

# HIV reverse transcriptase: Structural interpretation of drug resistant genetic variants from India

Abraham Joseph Kandathil<sup>1,5</sup>, Agnel Praveen Joseph<sup>2,5</sup>, Rajesh Kannangai<sup>1</sup>, Narayanaswamy Srinivasan<sup>2\*</sup>, Oriapadickal Cherian Abraham<sup>3</sup>, Susanne Alexander Pulimood<sup>4</sup>, Gopalan Sridharan<sup>4</sup>

<sup>1</sup>Departments of Clinical Virology, Christian Medical College, Vellore, India; <sup>2</sup>Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India; <sup>3</sup>Internal Medicine and <sup>4</sup>Dermatology, Christian Medical College, Vellore, India; <sup>5</sup>Joint first authors; Narayanaswamy Srinivasan - Email: ns@mbu.iisc.ernet.in; Phone: +91-416-228 2070; Fax: +91-416-2232035; \*Corresponding author

Received June 05, 2009; Revised June 18, 2009; Accepted July 15, 2009; Published August 20, 2009

## Abstract:

The reverse transcriptase (RT) enzyme is the prime target of nucleoside/ nucleotide (NRTI) and non-nucleoside (NNRTI) reverse transcriptase inhibitors. Here we investigate the structural basis of effects of drug-resistance mutations in clade C RT using three-dimensional structural modeling. Apropos the expectation was for unique mechanisms in clade C based on interactions with amino acids of p66 subunit in RT molecule. 3-D structures of RT with mutations found in sequences from 2 treatment naïve, 8 failed and one reference clade C have been modeled and analyzed. Models were generated by computational mutation of available crystal structures of drug bound homologous RT. Energy minimization of the models and the structural analyses were carried out using standard methods. Mutations at positions 75,101,118,190,230,238 and 318 known to confer drug resistance were investigated. Different mutations produced different effects such as alteration of geometry of the drug-binding pocket, structural changes at the site of entry of the drug (into the active site), repositioning the template bases or by discriminating the inhibitors from their natural substrates. For the mutations analyzed, NRTI resistance was mediated mainly by the ability to discriminate between inhibitors and natural substrate, whereas, NNRTI resistance affected either the drug entry or the geometry of the active site. Our analysis suggests that different mutations result in different structural effects affecting the ability of a given drug to bind to the RT. Our studies will help in the development of newer drugs taking into account the presence of these mutations and the structural basis of drug resistance.

**Key words:** Drug resistance, Genetic mutations, HIV-1, Reverse transcriptase, Structural modeling

## Background:

Human immunodeficiency virus type -1 (HIV-1) belongs to the family *Retroviridae*. Like all the other members of this family it requires the enzyme reverse transcriptase (RT) to reverse transcribe the genomic RNA into DNA. The HIV-1 RT is the prime target of nucleoside reverse transcriptase inhibitors (NRTIs) and non-nucleoside reverse transcriptase inhibitors (NNRTIs) which are used in antiretroviral therapy (ART). In India, a combination from NRTIs and NNRTIs are used as first line drugs to treat HIV-1 infections.

The RT enzyme has two polymerase activities, namely, RNA-dependent DNA polymerization and DNA-dependent DNA polymerization [1]. The RT is a heterodimer consisting of two subunits p66 and p51 [1, 2]. The p66 subunit consists of 560 amino acids whereas the p51 subunit consists of 440 amino acids. The additional 120 amino acids in p66 are responsible for the ribonuclease H (RNase H) activity [2]. The p66 subunit is responsible for the enzymatic activity of the RT and contains the DNA binding groove and the active site. The p66 has an 'open' conformation to enable it to 'grasp' the template. The overall shape of the p66 subunit has been compared to that of a right hand [2]. It is composed of the "fingers" (residues 1-85, 118-155), "palm" (86-117, 156-237) and "thumb" (238-318). The "palm" subdomain is the location of the catalytic aspartate residues (D110, D185, D186) [2]. In addition, it also has the 'connection' (319-426) and the RNase H. The p51 subunit functions more as a scaffold for p66 and has a closed conformation. It has no enzymatic activity [1-3].

The RT of HIV-1 can develop mutations that can hamper interactions to NRTIs and NNRTIs which translates to resistance to RT inhibitors.

The mechanisms of resistance of many of these mutations have been studied and crystal structures of these mutant RTs are available in protein data bank (PDB). In this study, we report our studies to elucidate structural basis of plausible mechanisms of resistance conferred by certain mutations in clade C HIV-1 strains observed in Indian patients. We have generated three-dimensional structural models to postulate mechanisms of resistance induced by these mutations.

