CHAPTER I
INTRODUCTION

1.1 Acquired Immune Deficiency Syndrome (AIDS)

1.1.1 History

In 1981, the first article related to Acquired immune deficiency syndrome (AIDS) was published by Michael Gottlieb and coworkers [1]. The article reported about a random increase in pneumocystis carinii pneumonia (PCP), a rare lung infection. The Centers for Disease Control and Prevention (CDC) noticed that there was an unusual high number of requests for the drug that treated PCP. A short while later, another article reported eight outbreaks of Kaposi's Sarcoma (KS) in young homosexual males in New York. This was surprising because Kaposi's Sarcoma was a rare form of cancer that normally showed up in older people. At that time, the medical community realized that a new disease was probably heading their way. In 1982, the term Acquired Immune Deficiency Syndrome or AIDS was named by CDC. The causative agent of AIDS was later identified in 1983 by Luc Montagnier of France and Robert Gallo of the United State and named human immunodeficiency virus (HIV).

1.1.2 The Global HIV epidemic today

Since its first reported in 1981 with a small number of patients (there were 1,600 people diagnosed with the disease and almost seven hundred deaths), AIDS has now become a major epidemic with more than 39.5 million people worldwide living with HIV (Table 1.1). In 2006 alone, there were 4.3 million new HIV-1 infections and 2.9 million AIDS deaths. These estimates mask the dynamic nature of this evolving epidemic in relation to temporal changes, geographic distribution, magnitude, viral diversity, and mode of transmission. Today, there is no region of the world untouched by this pandemic. Although promising developments have been seen in recent years in global efforts to address to effective treatment and prevention programs, the number of people living with HIV continues to grow as does the number of deaths due to AIDS.

Table 1.1 Global summary of the AIDS epidemic in 2006 (data as of December 2006), the number is an average while the number in parenthesis is for range. The data was taken from the jointed United Nations programme on HIV/AIDS (UNAIDS) 2006 [2].

		Total (million)
People living with HIV	Total	39.5 (34.1 – 47.1)
	Adults	37.2 (32.1 – 44.5)
	Children under 15 years	2.3 (1.7 – 3.5)
People newly infected with HIV	Total	4.3 (3.6 – 6.6)
	Adults	3.8 (3.2 – 5.7)
	Children under 15 years	0.53 (0.41- 0.66)
AIDS deaths	Total	2.9 (2.5- 3.5)
	Adults	2.6 (2.2 -3.0)
	Children under 15 years	0.38 (0.29 – 0.50)

1.1.3 HIV/AIDS epidemic in Thailand

Thailand is one of the countries with a long standing AIDS epidemic. As reported by the Bureau of Epidemiology (December 31, 2006), the cumulative number of AIDS patients was approximately 307,114 cases and 85,459 AIDS death cases [3]. Within the Thai population, the percentage of AIDS cases were found to be 25.84 % in 30-34 age groups followed by 24.26 % in 25-29, 17.31 % in 35-39, 9.42 % in 40-44 , 8.59 % in 20-24, and less than 4.18 % in 0-14 age groups. For 15-19 age groups, the rate of AIDS cases in young women was higher than that of men. Most of AIDS cases were reported among labor groups, whose occupations include 46.43% laborers (general employees, industry employees, track driver and labors) and 20.81% agriculture. In 2007, the Thai Working Group or A2 Team estimated that the cumulative number of HIV cases included 1,102,628 people, while the cumulative AIDS deaths stood at 558,895 cases, and the remaining people living with HIV alive (PLWHA) who require medical attention amounted to 546,578 cases. Likewise in the global situation, it is obvious that the number of HIV-1 infected patients has continued to rise in Thailand as well.

1.2 Human Immunodeficiency Virus (HIV)

As aforementioned human immunodeficiency virus or HIV is the virus that causes AIDS. HIV cannot grow or reproduce on its own. In order to make new copies of itself, it must infect the cells of a living organism. HIV belongs to a special class of viruses called retroviruses. Within this class, HIV is placed in the subgroup of lentiviruses. Other lentiviruses include Simian immunodeficiency (SIV), Feline immunodeficiency virus (FIV), Visna and Caprine arthritis-encephalitis virus (CAEV), which cause diseases in monkeys, cats, sheep and goats, respectively [4].

There are 2 distinct major strains of the virus i.e. HIV-1 and HIV-2. HIV-2 has effects like those of HIV-1, but it is only found in the certain areas, normally in West Africa [5]. There are a few differences between the two strains. HIV-2 is harder to spread as it is only transmitted through sexual intercourse from woman to her child when it is in the womb whereas HIV-1 can also be transmitted through this and via blood to blood contact. Moreover, patients infected by HIV-2 have a lower virus density in their bloodstream. It is likely that HIV-1 is more common and is more often found so when people refer to HIV without specifying the type of virus they mean HIV-1.

1.2.1 HIV-1 Structure

HIV exists as roughly spherical particles (sometimes called virions). The surface of each particle is studded with a lot of little spikes. An HIV particle is around 100-120 nm in diameter [6]. Obviously, HIV particles are much too small to be seen through an ordinary microscope. However, they can be seen clearly with an electron microscope. HIV particles surround themselves with a coat of fatty material known as the viral envelope (or membrane). Projecting from this are around 72 little spikes, which are formed from the proteins gp120 and gp41. Below the viral envelope is a layer called the matrix, which is made from the protein p17. The viral core (or capsid) is usually bullet-shaped and is made from the protein p24. Inside the core are three enzymes required for HIV replication called reverse transcriptase (RT), integrase (IN) and protease (PR). Also held within the core is HIV's genetic material, which consists of two identical strands of RNA.

