CHAPTER 1

INTRODUCTION

Five cases of *Pneumocystis carinii* Pneumonia (PCP), a rare opportunistic infection in immune compromised persons, were reported in previously healthy and young gay men from Los Angeles in 1981 (Gottlieb, 2006). A set of men were then reported to develop Kaposi’s sarcoma. Several patients reported for PCP and KP. The committee set by U.S. CDCP recognized this outbreak and declared the outbreak as Acquired Immune Deficiency Syndrome (AIDS) (1982).

One year later in 1983, Robert Gallo and Luc Montagnier reported in Science separately that the causative agent for AIDS might be a retrovirus (Gallo, Sarin et al., 1983); (Barre-Sinoussi, Chermann et al., 1983). Gallo called the virus HTLV-III as its shape was similar to Human T Lymphotropic Viruses (HTLVs). Unlike Gallo’s finding, Montagnier found that proteins present in the core of this newly isolated virus were totally different from HTLV-1 immunologically. His group termed this virus as Lymphadenopathy-Associated Virus (LAV). In the subsequent year (1984) Robert Gallo and Jay Levy independently established the association of this virus with AIDS. Later on, in 1985, an International Virus Taxonomy Consortium decided to name this virus as Human Immunodeficiency Virus type 1, or HIV-1 (Alizon, Sonigo et al., 1984; Wain-Hobson, Sonigo et al., 1985; Wain-Hobson, Alizon et al., 1985). In 2008, Nobel Prize in medicine for this discovery was given to Montagnier and Francoise Barre-Sinoussi.
1.1 Worldwide AIDS prevalence

As of 2011 report, 34 million people were living with HIV, indicating 17% increase in HIV infected individuals in last 10 years. 2.7 million Individuals got infected with HIV in 2010. 1.76 million People have died in 2010 because of AIDS. Introduction of HAART in low and middle income countries has helped to reduce not only the number of deaths due to AIDS in patients but also infection in children. Sub-Saharan Africa region is still heavily affected by HIV epidemic. 50% of total infected population living with HIV is women.

Table 1.1: Statistics of AIDS epidemic from 2001-2010 (UNAIDS fact sheet 2011, World AIDS day report 2011)

<table>
<thead>
<tr>
<th></th>
<th>2001</th>
<th>2005</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>People living with HIV</td>
<td>28.6 million</td>
<td>31.0 million</td>
<td>32.3 million</td>
<td>32.9 million</td>
<td>34 million</td>
</tr>
<tr>
<td></td>
<td>[26.7-30.9 million]</td>
<td>[29.2-32.7 million]</td>
<td>[30.4-33.8 million]</td>
<td>[31.0-34.4 million]</td>
<td>[31.6-35.2 million]</td>
</tr>
<tr>
<td>New HIV infections</td>
<td>3.15 million</td>
<td>2.81 million</td>
<td>2.74 million</td>
<td>2.72 million</td>
<td>2.67 million</td>
</tr>
<tr>
<td></td>
<td>[2.96-3.33 million]</td>
<td>[2.63-2.97 million]</td>
<td>[2.52-2.93 million]</td>
<td>[2.48-2.93 million]</td>
<td>[2.46-2.90 million]</td>
</tr>
<tr>
<td>AIDS-related deaths</td>
<td>1.85 million</td>
<td>2.22 million</td>
<td>2.04 million</td>
<td>1.89 million</td>
<td>1.76 million</td>
</tr>
<tr>
<td></td>
<td>[1.67-2.16 million]</td>
<td>[2.07-2.48 million]</td>
<td>[1.87-2.21 million]</td>
<td>[1.72-2.05 million]</td>
<td>[1.59-1.91 million]</td>
</tr>
<tr>
<td>New infections in children</td>
<td>550 000 [490 000-620 000]</td>
<td>540 000 [480 000-600 000]</td>
<td>460 000 [400 000-510 000]</td>
<td>430 000 [370 000-490 000]</td>
<td>390 000 [340 000-450 000]</td>
</tr>
</tbody>
</table>
Table 1.2: Worldwide AIDS prevalence (The 2011 UNAIDS World AIDS Day report)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-Saharan Africa</td>
<td>22.9 million [21.6–24.1 million]</td>
<td>1.9 million [1.7–2.1 million]</td>
<td>1.2 million [1.1–1.4 million]</td>
<td>5% [4.7%–5.2%]</td>
</tr>
<tr>
<td>Middle East and North Africa</td>
<td>470 000 [350 000–570 000]</td>
<td>69 000 [40 000–73 000]</td>
<td>35 000 [25 000–42 000]</td>
<td>0.2% [0.2%–0.3%]</td>
</tr>
<tr>
<td>South and South-East Asia</td>
<td>4.0 million [3.5–4.5 million]</td>
<td>270 000 [230 000–340 000]</td>
<td>250 000 [210 000–280 000]</td>
<td>0.3% [0.3%–0.3%]</td>
</tr>
<tr>
<td>East Asia</td>
<td>790 000 [680 000–1.1 million]</td>
<td>88 000 [48 000–160 000]</td>
<td>56 000 [40 000–76 000]</td>
<td>0.1% [0.1%–0.1%]</td>
</tr>
<tr>
<td>Oceania</td>
<td>54 000 [48 000–62 000]</td>
<td>3300 [2400–4200]</td>
<td>1600 [1200–2000]</td>
<td>0.3% [0.2%–0.3%]</td>
</tr>
<tr>
<td>Latin America</td>
<td>1.5 million [1.2–1.7 million]</td>
<td>100 000 [73 000–140 000]</td>
<td>67 000 [45 000–95 000]</td>
<td>0.4% [0.3%–0.5%]</td>
</tr>
<tr>
<td>Caribbean</td>
<td>200 000 [170 000–220 000]</td>
<td>12 000 [9400–17 000]</td>
<td>9000 [6900–12 000]</td>
<td>0.9% [0.8%–1.0%]</td>
</tr>
<tr>
<td>Eastern Europe and Central Asia</td>
<td>1.5 million [1.3–1.7 million]</td>
<td>160 000 [110 000–200 000]</td>
<td>90 000 [74 000–110 000]</td>
<td>0.9% [0.8%–1.1%]</td>
</tr>
<tr>
<td>Western and Central Europe</td>
<td>840 000 [770 000–930 000]</td>
<td>30 000 [22 000–39 000]</td>
<td>9900 [8900–11 000]</td>
<td>0.2% [0.2%–0.2%]</td>
</tr>
<tr>
<td>North America</td>
<td>1.3 million [1.0–1.9 million]</td>
<td>58 000 [24 000–130 000]</td>
<td>20 000 [16 000–27 000]</td>
<td>0.6% [0.5%–0.9%]</td>
</tr>
</tbody>
</table>

Source: The 2011 UNAIDS World AIDS Day report
1.2 AIDS prevalence in India

More than 60% of the world’s population resides in Asia and it ranked second in countries with the highest number of people living with HIV-AIDS after the sub-Saharan African region. Demographically, India is the second largest country in the world. It has roughly half of Asia’s HIV prevalence. There are 23.95 lac of people living with HIV infection in India. In the year 2009, 1.2 lac people newly infected were reported. 3.5% of all infections were reported in children under 15 years of age, while majority of the cases (83%) were in age group 15-49 years. Women constitute 39% (9.3 lac) of HIV infection. The southern Indian states (Tamil Nadu – 1.5 lac, Maharashtra – 4.2 lac, Andhra Pradesh – 5 lac and Karnataka – 2.5 lac,) are highly affected. 55% of total cases were reported from these states. States from north India constitute 22% of total HIV infections (West Bengal, Bihar, Uttar Pradesh & Gujarat). 50,000–1 lac cases were reported in each of the four states (Madhya Pradesh, Odisha, Punjab & Rajasthan nearly 12% of total infections). The data of annual projection indicated a gradual decline in number of new infection by 50% during the last decade. The primary drivers of HIV transmission are female sex workers, men who have unprotected sex with men and injected drug users.
Figure 1.1: AIDS prevalence in India (NACO report 2010)
1.3 HIV Taxonomy and Classification

HIV is a member of Retroviridae family, subfamily Lentivirinae and Lentivirus genus in primate lentivirus group. Lentiviruses (lenti, Latin for slow) are characterized by slow growth and long asymptomatic incubation period. HIV-1 is the result of zoonotic transmission of SIVcpz in Chimpanzees in West Central Africa. HIV is classified as HIV-1 and HIV-2 on the basis of sequence deviation exceeding 50% and on the presence of vpx gene in HIV-2. HIV-2 has 40-60% sequence similarity with HIV-1 and was isolated from some AIDS patients of West Africa in 1986. HIV-2 is less predominant than HIV-1, which is found across the world. HIV-2 is mainly concentrated in West African part of the globe, with some cases also identified in America and Western Europe. The clinical manifestation associated with both etiological agents of AIDS is indistinguishable. However, epidemiology studies suggest that the incubation period of disease development is longer for HIV-2 than HIV-1. In addition, HIV-2 maintains low level of viremia and shows lower rates of transmission and thereby remains geographically localized compared to HIV-1 (Alizon & Montagnier, 1987; Yokoyama & Gojobori, 1987; Alizon, Sonigo et al., 1984; Biberfeld, Brown et al., 1987).

Based upon differences in envelope gene (env) three groups of HIV-1 were reported. These three groups are the “major” group M, the “outlier” group O and the “new or non-M and non-O” group N.

Recently, a new Human Immunodeficiency Virus closely resembling SIVgor has been isolated from a Cameroonian woman.

This virus isolate doesn’t show evidence of recombination with any of the HIV-1 lineages and it is distinct from other HIV-1 groups M, N and O; hence it was designated as HIV-1 group P.

Within HIV-1 groups, M accounts for more than 90% infections worldwide.

Group M is further subclassified into nine genetically distinct subtypes (clades): A, BC, D, F, G, H, J, and K. These subtypes show 25-30% amino acid differences in their env sequence due to the error prone Reverse Transcriptase activity. Within A and F
subtypes, sub clusters have been identified and designated as sub-subtypes A1, A2 and F1, F2 respectively which are more closer to each other than to other subtypes (Biberfeld Brown et al., 1987; Thomson Delgado et al., 2002).

Further, there are circulating recombinant forms of HIV, generated by mixing of genetic material of two subtypes in cells of infected individual. CRFs were first identified by full length sequence analysis of virus isolated from Thailand and Central Africa. The distribution of HIV-1 and CRFs is not even in the world, with subtypes A and C being widespread, subtype A predominates in West and Central Africa. Subtype C is the major cause of AIDS in Asian countries like India and countries of South and East Africa. Subtype C constitutes nearly half of world’s HIV infections. In United States, Australia, Europe and Japan subtype B is prevalent. In East and central Africa subtype D is prevalent. CRF A/E though originated in Central Africa, has spread in South-East Asia. Central Africa, Eastern European countries and South America reported prevalence of subtype F.

In Central Europe, West and East Africa Subtype G and CRF A/G have been reported. Subtype H and J were found in central Africa and Central America respectively. Subtype K is restricted in Democratic Republic of Congo and Cameroon. Many Unique Recombinant Forms (URFs) also exist among HIV isolates.

These arise in a similar way as the CRFs, when two or more strains undergo recombination within an individual host to form a mosaic genome. These URFs globally occur at a very high frequency. However, in contrast to CRF’s they have not spread beyond their initial host. These genetically diverse populations which have evolved in the absence of selection pressure in the HIV-1 infected individuals are also called quasispecies (Vidal Mulanga-Kabeya et al., 2000; Vidal Peeters et al., 2000; Leitner, Halapiet al., 1993; Leitner Escañilla et al., 1995).
Figure 1.2: Neighbor joining phylogenetic tree of full length genome sequence of representative HIV-1 isolates (Adapted from Thomson Perez-Alvarez et al., 2002)
Figure 1.3: The frequency of each HIV-1 subtype and recombinant forms was estimated in each country based on published findings (Adapted from (Arien, Vanhamet al., 2007)).
1.4 HIV pathogenesis

A gradual and slow deterioration of the immune system is a characteristic of HIV infection. Although the clinical manifestation of HIV-1-associated immune system dysfunction varies from individual to individual, the natural course of HIV-1 infection generally involves three phases: acute phase, chronic phase and clinical apparent disease. Acute phase, lasting for 6-12 weeks, is characterized by flu-like symptoms, peak virus load and decline in number of CD4+ cells within the peripheral circulation. Chronic asymptomatic phase, lasting for an average seven to ten years, is characterized by a slow and persistent decrease in number of T lymphocytes thereby affecting integrity of immune system.

