UNIVERSITÀ DEGLI STUDI DI SALERNO

UNIVERSITÀ DEGLI STUDI DI SALERNO
Dipartimento di Farmacia

Dottorato di Ricerca
in **Scienze del Farmaco**
Ciclo XXIX — Anno accademico 2016/2017

Tesi di Dottorato in

Design and synthesis of new integrase inhibitors

Dottorando **Tutore**

Dott. *Antonia Spensiero* **Chiar.mo Prof.** *Pietro Campiglia*

Coordinatore: Chiar.mo Prof. *Gianluca Sbardella*

Index

Figure captions

Chapter I:

Chapter II:

Chapter IV:

Table captions

Chapter III:

Chapter IV:

Abstract

The viral enzyme integrase (IN) is essential for the replication of human immunodeficiency virus type 1 (HIV-1) and represents an important target for the development of new antiretroviral drugs. In this PhD project, we focused on the N-terminal domain of integrase (NTD) for the development and synthesis of a library of overlapping peptide sequences, with specific length and specific offset covering the entire native protein sequence NTD IN 1-50. The most potent fragment, VVAKEIVAH (peptide **18**), inhibits the HIV-1 IN activity with an IC_{50} value of 4.5 μM. Amino acid substitution analysis on this peptide revealed essential residues for activity and allowed us to identify two nonapeptides (peptides **24** and **25**), that show a potency of inhibition similar to peptide **18**. Interestingly, peptide **18** does not interfere with the dynamic interplay between IN subunits, while peptides **24** and **25** modulate these interactions in different manners. In fact, peptide **24** inhibits the IN-IN dimerization, while peptide **25** promotes IN multimerization, with IC_{50} values of 32 and 4.8 μM, respectively. In addition, peptide **25** has shown to have selective anti-infective cell activity for HIV-1. Moreover, the NMR analysis showed an alpha helix conformation of peptide **25**, which could be essential for the interaction with IN. These results indicated peptide **25** as a hit for further development of new chemotherapeutic agents against HIV-1. In addition, we observed that the peptide **5**, EKYHSNWRAM, conveniently conjugated with the cell-penetrating fragment TAT, inhibits replication of HIV-1 and HIV-2 in infected MT-4 cells.

Keywords: HIV-1, integrase, N-terminal domain, peptides, inhibitors.

Abbreviations

IN, integrase; HIV-1, human immunodeficiency virus type 1; PIC, preintegration complex; NTD, N-terminal domain; CCD, catalytic core domain; CTD, C-terminal domain; INSTIs, integrase strand transfer inhibitors (INSTIs); HOAt, 1-Hydroxy-7-azabenzotriazole; HOBt, hydroxybenzotriazole; HBTU, N,N,N′,N′-tetramethyl-O-(1H-benzotriazol-1-yl)uraniumhexafluorophosphate; DIEA N,N-diisopropylethylamine; DMF, dimethylformamide.

CHAPTER I:

The Continuing Evolution of AIDS: HIV-1 integrase as a target for antiretroviral therapy

1. Acquired immune deficiency syndrome (AIDS)

1.1 Global distribution of HIV-1 infection

Acquired immune deficiency syndrome (AIDS), is induced by HIV infection and consists in a profoundly damage of immune system.

Following the discovery of the HIV as the causative agent of AIDS, many advances were made in a short period of time, including the understanding of the modes of transmission; the sequencing of the HIV-1 genome;[1] elucidation of the main cell targets, CD4 T cells and macrophages;[2] the genomic heterogeneity of HIV with the innumerable micro variants within a single patient;[3] identification of a SIV-monkey model to study and monitoring the epidemic events.[4] Human immunodeficiency virus (HIV), is a retrovirus known for immune system degradation, transmitted by direct contact with infected blood, sexual contact or by direct transmission of mother to child. The HIV virus has remained a predominant health care concern for the past 30 years.[5] Around 35 million people have died from AIDS-related illnesses since the first cases of HIV were reported, and, overall, 78 million have been infected with HIV, according to UNAIDS, the Joint United Nations Programme on HIV/AIDS. Thanks to a dramatic scale-up of antiretroviral treatment, AIDS-related deaths have fallen by 45% to1.1 million in 2015 from a peak of 2 million in 2005. Globally, 36.7 million people were living with HIV in 2015, the majority in poor and middle-income countries. As of June 2016, around 18.2 million were receiving antiretroviral therapy, up from 15.8 million at the same time last year, and 7.5 million in 2010 (*figure 1.1*).

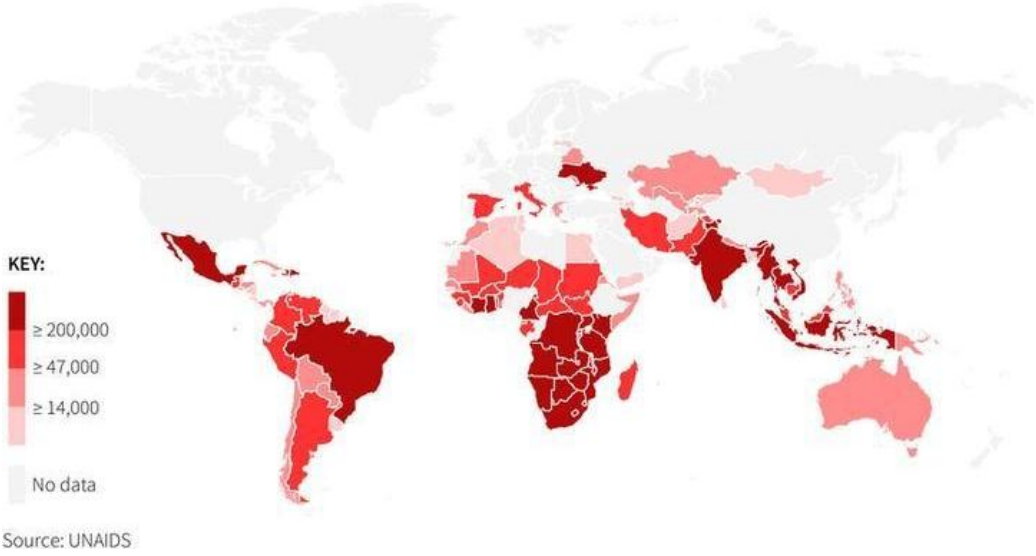


Figure 1.1 *Population of people living with HIV in 2015. Adapted from: http://chartsbin.com/view/11852.*

Last year, around 2.1 million people became newly infected with HIV. While the number of adults acquiring the virus is not falling, huge advances have been made in tackling mother to child transmissions. In the past five years, new infections among children halved 150.000 in 2015, down from 290.000 in 2010. UNAIDs warns that young women are particularly at risk of infection. In sub Saharan Africa, adolescent girls face a triple threat: a high risk of HIV infection, low rates of HIV testing and difficulty sticking with HIV treatment. The first approved therapy involved the use of multiple drugs (HAART, Highly Active Anti-Retroviral Therapy),[6] which transform the HIV infection from lethal to chronic. Despite this, due to drug resistance that are highlighted below, the treatment regimen has not allowed the complete eradication of the disease itself. The treatment only reduced the number of new infections and subsequent deaths, but despite these promising advances, HIV infection remains a looming problem on global health and an important goal for scientific research.

1.2 HIV Genome

The HIV virus belongs to the retroviruses family and is a member of the genus of lentiviruses, which indicates a long incubation period. The essential genes for replication and then for virus activity are: gag, pol, and env. The gag gene encodes the nucleocapsid protein; the pol gene encodes the enzyme reverse transcriptase, protease and integrase; Finally, env encodes for the proteins of the pericapside.[7,8] The HIV genome also contains, at the sides of the env gene, six more genes that code for regulatory proteins and accessory: tat (transactivation of Transcription), rev (Regulatory of Virus), vpu (Viral Protein U), nef (Negative Factor) , vif (viral infectivity factor) and vpr (Viral Protein R), involved in different steps in regulation and HIV pathogenesis. A conceptual representation of the virus architecture is represented in figure 1.2.

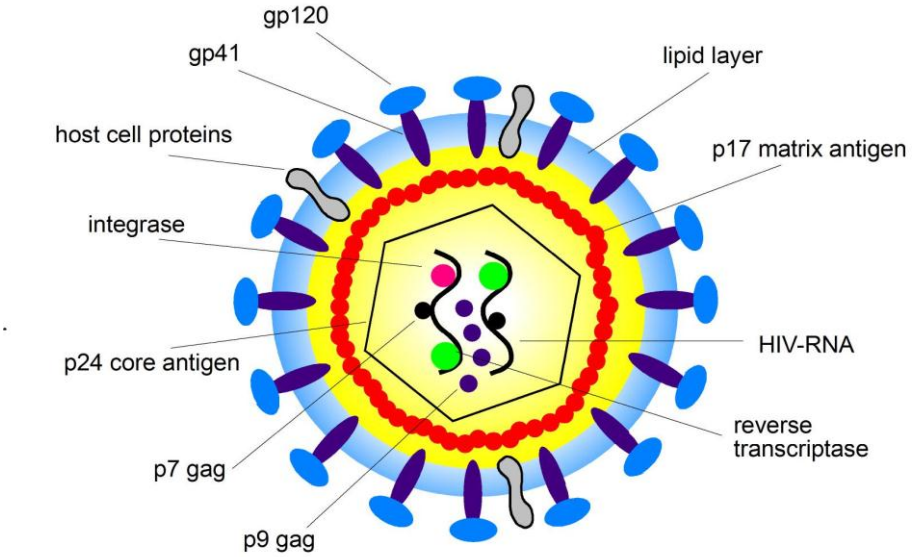


Figure 1.2 *HIV virus structure.*
Adapted from: http://healthfavo.com/hiv-virus-structure-anatomy-picture-reference.html.

1.2.1 Replicative cycle of HIV-1

Replicative cycle of the virus can be divided into different stages: penetration of the virus within the host cell, reverse transcription of the viral genome and introduction into the host cell genome by the integrase enzyme, the production of second messengers and, finally, assembly of new infectious virions during the late phase. The infection begins when the envelope (Env) glycoprotein gp120 binds the receptor CD4 and the co-receptor CC-chemokine receptor 5 (CCR5) (step 1), leading to fusion of the viral and cellular membranes and entry of the viral particle into the cell (step 2). The role of the CD4 receptor in HIV cell entry was identified shortly after the isolation of HIV.[9] The conformational changes in the gp120 protein leads to exposing an epitope which allows binding to a co-receptor, such as the chemokine receptors CCR5 or CXCR4, which are the most co-receptors used by HIV in vivo. The importance of this co-receptor binding was empathized by the compromised HIV-1 infectivity in the individuals harbouring mutant CCR5 proteins.[10] Partial core shell uncoating (step 3) facilitates reverse transcription (step 4). This step is catalyzed by the viral enzyme reverse transcriptase (RT), which transcribes the viral single-stranded RNA genome which is converted into double-stranded DNA (dsDNA). Together with viral and host proteins the dsDNA forms the pre-integration complex (PIC), which is guided to the nuclear pore. This is an essential step in the HIV replication cycle of HIV because it prepares the viral genome for the subsequent integration into the host chromosome. Following of displacement into the cell nucleus (step 5), PIC-associated viral integrase forms the integrated provirus (step 6).

Proviral transcription (step 7), mediated by host RNA polymerase II (RNA Pol II) and positive transcription elongation factor b (P-TEFb),

yields viral mRNAs of different sizes, the larger of which requires energy-dependent export to leave the nucleus via host protein CRM1 (step 8). mRNAs is used as templates for protein production (step 9) and genome-length RNA is incorporated into viral particles with protein components (step 10). Viral-particle budding (step 11) and release (step 12) from the cell is mediated by ESCRT (endosome sorting complex required for transport) complexes. Each step in the HIV-1 life cycle is a potential target for antiviral intervention;[6] The sites of action of clinical inhibitors and cellular restriction factors are indicated in the figure 1.3.

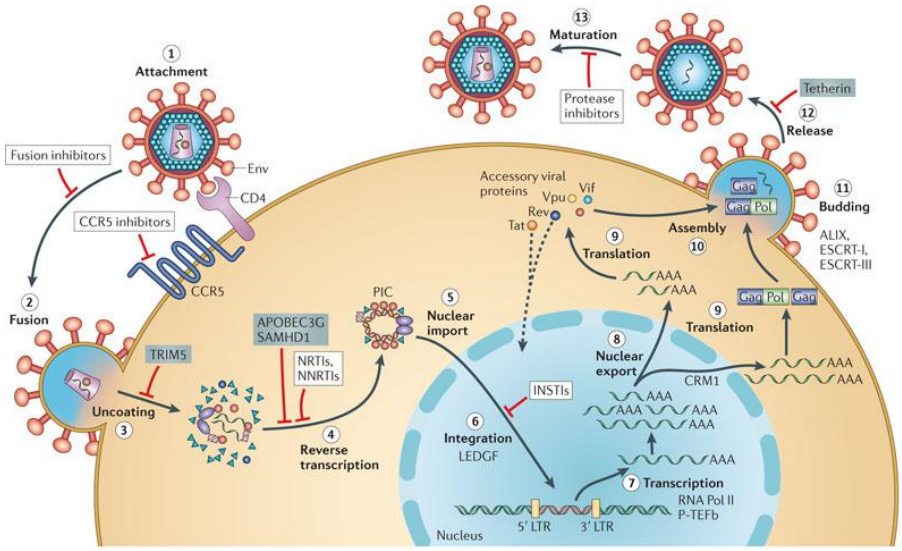


Figure 1.3 *Schematic overview of the HIV-1 replication cycle.*
Adapted from: Engelman A and Cherepanov P. The structural biology of HIV-1: mechanistic and therapeutic insights. Nature Reviews Microbiology. ***2012****; 10, 279-290.*

1.3 Antiretroviral Drugs

The significant advancement in the understanding of HIV replication and its pathogenesis has helped in the identification of various pharmacological targets. The first anti-HIV agent to be licensed for clinical use was Zidovudine (AZT), in 1987. Since then, several anti-HIV compounds have been approved for the treatment of AIDS by the FDA and EMEA (European Medicine Agency) for treating HIV infections.[11] These compounds fall within different categories depending on the target within the HIV replicative cycle they interact with.

1.3.1 The Entry and Fusion Inhibitors

These inhibitors intercept the viral replication at the entry of the viral core into the cytosol of the host cell. The group of entry inhibitors can be subdivided into classes of agents that act at different stages of entry: attachment and CD4 binding, co-receptor binding, and fusion. In the recent years, important progress has been made in understanding the HIV-1 entry process in which the viral and cellular membranes are fused, resulting in the subsequent delivery of the viral genome into the host cell. These studies have led to the formulation of new approaches for therapeutic intervention. One of the first and clinically most advanced drug emerged from this effort is the fusion inhibitor T20. T20 acts by freezing a transient structural intermediate of the HIV-1 fusion process, thus blocking an essential step in viral entry. With phase III clinical trials already well underway, the success of T20 indicates that targeting the viral entry process will soon be an important component of antiretroviral therapy. Drug resistance mutations are usually located in

the ENF binding site on gp41 (direct resistance) or confer resistance indirectly via mutations in other regions of gp41 and even in gp120.[12] Currently, only antagonists that block CCR5 binding (Maraviroc) and fusion (Enfuvirtide) have been approved by the FDA for treatment of HIV infected patients, although strategies to inhibit other aspects of HIV entry are under development.[13] In addition to Maraviroc, other CCR5 and also CXCR4 inhibitors are being investigated.[14]

1.3.2 The Reverse Transcriptase (RT) Inhibitors

The Reverse Transcriptase (RT) inhibitors interfere with the generation of a DNA copy of the viral genome. RT functions as a heterodimer and catalyzes the conversion of the single-stranded genomic RNA into double-stranded DNA (with duplicated long terminal repeats), which is integrated into cellular DNA by the viral integrase.[15] There are two classes of RT inhibitors, that exhibit different mechanism of action. First of them is the group of nucleoside/nucleotide reverse transcriptase inhibitors (NRTIs). These are nucleoside and nucleotide analogues that are incorporated by the viral RT into the newly synthesized DNA strand. They are inactive in their parent forms and require successive phosphorylation steps by host cell kinases and phosphotransferases to form deoxynucleoside triphosphate (dNTP) analogs responsible of viral inhibition. In their respective triphosphate (TP) forms, NRTIs compete with their corresponding endogenous dNTPs for incorporation by HIV RT. Once incorporated, they assist as terminators chain of viral RT, thus, acting early in the viral replication cycle by inhibiting a critical step of proviral DNA synthesis prior to integration into the host cell genome.[16]

The currently approved NRTIs are Zidovudine, Lamivudine, Didanosine, Zalcitabine, Stavudine, Abacavir, and Emtricitabine.

The second group of RT inhibitors is the Non-nucleoside reverse transcriptase inhibitors (NNRTIs). They are small molecules that carry out the inhibition of RT by binding to a hydrophobic pocket in the proximity of the active site of the enzyme. After the inhibitor is bound, it impairs the flexibility of the RT resulting in its inability to synthesize DNA. Mutations may induce resistance to NNRTIs resulting in reduced affinity of the inhibitor with the protein. Usually, a single mutation selected by one NNRTI is sufficient to confer complete resistance to all compounds of the drug class.[17]

1.3.3 The Protease Inhibitors (PIs)

Protease Inhibitors (PIs) interfere with the process of forming new infectious viral particles. The viral protease is involved in virion maturation. Protease targets the amino acid sequences in the gag and gag–pol polyproteins, which must be cleaved before nascent viral particles (virions) can mature. Cleavage of the gag polyprotein produces three large proteins (p24, p17, and p7) that contribute to the structure of the virion and to RNA packaging, and three small proteins (p6, p2, and p1) of uncertain function. PIs are small molecules that bind to the active site of the protease and compete with its natural substrates. PIs contain a synthetic analogue of the amino acid sequence of the gag–pol polyprotein that is cleaved by the protease. PIs prevent cleavage of gag and gag–pol protein precursors in acutely and chronically infected cells, arresting maturation and hence blocking the infectivity of nascent virions.[18]

The resistance of HIV against PIs can be achieved by two mechanisms. The first one involves the exchange of amino acids in the protease such that the affinity to the inhibitor is decreased while the natural substrates can be bound efficiently as opposed to the synthetic analogues.[18] Modifications of the affinity to the natural substrate alter also the efficiency of the protease. Thus, the second mechanism introduces compensatory mutations aiming at re-establishing the efficiency of the enzyme while maintaining resistance against the inhibitor. These mutations can occur both in the protease or in its substrate, at cleavage sites.[19] The four approved HIV-protease inhibitors are based on amino acid sequences: Indinavir, Nelfinavir, Ritonavir and Saquinavir.

1.4 Highly Active Antiretroviral Therapy (HAART) and resistance effect to drugs

The HIV virus develops resistance against individual drugs and inhibitors. This problem required a new pharmaceutical strategy. An approach of combination therapy, which involved combining several antiretroviral compounds, was developed. This approach benefited the most from the development of drugs in NNRTIs and PI classes. Combination therapy can block the resistance effect more effectively for two reasons; first, multiple mechanisms are required for resistance to occur to all drugs in the regimen and second; multiple drugs suppress viral replication more effectively than single agents. This marked the beginning of the era of highly active antiretroviral therapy (HAART) in 1995. HAART combines a minimum of three drugs from at least two different drug classes targeting distinct proteins.[17] A typical HAART treatment combines two NRTIs plus either one PI or one NNRTI.[20]

Combinations of antiretrovirals create multiple barriers to the HIV replication process. This helps to keep the number of offspring low and reduce the possibility of a superior mutation.

In 2006, it was reported that the number of HIV related deaths declined as compared to the pre-HAART. Consequently, multidrug regimens are necessary for successful treatment. Since each HAART agent has its own unique adverse effect profile, selecting a regimen with a proper profile may be difficult. For example, certain PIs produce adverse metabolic effects that may increase the risk of developing cardiovascular disease. On the other hand, NNRTI-based therapies may result a different side effect profile. Once this therapy is initiated, it should never be stopped.

In 2006, FDA approved the combinations of antiretrovirals (for example, Atripla). These are multiple antiretroviral drugs combined into a single pill, helping to increase adherence and thus reducing potential development of viral resistance to the drugs. This may result in longer term effectiveness of the regimen. They may combine different classes of antiretrovirals or contain only a single class. Another milestone in HAART was the discovery that the protease inhibitor ritonavir interferes with the liver enzyme cytochrome P450.[21]

This enzyme is involved in the metabolic processing of most protease inhibitors. Thus, the use of a small dose of ritonavir inhibits the liver enzyme, and helps to maintain optimal levels of other protease inhibitors in the patient's circle for a longer period of time. The boosting of protease inhibitors with ritonavir is standard as of 2001 following the introduction of Kaletra (LPV+RTV) and is usually denoted by PI/r. Despite the increasing concerns regarding antiretroviral resistance, the death rate among HIV-infected people continued to decline.[22]

HAART suffers from certain limitations, despite of its success. HAART therapy is highly effective in delaying the onset of AIDS but its clinical utility is limited by viral resistance, non-adherence to therapy, and drug toxicity. Thus, the development of new inhibitors, targeting a distinct step in the retroviral life cycle, remains essential to reduce side effects and selection of drug-resistant viruses. In this context, integration of the proviral DNA into the host cell genome, a process carried out by a specific viral enzyme, integrase, has been recently recognised as site for therapeutic intervention. Development of new chemical entities able to interrupt this process, would allow to identify key biochemical sites of HIV-1 Integrase and increase the potency and selectivity of new Integrase inhibitors.

1.5 References

1. Sanchez-Pescador R, Power MD, Barr PJ, Steimer KS, Stempien MM, Brown-Shimer SL, Gee WW, Renard A, Randolph A, Levy JA, et al. Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2). *Science*. **1985**; 227 (4686):484-92.

