# Literature Review

## CHAPTER 1

### Global epidemiology Report of HIV

United Nations Acquired Immune Deficiency Syndrome (UNAIDS) global report 2012 approximates 34 million people have been infected by Human Immune deficiency Virus (HIV) by the end of 2011 [UNAIDS 2012]. Though the report shows that the global trend of new HIV infections (figure 2) and HIV-related deaths (figure 3) per year is declining, the current number of HIV infections is the highest since 1990 (figure 1). There is a huge difference in the HIV prevalence by age and sex, among the countries around the global. Figure 4 shows the distribution of HIV prevalence by regions all over the world. Sub-Saharan Africa region is the most HIV aggravated region; there are 23.5 million HIV positive people. UNAIDS estimates that approximately 1 in every 20 adults are HIV infected (UNAIDS). This is 25 or more times the HIV prevalence in any other region of the world. Countries in Sub-Saharan Africa also have varying HIV prevalence (figure 5). The next severely affected regions are Caribbean and Eastern Europe, Central Asia, South Asia, South-East Asia and East Asia. UNAIDS targets to decline the HIV transmission by 50%, provide anti-retroviral drugs to 15 million HIV positive people and increase the annual global investment to US $ 24 billion in low and middle-income countries by 2015.

### Introduction to HIV

HIV is a retrovirus, which has two strands of Ribo Nucleic Acid (RNA) as the genetic material. The most common modes of the virus infection are sex transmission, intravenous drug usage, transfusion of infected blood and mother to child transmission at birth. HIV enters CD4+ expressing immune cells like T-lymphocytes and, very quickly, replicates in it. The immune cells burst, releasing thousands of new HIV viruses; each one capable of targeting new immune cell to begin cycle again. The HIV patient’s immune system is disabled to defend foreign particles when the virus destroys huge amount of immune cells. Opportunistic infections by other diseases take the chance of week immune system to attack the human body and eventually kill the patient.

### Origin and Evolution of HIV

In 1981, the first patient of HIV AIDS was identified in the USA. In 1983, the virus was first isolated at “Institut Pasteur” in Paris (Alvine, Kien). Researches on the causative agent of the disease – HIV, implied that it is a retrovirus. Exploration of the retrovirus led researchers to find similar type of retrovirus in non-human primates that were then called as Simian Immunodeficiency Virus (SIV). Researches have established that HIV in human was a result of zoonotic transmission of SIV from non-human African wild primates (Bailes et al. 2002). About 40 different primates, in Africa, are infected with SIV and some are harboring some than one strain of SIVs. Phylogenetic analysis of SIVs from African non-human primates and two HIVs (HIV -1 and HIV -2) in human provided remarkable understanding of viral transmission and evolution from non-human primates to human.

Researchers have now established that HIV-1 and HIV-2 transmission to human are independent and their source are different. Discovered in 1986 AD, HIV-2 is transmitted from Sooty mangabey monkeys (Cercocebus atys) and the HIV-2 prevalence is also high in the geographical location of these monkeys in West Africa (Santiago et al., 2005). Sooty mangabey monkeys are naturally infected by a strain of SIV that is very close to HIV-2 (Hirsch et al. 1989) and the phylogenetic analysis of all HIV-2 strains show that it closely groups with the SIVsmm strains (Geo et al 1992, 1994). SIVsmm does no harm to its host monkeys and must have modified to produce multiple strain and subsequent multiple zoonotic transmissions from sooty mangabey monkeys to human (Hahn et al 2000). Although HIV-2 subtypes A to G are identified in human, it is assumed that more subtypes were introduced into human (Gurtler 2004) but are lost for low adaptation fitness (Damond et.al 2004).

Initial researches show that chimpanzees (Pan troglodytes) are the source of HIV-1 infection to human (Peeters et al. 1989; Huet et al. 1990) but the lack of enough evidence from vast number of other chimpanzees being test showed negative results and then subsequently discarded the idea (Vanden Haesevelde et al. 1996). Years later, in 1999, another chimpanzee was tested positive for SIV close to HIV-1. Four chimpanzee subspecies were analyzed with mitochondrial DNA (Groves 2001). The subspecies were: western (Pan troglodytes verus), Nigerian (Pan t. ellioti), central (P. t. troglodytes) and eastern (Pan t. schweinfurthii) chimpanzees (Gagneux et al. 1999). Retrospective research showed that all chimpanzees that are tested negative were subspecies *P.t versus* (Prince et al. 2002; Switzer et al. 2005) and those tested positive were *P.t troglodytes or P.t. schweinfurthii* (Geo et.al 1999, Corbet et al 2000, Santiago et al. 2003; Worobey et al. 2004; Keele et al. 2006, 2009; Van Heuverswyn et al. 2007). These studies and evidence from faecal samples (Keele et al 2006) confirmed the source of HIV-1 as *Pan troglodytes*.

### Structure of HIV

HIV is spherical in shape and the size is approximately 0.0001mm or 1/12th the size of *Escherichia coli*. The viral structure can be divided into outer viral envelope and the inner viral core.