During this period, HIV-1 replication reaches a steady level known as setpoint. Clinically apparent disease or AIDS onset is characterized by drastic reduction in CD4+ cells to 200 cells per mm$^3$ from the normal level of 800-1200 cells per mm$^3$. The CD4+ T-cell depletion is associated with high viral turnover and is continued by the progressive loss of T-cell mediated immunity. It has been suggested that apoptosis may be one of the mechanisms that is responsible for T-cell depletion in HIV-1 infection. Abnormal levels of apoptosis have been shown in CD4+ T-cells from HIV-1 infected individuals.

According to the CDC, AIDS patients suffer from opportunistic infections. Lungs, esophagus and bronchi get easily infected by yeast. *Pneumocystis pneumonia*, toxoplasmosis, Kaposi’s sarcoma, CytoMegalovirus (CMV), *Salmonella* infections and tuberculosis occur in AIDS patients. In addition, the patients later develop cryptosporidium and/or coccidiomycosis and die. Time of AIDS onset varies from individual to individual from 2-15 years (Haseltine 1988; Detels, English et al., 1988).
Figure 1.4: HIV-1 /AIDS disease progression (Adapted from Sewellet al, 2000).
1.5 Life cycle of HIV-1

The life cycle of HIV-1 begins with viral attachment to cells. The gp120 protein of virus attaches to CD4 receptor followed by binding to chemokine receptor CXCR4 or CCR5. The HIV gp120 has high affinity for binding to CD4. The CD4-gp120 complex causes conformational change exposing site for co-receptor binding. Gp41, a transmembrane glycoprotein of env has fusion peptide. The peptide embeds into the cell membrane. The viral membrane fuses with cell membrane and viral core is released into the cell. Uncoating of the core exposes a viral nucleoprotein complex, containing the RNA genome, RT, and other proteins into the cytoplasm (Loftin Kienzle et al., 2011; Wu & Yoder, 2009). The viral Reverse Transcriptase then reverse transcribes genomic RNA to a linear double stranded DNA. The viral dsDNA then translocates into the nucleus and integrates into the host cell genome, where it is called as ‘proviral’ DNA. The viral DNA is integrated into a host chromosome by integrase enzyme.

The integrated proviral DNA is transcribed into mRNA. The mRNA is spliced into small pieces of mRNA. In cytoplasm, they are translated into the regulatory proteins tat and rev. In the nucleus, Rev protein accumulates. These protein bind to viral mRNAs and causes translocation of unspliced RNAs out of the nucleus. At this stage, the full-length mRNA is translated to produce structural proteins gag and env. The full-length RNA is actually the virus genome; binding of gag to full length RNA is followed by packaging into new virus particles. There is a difference in packaging of HIV-1 and HIV-2. Unlike HIV-1, HIV-2 preferably binds to mRNA that was used to create the gag protein itself. The plasmamembrane of the host cell is the site for assembly of virions. The translocation of env polyprotein (gp160) from the endoplasmic reticulum to the Golgi complex causes cleavage of env by proteases to form glycoprotein gp41 and gp120.

In the plasma membrane, newly formed env proteins are translocated. The env gp120 attaches to membrane with help of gp41. Gag (p55) and gag-Pol (p160) poly proteins with the HIV genomic RNA are incorporated near plasmamembrane. All proteins are incorporated to form immature virus particle. Protease cleaves the gag, pol poly
proteins into functional proteins and fully mature virus is the formed. The mature virus is then able to infect another cell Haseltine1988; Huber, McCoy et al., 1989.

Figure 1.5: Life cycle of HIV-1

The HIV binds to the CD4 receptor on the cell surface with viral gp120 protein and then the viral protein gp41 fuses with one of the coreceptor CCR5 or CXCR4 or utilizes other membrane glycoproteins for the fusion and attachment. The viral RNA is then inserted into the cell cytoplasm. The RNA is then transcribed by the viral enzyme reverse transcriptase to DNA which is then integrated into the host genome. The viral proteins are then synthesized and immature virus particles are formed. The assembling of the virus particle in mature virion takes place at the cell membrane. The mature virus then buds out from the cell membrane and released into circulation and thus its ready to infect another cell.
1.6 HIV structure

Unlike retrovirus, HIV-1 has distinct structure. It has spherical shape and size of HIV-1 is 120 nm in diameter (around 60 times smaller than a red blood cell).

Two copies of single-stranded RNA are embedded into conical capsid. The viral p24 protein forms the capsid. The RNA is 9749 nucleotides long. The single-stranded RNA is engulfed by p6, p7, the nucleo-capsid proteins and Reverse Transcriptase and Integrase, Protease enzymes that are required for the development of the virion. The individual enzymes, Protease, Reverse Transcriptase (RT), and Integrase (IN), are cleaved from Pr160GagPol by the viral PR.

The p6 and p7 proteins of nucleocapsid associate with the genomic RNA (one molecule per hexamer). It protects the RNA from digestion by nucleases. Capsid is surrounded by matrix. Matrix ensures the integrity of the virion particle. Other viral proteins like Vif, Vpr, Nef, p7 and viral Protease are also enclosed within virion particles. Budding of capsid from the host cell initiates formation of envelope, taking some of the host-cell membrane with it.

The envelope includes the glycoproteins gp120 and gp41. In 2006, structure of env using cryo-electron microscopy was reported. According to this model, the env spike is formed by 3 copies of gp120-gp41 heterodimers. Soon thereafter, a single-stalk "mushroom model was proposed."
The mushroom model suggested a head consisting of a trimer gp120s and agp41 stem, which appears as a compact structure with no obvious separation between the three monomers, anchoring it to the envelope (Frankel & Young, 1998). More recently, further evidence backing up the heterodimer trimer-based model have been reported.

Components of virus structure are described below in brief.

Figure 1.6: Schematic representation of HIV-1 virion structure (Source NIAID website www.niaid.nih.gov).
1.6.1 HIV-1 Long Terminal Repeat

The viral genome has sequences located at 5' and 3' ends which serve as a viral promoter known as Long Terminal Repeat or LTR. The 5' LTR and 3' LTR have the same nucleotide sequence but different functional properties. A major functional difference between the two LTRs is the apparent low rate of transcription starting in the 3' LTR in comparison with the 5' LTR. The HIV-1 LTR is approximately 650 bp in length and comprised of U3R and U5 regions. The basal or core promoter (nt -78 to -1), the trans-activation response (TAR) element found within R (nt +1 to +60), core enhancer (nt -105 to -79) and a modulatory region (nt -454 to -106) form the functional regions. TATA box which binds to TATA binding protein (TBP), is located just upstream of the transcription start site (+1).

Three Sp1 binding sites in the core promoter and two NFkB enhancer motifs are the key elements involved in the regulation of HIV-1 transcription. The 5' end of the U3 region has a fourth Sp1 binding site in addition to USF and NFkB sites, which is essential for HIV-1 LTR negative sense transcription. Upon integration into the host genome, LTR functions like a host promoter. The cellular host proteins bind to a variety of cis-acting DNA sequences located within the LTR. LTR-mediated transcription is initiated by binding of cellular machinery to the HIV-1 core promoter. This initiation of transcription leads to formation of short transcripts after which the viral transactivator protein (Tat) recruits PTEFb complex to transcribe a full-length transcript (Rice & Mathews, 1998; Kogan & Rappaport, 2011; Palmieri & Trimboli et al., 2003; Pumfrey Deng et al., 2003).
Figure 1.7. The structure of HIV-1 5'-long terminal repeat (LTR). (Adapted and modified from Tripathy, Abbasetal., 2011)
1.6.2 Gag And Gag-Pol Precursor Polyproteins

The 55kD gag is cleaved to produce nucleocapsid (NC), matrix (MA), capsid (CA), p1, p2, and p6Gag proteins. The assembly and budding of immature viral particles is mediated by gag, which polymerize to form a shell-like layer associated with the inner surface of virus membrane during budding. Pol is a 160kDa gag-pol fusion protein (Pr160-gag-pol). The gag-pol precursor polyprotein contains the gag domain plus the integrase, protease and Reversetranscriptase domains. The viral enzymes catalyze a number of essential steps in HIV-1 replication. gag and gag-pol precursor proteins are cleaved by viral protease. It is thought that gag-pol and gag associates in infected cells, before they are cleaved into their constituent proteins by protease component of gag-pol during or shortly after, budding. This process of maturation results in dramatic morphological changes and formation of an infectious virus (WattsDang et al., 2009).

Figure 1.8: gag-pol precursor polyproteins Source J.M. Watts, Nature, 2009, 460(7256): 711–716
1.6.3 Matrix

Matrix is a 132 aminoacid myristoylated structural protein that plays a major role in assembling virus structure. It is also required for virion infectivity. MA is also a part of the pre-integration and reverse transcription complexes (PIC and RTC) and being a part of the PIC, it is involved in import of PIC into the nucleus. It also contains nuclear export signal (NES) to direct the unspliced viral RNA to plasmamembrane (Chukkapalli & Ono, 2011; Ganser-Pornillos Yeager et al., 2012).

1.6.4 Capsid (CA)

Capsid is the core of the virus particle. Capsid is the largest gag protein, present next to matrix in gag precursor polypeptide. It is the highly conserved sequence of the gag protein. CA is largely helical. It consists of two helical domains, N-Terminal Domain (NTD) and C-Terminal Domain (CTD) which are separated by a flexible linker inter domain. The C-terminal domain of capsid is important for Pr55Gag multimerization due to its ability to form dimers. There are around 2000 molecules of capsid per virion. In the viral replication cycle immediately after reverse transcription, capsid protein disassembles to release the PIC. The C-terminal domain of capsid is necessary for its function in viral assembly (Fassati 2012; Bocanegra Rodriguez-Huete et al., 2012).

1.6.5 Nucleocapsid (p7 Gag)

The nucleocapsid (NC) is 60-90 amino acids long. Nucleocapsid coat the genomic RNA inside the virion core. The primary function of this protein is to bind specifically to the packaging signal and deliver full-length viral genome into the assembling virion. NC has high affinity to genomic RNA. Cysteine and lysine residue motifs are spaced (CCHC). This CCHC motif resembles other short cysteine and histidine-containing structures called "zinc fingers," that coordinate Zn<sup>++</sup> ions. The CCHC motif mediates binding of this protein to nucleic acids i.e. RNA capsidation. The lysine or arginine residues are present in clusters following the CCHC motifs.
which are involved in interaction with "packaging sequences, "ψ" near the 5’ end of retroviral genomic RNAs. NC also contains sequences called "assembly domains", which are essential for assembly or budding of the virion. NC protein is also known to induce a conformational change in viral RNA dimer, making its structure physically more stable (Ganser-Pornillos Yeager et al., 2012).

1.6.6 p6 gag

The C-terminal end of gag encodes the 52 amino acid protein p6. This protein has high proline content and is shown to be involved in the formation and release of virus particles from cell surface. It is also found to be important for incorporation of Vpr during viral assembly (Mervis Ahmad et al., 1988).

