UNIVERSITÀ DEGLI STUDI DI ROMA "TOR VERGATA"

FACOLTA' DI MEDICINA E CHIRURGIA

DOTTORATO DI RICERCA IN MICROBIOLOGIA ED IMMUNOLOGIA MEDICA

XX CICLO

Computational studies of mutations associated to resistance in HIV-1 macromolecular targets and implications in rational design of novel antiviral agents

Coordinatore   *Candidata*
Prof. Enrico Garaci   Anna Artese
Tutors
Prof. Carlo Federico Perno
Prof. Stefano Aquaro

anno accademico 2007/2008

Abstract

In order to discover novel selective anti-HIV resistance-evading drugs, a theoretical study was carried out combining structural analysis of RT crystallographic models, clinical data about RT conserved residues and an innovative computational method based on GRID maps. Such analysis allowed to reproduce clinical results and to highlight the consequences of the mutations in the recognition step. Moreover the computational approach generated a pharmacophore model useful for the design of novel RT inhibitors.

The presence of the I135T polymorphism in NNRTI-naive patients significantly correlated with the appearance of K103N in cases of NNRTI failure, suggesting that I135T may represent a crucial determinant of NNRTI resistance evolution. Molecular Dynamics simulations (MD) showed that I135T can contribute to the stabilization of the K103N-induced closure of the NNRTI binding pocket by reducing the distance and increasing the number of hydrogen bonds between 103N and 188Y.

In addition the influence of two drug resistance-associated mutations, L33F and L76V, of HIV-1 PR has been evaluated with respect to lopinavir (LPV) and atazanavir (ATV) molecular recognition. The evaluation of the interaction energies after the MD revealed that L33F substitution is related to reduced host/guest interactions, decreased affinity and to a dimer destabilizing effect for both PR inhibitors. In presence of L76V mutation, LPV showed a lowered binding affinity and a reduced hydrogen bonding network, while ATV complexes revealed a more productive binding affinity, increased host/guest interactions and dimer stabilizing effects, in agreement with hyper susceptibility data.

With the aim to estimate the stability of its 6-helix bundle, the gp41 conformational properties were investigated in presence of V38A and N140I, known enfuvirtide resistance-associated mutations. MD showed that the co-presence of V38A+N140I abolished the interaction between residue 38 and 145 important for the 6-helix-bundle stabilization.

Key-words

Riassunto

Al fine di identificare nuovi farmaci anti-HIV capaci di superare i problemi legati alla resistenza, è stato condotto uno studio teorico combinando l'analisi strutturale sui modelli cristallografici della trascrittasi inversa (RT), i dati clinici relativi ai residui conservati dell'RT ed un'innovativa metodica computazionale basata sulle mappe di GRID.

Tale analisi ha permesso di riprodurre i risultati clinici e di evidenziare le conseguenze delle mutazioni nella fase di ricognizione. Inoltre l'approccio computazionale ha portato all'identificazione di un modello farmacoforico utile per la progettazione di nuovi inibitori dell'RT.

E' stato riscontrato che la presenza del polimorfismo I135T nei pazienti NNRTI-naïve correlasse in modo significativo con la mutazione K103N nei casi di fallimento agli NNRTI, suggerendo così che la sostituzione I135T rappresenti un punto cruciale per l'evoluzione della resistenza agli NNRTI. Le simulazioni di dinamica molecolare (MD) hanno mostrato che la mutazione I135T contribuisce alla stabilizzazione della chiusura della tasca di legame degli NNRTI indotta dalla K103N in seguito alla riduzione della distanza ed all'aumento del numero di legami idrogeno tra l'Asn103 e la Tyr188.

Inoltre è stata valutata l'influenza di due mutazioni associate a resistenza, L33F e L76V, presenti a livello della proteasi (PR) di HIV-1 rispetto alla ricognizione molecolare del Lopinavir (LPV) e dell'Atazanavir (ATV). L'analisi delle energie di interazione ottenute in seguito alla MD ha rivelato che la mutazione L33F determina una riduzione delle interazioni tra il ligando ed il recettore, dell'affinità di legame e della stabilità del dimero per entrambi gli inibitori della PR. In presenza della mutazione L76V, il LPV ha mostrato una minore affinità di legame ed un ridotto network di legami idrogeno, mentre i complessi con l'ATV hanno rivelato una migliore affinità, un effetto stabilizzante a livello dell'interfaccia del dimero e più efficaci interazioni ligando-recettore, in accordo con i dati di ipersuscettibilità.

Al fine di valutare la stabilità del 6-helix bundle, sono state studiate le proprietà conformazionali della glicoproteina gp41 in presenza delle mutazioni associate a resistenza all'enfuvirtide V38A ed N140I. Le simulazioni di MD hanno mostrato che la copresenza delle mutazioni V38A+N140I è in grado di abolire l'interazione stabilita tra i residui 38 e 145, che risulta fondamentale per la stabilizzazione del 6-helix bundle.

Parole-chiave

Summary

1. Introduction

1.1. *Pathogenesis of HIV-1 infection*

In 1981 the Centers for Disease Control (CDC) reported an unusual occurrence of *Pneumocystis carinii* pneumonia (PCP) in five young homosexual men from Los Angeles [1]. Other unusual cases of diseases indicative of cellular immunodeficiency occurring in young, previously healthy homosexual men, were described later [2]. Initially, more than 90% of these cases occurred in homosexual and bisexual men, and therefore the term "gay-related immunodeficiency" (GRID) was coined. However, by 1982, hundreds of similar cases of an acquired immunodeficiency syndrome (AIDS) were reported in intravenous drug users, haemophiliacs, blood-transfusion recipients, heterosexual adults and infants born by mothers with the syndrome. Moreover the increase in the incidence of Kaposi's sarcoma (KS) in young gay men highlighted that AIDS was an emerging new syndrome [3]. An infectious etiology as the cause of AIDS was suggested by geographic clustering of cases and the depletion in numbers and function of T-helper cells suggested that a lymphotropic agent might be responsible [4]. The prominent immunosuppression that characterized AIDS recalled certain retroviral infections in animals. Human T-cell leukaemia virus type 1 (HTLV-1) was known to be tropic for T-helper cells and was a candidate for searching the cause of AIDS [5]. A human T-lymphotropic retrovirus immunologically related to HTLV-I was detected in lymph node tissue from a French patient who was at risk for AIDS [6]. This virus, later named "lymphadenopathy-associated virus" (LAV), proved to be highly related to the virus subsequently isolated at the National Cancer Institute, termed HTLV-III [7], and in San Francisco, termed AIDS-associated retrovirus (ARV) [8]. These viruses were consequently cloned and characterized [9] and are now known as human immunodeficiency virus type 1 (HIV-1). Molecular and sieroepidemiologic studies have consistently demonstrated a great correlation between infection with HIV-1 and subsequent development of AIDS [10]. All known HIV-1 sequences are closely related to simian immunodeficiency virus (SIV) sequences harboured by the *Pan troglodytes troglodytes* subspecies of chimpanzees [11]. In fact, zoonotic transmission of retroviruses from chimpanzees to humans is the most plausible explanation for the genesis of the HIV epidemic [11].

In spite of the rapid scientific advances that have been made, the HIV/AIDS epidemic has continued largely unchecked in most parts of the world. Prevention efforts and the

availability of antiretroviral therapies have blunted the expansion of the epidemic in the USA and Western Europe; however the diffusion of the epidemic continues in Africa and Southeast Asia. Approximately 300 cases of AIDS had been reported by the end of 1982; by the end of 2000, nearly 22 million individuals had died of AIDS, and 36 million were living with HIV-1 infection worldwide.

1.1.1. Epidemiology

The global HIV pandemic ranks among the greatest infectious disease in human history. During 2000 alone, nearly 6 million individuals worldwide acquired HIV infection, and 3 million deaths due to HIV infection occurred; these statistics highlight the continued expansion of the epidemic on a global scale [12]. Phylogenetic analysis of HIV proviral sequences reveals three major genetic groups of HIVs, the M (major), N (new), and O (outlier), that are close related to SIV sequences. These groups are classified into clades, or subtypes, according to their degree of genetic similarity [13]. Subtype B viruses comprise the overwhelming majority of viruses isolated in the USA, Europe, Australia and South America [14]. The dynamics of the spread of different subtypes of HIV through populations is determined by the genetics of the population as well as by viral properties. Zoonotic transmission of a retrovirus from chimpanzees to humans in sub-Saharan Africa is the most plausible explanation for the genesis of the HIV epidemic [11]. Currently, the heaviest burden of the epidemic is concentrated in southern Africa, India, and southeast Asia. Although only 12% of the world's population lives in sub-Saharan Africa, 70% of individuals who acquired HIV infection in 2000 live there. Of all AIDS deaths since the beginning of the AIDS epidemic, 78% occurred in sub-Saharan Africa, but the epidemic is expanding very rapidly in South Africa, Asia and eastern Europe, as shown in figure 1.
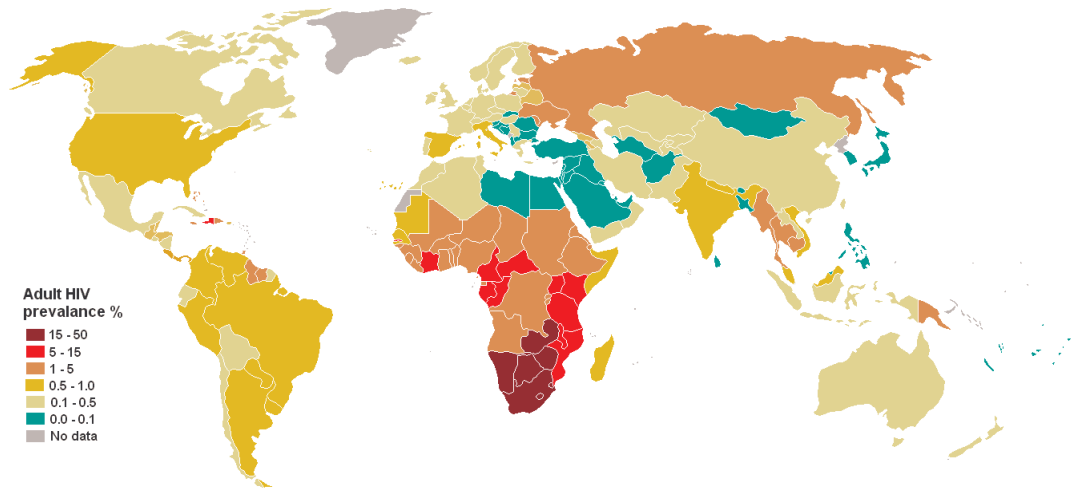
Figure 1. Adult HIV-1 world prevalence percentage updated as of November 2007 [15].

Nearly 1% of the adult population in India is infected with HIV. Although this seroprevalence is lower than that in many affected countries, it makes India the country with the largest number of HIV-infected individuals (approximately 4 million). Most HIV infections in India have occurred during the 1990s, but recently morbidity and mortality are increasing, with nearly 14% of women tested positive for HIV [16]. Current estimates place the burden of HIV infections at 500,000 in China, and this number appears to be growing rapidly. Molecular epidemiologic studies indicate that about 50% of viral isolates originated from injection-drug users in Thailand, and one third of isolates originated from injection-drug users in India. In the United States , between 600,000 and 800,000 individuals are infected with HIV [12]. The availability of therapeutic regimens in developed countries has led to substantial reductions in the death rate due to AIDS [17]. Unfortunately, no such benefit has reached the developing world, where AIDS-related mortality continues to exact a devastating toll. The nine African nations with an HIV seroprevalence of greater than 10% will experience a decline in average life expectancy of 17 years by the year 2015; instead of an average life expectancy that would have been on a trajectory to reach 64 years in the absence of HIV, it will decline to only 47 years [12]. Infant mortality is increasing in these countries as well, largely as a direct result of HIV infections. Finally, most developing countries do not have adequate resources to provide for a growing population of AIDS orphans; 20% to 30% of all children younger than 15 years had been orphaned by the year 2000 in nine sub-Saharan countries [18].

1.1.2. Transmission

The three main transmission routes of HIV are sexual contact, exposure to infected body fluids or tissues, and from mother to fetus during perinatal period. It is possible to find HIV in the saliva, tears and urine of infected individuals, but there are no recorded cases of infection by these secretions, and the risk of infection is negligible.

The majority of HIV infections are acquired through unprotected sexual relations between partners, one of whom has HIV. Heterosexual intercourse is the primary mode of HIV infection worldwide [19]. Sexual transmission occurs with the contact between sexual secretions of one partner with the rectal, genital or oral mucous membranes of another. Unprotected receptive sexual acts are riskier than unprotected insertive sexual acts, with the risk for transmitting HIV from an infected partner to an uninfected partner through unprotected anal intercourse greater than the risk for transmission through vaginal intercourse or oral sex. Oral sex is not without its risks as HIV is transmissible through both insertive and receptive oral sex [20]. The risk of HIV transmission from exposure to saliva is considerably smaller than the risk from exposure to semen; contrary to popular belief, one would have to swallow gallons of saliva from a carrier to run a significant risk of becoming infected [21]. Approximately 30% of women in ten countries representing "diverse cultural, geographical and urban/rural settings" report that their first sexual experience was forced or coerced, making sexual violence a key driver of the HIV/AIDS pandemic [22]. Sexual assault greatly increases the risk of HIV transmission as protection is rarely employed and physical trauma to the vaginal cavity frequently occurs which facilitates the transmission of HIV [23].

Sexually transmitted infections (STI) increase the risk of HIV transmission and infection because they cause the disruption of the normal epithelial barrier by genital ulceration and/or microulceration; and by accumulation of pools of HIV-susceptible or HIV-infected cells (lymphocytes and macrophages) in semen and vaginal secretions. Epidemiological studies from sub-Saharan Africa, Europe and North America have suggested that there is approximately a four times greater risk of becoming infected with HIV in the presence of a genital ulcer such as those caused by syphilis and/or chancroid. There is also a significant though lesser increased risk in the presence of STIs such as gonorrhea, Chlamydial infection and trichomoniasis which cause local accumulations of lymphocytes and macrophages [24].

Transmission of HIV depends on the infectiousness of the index case and the susceptibility of the uninfected partner. Infectivity seems to vary during the course of illness and is not constant between individuals. An undetectable plasma viral load does not necessarily indicate a low viral load in the seminal liquid or genital secretions. Each 10-fold increment of blood plasma HIV RNA is associated with an 81% increased rate of HIV transmission [24; 25]. Women are more susceptible to HIV-1 infection due to hormonal changes, vaginal microbial ecology and physiology, and a higher prevalence of sexually transmitted diseases [26]. People who are infected with HIV can still be infected by other, more virulent strains.

HIV can also be transmitted through exposure to infected body fluids. This transmission route is particularly relevant to intravenous drug users, hemophiliacs and recipients of blood transfusions and blood products. Sharing and reusing syringes contaminated with HIV-infected blood represents a major risk for infection with not only HIV, but also hepatitis B and hepatitis C. Needle sharing is the cause of one third of all new HIV-infections and 50% of hepatitis C infections in North America, China, and Eastern Europe. The risk of being infected with HIV from a single prick with a needle that has been used on an HIV-infected person is thought to be about 1 in 150 (Table 1) [27].

Exposure route	Estimated infections per 10,000 exposures to an infected source
Blood transfusion	9,000
Childbirth	2,500
Needle-sharing injection drug use	67
*Receptive anal intercorse**	50
Percutaneous needle stick	30
*Receptive penile-vaginal intercourse**	10
*Insertive anal intercourse**	6.5
*Insertive penile-vaginal intercourse**	5
*Receptive oral intercorse**	1§
*Insertive oral intercourse**	0.5§
* assuming no condom use § Source refers to oral intercourse performed on a man	

Table 1. Estimated per act risk for acquisition of HIV by different exposure routes [27].

Post-exposure prophylaxis with anti-HIV drugs can further reduce that small risk. Health care workers (nurses, laboratory workers, doctors etc) are also concerned, although more rarely. This route can affect people who give and receive tattoos and piercings. Universal precautions are frequently not followed in both sub-Saharan Africa and much of Asia because of both a shortage of supplies and inadequate training. The World Health Organization (WHO) estimates that approximately 2.5% of all HIV infections in sub-Saharan Africa are transmitted through unsafe healthcare injections. Because of this, the United Nations General Assembly, supported by universal medical opinion on the matter, has urged the nations of the world to implement universal precautions to prevent HIV transmission in health care settings [28].

The risk of transmitting HIV to blood transfusion recipients is extremely low in developed countries where improved donor selection and HIV screening is performed. However, according to the WHO, the overwhelming majority of the world's population does not have access to safe blood and "between 5% and 10% of HIV infections worldwide are transmitted through the transfusion of infected blood and blood products" [29].

The transmission of the virus from the mother to the child can occur *in utero* during the last weeks of pregnancy and at childbirth. In the absence of treatment, the transmission rate between the mother to the child during pregnancy, labor and delivery is 25%. However, when the mother has access to antiretroviral therapy and gives birth by caesarean section, the rate of transmission is just 1% [30]. A number of factors influence the risk of infection, particularly the viral load of the mother at birth (the higher the load, the higher the risk). Breastfeeding increases the risk of transmission by 10–15%. This risk depends on clinical factors and may vary according to the pattern and duration of breast-feeding. Some recent studies have shown that antiretroviral drugs, caesarean delivery and formula feeding reduce the chance of transmission of HIV from mother to child [31]. Current recommendations state that when replacement feeding is acceptable, feasible, affordable, sustainable and safe, HIV-infected mothers should avoid breast-feeding their infant. However, if this is not the case, exclusive breast-feeding is recommended during the first months of life and discontinued as soon as possible [12]. In 2005, around 700,000 children under 15 contracted HIV, mainly through MTCT, with 630,000 of these infections occurring in Africa [32]. Of the estimated 2.3 million (1.7-3.5 million) children currently living with HIV, 2 million (almost 90%) live in sub-Saharan Africa [12]. Prevention strategies are well known in developed countries, however, recent epidemiological and behavioral studies in Europe and North America have suggested that a substantial minority

of young people continue to engage in high-risk practices and that despite HIV/AIDS knowledge, young people underestimate their own risk of becoming infected with HIV [33]. However, transmission of HIV between intravenous drug users has clearly decreased, and HIV transmission by blood transfusion has become quite rare in developed countries.

1.1.3. Clinical features of HIV infection

HIV primarily infects vital cells in the human immune system such as helper T cells (specifically CD4+ T cells), macrophages and dendritic cells. HIV infection leads to low levels of CD4+ T cells through three main mechanisms: firstly, direct viral killing of infected cells; secondly, increased rates of apoptosis in infected cells; and thirdly, killing of infected CD4+ T cells by CD8 cytotoxic lymphocytes that recognize infected cells. When CD4+ T cell numbers decline below a critical level, cell-mediated immunity is lost, and the body becomes progressively more susceptible to opportunistic infections [34].

The clinical features of HIV infection vary depending on the stage of disease. In the acute and early stages of the disease, symptoms resulting from immune hyperactivation predominate; however, with the depletion of CD4+ T-cells that characterizes disease progression, an evident cellular immunodeficiency and its many complications develop. The CD4+ T-cell count is an excellent marker of the degree of immunodeficiency and determines the immediate risk of opportunistic infections and other AIDS-related complications [35]. The plasma viral load is the most valuable prognosticator of HIV infection that is currently available, because it not only predicts disease progression independent of the CD4+ T-cell count, but also predicts the trajectory of the CD4+ T-cell count [36]. Although more than half of primary HIV infections are accompanied by symptoms, most of these infections are not recognized in the acute setting. The differential diagnosis of primary HIV infection is extremely broad, given the variable presentations associated with the syndrome. Fever, rash, pharyngitis and lymphadenopathy are common during HIV infection, as they are in many "flu-like" and "mononucleosis-like" illness [37]. Myalgias, arthralgias, diarrhea, nausea, vomiting, headache, hepatosplenomegaly, weight loss, thrush and neurological symptoms also may occur. Among patients who present symptoms during primary HIV infection, the mean duration of symptoms is 3 weeks [38].

The natural history of HIV infection includes a long period (i.e., approximately 10 years) of clinical latency between the time of primary infection and the development of symptoms indicative of advanced immunodeficiency. Even during the period of clinical latency, many patients experience fatigue as well as generalized lymphadenophaty. Syndromes

indicative of depressed cell-mediated immunity in HIV infection that are not AIDS defining generally begin to appear as the CD4+ T-cell count falls below 500 cells/μL. Examples include oropharyngeal and recurrent vulvovaginal candidiasis, bacillary angiomatosis, multidermatomal herpes zoster, listeriosis, pelvic inflammatory disease, oral hairy leukoplakia associated with EBV, cervical dysplasia, unexplained fever or diarrhea lasting more than 1 month, idiopathic thrombocytopenic purpura and peripheral neuropathy [39].

As the CD4+ T-cell count decreases below the level of approximately 200 cells/μL, the loss of integrity of cell-mediated immune responses allows ubiquitous environmental organisms with limited virulence (e.g., *Pneumocystis carinii* and *Mycobacterium avium*) to become life-threatening pathogens. HIV-associated diseases indicative of severe impairment of cell-mediated immunity, but not included in the current case definition of AIDS, include chronic microsporidiosis, gastrointestinal infection with *Cyclospora cayetanensis*, disseminated *Penicillium marneffei* infection, cerebral or disseminated *Trypanosoma cruzi* infection, relapsing or chronic visceral leishmaniasis, anal carcinoma, EBV-positive cases of leiosarcoma, leiomyosarcoma and Hodgkin's disease [40]. All of these associated illnesses are now known to be caused by infection with or reactivation of opportunistic organisms. In this regard, human papillomaviruses EBV and human herpesvirus-8 appear to play causative roles in the AIDS-defining malignancies, cervical, cancer, lymphomas and Kaposi's sarcoma, respectively. The degree to which an infection or reactivation syndrome can be considered opportunistic can be inferred by the CD4+ T-cell counts at which it occurs [35]. The incidence of AIDS-defining illnesses in industrialized nations has decreased sharply since 1996. In the USA, the AIDS mortality rate decreased 68% between 1996 and 1999 [41]. These decreases in AIDS-associated mortality can be attributed to a number of factors, including continued progress in prophylaxis against opportunistic infections, improved access to health care and improved antiretroviral therapies.

Several host and virologic factors determine the extremely variable rates of disease progression that are observed among HIV-infected individuals. A small percentage of HIV-infected individuals experience no evidence of disease progression over a prolonged period. Although a minority of cases of long-term non-progressive HIV infection may be associated with attenuated strains of HIV [42], most data suggest that viral attenuation is rare among long-term non-progressors, and that host factors play a crucial role in determining the state of non-progression [43].

Host genetic factors strongly influence the rate of disease progression in HIV infection; genetic polymorphisms in the HIV coreceptors and their ligands may also significantly affect the course of the progression [44]. Genetic polymorphisms in the chemokine-chemokine receptor axis genes represent only one of a number of possible mechanisms responsible for resistance to HIV infection or delayed disease progression [45].

Also quantitative as well as qualitative aspects of the host immune response that is mobilized against HIV are important determinants of the rate of disease progression as well [46].

At last the morphologic abnormalities of lymphoid tissue associated with HIV disease progression are important determinants of immunodeficiency [47].

1.1.4. Organization of viral genome

The genome and proteins of HIV have been the subject of extensive research since the discovery of the virus in 1983. The new and unusual syndrome was characterized by generalized lymphadenopathy and opportunistic infections, with a marked depletion of the CD4+ T-lymphocyte subset in the peripheral blood. The disease was first brought to the attention of the general medical community in June 1981, when the Centers for Disease Control described five California men with severe immunodeficiency in the *Morbidity and Mortality Weekly Report* [48]. This notification was followed by several reports describing male homosexual and intravenous drug users with impaired immune systems and T-lymphocytes that responded poorly to antigen and mitogen stimulation in functional assays. Within several months, it became clear that a similar immunodeficiency disease was also affecting other groups, including hemophiliacs, blood transfusion recipients, sexual partners and/or children of different risk groups, and that the new disease was transmitted by a novel pathogen in contaminated blood or following sexual intercourse with an affected individual. In 1983, scientists at the Pasteur Institute recovered an agent from the lymph nodes of an asymptomatic individual who presented a generalized lymphadenopathy of unknown origin [49].

The new retrovirus, associated with AIDS in the USA, Europe and central Africa and exhibiting morphologic and genetic characteristics typical of the *Lentivirus* genus, was named human immunodeficiency virus, or HIV [50] (and subsequently HIV-1). In 1986, a related, but immunologically distinct and less pathogenic human retrovirus (now called HIV-2), was recovered from individuals residing in several west African countries such as Senegal, Ivory, Coast and Guinea-Bissau [51].

One of the first features the Pasteur Institute group noted about HIV-1 was its particle-associated reverse transcriptase, a property that placed the new agent in the retrovirus family. This was consistent with EM analyses of particles released from infected cell cultures, which revealed 100- to 120-μm enveloped virions, similar in size and morphology to other retroviruses. The mature HIV-1 particles contained a cone-shaped cylindrical core or nucleoid reminiscent similar to that of visna virus [52]. The cloning and sequencing of proviral DNA indicated that HIV-1 not only possessed a genomic organization related to other replication-competent retroviruses, but placed it, taxonomically, in the *Lentivirus* genus [53]. As their name suggests, *Lentiviruses* were known to cause slow, unremitting disease in sheep, horses and cattle, and to target various lineages of hematopoietic cells, particularly lymphocytes and differentiated macrophages. It was originally thought that HIV-1, like other replication competent retroviruses, would be genetically homogeneous. However, as proviral DNAs corresponding to HIV-1 isolates from Europe, North America and Africa became available and were compared, their extensive genetic heterogeneity became apparent [54]. No two HIV-1 isolates were identical. When subjected to nucleotide sequence analysis, even HIV-1 samples recovered from a single individual exhibited significant heterology [55]. The term *quasi-species* was subsequently coined to describe the heterogeneous and changing population of virus present in an HIV-1 infected individual [56]. Distinct genetic subtypes or clades of HIV-1 were consequently defined and classified into three groups: M (major), O (outlier) and N (non-M or O). The M group of HIV-1, which includes over 95% of the global virus isolates, consists of at least eight discrete clades (A, B, C, D, E, F, G, H, and J), based on the sequence of complete viral genome [57]. Members of HIV-1 group O have been recovered from individuals living in Cameroon, Gabon and Equatorial Guinea; their genomes share less than 50% identity in nucleotide sequence with group M viruses [58]. The more recently discovered group N HIV-1 strains have been identified in infected Cameroonians, fail to react serologically in standard whole-virus enzyme-linked immunosorbent assay (ELISA) [59].

HIV is different in structure from other retroviruses. It is around 120 nm in diameter (120 billionths of a meter; around 60 times smaller than a red blood cell) and roughly spherical.

HIV-1 is composed of two copies of single-stranded RNA enclosed by a conical capsid comprising the viral protein p24, typical of *Lentiviruses*. The RNA component is 9749 nucleotides long. This is in turn surrounded by a plasma membrane of host-cell origin. The single-strand RNA is tightly bound to the nucleocapsid proteins, p7 and enzymes that are indispensable for the development of the virion, such as reverse transcriptase and integrase.

The nucleocapsid (p7 and p6) associates with the genomic RNA (one molecule per hexamer) and protects the RNA from digestion by nucleases. A matrix composed of an association of the viral protein p17 surrounds the capsid, ensuring the integrity of the virion particle. Also enclosed within the virion particle are Vif, Vpr, Nef, p7 and viral protease. The envelope is formed when the capsid buds from the host cell, taking some of the host-cell membrane with it. The envelope includes the glycoproteins gp120 and gp41 (figure 2).
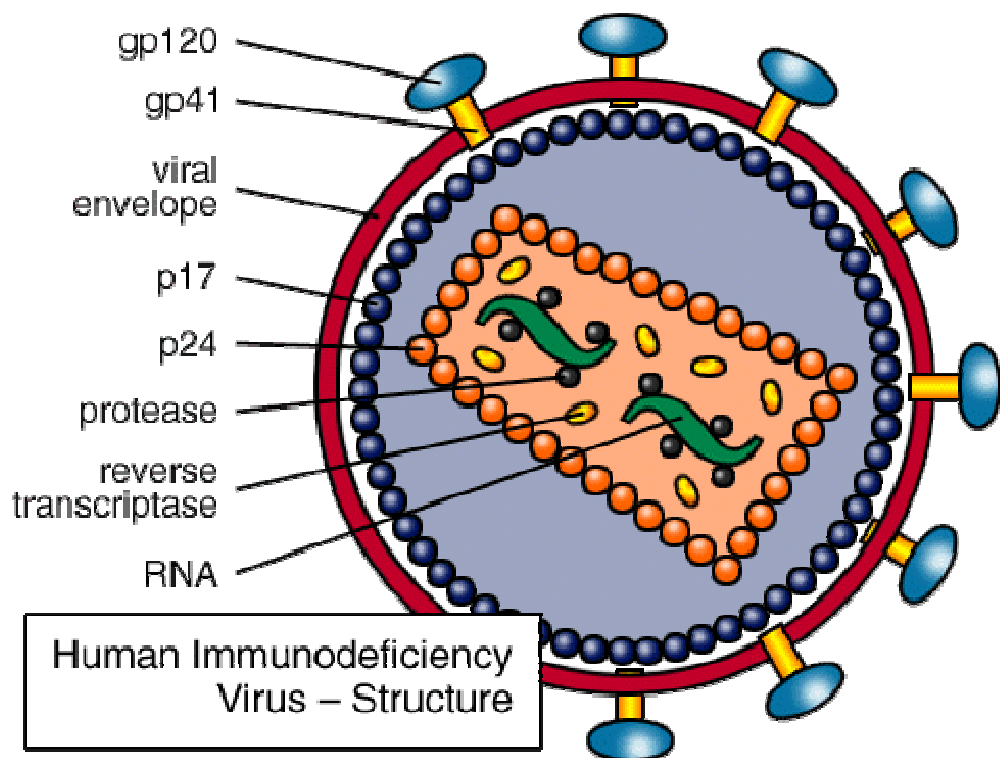
Figure 2. HIV-1 structure. In this figure are represented the viral envelope with its transmembrane and fusion glycoproteins (gp120 and gp41), the matrix (p17) and the core (p24) proteins, the viral nucleic acid (RNA) and the enzymes involved in different cleavage events (PR) and in reverse transcription step (RT) [60].

HIV has several major genes coding for structural proteins that are found in all retroviruses, and several nonstructural ("accessory") genes that are unique to HIV. The Gag gene provides the basic physical infrastructure of the virus, and Pol provides the basic mechanism by which retroviruses reproduce, while the others help HIV to enter the host cell and enhance its reproduction. Though they may be altered by mutation, all of these genes except Tev exist in all known variants of HIV (figure 3) [61].

Gag (for "Group-specific Antigen"): encodes for p24, the viral capsid, p6 and p7, the nucleocapsid proteins, and p17, a matrix protein.

15

Pol: encodes for the specific viral enzymes, reverse transcriptase, integrase, and protease. The last one cleaves the proteins derived from Gag and Pol into functional proteins.

Env (for "Envelope"): encodes for the polyproteic precursor gp160, that, after glycosylation, is cleaved into gp120 and gp41, proteins embedded in the viral envelope which enable the virus to attach to and fuse with target cells.

Tat (for "Transactivation gene"): after its synthesis, it enters in the cellular nucleus and acts as transactivating factor of proviral genome transcription linking the Tar sequence of the originating mRNAs.

Rev (for "Regulatory of Virus Gene"): regulates the sequence in the production of specific-virus RNAs, protecting them from splicing process.

Nef (for "Negative Factor Gene"): through its binding to CD4 and HLA-I molecules, it facilitates their translocation and degradation.

Vif (for "Viral Infectivity Factor Gene"): is important during the core assemblage and enhances viral infectivity.

Vpr (for "Viral Protein R Gene"): favors the intranuclear shift of the viral nucleoproteic complex for the subsequent integration of the provirus in cellular genome.

Vpu (for "Virus Protein U"): through its binding to CD4 molecules, it facilitates the translocation of the product of Env gene towards the cellular membrane.

Tev: this gene is only present in few HIV-1 isolates. It results from the fusion of parts of Tat, Env, and Rev genes, and it encodes for a protein with some of the properties of Tat, but little or none of the properties of Rev.
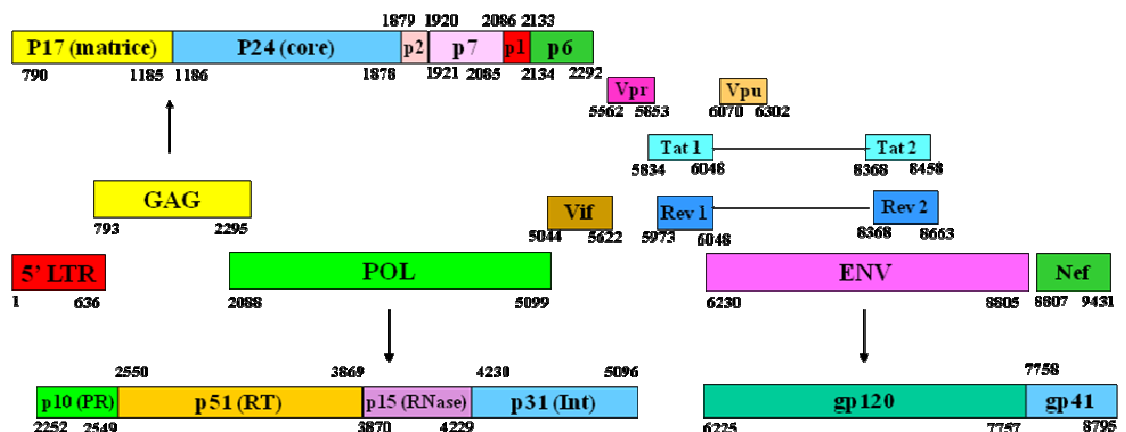

Figure 3. HIV-1 encoded proteins. The location of the HIV-1 genes, the relative positions and the processed mature viral proteins are indicated [62].

The HIV-1 genome is unique in having a novel central polypurine tract (PPT) in addition to a traditional U3-proximal PPT, both of which are used for the synthesis of plus-strand viral DNA [63]. The viral RNA folds into two additional complex structures (TAR and RRE), which are involved in RNA synthesis and transport, respectively. The TAR has also been reported to be required for the efficient initiation of reverse transcription [64].

1.1.5. The HIV-1 replication cycle

The main cellular targets for HIV-1 are the CD4+ T-helper/inducer subset of lymphocytes, CD4+ cells of macrophage lineage and some populations of dendritic cells.

The HIV-1 replicative cycle begins with adsorption of virus particles to CD4 molecules on the surface of susceptible cells (figure 4). While binding of virions to CD4 is essential for HIV infectivity, their subsequent interaction with a coreceptor, one of which are now known to be the seven membrane-spanning CC or CXC families of chemokine receptors, is required for membrane fusion and entry [65; 66]. Although multiple chemokine receptors exhibit activity in fusion/entry assays with HIV-1 gp120, the two most important are CXCR4 and CCR5.

Unlike other enveloped viruses, which use receptor-mediated endocytosis to enter cells, HIV-1 and most other retroviruses directly fuse with the plasma membrane of a susceptible cell. Following entry, subviral particles are partially uncoated in the cytoplasm and initiate the reverse transcription of their viral RNA genomes. The partially double-stranded DNA reverse transcription product is transported through the cytoplasm and into the nucleus as a component of a nucleoprotein-preintegration complex (PIC) containing a subset of the Gag and Pol proteins [67; 68]. This process very likely depends on host factors that actively transport the PICs from the plasma membrane to the nuclear pore. These "early" steps in the retrovirus replication cycle require the participation of multiple cellular factors for their successful completion. Following the import of the PIC into the nucleus, full length linear copies of the reverse transcript are integrated into the chromosomal DNA of the infected cell, a step required for the efficient viral RNA synthesis and infectious particle production [69]. The *Lentiviruses* are unique among Retroviruses in generating PICs that are actively transported by the nuclear import machinery into the interphase nucleus of non-dividing cells arrested in the G_1 phase of the cell cycle. In activated T lymphocytes, integrated copies of HIV DNA serve as templates for RNA polymerase II (Pol II)-directed viral RNA synthesis. The coordinated interaction of the HIV-encoded Tat protein and some cellular transcriptional transactivator proteins with the Pol II transcriptional apparatus ensures the

production of high levels of viral RNA. Unspliced or partially spliced HIV transcripts are exported from the nucleus to the cytoplasm by a unique transport mechanism mediated by the shuttling virus-encoded Rev protein. The subsequent translation of the gp160 Env precursor occurs in the endoplasmic reticulum, while the Gag and Gag-Pol polyproteins are synthesized on free cytoplasmic ribosomes and transported to the plasma membrane. The Gag and Gag-Pol polyproteins, in association with dimer of genomic RNA, condense at the plasma membrane to form an electron dense "bud", that gives rise to a spherical immature particle containing the mature gp41 transmembrane and the gp120 surface envelope glycoproteins. The proteolytic process carried out by the HIV PR during or immediately after particle release generates the cone-shaped nucleoid characteristic of mature HIV virions.
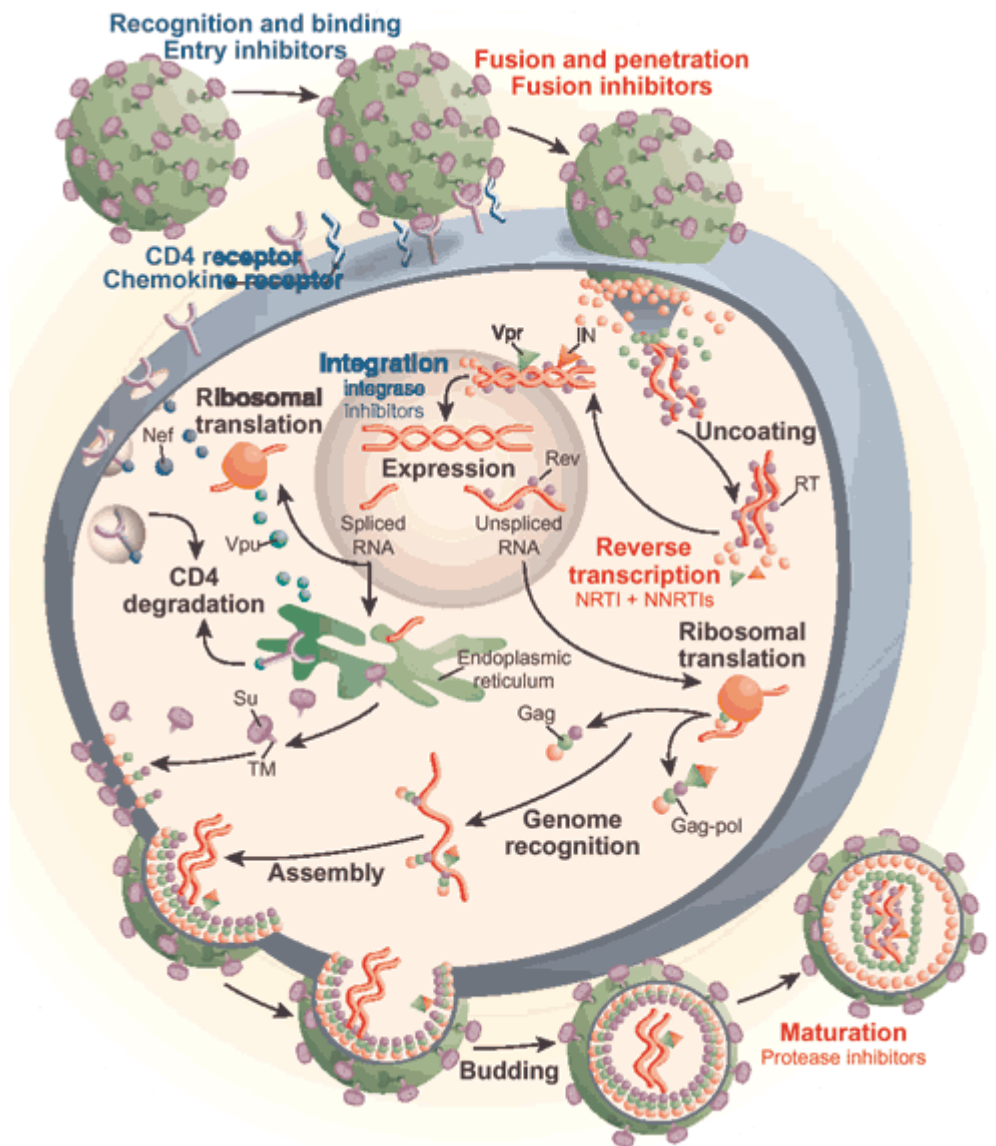
Figure 4. The HIV-1 replication cycle [70].

1.2 The reverse transcriptase (RT)

Retroviral replication requires the conversion of single-stranded genomic RNA into double-stranded proviral DNA that is later integrated into the host chromosome. Such a process is a central step in the life cycle of HIV-1 and is mediated by the virus encoded reverse transcriptase (RT). HIV-1 RT is a multifunctional enzyme that possesses RNA-dependent and DNA-dependent DNA polymerase activities as well as an RNase H activity that specifically degrades the RNA strand of RNA/DNA hybrids [71]. Its structure is a heterodimer consisting of two subunits, termed p66 and p51. Based on its crystal structure, the enzyme has been divided into several regions, as shown in figure 5. These include *fingers*, *thumb*, *palm* and *connection* subdomains, that are present in both p66 and p51, as well as the RNase H domain, found at the carboxy-terminus of p66 [72].
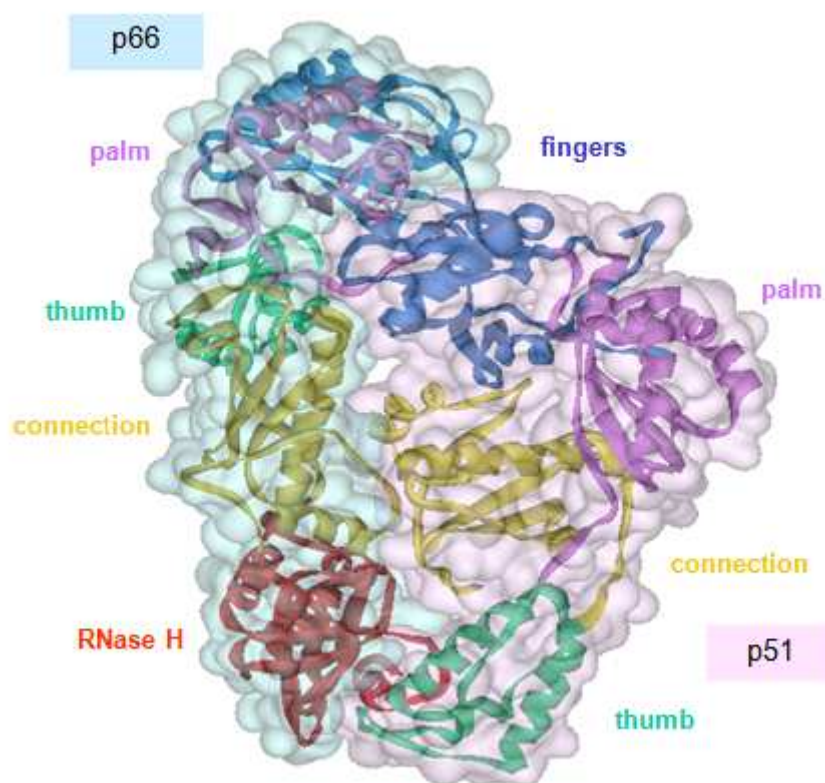
Figure 5. Structure of HIV-1 reverse transcriptase. The p66 and p51 subunits are showed in cyan and pink, respectively; the four subdomains *palm*, *thumb*, *fingers* and *connection* and the RNase H domain are colored in purple, green, blue, gold and red, respectively [73].

The larger subunit p66 contains both polymerase and RNase H activities, while the smaller subunit p51 lacks these functions, in context of the heterodimer [74]. However, both the p66 and p51 monomers are functionally inactive when dissociated from each other. Despite the evidence that p51 shares an identical amino acid sequence with the N-terminal portion of p66,

the two subunits assume different global folding patterns in the formation of the asymmetric dimer. Structural determination through X-ray crystallography has revealed that the p66 subunit of HIV-1 RT has its polymerase domain in an "open" conformation, with its subdomains forming a large cleft which well accommodates DNA. In contrast, the p51 subunit assumes a compact folded conformation that causes the active site residues in this subunit to be buried and therefore non-functional. The two subunits interact mainly via their *connection* subdomains; additional contacts between the *thumb* subdomain of p51 and RNase H subdomain of p66 are also crucial [75]. The RNase H domain of HIV-1 RT has been crystallized in free form and has a structure that is similar to that of *E. Coli* RNase H [72]. However the isolated form of RNase H is not active; restoration of its activity requires the presence of p51, indicating that additional RT residues contribute to RNase function. In fact, subunit-selective amino acid substitutions demonstrated that residues located in the *thumb* of p51 and p66 are important for RNase H cleavage [76].

1.2.1. RT polymerase site

All DNA polymerases, including RT, have an absolute requirement for the 3′-OH group of primers during DNA polymerization. The DNA duplex binds along a groove, about 60 Å in length, stretching from the polymerase active site to the RNase H active site (figure 6).
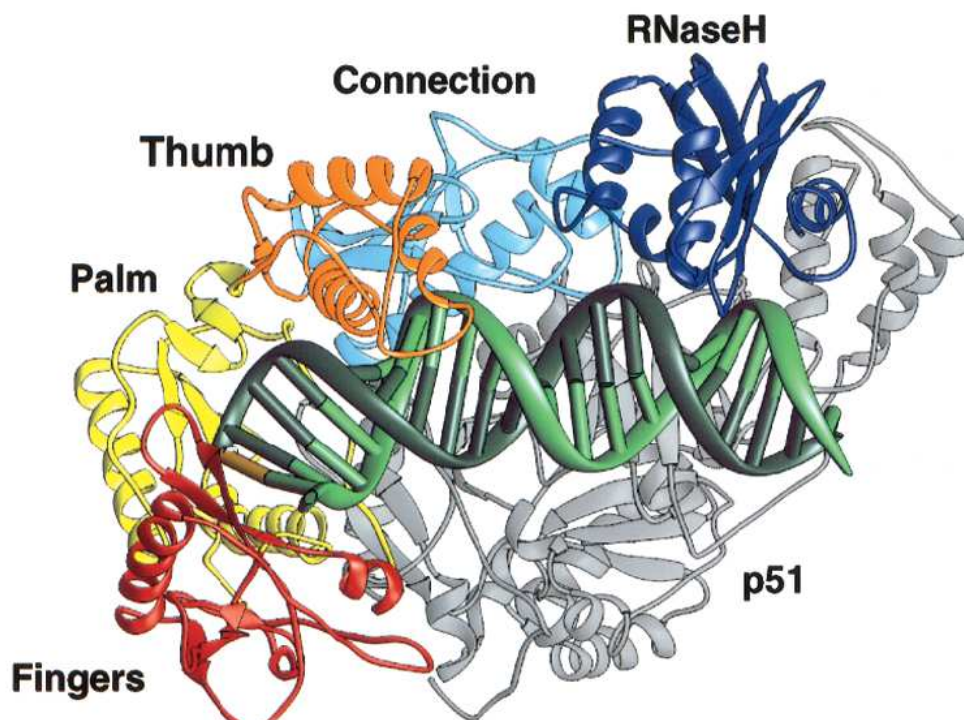

Figure 6. The structure of the catalytic site of HIV-1 RT. The p51 subunit is shown in grey, while in p66 the *thumb* subdomain is represented in orange, *palm* in yellow, *fingers* in red, *connection* in cyan and RNase H domain is shown in blue [77].

The DNA has an A-like structure in the region of the polymerase active site, but between base pair n-5 (the site of the covalent cross-link to the *thumb*) and n-11 (figure 7), the structure becomes gradually more B-like, but the major groove remains unusually deep for B DNA and the base pairs retain some tilt with respect to the helix axis. The 5′-template overhang bends away from the duplex and extends across the face of the *fingers* subdomain. The DNA duplex contacts mainly the p66 subunit of the RT heterodimer; the only direct interactions with bases occur in the minor groove. There are van der Waals contacts with Pro157 and Met184 and with Ile94, and hydrogen bonds with Tyr183 [77].
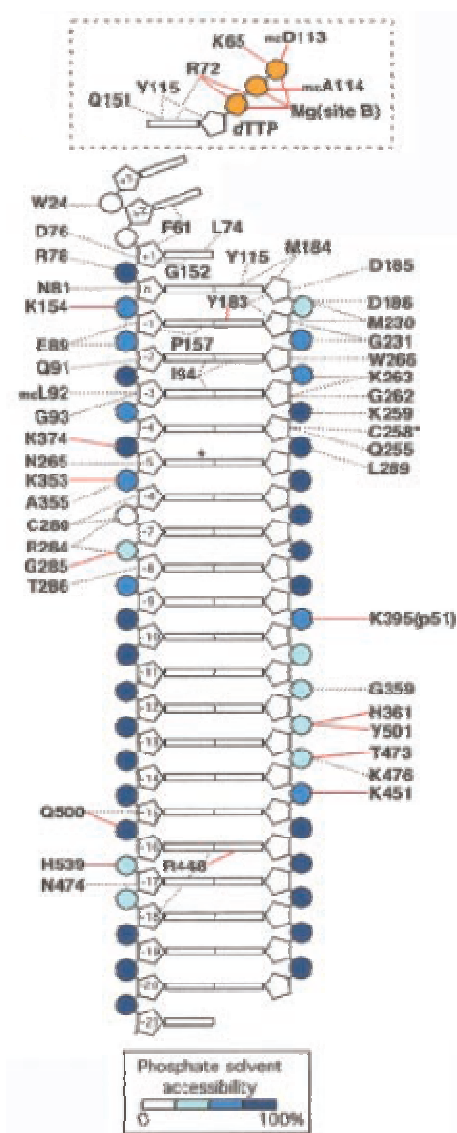
Figure 7. Interactions between RT and its substrates. Assigned hydrogen bonds are showed as solid red lines; other interactions are dashed black lines. Main-chain contacts are abbreviated "mc". Asterisks indicate the position of the disulfide cross-link. The numbering scheme for the template-primer is shown along the sugar phosphate backbone. The base pair containing the 39-primer terminus is labeled n. Bases in the 59-template extension are numbered n11 to n13, where n11 is the templating nucleotide. Base pairs in the duplex are numbered from n to n-20, and there is a single-base overhang at n-21 at the 39 end of the template strand [77].

The catalytic triad Asp110, Asp185 and Asp186 is located near to the 3′ primer terminus. The side chains of these residues are carboxy groups able to induce the DNA nucleophilic attack during the polymerization step.

Between nucleotides n and n-3, the sugar-phosphate backbone of the template strand contacts residues in the *fingers* and *palm* through van der Waals interactions. Between nucleotides n-1 and n-4 of the primer strand and n-5 and n-8 of the template strand, the minor groove faces the *thumb*. The primer strand contacts the loop between *palm* and *thumb* (Met230 and Gly231: the "primer grip"), and both strands interact with residues in helices H and I (figure 8).
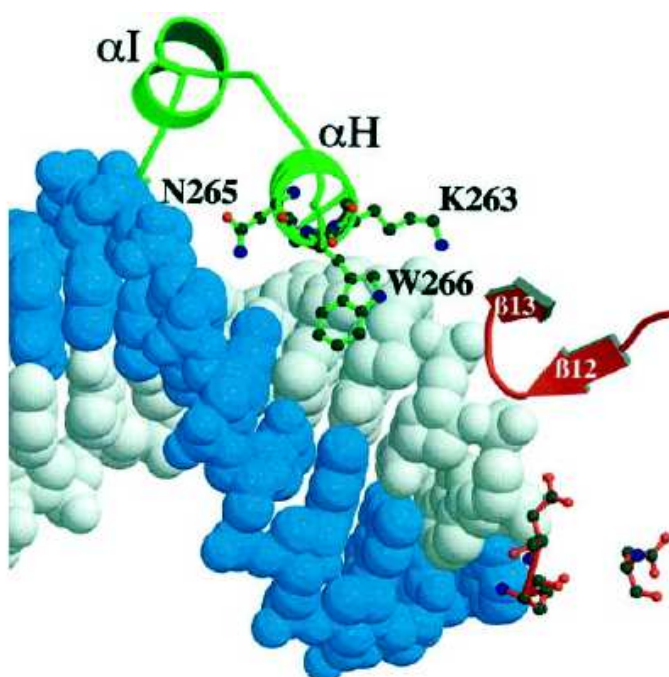
Figure 8. "Helix clamp" area constituted by the two antiparallel helices, αH (residues 253-270) and αI (residues 277-287) [77].

Near the RNase H active site, the position of the template backbone appears to be close to the expected catalytic configuration for an RNA strand, even though a DNA duplex is not the correct substrate for the ribonuclease activity. In particular, there is a Mg ion bound by Asp443, Asp549, and phosphate n-17, which lies between Gln478 and His539, essential for activity in homologous enzymes. Also Arg448 inserts into the minor groove in the same region. Binding of the template-primer and dNTP induces significant conformational changes in parts of p66. The outer part of the *fingers* domain shifts significantly with respect to its base, bringing the "fingertips" closer to the *palm*. The *palm* moves slightly toward the core and the tyrosines 181 and 188, which rotate in response to NNRTI binding, present

unperturbed conformations with respect to the unbound RT. The closure of the outer part of *fingers* subdomain causes a distortion of the hydrophobic core, allowing some residues to contact with the incoming nucleoside triphosphate (dNTP). Its triphosphate moiety is coordinated by Lys65, Arg72, main-chain-NH groups of residues 113 and 114 and two metal ions (Mg). The guanidinium group of Arg72, which lies flat against the dNTP base, gives hydrogen bonds to the α-phosphate and the ε-amino group of Lys 65 donates hydrogen bonds to the γ-phosphate. Both of these side chains move into position as a result of the *fingers* closure. The 3′-OH of the dTTP projects into a small pocket lined by the side chains of Asp113, Tyr115, Phe116 and Gln151, and the peptide backbone between 113 and 115; moreover it accepts a hydrogen bond from the main-chain-NH of Tyr115 (figure 9) [77].
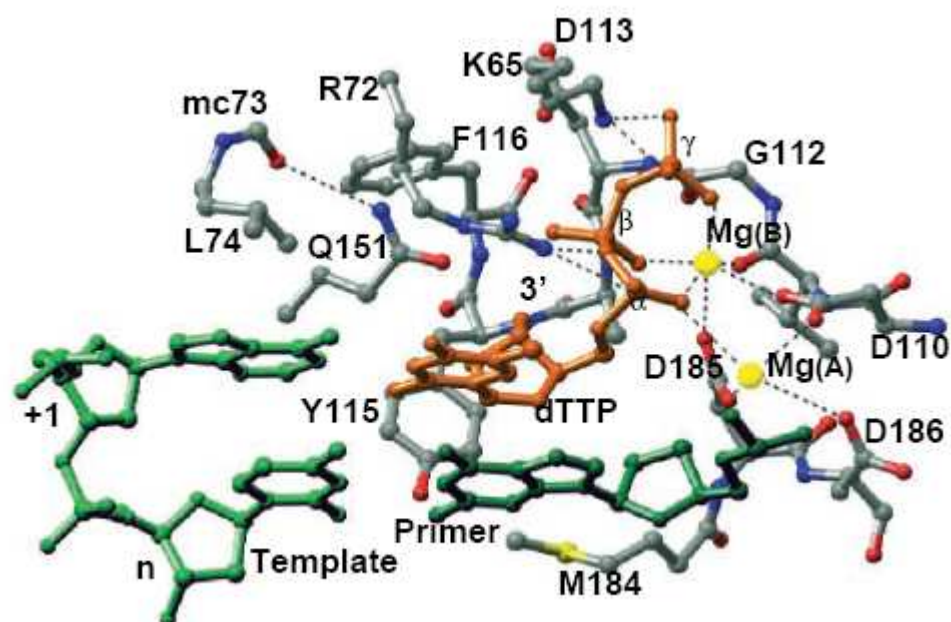
Figure 9. Stereo view of polymerase active site. Template, primer strand and dTTP are shown in light green, dark green and gold, respectively; Mg ions are represented as yellow spheres; H-bonds and metal ligand interactions are indicated with dashed lines [77].

