THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (Ph.D.)

Kinetic analysis of retroviral proteases

by

HELGA EIZERT (ROSMANN)

Thesis supervisor: Prof. Dr. József Tőzsér

UNIVERSITY OF DEBRECEN
DOCTORAL SCHOOL OF MOLECULAR CELL AND IMMUNE BIOLOGY

DEBRECEN, 2013

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AIDS	acquired immunodeficiency syndrome
AMV	avian myeloblastosis virus
ASLV	avian sarcoma-leukosis virus
BFV	bovine foamy virus
BLV	bovine leukemia virus
CA*	capsid protein
$D_{1/2}$	50% loss in enzymatic activity
EDTA	ethylene-diamino-tetraacetic acid
EFV	equine foamy virus
EIAV	equine infectious anemia virus
env	gene that encodes envelope protein of retroviruses
FeLV	feline leukemia virus
FFV	feline foamy virus
FIV	feline immunodeficiency virus
FVs	foamy viruses
gag	gene that encodes structural proteins of retroviruses
GALV	gibbon-ape leukemia virus
HAART	highly active antiretroviral therapy
HERV-K	human endogenous retrovirus type K
HFV	human foamy virus
HIV	human immunodeficiency virus
HPLC	high-performance/pressure liquid chromatography
HTLV	human T-lymphotropic virus
ICTV	International Committee on Taxonomy of Viruses
IN*	integrase
IPTG	isopropyl-D-thio-galactopyranoside
MA*	matrix protein
MAV	myeloblastosis associated virus
MBP	maltose binding protein
MMLV	Moloney murine leukemia virus

MMTV	mouse mammary tumor virus
MPMV	Mason-Pfizer monkey virus
NC*	nucleocapsid protein
PERV	porcine endogenous retrovirus
PI	protease inhibitor
PR*	wild-type protease
PR_{Q8R}	PR with Q8R mutation
$PR_{Q8R /T28D}$	PR with Q8R and T28D double mutations
pol	gene that encodes enzymes of retroviruses
PFV	prototype foamy virus
RH	RNase H
RNA	ribonucleic acid
RP	reversed-phase
RSV	Rous sarcoma virus
RT*	reverse transcriptase
RV	retrovirus
SFV	simian foamy virus
SIV	simian immunodeficiency virus
SP211	Val-Ser-Gln-Asn-Tyr-Pro-Ile-Val-Gln oligopeptide substrate
SU*	surface protein
TM*	transmembrane protein
TF	transframe protein
TFA	trifluoroacetic acid
Tris	2-amino-2-hydroximetil-1,3 -propane
WDSV	walleye dermal sarcoma virus
WHO	World Health Organization
Xaa	amino acid residue in general

---

* Nomenclature of viral proteins is that of Leis et al. (1988). Other proteins are indicated as p followed by the molecular weight of that protein in kDa (p1, p2, p6).
Nomenclature of substrate residues and enzyme subsites is based on Schechter and Berger (1967). The amino acid residues occupy adjacent subsites, those towards the NH_2-end occupying subsites S_1, S_2, etc., those towards the COOH-end occupying subsites S'_1, S'_2, etc. The positions of the residues in the peptide based to the subsites they occupy are P_1, P_2, etc., respectively P'_1, P'_2, etc.

1. INTRODUCTION

Retroviruses were originally studied early in the 20th century because it was discovered that certain avian retroviruses could induce neoplastic disease very efficiently in birds. The intense study of retroviruses has taken on since the discovery of pathogenic human retroviruses including human immunodeficiency virus 1 (HIV-1), HIV-2, and human T-lymphotrophic virus 1 (HTLV-1). HIV-1 protease (PR) is essential for viral replication and has proved to be an effective target for antiviral drugs to treat acquired immunodeficiency syndrome (AIDS). However, the long term effectiveness of current protease inhibitors (PIs) as therapeutic agents is limited by the rapid development of drug-resistant variants of the protease. Residues that confer inhibitor resistance to HIV-1 PR are frequently seen in equivalent position of other retroviral PRs. Therefore, understanding the specificity differences of PRs may help the design of inhibitors effective against the mutant HIV-1 PR forms appearing in resistance. Previously, the substrate specificity of HIV-1, HIV-2, equine infectious anemia virus (EIAV), Moloney murine leukemia virus (MMLV) and avian myeloblastosis virus (AMV) proteases were studied on P4-P3' (Val-Ser-Gln-Asn-Tyr↓Pro-Ile-Val-Gln) substituted peptides (the arrow indicate the size of cleavage). Comparative study of the substrate specificity of additional retroviral proteases (mouse mammary tumor virus (MMTV), Mason-Pfizer monkey virus (MMPV), HTLV-1, bovine leukemia virus (BLV), human foamy virus (HFV), walleye dermal sarcoma virus (WDSV)) was expected to facilitate our understanding of the fundamental interactions between HIV-1 MA/CA type 1 cleavage site substrate and different retroviral PRs. The characterization of the specificity and structure of the proteases was aimed to complement the studies of retrovirus family. Comparison with each other expanded these studies in a novel direction in order to fully understand the common characteristics of retroviral protease specificity as well as their differences. Molecular modeling was used to help the interpretation of the observed specificity changes.

Spumaviruses are endemic in most non-human primates, comprise a special type of retrovirus that has developed a replication strategy combining features of both retroviruses and hepadnaviruses. A unique feature of foamy viruses (FVs) includes an apparent apathogenicity in natural hosts as well as zoonozically infected humans. It appears that some inherent properties of FV vectors set them favorably apart from orthoretroviral vectors and ask for additional basic research on the viruses as well as on the application in gene therapy (Lindemann and Rethwilm,

2011). Retroviral proteases are only active in a dimeric state, which raises the question of how FV PR activity is regulated. Recombinant prototype foamy virus (PFV) and simian foamy virus (SFV) PR-RT domains are predominantly monomeric in solution, but appear to have some proteolytic activity that is enhanced by high salt conditions (Hartl et al., 2008, 2010a and 2010b). One publication suggests that recombinant FV PR is able to form transient dimers and therefore escapes detection by traditional methods (Hartl et al., 2010b). Others proposed that PR dimerization and activity might be regulated at the precursor protein level by the IN (integrase) oligomerization domains, favoring protein-protein interactions (Lee et al., 2011). However, strong evidence has been provided in favor of the former by identifying a nucleic acid motif (PARM for protease-activating RNA motif) that is bound by the Pol precursor protein and regulates PR activity (Hartl et al., 2011). The foamy virus PR is essential for viral infectivity, since mutation of the active site Asp residues resulted in non-infectious virions (Konvalinka et al., 1995b), as previously found for HIV-1 PR (Kohl et al., 1988). Previously, HFV PR was cloned in fusion with maltose binding protein (MBP) and characterized the fusion protein (Fenyőfalvi et al., 1999). Later the purification protocol was improved (Boross et al., 2006), and the modifications allowed us to prepare and characterize purified, processed and active HFV PR. Comparison of the processed and fusion forms of the wild-type and mutant (S25T) PRs suggested that the fusion forms can be used instead of the processed enzymes for comparative studies (Boross et al., 2006). The obtained catalytic constants for HFV PR were much lower than those were previously determined for various other retroviral proteases coded on *pol* genes (Tőzsér et al., 1991; Tőzsér et al., 1993; Louis et al., 1999), but similar to those obtained previously with *gag*-encoded avian retrovirus PR (Tőzsér et al., 1996). The pH optimum of HFV PR (Fenyőfalvi et al., 1999) was much higher than those values published for HIV-1 PR (Szeltner and Polgár, 1996b; Hyland et al., 1991; Darke et al., 1989), depending on the used substrate, ionic strength and other experimental conditions. Furthermore, the dimer stability of HFV PR (Fenyőfalvi et al., 1999) was much lower (half-maximal enzyme activity was obtained at 0.75 M as compared with 1.85 M denaturant concentration) than that of HIV-1 PR (Wondrak et al., 1996). We have introduced some mutations close to the catalytic aspartates for exploring the role of certain residues for these unusual features. We were interested in determining whether, similar to orthoretroviruses, the mutated amino acids (present in HIV-1 PR sequence at the same positions) will improve the pH optimum and stability of spumavirus proteases. Understanding the role of these

residues may help in understanding the unusual features of the foamy PR and maybe the different replication cycle and biology of foamy viruses.

2. THEORETICAL BACKGROUND

2.1 Retroviruses in general. Classification of retroviruses

Retroviruses are enveloped, positive-strand RNA viruses with a unique morphology and means of replication. These viruses are distinguished by their use of viral-encoded reverse transcriptase to copy virion genomic RNA into proviral DNA, which is integrated into the infected host cell genome. Thus, there is a reverse flow of information from RNA to DNA during the virus life cycle (hence “retro” for backwards). Retroviral virions measure approximately 80-100 nm in diameter. They are surrounded by a lipid envelope derived from the infected cell upon budding of the virus. Virion morphology, as ascertained via electron microscopy, has been used for classification of different retroviruses. The genome within the core consists of two, usually identical, single-stranded positive strand RNA molecules. The size of the genome ranges from 7 kb for avian leucosis viruses to 12 kb for human foamy viruses.

The retroviruses can be classified by the disease they cause, tissue tropism and host range, virion morphology, and genetic complexity (Coffin 1992; Coffin 1996). The *oncoviruses* include retroviruses that can easily immortalize or transform target cells. The *lentiviruses* are viruses associated with neurological and immunosuppressive diseases. *Spumaviruses*, represented by the human foamy virus (HFV), cause a distinct cytopathological effect but do not seem to cause clinical diseases. Retroviruses are further subdivided into seven groups defined by evolutionary relatedness. Their classification is based on the latest report of the International Committee on Taxonomy of Viruses (ICTV), as detailed at the ICTV database (Fig. 1).

All retroviruses contain three major coding domains with information for virion proteins: *gag*, which directs the synthesis of internal virion proteins that form the matrix (MA), the capsid (CA), and the nucleoprotein (NC) structures; *pol*, which contains the information for the reverse transcriptase (RT) and integrase (IN) enzymes; and *env*, from which are derived the surface (SU) and transmembrane (TM) components of the viral envelope protein. An additional, smaller, coding domain present in all retroviruses is *pro*, which encodes the virion protease (PR).

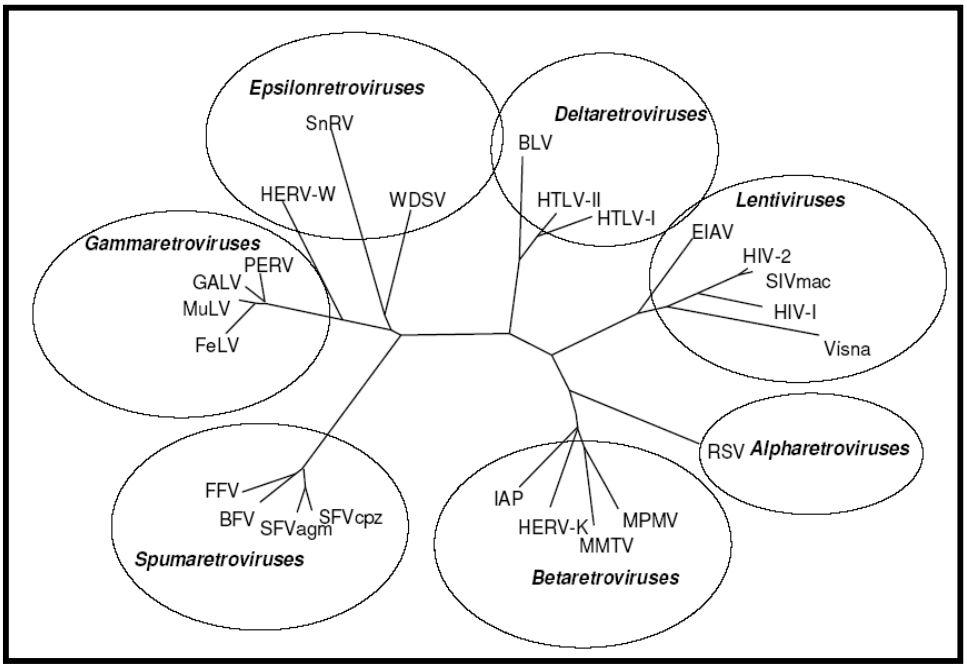


Figure 1: Retrovirus classification of Family *Retroviridae*

i. **Subfamily *Orthoretrovirinae*** (6 genere)
Genus Alpharetrovirus (9 Species), e.g. Rous sarcoma virus (RSV)
Genus Betaretrovirus (5 Species), e.g. MMTV
Genus Deltaretrovirus (4 Species), e.g. BLV
Genus Epsilonretrovirus (3 Species), e.g. WDSV
Genus Gammaretrovirus (17 Species), e.g. MMLV
Genus Lentivirus (9 Species), e.g. HIV-1
ii. **Subfamily *Spumaretrovirinae*** (1 genus)
Genus Spumaretrovirus, e.g. HFV
In the examples the typical species are indicated in each genus.
(Based on: www.wikipedia.org)

Alpharetroviruses, betaretroviruses and gammaretroviruses are genetically simple, encoding only MA, CA, NC, PR, RT, IN, and Env proteins. Deltaretroviruses, epsilonretroviruses, lentiviruses and spumaviruses are considered complex because they encode a number of ancillary proteins in addition to the proteins listed above, that often play an important role in gene regulation (Coffin 1992) (Fig.2). The single-stranded RNA genome found in virions is shorter than the viral double-stranded DNA genome. The sequence that is lost during RNA transcription is replaced during reverse transcription.The HIV reverse transcriptase also has ribonuclease activity (RNase H) that degrades the viral RNA during the synthesis of cDNA, allowing completion of the newly synthesized DNA.

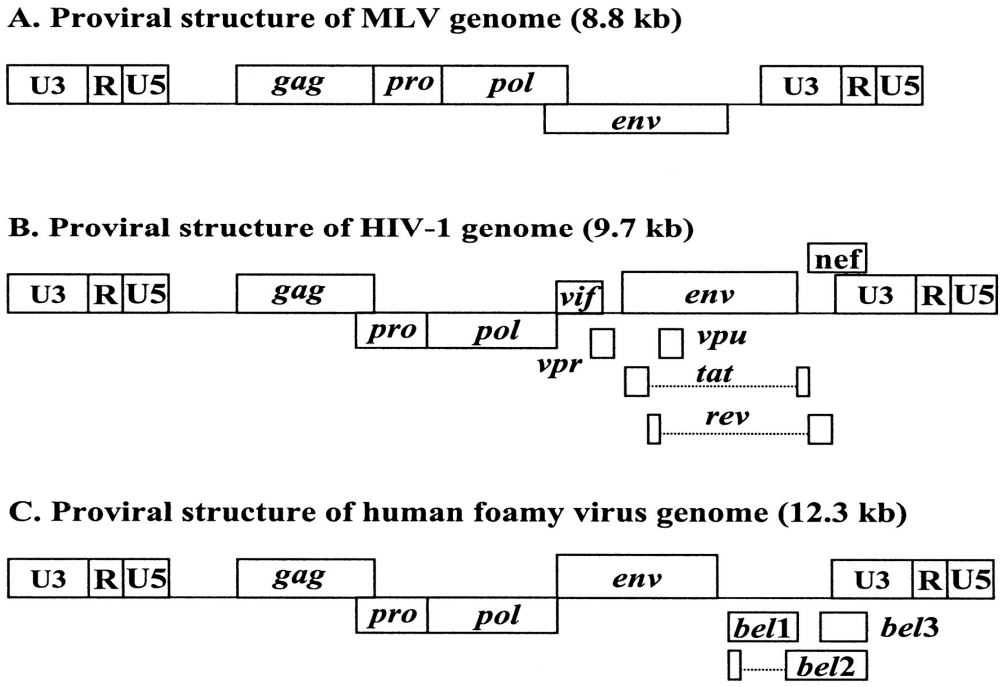


Figure 2: Genomic organization of the gammaretroviral MLV, lentiviral HIV-1 and spumaretroviral HFV

R: repeat sequence, U3: promoter/enhancer, U5: reverse transcription/integration. There are additional regulatory nonvirion proteins involved in the replication: vif, nef, vpr, vpu, tat, rev. They are specific additional proteins for HIV. Vif: viral infectivity factor, nef: negative regulatory factor, vpr: viral protein r, plays a role in regulating nuclear import of the HIV-1 pre-integration complex, vpu: viral protein u, is involved in viral budding enhancing virion release from the cell, tat: trans-activator of transcription, a regulatory protein that drastically enhances the efficiency of viral transcription, rev: regulator of virion expression. For HFV we have three bel open reading frames. These are: bel1, bel2 and bel3. Bel-1 (Tas) encodes a transcriptional trans-activator protein of the LTR promoter. Bel-2 (bet protein) is a viral antagonist for the host cells' innate defense system, by neutralization of cellular APOBEC3 protein functions. Bel-3 does not have a defined role in the viral life cycle.

2.2. Retroviruses: biological importance

One common feature of six of the seven genera of retroviruses is that most viral isolates are pathogenic either in the natural host or in accidentally infected hosts. In some groups, including viruses such as human T cell lymphotropic virus (HTLV, a deltavirus), viral gene expression and replication are very poor in both tissue cultures and infected humans. However, even viruses that replicate very poorly, such as HTLV, can be pathogenic. HTLV causes highly malignant T-cell lymphomas in a small proportion of infected humans (Cann et al., 1996). In other cases, such as the lentiviruses (e.g. HIV), infected cells can be killed rapidly. Lentiviruses could replicate at high levels without pathogenic consequences in their natural hosts, whereas in other experimentally or zoonotically infected hosts some level of replication could be lethal. By contrast, viruses such as

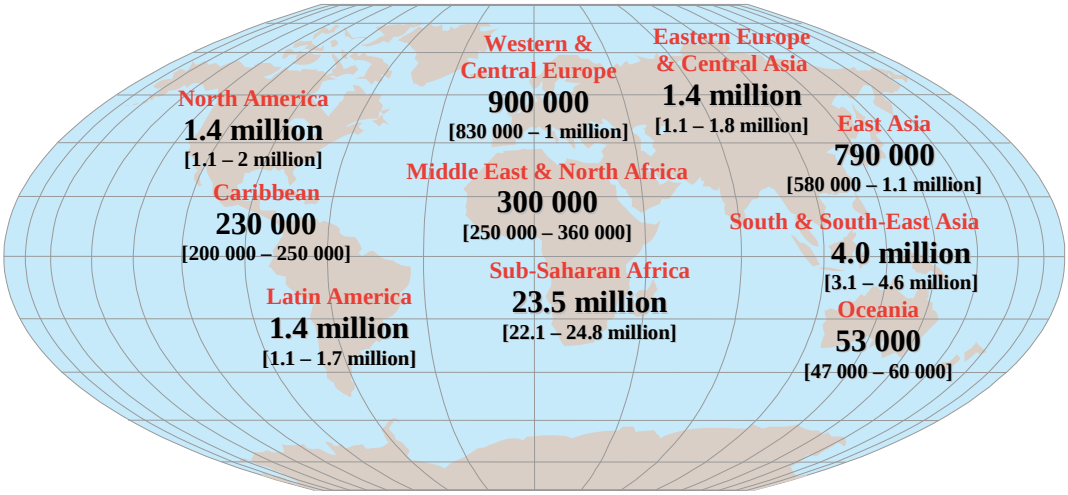
avian leucosis virus or murine leukemia virus often induce malignancies, immunodeficiencies or neuropathologies in their natural hosts. The one exception to the pathogenicity of retroviruses is the spumavirus genus. Although foamy viruses (FVs) are often highly cytopathic in tissue culture, they are not clearly associated with any disease. FVs are endemic in many vertebrates, including cows, domestic and feral cats, horses, and all primates examined other than humans. It is remarkable that the FV distribution mirrors that of the lentiviruses. All groups that are naturally infected with a lentivirus are also infected with a FV (Linial 2000). That gives a "plus" for studying the HFV PR and the importance to know its function better. The molecular and immunological underpinnings of FV-host interactions are just beginning to be understood, but these have much to teach us about retroviral adaptation (Linial 2000). EIAV is a lentivirus with strong homology to the immunodeficiency viruses (Stephens et al., 1986). It is an important model for HIV (Weber et al., 1993). Mason-Pfizer monkey virus (MPMV), the prototype type D retrovirus induces an AIDS-like syndrome in non-human primates that is distinct from that caused by simian immunodeficiency virus (Hruskova-Heidingsfeldova et al., 1995). The mouse mammary tumor virus (MMTV) is a B-type retrovirus and is most frequently associated with the induction of breast cancer in mice. BLV is the etiologic agent of enzootic bovine leukosis, a disease characterized by the occurrence of clonal lymphoid tumors of B-cell origin (Ghysdael et al., 1984). Walleye dermal sarcoma virus (WDSV) is a piscine retrovirus associated with skin tumors in walleyes (Bowser et al., 1988).