2. Harper ME, Marselle LM, Gallo RC, Wong-Staal F. Detection of lymphocytes expressing human T-lymphotropic virus type III in lymph nodes and peripheral blood from infected individuals by in situ hybridization. *Proc. Natl. Acad. Sci U S A.* **1986**; 83(3):772-6.

3. Starcich BR, Hahn BH, Shaw GM, McNeely PD, Modrow S, Wolf H, Parks ES, Parks WP, Josephs SF, Gallo RC, et al. Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS. *Cell.* **1986**; 45 (5):637-48.

4. Paul M, Sharp and Beatrice H. Origins of HIV and the AIDS Pandemic Hahn. *Cold Spring Harb Perspect Med.* **2011**; 1(1): a006841.

5. Metifiot M, Marchand C, Pommier Y. HIV integrase inhibitors: 20-year landmark and challenges. *Adv Pharmacol.* **2013**; 67: 75-105

6. Arts EJ and Hazuda DJ. HIV-1 Antiretroviral Drug Therapy. *Cold Spring Harb Perspect Med.* **2012**; 2(4): a007161.

7. Frankel AD, Young JA. HIV-1: fifteen proteins and an RNA. *Annu Rev Biochem.* **1998**; 67:1–25.

8. Seelamgari A, Maddukuri A, Berro R, de la Fuente C, Kehn K, Deng L, Dadgar S, Bottazzi ME, Ghedin E, Pumfery A, Kashanchi F. Role of viral regulatory and accessory proteins in HIV-1 replication. *Front Biosci.* **2004***;* 9:2388–2413.

9. Dalgleish AG, Beverley P C L, Clapham PR, Crawford DH, Greaves MF and Weiss RA. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature.* **1984;** 312, 763-767.

10. Berger1 EA, Murphy PM and Farber JM. Chemokine receptors as hiv-1 coreceptors: roles in viral entry, tropism, and disease annual. *Review of Immunology.* **1999;** 17: 657-700.

11. Broder SMD. The development of antiretroviral therapy and its impact on the HIV-1/AIDS pandemic. *Antiviral Res*. **2010**; 85(1): 1.

12. Miller MD, Hazuda DJ. HIV resistance to the fusion inhibitor enfuvirtide: mechanisms and clinical implications. *Drug Resist Updat*. **2004**; 7(2):89-95.

13. Tilton JC, Doms RW. Entry inhibitors in the treatment of HIV-1 infection. *Antiviral Res*. **2010**; 85(1):91-100.

14. Esté JA, Telenti A. HIV entry inhibitors. *Lancet.* **2007**; 370(9581):81-8.

15. Powers R, Clore GM, Stahl SJ, Wingfield PT, Gronenborn A. Analysis of the backbone dynamics of the ribonuclease H domain of the human immunodeficiency virus reverse transcriptase using nitrogen-15 relaxation measurements. *Biochemistry*. **1992**; 31 (38):9150–9157

16. Cihlar T, Ray AS. Nucleoside and nucleotide HIV reverse transcriptase inhibitors: 25 years after zidovudine. *Antiviral Res*. **2010** J; 85(1):39-58.

17. Clavel F, Hance AJ. HIV drug resistance. *N Engl J Med*. **2004**; 350(10):1023-35.

18. Flexner C. HIV-protease inhibitors. *N Engl J Med*. **1998**; 338(18):1281-92.

19. Nijhuis M1, van Maarseveen NM, Lastere S, Schipper P, Coakley E, Glass B, Rovenska M, de Jong D, Chappey C, Goedegebuure IW, Heilek-Snyder G, Dulude D, Cammack N, Brakier-Gingras L, Konvalinka J, Parkin N, Kräusslich HG, Brun-Vezinet F, Boucher CA. A novel substrate-based HIV-1 protease inhibitor drug resistance mechanism. *PLoS Med.* **2007**; 4(1):e36.

20. Dybul M, Bolan R, Condoluci D, Cox-Iyamu R, Redfield R, Hallahan CW, et al. Evaluation of initial CD4+T cell counts in individuals with newly diagnosed human immunodeficiency virus infection, by sex and race, in urban settings. *Journal of Infectious Diseases*. **2002**; 185:1818–1821.

21. Kumar GN, Rodrigues AD, Buko AM and Denissen JF. Cytochrome P450-mediated metabolism of the HIV-1 protease inhibitor ritonavir (ABT-538) in human liver microsomes. *Journal of Pharmacology and Experimental Therapeutics*. **1996**; 277 (1) 423-431;

22. Eras C, Riffenburgh NF, Robert H, Wegner S, Agan BK, Tasker SA, Spooner KM, Armstrong AW, Fraser S, Wallace MR. Epidemiology and Social Science Comparisons of Causes of Death and Mortality Rates Among HIV-Infected Persons: Analysis of the Pre-, Early, and Late HAART. *JAIDS.* **2006**; 41(2):194-200.

CHAPTER II:

The HIV-1 integrase as molecular target

2.1 Integrase and integration: biochemical activities of HIV-1 integrase

2.1.1 Integrase protein

As mentioned above, the human immunodeficiency virus (HIV) that causes acquired immunodeficiency syndrome (AIDS), belongs to the family of Retroviridae, genus Lentivirinae. Lentiviruses represent a kind of slow viruses with a long incubation period (months, even years) and a propensity to induce a wide range of diseases in different animal species. According to the recent classification of the International Committee on Taxonomy of Viruses (ICTV), the genus Lentivirus consists of nine species, seven lentivirus animal and two human lentivirus. At the same time there exist different species of integrase. For example, PFV IN is significantly longer, comprises 392 residues and ASV IN is decoded into 323 amino acids and is only in post-translational phase modified polypeptide comprising 286 residues, which is the active form.[1]

Although the enzyme encoded by avian sarcoma virus (ASV) has been studied long before, it is clear that is more interesting to study HIV IN.[2]

The integrase of HIV-1 is structurally characterized by a single polypeptide chain consisting of 290 amino acid residues and consists of three clearly identifiable domains and inter-domain linker.[3]

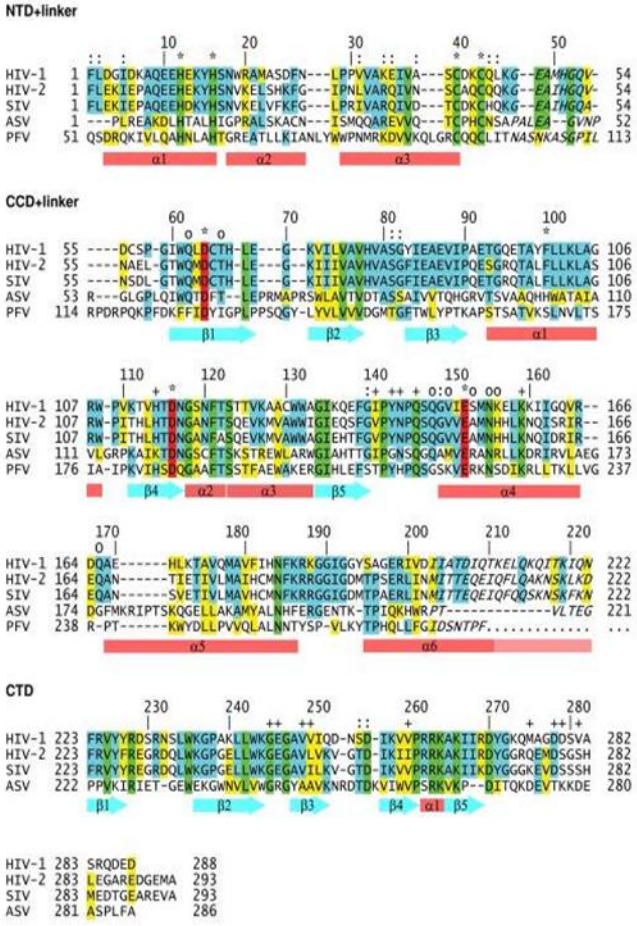


Figure 2.1 *Amino acid sequence alignment of retroviral integrase.*
Adapted from: Jaskolski M, Alexandratos JN, Bujacz G and Wlodawer A. Piecing together the structure of retroviral integrase, an important target in AIDS therapy. FEBS J. ***2009****; 276(11): 2926–2946.*

As shown in the figure 2.1, the N-terminal domain (NTD) of HIV-1 IN contains the residues from 1 to 46, followed by a linker consisting of from 47-55 residues. The catalytic domain (CCD) contains 56-202 residues, and is followed by a linker sequence which comprises the amino acids from 203 to 219. Finally, the C-terminal domain (CTD) contains 220-288 residues. The number of domain residues of HIV-2 and SIV enzymes is approximately equal, which differs for ASV IN.

For the PFV IN, it is possible that there is an additional domain consists of approximately 50 residues preceding the NTD domain.

Analyzing for individual domains the ratio of the percentage identity/similarity are obtained the following data: for NTD 55/76% comparing HIV-1 IN to SIV, 26/46% comparing ASV IN; for the CCD are respectively 61/77% and 27/46% and for the CTD are 53/68% and 14/25% respectively. It is know that the lower sequence conservation is for the C-terminal domain.[4]

However, it should be emphasized, that the domains are characterized by a highly conserved sequence in all retroviral integrase:

• CCD has three highly conserved amino acid residues (DDE motif)

• NTD sequence corresponds to HHCC residues.

2.1.2 The role of integrase in HIV-1 replication

Integrase (IN) is a key enzyme for the integration of the HIV-1 genome into the host cell chromosome and, therefore, a very promising target for anti-AIDS drug design. IN performs two essential catalytic reactions.[5]

Following reverse transcription, the viral cDNA is primed for integration in the cytoplasm by integrase-mediated trimming of the 3′-ends of the viral DNA (*figure 2.2*). This step is referred to as 3′-Processing. It requires both fully functional integrase and the integrity of the last 10–20 base pairs at both ends of the viral cDNA. 3′-processing consists of the endonucleolytic cleavage of the 3′-ends of the viral DNA.

This cleavage occurs immediately 3′ to a conserved CA dinucleotide motif. Alterations of this sequence prevent integrase from catalysing 3′-processing. This reaction generates CA-3′-hydroxyl DNA ends, which are the reactive intermediates required for Strand Transfer.

Following 3′-processing, integrase remains bound to the viral cDNA as a multimeric complex that bridges both ends of the viral DNA within the PIC.

The PIC contains both viral and cellular proteins in addition to the integrase–DNA complexes. The viral proteins reverse transcriptase (RT), matrix (Ma), nucleocapsid (Nc) and Vpr can contribute to the transport of PICs through the nuclear envelope. Some cellular proteins, packaged within PICs, can bind to integrase and stimulate the enzymatic activities of integrase. These proteins include interactor 1 (INI1)16 (the first integrase-binding protein discovered), lens epithelium derived growth factor (LEDGF, also known as p75);[6] embryonic ectoderm-development protein 18 and heat shock protein 60 (HSP60).[7] Promyelocytic leukaemia protein (PML) also co-localizes and co-migrates with PICs20. Two cellular proteins, high-mobility group protein A1 (HMGA1, also known as HMG1(Y)) and barrier to auto-integration factor (BAF), regulate integration by binding to DNA directly. HMGA1 stimulates integrase activity;[8,9] BAF stimulates intermolecular integration and suppresses auto-integration. By contrast to other lentiviruses, such as the oncoretroviruses murine Moloney virus and Rous sarcoma virus, which require mitotic nuclear-envelope breakdown to access the chromosomes of infected cells, HIV-1 PICs are able to cross the nuclear envelope. The karyophilic property of the PICs enables HIV to replicate in not proliferative cells, such as macrophages.[10] Once in the nucleus, integrase catalyzes the insertion of the viral cDNA ends into host chromosomes. This Strand Transfer reaction consists of the ligation of the viral 3′-OH DNA ends (generated by 3′-processing) to the 5′-DNA phosphate of a host chromosome.

Integrase can also catalyse the reverse reaction, referred to as Disintegration.[11] Physiological integration requires the concerted joining of both ends of the viral cDNA on opposite DNA strands of the target (acceptor DNA) host chromosome with a canonical fivebase-pair stagger. The five-base stagger indicates that each viral cDNA end attacks the chromosomal DNA across its major groove. Completion of integration requires ligation of the 5′-end of the viral DNA. This last step of integration can only take place after trimming of the last two nucleotides at the proviral DNA 5′-ends and extension (gap filling) from the 3′-OH remodelling, and transcription complexes, such as those bound to integrase in the PICs (and described above), are implicated in the selection of the HIV integration sites within transcribing genes. The fully integrated viral genome is also referred to as the provirus.

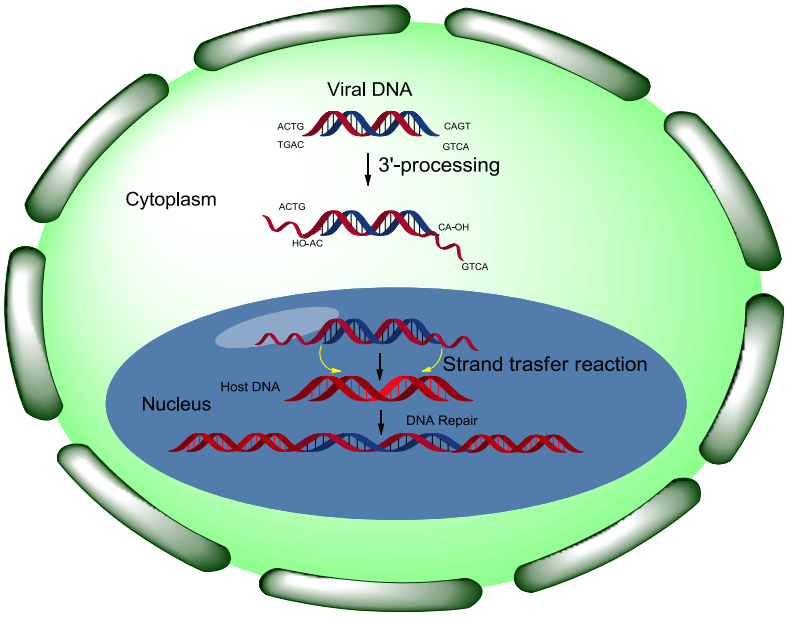


Figure 2.2 *The role of integrase in the life cycle of HIV virus.*

2.1.3 HIV integrase structure and functional domains

Integrase is generated during virus maturation by cleavage of the Pol polyprotein by HIV protease. HIV integrase is a 32-kDa protein comprising three structural domains: the N-terminal domain (NTD), the catalytic core domain (CCD) and the C-terminal domain (CTD) (*figure 2.3*). The atomic structure of each of these domains has been determined by X-ray diffraction or solution nuclear magnetic resonance (NMR).

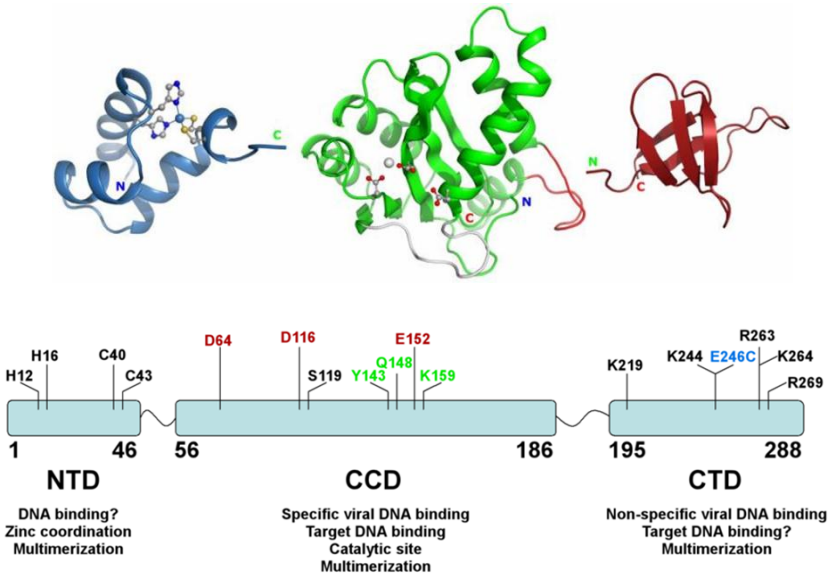


Figure 2.3 *Functional domain of integrase.*
Adapted from: Kessl JJ, McKee CJ, Eidahl JO, Shkriabai N, Katz A and Kvaratskhelia M. HIV-1 Integrase-DNA Recognition Mechanisms. Viruses. ***2009****; 1(3), 713-736.*

The figure 2.4 shows the structures of the CCD in association with the CTD,[12] and of the NTD with the CCD.[13] CCD, which encompasses residues 50–212, forms a dimer in all the structures examined. It is, structurally, remarkably similar to other retroviral integrases (MLV and avian sarcoma virus (ASV)), to the Tn5, to RNase H, to the Holiday junction recombinase RuvC35 and to the PIWI domain of Argonaute,[14]

the RNase associated with Dicer in RNA interference. This family of DNA processing enzymes (polynucleotide transferases) contains a canonical three amino acid, which formed in HIV-1 integrase the catalytic triad D64, D116 and E152. These residues are highly conserved in all integrases and retrotransposases (DDE motif).[15] Mutation of any of these three acidic residues abolishes integrase's enzymatic activities and viral replication. The two D64 and D116 residues form a coordination complex (chemical bonds) with a divalent metal (Mg^{2+} or Mn^{2+}). Because a second metal has been observed in an ASV integrase crystal structure,[16] and because of the two-metal structure for polynucleotide transferases,[17] it has been proposed that a second metal (Mg^{2+} or Mn^{2+}) can be coordinated between D116 and E152 once HIV-1 integrase binds its DNA substrate(s).[18] It is therefore likely that the metal(s) coordinate(s) integrase and the phosphodiester backbone of the DNA substrate(s) during the 3′-processing and strand-transfer steps. In most structures, the CCD contains a short disordered loop (encompassing residues 141–150), the structure of which can be stabilized by DNA. Although the CCD contains the enzyme catalytic site(s), in the absence of the NTD and CTD it can only catalyze the disintegration reaction the reverse of the strand-transfer reaction in vitro.[19] Disintegration is the only reaction catalyzed by the isolated CCD. To catalyze 3′-processing and strand transfer, the CCD needs both the NTD and CTD in a dimeric complex.[20] The NTD encompasses residues 1–50 and contains an HHCC motif that is common to all retroviral integrases. Binding of one Zn^{2+} atom to the HHCC motif stabilizes the folding of the NTD domain and is required for integrase activity. Single mutations of any of these four residues reduce integrase enzymatic activity.[21] The NTD dimer interface is different in the crystal structures and the solution NMR structure (not

shown), which is indicative of multiple arrangements of the integrase multimers. The NTD is the preferential binding region for two cellular transcription factors in the PICs, INI1 and LEDGF/p75. The CTD, which encompasses residues 212–288, has an overall SH3 fold. It binds DNA nonspecifically and is required for integrase 3′-processing and strand-transfer activities. The CTD binds the cellular embryonic ectoderm development protein as well as RT, and this interaction seems to be required for reverse transcription. The figure 2.4 shows the structures of the CCD both with the NTD and the CTD. Together these two structures indicate the possibility that the NTD is positioned between the CCD and CTD, next to the extended α-helix joining the CCD and the CTD. The solution NMR dimer interfaces for the NTD and the CTD, which are different from those observed in the crystallographic structure might be used in higher-order complexes (tetramers and/or octamers), which have been proposed to correspond to the active enzyme.

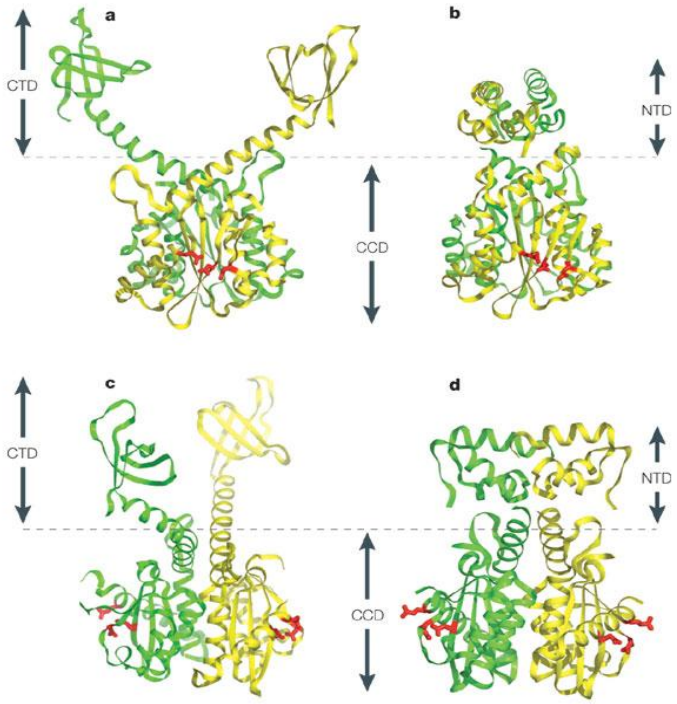


Figure 2.4 *The structures of the CCD in association with the CTD, and of the NTD with the CCD. a and b | Side views showing the catalytic acidic triad in red (the canonical DDE motif consisting of residues D64, D116, E152; BOX 1, figure part a) in the catalytic core domain (CCD) of integrase. The two subunits of the dimer are shown in yellow and green. c and d | Front views of the same structures (after 90° anticlockwise rotation of panels a and b, respectively). a and c | Structure of the CCD–carboxy-terminal domain (CTD) dimer; (PDB codes: 1EXQ & 1EX4). b and d | Structure of the amino-terminal domain (NTD)–CCD (PDB code: 1K6Y). Combining the structures (a with b; c with d) indicates the positioning of each NTD into the cavity between the CCD and CTD in the full-size integrase dimer. The functional structure of integrase is probably tetrameric, and would therefore involve another dimer interface (unknown, and therefore not represented here).*
Adapted from: Pommier Y, Johnson AA and Marchand C. Integrase inhibitors to treat HIV/Aids. Nature Reviews Drug Discovery. ***2005****; 4, 236-248.*

2.1.4 IN interactions with host cell proteins

Purified IN proteins catalyzes 3′ processing and DNA strand transfer activities in vitro,[22,23] indeed, the results obtained by numerous studies indicate that cell proteins play important roles, during virus infection. In particular, a crucial role for the IN-interacting protein lens epithelium-

derived growth factor (LEDGF)/p75 in HIV-1 replication and integration has highlighted.[24] However, the function of LEDGF/p75 in viral replication is unknown.