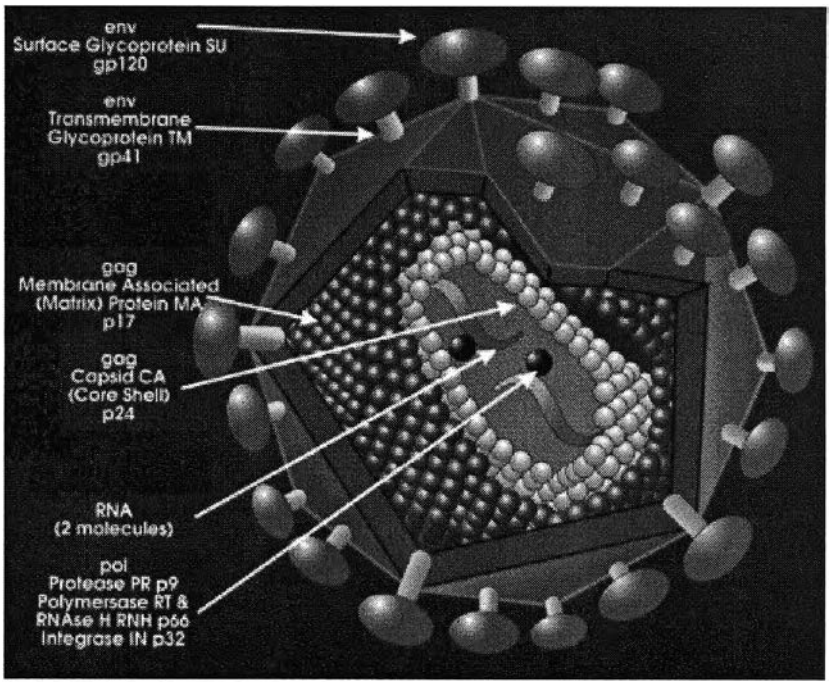


Figure 1.1 Structure of HIV-1 particle [7].

1.2.2 HIV-1 Replication

To produce new viral particles, HIV-1 replication (Figure 1.2) occurs by the following steps.

1. Binding and Fusion

HIV begins its life cycle when it binds to a CD4 receptor and one of two co-receptors on the surface of a $CD4^+$ T-lymphocyte. The virus then fuses with the host cell. After fusion, the virus releases RNA, its genetic material, into the host cell.

2. Reverse Transcription

An HIV enzyme called reverse transcriptase converts the single- stranded HIV RNA to double-stranded HIV DNA.

3. Integration

The newly formed HIV DNA enters the host cell's nucleus, where an HIV enzyme called IN "hides" the HIV DNA within the host cell's own DNA. The integrated HIV

DNA is called provirus. The provirus may remain inactive for several years, producing few or no new copies of HIV.

4. Transcription

When the host cell receives a signal to become active, the provirus uses a host enzyme called RNA polymerase to create copies of the HIV genomic material, as well as shorter strands of RNA called messenger RNA (mRNA). The mRNA is used as a blueprint to make long chains of HIV proteins.

5. Assembly

An HIV enzyme called protease cuts the long chains of HIV proteins into smaller individual proteins. As the smaller HIV proteins come together with copies of HIV's RNA genetic material, a new virus particle is assembled.

6. Budding

The newly assembled virus pushes out (buds) from the host cell. During budding, the new virus steals part of the cell's outer envelope. This envelope, which acts a a covering, is studded with protein/sugar combinations called HIV glycol-proteins. These HIV glycoproteins are necessary for the virus to bind CD4 and co-receptors. This new copies of HIV can now move on to infect other cells.

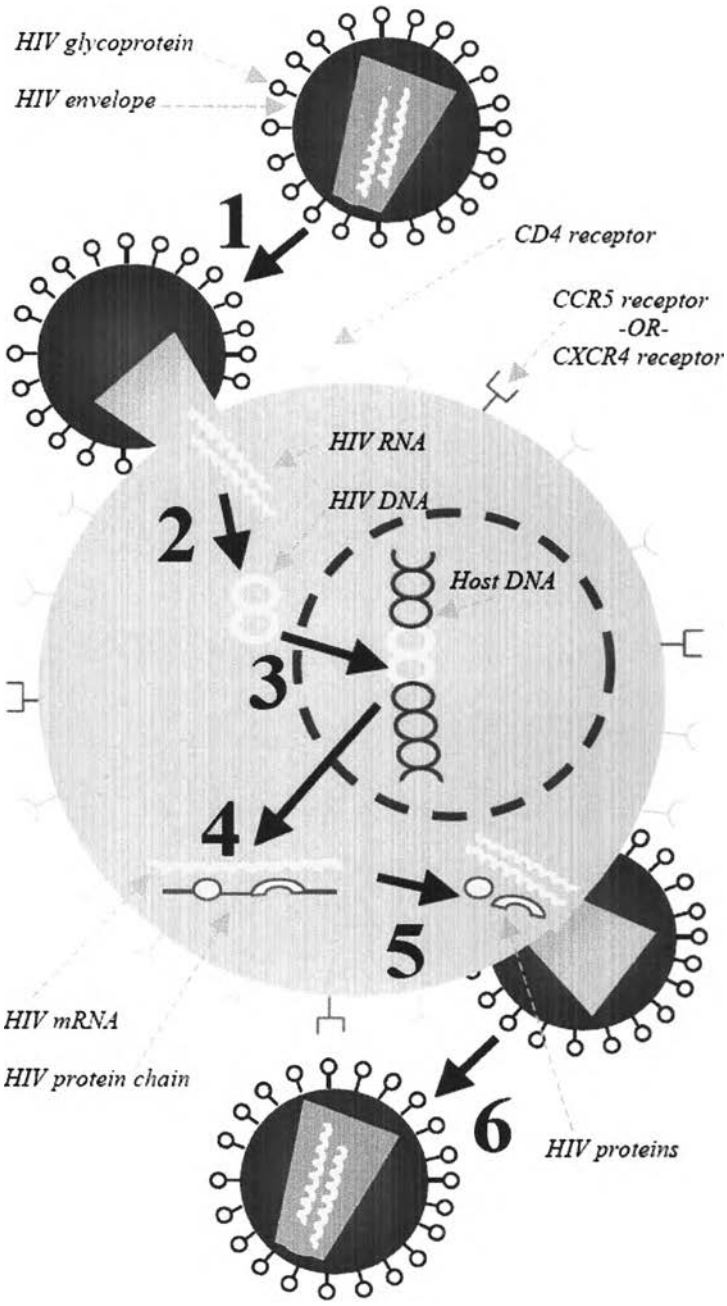


Figure 1.2 HIV-1 replication cycle [8].

