components have been opted for the CoMFA models of HIV-1 RT mutant strains K103N, Y181C and Y188L. These models have shown Q^2 values from 0.735 to 0.615 and non-cross validated r^2 values from 0.930 to 0.904 and predictive r^2 (r^2_{pred}) values from 0.967 to 0.852 (Table 3). In all CoMFA models the steric field has been found to be

Table 3 Goodness of fit of CoMFA and CoMSIA models for the HIV-1 RT inhibitory activities of pyridones against Wild type and mutant strains

	CoMFA				CoMSIA			
	Wild	K103N	Y181C	Y188L	Wild	K103N	Y181C	Y188L
r_{ncv}^2 ^a	0.947	0.930	0.904	0.913	0.909	0.880	0.886	0.852
SEE ^b	0.238	0.363	0.361	0.279	0.314	0.481	0.426	0.402
F -test ^c	269.67	240.97	173.96	167.39	227.12	132.25	119.57	93.87
Q^2 ^d	0.786	0.735	0.629	0.615	0.809	0.727	0.657	0.609
SEP ^e	0.576	0.646	0.689	0.608	0.545	0.639	0.670	0.576
r_{pred}^2 ^f	0.741	0.967	0.852	0.897	0.616	0.907	0.763	0.802
r_{boot}^2 ^g	0.960	0.957	0.933	0.919	0.929	0.916	0.905	0.884
SEE_{boot}^h	0.247	0.282	0.303	0.268	0.330	0.398	0.357	0.320
r_{LHO}^2 ⁱ	0.529	0.659	0.511	0.517	0.535	0.582	0.556	0.528
SD_{LHO}^j	0.068	0.055	0.079	0.071	0.088	0.081	0.071	0.054
r_{5cv}^2 ^k	0.680	0.738	0.517	0.543	0.555	0.668	0.588	0.561
SD_{5cv}^l	0.085	0.034	0.053	0.040	0.062	0.033	0.051	0.039
<i>Field contributions</i>								
Steric	0.653	0.659	0.611	0.581	0.554	0.537	0.513	0.315
Electrostatic	0.347	0.341	0.389	0.419	—	—	—	0.425
Hydrophobic					0.446	0.463	0.487	0.260

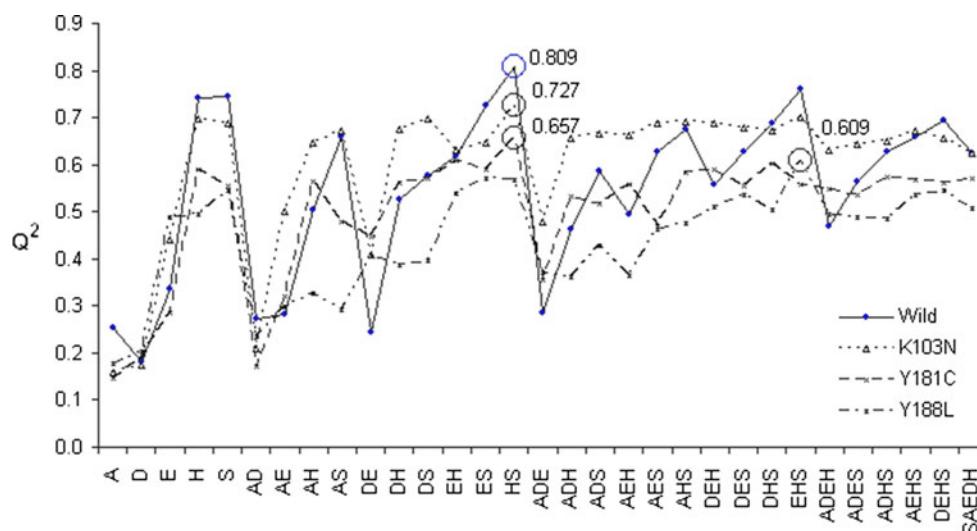
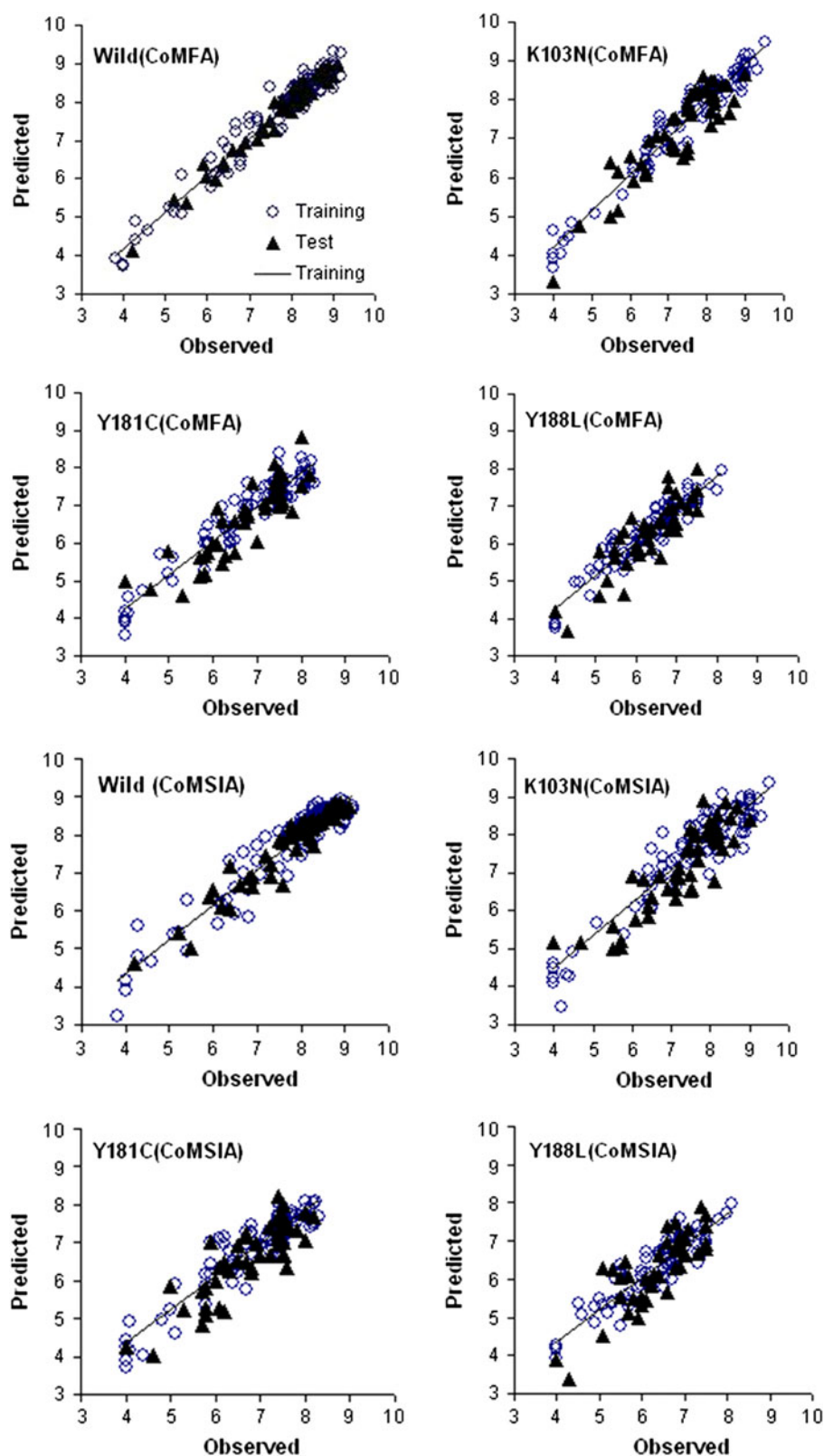
^a The non-cross validated correlation coefficient. For all models optimal number of PLS components are five^b Standard error of estimate^c Ratio of r^2 explained to unexplained = $r^2/(1 - r^2)$ ^d Leave-one-out predicted cross-validated correlation coefficient^e Standard error of prediction^f External predicted correlation coefficient for test set of compounds^g Average of correlation coefficient for 100 samplings using bootstrapped method^h Average standard error of estimate for 100 samplings using bootstrapped methodⁱ Average cross-validated correlation coefficient for 50 runs using leave-half-out (LHO) group^j Standard deviation of average cross-validated correlation coefficient for 50 runs^k Average cross-validated correlation coefficient for 50 runs using five cross-validation group^l Standard deviation of average cross-validated correlation coefficient for 50 runs**Fig. 3** Plots of Q^2 values of CoMSIA models resulted from single and multiple field combinations for pyridone analogues as HIV-1 RT inhibitors against wild and mutant strains. On the x-axis the fields are identified with single letters as A for HB-Acceptor, D for HB-donor, E for electrostatic, H for hydrophobic and S for steric. For each viral strain the field combinations which led to the model with optimum Q^2 is marked with circle