2.3. Orthoretroviruses

The study of orthoretroviruses has been intensified since the discovery of pathogenic human retroviruses. This is the case of HIV. Because of its association with acquired immunodeficiency syndrome, human immunodeficiency virus has been studied intensively. HIV protease has been targeted as a therapeutic intervention point in the treatment of AIDS because of its essential role it plays in viral maturation.

2.3.1. Human immunodeficiency virus and anti-HIV treatment

AIDS (acquired immunodeficiency syndrome) is a worldwide epidemic caused by infection with HIV, a human retrovirus. According to the UNAIDS 2011 report (see Fig. 3), 34 million people worldwide are living with HIV, of which more than one million are in the European countries.



Total: 34 million (31.6 – 35.2 million)

Figure 3: UNAIDS global report. Adults and children estimated to be living with HIV (2011)

Proteolysis occurs at many points of the retroviral life cycle, and these events can be considered as targets for chemotherapy (Tőzsér and Oroszlán, 2003). The most well-known proteolytic action in the retroviral life cycle is the processing of the Gag and Gag-Pro-Pol polyproteins with the virally encoded PR at the late phase of viral infection. Protease inhibitors, together with RT (reverse transcriptase) inhibitors, are important components of the drug combinations currently used to treat HIV patients. The use of these drugs in combination (a treatment termed highly active antiretroviral therapy, HAART) is available to extend and improve the lives of patients afflicted with AIDS. HAART reduces viral load and prevents many opportunistic infections associated with AIDS (Tőzsér 2003). The status of anti-HIV treatment is never static, but constantly changing directed to meet the requirements (Tandon and Chhor, 2005). The current combination therapy substantially reduced morbidity and mortality in HIV-infected patients. However, these drugs do not allow viral eradication; therefore their long-term use is required, allowing the development of resistance in a large portion of patients. Furthermore, several adverse metabolic side effects have been observed associated with the therapy.

The successful development of about 20 new anti-HIV drugs during the last two decades is ample proof that antiviral discovery has truly come of age and that selective antiviral drugs can achieve important clinical benefits. Among the approved drugs, nine compounds are inhibitors of the PR (Pauwels 2006).

2.3.2. Life cycle of orthoretroviruses

The infection of HIV begins with the recognition of viral envelope glycoprotein by the cell surface receptors CD4 and other coreceptors on the host cells (Maddon et al., 1986; Deng et al., 1996). After the virus fuses with the host cell membrane, HIV releases the viral genetic material into the cytoplasm of the host cell. The viral RNA is first reverse transcribed into DNA by reverse transciptase; therefore DNA replicates into double strands, and then it is inserted into the genome of the host cell (mediated by integrase). The viral genome could also replicate with the host cell genome. The most important enzyme for this doctoral study is the protease, which is located at the upstream of Pol in Gag-Pro-Pol polyprotein. The *gag* and *pol* genes are usually in different translational reading frames, yet both are translated from the same unspliced viral mRNA. Most of the time just the Gag polyprotein is translated, but in a fraction of time (e.g. for HIV-1 about 5%) a larger Gag-Pol polyprotein is synthesized owing to translational frameshifting. HTLV-1 and BLV are members of deltaretrovirus genus and their PR is coded through a mechanism of expression in separate *gag-pro* and *gag-pro-pol* open reading frames. In the case of MMLV (gammaretrovirus) there is a stop codon (AUG) between the two genes, which are in the same reading frame. WDSV shares the same characteristics as MMLV, in this regard. These fish viruses are unique in their natural replication circumstances as it occurs in the cold, near 4°C. Avian C type viruses (e.g. RSV) code the PR also in frame with the *gag* gene but without the need of stop codon suppression, therefore their PR is made in equivalent amounts to the structural proteins, unlike the other retroviruses where the PR is synthesized just 5-10% of the Gag.

The virus particle packing is a self-assembly course under the direction of the Gag precursor polyproteins (Gheysen et al., 1989; Luban and Goff, 1994). When the uncleaved viral precursor polyproteins, viral RNA and other elements are packed into the viral particles and released from the infected cells, they are immature and have no infectious ability. The virus only becomes infectious after the PR cleaves the Gag and Gag-Pol into functional proteins.

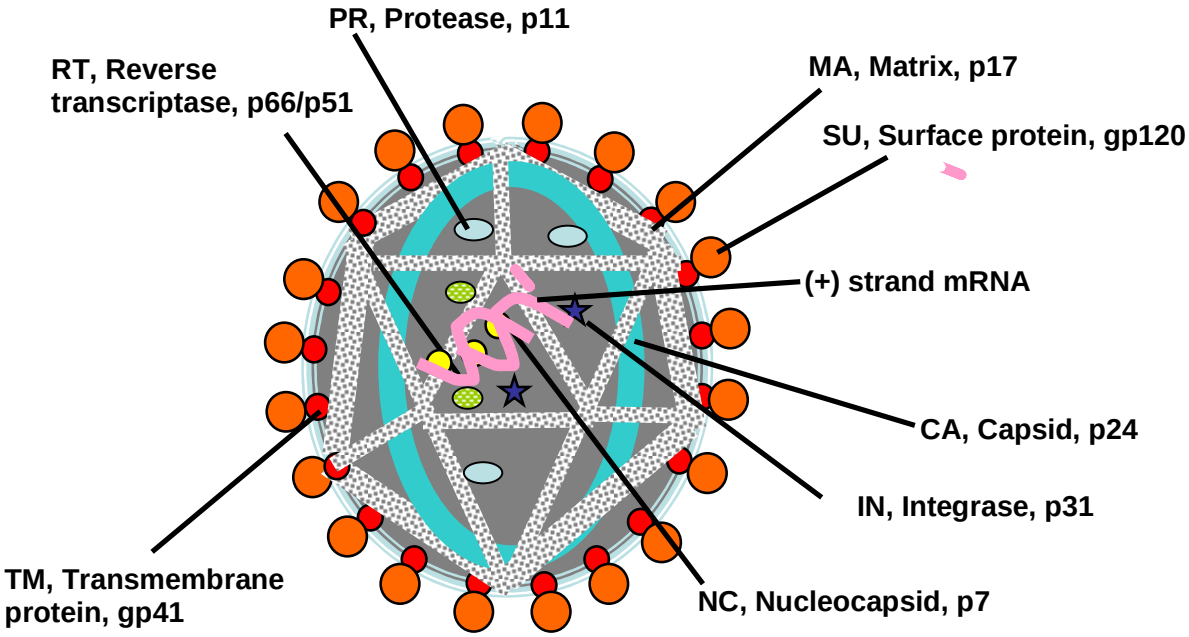


Figure 4: Cross section of a retrovirus

The Env polyprotein components: SU and TM proteins are out of the lipid bilayer as knobs. The Gag polyprotein: MA, CA and NC are internal structural proteins. Retroviral ribonucleoproteins are encased within a protein shell, built from the capsid protein to form an internal core, which can have different shapes, typically icosahedral, but HIV core has a conical shape. In the core, along with the RNA that carries the virus's genetic information, are enzymes known as reverse transcriptase, integrase and protease.

Besides these primary proteins, HIV also encodes some accessory proteins, which regulate the viral infection, replication and maturation. All the retroviruses follow a similar life cycle including entry, reverse transcription, integration, translation, assembly and budding (Figure 5). The whole life cycle is usually divided into two phases: the early phase ends with the integration of the viral genetic information into the chromosome of the host cell, while the late phase includes viral protein expression and virus maturation (Turner and Summers, 1999; Freed 2001; Tőzsér 2003).

For most retroviruses, assembly and budding occur at the cell surface. The Gag and Gag-Pro-Pol polyproteins are assembled together with the envelope proteins and the viral genomic RNA at the surface of the infected cell. The viral RNA that will be packaged is associated with Gag. Once all of the virus components are localized to the cell membrane, assembly into virions occurs. The assembled "immature" particles bud from the membrane, followed by "maturation" to a morphologically distinct form with a condense core.

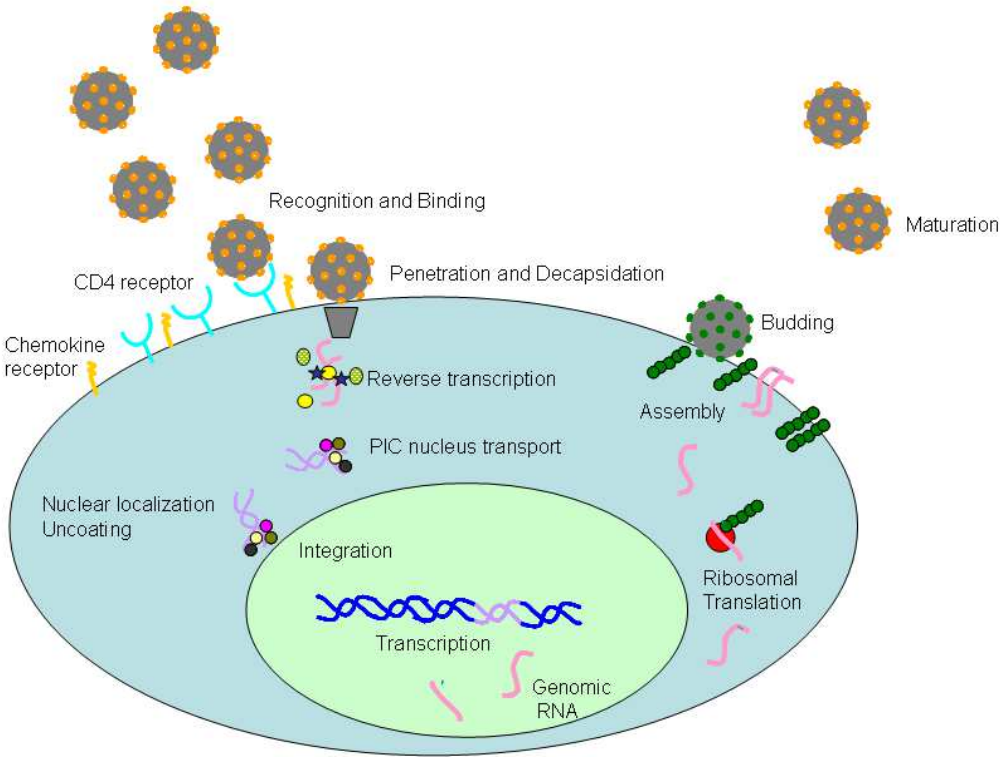


Figure 5: General features of HIV life cycle

The cycle starts by the interaction between receptor and envelope protein on the surface of the target cell. This leads to fusion and entry of the cone shape core. Reverse transcription then generates a double-stranded DNA copy of the RNA genome. The provirus is transported into the nucleus and integrated into chromosomal DNA. This process is mediated by integrase. In the late phase of the cycle, the viral DNA is transcribed by the cellular RNA polymerase II. The RNA is used as template for protein synthesis. Virion proteins and progeny RNA assemble at the plasma membrane. Progeny "immature" virus is released by a process of budding followed by protease mediated "maturation" into infectious virus.

2.4. Spumaviruses

Foamy viruses (FVs) are the sole members of the *Spumaretrovirinae* subfamily. Spumaviruses are complex retroviruses whose replication cycle resembles that of the hepadnaviruses, which exploits a late RT step and the utilization of different promoters for RNAs encoding virion-associated proteins and accessory proteins. They have been isolated from a wide range of mammalian hosts, including primates, cats, cows, and horses, and are characterized by a very broad tropism in vitro, allowing the infection of a great variety of different cell types and species (Meiering and Linial, 2001). HFV has been renamed simian foamy virus SFVcpz (hu) to more clearly indicate that the original HFV isolate is a chimpanzee (cpz) FV isolated from a human-derived cell culture (Herchenroder et al., 1994). In cell culture, FV replication often results

in the formation of vacuolated, multinucleated syncytia with a characteristic foamy appearance (Rustigian et al., 1955).

Feature of FVs not shared by other retroviruses (some will be noted in the life cycle part and retroviral protease part) is the presence of an internal promoter for expression of the regulatory *bel* (between env and LTR) gene (Fig. 2). Bel-1 activates transcription of the LTR of HFV and HIV (Rethwilm et al., 1991; Keller et al., 1991). In contrast to orthoretroviruses, the FV PR cleaves the cognate Gag protein only once prior to or during budding (Flügel and Pfrepper, 2003).

The fact that we are dealing with a group of viruses with no known pathogenicity raises questions about the rationale for studying FVs. The characteristics of FV infection, such as lack of pathogenicity, broad host range, and wide tissue tropism, persistence in the presence of neutralizing antibody, and the large genome size, make FV attractive vectors for gene therapy (Lindemann and Rethwilm, 2011; Trobridge and Russell, 2004; Mergia and Heinkelein, 2003; Meiering and Linial, 2001; Linial 2000). However, such an application requires detailed knowledge of the replication strategy as well as the enzymes of this virus. Since the FVs have properties that are most distinct from other members of the retrovirus family (unique replication strategy), it seems of particular interest to examine this valuable new system (Yap et al., 2008).

2.4.1. Life cycle of spumaviruses

Foamy viruses display a number of remarkable differences in their life cycles compared to conventional retroviruses. Interestingly, some of the unique features of the FV replication strategy show strong similarity to those of hepatitis B virus (by a late-occurring reverse transcription step and packaging DNA into the virions). Figure 6 shows an outline of HFV replication, with the unique aspects in red text and arrows, and non-unique stages in black. For example, unlike in other members of the retrovirus family, reverse transcription of FVs occurs predominantly in the producer cells (late step), not in the target cells. FV particles contain both viral RNA and DNA at ratio of about 5:1. Studies on the kinetics of RT strongly suggest that the functionally relevant nucleic acid within virions is DNA (Yu et al., 1996; Moebes et al., 1997; Yu et al., 1999; Roy et al., 2003).

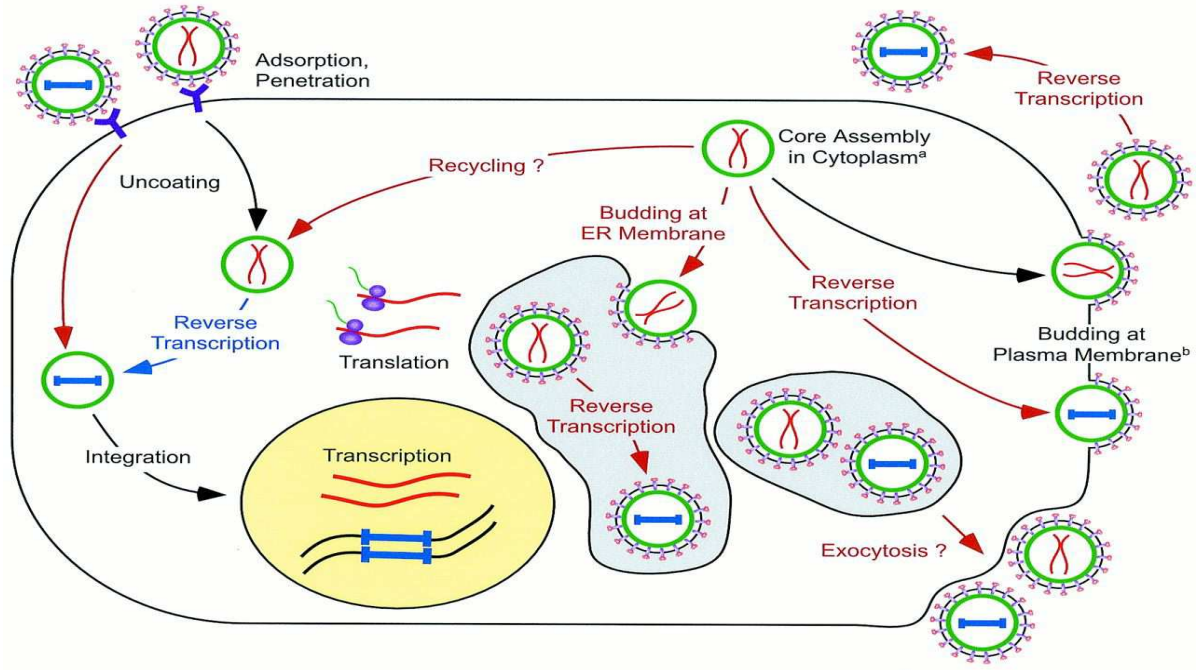


Figure 6: Life cycle of HFV

(Based on: Meiering and Linial, 2001)

Furthermore, the FV Pol protein is expressed as a separate protein (using a separate mRNA) instead of the Gag-Pol fusion protein found in the orthoretroviruses (Flügel and Pfrepper, 2003; Bodem et al., 1996; Enssle et al., 1996; Jordan et al., 1996; Löchelt and Flügel, 1996; Yu et al., 1996). Foamy virus genome contains the *pro* gene that codes for PR in the *pol* reading frame. The FV PR domain starts at the first Met of the *pol* reading frame (Löchelt and Flügel, 1996). Only the integrase domain is cleaved off from Pol resulting in a mature reverse transcriptase harboring the protease domain at the N-terminus (PR-RT) (Hartl et al., 2010). The FV Gag protein is unusual because it is not clearly divided and processed into matrix, capsid, and nucleocapsid proteins, as is observed for conventional retroviruses. During particle assembly and release, proteolytic processing of FV Gag at a single C-terminal site seems to take place, resulting in FV virions composed of the Gag precursor (p71Gag) and its processing product (p68Gag) (Flügel and Pfrepper, 2003). Further, FV particles resemble to the immature forms of conventional retroviruses, also suggesting the absence of Gag cleavage and consequent CA rearrangement in the extracellular phase of the life cycle (Morozov et al., 1997). The secondary FV Gag processing sites by the FV PR suggests that they play an important role during the early steps in FV replication, involving a cellular and FV protease-dependent disassembly pathway during entry into the target cell (Pfrepper et al., 1999;

Lehmann-Che et al., 2005). Some reports suggest that further FV Gag processing by the viral and cellular proteases occurs upon target cell entry and is essential for productive infection, probably by controlling capsid disassembly (Zemba et al., 1998; Lehmann-Che et al., 2005). After attachment of the virus to the cell membrane, the virus particle may gain entry into the cell by viropexis (engulfment of the virion by the cell). During this process the envelope of the virus and the plasma membrane fuse and the virion is carried into the cell within a vacuole (Hooks and Gibbs, 1975).