1.2.3 HIV-1 Treatment

The goal of antiretroviral treatment is to decrease the morbidity and mortality that is generally associated with HIV-1 infection. Currently, 21 antiretroviral drugs [9] belonging to four classes are approved by the US Food and Drug Administration (FDA)

(Table 1.2). Enfuvirtide, a fusion inhibitor, targets a gp41 region of the viral envelope and stops the fusion process before the cell is infected. Its use is reserved for treatment of heavily drug-experienced patients since it can help overcome existing drug resistance [10, 11]. Eight nucleoside/nucleotide analogues and three non-nucleoside reverse transcriptase inhibitors inhibit viral replication after cell entry but before integration. Eight protease inhibitors prevent the maturation of viral resulting in production of non-infectious particles. The recently approved darunavir (June, 2006) is the first of its class that retains activity against viruses with reduced susceptibility to protease inhibitors.

Initially, treatment for infected individuals consisted of a single inhibitor class, mono-therapy. This had a modest impact on plasma HIV-1 RNA levels, but owing to the rapid emergence of drug-resistant virus strains and the inevitable development of serious immunodeficiency, mono-therapy has become obsolete [12, 13]. Consequently, combination therapy which consists of three or more drugs from two or more different inhibitor classes termed highly active antiretroviral treatment (HAART) was put into use. As a result, a drastic decrease in patient mortality has been observed [14]. Despite a dramatic increase in patient lifespan under HAART and successful suppression of HIV-1 RNA plasma levels for a prolonged period of time, the news has not been all good. Studies have shown that replication-competent viruses can be isolated from peripheral blood mononuclear cells and semen from infected patients who have maintained negligible plasma HIV RNA levels for up to two years. This means that the viral replication is incompletely suppressed using HAART [15, 16]. Other challenges facing the HAART include drug resistance, patient adherence, and toxicity [17]. Drug resistance is a consequence of incomplete suppression of viral replication. The high mutation rate of HIV causes mutations to accumulate, leading to diminished drug efficacy and gradually rendering each component of the regimen inactive [18]. Toxicity is another major concern in the administration of HAART and is yet one of the factors that influence adherence and thereby drug resistance. Given these serious limitations, the ongoing search for new, milder, and more varied clinical trials target almost all of the steps in the viral replication cycle including virus adsorption (gp120), viral co-receptors (CCR5 and CXCR4), viral fusion (gp41), nucleocapsid protein (NCP7), transcription factor, and IN.

Table 1.2 Antiretroviral drugs currently approved by FDA [9].

	Entry	Reverse transcriptase			Protease
Single compound tablets		Nucleoside	Non-nucleoside	Nucleotide	
	Enfuvirtide	Abacavir	Delaviridine	Tenofonir	Amprenavir
		Didanosine	Efavirenz		Atazanavir
		Emtricitabine	Nevirapine		Darunavir
		Lamivudine			Nelfinavir
		Stavudine			Ritronavir
		Zalcitabine			Saquinavir
		Zidovudine			Tripanavir
Fixed-dose combination tablets		Abacavir/lamivudine (Epzicom)			Lopinavir/ritonavir
		Zidovudine/lamivudine (Combivir)			
		Tenofovir/emtricitabine (Truvada)			
		Abacavir/lamivudine/zidovudine (Trizavir)			
		Tenofovir/emtricitabine/efavirenz (Atripla)			

1.3 HIV-1 Integrase (IN)

The *pol* gene of HIV-1 encodes three essential enzymes, RT, PR, and IN (Figure 1.3). These three enzymes are essential for viral replication (see 1.2.2 HIV-1 Replication) and can be exploited for the development of antiviral chemotherapeutic agents [19]. The discovery and development of clinically useful inhibitors of RT and PR for the treatment of AIDS have suggested that the therapeutic methodology aimed at targeting key viral enzyme is viable approach in antiviral chemotherapy. While RT and PR enzymes have received extensive attention, IN has received much less consideration. The function of HIV-1 IN is to catalyze the insertion of newly reverse transcribed double strand viral DNA into host DNA where it can be replicated using host machinery. In this thesis, special attention is given to HIV-1 IN.

