CHAPTER-1

Non-Nucleoside Reverse Transcriptase and Integrase Inhibitors: An Overview
NON-NUCLEOSIDE REVERSE TRANSCRIPTASE AND INTEGRASE INHIBITORS: AN OVERVIEW

1.1 Introduction:

Acquired Immune Deficiency Syndrome (AIDS) symptoms were first observed in Pneumocystis carinii pneumonia infected gay patients in 1981 by the doctors of U.S. Centers for Disease Control (CDC) [1]. These cases were among young, healthier, homosexual men showing a group of rare diseases like Kaposi sarcoma, Pneumocystis carinii pneumonia, and persistent lymphadenopathy. They had a common immunological deficiency in cell-mediated immunity, the circulating CD4⁺ (Cluster of Differentiation No.4) T cells. Later on it was found that a retrovirus called Human immunodeficiency virus (HIV) is the causative agent in the transmission and development of AIDS [2]. In the beginning, the CDC did not have an official name for the disease, calling them as with name of disease that were associated with it for e.g., lymphadenopathy or Kaposi's Sarcoma and Opportunistic Infections. Before “Acquired immunodeficiency disease”, had been coined for these kind of symptoms various different names were proposed like GRID (Gay-related immune deficiency) 4H disease as the individuals infected were Haitians, homosexuals, hemophiliacs, and heroin users. By the time, depth of knowledge regarding this disease increased and it was found that aids was not isolated only from the homosexual community, the term GRID became misleading and “AIDS” was introduced in a meeting in July 1982. Since then HIV epidemic has been observed throughout the world especially in developing and underdeveloped countries. Epidemiological data suggest that the disease is spreading at an alarming pace and it is estimated that globally, approximately 5 million
people are newly infected with the virus. Global trends in HIV infection rates in different age groups from the year 2010 to 2013 are shown in Fig.1.1[3].

Figure.1.1 Trend in HIV infection rates (2010-2013).

In the year 2014, UNAIDS reported that India has the third-highest number in the world living with 2.1 million HIV infected people accounting for about four out of 10 people infected with the deadly virus in the Asia—Pacific region (Fig 1.2).

Figure.1.2 Present scenario of HIV infection in Asian countries (2014).
The spread of HIV in India has been varied, with much of India having a low rate of infection and the epidemic being most extreme in the southern half of the country and in the far north-east. The highest HIV prevalence rates are found in Maharashtra, Andhra Pradesh, Karnataka, Manipur and Nagaland [4].

It has been seen that HIV-2 is non-pandemic and is largely restricted to West Asia, suggesting that it is less efficiently transmitted than HIV-1. HIV-2 is less pathogenic than HIV-1 because HIV-2 infected persons seems to remain symptom free longer than persons infected with HIV-1. Studies have shown that HIV-2 positive persons are at decreased risk for acquiring HIV-1. Because of this difference in viral strains it has not been possible to have a single vaccine that could help in eradicating the disease at global level [5].

1.2 Objective

The objectives of this chapter are

- To discuss about the pathogenesis & HIV life cycle
- Talk about the therapeutic targets of HIV-1
- Briefly explore the reverse transcriptase and integrase enzymes along with their inhibitors followed by drug resistance.

1.3 HIV-1 Pathogenesis

The clinical features of HIV infection are acute primary infection syndrome, asymptomatic infection and symptomatic HIV infection and AIDS. In acute primary infection symptoms are flu-like illness, with high levels of virus replication until the infection is brought under
immune control. At this infection stage viral replication cycle is fast and doubles in ten hours time and peak viremia is observed at 21 days [6]. Virus load drops rapidly once the host immune response is initiated. Patients with asymptomatic infections have no outward sign of disease, although there is a slow decline in the CD4 count and very active viral replication. This stage can persist for more than ten years. In symptomatic infection patients develop AIDS like symptoms and as disease progress result in death due to opportunistic infections.

It has not been possible as yet, to completely understand the relationship between HIV replication and HIV pathogenesis. It is generally believed that HIV pathogenesis initiates from the time of HIV infection and initiation of HIV replication. The viral proteins expressed interact with host proteins to perform their respective functions in viral replication cycle. An understanding of the dynamic interplay of host cells and virus responsible for viral replication and virus survival is the need of the hour. Until then, the clinical situation is that the adequate control of viral replication is not sufficient for eradicating virus from the host.

1.4 HIV-1 Life Cycle

A schematic representation of the virus life cycle is shown in Fig 1.3. The entire life cycle can be broadly classified into early phase and late phase. The early phase begins with CD4 recognition and involves events up to and including integration of the proviral DNA into the host genome, and the late phase includes all events from transcription of the integrated DNA to virus budding and maturation. As a result of concerted effort by many research
groups all biochemical steps in the virus life cycle have been thoroughly understood. This has led an array of druggable targets [7].

Figure 1.3 Life cycle of HIV-1.

