1. Introduction

Human immunodeficiency virus (HIV) belongs to the family Retroviridae and the subfamily lentivirinae. The other medically important members in this family include human T-lymphotropic viruses I and II. The HIV is of two types and are designated HIV-1 and HIV-2 (1-3). HIV-1 and 2 are subdivided based on genetic differences and are divided into groups and clades/subtypes (1-3). Although the modes of transmission in humans are the same for HIV-1 and HIV-2, the disease progression differs for HIV-1 and HIV-2 infected individuals (4). HIV-1 infection is associated with a faster disease progression than HIV-2 infection (4). Globally there are 33 million HIV infected individuals according to the figures released by the Joint United Nations Program on HIV/AIDS (UNAIDS) (5). India has about 3 million HIV infected individuals as per the latest National AIDS Control Organization (NACO) surveillance reports with an estimated prevalence of 0.36% in the general population (6).

Studies have shown that ART has helped to increase the survival rate of individuals infected with HIV. Porter et al have shown that the proportion of person-time on ART has shown an increase from 22% in 1997 to 57% in 2001(7). Effective use of ART has increased the mean survival time of HIV infected individual significantly altering the lower survival time in the natural history of HIV infection (8).

Reports from Africa and Asia have shown similar effects of ART on survival rate as compared to the pre-ART era (9, 10). A HIV infected individual is put on combination of drugs which target various proteins of HIV whereby the production of mature viral particles in the infected individual are inhibited (11-14). This combination of drugs is called antiretroviral therapy (ART) has been found to be effective in controlling HIV infections than administration of only one drug (15, 16). The health agencies like NACO and other non-
governmental organizations (NGOs) are taking a number of steps to control the spread of HIV infections in India. These steps have helped in bringing about a decline in the number of HIV infections in the country (17, 18). One of the steps taken to tackle HIV infections is to increase the accessibility of antiretroviral drugs by providing them free of cost through the National AIDS Control Program Phase III (NACP III) by NACO (www.naco.org). Earlier WHO/NACO had initiated the “3 by 5” motto (treating 3 million by 2005). At present ART has helped in reducing the morbidity and mortality associated with HIV infections and has also helped in increasing the life span. Hence, India has a considerable number of people living with HIV and AIDS (PLWHA). This wide use of ART will have an overall beneficial effect in the control of the spread of HIV infection but the success of this will largely depend on careful use of ART regimens by physicians and compliance of the infected individual (19, 20). Presently, it is established that HIV is quickly capable of acquiring drug resistance by mutation under the influence of drugs when not used in appropriate combinations and uninterrupted treatment. Drug resistance and transmission of drug resistant strains will be a major threat for the success of the NACO initiated program (21). One of the primary requirements would be monitoring emergence of drug resistance is to periodically adjust the treatment regimens avoiding ineffective drug combinations.

HIV like all the other members of the family Retroviridae requires the enzyme reverse transcriptase (RT) to reverse transcribe its genomic RNA into DNA. The HIV-1 RT is the prime target of nucleoside/nucleotide reverse transcriptase inhibitors (NRTIs) and non-nucleoside reverse transcriptase inhibitors (NNRTIs) which are used in ART. Another target of the ART regimen is the HIV-1 protease (Pr). The group of drugs that target the protease
are known as protease inhibitors (PIs). In India, a combination from NRTIs and NNRTIs are used as first line drugs to treat HIV-1 infection (22, 23).

The methods employed to determine HIV-1 drug resistance can be divided into genotypic and phenotypic assays. Phenotypic assays involve the use of cell cultures to propagate HIV-1 containing inserted amplicons in the presence of varying concentration of the drugs. These amplicons are derived by amplification of the RT and Pr gene of the predominant HIV-1 strain from the infected individual. In contrast, a genotypic assay involves analysis of changes in amino acid sequences in the regions of HIV-1 that are targeted by ART. Thus this method detects resistance by observing for mutations. Both the methods have their individual advantages and disadvantages (24, 25).

There are a number of web based public access databases like the most widely used one, the HIV-1 Stanford drug resistance database. These contain correlating information between genotypic and phenotypic assays. The databases upon submission of the HIV-1 RT and Pr nucleotide sequence will give a list of mutations and polymorphisms and their contribution to drug resistance. Previous studies have shown varying levels of agreement among these different databases used for genotypic HIV-1 drug resistance testing interpretation (26, 27).

The most common transmitted drug resistance mutation detected among the treatment naive population are those conferring resistance to NRTIs (28). Numerous polymorphisms and mutations have been reported which are known to contribute to drug resistance in clade B strains (29). There are a few reports from India on such changes in the clade C strains (30,
These reports from ART naïve individuals in India have shown the Pr gene to be more polymorphic when compared to the RT gene (30, 31). In this thesis mutations were defined as differences from the wild type consensus B amino acid sequences which conferred resistance to the strains while all other amino acid changes not associated with known drug resistance were defined as polymorphisms. Mutations that confer resistance can be classified as primary and secondary mutations (16). Primary mutations result in reduced drug susceptibility directly whereas secondary mutations reduce drug susceptibility or improve replication fitness of the virus in conjunction with primary mutations (16).