The viral envelope consists of the outermost lipid bilayer membrane, derived from the host immune cell membrane during budding out from the cell. HIV has genetic code for all its structure parts, but the lipid bilayer. Several host protein molecules are embedded on the membrane. Besides these proteins, the viral protein “Env” are also present on the surface. Env proteins form spike like structure that emerges out from the viral lipid membrane. Env consists of three glycoprotein (gp) 120 that make trimeric structure with three gp41 stems that anchor the viral envelop (Zhu et al. 2006, Subramaniam et al. 2006, Zanetti et al. 2006, Zhu et al. 2008) (Figure 7- <http://www.niaid.nih.gov>).

The inner core consists of matrix, capsid – made up of p24 viral proteins and nucleocapsid. Inside the capsid, there are two single strands of viral RNA molecules that codes for all the viral proteins and structures.

### HIV genome structure and organization

The length of HIV-1 HXB2 (HIV reference sequence) genome is about 9719 bps and HIV-2 BEN genome is about 10359 bps (figure-8 www.hiv.lanl.gov). On overall, HIV genome produces 15 proteins. The 5’ and 3’ ends consist of long terminal repeats (LTR) or also known as Untranslated Terminal Repeats (UTR). The 5’ – gag – pol – env – 3 ’ are the three genes that code for most of the structural proteins and necessary enzymes for viral replication. The structural elements and genes of HIV genome and their function or product are described below as they appear from 5’ to 3’.

**LTR or UTR**: The long terminal repeat is the most conserved region in the HIV -1 genome [Pillart et al 2004, Kuiken et al 2009]. It promotes viral packing before budding out, regulates and promotes transcriptional activation, binds primer for reverse transcription and dimerizes two single strands of RNA [Lu et al 2011]. It has splice donor site contributing to the packing of the full length viral RNAs [Mann et al 1985, McBridge et al 1996]. It interacts with the nucleocapsid domains to incorporate the full-length genome during the assembly of new virions [Berkhout 1996]. Researchers have proposed that the 5’-UTR adopts multiple secondary structures and regulates mutiple RNA-dependent functions during viral replication. The nucleotides from 105-115 of 5’-UTR region can bind to oligoribonucleotides having *gag* start codon (AUG) to disrupt pre-existing stem loop and form a heteroduplex [Spriggs 2008].

**GAG**: The gag region spans from 790 to 2292 bp numbering in HXB2 genome at reading frame 1. The translational product of Gag is a precursor Pr55 that has all the building blocks to form a fully infectious virion, even in the absence of other viral products [Wang et al 1993]. The viral protease enzyme cleaves the gag precursor to yield the necessary structural proteins of mature virion [Gheysen et al 1989, Hunter 1994, Coffin, Swanstrom and Wills 1997, Freed 1998,]. The protease enzyme comes in to nascent virions as a component of the Gag-pol polyprotein, which is, later, produced by ribosomal frameshifting between the overlapping gag and pol genes [Lillehoj et al 1988, Shioda et al 1990, Schneider et al 1997]. Protease cleaves gag precursor forming three structural proteins: matrix, capsid and nucleocapsid. The matrix protein is always at the N-terminal and p6 at the C-terminal of gag precursor with two spacer regions –p1,p2 (figure 8) which separate capsid from nucleocapsid and nucleocapsid from p6 [Mervis et al 1988, Henderson et al 1992].

**POL**: The pol region spans from 2085 to 5096 bps numbering in HXB2 genome at reading frame 3 (figure 8). Pol produces the most necessary enzymes - protease, reverse transcriptase and integrase for viral replication. These enzymes are the targets for drug development to find a way to stop HIV replication. Many drugs have been developed to interfere the functions of protease and reverse transcriptase. The protease enzyme processes the gag-pol precursor polyprotein to produce these enzymes. The protease enzyme cleaves the gag polyprotein to form the viral structure proteins. The reverse transcriptase enzyme produces the cDNA from the viral RNA. The intergrase enzyme integrates the proviral cDNA into the host genome.

**ENV**: Env spans from 6225 to 8795 bps numbering in HXB2 at reading frame 3. It produces viral glycoprotein as a precursor (gp160), which is processed to produce glycoprotein 120 (gp120) and glycoprotein 41 (gp41). Gp120 is a non-covalent complex of external protein and gp41 is a trans-membrane protein. They are organized to form trimeric complexes on the surface of HIV and mediate the HIV entry into the host cell [Allan et al 1985, Robey et al. 1985, Veronese et al 1985]. The exposed external complex gp120 binds to the CD4 receptor on the host immune cell. This triggers the conformational change in it and binds to one of the two chemokine receptor, CCR5 or CXCR4, to fuse with the host cell [Dalgleish eta al 1985, Feng et al 1996, Deng et al 1996, Choe et al 1996, Dragic et al 1996].