1.5 Therapeutic Targets:

There are many events in the HIV replication cycle that are potential targets for therapeutic intervention as shown in Fig 1.4[8]. Majority of the targets have been successfully exploited to obtain one or more drugs in clinical usage. Majority of these drugs are inhibitors of reverse transcriptase, integrase and protease. Since HIV-RT and integrase inhibitors are the central theme of the present dissertation, these are articulated in detail so as to enable better appreciation of the results discussed in chapters 2, 3, 4 & 5.
1.6 HIV-1 Reverse Transcriptase Enzyme:

Reverse transcription is the conversion of an RNA genome into a double stranded DNA molecule. This process was first reported by Baltimore, Temin and Mizutani for RNA tumor viruses. All viruses that depend for their replication on reverse transcription have been grouped in the family of the Retroviridae. Reverse Transcriptase (RT) is a key enzyme, packaged within the HIV-1 virion capsid, which plays an essential role in the reverse transcription of the HIV-1 genome [7, 8]. The enzyme RT is used by retroviruses to transcribe their single stranded RNA genome into single stranded DNA and to subsequently
construct a complementary strand of DNA, providing a DNA double helix capable of integration into the host cell genome.

Functional HIV-RT is a heterodimer containing sub units of p66 and p51 as shown in Fig. 5. The upper and lower portions of the dotted line correspond to the p66 and p51 domains respectively. The dotted line corresponds to the interacting domains. The HIV-1 reverse transcriptase exhibits both N-terminal DNA polymerase domain (440 residues) and C-terminal RNaseH domain (120 residues) activities to complete the reverse transcription. The x-ray co-crystal structure of the enzyme HIV-1 RT with ds-DNA has revealed that a “template grip” could be designated to portions of the p66 palm and finger that are closely associated with nucleotides of the template strand (β3, β2- β3 loop, α2 of the fingers and β6-α5 loop in the palm, Fig 1.5)[9-11].

The NNRTI-BP may normally function as a hinge between the palm and thumb sub-domains of p66 subunit of HIV-1 RT and the mobility of the thumb may be important to facilitate template/primer (T/P) translocation during DNA polymerization. It has been suggested that the binding of NNRTIs may restrict the mobility of the thumb sub-domain (the “arthritic thumb” model) thus slowing or preventing T/P translocation and thereby inhibiting facile elongation of nascent viral DNA [12-16]. A recent steered molecular dynamics study on NNRTI binding supports this hypothesis [13].
Figure 1.5 The three-dimensional structure of the unliganded HIV-1 RT with the numbered indications of the structural elements. Finger domains are indicated in blue, the thumbs in yellow, the palm domains in green, the connections in red and the RNaseH domain in purple.

1.7 HIV-RT Inhibitors

The existing inhibitors of HIV-RT can be subdivided into two major classes: first, termed nucleoside reverse transcriptase inhibitors (NRTIs), which binds competitively and co-valently at the enzyme active site, and inhibits polymerization [15]. The NRTIs compete with the dNTP substrate, and get incorporated into the primer strand. The NRTIs lack the 3-OH group, causing DNA chain termination. The second class of inhibitors comprises non-
nucleoside reverse transcriptase inhibitors (NNRTIs), which binds non-competitively and non-covalently at allosteric site on the enzyme different from the enzyme active site [17, 18]. NNRTIs bind to the same hydrophobic pocket that is located between the β-sheets of the palm and at enzyme-substrate complex more efficiently than enzyme itself and change the conformation of enzyme to inhibit polymerization (Fig 1.5)[19]. More importantly they are specific to the HIV-1 RT and therefore are less likely to be toxic to the host cell. It is surmised that specificity is because of the presence of a flexible hydrophobic pocket in the HIV-1 RT, which is absent in the reverse transcriptase from other viruses. Inhibitors of the NNRTI class have attracted the attention of medicinal chemists because of their highly specific binding to the HIV-1 reverse transcriptase; some important scaffolds with one example each are discussed below.

*Dipyridodiazepinones (Nevirapine analogues):*

In the year 1988, Nevirapine got the first regulatory approval as HIV-1 NNRTI. Nevirapine has a potent activity against wild-type RT, with a 50% inhibition at 80 nM concentration (Fig 1.6) [20]. Along with this, it shows good metabolic stability, good bioavailability and crosses the blood–brain barrier easily [21].

![Figure 1.6 The structure of nevirapine and its analogues.](image-url)
Different substituted analogues were made to overcome the drug resistance problem but none showed promising activity against the mutants. Among these, the dipyrido[2,3-b:2′,3′-c]diazepinones[22], pyridazinobenzodiazepines [23] and pyridobenzodiazepines [20] showed an improved anti-HIV potency in resistant cell lines. As a class, oxazepinones showed poor anti HIV-1 activity than the diazepinones due to poor solubility.

