8. Summary and Conclusion

In the study reported in this thesis certain related aspects of HIV-1 drug resistance were investigated. It included sequencing the RT and Pr genes to detect mutations, assessment of the mutation effects on drug susceptibility of strains characterized from both treatment naive (n=101) and treatment failed (n=19) individuals. The effect of nucleotide mutations on amino acid changes and its bearing on drug susceptibility i.e., genotypic assay was achieved by using relevant web based databases designed to give information on HIV drug resistance. The amino acid sequences from selected strains (n=23) were taken for modelling to identify the specific effects of the mutations on the functioning of the RT (n=10) and Pr (n=13) molecules. The different genotypic algorithms were also applied on the same set of sequences to investigate any differences in interpretation of the effects of mutation on the susceptibility of the drug. HIV-1 strains have been shown to evolve showing difference in coreceptor usage i.e., CCR5 receptor utilizing R5 strains (primarily macrophage tropic, non syncitium inducing; NSI) to CXCR4 utilizing X4 strains (primarily dual tropic infecting CD4 cells and/or macrophages, syncitium inducing; SI) in a given individual during the course of the infection. Natural history studies prior to use of ART in western countries have shown a predominance of X4 strains in AIDS patient with low CD4 counts. This phenomenon is most marked for subtype B. Furthermore, coreceptor CCR5 blocking therapeutic agent Maraviroc, is now licensed for treating drug experienced individuals. The data on coreceptor usage is preliminary on subtype C which is predominant in India. Hence, an investigation was also carried out to characterize coreceptor usage based on examining the V3 sequences of 47 strains (subtype C) obtained from HIV-1 infected individuals to give information on X4 strains among the presently circulating strains. A study carried out a decade ago at National Institute of Virology, Pune had shown all type subtype C strains from individual to be NSI strains.
Among the strains from the treatment naive individuals hundred strains were identified as clade C and the one strain as clade A1 based on Pr and RT sequence data using the REGA HIV-1 Subtyping Tool - v2.0 and phylogenetic analyses. Amplification and sequencing of the reverse transcriptase gene revealed hundred and one strains with amino acid substitutions at positions 35, 39, 245 and 272. Amino acid substitutions were also seen at positions 60 (n=95), 122 (n=96), 173 (n=97), 200 (n=87), 207 (n=97), 291 (n=95) and 293 (n=94). The M41L amino acid substitution was seen in the RT of one of the strains and is known to cause high level resistance when it occurs along with T215Y. One of our patients had the Y181C mutations which is a mutation that is most frequently selected by nevirapine. The Y181C mutation has been shown to reduce resistance to Zidovudine. Two other NNRTI resistance conferring mutations were seen in one patient each. One of them was A98G, which is known to confer low level of resistance to NNRTI. The strain with the A98G substitution also had L210F/L substitution. The L210W substitution is a part of the TAMs. The other was V179E which causes potential low level resistance to all the NNRTIs based on the scores given by the Stanford HIV drug resistance database. Six strains had amino acid substitutions at position 101 of which two had K101KN and two strains each showed K101Q and K101R. K101N is known to cause resistance to nevirapine and delavirdine. However, K101Q/R does not affect susceptibility to NNRTIs based on the report from the Stanford drug resistance database.

Amino acid substitutions were also seen at positions conferring resistance in the RT. These substitutions were however different from those known to confer drug resistance in clade B e.g., L100F/L (n=1), K103R (n=2), V179I (n=5). The L100I substitution confers intermediate resistance to the NNRTIs. The K103R substitution does not confer resistance by itself but in association with V179D/E can cause resistance to NNRTIs. Five percent of the
reverse transcriptase sequenced showed 179E substitution and 1% had V179E. We could cautiously assume that the genetic barrier needed for the development of resistance would be lesser in the strain with these mutations.

All the strains sequenced for the protease gene showed the amino acid substitution H69K. Other common sites of amino acid substitution were at amino acid positions 15 (n=91), 19 (n=96), 36 (n=99), 41 (n=92), 63 (n=94) and 93 (n=96). It has been reported in clade B that the prevalence of substitutions at amino acid positions 36, 63 and 93 increases in patients harboring strains with multiple PI resistance mutations.