## Methodology:

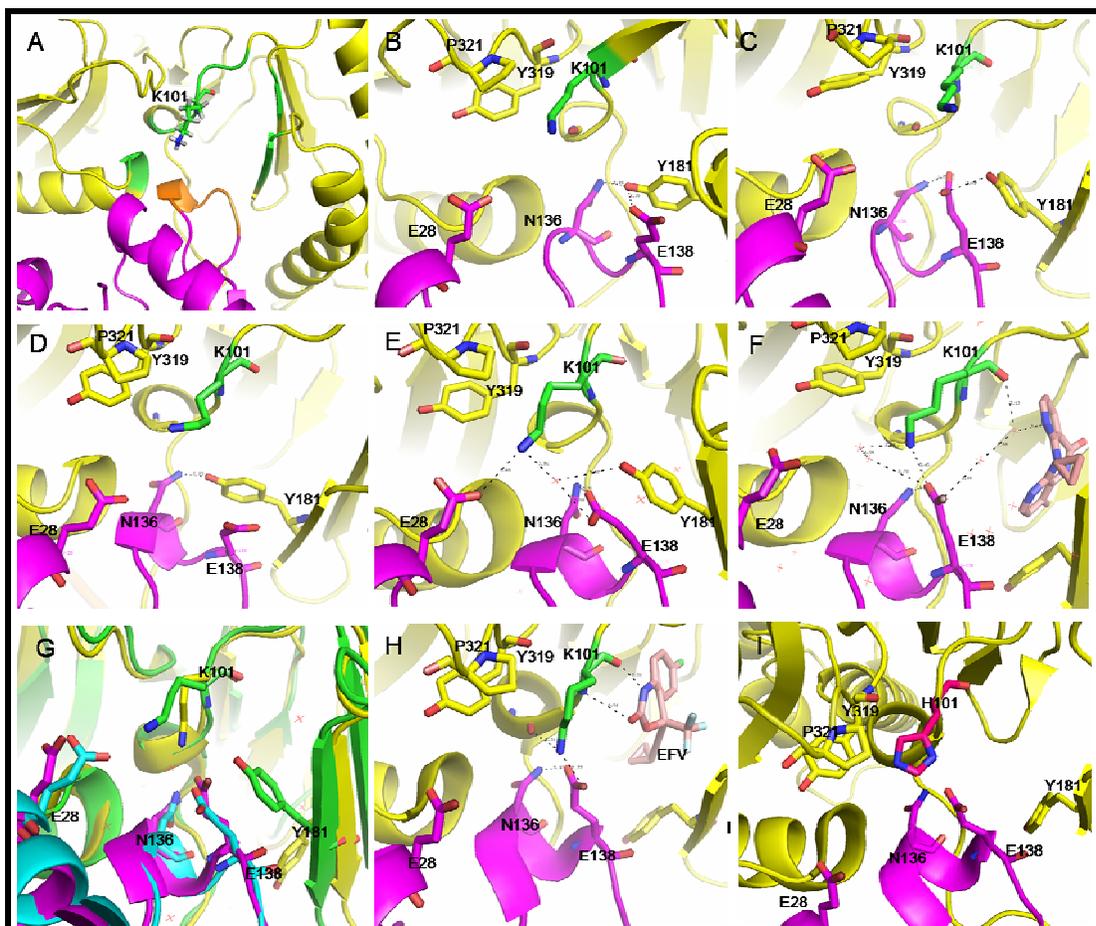
Blood samples were collected from HIV-1 infected treatment naïve individuals and treatment experienced patients who were failing ART. They had come to the Clinical Virology department of a tertiary care hospital in south India for CD4+ T- cell estimation and/or HIV-1 viral load during the years 2005-2007. The *pol* gene of these strains was amplified by an in-house PCR, this sequence data was used only for the development of the models [4, 5]. This derived work was part of another study wherein we looked at HIV-1 drug resistance. It involved amplification of the HIV-1 protease and RT gene to obtain an 1800 bp amplicon.

RT sequences from 10 strains (8 from treatment failure and 2 from treatment naïve groups) (Table 1 in supplementary material) and one reference clade C (AAY23520.1) was selected to generate three dimensional models. The treatment naïve group was selected based on the presence of unusual mutation/mutations not seen in clade B strains based on results obtained from the Stanford HIV-1 drug resistance database. The 3D modeling was done for these variants, some known to confer drug resistance in Clade B and some are newly found in clade C. The data on mutational patterns were based on information

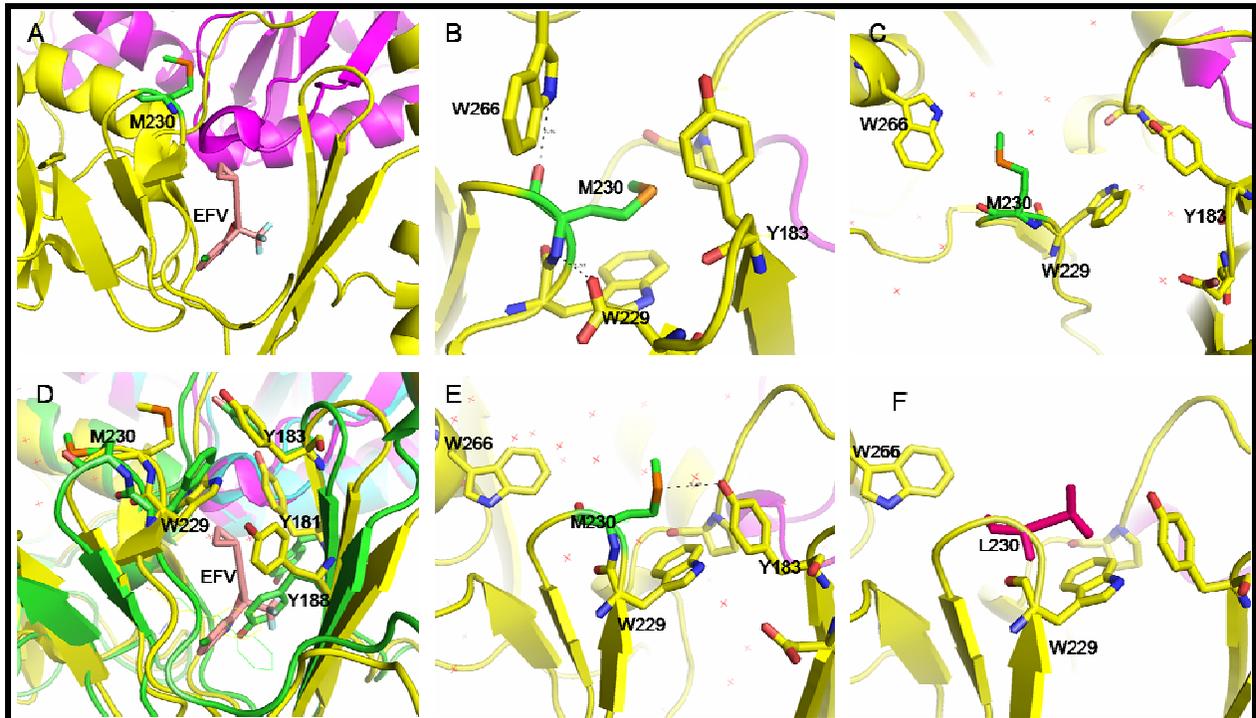
got from the Stanford HIV drug resistance database. The nucleotide sequences were translated using the ExPasy translate tool (<http://www.expasy.ch/tools/dna.html>). The translated sequences were aligned using clustalw (<http://www.ebi.ac.uk/clustalw/>).