Benzoxazinone (Efavirenz analogues):

Figure 1.7 Benzoxazinone analogues

Efavirenz is a Benzoxazinone derivative, approved by the FDA in the year 1998 (Fig 1.7). It is a very potent inhibitor against a wide range of mutant HIV-1 RTs. However, it gave EC$_{95}$ >1.5 µM [24]. Most of the mutants are inhibited by efavirenz at lower concentrations. In spite of that efavirenz shows poor activity against K103N mutant. To overcome this problem two series of efavirenz analogues have been identified: one in which the cyclopropane ring has been replaced by small heterocycles and acetylenic side chain has been replaced by alkyl groups. A number of derivatives belonging to these series showed equal potency to efavirenz against both wild-type virus and K103N mutant [25].


*Bis(heteroaryl)-piperazine derivatives (Delavirdine analogues)*:

![Image of Bis(heteroaryl)-piperazine derivatives](image.png)

**Figure 1.8** General structure of Bis(heteroaryl)-piperazine derivatives.

Bis(heteroaryl)piperazine (BHAP) was discovered by the Upjohn laboratories as NNRTIs (Fig 1.8). Delavirdine is a BHAP analogue which was given FDA approval in April 1997 for therapeutic use against HIV-1. Atevirdine mesylate (U87201E) is another compound which was chosen for clinical evaluation [26]. Delavirdine shows ED50 value of 0.066±0.137 µM against wild type HIV-1[27].

*Diarylpyrimide (Rilpivirine analogues)*:

![Image of Diarylpyrimide analogues](image.png)

**Figure 1.9** General structure of diarylpyrimide analogues.

Rilpivirine is a new generation NNRTI compound which was approved by the FDA for anti HIV-1 therapy in 2011(Fig 1.9). A new class of compounds, diarylpyrimide (DAPY), were
identified by scaffold hopping. Here, the central triazine ring of diaryltriazine (DATA) compounds was replaced by a pyrimidine moiety [28]. This new class was more useful against drug resistant HIV-1 strains than other DATA analogues.

1-(2-hydroxyethoxymethyl)-6-(Phenylthio)thymine (HEPT): 

![General structure of HEPT analogues.](image)

**Figure 1.10** General structure of HETP analogues.

The HEPT analogue (TS-II-25) was the first NNRTI that exclusively interact with an allosteric site located at some 10 Å from the catalytic site of the HIV-1 reverse transcriptase. The interaction between Tyr181 of RT and 6-benzyl ring of the inhibitors would play an important role for activity (Fig 1.10)[29]. Research efforts by the Triangle Pharmaceuticals led to identification of MKC-442 which has moved quickly to phase III clinical trials. The HEPT analogues are disreputable for resistance after mutations particularly in case of K103N and Y181C in the HIV-1 RT. Therefore, to overcome this problem several new analogues of MKC-442 were synthesized e.g., TNK-6123 and GCA-186 [30]. Among them, compound TNK-6123 showed promising activity due to the presence of cyclohexylsulfanyl group which provides more flexibility at mutated drug-binding site.
**Dihydroalkoxybenzyloxopyrimidines (DABO):**

![General structure of DABO analogues.](image)

R1 and R3 are aliphatic substituents  
R2 is aliphatic/ aromatic substituents  
X = S, SO2, NH, CH2

**Figure 1.11** General structure of DABO analogues.

The dihydroalkoxybenzyloxopyrimidines (DABOs) and their thio analogues (S-DABOs) are structurally related to the HEPT derivatives. SAR studies of DABO have shown that best activities are achieved with bulky substituents for R2, such as sec-butyl, cyclopentyl or cyclohexyl (Fig 1.11)[31]. Replacing oxygen linker with sulphur at X position led to the S-DABO compounds which showed improvement in anti-HIV-1 activity [32]. In this series, 5-isopropyl-2-[(4’-methoxy-phenyl-carbonylmethyl)thio]-6-(1-naphthylmethyl) pyrimidin-4(3H)-one exhibited most promising inhibitory activity against HIV-1 with IC$_{50}$ 0.030µM, CC$_{50}$ 203 µM with selectivity index 6766 [33].

**PETT / Urea-PETT:**

![PETT / Urea-PETT](image)

R/R$_1$ = aliphatic / aromatic substituents  
R$_2$ = aromatic substituents  
R$_2$/R$_2$R$_3$ = aliphatic substituents
Phenylethylthiazolylthiourea (PETT) derivatives were designed by systematic arrangement of the molecular architecture of TIBO HIV-1 RT inhibitors (Fig 1.12) [34, 35]. The lead compound of this series is LY73497. After structural optimization, trovirdine (N-[2-(2-pyridyl)ethyl]-N-[2-(5-bromopyridyl)] thiourea hydrochloride) was developed which is currently in clinical trials for potential use in the treatment of HIV-AIDS. This compound displayed IC$_{50}$ against wild-type RT at 5.0 nM. Furthermore, the development of NNRTIs in the urea analogue series of PETT was investigated. The urea-PETT compounds may have better toxicological and pharmacokinetic profile than PETT derivatives. A cis configured cyclopropyl linker was used to improve the activity of urea-PETT analogues [36].