The major homology region and zinc finger(s), which are found in CA and NC, respectively, in conventional retroviruses, are also absent in FV Gag (Maurer et al., 1988). FV Gag proteins lack the typical Cys-His motifs of orthoretroviral Gag proteins with implicated function in different steps of retroviral replication including genome packaging (Swanstrom and Wills, 1997). Instead, the C-terminus of the larger Gag processing product in all FV species is rich in glycine (G) and arginine (R) residues (three such boxes) (Schliephake and Rethwilm, 1994). It was recently proposed that a GR-box (boxes rich in glycines and arginines) plays a role in the encapsidation of foamy viral Pol protein (Lee and Linial, 2008). Finally, the requirement of FV Env coexpression for Gag membrane association and budding suggests the absence of a membrane-targeting domain in FV Gag that is normally located in the retroviral MA subunit (Baldwin and Linial, 1998; Fischer et al., 1998; Pietschmann et al., 1999).

2.5. Retroviral proteases

Retroviral proteases are key enzymes in viral propagation and are initially synthesized with other viral proteins as polyprotein precursors that are subsequently cleaved by the viral protease at specific sites to produce mature, functional units (Dunn et al., 2002).

2.5.1. Structure and function

Retroviral proteases are members of the aspartic protease subfamily of endoproteolytic enzymes, with the conserved -Asp-Thr/Ser-Gly- (DT/SG) motif that is the essential part of the catalytic center. Active retroviral proteases are homodimers with 99-138 amino acid residues in each subunit. These enzymes catalyze peptide bond hydrolysis through an acid-base mechanism mediated by the two conserved catalytic aspartic acid residues of the symmetric homodimer (Fig.7). One Asp (Asp-COO-) which is unprotonated acts as the general base to activate the attacking water (H-OH) in the active site. The activated water in turn attacks the carbonyl C of substrate at the

scissile bond to form a tetrahedral intermediate. The other carboxylate of aspartate in protonated state acts as an acid to polarize the carbonyl bond. The intermediate is broken into carboxylic acid and amine products as a result of the protonation of the scissile amide N and unstable bonds.

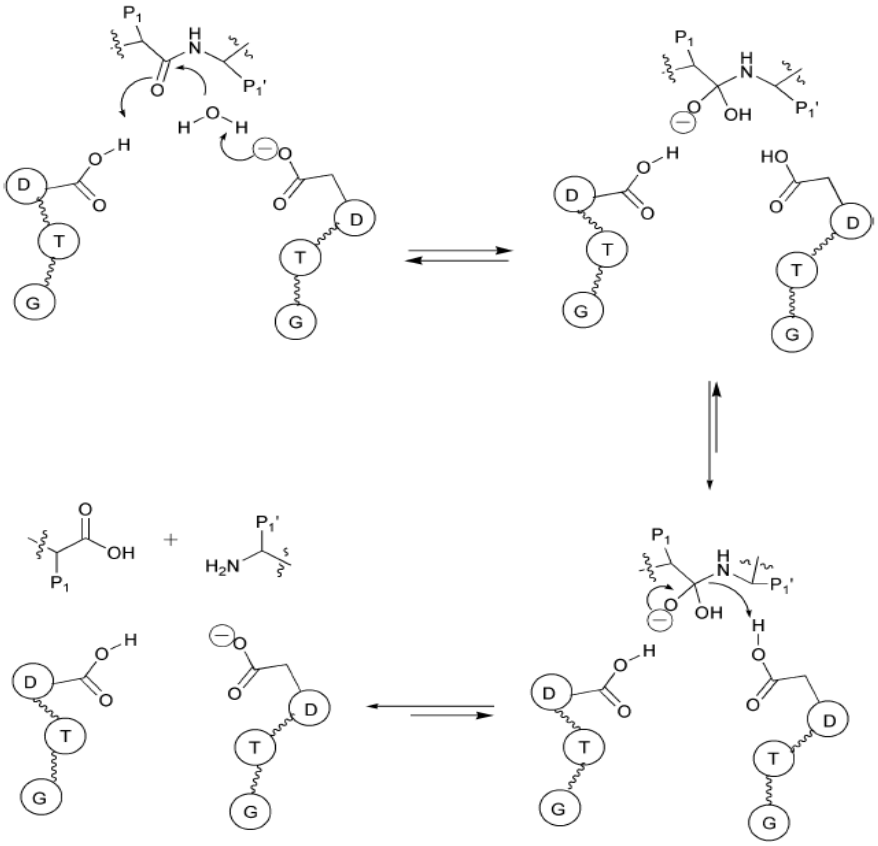


Figure 7: The proposed "general-catalytic mechanism" for aspartic PR

(Based on: Brik and Wong, 2003)

Each monomer folds into a compact structure and has four structural elements: two distinct hairpin loops, a wide loop containing the catalytic aspartic acid and a short α helix near the C terminus (Fig. 8). In addition to the four core structural elements, the amino and carboxyl termini in a dimer form a four-stranded β-sheet interface. The amino acid sequences of retroviral proteases are significantly similar, particularly in the locations that are important in preserving both structure and function. As can be seen in Figure 8 and 9, the retroviral PRs share some common features in amino acid sequence and tertiary structure.

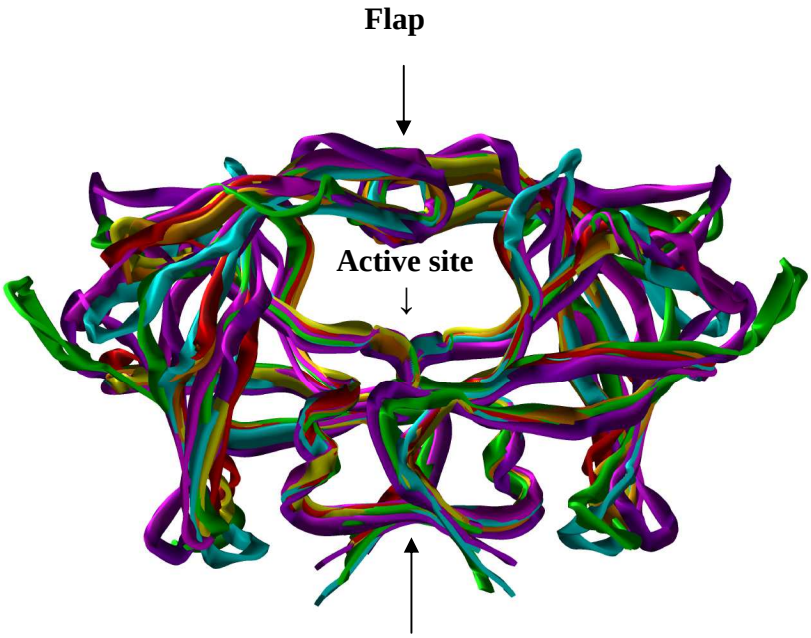


Figure 8: Superposition of seven retroviral proteases (crystal structures) shown in ribbon

(HIV-1 PR in red, HIV-2 PR in yellow, SIV PR in orange, EIAV PR in pink, FIV PR in blue, MAV PR in green, HTLV-1 PR in violet) (The figure was created by Péter Bagossi)

The overall three dimensional (3D) structures of PRs are very similar in the central region. They share conserved structural motifs at the conserved triplet (Asp-Thr/Ser-Gly) at the active site, the flap region, the dimer interface. Unlike the pepsin-like PR, which has only a single flap, the active retroviral PR possesses two flap regions, one from each monomer. The flap clinches a substrate into its active site cavity and releases products out of the active site, so it has to be fairly flexible. The flaps are important factors in determining the substrate specificity of the proteases.

Unlike with other retroviruses, FV PR of low molecular mass is not detectable in purified virions. Another special feature is that the catalytic center of the active dimers of HFV PR consists of DSG instead of DTG of most orthoretrovirus proteases. The reported equine foamy virus (EFV) PR sequence contains a Thr residue at this position indicating that this position can be occupied by either Ser or Thr in FV PRs. In avian retroviruses and retrotransposons, the Thr is also replaced by a Ser residue. Also there is a DSQ catalytic center of cat foamy virus (FFV) PR, which is an unprecedented feature of this enzyme.

Sequence analysis of retroviral proteases, including the genetic distances between them suggests viral evolution under selective pressure. Sequence alignment of retroviral proteases showed

that several unusual residues can be found around the active triplet: the residue corresponding to His22 (in HFV) is Leu in all other studied retroviral PRs. There is an exception for those of foamy viruses in which aromatic residues Phe (in EFV) or Tyr (FFV, BFV) also can be found. Residues corresponding to Leu10, Ala28, Val82, Ile84 in HIV-1 PR (Leu10, Ala27, Val90, Trp92 for HFV PR) have similar character in HFV PR, while the polar character of His22 may disturb the hydrophobic shell located near the catalytic aspartates.

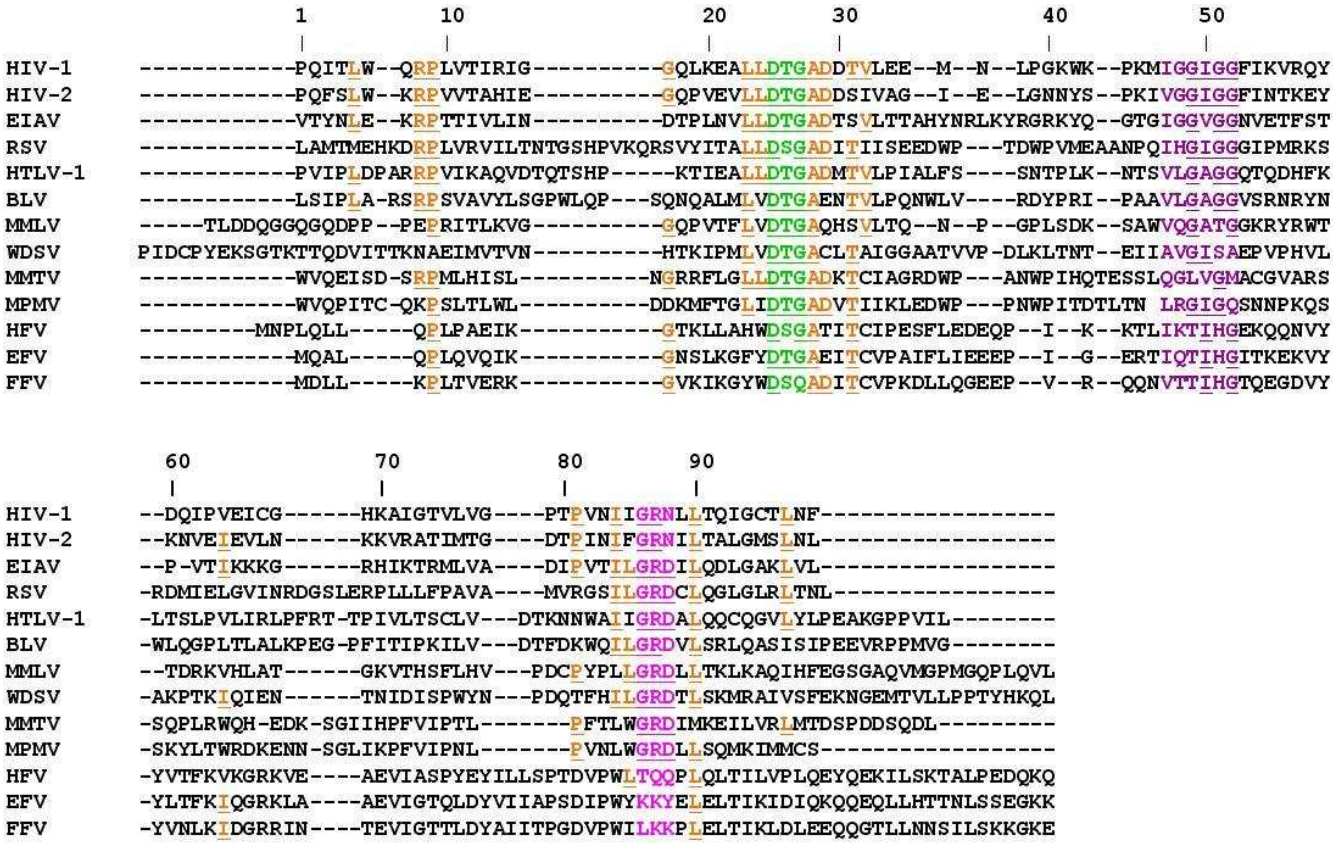


Figure 9: Sequence alignment of retroviral proteases

Coloured conserved regions: green for catalytic region (DTG), dark violet for flap, and pink for dimer interface (GRN). Identical residues are shown in orange and/or underlined.

A molecular model suggested that the protease containing Ser at the active site triplet instead of Thr may form less stable dimers (Bagossi et al., 1996). When spumaretroviral PRs are aligned with other retroviral PRs, in general, a very low degree of sequence identity between FV PR and retroviral PRs is obtained (Löchelt and Flügel, 1995) (Fig. 9); conserved features like the so-

called flap region cannot be identified. But the PR of the HFV is also essential for viral infectivity (Konvalinka et al., 1995).

Crystal structures have been determined for HIV PR in the absence and presence of inhibitors, as reviewed in Wlodawer and Erickson (1993). Meanwhile, many other retroviral PR structures have been solved by X-ray crystallography, including those of ASLV (avian sarcoma-leukosis virus) (Miller et al., 1989), RSV (Jaskolski et al., 1990), HIV-2 (Mulichak et al., 1993; Tong 1995), SIV (simian immunodeficiency virus) (Rose et al., 1993; Zhao et al., 1993), FLV (feline leukemia virus) (Wlodawer et al., 1995), EIAV (Gustchina et al., 1996), and HTLV-1 (Li et al., 2005).

2.5.2. Type of- and natural cleavage sites found in retroviral proteases

Classification of the cleavage sites found in HIV-1 has been published by Oroszlán and Luftig (1990), based on scissile bonds and the surrounding sequences. The cleavage sites of retroviral proteases have been classified as type 1, which contains an aromatic amino acid at P1 and Pro at P1', respectively, and type 2, which has mainly hydrophobic residues but no Pro at the site of cleavage (Table1).

Table 1: Classification of natural cleavage sites of HIV-1 PR (Based on: Louis et al., 2000)

Positions:	P4	P3	P2	P1		P1'	P2'	P3'	P4'	Cleavage site
Type 1:	Ser	Gln	Asn	Tyr	↓	Pro	Ile	Val	Gln	**MA/CA**
	Ser	Phe	Asn	Phe	↓	Pro	Gln	Ile	Thr	**TF/PR**
	Thr	Leu	Asn	Phe	↓	Pro	Ile	Ser	Pro	**PR/RT**
Type 2:	Ala	Arg	Val	Leu	↓	Ala	Glu	Ala	Met	**CA/p2**
	Ala	Thr	Ile	Met	↓	Met	Gln	Arg	Gly	**p2/NC**
	Pro	Gly	Asn	Phe	↓	Leu	Gln	Ser	Arg	**NC/p6**
	Arg	Lys	Ile	Leu	↓	Phe	Leu	Asp	Gly	**RT/IN**

Due to the rapid evolution of retroviruses, the cleavage sites may contain optimized sequences. Some of them should provide rapidly cleavable sites, while others should be poor substrates at other sites to provide the sequential order of cleavages for proper particle maturation

(Wiegers et al., 1998). *In vitro*, the order of HIV-1 Gag cleavage has been shown to be as follows: p2/NC, MA/CA, and p1/p6, NC/p1, and CA/p2 (Pettit et al., 1998). Cleavage of CA/p2 and cleavage of NC/p1 are the two slowest cleavage events and are thus good candidates for threshold-defining events (Dauber et al., 2002).

Naturally occurring type-1 cleavage site sequences on Gag and Gag-Pol polyprotein are listed in Table 2. Despite the highly specificity in site selection, recognition sequences for the retroviral PR are diverse in amino acid composition

Table 2: Type 1 natural cleavage sites (-Phe/Tyr↓Pro-) of the studied retroviral proteases

Virus[a]	Cleavage site	Oligopeptide sequence
HIV-1	MA/CA	VSQNY↓PIVQ
	in p6	DKELY↓PLTSL
	TF/PR	VSFNF↓PQITL
	PR/RT	CTLNF↓PISP
HIV-2	MA/CA	EKGGNY↓PVQHV
	NC/p6	KPRNF↓PVAQV
EIAV	MA/CA	PSEEY↓PIMID
	NC/p9	QKQTF↓PIQQT
AMV	RT/IN	TFQAY↓PLREA
MPMV	p12/CA	PKDIF↓PVTET
MMTV	n/CA[b]	LTFTF↓PVVFMRR
MMLV	MA/p12	PRSSLY↓PALTP
	p12/CA	TSQAF↓PLRAG

[a]There is not known type 1 cleavage site for HTLV-1, BLV, WDSV, and HFV PR.
[b]In MMTV the small peptide (n) presumably located upstream of CA has not been identified.

2.5.3. Substrate specificity

Understanding the substrate specificity of HIV PRs is important for studying the molecular basis of drug resistance and development of new drugs. The naturally occurring cleavage site sequences on Gag and Gag-Pol polyprotein are listed in Table 1. For the optimal catalysis, the minimal length of substrates is 7 amino acids (Darke et al., 1988; Tőzsér et al., 1991; Tomasselli and Heinrikson, 1994; Coffin et al., 1997) and that is determined by the ability of substrate amino acids side chains to bind into eight individual subsites within the enzyme. Although the subsites are able to act somewhat independently in selection of amino acid side chains that fit into each pocket, significant interactions exist between individual subsites that substantially limit the number of cleavable amino acid sequences. The substrate of HIV PR binds to the dimer in an asymmetric way, as illustrated in Figure 10. Substrate sequences cleaved by protease are relatively diverse. Despite this diversity, some common features can be deduced from the analysis of natural viral sequences, enzyme kinetic and mutagenesis studies (Pettit et al., 1991; Poorman et al., 1991; Tőzsér et al., 1992; Coffin et al., 1997; Louis et al., 2000; Beck et al., 2002). The substrate specificity is determined by the overall shape and chemistry of the side chains of peptide substrates rather than the specific sequences (Konvalinka et al., 1990; Tőzsér et al., 1991a, 1991b; Griffiths et al., 1992; Tőzsér et al., 1992; Prabu-Jeyabalan et al., 2002; Ozer et al., 2006).

The subsite preference has also been studied by enzyme kinetics using oligopeptide substrates. In agreement with the natural cleavage site sequences, HIV PR prefers small hydrophobic residues at P2 and can accommodate various residues at P3 and P3' while the P2' preference depends on the amino acid of P1 and P1' (Konvalinka et al., 1990; Tőzsér et al., 1991a, 1991b; Tőzsér et al., 1992; Cameron et al., 1993).

To understand the similarities and differences among the specificities of retroviral proteases, and to better understand the structure-activity relationship molecular models were built for all of the studied enzymes (Tong et al., 1995; Gutschina et al., 1996; Kervinen et al., 1998; Wu et al., 1998; Mahalingam et al., 2002).