Three dimensional models were generated by mutating the available crystal structures of drug bound and unbound RT using the Biopolymer tool of Sybyl (Tripos Inc., St. Louis, MO). The protein databank codes of the crystal structures of bound and unbound forms of RT used in this study are 1IKW [6], 1VRT [7], 1DLO [8], 1HMV [9], 1RTJ [10] and 1RTD [11]. Energy minimizations were carried out by targeting the sites of mutations and neighborhood using the Maximin tool of Sybyl package. All the atoms that were within a

distance of 6Å (Angstrom) from the mutated residue (“HOT” region), were allowed to move during minimization. Kollman United force field [12] available in Sybyl package was used for energy minimization. A cutoff of 9Å was used for calculating non-bonded interactions. The inhibitor molecule was not included during energy minimization since the sites of most of the mutations are not located on the active site. We have used caricature modeling to postulate changes in the RT that could affect drug affinity. Mustang [13] and DALI [14, 15] were used for multiple and pair-wise structure superposition respectively. PyMOL (DeLano Scientific LLC, San Carlos, California) and Setor [16] were used to visualize modeled protein structures and to analyze interactions between amino acid residues.



**Figure 1:** (A) Location of Lys 101 in the unliganded RT structure, 1RTJ. p66 subunit is coloured in yellow while the p55 subunit is in pink (Same subunit colouring code is followed for the other figures). The part of the NNRTI entry site in the p66 subunit and the side chain of Lys 101 are coloured green. The part of drug entry site in the p55 subunit is coloured in orange. (B-F) Local hydrogen bond interactions involving K101 in unliganded RT structure 1DLO, 1HMV, 1RTD, 1RTJ and in Nevirapine bound RT 1VRT (G) Structural superposition of 1IKW (yellow) and 1RTJ (green) showing changes in residue conformations. (H) Interaction pattern involving K101 in Efavirenz (shown in orange) bound RT (1IKW) (I) Changes due to K101H mutation. Site of mutation (here and in the rest of the figures) is shown in pink. All the figures are rendered in PyMol.



**Figure 2:** (A) Location of M230 (green) in Efavirenz (in orange) bound RT (1IKW) (B) Interactions involving M230 in unliganded RT structures (1DLO and (C) 1RTJ. (D) RT structure that corresponds to open conformation (1RTJ, green) superposed on Efavirenz (in orange) bound RT (1IKW, yellow) (E) Interactions involving M230 in Nevirapine bound RT (1VRT) (F) Model for M230L mutation.

**Discussion:**

In this study we have studied eight mutations found in the Clade-C strains that confer resistance to both the classes of reverse transcriptase inhibitors viz., NRTIs and NNRTIs. Mutations that confer resistance to these two groups of drugs are distinct. We have mainly studied hydrogen bond interactions and van der Waals contacts involving the site of mutation. The changes in the local interaction patterns were then interpreted with the help of models for these mutations. The local interaction patterns found in the drug bound RT and absent in the unliganded protein, is expected to have some significance in stabilizing the drug bound form here we try to identify such interactions involving the site of mutation. Mutations like K101H, M230L and K238T confer resistance to NNRTIs by affecting drug entry into the binding pocket and by altering the geometry of the pocket. Other mutations like G190A and Y318F directly affect drug binding.

Mutations such as V75M, V118I and Q151M that confer resistance to NRTIs act by affecting certain interactions that help dNTP (deoxy Nucleotide Tri Phosphate) or the drug binding. This not only involves interactions between the drug and RT but also interactions between the drug and the template RNA.

**K101H**

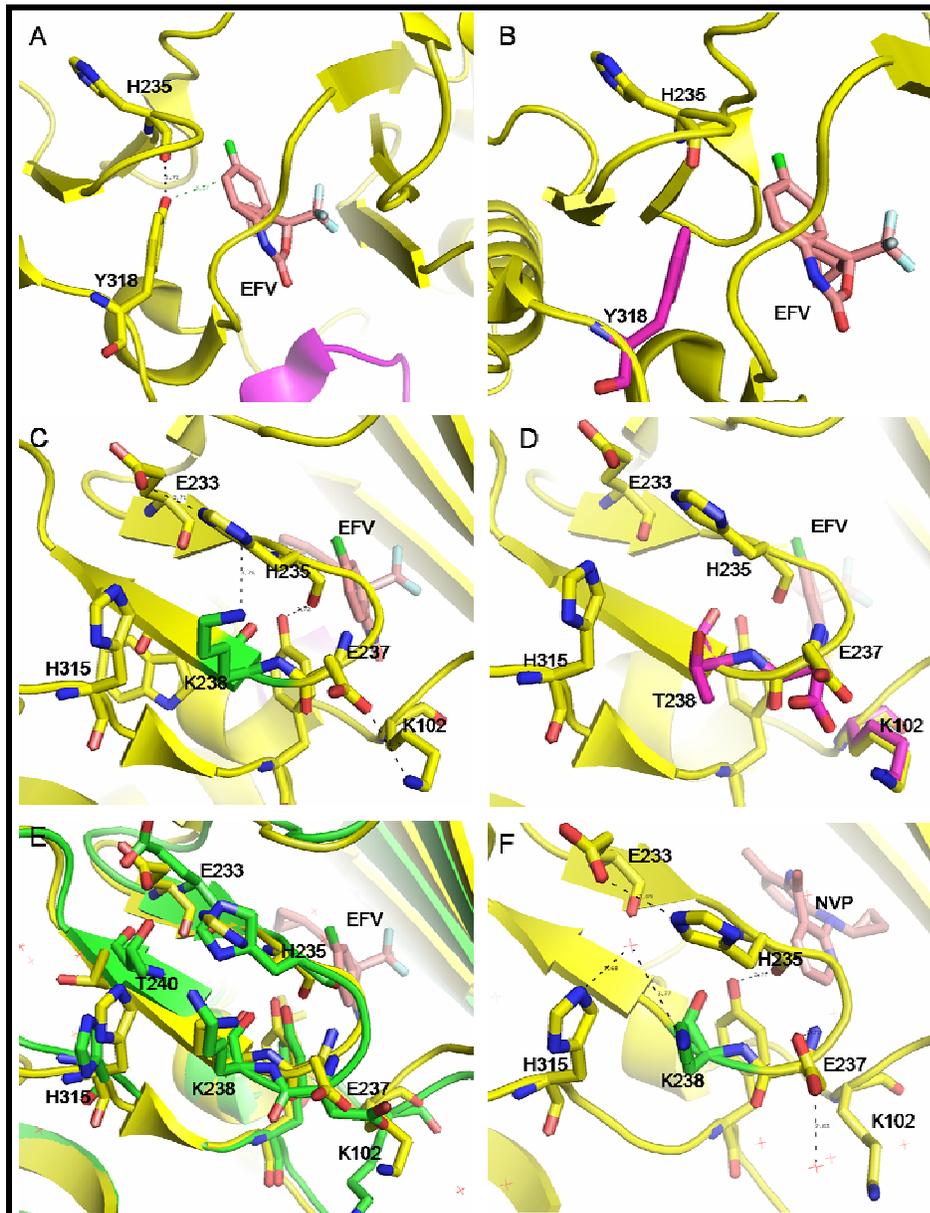
Figure 1A shows the location of Lys 101 in the wild type unliganded RT crystal structure (1RTJ). In the ligand unbound RT structure, the side chain of Tyr 181 points into the hydrophobic pocket where the