4,5,6,7-Tetrahydroimidazo[4,5,1-jk][1,4]benzodiazepin-2(1H)-one and -thione (TIBO) Derivatives

![Figure 1.13 General structure of TIBO derivatives.](image)

The 4,5,6,7-tetrahydro-5-imidazo[4,5,1-jk][1,4]-benzodiazepin-2(1H)-one and -thione (TIBO) analogues were discovered by rational screening of 600 compounds containing...
different chemical classes followed by structural modifications (Fig 1.13)[37]. It was observed that a thione group is important for anti HIV-1 activity. The most potent compound showed an EC\textsubscript{50} value of 0.042 µM in MT-4 cells. HEPT and TIBO were the first new category of HIV-1 NNRTIs. The most potent molecule among all TIBO derivatives (tivirapine) [38] was not endorsed further for clinical use due to its higher toxicity.

**Thiazolobenzimidazole:**

![General structure of TBZs](image)

Figure 1.14 General structure of TBZs.

Chimirri in 1991 reported 1-H, 3H-thiazolo[3,4-a]benzimidazoles (TBZs) as HIV-1 NNRTI agents (Fig 1.14)[39, 40]. The most active compound in this series showed EC\textsubscript{50} 0.356 µM confirmed a potent inhibitor of HIV-1 RT. Another potent compound was trans-1-(2,6-difluorophenyl)-3-methyl-1H, 3H-thiazolo[3,4-a]benzimidazole had an EC\textsubscript{50} 0.3±0.01 µg/ml against the HIV-1 in MT-4 cell lines.
1.8 Resistance to HIV-1 Non-Nucleoside Reverse Transcriptase Inhibitors

The rapid replication of HIV and its inherent genetic variability has lead to the generation of viral variants that exhibit drug resistance inevitably [41]. The emergence of resistance towards currently used anti-HIV drugs (NRTIs, NNRTIs, and Protease inhibitors) is an important determinant in the eventual drug failure, new drug development strategies are attempting at circumventing the virus-drug resistance problem by focusing on either novel targets (other than the reverse transcriptase and protease) or new compounds capable of suppressing HIV strains that are resistant to the currently used reverse transcriptase or protease inhibitors.

Though NNRTIs represent an important component of anti-HIV chemotherapy, clinical utility of the NNRTIs is adversely affected by the emergence of drug-resistant HIV-1 variants. Common RT mutations that confer resistance to NNRTIs include L100I, K103N, V106A/I/L, Y181C, and G190A/T/V [42, 43]. Different mutations have very different effects on different inhibitors. Mutations (located in the reverse transcriptase) that are associated with resistance (or reduced susceptibility) to the NNRTIs are listed in Table 1.

The NNRTIs interact with a specific “pocket” site of the HIV-1 RT, which is closely associated with, but distinct from, the substrate binding site. NNRTIs are notorious for rapidly eliciting resistance, resulting from mutations at the amino acid residues that surround the NNRTIs binding site of HIV-1 RT as shown in the Table 1.1 & Fig 1.15. However, emergence of NNRTI-resistant HIV strains can be prevented if the NNRTIs are combined with NRTIs and used from the beginning at sufficiently high concentrations.
Table 1.1 Principle RT mutations involved in resistance to NNRTIs

<table>
<thead>
<tr>
<th>RT mutations</th>
<th>Codon change</th>
<th>Amino acid change</th>
</tr>
</thead>
<tbody>
<tr>
<td>A98G</td>
<td>GCA→GGA</td>
<td>Ala→Gly</td>
</tr>
<tr>
<td>L100I</td>
<td>TTA→ATA</td>
<td>Leu→Ile</td>
</tr>
<tr>
<td>K101E</td>
<td>AAA→GAA</td>
<td>Lys→Glu</td>
</tr>
<tr>
<td>K103N</td>
<td>AAA→AAC</td>
<td>Lys→Asn</td>
</tr>
<tr>
<td>V106A</td>
<td>GTA→GCA</td>
<td>Val→Ala</td>
</tr>
<tr>
<td>V108I</td>
<td>GTA→ATA</td>
<td>Val→Ile</td>
</tr>
<tr>
<td>E138K</td>
<td>GAG→AAG</td>
<td>Glu→Lys</td>
</tr>
<tr>
<td>V179D</td>
<td>GTT→GAT</td>
<td>Val→Asp</td>
</tr>
<tr>
<td>Y181C</td>
<td>TAT→TGT</td>
<td>Tyr→Cys</td>
</tr>
<tr>
<td>C181I</td>
<td>TGT→ATT</td>
<td>Cys→Ile</td>
</tr>
<tr>
<td>Y188C</td>
<td>TAT→TGT</td>
<td>Tyr→Cys</td>
</tr>
<tr>
<td>Y188H</td>
<td>TAT→CAT</td>
<td>Tyr→His</td>
</tr>
<tr>
<td>H188L</td>
<td>CAT→CTT</td>
<td>His→Leu</td>
</tr>
<tr>
<td>G190A</td>
<td>GGA→GCA</td>
<td>Gly→Ala</td>
</tr>
</tbody>
</table>
Figure 1.15: Schematic view of the V-loop with two “horizontal axes” that are formed by the non nucleoside binding sites Tyr188 and Tyr188. The two aspartyl residues Asp185 and Asp186 of the catalytic triad (Asp110, Asp185, Asp186) are indicated. The strong hydrogen bonds of the peptide backbone that maintain the constrained loop are between Tyr181 and Tyr188, and within the internal cage. Rotations of the side chains of Tyr181 and Tyr188 help to create a cavity to accommodate the NNRTIs during their entry process into the pocket.