1.6.7 Envelope glycoprotein

The envelope protein of HIV-1 play important role in the virus replication cycle. During the entry, the fusion of virus to cell membrane is mediated by env. The env glycoprotein complex is present on the virion surface as highly glycosylated spikes. Cellular proteases cleave the env into gp120 and gp41. The glycoprotein complex is formed by heteromerization of the surface (SU) gp120 and the transmembrane (TM) gp41. The gp120/41 form a heterotrimeric spike structure which is embedded into lipid membrane of virions. The first step of infection process begins by binding of env complex to receptor and co-receptor on the surface of target cells. Cell tropism of the virus is mostly determined by binding affinity between chemokine CCR5 or CXCR4 receptors and gp120. HIV-1 isolates binds CCR5 of macrophages for entry hence called as R5 viruses; the primary CD4+ T-cells but not T-cell lines get infected by the HIV-1. X-4 tropic viruses infect T cell lines and primary T-cells by binding to CXCR4 receptors. R5X4 or “dual-tropic” strains can enter cells via either CCR5 or CXCR4 (Checkley Luttge et al., 2011; Arrildt Joseph et al., 2012).
1.6.8 GP120

High degree of variability was observed in env gene sequences from several HIV-1 isolates. gp120 comprises five variable domains (V1–V5) and constant domains (C1–C5). The Cys residues are conserved and present in abundance throughout gp120 and gp41. The disulfide bonds present in Cys residues give strength to the tertiary structure of env protein. Typically, there are 18 Cys residues in gp120. 9 disulfide bridges are formed by Cys residues. Two disulfide bonds are present between V1 and V2 and the larger loop is formed by two small loops; V3 and V4 loops are also delimited by a disulfide bond. The overall distribution of Cys residues and the number of these residues may affect env structure and antigenicity. N-linked glycans constitute 50% of molecular mass of gp120, O-linked glycans are present in relatively small amount. Nearly 20–35 N-linked glycosylation sites are present in gp120 and 3–5 N-linked glycosylation sites are there in gp41. The glycosylation sites mask env from
recognition by immune cells and also contribute to env folding, and binding of virus to the host cell surface (Putney Rusche et al., 1990; McKeating & Willey, 1989; Moore & Nara, 1991).

1.6.9 Glycoprotein 41 (TM)

The fusion of virus with host cell membrane is mediated by the gp41 transmembrane glycoprotein. The gp41 subunit is of 345 amino acid long protein. It is organized into three major domains: an ectodomain, a TMD and a C-terminal CT. The ectodomain has a fusion peptide at N-terminal hydrophobic domain, a polar region, and α-helical heptad repeat regions HR1 and HR2 (also known as N-helix and C-helix, respectively), and the membrane-proximal external region (MPER) rich in Trp. A disulfide linkage connects HR1 and HR2 within a hydrophilic loop. The interaction between HR1 and HR2 facilitates fusion. The fusion peptide is present inside the gp120/gp41 ternary complex. Binding of gp120 to CD4 and a co-receptor lead to conformational changes in the structure of gp41. This change exposes the peptide. The conformational change causes penetration of peptide into membrane of cell. Destabilization of membrane is followed by formation of pore into the membrane. A stable six-helix bundle is formed by HR1 and HR2 motifs. The cellular and viral membranes are brought close enough by the six helix bundle for fusion to occur (Postler & Desrosiers, 2013; Pan, Liu et al., 2010; Gelderblom Ozel et al., 1989).
1.6.10 Reverse Transcriptase

Reverse Transcriptase enzyme transcribes the RNA genome into DNA. This DNA is single stranded. A complementary strand of DNA is synthesized and DNA double helix is formed. In the host cell chromosome, the double stranded DNA is incorporated by integrase. A heterodimer of HIV-1 RT has subunits of 66 kDa (p66) and 51 kDa (p51). The N-terminal polymerase domain is relatively longer (440 residues) compared to the C-terminal RNase H domain (120 residues) Protease cleaves p66 into a 55 kDa protein p55 (Lightfoot Coligan et al., 1986).

Reverse transcription of HIV-1 genome is catalyzed by Reverse Transcriptase as described in following steps:

1. First, annealing of tRNA to primer binding site present in LTR initiates the process of synthesis of the negative strand DNA.

2. RNase H activity of RT leads to degradation of RNA of DNA-RNA duplex.

3. Following the degradation of RNA, first strand transfer takes place. The newly formed DNA joins to the 3’ end of the viral genome by identical repeat (R) sequences followed by synthesis of minus strand. All template RNAs are digested by RNase H except the polypurine tract that is used as primer for the synthesis of DNA.

4. The newly formed PBS DNA (second strand) is then joined to the PBS DNA from the first strand resulting in formation of second DNA strand leading to synthesis of DNA which is now double stranded having LTRs at both ends.

The HIV-1 RT, unlike mammalian DNA polymerases does not have a 3’-5’ exonuclease activity for proof-reading of newly synthesized strand. So, it fails to check the correct matching of base pairs during the synthesis of double stranded DNA from the RNA genome leading to mutations in virus (Cheng Dutschman et al., 1987; Sarafianos, Marchandet al., 2009).
Figure 1.10: RT structureRibbonrepresentation of HIV-1RT in a complex with nucleic acid. The fingerspalm thumbconnection, and RNaseHsubdomains of the p66subunit are shownin bluered, greenyellow, andorange, respectively. The p51 subunit is shown in darkbrown. Thetemplate andprimerDNA strands are shown in light gray and dark gray, respectively. (adapted from (SarafianosMarchand et al., 2009))

1.6.11 Protease (PR, p10)

HIV protease belongs to asparticproteases family (position 25 bears aspartic acid). It exists as homodimer. It has two identicalsubunits each 99 amino acids long (Navia, Fitzgerald et al., 1989; WlodawerMiller et al., 1989) The substrate-binding cleft constitute centre of the protein. This cleft recognizes different sequences for sequence in the gagand Gag-Pol proteins. How exactly the protease enzyme activated, is notknown, although it is reported that formation of dimer of Gag-Pol precursorproteins is required for activation of enzyme HIV protease plays important
role in viralmaturation and infectivity. One-third of the 99 aminoacids are known to show polymorphism. The proteaseenzyme is also encoded by polregion of HIV-1 genome. The viralgag and gag-pol precursor proteins are cleaved by PR to form infectious virus(KohlEminiet al., 1988). The structural viral proteins like capsid (p24, CA) matrix (p17, MA), and the p6 protein and the twospacer proteins p2 (SP1) andp1 (SP2), nucleocapsid (p7, NC), are encodedby gag precursorprotein (Fig. 1A). A −1 ribosomal frameshift event leads to formation of gag-pol polyprotein. The frame shifts occur at a frequency of about 5–10%. MA, CA,p2, NC, the transframe protein (TFP) are encoded by the polyprotein. It is also known to encode viral enzymes protease, Reverse Transcriptase and integrase(JacksPower et al., 1988).

Figure 1.11: HIV protease enzyme. (source :BonomiBarducci et al., 2010)
1.6.12 Integrase

One of the most important enzymes, integrase binds to the double-stranded viral DNA. It is known to facilitate integration of viral DNA into the cellular genomic DNA of the infected host. HIV DNA is usually inserted into active transcription units of host cell genome. Integration is the final step of HIV infection. IN is encoded by the HIV-1 pol gene, immediately 3’ of the RT/RNaseH coding sequence. IN enzyme is a 32kDa protein comprising of 288 amino acids. The Integrase protein has three structural and functional domains (Fig. 2). Cleavage of the gag-pol precursor polypeptide by protease leads to formation of functional integrase.

The N-terminal domain (NTD), the catalytic core domain (CCD) and the C-terminal domain (CTD) constitute functional domains of integrase. The monomers of integrase come together to form an IN homo-dimer. The functional tetrameric or higher order IN complexes are formed by further association of dimers. Each newly formed virion has roughly 50-100 copies of integrase enzyme (Asante-Appiah & Skalka, 1997; Shin, Taddeo et al., 1994; Engelman & Bushman et al., 1993).
Figure 1.12: HIV integration (Source(MarchandMaddali et al., 2009))
1.6.13 **Regulator of Viral Expression (Rev), p19**

The HIV-1 **rev response element** (RRE) is a highly structured RNA. It is ~350 base pair long. It is present in env coding regions of partially spliced and unspliced viral mRNAs. RREs help in translocation of HIV RNAs from nucleus to cytoplasm. Once exported into the cytoplasm, the RNAs are translated to form proteins required for virion packaging. A single promoter is present in HIV genome. Multiple reading frames and alternative splicing in HIV genome are used to produce 15 proteins from a single pre-mRNA species. A single 9 kilobase (kb) pre-mRNA is produced upon transcription. This pre-mRNA has multiple splice sites and nuclear retention signals.

During early events of life cycle of virus, the pre-RNA is completely spliced into 2kb messages. These are free of RRE. Translocation of spliced mRNAs from the nucleus to the cytoplasm occurs by nuclear export pathways. The HIV-1 Rev protein is encoded by a small 2kb message. Rev has nuclear localization sequence which causes migration of Rev into the nucleus. This phase of the virus is not dependent on either Rev or RRE (Cullen & Malim, 1991).

During late stages of virus life cycle, unspliced (9kb) or partially spliced (4 kb) messages containing the RRE encode for the viral proteins. RNAs with introns are retained in the nucleus as they have sequences for retention and splicing. In nucleus, these RNAs are further spliced or degraded. However, once Rev is produced, the longer messages can be exported to the cytoplasm via a Rev-dependent export pathway. Multiple Rev molecules assemble on RRE and initiate translocation of the RNAs. The proteins exportin-1 (XPO1/CRM1) and Ran-GTP are deposited once the Rev-RRE complex is formed. Nuclearexport sequence present in Rev induces recruitment of this export machinery. Translocation of Rev-RRE-Xpo1/RanGTP complex to the cytoplasm initiates translation of RNAs to form structural and function proteins (Wong-Staal & Haseltine, 1992).
1.6.14 Transactivator protein (Tat)

Tat is a small (14 kDa) protein. It is essential for viral replication and progression of AIDS. Tat gene is transcribed early in the replication cycle from multiply spliced transcripts. Tat protein is not packaged into the virions. It exists in two different length forms: 86-87 residues long form, predominantly existing in subtype B isolates and 99-101 residues long form, predominantly found in clinical isolates from all HIV subtypes except subtype D. It is synthesized from an mRNA consisting of two coding exons. The full length, two exon form is 101 amino acids in length. Amino acids 1–72 and (in most strains of HIV-1) are encoded by first exon while the second exon encodes amino acids 73–101.

The most important role of Tat is transactivation of transcription. TAR RNA is encoded downstream of HIV LTR. Joining of Tat to TAR is mediated by the basic domain of the protein. Peptides derived from Tat showed high affinity for binding to TAR RNA. Tat functions as an HIV gene regulator. Apart from gene regulator, Tat has been reported to have role as immunosuppressor. Virus replication in quiescent T-cells can be initiated by Tat. Tat protein also acts as chemokine analog extracellularly. Tat is known to activate the expression of certain chemokine receptors on lymphocytes and monocytes/macrophages, thereby facilitating HIV-1 entry. It is found that cystein rich residues are important for such diverse functions of Tat (Cullen 1990; Karn & Graeble, 1992).
1.6.14.1 Role of Tat in viral transcription

Tat performs the most critical step of the viral life cycle – transcription from the LTR promoter. In the absence of Tat, transcription can initiate from the LTR but RNA polymerase II gets stalled on the viral template resulting in truncated transcripts. These truncated transcripts cannot support viral replication. Tat mediates the function of transcription elongation to synthesize full length viral transcripts.

1.6.14.2 TAR-dependent Tat transactivation

Tat binds to the transactivation response element (TAR), an initial transcript of the viral genome from +1 to +59 nucleotides that folds into a stemloop structure. After binding to TAR RNA, Tat by directly interacting with cyclin T1 recruits pTEF-b complex (a complex of CyclinT1 and CDK9). The recruitment of pTEF-b removes the transcription elongation block posed by DRB sensitivity inducing factor (DSIF) and a negative elongation factor (NELF).

The kinase partner of pTEF-b CDK9 hyperphosphorylates the C-terminal domain of RNA polymerase II at Ser2 and Ser5. Along with the phosphorylation of the RNAPII, kinase activity of pTEF-b also acts on DSIF and NELF (Peterlin and Price, 2006; Zhou and Yik 2006). All these phosphorylation switches “on” the transcription elongation. NELF and DSIF phosphorylation removes the suppression on viral transcription elongation releasing RNAPII from promoter proximal pausing.