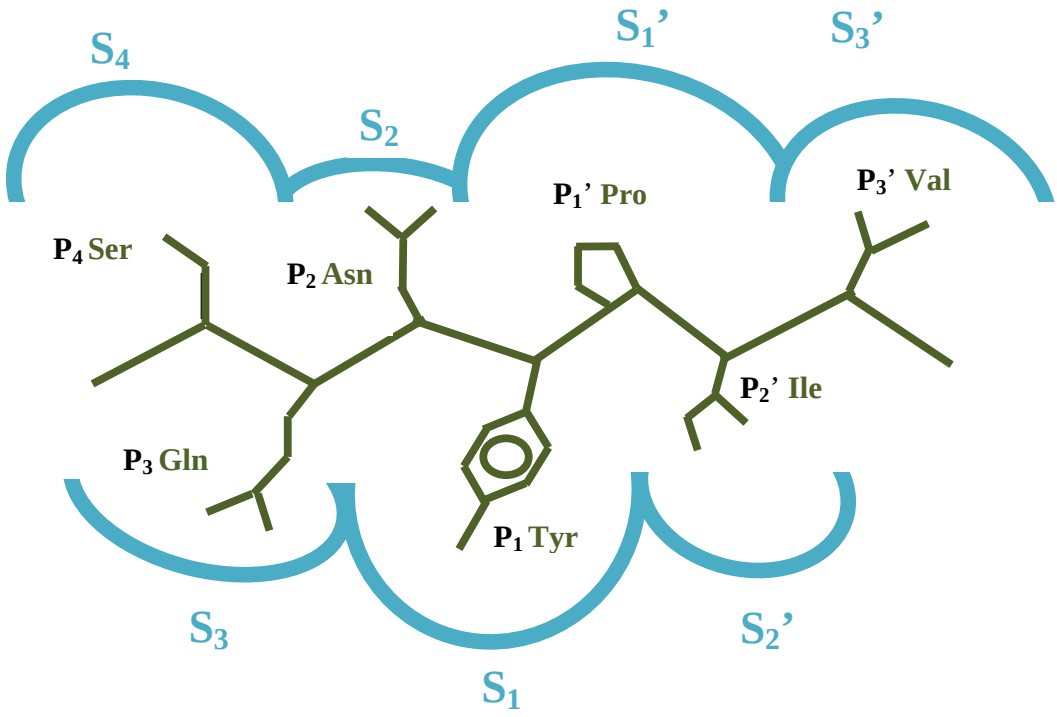


Figure 10: Schematic representation of HIV-1 MA/CA substrate (P4-P3') bound to HIV-1 PR (S4-S3') subsites

The relative size of each subsite is indicated approximately by the area enclosed by the curved line around each substrate side chain. The active site of the enzyme is composed of seven subsites located on both sides of the catalytic site. The P positions on the substrate are counted from the site of cleavage and thus have the same numbering as the subsites they occupy. (Based on nomenclature of Schechter and Berger, 1967)

2.5.4. Mutations in retroviral proteases. Resistance to protease inhibitors

Besides the natural polymorphisms, the emergence of drug resistance is very severe and it is probably the main factor leading to the failure of current treatments of HIV. It is estimated that over 70% of HIV-1 infected individuals harbor drug resistant virus and nearly 5-10% of them reveal resistance to all of the current RT and PR inhibitors (Yu et al., 2005). What makes it even worse is that virus strains carrying drug resistant mutations are transmitted directly to newly infected individuals, which correspond to about 10-15% of the total newly infected cases (Hirsch et al., 1998; Wainberg and Friedland, 1998). The mutations in PR alter single or multiple residues which can cause multidrug resistance and cross-resistance (Figure 11). Mutations in 45 out of 99 residues of the PR have been associated with the treatment with protease inhibitors (PIs) (Zhang et al., 1997; Hertogs et al., 2000; Wu et al., 2003; Johnson et al., 2008). In some cases, the observed structural changes in mutations are in agreement with kinetic and stability changes. The selection of

mutations in cleavage sites raises the possibility of simultaneous changes in protease specificity, and a potential role of altered specificity in the development of drug resistance. Furthermore, it has been shown that various drug-resistant viruses produce different patterns of accumulated, partially processed polyproteins (Zennou et al., 1998). This observation suggests that particular mutations in the protease can differentially affect cleavage site recognition and, by association, the substrate preference of the protease.

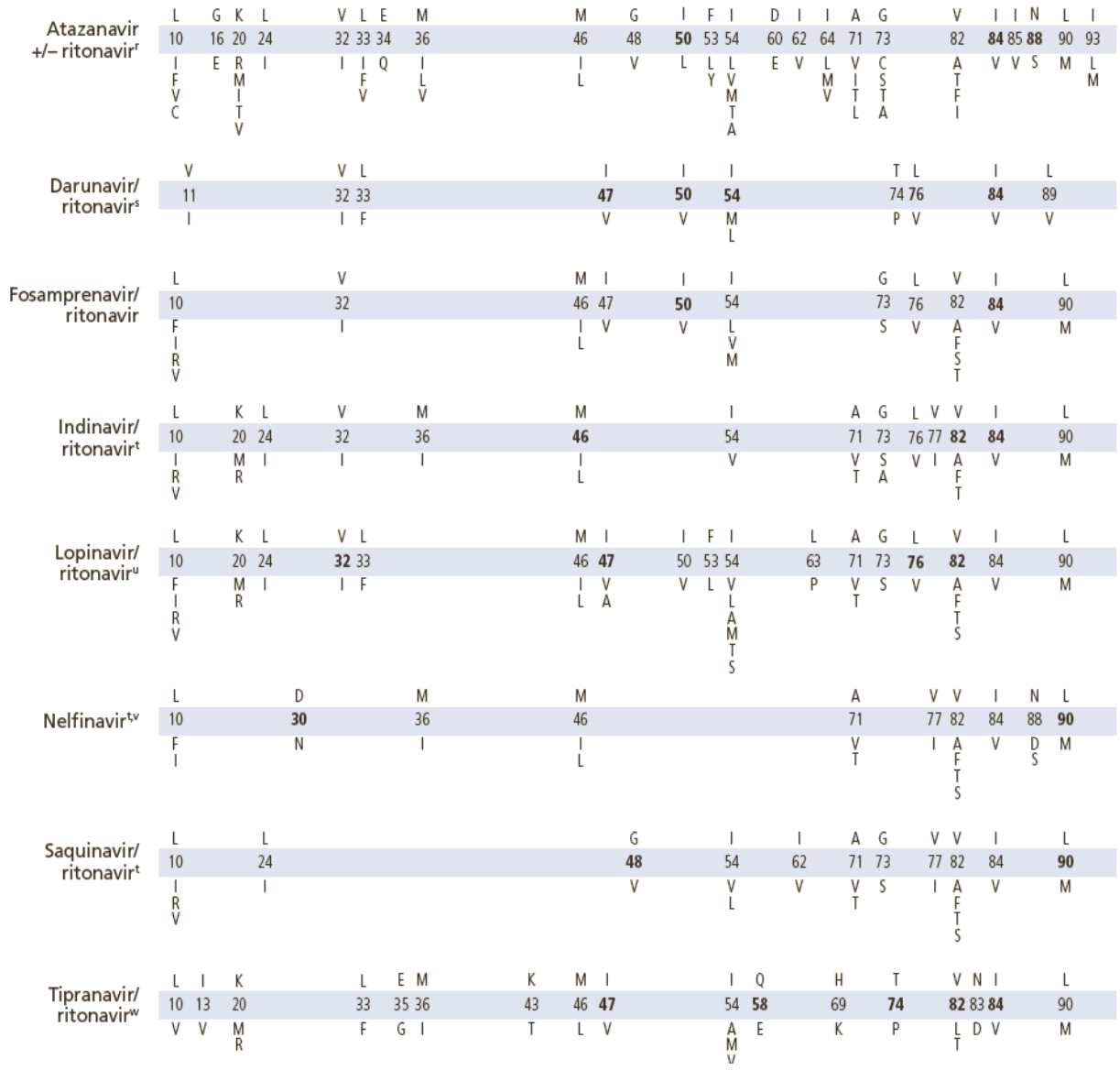


Figure 11: Mutations in the protein sequence associated with resistance to protease inhibitors

(Based on: Johnson et al., 2009)

2.6. Scope of the study

Foamy virus is a model system for gene transfer. FV vectors were shown to be efficient gene delivery vehicles for different scientific and therapeutic approaches in a variety of different target cells (Rothenaigner et al., 2009). In comparison with HIV-1 and MLV-based vectors they are capable of similar or higher gene transfer efficiency (Leurs et al., 2003). Our aim was to analyse the structure-function relationship in the foamy protease because knowing it better would help in understanding the different replication cycle of the virus. Recently, Hartl et al. demonstrated that FV PR is an inactive monomer with a very weak dimerization tendency and postulated PR activation through dimerization. They identified a specific protease-activating RNA motif (PARM) located in the *pol* region of viral RNA which stimulates PR activity *in vitro* and *in vivo*, revealing a novel and unique mechanism of retroviral protease activation (Hartl et al., 2011).

In the absence of a crystal structure for the HFV PR, a model was built in order to understand the molecular basis for the unusual parameters (lower dimer stability and higher pH optimum in comparison with HIV-1 PR). The amino acid sequence of HFV PR was aligned with the sequences of HIV-1 PR, to find the best starting structure for building the model and constructing the mutants. The specificity and mutation tolerability of the wild-type FV PR has not been characterized in detail. Previously, the Ser of the active site triplet of the enzyme was changed to Thr in an attempt to enhance the activity and stability of the enzyme (S25T; the Ser is the 25^{th} residue of the native enzyme). The increased dimer stability and pH optimum of the S25T enzyme as compared to the wild-type (wt) HFV PR initiated us to design and test further mutations in the spatial proximity of the active site.

Retroviruses are associated with human and animal diseases; therefore its proteases are potential targets for chemotherapy. The HIV-1 PR has proved to be the most effective target of antiviral therapy. However, the application of PR inhibitors was largely limited by rapid development of drug resistance variants (Tamalet et al., 2000). Comparative studies of divergent members of the retroviral protease family are a promising approach not only to recognize general and specific features of the PR, but to discover the mutational capacity of the PR. Several of the mutations causing drug resistance of HIV-1 PR introduce residues into the substrate binding sites found in equivalent position of other retroviruses (Ridky and Leis, 1995; Louis et al., 1999). Previously, a large series of peptides containing single amino acid substitutions in the P4-P3' (nomenclature is according to reference (Schechter and Berger, 1967)) region of the Val-Ser-Gln-

Asn-Tyr↓Pro-Ile-Val-Gln oligopeptide was used to characterize the specificity of the protease of various retroviruses including that of HIV-1, HIV-2, EIAV, MMLV and AMV. To compare the specificity of the retroviral enzymes we determined the relative activities for the substituted peptides by dividing it with the unmodified peptide. Previously, substrate specificity for P2 site was studied on a representative set of retroviral proteases, which included at least one member from each of the seven genera of *Retroviridae* (Bagossi et al., 2005). Here we complete the study with investigation of the specificity of P1, P3 and P4 positions using the same protease set including HIV-1, HIV-2, EIAV, MMLV, AMV, MPMV, MMTV, HTLV-1, BLV, HFV and WDSV proteases. A major reason why the specificity of various retroviral proteases is useful is the need to develop efficient PR inhibitors for clinical use, from which multiple viral strains and even different retroviruses can not escape by mutations.

Aim 1: HFV PR characterization (wild-type and mutants)

1. To replace and explore the role of some amino acid residues being close to the catalytic aspartates in the higher pH optimum and in the lower dimer stability of human foamy virus protease as compared to the classical retroviral PR (e.g. HIV-1 PR). For that analysis we constructed mutants based on the conserved amino acids in HIV-1 PR sequence at the same positions. The mutants were designed based on sequence alignments and on the molecular model of HFV PR for structure-function studies.
2. To study the pH optimum values of wild-type, single- (Q8R, H22L, S25T, T28D) and double- (Q8R/T28D, H22L/T28D) mutant foamy proteases in order to examine the role of the residues in the vicinity of catalytic aspartates.
3. To characterize stability of the foamy PRs dimer using urea denaturation, at two pH values (6.0 and 7.2).

Aim 2: Substrate specificity characterization of different retroviral proteases

1. To extend the specificity studies on a panel of retroviral proteases using oligopeptide substrates containing systematic mutations (on P1, P3, and P4 positions) in the naturally occurring cleavage site found between the matrix and capsid proteins of HIV-1.
2. To compare and understand the molecular basis of the different specificity among retroviruses using molecular models.

3. MATERIALS AND METHODS

3.1. Retroviral proteases

Purification procedures and the expression forms of retroviral proteases were described previously: HIV-2 PR (chemically synthesized, purified and refolded) (Tőzsér et al., 1992), MPMV PR (the shortest form, recombinant 12kDa, self-processed from fusion protein) (Zabransky et al., 1998), MMTV PR (recombinant, GST-fusion protein, processed with factor Xa and purified) (Menéndez-Arias et al., 1992), MMLV PR (recombinant, processed from MBP fusion protein, hexahistidine tags) (Fehér et al., 2004), HTLV-1 PR (recombinant, purified from inclusion bodies, stabilized from autoproteolysis) (Louis et al., 1999), BLV PR (recombinant, purified from inclusion bodies, self-pocessed from MBP fusion protein) (Zahuczky et al., 2000), WDSV PR (recombinant, purified from inclusion bodies, processed from a fusion form) (Fodor and Vogt, 2002), and HFV PR (recombinant, wild-type and mutants MBP-fusion proteins) (Q8R, H22L, S25T, T28D, Q8R/T28D, H22L/T28D) (Fenyőfalvi et al., 1999). The activity of the MBP fusion form of HFV PR was similar to that of the processed form (Boross et al., 2006).

Proteases that showed substantial self-processing ability were purified after large scale expression (500 ml) by ion-exchange chromatography (e. g. BLV PR, MPMV PR), whereas other proteases were purified in the fusion forms by using (amylose-) affinity chromatography as described previously for human foamy virus PR (Fenyőfalvi et al., 1999). All of the enzymes were at least 90% pure based on Coomassie staining on SDS-polyacrylamide gels (using 10-20% gradient gels).

Active site titration for the HIV-2, HTLV-1, BLV, and MMLV PRs was performed as described previously (Tőzsér et al., 1992, Zahuczky et al., 2000, Fehér et al., 2004). Active site titration for MPMV PR was performed using the phenylstatine containing inhibitor PYVPstAMT. Comparison of the protein content determined by Bradford assay (Bio-Rad) with the active-site values suggested that the folding of the PR was efficient. We didn't perform the active site titration for MMTV, WDSV and HFV PRs, due to the lack of sufficiently potent inhibitors.

3.2. Oligopeptides were synthesized by solid-phase peptide synthesis on a model 430A automated peptide synthesizer (Applied Biosystems, Foster City, CA, USA) or a semiautomatic Vega peptide synthesizer (Vega-Fox Biochemicals Div., Newbery Energy Corp., Tucson, Arizona) as described

previously (Tőzsér et al., 1992; Menéndez-Arias et al., 1992 and 1993; Fenyőfalvi et al., 1999). Prior to use, all the oligopeptides were purified by RP-HPLC (reverse-phase high performance liquid chromatography) (Copeland and Oroszlán, 1988). Their purity and peptide concentration was assessed after determination of their amino acid composition by amino acid analysis with a Beckman 6300 amino acid analyzer. Stock solutions and dilutions were made in distilled water (or in 5 mM dithiotreitol for the Cys-containing peptide). The following peptides were used in cleavage reactions: HIV-1 MA/CA (VSQNY↓PIVQ) and analogs (single amino acid substitutions in the P1, P3 and P4 positions) (Table 3), and a HFV oligopeptide substrate (SRAVN↓TVTQS). Peptides were obtained from Dr. Stephen Oroszlán and from Dr. Terry D. Copeland (Molecular Virology and Carcinogenesis Laboratory, NCI-FCRDC, Frederick, MD, USA).

Table 3: Set of nonapeptides containing single amino acid substitutions in the peptide representing the HIV-1 MA/CA cleavage site

VSQNY↓PIVQ the sequence of the original peptide

VSQN**G**↓PIVQ	VS**G**NY↓PIVQ	V**G**QNY↓PIVQ
VSQN**A**↓PIVQ	VS**A**NY↓PIVQ	V**A**QNY↓PIVQ
VSQN**V**↓PIVQ	VS**V**NY↓PIVQ	V**V**QNY↓PIVQ
VSQN**I**↓PIVQ	VS**L**NY↓PIVQ	V**I**QNY↓PIVQ
VSQN**L**↓PIVQ	VS**D**NY↓PIVQ	V**L**QNY↓PIVQ
VSQN**W**↓PIVQ	VS**K**NY↓PIVQ	V**T**QNY↓PIVQ
VSQN**C**↓PIVQ	VS**F**NY↓PIVQ	V**M**QNY↓PIVQ
VSQN**M**↓PIVQ	VS**N**NY↓PIVQ	V**D**QNY↓PIVQ
VSQN**S**↓PIVQ		V**K**QNY↓PIVQ
VSQN**D**↓PIVQ		V**P**QNY↓PIVQ
VSQN**K**↓PIVQ		V**F**QNY↓PIVQ
VSQN**P**↓PIVQ		V**N**QNY↓PIVQ
VSQN**F**↓PIVQ		

The substituted amino acid in the peptide is shown **bold** in pink. The substituted amino acids are hydrophobic: A (alanine), V (valine), I (isoleucine), L (leucine), F (phenylalanine), W (tryptophan), C (cysteine), M (methionine), P (proline), G (glycine), hydrophylic: S (serine), T (threonine), N (asparagine), Q (glutamine), Y (tyrosine), H (histidine) or charged: D (aspartic acid), K (lysine), R (arginine).

3.3. Study of protease activity and measurements of relative activity

Protease assays were performed at 37°C using purified retroviral proteases and chemically synthesized oligopeptides (0.4 mM) in 0.25 M potassium phosphate buffer, pH 5.6, containing 7.5% glycerol, 5 mM dithiothreitol, 1 mM EDTA, 0.2% Nonidet P-40 and 2 M NaCl. The reaction mixtures (20 µl: 5 µl enzyme + 5 µl substrate + 10µl incubation buffer) were incubated at 37°C for 1 hour (HIV-2, MPMV, MMTV, MMLV), 20 hours (HTLV-1, BLV) and 24 hours (WDSV, HFV) and were stopped by the addition of 9 volumes 1% trifluoroacetic acid (TFA) then injected onto a Nova-Pak C_{18} reversed-phase chromatography column (3.9 x 150 mm, Waters Associates Inc.) using an automatic injector. Substrates and the cleavage products were separated using acetonitrile gradient (0-100%) in water in the presence of 0.05% TFA. Relative activities were calculated from the molar amount of peptides cleaved per unit time at less than 20% substrate turnover, by dividing the activity on a given peptide by the activity on the reference substrate, as described (Blaha et al., 1992; Bagossi et al., 2005). The relative activities for the HIV-1 PR (Tőzsér et al., 1992), EIAV PR (Weber et al., 1993), and AMV PR (Tőzsér et al., 1996) have been reported previously. Previous studies also indicated a strong correlation between relative activities and the specificity constants (Tőzsér et al., 1996); therefore, the determined activity values can be considered as a measure of the k_{cat}/K_m values.

Kinetic parameters were determined by fitting the data obtained at < 20% substrate hydrolysis to the Michaelis-Menten equation by using the Fig.P program (Fig. P Software Corp.). Measurements were performed in duplicate and the average values were calculated. The standard errors of the kinetic parameters were below 20%.

3.3.1. High Pressure Liquid Chromatography (HPLC) methods

The kinetics of oligopeptide hydrolysis was determined using a high-performance liquid chromatography (HPLC) method, which permitted the separation of substrate from split products (Fig. 12). Peptides were detected at 206 nm, peaks were integrated, and than the kinetic parameters were determined. The integrated values were converted to values relative to our reference substrate (SP-211: VSQNY↓PIVQ). Cleavage products (VSQNY, PIVQ) of the hydrolysis of SP-211 (VSQNY↓PIVQ) were identified by their amino acid composition to confirm the site of cleavage with HIV-1 PR (Tőzsér et al., 1992), while the VSQNY cleavage product of the other substrates was identified on the basis of its retention time, which was found to be identical to that obtained with HIV-1 PR.