NNRTIs bind (Figure 1B-E). This conformation is stabilized by the hydrogen bonds formed by the side chain hydroxyl group of Tyr 181 with the side chain N-H group of Asn 136 of the p55 subunit (1DLO, 1RTD) and with side chain carboxyl of Glu 138 of the p55 subunit (1HMV) (Figure 1B-D). The side chain of Lys 101 interacts with Pro 321 through van der Waals forces. In some structures (1DLO, 1RTD), it also interacts with the side chain carboxyl of Glu 28 through a weak salt bridge (Figure 1B, D, E). In the template/primer bound RT structure (1RTD), Lys 101 side chain makes van der Waals contacts with the side chain of Tyr 319 (Figure 1D). In the structure of RT (1RTJ), which largely corresponds to open conformation (Figure S1 in supplementary material), Lys 101 forms a salt bridge with Glu 28. It interacts with Glu 138 through water mediated hydrogen bonds (Figure 1E).

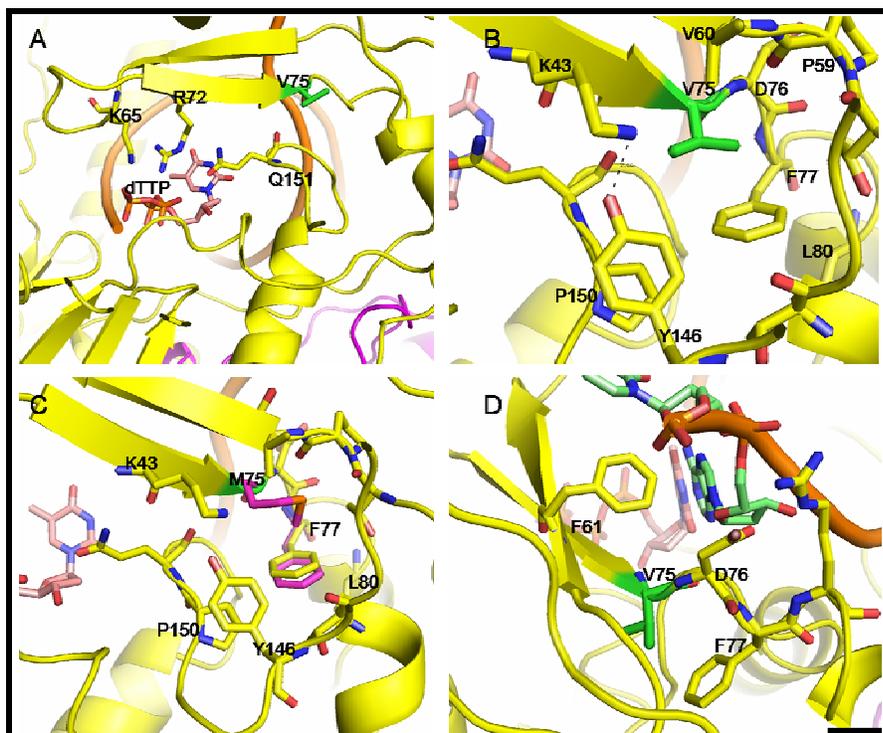
In efavirenz and nevirapine bound RT structures (1IKW, 1VRT), Lys 101 forms salt bridge with Glu 138. The side chain carboxyl of Glu 138 also interacts with Gln136 through a NH...O hydrogen bond (Figure 1F, H). Efavirenz directly interacts with the main chain carbonyl and amino groups of Lys 101 through hydrogen bonds. Nevirapine forms water mediated hydrogen bonds with the main chain amino group and carboxyl groups of Lys 101 (Figure 1F, H). The conformational changes that occur in NNRTI binding pocket upon ligand binding are shown in Figure 1G. As a result of the mutation Lys 101 → His (K101H), the salt bridge that Lys 101 forms with Glu 138 is lost (Figure 1I) due to the shorter sidechain of Histidine. This might alter the geometry of the binding pocket. The loss of this salt

bridge might also affect the hydrogen bond interaction between Glu 138 and Gln 136 and their side chains are likely to become more flexible. This increases the possibilities of interaction between Glu 138/Gln 136 and Tyr 181, thereby affecting the inhibitor entry into the

pocket. The direct or water mediated interactions with Efavirenz and Nevirapine also get affected because of changes in the side chain conformation.



**Figure 3:** (A) Interactions involving Tyr 318 in Efavirenz (orange) bound RT (1IKW) and (B) Changes due to Y318F mutation. (C) Local interaction network involving Lys238 in Efavirenz bound RT structure (1IKW). (D) Changes due to K238T mutation (E) Comparison of side chain conformations of residues in the region involving Lys238, of Efavirenz bound RT, 1IKW (yellow) and of the unliganded form, 1RTJ (green) (F) Interactions involving K238 in Nevirapine bound RT (1VRT).