Phosphorylation of RNAPII CTD turns it into a processive elongating polymerase. Intense research in this area has delineated the complex intricacies involved in the HIV-1 transcription. pTEF-b exists in mainly two forms in the cell. Majority of it is found in the form of inactive pTEF-b and small pool of core active pTEF-b also exists. The inactive form termed 7SK snRNP contains pTEF-b, 7SK snRNA, 7SK capping enzyme, MEPCE, LARP7, HEXIM1/2 and SART3 (Nguyen et al., 2001; Yanget al., 2001; Yik et al., 2003; Jeronimo et al.,
Tat binds to the TAR element and then recruits pTEF-b by associating with one of its subunit Cyclin T1 (Weiet al., 1998; Karn, 1999; Price, 2000; Bannwarth and Gatignol, 2005). Tat also binds to 7SK snRNA component of inactive pTEF-b complex and displaces HEXIM1 from the complex liberating the active form (Muniz et al., 2010). This active pTEF-b then phosphorylates RNAPII CTD, NELF and DSIF thus preparing them for the elongation step (Ivanovet al., 2000; Zhou et al., 2000; Fujinaga et al., 2004; Chen et al., 2009). Active pool of pTEF-b is called the Super elongation complex (SEC). SEC contains pTEF-b that needs no activation by Tat and is readily delivered to the paused RNAPII. In SEC, pTEF-b is present along with transcription elongation factors and coactivators like AF9, AFF4 ENL, ELL and PAF1 complex (He et al., 2010; Lin et al., 2010; Sobhian et al., 2010; He and Zhou, 2011).

Thus, the role of Tat in transcription elongation is fairly clear, however, its role in transcription initiation is debatable although evidence exists in literature about its involvement in transcription initiation through binding to RNAPII, TFIIB, TFIID and TFIH (Kashanchi et al., 1994; Wu-Baer et al., 1995; Parada and Roeder, 1996; Veschambre et al., 1997). Tat is known to induce chromatin remodelling by recruiting an plethora of histone acetyl transferases (HATs) and chromatin remodelling complexes at the LTR promoter. HATs that are recruited to the HIV-1 promoter are p300/CBP, CAF, GCN5 and Tip60. As mentioned earlier in this chapter, remodeling is needed at nuc-1 of the LTR. Many complex regulations operate at this step to accomplish this remodeling.

HATs recruited by Tat open up the chromatin at nuc-1 so that the elongating polymerase can extend the transcripts rather than getting stalled at the nucleosome (Hottiger and Nabel, 1998; Benkirane et al., 1998; Marzio et al., 1998; Col et al., 2001; Lusic et al., 2003). Tat engages all the five HATs (p300/CBP, P/CAF, GCN5, Tip60 and TAFII250) to establish the virus in the host cell. Tat interacts with and recruits p300/CBP to the LTR promoter (Marzio et al., 1998; Wong et al., 2005) where it acts as a coactivator of Tat. Tat actually forms a ternary complex with
p300 and PCAF that increases its affinity for pTEF-b/CDK9 complex (Benkirane et al., 1998; Deng et al., 2001). Apart from its histone acetylation activity, p300 acetylates viral integrase and Tat. This acetylation at Lys 50 and Lys 51 of Tatis essential for its activity (Ott et al., 1999). Acetylated integrase regulates integration. Vpr also requires help of p300 for its transcriptional activation activity. Thus, p300 participates at multiple stages of HIV-1 lifecycle. PCAF is another HAT that participates extensively in HIV lifecycle. PCAF interacts with Lys 50 acetylated Tat to serve as a coactivator of HIV transcription. Like p300, PCAF also activates transcription from LTR promoter by acting in synergy with Tat protein (Dorr et al., 2002; Bres et al., 2002). GCN5 also acetylates Tat at same residues K50 and K51 as p300 and mediating same function of LTR activation (Col et al., 2001). Tat also interacts with TAFII250 (Weissman et al., 2001) to repress transcription from TAFII250 dependent MHC-I genes (Weissman et al., 1998). Tip60, an HAT originally identified as a Tat interacting cellular protein also assists Tat in transactivation of the LTR promoter (Kamine et al., 1996).

Tat also interacts with chromatin remodeling complexes like SWI/SNF that use the energy from the hydrolysis of ATP to remodel the nucleosomes (Mahmoudi et al., 2006). Integrase interacting protein (INI-1 or SNF5), BRM and BRG1 are the core subunits of SWI/SNF complex, which participate in expression of many eukaryotic genes. Nuc-1 at the LTR requires remodeling via SWI/SNF to initiate transcription. INI-1 and BRG1 interact with Tat to recruit SWI/SNF remodeling complex at Nuc-1 and to activate the transcription from the LTR promoter (Treand et al., 2006; Bukrinsky, 2006; Mahmoudi et al., 2006).

Tat mediated LTR activation also removes the repressive factors occupying the nuc-1 at the LTR like HDAC1 (histone deacetylase 1) which recruited to the HIV promoter by a variety of cellular factors - YY1, LSFe-Myc, SP1, TIP2, CBF1 and NFκB p50 - creating an expressive chromatin environment at the LTR (Coul et al., 2000; Imai and Okamoto, 2006; Williams et al., 2006; Jiang et al., 2007; Marban et al., 2007; Tyagi and Karn, 2007).
1.6.14.3 TAR independent Tat transactivation

As mentioned earlier, Tat primarily mediates the transactivation of the LTR by binding to the TAR RNA element, but reports also exist in literature regarding Tat functioning independently of the TAR element to activate the LTR promoter. TAR-independent transactivation is well known in glial cells. TAR-independent mechanism requires the NFκB and SP1 sites of the LTR. (Bagasra et al., 1992; Taylor et al., 1992; Taylor et al., 1994; Kim and Panganiban, 1996; Niikura et al., 1996; Yang et al., 1997).

1.6.14.4 Role of Tat in cellular genemodulation

To make the cellular environment conducive for virus propagation, Tat functions to modulate the expression of cellular genes at transcriptional, post-transcriptional, and translational levels (Buonaguro et al., 1992; Zauli et al., 1995; Li-Weber et al., 2000; Lim and Garzino-Demo, 2000; Kota-Gomez et al., 2002; de la et al., 2002; Coiras et al., 2006). At transcriptional level, there are reports in literature about both TAR-dependent (Buonaguro et al., 1994) and independent mechanisms (Scala et al., 1994; Thatikunta et al., 1997) of Tat-mediated cellular genemodulation.

1.6.14.5 Role of Tat in viral splicing

Tat participates in viral splicing by interacting with cellular splicing cofactors like Tat-SF1 (HIV-1 Tat specific factor 1), CA150 (Jablonskiet al., 2010). These two factors are key participants in splicing and transcription (Zhou and Sharp, 1996; Sune and Garcia-Blanco, 1999; Smith et al., 2004; Sanchez-Alvarez et al., 2006). Acetylated Tat has been shown to bind to splicing regulator p32 which is a cofactor of splicing factor ASF/SF2 (Berro et al., 2006). It selectively regulates viral splicing by shutting down levels of one mRNA and increasing the levels of the other.
1.6.14.6 Role of Tat in reverse transcription

Studies with HIV-1 viruses lacking tat genes showed inefficient reverse transcription initiation, thus illustrating role of Tat in reverse transcription (Harrich et al. 1997). Later studies have shown that Tat protein regulates reverse transcription mainly at two stages – it helps in the placement of tRNAlys primer at the primer binding site of the HIV genome. At later stages of viral replication, it represses RT activity and thus the DNA polymerisation (Kameoka et al., 2001; Kameoka et al., 2002).

1.6.15 Viral infectivity factor (Vif)

Vif is a 192 residue (22.5 kDa) protein required for the production of highly infectious mature virions. Cells infected with HIV carrying a mutant Vif gene produce normal virions but are 1000-fold less infectious than the wild type virions. Host restriction factors like APOBEC3G and APOBEC3F are cytidine deaminases, which restrict replication of Vif deficient HIV-1.

Both of these host proteins introduce G to A hyper mutation in HIV-1 genomic DNA which blocks the viral replication. The viral Vif protein counteracts their activity. Vif is also shown to prevent encapsidation of these two factors. Additionally, it was observed that translation of APOBEC3G could be blocked by Vif. Several studies have shown that interaction of vif with APOBEC3G leads to polyubiquitination and degradation of APOBEC3G by protease (Niewiadomska & Yu, 2009; Wissing Galloway et al., 2010; Batisse Guerrero et al., 2012).

1.6.16 Viral Protein R (Vpr), p12

Vpr is a 96 amino acids 14-kDa protein. Vpr has been shown to play role in mediating nuclear import of the HIV-1 pre-integration complex. It facilitates replication of virus in cells not undergoing division such as macrophages. A hydrophobic central core domain, with three α-helices (H1 a.a17–33, H2 a.a38–50 and H3 a.a55–77) form vpr. The helices are linked to each other by loops. Two flexible N- and C-terminal domains
are present surrounding the helices. N terminal domain has positive charge and C terminal domain is negatively charged[10].

By contrast with other HIV- auxiliary proteins, Vpr is present in high abundance. High copy number of vpr is found in virus particles. The presence of Vpr in cytoplasm of newly infected cells indicates that it may have specific roles in the initial post-entry events of viral replication. Vpr is also known to induce G2 cell cycle arrest leading to apoptosis in cells undergoing cell division resulting in immune dysfunction. In cytoplasm, Vpr inhibits transcriptional activators of inflammatory cytokines/proteins and act as immunosuppressant. HIV-2 comprises a Vpr protein and a Vpx protein (Viral Protein X). In HIV-2, functions of Vpr are split between vpr and vpx. Vpr has a role in cell cycle arrest and the Vpx protein implicated in nuclear import in HIV-2 (Zhao Li et al., 2011; Planelles & Benichou, 2009).

1.6.17 Viral protein U (Vpu), p16

It is an integral transmembrane protein. It is inserted into membranes of infected cells. The protein consists of an N-terminal hydrophobic domain and a hydrophilic C terminal domain. The N terminal domain functions both as signal peptide and membrane anchor. The hydrophilic C-terminal domain protrudes into the cytoplasm. Two highly conserved serine residues are present in the cytoplasmic domain of Vpu. These residues are phosphorylated by the ubiquitous protein kinase CK-2. Phosphorylation is thought to be important for at least one of the biological functions of Vpu (Subbramanian & Cohen, 1994).

1.6.18 Negative factor (Nef)

Nef, is a 27–35 kDa protein. It is myristoylated. The nef gene is present at 3’ end of HIV genome which partly overlaps with 3’ end of env gene and 5’ end of 3’ LTR. It expresses abundantly during the initial events of replication. Following post translational events like phosphorylation and myristoylation, nef is transported to cell membrane. Nef plays several important functions like suppression of immune cells,
MHC-I complex, activation of PAK2. It is also known to enhance infectivity (Trono & Wang, 1997).

Nef has ~70 residues of N-terminal and approximately 120-residue long core domain which is conserved. The first 22 amino acids of N-terminal are positively charged. These serve as membrane anchor.

The core domain exhibits tertiary structure. The α-β domain of core has central anti-parallel β sheet of four strands (βA–βD). Two anti-parallel α helices (αA and αB) are attached to betasheets. Two short α helices (αC and αD) are also present at C terminal (Luo & Garcia, 1997; Ahmad & Venkatesan, 1988).

HIV-1 Nef has been characterized structurally using NMR and X-ray crystallography. Although the structure of full-length Nef is not known so far, but structural data is available for its component fragments covering the whole polypeptide chain. HIV-1 Nef possesses a genetically diverse and structurally flexible N-terminal arm of ~70 residues followed by a well-conserved and folded core domain of ~120 residues. The N-terminal arm of Nef (residues 1–70) contains many positive charges in the first 22 amino acids that serve, together with a myristoyl group attached to residue G2, as membrane anchor (Welker et al., 1998).

The core domain is the only part of Nef that adopts a stable tertiary fold. This fragment has been characterized for HIV-1 Nef in its free form, by both NMR (Grzesiek et al., 1996a) and crystallography. It was also crystallized in Src homology (SH) 3 domain bound form (Arold et al., 1997; Lee et al., 1996). The core of Nef forms an α-β domain in which a central anti-parallel β sheet of four strands (βA–βD) is flanked N-terminally by two long anti-parallel α helices (αA and αB) and terminally by two short α helices (αC and αD).