Fig. 4 The plots of experimental pIC_{50} values versus CoMFA/CoMSIA training (circle) and the test set (filled triangle) predicted pIC_{50} values of pyridones for different HIV-1 RT strains



the major contributor (58.1–65.9 %) in explaining the activity of the compounds followed by the electrostatic field (Table 3).

In the CoMSIA study also, five components were found to optimally explain the wild type HIV-1 RT inhibitory activity of the compounds. It has shown Q^2 value of 0.625

Table 4 Gist of CoMFA and CoMSIA contours in the molecular regions of pyridones with reference to the HIV-1 RT inhibitory activities against wild type and mutant strains

Contour region ^a	Field ^b	CoMFA ^c				CoMSIA ^c			
		Wild	K103N	Y181C	Y188L	Wild	K103N	Y181C	Y188L
R ₁	S								
	E								
	H					UF			
R ₂	S	FA	FA	FA	FA	FA	FA	FA	FA
	E	+Ve	+Ve	+Ve	+Ve				+Ve
	H					UF			UF
R ₃	S	UF	UF	UF		UF	UF	UF	
	E								
	H					FA	FA	FA	FA
R ₄	S	FA	FA	FA	FA	FA	FA	FA	FA
	E								
	H							UF	UF
X	S								
	E	−Ve	−Ve	−Ve	−Ve				−Ve
	H						UF	UF	

^a specifies the region of contour in the molecular space identified by varying positions (Fig. 1)

^b S, E and H represent steric, electrostatic and hydrophobic fields, respectively

^c FA for favorable; UF for unfavorable; +Ve for electropositive; −Ve for electronegative

and a non-cross validated r^2 value of 0.878. Taking cue from the wild type model, to facilitate comparison, five components were considered as optimum for the CoMSIA models of K103N, Y181C and Y188L strains. Here, an examination of the models has indicated that all five fields of CoMSIA are not equally important for explaining the activity of the compounds. Hence, single as well as different combinations of CoMSIA fields have been assessed to determine their significance in explaining the variance in the activity. The Q^2 values from this study have indicated the primacy of steric, hydrophobic and/or electrostatic fields in explaining the variance in the HIV-1 RT inhibitory activities of these compounds (Fig. 3). Nevertheless, we can not deny the significance of H-bond donors/H-bond acceptors in drug-receptor interactions. Notwithstanding this fact, these fields are not incorporated into the models due to diminished statistical significance.

For the HIV-1 RT wild type, K103N and Y181C strains, the optimum CoMSIA models have emerged from steric and hydrophobic fields with Q^2 values of 0.809, 0.727 and 0.657, respectively (Fig. 3, Table 3). In case of Y188L strain, the CoMSIA model has involved electrostatic field together with steric, hydrophobic fields to optimally explain the activity of compounds ($Q^2 = 0.609$) (Fig. 3, Table 3). All CoMFA and CoMSIA models are found to be robust in bootstrapping and different cross-validation tests (Table 3). All models have satisfactorily predicted the activities of training and test set compounds. The plots of

observed versus predicted activities of training and test sets from the CoMFA and CoMSIA models are shown in Fig. 4. While Fig. 2 provides poses of putative bioactive conformations of the study, to maintain brevity CoMFA/CoMSIA alignments and the predictions from the models are given in the supplementary information.

Contour map analysis

The major contours surfaced in the molecular regions of pyridones and their influence on the HIV-1 RT inhibitory activities of wild type and mutants are summarized in Table 4. For the convenience of discussion, the contours encircling and/or in the vicinity of pyridone ring are identified and addressed with its varying groups R₁, R₂ and R₃ (Fig. 1). Similarly, the contours positioned in the vicinity of aryl moiety and linker (X) region are discussed using R₄ and linker (X), respectively (Fig. 1).

CoMFA steric

The CoMFA steric contour maps of pyridones against wild type and mutant strains of HIV-1 RT are shown in Fig. 5. In this, green and yellow colour contours respectively indicate the sterically favourable and unfavourable sub-structure space for the activity. In the wild type, the presence of green contour at the aryl moiety (Fig. 5, Table 4) indicates sterically favourable scenario for substituent

groups in this region leading to better inhibition as exemplified by the compounds **19**, **119** (Table 1) and many others.

Also the presence of a green contour in the vicinity of R_2 (Fig. 5; Table 4) indicates the favourability of this position for bulky groups. In several high active analogues (Compounds **84**, **102**, **120**, **126** etc., Table 1) this position is satisfied by different heterocycles such as imidazole, piperidine and tetrazole moieties. Apart from these, two large yellow contours have situated in the proximity of R_3 (Fig. 5, Table 4) and a small yellow contour has enclosed linker (X) moiety (Fig. 5, Table 4). This speaks in favour of minimum steric crowding at these positions. For the compounds under investigation, R_3 is represented by halogens, primary and secondary amines and alkyl groups. An examination of the contour surrounding this region and the substituent groups contained therein indicate the choice towards iodo- or like groups for R_3 in order to improve the activity. The CoMFA steric contour maps of K103N and Y181C have shown almost same trend as that of wild type (Fig. 5, Table 4). Y188L has shown some difference with the wild due to missing contours in the vicinity of R_3 region (Table 4).