Part of the catalytic site, the YMMD motif (Tyr183-Asp186) result crucial in polymerase activity, as well as the single substitution Met184Pro is related to a drastic decrease of viral replication.

Several mutagenesis studies showed that some residues of the primer grip region (residues 226-235) have a pivotal role both in catalytic activity and in heterodimer association [78]. Analyzing the RT conformation in presence of the primer, Trp266 was found to be essential in the interactions with the nucleic acid. In fact the mutation Trp266Glu causes a repulsion

with the phosphate of the nucleic acid due to the negative charge present on the glutamate side chain.

Even if it is not directly involved in catalytic activity, p51 subunit resulted crucial in dimer stabilization and RT functions. Its *thumb* subdomain is responsible to guarantee RT activities through fundamental interactions with the RNAse H domain (figure 10). Site-directed mutagenesis analysis confirmed such a data, showing that Asn255Asp, Asn265Asp and Cys280Ser mutations are related to an incorrect polymerization process, a reduced DNA affinity and a decreased RNase H activity, respectively [79].
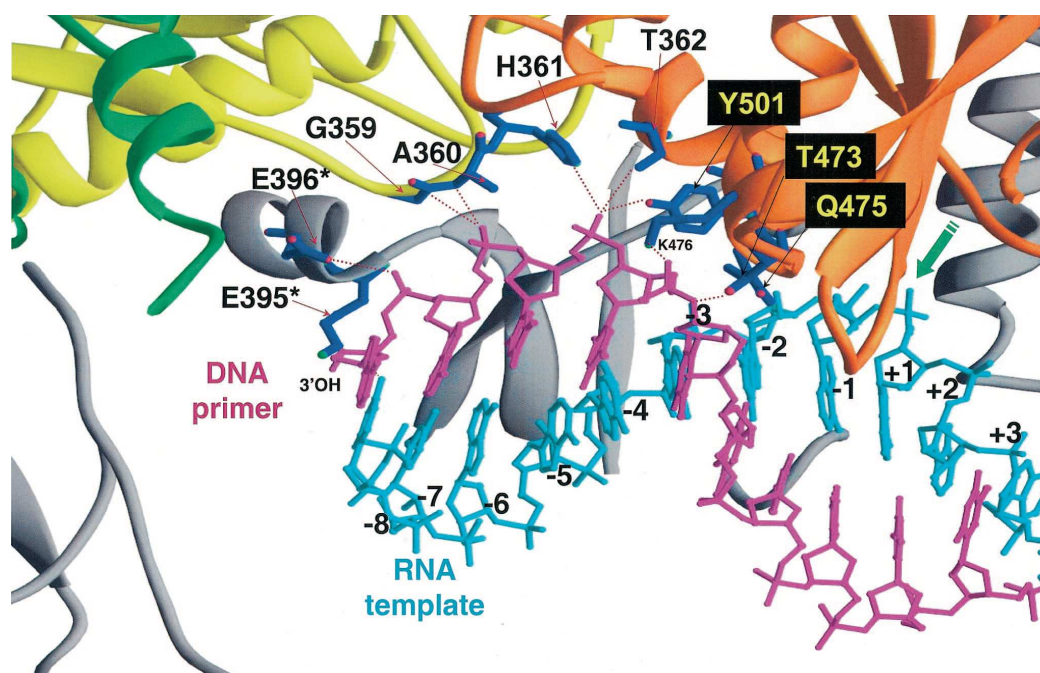
Figure 10. Diagram of the structure of HIV-1 RT in complex with an RNA-DNA template-primer. The figure shows the portion of the structure near the RNase H active site; the contacts between the nucleic acid and the RNase H domain and the *connection* domain are indicated. The RNA template strand is shown in turquoise; the DNA primer strand is shown in purple. The scissile phosphate is designated with a green arrow pointing to the phosphate. The RNase H domain is represented in orange; the *connection* subdomain is shown in yellow, while the *thumb* in green. The p51 subunit is in gray. Amino acids in the p51 subunit that contact the nucleic acids are marked with an asterisk next to the position number. Amino acids contacting the nucleic acid are labeled. Contacts between individual amino acids and the nucleic acid are shown in red. The RNA strand is numbered relative to the site of cleavage [79].

1.2.2. Heterodimer interface area

The RT heterodimer represents the biologically relevant form of the enzyme; the monomeric subunits show only low catalytic activity. Structural analysis reveals three major contacts between p66 and p51 subunits, with most of the interactions surfaces being largely hydrophobic [80]. The three contacts comprise an extensive dimer interface that includes the *fingers* subdomain of p51 with the *palm* of p66, the *connection* subdomains of both subunits and the *thumb* subdomain of p51 with the RNase H domain of p66.

Two highly conserved residues of the *fingers* subdomain of p51, Trp24 and Phe61, strictly interact with several amino acids of the *connection* subdomain, including Glu399, Trp402, Thr403, Trp406 and Trp410 (figure 11). Even if they are not located at the interface region, the specific mutagenesis of both residues of p51 subunit dramatically reduces the interactions between p51 and p66 subunits, decreasing the heterodimer stability [81].
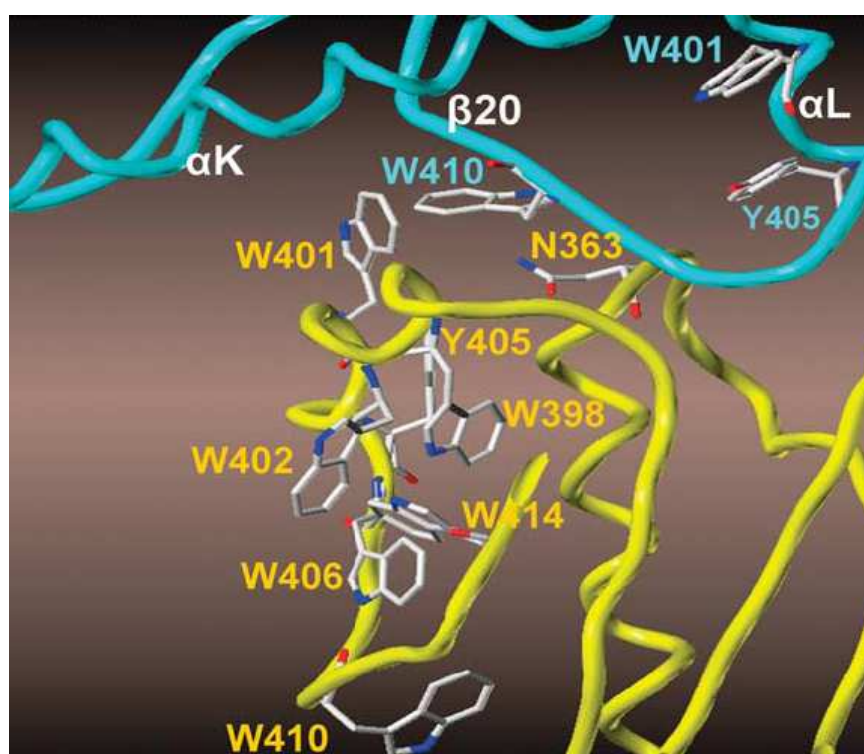
Figure 11. Representation of the most important residues of both RT subunits involved in the interface interactions stabilizing the heterodimer enzyme [81].

The mutations Trp402Arg and Trp406Arg, located in the *connection* subdomain that contacts the p51 *connection* domain in the heterodimer, were found to reduce RT stability. In fact the appearance of a basic residue in both codons 402 and 406 suggests a charge interaction with an acid residue in p51 or, alternatively, an increase in electrostatic potential between the surfaces at the *connection* domain interface [81].

Another critical region resulted essential for RT proper dimerization contains a leucine hepta-repeat motif from leucine 282 to leucine 310 [82]. This area showed an important role in protein-protein interactions required for dimerization [83]. After site-directed mutagenesis of these leucine residues, L289K of p66 resulted unable to dimerize with itself and wild-type or L289K of p51.

Also the Leu234Ala substitution was found to inhibit RT dimerization due to the capability of this residue, not placed at the dimer interface, to indirectly affect contacts between Pro95 in the *palm* of p66 with residues in the β7-β8 loop of p51.

The loop between β-strands 7 and 8 in HIV-1 RT is highly conserved in amino acid sequence. This sequence, SINNET (residues 134–139), is relatively conserved in drug-naïve HIV-1 clones/isolates with the exception of Ile135, which is sometimes substituted by threonine or valine.

Some structural analysis indicated that the Asn136 residue in p51 is placed centrally in the β7-β8 loop at the p51-p66 dimer interface. There appears to be a key inter-subunit contact between the side-chain of Asn136 in p51 and the backbone of residues Pro95 and His96 at the carboxy-terminal end of the β5-strand in p66, which is consistently at a distance of less than 3.0 Å (figure 12). In several structures, the prolyl ring and backbone of Pro95 are close enough (3.9 Å on average) to interact with the side-chain of Trp229 in the polymerase primer grip [84].
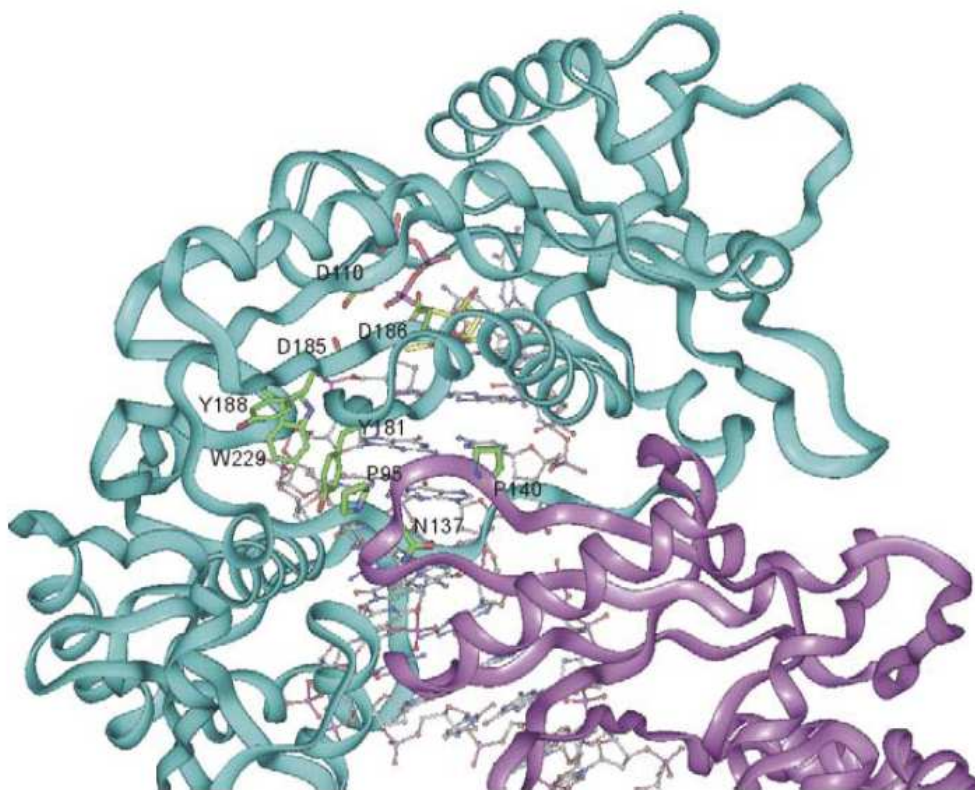
Figure 12. Detail of the HIV-1 RT subunit interface showing as sticks the side chains of Asn137 and Pro140 in the β7-β8 loop of p51 (pink ribbon) and of Pro95 in the p66 subunit (cyan ribbon). Carbon atoms of the DNA template-primer are colored grey whereas those of the incoming nucleotide triphosphate are colored yellow [84].

Alanine substitution of Asn136 in the β7–β8 loop of p51 destabilized the heterodimeric RT, rendering it sensitive to degradation by the viral PR, and significantly reduced viral infectivity [81]. Also the mutations of four amino acids on the β7-β8 loop with alanine reduced the DNA binding ability of the enzyme and similar results were obtained when four amino acids were deleted from the loop [85].

1.2.3. NNRTI binding pocket region

The polymerase activity of the catalytic triad can be considerably influenced by conformational changes occurring in an allosteric site known as NNRTI binding pocket (NNRTI-BP). The NNRTI-BP is situated between the β6-β10-β9 and β12-β13-β14 sheets in the *palm* subdomain of the p66 subunit, approximately 10Å from the RT DNA polymerase aspartic acid catalytic triad [86]. The NNRTI-BP is predominantly hydrophobic in nature with substantial aromatic character (Tyr181, Tyr188, Phe227, Trp229, and Tyr232), but also contains several hydrophilic residues (Lys101, Lys103, Ser105, Asp192, and Glu224 of the p66 subunit and Glu138 of the β7-β8 loop of the p51 subunit). A probable solvent accessible entrance to the NNRTI-BP is located at the p66/p51 heterodimer interface, ringed by residues

Leu100, Lys101, Lys103, Val179, and Tyr181 of the p66 subunit and Glu138 of the p51 subunit (figure 13) [86].
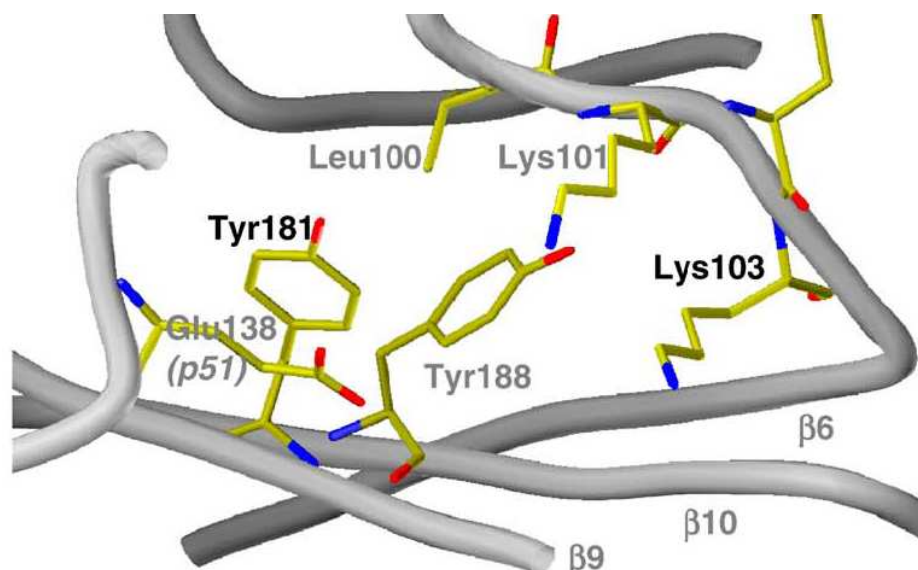

Figure 13. The NNRTI-BP of wild-type RT in the unbound conformation [87].

Interestingly, in the absence of ligand, the side chains of Tyr181 and Tyr188 of p66 point into the hydrophobic core and, as a consequence, the NNRTI-BP does not exist in the free enzyme [88]. Instead, a surface depression is present at a location which is equivalent to the putative entrance of the pocket. NNRTI-binding to HIV-1 RT causes the side chains of both Tyr181 and Tyr188 to rotate away from their positions in the hydrophobic core thereby creating a space to accommodate the ligand [88]. The major difference in the location of the secondary structural elements that form the pocket between the structures of HIV-1 RT with and without NNRTI is a differential twisting (about 30°) of the β12-β13-β14 sheet, which results in an expansion of the NNRTI-BP, as reported in figure 14.
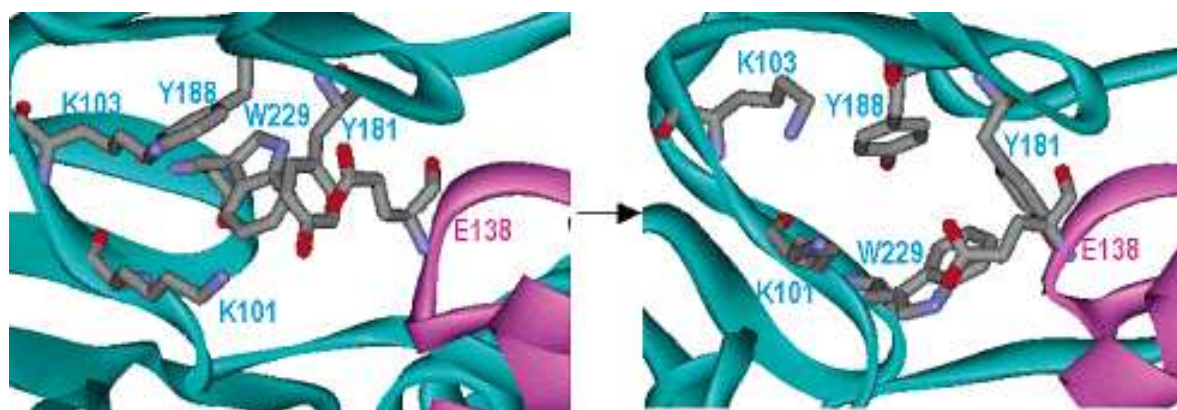

Figure 14. Closed (left) and open (right) NNRTI-BP conformations. The p66 and p51 subunits are shown as cartoons and colored in cyan and magenta, respectively [88].

Comparison of the available structures of HIV-1 RT indicates that the binding of an NNRTI causes both short-range and long-range distortions of the HIV-1 RT structure. The short-range distortions include the conformational changes of the amino acids and/or structural elements that form the NNRTI-BP, such as the re-orientation of the side chains of Y181 and Y188, and the displacement of the β12-β13-β14 sheet. The long-range distortions involve a hinge-bending movement of the p66 *thumb* subdomain that results in the displacements of the p66 *connection*, the RNase H domain, and the p51 subunit relative to the polymerase active site [86].

An interesting aspect of the NNRTI-BP is that it is situated close to the subunit-subunit interface, and the entrance of the pocket is composed of residues from the p66 (Leu100, Lys101, Lys103, Val179 and Tyr181) and p51 (Glu138) subunits that also form part of the RT dimer interface.

Several studies have evaluated the ability of NNRTI to impact on the dimeric structure of HIV-1 RT and surprising results have been obtained in this regard [89].

K103N is a clinically relevant mutation that hampers the binding of and confers resistance to many classes of structurally diverse NNRTIs inhibitors. Some crystallographic studies with the K103N mutant enzyme have revealed a notable absence of specific interactions involving the Asn103 side chain. In the presence of the K103N mutation, the formation of a good hydrogen bond between the phenol oxygen of Tyr188 and the side-chain carboxamide of Asn103 was early suggested to hamper the rotation of the tyrosines toward the polymerase active site, thereby increasing the stabilization of the closed-pocket form of this mutant enzyme [90].

1.3. The protease (PR)

Several studies demonstrated that the processing function is provided by a viral rather a cellular enzyme, and that proteolytic digestion of the Gag and Gag-Pol precursors is essential for virus infectivity [91]. Sequence analysis of retroviral PRs indicated that they are related to cellular "aspartic" proteases such as pepsin and rennin. Like these cellular enzymes, retroviral PRs use two apposed Asp residues at the active site to coordinate a water molecule that catalyzes the hydrolysis of a peptide bond in the target protein.

HIV-1 PR is a homodimer in which each monomer contains 99 amino acid residues. X-ray crystallographic data from HIV-1 PR [92] indicate that two monomers are held together in part by a four-stranded antiparallel β-sheet derived from both N- and C-terminal ends of each monomers and the C-terminus of the first monomer and the N-terminus of the second one are separated by less than 5 Å (figure 14).
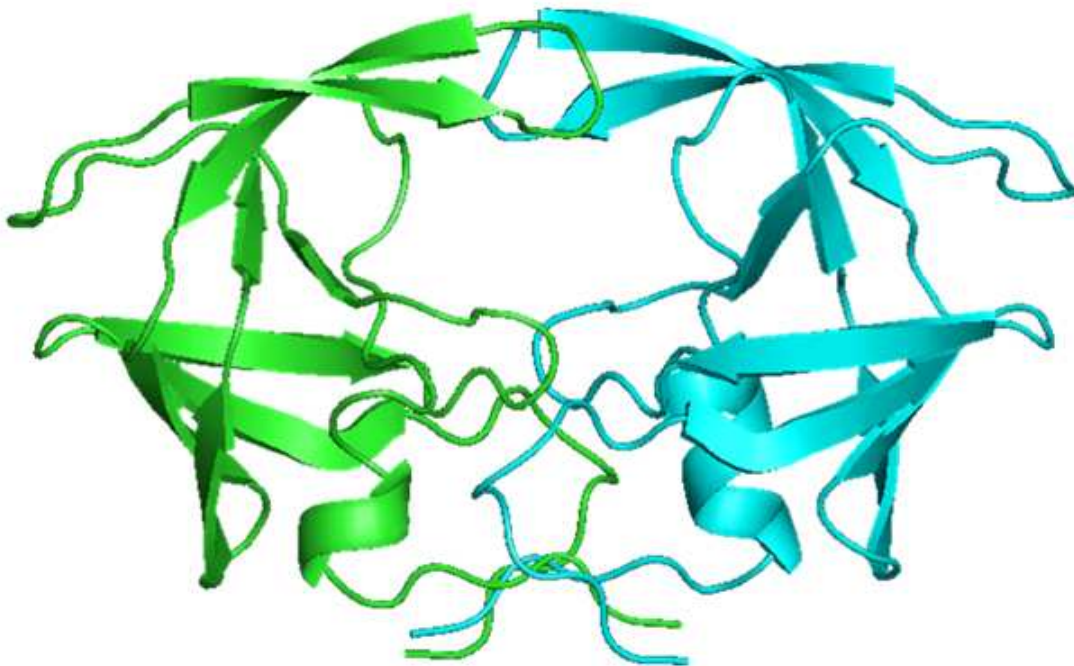

Figure 14. Tridimensional structure of HIV-1 protease. The two monomers are represented as cartoon and colored in green and cyan [Artese et al., unpublished data].

The substrate binding site is located within a cleft formed between the two monomers. Like their cellular homologs, the HIV-1 PR dimer contains flexible "flaps" that overhang the binding site and may stabilize the substrate within the cleft; the active site Asp residues lie in the center of the dimer.

The first cleavage events catalyzed by retroviral PRs during or immediately after virion release from the cell serve to liberate PR from the Gag-Pol precursors. It is still unclear

whether this initial cleavage event takes place in cis (intramolecularly) or in trans (intermolecularly); in either case, following the release of PR, the dimeric enzyme cleaves a number of sites in both Gag and Gag-Pol. These cleavage events lead to obtain functional proteins essential for maturation of infectious HIV particles. Inactivation of this enzyme causes the production of immature, noninfectious viral particles and hence blocks further HIV infection [92].

1.3.1. PR active site

As a member of the aspartyl protease family, the HIV-1 PR is composed of conserved residues, the so-called binding triads: D25(D25′)-T26(T26′)-G27(G27′), of which D25 and D25′ are known to be active residues [93]. The active cylindrical site presents a length of 23 Å and a diameter of 6-8 Å. Through interaction with this site, an inhibitor competes for binding with a normal substrate.

Residues 1-27 and 60-99 in each monomer form the region known as "core", while residues 28-59 constitute the "flap" region, a flexible area shown to be pivotal for PR activity.

Crystal structures of a functional HIV-1 PR tethered dimer (PDB codes 1LV1 and 1G6L) showed that the two flaps are in "closed conformation" even when no ligands are bound in the active-site cavity. HIV-1 PR catalyses the peptide hydrolysis reaction through nucleophilic attack by a water molecule on the carbonyl carbon of the scissile peptide bond [94]. This water molecule is believed to be the one that is hydrogen-bonded to both of the catalytic aspartates in the unliganded enzyme structure. When the substrate peptide moves into the active site of the enzyme and is oriented suitably, the water molecule attacks the carbon atom of the scissile peptide bond. The carbon atom then becomes tetravalent, with two oxygens attached to it.

This leads to a highly strained structure. When the entire substrate is enveloped by the enzyme, as is the case here, such a strained substrate structure would be stabilized by many different interactions with the enzyme. Stabilization of such a reaction intermediate by the enzyme is the key to rate enhancement in a catalytic reaction [95]. The dynamics of the flaps are very important for the functioning of this enzyme [96]. However, the flap dynamics (flexibility) are lost in the present hexagonal crystal, where flaps from symmetry-related molecules are in van der Waals contact. In the crystal structures of complexes between substrate oligopeptides and the inactive mutant D25N HIV-1 protease, the substrate is retained as a regular peptide [97]. It may be seen that the major change is in the orientation of the scissile amide proton relative to the catalytic aspartates. There is a strong hydrogen bond

between the scissile nitrogen atom and the outer oxygen of one of the aspartates (O. . . N distance 2.80 and 2.90 Å); this change is caused largely by the decrease in the torsion angle about the scissile C–N bond (figure 15).
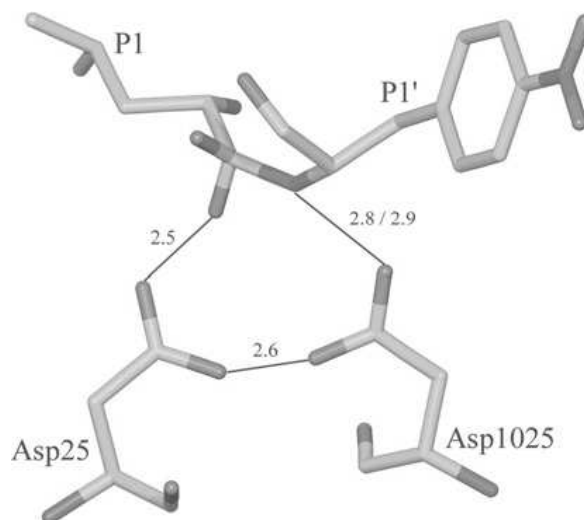

Figure 15. Hydrogen-bonding interactions at the catalytic site [98].

Special attention was given to the protonation state of Asp25 and Asp25′ in the active site because they occur in three possible states (diprotonated, monoprotonated, and deprotonated) depending on the inhibitor bound. So protonation of D25/D25′ was assigned into six different ionizable states, including unprotonation, four monoprotonated states for D25′ or D25, and diprotonation. The two Asp25 residues in protease can adopt two possible configurations, labeled as "up" and "down" based on the position of proton on OD1 oxygen or OD2 oxygen atom of D25 or D25′.

Based on the distance criteria alone, there are two strong hydrogen bonds involving catalytic aspartates: (a) between the O1 atom of the tetrahedral intermediate and the outer oxygen atom (OD1) of one of the catalytic aspartates (distance 2.5 Å), and (b) between the substrate

nitrogen and the outer oxygen (OD1∗) of the second catalytic aspartate (distance 2.8 Å). In

contrast, in the inactive enzyme complex, there is only one hydrogen bond at the catalytic centre, and that is between the scissile carbonyl oxygen and the outer carboxyl oxygen of the catalytic aspartate. One water molecule in the active site, popularly known as the flap water,

links the reaction intermediate to the protein through four hydrogen bonds, two to the flap amide groups of Ile-50 and Ile-1050, and two to the CO groups of residues P2 and P1′ from the reaction intermediate (figure 16) [98].
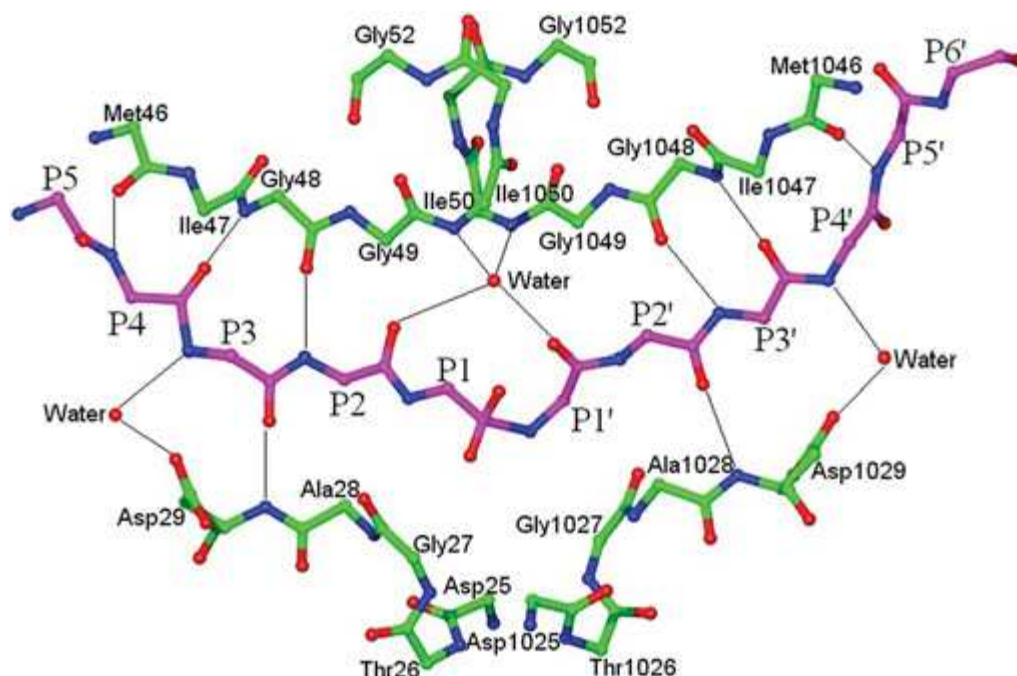

Figure 16. Hydrogen-bonding interactions between main-chain atoms of the substrate and the PR molecule [98].

The NH groups of the residues at positions P3 and P4′ are hydrogen bonded to water molecules, which in turn are hydrogen-bonded to the carboxy oxygen atom of side-chain Asp-29/1029 from the enzyme.

Several residues contribute to the PR active site: Arg8, Leu23, Asp25-Thr-Gly-Ala-Asp-Asp30, Val32, Ile47-Gly-Gly-Ile50, Phe53, Thr80-Pro-Val82 and Ile84. Within such a pocket a wide distribution of positive and negative charges is localized and different hydrophobic elements are present. Crystallographic data showed that these residues are involved in van der Waals interactions with the substrate, forming the binding site. Many residues are able to create more than a single sub-site; some are involved only in contacts with the longest side chains of the substrate at specific positions. The bulky aromatic side chains of Tyr or Phe placed at P1 or P1' are introduced between the PR side chains of Pro81 and Val82, while the side chain of Phe at P3 is located between PR residues Pro81′ and Phe53.

The importance of the Asp25 residue for PR function was confirmed in studies showing that it is required for catalytic activity and viral infectivity [99]. Subsequent studies showed that co-expression of the wild-type proviral DNA with increasing amounts of the mutant proviral

DNA bearing the D25N mutation results in a concomitant decrease in the proteolytic activity monitored by *in vivo* viral polyprotein processing [100]. This suggested that the PR domain bearing the D25N mutation is capable of adopting a native-like fold to sequester the wild-type PR via heterodimer formation and is related to a drastic decrease in PR inhibitors binding affinity.

More than 20 mutations, located in the active site, the flap region, dimerization interface, and double-strand β-sheet, have been identified in clinical and *in vitro* isolates exposed to PR inhibitors [101]. Among all these mutations, the mutations of six residues can confer resistance to almost all protease drugs, including L10, M46, I54, V82, I84, and L90, which are located in different regions of protease, including the active site (V82 and I84), the flap region (M46 and I54), and the dimerization interface (L10 and L90) [102]. The mutations of V82 and I84 in the active site have more significant effect on the binding affinities than the other four residues [103]. The double mutation of V82 and I84 impair affinities of all available drugs significantly. For example, the V82F/I84V double mutant has a 700-fold increase in K_i for ritonavir, 100-fold increase in K_i for amprenavir, a 79-fold increase in K_i for indinavir, 86-fold increase in K_i for nelfinavir, and 19-fold increase in K_i for saquinavir [104].

1.3.2. Heterodimer interface area

The dimer of PR is maintained by interactions between the two subunits, including the terminal residues (1-4 and 96-99), the tips of the flaps (50, 51), Asp29, Arg87 and Arg8′ and residues in and surrounding the active site (residues 24-27) [105].

Upon its intramolecular maturation at its N terminus, the protease forms a stable dimer concomitant with the formation of the terminal β-sheet structure and a very low equilibrium dimer dissociation constant (K_d < 10 nM) [106]. Some CH-π interactions stabilizing the two subunits at the termini are essential for PR dimer stability. Moreover the hydroxyl oxygen of Thr26 and the carbonyl of Leu24 form a hydrogen bond that is critical for dimerization (inter-oxygen distance equal to 2.7 Å) [107], while the distance between NH_2 and OD_2 of Arg8 and Asp29′ must be equal to 2.8 Å.

NMR studies of the inhibitor-free PR and PRD25N suggest that the principal difference between these two proteins resides in the P1 loop and flap regions, which are known to be relatively mobile. These results indicate that the active site Asp residues make a large and specific contribution to the stabilization of the monomer fold and dimerization. Other PR mutations related to a destabilizing dimer effect and to a reduced sensitivity to the current PR inhibitors are: Arg8Gln, Leu10Arg, Ile50Val and Leu90Met.

Arg8 results fundamental in the formation of the subsite S3 and S3′ and is involved in a ionic interaction with the Asp at P3′. The mutant R8Q is not allowed to establish this interaction and causes a reduction in the viral growth.

The L10R substitution is identified as being strongly associated with indinavir resistance in clinical isolates due to the loss of hydrophobic interactions [108].

Ile50 is located in the flap region of HIV-1 PR. The conserved sequence domain of the flap begins at position 47 of the HIV-1 sequence and extends through the glycine at position 52. Sequences between these points align in length and in a pattern: aliphatic-Xaa-Gly aliphatic-small-Gly. In the HIV-1 PR, these residues combine to give a short stretch of β-sheet followed by a turn that ends with the conserved glycine at position 52.

Ile50 can participate in both the S1/S1′ and S2/S2′ subsites, although in an asymmetric way. The δ carbon can be involved in hydrophobic packing within the flap or extend out toward solvent, with the γ carbons making contacts with bound inhibitors in either orientation. Thus, while a valine substitution would reduce some of the internal packing, it would not preclude continued interaction with substrate, consistent with the near wild-type level of activity that was observed with the Ile50Val mutant. However, the loss of the internal packing of the d carbon may provide more flexibility to the remaining atoms of the side chain. This explanation could account for the appearance of Ile50Val as a substitution associated with at least partial resistance to some PR inhibitors [109].

The single mutation L90M, usually associated to saquinavir treatment, exhibits the impairment in viral fitness and a less processing *in vitro* with respect to wild-type PR and causes also a significant decrease in replicative capacity [110].

1.4. The glycoproteins of HIV-1 envelope: gp41 and gp120

The HIV Env glycoproteins play a major role in the virus life cycle. They contain the determinants that interact with the CD4 receptor and coreceptor, and they catalyze the fusion reaction between the lipid bilayer of the viral envelope and the host cell plasma membrane. In addition, the HIV Env glycoproteins contain epitopes that elicit immune responses that are important from both diagnostic and vaccine development perspectives.

The HIV Env glycoprotein is synthesized from the single spliced Vpu/Env bicistronic mRNA; translation occurs on ribosomes associated with the rough endoplasmic reticulum (ER). The 160-kd polyprotein precursor (gp160) is an integral membrane protein anchored to cell membranes by a hydrophobic stop-transfer signal in the domain destined to be mature

transmembrane Env glycoprotein, gp41. The gp160 is cotranslationally glycosylated, rapidly associates with the host chaperone BiP/GRP78, forms disulfide bonds and undergoes oligomerization in the ER [111]. The gp160 is transported to the Golgi, where is proteolytically cleaved by cellular enzymes to the mature surface (SU) glycoprotein gp120 and transmembrane (TM) glycoprotein gp41 (figure 17).
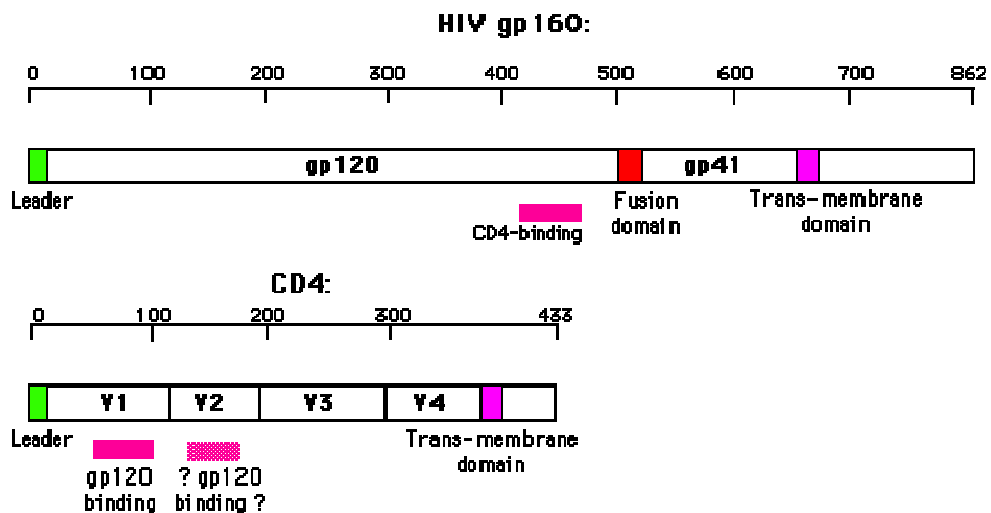

Figure 17. Linear representations of the HIV-1 Env gp160 glycoprotein and CD4 sequence [112].

The cellular enzyme responsible for cleavage of retroviral Env precursors following a highly conserved Lys/Arg-X-Lys/Arg-Arg motif is furin or a furin-like protease, although other enzymes may also catalyze gp160 processing. Cleavage of gp160 is required for Env-induced fusion activity and virus infectivity [113]. Subsequent to gp160 cleavage, gp120 and gp41 form a non-covalent association that is critical for transport of the Env complex from the Golgi to the cell surface. The HIV Env glycoprotein complex, in particular the SU (gp120) domain, is very heavily glycosylated; approximately half the molecular mass of gp160 is composed of oligosaccharide side chains. During transport of Env from its site of synthesis in the ER to the plasma membrane, many of the side chains are modified by the addition of complex sugars. Following its arrival at the cell surface, the gp120-gp41 complex is rapidly internalized. Several studies have demonstrated that a Tyr-X-X-Leu sequence in the gp41 cytoplasmic tail is partially responsible for this rapid internalization.

The mechanism by which the Env glycoproteins are incorporated into budding virus particles remains incompletely characterized. The Env glycoproteins are actively recruited into virions via direct interactions between Env and the matrix (MA):

1. Mutations in the HIv-1 MA can block HIV-1 Env incorporation [114].

36

2. HIV-1 Env directs basolateral budding of Gag in polarized epithelial cells [115].

3. A direct binding between HIV-1 MA and peptides derived from the gp41 cytoplasmic tail was detected *in vitro* [116].

4. A single amino acid change in MA reverses an Env incorporation defect caused by a small deletion in the gp41 cytoplasmic tail [117].

Thus it is clear that the gp41 cytoplasmic tail plays a crucial role in Env incorporation.

A primary function of viral Env glycoproteins is to promote a membrane fusion reaction between the lipid bilayers of the viral envelope and host cell membranes. This membrane fusion event enables the viral core to gain entry into the host cell cytoplasm. A number of regions in both gp120 and gp41 have been implicated, directly or indirectly, in Env-mediated membrane fusion. Studies of the HA$_2$ hemagglutinin protein of the *Orthomyxoviruses* indicated that a highly hydrophobic domain at the N-terminus of this protein, referred to a fusion peptide, plays a critical role in membrane fusion [118]. Mutational analyses demonstrated that an analogous domain was located at the N-terminus of the HIV-1 TM glycoprotein. Non hydrophobic substitutions within this region of gp41 greatly reduced or blocked syncytium formation and resulted in the production of non infectious progeny virions [119]. C-terminal to the gp41 fusion peptide are two amphipatic helical domains which play a central role in membrane fusion. Mutations in the N-terminal helix, which contains a Leu zipper-like heptad repeat motif, impair infectivity and membrane fusion activity [120]. The structure of the ectodomain of HIV-1 and SIV gp41, the two helical motifs in particular, has been the focus of structural analyses in recent years by means of x-ray crystallography or NMR spectroscopy. These studies obtained similar trimeric structures, in which the two helical domains pack in an antiparallel fashion to generate a six-helix bundle, as shown in figure 18.
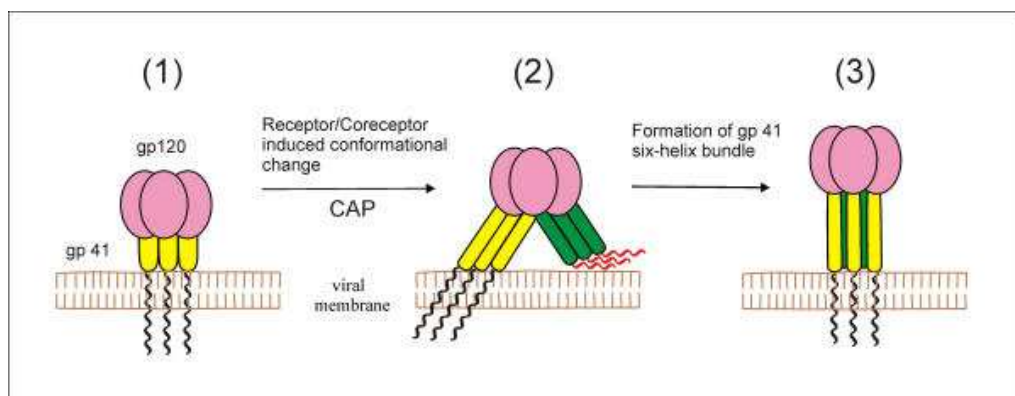

Figure 18. The formation process of gp41 six-helix bundle mediated by receptor/coreceptor conformational changes [121].

The N-helices form a coiled-coil in the center of the bundle, with the C-helices packing into hydrophobic grooves on the outside. Peptides corresponding to the helical domains presumably inhibit fusion by interacting with their complementary binding partner on gp41 and preventing gp41 itself from adopting the six-helix bundle structure (figure 19).
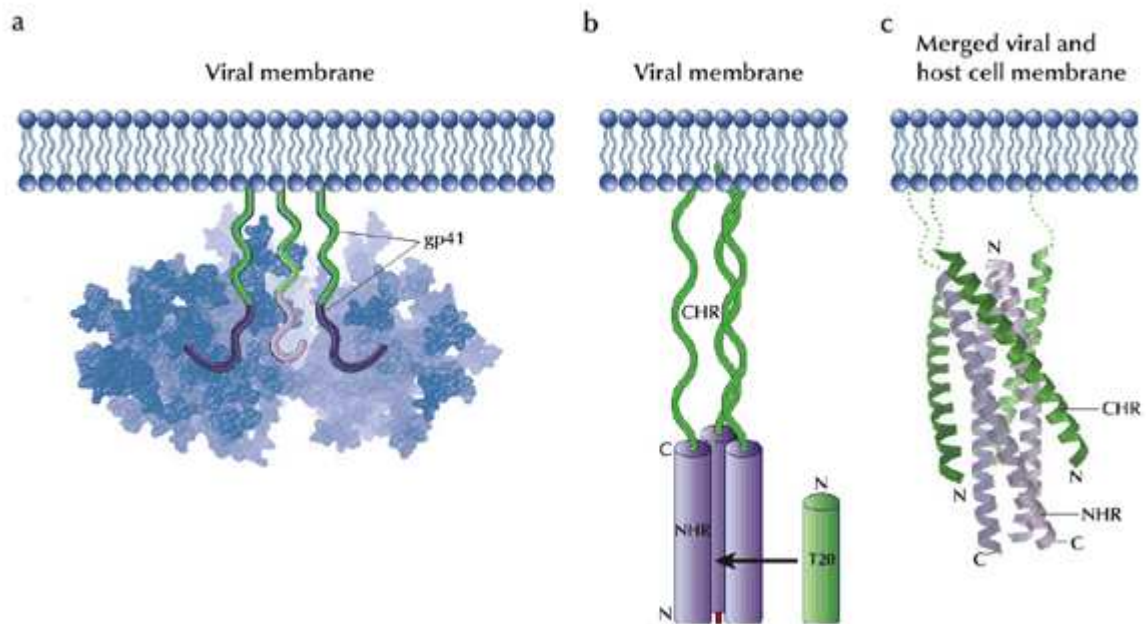

Figure 19. The fusion process of gp41 glycoprotein. (**a**) After gp41 conformational changes, (**b**) N-helices form a coiled-coil in the center of the bundle, with the C-helices packing into hydrophobic grooves, where T-20 can interact with and (**c**) the consequent formation of the six-helix bundle leads to the fusion reaction [122].

Binding studies using native HIV-1 Env indicated that gp41 could interact with C-helix-derived peptide only after CD4 binding [123]. These results suggest that the N- and C- helices undergo conformational changes (exposure) following CD4 binding, and that these rearrangements are required for membrane fusion.

Whereas gp41 appears to directly interact with the lipid bilayer to catalyze the membrane fusion reaction, a variety of domains within gp120 are involved in activating Env fusogenicity. The V_3 loop of gp120, which elicits isolate-specific neutralizing antibodies, is an essential player in the membrane fusion reaction (figure 20). Mutations throughout the V_3 loop of HIV-1 block syncytium formation and virus infectivity without perturbing the processing, transport and CD4 binding properties of gp120 [124]. The importance of V_3 in membrane fusion appears to be a consequence of its role in gp120-coreceptor interactions. In addition to V_3, the V_1/V_2 region participates in some manner in the fusion reaction, as

confirmed by the observations that mutations in these variable loops impair fusion and antibodies that bind this region can neutralize virus infectivity [125].
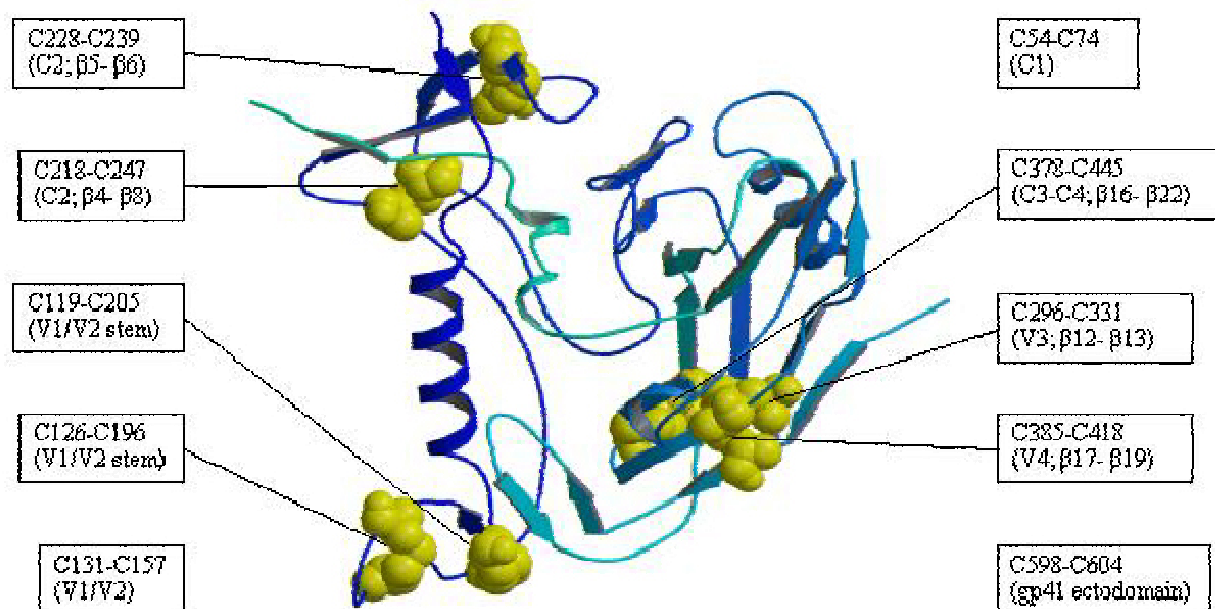

Figure 20. HIV-1 gp120 structural domains and loops [126].

The high affinity CD4 binding site for gp120 has been localized to a small segment of the N-terminal extracellular domain. The cytoplasmic domain of CD4 is not required for its viral receptor function, and mutations that disrupt CD4 internalization do not block HIV entry [127]. CD4 binding determinants in Env map to the C_3 and C_4 domains of gp120; CD4 binding serves not only to promote virion attachment to the target cell, but also to induce conformational changes in gp120 and gp41 required to activate Env fusogenicity. The recent crystallization of a gp120 "core" domain (an unglycosylated gp120 derivative lacking the V_1/V_2 and V_3 loops and the N- and C- termini) complexed with fragments of Cd4 and a neutralizing antibody has further refined the understanding of the gp120-CD4 interaction [128]. The core structure reveals two major domains connected by a so-called bridging sheet. The latter is composed of a four-stranded, antiparallel β-sheet derived from sequences in the V_1/V_2 stem and portions of C_4. A CD4 residue (Phe43), shown to be critical for gp120 binding, occupies the opening of the CD4-binding cavity, while many of the gp120 amino acids crucial for CD4 binding line the opening of the cavity [128]. The HIV-1 Env glycoprotein binds CD4 not only during the early phase of infection, but also during the transport of gp160 to the cell surface [129]. As a consequence of this intracellular association,

CD4 is down-modulated from the cell surface, and Env-expressing cells are partially resistant to further infection ("superinfection interference") [130]. Moreover the HIV-1 accessory protein Vpu degrades the CD4 component of gp160-CD4 complexes in the ER, thereby releasing Env for transport to the cell surface.

In concert with CD4, secondary receptors or coreceptors expressed on human cells were found to function in the membrane fusion process. Several studies demonstrated that members of the G protein-coupled receptor superfamily of seven-transmembrane domain proteins provided the long-sought coreceptor function. So the α-chemokine receptors CXCR4 and CCR5 were identified as the primary coreceptors respectively for TCL-tropic and M-tropic isolates of HIV-1 [131]. Several domains within gp120 directly or indirectly function in Env-coreceptors interactions. Consistent with its influence on HIV-1 tropism, the V3 loop plays a major role: single amino acid changes or deletions in V3 shift or impair coreceptor usage, V3 peptides interact with CXCR4 and antibodies to V3 block gp120-CCR5 binding. The V1/V2 region also helps determine coreceptor usage. In addition to the involvement of variable regions in coreceptor interaction, highly conserved portions of gp120 exposed upon CD4 binding appear to interact with the coreceptors. The observation that CD4 induces conformational changes in gp120 that enhance interaction with coreceptors suggests that events leading up to membrane fusion occur sequentially. First gp120 interacts with CD4, then a ternary complex forms between coreceptor, CD4 and gp120, and finally conformational changes take place in gp41 that trigger membrane fusion (figure 21) [128].
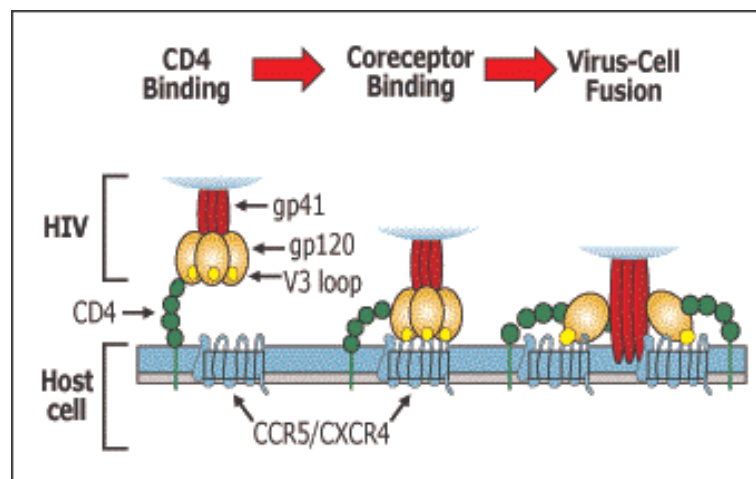

Figure 21. Sequential steps leading up to membrane fusion reaction [132].

1.5. Antiretroviral drugs

The development of new antiretroviral drugs is a dynamic process that is continuously focused, at the basic level, to the identification of new molecular targets for chemotherapeutic intervention and at the clinical level to bypass the problems (i.e., adherence, tolerability, toxicity, virus-drug resistance) related to existing drugs [133]. The most recent approaches toward antiretroviral drugs include the following classes:

❖ cellular CD4 receptor down-modulators;

❖ virus attachment inhibitors;

❖ CXCR4 and CCR5 antagonists;

❖ virus-cell fusion inhibitors;

❖ reverse transcriptase inhibitors:

  ➢ Nucleoside Reverse Transcriptase Inhibitors (NRTIs);

  ➢ Nucleotide Reverse Transcriptase Inhibitors (NtRTIs);

  ➢ Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs);

❖ HIV integrase inhibitors;

❖ transcription (transactivation) inhibitors;

❖ HIV protease inhibitors;

❖ HIV Ribonuclease H inhibitors.

The inhibition of the viral entry into the cells may be considered an attractive approach toward prevention of HIV infection. Cyclotriazadisulfonamide (CADA) was recently shown to inhibit HIV infection by down-modulation of the cellular CD4 receptor (figure 22) [134].
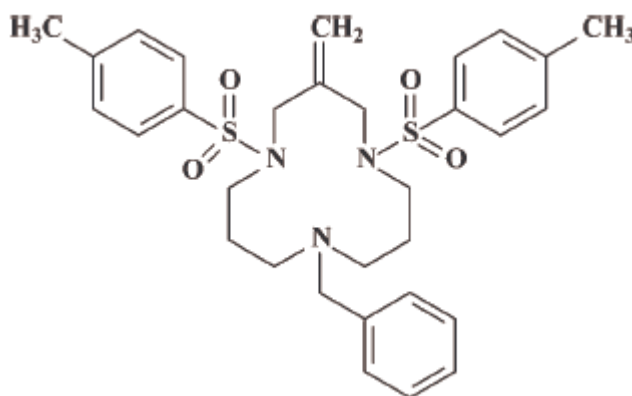
Figure 22. Chemical structure of cyclotriazadisulfonamide (CADA) [134].

CADA is specific for the CD4 receptor and is assumed to down-regulate CD4 expression at the post-translational level.

The plant lectins derived from GNA (Snowdrop) and HHA (Amaryllis) were shown to interrupt the viral entry process by interfering with the viral envelope glycoprotein gp120. These plants represent a unique class of anti-HIV agents with an entirely novel HIV drug resistance profile; they are endowed with a number of interesting properties (i.e., availability, formulation, stability, efficacy, safety) that make them primary candidate drugs to be considered for potential topical use as microbicides for the prevention of the sexual transmission of HIV infection [135].