The LEDGF/p75 belongs to the hepatoma-derived growth factor (HDGF) related protein (HRP) family, which is defined by the amino acid sequence conservation of an N-terminal Pro-Trp-Trp-Pro (PWWP) domain (*figure 2.5*).[25] The binding occurs through a conserved IN-binding domain (IBD) found within the C-terminal portion of the larger p75 LEDGF splice variant. The IBD is essential for stimulation of IN activity in vitro and for LEDGF/p75 function during HIV-1 infection.[26] LEDGF/p75 might therefore act as a critical costimulator of IN activity.[27] Ectopically expressed HIV-1 IN is degraded by the proteasome in human cells, and LEDGF/p75 significantly increases its stability. The HIV-1 PIC can be degraded by the proteasome, so the IN-LEDGF/p75 interaction might help maintain PIC integrity during infection. The functional HIV-1 and feline immunodeficiency virus PICs were recovered from cytoplasmic extracts of infected cells using anti-LEDGF antibodies.[28]

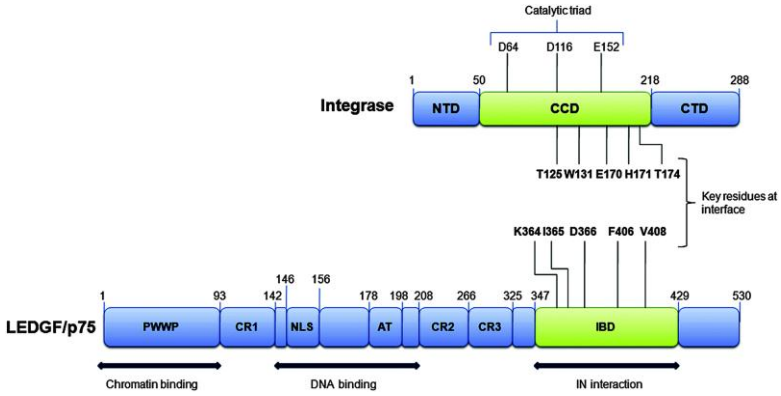


Figure 2.5 *Region of the CCD interaction with IBD.*
Adapted from: Reddy KK, Singh P and Singh SK. Blocking the interaction between HIV-1 integrase and human LEDGF/p75: mutational studies, virtual screening and molecular dynamics simulations. Mol. BioSyst. ***2014****;10, 526-536.*

Alternatively, LEDGF/p75 might function as an obligate chromatin acceptor for the PIC. This hypothesis indicates that LEDGF/p75 intimately associates with chromatin, and its N-terminal PWWP domain and AT-hook (ATh) DNA-binding motifs, which mediate chromatin binding, are required for HIV-1 infection.[29]

It was identified the NMR structure of the integrase-binding domain (IBD) in LEDGF and identified amino acid residues essential for the interaction. The IBD is a compact right-handed bundle composed of five alpha-helices. Based on folding topology, the IBD is structurally related to a diverse family of alpha-helical proteins that includes eukaryotic translation initiation factor eIF4G and karyopherin-beta. LEDGF residues essential for the interaction with IN were localized to interhelical loop regions of the bundle structure. Interaction-defective IN mutants were previously shown to cripple replication although they retained catalytic function. The initial structure determination of a host cell factor that tightly binds to a retroviral enzyme lays the groundwork

for understanding enzyme-host interactions important for viral replication.

2.2 Integrase inhibitors

Until now, only inhibitors targeting the catalytic site of IN with a specific effect on strand-transfer process (INSTIs) have been identified and developed.[30] In 2007, the INSTI Raltegravir became the first IN inhibitor approved for use in the treatment of HIV infected patients.[31] More recently, the FDA also approved Stribild, a single-tablet regimen HIV medication containing four drug combination of Elvitegravir, another HIV INSTI, Cobicistat, a CYP3A inhibitor, and Emtricitabine and Tenofovir DF, both HIV nucleoside analogue reverse transcriptase inhibitors (*figure 2.6).*[32]

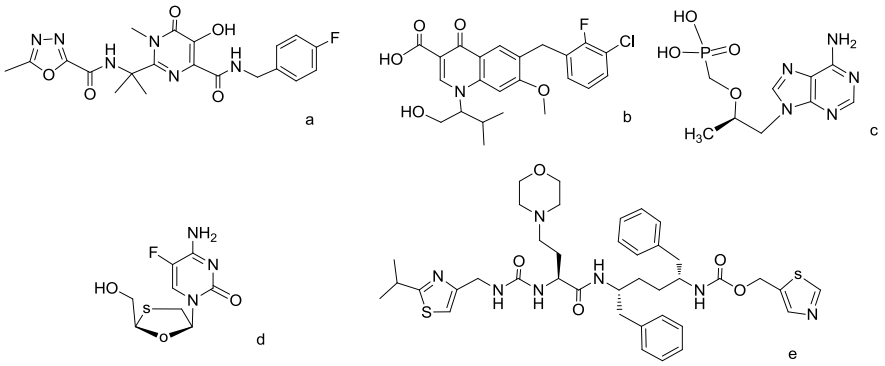


Figure 2.6 *Conventional drugs approved by the FDA for the treatment in patients with HIV. a) Raltegravir; b) Elvitegravir; c) Tenofovir; d) Emtricitabine; e) Cobicistat.*

However, the emergence of the resistance to these drugs gives raise to the pressing need of novel IN inhibitors.[33] Absence of information on the structures of a full-sized enzyme, on its complex with DNA and also on

the PIC composition and operation complicates and slows the search for novel IN inhibitors.

To overcome these drawbacks, targeting allosteric sites of the protein including interaction site of IN with cellular co-factor essential for integration, for example the LEDGF/p75-IN interaction, or oligomerization/multi oligomerization sites might represent alternative approaches to IN inhibition.[34]

2.2.1 Integrase allosteric inhibitors

The design and development of compounds targeting integrase in a different way open a route to bypass the cross-resistance problematic of INSTIs.

Recently, a structure-based design approach resulted in the discovery of 2-(quinolin-3-yl)acetic acid derivatives.[35] These first class of IN inhibitors are named "LEDGINs" since these compounds bind in the LEDGF/p75 binding pocket of IN and block the interaction of LEDGF/p75 with IN. LEDGINs likely also affect the catalytic activity of IN, since LEDGF/p75 binding allosterically modulates integrase activity. As a consequence, LEDGINs potently inhibit HIV replication in cell culture.[36,37,38] Other studies have shown that Ledgin show an additive effect with INSTI. The binding site of the LEDGF/p75 is localized at the interface of the IN monomer-monomer, involving both of amino acid residues. So it was assumed its indirect role to promove oligomerization of IN.[39,40]

In fact, a study of LEDGIN-7, which mimics the amino acid sequence 365-368 of LEDGF/p75, showed effective inhibition of tetramerization of IN.

Other peptides derived from LEDGF/p75, which have amino acid sequence 353-378, 361-370, 402-411, hexhibited in vitro inhibitory activity for both binding IN-LEDGF/ p75 and the tetramerization of IN. In particular, this effect has been suggested since the peptides showed inhibition in the presence and in the absence of LEDGF/ p75.

Therefore, the formation of the tetramer, represent a very interesting target for the development of new allosteric inhibitors.

2.2.2 Integrase tetramerization inhibitors

The first approach to the inhibition of integrase oligomerization process resulted the development of two peptides derived from the interface of the dimer IN CCD: INH1 (ATGQETAYFLLKLAGKA) and INH5 (DQAEHLKTAVQMAVFIHNYKA).

The first reproduces the amino acid sequence of the helix α1 (aa 93-107), the second reproduces the α5 helix and part of the loop which separates the α4 and α5 helices (aa167-187) (*Figure 2.7*). Experimental studies confirmed the dissociation of the tetramer by a INH1 and INH5, resulting in loss of enzymatic IN.[41]

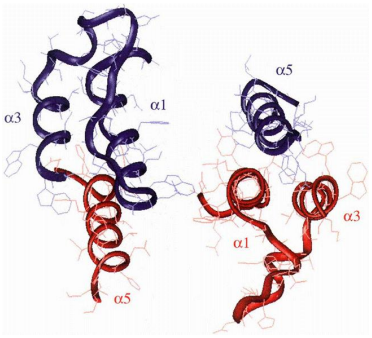


Figure 2.7 *Structure of the catalytic domain dimerization of HIV-1 IN. The two monomers are shown in red and blue. The propellers a1 and α5, from which the inhibitory peptides INH1 and INH5 respectively, are indicated with the α3 helix. Adapted from: Maroun RG, Gayet S, Benleulmi MS, Porumb H, Zargarian L, Merad H, Leh H, Mouscadet JF, Troalen F, Fermandjian S.*

Peptide inhibitors of HIV-1 integrase dissociate the enzyme oligomers. Biochemistry. ***2001****; 40: 13840-13848.*

Other studies, have led to the identification of two peptides, NL6 (TAYFLLKLAGRW) NL9 (ACWWAGIKQEF), with low values of IC_{50} for IN inhibition. These derived from the α1 helical domain (NL6) and α3 helical domain respectively (NL9) (*figure 2.8*) .

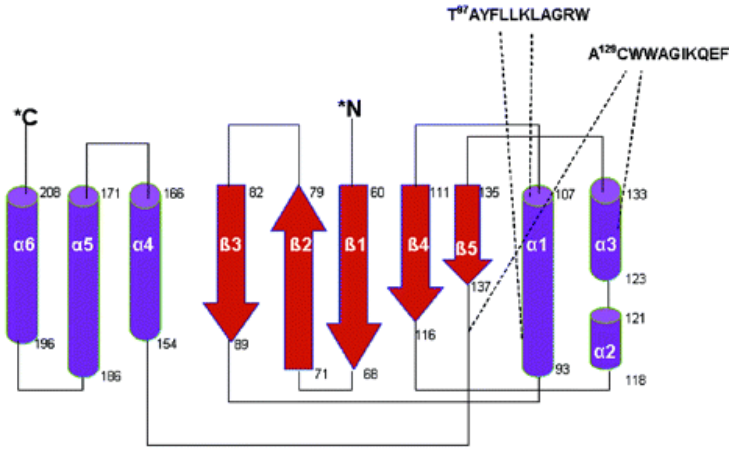


Figure 2.8 *IN secondary structure. The peptides derived from α1 (NL6) and α3 (NL9). Adapted from: Al-Mawsawi LQ, Neamati N. Allosteric inhibitor development targeting HIV-1 integrase. Chem. Med. Chem.* ***2011****; 6: 228-41.*

To determine the minimal active sequence, based on these sequences have been designed other peptides, only NL6-5 (YFLLKL), which contains the 6 central residues of NL6, has retained inhibitory activity. Furthermore they were synthesized analogues of NL6 and NL9 peptides with the amino acid sequence is inverted (RNL6, RNL9) and with the replacement of L-amino acids with D-amino acids (DNL6, RDNL6, DNL9, RDNL9). However, only analogues of NL6 have yet shown inhibitory activity.[42]

2.2.3 Integrase inhibitors derived from N-terminal domain

The N-terminal integrase plays an important role in the formation of tetramer integrase process. Therefore, this domain represents an important target for the design of drugs for inhibition of IN oligomerization process. In this context, compounds that target the IN N-terminal domain would be effective at disrupting IN function through a range of mechanisms, including allosteric or oligomerization inhibition. Recently, considering the amino acid sequence (1-50) of the N-terminal domain of integrase, the peptides NZ-1, NZ-2, N-ITFcc and NZ-4 were synthesized. Biological assays showed that peptide N-ITFcc, corresponding to the third α-helix (FNLPPVVAKEIVAS), inhibits viral replication almost 80%, indicating a possible use as an antiviral drug.[43]
Thus, we considered this domain an important starting point for the identification of peptide inhibitors in order to clarify some points on IN function in viral replication, such as the elucidation of HIV-1 IN polymerization state, or its potentiality in the structure-based design.
In my PhD project, we present the design and synthesis of peptides targeting the integration events by directly inhibiting IN or its multimerization process. The synthesized peptides were assayed in vitro for their ability to inhibit IN strand transfer activity and for the capability to inhibit IN dimerization or to promote IN multimerization. Finally, the most potent compounds, conveniently conjugated with cell-penetrating fragment Tat, were assayed in MT-4 cells for determining anti-HIV infective activity.

2.3 References

1. Katz RA, and Skalka AM. A C-terminal domain in the avian sarcoma-leukosis virus pol gene product is not essential for viral replication. *J. Virol.* **1988**; 62 (2): 528-533.

2. Grandgenett DP, Ajaykumar CV, Schiff RD. A 32,000-Dalton nucleic acid-binding protein from avian retravirus cores possesses DNA endonuclease activity. *Virology.* **1978;** 89 (1): 119-132.

3. Bushman FD, and Wang B. Rous sarcoma virus integrase protein: mapping functions for catalysis and substrate binding. *J. Virol.* April **1994**; 68 (4): 2215-2223.

4. Rhee SY, Liu TF, Kiuchi M, Zioni R, Gifford RJ, Holmes SP and Shafer RW. Natural variation of HIV-1 group M integrase: Implications for a new class of antiretroviral inhibitors. *Retrovirology.* **2008;** 20085: 74.

5. Delelis O, Carayon K, Saïb A, Deprez E and Mouscadet JF. Integrase and integration: biochemical activities of HIV-1 integrase. *Retrovirology*. **2008;** 20085: 114.

6. Cherepanov P, Maertens G, Proost P, Devreese B, Van Beeumen J, Engelborghs Y, De Clercq E, Debyser Z. HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J.Biol. Chem*. **2002**; 278: 372–381.

7. Parissi V, Calmels C, De Soultrait VR, Caumont A, Fournier M, Chaignepain S and Litvak S. Functional interactions of human immunodeficiency virus type 1 integrase with human and yeast HSP60. *J. Virol*. **2001;** 75: 11344–11353.

8. Hindmarsh P, Ridky T, Reeves R, Andrake M, Skalka AM and Leis J. HMG protein family members stimulate human immunodeficiency virus type 1 and avian sarcoma virus concerted DNA integration in vitro. *J. Virol*. **1999**; 73:2994–3003.

9. Farnet C and Bushman FD. HIV-1 cDNA integration: requirement of HMG I (Y) protein for function of preintegration complexes in vitro. *Cell*. **1997;** 88: 483–492.

10. Bukrinsky MI, Sharova N, Dempsey MP, Stanwick TL, Bukrinskaya AG, Haggerty S and Stevenson M. Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. *Proc. Natl. Acad. Sci.* **1992;** 89:6580–6584.

11. Chow SA., Vincent, KA, Ellison V and Brown PO. Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. *Science.* **1992;** 255: 723–726.

12. Chen JC, Krucinski J, Miercke LJ, Finer-Moore JS, Tang AH, Leavitt AD, Stroud RM.. Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. *Proc. Natl Acad. Sci.* **2000**; 97:8233–8238.

13. Wang JY, Ling H, Yang W and Craigie R. Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. *EMBO J*. **2001**; 20:7333–7343.

14. Yang W and Steitz TA. Recombining the structures of HIV integrase, RuvC and RNase H. *Structure*. **1995;** 3:131–134.

15. Song JJ, Smith SK, Hannon GJ and and Joshua-Tor L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science.* **2004;** 305:1434–1437.

16. Rice PA and Baker TA. Comparative architecture of transposase and integrase complexes. *Nature Struct. Biol*. **2001;** 8:302–307.

17. Bujacz G, Alexandratos J, Wlodawer A, Merkel G, Andrake M, Richard A, Skalka K and AM. Binding of different divalent cations to the active site of avian sarcoma virus integrase and their effects on enzymatic activity. *J. Biol. Chem*. **1997**; 272: 18161–18168.

18. Beese LS and Steitz TA. Structural basis for the 3'–5' exonuclease activity of Escherichia coli DNA polymerase I: a two metal ion mechanism. *EMBO J.* **1991;** 10: 25–33.

19. Grobler JA, Stillmock K, Hu B, Witmer M, Felock P, Espeseth AS, Wolfe A, Egbertson Melissa, Bourgeois M, Melamed J, Wai JS, Young S, Vacca J and Hazuda DJ. Diketo acid inhibitor mechanism and HIV-1 integrase: implications for metal binding in the active site

of phosphotransferase enzymes. *Proc. Natl Acad. Sci*. **2002;** 99: 6661–6666.

20. Marchand C, Johnson AA, Karki RG, Pais GC, Zhang X, Cowansage K, Patel TA, Nicklaus MC, Burke TR Jr, Pommier Y. Metal-dependent inhibition of HIV-1 integrase by β-diketo acids and resistance of the soluble double-mutant (F185K/C280S). *Mol. Pharmacol*. **2003;** 64: 600–609.

21. Chow SA and Brown PO. Juxtaposition of two viral DNA ends in a bimolecular disintegration reaction mediated by multimers of human immunodeficiency virus type 1 or murine leukemia virus integrase. *J. Virol.* **1994;** 68: 7869–7878.

22. Engelman A, Bushman FD and Craigie R. Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. *EMBO J*. **1993;** 12: 3269–3275.

23. Engelman A, and Craigie R. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. *J. Virol*. **1992;** 66, 6361–6369.

24. Ming-Chieh S, Nidhanapati K, Raghavendra N, Vandegraaff J, Daigle SH, Kellam P, Cherepanov P, and Engelman A. LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration *Genes Dev*. **2007**; 21(14): 1767–1778.

25. Izumoto Y, Kuroda T, Harada H, Kishimoto T, Nakamura H. Hepatoma-derived growth factor belongs to a gene family in mice showing significant homology in the amino terminus. *Biochem Biophys Res Commun* **1997**; 238(1):26-32.

26. Cherepanov P, Devroe E, Silver PA, Engelman A. Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. *J Biol Chem* **2004**; 279(47):48883-92.

27. Vandegraaff N and Engelman A. Molecular mechanisms of HIV integration and therapeutic intervention. *Expert Rev Mol Med* **2007;** 9(6):1-19.

28. Llano M, Vanegas M, Fregoso O, Saenz D, Chung S, Peretz M, Poeschla EM. LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes. *J Virol.* **2004**; 78(17):9524-37.

29. Llano M, Vanegas M, Hutchins N, Thompson D, Delgado S, Poeschla EM. Identification and characterization of the chromatin-binding domains of the HIV-1 integrase interactor LEDGF/p75. *J Mol Biol.* **2006;** 360(4):760-73

30. Esposito F and Tramontano E. Past and Future. Current drugs targeting HIV-1 Integrase and Reverse Transcriptase-associated Ribonuclease H activity: single and dual active site inhibitors. *Antiv. Chem Chemother*. **2013;** 23, 129-144.

31. Evering, TH and Markowitz M.. Raltegravir (MK-0518): an integrase inhibitor for the treatment of HIV-1. *Drugs Today*. **2007**; 43, 865-877.

32. Brinson, C and. Stribild. A Single Tablet Regimen for the Treatment of HIV Disease. *Comb. Prod. The.* **2013**; 3, 1-8.

33. Mesplède T, Quashie PK, Wainberg MA. Resistance to HIV integrase inhibitors. *Curr Opin HIV AIDS*. **2012**; 7, 401-408.

34. Long YQ, Huang SX, Zawahir Z, Xu ZL, Li H, Sanchez TW, Zhi Y, De Houwer S, Christ F, Debyser Z, Neamati N. Design of Cell-Permeable Stapled Peptides as HIV-1 Integrase Inhibitors. *J. Med. Chem.* **2013;** 56, 5601-5612.

35. Christ F, Voet A, Marchand A, Nicolet S, Desimmie BA, Marchand D, Bardiot D, Van der Veken NJ.; Van Remoortel B, Strelkov SV, De Maeyer M, Chaltin P, Debyser Z. Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. *Nat. Chem. Biol.* **2010**; 6: 442−448.

36. Christ F, Shaw S, Demeulemeester J, Desimmie BA, Marchand A, Butler S, Smets W, Chaltin P, Westby M, Debyser Z, Pickford C. Small-molecule inhibitors of the LEDGF/p75 binding site of integrase block HIV replication and modulate integrase multimerization. *Antimicrob. Agents Chemother.* **2012**; 56: 4365-4374.

37. Kessl JJ, Jena N, Koh Y, Taskent-Sezgin H, Slaughter A, Feng L, de Silva S, Wu L, Le Grice SF, Engelman A, Fuchs JR, Kvaratskhelia M. A multimode, cooperative mechanism of action of allosteric HIV-1 integrase inhibitors. *J. Biol. Chem*. **2012**; 287: 16801-16811.

38. Tsiang M, Jones GS, Niedziela-Majka A, Kan E, Lansdon EB, Huang W, Hung M, Samuel D, Novikov N, Xu Y, Mitchell M, Guo H, Babaoglu K, Liu X, Geleziunas R, Sakowicz R. New class of HIV-1 integrase (IN) inhibitors with a dual mode of action. *J. Biol. Chem*., **2012**; 287: 21189−21203.

39. Al-Mawsawi LQ, Christ F, Dayam R, Debyser Z, Neamati N. Inhibitory Profile of a LEDGF/p75 Peptide against HIV-1 integrase: insight into integrase-DNA complex formation and catalysis. *FEBS Lett.* **2008**; 582: 1425-1430.