CoMFA electrostatic

Figure 6 shows CoMFA electrostatic contour map for the wild type and mutant strains of HIV-1 RT inhibition by pyridone derivatives. In this, the areas where electronegative groups increase the activity are represented in red, and the areas where electropositive groups increase the activity are represented in blue. In Fig. 6, the blue contour in the vicinity of R_2 group of pyridone ring (Fig. 6, Table 4) reflected the favourability of electropositive substitution at this position for the activity (e.g., piperidine/morpholin moieties in compounds 101, 102 and 143; thiazole/tetrazole moieties in compounds, 60, 120, 126; trazole, imidazole and pyrrole moieties in compounds 84, 92 130 etc., Table 1). In these compounds, tetrazole moiety is rather neutral. As discussed in CoMFA steric, R_2 is also choice region for steric field. Moreover, CoMFA model has indicated steric field as major contributor to the activity. Probably, tetrazole, morpholine and thiazole moieties are making significant contribution due to their steric fields. However, the dataset has good number of compounds with electropositive moieties (trazole, imidazole and pyrrole moieties) as R_2 substituents. This has led to the development of clear electropositive contour in this region. The absence of electropositive and/or steric groups in place of R_2 made compounds 30, 35 etc. as less activity (Table 1). The red contour near to the linker X suggested in favour of electronegative moieties (X is O or CO; compounds **2**, **16**, **56** etc., Table 1) for this position (Fig. 6, Table 4). The

molecular inferences from electrostatic contour maps of wild type and K103N are almost same and therefore conveyed same kind of information (Fig. 6). Also, the electrostatic contour maps of Y181C and Y188L shared commonalities between them (Fig. 6). However, they differed with wild type (also K103N) only in their prominence. Hence, to a great extent the inferences emerged for the wild type also hold good for the three mutant strains. Moreover, for all, a blue contour is located near pyridone moiety indicating the significance of electropositive ring system for the activity.

CoMSIA steric

The steric contributions identified from the CoMFA and CoMSIA analysis are in general agreement with each other (Table 4) (CoMSIA steric contour maps are given in supplementary material). Here also the sterically favourable areas are indicated in green contour and the disfavoured regions are indicated in yellow contour. For wild type, green contours existed close to R_2 and R_4 regions (Table 4) indicating the favourability of bulky groups for activity. The presence of thiazole, imidazole, tetrazole and like moieties appears to be favourable for R_2 for the activity (e.g., compounds **60**, **93**, **120**, **126** etc., Table 1). The substituted aryl moiety has satisfied the steric requirements of the green contour corresponding to R_4 region. Most of the compounds in the dataset share this feature (e.g., compounds **60**, **84**, **93**, **126** etc.) (Tables 1, 4). Apart from the foregoing, a yellow contour was found near R_3 . It indicates that bulky groups in this region reduce the activity (compounds, **29**, **30**, **37** etc., Table 1). The CoMSIA steric contour maps of K103N and Y181C indicated comparable steric requirements as that of the wild type (Table 4). The steric contour map of Y188L showed no significant field in the vicinity of R_3 region.

CoMSIA hydrophobic

The CoMSIA hydrophobic contour maps are shown in Fig. 7 with yellow contours for hydrophobic favourable regions and gray contours for hydrophobic unfavourable regions. In the map for wild type, yellow contour near R_3 position indicate the favourability of this region for hydrophobic groups. In high active compounds this space is occupied by the groups like iodo-, dimethyl amine (compounds **19**, **54**, **55**, **59** etc., Table 1). Probably, the hydrophobic component of these groups along with other characteristics may be suitable for this position. Apart from this, a large gray contour encircled both R_1 and R_2 positions on the pyridone moiety (Fig. 1). This indicates the unfavourable environment for hydrophobic moieties in the vicinity of one side of the pyridone marked by R_1 and R_2 groups.

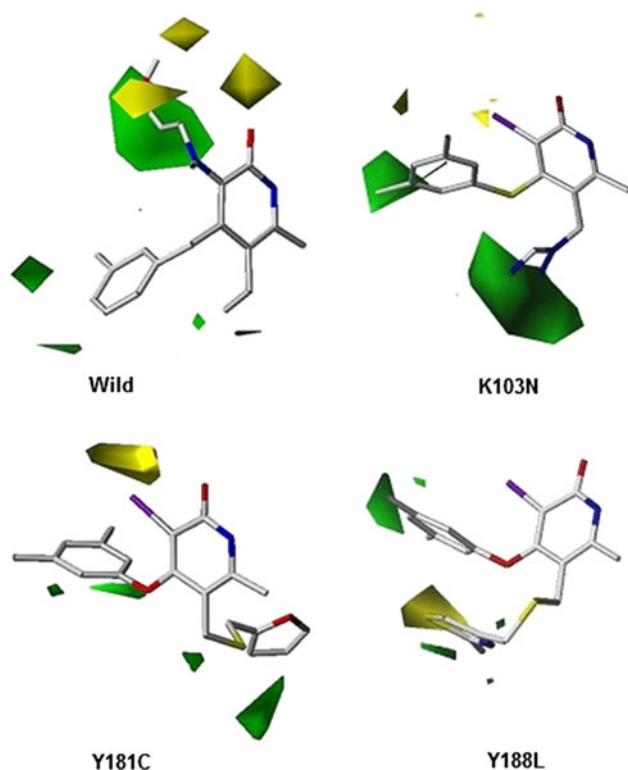


Fig. 5 CoMFA 3D-QSAR steric field contour maps emerged for pyridones as HIV-1 RT inhibitors against wild type and mutant strains K103N, Y181C and Y188L of the virus. Colour code: *green* contours indicate sterically favourable regions and *yellow* contours indicate sterically unfavourable regions for the activity. In this and subsequent contour maps the poses of compounds embedded in them correspond to the putative bioactive conformations of compounds **44** (Wild type), **126** (K103N), **61** (Y181C) and **60** (Y188L) identified from the docking experiments. In this and subsequent contour map figures, the poses shown have provided optimum visibility of the fields

In case of mutant strains, the CoMSIA hydrophobic contour map of Y188L has shown same nature of fields with respect to R_2 and R_3 of wild type. It has also shown to some extent hydrophobically unfavourable region in the vicinity of aryl moiety (Fig. 7, Table 4). This further emphasized the importance of steric field in the aryl region compared to other types. The contour maps of K103N and Y181C indicated the need for hydrophobic environment around R_3 substituent space as in case of wild type and Y188L strain. Similar to Y188L strain they showed unfavourable environment for hydrophobic groups in the linker region (Fig. 7). The Y181C showed a gray contour in the vicinity of aryl moiety (Fig. 7, Table 4). As in case of Y188L, it may be interpreted that steric field is vital in this region compared to other types.

CoMSIA electrostatic

In CoMSIA analysis, the electrostatic field showed prominence only in the model of Y188L (Fig. 8, Table 4). In

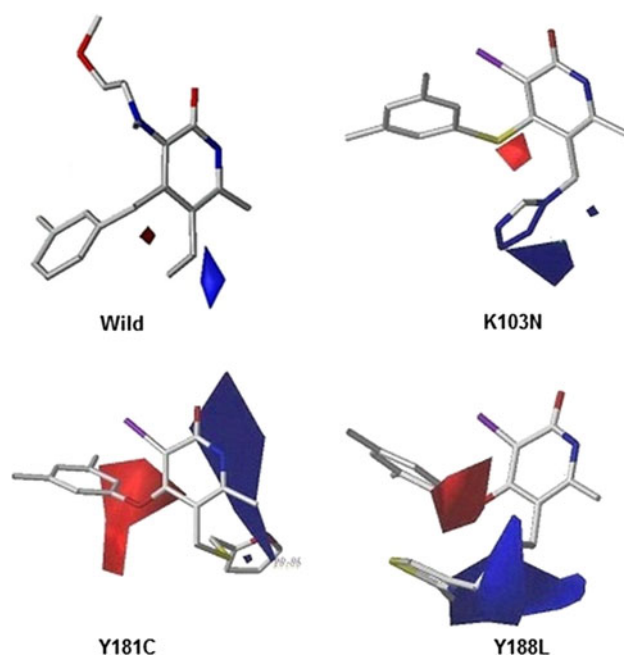


Fig. 6 CoMFA 3D-QSAR electrostatic field contour maps emerged for pyridones as HIV-1 RT inhibitors against wild type and mutant strains K103N, Y181C and Y188L of the virus. Colour-code: *blue* contours indicate regions favourable for electropositive groups and *red* contours indicate regions favourable for electronegative groups