Also cyanovirin-N, a 11 kDa protein originally isolated from the cyanobacterium *Nostoc ellipsosporum*, is a potential microbicide to prevent AIDS transmission (figure 23). It has a high affinity for gp120 and is able to impair both CD4-dependent and -independent binding of gp120 to the target cells, to block CD4-induced binding of gp120 with CXCR4, and to dissociate bound gp120 from target cells [136].

(H₂N)Leu—Gly–Lys—Phe—Ser—Gln—Thr—Cys—Tyr—Asn—Ser—Ala—

—Ile—Gln—Gly—Ser—Val—Leu—Thr—Ser—Thr—Cys—Glu—Arg—Thr—Asn—Gly–Gly—Tyr—Asn—Thr–Ser—

—Ser—Ile—Asp—Leu—Asn—Ser—Val—Ile—Glu—Asn—Val—Asp—Gly—Ser—Leu—Lys—Trp—Gln—Pro—Ser—

—Asn—Phe—Ile—Glu—Thr—Cys—Arg—Asn—Thr—Gln—Leu—Ala—Gly—Ser—Ser—Glu—Leu—Ala—Ala—Glu—

—Cys—Lys—Thr—Arg—Ala—Gln—Gln—Phe—Val—Ser—Thr—Lys—Ile—Asn—Leu—Asp—Asp—His—Ile—Ala—

—Asn—Ile—Asp–Gly—Thr—Leu—Lys—Tyr—Glu(COOH)

Figure 23. Chemical structure of cyanovirin-N [136].

Recently, a new class of HIV-1 attachment inhibitors [prototype: 4-benzoyl-1-[4-methoxy-1H-pyrrolo[2,3-b]-pyridin-3-yl)oxoacetyl]-2-(R)-methylpiperazine (BMS-378806)] has been identified that interferes with CD4-gp120 interactions (figure 24). Such a compound binds directly to gp120 at a stechiometry of approximately 1:1, with a binding affinity similar to that of soluble CD4 [137].
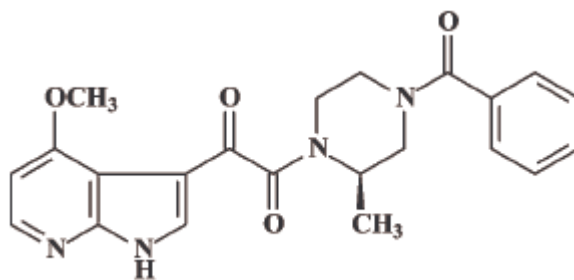
Figure 24. Chemical structure of BMS-378806 [137].

Various low-molecular-weight CXCR4 and CCR5 antagonists have been identified, the prototype of the CXCR4 antagonists being the bicyclam AMD3100 (figure 25). This compound is truly specific for CXCR4; it does not interact with any other CXCR or CCR receptor and blocks TCL-strains HIV-1 replication through CXCR4 antagonization. SAR studies with bicyclam analogues have revealed that the bis-macrocyclic structure is not essential for anti-HIV activity; i.e., AMD3465, an AMD3100 derivative in which one of the cyclam rings is replaced by a pyridinylmethylene amine moiety, proved to be more potent of AMD3100 (figure 25) [138].
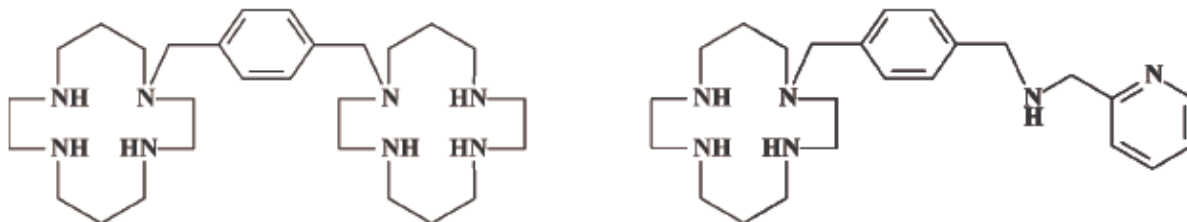
Figure 25. Chemical structures of AMD3100 (left) and AMD3465 (right) [138].

Of the CCR5 antagonists, the quaternary ammonium derivative TAK-779 was the first non-peptidic molecule shown to block the replication of the M-tropic HIV-1 strains by interaction with CCR5 (figure 26). The binding site for this derivative has been identified within specific transmembrane (TM) helices of CCR5 [139].
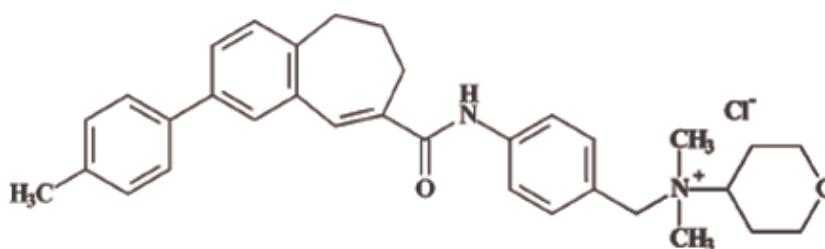
Figure 26. Chemical structure of TAK-779 [139].

SCH 351125 (SCH-C) (figure 27) was the first CCR5 antagonist to be advanced to clinical studies. Such a compound has potent activity *in vitro* against primary HIV-1 isolates using CCR5 as their cell entry coreceptor. It has been proposed that the binding of this molecule to the TM domain of CCR5 may disrupt the conformation of its extracellular domain, specifically the second extracellular loop, so that it can no longer interact with the crown of the V_3 loop of gp120 [140].
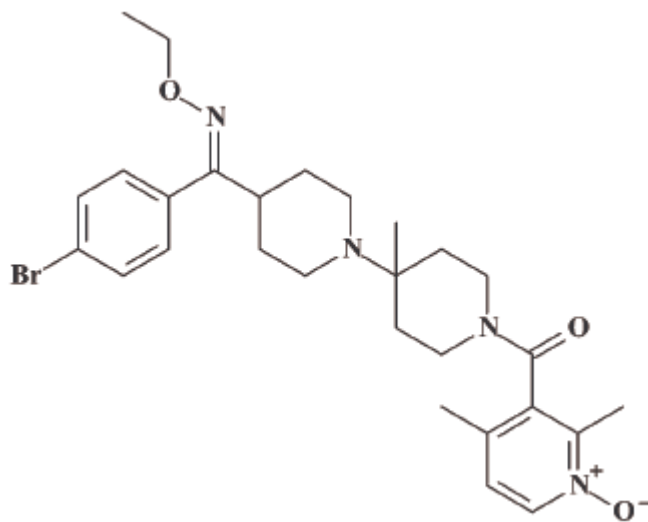

Figure 27. Chemical structure of SCH 351125 (SCH-C) [140].

UK-427857 (figure 28), another CCR5 antagonist, has recently been selected as a clinical development candidate drug for the treatment of HIV infection. It was found to exhibit an excellent potency against CCR5 isolates, while being inactive against CXCR4-tropic viral isolates [141].
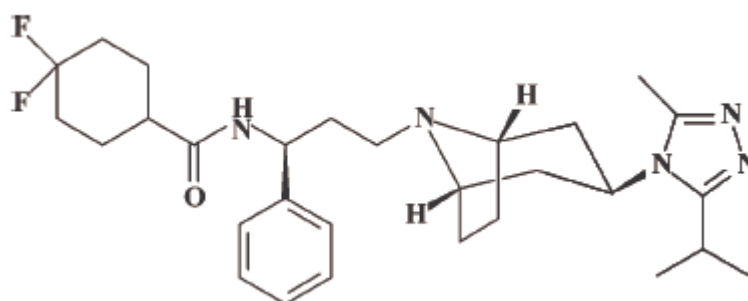
Figure 28. Chemical structure of UK-427857 [141].

The interaction of HIV-1 envelope gp120 glycoprotein with the coreceptors CXCR4 or CCR5, respectively, is followed by a spring-loaded action of gp41 that then anchors through its amino terminus into the target cell membrane and, in doing so, initiates the fusion of the viral envelope with the cellular plasma membrane. At the onset of the fusion process, hydrophobic grooves on the surface of the coiled coil gp41 ectodomain become available for binding with extraneous inhibitors, such as enfuvirtide (T-20, DP-178, pentafuside, Fuzeon), a synthetic, 36-amino-acid peptide corresponding to residues 127-162 of gp41 or residues 643-678 of the gp160 precursor (figure 29) [142].
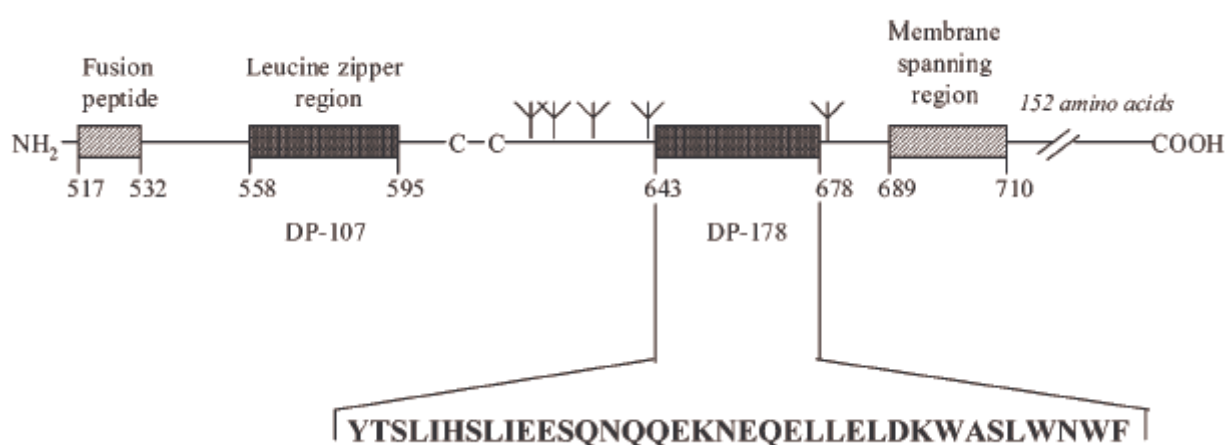
Figure 29. Chemical structure of T-20 (enfuvirtide) [142].

In HIV-infected adults, at the highest dose (100 mg, twice daily) enfuvirtide achieved a 1.5- to 2.0-fold reduction in plasma HIV RNA by the 15th day, thus providing proof of concept that HIV fusion inhibitors are able to reduce virus replication *in vivo*.

Recently, a new strategy targeted at gp41 and aimed at inhibiting virus-cell fusion has been described. This strategy is based on α-helix mimicry. Tris-functionalized 3,2′,2″-terphenyl derivatives (figure 30) could serve as effective mimics of the exposed N-helical regions of the transient gp41 intermediate and thus potentially trap this structure prior to the six-helix bundle formation, which is required for virus-cell fusion [143].
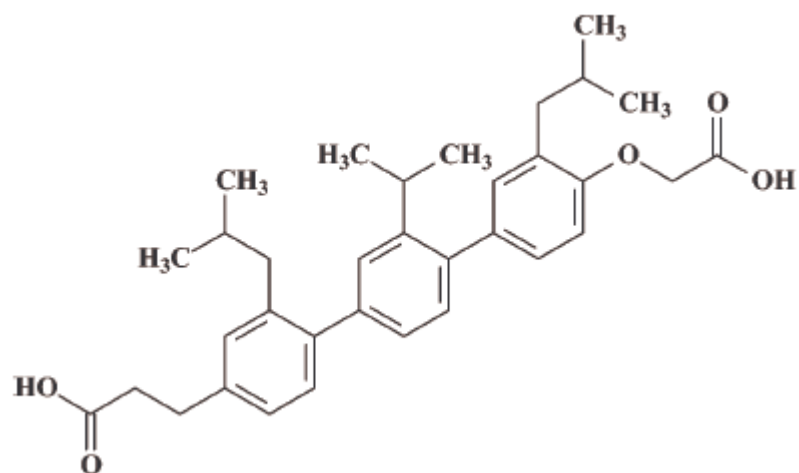
Figure 30. Chemical structure of the terphenyl derivative [143].

The nucleoside reverse transcriptase inhibitors (NRTIs) are analogs of the natural substrate (dNTP) and, by a competitive mechanism with its binding site (dTTP), they behave as chain terminators.

To enact such a mechanism, these inhibitors must necessarily have a free OH group in 3′ position (figure 31) [144].
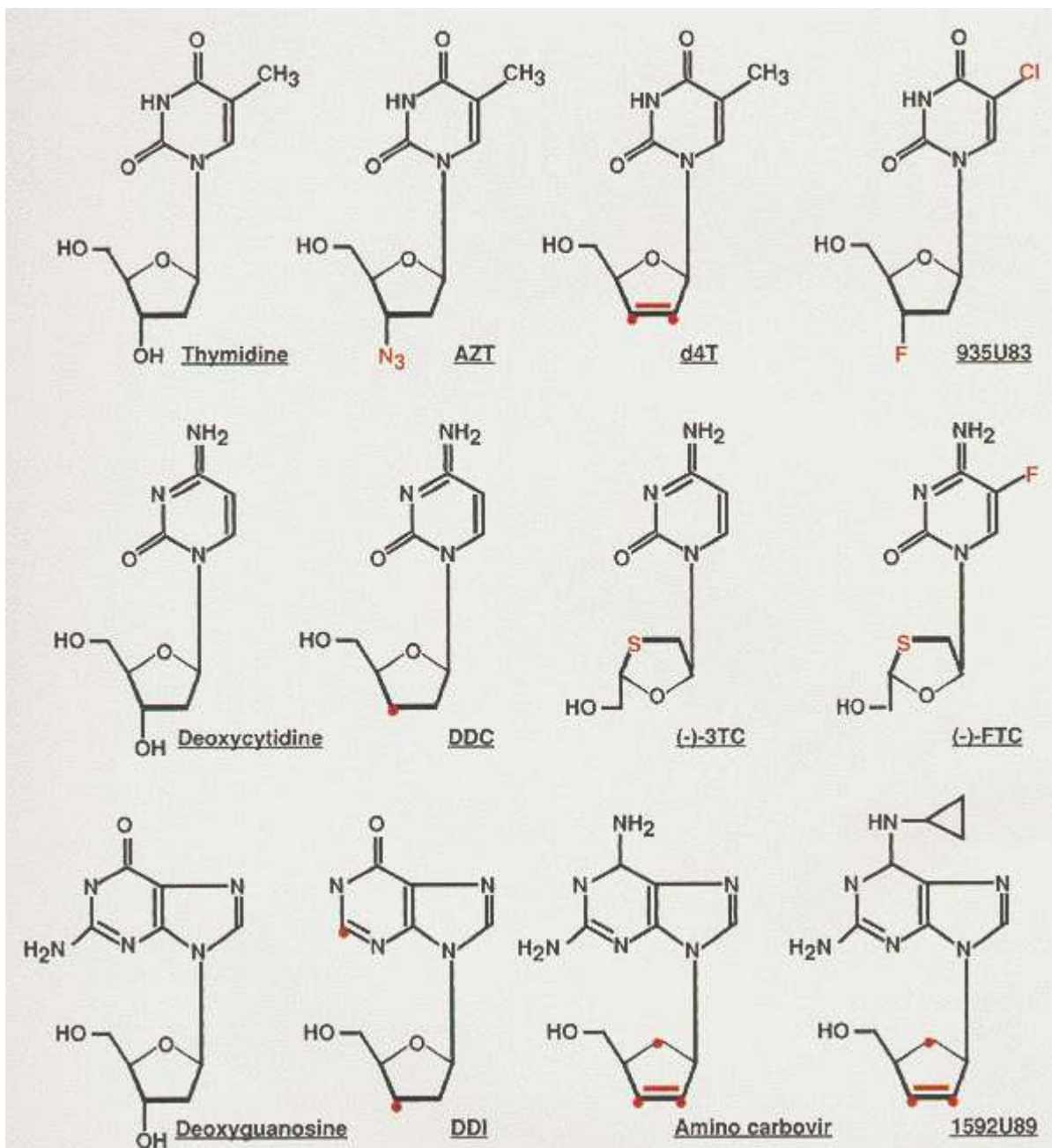
Figure 31. Chemical structures of several NRTIs. These molecules can be classified as analogs of thymidine, deoxycytidine and deoxyguanosine [144].

Subsequently they are converted in triphosphate derivatives by a cellular thymidine kinase, which activates them to block the natural dNTPs binding to DNA. In figure 32 is reported as example of activation the AZT tri-phosphorilation process [145].
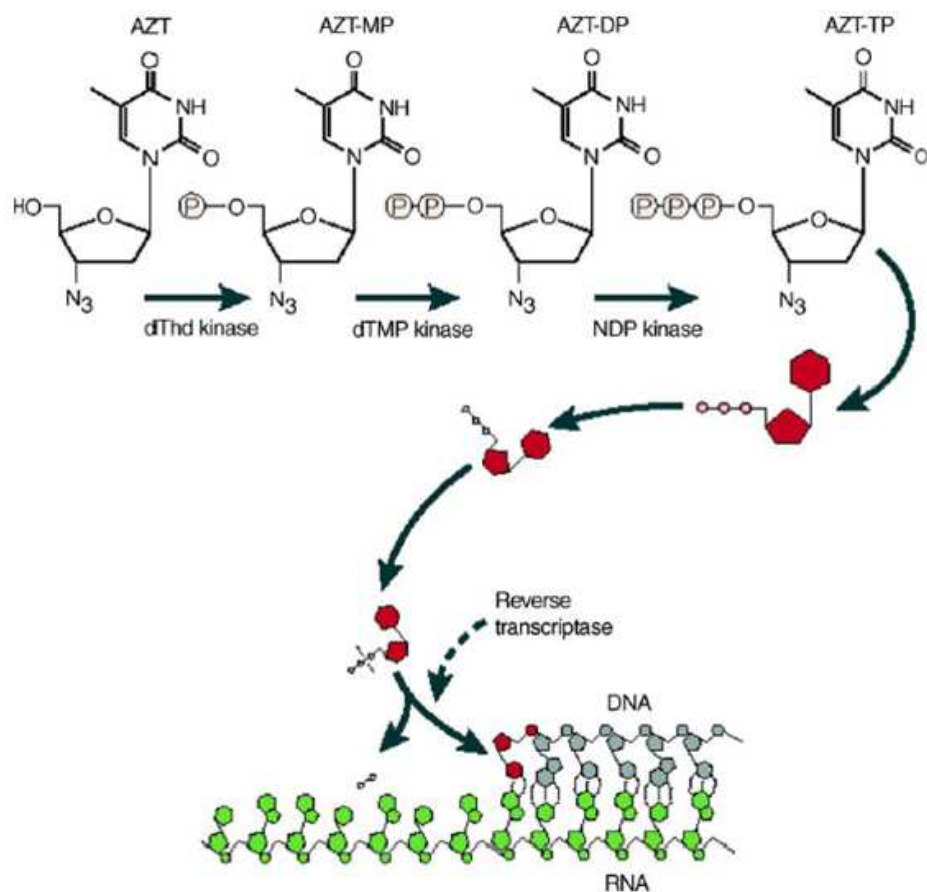
Figure 32. AZT tri-phosphorilation activation process mediated by cellular kinases [145].

Emtricitabine (2′,3′-dideoxy-3′-thia-5-fluorocytidine, (-)-FTC, previously referred to as Coviracil and now marketed as Emtriva) (figure 33) is the seventh 2′,3′- dideoxynucleoside analogue that has been officially approved for the treatment of HIV infections [146].
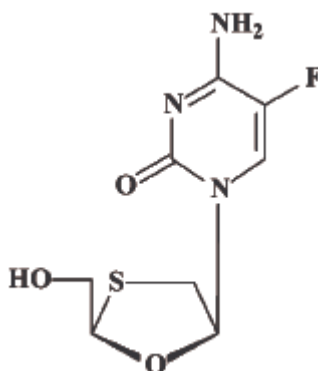
Figure 33. Chemical structure of emtricitabine [146].

Emtricitabine has been considered an "ideal drug candidate" because it shows synergism with other antiretrovirals, excellent tolerability, a long intracellular half-life, and, in comparison with lamivudine, 4- to 10-fold higher *in vitro* potency against HIV.

In contrast to the nucleoside reverse transcriptase inhibitors (NRTIs), the nucleotide reverse transcriptase inhibitors (NtRTIs) such as adefovir [9-(2-phosphonylmethoxyethyl) adenine (PMEA)] and tenofovir [(R)-9-(2- phosphonylmethoxypropyl)adenine (PMPA)] (figure 34) are already equipped with a phosphonate group and, therefore, only need two phosphorylation steps to be converted to the active metabolites (PMEApp and PMPApp, respectively). The metabolites then serve as alternative substrates (with respect to the natural substrate dATP) in the RT reaction, where upon their incorporation they act obligatorily as chain terminators [147].
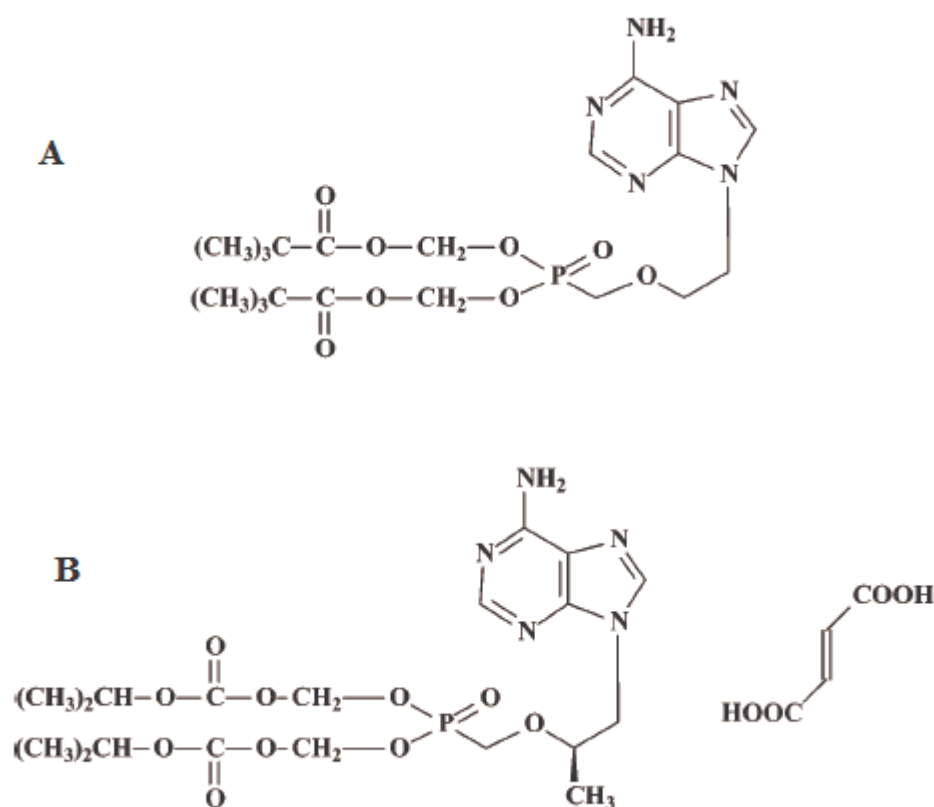

Figure 34. Chemical structures of (**A**) adefovir and (**B**) tenofovir [147].

More than 30 structurally different classes of compounds have been identified as NNRTIs, namely, compounds that are specifically inhibitory to the replication of HIV-1 and that are targeted at a specific, allosteric (i.e., non-substrate binding) site of the reverse transcriptase. Three NNRTIs [i.e., nevirapine (Viramune), delavirdine (Rescriptor), and efavirenz (Sustiva,

Stocrin] have so far been formally licensed for clinical use in the treatment of HIV-1 infections (figure 35) [148].
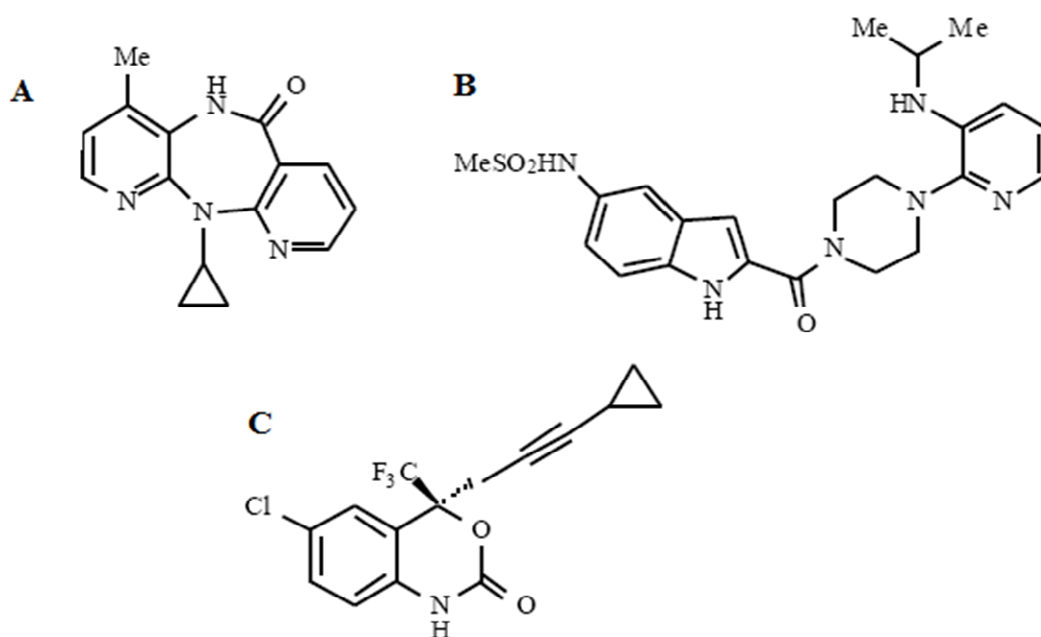

Figure 35. Chemical structures of the three approved NNRTIs (**A**) nevirapine, (**B**) delavirdine and (**C**) efavirenz [148].

Another NNRTI, the thiocarboxanilide UC-781 (figure 36), has been recognized as a retrovirucidal agent capable of reducing the infectivity of HIV-1 virions and therefore seems to be an ideal candidate microbicide [149].
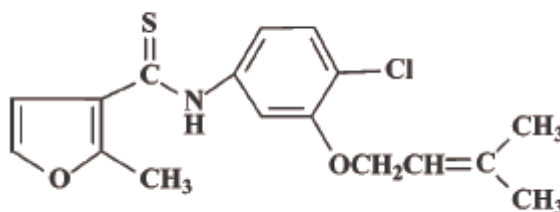

Figure 36. Chemical structure of the thiocarboxanilide UC-781 [149].

The "first-generation" NNRTIs are notorious for rapidly eliciting virus drug resistance, especially when used as monotherapy. The most common mutations occurring in the clinical setting following the use of NNRTIs are K103N and Y181C. Therefore, attempts have been made to develop "second-generation" NNRTIs that are resilient to such drug resistance

mutations. These attempts have led to the identification of a number of compounds that are indeed resilient to the K103N and/or Y181C mutations. Capravirine (figure 37) is such a compound that retained activity against HIV-1 strains carrying the K103N mutation in their RT; it has proceeded to phase II/ III clinical trials [150].
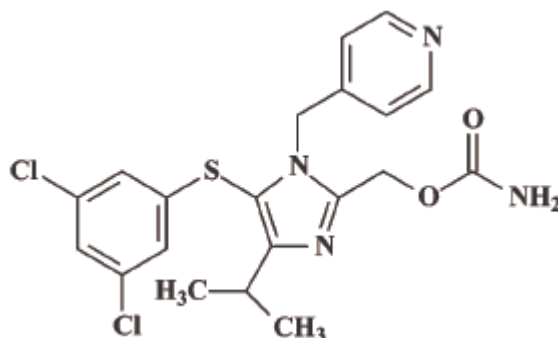

Figure 37. Chemical structure of capravirine [150].

Integration of the proviral DNA into host cell chromosomal DNA is an essential step in the viral replication cycle. This process is mediated by the viral integrase, and because there is no cellular homologue for this enzyme, it has been considered an attractive target for HIV therapeutics.

S-1360 (figure 38) actually represents the first integrase inhibitor to reach clinical studies. Its mechanism of action is based on an interaction between the carboxylate group of the diketo acids or the isosteric heterocycle in the two other compounds and metal ion(s) in the active site of the integrase, resulting in a functional sequestration of these critical metal cofactors [151].
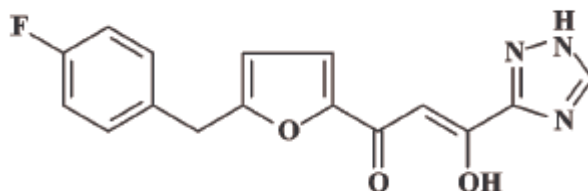

Figure 38. Chemical structure of S-1360 [151].

Recently, an entirely new class of HIV integrase inhibitors was identified, namely, that of the 5H-pyrano-[2,3-d:-6,5-d′] dipyrimidines (PDPs). The most potent congener of this series is

51

the 5-(4-nitrophenyl)-2,8-dithiol-4,6-dihydroxy-5H-pyrano[2,3-d:-6,5-d′] dipyrimidine (V-165), reported in figure 39 [152].
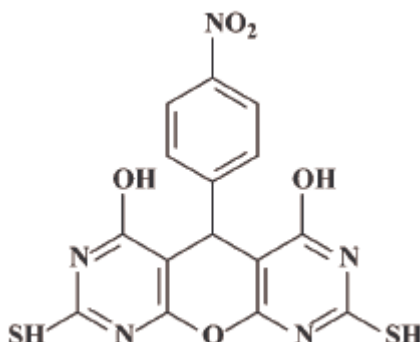

Figure 39. Chemical structure of the dipyrimidine V-165 [152].

At the transcriptional level, HIV gene expression may be inhibited by compounds that interact with cellular factors (such as NF-κB) that bind to the LTR promoter and that are needed for basal-level transcription. A number of compounds have been reported to inhibit HIV-1 replication in both acutely and chronically infected cells, through interference with the transcription process that could partially be attributed to inhibition of Tat or other transactivators. A 6-aminoquinolone, WM5 (figure 40), was recently shown to inhibit HIV-1 replication in acutely infected as well as chronically infected cells. This aminoquinoline was found to efficiently bind to TAR RNA and to inhibit Tat-mediated LTR-driven transcription [153].
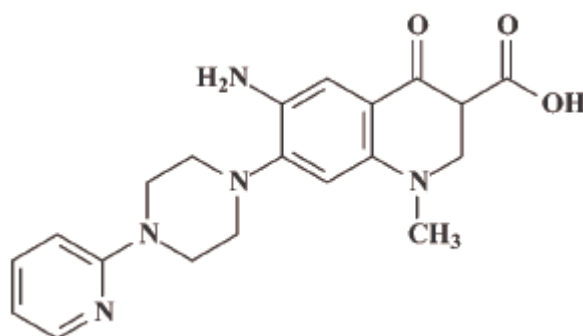

Figure 40. Chemical structure of the 6-aminoquinolone WM5 [153].

Following transcription of the unspliced (or partially spliced) HIV mRNA, this mRNA has to be transported from the nucleus into the cytoplasm in order to be translated to viral proteins. This export is promoted by the HIV-1 Rev (regulator of expression of viral proteins). Nuclear

export of Rev is mediated by its leucine-rich nuclear export signal (NES). NES uses the export factor CRM1 to export viral mRNA from the nucleus to the cytoplasm. This process can be blocked by a small molecular-weight molecule, PKF 050-638 (figure 41), that specifically inhibits CRM1-NES complex formation and, hence, Rev-mediated nuclear export [154].
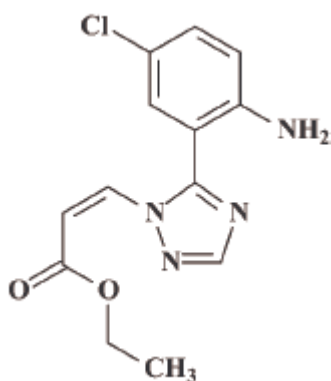

Figure 41. Chemical structure of PKF 050-638 [154].

The HIV protease is responsible for the cleavage of the Gag and Gag-Pol precursor polyproteins to the structural proteins (p17, p24, p7, p6, p2, p1) and the viral functional enzymes. All PR inhibitors (PIs) that have been licensed for the treatment of HIV infections, namely, saquinavir (Fortovase, Invirase), ritonavir (Norvir), indinavir (Crixivan), nelfinavir (Viracept), amprenavir (Agenerase) (and its phosphate derivative, Lexiva), lopinavir (Kaletra, lopinavir with ritonavir at a 4:1 ratio), and atazanavir (Reyataz), share the same structural determinant or scaffold, i.e., a hydroxyethylene (instead of the normal peptidic) bond, which makes these compounds peptidomimetic but non-scissile substrate analogues for the HIV PR (figure 42) [155].
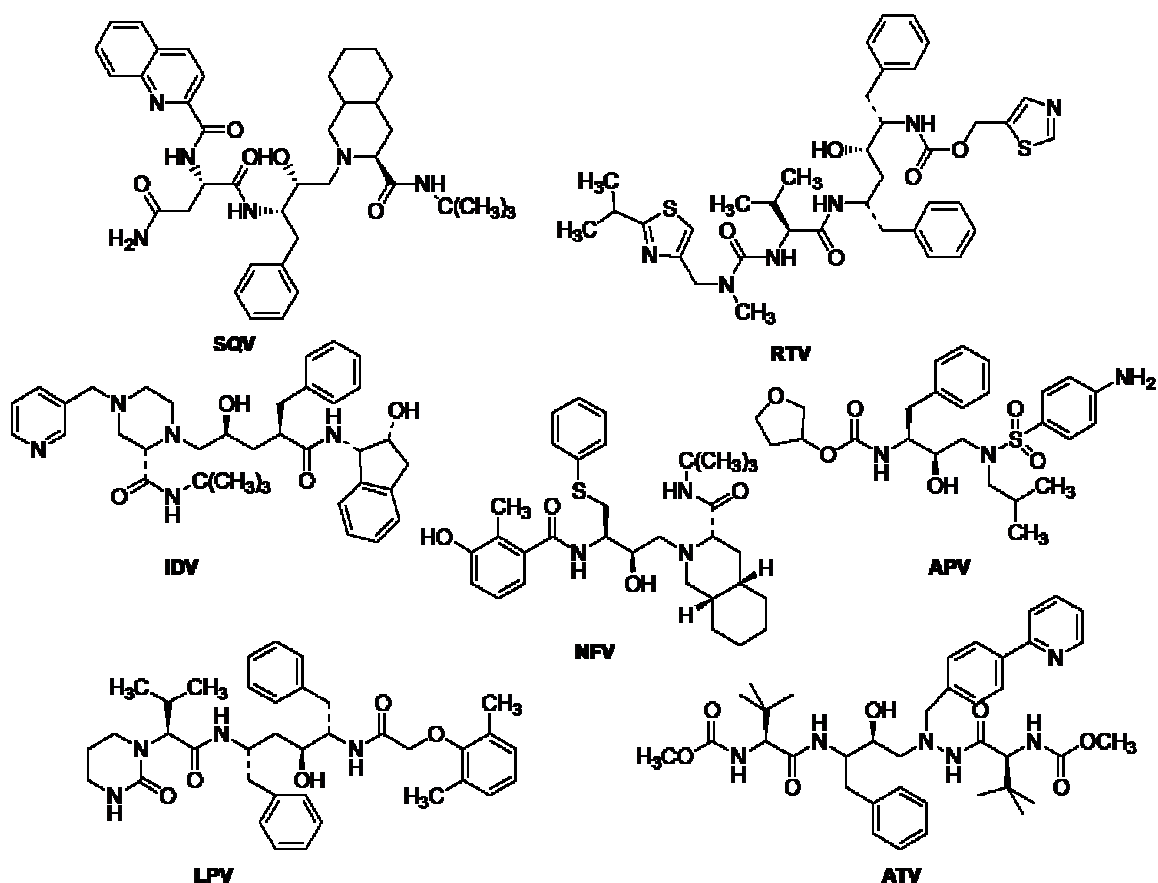
Figure 42. Chemical structures of the PR inhibitors licensed for the treatment of HIV-1 infections [156].

The aza-dipeptide analogue atazanavir, the latest and seventh PI to be approved for clinical use, combines a favorable resistance profile distinct from that of the other PIs, with a favorable pharmacokinetic profile allowing once-daily dosing [157].

Under clinical development is TMC 114, a structural analogue of TMC 126 (UIC-94003) (figure 43). Such a compound is a peptidomimetic PI containing a bis-tetrahydrofuranyl urethane and 4-methoxybenzenesulfonamide (and thus structurally related to amprenavir). It was reported to be extremely potent against a wide spectrum of HIV-1 strains [158].
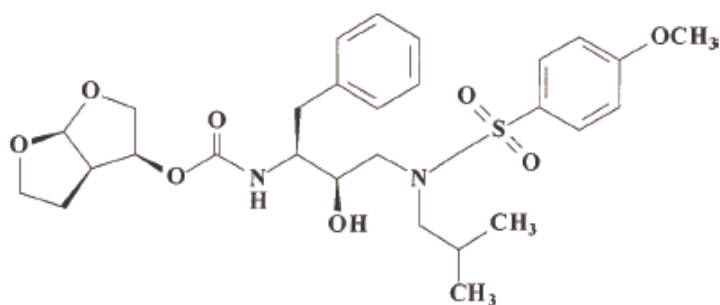

Figure 43. Chemical structure of TMC 114 [158].

Tipranavir (PNU-140690) (figure 44) is a non-peptidomimetic inhibitor of the HIV PR which, as could be expected, shows little cross-resistance to the peptidomimetic PIs. it might be useful in drug combination regimens with other antiretroviral agents for patients who already failed on other PI-containing drug regimens [159].
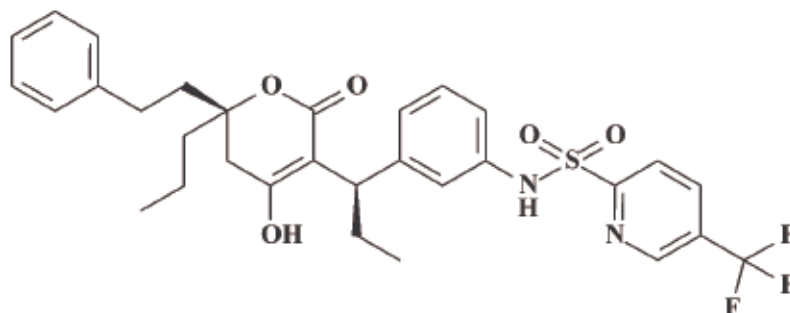

Figure 44. Chemical structure of the non-peptidomimetic PR inhibitor tipranavir [159].

During the reverse transcription process, that converts the single-stranded viral RNA into double-stranded (pro)viral DNA, a RNA·DNA heteroduplex (hybrid) is formed. The RNA strand of this heteroduplex must be cleaved by the RNase H domain before the remaining (-)-DNA strand can be duplicated to give the double-stranded (pro)viral DNA that is then integrated into the host cell genome. Several recent reviews have addressed the possibility of inhibiting RNAse H activity by small-molecule inhibitors. It appears feasible to specifically inhibit HIV-1 RNase H by, for example, novel diketo acids, such as 4-[5-(benzoylamino)thien-2-yl]-2,4-dioxobutanoic acid, and N-hydroxyimides, such as 2-hydroxy-4H-isoquinolone- 1,3-dione (figure 45) [160].

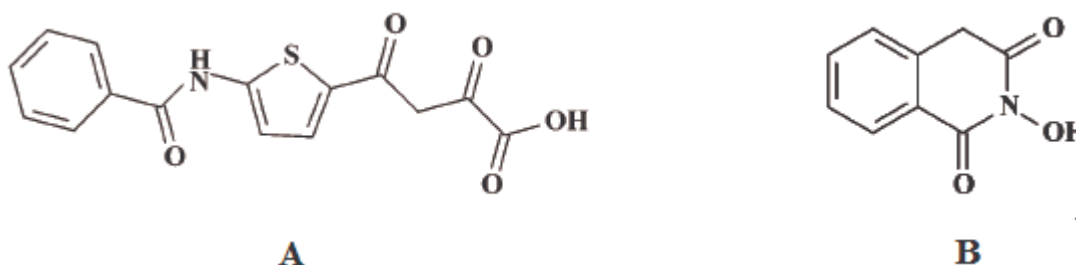**Figure 45.** Chemical structures of (**A**) the 4-[5-(benzoylamino)thien-2-yl]-2,4-dioxobutanoic acid and (**B**) the -hydroxy-4H-isoquinolone- 1,3-dione [160].

Whether both compounds are able to inhibit HIV replication in cells remains to be established.

1.6. The development of mutations associated to resistance

The use of combinations of antiretroviral drugs has proven remarkably effective in controlling the progression of HIV disease and prolonging survival, but these benefits can be compromised by the development of drug resistance. Resistance is the consequence of mutations that emerge in the viral proteins targeted by antiretroviral agents and the transmission of drug-resistant strains complicates further efforts to control viral replication [101]. The combination of antiretroviral drugs used for the treatment of HIV infection is called highly active antiretroviral therapy (HAART) and its regimens generally comprise three antiretroviral drugs, usually two NRTIs and either PR inhibitors or a NNRTI drug.

To fully understand the development of HIV drug resistance, two concepts have to be considered. First, HIV infection is characterized by high levels of virus production and turnover. Because the half-life of the infected cells is remarkably short (one to two days), the maintenance of this steady-state requires that HIV infect new target cells at a very high rate [161; 162].

Second, the viral population in an infected person is highly heterogeneous. In addition, the reverse transcription of viral RNA into DNA is prone to error, introducing on average one mutation for each viral genome transcribed. Thus, the high rate of HIV infection, combined with the high mutation rate that occurs during each cycle of infection, causes that patients have a complex and diverse mixture of viral *quasi-species*. These mutations can confer different advantages to the virus, such as a decrease of its susceptibility to an antiviral agent and the selection of the corresponding *quasi-species* (figure 46) [163].
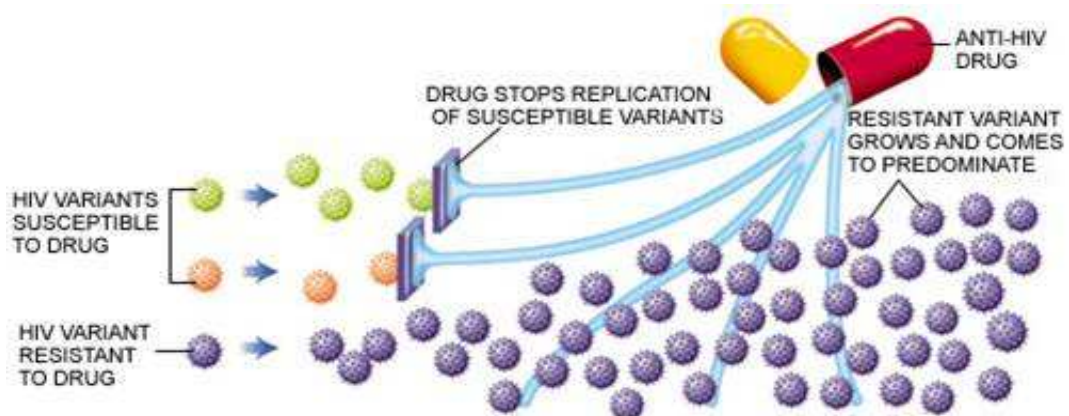

Figure 46. Mechanism of growth and predominance of resistant variants after pharmacological pressure [163].

In some cases the substitution of single amino acids is sufficient to produce high levels of resistance, but in other cases the gradual accumulation of additional mutations is required [164]. In context of the HAART therapy, resistance is most often the consequence, not the cause, of initial treatment failure, and a variety of mechanisms have been identified for the different classes of drugs. Resistance to nucleoside and nucleotide analogs is based on two distinct mechanisms:

1. the impairment of the incorporation of the analogue into DNA;
2. the removal of the analogue from the prematurely terminated DNA chain.

The incorporation of a nucleoside analogue into drug-sensitive viruses results in the termination of the viral DNA chain. Mutations in drug-resistant viruses, such as K65R, Q151M, M184V prevent the incorporation of the nucleoside analogue into the growing viral DNA chain, as shown in figure 47.
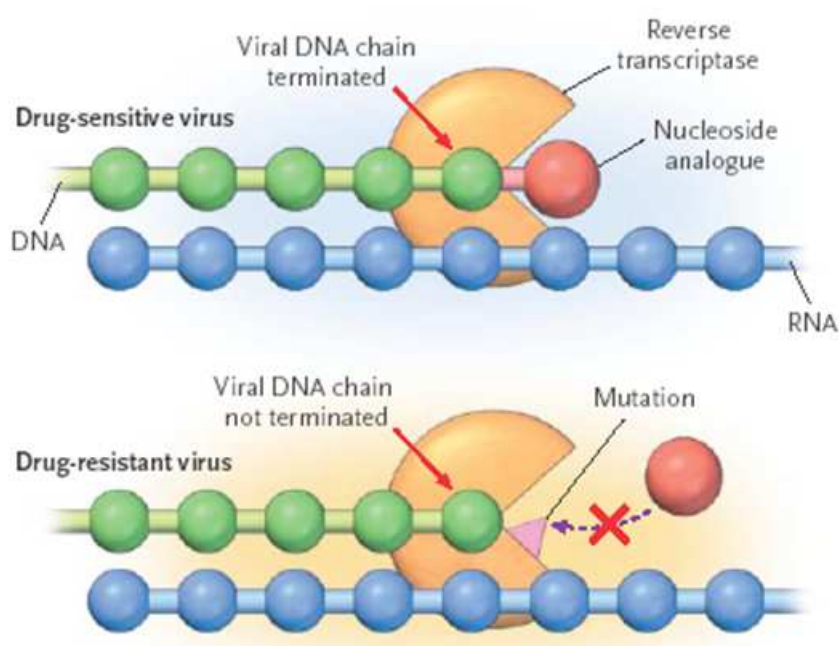

Figure 47. Mechanism of resistance by interference with the incorporation of a nucleoside analogue [101].

Removal of the nucleoside analogue from the terminated DNA chain is associated with a group of mutations commonly defined "thymidine analogue mutations". These mutations are most frequently selected after the failure of drug combinations including thymidine analogues, such as zidovudine and stavudine, but they can promote resistance to almost all nucleoside analogues. These substitutions, such as Y115F, M184V, L210W, promote resistance by fostering ATP- or pyrophosphate-mediated removal of nucleoside analogues

from the 3′ end of the terminated strand. ATP or pyrophosphate do not participate in the DNA-polymerization process, but, in presence of the thymidine analogue mutations, it can attack the phosphodiester bond that links the analogue to DNA, resulting in removal of the analogue (figure 48) [101].
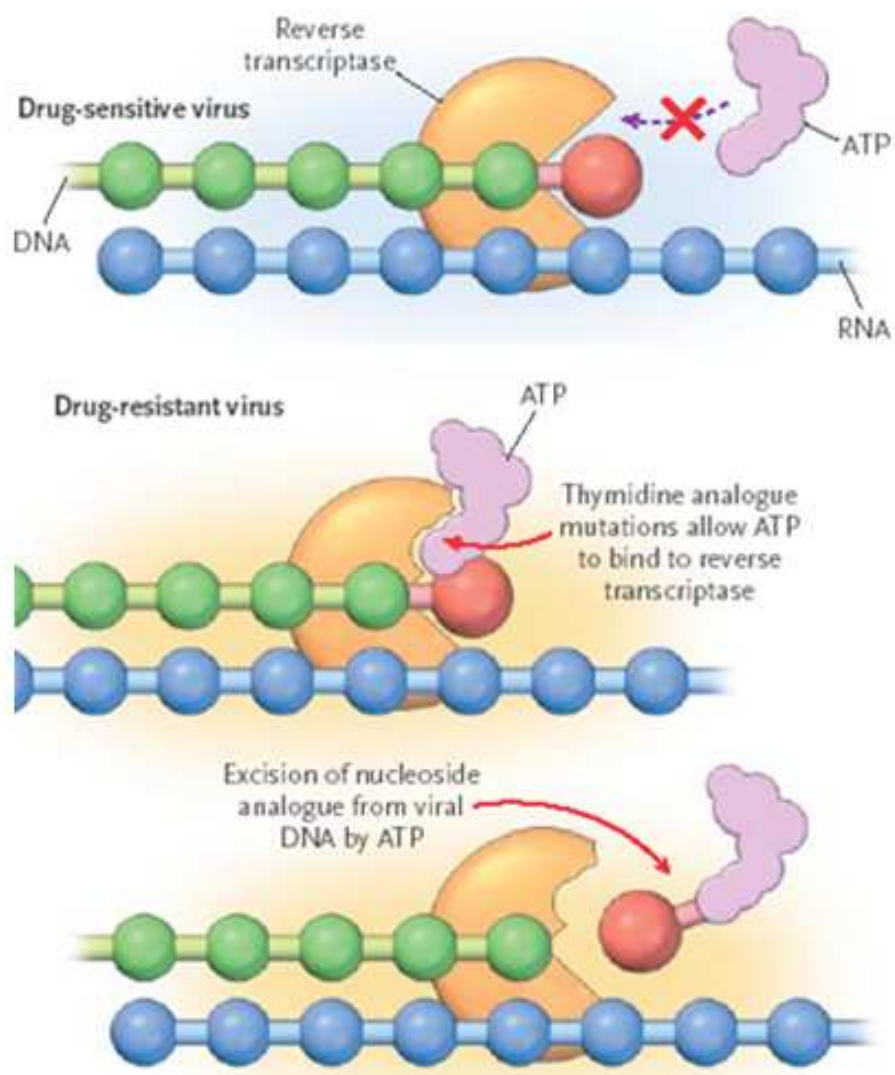
Figure 48. Mechanism of resistance by ATP-mediated excision of the nucleoside analogue [101].

Resistance associated to the failure of treatment with non-nucleoside RT inhibitors is caused by several mutations located in the hydrophobic pocket (NNRTI-BP) targeted by these compounds. These mutations, such as K103N, V106A, Y181C, Y188C are drug-dependent and reduce the affinity of the drug, allowing DNA polymerization to proceed normally (figure 49) [165].
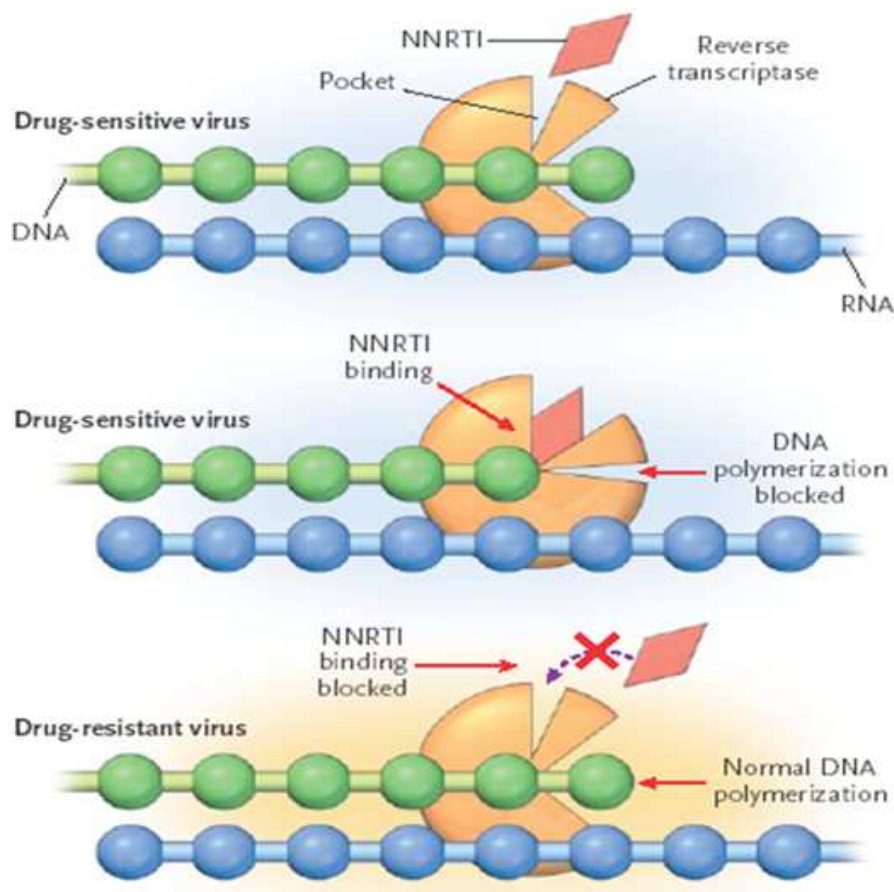
Figure 49. Mechanism of resistance of HIV to NNRTIs [101].

Protease inhibitors have a chemical structure that mimics that of the normal peptides that are normally recognized and cleaved by the PR. These compounds show a strong affinity for the active site of the HIV PR and inhibit the catalytic activity of the enzyme in a highly selective mode. Resistance to PIs is the consequence of amino acid substitutions occurring either inside the substrate-binding domain or at distant sites [166]. These mutations, such as G48V, I50V, V82A, I84V, directly or indirectly, can modify the number and the nature of the points of contact between the inhibitors and the PR, thus reducing their affinity for the enzyme. Resistance mutations in the PR result in a final enlargement of the catalytic site, as shown in figure 50.
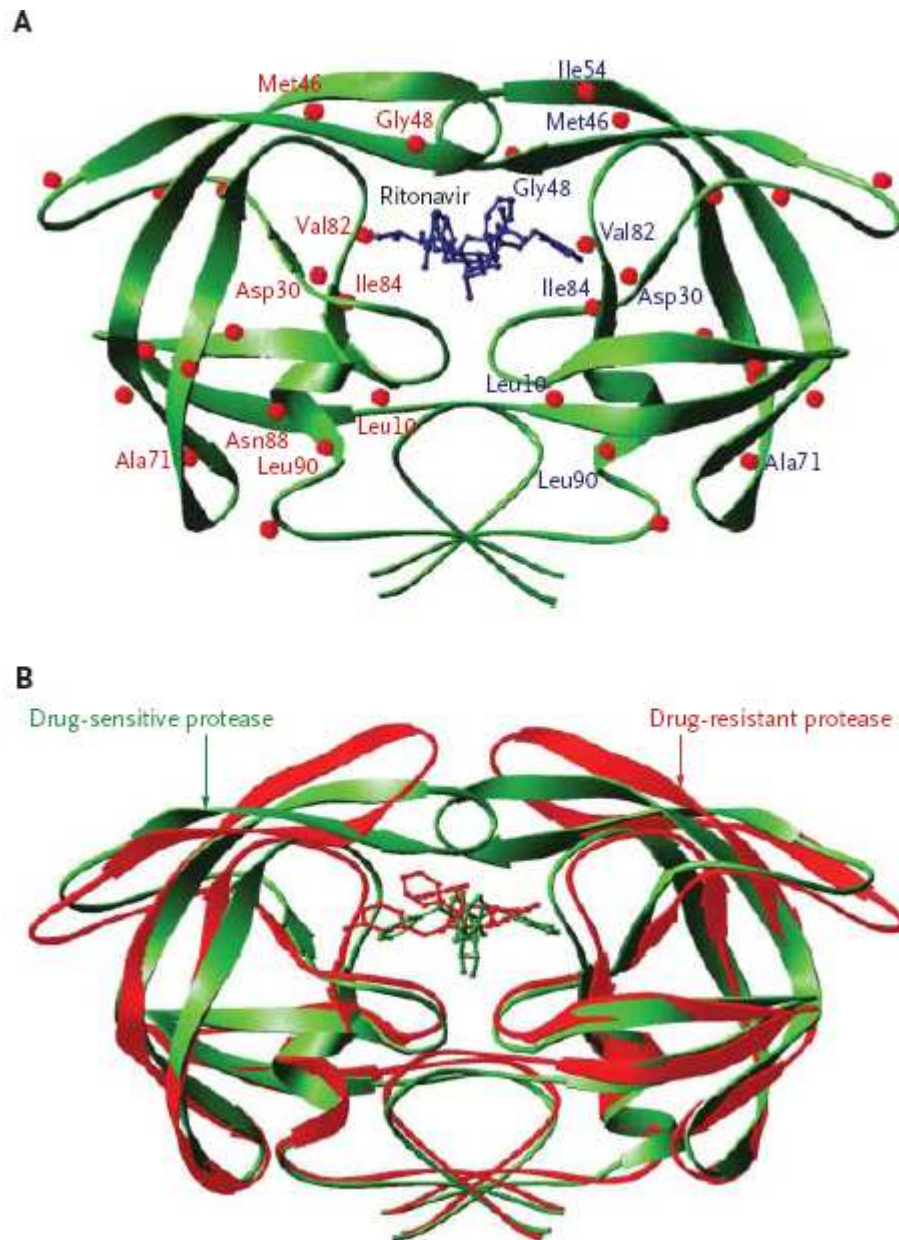
Figure 50. HIV-1 PR dimer binding with ritonavir showing the sites of residues mostly involved in resistance to PIs (**A**). The PR cavity with lopinavir in the context of either the sensitive PR (green) or the resistant PR (red) (**B**) [101].