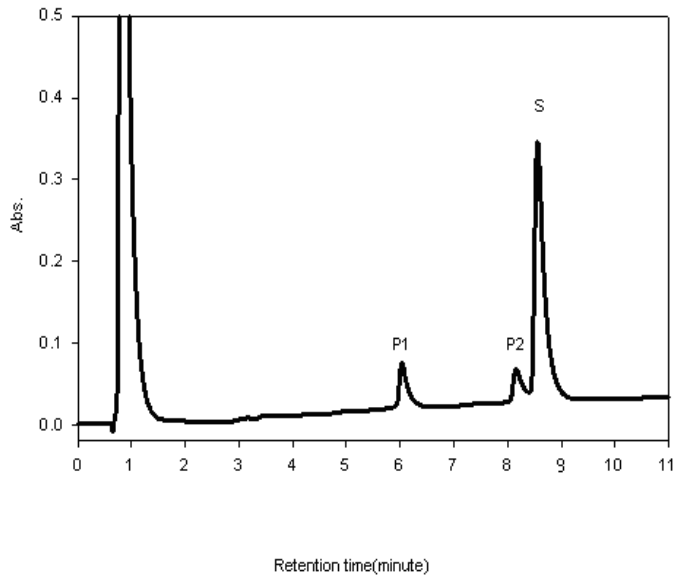


Figure 12: Reversed-phased chromatograph of a reaction mixture containing the reference substrate SP211 (VSQNY↓PIVQ), retroviral PR and cleavage products

P1 peak: VSQNY, P2 peak: PIVQ and S uncleaved substrate peak: VSQNY↓PIVQ

3.4. Preparation of human foamy virus protease (HFV PR) mutants

The pMBP-HFV PR clone was used as a template for mutagenesis (Fenyőfalvi et al., 1999). Cloning of the wild-type and S25T mutant HFV PR in fusion with maltose binding protein (MBP) was described previously (Boross et al., 2006). The sequences of primers used for introducing the mutations are listed in the Table 4. For single mutants (S25T, T28D, Q8R and H22L) the wild-type residues were exchanged conform the protocol. The second mutation for the double mutants (Q8R/T28D, H22L/T28D) was introduced into the single T28D mutant with the proper oligonucleotide pairs. Mutants were generated by the Quick-Change mutagenesis protocol (Stratagene, La Jolla, CA, USA) with the appropiate oligonucleotide pairs from Sigma-Genosys (The Woodlands, TX, USA). Mutations were verified by DNA sequencing performed with ABI-Prism dye terminator cycle sequencing kit and a model 373A sequencer (both from Applied Biosystems, Foster City, CA, USA).

Table 4: The DNA sequences of primers for introducing mutations

Mutant	5'-3' primer sequences (number of nucleotide)
Q8R	Forward: CTT CAG CTG TTA CGG CCG CTT CCG GCG (27 nt) Reverse: CGC CGG AAG CGG CCG TAA CAG CTG AAG (27 nt)
H22L	Forward: GGG ACT AAA TTG TTA GCC CTC TGG GAT TCA GGG GCA AC (38 nt) Reverse: GT TGC CCC TGA ATC CCA GAG GGC TAA CAA TTT AGT CCC (38 nt)
S25T	Forward: G TTA CGC CAC TGG GAT ACA GGG GCA ACA ATA AC (32 nt) Reverse: GT TAT TGT TGC CCC TGT ATC CCA GTG GGC TAA C (32 nt)
T28D	Forward: CC CAC TGG GAT TCA GGG GCA GAC ATA ACT TGT ATT CCT GAA AG (43 nt) Reverse: CT TTC AGG AAT ACA AGT TAT GTC TGC CCC TGA ATC CCA GTG GG (43 nt)
H22L/T28D	Forward: GGG ACT AAA TTG TTA GCC CTC TGG GAT TCA GGG GCA GAC (39 nt) Reverse: GTC TGC CCC TGA ATC CCA GAG GGC TAA CAA TTT AGT CCC (39 nt)

Mutated positions are indicated by underlined letters.

3.5. Expression of the wild-type and mutant HFV PRs

Protein expression was induced by the addition of 1 mM IPTG for 5 hours to *Escherichia coli* BL21(DE3) harbouring the plasmid encoding the wild-type or mutant MBP-HFV PR fusion proteins in 500 ml cultures. The culture was grown up at 37°C to an absorbance (at 600 nm) of 0.7-1.0, in Luria-Bertani medium containing 100 μg/ml ampicillin. After expression, cells were collected by centrifugation (2000g for 10 minutes at 4°C). After removal of the supernatant, 25 ml lysis buffer (50 mM Tris, pH 7.2, 1 mM EDTA, and 100 mM NaCl) was added to the pellet. Cells were disrupted by freezing-thawing folowed by sonication on ice. Samples were centrifuged at 9000g for 15 minutes at 4°C. The supernatant was loaded on a column containing amylose resin (25 ml) applied to ÄKTApurifier automated liquid chromatography system (Amersham Pharmacia Biotech AB, Uppsala, Sweden). The column was washed with three volumes of lysis buffer. The fusion protein was eluted with the same lysis buffer containing 20 mM maltose. Protein concentration of the fractions was determined by Bradford spectrophotometric method (Bradford 1976) using bovine serum albumin (BSA) as a standard protein. Purity of the protease was assesed by SDS-PAGE, using 10-20% gradient gels. Molecular mass markers (BioRad, Precision Plus Protein Dual Color Standard, Amersham Biosciences, Piscataway, NJ, USA) were used for comparison. The mutant proteases were estimated to be 90% pure based on Coomassie staining.

3.6. Proteolytic assay for HFV PR activity

Kinetic parameters were determined in 50 mM MES, 100 mM Tris, 50 mM acetate, 1 M NaCl, pH 6.3 buffer (META). The pH of the buffer system was sensitive to temperature; therefore it was adjusted at 37°C. The reaction mixtures contained 1.4-22 µM purified enzyme and 0.2-1.0 mM substrate (SRAVN↓TVTQS) and they were incubated for 1 h. The reactions were stopped by the addition of 180 µl 1% trifluoroacetic acid, and an aliquot was analyzed by reversed-phase HPLC as described by Fenyőfalvi et al. (1999). Cleavage products of PR-catalyzed hydrolysis were previously identified by amino acid analysis for wild type HFV PR (Fenyőfalvi et al., 1999) and mutant enzymes produced the same cleavage fragments as indicated by identical retention times. The same amount of purified wild-type (chromatogram a) and mutant (chromatogram b) enzymes were assayed with the foamy oligopeptide substrate SRAVN↓TVTQS to demonstrate their similar specific activity (Figure 13) (Boross et al., 2006). The Fig. P program (Fig. P Software Corp., Durham, NC, USA) was used for determination of the kinetic parameters (K_m and k_{cat}).

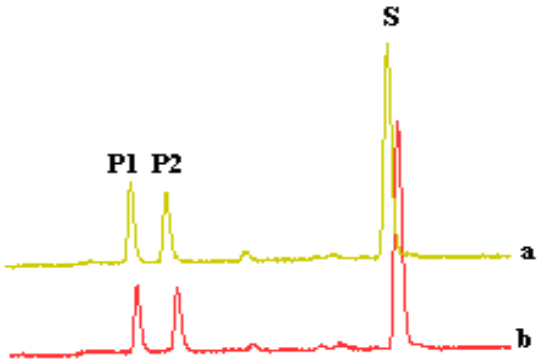


Figure 13: Reversed-phased chromatograph of a reaction mixture containing the foamy (SRAVN↓TVTQS) substrate, foamy PRs (wild-type and mutant) and cleavage products

P1 peak: SRAVN, P2 peak: TVTQS, S: uncleaved substrate peak: SRAVN↓TVTQS (Based on: Boross et al., 2006)

3.7. Determination of pH optimum of HFV PRs

The pH optimum of the enzymes was determined in META buffer with pH in the range of 3-9. The reaction mixtures were incubated for 1 hour and analyzed by HPLC method. In all cases, the chromatography was run at room temperature. Symmetrical bell-shaped pH optimum curves were fitted by nonlinear regression module of SigmaPlot program (Systat Software, Inc., Point Richmond, CA, USA) (Polgár et al., 1994).

3.8. Urea denaturation assay of HFV PRs

The denaturing effect of urea was measured using META buffer having pH 6.0 or 7.2 in the presence of increasing concentration of urea (0-6 M). PR activity was measured by injecting the cleavage reactions on HPLC column. The UC_{50} values at half-maximal velocity were obtained by plotting the initial velocities against urea concentration. Sigmoidal urea denaturation curves were fitted by the nonlinear regression module of SigmaPlot software (Jandu et al., 1990; Szeltner and Polgár, 1996a).

3.9. Sequence alignment and molecular modeling of retroviral proteases

Structure-based alignment of the HIV-1, HIV-2, SIV, EIAV, FIV and RSV proteases was used as a template for the alignment of BLV, HTLV-1, MPMV, MMTV, MMLV, WDSV and HFV protease sequences, as shown in Fig. 9. The initial multiple sequence alignment was made by ClustalW (Thompson et al., 1994), structural alignment was made by Whatif (Vriend, 1990), followed by manual corrections based on the structural alignment. The phylogenetic tree was made by ClustalW and Phylip (Felsentein, 1989). In the building of the molecular models we have used crystal structures of known proteases available at the time of modeling as templates (HIV-1 (Mahalingam et al., 2002), HIV-2 (Tong et al., 1995), EIAV (Gutschina et al., 1996), FIV (Kervinen et al., 1998) and RSV (Wu et al., 1998) (Protein Data Bank accession codes (and resolutions): 1KIT (1.2 Å), 1IDA (1.7 Å), 1FMB (1.8 Å), 4FIV (1.8 Å), and 1BAI (2.4 Å), respectively). We generated three simultaneous models for each protease. RSV PR structure was used to interpret the AMV PR results because AMV and RSV proteases differ in only two residues, which are not expected to be involved in the enzyme-substrate interactions (Tőzsér et al., 1996). Homologous model of HFV was built from the crystal structure of HIV-1 PR complexed with an inhibitor (PDB code: 7HVP) by Modeller (Sali and Blundell, 1993). The program allows the use of multiple template crystal structures as an input and can create multiple homologous models as an output. A model of VSQNY↓PIVQ oligopeptide was docked into the substrate binding site of each retroviral protease model in forward and reverse directions, and a water molecule was also placed in a conserved position between the flaps of the protease and the substrate. The minimization and analysis procedure were applied as described previously (Bagossi et al., 2005) with the help of Sybyl program package (Tripos Inc., St. Loius, MO, USA) run on Silicon Graphics Fuel computer graphics system. Volume of the amino acid residues was retrieved from the literature (Zamyatin, 1972). The residues forming the subsites (not shown) were

predicted previously for HIV-1 and HIV-2 PRs (Tőzsér et al., 1992), EIAV (Weber et al., 1993), AMV (Tőzsér et al., 1996), MMLV (Boross et al., 1999), and HTLV-1 (Tőzsér et al., 2000) PRs and the corresponding residues in BLV, MMTV, MPMV, HFV, and WDSV PRs were obtained from the sequence alignment in Fig. 9 and verified using the crystal structures and homologous models.

4. RESULTS

4.1. CHARACTERIZATION OF HFV PROTEASES

4.1.1. Site-directed mutagenesis and molecular modeling of HFV PRs

Three-dimensional structure analysis of several retroviral PRs by either NMR or X-ray crystallography shows that despite large differences in the amino acid sequences, the global folds of these proteases are rather similar (Wlodawer et al., 1989; Jaskolski et al., 1990; Wlodawer and Gustchina 2000; Dunn et al., 2002). Many of the effects on the kinetics observed in the present study (by using various substitutions) are predictable or explainable by molecular modeling. Modeling is able to describe binding capabilities rather than catalytic efficiency. Mutant forms of human foamy virus (HFV) protease were designed (Q8R, H22L, S25T, T28D) based on the molecular model of HFV protease toward the classical retrovirus (e. g. HIV-1) consensus sequence to explore the role of these residues in the higher pH optimum and/or the lower dimer stability of HFV PR as compared to the classical retroviral PRs. We have examined the possible hydrogen bond forming residues in the catalytic aspartates of HIV-1 PR (see in Fig. 14).

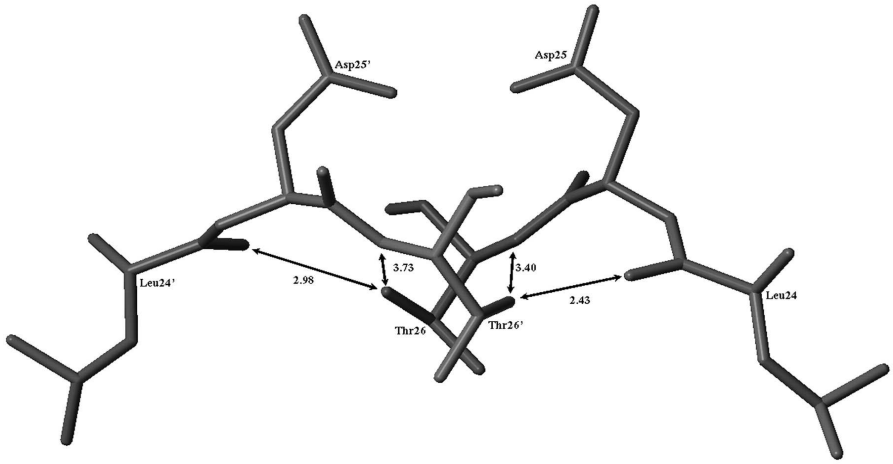


Figure 14: Hydrogen bonds around the catalytic aspartates (‘fireman's grip’) in the crystal structure of HIV-1 PR

Arrows indicate the possible hydrogen bonds, and the distances in Angstrom units are also provided.

In Fig. 9 it can be seen that several unusual residues can be found around the active site triplet of spumavirus enzyme: the residue corresponding to His22 is Leu in all other retroviral PRs (Leu23 in HIV-1 PR) except those of foamy viruses in which aromatic residue (Phe or Tyr) also

may be present. All retroviral proteases contain a hydrophobic patch around the catalytic aspartates that may help to isolate the catalytic residues from the aqueous environment for maintaining the proper catalytic power and it also may contribute to the interaction energy of the dimer by providing hydrophobic contacts between the monomers. As we see in Fig. 15 this hydrophobic cluster is formed by Leu10, Leu23, Ala28, Val82 and Ile84 residues, in case of HIV-1 PR.

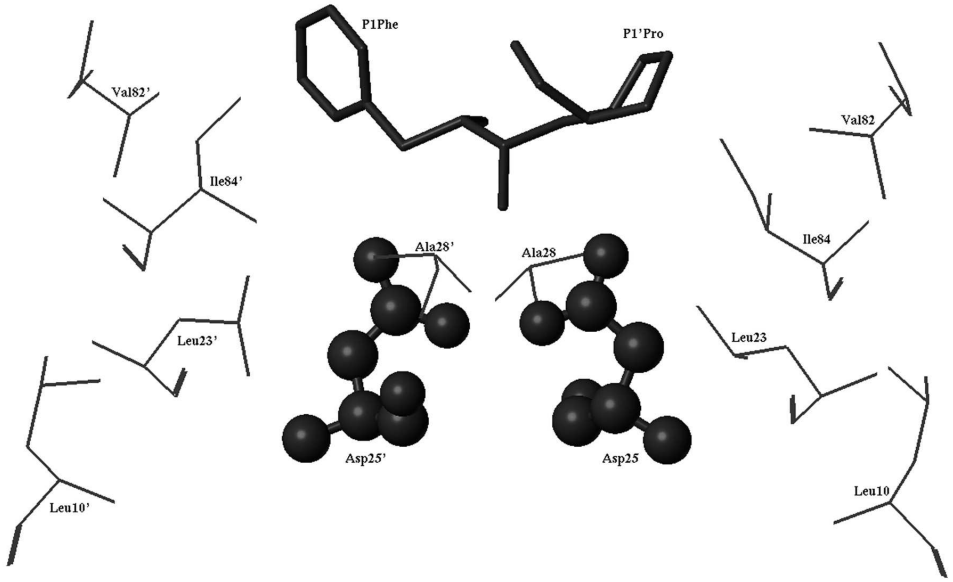


Figure 15: Hydrophobic residues being around the catalytic aspartates in the crystal structure of HIV-1 PR

Catalytic aspartates are shown with ball and stick representation. Hydrophobic residues around them are shown as lines (enzyme) or tubes (inhibitor).

These residues interact with the hydrophobic P1 and P1' side chains of a substrate or an inhibitor making hydrogen bonds. In case of HFV PR the hydrophobic cluster is made by amino acids with similar character (Leu10, Ala27, Val90, and Trp92) with one exception of the His22; its polar character may disturb the hydrophobic shell and also may decrease the interaction energy between the hydrophobic residues in the dimer interface. The hydrophobic side chains are placed in the protein interior, the polar side chains in locations more exposed to solvent.

The HFV PR contains a Ser in the active site triplet, similar to the PR of RSV, as compared with Thr in HIV-1 PR and most other retroviruses. Substitution of Ser to Thr in HIV-1 PR (T26S)

decreased the catalytic efficiency (Konvalinka et al., 1995a; Rose et al., 1995); but the reverse-substitution of Ser to Thr in case of *gag*-encoded avian retrovirus PR (S38T) substantially increased its activity (Arad et al., 1995). By comparing the kinetic values obtained for the HIV-1 PR mutant (T26S) (Konvalinka et al., 1995; Rose et al., 1995) and the values for linked wild-type and T26S/T26S HIV-1 homodimers (Bagossi et al., 1996), and supported by our molecular modeling studies, we have suggested that the PR containing Ser at the active site triplet instead of Thr may form less stable dimers (Bagossi et al., 1996).

The residue corresponding to Thr28 is Asp/Glu/Gln in the majority of other retroviral PR (Asp29 in HIV-1 PR). The electrostatic interaction between the Arg8 of the first monomer and the Asp29 of the second monomer of HIV-1 PR contributes significantly to the dimer stability of the enzyme (Lapatto et al., 1989; Loeb et al, 1989; Miller et al., 1989; Wlodawer et al., 1989; Wlodawer and Erickson, 1993); therefore the corresponding Gln8 and Thr28 residues were also selected for mutational analysis. We assumed that changing these residues to the charged ones can be found in HIV-1 PR may stabilize the dimer of HFV PR. The double mutant enzymes containing Gln8→Arg and Thr28→Asp mutations may completely regenerate the ion-pair as we see in Fig. 16, while Gln8 with Thr28 cannot form an ion-pair. To study the role of these residues in stabilization of the enzyme structure two single mutants (Q8R, T28D) and one double mutant were constructed (Q8R/T28D).