A proline rich sequence (P69xxPxxPxxP78) is located upstream of αA. Nef also interacts with the SH3 domain of the guanine nucleotide exchange factor Vav (Renkema, et al., 1996). The core domain also mediates oligomerization of HIV-1 Nef (Arold et al. 2000). Nef dimers and trimers have been observed in vivo and in vitro but their role remains to be elucidated. Oligomerization is commonly used to trigger events in cellular signalling and endocytosis.
The quarternary structure of Nef might help to functionally connect cellular factors or to increase Nef’s affinity for oligomeric targets. Nef has two large flexible regions that, together, comprise 50% of the polypeptide chain: the N-terminal arm and a 30-amino-acid loop (residues 148–178) that projects out of the core domain. This loop contains three binding motifs, di-Leu internalization motif (E/D160xxxLL165) (Greenberg et al., 1998), diacidic sequence 155EE and an acidic cluster D174-E178 which are involved in linking Nef with endocytic pathways. The diacidic sequence (155EE) was also found to be essential for Nef and β-COP association (Piguet V. et al., 1999) and Cterminally located acidic cluster (D174–E178) is also required to colocalize Nef with both AP (adapter protein) complexes (Mangasarian et al., 1999) and v-ATPase.

1.7 Anti-HIV drugs and their targets

To date, FDA has approved >20 drugs for ART treatment. T20 and Maraviroc are the only two drugs that act at entry step. T20 (Enfuvirtide) target gp41 and thus block fusion. CCR5 antagonism is displayed by Maraviroc. At present Maraviroc is the only anti-HIV drug that targets host cell protein.

The enzymes RT, PR and IN of HIV-1 are targeted by other ART drugs. Several drugs have been reported that target RT and PR. These drugs are given in combination as a standard treatment regime.

Depending on mode of action, RT inhibitors are subdivided into the nucleoside/nucleotide-analog RT inhibitors (NRTIs) and non-nucleoside/nucleotide -analog RT inhibitors (NNRTIs).

PR inhibitors (PIs) act on proteas enzyme and block its function. Recently Raltegravir, an Integrase Inhibitor was approved in 2007 and is widely used clinical practice (LeGrice, 2012). The antiretroviral drugs helped to extend life of patientssignificantly. The highly active antiretroviral therapy (HAART) comprises a combination of three to four drugs in therapeutic regimens. These classes of anti-HIV drugs are discussed below in detail.
Figure 1.13: Various steps targeted by anti-HIV-1 drugs. (Source: [Hazuda Iwamoto et al., 2009])
1.7.1 Entry inhibitors

As the name suggest, this class comprises drug that inhibit viral entry. The early event of life cycle begins at entry step. The process of entry involves several steps and is a complex event. Several compounds have been identified that can interfere in the entry of virus into the cell. These drugs act on any step during entry such as attachment of gp120 to CD4 receptor and binding, co-receptor binding and fusion (Meanwell & Kadow, 2003).

1.7.1.1 Drugs blocking the gp120–CD4 interaction

Gp120 is an essential component of entry process. Though heterogeneity in gp120 protein is observed, the binding site of gp120 is conserved among virus strains. Hence, this conserved region is of interest to target virus. Soluble CD4 (sCD4) had been evaluated for its efficacy to inhibit attachment. Recombinant sCD4 have been shown to be effective against laboratory adapted strains but are found to have low to moderate activity against primary HIV isolates. In vivo data indicated that sCD4s have poor efficacy (Simmons, Reeves et al., 2000).

Polyanionic compounds have divergent structure and size. Carraguard, PRO 2000, sulfates (dextran and cellulose sulfate) have been shown to block interaction of gp120 to CD4 thereby inhibiting viral entry. Early reports using these compounds raised possibility of potential use of these inhibitors. Low efficacy and toxicity in vivo have been reported for these compounds when administered systemically. Some of these compounds are being evaluated in Phase II and III studies as microbicidal candidates for preventing sexual transmission of HIV. BMS-488043 is another investigational new drug. BMS488043 binds to gp120 and blocks attachment of the virus to CD4 receptor of cells. In phase 1 study this molecule has shown promising results with good bioavailability and safety efficacy (Kuritzkes, 2009).
1.7.1.2 HIV-1 gp41 inhibitors

During fusion, HIV-1 gp41 plays an essential role. N-HR coiled coil structure has prominent hydrophobic cavity on its surface. This region has been evaluated as a target. The domains present on N and C terminal regions of gp41 were then studied as a target for fusion inhibitors (Pan Liu et al., 2010).

In early 1990s, Jiang et al. reported isolation and identification and anti-HIV properties of SJ-2176 peptide from the HIV-1 gp41 C-HR region. This peptide inhibits HIV-1 infection in a low nM range (Pan Liu et al., 2010; Jiang & Lin, 1995). An analogue of this peptide, Enfuvirtide (T20) was developed by Trimeris Ltd jointly with Roche. T20 binds to the N-HR region of gp41 (Pan Liu et al., 2010; Cervia & Smith, 2003). T-1249 belongs to second generation of these inhibitors.

It was developed by Trimeris and Roche. T-1249 is a more potent HIV-fusion inhibitor. T-1249 is active against most of T20 resistant strains. However, the clinical development was discontinued due to safety and efficacy issues (Pan Liu et al., 2010; Eron Gulick et al., 2004; Martin-Carbonero, 2004). Betulinic acid is a terpenoid isolated from the leaf extract of Syzigium claviflorum. It was found to possess potent activity against HIV-1 (IC50 of 1.4 uM). Several derivatives of betulinic acid have been developed. These agents mostly interfere with functions of gp41 during fusion (Aiken & Chen, 2005b).

Side effects/adverse effects: lymphadenopathy and bacterial pneumonia, Septicemia
1.7.1.3 Drugs blocking the gp120–co-receptor interaction

During infection, the gp120 envelope protein of HIV-1 binds to the CD4 receptor. The binding initiates conformational changes and the co-receptor-binding sites get exposed. Further interaction between gp120 and one of the two chemokine receptors, CCR5 or CXCR4, results in the final changes in structure that are required for efficient membrane fusion. HIV-1 gp41 unfolds, and gets embedded in the cell membrane and fusion process completes. The CCR5 or CXCR4 receptors functions as one of the HIV-1 co-receptors.

These co-receptors can be considered as important target for developing novel anti-HIV agents. Ligands for CCR5 like CCL3 (MIP-1α), CCL4 (MIP-1β), and CCL5 (RANTES) have been shown to inhibit HIV-1 infection. The antiviral effect of these chemokines is mainly due to blocking of en binding to CCR5. This may lead to internalization of the co-receptor. Chemokine ligands exert undesirable agonist effects. RANTES derivatives, like AOP-RANTES, NNY-RANTES, and PSC-RANTES have been shown to possess affinity for CCR5 and compete with gp120. These derivatives of RANTES act as competitive antagonists of HIV infection (Schols 2004; Hartley & Offord, 2005).

Many small-molecule acting like antagonists have been developed. 3 agents have been studied widely. Aplaviroc (GW873140) was tested in phase IIb studies. However, adverse effects like idiosyncratic hepatotoxicity resulted in discontinuation in 2005 (Crabb 2006; Lalezari Thompson et al., 2005). Vicriviroc is currently being evaluated in phase III clinical trials (Strizki Tremblay et al., 2005; Caseiro Nelson et al., 2012). FDA, in 2007 has approved Maraviroc (UK-427857) for treating infection of drug resistant HIV strains (Maeda Das et al., 2012).

T-22, ALX40-4C, T-140, and T-134 are polypeptides that mimic natural ligands of CXCR4 or CXCL12. These peptides display high specificity of binding to CXCR4 receptor and prevent gp120 from binding (Fujii & Tamamura, 2001). AMD3100 is a bicyclamalog. It is shown to possess strong inhibitory activity against CXCR4 tropic (X4) strains of HIV in vitro. However AMD3100 could not succeed in
Several strategies have been reported in the literature for the inhibition of HIV infection and replication. They are listed below in Table 1.3

Table 1.3: Strategies reported in the literature for inhibition of HIV-1 infection or replication.

<table>
<thead>
<tr>
<th>Step in viral life cycle</th>
<th>Mechanism of inhibition</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viral entry</td>
<td>ASRs that target CCR5 and CXCR4 inhibit their gene expression.</td>
<td>Lamothe et al, 2000</td>
</tr>
<tr>
<td>Pre integration</td>
<td>Psi-gag and U3-5 5' UTR targeted ASRs are copackaged with the genomic RNA and inhibits RT in incoming virions.</td>
<td>Veres et al, 1996</td>
</tr>
<tr>
<td>Nuclear export of viral transcripts</td>
<td>Rev targeted ASRs inhibit interaction of Rev with viral transcripts.</td>
<td>Cohli et al, 1994</td>
</tr>
<tr>
<td>RNA Decoys</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viral transcription</td>
<td>RNA decoy that target Tat inhibit Tat regulated transcription.</td>
<td>Browning et al, 1999</td>
</tr>
<tr>
<td>Ribozymes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viral transcription</td>
<td>Tat targeting Ribozymes inhibit Tat regulated transcription.</td>
<td>Sun et al, 1995</td>
</tr>
<tr>
<td>Nuclear export</td>
<td>Ribozymes that target RRE and Rev encoding regions cleave viral RNAs</td>
<td>Zhou et al, 1994</td>
</tr>
<tr>
<td>Translation of viral Proteins</td>
<td>Ribozymes that target U5 cleaves off the 5' cap structure localized on viral mRNAs</td>
<td>Yu et al, 1993</td>
</tr>
<tr>
<td>Viral assembly</td>
<td>Psi sequence recognizing Ribozymes cleave HIV-1 RNAs before packaging.</td>
<td>Sun et al, 1994</td>
</tr>
<tr>
<td>Viral assembly</td>
<td>Ribozymes that cleave Gag encoding transcripts inhibit the formation of multimeric gag and Env complexes.</td>
<td>Sarver et al, 1990</td>
</tr>
<tr>
<td>Step in viral life cycle</td>
<td>Mechanism of inhibition</td>
<td>Reference</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Viral assembly</td>
<td>Ribozymes that target gp120-encoding region cleave different conserved sequences in gp120 sequence.</td>
<td>Yamada et al, 1994</td>
</tr>
<tr>
<td>Virus release</td>
<td>Ribozymes that cleave Nef encoding region, inhibit down modulation of CD4 and MHC I.</td>
<td>Larsson et al, 1996</td>
</tr>
</tbody>
</table>

**RNA aptamers**

| Pre integration of viral RNA transcripts | RNA aptamers that target RT enzyme acts as template analogues and inhibit reverse transcription. | Joshi et al, 2002 |
| Nuclear export of viral RNA transcripts | RNA aptamers that target Rev protein interfere with its function and inhibit the export of unspliced and singly spliced viral mRNAs. | Kjems et al, 2000 |