40. Demeulemeester J, Tintori C, Botta M, Debyser Z, Christ F. Development of an AlphaScreen-Based HIV-1 integrase dimerization assay for discovery of novel allosteric inhibitors. *J. Biomol. Screen.* **2012**; 17: 618-628.

41. Maroun RG, Gayet S, Benleulmi MS, Porumb H, Zargarian L, Merad H, Leh H, Mouscadet JF, Troalen F, Fermandjian S. Peptide inhibitors of HIV-1 integrase dissociate the enzyme oligomers. *Biochemistry*. **2001**; 40: 13840-13848.

42. Al-Mawsawi LQ, Neamati N. Allosteric inhibitor development targeting HIV-1 integrase. *Chem. Med. Chem*. **2011**; 6: 228-41.

43. Masuda T, Kannagi M. European Patent Specification, International application number: PCT/JP2000/002698, International publication number: WO 2000/070035 (23.11.**2000** Gazette 2000/47).

CHAPTER III:

Design, synthesis, biological evaluation and conformational studies of peptides derived from N-terminal domain HIV-1 integrase

The main goals of this project were been the design and synthesis of peptides and peptidomimetics able to inhibit the activity of IN. Subsequently, the development of these compounds was based on: a) elaboration of a pharmacophoric model concerning the activity of the studied peptides; through the combination of synthetic methods and conformational analysis, b) optimization of the biological potency and selectivity to identify the lead compounds by multiple cycles of design-synthesis-biological evaluation, c) design and synthesis of new molecules with reduced conformational freedom, in which the postulated pharmacophoric conformation would be highly probable.

These studies led to the design of peptides able to modulate the activity of integrase.

3.1 Results

3.1.1 Design of peptides derived by the sequence NTD IN 1-50

Starting from HIV-1 IN NTD, we synthesized an overlapping peptide library to generate a library of peptide sequences of specific length and specific offset, to cover the entire native protein sequence NTD IN 1-50 (peptides **1-15**, *figure 3.1*). The solution structure of NTD has been solved by 3D and 2D NMR spectroscopy, classifying NTD as an "helix-turn-helix" structure. Indeed, it consists of three α helices, which range from residues 6–15, 19–25 and 31–39, stabilized by a Zn^{2+} ion coordinated with the HHCC motif. Ile^{5}, Ala^{8}, Leu^{28}, Ala^{33} and Ile^{36} form part of the hydrophobic core, which explains their importance for structural integrity.[1] As showed in figure 3.1, peptide **2** and peptides **5**, **6**,

7 were included into the 1[rd] α-helix and 2[rd] α-helix respectively, while peptide **11** forms the 3[rd] α-helix of the NTD, and incorporates the Cys[40] of the HHCC motif. In this section we focused on peptide **11.**

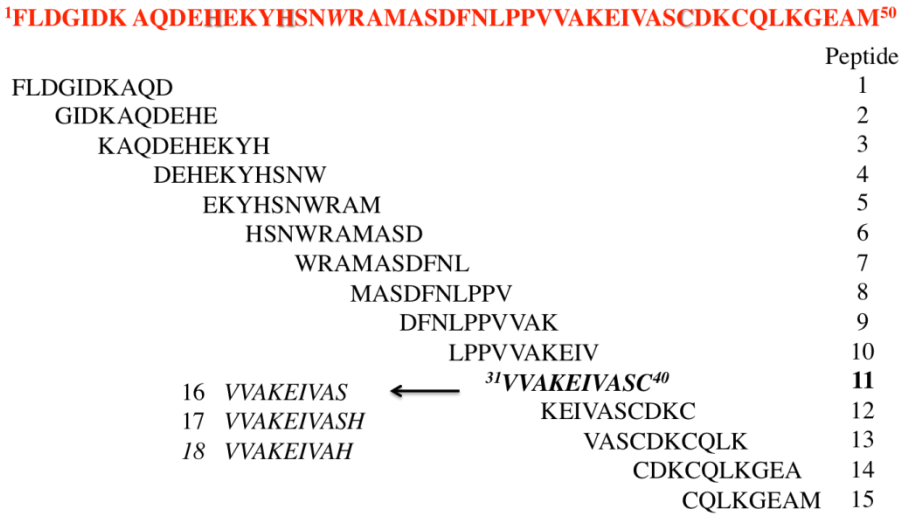


Figure 3.1. *Sequences of the designed peptides in this study.*
*Adapted from: Sala M, Spensiero A, Esposito F, Scala M C, Vernieri E, Bertamino A, Manfra M, Carotenuto A, Grieco P, Novellino E, Cadeddu M, Tramontano E, Schols D, Campiglia P, Gomez-Monterrey I M. Development and Identification of a Novel Anti-HIV-1 Peptide Derived by Modification of the N-Terminal Domain of HIV-1 Integrase. Frontiers in Microbiology, Virology. **2016**; 7, 845.*

The relevance of both a α-helix structure and a Zn^{2+} chelating group on the inhibitory activity was assessed by the synthesis of three peptides derived from peptide **11**, in which the Cys residue was deleted (peptide **16**) or substituted by His amino acid a residue also involved in the Zn^{2+} coordination (peptide **17**).[2,3] The same His was also used to replace the C-terminal Ser residue of peptide **16** leading to peptide **18**.

3.1.2 Biochemical studies: HIV-1 IN catalytic activity inhibition (peptides 1-18)

The synthesized peptides **1** to **18** were tested *in vitro* by the company Express Biotech International, for their ability to inhibit the HIV-1 IN. Results showed that peptide **18** was the most potent of the series and inhibit the HIV-1 IN activity with an IC_{50} values of 4.5 µM (*Table 3.1*). Peptides **9** and **12** containing the VVAK and VA fragments of peptide **18**, respectively, and peptide **2**, which includes the 6-13 fragment of the first α-helix, maintained a certain inhibitory activity. In this assay, we use Raltegravir as a controls.

Table 3.1 Sequence, analytical data and effect of peptides **1-18** on HIV-1 IN activities.

Peptide	Sequence	HPLC k'	ESI-MS	IN inhibition IC_{50}(μM)
1	^{1}FLDGIDKAQD	4.70	1163.1	NA
2	^{4}GIDKAQDEHE	5.03	1182.5	162±9
3	^{7}KAQDEHEKYH	4.63	1326.7	NE
4	^{10}DEHEKYHSNW	4.75	1385.5	NA
5	^{13}EKYHSNWRA	5.02	1363.7	NE
6	^{16}HSNWRAMASD	2.94	1215.5	NE
7	^{19}WRAMASDFNL	2.30	1251.7	NE
8	^{22}MASDFNLPPV	1.30	1131.6	NA
9	^{25}DFNLPPVVAK	3.70	1140.6	213±12
10	^{28}LPPVVAKEIV	4.40	1105.8	NA
11	^{31}VVAKEIVASC	2.12	1059.7	NA
12	^{34}KEIVASCDKC	3.06	1136.5	477±23
13	^{37}VASCDKCQLK	2.80	1135.5	NA
14	^{40}CDKCQLKGEA	2.60	1135.4	NA
15	^{43}CQLKGEAM	3.70	920.3	NA
16	^{31}VVAKEIVAS	2.50	956.5	NA
17	VVAKEIVASH	5.22	1094.0	NA
18	VVAKEIVAH	3.50	1007.0	4.5±0.9
Ral.				0.18±0.01

All peptide are acetylated and amidate at N-terminal and C-terminal respectively.
k' [(peptide retention time – solvent retention time)/solvent retention time]. Peptide concentration μM.
NE = No effect. The peptide demonstrated no significant effect in the IN assay, where the percent activity for the compounds was found to be 100% ± 75%
NA = Not Applicable, 50% inhibition by the compound was not achieved

3.1.3 Design of peptides 19-21 containing the Phe residue and *Ala- scanning approach (peptides 22-28)*

Base on the ability of inhibiting HIV-1 IN showed by peptide **18**, we synthesized three peptides, in which Val at position 1, 2, and 7 was substituted with Phe (peptides **19**, **20**, and **21**, respectively). Indeed, Phe and Val differ in participating helical assembly of the peptide molecules,

and have different hydrophobicity degree.[4,5] In addition, Phe could be involved in additional stacking interactions with the viral target. Subsequently, we established the contribution of the various amino acid residues to the antiviral activity of peptide **18** (peptides **22-28**) through an *L*-Ala scanning analysis.

3.1.4 Biochemical studies: HIV-1 IN catalytic activity inhibition (peptides 19-28)

The results of the biochemical evaluation of peptides **19-28** are summarized in Table 3.2 Replacement of the Val residue for the bulky-aromatic Phe led to a strong reduction of the activity of the corresponding peptides **19-21** compared with **18**, indicating the importance of this ramified amino acid on the HIV-1 IN inhibition capability. This result was also confirmed with the data from the Alanine scanning study. In fact, the substitution of Val^1 or Val^2 or Val^7 residues by Ala produced a dramatic loss of the HIV-1 IN inhibitory activity of the corresponding analogues (**22**, **23**, and **27**). The most interesting result was obtained with peptides **24** and **25**. In these cases, the substitution by Ala of Lys^4 or Glu^5, respectively, produced only a weak decrease of HIV-1 IN inhibitor ability if compared to the hit peptide, suggesting that a lack of polar side chains in these positions or the introduction of low hindrance, lipophilic features is well accepted.

Table 3.2 Sequence, analytical data and effect of peptides **19-28** on HIV-1 IN activities.

Peptide	Sequence	HPLC k'	ESI-MS	IN inhibition IC_{50} (μM)
18	VVAKEIVAH	3.50	1007	4.5±0.9
19	FVAKEIVAH	4.40	1054.4	NA
20	VFAKEIVAH	4.50	1054.5	375±19
21	VVAKEIFAH	4.20	1054.2	504±22
22	AVAKEIVAH	3.90	978.6	NA
23	VAAKEIVAH	4.20	978.4	NA
24	VVAAEIVAH	4.84	949.5	20.8±1.3
25	VVAKAIVAH	4.05	948.6	13.0±1.0
26	VVAKEAVAH	1.39	964.5	NA
27	VVAKEIAAH	3.90	978.5	NA
28	VVAKEIVAA	4.89	939.5	NA
Ral.				0.24±0.02

All peptide are acetylated and amidate at N-terminal and C-terminal respectively
k' [(peptide retention time – solvent retention time)/solvent retention time]. Peptide concentration μM
NA = Not Applicable, 50% inhibition by the compound was not achieved (as example see the dose-response curve of peptide 19 in SI)

3.1.5 Conformational Studies on peptides 22-28

Circular Dichroism was used to determine the conformation of peptide **18** and its analogues (**22-28**). Analysis of spectra acquired in water solution indicated that all peptides are random coil in water (data not shown). In contrast, spectra acquired in 50% TFE/water solution (TFE, trifluoroethanol) indicate that peptides have high tendency to fold as α helix in fluoroalcohol solution (*figure 3.2*). In particular, all peptides have similar high helical content (about 40%).

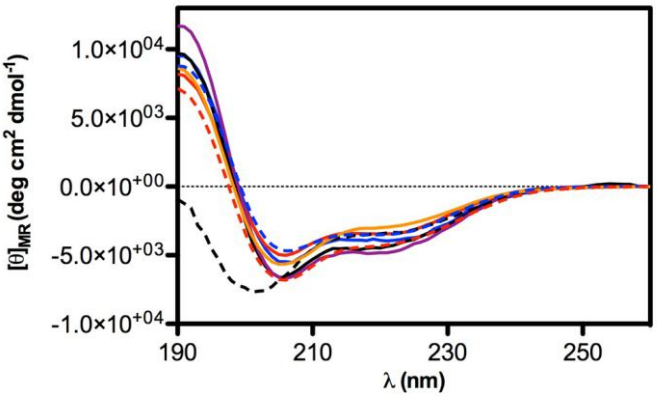


Figure 3.2 *CD Spectra of selected peptides in TFE/water 50% solution. Peptide are represented in:* ***18*** *(solid blue line),* ***22*** *(solid red line),* ***23*** *(solid black line),* ***24*** *(solid purple line),* ***25*** *(dashed black line),* ***26*** *(solid orange line),* ***27*** *(dashed blue line),* ***28*** *(dashed red line). Adapted from: Sala M, Spensiero A, Esposito F, Scala M C, Vernieri E, Bertamino A, Manfra M, Carotenuto A, Grieco P, Novellino E, Cadeddu M, Tramontano E, Schols D, Campiglia P, Gomez-Monterrey I M. Development and Identification of a Novel Anti-HIV-1 Peptide Derived by Modification of the N-Terminal Domain of HIV-1 Integrase. Frontiers in Microbiology, Virology.* ***2016****; 7, 845.*

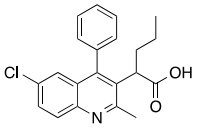
3.1.6 IN-IN dimerization and IN multimerization studies on peptide 18, 24 and 25

Taking in consideration that these peptides represent fragments of NTD domain, which is involved in different and essential process of the IN activity,[6,7] we wondered whether they were able to inhibit the IN-IN dimerization or stabilizing the IN in the multimeric form. In fact, it has been reported that allosteric compounds inhibit IN activity binding a site that is different from CCD binding site and working as disruptors of the IN dimer formation.[8] In collaboration with Prof. Enzo Tramonano at University of Cagliari, to verify whether active peptides **18**, **24** and **25** can modulate the dynamic interplay between IN subunits, we tested them using an HTRF-based IN subunit exchange assay. In this assay two IN preparation are used: a first His-tagged IN preparation, and a second

Flag-tagged IN preparation. In both preparations the purified IN is at equilibrium as a dimer; such dimers go through monomers exchange.

Hence, in the assay the monomer exchange of the two different preparations leads to the formation of a dimer with one His-tagged subunit and one Flag-tagged subunit. The formation of this "hybrid" dimer can be monitored in the HTRF-based assay as described.[9,10] When the assay is performed in the presence of a compound inhibiting dimerization, the HTRF signal decreases as the concentration of compound increases, whereas in the presence of a compound promoting IN multimerization (association of dimers) the HTRF signal increases as the concentration of the compound increases. We found that peptides **24** and **25** modulate the interactions between the IN subunits in different manners using the allosteric inhibitor LEDGIN-6 as a control (*Table 3.3*). In fact, peptide **24** was able to inhibit the IN-IN dimer formation with an IC_{50} value of 32 μM, but did not affect IN multimerization, while peptide **25** promoted IN multimerization, with a concentration of compound stimulating the Multimerization Increase by 50% (MI50 value) of 4.8 μM, but did not inhibited IN multimerization. While the reason for this different behavior was not clear, it demonstrates that IN multimerization can be differently modulated by interfering with the NTD sequence. Differently, peptide **18** was not able to modulate the dimer/multimer formation. Hence, the three peptides showed three different modes of IN inhibition in biochemical assays.

Table 3.3 Effects of peptides **18**, **24**, and **25** on the HIV-1 IN dimerization and multimerization process.

Compound	Sequence	IN-IN subunit exchange [a]IC_{50} (μM)	IN-multim. [c]MI_{50} (μM)
18	[31]VVAKEIVAH	>100 (100%)[b]	>100 (100%)
24	VVAAEIVAH	32 ± 3	>100 (100%)
25	VVAKAIVAH	>100 (100%)	4.8 ± 0.4
LEDGIN-6		>100 (100%)	10 ± 1

[a] Compound concentration required to inhibit the HIV-1 IN-IN subunit exchange by 50%.
[b] Percentage of control measured in the presence of 100 μM concentration.
[c] Compound concentration required to inhibit the multimerization increase by 50%.

3.1.7 Cell-based HIV replication assay on the most active peptides 18, 24 and 25

In collaboration with Prof. Dominique Schols at University of Leuven, the most potent peptides **18**, **24** and **25**, conveniently conjugated with the cell-penetrating fragment TAT, were evaluated for their capability to inhibit the HIV-1 replication in $CD4^+$ MT-4 cells (*Table 3.4*). The conjugated peptides **18** and **24** were devoid of antiviral activity when evaluated against HIV replication. In contrast, **25** showed marked inhibitory potency with good selectivity against HIV-1 (IC_{50} value in the low micromolar range, accordingly with the results on IN catalysis and dimerization, and selective index of 5.4), but not against HIV-2.

Table 3.4. Evaluation of the synthesized peptides against HIV-1 and HIV-2 replication in MT-4 cell cultures.

Peptide	HPLC k'	ESI-MS	EC_{50} [a] (μM)		CC_{50} [b] (μM)
			HIV-1	HIV-2	MT-4
TAT-18	4.31	2568.05	> 105	> 105	105
TAT-24	3.56	2510.96	> 50	> 50	> 50
TAT-25	4.50	2510.02	4.0 ±	> 21.4	21.4 ± 2.0
AMD3100			10.73±	10.56±	> 1000

[a] 50% Effective concentration, or compound concentration required to inhibit HIV-induced cytopathogenic effect in MT-4 cell culture.
[b] Compound concentration required to reduce by 50% MT-4 cell viability.
*GRKKRRQRRRPQ: TAT

3.1.8 Structure-activity relationship studies on peptide 25

During my PhD project, the structure-activity relationship of the peptide **25** was examined, focusing on the following points:

Evaluation of the role played by each aminoacid residue configuration for interaction of the peptide with IN (D-scan analysis). It's well known peptides may adopt secondary structures, which are responsible for their receptor affinity and biological activity. Although specific synthetic modifications of particular residues may reveal information on the receptor-bound conformer, combinatorial scanning strategies, in which each residue in a sequence is systematically modified, offer enhanced means for identifying the location of key residues and the global structural requirements for peptide ligand binding and activity.[11]

Information on the importance of native residues and their systematic modifications are respectively obtained from observed losses and improvements of activity on evaluation of the analog libraries from peptide scanning. In particular, the importance of amino acid configuration on biological activity can be determined by systematic substitution of each residue with its enantiomer (typically wild-type L- for D-amino acid). Consequentely, to have insights about the importance side-chain chirality of aminoacid residue, as well as information about certain conformations that might be important for peptide–integrase interactions, a D-scan analysis of peptide **25** was performed. We synthesized a new library of peptides (**29-37,** *Table 3.5*) containing the corresponding D-aminoacid and evaluated for their capability to inhibit IN activity.

Study of the importance of amino acid sequence. In order to verify if the amino acids sequence, rather than the amino acid composition, is important for activity, we synthesized the scramble peptide (peptide **38,** *Table 3.5*).

In addition, we synthesized peptide **39** containing the residues 25-39 of the 3rd alpha helix of N-teminal domain of integrase. This fragment, known in literature for its activity of viral replication inhibition, includes the sequence of peptide 25.[12]

So peptide **39** could be useful to a comparison study.

Identification of minimum active sequence through deletion at N- and C-terminal. Truncation library can be used to identify the shortest aminoacid sequence needed for the peptide activity. The truncation process is carried out via a systemic reduction of residues from each flank of the original peptide. For this reason, we synthesized

peptides **40-43** (*Table 3.5*) obtained from deletion of two amino acids at the N-terminal and C-terminal of peptide **25**.

3.1.9 Biochemical studies: HIV-1 IN catalytic activity inhibition and IN-IN dimerization and IN multimerization studies on peptides 29-39

The synthesized peptides (**29-39**) were tested on HTRF assay to evaluate the catalytic activity in presence of LEDGF/p75 protein. In this assay, Raltegravir was used as a control. Furthermore, to verify whether peptides can modulate the dynamic interplay between IN subunits, we tested them using an HTRF-based IN subunit exchange assay. As shown in Table 3.5 all peptides are not able to modulate the dimer/multimer formation and neither of peptides inhibits catalytic activity of integrase as peptide **25**. The results obtained, confirmed peptide **25** as a hit for further development of new chemotherapeutic agents against HIV-1.[13] Studies to evaluate the antiviral activity of trunks peptides (**40-43**) are still in progress.

Table 3.5 Effects of peptides **1, 4,** and **5** on the HIV-1 IN dimerization and multimerization process, and effects on the HIV-1 IN LEDGF-dependent and –independent catalytic activity.

Peptide	Sequence	HPLC k'	ESI-MS	IC_{50} IN inhibition (μM)	[a] IC_{50} IN-IN (μM)	[b] MI_{50} (μM)
25	VVAKAIVAH	3.50	948.6	13.0 ± 1.0	>100	4.8 ± 0.4
29	vVAKAIVAH	4.40	948.6	>100 (100%)	ND	ND
30	VvAKAIVAH	4.50	948.6	>100 (52%)	ND	ND
31	VVaKAIVAH	4.20	948.6	>100 (82%)	ND	ND
32	VVAkAIVAH	3.90	948.6	>100 (59%)	ND	ND
33	VVAKaIVAH	4.20	948.6	>100 (100%)	ND	ND
34	VVAKAiVAH	4.84	948.6	>100 (59%)	ND	ND
35	VVAKAIvAH	4.05	948.6	94 ± 4.0	ND	ND
36	VVAKAIVaH	1.39	948.6	>100 (55%)	ND	ND
37	VVAKAIVAh	3.90	948.6	>100 (61%)	ND	ND
38	AKVAVAHIV	4.89	948.6	89.5 ± 0.5	ND	ND
39	DFNLPPVVAKAIVAH	5.67	1630.90	10.5 ± 1.5	66.5 ±	>100
40	VVAKAIV	3.45	739.94	-	-	-
41	VVAKA	3.12	527.64	-	-	-
42	AKAIVAH	3.56	749.89	-	-	-
43	AIVAH	3.76	550.64	-	-	-
LEDGIN-6	-			9 ± 2	>100	10 ± 1
RAL	-			0.058± 0.002	>100	>100

[a] Compound concentration required to inhibit the HIV-1 IN-IN subunit exchange by 50%.
[b] Compound concentration required to inhibit the multimerization increase by 50%.
All peptide are acetylated and amidate at C terminal and N terminal respectively
k' [(peptide retention time – solvent retention time)/solvent retention time]. Peptide concentration μM

3.1.10 NMR analysis on peptides 25 and 39

In collaboration with Dr. M. Angeles Jimenez at University of Madrid, we evaluated the conformation of peptide **25** and peptide **39** by NMR analysis. We found that two peptides in aqueous solution are within the random coil range ($|\Delta\delta C\alpha| \leq 0.5$ ppm, and $|\Delta\delta C\beta| \leq 0.5$ ppm), and the $\Delta\delta H\alpha$ values are only slightly outside the random coil range ($|\Delta\delta H\alpha| \leq$ 0.05 ppm). While, in the presence of TFE the magnitudes of $\Delta\delta H\alpha$, $\Delta\delta C\alpha$ and $\Delta\delta C\beta$ increase significantly, being $\Delta\delta H\alpha$ and $\Delta\delta C\beta$ negative for peptide **39** and for residues 6-15 of peptide **39**, and the corresponding $\Delta\delta C\alpha$ positive.