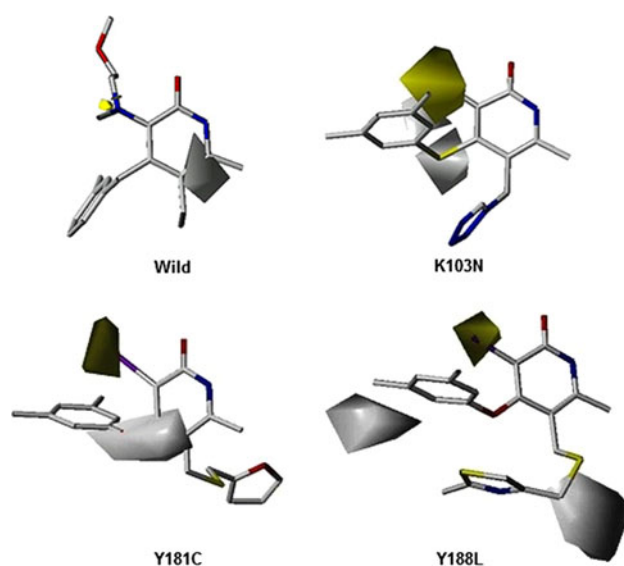


Fig. 7 CoMSIA 3D-QSAR hydrophobic field contour maps emerged for pyridones as HIV-1 RT inhibitors against wild type and mutant strains K103N, Y181C and Y188L of the virus. Colour code: *yellow* contours indicate regions favourable for hydrophobic groups and *gray* contours indicate regions unfavourable for hydrophobic groups

this electrostatic map (Fig. 8) blue contours indicate the electropositive favourable regions and red contours point electronegative favourable regions. The presence of a blue contour near R_2 region (Fig. 8) suggested in favour of

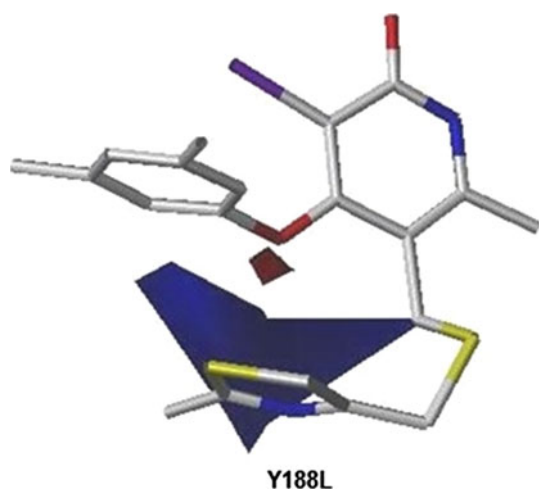


Fig. 8 CoMSIA 3D-QSAR electrostatic field contour maps emerged for pyridones as HIV-1 RT inhibitors against mutant Y188L of the virus. Colour-code: *blue* contours indicate regions favourable for electropositive groups and *red* contours indicate regions favourable for electronegative groups

electropositive groups for the activity. In high active compounds this position is satisfied by thiazole, tetrazole, triazole moieties (Compounds **60**, **126**, **130** etc., Table 1). Part of the discussion made for CoMFA electrostatic holds good here too. Introduction of small alkyl chains in this electropositive region reduced the activity. Also compounds with any other alkyl groups like dimethyl phenyl or any aliphatic chain in this region have shown reduced activity. The linker X region showed a red contour signifying the importance of electronegative environment in this area (Fig. 8). It favoured groups like oxo and carbonyl for improved activity (Compounds, **2**, **60**, **92**, **102** etc., Table 1). Also compounds with CH₂ linker showed less activity (compounds **32**, **46** etc., Table 1). Briefly, the

combined findings of the foregoing CoMFA and CoMSIA study are schematically summarized in Fig. 9.

Protein informatics

Closely following the CoMFA/CoMSIA studies, the protein informatics of HIV-1 RT wild type and mutants were independently attempted with a view to comprehend the changing characteristics of the ligand binding pocket on mutation. Drug resistance or failure may be viewed as an outcome of changed information content of the targeted protein which is programmed to bind the incoming entity/chemical species. In case of mutation(s) in protein, the changed scenario is primarily due to specific alterations in the backbone residues which hold the information necessary for binding the incoming molecule. An examination of the protein–ligand, that is HIV-1 RT–NNRTI co-crystal structures of wild type indicated that the residues 103, 181 and 188, along with several others, as part of the apparatus of NNIBP in chain-A. Comparison of aligned FASTA sequences of chain-A of 2BAN and mutant strains 3MED, 1JKH and 2YNF revealed that besides 103 in 3MED (K103N, lysine-103 by asparagine), 181 in 1JKH (Y181C, tyrosine-181 by cysteine) and 188 in 2YNF (Y188L, tyrosine-188 by leucine), all three strains showed 172 (K172R, lysine-172 by arginine), 280 (S280C, serine-280 by cysteine), 461 (K461R, lysine-461 by arginine), 468 (P468T, proline-468 by threonine), 471 (N471D, asparagine-471 by aspartic acid), 512 (K512Q, lysine-512 by glutamine) and 559 (I559V, isoleucine-559 by valine) as additional mutation points. However, none of these additional residues are part of the NNIBP. Among these, residue-172 is located near the binding pocket (at about 14 Å distance) and the remaining ones (residues 280, 461, 468, 471, 512 and 559) are fairly away from the pocket (at more

Fig. 9 Composite depiction of CoMFA/CoMSIA fields in the vicinity of pyridones for activity against wild type and mutant strains of HIV-1 RT

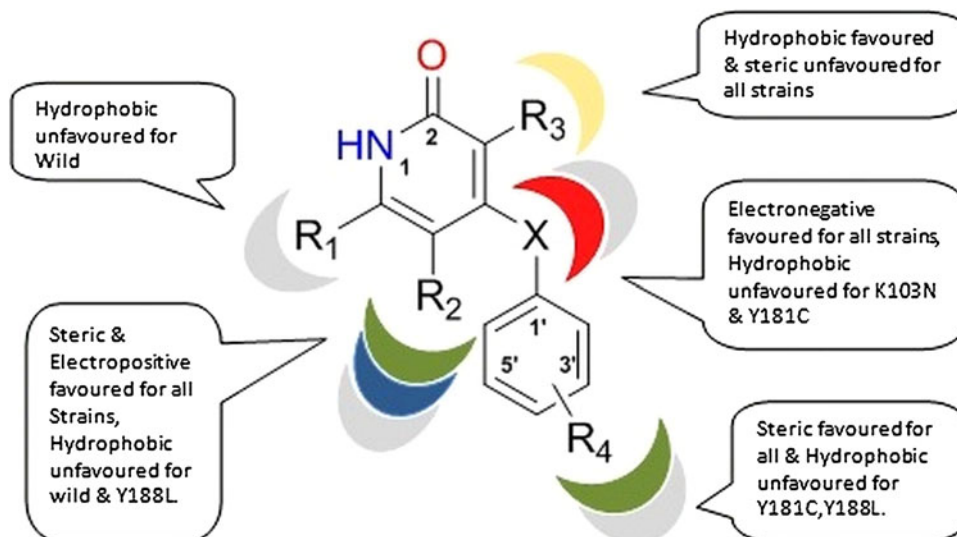


Table 5 The physicochemical properties and single point mutation effect of residues