**Figure 4** (A) Location of V75 with respect to dNTP (orange) binding site. Some residues in the vicinity of V75 involved in dNTP binding are also shown. (B) Interaction between Val 75, Phe 77 and neighbors (C) Model for V75M mutation (D) Residues in the vicinity of V75 involved in template RNA binding.

#### M230L

The position of M230L mutation site in Efavirenz bound RT structure is shown in **Figure 2A**. In the ligand unbound RT structures in the closed conformation (1DLO,1HMV), Met 230 interacts with the side chain of Pro 95 through van der Waals forces (**Figure 2B**). The side chain carboxyl of Asp 186 and the side chain amino group of Trp 266 forms hydrogen bonds with the main chain amino and carbonyl group of Met 230, respectively. The sulphur atom of Met 230 interacts with the  $\pi$  clouds of Tyr 183 and Trp 229 (<5Å distance). In ligand unbound RT structure close to open conformation (1RTJ), Met 230 interacts with Trp 266 side chain through sulphur- $\pi$  ( $\pi$ ) interaction (**Figure 2C**).

In RT bound to Efavirenz and Nevirapine (1IKW, 1VRT), Met 230 sulphur atom is located close to the side chain of Tyr 183 (**Figure 2D, E**). The side chain aromatic rings of Tyr 183 and Trp 229 interact through  $\pi$ - $\pi$  stacking. NH of Trp 229 can also interact with the  $\pi$  cloud of Tyr 183 (**Figure 2E**).

When NNRTI enters the hydrophobic pocket, the side chains of Tyr 181 and Tyr 188 move upward, giving space for the incoming inhibitor. This movement is accompanied by changes in side chain conformations of Trp 229, Tyr 183 and Phe 227. The side chain conformations of Tyr 183 and Trp 229 are stabilized by their interactions among themselves as well as with Met 230. The conformational changes of the residues that occur in the NNRTI binding pocket are shown in **Figure 2D**. The mutation M230L results

in the loss of these interactions, as leucine has a smaller size compared to methionine and can thus affect the entry of the inhibitor (**Figure 2F**). The interaction with Trp 266 which is in a loop joining palm and thumb sub-domains might have a role in stabilizing the structure of RT at different conformations. Thus, this mutation could also alter the flexibility of RT.

#### G190A

Gly 190 in NNRTI bound RT structures makes van der Waals contacts with the drugs Nevirapine and Efavirenz. When Ala replaces Gly as a result of the mutation, the  $\alpha$ -carbon in the Ala side chain, is at a distance of 2.46 Å and 2.58Å from one of the fluorine atoms of Efavirenz and from the cyclo propyl group of nevirapine respectively. These are too short distances and would result in steric clash which can seriously affect the binding of the drug.

#### Y318F

The site of Y318F mutation in Efavirenz bound RT (1IKW) is shown in **Figure 3**. The O-H group in the side chain of Tyr 318 in the NNRTI bound complex (1IKW) interacts with the main chain carbonyl of His 235 through O-H...O hydrogen bond. This interaction helps to hold the side chain of Tyr 318 away from the NNRTI binding site (**Figure 3A**).

The mutation Y318F results in the loss of this interaction as Phe lacks a hydrogen bonding group in its sidechain. This will make Phe 318 side chain more flexible, thereby affecting inhibitor binding. The side

chain hydroxyl of Tyr 318 is at a distance of 3.3Å from Efavirenz and there are favourable van der Waals contacts between them. These interactions are also lost as a result of this mutation (**Figure 3B**).

#### K238T

The interaction network involving Lys 238 shown in **Figure 3C**, might have a role in maintaining the geometry of the NNRTI binding pocket. The mutation K238T would disrupt this interaction network and cause changes in the binding pocket (**Figure 3D**). RT structure in the open conformation (1RTJ, green) superposed on Efavirenz bound RT (1IKW, yellow) is shown in **Figure 3E**. In the NNRTI bound RT structures, Lys 238 interacts with His 235 through van der Waals contacts. Lys 238 side chain makes van der Waals contacts, also with His 315. The side chain amino group forms a salt bridge with the side chain carboxyl of Asp 237 (weak in the case of Nevirapine bound RT (1VRT, **Figure 3F**)). The change in His 235 side chain conformation, as a result of the mutation, could also affect Tyr 318 - His 235 interaction.

#### V75M

The location of Val 75 with respect to dNTP binding site is shown in **Figure 4A**. In the RNA template/primer bound RT (1RTD), Val 75 interacts with Phe 77, Val 60 and Lys 73 through van der Waals contacts. Phe 77 makes extensive van der Waals contacts with the neighboring residues (**Figure 4B**). These interactions help to hold the antiparallel beta sheet ( $\beta$ - $\beta$ ) and the loop involving residues 64-70 over the bound dNTP. van der Waals contacts between Phe 77 and Pro 150 help to position Gln 151 that interacts directly with the dNTP.

The structural model for V75M mutation shows (**Figure 4C**) that terminal methyl group of Met 75 side chain is close to Phe 77. The side chain of Phe 77 also moves to avoid steric clashes and to facilitate formation of sulphur- $\pi$  interaction between these two residues. This affects the structural environment of the neighborhood. The hydrogen bond interaction between Lys 73 and Tyr 146 that help to hold the  $\beta$ - $\beta$  loop in position might be affected. The residues in the vicinity of Val 75 and Phe 77 such as Phe 61, Asp 76 and Arg 78 (**Figure 4D**) interact with the template strand. These interactions could also be affected as a result of the mutation and the position of template bases might be altered.