HIV-1 enters target cells through a complex sequence of interactions between the HIV envelope glycoproteins (gp120 and gp41) and specific cell-surface receptors. Enfuvirtide, a 36-amino acid peptide derived from HR2, destabilizes this process by binding to HR1 and blocks the infectivity of HIV-1. Viral resistance to enfuvirtide usually depends on mutations placed in a stretch of 10 amino acids within HR1. Interestingly, substitutions in residues in gp41 outside HR1, and even changes in gp120, appear to be associated with significant differences in the susceptibility of the virus to enfuvirtide [167].

2. Goals of the work

The combined use of different antiretroviral drugs has substantially improved the clinical management of HIV-1 infection in terms of delaying disease progression, prolonging survival, and improving quality of life. Nevertheless, when antiretroviral therapy fails to be fully suppressive, new viral variants emerge, thus allowing HIV-1 to become resistant to one or more drugs by accumulating mutations, either alone or in multiple and complex patterns. For such a reason, understanding the genetic basis of resistance and cross-resistance is essential for optimizing the use of existing drugs and for designing new antiviral agents.

The research activities during my PhD course have been focused on three different areas related to HIV-1 therapy.

A new RT inhibitor pharmacophore model, useful for further drug design studies, has been defined applying the innovative GBPM computational method to 96 crystallographic models of HIV-1 reverse transcriptase.

With the purpose to elucidate the role of the novel reverse transcriptase mutation I135T in presence of the more common K103N, associated to NNRTIs resistance, molecular dynamics simulations were performed in order to evaluate the final effect onto the stabilization of the closed NNRTI-BP conformation.

The influence of two drug resistance-associated mutations, L33F and L76V, of HIV-1 protease has been evaluated with respect to lopinavir and atazanavir molecular recognition. With the aim to estimate the stability of its 6-helix bundle, the fusion glycoprotein gp41 conformational properties have been investigated in presence of V38A and N140I, known enfuvirtide resistance-associated mutations.

3. GBPM analysis

Drug resistance to RT approved inhibitors is a major problem affecting their clinical efficacy and is associated to different pathways of mutations. However, some residues were found to be highly conserved because of their essential role in heterodimer stabilization and in catalytic activity. Moreover the RT has been recently crystallized un presence of several antiretroviral drugs. A number of HIV-1 RT models are deposited into the Protein Data Bank (PDB) [168].

Recently a new computational method for creating pharmacophore models starting from detailed macromolecular structures was carried out, such as those obtained by crystallographic measurements. The method called GBPM (GRID based pharmacophore model) is based on the well established computational approach invented by Prof. Sir Peter Goodford and implemented in the GRID software [169]. The GBPM method has been successfully developed and applied for studying complexes of difference chemical nature [170].

In this section a study combining structural analysis of HIV-1 RT crystallographic models, clinical data about RT conserved residues and the computational method based on GRID maps is reported. Such analysis allowed to validate clinical results and to design by this computational approach new pharmacophoric models helpful in the drug discovery of more effective anti-RT inhibitors.

3.1. Materials and methods

The GRID-based pharmacophore model is created in a 6-step procedure, as shown in figure 51 [170]. The first one has the aim to carry out the PDB file pre-treatment and to obtain three different structures.

The second step performs the GRID calculation [169] with a certain probe onto the three subunit models.

In the third step the GRAB utility [169] is used to energetically compare two generic maps, keeping the same matrix dimensions, and to generate a resulting map that gives information about interaction areas. The fourth step is dedicated to the identification of most relevant interaction points of the obtained map using the MINIM utility [169]. With the aim to obtain a suitable model, these operations should be repeated using at least three different probes: the hydrophobic probe (DRY), an hydrogen bond acceptor (O) and an hydrogen bond donor (N1). In the fifth step the information obtained from the different probe experiments are simply merged into a preliminary pharmacophore model.
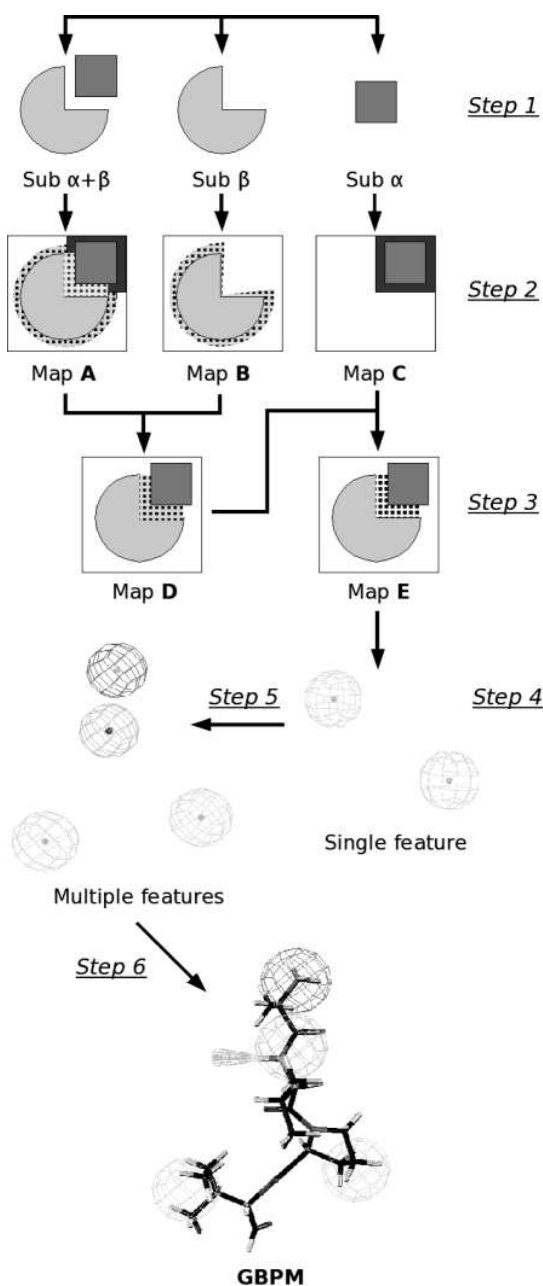
Figure 51. Graphical flow chart of the GBPM starting from a PDB complex. The bottom figure represents a generic feature-based pharmacophore model [170].

In this analysis GBPM analysis was carried out just up to the fifth step of the procedure, in order to validate clinical data, highlighting the most involved residues in the recognition areas.

To evaluate the most important interactions at reverse transcriptase drug sites, the program LigPlot [171] was used to select the most involved RT residues.

3.2. Results and discussion

3.2.1. Analysis of RT crystallographic models

The selection and the classification of the crystallographic models for the computational study was performed taking into account two different aspects.

The first one was related to the completeness of the available crystallographic model, i.e. number of mutations and deletions with respect to the wild type (WT) sequence.

The second aspect considered was the conformational effect onto the RT. The enzyme is conformationally very sensitive to the presence of nucleic acids and ligands, i.e. inhibitors. So we decided to divide our training set in three sections: a) the unbound enzyme models; b) RT structures with nucleic acid templates; c) co-crystals of RT and antiretroviral agents (NNRTI).

In this study we applied GBPM method [170] to the three identified groups of HIV-1 reverse transcriptase models. Since both subunits of the enzyme are proteins, in the first application we have performed the GBPM analysis considering p66 as receptor and p51 as ligand and viceversa (figure 52).
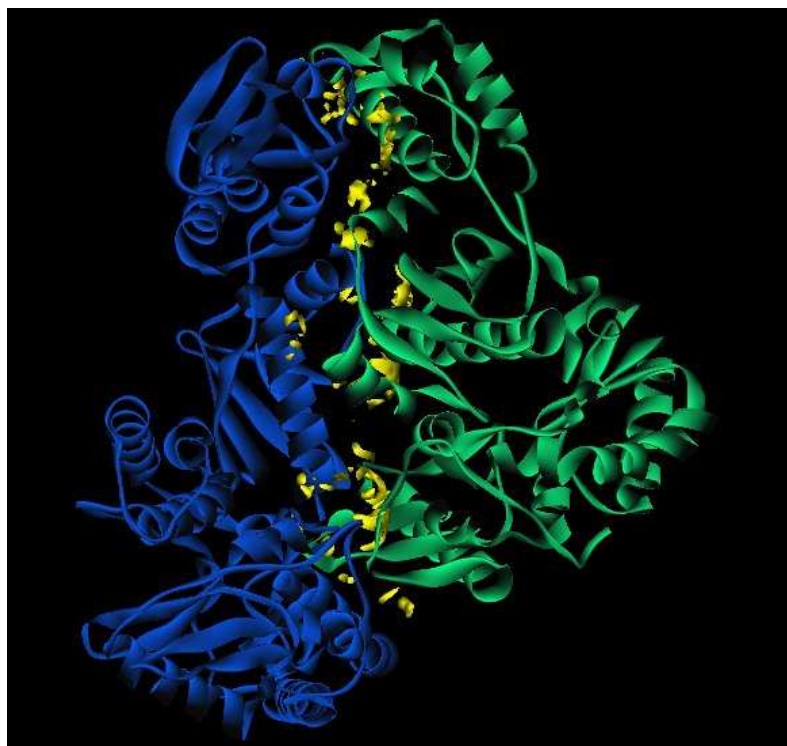

Figure 52. Representation of GBPM method applied to RT dimerization interface area. The subunit p66 is shown in blue cartoon and the subunit p51 is reported as green cartoon. In yellow are displayed the GRID maps obtained using the DRY probe (PDB model 2iaj) [Artese at al., unpublished data].

In the second case, in order to find out RT residues mainly involved in interactions with the nucleic acid, RT-template complex was used as α·β complex, the unliganded enzyme was identified as receptor and the nucleic acid was used as ligand (figure 53).
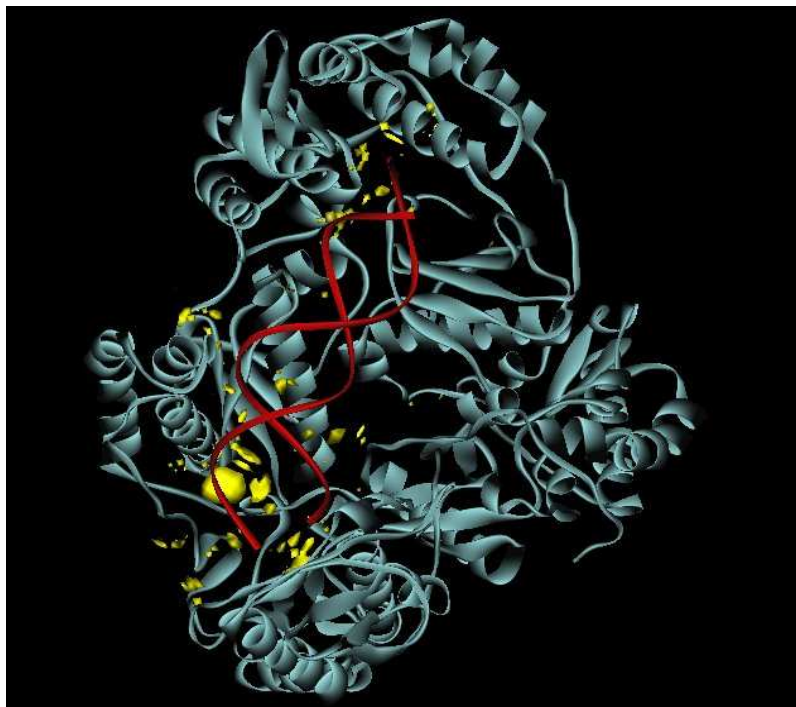
Figure 53. Representation of GBPM method applied to RT-nucleic acid complex. The enzyme is displayed as cyan cartoon, while the nucleic acid is showed in red ribbon. In yellow are displayed the GRID maps obtained using the DRY probe (PDB model 2hmi) [Artese at al., unpublished data].

In the third analysis we applied GBPM procedure to RT PDB models where the enzyme was complexed to different NNRTIs. In this study we identified the RT residues of the NNRTI-BP able to well recognize the inhibitor (figure 54).

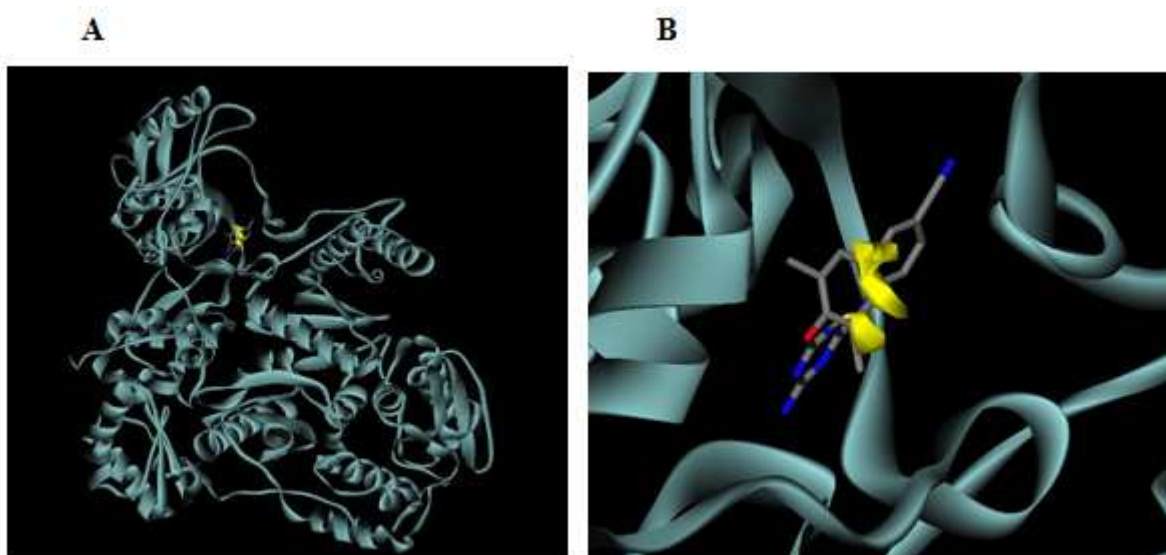**Figure 54.** Representation of GBPM method applied to an RT-NNRTI complexed model. In yellow are displayed the GRID maps obtained using the DRY probe (PDB model 1s9e). (**A**) RT complexed to the drug displaying all the enzyme; (**B**) The NNRTI binding pocket with the drug and the DRY maps around the aromatic portions [Artese at al., unpublished data].

These different crystallographic structures showed three dissimilar conformations, because of the presence of template or ligand.

All RT crystallographic models were downloaded from Protein Data Bank internet site [168]. In order to carry out superimposition of different PDB models, we used the Maestro program and we calculated the root mean square deviations (RMSD) using both the backbone and the α-carbon atoms. The crystallographic structures were selected considering the resolution of the models and classified on the basis of nucleic acid presence. We chose all PDB models with a maximum resolution value of 3.0 Å and we excluded those crystals that, compared with RT wild type, showed a large number of deletions and mutations.
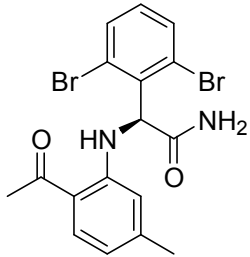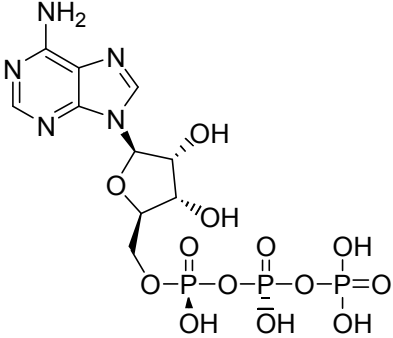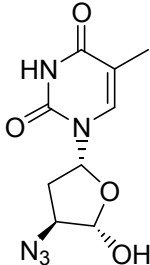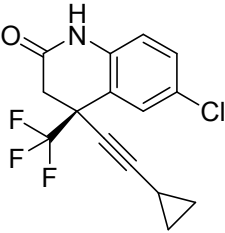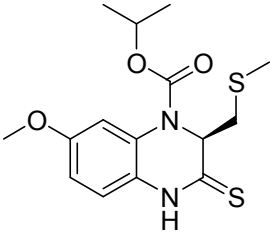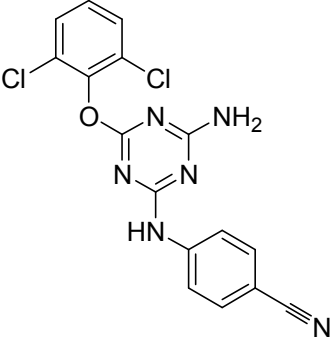
The structural analysis was carried out using 96 HIV-1 RT crystallographic models, as reported in Table 2.

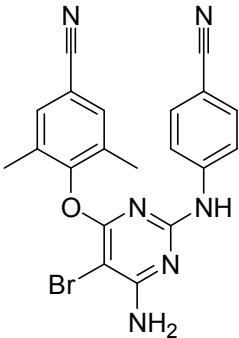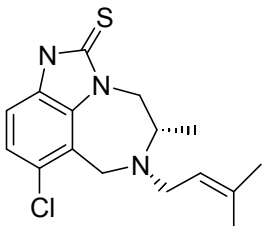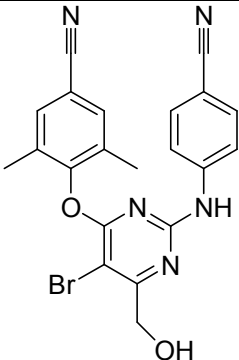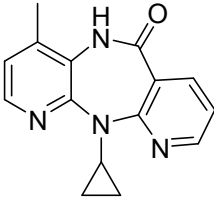| PDB	Res	Del	Mut	PDB	Res	Del	Mut	PDB	Res	Del	Mut
1tkz	2.81	43	11	1rt2	2.55	18	11	1c1c	2.50	24	11
1tl3	2.80	43	11	1klm	2.65	23	11	1n6q	3.00	2	15
1tkt	2.60	44	11	3hvt	2.90	5	13	1fko	2.90	18	12
1tl1	2.90	43	11	1bqm	3.10	4	12	1c0t	2.70	70	11
1tkx	2.85	25	11	1bqn	3.30	2	14	1c0u	2.52	22	11
1s1t	2.40	26	12	1tvr	3.00	2	15	1hni	2.80	2	14
1s1u	3.00	35	12	1uwb	3.20	2	14	1dtq	2.80	36	11
1s1v	2.60	26	13	1qe1	2.85	4	15	1dtt	3.00	20	11
1s1w	2.70	35	11	1rev	2.60	36	11	1hnv	3.00	2	14
1s1x	2.80	14	11	2i5j	3.15	8	14	1rt3	3.00	39	11
1jlq	3.00	27	11	1rtj	2.35	17	11	1rt4	2.90	24	11
1jla	2.50	13	11	2b5j	2.90	8	14	1rt6	2.80	22	11
1jlb	3.00	10	11	2ban	2.95	8	14	1rt7	3.00	25	11
1jlc	3.00	25	12	1rth	2.20	17	11	1rt5	2.90	24	11
1jle	2.80	34	11	1rti	3.00	17	11	2be2	2.43	8	14
1jlf	2.60	22	11	1vrt	2.20	35	11	2iaj	2.50	9	15
1jlg	2.60	29	12	1vru	2.40	34	11	1rt1	2.55	22	11
1lwe	2.81	17	11	2ic3	3.00	2	14	1t03	3.10	2	15
1lwc	2.62	53	11	1sv5	2.90	8	15	1n5y	3.10	2	15
1lw2	3.00	41	12	1tv6	2.80	22	14	1eet	2.73	18	14
1lwf	2.80	17	11	2b6a	2.65	2	14	1rtd	3.20	6	15
1jkh	2.50	9	11	1dlo	2.70	4	14	2opp	2.55	41	11
1lw0	2.80	47	11	1t05	3.00	6	15	2opr	2.90	21	11
1s6p	2.90	8	14	1r0a	2.80	2	15	2ops	2.30	24	11
1s6q	3.00	8	14	2hmi	2.80	2	14	1hqe	2.70	6	15
1s9g	2.80	18	14	2hnd	2.50	37	11	2opq	2.80	44	11
1suq	3.00	8	14	2hny	2.50	37	11	1hvu	4.75	6	14
1ep4	2.50	38	11	2hnz	3.00	43	11	1hmv	3.20	24	11
1fkp	2.90	19	12	1hpz	3.00	4	14	1ikv	3.00	4	14
1s9e	2.60	8	14	1hqu	2.70	4	15	1ikw	3.00	3	14
1fk9	2.50	41	11	1hys	3.00	7	14	1ikx	2.80	3	14
1c1b	2.50	24	11	1j5o	3.50	2	15	1iky	3.00	3	15

Table 2. Summary of the HIV-1 RT PDB models. PDB is the Protein Data Bank code, Res is the resolution in Å, Del and Mut are respectively the number of deletions and mutations in the p66 subunit with respect to the wild-type sequence [Artese at al., unpublished data].

The PDB models with the codes 1bqm, 1bqn, 1j5o, 1uwb, 2i5y, 1t03, 1n5y, 1rtd, 1hvu and 1hmv were excluded due to their bad resolution.

Among all available models, we selected 31 PDB with a number of deletions lower than 15 on the p66 subunit and some of them were co-crystallized with different ligands (Table 3).

	Ligand		Ligand
	1		2
	3		4
	5		6
	7		8
	9		10
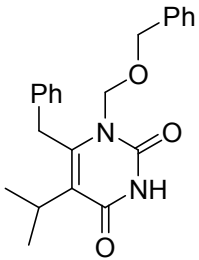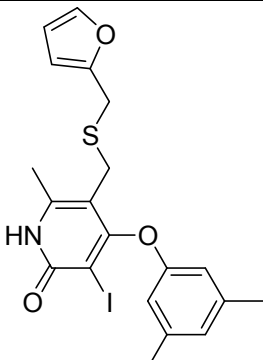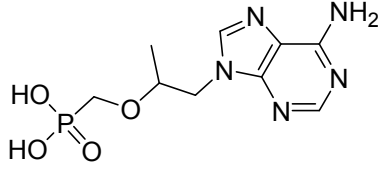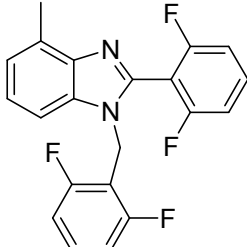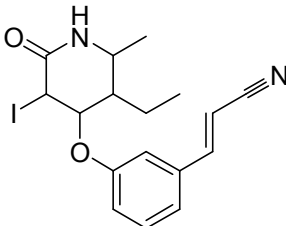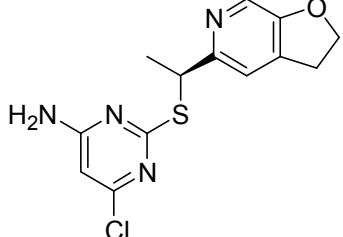
	11		**12**
	13		**14**
	15		**16**

Table 3. Chemical structures of ligands of the co-crystallized HIV-1 RT complexes used for the GBPM analysis [Artese at al., unpublished data].

The crystals 1hnv and 1tvr were complexed to the same inhibitor (TIBO) and presented the same resolution and number of deletions, so we chose the model with the lowest number of mutations (1hnv). As concerns the PDB models co-crystallized with the nevirapine, the choice of 1s1x was justified by the better resolution with respect to 1jlb.

Similarly we have proceeded for the models with the HBY inhibitor, selecting the best in resolution, i.e. 1hqu and discarding the 2ic3.

Among the identified models, some PDB were complexed to a class of NNRTIs with the same structural scaffold JANSSEN-R but different moieties, so we considered the three models (2b5j, 1sv5 and 1s9e) with the best resolution as examples of this class of drugs.

For the unbound enzyme we included in our study the models with the codes 1dlo, 1hqe and 1qe1 for a better resolution respect to the model 1hpz.

After the PDB selection step driven by the resolution, the number of deletions and mutations, we have divided our models into three groups characterized by similar RT folding properties. They are respectively related to unbound enzyme models (first group),

RT complexes with nucleic acid template (second group) and with NNRTI drugs (third group).

Since the emergence of resistance to HIV antiviral therapy represents the major limitation of a successful pharmacological approach [172], it is necessary to develop novel selective resistance-evading drugs, which bind to specific RT conserved sites inhibiting the enzyme in its main functions.

The list of the PDB models used for the analysis of the first group are reported in Table 4. In order to find out the most significant interactions at the p66-p51 subunits interface, only the protein residues were considered in this analysis.

PDB	Res	Ligand	PDB	Res	Ligand
1hnv	3.00	8	1hqe	2.70	-
1hni	2.80	1	1jkh	2.50	4
1jla	2.50	11	1sv5	2.90	7
2hmi	2.80	-	1suq	3.00	9
1n6q	3.00	3	1s1x	2.80	10
1hys	3.00	-	1dlo	2.70	-
1r0a	2.80	-	1qe1	2.80	-
1t05	3.00	13	2iaj	2.50	2
1s9e	2.60	6	1hqu	2.70	5
2be2	2.43	12	2b6a	2.65	14

Table 4. Crystallographic models used for RT-interface area structural analysis. PDB is the Protein Data Bank code, Res is the resolution in Å, Ligand is the drug co-crystallized with the enzyme [Artese at al., unpublished data].

With the aim to observe RT conformational changes as a result of the ligand binding, the RMSD analysis was performed with all PDB superimposed models. The maximum value was equal to 2.03 Å computed on RT α-carbons, 2.01 Å on the backbone and 2.03 Å on α-carbons in RT - DNA bound models, as reported in figure 55.
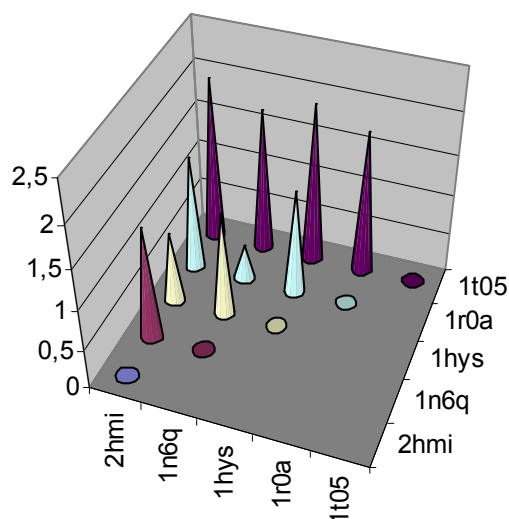
Figure 55. RMSD analysis of complexed RT-DNA models. On the x and z axis are reported the PDB codes of the analyzed models, while on the y axis is shown the value of the calculated RMSD expressed in Å [Artese at al., unpublished data].

The superimposition of the 5 PDB DNA-complexes highlighted few differences in the folding properties of the enzyme models.

For the second group (Table 5) we applied GBPM method to RT-template complexes to analyze HIV-1 RT conserved areas mainly involved in interactions with the nucleic acid.

PDB	Res	Ligand	Template
2hmi	2.80	-	DNA
1n6q	3.00	**3**	DNA
1hys	3.00	-	DNA/RNA
1r0a	2.80	-	DNA
1t05	3.00	**13**	DNA

Table 5. Crystallographic models used for RT-template structural analysis. PDB is the Protein Data Bank code, Res is the resolution in Å, Ligand is the drug co-crystallized with the enzyme, Template is the nucleic acid present in the model [Artese at al., unpublished data].

In order to analyze the residues of the NNRTI binding pocket (NNRTI-BP) responsible of the interactions with the drug, the method was applied onto 13 crystallographic models where RT was complexed to different inhibitors (Table 6).

PDB	Res	Ligand
1hnv	3.00	**8**
1suq	3.00	**9**
1jla	2.50	**11**
2b5j	2.90	**15**
1hni	2.80	**1**
1sv5	2.90	**7**
1ikw	3.00	**4**
1s1x	2.80	**10**
1ikx	2.80	**16**
1s9e	2.60	**6**
1hqu	2.70	**5**
2be2	2.43	**12**
2b6a	2.65	**14**

Table 6. Crystallographic models used for RT NNRTI binding pocket structural analysis. PDB is the Protein Data Bank code, Res is the resolution in Å, Ligand is the drug co-crystallized with the enzyme [Artese at al., unpublished data].

With the aim to select the most involved residues in the recognition step, we applied a geometric criterion using a maximum distance equal to 8 Å from the three analyzed targets and we finally compared them with respect to the identified conserved amino acids to carry out a statistical quantitative analysis.

3.2.2. Heterodimer interface area analysis

GBPM method was applied onto 20 HIV-1 RT unbounded crystallographic models with no inhibitors (Table 4) with the aim to identify residues mainly involved in heterodimer stabilization. In this analysis we adopted the DRY, N1 and O probes mimicking the hydrophobic, the hydrogen bond donor and the hydrogen bond acceptor, respectively (figure 56).
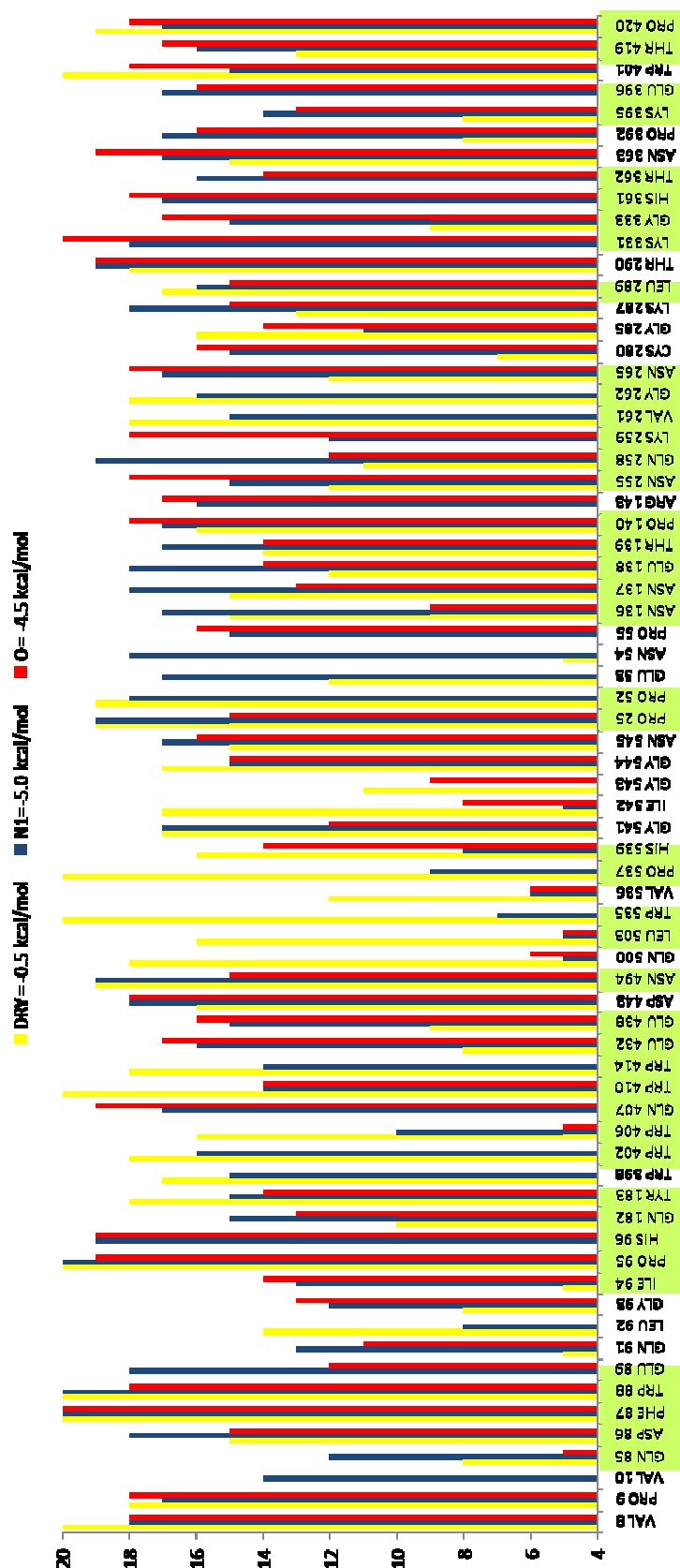
Figure 56. Statistical diagram deriving from analysis of unbounded PDB models after GBPM application. On the x axis are reported the most frequent and well mapped residues, on the y axis PDB models frequency [Artese at al., unpublished data].

A particular observation of the DRY experiments is for the aromatic residue Trp88, located onto the β5a strand [173]. It was found to be crucial for hydrophobic contacts at the interface area, emphasizing its important role in primer/template binding. Trp88 was able to establish interactions with Arg143 of p51, that was well mapped with O probe when p51 subunit was considered as receptor and resulted pivotal for whole protein stability (figure 57).
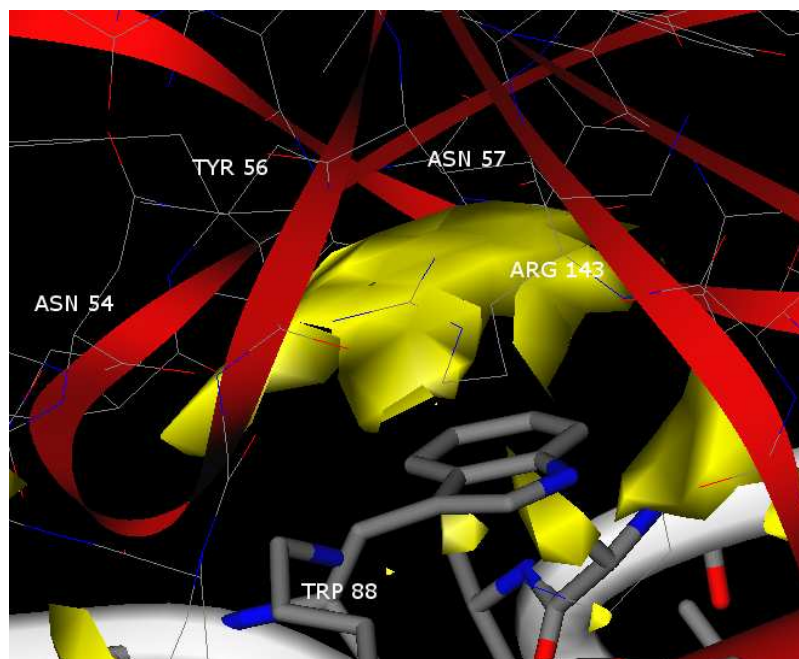
Figure 57. Interaction areas of p66 Trp88 mapped with DRY probe (yellow area) using PDB model 1hnv [Artese at al., unpublished data].

The LigPlot diagram reported in figure 58 clearly emphasizes the interaction of this residue in p66 and Arg143 in p51.
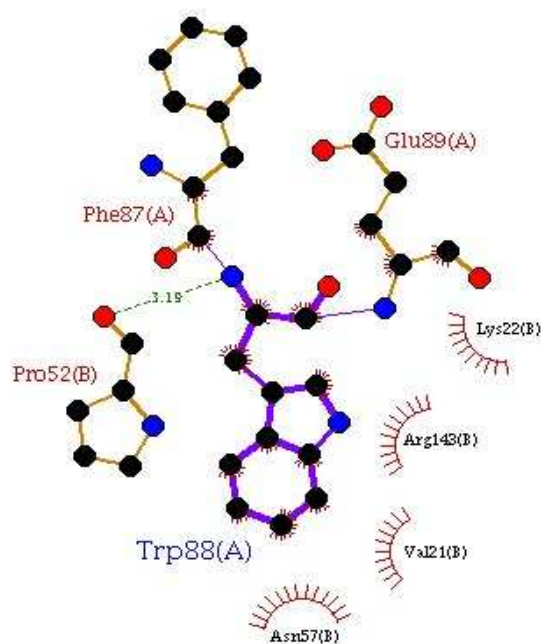
Figure 58. LigPlot schematic diagram relative to Trp88 of p66 using 1t05 PDB model. The green dashed line shows a hydrogen bond between Trp88 of p66 and Pro52 of p51, while hydrophobic contacts are showed with a red labelled arc [Artese at al., unpublished data].

Located in the *fingers* subdomain of p66 subunit, the high conserved Ile94 was well recognized with all the used probes, highlighting its importance in establishing contacts with the 2′-OH of RNA strand [174]. Analyzing maps obtained with O probe, two residues of p66, Pro95 and His96, resulted mainly involved in interactions with β7-β8 loop (Pro133-Pro140), that is crucial to stabilize RT heterodimeric form; this finding can be explained with the ability of these residues to make hydrogen bonds (p66 His96/p51 Asn136) [84].

Among all conserved residues identified by the DRY probe, Gln182 and Tyr183 were highlighted, confirming their essential role in RT polymerase activity [78]. Tyr183, that is part of YMDD motif (Y183-D186), was found to be really pivotal in the catalytic process, because of its hydrophobic interactions with the primer strand and its hydrogen bond with the cytosine O2 atom of nucleic acid. The mutation of Tyr183, well mapped with DRY probe, with non aromatic residues causes a remarkable decrease of the polymerase activity, confirming its importance in hydrophobic interactions [78].

The formation of the heterodimer is a two-step process. The first one involves a concentration-dependent association of the two subunits, followed by a slow isomerization/maturation step that involves interactions of the *fingers* and *thumb* of p51 with the *palm* and RNase H of p66, respectively. Specifically, residues 52–55 and 135–140

of the p51 *fingers* appear to interact with residues 85-96 of the p66 *palm*, while residues 280-295 of the p51 *thumb* subdomain appear to contact residues 536-545 of RNase H. In the first step of RT dimerization interactions between hydrophobic residues in the *connection* subdomains of p51 and p66 have been found to be crucial. This motif, that is well conserved, is believed to include residues W401-W410 of p66 and residues P392-W401 of p51. The *connection* subdomain is distinctive in having six tryptophan residues and a tyrosine between amino acid residues 398-414 [81].

The conserved residues 52-55 of p51 subunit resulted well mapped using all the probes, confirming their pivotal role in interactions with the p66 *thumb* subdomain. In fact the high conserved residues 85-93, located in the template grip of the *palm* subdomain, were well recognized with all the used probes, validating their essential role in interactions with the DNA sugar-phosphate backbone of the template strand [172] and in contacts with the p51 *fingers* [81].

Due to its ability to give and to accept hydrogen bonds, β7-β8 loop of p51 was well mapped using N1 and DRY probes, with a good recognition of residues Asn136, Asn137, Glu138, Thr139 and Pro140 (figure 59).
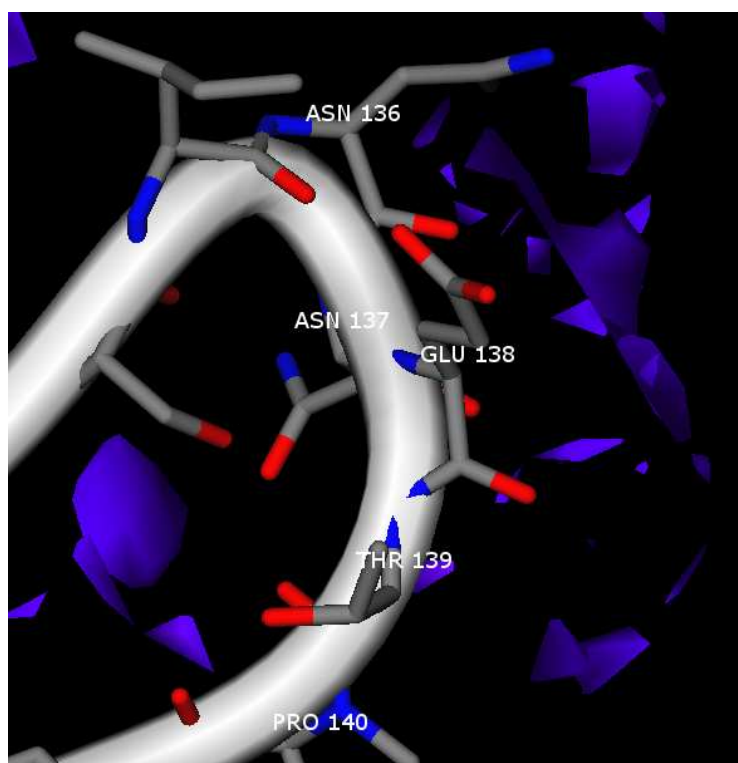
Figure 59. Interaction areas of p51 β7-β8 loop well mapped with N1 probe (blue areas), using PDB model 1hnv [Artese at al., unpublished data].

The residues of p51 found to make contacts with p66 resulted Asn136, Asn137, Glu138 and Pro140. This observation was confirmed by site-directed mutagenesis studies demonstrating the crucial role of Asn137 and Pro140 in order to maintain RT catalytic activity [84].

Asn136 of p51 was able to establish hydrogen bonds with His96 and Pro95 of p66 and this interaction was validated by the opposite probe properties when each subunit was considered as receptor. As shown in figure 60, Asn136 was well recognized with N1 probe when p51 was used as receptor and was placed in front of His96 of p66, well mapped with O probe when p66 was considered as receptor (hydrogen bond donor/acceptor interaction).

Figure 60. Interaction areas of p51 β7-β8 loop mapped with N1 probe (blue areas) in presence of p66 subunit using PDB model 1hnv [Artese at al., unpublished data].

The β7-β8 loop of p51 was found to interact with other residues of p66, as Tyr181, Gln182 and Tyr183, essential for RT polymerase activity and for NNRTI binding pocket formation [175]. GBPM assumed hydrogen bonds between Asn136 of p51 and His96 of p66 were also validated using LigPlot program [171], as reported in figure 61.
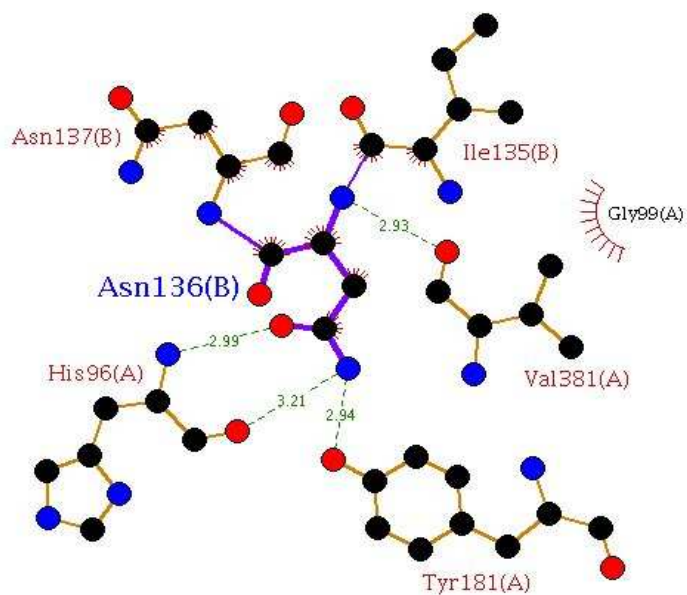
Figure 61. LigPlot schematic diagram relative to Asn136 of p51 using 1hnv PDB model. The green dashed lines show the hydrogen bonds between Asn136 of p51 and His96, Tyr181 and Val381 of p66, while hydrophobic contacts are showed with a red labelled arc [Artese at al., unpublished data].

Located on the *thumb* subdomain of p51, residues Cys280, Gly285, Lys287, Leu289 and Thr290 resulted well recognized in all the obtained maps, confirming their ability to establish favourable interactions with residues 536, 537, 539 and 541-545 of RNase H region [80], found to be crucial using p51 as receptor.

In perfect agreement with the literature [81], GBPM method allowed to identify the "tryptophan rich motif" at RT interface area using DRY probe (figure 62).
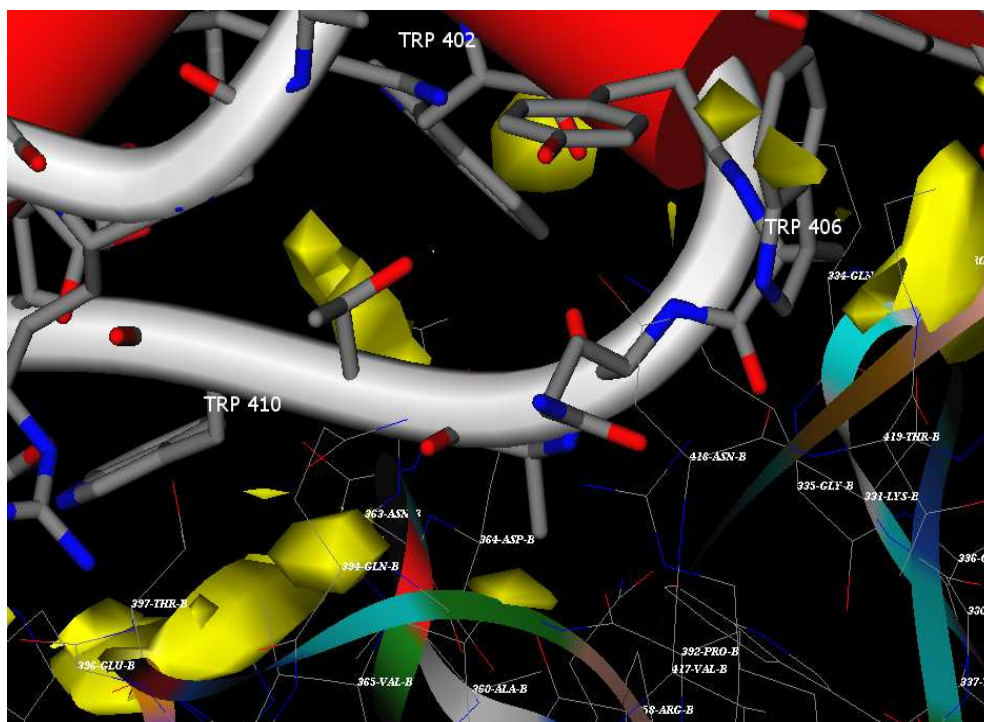
Figure 62. Interaction areas of p66 "tryptophan rich motif" well mapped with DRY probe (yellow areas) using PDB model 1hnv [Artese at al., unpublished data].

Such a result was confirmed by mutagenesis studies showing that, after mutations of Trp401 and Trp414 of p66 subunit, RT loses its activity [81].

Part of this same region, also Gln407 was found highly conserved and well mapped with N1 and O probes. This amino acid, together with tryptophan rich motif, is directly involved in the association between p66 and p51 subunits [176].

GBPM identification of highly conserved tryptophan rich motif in both subunits was well recognized with DRY probe. Using p51 subunit as receptor, Trp401, that is part of "tryptophan rich motif", was found to make hydrophobic contacts with p66 tryptophan residues already identified when p66 was considered as receptor (figure 63).
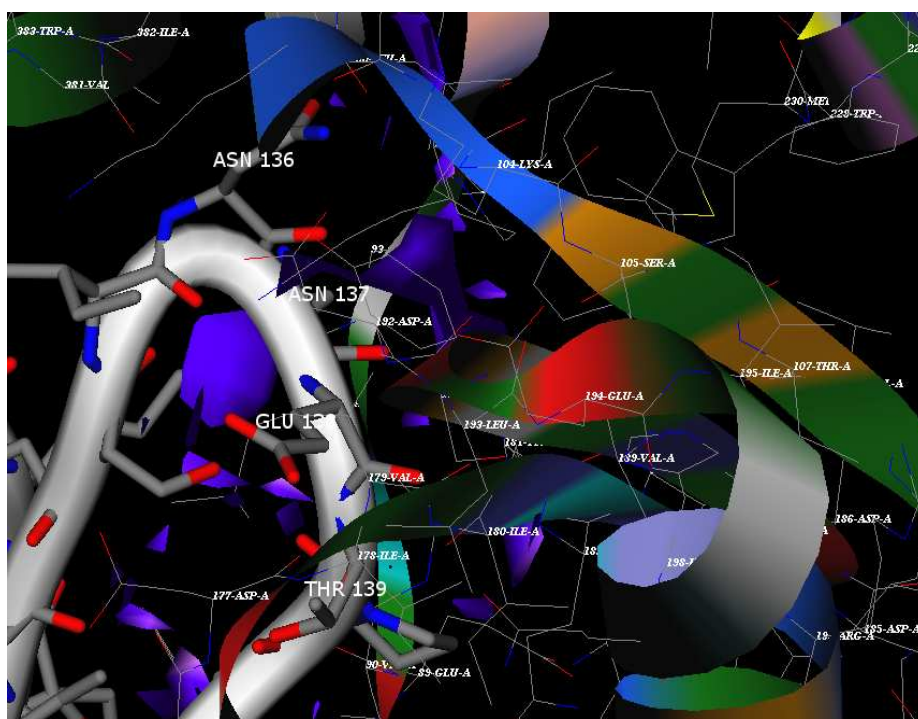
Figure 63. Interaction areas of p51 Trp401 mapped with DRY probe (yellow areas) in presence of p66 subunit using PDB model 1hnv [Artese at al., unpublished data].

A second region involved in heterodimer formation, crucial in protein-protein interactions, was represented by a repeat motif of 8 leucine residues (Leu282-Leu310).

After GBPM application, Leu289 of p51 resulted well recognized with DRY probe through a hydrophobic interaction with Phe61 of p66 and was found to be essential for the stability of the single subunits [177].

Located on the RNase H domain of the enzyme, the two conserved glutamic residues in position 432 and 438 were identified analyzing all the maps. Mutation of these amino acids impaired the *in vivo* p66 processing, confirming the critical role, especially for Glu438, in proper heterodimerization and function of virion-associated RT [178]. Always placed in this subdomain, residues Asp443 and Asn494, whose crucial role in the *in vivo* processing of a Pol precursor molecule to the p66/p51 heterodimer form was validated by mutagenesis experiments, resulted well recognized with all the used probes [179].

Two proline residues, Pro25 and Pro52, were well mapped by the hydrophobic probe, confirming their pivotal role in polymerase and RNase H activities [72; 76].

Part of the *thumb* subdomain of p51 subunit, the αH (Val254 to Ser268) is in contact with the RNaseH domain of p66, suggesting a substantial contribution of this region to dimer stabilization [172]. The analysis of GBPM obtained maps confirmed the important role of the conserved residues Asn255, Gln258, Lys259, Val261, Gly262 and Asn265, well recognized with all the probes.

In particular the maps obtained with O probe put in evidence the highly conserved residue Asn255, validating its importance in both enzyme dimerization and primer/template binding [75]. As reported in figure 64, Asn255 of p51 mapped with carbonylic oxygen was placed in front of Glu432 of p66, well recognized with N1, confirming maps complementarity.
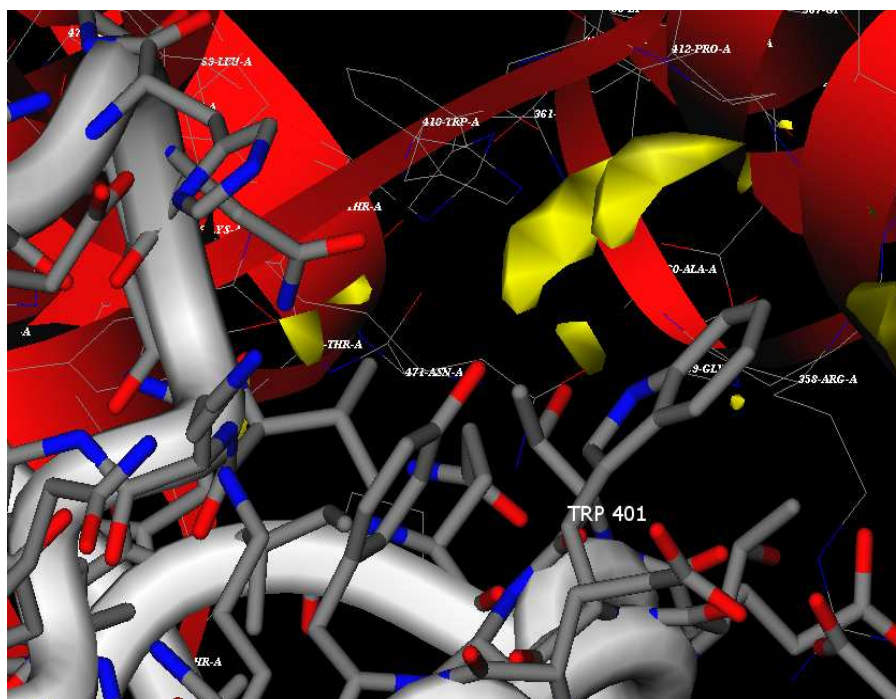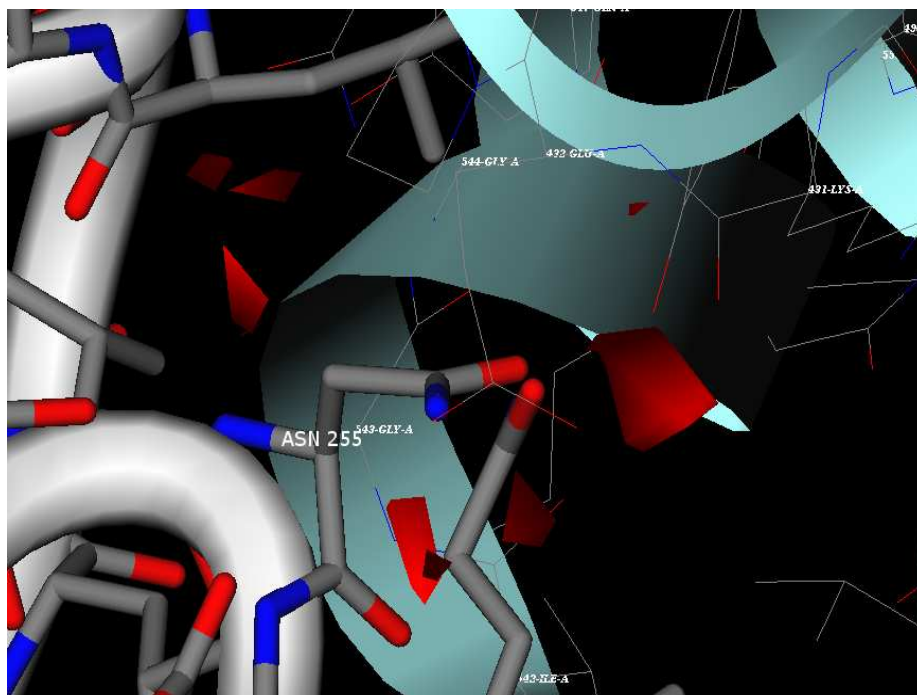

Figure 64. Interaction areas of p51 Asn255 mapped with O probe (red areas) in presence of p66 subunit using PDB model 1hnv [Artese at al., unpublished data].