**TAT**: Tat is a trans-activating factor for HIV gene expression. After infection of host cell, HIV proviral genome is integrated in to the host genome, which is regulated by cellular as well as the viral transcription regulatory factors. Tat is the primary transcriptional regulatory factor. An example of Tat action, is the control of RNA polymerase II elongation during transcription. In the absence of Tat, polymerase II disengages from the template DNA strand, terminating the transcription prematurely (Kao *et al*., 1987; Kessler & Mathews, 1992; Ratnasabapathy *et al*., 1990; Toohey & Jones, 1989). Sodroski *et al*. (1985) first explained the function of Tat. There are two forms of Tat – Tat-1 and Tat-2. Tat-1 is a minor form, which is 72 amino acids long; Tat-2 is the major form, which is 86 amino acids long. Tat is found at least at the lower level in the infected host cell, and located in the nucleus.

**REV**: Rev is another trans-activating factor for HIV gene expression. It is 19-kD phosphoprotein. Like Tat, it is also mainly localized in the nucleus of host cell, but cycles rapidly between the nucleus and cytoplasm as it promotes nuclear export of the transcriptional products. Rev binds at the Rev Responsive Element (RRE), which is an RNA element encoded within the env region of the virus.

**Virion Infectivity Factor (VIF)**: This gene is encoded at region 5041 to 5619 bp numbered in HXB2 at reading frame 1. It produces a protein of 220 amino acids and 23 kiloDalton. This protein promotes the viral infectivity to the host, but has no role in viral production. The protein is produced in the late stage of viral production [shreehy et al. 2002, Von Schwedler et al 1993] to suppress the innate antiviral activity of human immune cells [simon et al. 1998, Madani et al 1998]. Researcher show that human APOBEC3 family members – APOBEC3G and APOBEC3F are the two potent cytidine deaminases that attributes to introduction of excessive G → A mutations in the minus strand reverse transcripts [Harris et al 2003, Mangeat et al 2003, Lecossier et al 2003, Zhang et al 2003]. Vif protein prevents APOBEC3 members from hypermutating HIV reverse transcripts by inducing proteasomal degradation and exclusion from virions through recruitment of a cullin5 ECS E3 ubiquitin ligase complex [Conticello et al 2003, Marin et al 2003, Sheehy et al 2003, Stopak et al 2003, Mariani et al 2003, Mehle et al 2004, Wiegand et al 2004].

**Viral Protein R or VPR:** It is 96 amino acids long protein (14 kiloDaltons). This protein is packed into the nascent virions during budding out. Vpr plays role in entry of viral core into the host cell nucleus and thus localizes in the host cell nucleus after infecting the cell. Vpr arrests the cell cycle during the transfer from G2 to M phase [Jowett et al 1995, Rogel et all 1995] by preventing the activation of the p34cdc2/cyclin B complex [He et al 1995]. One research shows that vpr is also important for efficient viral replication in monocyte or macrophage cells, but in T-cells [Connor et al 1995].

**VPU or Viral Protein Unit:** This is the unique protein to HIV-1 and some other closely related simian immunodeficiency viruses (SIVcpz, SIV-GSN, SIV-MUS, Siv-MON, SIV-DEN). It is a 16 kiloDalton , 81 amino acids long protein. The Env precursor protein, gp160, makes stable intracellular complex with host cell CD4, reducing the CD4 expression at cell surface and processing of gp160 to gp120 and gp41. HIV protein vpu prevents the formation of Env-CD4 complex; increasing the gp160 processing but rapidly degrades CD4 in endoplasmic reticulum [Willey et al 1992]. Researchers have showed that vpu action of degrading CD4 could be blocked with peptide aldehyde or lactacystin, by interfering proteasome function of the protein [Schubert et al 1998].

**NEF:** This gene was formerly named F, 3’-ORF, or B-ORF. This accessory regulatory protein does not help the virus to increase infectivity but plays a role during the biogenesis of viral particles [Laguette et al 2009] and virulence [Piguet et al 1999, Prince et al 2002, Kirchhoff et al 2008]. The function of Nef is that it down regulates the production of major histocompatibility complex type 1 (MHC type 1) in the host cell [Schwartz et al 1996, Collin et al 1998, Cohen et al 1999]. This impairs the function of cytotoxic T lymphocyte cells to clear the infected cells [Collins et al 1998, Tomiyama et al 2002, Yang, et al 2002, Adnan et al 2006]. Nef also down regulates CD4 on host cell surface [Garcia et al 1991, Lama et al 1999] and modulates cellular activation [Baur et al 1994, Sawai et al 1994, Bodeus et al 1995, Saksela et al 1995, Smith et al 1996]. This shows that Nef plays role in host immune evasion.

### HIV life cycle

The life cycle of HIV starts from entry into the host immune cell to budding out nascent HIV virions for further infection to new cells or to new host. The mechanism of the viral entry into the immune cell involves HIV gp120 binding to the CD4 receptor of the host cell [Dalgeish et al 1984, Maddon et al 1986, McDougal et al 1986]. This binding brings conformational changes in gp120/gp41 complex [Sullivan et al 1998], to enable binding to a chemokine receptor either CCR5 or CXCR4 [Sattentau et al 1991, Moore et al 1992, Sattentau and Moore 1993, Berger et al 1999]. These binding assist the viral membrane fusion with the host cell membrane.