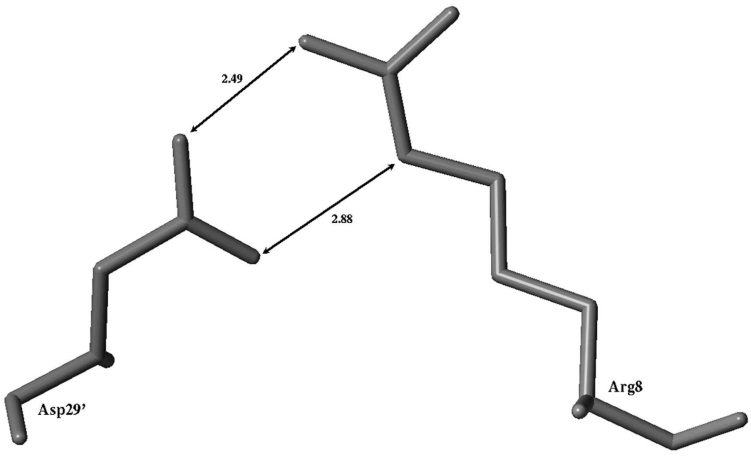


Figure 16: Ion-pair of Arg8 and Asp29′ in the crystal structure of HIV-1 PR

Arrows indicate the strong ionic/hydrogen-bond interaction, with distances in Angstrom units.

The five HFV PR mutants were chosen to conform to the sequence alignments of the two retroviral proteases (HIV-1 and HFV PRs); we replaced selected amino acid residues being close to the catalytic aspartates. The crystal structure of HIV-1 PR complexed with an inhibitor (PDB code: 7HVP) was the basis for the model of HFV PR by Modeller (Sali and Blundell, 1993). The program was used to build the initial model of all studied proteases. The amino acid sequence of HFV PR was aligned with the sequence of other retroviral proteases of known structure to determine the best starting structure for building the model. Numerous crystal structures of HIV-1 PR are available, for HFV PR no crystal structure exists. The HFV PR molecular modeling revealed 47% of sequence similarity with the HIV-1 PR in the substrate binding region (see in Table 5). The HFV and HIV-1 proteases have different lengths; HIV-1 PR is with 99 residues, while HFV PR has 125 residues (our fusion form is 141-amino-acid-long). In spite of the differences in length, we predicted that the foamy PR model share the conserved core region of HIV-1 PR (Figure 17).

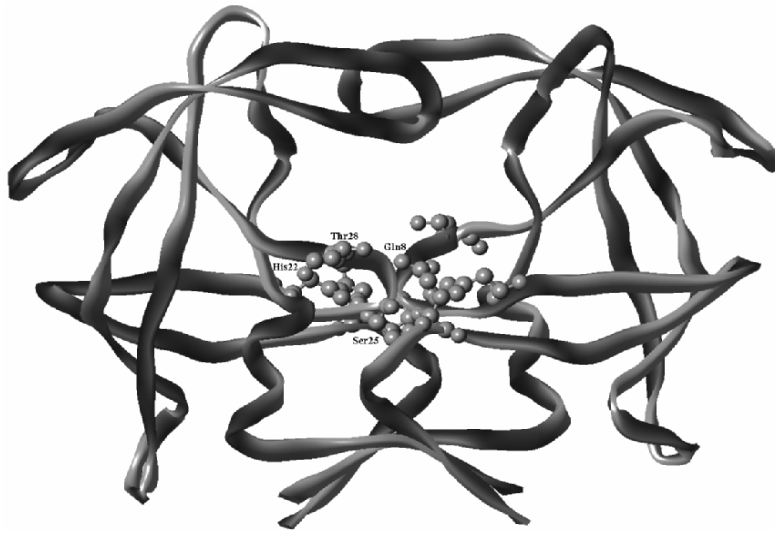


Figure 17: Molecular model of HFV PR

Ball and stick representation of the mutated residues in both monomers. The residues (Gln8, His22, Ser25 and Thr28), which were selected for mutagenesis study, are labeled only in the first monomer.

Sequence comparison of HIV-1 PR with HFV PR revealed a sequence identity of about 23% and similarity of about 30%. The extra amino acids of HFV PR may form longer loop structures on the surface of the molecule or may serve as a flexible linker between the PR and polymerase domains. The overall structure was expected to be similar to that of HIV-1 PR, but specific structural features of HFV PR remain unpredictable until a crystal structure of the foamy PR will be solved. It is possible

that the HFV PR structure significantly deviates from the known retroviral protease fold, which may cause, at least partially, the unusual features of the HFV PR. For example HTLV-1 PR crystal structure showed unexpected structural variations (Li et al., 2005).

The residues forming the subsites for Ser-Gln-Asn-Tyr↓Pro-Ile-Val-Gln were determined previously for HIV-1 PR (Tőzsér et al., 1992), and the corresponding residues in HFV PR were obtained from the sequence alignment and verified using the homologous model. The residues that are predicted to form the substrate binding sites in HFV PR are shown in Table 5.

Table 5: Comparison of residues forming the S1-S4 binding sites of two retroviral proteases

Amino acids [a] constructing the HIV-1 PR/HFV PR binding sites
S1 binding site: **Arg8/Gln8**, **Leu23/His22**, Asp25/Asp24, Gly27/Gly26, Gly49/Thr50, **Ile50/Ile51**, **Thr80/Thr88**, **Pro81/Asp89**, **Val82/Val90**, **Ile84/Trp92**
S2 binding site: **Ala28/Ala27**, **Asp30/Ile29**, **Val32/Cys31**, **Ile47/Ile48**, Gly48/Lys49, Gly49/Thr50, **Ile50/Ile51**, **Leu76/Tyr80**, **Ile84/Trp92**
S3 binding site: **Arg8/Gln8**, **Asp29/Thr28**, **Gly48/Lys49**
S4 binding site: **Asp30/Ile29**, **Met46/Leu47**, **Ile47/Ile48**, Gly48/Lys49, **Val56/Gln57**, **Gln58/Val59**, **Leu76/Tyr80**

[a] The amino acids that are different are shown in red. **Bold letters** show that side chain-side chain interaction may occur. Only the residues forming the S1-S4 subsites are given. Primed binding sites (like S1') have the same composition as the nonprimed one, but they are built from residues of the other subunit. The underlined letters are those which were substituted with the appropriate amino acids. Note that some amino acids can contribute to several subsites.

4.1.2. Purification of HFV PRs

Previous studies suggested that MBP is very effective to promote the solubility of polypeptides to which it is fused, compared to other commonly used proteins, like glutathione S-transferase and thioredoxin (Kapust and Waugh, 1999; Wang et al., 1999). Nevertheless, we previously observed that a substantial part of HFV-MBP fusion protein already formed aggregates, and purification of active unprocessed enzyme was unsuccessful in the absence of chaotropic agents, due to the increased aggregation. Previously, separation of the PR from the MBP was attempted by gel filtration in the presence of 4 M guanidine-HCl (or urea), but only residual activity was recovered (Fenyőfalvi et al., 1999). The loss of activity after processing from another type of HFV PR-fusion protein (containing

thioredoxin) has also been reported in the literature (Pfrepper et al., 1997). The latter expression/purification protocol allowed us to produce pure and active fusion enzyme. The fusion proteins were purified by affinity chromatography on amylose resin. Fractions showing high absorbance were collected and the fusion foamy virus protein was precipitated by ammonium-sulfate (4 M final concentration), on ice for 30 min. The precipitate was dissolved in lysis buffer immediately after incubation on ice because staying at -20°C made them less active. The purification protocol was successfully applied to the wild-type unprocessed enzymes and to the mutants, either (Figure 18).

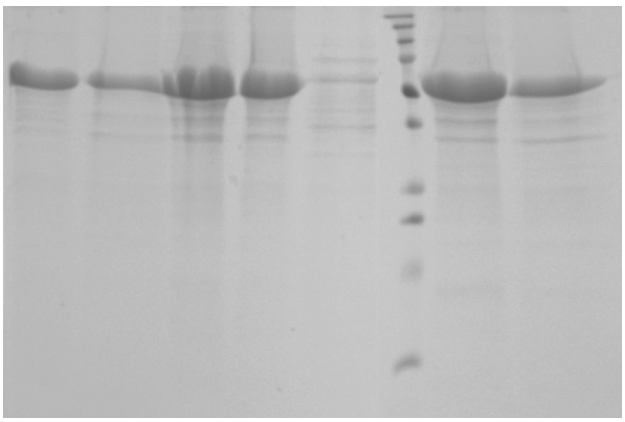


Figure 18: SDS-PAGE of purified wild-type and mutant HFV proteases

The wild-type (lane 1), H22L mutant (lane 2), Q8R mutant (lane 3), S25T mutant (lane 4), T28D mutant (lane 5), T28D/H22L mutant (lane 7), T28D/Q8R mutant (lane 8) HFV proteases were purified by affinity chromatography, than subjected to SDS-PAGE. Molecular weight marker (10, 15, 20, 25, 37, 50, 75, 100, 150, 250 kD) were used for comparison (lane 6).

4.1.3. Variation of catalytic constants between the mutants and wild-type foamy enzymes

It was previously found that the catalytic efficiencies (k_{cat}/K_m) for fusion and processed enzymes, for partially and completely purified enzymes are similar (Fenyőfalvi et al., 1999; Boross et al., 2006; Sperka et al., 2006).

The activity of mutant MBP-HFV PRs was compared with that obtained for the wild-type (wt) fusion enzyme (Table 6). The specificity constant of S25T mutant was the same as the wild-type, but the values of T28D, Q8R and H22L single mutants were 2.6, 2.7 and 10.8 times lower than that of the wild-type enzyme, respectively. The specificity constant of the double mutant: H22L/T28D

was also 4.2 times lower, but the corresponding value of Q8R/T28D mutant was 3.2 higher than that of wild-type fusion protein. This was mainly caused by the decreased K_m value, which may suggest that Arg-Asp ion-pair may stabilize the enzyme-substrate complex. There is a conserved ion-pair between the Asp29 from one monomer (HIV-1 PR numbering) and Arg 8' from the other monomer in most retroviral proteases and it contributes to the stability of the dimer (Lapatto et al., 1989; Loeb et al., 1989; Miller et al., 1989; Wlodawer et al., 1989; Wlodawer and Erickson, 1993; Manchester et al., 1994). However, these residues are Thr28 and Gln8 in HFV PR, respectively, and they cannot form an ion-pair. We assumed that changing these residues to the charged ones which can be found in HIV-1 PR may stabilize the dimer of HFV PR.

Table 6: Kinetic parameters determined for the wild-type and mutant HFV PRs in fusion with MBP for SRAVN↓TVTQS substrate

Enzyme	**K_m (mM)**	**k_{cat} (s^{-1})**	**k_{cat}/K_m ($mM^{-1}s^{-1}$)**
Wild-type	0.41±0.02	0.0053±0.0001	0.013±0.001
Q8R	1.3±0.3	0.0061±0.0008	0.0047±0.0012
H22L	-	-	0.0012±0.0001
S25T	0.26±0.04	0.0031±0.0001	0.012±0.002
T28D	0.35±0.06	0.0017±0.0001	0.0049±0.0009
Q8R/T28D	0.17±0.02	0.0071±0.0002	0.042±0.005
H22L/T28D	-	-	0.0031±0.0001

Most of the mutants showed k_{cat} values close to the wild-type value, implying that the folding capability of these mutants is likely similar to that of the wild-type. Unfortunately, there is no tight-binding inhibitor for HFV PR suitable for active site titration. Therefore, enzyme concentrations were determined by protein concentration measurements (Bradford method). These enzyme concentrations were used for calculation of the apparent k_{cat} values from the experimentally measured V_{max} values. The small variation of the appearent k_{cat} values of the mutant enzymes compared with that of the wild-type enzyme may reflect small variation of the true catalytic constant and in parallel small variation in the concentration of the active, correctly folded enzyme. The other possibility to get unchanged V_{max} value requires an increased "true" catalytic constant and a decreased folding capability or *vice versa*, which is unlikely.

4.1.4. pH optimum of HFV PRs

We have determined the pH profile of the wild-type and mutant fusion enzymes (Table 7). The lowest pH optimum was found to be about 6.2 in the case of Q8R mutant, while the highest was about 6.8 for T28D mutant. The pH optimum of HFV PRs is near neutrality.

Table 7: pK_1, pK_2 and pH optimum values for wild-type and mutant HFV PRs determined for MBP fusion proteins using substrate SRAVN↓TVTQS

Enzyme	pK_1	pK_2	pH optimum
Wild-type	5.61±0.16	7.00±0.15	6.30±0.11
Q8R	5.71±0.26	6.70±0.24	6.21±0.18
H22L	5.64±0.24	6.83±0.23	6.24±0.17
S25T	5.49±0.15	7.81±0.18	6.65±0.12
T28D	5.79±0.14	7.82±0.16	6.81±0.11
Q8R/T28D	5.59±0.15	7.58±0.17	6.59±0.11
H22L/T28D	6.01±0.29	6.98±0.29	6.50±0.20

These values are slightly higher than that of other retroviral enzymes (which all show maximal activity at more acidic pH) and might indicate that some cleavage events are delayed and occur in the infected cells (Flügel and Pfrepper, 2003). The measured pH optimum of H22L mutant was close to that of wild-type HFV PR. It is interesting to note that mutation of Ser to Thr in the active site triplet increased the pH optimum. However, the urea stability of S25T mutant increased at higher pH. The increased pH optimum might be a consequence of increased dimer stability (Table 8). The dimerization of HIV-1 PR is strongly dependent on the pH, forming less stable dimers at higher pH (Darke et al., 1994; Szeltner and Polgár, 1996). The pH optimum of all of the foamy proteases was nearly the same, close to neutrality. These features may be related to the distinct replication cycle of spumaviruses in cellular compartments (endoplasmic reticulum).

4.1.5. Urea denaturation of HFV PRs

Dimer stability of the wild-type and mutant forms of the enzymes have been compared by measuring their urea denaturation curves at two pH values (6.0 and 7.2) (Table 8). These values were located at the two sides of the bell-shaped curves on the activity versus pH plots. To get a

more profound effect, the pH values were chosen to have the highest possible distance between them allowing measurable activity on all mutant forms of the enzyme at both pH. The urea concentration leading to 50% loss in enzymatic activity ($D_{1/2}$ value) was the lowest in the case of wild-type, while S25T and T28D mutants were less sensitive against the effect of urea at both pH values. Dimer stability was increased by the Ser25 to Thr mutation (Table 8), in good agreement with the suggested role of this residue in dimerization. It was suggested that the threonine residue has an important role in "fireman's grip" formation between the two subunits of the homodimeric retroviral enzyme. A detailed study on the role of the "fireman's grip" in HIV-1 PR demonstrated that is crucial for stabilization of the PR dimer and for overall stability of the enzyme, but it is not absolutely required for activity (Strisovsky et al., 2000; Ingr et al., 2003). It was also shown that dimer stabilities of the Thr-containing HIV-1 and MAV proteases were about one order of magnitude higher than the corresponding Ser-containing mutant forms (Ingr et al., 2003).

The urea stability of the mutant H22L was increased at both pH, as expected. The T28D and the Q8R single mutants, constructed regarding the conserved amino acids in HIV-1 PR sequence at the same positions, showed higher stability against urea than the wild-type protease (Table 8).

Table 8: Dimer stability parameters for wild-type and mutant HFV PRs determined using MBP fusion proteins

Enzyme	**pH= 6.0 $D_{1/2}$ (M)**	**pH=7.2 $D_{1/2}$ (M)**
Wild-type	0.55±0.07	0.57±0.13
Q8R	0.61±0.12	0.73±0.17
H22L	0.69±0.14	0.79±0.13
S25T	0.83±0.11	1.15±0.11
T28D	0.86±0.11	0.92±0.13
Q8R/T28D	0.80±0.17	0.86±0.11
H22L/T28D	0.75±0.14	0.82±0.14

$D_{1/2}$ value for HIV-1 PR: 1.85 M (Wondrak et al., 1996)

We expected only marginal effect in the case of single mutants, when only half of the ion-pair was regenerated. T28D mutant was more stable than the double mutant Q8R/T28D enzyme, which suggests that other structural features may also play a role. We expected a more pronounced

effect in the case of double mutant. The stability values of these enzymes were separated into 2 groups: low stability group consisted of the Q8R mutant and the wild-type enzymes, while T28D and Q8R/T28D mutants had higher $D_{1/2}$ values at both pH values.

The stability of the enzymes correlated with the hydrogen-bond forming capability of these residue pairs. Only one hydrogen-bond can be formed between Gln-Thr and Arg-Thr residue pairs in contrast to Gln-Asp and Arg-Asp pairs, where two hydrogen bonds can be formed. While the wild-type HFV PR had the same sensitivity against urea at both pH values, mutant enzymes showed higher sensitivity against urea at pH 6.0 than at pH 7.2. Mutants with increased stability relative to the wild-type are likely to show greater catalytic activity.

Enzyme activity of the wild-type and the mutant proteins versus urea concentration and pH curves showed an increased stability and pH optimum of most mutants as compared to the wild-type HFV PR. These results suggests that residues found in the vicinity of the catalytic aspartates do not directly influence the pH optimum of HFV PR, but they act through the influence of the dimer stability of the enzyme.

4.2. SUBSTRATE SPECIFICITY OF RETROVIRAL PROTEASES

Previously, members of our laboratory and collaborators have characterized the substrate specificity of HIV-1, HIV-2, EIAV, and AMV proteases using an oligopeptide substrate set based on the naturally occurring type 1 cleavage site between MA and CA proteins of HIV-1 (Tőzsér et al., 1992; Tőzsér et al., 1996; Weber et al., 1993). These previous results were extended by the present study with HTLV-1, BLV, MPMV, MMTV, MMLV, WDSV and HFV proteinases, and in this way each genus of retroviruses was represented by at least one member. HTLV-1 and HFV PRs were not able to hydrolyze the unmodified peptide. HFV PR didn't hydrolyze the substituted peptides either, with the exception of P3-Val peptide; therefore, this enzyme was omitted from further analysis. We completed these studies with the investigation of the specificity of P1, P3 and P4 positions of the substrate using the eleven proteases. This work together with a previous one (Bagossi et al., 2005) took the advantage that the different retroviral proteases were mapped with the same peptide series in the same reaction conditions, in the same laboratory, therefore the results are easy to compare. Our previous studies established a strong correlation between the measured relative activities and the specificity constants (Tőzsér et al., 1996), therefore the relative activities determined was considered as measures of the specificity constants in this study.

4.2.1. Test of P1 specificity

Comparison of the specificity of divergent members of retroviral proteases using the type 1 MA/CA substrate series suggested that these PRs have many common features. All prefer hydrophobic residues at the P1 position, although the optimal size of the residue may depend on the residues forming S1 subsite and may also be a function of the residue at P3. Analysis of the amino acids at the P1 position of retroviral processing sites indicated Leu, Phe, Tyr, and Met being the most common preferred amino acid residues, suggesting the best fit of these to the sequence requirements of the P1 amino acid. Thus, the specificity of the amino acid side chain in the P1 position is best described as large, hydrophobic, and unbranched at the beta-carbon. The absence of the hydrophobic amino acids Ile and Val from the P1 position of cleavage sites suggested that the structure of the P1 amino acid side-chain at the beta-carbon is an important determinant of cleavage site specificity (Pettit et al., 1991). Curiously, while Leu appeared over 20% of the time (Table 1), the similar amino acids Val and Ile did not appear at all in case of HIV-1 PR.

Results with the set of relative activities obtained with the P1-modified substrates are provided in the Fig. 19. The variation of relative activities of the hydrolyzed P1 substituted peptides was surprisingly small. When the specificity of the S1 subsites of the proteases was studied, with the exception of BLV PR, the other enzymes showed P1 preference for aromatic residues (Phe, Tyr) which suggest that the size and hydrophobic nature of the S1 site is well conserved among retroviral proteases (Fig. 20). Nevertheless, the S1 binding site of the PRs appears to be a hydrophobic one in all studied cases. The Phe substitution of Tyr provided better specificity values for all enzymes, which might be due to the presence of the -OH group in Tyr aromatic ring, which can fit less in S1 subsite. Based on the molecular models it is predicted, that the S1 sites, being very close to the site of cleavage, should be filled by a hydrophobic side chain to obtain efficient cleavage. However, there are some fine specificity distinctions, in terms of whether the enzymes would also favor smaller residues at this position, especially Leu or Met. This creates a subgroup of the proteases, including MMLV, WDSV and BLV proteases; the last one showed the highest preference for Leu side chain at this position (Eizert et al., 2008). Nevertheless, the S1 binding site of the PRs appears to be a hydrophobic one in all studied cases.