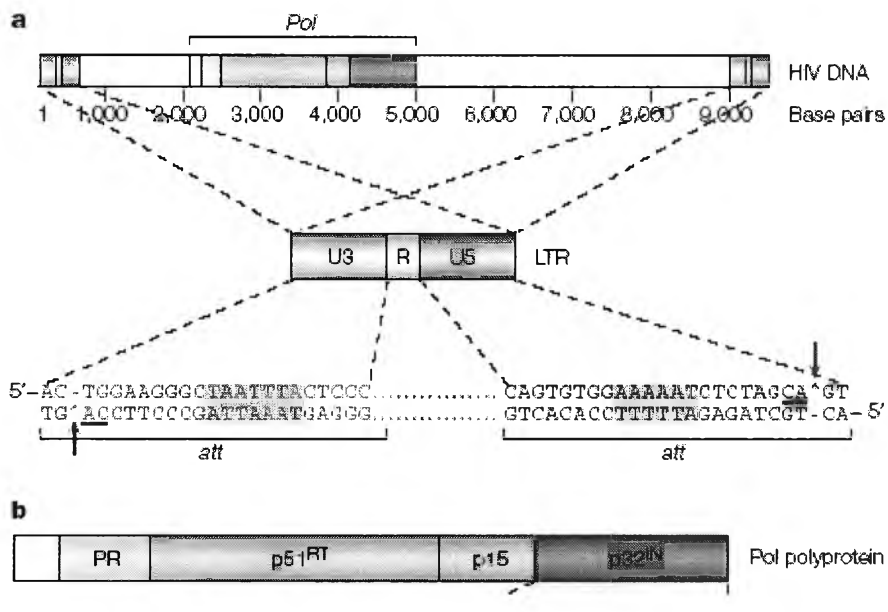


Figure 1.3 HIV provirus gene structure and IN domains [20].

1.3.1 HIV-1 IN structure

HIV-1 IN contains 288 residues and can be divided into three functional domains, N-terminus, catalytic core, and C-terminus, Figure 1.4. Structures of individual domain and the two fragments domain (N plus core and core plus C) have been solved by nuclear

magnetic resonance (NMR) spectroscopy and/or X-ray crystallography [21-26]. However, the complete structure containing all three domains is experimentally unavailable because of low solubility and the tendency of the enzyme to aggregate. The active form of HIV-1 IN is still unclear. The purified IN from a variety of retroviruses can be presented in dimeric, tetrameric, or oligomeric form in solution. Despite the confusion over the degree of multimerization, more recent studies indicate that a tetramer (a dimer of dimers) is probably required for the full integration reaction, with two of the active sites bound to DNA and the other two serving as structural support [27].

1.3.1.1 N-terminal domain

The N-terminal domain which comprises residues 1-50 is characterized by a conserved HHCC zinc binding motif [26]. The solution structure of this domain has been solved by NMR [28], and its function is to interact with the specific nucleotide sequences at the termini of the viral DNA [29].

1.3.1.2 Catalytic core domain

Catalytic core domain encompasses residues 50–212 and contains a highly conserved DDE (D64, D116 and E152) motif which is also found in other retroviral IN and bacterial transposases. Mutation of any of these three acidic residues abolishes IN's enzymatic activities and viral replication. The D64 and D116 residues form a coordination complex with a divalent metal (Mg^{2+} or Mn^{2+}). Because a second metal has been observed in an ASV IN crystal structure [30, 31], and because of the two-metal structure for polynucleotide transferases [32, 33], it has been proposed that a second metal (Mg^{2+} or Mn^{2+}) should be located between D116 and E152 once HIV-1 IN binds its DNA substrate(s) [34, 35]. The monomeric structure of catalytic core domain contains 5 β-sheets flanked by α-helices. In most structures, the catalytic core domain contains a short disordered loop (encompassing residues 140–149), the structure of which can be stabilized by DNA.

1.3.1.3 C-terminal domain

The C-terminal domain including residues 213-288 exhibits some similarity to Src homology domains (SH3). This domain is involved in nonspecific DNA recognition and may also function as a multimerization domain [36].

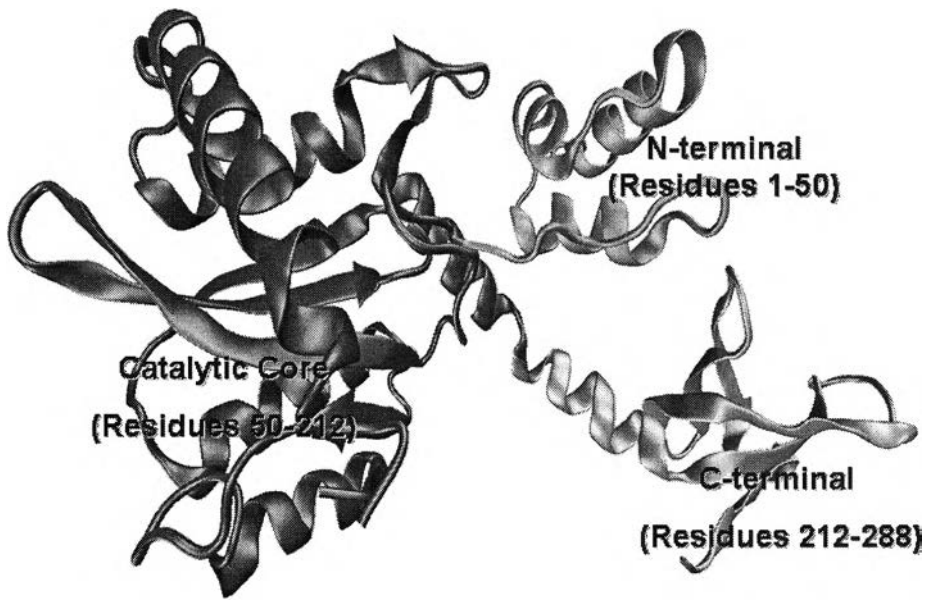


Figure 1.4 HIV-1 IN structure [37].

Although the catalytic core domain contains the enzyme catalytic site(s), in the absence of N and the C terminal domains, it can only catalyze the disintegration reaction, the reverse of the strand-transfer (ST) reaction *in vitro* [38]. To catalyze both 3′-processing and ST, the full-length protein is required.