There are six amino acid positions positions (10, 46, 54, 82, 84 and 90) that are associated with reduced susceptibility to more than one PI. Among our patients 12% had an amino acid change at position 82 (V82I). The V82I change has not been associated with any level of resistance to PIs, in contrast to V82 A/T/F/S. Reports from other centres in India have shown similar level of prevalence of this mutation. Amino acid change at position 10 of the protease was seen in 3% of the study sequence and this included the sequence from the clade A strain. According to the Stanford HIV drug resistance database the L10I/V substitution, which were seen in our strains, is associated with resistance to each of the PIs when present with other mutations. The well recognized E35D substitution was seen in 11% of the protease gene sequences in our study. The E35D substitution has been postulated to favour escape from the immune system in addition to conferring drug resistance. Based on the report from the Stanford HIV drug resistance database one of our strains had a E35G amino acid change, which according to the database is slightly more common in viruses from PI-treated
(particularly nelfinavir-treated) compared with untreated persons. This particular strain also had a T74S amino acid change which is associated with reduced susceptibility to nelfinavir.

All the treatment experienced patients had been exposed to NRTIs and NNRTIs. The most common mutation in the RT was the M184V mutation. This mutation is selected by lamivudine (3TC), emtricitabine (FTC) and abacavir (ABC) and confers high level resistance to 3TC. Patients receiving incompletely suppressive 3TC regimens usually develop M184V as their first mutation. The other mutations observed included the TAMs which are seen at amino acid positions 41, 67, 70, 210, 215 and 219 of the RT. The most commonly occurring TAMs was T215F followed by K219E/Q. These TAMs causes drastic reduction in susceptibility to AZT. They do not however drastically affect the action of 3TC although they can affect the synergism of 3TC along with AZT/d4T. The mutations conferring resistance to NNRTIs were more diverse with no particular mutation being prevalent in all the strains. Mutations observed included Y181C, Y188L which are close to the active site of the RT and K103N/S.

D30N was observed in one of the patients showing failure to PIs. D30N confers resistance to NFV only and does not show cross resistance to other PIs. The patient who had the D30N mutation also had mutations N88D and M46I. The M46I mutation was however not observed in the sequence of the sample collected on the second visit which was two years later. The viral load on the second visit showed a $\log_{10} 0.43$ reduction. This could be indicative of reduced viral fitness of the strain as indirectly indicated by reduced viral load. Alternately, there could have been amplification of a different quasispecies in the second sample.
L90M mutation was seen in the second patient and has been reported from patients treated with saquinavir, nelfinavir, indinavir and ritonavir. L90M has been implicated in clinical cross resistance to each of the PIs. In India PIs are not normally included in the first line HAART because of its cost. Hence, the mutation pattern in the protease may be different from previously published reports from the west.

Based on our findings, we have observed no alarming rise in drug resistance mutations or prevalence of known amino acid substitution in the treatment naïve population based on other reports from India. We however suggest the implementation of resistance testing to monitor treatment naïve population for early detection drug resistance. In this study we observed a Pr sequence with T74S and E35G substitution and a RT sequence with A98G, K101R and L210F/L substitutions. If a strain with major mutations associated with resistance to drugs used routinely in the country are found two years in a row with no evidence of previous treatment, sentinel surveillance should be considered as per the WHO recommendation. This will help to keep a tab on drug resistance. Further studies are needed to confirm if mutation causing resistance in clade B have the same effects/mechanism in clade C HIV-1 strains.

The SHDB and G2P databases differed in the sub-typing results. The SHDB identified all but one of the strains as clade C the G2P sub-typed them as clade C,D,H,CRF 08_BC and CRF 03_AB. The sub-typing results of the SHDB showed 100% concordance with the Rega HIV-1 sub-typing tool and hence can be considered a reliable tool for sub-typing of HIV-1 strains based on pol sequence (Protease:1-99 amino acids and Reverse Transcriptase: 1-300 amino acids).
The mean observed concordance between SHDB and G2P was 86% while between SHDB and DS was 43%. The highest concordance between SHDB and G2P was seen for NNRTIs for which it was 89.6% and the least was 81.8% which was observed for NRTIs. The PIs had the lowest level of concordance between SHDB and DS (20.9%) whereas the highest was seen for NRTIs (58.4%). When the level of concordance were seen based on exposure to ART, the G2P was found to have a better level of concordance (76.8%) to SHDB as compared to SHDB Vs DS (33.7%). However, when only the genotypic interpretation and not interpretation based on fold increase (phenotypic IC$_{50}$ values) was taken for DS the level of concordance between SHDB Vs DS for the ART experienced group and treatment naive group rose to 70% and 95.6%, respectively.