1.9 The HIV-1 Integrase Enzyme:

HIV-1 integrase (IN) is one of the key enzymes in integration of viral DNA into the host genome [44]. It is likely to transfer viral DNA into the host chromosome during retroviral
replication [45]. IN has no equivalent in the host cell. Therefore, integrase inhibitors have a high selective index since they do not make any hindrance in normal cellular processes [46].

1.10 Structure of Integrase Enzyme:

IN belongs to the family of polynucleotidyl transferases and composed of 288 amino acids to form the 32 kDa protein. Retroviruses encode with the POL gene with the 3’end encoding for IN.

![Figure 1.16: Structural domains of the HIV-1 integrase](image)

**Figure 1.16**: Structural domains of the HIV-1 integrase

IN contain three structurally independent, functional domains as shown in Fig. 1.16.

1. **The N-terminal domain (NTD)**: It includes amino acids 1 to 50 and contains two histidine residues (His12 and His16) and two cysteine residues (Cys40 and Cys43), these two amino acids are completely conserved and form a HHCC zinc-finger pattern [45, 47]. Single point mutations of any of these four amino acids reduce IN enzyme activity gradually. NTD has a primary role in multimer formation which requires zinc atom that stabilizes the fold [47].
Chapter 1  Non-nucleoside Reverse Transcriptase and Integrase Inhibitors: An Overview

(2) The catalytic core domain (CCD): This domain consists of 51-212 amino acids and contains the active site of IN but unable to catalyze integration in the absence of NTD and CTD (the C-terminal domain) [47]. It has three conserved negatively charged amino acids (D64, D116 and E152). These three amino acids make DDE motif that coordinate divalent metal ions (Mg\(^{2+}\) or Mn\(^{2+}\)) which are crucial for the catalysis of viral DNA integration [47,48]. CCD has six α-helices and five β sheets that are connected by flexible loops [47]. The flexible loops are allowed for conformational changes which are important for 3´processing of the viral DNA and strand transfer (STF) reactions. It is reported that substitution of any of the residues in the DDE motif significantly inhibits the IN activity [47].

(3) The C-terminal domain (CTD): Here, this portion is composed of 213–288 amino acids. CTD is the least conserved of the three domains. It binds to DNA non-specifically and interacts with NTD and CCD domains during integration [46, 47].

1.11 Role of Integrase Enzyme:

Multistep process are used in HIV-1 integration including 3´ endonucleolytic processing of proviral DNA ends (termed 3´ processing) and integration of 3´-processed viral DNA into the host cellular DNA (referred to as strand transfer). At the time of 3´ processing IN binds near to long terminal repeat (LTR) of the viral DNA and catalyzes endonucleotide cleavage. This confirms the elimination of a dinucleotide from each of the 3´ ends of the LTR. After that, cleaved DNA is used as a substrate for integration [44]. Strand transfer is a trans-esterification reaction including a direct nuleophilic attack of the 3´hydroxyl group of viral 3´-DNA ends on the phosphodiester backbone of the host DNA [49]. The integration
reaction ends with removing of unpaired dinucleotides from the 5'- ends of the viral DNA [45, 49]. However, Mg$^{2+}$ or Mn$^{2+}$ are required for 3'-processing and strand transfer steps for assembly of viral DNA by IN. Among them, magnesium is a preferred divalent cofactor for integration process.

1.12 HIV-1 Integrase Inhibitor:

*Integrase Strand Transfer Inhibitors (INSTIs):*

![Structure activity relationship of elvitegravir and raltegravir. A benzyl group in a hydrophobic pocket and a triad to chelate the two Mg$^{2+}$ ions.](image)

*Mg$^{2+}$ and Mn$^{2+}$ are significant cofactors for integration process (Fig. 1.17). Removal of these cofactors can cause functional impairment of IN. This idea gives researchers the opportunity to design and develop potent IN inhibitors (INIs). Indeed, all small molecule HIV-1 INIs are showing a structural shape that coordinates the two divalent ions at the catalytic site [50]. Present marketed drugs (Raltegravir and elvitegravir) bind to the catalytic site and make a coordinate complex with Mg$^{2+}$ ions. By this mechanism, they can directly compete with viral DNA for binding to integrase in order to inhibit 3’-end processing [51]. This inhibition is called strand transfer inhibition.*
1.13 Drug Resistance problem with INSTIs:

It has been identified that more than 60 variations of INSTI mutations cause in *in vitro* as well as *in vivo* resistance (Table 1.2). Therefore, inhibitors are less effective against the integrase enzyme to protect the human genome [44].