Researches show that a seven –transmembrane protein domain in T-cells fusion [Berson et al 1996] and CC CKR5 in Macrophages cells [Alkhatib et al 1996] serves as an accessory fusin factor for the viral fusion.

HIV fusion to the host cell and successfully release genome and integrate it to cell genome has always been the subject of interest. The technology of live cell imaging has substantially improved our understanding on the viral process of internalization, its steps of intracellular pathway/s through the dense cell cytoplasm to reach the nucleus and release the genome [Sun et al 2013]. Unlike most other viruses, HIV has pH-independent gp41 fusion protein [Wilen et al 2012], forms fusion pore [Melikyan et al 2008] and directly fuse with the plasma membrane [Marsh et al 2006, Wilen et al 2012]. The fusion process is initiated after coreceptor binding induction and subsequent exposure of hydrophobic gp41 and insertion of it into the host cell membrane [Wilen et al 2012].

Upon fusion, there are many events occurring in the journey of HIV to the cell nucleus. From cell entry to nucleus, HIV loses some of its proteins, exploits cellular factors, reverse transcribe is RNA genome to DNA preintegration form and ultimately integrate it into the cell genome [Cullen et al 2001]. McDonald et al [2002] were able to track the HIV virion pathway to the nucleus using GFP-labeled particles. The researched showed that HIV uses cytoplasmic dynein and the microtubules network to accumulate at the perinuclear region, often near the microtubules-organizing center. The viral movement is curvilinear in the cytoplasm. The same study showed that the reverse transcription process starts in the intact capsid and the required deoxynucleotides are imported to build up the DNA. The capsid proteins dissociates but the reverse transcription complex remains intact interaction with microtubules, on the way to microtubules organizing center, close to nucleus. Ultimately, the reverse transcription complex reach close to the nucleus; enters through nuclear pore and the integrease enzyme integrates the preintegration DNA into the host DNA [McDonald et al 2002].

HIV uses the cell mechanism of gene expression to transcribe the viral genes. HIV gene Tat promotes the transcription of the viral DNA [Ott et al 2011, Razooky and Weinberger 2011]. HIV viral production from the infected cell can be broadly explained in three steps: assembly, budding and maturation.

Upon production of all necessary viral particles, initiation of HIV virion assembly occurs at the plasma membrane. The Gag polyprotein mediates all the essential events in the assembly, including the binding to the plasma membrane. It also makes protein-protein interactions to create spherical particles at the membrane, concentrates the viral Env protein and packages two copies of capped and polyadenlylated full length viral genomic RNA [Johnson et al 22010]. The complete assembled virion particles are released from the plasma membrane by the host Endosomal Sorting Complexes Required for Transport (ESCRT) machinery [Morita and Sundquist 2004, Bieniasz 2009, Carlton et al 2009, Usami et al 2009, Hurley and Hanson 2010, Peel et al 2011]. The maturation of the nascent HIV virions begin concomitantly with budding. Maturation is driven by the viral PR enzyme cleavage of the Gag and Gag-Pro-Pol polyproteins. The processed polyproteins produces matrix, capsid, nucleocapsid, p6, protease, reverse transcriptase and integrase proteins [Hill et al 2005]. These processed proteins get rearranged to create the mature infectious virion, ready to next infection to start new life cycle [Mariani et al 2003].

### HIV types and subtypes, Subsubtypes and Circulating Recombinant Forms

HIV is hugely diverse virus. Its diversity is obtained from phylogenetic analysis of genomic region to genome wide sequence. It creates its diversity to adapt in different environment like host immune system and drug pressure.

Classification of HIV by group, subtype and subsubtype require a reference sequence with the criteria: a sequenced full-length genome, no recombination history, HIV genome sequence published in peer reviewed citation, isolated from recent samples, HIV is covered in major geographical distribution, HIV has no sign of hypermutation, HIV genome sequence is real sequence from a patient, HIV genome has no extreme indels, and the virus must be viable and intact (Leitner et al 2005).

There are two distinct types of HIV: HIV-1 and HIV-2. These viruses can be differentiated by their genome organization and phylogenetic relationship (Hahn et al 2000), pathogenesis, transmissibility and pattern of spread (De cock et al 1993, Kanki and De cock, 1994). A notable difference is the source of the infection to human. Evidences show that HIV-1 is zoonotic transmission to human from chimpanzee *(Pan troglodytes*) and HIV-2 from sooty mangabey (*Cercocebus atys*). The genome wide sequence of both HIV show that HIV-2 has an extra gene “vpx” which is lacking in HIV-1 genome (Henderson et al 1988, Kappes et al 1988, Tristem et al 1992, Bergamaschi et al 2009). Drug resistance clinical results show that non-nucleoside reverse transcriptase inhibitors (NNRTIs) anti-retroviral drugs are effective against HIV-1, but non-effective against HIV-2 reverse transcriptase (Hizi et al 1993).