Analyzing maps obtained using N1 and O probes, a statistically significant residue resulted Lys331. This amino acid is able to stabilize RT heterodimer through hydrogen bonds with the backbone of residues Trp402, Tyr405 and Gln407 of p66, confirming that mutations placed at this codon can negatively influence the dimerization step [180].

Part of the *connection* subdomain, another recognized residue was Gly333 of p51 subunit, well mapped with all the used probes. The variable Gly333 is located far from the polymerase active site, but between the *palm* region and the RT carboxy terminus. This residue is also positioned close to the base of the *thumb* region, which is involved in template-primer interactions. Thus, it is possible that changes at this codon alter the positioning of the *thumb* region and subsequently reposition the template-primer in the active site of the polymerase domain. In fact mutagenesis studies revealed that in patients taking AZT and 3TC polymorphism at codon 333 was responsible for facilitating dual resistance [181].

Two highly conserved residues of p51 subunit, His361 and Thr362, are known to contact the phosphate backbone of the DNA primer strand near to RNase H cleavage site. Mutagenesis analysis demonstrated that most of the mutations in the *connection* domain have relatively small effects on the virus titer; however the substitutions His361Ala and Thr362Ala decreased the virus titer respect to that of the WT virus. Moreover the mutation His361Ala also had a significant effect on RNase H cleavage, since His361 contacts the nucleic acid through the p66 subunit [79].

As validation of these data, both the analyzed residues were found well recognized with N1 probe, confirming their capability to establish hydrogen bonds with the template (figure 65).
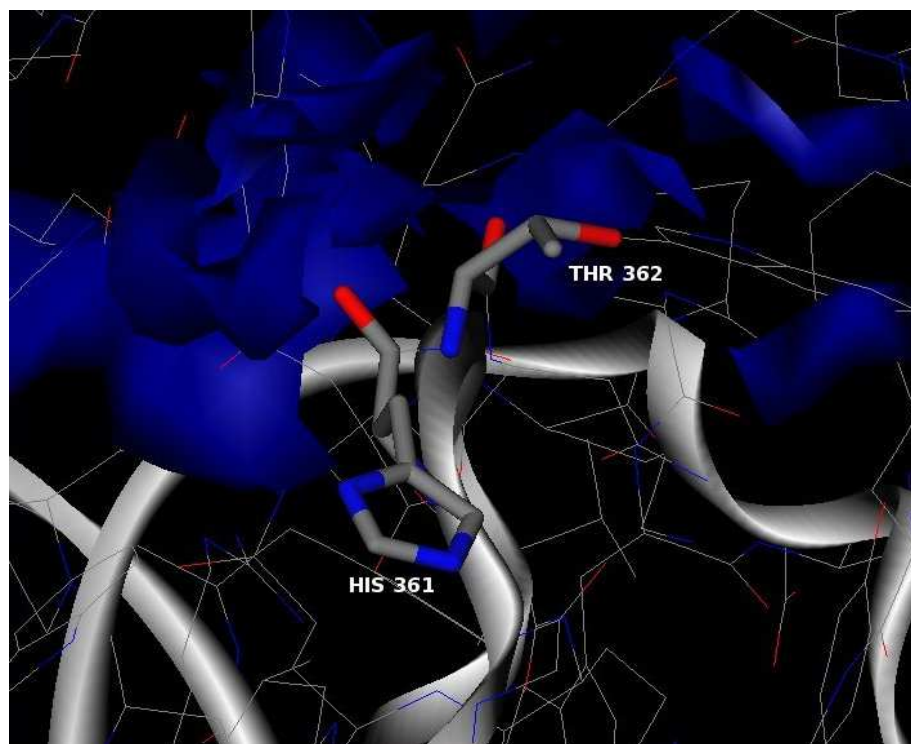
Figure 65. Interaction areas of p51 His361 and Thr362 well mapped with N1 (blue areas), using PDB model 1jkh [Artese at al., unpublished data].

Also the slightly variable Lys395 and Glu396, present in *connection* subdomain, were found to be pivotal for primer-template binding by hydrogen bonding interactions. Both residues are referred to as the RNase H primer grip and lie either in the p51 or p66 subunit of RT. *In vivo* and *in vitro* studies of the RNase H primer grip indicate that these amino acid residues are important for the proper binding and positioning of the nucleic acid to RT. Lys 395, well recognized with N1 probe, resulted crucial to form hydrogen bonds and

other interactions with the primer strand; in fact its substitutions decreased the viral replicative capacity less than two to three times compared to that of the WT.

Glu396 was well mapped with N1 and O probes, validating mutagenesis analysis in agreement with the observation that the single mutation Glu396Ala causes the loss of HIV-1 polymerase function due to the inefficiency of RNase H to recognize primer strand [181]. Mutational analysis of HIV-1 reverse transcriptase revealed that two adjacent residues of p51 subunit located at position 419 and 420 resulted crucial in forming properly folded and, hence, enzymatically active RT. It was observed that the double mutant 419/420 was associated to the loss of RNase H activity because these mutations are positioned close to the critical "hinge" region that connects the putative DNA polymerase and RNase H domains [182]. As validation of these experimental data, the two highly conserved residues resulted well mapped with the used probes, confirming their pivotal role in the interactions with the tryptophan rich motif of p66 subunit and with the hydrophobic amino acids Leu503 and Trp535 placed at dimer interface.

Other highly conserved residues part of the RNase H subdomain were Gln500 and His539 of p66 subunit, both well recognized with DRY probe. Gln500 and His539 are involved respectively in ribose and phosphate contacts to the RNA primer [183].

It was demonstrated that mutation His539Asn is responsible of a reduced RNase H activity and an increased AZT resistance respect to wild-type due to a reduction of RNA degradation, with a consequent increase of the time for AZT-MP to be excised from the terminated primer and for polymerization to resume on an intact template [184].

3.2.3. Analysis of RT-DNA complexes

In order to evaluate the most involved interaction areas between RT and nucleic acid templates, the GBPM method was applied for the 5 RT-template complexed PDB models using the same probes of the first analysis and taking into account the hydrophilic nature of nucleic acid.

In this analysis the GBPM maps were able to identify those residues mostly involved in interactions with the nucleic acid (figure 66).
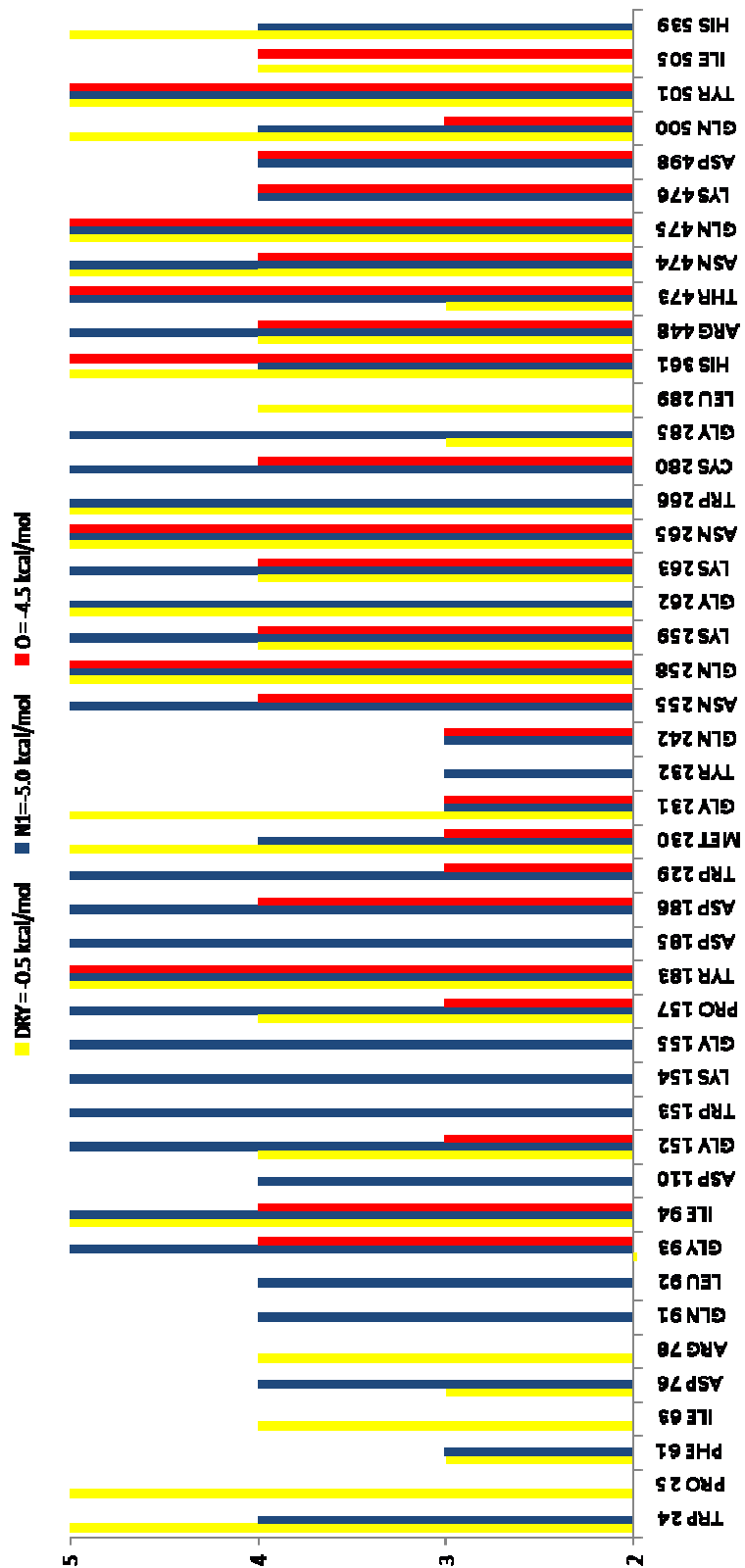
Figure 66. Statistical diagram deriving from analysis of 5 RT-template complexed PDB models after GBPM application. On the x axis are reported the most frequent and well mapped residues, on the y axis PDB models frequency [Artese at al., unpublished data].

The most important conserved residues in DNA recognition resulted: Trp24, Pro25, Phe61, Ile63, Asp76, Arg78, Gln91-Ile94, Asp110, Gly152-Gly155, Pro157, Tyr183, Asp185, Asp186, Trp229-Tyr232, Gln242, Asn255, Gln258-Lys263, Asn265, Trp266, Cys280, Gly285, Leu289, His361, Arg448, Thr473-Lys476, Asp498, Gln500, Tyr501, Ile505 and His539.

Some of these amino acids are well recognized within the minor groove of DNA, as reported in figure 67.
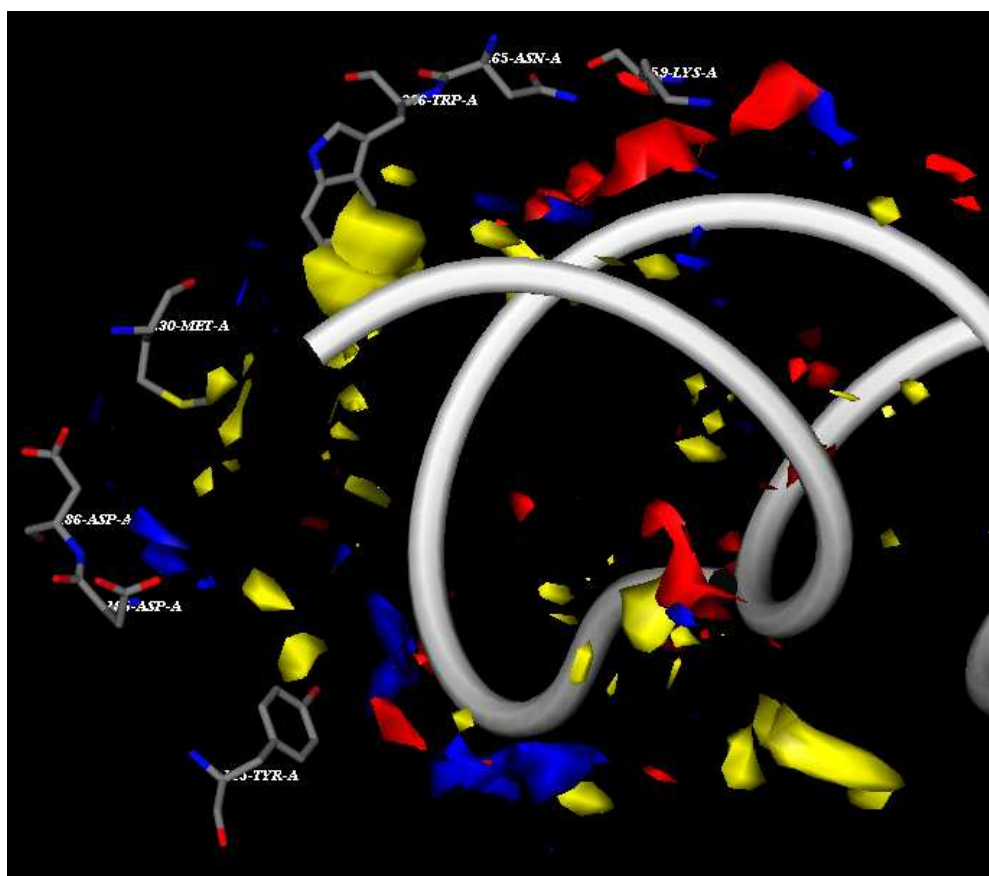
Figure 67. Minor groove DNA recognition with some highly conserved residues mapped with three different probes. In yellow are reported DRY maps, in blue N1 maps and in red O maps and the most important residues are showed in politubes [Artese at al., unpublished data].

Located on the *fingers* subdomain of p51, the aromatic amino acids Trp24 and Phe 61, involved in stabilization of reverse transcriptase, were well mapped with DRY probe, putting in evidence their role in hydrophobic interactions; in fact selective mutation of these residues on p51 to a glycine dramatically alters the stability of the RT-heterodimer [185]. Part of the same area, Pro25 was found in all DNA-RT models using DRY probe, confirming its importance in polymerase and RNase H activities [72].

Placed on the *fingers* subdomain of p66, amino acids Asp76 and Arg78 were well recognized using DRY and N1 probes, validating their important role in polymerase activity [172].

DRY maps highlighted also the conserved Trp229, that is part both of NNRTI binding pocket and of primer grip region. Mutagenesis studies confirmed the crucial role of Trp229 aromatic ring for RT activity, correct folding of the protein and NNRTI-BP formation [186].

The same probe was useful to recognize a strategical residue, Ile63, whose mutation is responsible to partially restore the DNA polymerase activity of a Trp229Tyr mutant RT [187]. Such a data puts in evidence that designing a drug able to interact with Ile63 a possible escape mechanism of HIV is not allowed.

Well mapped with the donor hydrogen bonds probe, residues Gln91-Gly93 were recognized within the template grip region, confirming their importance in positioning the template strand near the polymerase active site [188]. Also Ile94, already recognized at interface recognition area, resulted well mapped with all the used probes, highlighting its importance in establishing contacts with the 2′-OH of RNA strand [173]. In this region were also identified conserved residues Gly152-Gly155, well recognized with N1 probe, and Pro157, well mapped with all the used probes. Amino acids Gly152-Gly155 have an essential role in DNA interactions and RNA hydrogen bonding network [174]. Trp153 is located next to the conserved LPQG motif, at the juncture of the *palm* and *fingers* subdomains. Ala substitutions at this codon substantially hindered the polymerase function of the enzyme, impacting the geometry of the dNTP binding pocket [189].

Pro157, that constitutes the N-terminus of the αE of the *palm* subdomain, was found to be pivotal in maintaining the structural integrity of β8-αE connecting loop and the appropriate conformation of the template grip [190].

Using the hydrogen bonds donor N1 probe, residues Asp110, Tyr183, Asp185 and Asp186 were put in evidence in obtained maps, confirming the crucial interactions of these amino acids with DNA phosphates and nucleotide bases. In fact the side chains of aspartic residues establish hydrogen bonds with 3′ primer hydroxyl group, facilitating the nucleophilic attack during polymerization step, while Tyr183 was found to be pivotal in the catalytic process due to its ability to establish hydrophobic interactions with the primer strand [191].

YMDD motif conformation is highly influenced by the structure of primer grip region, formed by β12-β13 sheets in the p66 *palm* subdomain, in close proximity to nucleotides at

the 3′ primer terminus. This structural element is involved in template translocation. Mutagenesis experiments demonstrated that substitution of alanine residues in the β12-β13 hairpin (Trp229, Met230, Gly231 and Tyr232) in p66 alters both DNA polymerase and RNase H activities [78]. Such a result was confirmed by maps analysis, where the importance of the interactions between Met230 and Gly231, well recognized with all the used probes, and nucleic acid was highlighted. Other studies validated that mutations of Tyr232, well mapped with N1 probe, cause an incorrect RNase H cleavage due to the crucial role of this aromatic amino acid in π-π interactions with the template primer and in DNA-dependent DNA polymerase function [192].

Due to its ability to establish hydrogen bonds, the conserved Gln242 resulted well mapped with N1 and O probes, highlighting its essential role for RT structure and function [172].

Located on the polymerase domain of the *thumb* subdomain, residues Gln258, Lys259, Gly262 and Lys263 were well recognized in all obtained maps. It has been shown by alanine-scanning mutagenesis that these amino acids, that are part of the helix clamp, resulted critical for template-primer affinity [193]. In fact cross-linking of DNA to an RT with the Q258C mutation in both subunits decreased the stability of the complex and caused the dissociation of the p51 subunit from the cross-linked p66-DNA complex; this event is due to the disruption of the interactions between Gln258 side chain on the p51 subunit (*thumb* subdomain) and the RNase H subdomain of p66 [194].

Further site-directed mutagenesis studies demonstrated that some residues of the region defined by Trp251-Tyr271, highly conserved in treated patients [172], resulted essential for enzyme stability and catalytic activity, like Trp266, Asn255 and the aspartic residues of the active site [195].

These amino acids, especially located in p66 conserved regions, were found to interact, through Trp266, with DNA terminal sequence in the minor groove with hydrophobic contacts. This finding is in perfect agreement with structural data, thus validating GBPM method reliability and underlining Trp266 importance in interaction with the primer-template. This residue is part of the MGBT (Minor Groove Binding Track) region crucial to establish contacts with the DNA in the minor groove. The loss of these interactions is associated with a remarkable decreasing in DNA binding affinity and in frame shift process [195]. Mutagenesis analysis confirmed Trp266 pivotal role due to hydrophobic contacts with the nucleic acid, as found with GBPM method with the DRY probe (figure 68).
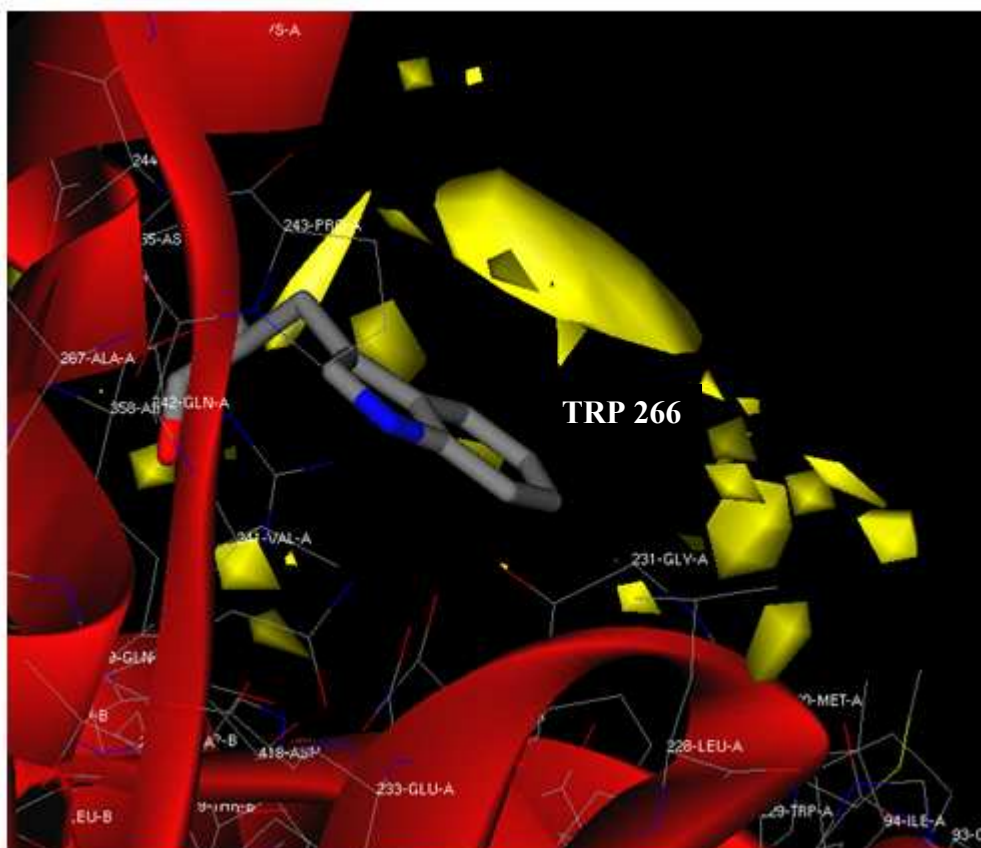
Figure 68. Interaction areas of p66 Trp266 well mapped with DRY probe (yellow area), using PDB model 2hmi [Artese at al., unpublished data].

Also residues Asn255 and Asn265, placed close to MGBT region, resulted well recognized with N1 probe, emphasizing their importance in interactions with the sugar of nucleic acid during translocation step [196] (figure 69).
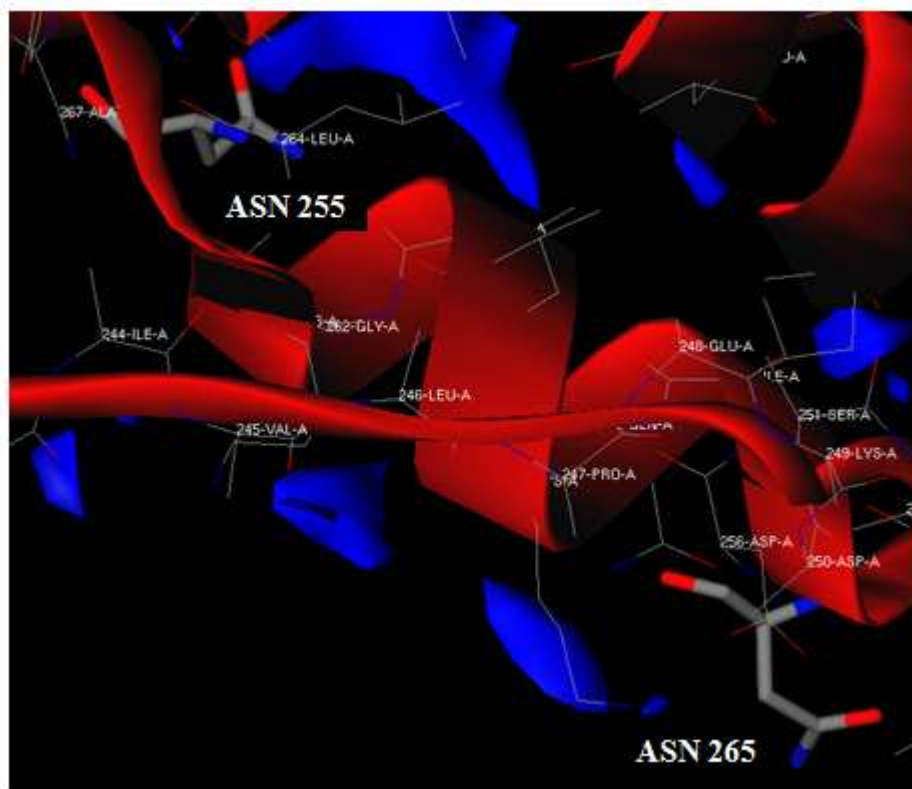
Figure 69. Interaction areas of p66 Asn255 and Asn265 well mapped with N1 probe (blue area), using PDB model 2hmi [Artese at al., unpublished data].

Located in the α helix I of the *thumb* subdomain of the polymerase domain, Cys280 is a highly conserved residue and resulted well mapped with N1 and O probes. This amino acid is involved both in DNA polymerase and RNase H activities; strand transfer assays and RNase H analysis showed that viruses with mutated Cys280 residue present biological disadvantages [197].

Placed in the leucine hepta-repeat motif (L282-L310), conserved residues Gly285 and Leu289 were recognized using N1 and DRY probes, respectively. Mutagenesis analysis validated the important role of Leu289 of p51 because, through a hydrophobic interaction with Phe61 of p66, resulted essential for the stability of the single subunits [177].

Located on the *connection* subdomain of p66 subunit, residue His361, that contacts the phosphate backbone of the DNA primer strand, was well mapped with all the used probes. It was recently observed that the mutation His361Ala decreased the virus titer to about 25% of that of the WT virus, suggesting that its interactions are crucial for RNase H cleavage and for virus infectivity [181].

In the RNase H subdomain are located Arg448 and Gln500, well recognized in all obtained maps. Both residues are responsible to establish crucial contacts with base, sugar and phosphates of RNA primer [182].

After GBPM application, RNase H primer grip residues Thr473-Lys 476, Tyr501 and Ile505 were identified with all the used probes, confirming some mutagenesis data showing their essential role in polymerization-independent RNase H activity, dimerization and DNA strand transfer [198].

The RNase H domain is folded into a 5-stranded mixed β-sheet flanked by 4 asymmetric α-helices; the strong conservation of structure suggests that this region is very stable. In the active site is located the well recognized Asp498, a crucial residue implicated in catalysis, whose replacement generates an unstable enzyme (figure 70) [190]. In addition His539, which is found in a six-residue loop connecting β5 and αE, also constitutes part of the active site; this residue resulted well recognized with N1 and DRY probes, confirming its important role in interactions with the phosphates of the RNA primer [199].
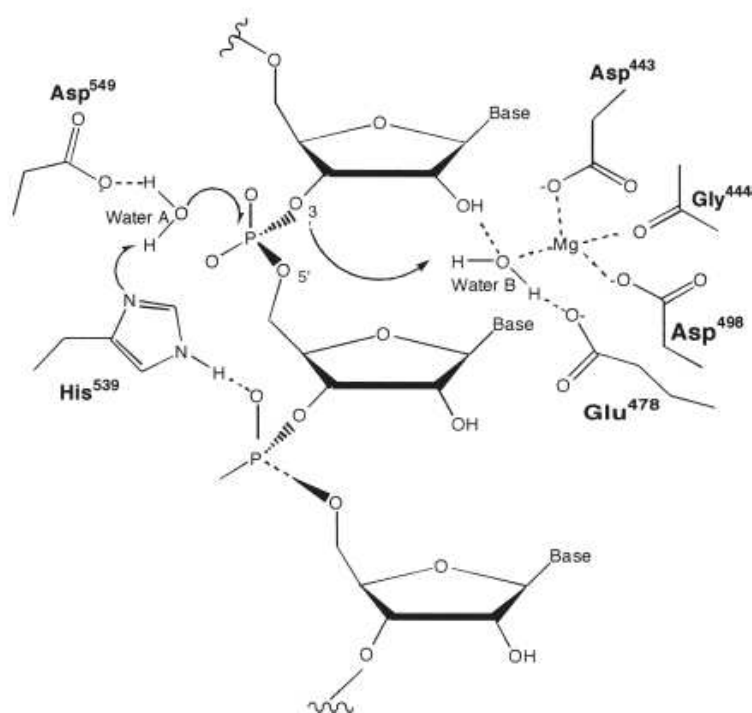
Figure 70. Proposed catalytic mechanism for HIV-1 RNase H based on studies of Haruki with E. Coli [190].

3.2.4. Analysis of RT-NNRTI binding pocket

After GBPM application, the obtained results were defined into a statistical study, as reported in figure 71.
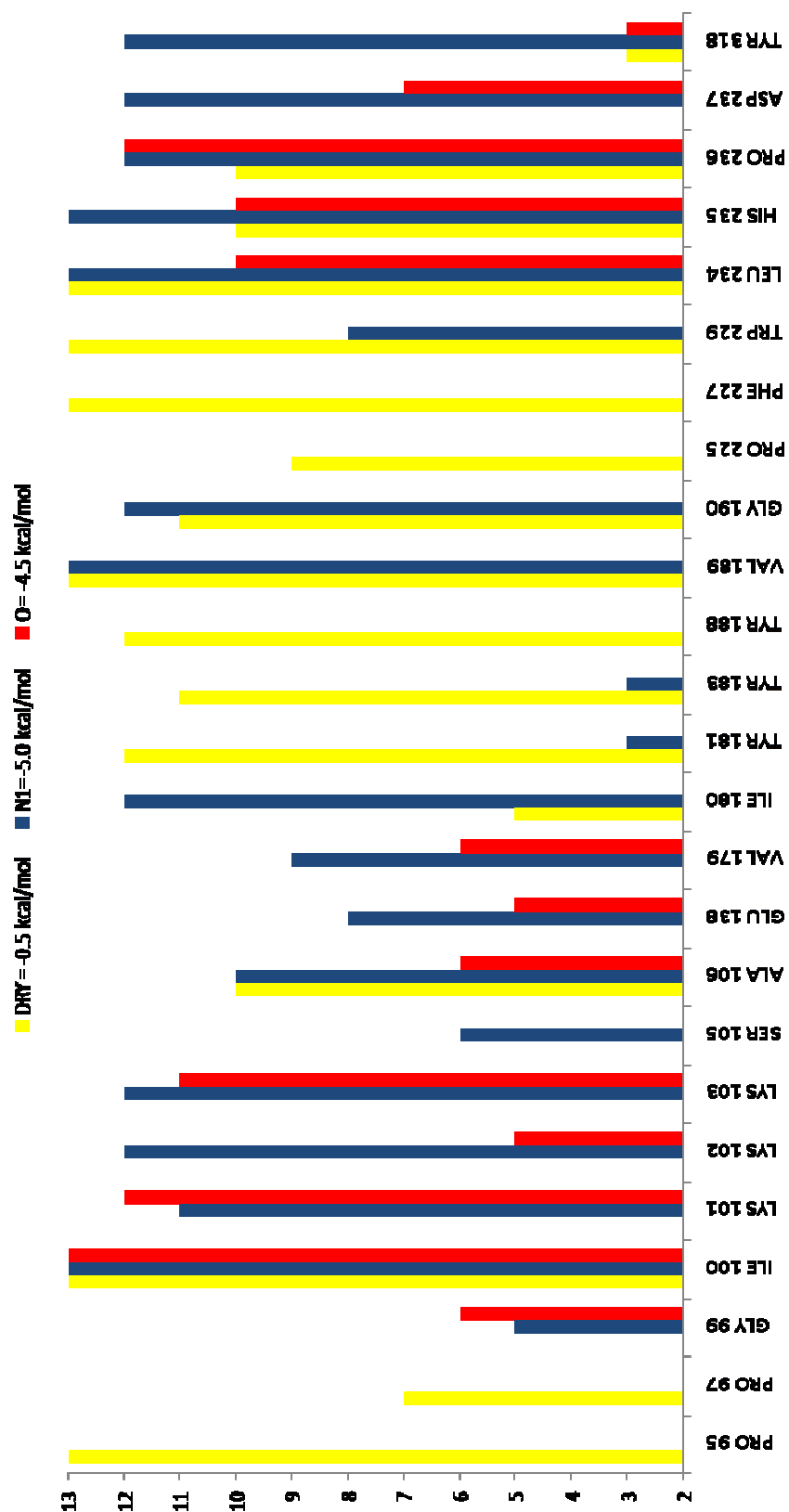
Figure 71. Statistical diagram deriving from analysis of NNRTI-RT complexed PDB models after GBPM application. On the x axis are reported the most frequent and well mapped residues, on the y axis PDB models frequency [Artese at al., unpublished data].

In all analyzed models, the most involved residues in NNRTI-BP recognition resulted Pro95, Pro97, Gly99, Leu100, Lys101, Lys102, Lys103, Ser105, Val106, Val179, Ile180, Tyr181, Tyr183, Tyr188, Val189, Gly190, Ser191, Pro225, Phe227, Trp229, Leu234, His235, Pro236, Asp237, Tyr318 of p66 subunit and Glu138 of p51 subunit. Most of these amino acids are reported in figure 72, that shows the binding pocket for a new NNRTI and the RT residues mainly involved in drug stabilization.
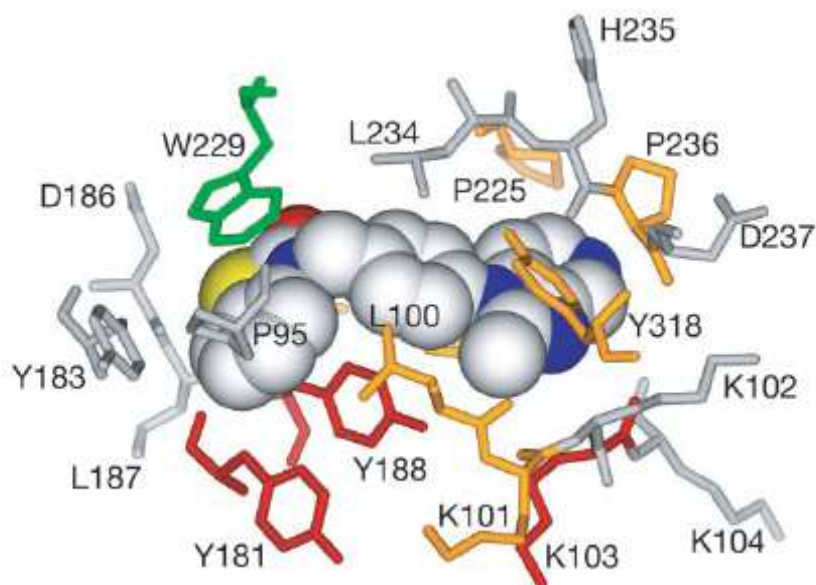
Figure 72. The binding pocket for a non-nucleoside inhibitor of second generation. Residues shown (stick representation) are those with atoms within 4.5 Å of the drug (space-filling representation). In red are residues that when mutated give high-level resistance to nevirapine, efavirenz, and/or delavirdine (Tyr-181, Tyr-188, and/or Lys-103) [200].

Previous conservation studies highlighted the crucial role of residues Pro95, Tyr181, Phe227, Trp229, Leu234 (figure 73) and Tyr318 in catalytic activity after pharmacological treatment with NNRTIs [172].
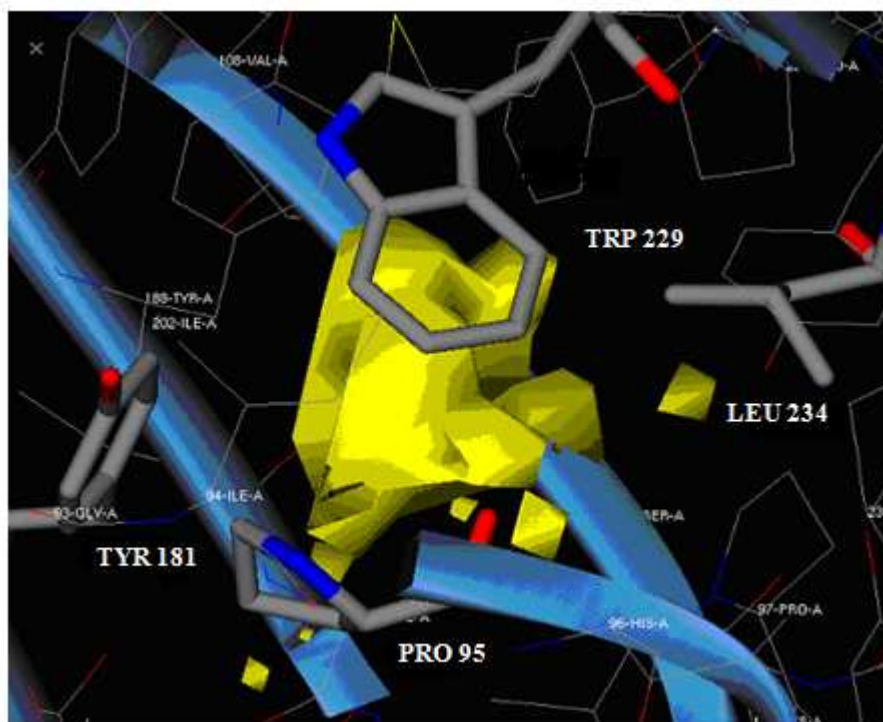
Figure 73. Interaction areas of Pro95, Tyr181, Trp229 and Leu234 of p66 well mapped with DRY probe (yellow areas) using PDB model 1s9e [Artese at al., unpublished data].

Site-directed mutagenesis analysis showed that the substitution Pro95Ala involves a drastic decreasing of RT RNA-dependent DNA polymerase function due to the lack of hydrogen bond between Pro95 of p66 and Asn137 of p51, with a consequent destabilizing dimer effect [84].

The obtained data were validated by GBPM maps using DRY probe, with a good recognition of the heterocyclic ring of Pro95 and of the aromatic moieties of the variable Tyr181 and more conserved Tyr188 [172].

In fact non-nucleoside drugs, in spite of their structural diversities, have in common a remarkable lipophilic nature and interact with RT through hydrophobic contacts with residues Tyr181, Tyr188, Pro225, Phe227, Trp229 and Pro236. NNRTI binding to RT causes the side chains of both Tyr181 and Tyr188 to rotate away from their positions by twisting about 30° in the hydrophobic core, thereby creating a space to accommodate the ligand [86]. The hydrophobic interactions that stabilize the enzyme-inhibitor complex take place among the aromatic side chain residues of Tyr181 and Tyr188 and π-electron donor groups of the NNRTIs.

As reported in literature, the single mutations Tyr181Cys and Tyr188Cys were found to be associated to resistance to NNRTIs of first generation and to nevirapine and efavirenz, respectively, due to the loss of the aromatic ring [201].

Part of the YMDD motif, the high conserved Tyr183 was well mapped using DRY probe, confirming its crucial role in the interactions with the drugs. In fact some mutagenesis studies showed that a Y183S mutation resulted in a 99% reduction in polymerase activity, and a subsequent analysis demonstrated that this same mutation caused a 77-fold decrease in the affinity for dNTP substrates [202].

More recent data have shown that Y183F mutation increases the fidelity of HIV-1 RT [203], and that the Y183A mutant is active on DNA templates but not on RNA templates [204].

Located in the β5b-β6 loop of the *palm* subdomain, the conserved residues Pro97, Gly99 and Ser105 resulted well mapped with DRY, N1 and O, and N1 probes, respectively, confirming their essential role in polymerase activity and in hydrophobic contacts with NNRTIs [205].

As reported in literature [86], mutations Leu100Ile and Val106Ala determine a decreasing of side-chains length and a distortion of binding pocket region with a complete loss of protein-drug contacts. As validation of these data, the high conserved Leu100 and the slightly variable Val106 were well recognized with all used probes, highlighting their crucial role in hydrogen bonding and hydrophobic interactions with NNRTIs.

Also the conserved Lys101 and the high conserved Lys102 are involved in direct and indirect contacts with non nucleoside drugs. In fact these inhibitors, through water molecules, are able to establish hydrogen bonds with carbonylic group of Lys101 [206], well mapped with N1 and O probes (figure 24), and with carbonylic and amide groups of Lys102 [207], well recognized with N1 probe.

In an external region of NNRTI-BP is located the high variable Lys103, that is responsible to stabilize NNRTI into the pocket through a hydrogen bond. N1 and O maps validated this data (figure 74), even in presence of mutation Lys103Asn, that causes the formation of a good hydrogen bond between the phenol oxygen of Tyr188 and the side-chain carboxamide of Asn103, consequently hampering the reorientation of Tyr181 and Tyr188 toward the polymerase active site [86].
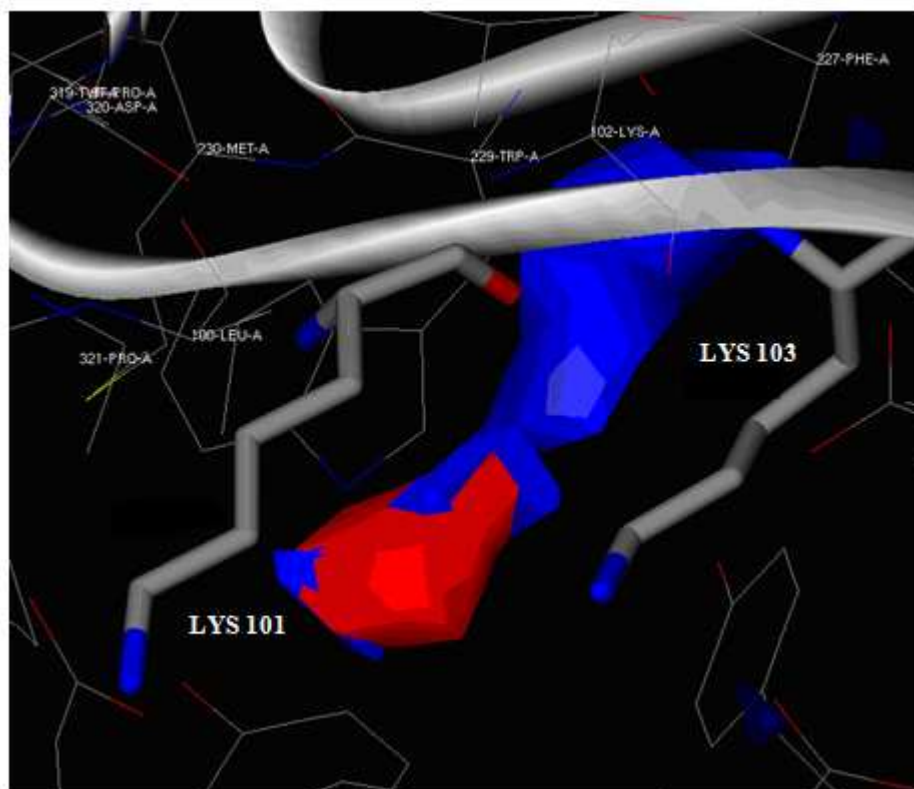
Figure 74. Interaction areas of Lys101 and Lys103 of p66 well mapped with N1 probe (blue area) and O probe (red area) using PDB model 1hnv [Artese at al., unpublished data].

Also Val179 is part of NNRTI-BP and is placed at dimerization interface. This residue is involved in a hydrogen bond with Lys103 and facilitates NNRTI incoming into the hydrophobic pocket [208]. Analyzing maps obtained using N1 and O probes, Val179 and Gly190, that is located in binding pocket β-sheet [209], resulted well recognized, validating the importance of these residues for interactions with the drugs. Amino acid substitutions at codon 190 of HIV-1 RT are frequently selected during HAART regimens and mutagenesis data confirmed that high-level resistance to nevirapine and moderate resistance to both stavudine and zidovudine were associated with G190S/A/E mutations [210].

Since the NNRTI-binding pocket of HIV-1 RT includes amino acids 100-110, 180-190 and 220-240, mutations in these RT regions can substantially decrease susceptibility to all NNRTIs [211]. Analyzing N1 and DRY maps, the high conserved and hydrophobic residues Ile180 and Val189 were identified, highlighting their importance in favourable contacts within the NNRTI binding pocket [205; 172].

Pro225 is one of the amino acids forming the mouth of the hydrophobic NNRTI-specific binding pocket of HIV-1 RT, and is highly conserved in the RTs of HIV-1, HIV-2 and simian immunodeficiency virus (SIV) [212]. This amino acid was well recognized using

DRY probe and such a data validates its essential role in lipophilic interactions with the drugs, as much as the P225H mutation in the recombinant HIV-1 resulted in a 2±6-fold decreases of sensitivity to all NNRTIs tested, except for BHAP U-90152, where there was a 2±5-fold increase in sensitivity [211].

Reverse transcriptase residues from 224 to 235 are known to be involved in positioning the primer terminus ("primer grip") [213]. The aromatic and conserved amino acids Phe227 and Trp229, well mapped with DRY probe, are pivotal in NNRTI-BP for their ability to establish π-stacking interactions with the drugs [200]. Mutagenesis studies confirmed the crucial role of Trp229 aromatic ring for RT activity, correct folding of the protein and NNRTI-BP formation [185].

Located in the primer grip area, Leu234 resulted well mapped using DRY probe. This residue is known to be essential in RT dimerization step as much as the substitution Leu234Ala inhibits p66/p51 association, causing long range conformational changes and consequently altering NNRTI-BP structure [80].

Several studies showed that His235 and Pro236, through its indolic ring, result crucial for establishing, respectively, hydrophobic interactions and hydrogen bonds with NNRTIs and for stabilizing the open conformation of binding pocket, in order to better accommodate the drug [200]. These data are in agreement with GBPM analysis, since His235 and Pro236 were found to be well mapped with all the used probes (figure 75).
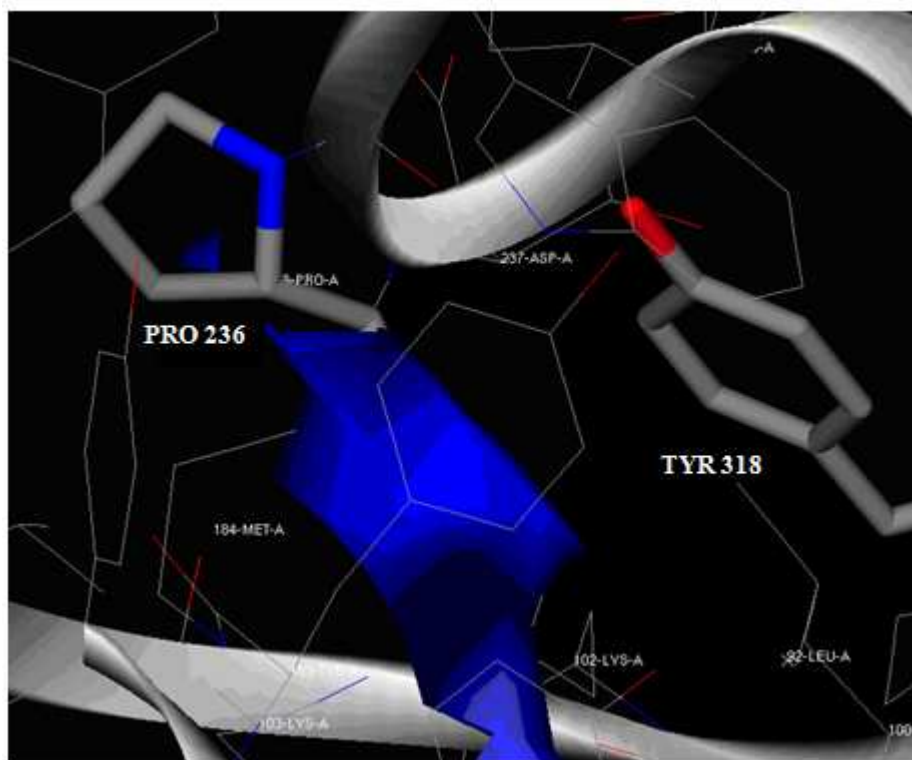
Figure 75. Interaction areas of p66 Pro236 and Tyr318 well mapped with N1 probe (blue area) using PDB model 1hnv [Artese at al., unpublished data].

N1 maps led us to identify another residue of the *palm* subdomain of p66, Asp237, whose role in drug resistance and/or viral fitness is currently unclear [214].

As shown in figure 75, the obtained map using the hydrogen bond donor probe highlighted also the slightly conserved residue Tyr318, known to interact with NNRTI-BP through its phenolic hydroxyl group of the side-chain. In fact recent mutagenesis studies confirmed the pivotal role of this amino acid, that, if mutated with a non-aromatic residue, involves the lack of RT activity and the development of NNRTI associated resistance [215].

The p51 subunit of RT is known to be involved in stabilization of the heterodimer and in the formation of the hydrophobic pocket at which the NNRTIs are targeted. The glutamic acid residue at position 138 of the p51 determines the sensitivity and resistance of the enzyme to the inhibitory effect of the TSAO derivatives due to its capability to establish hydrogen bonds with the drug [216]. These data were validated analyzing GBPM maps with O probe, that confirm the interactions of the poorly variable Glu138 and its functional and structural roles in NNRTI-BP.

3.2.5. Pharmacophore model and drug design

With the aim to obtain a pharmacophore model by the GBPM approach on the three crucial RT drugs sites, the Minim utility [169] was used and the points with minimum energetic values for all the adopted probes were identified. Subsequently, the obtained maps points were imported into Catalyst program, obtaining the starting features for the generation of pharmacophoric models. Each feature has a relative weight in function of single point energy. Thus, for all the analyzed PDB complexes, the hypothesis features number calculation was carried out and the model with the lowest value was selected. As a result, the RT crystallographic model complexed with tenofovir (TDF) (PDB code 1T05) was chosen and a first pharmacophore hypothesis was generated.

RT interface GBPM application led to the identification of a large mapping area of the β7-β8 loop of p51 subunit, known to interact with heterodimer stability interfering TSAO compounds (figure 76) [217].
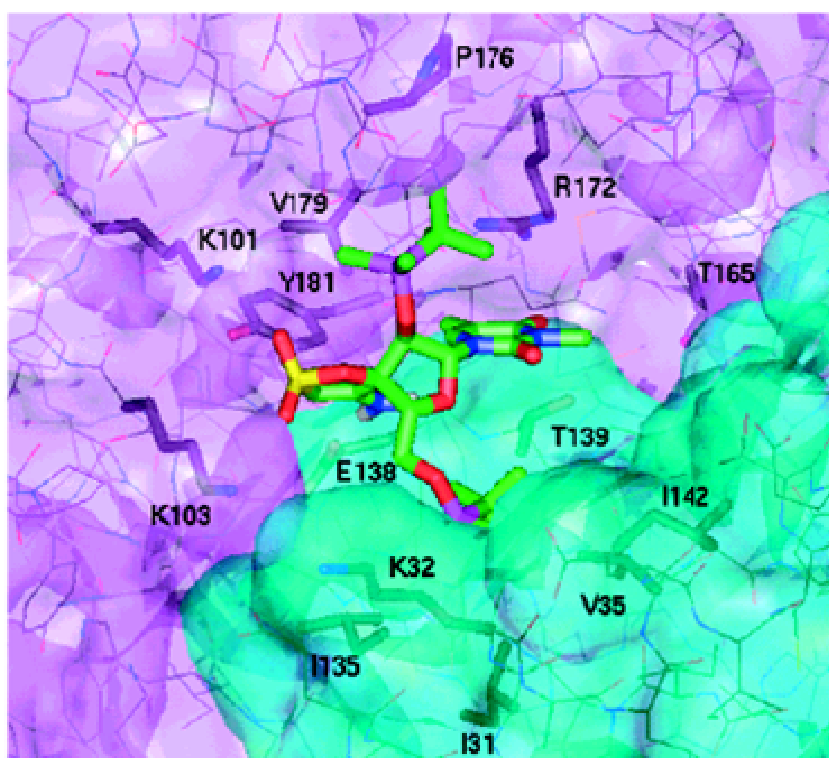
Figure 76. Interaction area of TSAO at RT dimer interface [217].

Due to the emergence of HIV resistant strains in presence of the single mutation E138K, an alternative approach can be represented by molecular hybridization therapy, based on the combination of a nucleoside analogue with a non-nucleoside inhibitor by a chemical spacer.

Some examples of molecular hybridization, shown in figure 77, are: TSAO-T-$(CH_2)_3$-AZT, TSAO-T-$(CH_2)_3$-d4T [218] and TSAO-T-$(CH_2)_3$-PFA [219].
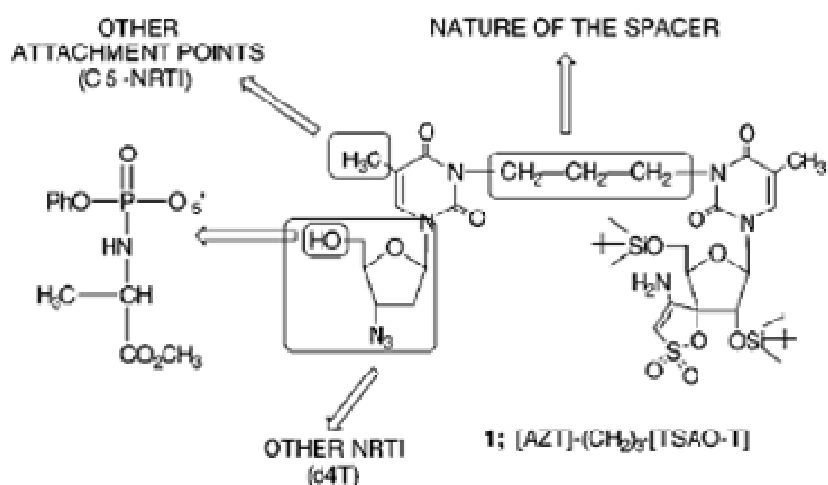
Figure 77. Examples of molecular hybridization approach therapy [218].

Since the crucial residues in TSAO-RT interactions, such as Ile31, Lys32, Val35, Lys101, Lys103, Ser134-Pro140, Ile142, Thr165, Arg172, Pro176, Val179 and Tyr181 [217] resulted well recognized with GBPM analysis, the pharmacophore model, screening a large interface region and the TDF binding site (figure 78), could lead to design a molecular hybrid linked with a spacer.
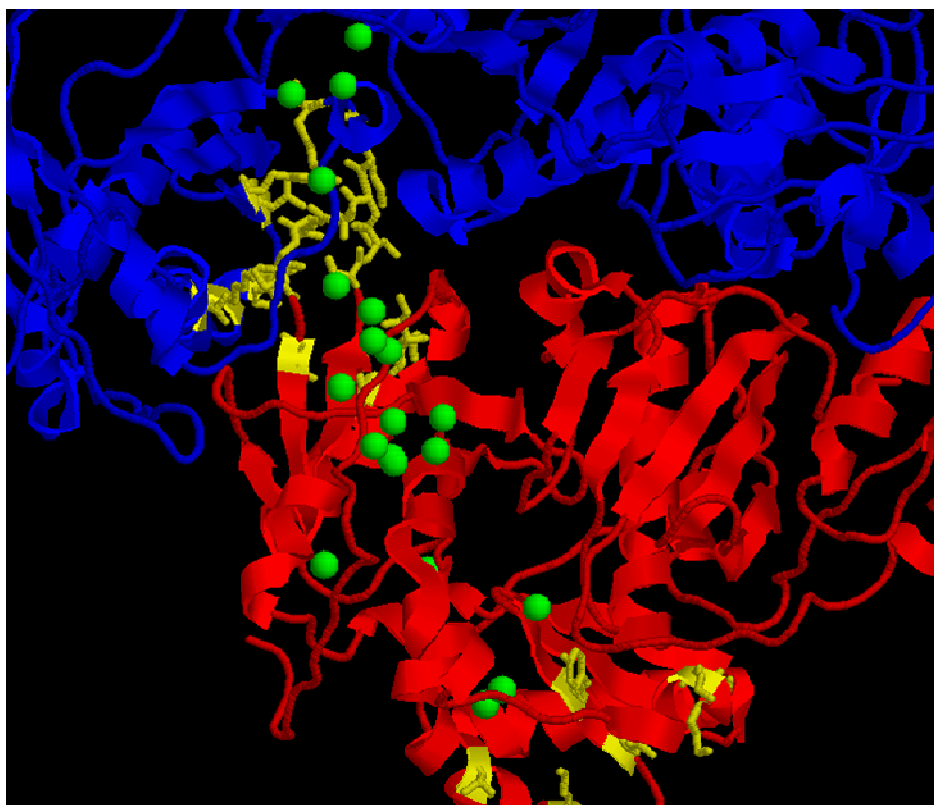
Figure 78. The RT GBPM pharmacophore model obtained using the PDB model 1T05. The p66 and p51 subunits are shown in cartoon and coloured in blue and red, respectively. The crucial residues in TSAO-RT interactions are represented in yellow sticks; the features of the pharmacophore model are displayed as green spheres [Artese at al., unpublished data].