The benefits of ART to an infected individual could be hampered by a number of factors with adherence to the regimen being one of them (19). Adherence is a variable process not only among HIV infected individuals but also in a given individual from time to time (32). Individuals with a high level of adherence have a greater chance of viral suppression than individuals who are adherent moderately or non adherent (32). Incomplete suppression can lead to emergence of drug resistant strains (16). There have been very few published reports of ART treatment failure from India (33, 34). Previously published reports from India have reported lamivudine resistance conferring mutation M184V as the most commonly observed mutation in the RT of HIV-1 strains in patients failing ART (33, 34). There is no evidence for any protease mutations in these reports. More studies from India would help to contribute to the global data on the common mutational patterns observed in clade C HIV-1 strains from India available in the drug resistance data base like the Stanford HIV database.

As mentioned earlier, poor compliance from the infected individual can thwart the benefits of ART (19). Failure of a particular regimen implies prescription of a new regimen containing
the next best combination of drugs by the physician to the individual. Prospective studies have shown that availability of drug resistance reports for infected individuals have allowed more effective regimens to be prescribed as compared to a control group whose physicians had no access to drug resistance reports (35, 36).

The RT enzyme has two polymerase activities, namely, RNA-dependent DNA polymerization and DNA-dependent DNA polymerization (24). The RT is a heterodimer consisting of two subunits p66 and p51 (24, 37). 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 (24, 37, 38). The p66 subunit is responsible for the enzymatic activity of the RT and contains the DNA binding groove and the active site. The overall shape of the p66 subunit has been compared to that of a right hand (37). It is conceptualized of the “fingers” (residues 1-85, 118-155), “palm” (86-117, 156-237) and “thumb” (238-318) based on X-ray crystallography studies. The p66 has an ‘open’ conformation to enable it to ‘grasp’ the template. The “palm” subdomain is the location of the catalytic aspartate residues (D110, D185, D186) (37). 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 (24, 37-39). The RT gene sequence of HIV-1 can develop non-synonymous 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) (www.rcsb.org).
HIV-1 protease is a homo-dimer with each protomer consisting of 99 amino acids. It is an aspartyl protease with sequence motif Asp-Thr-Gly at positions 25 - 27 in each protomer constituting the active site at the dimer interface. The HIV-1 protease is responsible for post-translational modification of gag and gag-pol polyprotein. The hydrophobic substrate cleft recognizes and cleaves 9 different peptide sequences to produce the matrix, capsid, nucleocapsid and p6 proteins from the gag polyprotein and protease, reverse transcriptase (RT) and integrase proteins from the gag-pol polyprotein. In the absence of protease, only immature virions are produced that are incapable of completing the replication cycle. The active site of the protease is surrounded by two hairpin structures or ‘flaps’ which serve as clamps. The protease is said to be in an ‘open’, ‘semi-open’ or ‘closed’ conformation depending on the positions of these flaps. The HIV-1 protease structure can be divided into the following substructures based on the Cantilever-Fulcrum model of flap opening. The residue numbers corresponding to the region concerned are given within parentheses: fulcrum (10-23), flap elbow (35-42), flap (43-58), flap tip (49-52) and cantilever (59-75). The HIV-1 protease inhibitors are examples of compounds developed by structure-based rational design. Mutations in the protease enzyme can reduce the efficacy of the protease inhibitors. Previous studies on non-clade C strains have shown that mutations can confer complete or partial resistance either by affecting the flap flexibility of the protease or by altering the drug binding site.

The human immunodeficiency virus uses co-receptors in addition to CD4 receptors for viral entry into target cells like T-cells. These most commonly used co-receptors are the chemokine receptors CCR5 and CXCR4. CCR5 co-receptors are used in the early stages of HIV-1 infection by strains designated as R5, but there is a switch to CXCR4 usage in the
later stage by strains designated as X4 (50). The switch occurs in about 50% of individuals infected with subtype B strains (51). This switch in co-receptor usage is not common among clade C strains (52). Previous data on co-receptor usage in HIV-1 clade C strains from India has shown the absence of this co-receptor switch (53). A recent report from Africa have however shown, an emergence of X4 utilizing clade C strains to the extent of 30% in infected individuals (54). The introduction of a new class of drugs, in the west, targeting CCR5 (55) makes it important to have data on co-receptor usage of the predominant subtype C HIV-1 strains, especially those circulating in India. Studies to determine co-receptor usage in HIV-1 strains from individuals who are ART naïve and those showing failure to ART from India are needed.

A review of literature revealed certain lacunae about clade C strains of HIV-1 circulating in India and hence this study was carried out after formulating a hypothesis and setting objectives to investigate the validity of the hypothesis

**Hypothesis**

Nucleotide variations in the reverse transcriptase and protease regions of the pol and env genes of HIV-1 determine drug resistance and influence co-receptor usage for clade C.