Isolation, characterization and sequence analysis of each type of HIV show that there exist genomic heterogeneity and variability among different isolates (Benn et al 1985, Hahn et al 1985, Wong-Stall et al 1985). Analysis of isolates within a patient and between patients demonstrates that intra-patient isolates are more related than inter-patient isolates (Hahn et al 1986). HIV-1 isolates from around the world and their phylogenetic sequence analysis reveals distinct subtypes that cluster together to form three groups: M (Main), O (Outlier) and N (non-M and non-O) (Robertson et al 2000, Leitner et al 2005), which represent three independent zoonotic transfection of SIV from chimpanzee to human (Hahn et al 2000, Sharp et al 2001).

Within each group, phylogeny of the genomic region with most variability allows classification of the virus to subtypes. An unrooted phylogenetic analysis of the HIV-1 isolates from group M using complete gag, pol, env and nef sequences or full length genome sequence analysis show distinct subtypes A – D, F – K (figure Buonaguro et al 2007). The phylogenetic clades of subtypes can be constructed from any part of the HIV-1 genome when the alignment is at least 300-500 bases long (Leitner et al 2005). Sequence length shorter than threshold from genomic regions under higher evolutionary pressure e.g. env V3 can reconstruct the phylogenetic clades distinguishing the subtypes, while regions under slower evolutionary change need long sequences to give reliable results (Leitner et al 2005). Further phylogenetic structure for subtype A and F have been identified leading to the classification of subsubtypes A1 and A2 for subtype A and F1 and F2 for subtype F (Gao et al 2001, Lietner et al 2005).

Two or more HIV subtypes infecting a single patient create inter-subtype recombinant forms called “unique recombinant forms” (McCutchan 2006). Isolation and identification of unique recombinant form from at least three epidemiologically unlinked patients and characterized by full-length genome sequencing is designed as circulating recombinant forms (CRFs) (Los Alamos Laboratory [[http://www.hiv.lanl.gov/content/hiv-db/CRFs/CRFs.html](http://www.hiv.lanl.gov/content/hiv-db/CRFs/CRFs.html" \t "_blank)]). There are huge numbers of CRFs known; the most prevalent are CRF01\_AG, CRF02\_AG in West Africa and CRF01\_AE in Southeast Asia. Country specific prevalent CRF may be different e.g. CRF01\_AG in Nepal and India (Shahid et al 2011). All discovered CRFs are documented in [http://www.hiv.lanl.gov/content/hiv-db/CRFs/CRFs.html](http://www.hiv.lanl.gov/content/hiv-db/CRFs/CRFs.html" \t "_blank).

Unlike group M, group O is endemic and largely confined to certain geographical region of Cameroon and neighboring West Central African region; group O only represents a small minority of HIV-1 strain in the region (Peeters et al 1997, Jaffe and Schochetman 1998, Janssens et al 1999). Group N is also confined to Cameroon and is found in limited isolates only (Simon et al 1998).

HIV-2, first isolated from patients in West Africa, exhibits HIV-1 AIDS like symptoms but shows seronegative to HIV-1 assay (Clavel et al 1986, Clavel et al 1987, Leys et al 1990, Gao et al 1992). Both viruses exhibit significant similarity at genome level (guyader et al 1987).

HIV-2 is classified to epidemic subtypes (A, B) and non-epidemic subtypes (C, D, E, F) [Hahn et al 2000, Tebit et al 2007]. Researchers expect that HIV-2 subtypes are independent cross-species transmissions of SIV to human population, very much like HIV-1 groups [Hahn et al]. Aguchi et al [2000] suggested a new HIV-2 subtype G. Some researches are done on subtypes A and B (Gao et al 1994, Chen et al 1997, Damond et al 2001, Damond et al 2002, Pieniazek et al 2004, Ishikawa et al 2004, Tebit et al 2007) and very little is known about the other subtypes. Widely spread subtypes A and B have high chance co-infecting a single person. The first HIV-2 CRF reported was CRF01\_AB, which is estimated to be recombined as between 1964 to 1973 [Iba et al 2010]. The discovery of the CRF01\_AB outside of subtype A and B epidemic area shows that it is ongoing global spread [Iba et al 2010].

### Anti retroviral drugs and HIV treatment

HIV infection can be confirmed using a serological assays or antigen-antibody assays [Laperche et al 2012, Rosenberg et al 2012, Workman et al 2009, Schupbach et al 2006, Fox et al 2011, Fiscus et al 2007, Pilcher et al 2005]. Upon HIV infection, the virus replicates very quick, destroying the host immune cells; and the patient has to undergo HIV treatment with the only available method, which is anti retroviral therapy (ART) [Gilks et al 2006].