This indicates that the two peptides form a helix structure in the presence of TFE. The populations of helix estimated on the basis of the averaged ΔδHα value observed in 30 % TFE are 77 % for peptide **25** and 56 % for peptide **39**. To visualise the structural features of the helices formed by peptides **25** and **39** in the presence of TFE, we performed structure calculations. We found that short peptide **25** shows an α-helix extending throughout all its residues, which is very well defined. Peptide **39** also exhibits a well-defined structure, but it displays a non-regular structure in segment 1-5, and an α-helix spanning residues 6-15 (*figures 3.3).*

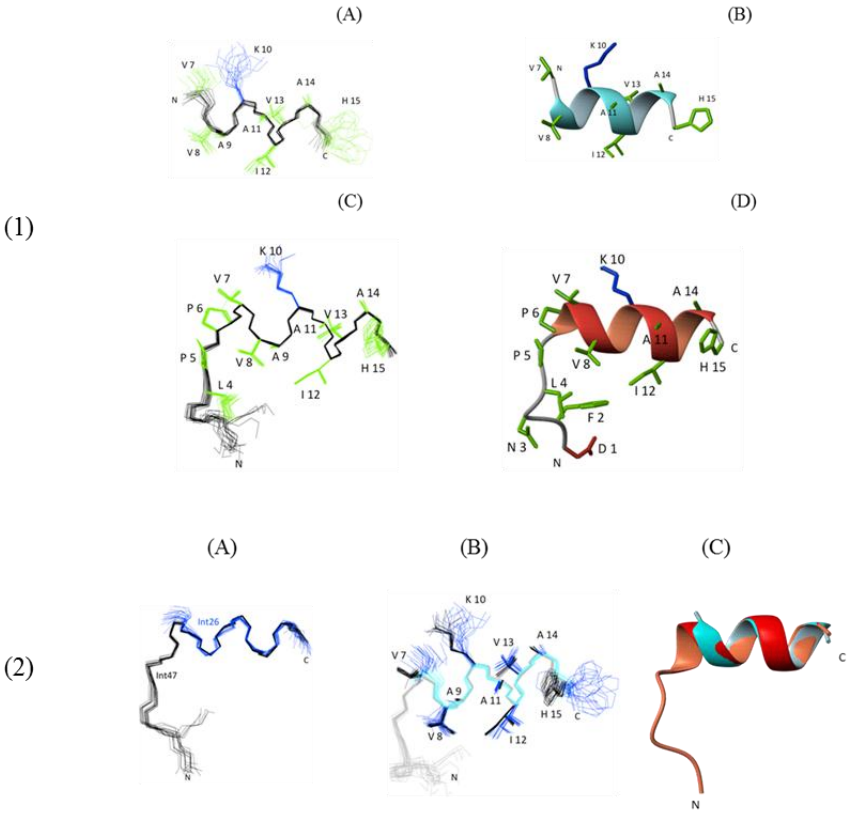


Figure 3.3 (1) NMR structures of peptides 25(A and B) and 39 (C and D) in 30 % TFE. (2) Overlay of the NMR structure of peptide 25 onto the NMR structure of peptide 39. (A) Backbone atoms for 25 are in blue and for 39 in black. (B) Backbone atoms for 25 are in cyan and side chains in blue, and for

39 backbone in grey and side chains in black. (C) Ribbon representation of the overlay of representative conformers of 25 (in cyan) and 39 (in red).

3.2 Improvement of pharmacokinetic properties on peptide 25

To improve pharmacokinetic properties of peptide **25,** so in order to identify the most active and at the same time the most stable peptide, we design cyclic peptides, using two different approaches. It's well known that cyclic peptides are more resistant to proteolysis than their linear counterparts due to the lack of exopeptidase cleavage sites.

3.2.1The head-to-tail cyclization approach

Head-to-tail peptide cyclization is well-know strategy to constrain the high flexibility of linear peptides inducing or stabilizing the bioactive conformation of peptides and allows us to have a peptide with minor possible conformers.[14] On the basis of these observations, we synthesized peptide **44** (*figure 3.4*). Studies to evaluate the antiviral activity of the cyclic peptide **44** are still in progress.

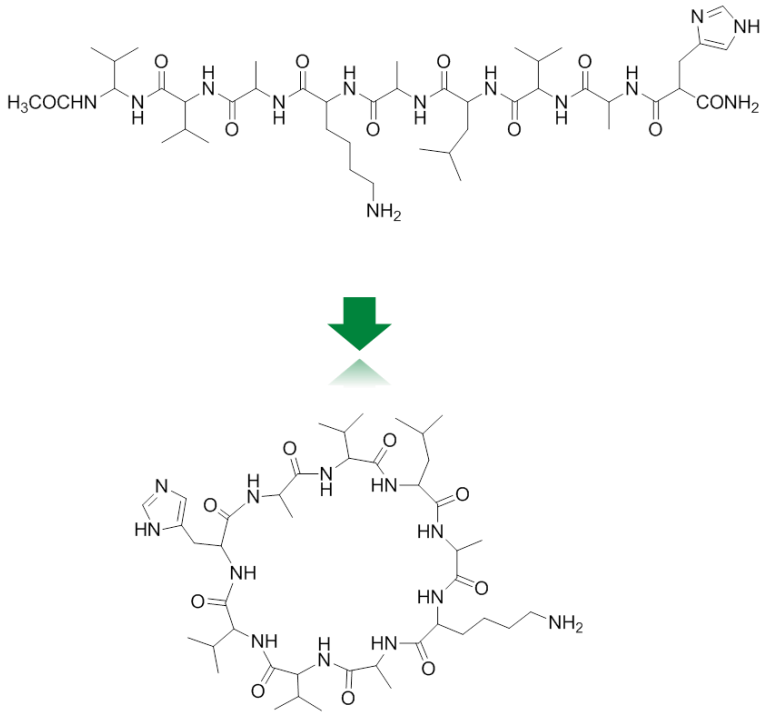


Figure 3.4 *The cyclic peptide **44**.*

*3.2.2 Insertion of a lactam bridge on peptide **25** (peptides **45-46**)*

An approach for stabilizing peptides is the incorporation of covalent or not covalent linkages between constituent aminoacid side chains. In this case we designed and synthesized analogues of peptide **25** characterized by the introduction of a lactam bridge resulting in peptides **45-46** (*figure 3.5, Table 3.6).*

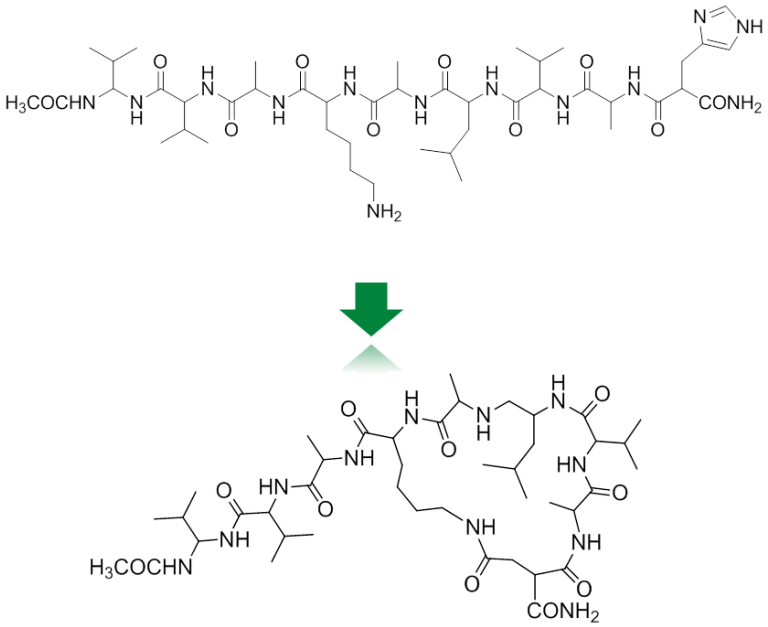


Figure 3.5 *Cyclic peptide **46** with the lactam bridge.*

In particular, we replaced the aminoacid residue Histidine at positions 9 with carboxylic (Asp or Glu) functions on the side chain.

Therefore the lactam bridge was obtained by the bond between the side chains of Asp or Glu with Lys in position 4. Studies of cyclic peptides are still in progress.

Table 3.6 *Cyclic peptides with lactam bridge*

Peptide	Sequence	HPLC k'	ESI-MS
45	VVA[KAIVAE]	4.35	880.09
46	VVA[KAIVAD]	3.46	866.06

3.3 Chemistry

3.3.1 General procedure for synthesis

The synthesis of peptides (**1-46**) was performed according to the solid phase strategy, using standard Fmoc methodology in a manual reaction vessel.[15] The first amino acid, was linked to the Rink resin previously deprotected by a 25% piperidine solution in N,N-dimethylformamide. The following protected amino acids were then added stepwise. Each coupling reaction was accomplished using a 3-fold excess of amino acid with HBTU and HOBt in the presence of DIPEA. The Nα-Fmoc protecting group was removed by treating the protected peptide resin with a 25% solution of piperidine in DMF. Kaiser test was performed step to step to confirm the complete removal of the Fmoc protecting group and the complete coupling reaction. The N-terminal Fmoc group was removed as described above and the peptides were acetylated adding a solution of Ac_2O/DCM. Finally, the peptides were released from the resin with trifluoroacetic acid (TFA/TES/H_2O) for 3 h. The resin was removed by filtration, and the crude peptides were recovered by precipitation with cold anhydrous ethyl ether to give a white powder and then lyophilized (*figure 3.6*).

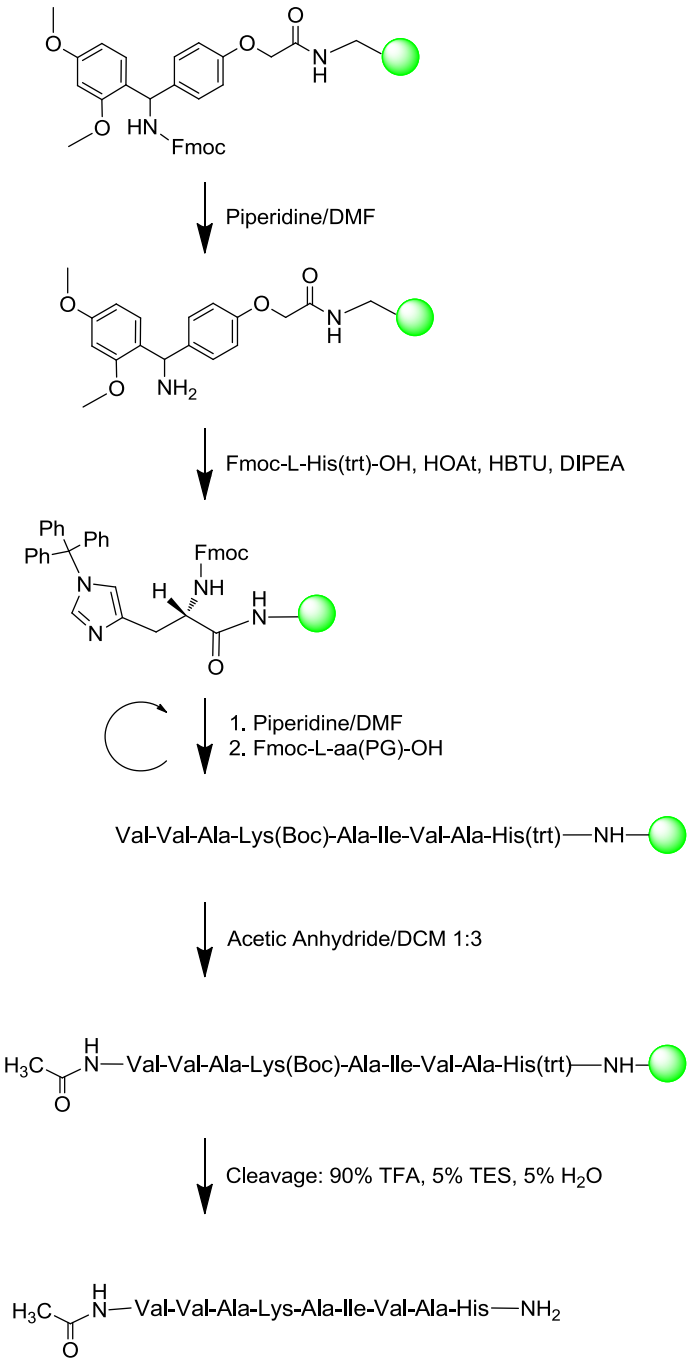


Figure 3.6 *Schematic representation of the peptide synthesis.*

3.3.2 Synthesis of head-to-tail cyclic peptide 44: side reaction during the sequence assembly

During the past years, great effort has been made to develop more efficient methods for the synthesis of cyclic peptides. The use of solid phase assembly on an acid labile linkers have allowed detachment of peptide from the polymer without sacrificing side-chain protection. In this context the 2-chlorotrityl resin and the 4-hydroxymethyl-3-methoxyphenoxybutyric acid linker are most used as solid support. An alternative method involves the use of the oxime resin.

Initially, we chosen the 2-chlorotrityl chloride resin for the synthesis of head to tail cyclic peptide, peptide **44**. The first N-α-Fmoc amino acid Fmoc-His(Boc)-OH and DIPEA were dissolved in dry dichloromethane containing, if necessary, a small amount of dry DMF (enough to facilitate dissolution of the acid). This latter, was added to the resin and stirred for 2-4 h. Then the loading test was carried out. Other Nα -Fmoc amino acids were sequentially coupled as previously described. The final cleavage with TFA 1% in DCM resulted in protected peptides.

A solution of the linear protected peptide in DMF was added at room temperature to a reaction flask containing a solution of N-hydroxybenzotriazole (HOBt), HBTU and DIPEA in DMF. The mixture was stirred for 24 h at room temperature and monitored by TLC.

However this strategy was failed. [16] The stretching of the sequence was monitored by test cleavage an LC-MS and we observed, during the second aminoacid deprotection, a premature cleavage of the C-terminal acid peptides. This could be interpreted as diketopiperazine formation and simultaneous cleavage from the Resin (*figure 3.7*).

In peptide synthesis the formation of diketopiperazine is a notorious side reaction at the dipeptide Fmoc-deprotection step in SPPS. (*figure 3.8*).[17]

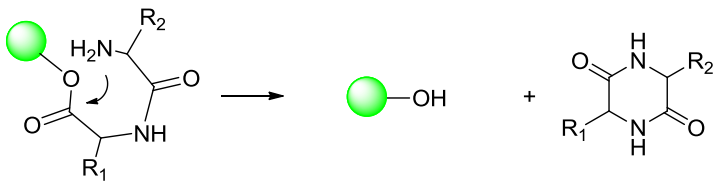


Figure 3.7 *Schematic representation of diketopiperazine formation.*

As described in literature, we hypothesized that during the base-induced deprotection of the second amino acid, the free amino function attacks the carbonyl group leading to diketopiperazine formation (nucleophile attack), followed by the resin cleavage.

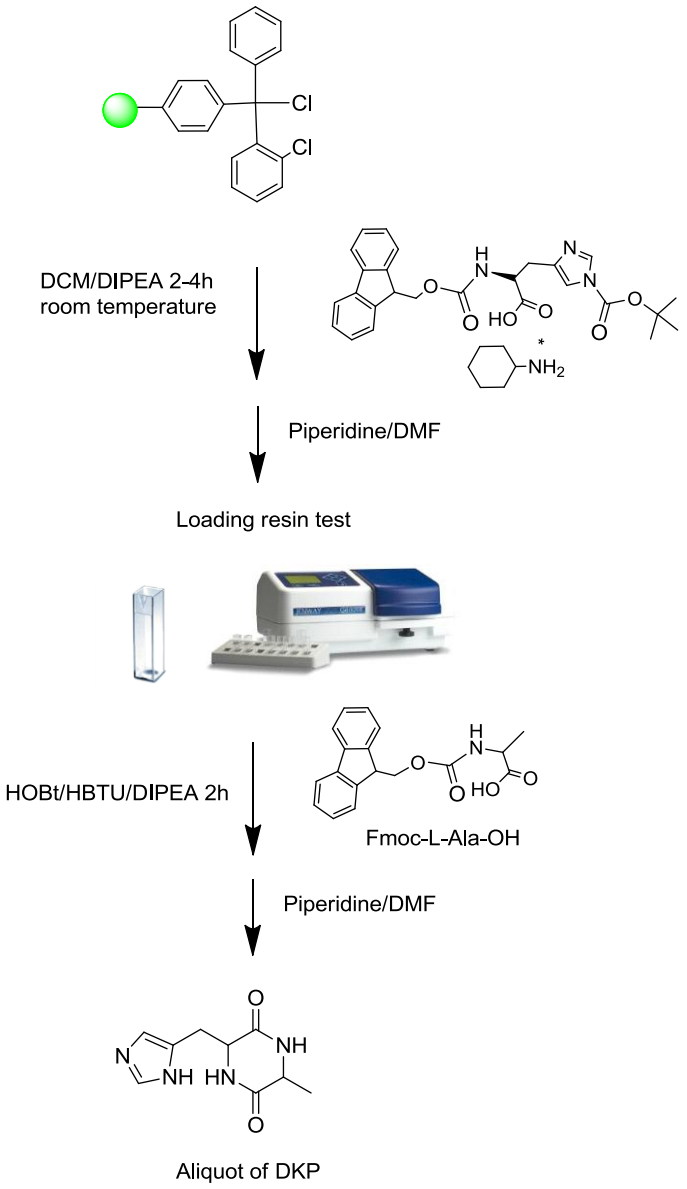


Figure 3.8 *Schematic representation of the cyclic peptide synthesis.*

To maintain the orthogonal strategy, and try to obtain the desired sequence in a good yield another approach was evaluated. So we

performed the peptide **44** synthesis using a base labile and orthogonally compatible linker, HMBA (4-Hydroxymethylbenzoic acid).

The cyclic peptide was synthesized on a TentaGel-S-NH_2 resin. Coupling reaction of HMBA Linker was performed using coupling reagents HBTU, HOBt and DIEA in DMF at 60 ºC. The first amino acid, Nα-Fmoc amino acid, was attached to the HMBA linker using HOBt/HBTU as activating agents and a catalytic amount of DMAP. In addition, after each step of deprotection and after each coupling step, Kaiser test was performed. N-terminal Fmoc group was removed as described above. Then, the peptide was released from the resin with NaOH/Dioxane for 15 min. To neutralize NaOH we used a 0.1 M HCl solution. The resin was removed by filtration, and the crude peptide was recovered by lyophilization.

In these conditions, we observed the formation of two peptides: the unprotected peptide and the desired one with a no detectable yield, indicating the detachment of Boc protecting groups present on the side chains, during the neutralization step (*figure 3.9*).

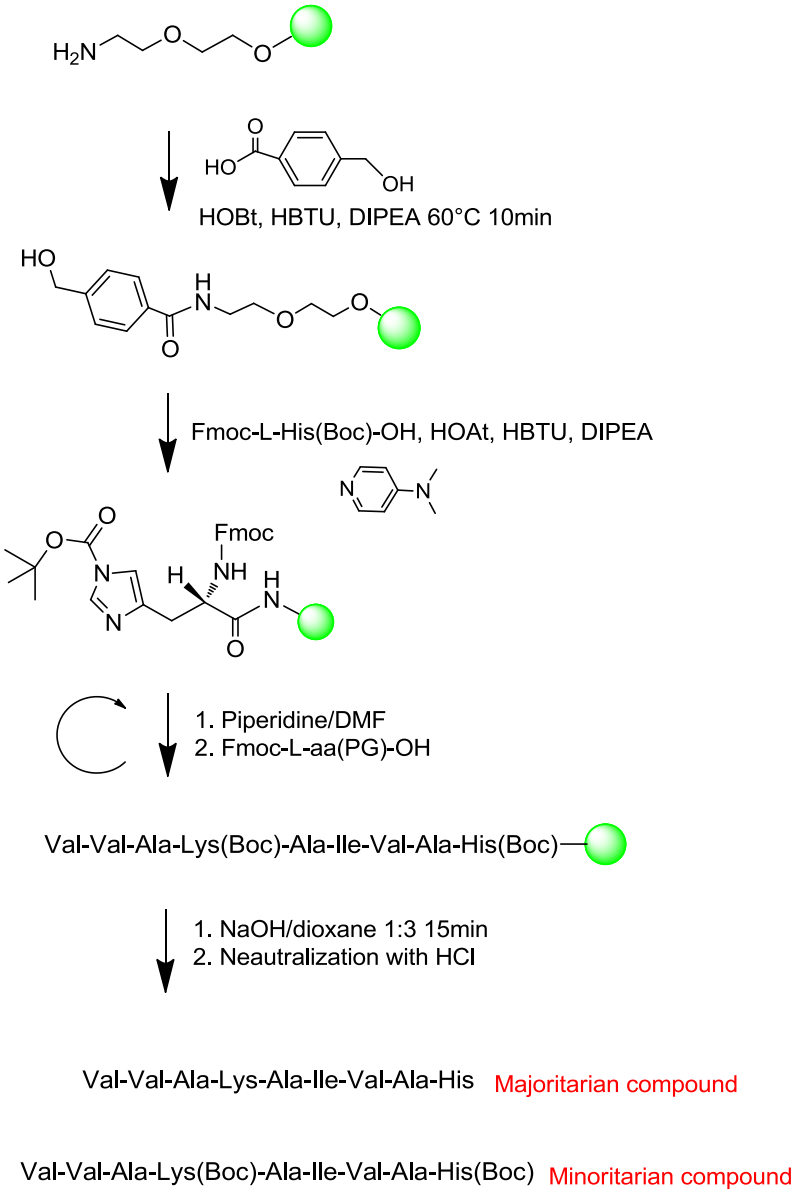


Figure 3.9 *Schematic representation of the cyclic peptide synthesis.*

Finally, Fmoc-His(Allyl)-Trityl resin was used to avoid the collateral reactions above discussed. The following protected amino acids were then added stepwise. Each coupling reaction was accomplished using a HBTU and HOBt as coupling reagents in the presence of DIEA. The Nα-Fmoc amino acids protecting groups were removed by treating the protected peptide resin with a 25% solution of piperidine in DMF.