#### V118I

The position of Val 118 in template/primer bound RT structure is in a hydrophobic environment with tight packing. In the template/primer and TTP bound RT (1RTD), Val 118 interacts with Leu 149 and Phe 160 through van der Waals contacts (**Figure S2A in supplementary material**). In the V118I mutant model (**Figure S2B in supplementary material**), we note changes in the side chain conformation of Tyr 115, Leu 149 and Ser 156 owing to the bulkier  $-\text{CH}_2-\text{CH}_3$  group in the sidechain of Ile compared to Val which has only  $-\text{CH}_3$ . Tyr 115 interacts directly with the dNTP. Ser 156 forms O-H...O hydrogen bond with the main chain carbonyl of Pro 150 and this interaction aid in positioning Gln 151 for dNTP binding. Thus the side chain conformational changes of these residues triggered by the presence of bulkier Ile in the place of Val 118 would affect dNTP binding.

#### Q151M

In the template/primer and TTP bound RT structure (**Figure S2C in supplementary material**), the side chain NH of Gln 151 interacts with the main chain carbonyl at Lys 43 through a N-H...O hydrogen bond. The side chain carbonyl of Gln 151 also forms hydrogen bonds

with side chain N-H of Arg 72 and one of the phosphate oxygens of dTTP. These interaction help directly or indirectly in TTP binding. The side chain of Gln 151 also forms van der Waals contacts with TTP and the side chain benzene rings of Tyr 115 and Phe 116.

The mutation Q151M (**Figure S2D in supplementary material**) would cause a loss of the hydrogen bond interactions due to apolar nature of Met sidechain. The dramatic loss of these interactions in the mutant could affect the closing of  $\beta$ - $\beta$  loop upon dNTP binding. Though the side chain sulphur of Met 151 can interact with the  $\pi$  cloud of Phe 116 the terminal carbon of Met 151 side chain is at a distance of 2.27Å from one of the phosphate oxygens of dTTP. This would lead to steric clashes and can hamper the binding of dNTP.

The mutation Q151M shows high resistance to AZT, D4T, and DDI [3]. The difference between these drugs and others like DDC, 3TC etc is that they interact with the template base through 2 hydrogen bonds (A-T base pairing) and not three. Also, in case of dNTPs the hydrogen bond that the 3'OH forms is an additional interaction that can stabilize the bound dNTP. So the mutations that affect certain interactions that stabilize the dNTP bound to the N site can be thought to be tools for discriminative selection.

#### Conclusion:

The mutations were primarily seen to affect drug efficacy by a number of ways such as altering the geometry of the drug binding pocket, perturbing the entry of the drug into the active site, repositioning the template bases and by discriminating the inhibitors from their natural substrates. These are similar to the mechanisms that are known for clade B strains [3, 17]. The mutations that were analyzed in this study conferred resistance to both NRTIs and NNRTIs (**Figure S3 in supplementary material**). Drugs belonging to these two classes have different mechanisms of actions and hence the sites of mutations on RT would also be expected to be distinct. NRTI resistance analyzed in this study was mediated either by discrimination between inhibitors and natural substrate or by template repositioning, whereas, NNRTI resistance affected the geometry of the active site, drug binding and entry. Sites of residue polymorphisms (not seen in Clade B strains) found in most of our Clade C strains are shown in **Figure S4 (see supplementary material)**. Similar to HIV-1 protease, RT is also a dynamic enzyme but unlike in the former the mutations do not greatly hamper the dynamics of the enzyme [18, 19]. Resistance in RT can broadly be seen as changes in interactions at the active site of the enzyme which can in turn affect its behavior to the inhibitors. Newer groups of drugs in this class would need to take into account these mutations to be effective. RT does present itself as a potent target [3]. There are two copies of RT in a HIV virion and each catalyzes about 20,000 nucleotide incorporations. Thus, there are approximately 5,000 chances for chain termination. We have studied only at a part of the protein and this may not reflect the behavior of the entire enzyme. However, since the study looks at changes in interactions of the amino acids on or near the drug binding site, it is likely that our approach is useful but within its limits.

#### Acknowledgement:

This research is supported to NS by Department of Biotechnology, Government of India as a part of institution-wide support.

#### References:

- [1] RW Shafer, *Clin. Microbiol. Rev.* **15**:247 (2002) [PMID: 11932232]

- [2] LA Kohlstaedt *et al.*, *Science* **256**:1783 (1992) [PMID: 1377403]
- [3] N Sluis-Cremer *et al.*, *Cell. Mol. Life Sci.* **57**:1408 (2000) [PMID: 11078020]
- [4] AJ Kandathil *et al.*, *Int J STD AIDS.* **20**:522 (2009) [PMID: 19625581]
- [5] AJ Kandathil *et al.*, *Indian J Med Microbiol.* **27**: 231 (2009) [PMID: 19584504]
- [6] J Lindberg *et al.*, *Eur. J. Biochem.* **269**:1670 (2002) [PMID: 11895437]
- [7] J Ren *et al.*, *Nat. Struct. Biol.* **2**:293 (1995) [PMID: 7540934]
- [8] Y Hsiou *et al.*, *Structure* **4**:853 (1996) [PMID: 8805568]
- [9] DW Rodgers *et al.*, *Proc. Natl. Acad. Sci. USA* **92**:1222 (1995) [PMID: 7532306]
- [10] R Esnouf *et al.*, *Nat. Struct. Biol.* **2**:303 (1995) [PMID: 7540935]
- [11] H Huang *et al.*, *Science* **282**:1669 (1998) [PMID: 9831551]
- [12] J Scott *et al.*, *J. Am. Chem. Soc.* **106**:765 (1984) (<http://pubs.acs.org/doi/abs/10.1021/ja00315a051>)
- [13] AS Konagurthu *et al.*, *Proteins* **64**:559 (2006) [PMID: 16736488]
- [14] L Holm, C Sander. *Science* **273**:595 (1996) [PMID: 8662544]
- [15] L Holm, J Park, *Bioinformatics* **16**:566 (2000) [PMID: 10980157]
- [16] SV Evans, *J. Mol. Graphics* **11**:134 (1993) [PMID: 8347566]
- [17] N Sluis-Cremer *et al.*, *Curr. HIV Res.* **2**:323 (2004) [PMID: 15544453]
- [18] V Hornak, C Simmerling, *Drug Discov. Today* **12**:132 (2007) [PMID: 17275733]
- [19] JR Collins *et al.*, *Nat. Struct. Biol.* **2**:334 (1995) [PMID: 7796268]