1.3.1.4 The Catalytic core domain complexed with inhibitor

In addition to the apo protein, the structure of the catalytic core domain of HIV-1 IN bound to an inhibitor 1-(5-chloroindole-3-yl)-3-hydroxy-3(2H-tetrazol-5-yl)-propenone (5CITEP) was solved by the X-ray crystallography (pdb code: 1QS4, Figure 1.5) [39]. In the crystal structure, 5CITEP is located in the center of active site and laid

between Asp64, Asp116 and Glu152. The keto-enol moiety of 5CITEP forms hydrogen bond with Glu152 while all four nitrogen atoms of tetrazole ring form hydrogen bonds with Asn155, Thr66, Lys159 and Lys156. Moreover, the nitrogen atom of indole ring is hydrogen bonded to Gln148. The solution of the co-crystal structure has been used for the study of HIV-1 IN-ligand interactions.

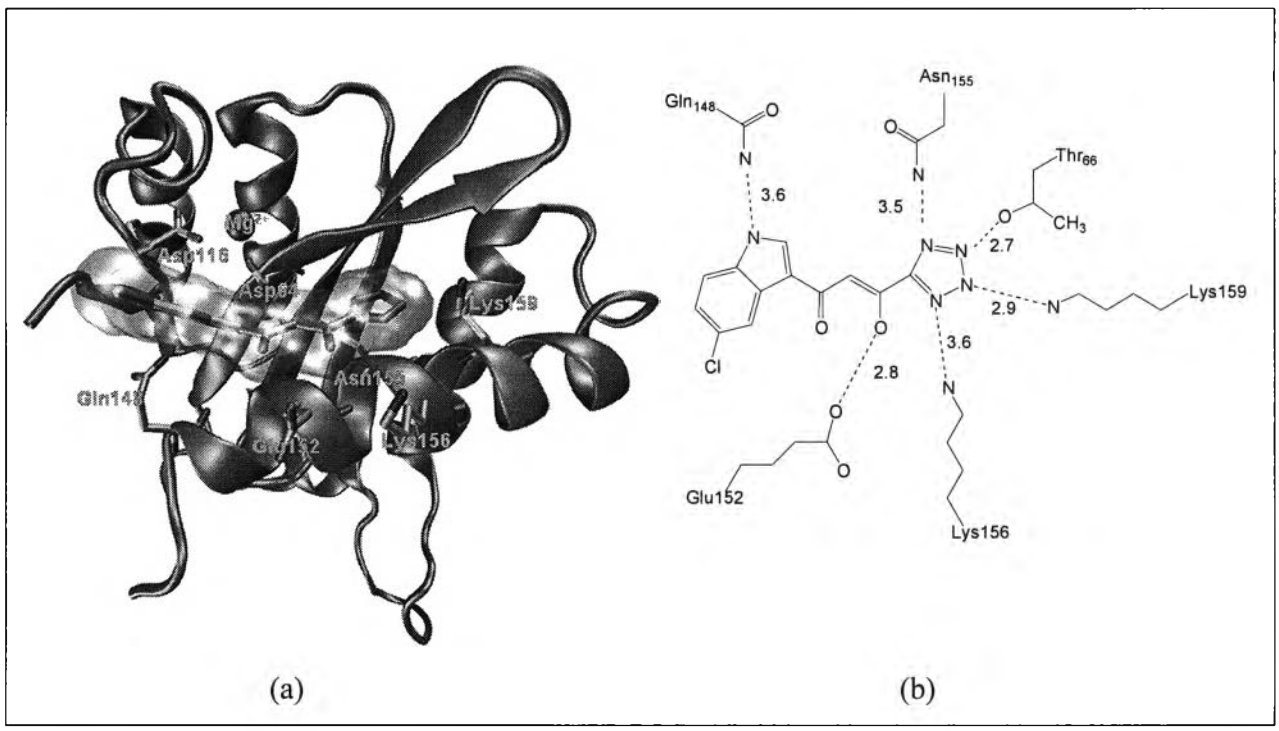


Figure 1.5 (a) X-ray co-crystal of HIV-1 IN-5CITEP and (b) schematic drawing of hydrogen bonds (estimated by heavy atom distances) between 5CITEP and surrounding amino acids [39].

1.3.2 HIV-1 IN mechanism and the possible drug targets

Basically, the process of retroviral integration has been described as two well-characterized catalytic steps referred to 3'-processing (the removal of two terminal nucleotides from the 3'-ends of the viral double-stranded DNA) and ST (transfer of viral DNA to the human chromosomal DNA) [40]. However based on the possible drug targets, it may be better described as four unique processes (see Figure 1.6) that can be targeted and potentially inhibited. These four processes include 1) assembly of a stable

complex between IN and specific viral DNA sequences at the end of the HIV-1 long terminal repeat regions (LTRs), 2) 3'-processing, 3) strand transfer, and 4) DNA gap repair and ligation [41].

1.3.2.1 Binding of IN to viral DNA

After the viral RT enzyme creates a double-stranded DNA product, IN then assembles at the ends of the viral DNA, binding to the HIV-1 LTR region. The viral LTR ends contain specific DNA sequences that are required for recognition by IN resulting in a stable viral DNA-IN binding complex [20]. This is a crucial initial step in the integration process (Figure 1.6). One class of IN inhibitors in preclinical development is the pyrano-dipyrimidines. These compounds have been shown *in vitro* to prevent viral DNA binding to IN, thus preventing proper assembly of viral DNA (step 1) and consequently any further steps of integration (see Figure 1.6).

1.3.2.2 First catalytic step: 3'- processing

During the 3'-processing reaction, IN removes two nucleotides from each viral cDNA end adjacent to a conserved 3'-CA sequence. This leads to the formation of a new recessed 3'-CA-OH end. This reaction occurs in the cytoplasm, within a large viral nucleoprotein complex, the preintegration complex (PIC). The PIC contains linear viral DNA and several viral proteins including matrix, RT, IN, and nucleocapsid (see Figure 1.6) [40]. Styrylquinolines, one class of IN inhibitors in preclinical development, have been described as potent *in vitro* 3'-processing inhibitors by directly competing for HIV LTR substrates [42].