The Allyl groups was removed according to strategy reported by Grieco et al.[18] The last Nα -Fmoc was removed and the macrocyclic lactam ring formation was mediated by addition of HBTU, HOBt and DIPEA for 2 h. The process was repeated if necessary. Subsequently, the peptide was released from the resin with TFA/H_2O (50:50) for 3 h. The resin was removed by filtration, and the crude peptide was recovered by precipitation with cold anhydrous ethyl ether to give a white powder and then lyophilized (*figure 3.10*).

Using this approach we obtained the desired cyclic peptide with a yield of 60%.

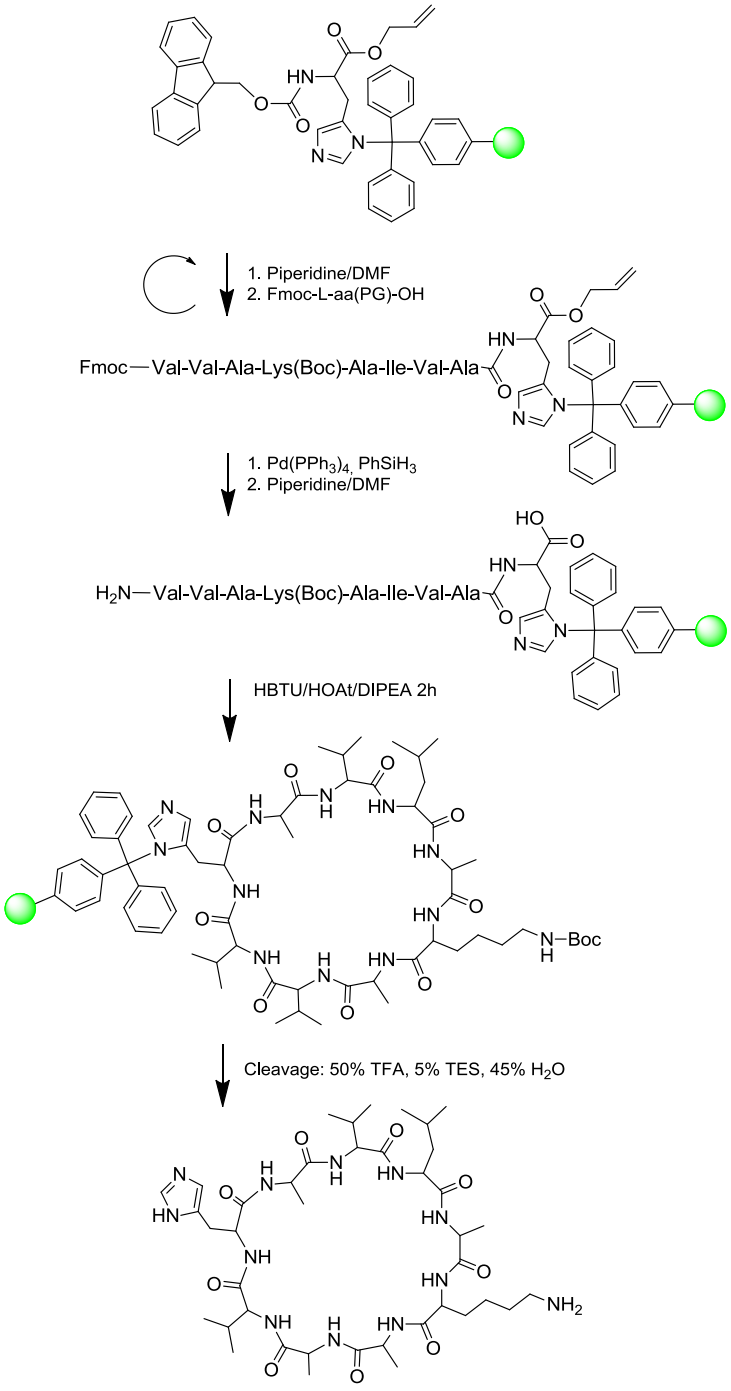


Figure 3.10 *Schematic representation of the cyclic peptide synthesis.*

3.3.3 Synthesis of lactam analogues (peptides 45-46)

The corresponding linear peptides were synthesized as described in the paragraph 3.1. The amino acids Nα-Fmoc-Asp(Allyl)-OH, Nα-Fmoc-Glu(Allyl)-OH and Nα-Fmoc-Lys(Alloc)-OH were used as lactam precursors. After linear assembly, the Nγ-Alloc and the Allyl groups were removed according to strategy reported by Grieco et al.[18] The macrocyclic lactam ring formation was mediated by addition of HBTU, HOBt and DIPEA for 2 h. The process was repeated if necessary (Kaiser test was used to monitor completion). The N-terminal Fmoc group was removed and the peptide was released from the resin as described above (*figure 3.11*).

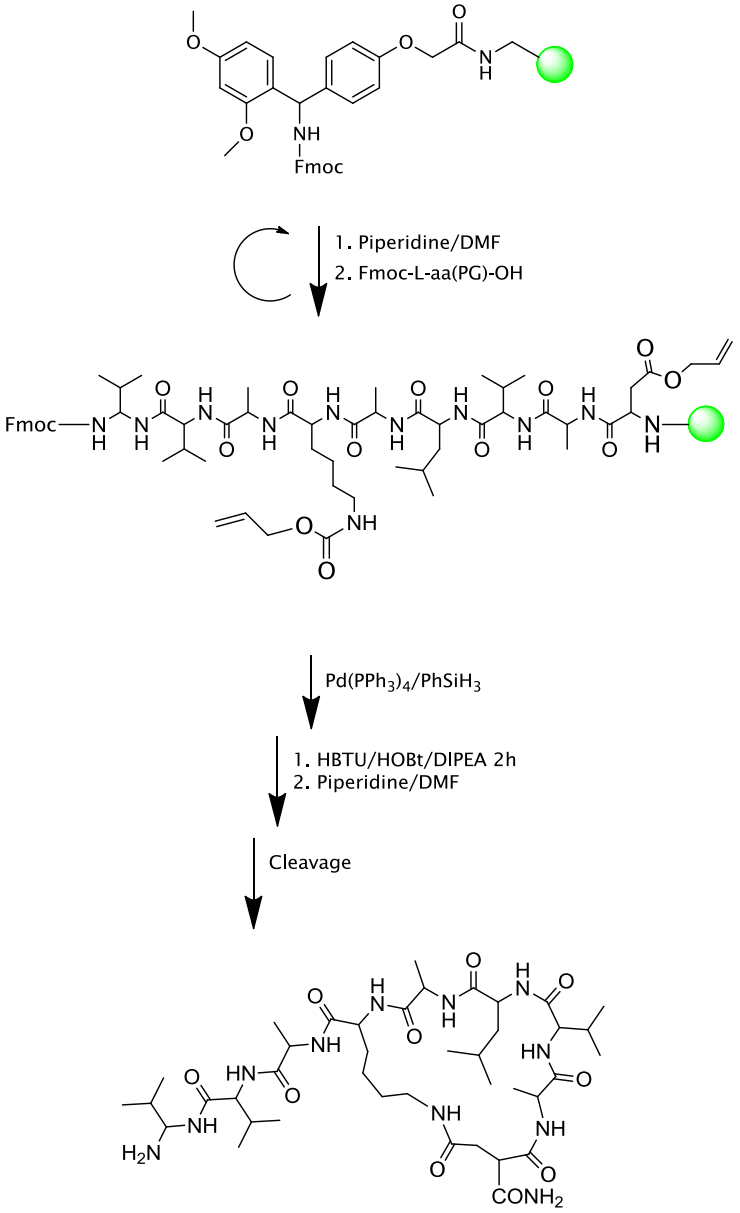


Figure 3.11 *Schematic representation of the lactam cyclic peptide synthesis.*

All crude peptides were purified by RP-HPLC on a preparative C18-bonded silica column using a Shimadzu SPD 10A UV/VIS detector, with detection at 215 and 254 nm. All peptides were characterized with analytical RP-HPLC and peptides molecular weights were determined by ESI mass spectrometry in a MS Thermo Quest Finnigan LTQ Advantage. All analogues showed >97% purity when monitored at 215 nm.

3.4 Discussion

A library of peptides (**1-15**) was developed from the NTD sequence (1-50) by using a mimotopic strategy. Preliminary results of the *in vitro* integrase assay indicated that the fragment VVAKEIVAH (peptide **18**), which includes a His residue instead of the natural Ser at position 39, reduced significantly the HIV-1 IN activity showing an IC_{50} value of 4.5 μM. Peptide **18** is included into the sequence of the 3rd α-helix of the NTD, which is located on the dimerization interface within the context of the crystal structure of this domain.[6] Accordingly, we observed significant changes in potency of this peptide when lipophilic and ramified residues involved in the interaction between subunits, such as Val^1, Val^2, Ile^6 and Val^7 are substituted. Ala scanning showed also that loss of a lysine positively charge or a glutamate negatively charge does not modify substantially the inhibitory activity of the resulting peptides (**24** and **25**, respectively). According to the circular dichroism results, the inhibitory activity of these compounds does not seem to be related to their ability to form helical structures. In fact, inactive peptides **22**, **23**, **26**, **27** and **28** all show similar helical content as **18** and **24**. On the other hand, compound **25** shows a significant reduction of helical tendency but

is approximately as active as **18**. Recently, an approach to develop new HIV-1 IN inhibitors that bind in a pocket different from the active binding site, lead to the discovery of new drugs that are able to interfere with the IN dynamic interplay, in two possible ways. In the first one, the molecules interfere with the IN dimer-dimer formation, while the second, the molecules promote the IN multimerization. In this case, IN is trapped in a multimeric form in which the movements of the individual subunits are restricted, and, as a consequence, also the catalytic process. We tested peptides **18**, **24**, and **25** on HTRF assays to determine their potential mechanism of actions and to verify whether they were capable to inhibit the HIV-1 IN. Peptide **24** was able to inhibit the IN-IN dimer formation, while peptide **25** stabilizes the multimeric form of IN. Differently, peptide **18** was not able to modulate the dimer/multimer formation. These data suggest that the modulation of IN dimerization and/or multimerization process could be determined by the net charge of the synthesized peptides, or more specifically by their ability to form intermolecular salt bridges between Glu^5 (peptide **24**) or Lys^4 (peptide **25**) residues and Lys or Asp/Glu residues of different subunits, respectively.[19,20] This hypothesis, however, is less probable for peptide **18**, in which the formation of an intramolecular salt bridge between adjacent residues Lys^4 and Glu^5 prevents their interactions with acid and basic residues of other subunits. In this case, the *in vitro* inhibitory activity of peptide **18** could be related to its ability of inhibiting IN-DNA interactions.

Interestingly, peptide **25**, conjugated with the cell-penetrating Tat fragment, is the only compound into this series able to inhibit the HIV-1 induced cytopathogenic effect in MT-4 cell. Our results indicate that the presence of a negative charge (Glu^5 side chain) is detrimental for the

inhibition of HIV replication in target CD4 T cells (analogues **18** and **24**), while the positively charged Lys^4 residue (peptide **25**) plays an important role in the potential antiviral activity. In addition, this peptide has shown to have potent and selective antiviral activity for HIV-1. Although the antiviral activity is about 5-fold lower than the concentration of its cellular toxicity, peptide **25** represents a suitable lead compound for the development of novel derivatives with improved toxicity profile.

Overall, our data confirm the hypothesis that peptides derived from the protein of interest may be used to identify domains which are important for catalytic activity, protein−protein interactions, and/or multimerization process.[21,22] Data obtained reflect also the difficulties in identifying inhibitor compounds whose biological target is the viral integration process.[23]

In fact, the modification or replacement of a single amino acid in the sequence of the most potent compound in this series (peptide **25)**, determines the diminution or loss of the target activity. The identification of the amino acid residues implicated in the inhibition of IN activity highlights the importance of studies that identify such residues and investigates possible roles for them.

In conclusion, this study shows that the modified 31-39 sequence of the N-terminal domain of HIV integrase (VVAKAIVAH, peptide **25**) is able to interfere with functional integration of the HIV-1 enzyme. This peptide conveniently conjugated (peptide **TAT-25**) inhibits the HIV activity in cell culture probably due to its ability to stabilize the IN multimeric form.

Moreover, we confirmed peptide **25** the most powerful of the synthesized series, employing of enantiomeric aminoacids, to evaluate

factors such as configuration, conformation and hydrogen-bonding for activity. Furthermore, we used NMR approach to investigate the conformation of the peptide. The results obtained showed a strong alpha helix conformation of the peptide **25** which could be essential for interaction with IN. This information will be used in future studies to design peptide **25** analogs acting as selective IN inhibitors.

Indeed, our peptide, one of the few in cell active peptides reported to date, could be considered as a suitable hit compound in the development of novel anti-HIV1 peptidomimetic agents with a mechanism of action different to that the approved anti-IN drugs.[24]

3.5 Experimental section

3.5.1 Material and methods

N^{α}-Fmoc-protected amino acids, Rink amide-resin, HOAt, HOBt, HBTU, DIEA, piperidine and trifluoroacetic acid were purchased from Iris Biotech (Germany). Peptide synthesis solvents, reagents, as well as CH_3CN for HPLC were reagent grade and were acquired from commercial sources and used without further purification unless otherwise noted.

3.5.1.1 Peptide synthesis

The synthesis of IN analogues was performed according to the solid phase approach using standard Fmoc methodology in a manual reaction vessel and automated microwave synthesizer.[25,26]

The first amino acid, N^{α}-Fmoc-Xaa-OH (N^{α}-Fmoc-Asp(OtBu)-OH, N^{α}-Fmoc-Glu(OtBu)-OH, N^{α}-Fmoc-His($N_{(im)}$trityl(Trt))-OH, N^{α}-Fmoc-Trp(Boc)-OH, N^{α}-Fmoc-Ala-OH, N^{α}-Fmoc-Leu-OH, N^{α}-Fmoc-Val-OH, N^{α}-Fmoc-Lys(Boc)-OH, N^{α}-Fmoc-Cys(Trt)-OH, N^{α}-Fmoc-Met-OH, N^{α}-Fmoc-Ser(tBu)-OH, was linked on to the Rink resin (100-200 mesh, 1% DVB, 0.59 mmol/g) previously deprotected by a 25% piperidine solution in DMF for 30 min. The following protected amino acids were then added stepwise. Each coupling reaction was accomplished using a 3-fold excess of amino acid with HBTU (3 eq.) and HOBt (3 eq.) in the presence of DIEA (6 eq.). The N^{α}-Fmoc protecting groups were removed by treating the protected peptide resin with a 25% solution of piperidine in DMF (1×5 min and 1x 25 min). The peptide resin was washed three times with DMF and the next coupling step was initiated in a stepwise

manner. The peptide resin was washed with DCM (3×), DMF (3×), and DCM (3×), and the deprotection protocol was repeated after each

coupling step. In addition, after each step of deprotection and after each coupling step, Kaiser test was performed to confirm the complete removal of the Fmoc protecting group, respectively, and to verify that complete coupling has occurred on all the free amines on the resin. The N-terminal Fmoc group was removed as described above, and the peptides were acetylated adding a solution of Ac_2O/DCM (1:3) shaking for 30 min. Finally the peptides were released from the resin with TFA/iPr_3SiH/H_2O (90:5:5) for 3 h. The resin was removed by filtration, and the crude peptide was recovered by precipitation with cold anhydrous ethyl ether to give a white powder and then lyophilized.

3.5.2 Synthesis peptide **44**

Peptide was synthesized on a Fmoc-His(Allyl)-Trityl resin (0.7 mmol/g), previously deprotected by a 25% piperidine solution in DMF for 30 min. The following protected amino acids were then added stepwise. Each coupling reaction was accomplished using as coupling reagent HBTU (3.9 eq, 0.5M), HOBt (4.0 eq, 0.5M) and DIEA (7.2 eq, 0,5M) in DMF. The Nα-Fmoc protecting groups were removed by treating the protected peptide resin with a 25% solution of piperidine in DMF (1 × 5min and 1 × 25min).

The Allyl group was removed according to the following procedure: 200 mg of peptide resin was washed with dichloromethane (DCM) under nitrogen and a solution of PhSiH3 (24 equiv) in 2 mL of DCM was added. Subsequently, a solution of Pd(PPh3)4 (0.25 equiv) in 6 mL of DCM was added and the reaction was allowed to proceed under nitrogen for 30 min. The peptide resin was washed with DCM (3x), DMF (3x) and DCM (4x), and the deprotection protocol was repeated (3x). The last Nα -Fmoc was removed and the macrocyclic lactam ring formation was mediated by addition of HBTU (6 equiv), HOBt (6 equiv) and DIPEA (12 equiv) for 2 h. The process was repeated if necessary (Kaiser test used to monitor completion). Subsequently the peptide was released from the resin with TFA/H2O (50:50) for 3 h. The resin was removed by filtration, and the crude peptide was recovered by precipitation with cold anhydrous ethyl ether to give a white powder and then lyophilized.

3.5.3 Synthesis of lactam analogues (peptides 45-46)

The corresponding linear peptides were synthesized as described in the paragraph 5.1.1. The amino acids Nα-Fmoc-Asp(Allyl)-OH, Nα-Fmoc-Glu(Allyl)-OH and Nα-Fmoc-Lys(Alloc)-OH were used as lactam precursors. After linear assembly, the Nγ-Alloc and the Allyl groups were removed according to the following procedure: 200 mg of peptide resin was washed with dichloromethane (DCM) under nitrogen and a solution of PhSiH3 (24 equiv) in 2 mL of DCM was added. Subsequently, a solution of Pd(PPh3)4 (0.25 equiv) in 6 mL of DCM was added and the reaction was allowed to proceed under nitrogen for 30 min. The peptide resin was washed with DCM (3x), DMF (3x) and DCM (4x), and the deprotection protocol was repeated (3x). The macrocyclic lactam ring formation was mediated by addition of HBTU (6 equiv), HOBt (6 equiv) and DIPEA (12 equiv) for 2 h. The process was repeated if necessary (Kaiser test used to monitor completion). The N-terminal Fmoc group was removed and the peptide was released from the resin as described above.

3.5.4 Purification and characterization

All crude peptides (**1-46**) were purified by RP-HPLC on a preparative C18-bonded silica column (Kinetex 5μ, 100Å, 100×21.2mm) using a Shimadzu SPD 10A UV−Vis detector, with detection at 210 nm and 254 nm. The column was perfused at a flow rate of 20 mL/min with solvent A (10%, v/v, water in 0.1% aqueous TFA), and a linear gradient from 5% to 90% of solvent B (80%, v/v, acetonitrile in 0.1% aqueous TFA) over 20 min was adopted for peptide elution. Analytical purity, was determined by RP-HPLC, using an autosampler, with dual wavelength absorbance detector with wavelength set to 215 nm, all of which were

obtained from Shimadzu. The mobile phases employed were: solvent A (water + 0.1% TFA) and solvent B (acetonitrile + 0.1% TFA); using a linear gradient from 5% to 90% B over 11 min, fitted with C-18 column Phenomenex, Ascentis ES-C18 column (5cm x 3.0mm, 2.7µm), at a flow rate of 0.800 mL/min (*figure 3.13*). ESI mass spectrometry was determined using a Finnigan LCQ Deca ion trap instrument, manufactured by Thermo Finnigan (San Jose, CA, USA), equipped with the Excalibur software for processing the data acquired. The sample was dissolved in a mixture of water and methanol (50/50) and injected directly into the electrospray source, using a syringe pump, which maintains constant flow at 5 µL/min. The temperature of the capillary was set at 220 °C

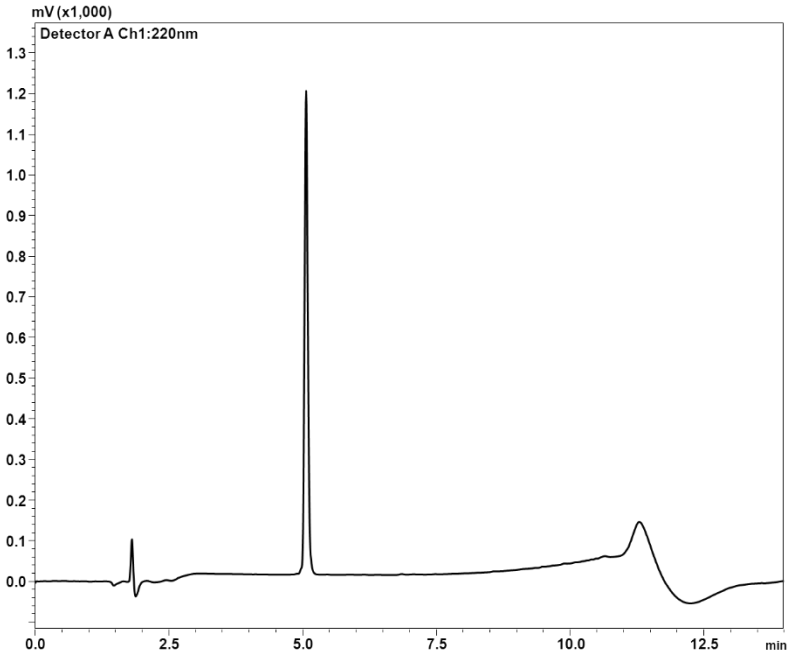


Figure 3.12 *Analytical HPLC spectrum of peptide* ***25***.