Thus tenofovir association with dimerization synthetic inhibitors (TSAO, BBNH), or with peptidic inhibitors, could be considered an interesting starting point in order to obtain new antiretroviral drugs. The identification of new hypothesis in agreement to structural and clinical data, mediated by the reduction of features number in all RT models according to energetic criteria, is in progress. Moreover the virtual screening of compounds able to fit with the pharmacophoric hypothesis generated from PDB model 1t05 is still running.

4. Molecular dynamics simulations

4.1. RT simulations

The development of NNRTIs resistance is related to several mutations and is associated to NNRTI therapy failure. These mutations affect NNRTI binding directly by altering the size, shape, and polarity of different parts of the NNRTI-BP or indirectly by affecting the access of NNRTIs to the pocket itself [220].

In order to analyze the effects of Lys103Asn, a well known mutation occurring onto the p66 subunit of HIV-1 RT, and Ile135Thr, a novel substitution present on the p51 subunit, molecular dynamics simulations (MDS) were performed.

4.1.1. Materials and methods

The computational work started with the analysis of the available crystallographic unbound models of RT. The 1DLO PDB model was judged as the most complete (971 residues, resolution = 2.70 Å) and selected for creating the starting geometries of the MDS, adding on each subunit missing amino acids. MDSs were carried out by using the GROMACS software [221]. The wild-type RT model and the two RT mutants with mutations K103N and K103N_I135T were analyzed, submitting to MDS the starting geometries obtained in the previous step by single-residue substitution. For each structure, the GROMOS96 53a6 force field with the explicit water implementation and an octahedral box were used, setting the box edge at a distance of 3.0 nm from the target. All studied mutants were energy minimized for 50.000 steps. The Linear Constraints Solver algorithm was adopted for all atoms to prevent bond distortions as suggested for time steps larger than 1 fs. After energy minimization, we performed three MDSs by using Berendsen's temperature and pressure coupling methods and particle mesh Ewald electrostatic treatment under coulomb-type [221]. In the first simulation, for 100 ps, RT was kept frozen and only water molecules were free to move; in the second one, water molecules and added residues were allowed to move for 500 ps, while the entire enzyme was still fixed; and in the last one, the whole enzyme was free to move for 1 ns. Several geometrical parameters were analyzed by adopting different GROMACS utilities, like g_hbonds, which computes and analyzes hydrogen bonds during the whole simulation. The default distance cutoff of 2.5 Å and the default angle cutoff of 60 degrees were used.

4.1.2. Results and discussion

Since RT NNRTI-BP is not a cavity, but a hydrophobic region situated in RT *thumb* subdomain, NNRTI-BP residues have to arrange in order to accommodate the drug. The key event in this process is represented by primer grip ß12-14 sheets rotation, that causes Trp229 shift from binding pocket and Tyr181-Tyr188 "flipping" toward polymerase active site. These structural changes make consequently catalytic aspartates to assume an improper geometry with the final inhibition of enzyme polymerase action [78]. The tridimensional localization of such groups of residues is shown in figure 79.

Figure 79. Tridimensional localization of primer grip residues (residues 226-235), reported as magenta polytube, and of p66 NNRTI-BP interface residues, represented with yellow polytube rendering [Artese at al., unpublished data].

With the aim to evaluate ß12-14 sheets rotation for all mutants, the distance between primer grip residues and p66 NNRTI-BP interface amino acids (Pro95, Leu100, Lys101, Lys103, Val179, Tyr181, Tyr188), was monitored, as reported in figure 79. A decreasing of the average distance for the RT double mutant K103N_I135T with respect to wild-type model (table 7, figure 79) was noticed.

Model	Average distance
WT	1.19
K103N	1.16
I135T	1.17
K103N_I135T	1.10

Table 7. Average distance analysis between primer grip residues and p66 NNRTI-BP interface amino acids calculated for WT and the studied RT mutants [Artese at al., unpublished data].
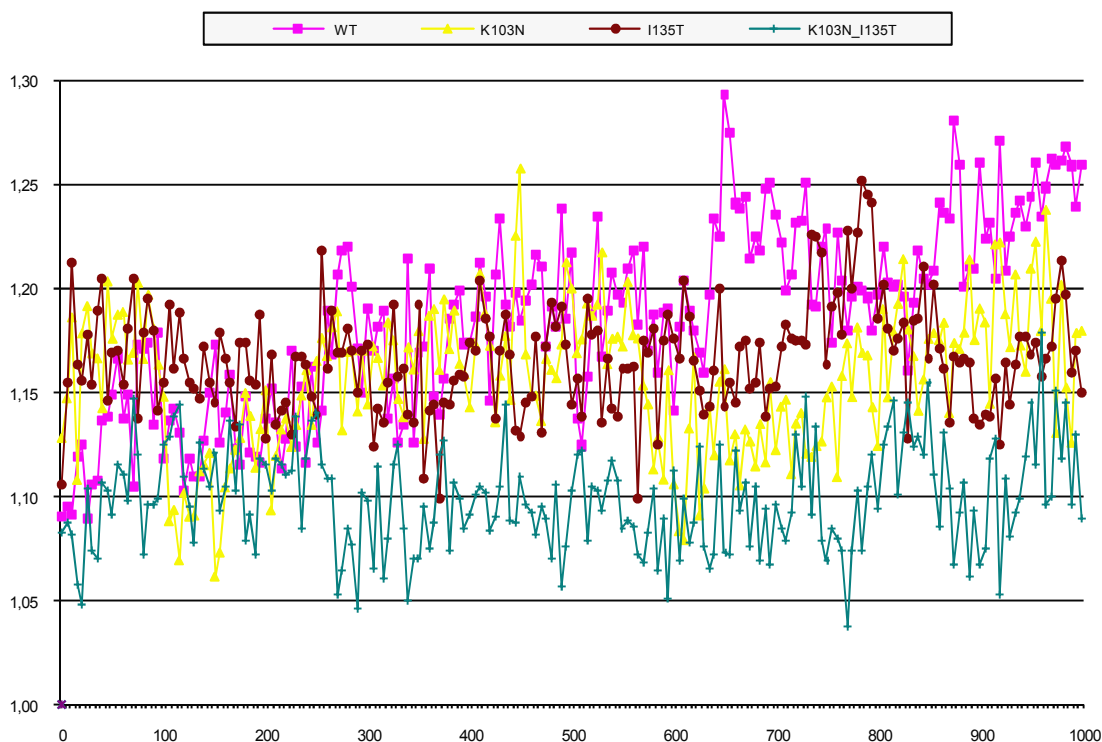
Figure 80. Distance analysis between primer grip residues and p66 NNRTI-BP interface amino acids for WT and RT mutants. On the x axis is reported simulation time measured in ps, on the y axis the distance expressed in nm [Artese at al., unpublished data].

As reported in figure 80, in initial MD steps such a distance was found very similar for all analyzed mutants, while in final steps a reduction for the double K103N_I135T with respect to WT was observed.

Subsequently a distance calculation analysis between Tyr188 and the catalytic aspartates was carried out in order to evaluate the accessibility of a generic NNRTI inhibitor to the binding pocket (table 8).

In NNRTI-BP formation, Tyr181 and Tyr188 have to flip toward catalytic aspartates, but in presence of K103N mutation is observed a high energetic barrier for creation of this pocket, with a RT closed conformation stabilization [88].

Model	*Average distance*
WT	1.03
K103N	1.06
I135T	1.06
K103N_I135T	1.09

Table 8. Average distance analysis between Tyr188 and catalytic aspartates calculated for WT and all RT mutants [Artese at al., unpublished data].

As shown in figure 81, in final steps of MD simulation the distance between Tyr188 and active polymerase site resulted greater for the double K103N_I135T with respect to WT, thus validating the higher drug accessibility to accommodate into binding pocket without mutation on residue 103.
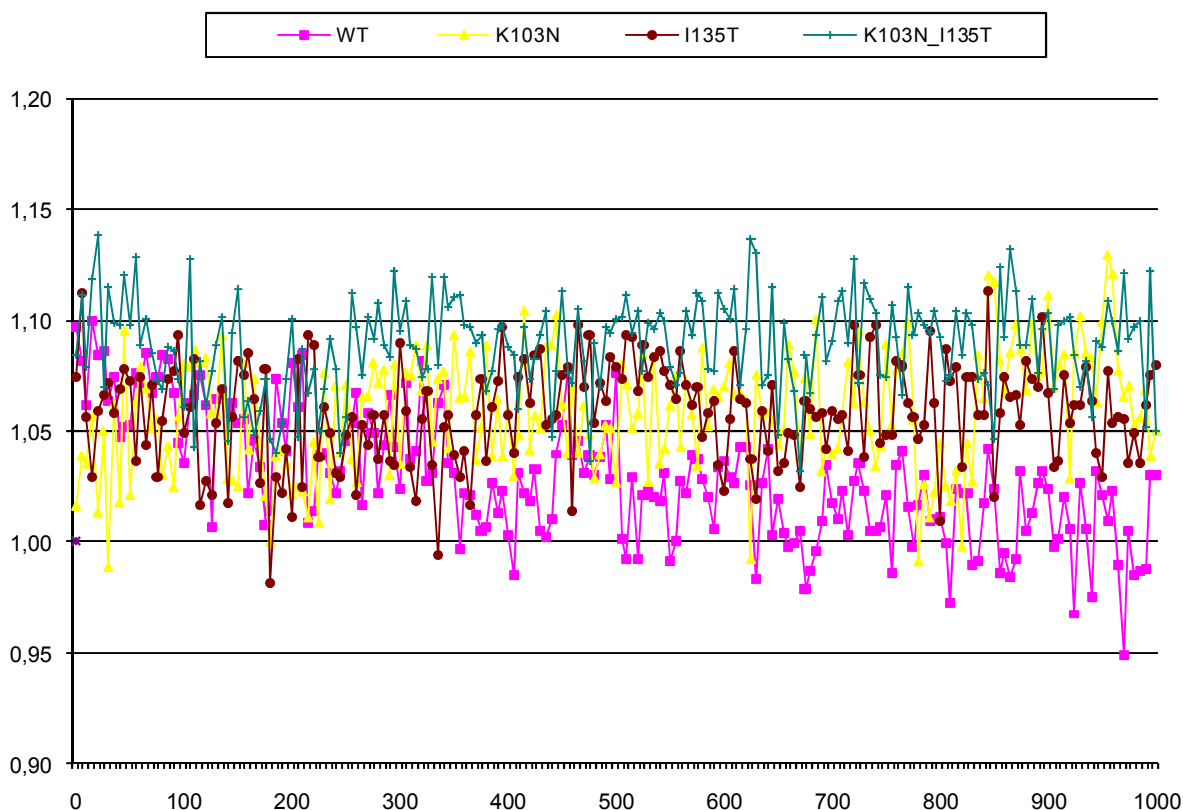
Figure 81. Distance analysis between Tyr188 and the catalytic aspartates for WT and RT mutants. On the x axis is reported simulation time measured in ps, on the y axis the distance expressed in nm [Artese at al., unpublished data].

The substitution of Lys to Asn at position 103 seemed to induce significant differences in NNRTI binding pocket structure with respect to WT. In K103N mutant RT long positively charged side chain of Lys103 is replaced by the smaller and uncharged side chain of Asn, with perturbation effects in total enzyme electrostatic charge. In K103N mutant the empty space derived from the Lys to Asn substitution was occupied by one water molecule, as reported in figure 82.
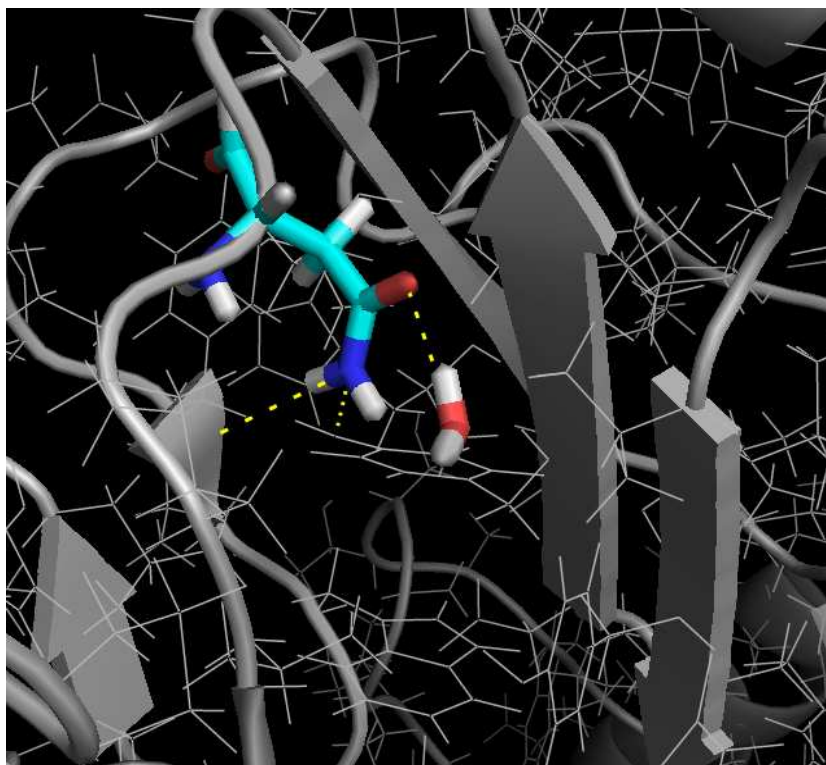
Figure 82. Hydrogen bond between Asn103 carbonylic group and a water molecule placed into NNRTI binding pocket. RT backbone is represented with cartoon rendering, while the side chains are displayed in wireframe [Artese at al., unpublished data].

A key hydrogen bond is established between the nitrogen atom of Asn103 side chain and the hydroxylic group of Tyr188 in the unbounded mutant K103N, as reported in figure 83.
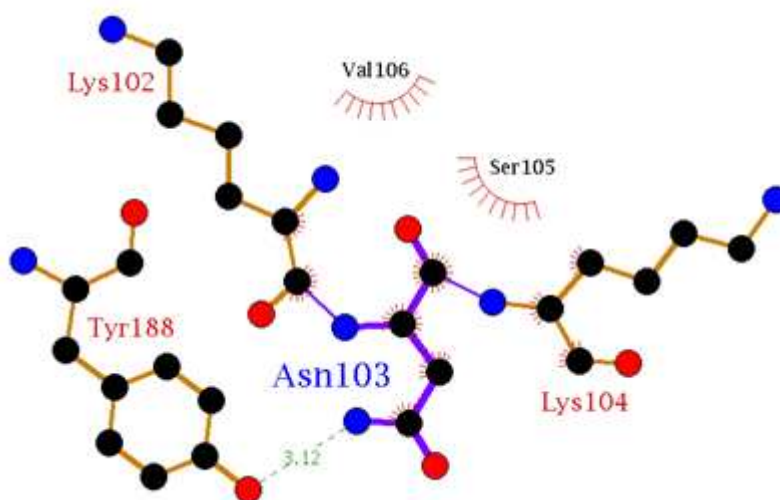
Figure 83. LigPlot schematic diagram relative to Asn103. The hydrogen bond between the nitrogen atom of Asn103 side chain and the hydroxyl group of Tyr188 in the unbounded mutant Lys103Asn is indicated with a green dashed line, while hydrophobic contacts are showed with a red labelled arc [Artese at al., unpublished data].

Another pivotal hydrogen bond is observed between the nitrogen atom of Asn 103 side chain and the carbonylic oxygen of Lys101. Such a hydrogen bonding network probably causes the stabilization of NNRTI-BP closed conformation in the mutant K103N [78].

As reported in literature [90], NNRTI-BP formation is hampered by Asn103-Tyr188 hydrogen bond and this interaction is absent in WT. With the aim to verify its presence, this hydrogen-bond for the trajectories of the studied mutants was monitored and it resulted formed only in presence of the K103N mutation, single or associated, while for WT no hydrogen-bonds were observed during the entire MD simulations (figure 84).
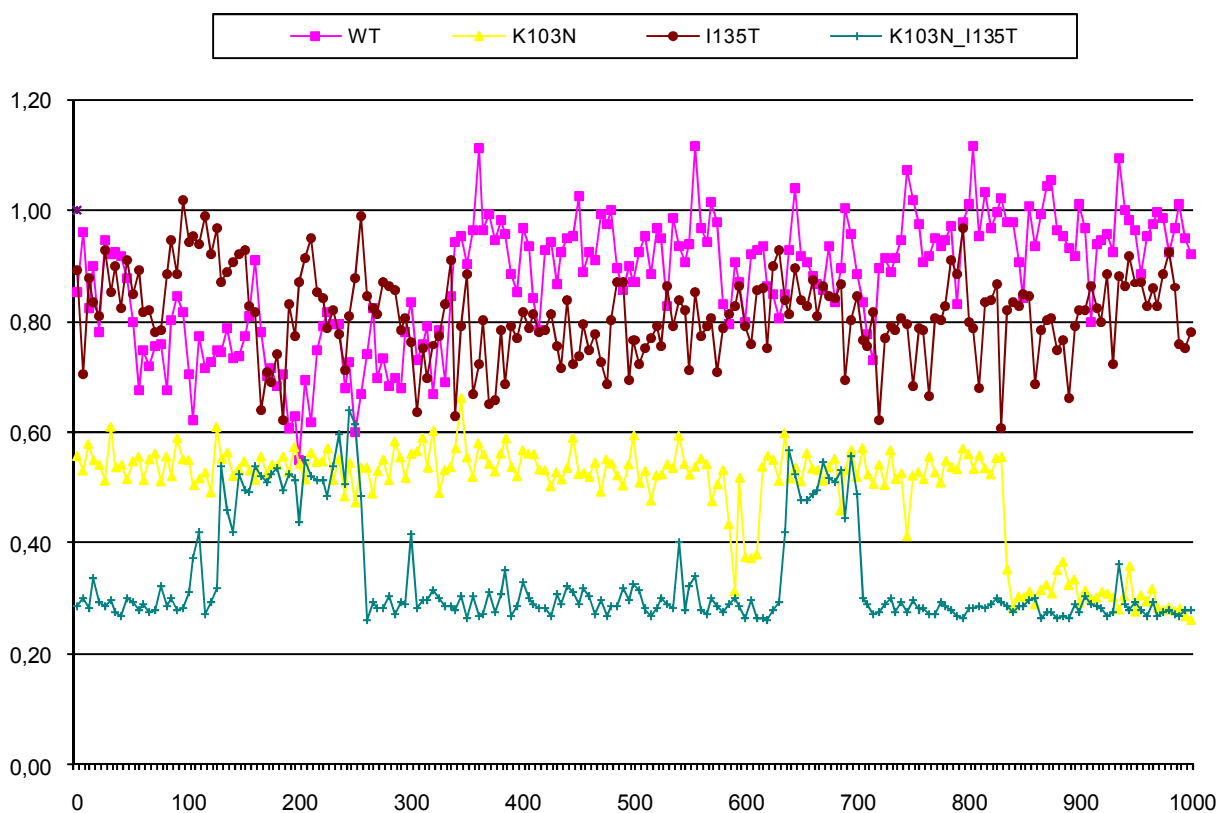
Figure 84. Distance analysis between the nitrogen atom of Asn103 side chain and the hydroxyl group of Tyr188 calculated for WT and RT mutants. On the x axis is reported simulation time measured in ps, on the y axis the distance expressed in nm [Artese at al., unpublished data].

Therefore the performed MDS showed that the phenol oxygen of Tyr188 and the side chain carboxamide of Asn103 are closer in the presence of Thr135 than they are in the presence of the WT Ile residue at position 135 (distance of 1.94 Å versus 2.90 Å) (figure 85) [222].
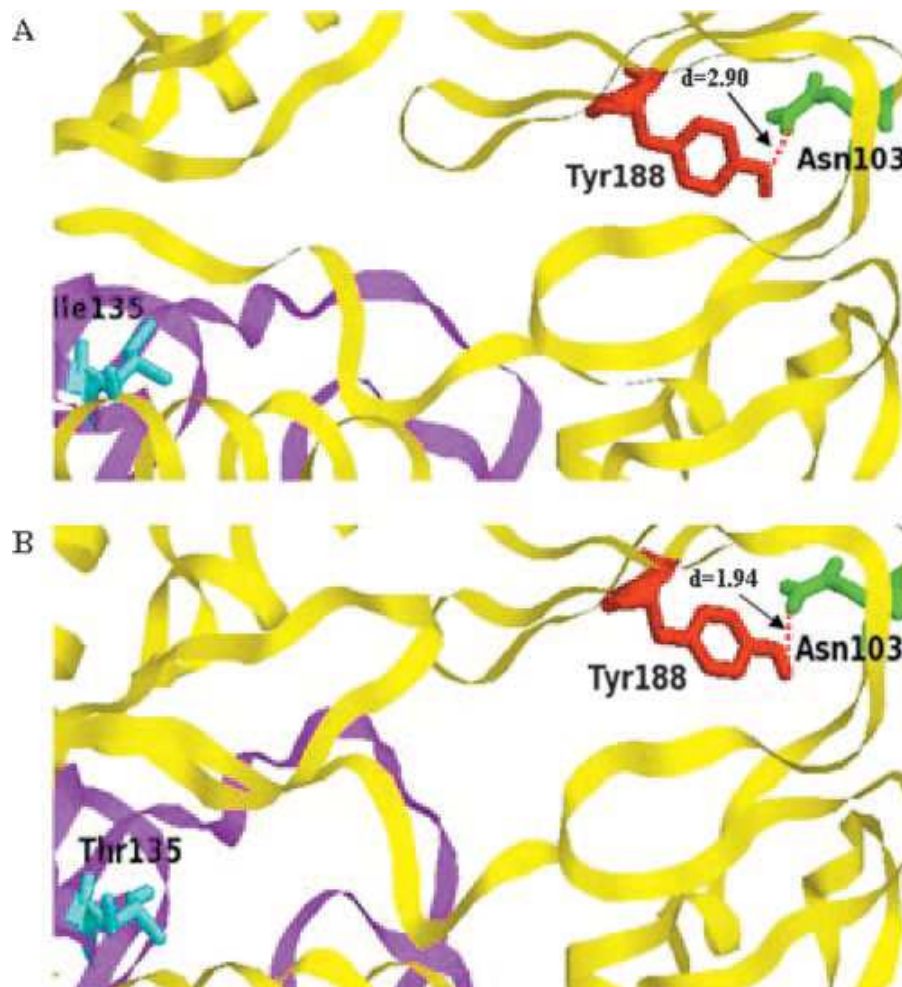
Figure 85. Insight into the NNRTI-BP showing the distance in angstroms (d) between NH_2 of Asn103 and the OH of Tyr188 in the presence of Ile135 (A) or Thr135 (B). Side chains of residues Asn103, Tyr188, and Ile/Thr135 are highlighted in wireframe. The p66 and the p51 subunits are represented as cartoons and coloured in yellow and magenta, respectively [222].

In addition, the number of occurrences of a hydrogen bond over 1 ns between Tyr188 and Asn103 in the presence of Thr135 was significantly higher than that in the presence of the Ile135 wild type (152 occurrences versus 18 occurrences; P<0.001). This suggests that I135T may favor the stabilization of the closed form of the NNRTI binding pocket induced by K103N.

Moreover, analyzing within the NNRTI binding pocket the energy related to the protein-protein electrostatic interaction at the heterodimer p66-p51 interface, it was observed that the mutation K103N itself determined a reduction of this energetic term with respect to the WT sequence (average Coul-SR energy, -38.004,6 versus -37.457,6 kJ/mol) and this average energy decreased much more with the presence of the double mutation I135T_K103N (average Coul-SR energy, -38.679,1 kJ/mol) [222].

4.2. PR simulations

Drug resistance is a major problem affecting the clinical efficacy of antiretroviral agents, including protease inhibitors (PIs), in the treatment of infection with HIV-1/AIDS.

For its crucial role in the HIV-1 life cycle, PR represents an important target for the antiretroviral therapy. To date seven PIs (indinavir, ritonavir, saquinavir, nelfinavir, amprenavir, lopinavir, and atazanavir) have been approved by the Food and Drug Administration (FDA) and are clinically available. Unfortunately, when antiretroviral therapy fails to be fully suppressive, viral variants with reduced susceptibilities to PIs can emerge [223]. Resistance to PIs is mediated by the appearance of PR amino acid substitutions (at positions either in direct contact with the inhibitor or at distant sites) that reduce the binding affinity between the inhibitor and the mutant PR enzyme. These amino acid substitutions, defined in the literature as major mutations, may deeply impair the PR catalytic activity and, consequently, the replication capacity of the virus. Restoration of the replication capacity is due to the presence of mutations defined in the literature as compensatory mutations [223]. Several studies have contributed to our current knowledge of the drug-related variants of HIV-1 PR. To date, mutations at 50 of 99 residues of PR have been related to one or more experimentally tested PIs, thus attesting to the high degree of flexibility of the PR enzyme; 22 of these residues are involved in resistance to the PIs used in clinical practice [223].

Consequently, the elucidation of the molecular recognition of the currently approved PIs in presence of known mutations responsible to confer resistance is critical to the development of superior inhibitors.

4.2.1. Materials and methods

For both PR inhibitors the two analyzed mutants were built, starting from crystallographic models, by residues replacing. Theoretical structures were energy minimized using the AMBER* force field [224] and the GB/SA water implicit solvation model [225]. Using the same force fields and environment, the optimized structures were submitted to 1 ns of Molecular Dynamics Simulation (MDS) at 300K, sampling 200 structures. These calculations were performed with the molecular modeling software MacroModel ver.7.2 [226]. According to Moline method [227], the interaction energies of all PR complexes were evaluated, identifying the most populated configuration. The contacts of the most probable configurations were analyzed using LigPlot program [171].

4.2.2. Results and discussion

In this analysis the influence of two drug resistance-associated mutations, L33F and L76V, of HIV-1 PR has been evaluated with respect to lopinavir and atazanavir molecular recognition. Lopinavir (LPV) is a PI developed from ritonavir (figure 86). Co administration with low-dose of ritonavir significantly improves the pharmacokinetic properties and hence the activity of LPV against HIV-1 PR. Some mutations have been shown to be associated with LPV treatment (E34Q, K43T, and K55R) and correlated with mutations associated with LPV resistance (E34Q with either L33F or F53L, or K43T with I54A) or clustered with multi-PI resistance mutations (K43T with V82A and I54V or V82A, V32I, and I47V, or K55R with V82A, I54V, and M46I) [223].
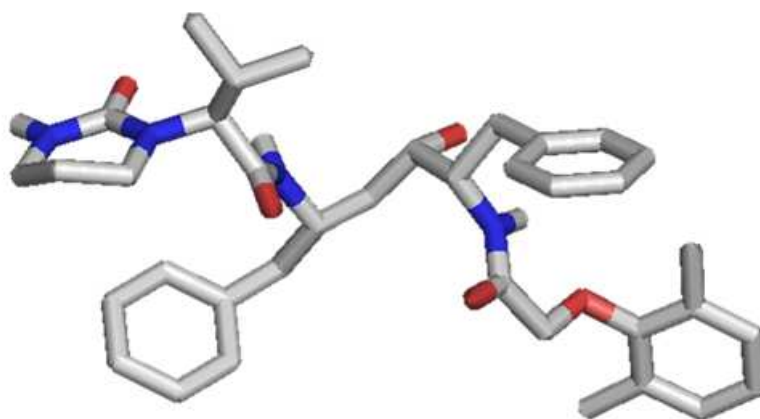
Figure 86. Tridimensional structure of lopinavir [Artese at al., unpublished data].

Atazanavir (ATV) is a second generation PI reporting for the first time an azapeptide structure (figure 87). The *in vitro* drug susceptibilities evaluation, carried out onto a large panel of ATV-naïve HIV clinical virus isolates, revealed a distinct resistance profile with respect to other PIs. Viruses isolates characterization from PI-naïve patients failing ATV therapy, highlighted a single I50L mutation of the protease in 100% of ATV resistant isolates [228].
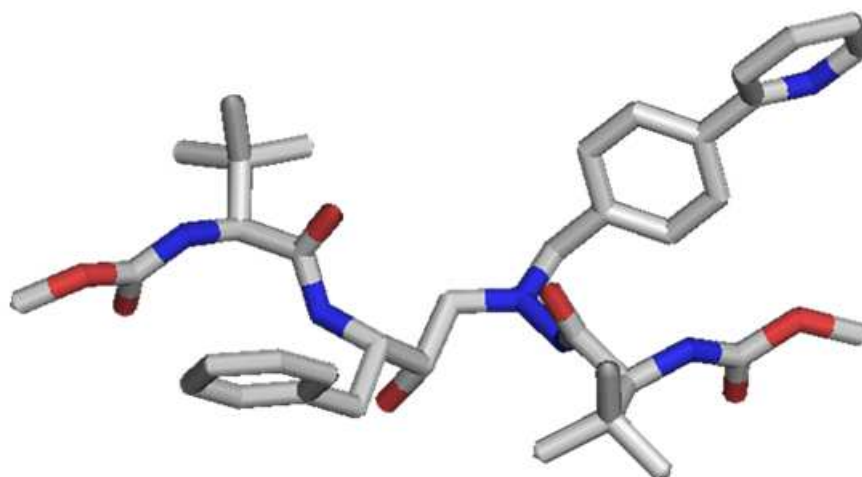
Figure 87. Tridimensional structure of atazanavir [Artese at al., unpublished data].

In order to analyze LPV-PR complex, due to the absence of mutations and to its low resolution (2.80 Å), the crystallographic PDB structure with the code 1MUI [229] was selected. Residues Leu33 and Leu76 are close with each other in the HIV PR, as shown in figure 88.
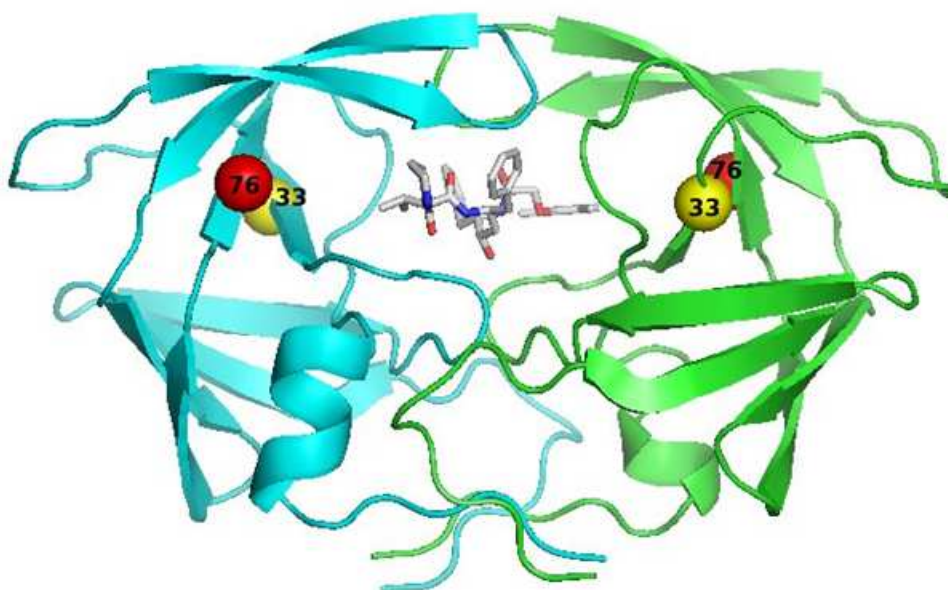
Figure 88. A ribbon model of HIV-1 PR with bound LPV showing the localization of residues 33 and 76 [Artese at al., unpublished data].

These amino acids directly interact by an hydrogen bond in both PR monomers, as reported in figure 89.
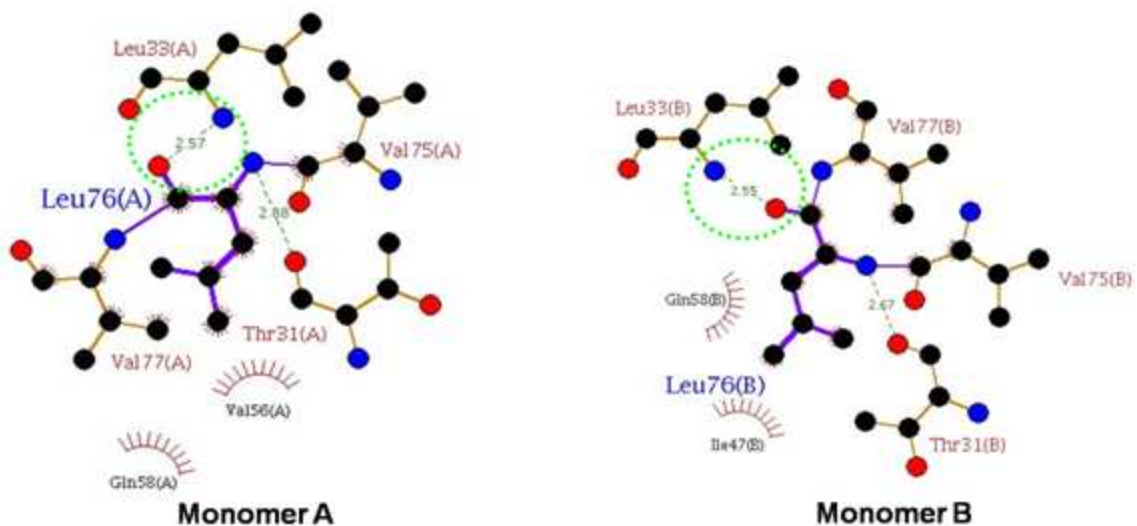
Figure 89. LigPlot schematic diagram showing residues Thr31, Leu33, Val75, Leu76 and Val77 in HIV-1 PR. Leu33 directly interacts with Leu76 through an hydrogen bond (dotted green line). The distance (Å) between atoms involved in hydrogen bonds is reported [Artese at al., unpublished data].

Using the program LigPlot [171] onto the crystallographic structure, LPV was found to establish 3 hydrogen bonds to the catalytic Gly27 and Asp29. The peptidic inhibitor revealed different non-bonded contacts with other PR residues, because of its hydrophobic nature (figure 90).
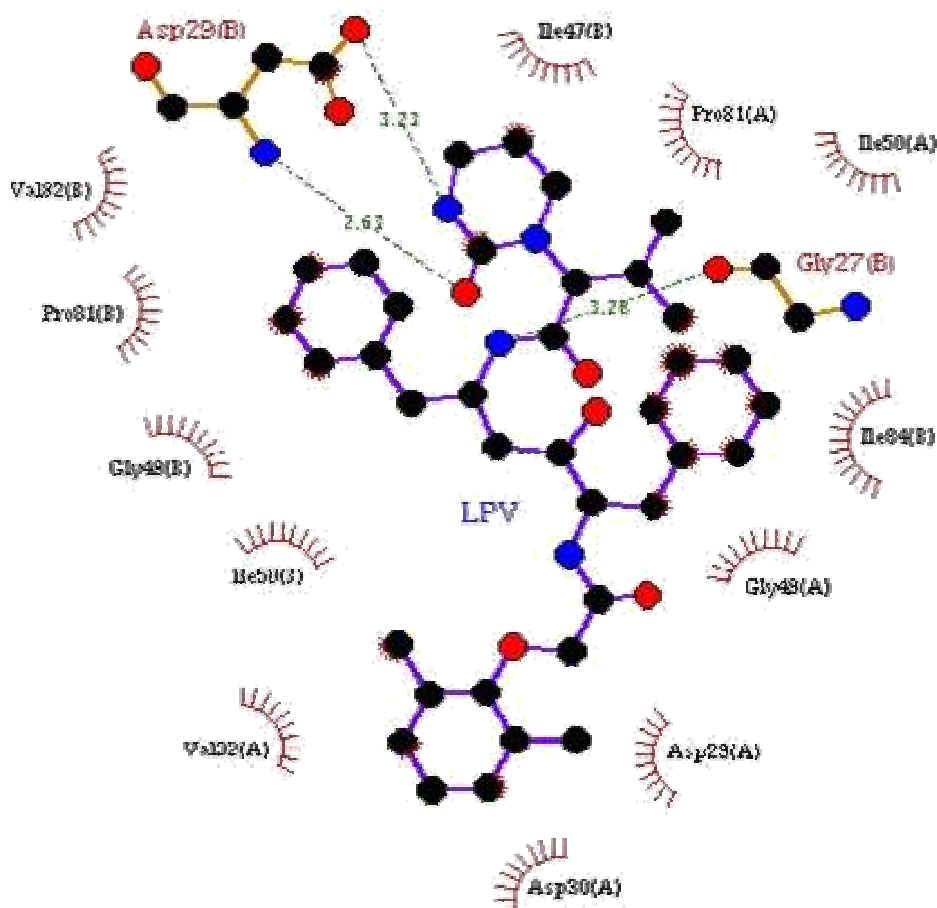
Figure 90. LigPlot schematic diagram showing LPV interactions in the crystallographic PR complex. The hydrogen bonds established by the PI are represented with green dashed lines, while hydrophobic contacts are showed with a red labelled arc [Artese at al., unpublished data].

In order to analyze ATV-PR complex, the crystallographic PDB model with the code 2O4K (1.60 Å) [230] was used, localizing the studied substitutions with respect to the azapeptide PI, as shown in figure 91.
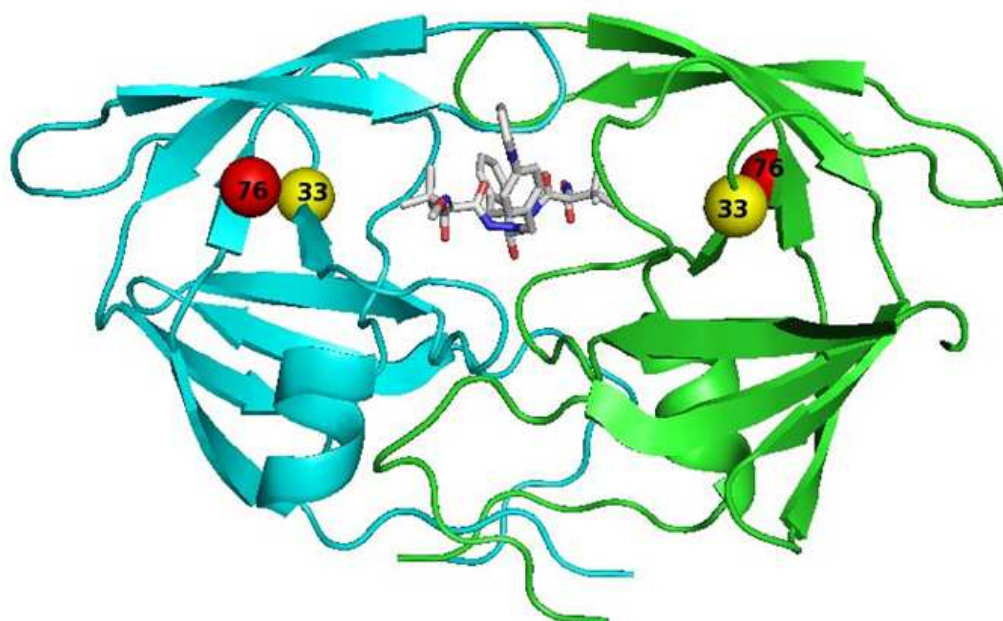
Figure 91. A ribbon model of HIV-1 PR with bound ATV showing the localization of residues 33 and 76 [Artese at al., unpublished data].

Also in ATV-PR complex, these amino acids directly interact by an hydrogen bond in both PR monomers, as reported in figure 92.

Monomer A   **Monomer B**
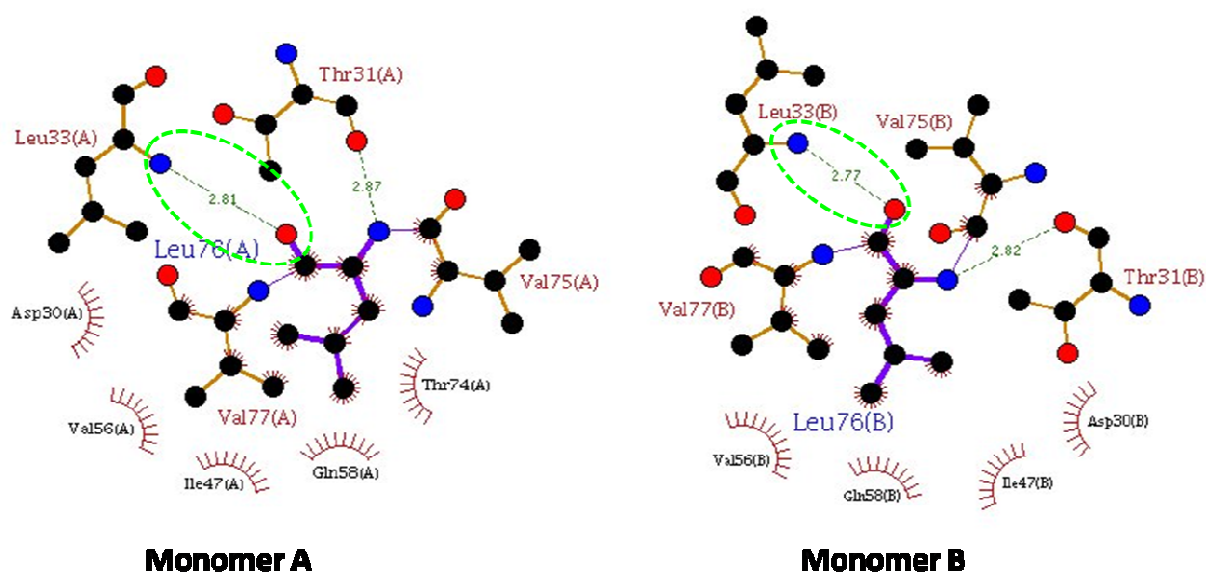
Figure 92. LigPlot schematic diagram showing residues Asp30, Thr31, Leu33, Ile47, Val56, Gln58, Val75, Leu76 and Val77 in HIV-1 PR. Leu33 directly interacts with Leu76 through an hydrogen bond (dotted green line). The distance (Å) between atoms involved in hydrogen bonds is reported [Artese at al., unpublished data].

Using LigPlot [171] onto the selected crystallographic structure, ATV revealed 6 hydrogen bonds to the catalytic Gly27 and Asp29 and to Gly48 in both chains. The azapeptide inhibitor

showed several productive non-bonded interactions due to the presence of its three aromatic rings, as reported in figure 93.
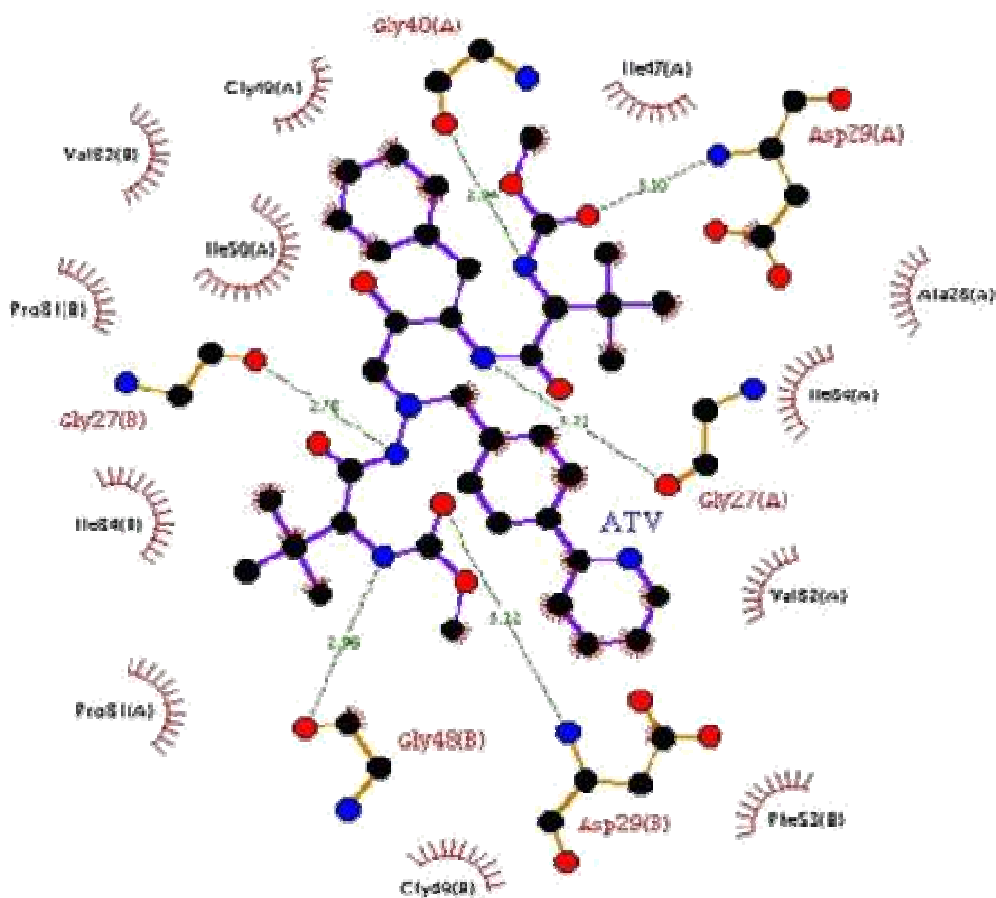
Figure 93. LigPlot schematic diagram showing ATV interactions in the crystallographic PR complex. The hydrogen bonds established by the azapeptide PI are represented with green dashed lines, while hydrophobic contacts are showed with a red labelled arc [Artese at al., unpublished data].

With the purpose to analyze the obtained theoretical complexes for both PR inhibitors, the interactions established between the studied PI within the enzyme active site were evaluated for WT and in presence of the two mutations.

Analyzing the most populated WT configuration of LPV-PR complex, 32 non-bonded contacts and 6 hydrogen bonds were observed, as shown in figure 94. The accommodation of LPV was well stabilized through a crucial hydrogen-bond with Arg8, found to be essential in PR catalytic site [231].

Figure 94. LigPlot schematic diagram showing LPV interactions in the obtained most populated WT-PR configuration. The hydrogen bonds established by LPV are represented with green dashed lines, while hydrophobic contacts are showed with a red labelled arc [Artese at al., unpublished data].

In LPV most populated configuration with the L33F mutation, 40 non-bonded contacts and 5 hydrogen bonds were identified. In this case LPV, with respect to WT, lost two hydrogen-bond with Gly27 present in the active site, and Arg8, with a resultant decreasing of its binding affinity (figure 95).
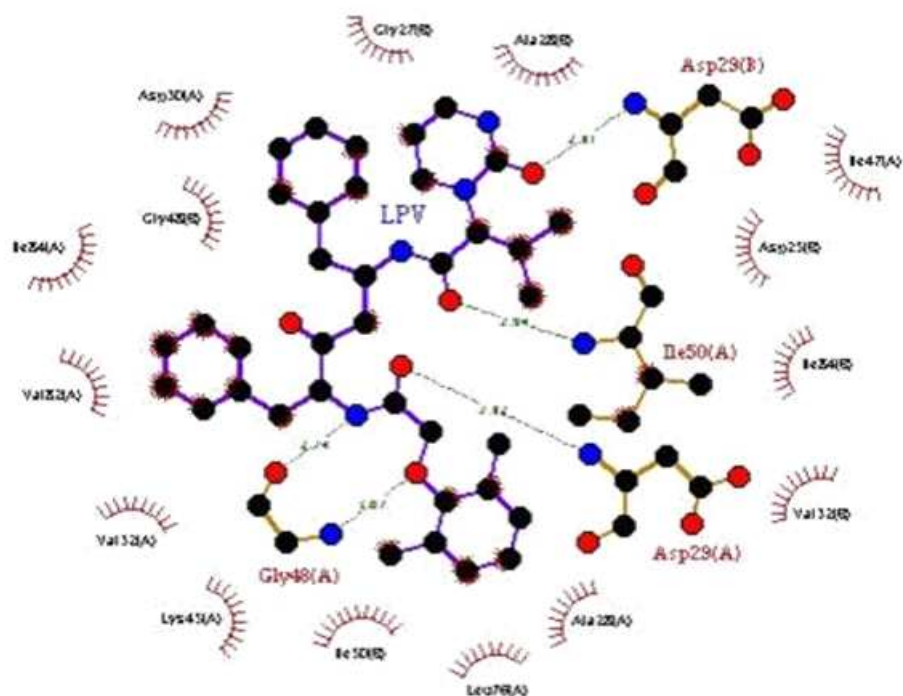
Figure 95. LigPlot schematic diagram showing LPV interactions in the obtained most populated L33F-PR configuration. The hydrogen bonds established by LPV are represented with green dashed lines, while hydrophobic contacts are showed with a red labelled arc [Artese at al., unpublished data].

Analyzing LPV most probable configuration in presence of the L76V mutation, although, with respect to WT, a higher number of non-bonded contacts was observed (55), such a mutation caused a reduction of hydrogen bond network (3) and destabilized the drug PR complex (figure 96).
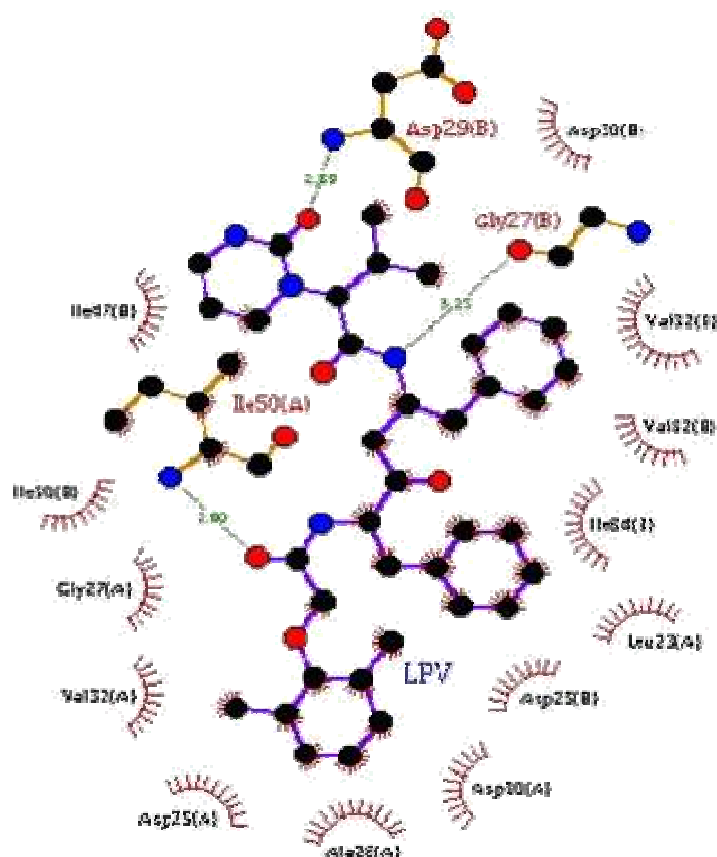
Figure 96. LigPlot schematic diagram showing LPV interactions in the obtained most populated L76V-PR configuration. The hydrogen bonds established by LPV are represented with green dashed lines, while hydrophobic contacts are showed with a red labelled arc [Artese at al., unpublished data].

In ATV-WT most populated configuration, the analysis of the interactions established by the azapeptide inhibitor showed 30 non-bonded contacts and 5 hydrogen bonds, as reported in figure 97.

A good stabilization of ATV through its hydrogen bonds with the catalytic residues and with Gly48, located in the crucial flap region, was noticed.
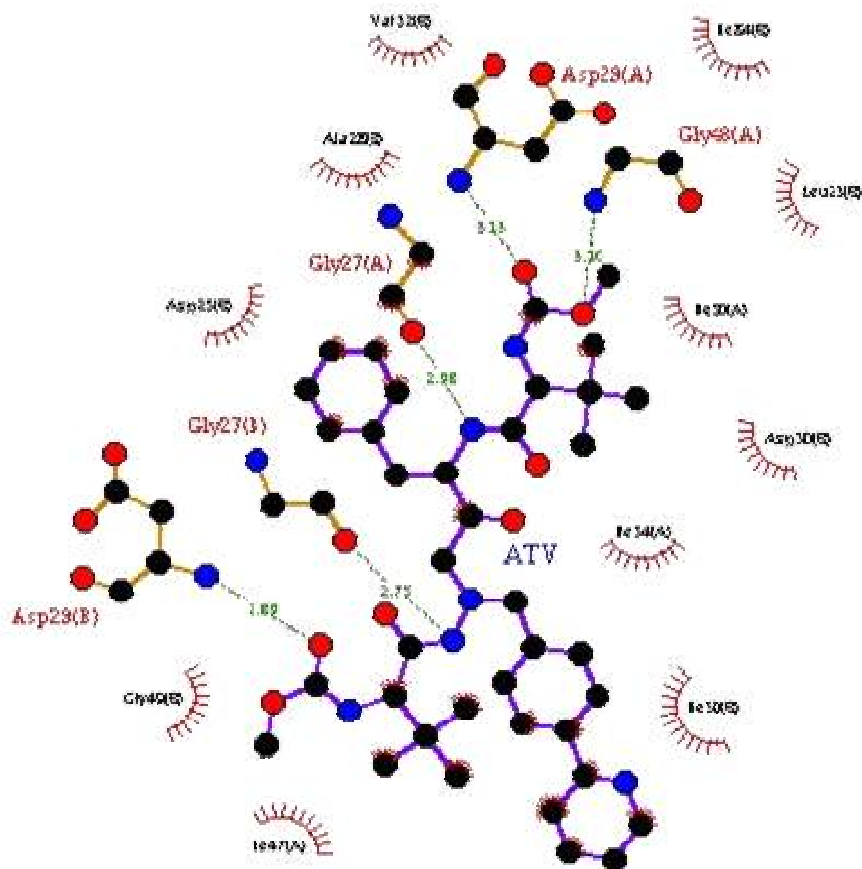
Figure 97. LigPlot schematic diagram showing ATV interactions in the obtained most populated WT-PR configuration. The hydrogen bonds established by ATV are represented with green dashed lines, while hydrophobic contacts are showed with a red labelled arc [Artese at al., unpublished data].