1.3.2.3 Second catalytic step: strand transfer

After the PIC is transported from cytoplasm through the nuclear pore into the cell's nucleus, the two newly processed 3'-viral DNA ends are inserted into opposite strands across a five base pair stretch of host target DNA. This joining reaction includes a coupled five base pair staggered cleavage of the target DNA and the ligation of the processed CA-3'-OH viral DNA ends to the 5'-*O*-phosphate ends of the target DNA. β-diketo acids (DKAs) and their DKA derivatives are examples of ST inhibitors that have

been shown *in vitro* and *in vivo* to effectively stop this process [34, 43]. Oxadiazoles are another class of IN inhibitors in early development that are described as nuclear import inhibitors [44]. These compounds have been shown *in vitro* to indirectly prevent the second catalytic step of integration by targeting and preventing the nuclear translocation of the HIV-1 PIC.

1.3.2.4 Gap repair and ligation of viral DNA to host DNA

The product of ST step is a gapped intermediate product in which the 5'-phosphate ends of the viral DNA are not attached to the 3'-OH ends of the host DNA. The integration reaction is completed by the removal of the two unpaired nucleotides at the 5'end of the viral DNA and the repair of the single-stranded gaps created between the viral and target DNA [40, 45]. Although IN may be involved in these repair reactions, it is not necessary because the host cell already has the machinery to carry out such processes so repair is probably accomplished by host cell DNA repair enzymes [38, 45]. Staggered strand transfer and gap repair result in the duplication of host cell sequences immediately flanking the inserted pro-viral DNA [40]. Host cellular DNA repair proteins have also been mentioned as possible attractive targets for anti-HIV-1 therapeutics. For the integration reaction, no source of energy (e.g., no adenosine tri-phosphate [ATP]) is needed and only divalent cations such as Mn^{2+} or Mg^{2+} are required for the catalytic activity [34, 35].

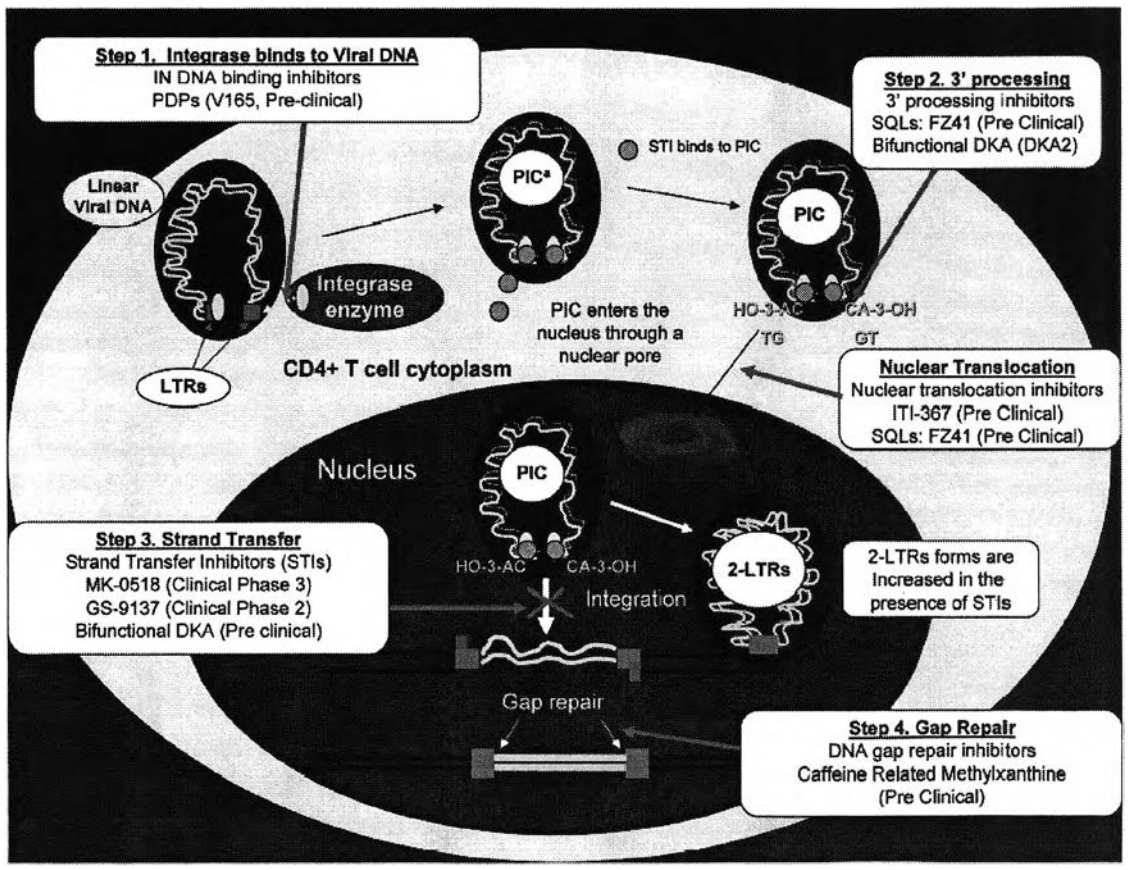


Figure 1.6 HIV-1 integration process and potential drug targets for inhibition.

1.3.3 Progress of HIV-1 IN inhibitors design

The first IN inhibitors were reported approximately 10 years ago [46, 47]. The developing of potent and effective inhibitors of HIV-1 IN has been hampered by the complex nature of integration process and often confounded by non concordance between in vitro biochemical and antiretroviral activity. Moreover, the wide shallow active site of IN and the unclear three dimensional organization of its domain in vivo present redoubtable challenges to the design of drugs against this enzyme. However, after a decade of research, much progress has been made and lead candidates are beginning to enter preclinical and clinical trials with promising outcomes (see Table 1.3). These developments suggest that IN inhibitors will be important part of future combination therapies.