3.5.5 IN Activity Inhibition assay

Express Biotech International Company performed this assay using its HIV-1 IN Assay kit (Xpressbio Life Science Products, USA. www.xpressbio.com). Briefly, Streptavidin coated 96-well plates were coated with a double-stranded HIV-1 LTR U5 donor substrate (DS) oligonucleotide containing an end-labeled biotin. Full-length recombinant HIV-1 integrase protein was then loaded onto this oligo substrate. Compounds were added to the reaction and then a different double-stranded target substrate (TS) oligo containing 3'-end modifications was added to the plate. The sequence of DS DNA and TS DNA substrates used in this assay were 5'-P ACC CTT TTA GTC AGT GTG GAA AAT CTC TAG CAG TGAA AAT CAG TCA CAC CTT TTA GAG ATC GTC A and 5' TGA CCA AGG GCT AAT TCA CT ACT GGT TCC CGA TTA GAS GA respectively.[27,28]

The HIV-1 IN cleaves the terminal two bases from the exposed 3'-end of the HIV-1 LTR DS and then catalyzes a strand-transfer reaction to integrate the DS into the TS. The products of the reaction were detected colorimetrically using an HRP-labeled antibody directed against the TS 3'-end modification. Percent inhibition in the IN assay was calculated and IC_{50} values were determined.

3.5.6 HTRF-based IN Subunit Exchange Assay

Full-length IN and LEDGF proteins were expressed in *E. coli* BL21(DE3) and were purified as described.[9,10,19] His- and FLAG-tagged INs, were mixed in 25 mM Tris (pH 7.4) buffer containing 150 mM NaCl, 2 mM $MgCl_2$, 0.1% Nonidet P-40, 1 mg/ml BSA. Test compounds were then added to the mixture. A mixture of anti-His6-XL665 and anti-FLAG-EuCryptate antibodies were then added to the reaction and, after

an incubation, the plate are read and the HTRF signal is calculated from the 665 /620 nm ratio.

3.5.7 HTRF LEDGF-dependent and -independent HIV-1 IN assay

The IN LEDGF/p75 dependent assay allow to measure the inhibition of 3'processing and strand transfer IN reactions in presence of recombinant LEDGF/p75 protein, as previously described. Briefly, 50 nM IN was preincubated with increasing concentration of compounds in reaction buffer containing 20 mM HEPES pH 7.5, 1 mM DTT, 1% Glycerol, 20 mM $MgCl_2$, 0.05% Brij-35 and 0.1 mg/ml BSA. To this mixture, 100 nM DNA donor substrate and 50 nM DNA acceptor substrate and 50 nM LEDGF/p75 protein (or without LEDGF/p75 protein) were added and incubated at 37 °C for 90 minutes. After the incubation, 4 nM of Europium-Streptavidine were added at the reaction mixture and the HTRF signal was recorded using a Perkin Elmer Victor 3 plate reader using a 314 nm for excitation wavelength and 668 and 620 nm for the wavelength of the acceptor and the donor substrates emission, respectively.

3.5.8 Circular dichroism (CD)

All CD spectra were recorded using a JASCO J710 spectropolarimeter at 20°C between λ=260–190 nm (1 mm path, 1 nm bandwidth, 4 accumulations, and 100 nm min^{-1} scanning speed). Measurements were performed with peptides in H_2O (0.100 mM, pH 7.4) or in in 50% TFE/water solution.

3.5.9 Cell-based assays

The HIV-1 strain NL4.3 was obtained from the AIDS Research and Reference Reagent Program (Division of AIDS, NIAID, NIH, USA). The HIV-2 strain ROD was obtained from the Medical Research Council (MRC, London, UK). The anti-HIV assay in MT-4 cells has been described previously.[29] MT-4 cells ($1x10^6$ cells/ml; 50 μl of volume) were pre-incubated for 30 min at 37°C with the test compounds in 96-well plates (Falcon, BD Biosciences). The laboratory HIV strains (NL4.3 and ROD) were added according to the 50% tissue culture infectious dose ($TCID_{50}$) of the viral stock. Cellular cytopathicity was scored microscopically in MT-4 cells 5 days post-infection and IC_{50}s were calculated spectrophotometrically using MTS/PES. For the latter viability assay, the Cell-Titer96 Aqueous One Solution Proliferation Assay (Promega, Leiden, The Netherlands) kit was used.

3.5.10 NMR analysis

The NMR analysis of the peptides 25 and 39, was carried out in aqueous solution (1 mg peptide in 500 microliters H2O / D2O 9: 1 v / v pH 3.0) and in a 30% solution of TFE (1 mg peptide in 500 microliters to 30% TFE-d3 in H2O / D2O 9: 1 v / v pH 3.0), using DSS as reference. The NMR spectra were obtained by recording the measurements at 5 ° C and 25 ° C on a Bruker AV-600 equipped with a cryoprobe.

3.5.10.1 NMR assignment

^{1}H chemical shifts were assigned by standard strategy based on 2D COSY, TOCSY and NOESY spectra.

^{13}C chemical shifts were assigned from 2D $^{1}H,^{13}C$-HSQC spectra recorded at natural ^{13}C abundance.

In the case of peptide 39, which contains two Pro residues, only the major species was assigned. Based on the difference between the ^{13}C chemical shifts for the Cβ and Cγ carbons ($\delta_{C\beta} - \delta_{C\gamma}$, ppm; 3.3 and 4.6 ppm for P5 & P6 in aqueous solution, and 3.5 and 4.2 ppm for P5 & P6 in 30 % TFE), the two X-Pro bonds of the major species of 39 are trans (values for trans X-Pro bonds are < 5 ppm, and for cis X-Pro bonds are about 10 ppm).

3.5.10.2 Structure calculations

To visualise the structural features of the helices formed by peptides 25 and 39 in the presence of TFE, we performed structure calculations. Upper limit distance restraints were derived from the NOE cross-peaks observed in 2D NOESY spectra recorded at 5ºC, and 25ºC. The automatic integration sub-routine of the Sparky program was employed to integrate the cross-peaks. Restraints for the ϕ and ψ dihedral angles were obtained from the $^{1}H\alpha$, $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts using the TALOS-n server. Structures were calculated using the standard iterative protocol for automatic NOE assignment implemented in CYANA 2.1 program. This protocol performs seven cycles of combined NOE assignment and structure calculation. 100 conformers were calculated per cycle. The NMR structure for each peptide corresponds to the

ensemble of the 20 conformers with the lowest target function values resulting from the last cycle.

3.6 References

1. Eijkelenboom APAM, Vandenent FMI, Vos A, Doreleijers JF, Hard K, Tullius TD. The solution structure of the amino-terminal HHCC domain of HIV-2 integrase: a three-helix bundle stabilized by zinc. *Current Biology*. **1997**; 7, 739-46

2. Wolfe SA, Grant RA, Elrod-Erickson M, Pabo CO. Beyond the 'recognition code': structures of two Cys2His2 zinc finger/TATA box complexes. *Structure (Camb)*. **2001**; 9, 717-723.

3. Liao SM, Qi-Shi Du QS, Meng JZ, Zong-Wen Pang ZW, Huang RB.The multiple roles of histidine in protein interactions. *Chemistry Central Journal*. **2013**; 7:44.

4. Azmi S, Srivastava S, Mishra NN, Tripathi JK, Shukla PK, Ghosh JK. Characterization of Antimicrobial, Cytotoxic, and Antiendotoxin Properties of Short Peptides with Different Hydrophobic Amino Acids at "a" and "d" Positions of a Heptad Repeat Sequence. *J. Med. Chem*. **2013**; 56, 924-939.

5. Javadpour MM and Barkley MD. Self-assembly of designed antimicrobial peptides in solution and micelles. *Biochemistry*. **1997**; 36, 9540-9549.

6. Cai M, Zheng R, Caffrey M, Craigie R, Clore GM, Gronenborn AM. Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. *Nat. Struct. Biol*. **1997**; 4, 567-77.

7. Engelman A and Craigie R. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. *J. Virol*. **1992**; 66, 6361-6369.

8. Christ F and Debyser Z. The LEDGF/p75 integrase interaction, a novel target for anti-HIV therapy. *Virology*. **2013**; 435, 102-10.

9. Esposito F, Tintori C, Ferrarese R, Cabiddu G, Corona A, Ceresola ER, Calcaterra A, Iovine V., Botta B, Clementi M, Canducci F, Botta M, Tramontano E. Kuwanon-L as a new allosteric HIV-1 integrase

inhibitor: molecular modeling and biological evaluation. *ChemBioChem.* **2015**; 16, 2507-2512.

10. Tintori C, Esposito F, Morreale F, Martini R, Tramontano E, Botta M. Investigation on the sucrose binding pocket of HIV-1 Integrase by

molecular dynamics and synergy experiments. *Bio org. Med. Chem. Lett.* **2015;** 25, 3013-3016.

11. Jamieson AG, Boutard N. et al. *Chem. Biol. Drug. Des.* **2013**; 81,148–165.

12. Masuda T, and Kannagi M. European Patent Specification, International application number: PCT/JP2000/002698, International publication number: WO 2000/070035 (23.11.2000 Gazette 2000/47).

13. Sala M, Spensiero A, Esposito F, Scala M C, Vernieri E, Bertamino A, Manfra M, Carotenuto A, Grieco P, Novellino E, Cadeddu M, Tramontano E, Schols D, Campiglia P, Gomez-Monterrey I M. Development and Identification of a Novel Anti-HIV-1 Peptide Derived by Modification of the N-Terminal Domain of HIV-1 Integrase. *Frontiers in Microbiology, Virology*. **2016**; 7, 845.

14. Hruby VJ and Bonner GG. Design of Novel Synthetic Peptides Including Cyclic Conformationally and Topgraphically Constrained Analogs. *Methods in Mol Biol* **1994;** 35, 201-240.

15. Atherto E and Sheppard RC. Solid-Phase Peptide Synthesis: A Practical Approach. *IRL Press*, Oxford, U.K. **1989**.

16. Chen X, Park R, Shahinian AH, Bading JR, Conti PS. Pharmacokinetics and tumor retention of 125I-labeled RGD peptide are improved by PEGylation. *Nucl Med Biol.* **2004**; 31, 11-19

17. Echalier C, Al-Halifa S, Kreiter A, Enjalbal C, Sanchez P, Ronga L, Puget K, Verdie´ P, Amblard M, Martinez J, Subra G. Heating and microwave assisted SPPS of C-terminal acid peptideson trityl resin: the truth behind the yield. *Amino Acids* **2013**; 45:1395–1403

18. Grieco P, Gitu PM, Hruby VJ. Preparation of 'side-chain-to-side-chain' cyclic peptides by allyl and alloc strategy: potential for library synthesis. *J Pept Res.* **2001;** 57, 250-256

19. Kessl JJ, Jena N, Koh, Y, Taskent-Sezgin H, Slaughter A, Feng, L, de Silva S, Wu L, Le Grice SF, Engelman A, Fuchs JR, Kvaratskhelia M.. Multimode, cooperative mechanism of action of allosteric HIV-1 integrase inhibitors. *J. Biol. Chem.* **2012**; 287, 16801-16811.

20. Lia X, Krishnana L, Cherepanovb P, Engelman A. Structural biology of retroviral DNA integration. *Virology.* **2011**; 15; 411, 194-205.

21. Michel F, Crucifix C, Granger F, Eiler S, Mouscadet JF, Korolev S, Agapkina J, Ziganshin R, Gottikh M, Nazabal A, Emiliani S, Benarous R, Moras D, Schultz P, Ruff M. Structural basis for HIV-1 DNA integration in the human genome, role of the LEDGF/P75 cofactor. *Embo J.* **2009**; 28, 980-91.

22. Li HY, Zawahir Z, Song LD, Long, YQ, Neamati N. Sequence-Based Design and Discovery of Peptide Inhibitors of HIV-1 Integrase: Insight into the Binding Mode of the Enzyme. *J. Med. Chem.* **2006**; 49, 4477-4486.

23. Lia X, Krishnana L, Cherepanovb P, Engelmana A. Structural biology of retroviral DNA integration. *Virology.* **2011**; 15; 411, 194-205.

24. Liao C, Marchand, C, Burke TR Jr, Pommier Y, Nicklaus MC. Authentic HIV-1 integrase inhibitors. *Future Med. Chem*. **2010**; 2, 1107-1122.

25. Malik L, Tofteng AP, Pedersen SL, Sørensen KK, Jensen KJ. Automated 'X-Y' robot for peptide synthesis with microwave heating: application to difficult peptide sequences and protein domains. *J. Pept. Sci.* **2010**; 16, 506-512.

26. Wang SS, Tam JP, Wang BS, Merrifield RB. Enhancement of Peptide Coupling Reactions by 4-Dimethylaminopyridine. *Int. J. Pept. Protein Res.* **1989**; 18, 459-467.

27. Hazuda D J, Hastings J C, Wolfe A L, Emini E A. A novel assay for the DNA strand-transfer reaction of HIV-1 integrase. *Nucleic Acid Res.* **1994**; 2, 1121-1122.

28. Engelman A, Bushman FD, Craigie R. Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. *EMBO J.* **1993**; 12, 3269-3275.

29. Vermeire K, Princen K, Hatse S, De Clercq E, Dey, K, Bell TW, Schols D. CADA, a novel CD4-targeted HIV inhibitor, is synergistic with various anti-HIV drugs in vitro. *AIDS.* **2004**; 18, 2115.

CHAPTER IV:

Development and identification of a novel anti-HIV peptide: design, synthesis and biological evaluation

Based on the preliminary screening data discussed in the third chapter, we focused on peptide **5** (EKYHSNWRAM), which showed a particular increases of the integrase activity in vitro assays (*figure 4.1*).[1] Moreover, peptide **5** is included in the second NTD α-helix of integrase, known in the literature to be responsible of binding with the viral DNA.[2]

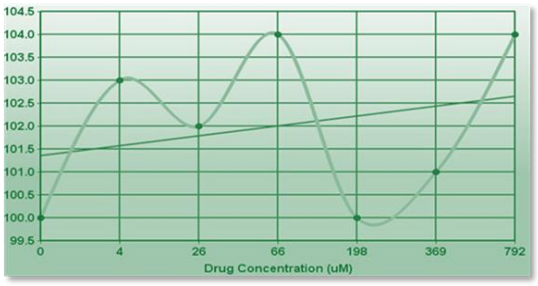


Figure 4.1. *Activity in vitro assay of peptide **5**.*

In this section, we discuss the evaluation of the activity of the peptide 5 in cell, considering that the translocation through the plasma membrane is a major limiting step for the cellular delivery of macromolecules. A promising strategy to overcome this problem consists in the chemical conjugation (or fusion) to cell penetrating peptides (CPP) derived from proteins able to cross the plasma membrane.[3] A large number of different cargo molecules such as oligonucleotides, peptides, peptide nucleic acids, proteins or even nanoparticles have been internalized in cells by this strategy.

4.1 Cell penetrating peptides

The integrity of cell membranes is critical to protect cells and for maintaining tissue homeostasis. Indeed the organization of cell

membranes prevents the leakage and the entrance of small molecules, proteins and genetic material in a not specific manner. The presence of cellular and tissue barriers is a strong drawback for drug delivery.

Recently, a new and efficient strategy is based on the use of CPPs (cell-penetrating peptides). The CPPs are peptides with a short sequence (fewer than thirty amino acids), able to cross the membrane at any point in a free manner. In this way, active molecules, antibodies, peptides, nanoparticles, nucleic acids, conjugated to a CPP, can be easily distributed, increasing their concentration, with an increase of the physiological activity.[4]

Therefore, the CPPs represent a revolutionary tool for the treatment of several human diseases, with a reduction of toxicity and invasiveness. The peptide TAT and Penetratine were the first CPPs discovered, and are still the most studied today.[5] The HIV-1 TAT (transactivator of transcription) gene codes for a 14-kDa protein and as its name suggests, it is a key activator of HIV-1 transcription. It is one of the first proteins to be expressed after infection occurs. Unlike typical transcription factors that are DNA binding proteins, TAT is a RNA binding protein that recognizes a specific sequence, TAR (*T*rans*a*ctivator *R*esponse Element), from the HIV-1 RNA molecule.[6]

TAT is a protein which consists a transcriptional activation domain containing a region enriched in cysteine and a hydrophobic motif enriched in arginine, that determines the binding of RNA protein. Park et al. defined the minimal peptide sequence (TAT 49-57) needed to cross membranes without losing efficiency. This region is enriched in basic aminoacids, lysine and arginine, and it is important for DNA binding. The peptides are relatively big molecules, so the transition from the external aqueous medium to the internal is difficult. Furthermore the

peptides, being of the amino acid polymers, have an electric charge, which does not allow the free passage through the cell membrane. Several studies show that the TAT peptide (GRKKRRQRRRPQ) can be used as cell penetrating peptides (CPP).[7]

4.1.1 Conjugation with fatty acids

Another approach used to improve the pharmacokinetics properties is the conjugation of the peptide to a lipid. In this way, the half-life of the drug is prolonged, because it can remain more time in the circulation.
Furthermore, the conjugated peptide to a lipid for various applications can be used, for example to increase their antibacterial activity.
For the conjugation of a fatty acid to peptide various approaches are available. Usually the lipids are conjugated to the N-terminal sequence of the peptide, or to the side chain of lysine residue. Moreover the cysteine residues in a peptide sequence can be conjugated to fatty acids, giving the corresponding thioesters derivatives. Typically, the derivatization involves long-chain fatty acids. The commonly used fatty acids are caprylic acid (C8), lauric acid (C12), myristic acid (C14), palmitic acid (C16) or stearic acid (C18).[8]

4.2 Results

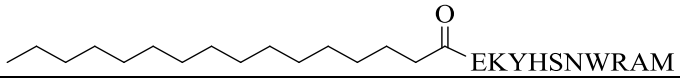
4.2.1 Design of peptides able to penetrate in cells

On the basis of the above considerations, in order to obtain compounds able to cross the cytoplasmic membrane, we conjugated peptide **5 with** two vectors (*Table 4.1*):

1. TAT sequence, GRKKRRQRRRPQ
2. Palmitic Acid, a saturated fatty acid to 16 carbon atoms.

In this way the peptides could be able to cross the membrane and enter into the cell.

Table 4.1. *Conjugated peptides.*

Peptide	Sequence
TAT-5	GRKKRRQRRRPQEKYHSNWRAM
PA-5	

4.2.2 Cell-based HIV replication assay on peptides 5, TAT-5 and PA-5

The synthesized peptides were evaluated for their capability to inhibit the HIV-1 and HIV-2 replication in MT-4 cells infected with the virus (*Table 4.2*). Only peptide 5 conjugated with fragment TAT (**TAT-5**), showed a significant inhibitory activity against HIV-1 and HIV-2 (IC_{50} value in the low micromolar range).

Table 4.2. Evaluation of the synthesized peptides against HIV-1 and HIV-2 replication in MT-4 cell cultures.

Peptide	HPLC K'	ESI-MS	EC_{50} [a] (µM)		CC_{50} [b] (µM)
			HIV-1	HIV-2	MT-4
5	5.02	1363.70	> 110	> 110	110± 0.6
PA-5	5.67	1558.47	> 103	> 103	103± 0.2
TAT-5	4.56	2924.37	27± 0.7	30± 0.3	113± 0.5
AMD3100			10.73±	10.56±	> 1000

[a] 50% Effective concentration, or compound concentration required to inhibit HIV-induced cytopathogenic effect in MT-4 cell culture.
[b] Compound concentration required to reduce by 50% MT-4 cell viability.
*GRKKRRQRRRPQ: TAT

So, according to the results obtained, the contribution of the various amino acid residues to the antiviral activity of peptide **TAT-5** was established through an L-Ala scanning analysis.

The compounds obtained, **47-55** were evaluated for their capacity to inhibit the replication of HIV-1 and HIV-2 in MT-4 cells infected with the virus. Unfortunately, the analysis of the data showed that neither of peptides inhibited the viral replication (*Table 4.3*).

Table 4.3. Evaluation of the synthesized peptides (**47-55**) against HIV-1 and HIV-2 replication in MT-4 cell cultures.