Analyzing ATV most probable configuration in presence of the L33F mutation, 27 non-bonded contacts and 8 hydrogen bonds were identified, with an increasing hydrogen bond network with respect to WT complex. Nonetheless, the L33F mutation caused a less productive binding energy due to the loss of ATV interactions with the catalytic aspartates (figure 98).
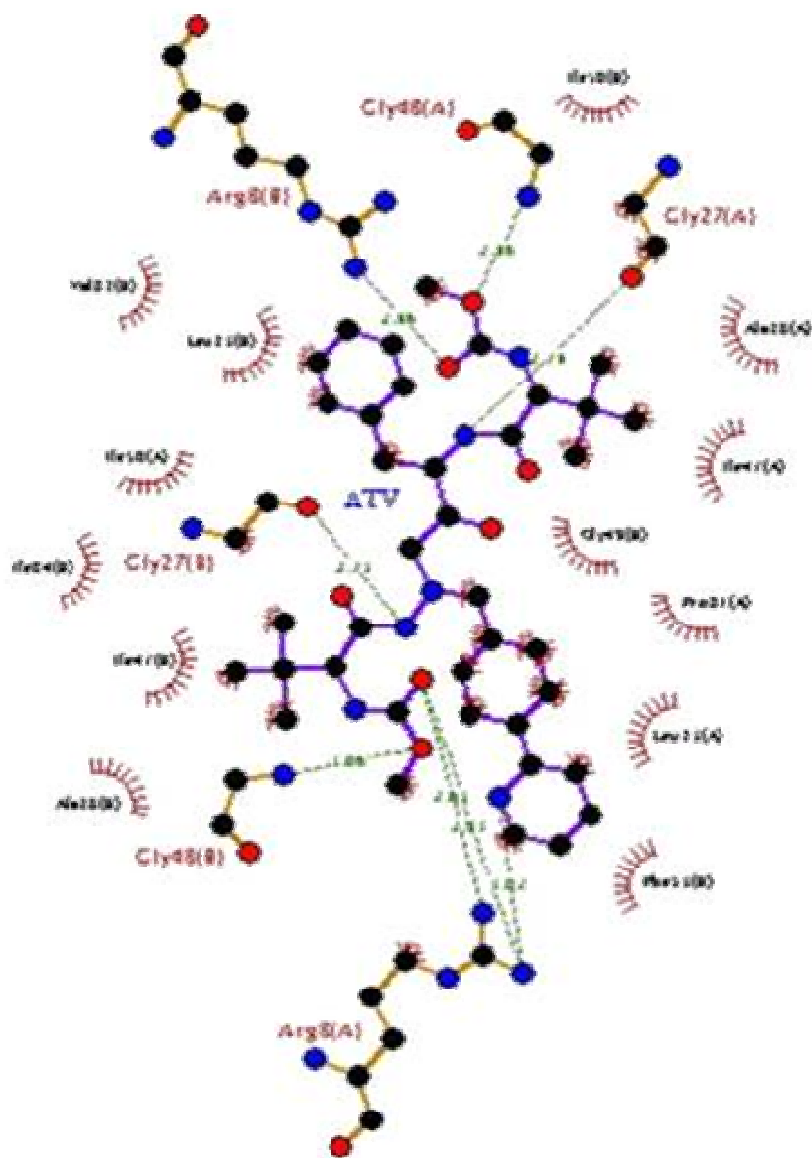
Figure 98. LigPlot schematic diagram showing ATV interactions in the obtained most populated L33F-PR configuration. The hydrogen bonds established by ATV are represented with green dashed lines, while hydrophobic contacts are showed with a red labelled arc [Artese at al., unpublished data].

Finally, analyzing ATV most populated configuration in presence of the L76V mutation, 28 non-bonded contacts and 6 hydrogen bonds were detected (figure 99). In this mutant, ATV showed hydrogen bonds with catalytic residues, but it resulted more stabilized through hydrogen bond with Gly48 and Ile50, residues placed in the flap region found to be crucial in substrate binding [105]. Such a profile is related to an increased drug affinity.
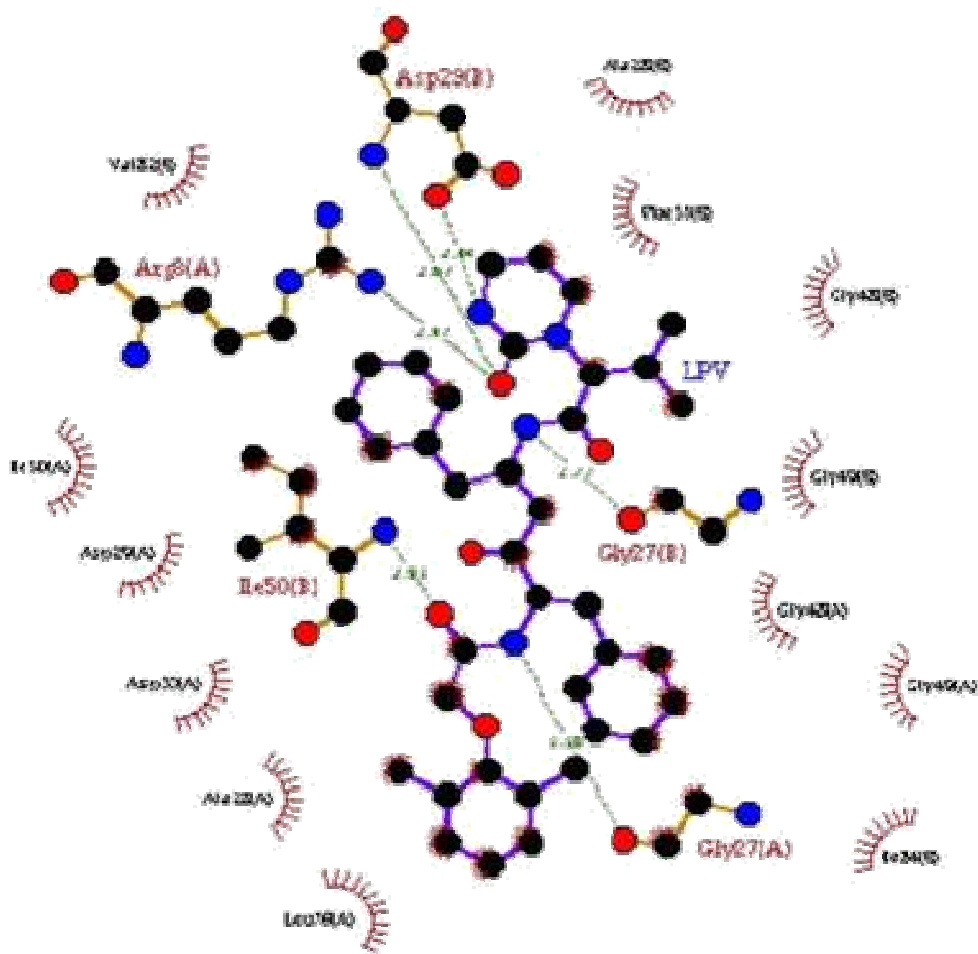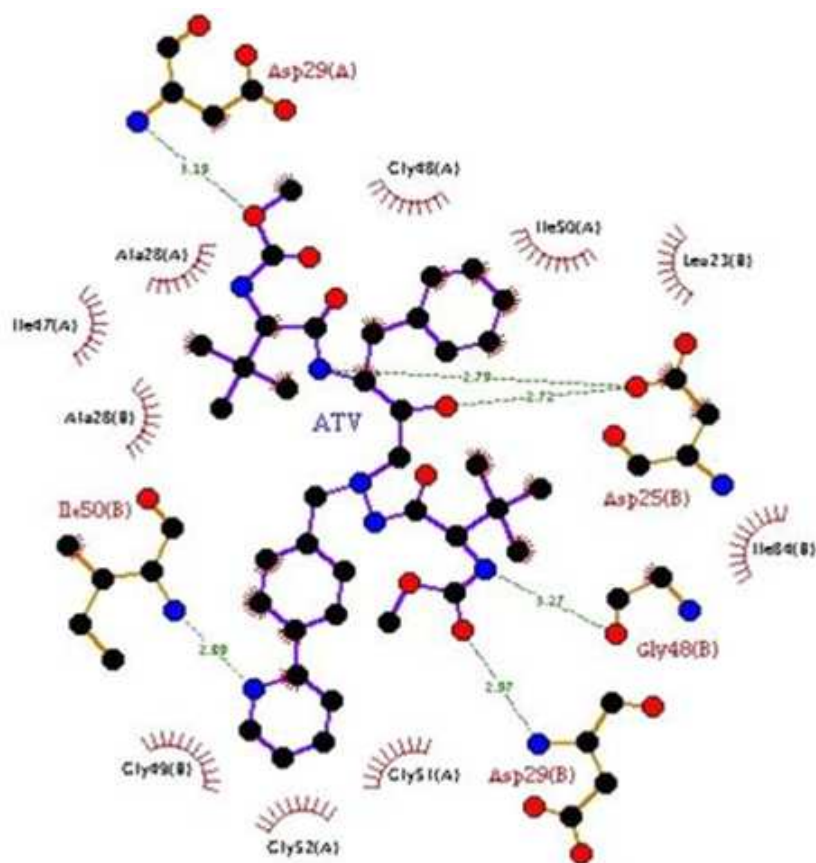
Figure 99. LigPlot schematic diagram showing ATV interactions in the obtained most populated L76V-PR configuration. The hydrogen bonds established by ATV are represented with green dashed lines, while hydrophobic contacts are showed with a red labelled arc [Artese at al., unpublished data].

As reported in literature, the dimer of PR is maintained by interactions between the two subunits, including the terminal residues (1-4 and 96-99), the tips of the flaps (50, 51), Asp29, Arg87 and Arg8′ and residues in and surrounding the active site (residues 24-27) [105]. Thus, in addition to the evaluation of the energetic and molecular recognition terms, the effects of the analyzed mutations were investigated onto the dimerization interface.

In LPV complexes, in presence of L33F mutation, a lower number of good contacts at PR dimer interface was found, as reported in table 9.

RES	WT		L33F		L76V	
	HB	*vdW*	*HB*	*vdW*	*HB*	*vdW*
1 A	1	1	1	1	1	4
1 B	1	1	0	2	2	0
2 A	2	1	1	2	1	8
2 B	1	5	0	1	1	1
3 A	2	2	2	2	2	3
3 B	2	3	2	3	2	2
4 A	0	3	0	2	0	1
4 B	0	1	0	1	0	2
96 A	2	6	3	3	4	5
96 B	4	5	4	7	3	4
97 A	2	5	2	4	2	5
97 B	2	4	2	7	2	2
98 A	3	6	3	1	4	1
98 B	2	1	3	1	4	3
99 A	3	15	0	7	1	2
99 B	1	7	1	7	1	18
# contacts	94		75		91	

Table 9. LPV number of interactions, including van der Waals contacts and hydrogen bonds, at the PR dimerization interface [Artese at al., unpublished data].

In ATV complexes, in presence of L33F mutation, a lower number of good contacts at PR dimer interface was found, while L76V mutant shows a profile quite similar to WT, with more productive interactions (table 10).

RES	WT		L33F		L76V	
	HB	*vdW*	*HB*	*vdW*	*HB*	*vdW*
1 A	1	1	1	0	1	0
1 B	1	4	1	0	1	1
2 A	0	3	0	4	2	3
2 B	1	8	0	3	2	2
3 A	2	2	2	2	2	9
3 B	2	10	2	6	2	2
4 A	0	3	0	1	0	3
4 B	0	0	0	1	0	1
96 A	2	5	3	6	3	3
96 B	3	4	2	8	5	10
97 A	3	14	2	6	2	4
97 B	2	5	2	1	2	3
98 A	2	8	2	4	4	1
98 B	2	4	3	3	3	1
99 A	1	10	1	8	2	7
99 B	1	9	1	4	1	12
# contacts	**113**		**79**		**94**	

Table 10. ATV number of interactions, including van der Waals contacts and hydrogen bonds, at the PR dimerization interface [Artese at al., unpublished data].

Resistance data in presence of the analyzed mutations and expressed as Z-scores index [232] clearly indicated that the substitution L33F is related to a more negative profile for both PR inhibitors in term of binding affinity with respect to WT (table 11).

PI	Z-score			$^{*}\Delta G_{bind}$		
	WT	*L33F*	*L76V*	*WT*	*L33F*	*L76V*
LPV	0.00	8.02	4.00	-29.05	-12.84	-14.15
ATV	0.00	7.25	-0.30	-16.16	-10.81	-23.97

Table 11. Correlation analysis between the Z-score index and the ΔG_{bind} (* values computed by MM-GBSA approach [233] calculated for the three PR studied complexes (WT, L33F and L76V) in presence of the two PR inhibitors (LPV and ATV) [Artese at al., unpublished data].

Such a result was further confirmed by the evaluation of interaction energy expressed as ΔG_{bind}. In fact this analysis revealed that the L33F mutation is associated to higher energetic values for both PR-LPV and PR-ATV complexes with respect to WT, indicating their decreased stability.

Moreover in LPV complexes, L33F substitution is related to reduced host/guest interactions and to a dimer destabilizing effect. In presence of L76V mutation, LPV showed a decreased binding affinity and a reduced hydrogen bonding network.

In ATV complexes, L33F mutation was related to an impaired dimer interface stabilization, to the loss of pivotal contacts in the active site and to a reduced binding affinity.

Interestingly in presence of L76V substitution, ATV revealed a more productive binding affinity, increased host/guest interactions and dimer stabilizing effects, in agreement with hyper susceptibility data, indicated by the negative Z-score value (-0.30).

4.3. gp41 simulations

The fusion glycoprotein gp41 consists of several domains, including an N-terminal fusion peptide, N- and C-terminal heptad repeat regions defined as HR1 and HR2, a transmembrane domain and a cytoplasmic tail. Enfuvirtide (T-20, Fuzeon) is the first fusion inhibitor (FI) approved for clinical practice. It interacts with HR1, preventing the formation of the six-helix bundle, and consequently the fusion process. To date, 18 mutations at 8 positions within the "ENF-target region" encompassing the amino acids 36-45 of HR1 have been associated with ENF-resistance [234]. Additional mutations outside the ENF-target region are also involved in ENF-resistance, and therefore leading to ENF virological failure.

4.3.1. Materials and methods

The X-ray crystallographic coordinates of HIV-1 gp41 deposited in the Protein Data Bank [168] with code 1IF3 were used for the structural analysis (figure 100). Starting from this model the mutants were generated by single residue replacement in all chains. Energy minimized starting structures were submitted to 1 ns of Molecular Dynamic Simulations (MDS) at 300K, storing 200 conformations and using AMBER* force field [224] and GB/SA water implicit model [225]. All simulations were performed by MacroModel ver. 7.2 [226]. To evaluate the most important interactions at interface of HR1 and HR2 subdomains, the programs LigPlot and DimPlot [171] were used to select the most involved gp41 residues. All molecular dynamics conformations were analyzed adopting a distance based descriptor computing the Boltzmann probability at 300K. Six dummy atoms averaging all the atoms of the residues 38 and 140 for each chain were generated and their distances with respect to residue Asn145 were measured.

4.3.2. Results and discussion

With the aim to estimate the stability of its 6-helix bundle, the gp41 conformational properties have been investigated in presence of V38A and N140I, known enfuvirtide resistance-associated mutations.
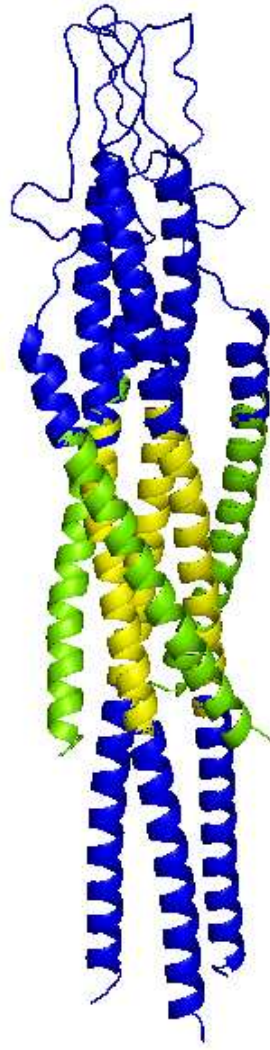
Figure 100. HIV-1 gp41 crystallographic model 1if3 shown as blue cartoon. The two crucial regions HR1 and HR2 are represented in yellow and green, respectively [Artese at al., unpublished data].

The stability of the six-helix bundle requires, in each hairpin, hydrophobic interactions between specific HR1 and HR2 residues [235]. One of them is just residue 38 that establishes van der Waals interaction with residue 145 (figure 101).
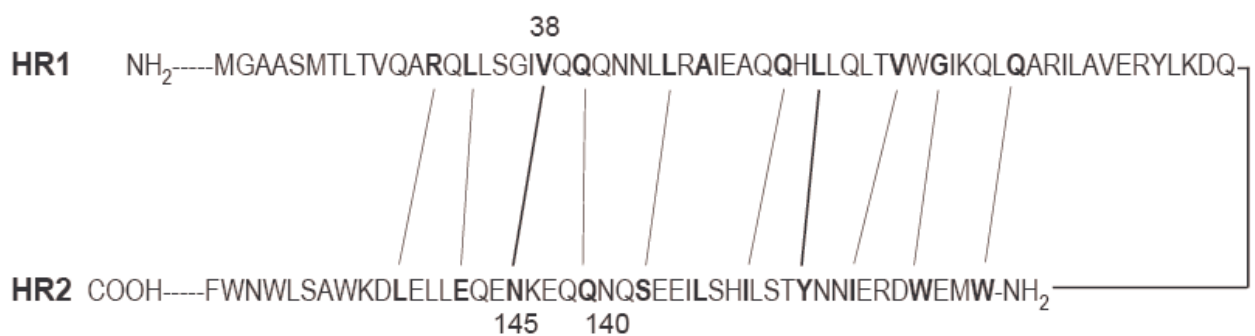
Figure 101. Primary amino acidic sequence of HR1 and HR2 residues of gp41 glycoprotein showing the crucial interactions between specific HR1 and HR2 residues [236].

By performing MDS, the impact of mutations at position 38 and 140 on the formation of this important interaction was investigated, by monitoring over 1 ns, every 5 ps (for 201 total observations), the distance and the frequency of occurrence of the van der Waals interaction between residues 38 and 145.

As reported in figure 102, in WT analyzed gp41 model, for the three glycoproteic hairpins, the interatomic distance between residues 38 and 145 showed values about equal to 5.85 Å, with a final stabilizing effect on the bundle.

Figure 102. Insight into the six-helix bundle in WT gp41 model. The glycoprotein is shown as cartoon and the three chains A, B and C are colored in yellow, light green and dark green, respectively. The gp41 residues involved in crucial interactions are represented in sticks and their interatomic distance, expressed in Å, is indicated with a cyan dashed line [236].

Such a data was confirmed evaluating the interactions of the WT most probable configuration obtained according to Moline method [227] and reported in figure 103.
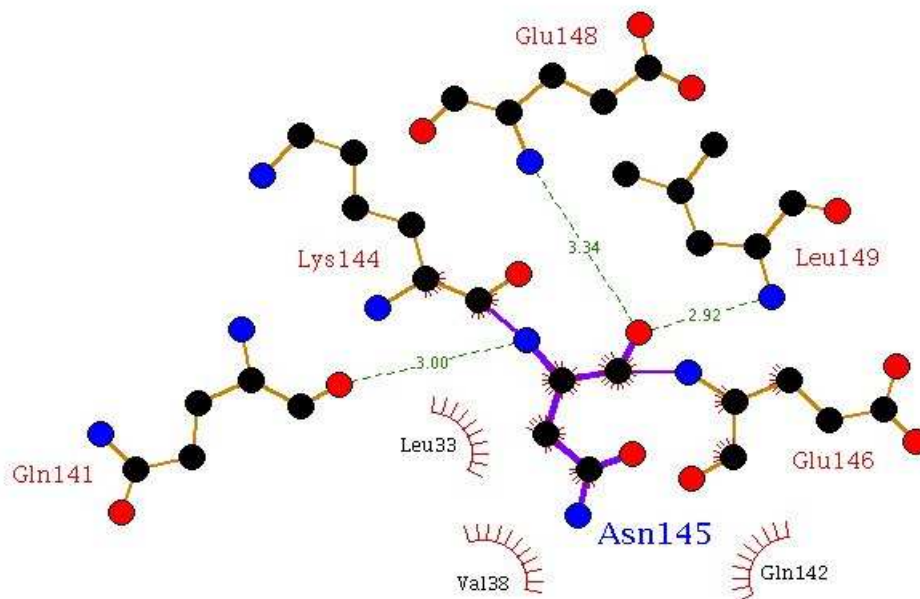
Figure 103. LigPlot schematic diagram showing Asn145 interactions in the obtained most populated WT-gp41 configuration. The hydrogen bonds established by the residue are represented with green dashed lines, while hydrophobic contacts are showed with a red labelled arc [Artese at al., unpublished data].

In WT gp41 model, residue Asn145 was found to establish two pivotal non-bonded contacts with Val38, improving the package of the six-helix bundle.

By contrast, in presence of the mutation V38A, a significant increasing of the evaluated distance was noticed for gp41 hairpin A (figure 104), explaining the absence of the hydrophobic contacts between residues 38 and 145 (figure 105). In fact, although in V38A mutant Asn145 showed four hydrogen bonds with residues Leu33, Glu141, Glu148 and Leu149, the lack of the crucial non-bonded contact with Val38 determined an impairment of the six-helix bundle stabilization.

Figure 104. Insight into the six-helix bundle in V38A gp41 mutant. The glycoprotein is shown as cartoon and the three chains A, B and C are colored in yellow, light green and dark green, respectively. The gp41 residues involved in crucial interactions are represented in sticks and their interatomic distance, expressed in Å, is indicated with a cyan dashed line [236].

Figure 105. LigPlot schematic diagram showing Asn145 interactions in the obtained most populated V38A-gp41 configuration. The hydrogen bonds established by the residue are represented with green dashed lines, while hydrophobic contacts are showed with a red labelled arc [Artese at al., unpublished data].

Interestingly, the substitution N140I caused an important reduction of the studied distance for two gp41 hairpins, with a more remarkable separation of Val38 from Asn145 (figure 106).
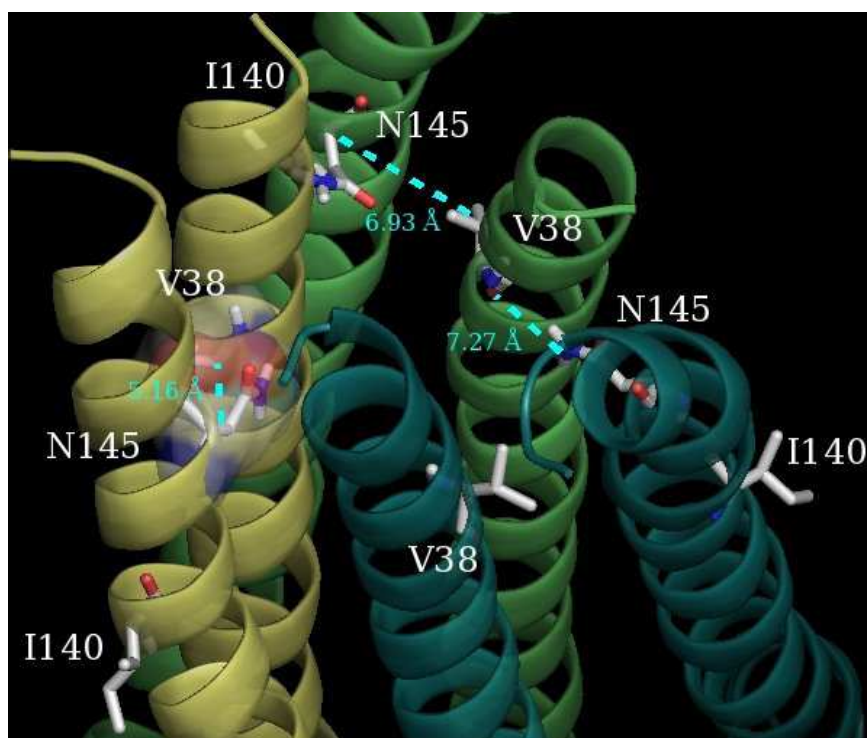
Figure 106. Insight into the six-helix bundle in N140I gp41 mutant. The glycoprotein is shown as cartoon and the three chains A, B and C are colored in yellow, light green and dark green, respectively. The gp41 residues involved in crucial interactions are represented in sticks and their interatomic distance, expressed in Å, is indicated with a cyan dashed line [236].

As reported in figure 107, analyzing the interactions of the most populated N140I configuration, Asn145 was found to form three hydrogen bonds with gp41 residues Gln141, Glu148 and leu149, but, in presence of this mutation, it lost one hydrophobic contact with Val38 with respect to WT.
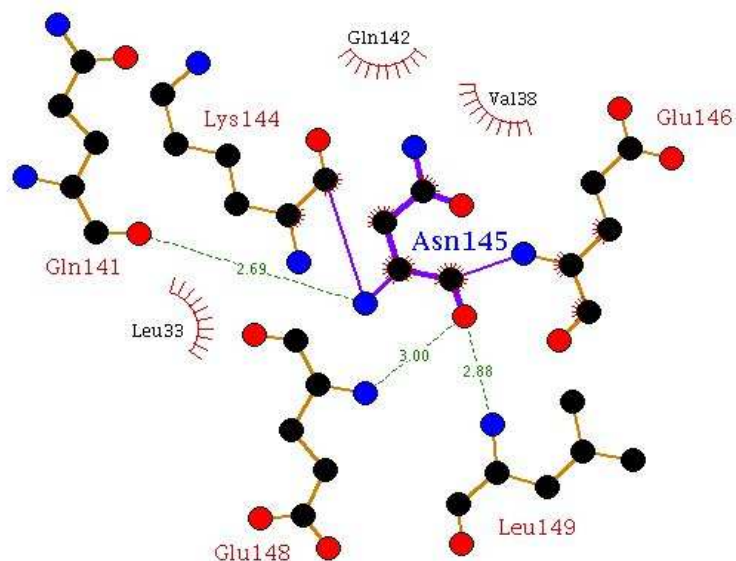
Figure 107. LigPlot schematic diagram showing Asn145 interactions in the obtained most populated N140I-gp41 configuration. The hydrogen bonds established by the residue are represented with green dashed lines, while hydrophobic contacts are showed with a red labelled arc [Artese at al., unpublished data].

Finally the MDS showed that in presence of the double mutation V38A_N140I, the distance between the residues Ala38 and Asn145 was increased in all three hairpins, as indicated in table 12 and in figure 108.

V38wt			V38A			N140I			V38A+140I		
A	B	C	A	B	C	A	B	C	A	B	C
98.0	85.3	100	0.5	94.5	62	97.5	1.0	2.5	0	14	4

Table 12. The frequency of occurrence of Van der Waals interactions between residues 38 and 145 monitored in each hairpin, over 1 ns, every 5 ps, for 201 total observations. A, B, C indicate the three hairpins in the six-helix bundle [236].
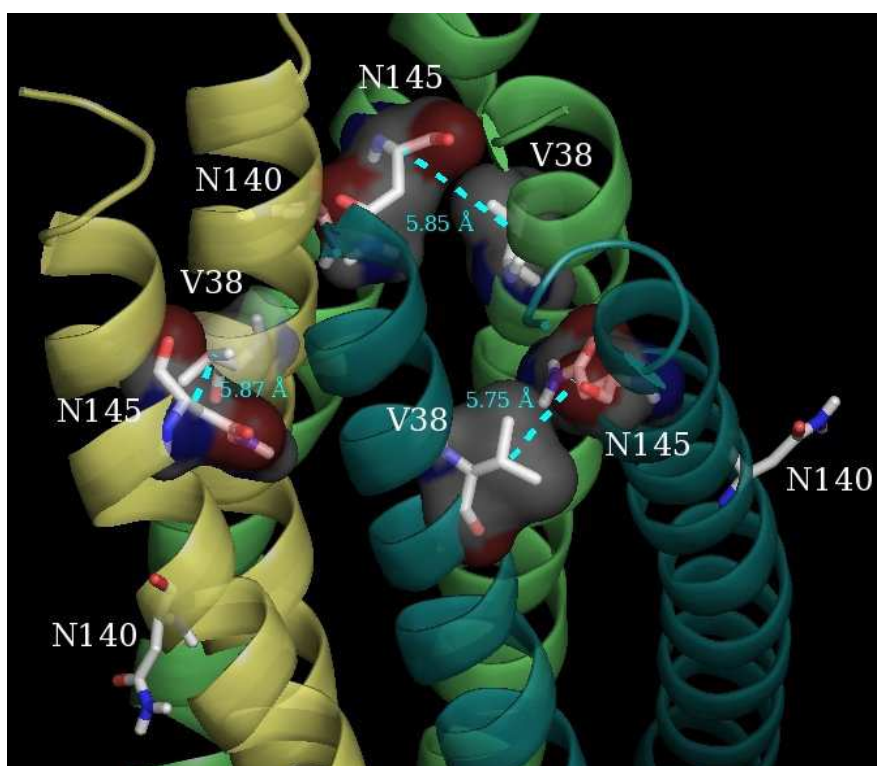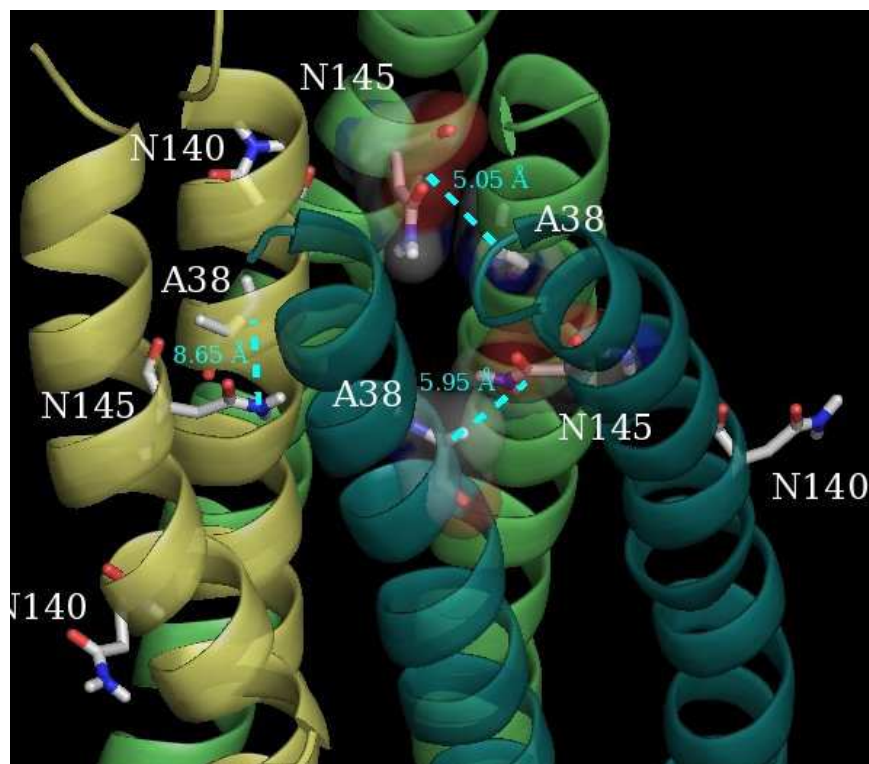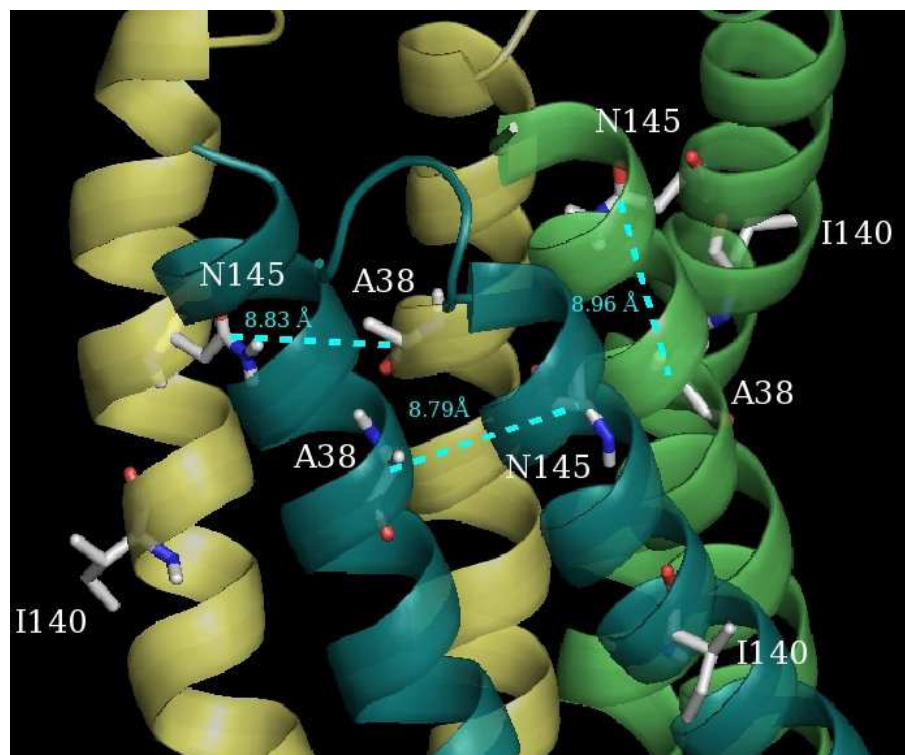
Figure 108. Insight into the six-helix bundle in gp41double mutant V38A_N140I. The glycoprotein is shown as cartoon and the three chains A, B and C are colored in yellow, light green and dark green, respectively. The gp41 residues involved in crucial interactions are represented in sticks and their interatomic distance, expressed in Å, is indicated with a cyan dashed line [236].

Such a result was further validated by the interactions analysis of the most probable V38A_N140I configuration, showing the lock of the non-bonded contacts between residues 38 and 145 (figure 109).

The decreasing of the analyzed distance significantly reduced the frequency of occurrence of the van der Waals interaction in all the three hairpins. Thus, in presence of V38A_N140I, the interaction between residues 38 and 145, important for the formation of the six- helix bundle, is drastically impaired.
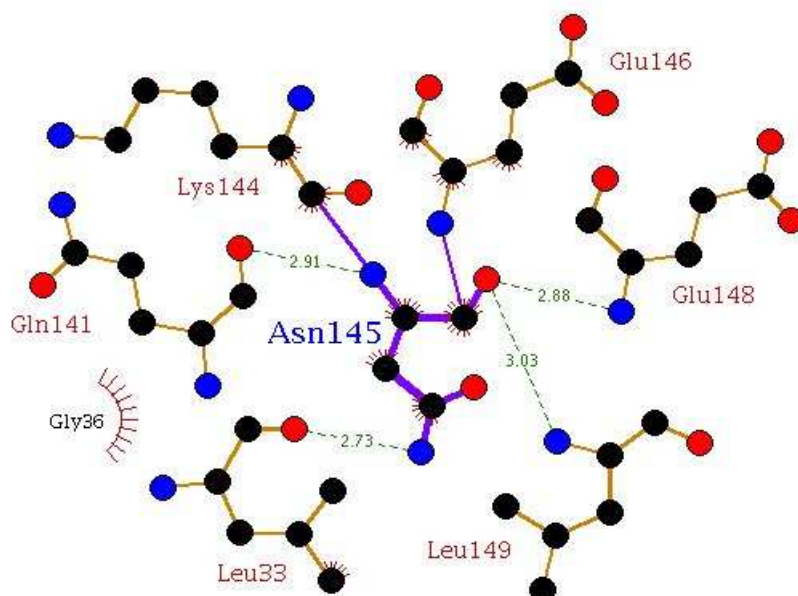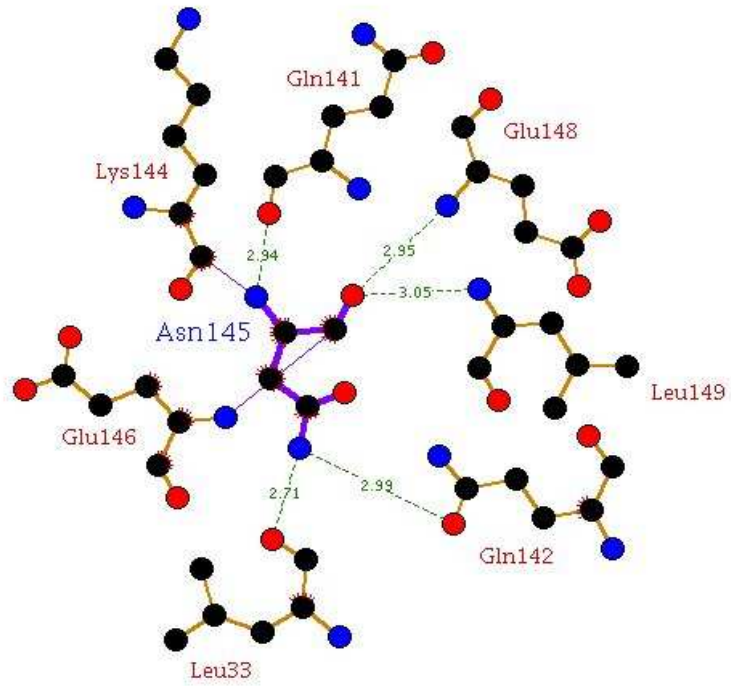
Figure 109. LigPlot schematic diagram showing Asn145 interactions in the obtained most populated V38A_N140I-gp41 configuration. The hydrogen bonds established by the residue are represented with green dashed lines, while hydrophobic contacts are showed with a red labelled arc [Artese at al., unpublished data].

5. Conclusions and further development

During my PhD course different computational studies of mutations associated to resistance in HIV-1 macromolecular targets have been performed with the aim to rationalize the design of novel antiviral agents.

In particular the role of the novel reverse transcriptase mutation I135T in presence of the more common K103N, associated to NNRTIs resistance, was elucidated by means of molecular dynamics simulations (MDS). The performed calculations showed that the phenol oxygen of Tyr188 and the side chain carboxamide of Asn103 resulted closer in the presence of Thr135 than they are in the presence of the WT Ile residue at position 135. In addition, the number of occurrences of a hydrogen bond over 1 ns between Tyr188 and Asn103 in the presence of Thr135 was significantly higher than that in the presence of the Ile135 wild type. This suggests that I135T may favor the stabilization of the closed form of the NNRTI binding pocket induced by K103N. Moreover, analyzing within the NNRTI binding pocket the energy related to the protein-protein electrostatic interaction at the heterodimer p66-p51 interface, it was observed that the mutation K103N itself determined a reduction of this energetic term with respect to the WT sequence and this average energy decreased much more with the presence of the double mutation I135T_K103N.

In the second MDS analysis, the influence of two drug resistance-associated mutations, L33F and L76V, of HIV-1 protease has been evaluated with respect to lopinavir and atazanavir molecular recognition. Resistance data in presence of the analyzed mutations and expressed as Z-scores index clearly indicated that the substitution L33F is related to a more negative profile for both PR inhibitors in term of binding affinity with respect to WT. Such a result was further confirmed by the evaluation of interaction energy expressed as ΔG_{bind}. Thus in LPV complexes, L33F substitution was related to reduced host/guest interactions and to a dimer destabilizing effect. In presence of L76V mutation, LPV showed a decreased binding affinity and a reduced hydrogen bonding network. In ATV complexes, L33F mutation was related to an impaired dimer interface stabilization, to the loss of pivotal contacts in the active site and to a reduced binding affinity.

Interestingly in presence of L76V substitution, ATV revealed a more productive binding affinity, increased host/guest interactions and dimer stabilizing effects, in agreement with hyper susceptibility data.

With the aim to estimate the stability of its 6-helix bundle, the fusion glycoprotein gp41 conformational properties have been investigated in presence of V38A and N140I, known

enfuvirtide resistance-associated mutations. The performed MDS showed that in presence of the double mutation V38A_N140I, the distance between the residues Ala38 and Asn145 was increased in all three hairpins. Such a result was further validated by the interactions analysis of the most probable V38A_N140I configuration. The decreasing of the analyzed distance significantly reduced the frequency of occurrence of the van der Waals interaction in all the three hairpins. Thus, in presence of V38A_N140I, the interaction between residues 38 and 145, important for the formation of the six- helix bundle, is drastically impaired.

Finally a new RT pharmacophore model, useful for further drug design studies, has been defined applying the innovative GBPM computational method to 96 crystallographic models of HIV-1 reverse transcriptase. The maps analyses confirmed the importance of the most involved RT residues in productive interactions at the three crucial drugs sites (dimer interface, polymerase binding site and NNRTI binding pocket). RT interface GBPM application led to the identification of a large mapping area of the β7-β8 loop of p51 subunit, known to interact with heterodimer stability interfering TSAO compounds. Since the crucial residues in TSAO-RT interactions resulted well recognized with GBPM analysis, the pharmacophore model, screening a large interface region and the tenofovir binding site, could lead to design a molecular hybrid linked with a spacer. Thus tenofovir association with dimerization synthetic inhibitors (TSAO, BBNH), or with peptidic inhibitors, could be considered an interesting starting point in order to obtain new antiretroviral drugs. The identification of new hypothesis in agreement to structural and clinical data, mediated by the reduction of features number in all RT models according to energetic criteria, is in progress. Moreover the virtual screening of compounds able to fit with the pharmacophoric hypothesis generated from PDB model 1t05 is still running.

7. References

1. Anonymous, *Pneumocystis pneumonia*: Los Angeles. MMWR Morb Mortal Wkly Rep., **1981**, *30*, 250-252.

2. Gottlieb, M.S.; Schroff, R.; Schanker, H.M.; Weisman, J.D.; Fan, P.T.; Wolf, R.A.; Saxon, A. *Pneumocystis carinii* pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. *N. Engl. J. Med.*, **1981**, *305*, 1425-1431.

3. Selik, R.M.; Haverkos, H.W.; Curran, J.W. Acquired immune deficiency syndrome (AIDS) trends in the United States, 1978-1982. *Am. J. Med.*, **1984**, *76*, 493-500.

4. Ammann, A.J.; Abrams, D.; Conant, M.; Chudwin, D.; Cowan, M.; Volberding, P.; Lewis, B.; Casavant, C. Acquired immune dysfunction in homosexual men: immunologic profiles. *Clin. Immunol. Immunopathol.*, **1983**, *27*, 315-325.

5. Gallo, R.C.; Sarin, P.S.; Gelmann, E.P.; Robert-Guroff, M.; Richardson, E.; Kalyanaraman, V.S.; Mann, D.; Sidhu, G.D.; Stahl, R.E.; Zolla-Pazner, S.; Leibowitch, J.; Popovic, M. Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science*, **1983**, *220*, 865-867.

6. Barre-Sinoussi, F.; Chermann J.C.; Rey, F.; Nugeyre, M.T.; Chamaret, S.; Gruest, J.; Dauguet, C.; Axler-Blin, C.; Vèzinet-Brun, F.; Rouzioux, C.; Rozenbaum, W.; Montagnier, L. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*, **1983**, *220*, 868-871.

7. Gallo, R.C.; Salahuddin, S.; Popovic, M.; Shearer, G.M.; Kaplan, M.; Haynes, B.F.; Palker, T.J.; Redfield, R.; Oleske, J.; Safai, B. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science*, **1984**, *224*, 500-503.

8. Levy, J.A.; Hoffman, A.D.; Kramer, S.M.; Landis, J.A.; Shimabukuro, J.M.; Oshiro, L.S. Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science*, **1984**, *225*, 840-842.

9. Shaw, G.M.; Hahn, B.H.; Arya, S.K.; Groopman, J.E.; Gallo, R.C.; Wong-Staal, F. Molecular characterization of human T-cell leukaemia (lymphotropic) virus type III in the acquired immune deficiency syndrome. *Science*, **1984**, *226*, 1165-1171.

10. Groopman, J.E.; Mayer, K.H.; Sarngadharan, M.G.; Ayotte, D.; Devico, A.L.; Finberg, R.; Sliski, A.H.; Allan, J.D.; Gallo, R.C. Seroepidemiology of human T-lymphotropic virus type III among homosexual men with the acquired immunodeficiency syndrome or generalized lymphadenopathy and among asymptomatic controls in Boston. *Ann. Intern. Med.*, **1985**, *102*, 334-337.

11. Gao, F.; Bailes, E.; Robertson, D.L.; Chen, Y.; Rodenburg, C.M.; Michael, S.F.; Cummins, L.B.; Arthur, L.O.; Peeters, M.; Shaw, G.M.; Sharp, P.M.; Hahn, B.H. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes. Nature*, **1999**, *397*, 436-441.

12. UNAIDS Joint United Nations Programme on HIV/AIDS. *AIDS epidemic update: December.* Geneva: UNAIDS, **2000**.

13.   Leitner, T. Genetic subtypes of HIV-1: In: Myers, G.; Foley, B.; Mellors, J., et al., eds. *Human retroviruses and AIDS 1996*. Vol III. Los Alamos, NM: Los Alamos National Laboratory, **1996**, 28-40.

14.   Essex, M. Human immunodeficiency viruses in the developing world. *Adv. Virus Res.*, **1999**, *53*, 71-88.

15.   http://www.abbott.com

16.   Gangakhedkar, R.R.; Bentley, M.E.; Divekar, A.D.; Gadkari, D.; Mehendale, S.M.; Shepherd, M.E.; Bollinger, R.C.; Quinn, T.C. Spread of HIV infection in married monogamous women in India. *JAMA*, **1997**, *278*, 2090-2092.

17.   Palella, F.J.; Delaney, K.M.; Moorman, A.C.; Loveless, M.O.; Fuhrer, J.; Satten, G.A.; Aschman, D.J.; Holmberg, S.D. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *N. Engl. J. Med.*, **1998**, *338*, 853-860.

18.   The White House Office of National AIDS Policy. *Report on the presidential mission on children orphaned by AIDS in sub-Saharan Africa: findings and plan of action.* Washington, DC: The White House Office of National AIDS Policy, **1999**.

19.   N'Galy, B. and Ryder, R.W. Epidemiology of HIV infection in Africa. *Journal of Acquired Immune Deficiency Syndromes*, **1988**, *1*, 551-558.

20.   Rothenberg, R.B.; Scarlett, M.; del Rio, C.; Reznik, D. and O'Daniels, C. Oral transmission of HIV. *AIDS* , **1998**, *12*, 2095-2105.

21.   Mastro, T.D.; De Vincenti, I. Probabilities of sexual HIV-1 transmission. *AIDS*, **1996**, 10 (Suppl A): S75-S82.

22.   World Health Organization, **2006**. WHO Multi-country Study on Women's Health and Domestic Violence against Women. Retrieved on 2006-12-14.

23.   Koenig, M.; Zablotska, I.; Lutalo, T.; Nalugoda, F.; Wagman, J.; Gray, R. Coerced first intercourse and reproductive health among adolescent women in Rakai, Uganda. *International Family Planning Perspectives*, **2004**, *30*, 156.

24.   Laga, M.; Nzila, N.; Goeman, J. The interrelationship of sexually transmitted diseases and HIV infection: implications for the control of both epidemics in Africa. *AIDS*, **1991**, 5 (Suppl 1): S55-S63.

25.   Tovanabutra, S.; Robison, V.; Wongtrakul, J.; Sennum, S.; Suriyanon, V.; Kingkeow, D.; Kawichai, S.; Tanan, P.; Duerr, A.; Nelson, K.E. Male viral load and heterosexual transmission of HIV-1 subtype E in northern Thailand. *J. Acquir. Immune. Defic. Syndr.*, **2002**, *29*, 275-283.

26.   Sagar, M.; Lavreys, L.; Baeten, J.M.; Richardson, B.A.; Mandaliya, K.; Ndinya-Achola, J.O.; Kreiss, J.K. and Overbaugh, J. Identification of modifiable factors that affect the genetic diversity of the transmitted HIV-1 population. *AIDS*, **2004**, *18*, 615-619.

27.   Smith, D.K.; Grohskopf, L.A.; Black, R.J.; Auerbach, J.D.; Veronese, F.; Struble, K.A.; Cheever, L.; Johnson, M.; Paxton, L.A.; Onorato, I.A. and Greenberg, A.E. Antiretroviral Postexposure Prophylaxis After Sexual, Injection-Drug Use, or Other Nonoccupational Exposure to HIV in the United States. *MMWR*, **2005**, *54*, 1-20.

28. Physicians for Human Rights (2003-03-13). HIV Transmission in the Medical Setting: A White Paper by Physicians for Human Rights. Partners in Health. Retrieved on 2006-03-01.

29. WHO, **2001**. Blood safety....for too few. Retrieved on 2006-01-17.

30. Coovadia, H. Antiretroviral agents: how best to protect infants from HIV and save their mothers from AIDS. *N. Engl. J. Med.*, **2004**, *351*, 289-292.

31. Sperling, R.S.; Shapirom, D.E.; Coombsm, R.W.; Todd, J.A.; Herman, S.A.; McSherry, G.D.; O'Sullivan, M.J.; Van Dyke, R.B.; Jimenez, E.; Rouzioux, C.; Flynn, P.M.; Sullivan, J.L. Maternal viral load, zidovudine treatment, and the risk of transmission of human immunodeficiency virus type 1 from mother to infant. *N. Engl. J. Med.*, **1996**, *335*, 1621-1629.

32. Berry, S. (2006-06-08). Children, HIV and AIDS. avert.org. Retrieved on 2006-06-15.

33. Dias, S.F.; Matos, M.G. and Goncalves, A.C. Preventing HIV transmission in adolescents: an analysis of the Portuguese data from the Health Behaviour School-aged Children study and focus groups. *Eur. J. Public Health*, **2005**, *15*, 300-304.

34. Buchbinder, S.P.; Katz, M.H.; Hessol, N.A.; O'Malley, P.M.; Holmberg, S.D. Long-term HIV-1 infection without immunologic progression. *AIDS*, **1994**, *8*, 1123-1128.

35. Masur, H.; Ognibene, F.; Yarchoan, R.; Shelhamer, J.H.; Baird, B.F.; Travis, W.; Suffredini, A.F.; Deyton, L.; Kovacs, J.A.; Falloon, J. CD4 counts as predictors of opportunistic pneumonias in human immunodeficiency virus (HIV) infection. *Ann. Intern. Med.*, **1989**, *111*, 223-231.

36. Mellors, J.W.; Munoz, A.; Giorgi, J.V.; Margolick, J.B.; Tassoni, C.J.; Gupta, P.; Kingsley, L.A.; Todd, J.A.; Saah, A.J.; Detels, R.; Phair, J.P.; Rinaldo, C.R. Jr. Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. *Ann. Intern. Med.*, **1997**, *126*, 946-954.

37. Schacker, T.; Collier, A.C.; Hughes, J.; Shea, T.; Corey, L. Clinical and epidemiological features of primary HIV infection. *Ann. Intern. Med.*, **1996**, *125*, 257-264.

38. Kinloch-de Loes, S.; De Saussure, P.; Saurat, J.H.; Stalder, H. ; Hirschel, B. ; Perrin, L.H. Symptomatic primary infection due to human immunodeficiency virus type 1: Review of 31 cases. *Clin. Infect. Dis.*, **1993**, *17*, 59-65.

39. Centers for Disease Control and Prevention. 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR Morb Mortal Weekly Rep.*, **1992**, *41*, 1-19.

40. Albert, J.; Abrahamsson, B.; Nagy, K.; Aurelius, E.; Gaines, H.; Nystrom, G.; Fenyo, E.M. Rapid development of isolate-specific neutralizing antibodies after primary HIV-1 infection and consequent emergence of virus variants which resist neutralization by autologous sera. *AIDS*, **1990**, *4*, 107-112.

41. Centers for Disease Control and Prevention. *HIV/AIDS surveillance report*. Atlanta: CDC, **1999**, *11*, 1-44.

42. Deacon, N.J.; Tsykin, A.; Solomon, A.; Smith, K.; Ludford-Menting, M.; Hooker, D.J.; McPhee, D.A.; Greenway, A.L.; Ellett, A.; Chatfield, C.; Lawson, V.A.; Crowe, S.; Maerz, A.; Sonza, S.; Learmont, J.; Sullivan, J.S.; Cunningham, A.; Dwyer, D.; Dowton, D.; Mills, J. Genomic structure of an attenuated quasi species of HIV-1 from a blood transfusion donor and recipients. *Science*, **1995**, *270*, 988-991.

43. Huang, Y.; Zhang, L.; Ho, D. Biological characterization of nef in long term survivors of human immunodeficiency virus type 1 infection. *J. Virol.*, **1995**, *69*, 8142-8146.

44. Carrington, M.; Nelson, G.W.; Martin, M.P.; Kissner, T.; Vlahov, D.; Goedert, J.J.; Kaslow, R.; Buchbinder, S.; Hoots, K.; O'Brien, S.J. HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science*, **1999**, *283*, 1748-1752.

45. Dean, M.; Carrington, M.; Winkler, C.; Huttley, G.A.; Smith, M.W.; Allikmets, R.; Goedert, J.J.; Buchbinder, S.P.; Vittinghoff, E.; Gomperts, E.; Donfield, S.; Vlahov, D.; Kaslow, R.; Saah, A.; Rinaldo, C.; Detels, R.; O'Brien, S.J. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science*, **1996**, *273*, 1856-1862.

46. Robinson, W.E.; Montefiori, D.C.; Mitchell, W.M. Antibody-dependent enhancement of human immunodeficiency virus type 1 infection. *Lancet*, **1988**, *1*, 790-794.

47. Pantaleo, G.; Graziosi, C.; Demarest, J.; Cohen, O.J.; Vaccarezza, M.; Gantt, K.; Muro-Cacho, C.; Fauci, A.S. Role of lymphoid organs in the pathogenesis of human immunodeficiency virus (HIV) infection. *Immunol. Rev.*, **1994**, *140*, 105-130.

48. CDC. *Pneumocystis pneumonia* - Los Angeles. *MMWR Morb Mortal Weekly Rep*, **1981**, *30*, 250-252.

49. Barrè-Sinoussi, F.; Chermann, J.C.; Rey, F.; Nugeyre, M.T.; Chamaret, S.; Gruest, J.; Dauguet, C.; Axler-Blin, C.; Vèzinet-Brun, F.; Rouzioux, C.; Rozenbaum, W.; Montagnier, L. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*, **1983**, *220*, 868-871.

50. Coffin, J.; Haase, A.; Levy, J.A.; Montagnier, L.; Oroszlan, S.; Teich, N.; Temin, H.; Toyoshima, K.; Varmus, H.; Vogt, P. Human immunodeficiency viruses. *Science*, **1986**, *232*, 697.

51. Clavel, F.; Guètard, D.; Brun-Vèzinet, F.; Chamaret, S.; Rey, M.A.; Santos-Ferreira, M.O.; Laurent, A.G.; Dauguet, C.; Katlama, C.; Rouzioux, C. Isolation of a new human retrovirus from West African patients with AIDS. *Science*, **1986**, *233*, 343-346.

52. Gonda, M.A.; Wong-Staal, F.; Gallo, R.C.; Clements, J.E.; Narayan, O.; Gilden, R.V. Sequence homology and morphologic similarity of HTLV-III and visna virus, a pathogenic lentivirus. *Science*, **1985**, *227*, 173-177.

53. Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Doran, E.R.; Rafalski, J.A.; Whitehorn, E.A.; Baumeister, K. Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature*, **1985**, *313*, 277-284.

54. Benn, S.; Rutledge, R.; Folks, T.; Gold, J.; Baker, L.; McCormick, J.; Feorino, P.; Piot, P.; Quinn, T.; Martin, M. Genomic heterogeneity of AIDS retroviral isolates from North America and Zaire. *Science*, **1985**, *230*, 949-951.

55. Saag, M.S.; Hahn, B.H.; Gibbons, J.; Li, Y.; Parks, W.P.; Shaw, G.M. Extensive variation of human immunodeficiency virus type-1 *in vivo*. *Nature*, **1988**, *334*, 440-444.

56. Meyerhans, A.; Cheynier, R.; Albert, J.; Seth, M.; Kwok, S.; Sninsky, J.; Morfeldt-Manson, L.; Asjo, B.; Wan-Hobson, S. Temporal fluctuations in HIV quasispecies *in vivo* are not reflected by sequential HIV isolation. *Cell*, **1989**, *58*, 901-910.

57. Myers, G. Tenth anniversary perspectives on AIDS. HIV: between past and future. *AIDS Res. Hum. Retroviruses*, **1994**, *10*, 1317-1324.