**Table 1.2** Common drug resistance mutations of INSTIs are expressed as fold change.

<table>
<thead>
<tr>
<th>Class</th>
<th>Mutations</th>
<th>Raltegravir</th>
<th>Elvitegravir</th>
<th>Dolutegravir</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>T66R</td>
<td>+</td>
<td>+++</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>E92Q</td>
<td>+</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>E138K</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Y143C</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Y143R</td>
<td>+++</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Q148H</td>
<td>++</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Q148K</td>
<td>+++</td>
<td>+++</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Q148R</td>
<td>+++</td>
<td>+++</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>N155H</td>
<td>+++</td>
<td>+++</td>
<td>-</td>
</tr>
<tr>
<td>Secondary mutations</td>
<td>T97A</td>
<td>G118R</td>
<td>G140S</td>
<td>R263K</td>
</tr>
<tr>
<td>---------------------</td>
<td>------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-</td>
<td>+++</td>
<td>++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Common drug resistance mutations of INSTIs are expressed as fold change. Fold change scores are designated as follows: <2, −; 2–10, +; 10–20, ++; 20–50, +++; 50–100, ++++; >100, +++++.

Primary mutations observed in loop region containing amino acid residues 140–149 which is positioned into the catalytic-core domain and responsible for protein folding and flexibility as well. Even though the exact role of this loop is not quite known but it is
thought to be important for DNA binding. Present drugs like raltegravir and elvitegravir share almost similar type of mutational problem. Further studies point out that raltegravir has a partial intestinal absorption. Thus, drug resistance issue is unable to control even by prescribing higher doses also. Moreover, other mutations are also involved for different drug like molecules [45]. A few mutations can boost the resistance to the drugs up to a large extent than other mutations. For example, one of the most common mutation pathway increases the resistance to raltegravir up to 100 times more than the second most common one [44]. To control the drug resistance, S/GSK1349572 is newly identified as an integrase inhibitor by ViiV/Shinongi which was in phase III clinical trials in 2011[52]. Recently it is reported that S/GSK1349572 developed resistance which is not been entirely categorized. However, it could be useful against some of the primary mutated strains.

1.14 Other binding sites of Integrase Enzyme and its inhibitors:

Since drug resistance issue could not be solved, researchers are looking for the allosteric sites of the enzyme to inhibit the activity of the enzyme and minimize the drug resistance problems as well. Different binding sites were identified other than the active site. Among them LEDGF/p75- integrase interaction site is promising one for further exploration since there are no reports of mutations in this area. Therefore, in this review, we mainly focused on this binding site to explore protein-protein interaction inhibitors to find out different structural scaffolds.

LEDGF/p75- integrase binding inhibitors:

Lens epithelial derived growth factor (LEDGF/p75) is a host protein which binds to integrase enzyme and induces viral replication. The mechanism of action is still exactly not known but it is believed that LEDGF/p75 helps in post integration step for incorporating
viral DNA into the active sites of the host genome. To stop this protein-protein binding interaction, recently different types of inhibitors are being developed and patented. Researchers are extremely interested to explore this target because of specificity and less prone to the drug resistance [53].

Structure and function of Lens epithelial derived growth factor (LEDGF/p75):

Figure 1.18 General structure of domain organization of LEDGF and HIV-1 integrase along with crystal structure (PDB ID: 2B4J) of their binding interactions.

HIV-1 virus contain limited genome, as a result it is depend upon cellular co-factors for provirus establishment and efficient replication in the host cell [54]. Lens epithelium-derived growth factor (LEDGF)/p75 is an imperative cellular co-factor for HIV-1 replication which is most likely the only molecule interacting with integrase whose knock-down harshly affects the HIV-1 integration levels. In the year 2003, it was originally recognized from over expressing HIV-IN cells by Cherepanov and co-workers [55]. This can suppress HIV-1 expression in infected cells. It is a 60kDa transcriptional co-activator encoded by PC4 and SFRS1 interacting protein 1(PSIP1) [56]. LEDGF/p75 is a part of the hepatoma-derived growth factor family (HDGF) and composed of 530 amino acids (Fig. 1.18). This protein emerge to function as a tethering factor, linking IN to the cellular
chromatin. The N-terminus of the protein composed of a Pro–Trp–Trp–Pro (PWWP) domain which is a nuclear localization sequence (NLS) [57]. This domain has a necessity for its interaction with chromatin whereas the C-terminal of LEDGF/p75 hold the integrase-binding domain which interacts with HIV-1 integrase (Fig.1.18). Crystallographic data identified the amino acids including K364, I365, D366, F406, and V408 of LEDGF/p75 are accustomed to make a relevant interactions with IN (Fig.1.18). The main role of this protein (i) evidenced via mutagenesis (ii) RNAi-mediated depletion (iii) trans-dominant over expression of the integrase binding domain (IBD) of LEDGF/p75 and (iv) knock out studies [58-59]. It is also connected with two proteins (Spt6 and Iws1) which are used to manage the gene expression and chromatin structure to form a stable complex and regulate the post integration steps of HIV-1 latency [59]. Iws1 plays a crucial role in the establishment of latent infection, whereas the purpose of Spt6 is to employ Iws1 and LEDGF/p75 to the silenced provirus and preserve histone occupancy at the HIV-1 promoter.