Peptide	Sequence	HPLC K'	ESI-MS	EC_{50} [a] (µM)		CC_{50} [b]
				HIV-1	HIV-2	MT-4
TAT-5	TAT-EKYHSNWRAM	4.56	2924.37	**27± 0.7**	**30± 0.3**	**113±**
TAT-47	TAT-AKYHSNWRAM	4.53	2866.34	> 10	> 10	> 50
TAT-48	TAT-EAYHSNWRAM	3.45	2867.28	> 10	> 10	> 50
TAT-49	TAT-EKAHSNWRAM	3.67	2832.27	> 10	> 10	> 50
TAT-50	TAT-EKYASNWRAM	3.57	2858.31	> 10	> 10	> 50
TAT-51	TAT-EKYHANWRAM	4.55	2908.37	> 10	> 10	> 50
TAT-52	TAT-EKYHSAWRAM	3.24	2881.35	> 10	> 10	> 50
TAT-53	TAT-EKYHSNARAM	4.42	2809.24	> 10	> 10	> 50
TAT-54	TAT-EKYHSNWAAM	4.89	2839.26	> 10	> 10	> 50
TAT-55	TAT-EKYHSNWRAA	3.79	2864.25	> 10	> 10	> 50
AMD3100	-			10.73±0.2	10.56±0.	> 1000

[a] 50% Effective concentration, or compound concentration required to inhibit HIV-induced cytopathogenic effect in MT-4 cell culture.
[b] Compound concentration required to reduce by 50% MT-4 cell viability.
*GRKKRRQRRRPQ: TAT

4.3 Chemistry

4.3.1 General procedure for synthesis of PA-5

The synthesis of peptide (**PA-5**) was performed according to the solid phase approach using standard Fmoc methodology in a manual reaction vessel.[9] The first amino acid, was linked to the Rink resin previously deprotected by a 25% piperidine solution in N,N-dimethylformamide (DMF) for 30 min. The following protected amino acids were then added stepwise. Each coupling reaction was accomplished using an amino acid with HBTU and HOBt in the presence of DIPEA. The N-terminal Fmoc group was removed with a 25% solution of piperidine in DMF, then the conjugation of palmitic acid was carried out.[10] Finally, peptide were released from the resin with trifluoroacetic acid (TFA/TES/H2O (90:5:5) for 3 h. The resin was removed by filtration, and the crude peptides were recovered by precipitation with cold anhydrous ethyl ether to give a white powder and then lyophilized (*figure 4.2*).

All crude peptides were purified by RP-HPLC on a preparative C18-bonded silica column using a Shimadzu SPD 10A UV/VIS detector, with detection at 215 and 254 nm. All peptides were characterized with analytical RP-HPLC and peptides molecular weights were determined by ESI mass spectrometry in a MS Thermo Quest Finnigan LTQ Advantage. All analogues showed >97% purity when monitored at 215 nm.

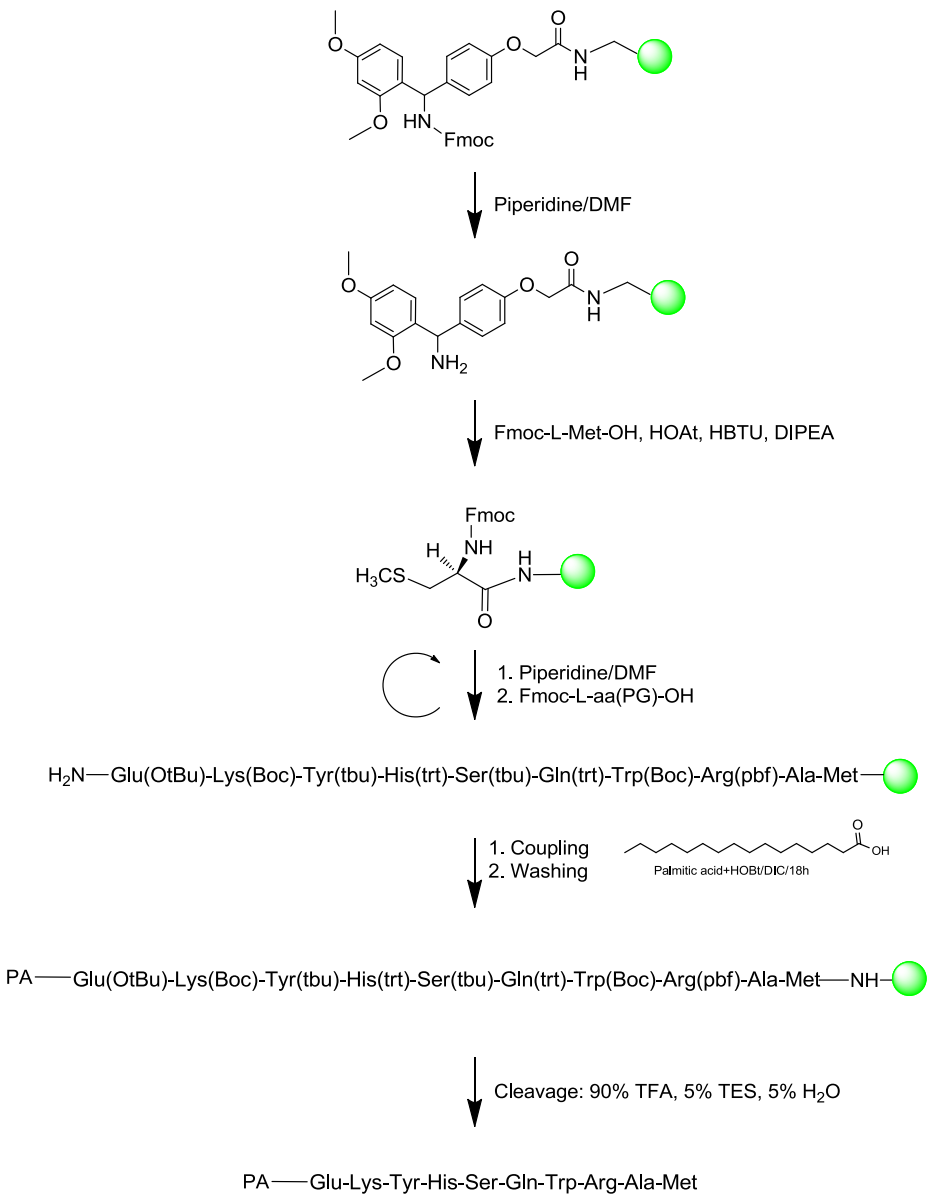


Figure 4.2. *Schematic representation of the peptide **PA-5** synthesis.*

4.3.2 Synthesis of peptide TAT-5, TAT-47-55

The synthesis of peptides conjugated to TAT fragment (**TAT-5** and from **TAT-47** to **TAT-55**) was carried out by microwave-assisted solid-phase synthesis, on a Rink Amide-ChemMatrix.

The advantages of the use of microwaves on chemical systems can be summarized as follows: (a) volumetric heating, without contact; (b) very high heating rate; (c) good interaction with many insulating materials and poor interaction with many nonferrous metals or gaseous products; and (d) convenient and clean heating system. The presence of microwaves greatly improves the rate of many chemical reactions, with a strong reduction of the reaction time, without altering the final yield. Reactions Chemistry by microwaves can also be carried out in dry conditions or in solvents with a high dielectric constant.[11] The synthesis by microwaves has found great success in organic chemistry, improving the yield and the completion of many reactions. The great success achieved in the organic field has also allowed the application in the chemistry of peptides (particularly the SPPS). Indeed, since the completion of a peptide synthesis requires the obtaining of two per cycle reactions (deprotection and coupling), the microwave energy is an efficient way to carry out these reactions. A frequent problem in peptide synthesis, in particular in long peptide sequences (30 or more amino acids) or hydrophobic is the aggregation of the chain. The energy derived by microwave is able to break the aggregation allowing the peptide to take on the favorable conformation at lower energy. Finally, the use of microwave can improve, in terms of yield and purity, the conventional solid phase synthesis, which instead takes place at room temperature.

4.4 Discussion

During my PhD project a library of peptides was developed from the NTD sequence (**1-50**) by using a mimotopic strategy. Preliminary results of the *in vitro* integrase assay indicated that the peptide **5** EKYHSNWRAM, increases significantly the HIV-1 IN activity. Peptide **5** is included into the sequence of the 2nd α-helix of NTD of integrase, which is known in the literature to be responsible of binding with the viral DNA. According to the antiviral activity in cell, the peptide **5** alone is not able to cross the cellular membrane, while **PA-5** is not active in cell. We hypothesized that the presence of fatty acid makes the compound very slightly soluble, preventing the correct penetration of the peptide.

Interestingly, peptide **5**, conjugated with the cell-penetrating Tat fragment, is the only compound of this series able to inhibit the HIV-1 and HIV-2 induced cytopathogenic effect in MT-4 cell. Although the antiviral activity is about 5-fold lower than the concentration of its cellular toxicity, peptide **TAT-5** represents a suitable lead compound for the development of novel derivatives able to inhibit of the viral replication.

Ala scanning performing on **TAT-5** showed that all substitutions modify substantially the inhibitory activity of the resulting peptides. Similarly to what happened with first series, these data still showed the difficulty of establishing a structure-activity relationship with peptides that target the viral integrase.

4.5 Experimental section

4.5.1 Material and methods

N^{α}-Fmoc-protected amino acids, HOAt, HBTU, DIEA, piperidine and trifluoroacetic acid were purchased from Iris Biotech (Germany). Rink Amide-ChemMatrix resin was purchased from Biotage AB (Sweden). Palmitic Acid was purchased from Sigma-Aldrich.

Peptide synthesis solvents, reagents, as well as CH_3CN for HPLC were reagent grade and were acquired from commercial sources and used without further purification unless otherwise noted.

4.5.1.1 Peptide synthesis by microwaves

The peptides **TAT-5** and from **TAT-47** to **TAT-55** were synthesized using a Biotage Syro Wave™ fully automated microwave and parallel peptide synthesizer or assembled on the Automated Microwave Peptide Synthesizer from Biotage AB (Initiator+ Alstra™).

Peptides were performed on a Rink Amide-ChemMatrix resin (150 mg, loading 0.4-0.6 mmol/g), previously deprotected with 25% piperidine/DMF (1x 3 min, 1x 10 min) at room temperature. The resin was then washed with DMF (4 x 4.5 ml). The following protected amino acids were then added on to the resin stepwise. Coupling reactions were performed using N^{α}-Fmoc amino acids (3.0 eq., 0.5 M), using as coupling reagent HBTU (3eq, 0.6M), HOAt (3eq, 0.5M) and DIEA (6eq, 2M) in N-methyl-2-pyrrolidone (NMP). All couplings were achieved for 10 min at 75 ºC (2x) and 2x45 min at RT for histidine and cysteine couplings to avoid the epimerization. After each coupling step, the Fmoc protecting group was removed as described above.

The resin was washed with DMF (4 × 4.5 ml) after each coupling and deprotection step. Finally peptides were released from the resin as described above.

4.5.1.2 Synthetic Procedure: conjugation of the peptide with palmitic acid

The synthesis of peptide **PA-5** was performed according to the solid phase approach using a standard Fmoc methodology in a manual reaction vessel.[9]

After deprotection of the Fmoc group, the conjugation of palmitic acid was carried out.[10] So we used 4 equivalent of palmitic acid; 4.4 equivalents of DIC (diisopropilcarbodiimide); 8 equivalent of HOBT; using DMF and DCM as the reaction solvent. These reagents were added under stirring keeping the reactor of the synthesis for 18 hours. Then, the peptides were washed several times with DCM (3x3 mL), DMF (2x3 mL), again with DCM (2x3 mL) and finally with MeOH (2x3 mL). To confirm the reaction of addition of palmitic acid, the Kaiser test was carried out.

At the end, we performed cleavage reaction using a mixture of 10 mL of 90% TFA, 5% of Et3SiH (TES) and 5 % of H_2O, stirring for 3 hours. Finally, the resin was removed from the solution by filtration and to the latter was added ethyl ether to give a white precipitate. The crude peptide was recovered through the use of a rotavapor, to obtain a white powder.

4.5.2 Purification and characterization

All crude peptides were purified by RP-HPLC on a preparative C18-bonded silica column (Kinetex 5μm, 100Å, 100×21.2mm) using a Shimadzu SPD 10A UV−Vis detector, with detection at 210 nm and 254

nm. The column was perfused at a flow rate of 20 mL/min with solvent A (10%, v/v, water in 0.1% aqueous TFA), and a linear gradient from 5% to 90% of solvent B (80%, v/v, acetonitrile in 0.1% aqueous TFA) over 20 min was adopted for peptide elution. Analytical purity, was determined by RP-HPLC, using an autosampler, with dual wavelength absorbance detector with wavelength set to 215 nm, all of which were obtained from Shimadzu. The mobile phases employed were: solvent A (water + 0.1% TFA) and solvent B (acetonitrile + 0.1% TFA); using a linear gradient from 5% to 90% B over 11 min, fitted with C-18 column Phenomenex, Ascentis ES-C18 column (5cm x 3.0mm, 2.7µm), at a flow rate of 0.800 mL/min. ESI mass spectrometry was determined using a Q-Tof Premier (ESI/Q-Tof). The sample was dissolved in a mixture of water and methanol (50/50) and injected directly into the electrospray source, using a syringe pump, which maintains constant flow at 5 µL/min. The temperature of the capillary was set at 220 °C

4.5.3 Cell-based assays

The HIV-1 strain NL4.3 was obtained from the AIDS Research and Reference Reagent Program (Division of AIDS, NIAID, NIH, USA). The HIV-2 strain ROD was obtained from the Medical Research Council (MRC, London, UK). The anti-HIV assay in MT-4 cells has been described previously.[12] MT-4 cells ($1x10^6$ cells/ml; 50 µl of volume) were pre-incubated for 30 min at 37°C with the test compounds in 96-well plates (Falcon, BD Biosciences). The laboratory HIV strains (NL4.3 and ROD) were added according to the 50% tissue culture infectious dose ($TCID_{50}$) of the viral stock. Cellular cytopathicity was scored microscopically in MT-4 cells 5 days post-infection and IC_{50}s were calculated spectrophotometrically using MTS/PES.

For the latter viability assay, the Cell-Titer96 Aqueous One Solution Proliferation Assay (Promega, Leiden, The Netherlands) kit was used.

4.6 References

1. Masuda T and Kannagi M. European Patent Specification, International application number: PCT/JP**2000**/002698, International publication number: WO **2000**/070035 (23.11.**2000** Gazette **2000**/47).

2. Sala M, Spensiero A**,** Esposito F, Scala MC, Vernieri E, Bertamino A, Manfra M, Carotenuto A, Grieco P, Novellino E, Cadeddu M, Tramontano E, Schols D, Campiglia P, and Gomez-Monterrey I M. Development and Identification of a Novel Anti-HIV-1 Peptide Derived by Modification of the N-Terminal Domain of HIV-1 Integrase *Frontiers in Microbiology, Virology*. **2016**, 7, 845.

3. Silhol M, Tyagi M, Giacca M, Lebleu B, Vivès E. Different mechanisms for cellular internalization of the HIV-1 Tat-derived cell penetrating peptide and recombinant proteins fused to Tat. *Eur J Biochem.* **2002**;269(2):494-501.

4. Sangho L, Ja-Hyun K, and Je-Min Choi. Use of Cell-Penetrating Peptides in Dendritic Cell-Based Vaccination *Immune Netw*. **2016**; 16(1): 33–43. doi: 10.4110/in.2016.16.1.33 PMCID: PMC4770098

5. Frankel AD, Pabo CO. Cellular uptake of the tat protein from human immunodeficiency virus. *Cell.* **1988**; 23;55(6):1189-93.

6. Feng S and Holland EC. HIV-1 tat trans-activation requires the loop sequence within tar. *Nature*. **1988**; Jul 14;334(6178):165-7.

7. Gump JM and Dowdy SF. TAT transduction: the molecular mechanism and therapeutic prospects. *Trends Mol Med.* **2007**;13(10):443-8.

8. Thomas PH. Immunogenic-ty of a synthetic peptide HBsAg: Enhancement by conjugation to a fatty acid carrier. *Molecular Immunology*. **1984**; 21 (1): 13-16

9. Atherton E and Sheppard RC. Solid-Phase Peptide Synthesis: A Practical Approach. *IRL Press, Oxford, U.K.* **1989**.

10. Hossein M, Ann R, Wei-Chiang. Water-soluble fatty acid derivatives as acylating agents for reversible lipidization of polypeptides. *FEBS Letters*. **1995**; 371(3): 283–286

11. Fini A, Breccia A. Chemistry by microwaves. *Pure appl. Chem*., **1999**; 71(4):573–579.

12. Vermeire K, Princen K, Hatse S, De Clercq E, Dey K, Bell TW, Schols D. CADA, a novel CD4-targeted HIV inhibitor, is synergistic with various anti-HIV drugs in vitro. *AIDS.* **2004**; 18(16):2115-25.

CHAPTER V:

Summary

5.1 Summary

The search for an effective therapy for the acquired immunodeficiency syndrome (AIDS) continues to be one of the biggest challenges for medicine. The success of highly active antiretroviral therapy (HAART) combinations, has improved patient long-term survival and tolerability to drug regimens. However, mutations of the HIV-1 genome confer resistance to those drugs in a large number of patients. Thus, the development of new inhibitors, targeting a distinct step in the retroviral life cycle, remains essential to reduce side effects and selection of drug-resistant viruses.

In this context, the viral enzyme integrase (IN), which is essential for the replication of human immunodeficiency virus type 1 (HIV-1), represents an important target for the development of new antiretroviral drugs.

The structure of IN consists of three domains: the N-terminal domain, the core domain that has the catalytic DDE motif which is required for its enzymatic activity and the C-terminal domain. The design and discovery of integrase inhibitors were first focused at targeting the catalytic site of IN with a specific effect on strand transfer. However, mutations of the HIV-1 genome confer resistance to those drugs in a large number of patients. For this reason the aim of my PhD project was the identification of new allosteric modulators of HIV-1 integrase (IN). We focused on the N-terminal domain (NTD) of HIV-1 IN which plays an important role in the formation of tetramer integrase process. Indeed, this domain represents an important target for the design of drugs for inhibition of IN oligomerization process. Thus, we synthesized a library of overlapping peptide sequences, with specific length and specific offset covering the entire native protein sequence NTD IN 1-50 (1-15). Then, we focused on peptide **5** and peptide **11** included in the 2rd α-

helix and the 3rd α-helix of N-terminal of IN respectively. Initially we synthesized three peptides derived from peptide **11**, in which the Cys residue was deleted (peptide **16**) or substituted by His amino acid a residue also involved in the Zn^{2+} coordination (peptide **17**). The same His was also used to replace the C-terminal Ser residue of peptide **16** leading to peptide **18**. We found that the most potent fragment, is peptide **18** (VVAKEIVAH), which inhibits the HIV-1 IN activity with an IC_{50} value of 4.5 μM. Amino acid substitution analysis on this peptide revealed essential residues for activity and allowed us to identify two nonapeptides (peptides **24** and **25**), that show a potency of inhibition similar to the one of peptide **18**.

To verify whether active peptides **18**, **24** and **25** can modulate the dynamic interplay between IN subunits, we tested them using an HTRF-based IN subunit exchange assay. Interestingly, peptide **18** does not interfere with the dynamic interplay between IN subunits, while peptides **24** and **25** modulated these interactions in different manners. Indeed, peptide **24** inhibited the IN-IN dimerization, while peptide **25** promoted IN multimerization, with IC_{50} values of 32 and 4.8 μM, respectively. In addition, peptide **25,** conjugated with the cell-penetrating Tat, has shown to have selective anti-infective cell activity for HIV-1. These results indicate that peptide **25** represents a suitable lead compound for the development of novel derivatives with improved toxicity profile.

To deepen on the structural requirements that determine the inhibitory activity of peptide **25** (VVAKAIVAH), we synthesized a new library containing the corresponding D-aminoacid (**29-37**, D-scanning analysis), while to verify if the position/sequence of amino acids is important for activity, we also synthesized the peptide **25** scramble (**38**). In addition, we synthesized peptide **39 (**DFNLPPVVAKAIVAH),

obtained from residues 25-39 of N-terminal sequence, which includes peptide **25** with addition of residues DFNLPP that may be essential to stabilize the conformation of the peptide.

The results obtained from the analysis HTFR confirmed that peptide **25** is the most powerful of the synthesized series. In addition to identify the shortest aminoacid sequence needed for the peptide activity we synthesized peptides **40-43** obtained from N- and C-terminal deletions of peptide **25**. Subsequently, in order to improve pharmacokinetic properties of peptide **25**, we design and synthesized cyclic peptides using different cyclization approaches (peptides **44, 45** and **46**); during the synthesis of peptide **44** we observed side reactions, solved with the use of Fmoc-His(Allyl)-Trityl resin, that has allowed to perform the cyclization reaction on the resin. So we obtained the desired cyclic peptide with a yield of 60%.

The studies to evaluate the antiviral activity of trunks peptides and cyclic peptides are still in progress. While the thesis is waiting for completion of some biological results, it is expected that truncated and cyclic peptides could offer additional information to guide next steps in this subject.

To evaluate the conformation of peptide **25** and peptide **39**, a NMR analysis was carried out. The NMR analysis showed an alpha helix conformation of peptide **25**, which could be essential for the interaction with IN.

Moreover, we considered to evaluate the activity of other peptides derived from the overlapping library of compounds, in particular we found that peptide **5** (EKYHSNWRAM), conveniently conjugated with the cell-penetrating fragment TAT, inhibits replication of HIV-1 and HIV-2 in infected MT-4 cells. Thus, contribution of the various amino

acid residues to the antiviral activity of peptide TAT-**5** was established through an L-Ala scanning analysis peptide (**47-55**). Unfortunately, the analysis of the data showed that none of the synthesized peptides inhibit viral replication and all substitutions modify substantially the inhibitory activity of the resulting peptides. In conclusion, the results obtained in this PhD project confirmed peptide **25** (VVAKAIVAH) as the most powerful of the synthesized series and the NMR analysis showed an alpha-helix conformation, which might be essential for the interaction with IN;

Moreover, we found that peptide **5** (EKYHSNWRAM), conveniently conjugated with the cell-penetrating fragment Tat, inhibits replication of HIV-1 and HIV-2 in infected MT-4 cells.

These results encourage us to pursue the path for design and the development of peptides derived from the N-terminal domain of IN, as potential anti-HIV agents.