Residue	State	Name	HI ^a	Polarity (charge) ^b	pKa	pKb	pKx	pI ^c	subSPEC ^d	$\Delta(\Delta G)^e$ (Kcal/mol)	Torsion angle favorability ^f	Overall Stability ^g	D(ΔG) ^h (Kcal/mol)
103	Wild	K(Lys)	−3.9	bp(+ve)	2.18	8.95	10.53	9.74					
	Mut	N(Asn)	−3.5	p(0)	2.02	8.80		5.41	−3.63444	−1.68	No	Yes	0.45
181	Wild	Y(Tyr)	−1.3	p(0)	2.20	9.11	10.07	5.66					
	Mut	C(Cys)	2.5	np(0)	1.96	10.28	8.18	5.07	−4.69518	−1.78	Yes	No	−12.33
188	Wild	Y(Tyr)	−1.3	p(0)	2.20	9.11	10.07	5.66					
	Mut	L(Leu)	3.8	np(0)	2.36	9.60		5.98	−7.53103	−1.24	Yes	No	−1.45
172	Wild	K(Lys)	−3.9	bp(+ve)	2.18	8.95	10.53	9.73					
	Mut	R(Arg)	−4.5	bp(+ve)	2.17	9.04	12.48	10.76	−4.06363	−0.62	No	No	−1.24
280	Wild	S(Ser)	−0.8	p(0)	2.21	9.15		5.68					
	Mut	C(Cys)	2.5	np(0)	1.96	10.28	8.18	5.07	NA	−0.33	No	No	−0.01
461	Wild	K(Lys)	−3.9	bp(+ve)	2.18	8.95	10.53	9.73					
	Mut	R(Arg)	−4.5	bp(+ve)	2.17	9.04	12.48	10.76	NA	0.04	Yes	No	−1.14
468	Wild	P(Pro)	−1.6	np(0)	1.99	10.60		6.3					
	Mut	T(Thr)	−0.7	p(0)	2.09	9.10		5.6	NA	−0.46	No	No	−0.40
471	Wild	N(Asn)	−3.5	p(0)	2.02	8.80		5.41					
	Mut	D(Asp)	−3.5	ap(−ve)	1.88	9.60	3.65	2.77	NA	0.37	Yes	No	−0.03
512	Wild	K(Lys)	−3.9	bp(+ve)	2.18	8.95	10.53	9.73					
	Mut	Q(Gln)	−3.5	p(0)	2.17	9.13		5.65	NA	−1.01	No	No	−5.01
559	Wild	I(Ile)	4.5	np(0)	2.36	9.60		6.02					
	Mut	V(Val)	4.2	np(0)	2.32	9.62		5.96	NA	−0.26 ⁱ			

^a Hydropathy Index; [48]^b Side-chain polarity (Side-chain charge, pH 7.4); *p* polar, *n* neutral, *np* nonpolar, *bp* basic polar, *ap* acidic polar^c pI is isoelectric point at pH 7.0^d Substitution position-specific evolutionary conservation: deleterious functional change score from gene ontology database; subPSEC ≤ -3.0 is said denote deleterious change; NA—score not available for the position; [44]^e Protein stability change upon single point mutations in the pdb. On point mutation if $\Delta(\Delta G)$ becomes less than zero then change leads to decrease stability otherwise increased stability; [45]^f Torsion angles ϕ and ψ distribution; ‘Yes’ for favourable and ‘No’ for unfavourable; [46]^g Protein stability upon point mutations; ‘Yes’ for stabilizing mutation and ‘No’ for destabilizing mutation; [46]^h Difference in free energy of unfolding between wild-type and mutant proteins; [46]ⁱ From protein sequence

than 30 Å distance) (spatial disposition of mutating residues is given in supplementary material). At this juncture it is worth mentioning that, it is also observed that different pdb structures available in the database for the same mutant strain differed to some extent in their non-binding pocket mutant residues. The physicochemical characteristics of all residues involved in the mutation are shown in Table 5. The hydropathy indices [48] and side-chain polarities of binding pocket residues 103, 181 and 188 have pointed out that the mutation of tyrosine (Y) to cysteine (C)/leucine (L) modifies the pocket characteristics more drastically as compared to that of from lysine (K) to asparagine (N) (Table 5).

The adjustment of phi (ϕ)/psi (ψ) angles of the protein backbone, changes in the free-energy of conformations,

free-energy of folding and other properties due to point mutations at specified locations of the wild type HIV-1 RT are predicted using bioinformatics tools namely PANTHER [44] I-Mutant 2.0 [45] and CUPSAT [46] (Table 5). These studies pointed out that mutation of lysine-103 into other residues is relatively more tolerated when compared to that of tyrosine-181 or tyrosine-188 to any other residue (Fig. 10). The point mutations at other residues showed different degrees of tolerance (data given supplementary material). All these results clearly demonstrated that even the single point mutations can lead to unfavorable torsion angles and decreased the overall stability (Table 5). The point mutation results from SIFT [43] also suggested that these changes culminate in proteins with the altered structure/function.

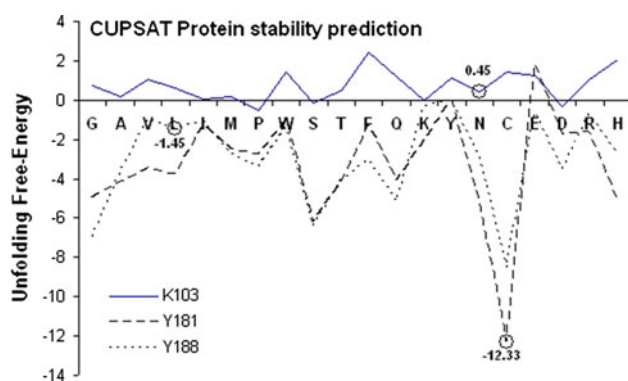


Fig. 10 The changes in the unfolding free-energy of 2BAN, representative of wild type, due to hypothetical point mutation at lysine-103 (K103), tyrosine-181 (Y181) and tyrosine-188 (Y188) with remaining amino acid residues. The unfolding free-energy represents the difference in free-energies of proteins' unfolding from pre- to post-point mutation with specific amino acid. If the change in the unfolding free-energy is more than zero then the new residue may be assumed as tolerated

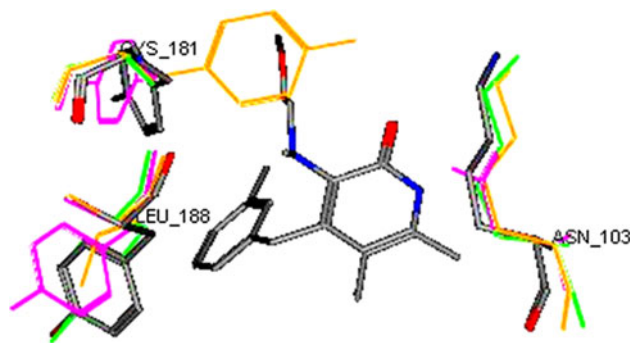


Fig. 11 Superposed binding pocket residues 103, 181 and 188 of wild type, K103N, Y181C, and Y188L strains from 2BAN, 3MED, 1JKH and 2YNF, respectively, along with compound **44**. The wild type residues and compound **44** are shown using default atom color codes namely carbon in gray, oxygen in red and nitrogen in blue. The residues of 3MED, 1JKH and 2YNF are shown in magenta, green and orange, respectively. The mutated residue of each strain is identified with its label. For clarity the hydrogen atoms are suppressed