The decision to start ART should depend on clinical, immunological and viral assessment of HIV in adults and adolescents [Gilks et al 2006]. Different parameters are checked to decide the ART treatment of a patient: CD4 cell count, pregnancy, co-infections with other viruses (HBV, HCV) and bacteria (TB), plasma HIV RNA level, body weight etc. However, International AIDS Society-USA recommends ART initiation regardless of CD4 count, in resource limited settings without readily access to proper clinical facilities [Thomson et al 2010]. The scale-up of ART program in the past decade has emerged drug resistant HIV virus [Boden et al 1999] circulating in the human population. The test for drug resistant HIV needs to be prioritized for the selection of drug regiment in ART [Hirsch et al 2000, Grant et al 2002, Hirsch et al 2003, Bansi et al 2010]. A person starting ART must consider important factors like adherence to the therapy for life long, toxicity, emergence of HIV drug resistance and subsequent change of drug regimen, drug interactions, the risk of viral transmission to uninfected healthy people etc [Thompson et al 2010].

The initial ART drug regimen (first line therapy) has a combination of two Nonnucleoside Reverse Transcriptase Inhibitors (NRTIs) and one Nucleoside and Nucleotide Reverse Transcriptase Inhibitors (NNRTIs). Upon first line ART drug failure, second line ART is recommended, which consists of change in a drug from NRTIs class and addition of a drug from Protease Inhibitor drug class.

**Nucleotide and Nucleotide Reverse Transcriptase Inhibitors (NNRTIs)**

NNRTIs are highly HIV-1 reverse transcriptase selective drug, and ineffective against HIV-2 reverse transcriptase. Three-dimensional structural information is the center for the development of NNRTIs drugs [Ren et al 1995, Sarafianos et al 2009]. The drugs in this class, bind at an allosteric site 10A from polymerase active site of the HIV-1 reverse transcriptase [Kohlstaedt et al 1992, Esnouf et al 1995, Ding et al 1995], close to substrate-binding site. The binding induces confirmational change in the enzyme, which distorts the catalytic aspartate triad and inhibits the function of the enzyme to convert RNA to cDNA. [Esnouf et al 1995, Balzarini 2004]. NNRTIs drug are administered along with other drugs in ART at the concentration lower than required to affect human cells [De Clercq 1996] leading to non-toxicity. NNRTIs drugs from first generation have inherent disadvantageous characteristics like cross-resistance and low genetic barrier [Adams et al2010] as they easily select for mutant viral strain with several degrees of drug resistance [Balzarini 2004]. Low genetic barrier indicates that the virus with one or two mutations can easily escape the drug. High genetic barrier indicates many mutations are required for the virus to escape the drug. Nevirapine and delavirdine are first generation NNRTIs drugs that easily loose their inhibitory potential against mutant viral strains. This has limited the use of first generation NNRTIs in the future. Etravirine [MInuto and Haubrich 2009], Efavirenz [Ren et al 2000], capravirine [Gewurz et al 2004] and dapivirine [Fletcher et al 2009] are second generation NNRTIs with higher genetic barrier and requires a markedly longer time period to obtain significant resistance them [Ghosn et al 2009Adams et al 2010].

**Nonnucleoside Reverse Transcriptase Inhibitors**

Reverse Transcriptase is the major target enzyme for the development of anti-AIDS drugs [Sarafianos et al 2004]. The enzyme produces cDNA from the viral template RNA. Nucleoside Reverse Transcriptase Inhibitors (NRTIs) are analogs of normal nucleotides but without 3’hydroxyl group. The drug is taken as unphosphorylated form, which cytokinases phosphorylates to form 5’ triophosphates [De Clarcq 2002, Ilina and Parniak 2008]. It leads to incomplete termination of cDNA synthesis and are the major anti retroviral drugs in ART regimen available in clinics. Reverse transcriptase possesses two characteristics: hydrolytic excision of nucleotide (primer-unblocking) blocking DNA synthesis (Arion et al 1998, Meyer et al 1999) and dynamic copy choice that is switching of the template RNA [Svarovskaia et al 200, Nikolenko et al 2004] while generating cDNA. There exists equilibrium between the rates of DNA synthesis and RNA degradation, which determines the template switching. The presence of NRTIs drugs terminates the DNA synthesis chain, breaking the equilibrium and increasing template switching. RNaseH degrades the viral RNA once the DNA synthesis stops; this leads to template and synthesized strand dissociation, terminating HIV-1 replication [Nikolenko et al 2004]. Researchers [Nikolenko et al 2004] postulated that reverse transcriptase has property to excise an incorporated NRTI and resume the DNA synthesis and subsequent RNA degradation by RNaseH. For e.g. NRTI drug 3′-azido-3′-deoxythymidine (AZT) related mutation in reverse transcriptase increases the rate of AZT excision allowing the continuation of DNA synthesis and RNA degradation at the end. Mutation in RNaseH leads to delay in RNA degradation and reverse transcriptase gets more time to excise the incorporated NRTI and resume DNA synthesis. Mutation in RNaseH gives the virus resistance to AZT and Stavudine/ 2,3-didehydro-2,3-dideoxythymidine(d4T) drugs but not to EFV and 3TC [Nikolenko et al 2004].