58. Subbarrao, S.; Schochetman, G. Genetic variability of HIV-1. *AIDS*, **1996**, *10*, S13-23.

59. Simon, F.; Mauclère, P.; Roques, P.; Loussert-Ajaka, I.; Muller-Trutwin, M.C.; Saragosti, S.; Georges-Courbot, M.C.; Barrè-Sinoussi, F.; Brun-Vèzinet, F. Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat. Med.,* **1998**, *4*, 1032-1037.

60. http:// www.medinet.hochiminhcity.gov

61. WainHobson, S. HIV genome variability *in vivo*, AIDS 3, **1989**, supp 1, 139.

62. Moulard, M; Decroly, E. Maturation of HIV envelope glycoprotein precursors by cellular endoproteases. *Biochim. Biophys. Acta*, **2000**, *1469*, 121-132.

63. Charneau, P.; Alizon, M.; Clavel, F. A second origin of DNA plus-strand synthesis is required for optimal human immunodeficiency virus replication. *J. Virol.*, **1992**, *66*, 2814-2820.

64. Harrich, D.; Ulich, C.; Gaynor, R.B. A critical role for the TAR element in promoting efficient human immunodeficiency virus type 1 reverse transcription. *J. Virol.*, **1996**, *70*, 4017-4027.

65. Choe, H.; Farzan, M.; Sun, Y.; Sullivan, N.; Rollins, B.; Ponath, P.D.; Wu, L.; Mackay, C.R.; LaRosa, G.; Newman, W.; Gerard, N.; Gerard, C.; Sodroski, J. The β-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell*, **1996**, *85*, 1135-1148.

66. Deng, H.; Liu, R.; Ellmeier, W.; Choe, S.; Unutmaz, D.; Burkhart, M.; Di Marzio, P.; Marmon, S.; Sutton, R.E.; Hill, C.M.; Peiper, S.C.; Schall, T.J.; Littman, D.R.; Landau, N.R. Identification of a major co-receptor for primary isolates of HIV-1. *Nature*, **1996**, *381*, 661-666.

67. Bowerman, B.; Brown, P.O.; Bishop, J.M.; Varmus, H.E. A nucleoprotein complex mediates the integration of retroviral DNA. *Genes Dev.*, **1989**, *3*, 469-478.

68. Brown, P.O.; Bowerman, B.; Varmus, H.E.; Bishop, J.M. Correct integration of retroviral DNA *in vitro*. *Cell*, **1987**, *49*, 347-356.

69. Schwartzberg, P.; Colicelli, J, Goff, S.P. Construction and analysis of deletion mutations in the Pol gene of Moloney murine leukemia virus: a new viral function required for productive infection. *Cell*, **1984**, *37*, 1043-1052.

70. Pomerantz, R.J.; Horn, D.L. Twenty years of therapy for HIV-1 infection. *Nat Med.*, **2003**, *9*, 867-873.

71. Gotte, M.; Li, X.; Wainberg, M.A. HIV-1 reverse transcription: a brief overview focused on structure-function relationships among molecules involved in initiation of the reaction. *Arch. Biochem. Biophys.*, **1999**, *365*, 199-210.

72. Kohlstaedt, L.A.; Wang, J.; Friedman, J.M.; Rice, P.A.; Steitz, T.A. Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science*, **1992**, *256*, 1783-1790.

73. Hsiou, Y.; Ding, J.; Das, K.; Clark Jr., A.D.; Hughes, S.H.; Arnold, E. Structure of unliganded HIV-1 reverse transcriptase at 2.7 Å resolution: implications of conformational changes for polymerization and inhibition mechanisms. *Structure*, **1996**, *4*, 853-860.

74. Le Grice, S.F.J.; Nass, T.; Wohlgensinger, B.; Schatz, O. Subunit-selective mutagenesis indicates minimal polymerase activity in heterodimer-associated p51 HIV-1 reverse transcriptase. *EMBO J.*, **1991**, *10*, 3905-3911.

75. Wang, J.; Smerdon, S.J.; Jager, J.; Kohlstaedt, L.A.; Rice, P.A.; Friedman, J.M. Structural basis of asymmetry in the human immunodeficiency virus type 1 reverse transcriptase heterodimer. *Proc. Natl. Acad. Sci. U.S.A.,* **1994**, *91*, 7242-7246.

76. Camarasa, M.J.; Velàzquez, S.; San-Fèlix, A.; Pèrez-Pèrez, M.J.; Gago, F. Dimerization inhibitors of HIV-1 reverse transcriptase, protease and integrase: a single mode of inhibition for the three HIV enzymes? *Antiviral Res.*, **2006**, *71*, 260-267.

77. Huang, H.; Chopra, R.; Verdine, G.L.; Harrison, S.C. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science*, **1998**, *282*, 1669-1675.

78. Ding, J.; Das, K.; Hsiou, Y.; Sarafianos, S.G.; Clark, A.D. Jr, Jacobo-Molina, A.; Tantillo, C.; Hughes, S.H.; Arnold, E. Structure and functional implications of the polymerase active site region in a complex of HIV-1 RT with a double-stranded DNA template-primer and an antibody Fab fragment at 2.8 Å resolution. *J. Mol. Biol.*, **1998**, *284*, 1095-1111.

79. Julias, J.G.; McWilliams, M.J.; Sarafianos, S.G.; Alvord, W.G.; Arnold, E.; Hughes, S.H. Mutation of amino acids in the connection domain of human immunodeficiency virus type 1 reverse transcriptase that contact the template-primer affects RNase H activity. *J. Virol.*, **2003**, *15*, 8548-8554.

80. Tachedjian, G.; Aronson, H.E.G.; Goff, S.P. Analysis of mutations and suppressors affecting interactions between the subunits of the HIV type 1 reverse transcriptase. *Proc. Natl. Acad. Sci. U.S.A.*, **2000**, *97*, 6334-6339.

81. Mulky, A.; Sarafianos, S.G.; Jia, Y.; Arnold, E.; Kappes, J.C. Identification of amino acid residues in the Human Immunodeficiency Virus type-1 reverse transcriptase tryptophan-repeat motif that are required for subunit interaction using infectious virions. *J. Mol. Biol.*, **2005**, *349*, 673-684.

82. Baillon, J.G.; Nashed, N.T.; Kumar, A.; Wilson, S.H.; Jerina, D.M. A leucine zipper-like motif may mediate HIV reverse transcriptase subunit binding. *New Biol.*, **1991**, *3*, 1015-1019.

83. Becerra, S.P.; Kumar, A.; Lewis, M.S.; Widen, S.G.; Abbotts, J.; Karawya, E.M.; Hughes, S.H.; Shiloach, J.; Wilson, S.H. Protein-protein interactions of HIV-1 reverse transcriptase: implication of central and C-terminal regions in subunit binding. *Biochemistry*, **1991**, *30*, 11707-11719.

84. Auwerx, J.; Van Nieuwenhove, J.; Rodrìguez-Barrios, F.; de Castro, S.; Velàzquez, S.; Ceccherini-Silberstein, F.; De Clercq, E.; Camarasa, M.J.; Perno, C.F.; Gago, F.; Balzarini, J. The N137 and P140 amino acids in the p51 and the P95 amino acid in the p66 subunit of human immunodeficiency virus type 1 (HIV-1) reverse transcriptase are instrumental to maintain catalytic activity and to design new classes of anti-HIV-1 drugs. *FEBS Lett.*, **2005**, *579*, 2294-2300.

85. Pandey, P.K.; Kaushik, N.; Singh, K.; Sharma, B.; Upadhyay, A.K.; Kumar, S.; Harris, D.; Pandey, V.N. Insertion of a small peptide of six amino acids into the β7–β8 loop of the p51 subunit of HIV-1 reverse transcriptase perturbs the heterodimer and affects its activities. *Biochemistry*, **2002**, *3*, 1-18.

86. Sluis-Cremer, N.; Temiz, N.A.; Bahar, I. Conformational changes in HIV-1 reverse transcriptase induced by nonnucleoside reverse transcriptase inhibitor binding. *Curr. HIV Res.*, **2004**, *2*, 323-332.

87. Das, K.; Lewi, P.J.; Hughes, S.H.; Arnold, E. Crystallography and the design of anti-AIDS drugs: conformational flexibility and positional adaptability are important in the design of non-nucleoside HIV-1 reverse transcriptase inhibitors. *Prog. Biophys. Mol. Biol.*, **2005**, *88*, 209-231.

88. Rodríguez-Barrios, F.; Gago, F. Understanding the basis of resistance in the irksome Lys103Asn HIV-1 reverse transcriptase mutant through targeted molecular dynamics simulations. *J. Am. Chem. Soc.*, **2004**, *126*, 15386-15387.

89. Tachedjian, G.; Orlova, M.; Sarafianos, S.G.; Arnold, E.; Goff, S.P. Nonnucleoside reverse transcriptase inhibitors are chemical enhancers of dimerization of the HIV type 1 reverse transcriptase. *Proc. Natl. Acad. Sci. U.S.A.*, **2001**, *98*, 7188-7193.

90. Rodrìguez-Barrios, F.; Balzarini, J.; Gago, F. The molecular basis of resilience to the effect of the Lys103Asn mutation in Non-Nucleoside HIV-1 reverse transcriptase inhibitors studied by targeted molecular dynamics simulations. *J. Am. Chem. Soc.*, **2005**, *127*, 7570-7578.

91. Crawford, S.; Goff, S.P. A deletion mutation in the 5′ part of the *pol* gene of Moloney murine leukemia virus blocks proteolytic processing of the *gag* and *pol* polyproteins. *J. Virol.*, **1997**, *71*, 1089-1096.

92. Navia, M.A.; Fitzgerald, P.M.D.; McKeever, B.M.; Leu, C.T.; Heimbach, J.C.; Herber, W.K.; Sigal. I.S.; Darke, P.L.; Springer, J.P. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature*, **1989**, *337*, 615-620.

93. Hou, T.; Yu, R. Molecular dynamics and free energy studies on the wild-type and double mutant HIV-1 protease complexed with Amprenavir and two Amprenavir-related Inhibitors: mechanism for binding and drug resistance. *J. Med. Chem.*, **2007**, *50*, 1177-1188.

94. Dunn, B.M. Structure and mechanism of the pepsin-like family of aspartic peptidases. *Chem. Rev.*, **2002**, *102*, 4431-4458.

95. Wolfenden, R. Conformational aspects of inhibitor design: enzyme-substrate interactions in the transition state. *Bioorg. Med. Chem.*, **1999**, *7*, 647-652.

96. Nicholson, L.K.; Yamazaki, T.; Torchia, D.A.; Grzesiek, S.; Bax, A.; Stahl, S.J.; Kaufman, J.D.; Wingfield, P.T.; Lam, P.Y.; Jadhav, P.K. Flexibility and function in HIV-1 protease. *Nat. Struct. Biol.*, **1995**, *2*, 274-280.

97. Jeyabalan, M.P.; Nalivaika, E.; Schiffer, C.A. Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes. *Structure*, **2002**, *10*, 369-381.

98. Kumar, M.; Prashar, V.; Mahale, S.; Hosur, M.V. Observation of a tetrahedral reaction intermediate in the HIV-1 protease-substrate complex. *Biochem. J.*, **2005**, *389*, 365-371.

99.   Kohl, N.E.; Emini, E.A.; Schleif, W.A.; Davis, L.J.; Heimbach, J.C.; Dixon, R.A.; Scolnick, E.M.; Sigal, I.S. Active human immunodeficiency virus protease is required for viral infectivity. *Proc. Natl. Acad. Sci. U.S.A.*, **1988**, *85*, 4686-4690.

100.   Babe, L.M.; Rose, J.; Craik, C.S. Trans-dominant inhibitory human immunodeficiency virus type 1 protease monomers prevent protease activation and virion maturation. *Proc. Natl. Acad. Sci. U.S.A.*, **1995**, *92*, 10069-10073.

101.   Clavel, F.; Hance, A.J. Medical progress: HIV drug resistance. *New Engl. J. Med.*, **2004**, *350*, 1023-1035.

102.   D'Aquila, R.T.; Schapiro, J.M.; Brun-Vezinet, F.; Clotet, B.; Conway, B.; Demeter, L.M.; Grant, R.M.; Johnson, V.A.; Kuritzkes, D.R.; Loveday, C.; Shafer, R.W.; Richman, D.D. Drug Resistance Mutations in HIV-1. *Topics HIV Med.*, **2002**, *10*, 21-25.

103.   Ohtaka, H.; Schon, A.; Freire, E. Multidrug resistance to HIV-1 protease inhibition requires cooperative coupling between distal mutations. *Biochemistry*, **2003**, *42*, 13659-13666.

104.   Klabe, R.M.; Bacheler, L.T.; Ala, P.J.; Erickson-Viitanen, S.; Meek, J.L. Resistance to HIV protease inhibitors: a comparison of enzyme inhibition and antiviral potency. *Biochemistry*, **1998**, *37*, 8735-8742.

105.   Sayer, J.M.; Liu, F.; Ishima, R.; Weber, I.T.; Louis, J.M. Effect of the active-site D25N mutation on the structure, stability and ligand binding of the mature HIV-1 protease. *J. Biol. Chem.*, **2008**, in press.

106.   Louis, J.M.; Ishima, R.; Torchia, D.A.; Weber, I.T. HIV-1 protease: structure, dynamics, and inhibition. *Adv. Pharmacol.*, **2007**, *55*, 261-298.

107.   Strisovsky, K.; Tessmer, U.; Langner, J.; Konvalinka, J.; Krausslich, H.G. Systematic mutational analysis of the active-site threonine of HIV-1 proteinase: rethinking the "fireman's grip" hypothesis. *Protein Sci.*, **2000**, *9*, 1631-1641.

108.   Melnick, L.; Yang, S.; Rossi, R.; Zepp, C.; Heefner, D. An *Escherichia coli* expression assay and screen for Human Immunodeficiency Virus protease variants with decreased susceptibility to Indinavir. *Antimicrob. Agents Chemother.*, **1998**, *42*, 3256-3265.

109.   Partaledis, J.A.; Yamaguchi, K.; Tisdale, M.; Blair, E.E.; Falcione, C.; Maschera, B.; Myers, R.E.; Pazhanisamy, S.; Futer, O.; Cullinan, A.B.; Stuver, C.M.; Byrn, R.A.; Livingston, D.J. *In vitro* selection and characterization of human immunodeficiency virus type 1 (HIV-1) isolates with reduced sensitivity to hydroxyethylamino sulfonamide inhibitors of HIV-1 aspartyl protease. *J. Virol.*, **1995**, *69*, 5228-5235.

110.   Maschera,B.; Furfine, E.; Blair, E.D. Analysis of resistance to Human Immunodeficiency Virus type 1 protease inhibitors by using matched bacterial expression and proviral infection vectors. *J. Virol.*, **1995**, *69*, 5431-5436.

111.   Earl, P.L.; Moss, B.; Doms, R.W. Folding, interaction with GRP78-BiP, assembly and transport of the human immunodeficiency virus type 1 envelope protein. *J. Virol.*, **1991**, *65*, 2047-2055.

112.   Cann, A.J. Principles of Molecular Virology. **1997**, *Academic Press*, 2nd Edition.

113. Freed, E.O.; Myers, D.J.; Risser, R. Mutational analysis of the cleavage sequence of the human immunodeficiency virus type 1 envelope glycoprotein precursor gp160. *J. Virol.*, **1989**, *63*, 4670-4675.

114. Dorfman, T.; Mammano, F.; Haseltine, W.A.; Göttlinger, H.G. Role of the matrix protein in the virion association of the human immunodeficiency virus type 1 envelope glycoprotein. *J. Virol.*, **1994**, *68*, 1689-1696.

115. Lodge, R.; Gottlinger, H.; Gabuzda, D.; Cohen, E.A.; Lemay, G. The intracytoplasmic domain of gp41 mediates polarized budding of human immunodeficiency virus type 1 in MDCK cells. *J. Virol.*, **1994**, *68*, 4857-4861.

116. Cosson, P. Direct interaction between the envelope and matrix proteins of HIV-1. *EMBO J.*, **1996**, *15*, 5783-5788.

117. Murakami, T.; Freed, E.O. Genetic evidence for an interaction between human immunodeficiency virus type 1 matrix and α-helix 2 of the gp41 cytoplasmic tail. *J. Virol.*, **2000**, *74*, 3548-3554.

118. White, J.M. Membrane fusion. *Science*, **1992**, *258*, 917-924.

119. Freed, E.O.; Myers, D.J.; Risser, R. Characterization of the fusion domain of the human immunodeficiency virus type 1 envelope glycoprotein gp41. *Proc. Natl. Acad. Sci. U.S.A.*, **1990**, *87*, 4650-4654.

120. Cao, J.; Bergeron, L.; Helseth, E.; Thali, M.; Repke, H.; Sodroski, J. Effects of amino acid changes in the extracellular domain of the human immunodeficiency virus type 1 gp41 envelope glycoprotein. *J. Virol.*, **1993**, *67*, 2747-2755.

121. Neurath, A.R.; Strick, N.; Jiang, S.; Li, Y.Y.; Debnath, A.K. Anti-HIV-1 activity of cellulose acetate phthalate: synergy with soluble CD4 and induction of "dead-end" gp41 six-helix bundles. *BMC Infect Dis.*, **2002**, *2*, 1-6.

122. Zwick, M.B.; Saphire, E.O.; Burton, D.R. gp41: HIV's shy protein. *Nature Medicine*, **2004**, *10*, 133-134.

123. Furuta, R.A.; Wild, C.T.; Weng, Y.; Weiss, C.D. Capture of an early fusion-active conformation of HIV-1 gp41. *Nat. Struct. Biol.*, **1998**, *5*, 276-279.

124. Freed, E.O.; Myers, D.J.; Risser, R. Identification of the principal neutralizing determinant of human immunodeficiency virus type 1 as a fusion domain. *J. Virol.*, **1991**, *65*, 190-194.

125. Sullivan, N.; Thali, M.; Furman, C. et al. Effect of amino acid changes in the V1/V2 region of the human immunodeficiency virus type 1 gp120 glycoprotein on subunit association, syncytium formation and recognition by a neutralizing antibody. *J. Virol.*, **1993**, *67*, 3674-3679.

126. Sanders, R.W.; Busser, E.; Moore, J.P.; Lu, M.; Berkhout, B. Evolutionary repair of HIV type 1 gp41 with a kink in the N-terminal helix leads to restoration of the six-helix bundle structure. *AIDS Res. Hum. Retroviruses*, **2004**, *20*, 742-749.

127. Bedinger, P.; Moriarty, A.; von Borstel, R.C. 2nd; Donovan, N.J.; Steimer, K.S.; Littman, D.R. Internalization of the human immunodeficiency virus does not require the cytoplasmic domain of CD4. *Nature*, **1988**, *334*, 162-165.

128. Kwong, P.D.; Wyatt, R.; Robinson, J.; Sweet, R.W.; Sodroski, J.; Hendrickson, W.A. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*, **1998**, *393*, 648-659.

129. Hoxie, J.A.; Alpers, J.D.; Rackowski, J.L.; Huebner, K.; Haggarty, B.S.; Cedarbaum, A.J.; Reed, J.C. Alterations in T4 (CD4) protein and mRNA synthesis in cells infected with HIV. *Science*, **1986**, *234*, 1123-1127.

130. Stevenson, M.; Meier, C.; Mann, A.M.; Chapman, N.; Wasiak, A. Envelope glycoprotein of HIV induces interference and cytolysis resistance in CD4+ cells: mechanism for persistence in AIDS. *Cell*, **1988**, *53*, 483-496.

131. Feng, Y.; Broder, C.C.; Kennedy, P.E.; Berger, E.A. HIV-1 entry cofactor. Functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science*, **1996**, *272*, 872-877.

132. Shattock, R.J.; Moore, J.P. Inhibiting sexual transmission of HIV-1 infection *Nature Reviews Microbiology*, **2003**, *1*, 25-34.

133. De Clercq, E. New anti-HIV agents in preclinical or clinical development. *Front. Med. Chem.*, **2004**, *1*, 543-579.

134. Vermeire, K.; Bell, T.W.; Choi, H.J.; Jin, Q.; Samala, M.F.; Sodoma, A.; De Clercq, E.; Schols, D. The anti-HIV potency of cyclotriazadisulfonamide analogs is directly correlated with their ability to down-modulate the CD4 receptor. *Mol. Pharmacol.*, **2003**, *63*, 203-210.

135. Balzarini, J.; Hatse, S.; Vermeire, K.; Princen, K.; Aquaro, S.; Perno, C.F.; De Clercq, E.; Egberink, H.; Vanden Mooter, G.; Peumans, W.; Van Damme, E.; Schols, D. Mannose-specific plant lectins from the *Amaryllidaceae* family qualify as efficient microbicides for prevention of human immunodeficiency virus infection. *Antimicrob. Agents Chemother.*, **2004**, *48*, 3858-3870.

136. Boyd, M.R.; Gustafson, K.R.; McMahon, J.B.; Shoemaker, R.H.; O'Keefe, B.R.; Mori, T.; Gulakowski, R.J.; Wu, L.; Rivera, M.I.; Laurencot, C.M.; Currens, M.J.; Cardellina, J.H., 2nd; Buckheit, R. W., Jr.; Nara, P.L.; Pannell, L.K.; Sowder, R.C., 2nd; Henderson, L.E. Discovery of cyanovirin-N, a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: potential applications to microbicide development. *Antimicrob. Agents Chemother.*, **1997**, *41*, 1521-1530.

137. Wang, T.; Zhang, Z.; Wallace, O.B.; Deshpande, M.; Fang, H.; Yang, Z.; Zadjura, L.M.; Tweedie, D.L.; Huang, S.; Zhao, F.; Ranadive, S.; Robinson, B.S.; Gong, Y.F.; Ricarrdi, K.; Spicer, T.P.; Deminie, C.; Rose, R.; Wang, H.G. H.; Blair, W.S.; Shi, P.Y.; Lin, P.F.; Colonno, R.J.; Meanwell, N.A. Discovery of 4-benzoyl-1-[(4-methoxy-1H-pyrrolo[2,3-b] pyridin-3-yl)oxoacetyl]-2-(R)-methylpiperazine (BMS-378806): a novel HIV-1 attachment inhibitor that interferes with CD4-gp120 interactions. *J. Med. Chem.*, **2003**, *46*, 4236-4239.

138. Schols, D.; Claes, S.; Hatse, S.; Princen, K.; Vermeire, K.; De Clercq, E.; Skerlj, R.; Bridger, G.; Calandra, G. Anti-HIV activity profile of AMD070, an orally bioavailable CXCR4 antagonist. *Antiviral Res.*, **2003**, *57*, A39.

139. Baba, M.; Nishimura, O.; Kanzaki, N.; Okamoto, M.; Sawada, H.; Iizawa, Y.; Shiraishi, M.; Aramaki, Y.; Okonogi, K.; Ogawa, Y.; Meguro, K.; Fujino, M. A small-molecule, nonpeptide CCR5

antagonist with highly potent and selective anti-HIV-1 activity. *Proc. Natl. Acad. Sci. U.S.A.*, **1999**, *96*, 5698-5703.

140. Strizki, J.M.; Xu, S.; Wagner, N.E.; Wojcik, L.; Liu, J.; Hou, Y.; Endres, M.; Palani, A.; Shapiro, S.; Clader, J.W.; Greenlee, W.J.; Tagat, J.R.; McCombie, S.; Cox, K.; Fawzi, A.B.; Chou, C.C.; Pugliese-Sivo, C.; Davies, L.; Moreno, M.E.; Ho, D.D.; Trkola, A.; Stoddart, C.A.; Moore, J.P.; Reyes, G.R.; Baroudy, B.M. SCH-C (SCH 351125), an orally bioavailable, small molecule antagonist of the chemokine receptor CCR5, is a potent inhibitor of HIV-1 infection *in vitro* and *in vivo*. *Proc. Natl. Acad. Sci. U.S.A.*, **2001**, *98*, 12718-12723.

141. Dorr, P.; Macartney, M.; Rickett, G.; Smith-Burchnell, C.; Dobbs, S.; Mori, J.; Griffin, P.; Lok, J.; Irvine, R.; Westby, M.; Hitchcock, C.; Stammen, B.; Price, D.; Armour, D.; Wood, A.; Perros, M. UK-427,857, a novel small molecule HIV entry inhibitor is a specific antagonist of the chemokine receptor CCR5. Presented at the 10th Conference on Retroviruses and Opportunistic Infections, Boston, MA, February 10-14, 2003; Abstract 12, p 62.

142. Kilby, J.M.; Hopkins, S.; Venetta, S.M.; DiMassimo, B.; Cloud, G.A.; Lee, J.Y.; Alldredge, L.; Hunter, E.; Lambert, D.; Bolognesi, D.; Matthews, T.; Johnson, M.R.; Nowak, M.A.; Shaw, G.M.; Saag, M.S. Potent suppression of HIV-1 replication in humans by T-20, a peptide inhibitor of gp41-mediated virus entry. *Nat. Med.*, **1998**, *4*, 1302-1307.

143. Orner, B.P.; Ernst, J.T.; Hamilton, A.D. Toward proteomimetics: terphenyl derivatives as structural and functional mimics of extended regions of an R-helix. *J. Am. Chem. Soc.*, **2001**, *123*, 5382-5383.

144. Hoffmann, C.; Rockstroh, J.K.; Kamps, B.S. HIV medicine **2006**, *Flying Publisher*, Paris.

145. Lewis, W. Nucleoside reverse transcriptase inhibitors, mitochondrial DNA and AIDS therapy. *Antivir. Ther.*, **2005**, 10, Suppl 2, M13-27.

146. Rousseau, F.S.; Kahn, J.O.; Thompson, M.; Mildvan, D.; Shepp, D.; Sommadossi, J.P.; Delehanty, J.; Simpson, J.N.; Wang, L.H.; Quinn, J.B.; Wakeford, C.; van der Horst, C. Prototype trial design for rapid dose selection of antiretroviral drugs: an example using emtricitabine (Coviracil). *J. Antimicrob. Chemother.*, **2001**, *48*, 507-513.

147. Margot, N.A.; Isaacson, E.; McGowan, I.; Cheng, A.; Miller, M.D. Extended treatment with tenofovir disoproxil fumarate in treatment-experienced HIV-1-infected patients: genotypic, phenotypic, and rebound analyses. *J. Acquired Immune Defic. Syndr.*, **2003**, *33*, 15-21.

148. De Clercq, E. The role of non-nucleoside reverse transcriptase inhibitors (NNRTIs) in the therapy of HIV-1 infection. *Antiviral Res.*, **1998**, *38*, 153-179.

149. Balzarini, J.; Naesens, L.; Verbeken, E.; Laga, M.; Van Damme, L.; Parniak, M.; Van Mellaert, L.; Annè, J.; De Clercq, E. Preclinical studies on thiocarboxanilide UC-781 as a virucidal agent. *AIDS*, **1998**, *12*, 1129-1138.

150. Ren, J.; Nichols, C.; Bird, L.E.; Fujiwara, T.; Sugimoto, H.; Stuart, D.I.; Stammers, D.K. Binding of the second generation non-nucleoside inhibitor S-1153 to HIV-1 reverse trancriptase involves extensive main chain hydrogen bonding. *J. Biol. Chem.*, **2000**, *275*, 14316-14320.

151. Yoshinaga, T.; Sato, A.; Fujishita, T.; Fujiwara, T. S-1360: *in vitro* activity of a new HIV-1 integrase inhibitor in clinical development. Presented at the 9th Conference on Retroviruses and Opportunistic Infections, Seattle, WA; February 24-28, 2002; Abstract 8, p 55.

152. Pannecouque, C.; Pluymers, W.; Van Maele, B.; Tetz, V.; Cherepanov, P.; De Clercq, E.; Witvrouw, M.; Debyser, Z. New class of HIV integrase inhibitors that block viral replication in cell culture. *Curr. Biol.*, **2002**, *12*, 1169-1177.

153. Parolin, C.; Gatto, B.; Del Vecchio, C.; Pecere, T.; Tramontano, E.; Cecchetti, V.; Fravolini, A.; Masiero, S.; Palumbo, M.; Palu, G. New anti-human immunodeficiency virus type 1 6-aminoquinolones: mechanism of action. *Antimicrob. Agents Chemother.*, **2003**, *47*, 889-896.

154. Daelemans, D.; Afonia, E.; Nilsson, J.; Werner, G.; Kjems, J.; De Clercq, E.; Pavlakis, G.N.; Vandamme, A.M. A synthetic HIV-1 Rev inhibitor interfering with the CRM1-mediated nuclear export. *Proc. Natl. Acad. Sci. U.S.A.*, **2002**, *99*, 14440-14445.

155. De Clercq, E. New Approaches toward Anti-HIV Chemotherapy. *J. Med. Chem.*, **2005**, *48*, 1297-1313.

156. Rusconi, S; La Seta Catamancio, S. HIV-1 protease inhibitors in development. *Expert Opin Investig Drugs*, **2002**, *11*, 387-395.

157. Haas, D.W.; Zala, C.; Schrader, S.; Piliero, P.; Jaeger, H.; Nunes, D.; Thiry, A.; Schnittman, S.; Sension, M. Therapy with atazanavir plus saquinavir in patients failing highly active antiretroviral therapy: a randomized comparative pilot trial. *AIDS*, **2003**, *17*, 1339-1349.

158. de Bèthune, M.; Wigerinck, P.; Jonckheere, H.; Tahri, A.; Maes, L.; Pauwels, R.; Erickson, J. TMC 114, a highly potent protease inhibitor (PI) with an excellent profile against HIV variants highly resistant to current PIs. Presented at the 41st Interscience Conference on Antimicrobial Agents and Chemotherapy (ICAAC), Chicago, IL, December 16-19, 2001; Abstract F-1677.

159. Turner, S.R.; Strohbach, J.W.; Tommasi, R.A.; Aristoff, P.A.; Johnson, P.D.; Skulnick, H.I.; Dolak, L.A.; Seest, E.P.; Tomich, P.K.; Bohanon, M.J.; Horng, M.M.; Lynn, J.C.; Chong, K.T.; Hinshaw, R.R.; Watenpaugh, K.D.; Janakiraman, M.N.; Thaisrivongs, S. Tipranavir (PNU-140690): a potent, orally bioavaible nonpeptidic HIV protease inhibitor of the 5,6-dihydro-4-hydroxy-2-pyrone sulfonamide class. *J. Med. Chem.*, **1998**, *41*, 3467-3476.

160. Shaw-Reid, C.A.; Munshi, V.; Graham, P.; Wolfe, A.; Witmer, M.; Danzeisen, R.; Olsen, D. B.; Carroll, S.S.; Embrey, M.; Wai, J.S.; Miller, M.D.; Cole, J.L.; Hazuda, D.J. Inhibition of HIV-1 ribonuclease H by a novel diketo acid, 4-[5-benzoylamino)thien- 2-yl]-2,4-dioxobutanoic acid. *J. Biol. Chem.*, **2003**, *278*, 2777-2780.

161. Ho, D.D.; Neumann, A.U.; Perelson, A.S.; Chen, W.; Leonard, J.M.; Markowitz, M. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature*, **1995**, *373*, 123-126.

162. Wei, X.; Ghosh, S.K.; Taylor, M.E.; Johnson, V.A.; Emini, E.A.; Deutsch, P.; Lifson, J.D.; Bonhoeffer, S.; Nowak, M.A.; Hahn, B.H. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature*, **1995**, *373*, 117-122.

163. Richman, D.D. How drug resistance arises. *Sci Am.*, **1998**, *279*, 1-88.

164. Coffin, J.M. HIV population dynamics *in vivo*: implications for genetic variation, pathogenesis, and therapy. *Science*, **1995**, *267*, 483-489.

165. Boyer, P.L.; Currens, M.J.; McMahon, J.B.; Boyd, M.R.; Hughes, S.H. Analysis of nonnucleoside drug-resistant variants of human immunodeficiency virus type 1 reverse transcriptase. *J Virol.*, **1993**, *67*, 2412-2420.

166. Condra, J.H.; Schleif, W.A.; Blahy, O.M.; Gabryelski, L.J.; Graham, D.J.; Quintero, J.C.; Rhodes, A.; Robbins, H.L.; Roth, E.; Shivaprakash, M. *In vivo* emergence of HIV-1 variants resistant to multiple protease inhibitors. *Nature*, **1995**, *374*, 569-571.

167. Derdeyn, C.A.; Decker, J.M.; Sfakianos, J.N.; Zhang, Z.; O'Brien, W.A.; Ratner, L.; Shaw, G.M.; Hunter, E. Sensitivity of human immunodeficiency virus type 1 to fusion inhibitors targeted to the gp41 first heptad repeat involves distinct regions of gp41 and is consistently modulated by gp120 interactions with the coreceptor. *J. Virol.*, **2001**, *75*, 8605-8614.

168. http://www.rcsb.org

169. Goodford, P.J. A computational procedure for determining energetically favourable binding sites on biologically important macromolecules. *J. Med. Chem.*, **1985**, *28*, 849-857.

170. Ortuso, F.; Langer, T.; Alcaro, S. GBPM: GRID-based pharmacophore model: concept and application studies to protein-protein recognition. *Bioinformatics*, **2006**, *22*, 1449-1455.

171. Wallace, A. C.; Laskowski, R. A. and Thornton, J. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Prot. Eng.*, **1995**, *8*, 127-134.

172. Ceccherini-Silberstein, F.; Gago, F.; Santoro, M.; Gori, C.; Svicher, V.; Rodrìguez-Barrios, F.; d'Arrigo, R.; Ciccozzi, M.; Bertoli, A.; d'Arminio Monforte, A.; Balzarini, J.; Antinori, A.; Perno, C.F. High sequence conservation of Human Immunodeficiency Virus type 1 reverse transcriptase under drug pressure despite the continuous appearance of mutations. *J. Virol.*, **2005**, *79*, 10718-10729.

173. Mellors, J.W.; Bazmi, H.Z.; Schinazi, R.F.; Roy, B.M.; Hsiou, Y.; Arnold, E.; Weir, J.; Mayers, D.L. Novel mutations in reverse transcriptase of human immunodeficiency virus type 1 reduce susceptibility to foscarnet in laboratory and clinical isolates. *Antimicrob. Agents Chemother.*, **1995**, *39*, 1087-1092.

174. Sarafianos, G.S.; Das, K.; Tantillo, C.; Clark Jr, A.D.; Ding, J.; Whitcomb, J.M.; Boyer, P.L.; Hughes, H.S.; Arnold, E. Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO Journal*, **2001**, *20*, 1449-1461.

175. Mulky, A.; Vu, B.C.; Conway, J.A.; Hughes, S.H.; Kappes, J.C. Analysis of amino acids in the beta7-beta8 loop of human immunodeficiency virus type 1 reverse transcriptase for their role in virus replication. *J. Mol. Biol.*, **2007**, *365*, 1368-1378.

176. Divita, G.; Restle, T.; Goody, R.S.; Chermann, J.C.; Baillon, J.G. Inhibition of human immunodeficiency virus type 1 reverse transcriptase dimerization using synthetic peptides derived from the connection domain. *J. Biol. Chem.*, **1994**, *269*, 13080-13083.

177. Goel, R.; Beard, W.A.; Kumar, A.; Casas-Finet, J.R.; Strub, M.P.; Stahl, S.J.; Lewis, M.S.; Bebenek, K.; Becerra, S.P.; Kunkel, T.A. Structure/function studies of HIV-1(1) reverse transcriptase: dimerization-defective mutant L289K. *Biochemistry*, **1993**, *32*, 13012-1318.

178. Navarro, J.M.; Damier, L.; Boretto, J.; Priet, S.; Canard, B.; Quérat, G.; Sire, J. Glutamic residue 438 within the Protease-sensitive subdomain of HIV-1 Reverse Transcriptase is critical for heterodimer processing in viral particles. *Virology*, **2001**, *290*, 300-308.

179. Mizrahi, V.; Brooksbank, R.; Nkabinde, C. N. Mutagenesis of the Conserved Aspartic Acid 443, Glutamic Acid 478, Asparagine 494, and Aspartic Acid 443 Residues in the Ribonuclease H

Domain of p66/p51 Human Immunodeficiency Virus type 1 Reverse Transcriptase. *J. Biol. Chem.*, **1994**, *269*, 19245-19249.

180. Tachedjian, G.; Aronson, H.E.; De Los Santos, M.; Seehra, J.; McCoy, J.M.; Goff, S.P. Role of residues in the tryptophan repeat motif for HIV-1 reverse transcriptase dimerization. *J. Mol. Biol.*, **2003**, *326*, 381-396.

181. Kemp, S.D.; Shi, C.; Bloor, S.; Harrigan, R.P.; Mellors, J.W.; Larder, B.A. A novel polymorphism at codon 333 of human immunodeficiency virus type 1 reverse transcriptase can facilitate dual resistance to Zidovudine and L-29,39-dideoxy-39-thiacytidine. *J. Virol.*, **1998**, *72*, 5093-5098.

182. Hizi, A.; Shaharaban, M. Functional analysis of novel selective mutants of the reverse transcriptase of human immunodeficiency virus type 1. *J. Biol. Chem.*, **1992**, *267*, 18255-18258.

183. Yi-Brunozzi, H.Y.; Le Grice, S.F.J. Investigating HIV-1 polypurine tract geometry via targeted insertion of abasic lesions in the (-)-DNA template and (+)-RNA primer. *J. Biol. Chem.*, **2005**, *280*, 20154-20162.

184. Nikolenko, G.N.; Palmer, S.; Maldarelli, F.; Mellors, J.W.; Coffin, J.M.; Pathak, V.K. Mechanism for nucleoside analog-mediated abrogation of HIV-1 replication: balance between RNase H activity and nucleotide excision. *Proc. Natl. Acad. Sci. U.S.A.*, **2005**, *102*, 2093-2098.

185. Depollier, J.; Hourdou, M.L.; Aldrian-Herrada, G.; Rothwell, P.; Restle, T.; Divita, G. Insight into the mechanism of a peptide inhibitor of HIV reverse transcriptase dimerization. *Biochemistry*, **2005**, *44*, 1909-1918.

186. Pelemans, H.; Esnouf, R.; De Clercq, E.; Balzarini, J. Mutational analysis of trp-229 of human immunodeficiency virus type 1 reverse transcriptase (RT) identifies this amino acid residue as a prime target for the rational design of new non-nucleoside RT inhibitors. *Mol. Pharmacol.*, **2000**, *57*, 954-960.

187. Pelemans, H.; Esnouf, R.; Min, K.L.; Parniak, M.; De Clercq, E.; Balzarini, J. Mutations at amino acid positions 63, 189, and 396 of human immunodeficiency virus type 1 reverse transcriptase (RT) partially restore the DNA polymerase activity of a Trp229Tyr mutant RT. *Virology*, **2001**, *287*, 143-150.

188. Ding, J.; Hughes, S.H.; Arnold, E. Protein-nucleic acid interactions and DNA conformation in a complex of human immunodeficiency virus type 1 reverse transcriptase with a double-stranded DNA template-primer. *Biopolymers*, **1997**, *44*, 125-138.

189. Sharma, B.; Kaushik, N.; Singh, K.; Kumar, S.; Pandey, V.N. Substitution of conserved hydrophobic residues in motifs B and C of HIV-1 RT alters the geometry of its catalytic pocket. *Biochemistry*, **2002**, *41*, 15685-15697.

190. Klarmann, G.J.; Smith, R.A.; Schinazi, R.F.; North, T.W.; Preston, B.D. Site-specific incorporation of nucleoside analogs by HIV-1 Reverse Transcriptase and the template grip mutant P157S. Template interactions influence substrate recognition at the polymerase active site. *J. Biol. Chem.*, **2000**, *275*, 359-366.

191. Harris, D.; Yadav, P.N.; Pandey, V.N. Loss of polymerase activity due to Tyr to Phe substitution in the YMDD motif of human immunodeficiency virus type-1 reverse transcriptase is compensated by Met to Val substitution within the same motif. *Biochemistry*, **1998**, *37*, 9630-9640.

192. Ghosh, M.; Williams, J.; Powell, M.D.; Levin, G.J.; Le Grice, S.F.J. Mutating a conserved motif of the HIV-1 Reverse Transcriptase palm subdomain alters primer utilization. *Biochemistry*, **1997**, *36*, 5758-5768.

193. Beard, W.A.; Stahl, S.J.; Kim, H.R.; Bebenek, K.; Kumar, A.; Strub, M.P.; Becerra, S.P.; Kunkel, T.A.; Wilson, S.H. Structure/function studies of human immunodeficiency virus type 1 reverse transcriptase. Alanine scanning mutagenesis of an alpha-helix in the *thumb* subdomain. *J. Biol. Chem.*, **1994**, *269*, 28091-28097.

194. Sarafianos, S.G.; Clark Jr., A.D.; Tuske, S.; Squire, C.J.; Das, K.; Sheng, D.; Ilankumaran, P.; Ramesha, A.R.; Kroth, H.; Sayer, J.M.; Jerina, D.M.; Boyer, P.L.; Hughes, S.H.; Arnold, E. Trapping HIV-1 reverse transcriptase before and after translocation on DNA. *J. Biol. Chem.*, **2003**, *278*, 16280-16288.

195. Beard, W.A.; Bebenek, K.; Darden, T.A.; Li, L.; Prasad, R.; Kunkel, T.A.; Wilson, S.H. Vertical-scanning mutagenesis of a critical tryptophan in the minor groove binding track of HIV-1 reverse transcriptase. Molecular nature of polymerase-nucleic acid interactions. *J. Biol. Chem.*, **1998**, *273*, 30435-30442.

196. Fisher, T.S.; Darden, T.; Prasad, V.R. Mutations proximal to the minor groove-binding track of human immunodeficiency virus type 1 reverse transcriptase differentially affect utilization of RNA versus DNA as template. *J. Virol.*, **2003**, *77*, 5837-5845.

197. Sevilya, Z.; Loya, S.; Duvshani, A.; Adir, N.; Hizi, A. Mutagenesis of Cysteine 280 of the reverse transcriptase of human immunodeficiency virus type-1: the effects on the Ribonuclease H activity. *J. Mol. Biol.*, **2003**, *327*, 19-30.

198. Rausch, J.W.; Lener, D.; Miller, J.T.; Julias, J.G.; Hughes, S.H.; Le Grice, S.F.J. Altering the RNase H primer grip of human immunodeficiency virus reverse transcriptase modifies cleavage specificity. *Biochemistry*, **2002**, *41*, 4856-4865.

199. Kern, G.; Pelton, J.; Marqusee, S.; Kern, D. Structural properties of the histidine-containing loop in HIV-1 RNase H. *Biophysical Chemistry*, **2002**, *96*, 285-291.

200. Pata, D.J.; Stirtan, W.G.; Goldstein, S.W.; Steitz, T.A. Structure of HIV-1 reverse transcriptase bound to an inhibitor active against mutant reverse transcriptases resistant to other nonnucleoside inhibitors. *Proc. Natl. Acad. Sci. U.S.A.*, **2004**, *101*, 10548-10553.

201. Smerdon, S.J.; Jager, J.; Wang, J.; Kohlstaedt, L.A.; Chirino, A.J.; Friedman, J.M.; Rice, P.A.; Steitz, T.A. Structure of the binding site for non-nucleoside inhibitors of the reverse transcriptase of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. U.S.A.*, **1994**, *91*, 3911-3915.

202. Lowe, D.M.; Parmar, V.; Kemp, S.D.; Larder, B.A. Mutational analysis of two conserved sequence motifs in HIV-1 reverse transcriptase. *FEBS Lett.*, **1991**, *282*, 231-234.

203. Bakhanashvili, M.; Avidan, O.; Hizi, A. Mutational studies of human immunodeficiency virus type 1 reverse transcriptase: the involvement of residues 183 and 184 in the fidelity of DNA synthesis. *FEBS Lett.*, **1996**, *391*, 257-262.

204. Harris, D.; Kaushik, N.; Pandey, P.K.; Yadav, P.N.S.; Pandey, V.N. Functional analysis of amino acid residues constituting the dNTP binding pocket of HIV-1 Reverse Transcriptase. *J. Biol. Chem.*, **1998**, *273*, 33624-33634.

205. Boyer, P.L.; Ferris, A.L.; Clark, P.; Whitmer, J.; Frank, P.; Tantillo, C.; Arnold, E.; Hughes, S.H. Mutational analysis of the fingers and palm subdomains of Human Immunodeficiency Virus Type-1 (HIV-1) reverse transcriptase. *J. Mol. Biol.*, **1994**, *243*, 472-483.

206. Ren, J.; Nichols, C.E.; Stamp, A.; Chamberlain, P.P.; Ferris, R.; Weaver, K.L.; Short, S.A.; Stammers, D.K. Structural insights into mechanisms of non-nucleoside drug resistance for HIV-1 reverse transcriptase mutated at codons 101 or 138. *FEBS J.*, **2006**, *273*, 3850-3860.

207. Esnouf, R.M.; Ren, J.; Hopkins, A.L.; Ross, C.K.; Jones, E.Y.; Stammers, D.K.; Stuart, D.I. Unique features in the structure of the complex between HIV-1 reverse transcriptase and the bis (heteroaryl) piperazine (BHAP) U-90152 explain resistance mutations for this non-nucleoside inhibitor. *Proc. Natl. Acad. Sci. U.S.A.*, **1997**, *94*, 3984-3989.

208. Shen, L.; Shen, J.; Luo, X.; Cheng, F.; Xu, Y.; Chen, K.; Arnold, E.; Ding, J.; Jiang, H. Steered molecular dynamics simulation on the binding of NNRTI to HIV-1 RT. *Biophys. J.*, **2003**, *84*, 3547-3563.

209. Bacolla, A.; Ih, C.K.; Rose, J.M.; Pirasli, G.; Warren, T.; Grygon, C.A.; Ingraham, R.; Cousins, R.B.; Greenwood, D.; Richman, D.; Chengli, Y.C.; Griffin, J. Amino acid substitutions in HIV-1 reverse transcriptase with corresponding residues from HIV-2. Effect on kinetic constants and inhibition by non-nucleoside analogs. *J. Biol. Chem.*, **1993**, *268*, 16571-16577.

210. Paolucci, S.; Baldanti, F.; Campanini, G.; Cancio, R.; Belfiore, A.; Maga, G.; Gerna, G. NNRTI-selected mutations at codon 190 of human immunodeficiency virus type 1 reverse transcriptase decrease susceptibility to stavudine and zidovudine. *Antiviral Research*, **2007**, *76*, 99-103.

211. Huang,W.; Gamarnik, A.; Limoli, K.; Petropoulos, C.J.; Whitcomb, J.M. Amino acid substitutions at position 190 of human immunodeficiency virus type 1 reverse transcriptase increase susceptibility to delavirdine and impair virus replication. *J. Virol.*, **2003**, *77*, 1512-1523.

212. Pelemans, H.; Esnouf, R.M.; Parniak, M.A.; Vandamme, A.M.; De Clercq, E.; Balzarini, J. A proline-to-histidine substitution at position 225 of human immunodeficiency virus type 1 (HIV-1) reverse transcriptase (RT) sensitizes HIV-1 RT to BHAP U-90152. *J. Gen. Virol.*, **1998**, *79*, 1347-1352.

213. Palaniappan, C.; Wisniewski, M.; Jacques, P.S.; Le Grice, S.F. ; Fay, P.J.; Bambara, R.A. Mutations within the primer grip region of HIV-1 reverse transcriptase result in loss of RNase H function. *J. Biol. Chem.*, **1997**, *272*, 11157-11164.

214. Stevens, M. ; Pannecouque, C. ; De Clercq, E. ; Balzarini, J. Inhibition of human immunodeficiency virus by a new class of pyridine oxide derivatives. *Antimicrobial Agents and Chemotherapy*, **2003**, *47*, 2951-2957.

215. Pelemans, H.; Esnouf, R.M.; Jonckheere, H.; De Clercq, E.; Balzarini, J. Mutational analysis of Tyr-318 within the non-nucleoside reverse transcriptase inhibitor binding pocket of human immunodeficiency virus type I reverse transcriptase. *J. Biol. Chem.*, **1998**, *273*, 34234-34239.

216. Jonckheere, H.; Taymans, J.M.; Balzarini, J.; Velazquez, S.; Camaras, M.J.; Desmyter, J.; De Clercq, E.; Annè, J. Resistance of HIV-1 reverse transcriptase against [2',5'-bis-O-(tert-butyldimethylsilyl)-3'-spiro-5"-(4"-amino-1",2"-oxathiole-2",2"-dioxide)](TSAO) derivatives is

determined by the mutation Glu138-->Lys on the p51 subunit. *J. Biol. Chem.*, **1994**, *269*, 25255-25258.

217.   Rodríguez-Barrios, F.; Pérez, C.; Lobatón, E.; Velázquez, S.; Chamorro, C.; San-Félix, A.; Pérez-Pérez, M.J.; Camarasa, M.J.; Pelemans, H.; Balzarini, J.; Gago, F. Identification of a putative binding site for [2',5'-bis-O-(tert-butyldimethylsilyl)-beta-D-ribofuranosyl]-3'-spiro-5"-(4"-amino-1",2"-oxathiole-2",2"-dioxide)thymine (TSAO) derivatives at the p51-p66 interface of HIV-1 reverse transcriptase. *J. Med. Chem.*, **2001**, *44*, 1853-1865.

218.   Velázquez, S.; Tuñón, V.; Jimeno, M.L.; Chamorro, C.; De Clercq, E.; Balzarini, J.; Camarasa, M.J. Potential multifunctional inhibitors of HIV-1 reverse transcriptase. Novel [AZT]-[TSAO-T] and [d4T]-[TSAO-T] heterodimers modified in the linker and in the dideoxynucleoside region. *J. Med. Chem.*, **1999**, *42*, 5188-5196.

219.   Velázquez, S.; Lobatòn, E.; De Clercq, E.; Koontz, D.L.; Mellors, J.W.; Balzarini, J.; Camarasa, M.J. Hybrids of [TSAO-T]- [Foscarnet]: the first conjugate of foscarnet with a non-nucleoside reverse transcriptase inhibitor through a labile covalent ester bond. *J. Med. Chem.*, **2004**, *47*, 3418-3426.

220.   Geitmann, M.; Unge, T.; Danielson, U.H. Interaction kinetic characterization of HIV-1 reverse transcriptase non-nucleoside inhibitor resistance. *J. Med. Chem.*, **2006**, *49*, 2375-2387.

221.   Berendsen, H.; Van der Spoel, D.; Van Drunen, R. GROMACS: a message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.*, **1995**, *91*, 43-56.

222.   Ceccherini-Silberstein, F.; Svicher, V.; Sing, T.; Artese, A.; Santoro, M.M.; Forbici, F.; Bertoli, A.; Alcaro, S.; Palamara, G.; d'Arminio Monforte, A.; Balzarini, J.; Antinori, A.; Lengauer, T.; Perno, C.F. Characterization and structural analysis of novel mutations in human immunodeficiency virus type 1 reverse transcriptase involved in the regulation of resistance to nonnucleoside inhibitors. *J. Virol.*, **2007**, *81*, 11507-11519.

223.   Svicher, V.; Ceccherini-Silberstein, F.; Erba, F.; Santoro, M.; Gori, C.; Bellocchi, M.C.; Giannella, S.; Trotta, M.P.; Monforte, A.; Antinori, A.; Perno, C.F. Novel human immunodeficiency virus type 1 protease mutations potentially involved in resistance to protease inhibitors. *Antimicrob. Agents Chemother.*, **2005**, *49*, 2015-2025.

224.   McDonald, D.Q.; Still, W.C. AMBER torsional parameters for the peptide backbone. *Tetrahedron Lett.*, **1992**, *33*, 7743-7746.

225.   Still, W.C.; Tempczyk, A.; Hawley, R.C.; Hendrickson, T. Semianalytical treatment of solvation for molecular mechanism and dynamics. *J. Am. Chem. Soc.*, **1990**, *112*, 6127-6129.

226.   Mohamadi, F.; Richards, N.G.J.; Guida, W.C. MacroModel: an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J. Comput. Chem.*, **1990**, *11*, 440-467.

227.   Alcaro, S.; Gasparrini, F.; Incani, O.; Mecucci, S.; Misiti, D.; Pierini, M.; Villani, C. A "quasi-flexible" automatic docking processing for studying stereoselective recognition mechanisms. Part I. Protocol validation. *J. Comput. Chem.*, **2000**, *21*, 515-530.

228.   Weinheimer, S.; Discotto, L.; Friborg, J.; Yang, H.; Colonno, R. Atazanavir signature I50L resistance substitution accounts for unique phenotype of increased susceptibility to other protease

inhibitors in a variety of human immunodeficiency virus type 1 genetic backbones. *Antimicrob. Agents Chemother.*, **2005**, *49*, 3816-3824.

229. Stoll, V.; Qin, W.; Stewart, K.D.; Jakob, C.; Park, C.; Walter, K.; Simmer, R.L.; Helfrich, R.; Bussiere, D.; Kao, J.; Kempf, D.; Sham, H.L.; Norbeck, D.W. X-ray crystallographic structure of ABT-378 (lopinavir) bound to HIV-1 protease. *Bioorg. Med. Chem.*, **2002**, *10*, 2803-2806.

230. Armstrong, A.A.; Muzammil, S.; Jakalian, A.; Bonneau, P.R.; Schmelmer, V.; Freire, E.; Amzel, L.M. Unique Thermodynamic Response of Tipranavir to Human Immunodeficiency Virus Type 1 Protease Drug Resistance Mutations. *J. Virol.*, **2007**, *81*, 5144-5154.

231. Weber, I.T.; Cavanaugh, D.S.; Harrison, R.W. Models of HIV-1 protease with peptides representing its natural substrates. *J. Mol. Struct.*, **1998**, *423*, 1-12.

232. Beerenwinkel, N.; Sing, T.; Lengauer, T.; Rahnenführer, J.; Roomp, K.; Savenkov, I.; Fischer, R.; Hoffmann, D.; Selbig, J.; Korn, K.; Walter, H.; Berg, T.; Braun, P.; Fätkenheuer, G.; Oette, M.; Rockstroh, J.; Kupfer, B.; Kaiser, R.; Däumer, M. Computational methods for the design of effective therapies against drug resistant HIV strains. *Bioinformatics*, **2005**, *21*, 3943-3950.

233. Lyne, P.D.; Lamb, M.L.; Saeh, J.C. Accurate Prediction of the Relative Potencies of Members of a Series of Kinase Inhibitors Using Molecular Docking and MM-GBSA Scoring. *J . Med. Chem.*, **2006**, *49*, 4805-4808.

234. Mink, M.; Greenberg, M.L.; Mosier, S. Impact of HIV-1 gp41 amino acid substitutions (position 36–45) on susceptibility to T20 (Enfuvirtide) *in vitro*; analysis of primary virus isolates recovered from patients during chronic Enfuvirtide treatment and site-directed mutants in NL4-3. *Antivir. Ther.*, **2002**, *7*, 17-18.

235. Chan, D.C.; Fass, D.; Berger, J.M.; Kim, P.S. Core structure of gp41 from the HIV envelope glycoprotein. *Cell*, **1997**, *89*, 263-273.

236. Svicher, V.; Aquaro, S.; D'Arrigo, R.; Artese, A.; Dimonte, S.; Alcaro, S.; Santoro, M.M.; Di Perri, G.; Lo Caputo, S.; Bellagamba, R.; Zaccarelli, M.; Visco-Comandini, U.; Antinori, A.; Narciso, P.; Ceccherini-Silberstein, F.; Perno, C.F. Specific Enfuvirtide associated mutational pathways in HIV-1 gp41 are significantly correlated with a gain of CD4 cell count despite virological failure. *J. Infect. Dis.*, **2007**, accepted.