Table 6 Some physical characteristics of selected HIV-1 RT crystal structures of wild type and mutant proteins

Strain	pdb	Free energy of folding ^a	ASA ^{a,b}	Total volume (packing) ^a	Isoelectric point ^c
Wild	2BAN	−872.50	46,927.2	142,214.5	7.887
K103N	3MED	−914.88	43,977.0	143,199.7	6.221
Y181C	1JKH	−876.07	45,024.3	151,237.8	7.310
Y188L	2YNF	−897.55	43,707.2	143,184.2	6.880

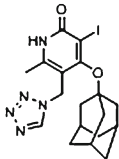
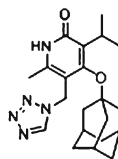
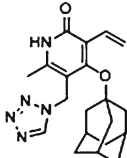
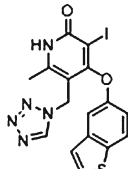
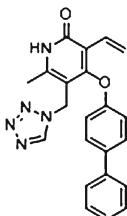
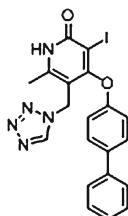
^a Computed from VADAR

^b Accessible surface area

^c Computed from [50]

The protein backbones of mutant crystals with respect to that of wild type have shown RMSD between 2.2 and 3.2 Å (RMSD: 2BAN vs. 3MED, 2.241 Å; 2BAN vs. 1JKH, 3.200 Å; 2BAN vs. 2YNF, 2.259 Å). Emphasizing the binding pocket residues 103, 181 and 188, Fig. 11 shows the deviations between wild type residues (2BAN with compound **44**) and the corresponding ones from the mutants (3MED, 1JKH and 2YNF). From the figure it is evident that the spatial disposition residue 103 is almost same even after mutation in wild type as well as mutant strain. But, in case of residues 181 and 188, they have considerably deviated on mutation from each other in wild type and mutant strains. In VADAR [47] the mutant protein crystal structures, in comparison to wild type, showed increased packing volumes and decreased accessible surface areas and free energy of folding (Table 6). Moreover, the isoelectric points [49, 50] of the mutant proteins are found to be less than that of wild type (Table 6). This shows that the mutants' character is tend to shift in acidic direction. In case of influenza virus also the resistant strains are reported to show smaller isoelectric points in comparison to the native ones [51]. Furthermore, in HIV-1 RT—NNRTI X-ray co-crystal structures and in docking experiments it is observed that the pyridone moiety preferably occupied locations near the residue-103 and the aryl moiety positioned near the residues-181 and -188. Since the mutation of lysine-103 to asparagine did not drastically alter the characteristics of the binding pocket, the compounds almost retained the activity against K103N strain. Nevertheless, in case of mutation of tyrosine-181 to cysteine and tyrosine-188 to leucine, the pocket region corresponding to these residues has changed characteristics considerably. This may have adversely affected the binding pocket environment of aryl moiety of pyridones resulting in considerable loss of activity against Y181C and Y188L strains. These results made clear that mutation of residues-181 and -188 are more deleterious to the protein character and function as compared to that of residue-103. The activity profile of pyridone derivatives (Table 2) against wild type, and three mutant strains is in clearly agreement with these observation. In the backdrop of changing geometric and physicochemical characteristics of the binding pocket, flexibility in the molecular frame may be important to tackle the drug resistance issue. It may be mentioned that rilpivirine, a flexible diarylpyrimidine analogue, is introduced to overcome the limitations of nevirapine, a structurally rigid molecule, in tackling the drug resistance viral strains [52, 53]. The high active compounds (**44**, **126**, **61** and **60**) taking different conformations in the binding domains of respective HIV-1 RT proteins also support this view. Additionally for pyridone derivatives it may be suggested that, alterations in the aryl moiety matching with the changed scenario of residue-181/-188 may make the

Fig. 12 Virtual compounds, their CoMFA/CoMSIA predicted activities and binding energies (B.E.—kcal/mol) profiles

									
VC1					VC2				
Wild	K103N	Y181C	Y188L		Wild	K103N	Y181C	Y188L	
9.33	9.47	8.42	8.18	pIC₅₀CoMFA	9.11	9.21	8.21	8.08	
9.38	9.39	8.51	8.24	pIC₅₀CoMSIA	9.00	8.98	8.28	8.14	
11.11	11.59	10.89	10.61	B.E.	10.01	10.40	10.05	10.00	
									
VC3					VC4				
Wild	K103N	Y181C	Y188L		Wild	K103N	Y181C	Y188L	
9.31	9.41	8.33	8.13	pIC₅₀CoMFA	9.23	9.21	8.31	8.12	
9.48	9.50	8.38	8.22	pIC₅₀CoMSIA	9.10	9.32	8.24	8.13	
11.28	11.75	10.87	10.44	B.E.	10.68	11.11	10.54	10.07	
									
VC5					VC6				
Wild	K103N	Y181C	Y188L		Wild	K103N	Y181C	Y188L	
9.11	9.17	8.37	8.17	pIC₅₀CoMFA	9.23	9.24	8.33	8.09	
9.19	9.38	8.49	8.33	pIC₅₀CoMSIA	9.31	9.28	8.47	8.19	
9.88	9.97	9.57	10.66	B.E.	10.08	10.67	10.56	9.73	
Training compounds with optimum observed activities against mutant strains									
Compound No		Ref. Strain	Wild	K103N	Y181C	Y188L			
126		K103N	8.30	9.51	8.10	7.80			
61		Y181C	7.60	9.00	8.30	7.50			
60		Y188L	8.22	9.00	8.20	8.10			

compounds compatible to multiple mutant strains leading to restoration of the activity.

The foregoing results from CoMFA/CoMSIA and protein informatics suggested that the aryl and R₂/R₃ regions of the compounds (Fig. 1) offer clear handle to come up with potential compounds against multiple viral strains.

These results are used in a concise exploration for the design of new pyridone derivatives against HIV-1 (Fig. 12). These molecules are due to modifications in the aryl and R₃ regions of pyridones (Fig. 1). The predictions of these compounds in CoMFA and CoMSIA models indicated that they may show the following activity (pIC₅₀)

profiles: wild type, 9.4–9.0; K103N, 9.4–9.1; Y181C, 8.4–8.2; and Y188L, 8.2–8.1 (Fig. 12). Furthermore, each molecule's solo in silico predicted activity profile against wild type and all three mutant strains is better than that of any one compound in the training/test set. Among the new molecules, **VC1** showed the best in silico activity against all (Wild, ~ 9.36 ; K103N, ~ 9.43 ; Y181C, ~ 8.46 ; and Y188L, ~ 8.21). In docking experiments with wild type and mutant pdbs, these compounds (Fig. 12) have shown docking scores comparable to the high active compounds of the current study. Thus these results provided clues for further exploration of pyridone derivatives as inhibitors of multiple HIV-1 strains.