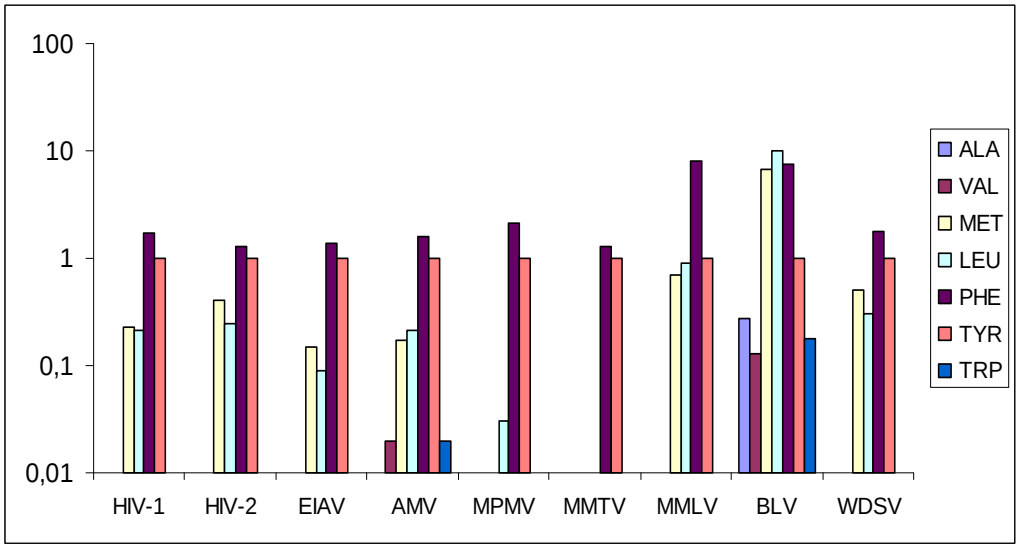


Figure 19: Comparison of relative activities obtained on Val-Ser-Gln-Asn-Xaa↓Pro-Ile-Val-Gln peptides for retroviral proteases

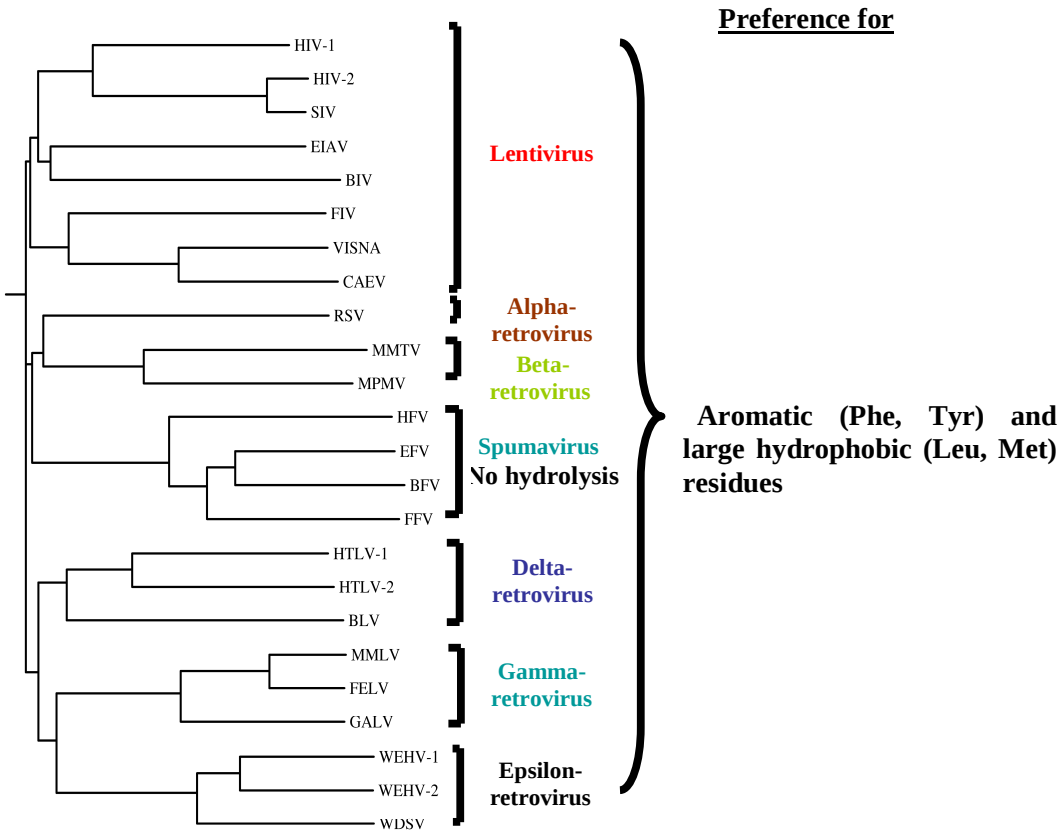


Figure 20: Phylogenetic tree of the retroviral proteases and specificity for P1 site

4.2.2. Test of P3 specificity

Subsite S3 is more open than S2 and can accept a variety of residues. Specificity in S3 is a function of the P1 residue: a large P1 side chain restricts the size of the P3 residue which can be accommodated (Tőzsér et al., 1996). Unlike the specificity of the S1 site, various residues were observed as the preferred ones when the S3 binding sites were mapped (Figure 21). The alpha- and betaretrovirus proteases preferred large hydrophobic residues, such as Phe and Leu at this position, similar to S1, while smaller hydrophobic residues, such as Val or polar residue such as Gln were preferred by the other enzymes.

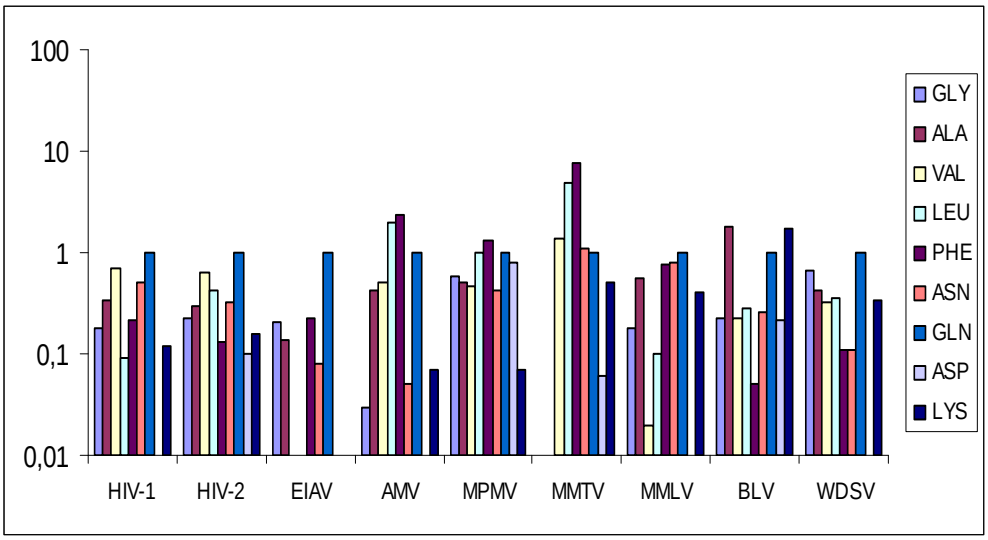


Figure 21: Comparison of relative activities obtained on Val-Ser-Xaa-Asn-Tyr↓Pro-Ile-Val-Gln peptides for retroviral proteases

The specificity at S3 also appears to correlate with the phylogenetic tree (Fig. 22): lentiviral proteases also showed a preference for the original polar residue (Gln) together with the preference for large hydrophobic residues, while MMLV and WDSW proteases were those preferring smaller hydrophobic or polar residues (Ala, Gln, Asn). BLV PR preferred small Ala and the polar Lys and WDSV PR the Gly at P3 in this substrate. Nevertheless, the size of the residue appears to be the main specificity determinant at this position. In this set of substituted peptides HFV PR was able to hydrolyze just one of them, the one containing Val at P3 position, unlike the case of the P2-modified substrates where many of the substitutions formed substrates of the enzyme (Bagossi et al., 2005). It should be mentioned that one of the deltaretroviral proteases, that of HTLV-1, was not

able to hydrolyze any of these peptides, in agreement with the lack of a type 1 cleavage site for HTLV-1 (see Table 2).

Crystal structures and molecular models of PRs suggested that S3 subsites are generally large. These pockets can accept various side chains, and the variation of activity is relatively small as compared to neighbourhood subsites, S4 and S2. The S3 subsite is similar to S4 in being partly exposed to solvent at the surface of the enzyme. Consequently, the P3 side chain may be positioned to interact with the hydrophobic internal residues of the enzyme.

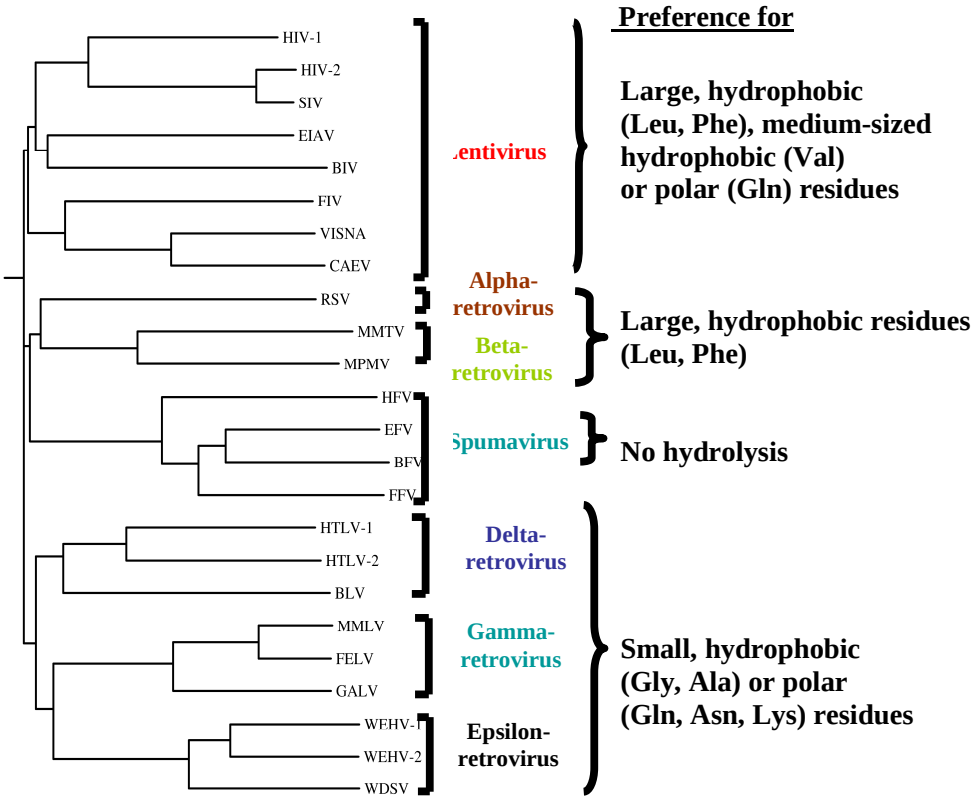


Figure 22: Phylogenetic tree of the retroviral proteases and the distinct specificity subgroups for P3 site

4.2.3. Test of P4 specificity

For the analysis of P4 variants, a different peptide (P4 Val) is chosen as the standard for measurement of relative activity to allow the inclusion of HTLV-1 PR, which can not cleave VSQNY↓PIVQ. Replacement of P4 Ser in the original peptide with different amino acids resulted in peptides which were more or less hydrolysable by the studied proteases, except for the HFV PR

which was not able to hydrolyze any of the peptides. Furthermore, different degree of selectivity was observed among the enzymes, for example proteases of deltaretroviruses (HTLV-1 and BLV PRs) were the most restrictive (4-6 noncleavable peptides), while WDSV PR showed low degree of variation (four fold) between the best and the worst substrates (Figure 23).

The complete mapping of P4 is provided in the Figure 23. Similarly to S3, various residues were found to fit preferably to the S4 sites of the PRs. In some cases small, even polar residues were preferred, as in case of primate lentiviral PRs, like HIV-1 PR, preferring Gly and Ser. In contrast, hydrophobic residues at P4 form better substrates for the non-primate lentiviral EIAV than for HIV PRs, due to the presence of the additional flap residues 50-52 that contribute to the S4 subsite.

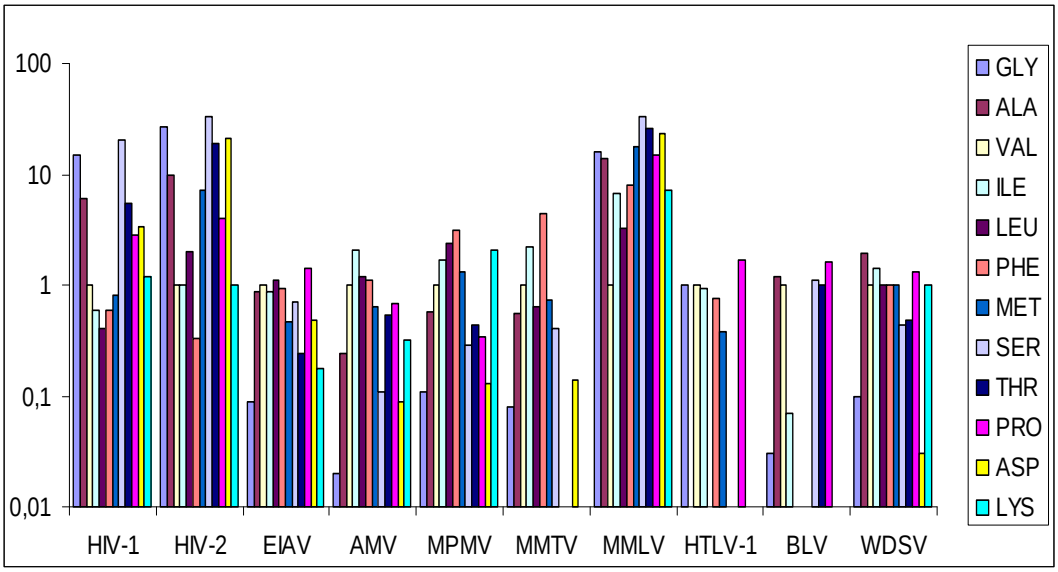


Figure 23: Comparison of relative activities obtained on Val-Xaa-Gln-Asn-Tyr↓Pro-Ile-Val-Gln peptides for retroviral proteases

Other enzymes were not very selective, able to accept various types of hydrophobic or polar residue. The preferred size of P4 is different: for AMV PR Ile substitution gave the best result, while for MPMV PR Phe substitution worked most efficiently. The best values were obtained with the unsubstituted substrate containing Asn and with the peptide containing Gly substitution. There are no well-defined pockets as compared to the internal binding sites.

The S4 subsite of retroviral proteases is close to the surface, the side chain may be partially exposed to the solvent. A side chain at P4 may either interact with the solvent molecules or may form interactions with residues of the binding pockets. A unique feature of the HIV-1 and other primate lentiviral proteases is that they contain Asp30, while most of the retroviral PRs contain uncharged residues at the equivalent position (eg. BLV: Asn, HTLV-1: Met, WDSV: Leu, FVs: Ile). It is interesting to note that the Asp30 to Asn mutation occurs frequently in HIV-1 PR with drug resistance to PR inhibitors.

However, unlike the most retroviral PRs, HIV-1 PR preferred the more hydrophilic residues Ser (and Thr) in this position. Nevertheless, the size of the preferred residue is a function of P2. As we know the largest variation of relative activities were obtained with the P2 substituted peptides (Bagossi et al., 2005).

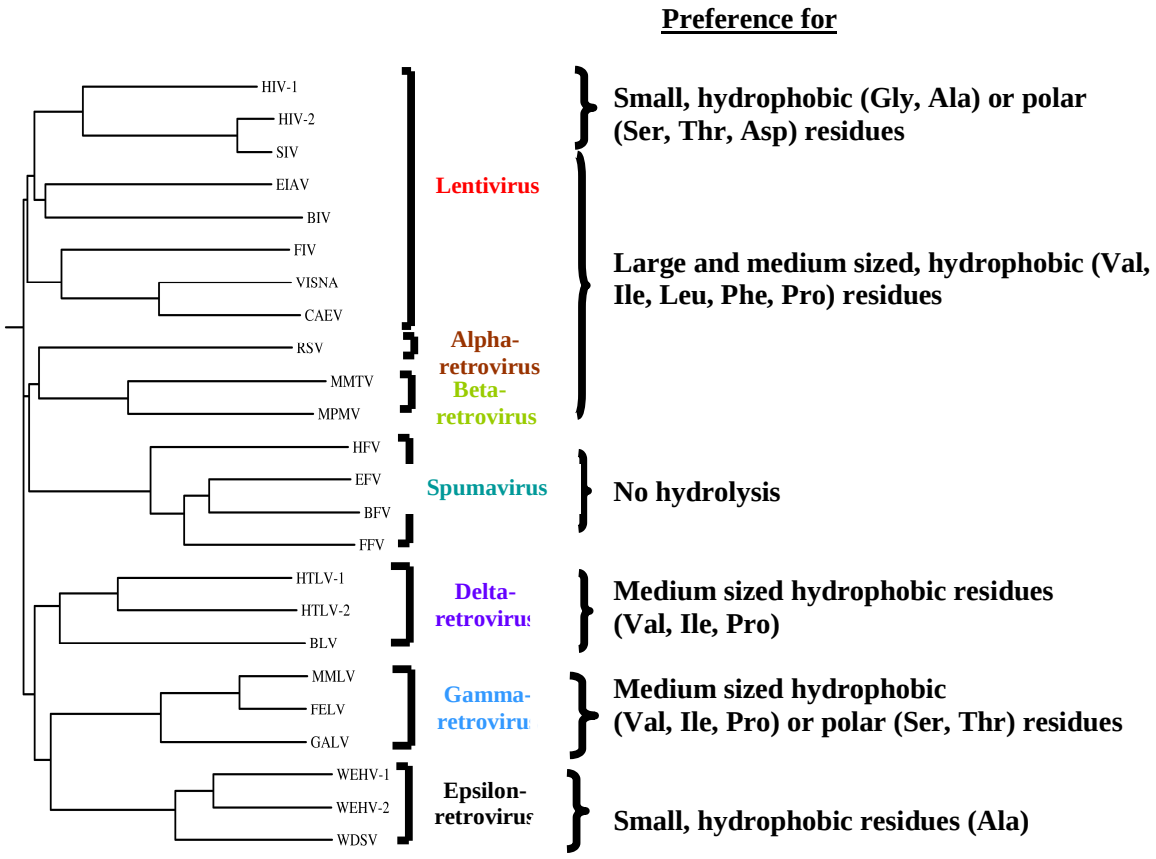


Figure 24: Phylogenetic tree of the retroviral proteases and the distinct specificity subgroups for P4 site

5. DISCUSSION

The virus family of *Retroviridae* consists of the two subfamilies *Orthoretrovirinae* and *Spumaretrovirinae* or FVs (foamy viruses). This discrimination is due to discovery of major differences in the replication strategies between the two subfamilies (Rethwilm 2003). These differences have been highlighted in a variety of reviews (Linial 1999; Rethwilm 2005; Linial 2007). Studying foamy viruses is fascinating because these viruses appear to do everything that is common to all other retroviruses differently. FVs have found a completely new way to propagate their genome, which show a surprisingly high evolutionary conservation (Rethwilm 2010). They do this extremely successfully because they are likely to be infectious non-pathogenically. The low catalytic efficiency of the HFV protease is in line with the lower specific activities of the HFV RT and IN as compared to those of other retroviruses (Pahl and Flugel, 1993; Kogel et al., 1995). As opposed to other retroviruses, in which the Gag-Pol fusion protein is present only in relatively smaller amount as compared to the Gag proteins, a significant amount of Pol is produced by HFV infected cells (Löchelt and Flügel, 1996), which may be required to have larger total concentrations of the active HFV enzymes. In *Orthoretrovirinae* the PR is created by autoprocessing of the Gag-Pol precursor protein and is subsequently present as a separate enzyme (Oroszlán and Luftig, 1990; Dunn et al., 2002). In contrast, the PR domain in FVs is not cleaved off from the RT at the N-terminus of the Pol precursor protein. Only the C-terminal IN is removed from Pol, thus leading to a mature PR-RT enzyme (Cartelierri et al., 2005). It was demonstrated that a PRshort (which is the separate PR domain) of simian foamy virus from macaques (SFVmac), as well as the full length PR-RT, exhibit proteolytic activity (Hartl et al., 2008, 2010). How FV PRs are activated is still unknown. In terms of its biochemical properties FV PR appears to be less active and characterized by lower processivity than, for example HIV-1 PR; HFV PR is able to process only two natural cleavage sites (Flügel and Pfrepper 2003) (Table 9), while HIV-1 PR cleaves at thirteen sites (Tőzsér and Oroszlán, 2003). The cleavage site junctions are foamy virus-specific and quite different from the corresponding sites of other retroviral PR. They surprisingly resemble the cleavage sites of the yeast retrotransposons Ty1 and Ty3 PRs, as evidenced by the hydrophobicity profiles (Merkulov et al., 1996). In line with these similarities, HFV PR was able to hydrolyse some Ty1 and Ty3 site-representing peptides (Fenyőfalvi et al., 1999); the specificity of HFV PR appeared to be more related to retrotransposon proteases than to other retroviral proteases. In this aspect, it is important

to note, that HFV is capable of efficient intracellular retrotransposition (Heinkelein et al., 2000; Heinkelein et al., 2003).