**siRNAs**

<p>| Viral entry | CCR5 and CXCR4 siRNAs decrease the expression of these receptors and impair gp120-chemokine co-receptor interaction. | Martinez et al, 2002 |
| Viral entry | siRNAs that target CD4 mRNAs decrease the levels of cell surface expression of CD4. | Novina et al, 2002 |
| Viral entry | siRNAs that target CD4 binding domain of gp120 inhibits the CD4-gp120 interaction | Park et al, 2002 |
| Pre integration | Viral LTR, Vif and Nef siRNAs destroy the viral genomic RNAs. | Das et al, 2004 |
| Pre integration | siRNAs that bind RT encoding region inhibits RT gene expression. | Surabhi et al, 2002 |
| Pre integration | siRNAs that target CA encoding region mediate the cleavage of pre-spliced viral RNA in the cytoplasm and prevent integration | Novina et al, 2002 |
| Viral transcription | siRNAs that target viral Tat encoding region inhibit Tat mediated transactivation. | Boden et al, 2004 |
| Viral transcription | NFκB p65 siRNA inhibits NFκB mediated activation of viral transcription. | Coburn et al, 2002 |
| Viral transcription | siRNAs that target 3′ UTR terminus of the Nef gene mediate the cleavage of all spliced and unspliced RNA produced from the provirus. | Capodici et al, 2002 |</p>
<table>
<thead>
<tr>
<th>Step in viral life cycle</th>
<th>Mechanism of inhibition</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viral transcription</td>
<td>siRNAs that target Rev transcript inhibit Rev mediated export of unspliced and singly spliced RNA transcripts.</td>
<td>Lee et al, 2002</td>
</tr>
<tr>
<td>Nuclear export of viral RNA transcripts</td>
<td>siRNAs that target Gag and Nef encoding regions mediate the cleavage of both spliced and unspliced produced from the provirus.</td>
<td>Jacque et al, 2002</td>
</tr>
<tr>
<td><strong>shRNA/miRNA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viral translation</td>
<td>shRNA that target Nef encoding region blocks the Nef mRNA stability and translation</td>
<td>Omoto et al, 2004</td>
</tr>
<tr>
<td><strong>Transdominant negative proteins (TNPs)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nuclear export of viral RNA transcripts</td>
<td>Sam68 mutants interfere in Rev function and thereby inhibit the export of viral unspliced and singly spliced mRNAs from nucleus to cytoplasm.</td>
<td>Reddy et al, 1999</td>
</tr>
<tr>
<td>Viral assembly</td>
<td>Tsg101 mutants inhibit the transport of the Gag polyprotein into multivesicular bodies.</td>
<td>Carter et al, 2002</td>
</tr>
<tr>
<td>Viral assembly</td>
<td>INI1 mutants interact with integrase domain of the Gag-Pol polyprotein and interfere with multimerization of Gag and Gag-Pol.</td>
<td>Yung et al, 2001</td>
</tr>
<tr>
<td>Viral maturation</td>
<td>Viral protease mutants inhibit the activation of viral protease.</td>
<td>Gottfredsson et al, 1997</td>
</tr>
<tr>
<td><strong>Fusion proteins (FPs)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre integration</td>
<td>SFv-nuclease associated with preintegration complexes cleaves proviral DNA after integration.</td>
<td>Tomaras et al, 2001</td>
</tr>
<tr>
<td>Viral transcription</td>
<td>Tat-Rnase H fusion protein interact with TAR and cleave all the viral transcripts.</td>
<td>Melekhovets et al, 1996</td>
</tr>
<tr>
<td>Cellular proteins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viral entry</td>
<td>Soluble CD4, a truncated form of cellular protein CD4 inhibits fusion event by binding to gp120 and thereby extending the distance to gp41 protein.</td>
<td>Eckert et al, 2001</td>
</tr>
<tr>
<td>Step in viral life cycle</td>
<td>Mechanism of inhibition</td>
<td>Reference</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td><strong>Intracellular antibodies (sFvs)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viral entry</td>
<td>gp120 targeting sFVs inhibit gp120-CD4 interaction.</td>
<td>Marasco et al, 1993</td>
</tr>
<tr>
<td>Viral entry</td>
<td>sFVs that target gp41 induced pre-hairpin intermediate inhibit the interaction of fusion peptide and cell membrane.</td>
<td>Lee et al, 2004</td>
</tr>
<tr>
<td>Reverse Transcription</td>
<td>HIV RT sFVs inhibit the reverse transcription.</td>
<td>Maciejewski et al, 1995</td>
</tr>
<tr>
<td>Viral transcription</td>
<td>Tat sFVs restrain Tat protein in the cytoplasm.</td>
<td>Mhashilkar et al, 1995</td>
</tr>
<tr>
<td>Nuclear export of viral RNA transcripts</td>
<td>Rev sFVs locate the Rev protein in the cytoplasm.</td>
<td>Duan et al, 1994</td>
</tr>
<tr>
<td>Viral assembly</td>
<td>sFVs that target CD4 binding region of gp120 interacts with Env and restrain it in endoplasmic reticulum.</td>
<td>Chen et al, 1994</td>
</tr>
<tr>
<td><strong>Monoclonal antibodies (Mabs)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viral entry</td>
<td>CCR5 and CXCR4 receptor Mabs inhibit the interaction of SU with co-receptors.</td>
<td>Dragic et al, 2001</td>
</tr>
<tr>
<td>Viral entry</td>
<td>Mabs that target extracellular loop on CCR5 inhibits the fusion and viral entry events.</td>
<td>Olson et al, 1999</td>
</tr>
</tbody>
</table>
1.7.2 Reverse Transcriptase Inhibitors

The N(t)RTIs and the NNRTIs interfere with reverse transcription process.

1.7.2.1 Nucleoside/Nucleotide Reverse Transcriptase Inhibitors

In 1985, 3-azidothymidine (AZT, Zidovudine) with \textit{in vitro} anti-HIV activity was identified as the first nucleoside inhibitor by Samuel Broder’s group at the National Cancer Institute together with collaborators from Burroughs-Welcome company (Mitsuya, Weinhold et al., 1985). Clinical studies using AZ have shown the first proof of concept that such agents can be used to treat infection in patients infected with HIV. This earmarked the development of novel therapeutic agent for treating HIV infected patients. In 1986 (Furman et al., 1986) it has been shown by Furman et al that AZT is metabolized to form triphosphate form and this phosphorylated molecule actually inhibit Reverse Transcriptase (RT).

Reverse Transcriptase is very important enzyme of HIV. It mediates formation of proviral DNA from RNA. Till date USFDA has approved 7 nucleosides and 1 nucleotide Reverse Transcriptase inhibitor for treating infection of HIV-1. eg Lamivudine (3TC), Abacavir (ABC), Didanosine (ddI), Stavudine (d4T), Tenofovir Disoproxil Fumarate [TDF] Zalcitabine (ddC) and Emtricitabine (FTC) (Le Grice 2012).

NRTIs have been derived from endogenous 2-deoxy-nucleosides and -nucleotides. These compounds are inactive in their native forms. Once in the cell, NRTIs get phosphorylated by cellular kinases and phosphotransferases and converted into deoxynucleoside triphosphate (dNTP) analogs. The dNTP analogs are the active form of NRTIs which inhibit viral infection. The triphosphate (TP) forms of NRTIs get incorporated by HIV RT easily compared to their corresponding endogenous dNTPs. The incorporation of dNTPs lead to termination of transcription activity and thus proviral DNA synthesis is halted. Thus NRTIs inhibit viral replication at early stage prior to integration step. Presently NRTIs form the backbone of ART. The highly active antiretroviral therapy (HAART), comprises 3 or more HIV inhibitors, mostly 2
NRTIs in combination with a NNRTI, protease inhibitor or, most recently, Integrase Inhibitor (Cihlar & Ray, 2010).

Patients receiving HAART reported gastrointestinal problems, headache, and fatigue, vomiting etc. The gastrointestinal complaints are easily treated symptomatically. Long-term side effects associated with use of NRTs include myelotoxicity, lactate acidosis, polyneuropathy and pancreatitis. Patients also suffer from lipoatrophy. Mitochondrial degeneration was also reported (Koczor & Lewis, 2010; Bogoch & Walmsley, 2009).

### 1.7.2.2 Nonnucleoside/nucleotide Reverse Transcriptase Inhibitors

NNRTIs bind to a hydrophobic pocket on the RT adjacent to the substrate-binding site. The binding of compound to RT leads to loss of activity of enzyme. The first NNRTIs to be identified were the 1-(2-2-hydroxyethoxymethyl)-6-(phenylthio)thymine (HEPT) and tetrahydroimidazo[4,5,1-3,4]benzodiazepin-2(1H)-one and -thione (TIBO) compounds. NNRTIs are active as such unlike NRTIs which require phosphorylation. The main characteristic of NNRTIs is the specificity with which they bind to RT. They exhibit specific activity towards HIV Reverse Transcriptase. NNRTIs are not active against other retroviruses.

Though NNRTIs differ in structure and chemical composition, they all have affinity for the same site of binding to the RT. This binding site is located in the palm domain of the p66 subunit of the heterodimeric protein between the $\beta 6-\beta 10-\beta 9$ and $\beta 12-\beta 13-\beta 14$ sheets. The binding site is at the basis of the $\beta 4-\beta 7-\beta 8$ sheet, at a distance of approximately 10Å from the catalytic site of the enzyme. This pocket consists of amino acids of the p66 subunit; the amino acids are hydrophilic or hydrophobic, aromatic innature and two amino acids of the p51 subunit. The binding pocket is lipophilic.

Nevirapine, Efavirenz and Delavirdine have been approved by US FDA for therapeutic use. Etravirine (TMC-125) is a new NNRTI that has potent *in vitro* anti-HIV activity.
against wildtype HIV-1 isolates, as well as isolates with key NNRTI resistance mutations (de Bethune 2010).

HIV rapidly develops resistance to NNRTIs. If virus becomes resistant to one of the NNRTIs, it shows resistance to all the compounds in the class.

Adverse effects: Liver problems, Gynecomastia

1.7.3 Protease Inhibitors

HIV protease is an important enzyme in the life cycle of virus. Protease cleaves viral gag and gag-pol precursor proteins. This cleavage leads to formation of mature infectious virus particles. The structure of HIV protease and its substrates have been studied widely. Specific protease inhibitors have been developed (PIs). These inhibitors have high affinity for binding viral protease but compared to natural substrates, these ligands require more space.

Saquinavir, Darunavir, Ritonavir, Atazanavir, Indinavir, Amprenavir, Tipranavir, Lopinavir, and Nelfinavir, have been approved by FDA for treatment of AIDS. Saquinavir was the first protease inhibitor approved by US FDA in 1995. Except Tipranavir, all are competitive peptidomimetic inhibitors. They mimic the natural substrate of the viral protease (Mastrolorenzo Rusconi et al., 2007).

A hydroxyl ethylene core present in these inhibitors is resistant to action of protease and thus protect from proteolytic cleavage. As a central scaffold, dihydropyrole ring is present in Tipranavir (Qiu & Liu, 2011). Ritonavir and Saquinavir have been used for treating infection. The bioavailability and half life of PIs have improved when it was found that Ritonavir helps in reducing the metabolism of concomitantly administered PIs through hepatic and intestinal cytochrome P450 3A4 inhibition (Mastrolorenzo Rusconi et al., 2007).

Side/adverse effects: gastrointestinal side effects, lipodystrophy and dyslipidemia, Cardiac arrhythmias
1.7.4 **Integrase Inhibitors**

Pol gene codes for HIV integrase. During replication of virus, the IN is responsible for catalyzing two vital reactions. Firstly, it modifies the 3’ end of proviral DNA (3’-end processing) and facilitates strand transfer (STF) into cellular genome. During strand transfer process, an integrated proviral DNA is formed by insertion and joining of viral DNA to DNA of cell. INSTIs have very high specificity and affinity towards IN. They carry out chelation of the divalent metal ions mediated by the catalytic triad, i.e. the DDE motif, in the CCD of IN. It has been found that the binding site gets exposed upon conformational change caused by DNA.

An IN inhibitor could bind to this site. Further studies revealed that DNA itself could be part of this binding site. The inhibitors have one or more aromatic lipophilic moieties in their integrase binding domain although many such groups can be accommodated in the binding site. These hydrophobic groups are thought to be required for anchoring INSTI to a hydrophobic pocket of binding site in IN (Shin Taddeo et al., 1994; Asante-Appiah & Skalka 1997). Several compounds having IN inhibitory activity were isolated from fungi. One of such compounds is Equisitin derived from Fusarium spp., which showed inhibition of STF with IC50 of 7 µM (Hazuda Blau et al., 1999).

A carboxylate group, β-hydroxyketo group, or diketo group constitute the active pharmacophores of Integrase Inhibitors. The β - diketo acid (DKA) motif is a prototypical pharmacophore. Several structural analogs have been identified like tricyclics styrylquinolone motifs, naphthylazo diazide compounds. Elvitegravir possess a quinolone carboxylic acid motif (Marchand Zhang et al., 2002).

The recent discovery of analogues of naphthamide by Merck initiated the next phase of Integrase Inhibitors. Most of these analogues have 8-hydroxy 1,6-naptheridine attached to a benzoylestrenuent (Zhuang Wai et al., 2003).