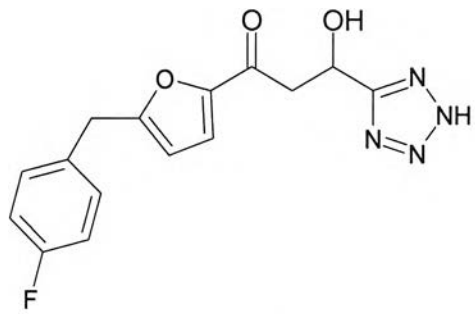
Numerous compounds have been designed, synthesized and evaluated for anti IN activity. Among all of the reported HIV-1 IN inhibitors, the most promising candidates are the DKA derivatives. These can be supported by the existing of several DKA compounds in clinical trials (Table 1.3, Figure 1.7). DKAs selectively inhibit the ST reaction of IN and exhibit antiviral activity against HIV-1 infected cells in a manner consistent with inhibition of integration. The mechanism of action of the DKA compound has been proposed to interact with the divalent catalytic metal ion which is required for the binding of DKA to HIV-1 IN.

Table 1.3 HIV-1 IN inhibitors: Phase of preclinical and clinical development.

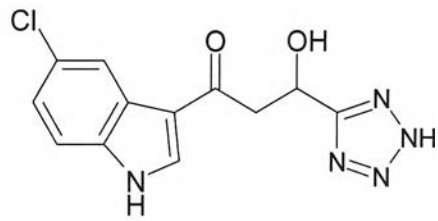
Phase of development	IN inhibitors	Compound specifics and mechanism of action
Preclinical	L-second generation compounds, Merck [48]	Diketo acid (DKA) derivative: Naphthyridine Carbox amide. Potent ST inhibitor, good safety profile
	Carbazole derivative NIID, Japan 7 compounds (CA-1 to CA-8) [49]	Carbazole Derivatives ST inhibitor, cytotoxic
	L870,810, derivative of S-1360 Glaxo SmithKline (GSK) [50]	DKA Potent ST inhibitor, (S-1360 2 trial stopped in 2005)
	Styrylquinoline derivatives, e.g., FZ41, Bioalliance Pharma [42, 51, 52]	Styrylquinolines. 3'- processing inhibitor selective resistance profile
	Pyrano-dypiriminidines, e.g., V-165 Rega Institute, Belgium [53].	Pyrano-dipyrimidines IN-DKA binding inhibitor. V-165 is the most potent of group. Prevents IN-Viral cDNA assembly, so 3'-processing never takes place. Selective resistance profile

Preclinical	Mycelium integrasone, Fungal polyketide, Merck [54, 55] Indolicidin, NIH [56]	Fungal Polyketide-Integrasone: ST inhibitors Antimicrobial Peptide: Indolicidine ST inhibitor and 3'-processing inhibitor. Natural peptide Dimeric & tetrameric analogues are more potent.
	Theophylline, caffeine-related methylxanthine, Thomas Jefferson University, Philadelphia [57]	Caffeine-related methylxanthines Host cell DNA repair protein inhibitors: inhibiting the last step of integration performed by DNA repair enzymes Post-strand transfer inhibitor
	ITI-367, Oxadiazol GWU [44]	Nuclear import inhibitor (new class) HIV-1 PIC nuclear import inhibitor 1 indirectly prevents ST since PIC translocation never takes place
Phase 2	GS-9137 (formerly JTK-303), Gilead [58]	Dihydroquinoline carboxylic acid Potent ST inhibitor Promising Phase 1 safety, tolerability and clinical data Phase 2 trials with GS-9137 now enrolling
Phase 3	L-900, 612, Now called MK-0518 Derivative of L870, 810, Merck [43]	Naphthyridine Carboxamide Derivative MK-0518 is featured compound. Potent ST inhibitor Promising phase 1 & 2 safety, tolerability and clinical data Phase 3 trials with MK-0518 now enrolling

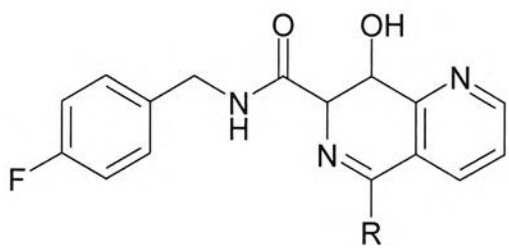
Figure 1.7 Chemical structures of HIV-1 IN inhibitors.