**Blocking of LEDGF/p75-IN interactions by peptides**

![Figure 1.19](image-url) Crystal conformations of Peptide analogues as a LEDGF/p75-IN interactions inhibitor.
To target the LEDGF/p75 interaction site, initial work focused on designing small peptides due to its linearity of amino acid sequence with target specific binding site (Fig. 1.19)[60]. It can interact with the LEDGF/p75 binding domain and change the oligomerization state which is used to affect the catalytic activity of IN indirectly. Hayouka and co workers described the design and synthesis of three LEDGF/p75 derived peptides (LEDGF/p75 353–378, 361–370 and 402–411) [61]. A similar type of peptide (LEDGF/p75 355–377) was able to compete with LEDGF/p75 for the binding to integrase and consequently inhibiting the cofactor-IN interaction with an IC$_{50}$ of 25 µM in a alpha screen based interaction assay [60]. The mechanism of action of small peptides (LEDGF/p75 361–370) was well explored by biophysical, biochemical and cellular assays [62]. The LEDGF/p75 derived peptide was cyclised further in a later study to increase its inhibitory activity [Fig.1.19, 63]. Furthermore, Rhodes and co workers used the cyclic peptide in order to identify new interactions at the interface of LEDGF/p75–IN interaction [63]. Herein, a cyclic peptide was produced by the LEDGF/p75 362–367 sequence and fused with three other amino acids which shows strong H-bond interaction with Gln168 of protein-protein interface [PDB ID: 3AVB, Fig.1.19]. From crystallography data of total 13 cyclic peptides, it was revealed that seven residues present in LEDGF/p75 make contacts with integrase binding with total surface area of 420 Å [63]. Therefore, specific size of cyclic peptides can provide new peptidomimetic drugs.
Blocking of LEDGF/p75-IN interactions by small molecule inhibitors

*N-acylhydrazone-based derivatives:*

![N-Acylhydrazone derivatives](image)

**Figure 1.20** Different types of N-acylhydrazone-based derivatives as LEDGINS.

N-Acylhydrazone is identified for anti tuberculosis and antiretroviral therapies [ref]. It is identified as a potential new scaffold by using LEDGF/p75 IBD-based pharmacophore models [64]. The central acylhydrazone moiety plays a decisive role in the activity. Compounds showns in Fig.20, with IC$_{50}$ of 1.3µM, 0.4µM and 2 µM respectively point out their exclusive selectivity towards LEDGF/p75 – IN interaction site [64]. It is also suggested that these compounds are solely binding to the LEDGF/p75-IN interaction site to disrupt the LEDGF/p75-IN complex formation. Molecular docking studies also showed their affinity to this allosteric binding site and revealed that incorporation of the hydroxyl group at 2 position forms an important H-bond with the target IN dimer interface which disrupt the LEDGF/p75-IN complex by locking the integrase into an inactive state [64].
Hydrazines, diazenes, and related compounds:

To explore the importance of nitrogen in central acylhydrazones linker, different hydrazine and diazenes related molecules were prepared to look at the affects of the two central nitrogen atoms by altering the position of the double bond [64]. It has been reported that LEDGF/p75-IN selective inhibitor 4-(2-(1-(2-bromophenyl)ethylidene)hydrazinyl)-N-(2-methoxyphenyl)-3-nitrobenzene sulfonamide block the protein-protein interaction with an IC$_{50}$ value of 6µM $\textit{in vitro}$. Similarly compound (IC$_{50}$ 3µM) also confirmed the importance of hydrazine into central portion of the leading molecules. In case of diazine linker (IC$_{50}$ 20µM) also researchers found substantial selectivity to LEDGF/p75-IN interaction site [64]. Additionally, hydrazine like moieties is also providing to inhibit the intermolecular interactions between LEDGF/p75 and integrase CCD domain.

Figure 1.21 Hydrazines, diazenes, and related compounds.
2-(tert-butoxy)-2-substituted acetic acid derivatives

Figure 1.22 2-(tert-butoxy)-2-substituted acetic acid derivatives as anti HIV-1 IN compounds.