Ralvitegravir (RAL, MK-0518), developed by Merck is clinically most successful IN inhibitor (Grinsztejn Nguyen et al., 2007; Markowitz Morales-Ramirez et al., 2006). Elvitegravir was later discovered by researchers at Japan (Al-Mawsawi Al-Safi et al., 2008; Shimura, Kodamaet al., 2008).
1.7.5 Budding and Maturation inhibitor

Compounds targeting gag, and its function during budding and maturation of virus particles are being developed. These inhibitors are known as maturation inhibitors. The small molecule 3-O-(3β, 3β-dimethylsuccinyl)-betulinic acid (DSB), also known as bevirimat (BVM), PA-457, or MPC-4326, is the prototype of this class of inhibitors. During maturation, proteolytic cleavage at the CA-SP1 junction is blocked by BVM. This leads to generation of virus that are found to be less or noninfectious. Additionally, it has been shown that during assembly of virus particles BVM gets incorporated in a gag-dependent manner (WenStern et al., 2006; LiGoila-Gaur et al., 2003).

A maturation inhibitor UK-201844 was developed by Pfizer. (Blair 2006). The mechanism of action of this compound is still not clear however it is thought to act by interfering gp160 processing, thereby generating non-infectious virus (Blair Cao et al., 2007).

Vpu, is encoded by HIV-1 but not HIV-2. Vpu is located in membranes and it is involved in many activities. From infected human cells, the budding of the virus and its release is mediated by Vpu. It has been shown that Vpu affects the release of newly formed virus and the infectivity. It is therefore a potential target for antiviral drugs. Many reports have shown that it is primarily localized in the membranes of the endoplasmic reticulum and Golgi compartments in infected cells. To find out inhibitors of Vpu ion channels, a compound library was developed by Biotron In monocyte-derived macrophages (MDM) these compounds displayed potent inhibition of HIV by blocking ion channels of vpu. BIT225 \{N-[5-(1-methyl-1H-pyrazol-4-yl)-naphthalene-2-carbonyl]- guanidine\} was identified as a novel compound. The ion channels of vpu are blocked by this compounds resulting in inhibition of infection (Ghoshet al., 2011; Richards & McCallister, 2008).
1.8 Immunotherapy

In addition to conventional ART, treatment strategies have been investigated that modulate immune function or response; however, they still lack a proof of benefit in clinic.

1.8.1 Interleukin-2 (IL-2, Aldesleukin, Proleukin™)

It is produced by activated T-cells. It causes propagation, survival of T-, B and NK cells. It is also known to induce cytokine release from these cells. It is licensed in Europe for the treatment of metastatic renal cell carcinoma. At the beginning of the 90s, IL-2 was already used intravenously in HIV-infected patients (Wood Montoya et al., 1993) but it is now given subcutaneously (Kovacs Vogelet al., 1996). In response to IL-2, there is an increase in memory T-cells followed by rise in number of naïve T-cells (Chun Engel et al., 1999; Carcelain Saint-Mezard et al., 2003). The CD4 nadir is predictive for the CD4 cell increase. It is still not clear whether the increase in CD4 cells on IL-2 has a clinical benefit.

1.8.2 Interleukin-12

NK cells and T-cells get stimulated by IL-12. Upon activation, these cells show a Th1-type immuneresponse. A randomized Phase I study with 100 ng/kg 2 x/week was conducted to see tolerance and efficacy of IL-12. IL-12 failed to activate lymphocyte subpopulations and did not produce antigen specific immune response. It did not reduce viral load (Jacobson Tang et al., 2006). Further development is therefore uncertain. Same is the case of IL-10 (Angel Jacobson et al., 2000).

1.8.3 Interleukin-7

Another promising candidate is IL-7. It is essential for T cell homeostasis and influences amongst other things the formation and maturation of CD4 cells. In two pilot studies, 6 and 16 HIV patients received different doses injected subcutaneously. In
both trials, good CD4 increases were observed together with good tolerability (Sereti Sklar et al., 2007; Sereti Herpinet al., 2001).

1.8.4 G-CSF and GM-CSF
These are useful in patients infected with HIV for a variety of reasons. The marketed forms of G-CSF are Filgastrim (Neupogen), Pegfilgastrim (Neulasta) and Lenogastrim (Granocyte) and that of GM-CSF are Sargramostim (Prokine) or molgramostin (Leucomax). FDA has approved G-CSF for treatment of prolonged neutropenia to reduce the risk of bacterial infections in AIDS patients. G-CSF may benefit patients receiving chemotherapy or drugs such as Gancyclovir or AZT. It has been found that there is marked reduction in bacterial infections in neutropenic HIV patients upon G-CSF treatment (Kedzierska Rainbird et al., 1998).

1.8.5 Interferons
Interferons have an antiretroviral effect. Because interferons have to be injected subcutaneously and have side effects which are not insignificant, they are not being pursued further in HIV medicine (Kedzierska & Crowe, 2001).

1.9 Bioactive compounds from marine organisms
The oceans represent an important resource for the discovery of novel bioactive compounds. The oceans cover more than 70% of the Earth’s surface and contain more than 3,00,000 species of plants and animals (Donia et al., 2003). Almost every class of marine organisms contains variety of molecules with unique structural features due to the physical and chemical conditions of the marine environment (Chin et al., 2006). The bioactive compounds are involved in biological functions of marine organisms such as communication, infection, reproduction and self-defense. More than 12,000 natural products have been isolated from marine algae,
sponges, coelenterates, ascidians, molluscsechinoderms and bryozoans (Matthee et al. 1999; Costantino et al. 2004). In

1950s, Spongothymidine and spongouridine were identified from Caribbean sponge *Cryptothecacrypta*, which are among the first compounds isolated from marine organisms. Their analogues Ara-A and Ara-C have potent antiviral activities and are in clinical trials (Chin et al. 2006). Identification of antiviral activity of nucleoside analogues spongothymidine and spongouridine provided a platform to screen nucleoside analogues for antiviral activity against HIV-1 virus and led to the identification of several of the nucleoside analogues with anti-HIV-1 activity. Recently, Ziconotide isolated from the cone snail *Conus magus* venom was approved to treat chronic pain. Ziconotide shows its effect by blocking N-type voltage-gated calcium channels (Schroeder et al., 2004).

Several marine organism-derived compounds and their analogues are under clinical trials for different diseases. Molecules with promising level of antiviral activity have been isolated from marine organisms following the bioassay-guided protocols. Cyanovirin-N, an antiviral lectin, was isolated from a bluegreen alga *Nostoc ellipsosporum*. It inhibits the replication of many viruses including Ebolavirus, HIV-1, Influenza Virus, Hepatitis C, Human Herpes Virus 6 and Measles Virus, etc. (Ziolkowska et al., 2006). Cyanovirin-N binds to HIV-1 surface protein gp120 and inhibits the viral entry.

Papuamides A, Avarone, and Microspinosamidewere isolated from different species of marine sponges and showed anti-HIV-1 activity. Marine bivalves belong to the class Bivalvia of Phylum Mollusca and are widely distributed in oceans. They are a good source of protein food and novel chemical structures. They are cultured to large numbers in coastal waters as a source of food and for identification of novel molecules. Myticin and *Aplysia* inhibitory peptides (AMIP), isolated from *Mytilus galloprovincialis*, show antimicrobial activity (Mitta et al., 1999). Bioactivity of some of the novel proteins such as class I metallothionein proteins isolated from *Crassostreavirginica* and Perninisolated from *Perna canaliculus* have not yet been characterized (Roesijadi et al., 1989; Scott et al., 2001).
Sulphated beta-galactans were isolated from clams, which inhibit HIV-1 replication in T cell lines (Amornrut et al, 1999).

1.10 Highly Active Antiretroviral Therapy

Soon after the introduction of AZT as mono therapy in 1987 to treat HIV infection, it was realized that mono therapy alone was not effective. The NRTIs d4T, ddI, and ddC were introduced in 1991 and 1994. However clinicians faced issues of efficacy similar to that of AZT. There were no clear guidelines about use of anti retroviral drugs, how to use them and what dose was required. In June 1996, a combination therapy “highly active antiretroviral therapy (HAART)” was reported by the World AIDS Conference in Vancouver (Aboulafia 1998; Shafer & Vuitton, 1999). This therapy was found to be more effective in treating HIV infected patients.

HAART comprising NRTIs and NNRTIs and protease inhibitors has been effectively used for the past decade. This has significantly helped to reduce and prolong viral load in infected individuals for several years. Life span of patients receiving HAART has improved significantly. HAART has now become a standard treatment regimen. Thus, it is believed that additional clinical benefits will be gained by introducing new anti-HIV drugs in HAART regimen. From the last several years, several studies are being carried out to develop new IN and PR inhibitors. Raltegravir and Maraviroc have been approved by FDA and these drugs were included in the portfolio of HAART. Though HAART has proven to be successful, prolonged use lead to the evolution of multi-drug-resistant virus strains. Patients were reported to have hepatotoxicity.

In many cases, inadequate PK, PD (pharmacokinetics and pharmacodynamics) have been reported. HAART is a costly treatment in the developing world, most of the patients have no access to this treatment as cost associated with HAART is limiting factor. People from Sub-Saharan Africa and Southeast Asia have been infected and virus is spreading rapidly especially among heterosexual women as there are no preventive measures and lack of effective antiretroviral therapy. In this region HIV transmission rate is very high. It is believed that sharing of information,
improving awareness and development of new and effective therapeutic strategies would help reduce or control the transmission of virus (MaltezDoroana et al., 2011b).

HIV infection mostly spreads through sexual intercourse. One of the most effective strategies would be use of microbicideto prevent sexual transmission. The rate of spread of epidemic necessitate the need for an efficient, safe and acceptable chemical and physical barrier methods like topical microbicidesthat can be controlled by females. It is realized that reduction in rate of transmission in the developing world would have huge effect AIDS pandemic. Hence many laboratories in the world are focusing on development of microbicidal agents. A singlemicrbicide with a moderate efficacy (50-60%) would be able to prevent several million cases of HIV transmission and infection in the world (Gupta 2011; Ramjee, 2011).

1.11 Microbicides to prevent HIV transmission and infection

Microbicides are designed for application at vaginal or rectal mucosae to inhibit or block early events in HIV infection and thereby prevent transmission of HIV. Microbicides can be formulated in different dosage forms e.g gels, creams films, or suppositories. The microbicidal agents may or may not possess contraceptive properties (spermicidal activity). Till date none of the microbicidal agent is proven to be effective. Several microbicidal candidates acting at mucosawherein HIV enters the epithelium are underdevelopment. Microbicides exert HIV inhibitory effect by blocking viral crossing of mucosal epithelia, or it may interfere in the attachment to Langerhans cells, DCs, macrophages, and CD4þ T-cells present at mucosa or sub mucosa. A microbicidal agent may also act by blocking the transfer of virus to peripherallymph nodes. Developing a safe and efficient microbicide is not an easy task.

Development of microbiciderequires support from the government bodies and funding agencies. Use of condoms, education or awareness about high risk behavior, early diagnosis and treatment of STDs may help to reduce HIV transmission. In spite of all these efforts, in developing countries, among women the rate of transmission is very high.
As of now, vaccine for HIV prevention is not available, in such scenario microbicides may help curtail new infections. In poor countries, the presently available prevention techniques are of far reach for women. An efficacious microbicide would enable women to protect themselves and their partners. Microbicides can easily be controlled by women unlike condoms. Use of microbicide by woman does not require the cooperation, consent or even knowledge of the partner.

Microbicides are sub classified as i) nonspecific and ii) specific microbicide. There are eleven microbicides presently being evaluated in clinical studies, and more than 50 are in preclinical development stage (Ramjee2011; PozzettoDelezay et al., 2012).

1.11.1 Non specific microbicides

Non specific microbicides are mostly comprised of detergents like nonoxynol9. Initially N-9 was found to have very good microbicidal activity in vitro. Further studies in vivo indicated that upon N-9 application, there was increase in HIV-1 transmission (Wilkinson Tholandi et al., 2002). Later on it was found that N-9 induced secretion of proinflammatory cytokines which damaged epithelial mucosa and thereby facilitated transmission of virus (Fichorova Tucker et al., 2001). PRO 2000 is a naphthalene sulfate polymer. It binds to CD4 receptors (Rusconi, Moonis et al., 1996). It was found to be safe and with no significant side effects (Keller Guzman et al., 2007) in clinical trials. Recent reports revealed that 30% protection could be achieved by application of pro 2000 for a period of 3 years (Cohen 2009).