Edited by P. Shapshak

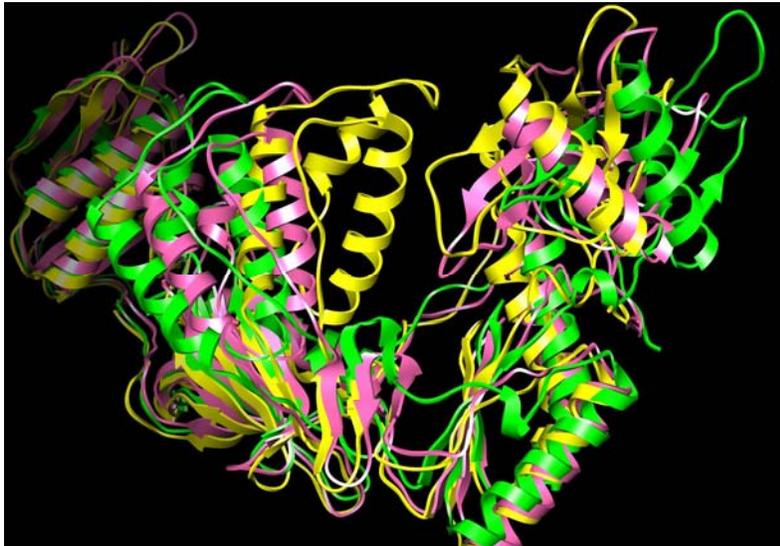
Citation: Kandathil *et al.*, Bioinformation 4(1): 36-45 (2009)

**License statement:** This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

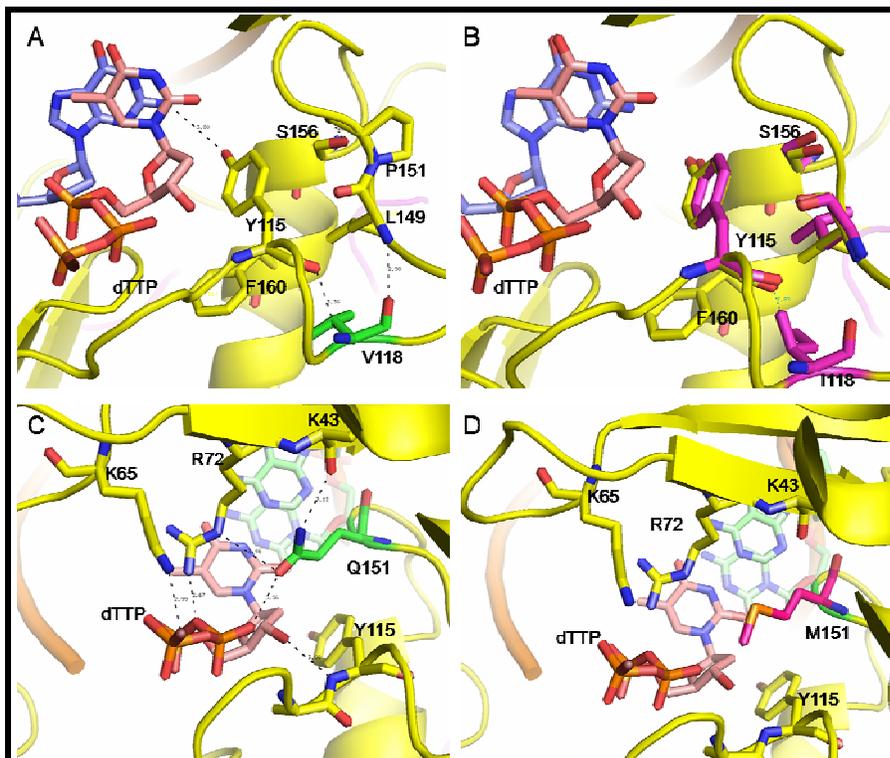
**Supplementary material**

**Table 1:** List of Genbank IDs of the sequences and mutations observed in the strains that were used to generate the models in this study