**Protease Inhibitors**

HIV protease enzyme mediates the maturation of newly formed HIV particles by proteolytic cleavage of gag and gag-pol precursors [Kohl et al 1988]. HIV is a logical target for the treatment of HIV infection. Inhibition of the enzyme function could stop the formation of new virus progenies. The detail knowledge of protease function and its three-dimensional structure has opened numerous approaches to rational design of inhibitors. Protease inhibitors are designed to bind specifically to the virus protease with high affinity although they occupy more space than the natural substrates [Wensing et al 2010]. There are nine protease inhibitors approved for clinical use: amprenavir, lopinavir, atazanavir, tipranavir, darunavir, indinavir, nelfinavir, saquinavir and ritonavir [show figure]. Except tipranavir, all protease inhibitors are competitive peptidomimetic inhibitors. They have hydroxyethylene core, which prohibits cleavage of the protease inhibitor by the HIV-1 protease [Craig et al 1991, Vacca et al 1994, Kempf et al 1995, Sham et al 1998, Koh et al 2003, Partaledise et al 1995, Robinson et al 2000]. Tipranavir has dihydropyrone ring instead of peptidomimetic hydroxyethylene core [Turner et a 1998] and can inhibit HIV isolates replication that are resistant to other protease inhibitors [Larder et al 2000, Back et al 2000].

The first generation protease inhibitors used are Saquinavir, Ritonavir, Indinavir and Nelfinavir. The limitation of first generation protease inhibitors was low bioavailability and short half-life, often required to multiple daily dose in-take. This reduced the patient adherence to the therapy, and subsequently increased drug failure [Bartlett et al 2001, 2006]. Two approaches were developed to fight problem of resistance selection by first generation inhibitors: first, increase the level of protease inhibitors in the plasma in combination with low-dose ritonavir and second, develop a novel protease inhibitor with high potency against known protease inhibitor resistant HIV variants [Wensing et al 2010]. The second generation of protease inhibitor therapy is the combination of first generation protease inhibitors with ritonavir. Ritonavir reduced the metabolism of first generation inhibitors, thus, increasing their level in the plasma [Zeldin and Petrus 2004, Youle 2007]. Patients with extensive drug resistant HIV have limited therapeutic options. They are put under double-boosted protease inhibitor therapy to increase the genetic barrier of HIV to drug failure [Ribera et al 2006, Gilliam et al 2006, Chetchotisakd et al 2007, Petersen et al 2007, Von Hentig et al 2007, Monosuthi et al 2008, Stebbing et al 2009].

Incorporation of protease inhibitor in ART has resolves many AIDS defining illness. However, patients consuming the inhibitors have developed lipodystrophy and hyperlipidemia. This is the side effect of the inhibitor as, besides inhibiting protease enzyme, it also inhibits degradation and secretion of apolipoprotein B; but in the presence of oleic acid, protease inhibitor increases the secretion of apolipoprotein B [Liang et al 2001].

**Integrase Inhibitors**

HIV integrase catalyzes the function of integrating the reverse transcribed DNA to the host genome. The enzyme cleaves two nucletides from 3’ end of the DNA, soon after the reverse transcription step [Sherman and Fyfe 1990, LaFemina et al 1991, Bushman and Craigie 1991]. The preintegration complex, formed after reverse transcription, is transported to the nucleus; integrase catalyzes the strand transfer process, in which it creates staggered nick in host chromosome and joins 3’ end of viral DNA of 5’ end of host chromosome [Grobler et al 2002].

Although the antiretroviral drug regimen, consisting of two NRTIs along with a NNRTIs or protease inhibitors (also called Highly Active Anti Retroviral Therapy - HAART), could suppress the viral replication, it is difficult for the patients to well-accept the drug regimen. It is expensive, leads to multiple drug resistance and requires high adherence [Cohen 2002]. This has diverted the search for drugs that targets third enzyme, the integrase. Researchers have demonstrated the integrase inhibitors feasibility and efficacy in Rhesus Macaques [Hazuda et al 2004].

Integrase inhibitors developed usually targets the strand transfer complex of the enzyme bound to the viral DNA, with the 3'-end [dinucleotide](http://europepmc.org/abstract/MED/18565342/?whatizit_url_Chemicals=http://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI%3A47885" \t "_blank) already cleaved. X-ray structure of the integrase enzyme has revealed the active site model of the enzyme complexed with the DNA [Chen et al 2008]. This model has application for designing wide range of potential integrase inhibitors. Diketo acids (DKA) and its derivatives are the highly potent integrase inhibitors for strand transfer inhibition [Hazuda et al 2000]. A very small concentration of DKA is required for effective integrase inhibitor. A drug 5-CITEP, derived from DKA, can inhibit integrase function at micromolecular concentration [Pluymers et al 2002]. Another compound, 4-Aryl-2, 4-dioxobutanoic acids, inhibits integrase activity at nanomolar concentration range [Wai et al 2000,].