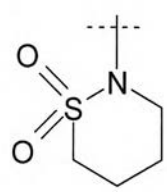
S-1360



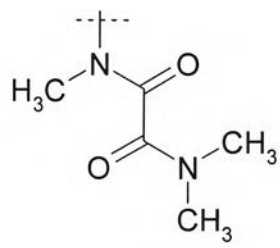
5CITEP



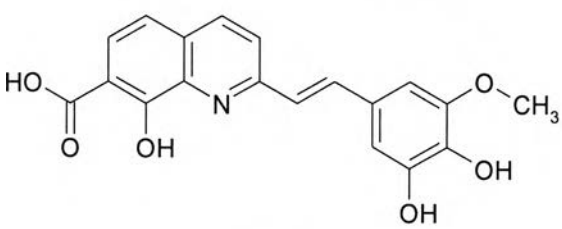
Naphthyridine Carboxamides



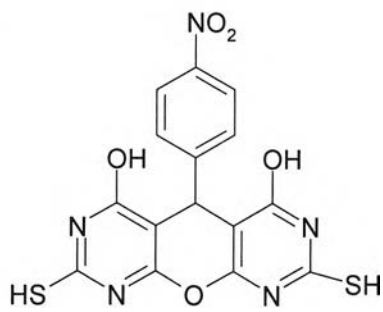
L870,810



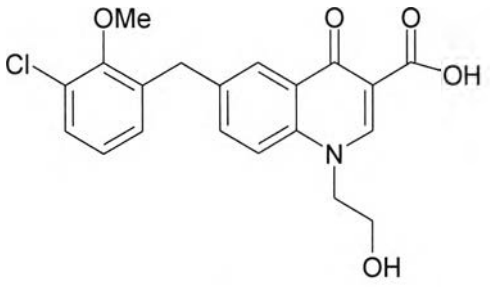
L870,812



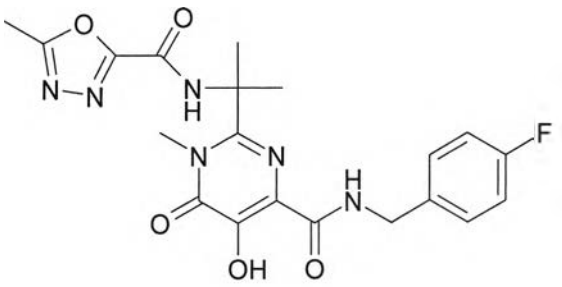
FZ41



V165



JTK303



MK-0518

1.3.4 Outlook and Future Challenges of HIV-1 IN

To date, the best IN inhibitor compounds suited for clinical evaluation and development are the DKAs and their derivatives such as the naphthyridine compounds (e.g., MK-0518) and the dihydroquinoline carboxylic acid compounds (GS-9137). The two lead ST inhibitors, MK-0518 and GS-9137, are currently in clinical trials in human volunteers. The pyrano-dipyrimidines (IN binding inhibitor) are the second group of IN inhibitors suited for clinical evaluation and development. Other promising preclinical classes of compounds are the styrylquinolines that are 3'-processing inhibitors whereas IN nuclear import inhibitors and the bifunctional DKAs that are described as both 3'-processing inhibitors and ST inhibitors. The specificity and precise mechanism of action of IN during HIV replication could make IN inhibitors a very unique and potent new class of antiretroviral agents.

1.4 Research Inspiration

Three essential enzymes, RT, PR and IN, are required for HIV replication. The development of effective drugs targeting RT and PR has demonstrated potential effectiveness of antiviral therapy for the treatment of AIDS. Currently, combination therapy or highly active antiretroviral treatment (HAART), comprising at least two anti HIV drugs, has become the standard treatment of AIDS or HIV infected patients. Although HAART drastically decreases viral spread and provides a significant improvement in the life expectancy of HIV/AIDS patients, the rapid emergence and spread of drug resistant HIV-1 strains and serious long term toxic effects of available anti retroviral drugs has generated demand for the design of new drugs with novel

mechanisms of action, targeting critical steps in the retroviral replication processes. IN which plays a crucial role in the viral replication by catalyzing the incorporation of the reverse transcribed viral DNA into host chromosome has become an attractive target for interrupting progression of HIV replication. Because there are no known similar enzymes in human hence specific inhibitors of HIV-1 IN are expected to have minimal side-effects as compared to other antiviral agents. IN inhibitors thereby would become a potential additive to HAART or a salvage therapy for patients resistant to currently available anti-HIV drugs.

In spite of its obvious importance, IN is a difficult molecular target for the application of structure based drug discovery. This is due to its shallow substrate binding site located on the surface of the protein, its role in the formation of a multimeric complex in pre-integration complexes, and the lack of a full-length structure for enzyme in absence or presence of DNA substrate. Moreover, unlike RT and PR, limited structural information on IN interaction with inhibitors has been published. Only one crystal structure of the IN core domain complexed with inhibitor is available. Although, over the last few years several classes of compounds have been reported to inhibit IN, many of them were either cytotoxic or inactive in blocking viral replication. Moreover, none of HIV-1 IN drug is available clinically yet. Therefore, the search for selective, potent and inexpensive IN inhibitors is substantially required. In order to accomplish this goal, the understanding of structural and dynamical behaviors of protein in solution including protein-ligand interaction is necessary. As computer aided drug design can provide such information, several computational techniques were employed in this thesis with the hope that the obtained information could facilitate the development of highly potent and selective HIV-1 IN inhibitors.

1.5 Research Goals

In this thesis, two major computational approaches, three-dimensional Quantitative Structure Activity Relationship (3D-QSAR) and hybrid Quantum Mechanical/Molecular Mechanical (QM/MM) Molecular Dynamics (MD) simulations were applied to accomplish the following aims:

1. 3D-QSAR techniques were used to construct single 3D-QSAR models describing the relationships between biological activities and physicochemical properties of diverse classes of HIV-1 IN inhibitors. The obtained information would be very helpful for better understanding of structural requirements of HIV-1 IN inhibitors. Although 3D-QSAR studies of HIV-1 IN inhibitors were reported [59-64], many of them investigated for only one or a few classes of compounds. For 3D-QSAR studies including several classes of HIV-1 IN inhibitors, an acceptable single 3D-QSAR model cannot be derived and all compounds must be categorized into at least two groups to obtain an acceptable single model. This is inconvenience for predicting the activities of many unknown compounds. Therefore, a single 3D-QSAR model which can predict the activities of many test compounds is generated in this work.

2. QM/MM MD simulations were carried out to investigate structural, dynamical and thermodynamical behaviors of HIV–1 IN and its complex with DKAs, the highly active class of inhibitor. Although, classical MD simulations of both free and complexes of HIV-1 IN with inhibitors were reported [65-69], no QM/MM MD simulation of HIV-1 IN is reported. In theoretical, the higher accuracy method, QM/MM is expected to better describe the structural properties of protein-ligand complexes.