Conclusions

The CoMFA/CoMSIA models have reasonably well explained the HIV-1 RT inhibitory activities of pyridone analogues and there are no serious outliers in the datasets. However, each strain has shown a few analogues as deviation points in the analysis. These deviations are mainly due to subtle changes in the spatial orientation of one or more groups positioned around the pyridone scaffold which in turn reflected in the respective molecular fields. Exclusion of 1–2 such compounds from each dataset would have improved the overall statistics of the models. For example, in case of K103N, compound 108 has been predicted high ($pIC_{50}(\text{Obs}) = 6.80$; $pIC_{50}(\text{CoMFA}) = 7.59$; $pIC_{50}(\text{CoMSIA}) = 8.06$). Exclusion of this compound has improved the overall statistics of the CoMFA and CoMSIA models (before exclusion: CoMFA, $r^2 = 0.930$, $q^2 = 0.735$; CoMSIA, $r^2 = 0.880$; $q^2 = 0.727$; after exclusion: CoMFA, $r^2 = 0.936$, $q^2 = 0.743$; CoMSIA, $r^2 = 0.897$; $q^2 = 0.758$). Here, the high flexibility of compound 108 has allowed it take a pose near to that of the most active analogue (compound 126). As a result this compound has been predicted to some extent higher even though its observed activity is low. This kind of misorientations does not come in the gamut of outliers. However, to maintain the integrity of the dataset, no compound with clear-cut activity was excluded from the study.

This study broadly indicated that while steric, hydrophobic and electropositive character of groups (R_2 and R_3) surrounding the pyridone moiety are important for broad-spectrum activity, the steric field around the aryl moiety (R_4) is important for the same. The R_2 region is favorable for steric as well as electropositive groups to result in broad-spectrum activity. The field emerged for the R_3 region indicated the importance of hydrophobicity and minimum steric influence for the activity. The R_4 representing the aryl moiety of the pyridones strongly favored steric nature for activity against all strains. The results

uniformly suggested that in these compounds an electro-negative linker (X) is preferred for broad-spectrum activity. The models did not offer much information concerning the R_1 region.

The Protein informatics has indicated the deleterious nature of mutations, especially occurring at residues 181 and 188. Flexibility in the molecular frame can come in a big way to combat the drug resistance. Compounds designed taking the binding pocket alterations into consideration, especially compensating the changes due to mutation of residues 181 and 188, help in increasing the activity against multiple strains. The in silico experiments with the newly designed compounds offered support to these findings. Thus, it is expected that the study will open up fresh avenues for hitherto unexplored molecules emerged from the investigation.

Acknowledgments Authors thank Dr. Shreekanth Deshpande and V. Murugesan for helpful discussion and Ms. Neelam Mishra for technical help. Two of the authors, UD and SV, thank CSIR and DST, New Delhi, respectively, for senior research fellowships. CDRI Communication No.8492

References

1. Global report: UNAIDS report on the global AIDS epidemic (2012) http://www.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2012/gr2012/20121120_UNAIDS_Global_Report_2012_en.pdf. Accessed on 9 April 2013
2. Imami N, Gotch F (2003) Twenty years of HIV-1 research: what the future holds. *Nat Immunol* 4:501
3. [http://www.ncbi.nlm.nih.gov/genome/?term=txid11676\[Organism:exp\]](http://www.ncbi.nlm.nih.gov/genome/?term=txid11676[Organism:exp])
4. Clercq ED (1993) HIV-1 specific RT inhibitors: highly selective inhibitors of human immunodeficiency virus type 1 that are specifically targeted at the viral reverse transcriptase. *Med Res Rev* 13:229–258
5. Spence RA, Kati WM, Anderson KS, Johnson KA (1995) Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors. *Science* 267:988–993
6. Esnouf R, Ren J, Ross C, Jones Y, Stammers D, Stuart D (2006) Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors. *Nat Struct Biol* 2:303–308
7. Figueiredo A, Moore Katie L, Mak J, Sluis-Cremer N, de Bethune MP, Tachedjian G (2006) Potent nonnucleoside reverse transcriptase inhibitors target HIV-1 Gag-Pol. *PLoS Pathog* 2:1051–1059
8. De Béthune MP (2010) Non-nucleoside reverse transcriptase inhibitors (NNRTIs), their discovery, development, and use in the treatment of HIV-1 infection: a review of the last 20 years (1989–2009). *Antivir Res* 85:75–90
9. Ren J, Stammers DK (2008) Structural basis for drug resistance mechanisms for non-nucleoside inhibitors of HIV reverse transcriptase. *Virus Res* 134:157–170
10. Petropoulos CJ, Parkin NT, Limoli KL, Lie YS, Wrin T, Huang W, Tian H, Smith D, Winslow GA, Capon DJ, Whitcomb JM (2000) A novel phenotypic drug susceptibility assay for human immunodeficiency virus type 1. *Antimicrob Agents Chemother* 44:920–928