Susceptibility profile to etravirine (ETR), tipranavir (TPV) and darunavir (DRV) were shown by SHDB only. This however does not interpret sensitivity to zalcitabine (ddC).

The databases varied in the mutations at two amino positions 3 and 37 for the PR and four amino acid positions 122, 214, 272 and 277 for RT. The SHDB compares the submitted sequence to a consensus clade B sequence whereas G2P uses HXB2 as the reference strain for alignment of the submitted sequence. This could explain the consistent difference at positions 3 and 37 for the PR and amino acid positions 122,214,272 and 277 for RT. The difference in alignment strain used by the two databases however did not lead to non-recognition of any major drug resistance mutation. The SHDB database also makes use of non-clade B strains to generate its output as per the release notes on the SHDB website. Hence, we have used the SHDB as a ‘gold standard’ for this comparison. We do not however have phenotypic data for the strains included in this study and hence not in a position to assign a particular algorithm as superior. The algorithms need to be updated regularly so as to
increase their utility. These results also show a possible need for subtype specific algorithm for interpretation of HIV-1 genotypic drug resistance.

The RT mutations were primarily seen in our in-silico studies 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. The mutations that were analyzed in this study conferred resistance to NRTIs and NNRTIs. 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. The altered geometry of the active site, lack of drug binding and entry into the hydrophobic pocket resulted in NNRTI resistance. 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.

In the HIV-1 protease molecule a majority of the polar residues in HIV-1 protease have flexible side chains. These residues can act in a concerted fashion by forming hydrogen bonds to stabilize the ‘closed’, ‘semi-open’ and ‘open’ conformations of the protease. Water mediated interactions also help to stabilize the different side chain conformations. The drug resistance mutations were found to alter the interaction network that holds the protease in closed conformation. Mutations were seen to confer resistance by a number of ways that included alterations in the geometry of the active site, weakening of interactions that stabilise the closed conformation and also affecting dimer interactions. A large number of mutations that we have analyzed could increase the flexibility of the flaps, which in a drug bound
protease act as clamps to hold the inhibitor in place. These mutations thus favour the partially open conformation by increasing the flexibility of the flaps. Thus, inhibitors with a very low dissociation constant will be more effective. The residues like Gln 58, Asn 83, Asn 88 and Gln 92 and their interactions with the neighbours, which are important for the transition from closed to the open state can be targeted in development of newer drugs to be used in ART. Drugs that interfere with dimer formation if administered with other inhibitors can make ART very effective since some of the resistance mutations also stabilize the interactions at the interface, which is required for the dimer stability when the flexibility is increased. These observations tailored with molecular dynamics simulations will help in the development of newer inhibitors to combat strains which have multiple mutations.

The majority of the study population strains were found to utilize CCR5. There were observed discrepancies in the strains found to utilize CXCR4 by Geno2Pheno and Web PSSM. While Geno2Pheno identified 3/47 (7%) strains as CXCR4 utilizing the Web PSSM identified 1 strain. The strain was identified as CXCR4 utilizing by both the methods. There was one CXCR4 strain from the ART failure group using Geno2Pheno which was however not detected using Web PSSM C matrix. The crown motif sequence was the same for both the R5 and the X4 strains. Previous studies have shown that the crown motif (GPGQ) varies between X4 and R5 strains among clade C strains.

Previous report from Africa has stated a 100% concordance of Web PSSM with a phenotypic assay in detecting CXCR4 strains among clade C strains. Hence, it would be appropriate to state that the strain reported is CXCR4 utilizing. This is the first report of clade C X4 strains from India. This is an indication of the two decades long epidemic with subtype C. One of
the limitations of this study is that we have not tested these strains with a phenotypic assay. We have used only genotypic tools due to lack of appropriate culture facilities for HIV propagation in our laboratory.

The findings are important with the introduction of a CCR5 antagonist for the treatment of HIV-1 infections. Hence, in such scenarios it would be imperative to determine the coreceptor utilized by the strains before administration of these classes of drugs. Strains resistant to maraviroc have been detected in patients elsewhere in whom they have been administrated. These resistant strains were CXCR4 strains which had emerged from a preexisting population in the patient. It would be important to monitor the scenario when this group of drugs is introduced in India.