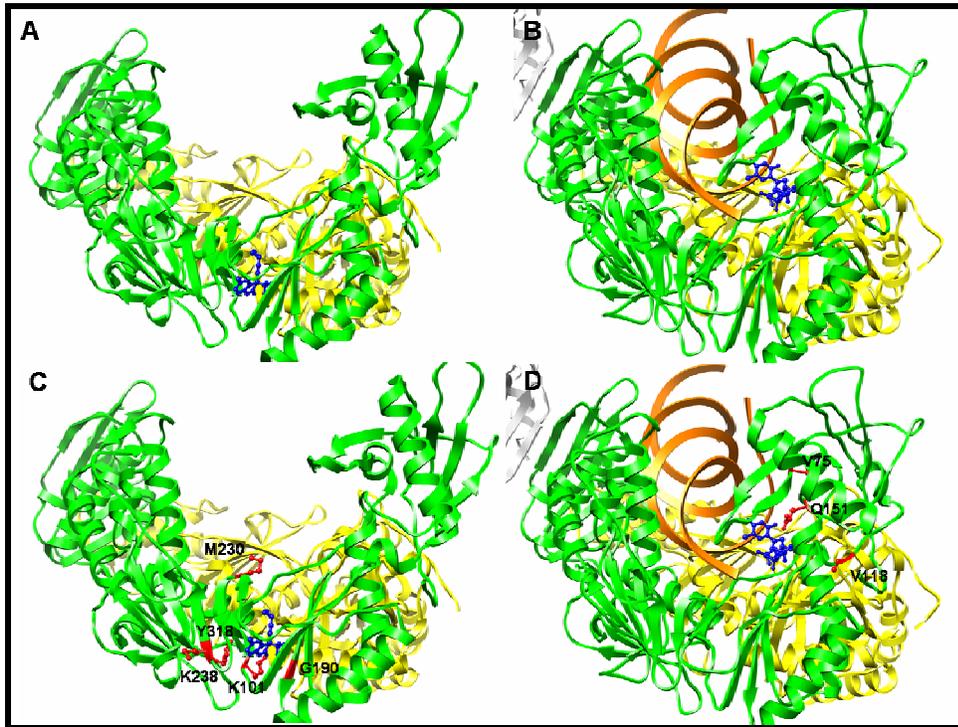
GenBank ID	Subtype (% similarity)	NRTI resistance Mutations	NNRTI resistance mutations	Other mutations
EU037772	C (92.9%)	D67N,K7R, V118I, M184V, T215F, K219E	K101H, Y181C, G190A	V35T, T39D, V60I, L120FL, D121Y, K122E, I135T, S162A, K166R, K173A, D177E, T200A, Q207E, R211K, L228H, V245Q, A272P, V292I, I293V
EU037773	C (93.0%)	M184V	Y188L, Y318F	V35M, E36A, T39D, S48T, V90FILV, P95LPQR, D121Y, K122E, D123E, K126KN, K173A, D177E, T200A, Q207E, R211K, V245Q, A272P, K275I, K277R, K281R, T286A, T290APST, E291D, V292I, I293V, E297A, R307K, S322A
EU037774	C(93.0%)	M184V, T215F	K103N, V108I, M230L	V35T,E36A,T39E,S48T,V60I,D121Y, K122E,I135T,S162A,K166R,K173A,D177E,T200A,E203D,Q207E,H221Y,V245Q,A272P,K277R,K281R,E291D,V292I,I293V, E297A, H315R
EU037775	C(93.0%)	D67N, K70R, V118I, M184V, T215F, K219E	K101H, Y181C, G190A	V35T, T39D, N57KN, V60I, K73N, D121Y, K122E, I135T, S162A, K166R, K173A, D177E, T200A, Q207E, R211K, D218E, L228H, V245Q, A272P, V292I, I293V
EU037776	C(94.8%)	M184V	Y188L	T12S, K14R, I15V, L19T, E35D, M36I, R41K, L63P, H69K, L89M, I93L V35M, T39D, S48T, D121Y, K122E, D123E, K173A, D177E, T200A, Q207E, R211K, F214L, V245Q, A272P, K277R, T286A, E291D, V292I, I293V, E297A
EU037777	C(92.4%)	M41L, D67N, T69D, K70R, V75M, M184V, T215F, K219Q	K103S, G190A	I31L, V35T, T39N, S48T, V60I, D121Y, K122E, T139K, I142V, S162A, K173A, Q174K, D177E, Q197H, T200A, Q207E, H208Y, R211K, L228H, V245K, A272P, E291D, V292I, I293V
EU037778	C(92.6%)	K65KMR,K70KM R, V75I, F77L, Y115F, F116Y, Q151M, M184V	Y188L	V35T, T39N, S48T, V60I, S68G, V106I, D121Y, K122E, I135V, S162A, K173A, D177E, I178V, G196E, T200A, Q207E, Y232H, V245E, A272P, K277R, T286A, E291D, V292I, I293V,
EU037779	C(94.1%)		K103N, V108ILV, K238IKRT	Q23LPQR, V35I, E36A, T39D, S48T, V60I, V90FILV, D121Y, K122E, K173T, Q174K, D177DE, T200A, Q207E, V245Q, A272P, V276I, K281R, T286A, E291D, V292I, I293V
EU0110067	C(94.4%)		Y181C	V35T, E36A, T39D, S48T, V60I, K104R, D121Y, K122E, S162A, K173T, D177E, T200A, Q207E, R211K, D237N, V245Q, D250E, A272P, E291D, V292I, I293D
EU0110054	C(93.3%)		K103R	V35T, T39I, S48T, V60I, T139A, E169D, K173A, D177E, T200A, Q207E, R211K, V245Q, A272P, K277R, K281R, T286A, E291D, V292I, I293V, E297A,



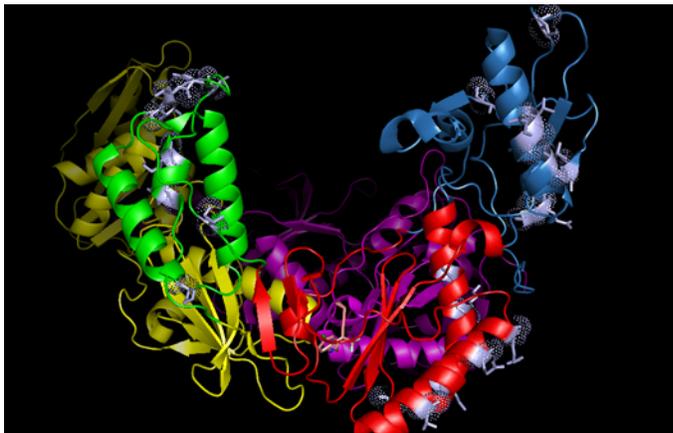
**Figure S1:** Comparison of structures of HIV-1 RT (1DLO-yellow, 1RTJ-pink, 1RTD-green) showing conformational changes from a closed to a DNA bound form.



**Figure S2:** (A) Contacts between Val 118 and the neighbors. dTTP bound at the N site is shown in orange. (B) Changes in the side chain conformations due to V118I mutation. (C) Interactions involving Q151 and neighbors in template/primer bound RT. (D) Model for Q151M mutation



**Figure S3:** (A) Structure of NNRTI (Efavirenz , shown in blue) bound HIV-1 RT (1IKW) showing the site of inhibitor binding. (B) Structure of NRTI (dTTP, shown in blue) bound HIV-1 RT (1RTD) showing the site of inhibitor binding. (C) Structure of Efavirenz bound RT showing the sites of NNRTI resistant mutations (residues indicated in red) analysed in this study. (D) Template/primer and dTTP bound form of HIV RT showing the sites of NRTI resistant mutations (in red), discussed in this study



**Figure S4:** Structure of Efavirenz (orange) bound RT, showing the sites of common residue polymorphisms in Clade C, highlighted in light blue color. The thumb, palm and finger domains are shown in green, red and blue respectively.