11. Young SD, Britcher SF, Tran LO, Payne LS, Lumma WC, Lyle TA, Huff JR, Anderson PS, Olsen DB, Carroll SS (1995) L-743, 726 (DMP-266): a novel, highly potent nonnucleoside inhibitor of the human immunodeficiency virus type 1 reverse transcriptase. *Antimicrob Agents Chemother* 39:2602–2605
12. Byrnes VW, Sardana VV, Schleif WA, Condra JH, Waterbury JA, Wolfgang JA, Long WJ, Schneider CL, Schlabach AJ, Wolanski BS, Graham DJ, Gotlib L, Rhodes A, Titus DL, Roth E, Blahy OM, Quintero JC, Satszewski S, Emini EA (1993) Comprehensive mutant enzyme and viral variant assessment of human immunodeficiency virus type 1 reverse transcriptase resistance to nonnucleoside inhibitors. *Antimicrob Agents Chemother* 37:1576–1579
13. Benjahad A, Croisy M, Monneret C, Bisagni E, Mabire D, Coupa S, Poncelet A, Csoka I, Guillemont J, Meyer C, Andries K, Pauwels R, de Bthune MP, Himmel DM, Das K, Arnold E, Nguyen CH, Grierson David S (2005) 4-Benzyl and 4-Benzoyl-3-dimethylaminopyridin-2(1H)-ones: in vitro evaluation of new C-3-amino-substituted and C-5, 6-alkyl-substituted analogues against clinically important hiv mutant strains. *J Med Chem* 48:1948–1964
14. Guillemont J, Benjahad A, Oumouch S, Decrane L, Palandjian P, Vernier D, Queguiner L, Andries K, de Bthune MP, Hertogs K, Grierson David S, Nguyen CH (2009) Synthesis and biological evaluation of C-5 methyl substituted 4-arylthio and 4-aryloxy-3-iodopyridin-2(1H)-one type anti-HIV agents. *J Med Chem* 52:7473–7487
15. Benjahad A, Oumouch S, Guillemont J, Pasquir E, Mabire D, Andries K, Nguyen CH, Grierson DS (2006) Structure-activity relationship in the 3-iodo-4-phenoxy-pyridinone (IOPY) series: the nature of the C-3 substituent on anti-HIV activity. *Bioorg Med Chem* 17:712–716
16. Garg R, Gupta SP, Gao H, Babu MS, Debnath AK, Hansch C (1999) Comparative quantitative structure-activity relationship studies on anti-HIV drugs. *Chem Rev* 99:3525–3601
17. Pungpo P, Saparpakorn P, Wolschann P, Hannongbua S (2006) Computer-aided molecular design of highly potent HIV-1 RT inhibitors: 3D QSAR and molecular docking studies of efavirenz derivatives. *SAR QSAR Environ Res* 17:353–370
18. Carlsson J, Boukharta L, Aqvist J (2008) Combining docking, molecular dynamics and the linear interaction energy method to predict binding modes and affinities for non-nucleoside inhibitors to HIV-1 reverse transcriptase. *J Med Chem* 51:2648–2656
19. Prabhakar YS, Solomon VR, Rawal RK, Gupta MK, Katti SB (2004) CP-MLR/PLS directed structure-activity modeling of the HIV-1RT inhibitory activity of 2, 3-diaryl-1, 3-thiazolidin-4-ones. *QSAR Comb Sci* 23:234–244
20. Rawal RK, Prabhakar YS, Katti SB, Clercq ED (2005) 2-(Aryl)-3-furan-2-ylmethylthiazolidin-4-ones as selective HIV-RT inhibitors. *Bioorg Med Chem* 13:6771–6776
21. Rawal RK, Prabhakar YS, Katti SB (2007) Molecular surface features in modeling the HIV-1 RT inhibitory activity of 2-(2, 6-disubstituted phenyl)-3-(substituted pyrimidin-2-yl)-thiazolidin-4-ones. *QSAR Comb Sci* 26:398–406
22. Rawal RK, Tripathi R, Katti SB, Pannecouque C, Clercq ED (2008) Design and synthesis of 2-(2, 6-dibromo-phenyl)-3-heteroaryl-1, 3-thiazolidin-4-ones as anti-HIV agents. *Eur J Med Chem* 43:2800–2806
23. Murugesan V, Prabhakar YS, Katti SB (2009) CoMFA and CoMSIA studies on thiazolidin-4-one as anti-HIV-1 agents. *J Mol Graph Model* 27:735–743
24. Murugesan V, Tiwari VS, Saxena R, Tripathi R, Paranjape R, Kulkarni S, Makwana N, Suryawanshi R, Katti SB (2011) Lead optimization at C-2 and N-3 positions of thiazolidin-4-ones as HIV-1 non-nucleoside reverse transcriptase inhibitors. *Bioorg Med Chem* 19:6919–6926
25. Rawal RK, Murugesan V, Katti SB (2012) Structure-activity relationship studies on clinically relevant HIV-1 NNRTIs. *Curr Med Chem* 19:5364–5380
26. Cramer RD III, Patterson DE, Bunce JD (1988) Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J Am Chem Soc* 110:5959–5967
27. Klebe G, Abraham U, Mietzner T (1994) Molecular similarity indices in a comparative analysis (COMSIA) of drug molecules to correlate and predict their biological activity. *J Med Chem* 37:4130–4146
28. Himmel DM, Das K, Clark AD Jr, Hughes SH, Benjahad A, Oumouch S, Guillemont J, Coupa S, Poncelet A, Csoka I, Meyer C, Andries K, Nguyen CH, Grierson DS, Arnold E (2005) Crystal structures for HIV-1 reverse transcriptase in complexes with three pyridinone derivatives: a new class of non-nucleoside inhibitors effective against a broad range of drug-resistant strains. *J Med Chem* 48:7582–7591
29. Lansdon EB, Brendza KM, Hung M, Wang R, Mukund S, Jin D, Birkus G, Kutty N, Liu X (2010) Crystal structures of HIV-1 reverse transcriptase with etravirine (TMC125) and rilpivirine (TMC278): implications for drug design. *J Med Chem* 53:4295–4299
30. Ren J, Nichols C, Bird L, Chamberlain P, Weaver K, Short S, Stuart DI, Stammers DK (2001) Structural mechanisms of drug resistance for mutations at codons 181 and 188 in HIV-1 reverse transcriptase and the improved resilience of second generation non-nucleoside inhibitors. *J Mol Biol* 312:795–805
31. Chong P, Sebahar P, Youngman M, Garrido D, Zhang H, Stewart EL, Nolte RT, Wang L, Ferris RG, Edelstein M, Weaver K, Mathis A, Peat A (2012) Rational design of potent non-nucleoside inhibitors of HIV-1 reverse transcriptase. *J Med Chem* 55:10601
32. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* 30:2785–2789
33. Clark M, Cramer RD III, Van Opdenbosch N (1989) Opdenbosch. Validation of the general-purpose tripos 5.2 force field. *J Comput Chem* 10:982–1012
34. Barakat MT, Dean PM (1990) Molecular structure matching by simulated annealing. I. A comparison between different cooling schedules. *J Comput Aided Mol Des* 4:295–316
35. SYBYL, version 7.3 (2006) Tripos Associates, St. Louis, MO
36. Klebe G, Abraham U, Mietzner T (1994) Molecular similarity indices in a comparative analysis (COMSIA) of drug molecules to correlate and predict their biological activity. *J Med Chem* 37:4130–4146
37. Viswanadhan VN, Ghose AK, Revenkar GR, Robins RK (1989) Atomic physicochemical parameters for three dimensional structure directed quantitative structure-activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring antibiotics. *J Chem Inf Comput Sci* 29:163–172
38. Klebe G (1994) The use of composite crystal-field environments in molecular recognition and the de novo design of protein ligands. *J Mol Biol* 237:212–235
39. Stahle J, Wold S (1988) 6 multivariate data analysis and experimental design in biomedical research. *Prog Med Chem* 25:291–338
40. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
41. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98

42. MOE: The Molecular Operating Environment from Chemical Computing Group Inc., 1255 University St., Suite 1600, Montreal, Quebec, Canada H3B 3X3. <http://www.chemcomp.com>
43. Ng PC, Henikoff S (2003) SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res* 31:3812–3814
44. Mi H, Guo N, Kejariwal A, Thomas PD (2007) PANTHER version 6: protein sequence and function evolution data with expanded representation of biological pathways. *Nucleic Acids Res* 35(Database issue):D247–D252
45. Capriotti E, Fariselli P, Casadio R (2005) I-mutant 2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res* 33(Web Server issue):306–310
46. Parthiban V, Gromiha MM, Schomburg D (2006) CUPSAT: prediction of protein stability upon point mutations. *Nucleic Acids Res* 34:239–242
47. Willard L, Ranjan A, Zhang H, Monzavi H, Boyko RF, Sykes BD, Wishart DS (2003) VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res* 31:3316–3319
48. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157:105–132
49. Sillero A, Maldonado A (2006) Isoelectric point determination of proteins and other macromolecules: oscillating method. *Comput Biol Med* 36:157–166
50. Calculation of protein isoelectric point. <http://isoelectric.ovh.org/>
51. Krenn BM, Egorov A, Romanovskaya-Romanko E, Wolschek M, Nakowitsch S, Ruthsatz T, Kiefmann B, Morokutti A, Humer J, Geiler J, Cinatl J, Michaelis M, Wressnigg N, Sturlan S, Ferko B, Batishchev OV, Indenbom AV, Zhu R, Kastner M, Hinterdorfer P, Kiselev O, Muster T, Romanova J (2011) Single HA2 mutation increases the infectivity and immunogenicity of a live attenuated H5N1 intranasal influenza vaccine candidate lacking NS1. *PLoS ONE* 6:e18577
52. Ren J, Milton J, Weaver KL, Short SA, Stuart DI, Stammers DK (2000) Structural basis for the resilience of Efavirenz (DMP-266) to drug resistance mutations in HIV-1 reverse transcriptase. *Structure* 8:1089–1094
53. Das K, Bauman JD, Clark AD Jr, Frenkel YV, Lewi PJ, Shatkin AJ, Hughes SH, Arnold E (2008) High-resolution structures of HIV-1 reverse transcriptase/TMC278 complexes: strategic flexibility explains potency against resistance mutations. *PNAS* 105:1466–1471