Different available crystal structures (CCD domain of HIV-IN) based pharmacophore studies explored a new class of 2-(tert-butoxy)-2- substituted acetic acid derivatives (CX014442 & CX05045) [65]. These data confirm the importance of steric and electronic features of active molecules which can bind to the hydrophobic pocket of LEDGF/p75-IN dimer (CCD domain) interface. Furthermore, Co-crystals confirmed the initial pharmacophore. Additionally, provide the importance of Integrase Ala128 which makes extensive Van der Waals contacts with LEDGF/p75 ( Ile365 and Leu368), this features help the researchers to promptly design more potent LEDGINs with improved biological activities, such as CX05045 [65] and CX014442 [65], allowing for a complete antiviral profiling of this compound class. CX014442 is the first LEDGIN reported to display antiviral activity in the low nanomolar range, EC<sub>50</sub> IIIB value 6.97 µM and high selectivity,
SI value 1391. The exclusive part of their mechanism of action is indeed not only blocks provirus integration but also affects the infectivity of the residual progeny virus. Recently the 2- (tert-butoxy)-2-substituted acetic acid derivatives are the best studied LEDGINs compounds which are in advanced preclinical development.

*Benzo (1,3)dioxole-4-carboxylic acid derivatives:*

![Figure 1.23 General structures with derivatives of Benzo (1,3) dioxole-4-carboxylic acid.](image)

Benzo (1,3)dioxole-4-carboxylic acid derivatives shows promising inhibitory action as a LEDGINS. Crystal structure of the its analogue verify that 1H-inden-2(3H)-one occupied the same position in the IN pocket (PDB ID 3ZT3) as residue Ile365 in the LEDGF loop [66]. Also, the carboxylic acid of the compound makes a virtually identical interaction to Asp366 of the LEDGF loop . Presence of this interaction directly reflected in its biological activity against integrase enzyme (in the AS assay 270 µM). This key interactions were used to develop a new series of analogues. It has been identified that the reduction of the carbonyl group of lactone ring to hydroxyl can improve the activity (in the AS assay 200 µM) to some extent. Further assessment suggested that a seven membered ring instead of...
five member lactone ring would more effectively packed into the binding pocket. For example, researchers were synthesized several 2,3,4,5-tetrahydro-1H-benzo[e][1,3]diazepine-1-one based analogues which were showed activity in the AS assay at 110 µM, but due to the higher cellular toxicity, actual activity could not be determined. In addition, it was found that ring-opened structure provide significantly better result with lower toxicity. Therefore, researchers synthesized a number of analogues of this series. Finally, few compounds were obtained with excellent activity. Among them some of the compounds showed target specific inhibitory activity at 8µM which is the best molecules in that series. Co-crystal structure (PDB ID: 3ZSO) of this compound point out the amine of 11 sits deeper in the LEDGF/p75 hydrophobic pocket and makes a hydrophobic interaction as well as hydrogen bond with the HIV IN backbone (the carbonyl of Gln168) which may useful to overcome drug resistance.

_Catechol derivatives:_

![Figure 1.24 Catechol analogues.](image)

To design a new set of small molecules scaffold hopping approached was used by the researchers to inhibit the activity of catalytic site and LEDGF/p75–IN interaction subsequently. To get a new scaffold molecular hybridization was used by adding salicylate and catechol moieties to form 2,3-dihydroxybenzamide[77]. This active scaffold
spectacularly inhibited the interaction between IN and the LEDGF/p75 cofactor. The sample example, N-(cyclohexylmethyl)-2,3-dihydroxy-5-(piperidin-1-ylsulfonyl)benzamide, inhibited the IN–LEDGF/p75 interaction with an IC50 value of 8 µM [67]. This compound was found to develop a strong interactions with IN residues in the LEDGF/p75 binding site without inducing any cytotoxicity in normal cell lines. Therefore, further structural modification is required for synthesized compounds to get better in vitro activities as well as antiviral activity.

1.15 Conclusion:

The investigation for new ways to improve the treatment of HIV-1 affected patients is a continuous process. Development of a successful Anti retroviral therapy was accomplished when NNRTIs and INIs were discovered. The above mentioned examples exhibit that there is a wonderful range in the structural diversity of compounds that are active against HIV-1 by inhibiting the reverse transcriptase and integrase activities. NNRTIs emerge to be hopeful therapy against HIV-1, when it is used in combination with other anti HIV-1 drugs such as nucleoside RT inhibitors integrase and protease inhibitors. The common problem with first generation of NNRTIs and INIs is rapid drug resistance due to mutations. To mitigate this drug resistance, second generation of NNRTIs (ex.Rilpivirine, Etravirine) and INIs (MK-2048,S/GSK1349572) are developed. Although the leading molecules mentioned above are promising for the development and has a long way to go to find out the efficacy, safety and narrowly focused mechanism of actions.