This trial was very encouraging for scientific community working on development of microbicide. Apart from inhibiting virus attachment to CD4 receptor, PRO 2000 has been shown to block transmission of virus from DCs to CD4T-cells (Teleshova Chang et al., 2008; Teleshova Kenney et al., 2006). ACIDFORM, which act by modifying pH of vaginal fluid, is undergoing phase II trial. Phase 1 clinical trials for safety of ACIDOFORM have been successful (Williams, Newman et al., 2007). Invisible condom is being evaluated in phase II/III clinical trials. Invisible condom is composed of a
polyoxyethylene-polyoxypropylene copolymer. It forms a layer like a barrier over mucosa and protects the mucosa from entry/attachment of virus.

A 14 day intravaginal application study was successfully conducted to evaluate safety of invisible condom (Trottier Omar et al., 2007) Carrageenan and cellulose sulfate are known to inhibit HIV entry. These two compounds failed to prevent HIV-1 transmission (Skoler-Karpoff Ramjee et al., 2008). Hence further development of these compounds as microbicides was stopped. Viva Gel has displayed potent activity against HIV and HSV. Both in vitro and in vivo studies, it has been shown to be effective. In clinical trials, it was found to be safe (Rupp Rosenthal et al., 2007). Viva Gel is currently in phase I/IINearly 50 compounds are being tested in preclinical phase.

1.1.1.2 Specific Microbicides

Typically a specific microbicide is expected to inhibit infection during early stages of viral life cycle. These agents exert their anti-HIV effect before the incorporation of viral DNA in host cell DNA. One of the approaches to inhibit virus infection once the virus has entered the cell, is to use a microbicide having RT inhibitory activity or they should be able to inhibit integrase activity. Entry inhibitors are effective at the stage of virus attachment and fusion (Kelly & Shattock, 2011b).

1.1.1.3 Microbicides Targeting Cellular Side

Monoclonal antibodies (mAbs; TNX355) have been developed that bind to CD4 receptors. These mAbs block interaction of CD4 receptor with gp120. Another strategy would be to reduce or down regulate the expression of CD4 receptors. Cyclo tri-aza–di-sulfonamide (CADA) molecules have been reported to down-regulate the CD4 receptor. This down regulation of CD4 receptor is reversible. Cells become temporarily resistant to virus infection. mAbs (PRO-140), small molecules (SCH-CSCS-D, and UK427857) and modified chemokinestarget
chemokine receptors like CCR5, CXCR4. The modified ligands of these receptors have been shown to potently inhibit HIV infection.

These agents also cause prolonged down regulation of receptors. However production of mAbs or biological molecules is a costly affair and there would be many difficulties in scaling up stability and delivery or administration of these agents. Hence small molecules are preferred choice for development of microbicide candidates (Doncel & Clark, 2010).

1.11.4 Inhibitors acting on HIV env

The HIV envelope is formed by association of a trimer of gp120 with a trimer of gp41. A study involving 3 mAbs (2F5, b12, 2G12) in combination to see effect upon gp120 challenge in macaques, indicated that these antibodies exerted synergistic effect and protected monkeys to some extent. Also, the CD4-Ig-2 which has extended half life compared to sCd4 has been shown to prevent adsorption of virus to the cell-associated CD4 receptor by binding to gp120 (Ruprecht Ferrantelli et al., 2003).

Small molecules like BMS-806 inhibitor and its analogs are known to have affinity for CD4 binding domain of gp120 (Olson & Maddon, 2003; Reeves & Piefer, 2005). Carbohydrate-binding agents (CBA) showed promising activity as HIV microbicides. The gp120 and gp41 proteins are made up of glycans. This glycan structure helps HIV to shield from immune cells.

The glycan structure is present in all HIV isolates. Nearly 50% of glycans are high-mannose type. They are present over the surface of envelop proteins. These glycan molecules can be targeted by small molecules having affinity towards glycans. CBAs exert broad neutralizing effect on variety of HIV isolates including HIV-2 e.g. cyanovirin-N, plant lectins.90–94 Cynovirin when applied intravaginally or rectally protected monkeys from SHIV infection in microbicid trial. CBA comprises class of molecules that can bind or target multiple glycans on HIV envelope. Therefore, it is postulated that CBA will not be or to less extent susceptible to loss of activity or emergence of drug resistant strains because of mutations in gp120 compared to mAbs.
such as 2G12 which targets an epitope or specific glycan on gp120. Efforts are being made to produce these molecules from natural flora or streptococci to develop and maintain microbicide like environment invagina.

The functional two-domain CD4 have been expressed in lactobacillus jensenii. This will help inhibiting or reducing HIV infection. A live microbicide peptide secreted commensal vaginal bacteria has been reported by Rao et al. Small molecules like artificial lectins that bind mannose containing glycans of gp120 have been developed. These artificial lectins mimic the resemble and display properties of lectins. Since many cells in humans have glycoproteins expressed on cell membrane, there could be undesirable side effects associated with use of CBAs (Huskens & Schols, 2012).

1.11.5 Post-entry inhibitors

Once the virus enters the cell, the drug targeting early events of life cycle may become ineffective. Further production of virus particles and infection can be prevented by use of RT inhibitors and Integrase Inhibitors. HIV RT is widely studied as target for anti-HIV compounds. Most of the RT inhibitors have been evaluated or presently being studied in preclinical and clinical trials for their use as microbicide candidate. The NRTI Tenofovir is in advance stages of development as a potential microbicide. Study involving Tenofovir gels (0.1% and 0.3%) for safety and efficacy has been successful. Tenofovir is presently being evaluated in phase II study.

In this study, tenofovir gel is applied topically however trace amount of tenofovir have been detected in serum. Clinical significance of this finding is still not clear. NNRTIs possess very high therapeutic index and display specific activity towards HIV RT (Van & Szpir, 2012; Chen Zhan et al., 2011). They are active as such unlike NRTIs which require modification by cellular enzymes to show their effect.

TMC-120 and UC-781 are lipophilic and bind tightly to HIV RT. These two compounds are in most advanced stage in clinical trials. In vitro experiment involving co-culture of monocytes derived dendritic cells and T-cells demonstrated that UC-781
and TMC-120 inhibit HIV infection. The safety index of these compounds were also good. UC-781 effectively inhibited transmission and infection of genital tissue. Results with TMC-120 were similar. It is thought that vaginal ring sustained release dosage form of TMC-120 may help in preventing HIV transmission and infection. TMC-120 gel prevented transmission and infection of HIV in huSCID mice.

Lipophilic nature of these compounds may result in low oral bioavailability and limited systemic absorption, although after high intravaginal dosing, TMC-120 at very low concentration may appear in plasma. Presence of drug at very low levels in plasma may result in emergence of drug resistant viruses. PETT derivatives (MIV-150) and DABO derivatives are promising candidates for microbicide development. Lipophilic compounds (for example UC781, TMC120) have tendency to remain at sites in vaginal tissues susceptible to HIV infection and thus remain at the site of action for prolong period. A microbicide formulation of such compounds will have long lasting effect. This would allow use of microbicide before sexual intercourse (D'Cruz & Uckun, 2006).
Figure 1.14: Viralentry and dissemination during the sexual transmission of HIV-1. (A) Interactions of HIV-1 envelope glycoproteins CD4, and CCR5 or CXCR4 co-receptors trigger fusion and entry of HIV-1. (B) Outline of the sequence and time course of events involved in viral dissemination (Fox & Fidler, 2010)
1.11.6 Resistance by exposure to microbicides

Getting an FDA approval for formulation of microbicide containing two or more drugs is a very difficult task. FDA states that microbicides having more than one active ingredient should show synergistic effect and superior clinical effects over each individual component. This would require huge investment for conducting clinical trials and makes the clinical study lengthy. A microbicide formulation with only one active ingredient poses high risk of developing drug resistance compared to a microbicide formulation with 2 or more active components targeting different stages of viral life cycle. Studies have shown that at high drug concentration, drug crosses the epithelial mucosa and enters into systemic circulation and low doses in serum/blood may lead to development of drug resistant strains. Inappropriate and inconsistent use of microbicides for long period of time by persons with high exposure risk may also cause development of drug resistant virus strains (tenofovir mutation K65R), which could thereby impact future use of this highly desired life-saving drug (McGowan Hoesley et al., 2013; Obiero Mweteha et al., 2012).
Table 1.4: Microbicides and PrEP candidates in ongoing clinical trials

<table>
<thead>
<tr>
<th>Product</th>
<th>Mechanism of action</th>
<th>Phase of Clinical Trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRO2000 (Gel)</td>
<td>virus entry</td>
<td>Phase 3</td>
</tr>
<tr>
<td>Viread (Oral)</td>
<td>NtRTI</td>
<td>Phase 3</td>
</tr>
<tr>
<td>Tenofovir</td>
<td>NtRTI</td>
<td>Phase 2B</td>
</tr>
<tr>
<td>Tenofovir (Gel)</td>
<td>NtRTI</td>
<td>Phase 2</td>
</tr>
<tr>
<td>Dapivirine (Gel)</td>
<td>NNRTI</td>
<td>Phase ½</td>
</tr>
<tr>
<td>VivaGel (Gel)</td>
<td>Virus entry</td>
<td>Phase 3</td>
</tr>
<tr>
<td>Acidform</td>
<td>Vaginal defcense enhancer</td>
<td>Phase 1</td>
</tr>
<tr>
<td>UC-781(Gel)</td>
<td>NNRT</td>
<td>Phase 1</td>
</tr>
</tbody>
</table>
1.12 AIM AND OBJECTIVE OF STUDY

The present therapy i.e., Highly Active Antiretroviral Therapy aims at reducing the viral load and thereby bringing CD4 counts to normal level. NRTIs, NNRTIs, PR inhibitors, one Integrase Inhibitor and fusion inhibitors are components of HAART. Though HAART has been proven to be successful it is not stillable to remove the virus from the body. Use of HAART is associated with several undesirable side effects. Low bioavailability, development of drug resistant viruses and poor pharmacological profile are some of the issues of concern. High treatment cost is another factor that makes the HAART far reachable to poor people in developing countries.

Due to unavailability of treatment, sub-Saharan Africa and Southeast Asia are the most widely affected by infection of HIV. In these regions, number of women infected with HIV is much more.

To reduce the rate of transmission and number of cases of new infection, behavioral awareness education program needs to be conducted. However such efforts are not complete and may not have huge impact on infection cases. This necessitates development of effective prevention strategies. Use of microbicides is one of the important preventive strategies for prevention of spread of virus and disease. Till date there is no cure for AIDS. Hence efforts should be made to find out new anti-HIV molecules with improved activity than the existing drugs so as to eradicate the virus. Natural resources like plants marine organisms have been the source of new therapeutic agents from several years.

Knowledge of Organic chemistry has helped immensely to produce or synthesis these agents and their analogs with better pharmacological or pharmaceutical properties. The scaffolds of natural products have been used to design and develop specific and potent drugs. Libraries of compounds have been developed based on alkaloids, polyketides, terpenoids and flavonoids. Natural products buchapine and mallotojaponin have shown very promising anti-HIV activity. However, these molecules have not been completely explored as anti-HIV lead molecules. We
intend to assess anti-HIV-1 activity of analogues of mallatojaponin (dimeric phloroglucinols) and buchapine (quinoline 2,4 diols) in various cell lines against HIV-1 isolate. Most potent compounds will be further assessed to determine their mechanism of action of anti-HIV-1 activity. Most potent compounds will be further evaluated for efficacy to be used as HIV-1 microbicides.
1.13 PLAN OF WORK

- Review of literature
- Procurement of proposed compounds
- Evaluation of Cytotoxicity of compounds
- Evaluation of anti-HIV-1 activity of compounds using TZM-bl cell lines
- Assessment of anti-HIV-1 activity of active compounds in human PBMCs
- Determination of mechanism of action of lead compounds
- Evaluation of lead compounds as potential HIV-1 microbicides