Some integrase inhibitors with weak antiviral activity have been reported. Styrylquinolines inhibit 3′ processing capacity of integrase [Zouhiri et al 2000, Pommier et al 2005]. Another compound, thiazolothiazepine, also possesses weak [Neamati et al 1999, Pommier et al 2005].

Integrase inhibitors are added in the drug regimen for patients with viral quasispecies resistant to most of the reverse transcriptase and protease inhibitors.

**Entry Inhibitors**

HIV gp120 binds to the CD4 receptor [Dalgleish et al 1984, Sattentau and Weiss 1988] and CCR5/CXCR4 co-receptors [Moore et al 2004] to enter into the host cells. HIV entry inhibitors are designed either to inhibit the binding of gp120 to CD4 or co-receptors CCR5 or CXCR4.

HIV gp120 binding to CD4 forms an unglycolylated cavity in the gp120. CD4 phenylalanine residue 43 is the only residue that binds to this cavity and provides significant energy of about 23% of the total energy of CD4-gp120 binding [Kwong et al 1998, Madani et al 2004]. This cavity has been the primary target for designing small molecules that could bind it and inhibit HIV entry [Kwong et al 1998, 2000, Wyatt et al 1998]. Some of the molecules that inhibit gp120-CD4 binding are: PRP-542 [Jacobson et al 2000], TNX-355 [Moore et al 1992, Kuritzkes et al 2004], CADA [Vermiere et al 2002, 2003], BMS-806 [Madani et al 2004, Lin et al 2003, Guo et al 2003].

Another target of entry inhibitor is the gp120 conformational change to bind to co-receptors CCR5 or CXCR4. For CCR5 using virus, gp120 recognizes the N-terminal domain or the second extracellular loop of CCR5 co receptor [Wu et al 1997, Dragic et al 1998, Zhang et al 2007]; whereas for CXCR4 using virus, the gp120 only recognizes the second extracellular loop [Picard et al 1997]. Unavailability of co receptors for gp120 binding prevents HIV entry into host cell. Co receptor antagonists are designed that binds specifically to targeted co receptors, making it unavailable to HIV gp120 binding.

Some of the CCR5 antagonists developed are TAK-779 [Baba et al 1999], TAK-652 [Baba et al 2005], vicriviroc [Strizki et al 2005, Schurmann et al 2007, Gulick et al 2007], AD101 (SCH-350581) [Tsamis et al 2003], Maravirok (UK-427857) [Rosario et al 2005, De Clercq 2005, Fatkenheuer et al 2005, Rosario et al 2006, Wheeler et al 2007], Aplaviroc (GW-873140) [Nichols et al 2007, Latinovic et al 2009], PRO-140 [Trkola et al 2001].

It has been observed that HIV is capable of co receptor switch from CCR5 to inclusion or exclusively CXCR4 [Esbjornsson et al 2010] in late-stage of disease progression. HIV can use CXCR4 co receptor as the next door to enter the host cell. Therefore, the development of CXCR4 antagonists is essential.

Some of the CXCR4 antagonists either for clinical use or in clinical trial stage are AMD3100 [Donzella et al 1998], AMD070 [Schols et al 2003, Stone et al 2007], KRH-1636 [Ichiyama et al 2003] and KRH-2731 [Murakami et al 2004].

HIV envelope protein gp41 is also a possible target for drug design. Three different strategies for develop anti HIV molecules have been reported based on gp41 coiled coil region structural information. In near future, drugs targeting gp41 might be available [Jiang and Debnath et al 2000].

**Fusion Inhibitors**

Binding of gp120-CD4 and subsequently to co receptor change the confirmation in the viral envelope that shifts the gp41 from a non-fusogenic to a fusogenic state. This change in state drives HIV and host cell fusion. The exposed gp41 N-terminal domain gets inserted into cellular membrane through fusion peptide. The heptad regions HR1 and HR2 changes the free energy associated with six-helix bundle to make it thermo stable and this change in free energy is necessary for the fusion pore formation [Weiss 2003, Briz et al 2006] and entry into the host cell.

The design of fusion inhibitors is based on the targeting the heptad regions HR1 or HR2. Binding of the inhibitor to either HR1 or HR2 of gp41 makes it unable to make fusion pore. Enfuvirtide [Duffalo and James 2003, Poveda et al 2005] is a synthetic peptide, approved for clinical use in 2003 (commercial name-Fuzeon) [Robertson 2003], which can bind to HR1 region of gp41 [Wild et al 1993]. T-1249 is a second-generation fusion inhibitor drug that target different HR1 region of gp41 than Enfuvirtide [Kilby and Eron 2003]. T-1249 is active against Enfuvirtide resistance HIV strains, HIV-2 and SIV. However, it is discontinued in 2004 for clinical use [Briz et al 2006, Melby et al 2007]. Sifuvirtide is another HIV fusion inhibitor peptide under research [Wang et al 2009].

### Drug Resistance in HIV