The wild-type HFV PR showed equally high activities on substrates containing Ala, Val or Cys at the P2 site (Bagossi et al., 2005), but didn't hydrolyse peptides with larger P2 residues or the original peptide (SP211). That is because the P1 position of spumaviral cleavage sites had a marked preference for hydrophilic residues (Flügel and Pfrepper, 2003), in contrast to those of most retrovirus PRs. The P1 and P4 substituted versions of HIV-1 MA/CA substrate were not a good substrate for HFV proteases even at 24 h incubation. P3 Val and P2 Val (Bagossi et al., 2005) were cleaved by foamy protease. The predominant residues at the foamy virus cleavage sites are (at the P2 and P2' positions) Val or Ile (Pfrepper et al., 1999). The S2 site of foamy enzyme appeared to be one of the smallest ones among the studied eleven retroviral proteases (Bagossi et al., 2005). At most (P3, P4, and P4') positions a relatively high variability of the flanking residues is observed (Flügel and Pfrepper, 2003).

Table 9: Naturally occurring cleavage sites in HFV Gag and Pol polyproteins

Gag↓p4	SRAVN↓TVTQS
RT↓IN	SYVVN↓CNTKK

(Based on: Flügel and Pfrepper, 2003)

HFV PR not only has an interesting specificity for substrates, but also its dimer stability is lower, pH optimum is higher than the HIV-1 PR and other retroviral proteases (Fenyőfalvi et al., 1999; Boross et al., 2006). The HFV PR contains a Ser in the active site triplet, similar to the proteases of alpharetroviruses (AMV), as compared with Thr in the PR of HIV and most other retroviruses. The specificity constant of HFV PR did not alter when Ser25 was changed to Thr (as in case of HIV-1 PR), hence the presence of Ser in the active site triplet does not appear to be an important determinant for the low catalytic efficiency of the enzyme (Boross et al., 2006). A selection of the mutants was made to study and explore the potential contribution of the given residues to the unusual parameters. To further understand the specificity of HFV PR, and its tolerance to mutations, several residues surrounding the active site were mutated to the corresponding HIV-1 PR residue, individually or in combination. Our results showed that several features of HFV PR were substantially different from those of other retroviral proteases. The pH

optimum of wild-type HFV PR was higher than that of HIV-1 PR. All mutants had the same or higher pH optimum than that of wild-type HFV PR. Interestingly, the studied mutants showed the same or higher stability against urea at both pH values, suggesting that during evolution HFV PR did not evolve to maximize the dimerization energy, as compared with HIV-1 PR. While the wild-type HFV PR had the same sensitivity against urea at both pH values, mutant enzymes showed higher sensitivity against urea at pH 6.0 than at pH 7.2. The overall results of these mutational studies suggest that requirements of HFV PR structure may differ from that of other retroviral protease structures, in response to a different selective pressure caused by the different replication strategy of foamy viruses. Since it was shown that the ProPol polyprotein is not efficiently cleaved between PR and RT (Flügel and Pfrepper, 2003), it is also possible that C-terminal flanking sequences, including the RT of Pol polyprotein, may modify the features of HFV PR, for example by providing additional dimerization interfaces, which are not provided after maturation of other viral proteases (Sperka et al., 2006).

There is data of a SFVmac (macaque) PRshort NMR structure and dynamics model published by Hartl et al. (2008). Mason-Pfizer monkey virus PR also exists as monomer in solution (Veverka et al., 2001, 2003). For this retroviral enzyme, the formation of an intra-molecular disulfide bridge between a cysteine near the N-terminus and a second one near the C-terminus is proposed to be the activating mechanism for dimerization (Veverka et al., 2003; Zabranska et al., 2007). The monomeric state of SFVmac PR may be the key in the regulation of the viral life cycle. How SFVmac PR is activated is still an unresolved issue, and viral or even cellular factors may be necessary for PR dimerization and activation. The analysis show that SFVmac PR lacks several important structural features necessary for dimer formation as observed for other retroviral proteases.

Interestingly enough, phylogenetic analysis showed closer relationship of foamy virus proteases to cellular aspartyl proteases, than any other group of retroviral protease family. Aspartyl proteases are usually active at acidic pH, since the catalitically competent enzyme has one protonated and one deprotonated Asp at the active site (Fig. 7), and this can be typically achieved at an acidic pH. It seems that the fold of cellular aspartyl protease allows broader pH optimum range (from 2-4 of pepsins to 6.5-7.5 of renin) than the retroviral protease fold (4.5-7.0). Suprisingly, similar trends in the substrate specificity can be seen for both the cellular and the retroviral aspartyl proteases: for example

gastric pepsins have broad specificity to digest several proteins in the stomach, while human renin has very narrow specificity to cleave only one peptide bond in angiotensinogen in the physiological pH.

It is important to note that while HIV has a high mutation rate, HFV do not utilize the error-prone reverse transcriptase to generate substantial sequence diversity, since the foamy viruses are able to reverse transcribe their genome late in infection before the virus buds from cellular membranes (Heinkelein et al., 2003; Roy et al., 2003). As a consequence, HFV has not undergone the extensive mutational changes that might have optimized the HIV-1 PR for both high catalytic efficiency as well as flexibility in tolerating mutations under selective pressure.

Proteases of each of the retrovirus subgroup possess a set of subsites with variable specificity. However, many times PRs belonging to viruses which are close in the phylogenetic tree appear to have common specificity determinants. For example, EIAV, AMV and MMTV share very similar specificity at S1, S2 (Bagossi et al., 2005) and S4 subsites. Furthermore, the specificity of primate lentiviral proteases from which HIV-1 and HIV-2 were tested in this study, appear to possess a unique set of subsite specificity.

We have mapped the specificity of retroviral PRs using a series of oligopeptides having amino acid substitutions in the P1, P3, and P4 positions. HFV PR was unable to hydrolyze these peptides, except for one (P3 Val). A schematic representation of the specificities is provided in Fig. 25 as a summary. The primate lentiviral HIV-1 and HIV-2 PRs have a substantially different specificity relative to that of EIAV PR, in the hydrophobicity of the S2 and S4 subsites as well as the size of S4 subsite. The alpharetroviral AMV and betaretroviral MMTV PRs appear to show very similar specificity. All their substrate binding sites are hydrophobic and large, except for the small S2 pocket. The deltaretroviral BLV PR and HTLV PR have large hydrophobic S1 and S2 pockets and smaller S3 and S4 subsites. Only the S3 subsite in case of BLV has some polar character. The gammaretroviral MMLV PR has similar characteristics to that of the deltaviruses proteases, with the exception of the S4 susbite that appears to be small and hydrophilic. The epsilonretroviral WDSV PR has a similar specificity to that of the deltaretroviral PRs, except that its S2 and S3 are more hydrophilic. Distant subsites (like S4) in HIV-1 PR and other retroviral enzymes demonstrate higher variability than the inner subsites (S1, S2, S3) that are located closer to the cleavage site (more internal and well-defined pockets). As we go further from the cleavage site the preference of the substrate side chains looks wider and the specificity grade is narrower. In conclusion, the specificity patterns of the subsites agree with the evolutionary relationship among the PRs as

represented by the phylogenetic tree (Fig. 20, 22 and 24). Comparative studies of retroviral proteases indicated that the enzymes retained common core specificity (Tőzsér 2010).

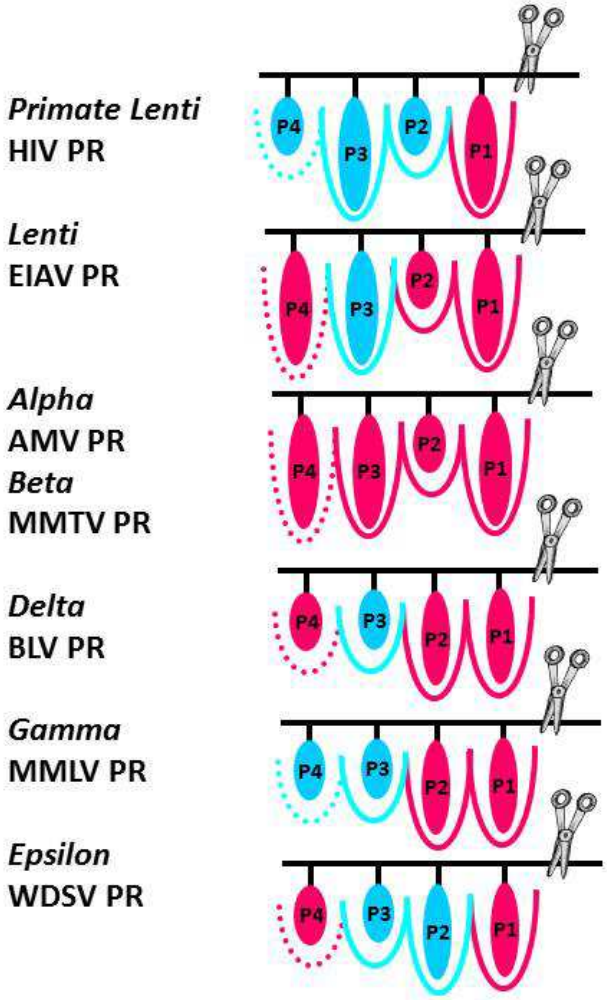


Figure 25: Schematic representation of the preferred residues of the S4 to S1 subsites of representative retroviral proteases

The size of the oval objects represents the substrate amino acid side chains. These approximate the size of the most-preferred residues. Red objects represent sites that prefer predominantly hydrophobic residues, while blue objects represent sites that do not discriminate based on hydrophobicity. Dashed lines for the P4 subsites indicate that these pockets are less defined than the other ones, due to their proximity to the protein surface.

Naturally occurring type 1 cleavage site sequences of the studied retroviruses have P1-P4 residues that are in good agreement with the findings of the S1-S4 mapping study. For example the best P1 residue (Tyr, Phe) in the mapping study of proteases appears naturally in type 1 cleavage sites of all retroviral proteases having type 1 sites (Table 2). The type 1 cleavage site is important for several reasons. No other protease, except pepsin, is known to act at the imino side of a Pro residue. Proline residues, especially after Tyr or Phe in the sequence (as the case in type 1 cleavage sites) have a relatively high probability of forming the *cis* isomer rather than the *trans* isomer of the preceding peptide bond (MacArthur and Thornton, 1991). Conformational selectivity of the HIV-1 PR towards the *trans* isomer of the cleaved peptide bond was demonstrated by NMR and kinetic studies (McCornack et al., 1997; Vance et al., 1997). In P3 position various amino acids are present in the cleavage sites, which correlate more or less to the mapping studies of proteases. In P4 position Ser and Thr are the most frequent amino acids in the cleavage sites (almost 50%). Discrepancies between the most preferred P3/P4 residues and those observed in the naturally occurring cleavage sites might be due to the different sequence contexts, which have been shown to have a profound effect on the subsite preference of HIV-1 PR, as well as to possibility that not all retroviral cleavage sites are optimized evolutionary for rapid processing (Fehér et al., 2002). We must note that residues forming the substrate binding sites are substantially more conserved in comparison with the full length of retroviral protease sequences. This could be a consequence of the selective evolutionary pressure to maintain these residues compared to those that are less important for the structure and activity of the enzyme.

The 3D structures of seven proteases from HIV-1, HIV-2, SIV, RSV, FIV, EIAV and HTLV-1 were available at the time of our study. The alignment of the primary and secondary structures of all retroviral proteases suggested a single domain of the cellular aspartic proteases. Comparison of the predicted PR structures of HIV-1, FIV and EIAV with crystal structures proved that the models were basically correct in prediction of the substrate binding sites (Weber 1991; Wlodawer et al., 1995; Gustchina et al., 1996). Therefore, the molecular modeling of proteases can serve as an important tool in the absence of crystal structures. Based on the molecular models, the S1 binding sites of RPs are usually bulky hydrophobic pockets. The S1 subsite is large, hydrophobic, and well conserved among retroviral PRs, which is in agreement with the results. Most of the proteases have a hydrophobic environment with Ile, Val and Leu amino acids on the tip of the flaps. But BLV and MMLV proteases have the small Ala, which explain the large and

hydrophobic P1 Leu, Phe and Tyr preferences at this subsite. Models for the S3 binding sites are usually large hydrophobic residues, but unlike S1 subsite, various residues were preferred by different retroviral proteases. With the exception of WDSV PR model, where the S3 subsite binding pocket was predicted to be relatively large and less suitable for small residues at P3 (Gly is preferred at P3), the other molecular models correlated well with the results. The S4 subsite lies at the protease surface and lacks the well-defined pockets.

It is of interest to know, that HIV-1 inhibitors typically do not inhibit the other retroviral proteases, with exception of MMLV PR which was found to be inhibited well (Fehér et al., 2006; Sperka et al., 2007). So there is an apparent contradiction between the somewhat conserved specificity and almost complete lack of inhibition which might be due to the fact that inhibitors are typically rigid, docking molecules, while substrates are more adaptable, flexible structures. The same phenomenon appears to be also critical in drug resistance.

6. SUMMARY

During my thesis work I had the opportunity to study retroviral proteases (HIV-1, HIV-2, EIAV, AMV, MMLV, MMPV, MMTV, BLV, HLTV-1, WDSV and HFV) and to compare their substrate specificity. HFV PR showed a low catalytic activity on type 1 substrates; just one peptide was cleaved (P3 Val). We have characterized some foamy PR mutants by constructing them based on the conserved amino acids in HIV-1 PR sequence at the same positions. Mutations were made in the vicinity of the catalytic aspartates of HFV PR. We built a molecular model for HFV PR. The mutations in HFV PR resulted in wild-type-like or even higher pH optimum. Similar results were found for stability against urea at both pH values studied (6.0 and 7.2). HFV PR showed not to be as sensitive towards mutations as other retroviral proteases, especially HIV-1 PR. Our mutational results suggest that requirements of FV PR structure may differ from that of other retroviral protease structures. It is possible that during evolution FV PR did not evolve to maximize the dimerization energy, as compared with HIV-1 PR. Knowing better the spumavirus enzymes and their replication strategy will help in the development and application of retroviral vectors (based on this non-pathogenic virus) in gene therapy.

We have examined the ability of 34 oligopeptides with single amino acid substitutions in the P1, P3, and P4 positions of the HIV-1 Gag cleavage site (MA/CA: Val-Ser-Gln-Asn-Tyr↓Pro-Ile-Val-Gln) to support cleavage by the mentioned retroviral PRs. The specificity of proteases of eleven retroviruses representing each of the seven genera of *Retroviridae* was studied using a series of oligopeptides. This system allowed us to examine the relative rates of cleavage under typical conditions (which include the low pH and high salt concentration) used to detect the cleavage of peptides. Molecular models for all studied proteases were built, and they were used to understand the specificity similarities and differences between retroviral proteases and for interpretation of the results. We have classified the processing sites into groups defined by the size and nature of the preferred amino acid residues at the P1, P3, and P4 positions. Because many of the mutations occurring in drug resistance (in the therapy against AIDS) produce residues that can be found in other retroviral proteases we tried to find an agreement between the amino acid preferences in a given position and their naturally occurring type 1 cleavage sites sequences. The specificity distinctions of the proteases correlated well with the phylogenetic tree of retroviruses prepared solely based on the PR sequences.

Comparative study of retroviral proteases is expected to contribute to our understanding of the general and specific features of the PR, and to help to discover the mutational capacity of the HIV-1 PR. Knowledge of the substrate specificity of a variety of retroviral proteases constitutes an essential step toward the rational design of broad spectrum inhibitors, from which viruses can not escape by mutations.

7. REFERENCES

Achong, B. G., Mansen, P. W., and Epstein, M. A. (1971) A new human virus in cultures from a nasopharyngeal carcinoma. *J. Pathol.* **103**(2), 18.

Akkina, R. K., Walton, R. M., Chen, M. L., Li, Q. X., Planelles, V., and Chen, I. S. (1996) High-efficiency gene transfer into CD34+ cells with a human immunodeficiency virus type 1-based retroviral vector pseudotyped with vesicular stomatitis virus envelope glycoprotein G. *J. Virol.* **70** (4), 2581-2585.

Arad, G., Chorev, M., Shtorch, A., Goldblum, A., and Kotler, M. (1995) Point mutation in avian sarcoma leukaemia virus protease which increases its activity but impairs infectious virus production. *J. Gen. Virol.* **76**(8), 1917-1925.

Bagossi, P., Cheng, Y. S., Oroszlán, S., and Tőzsér, J. (1996) Activity of linked HIV-1 proteinase dimers containing mutations in the active site region. *Protein Eng.* **9** (11), 997-1003.

Bagossi, P., Cheng Y. S., Oroszlán, S., and Tőzsér, J. (1998) Comparison of the specificity of homo- and heterodimeric linked HIV-1 and HIV-2 proteinase dimers. *Protein Eng.* **11** (6), 439-445.

Bagossi, P., Sperka, T., Fehér, A., Kádas, J., Zahuczky, G., Miklóssy, G., Boross, P., and Tőzsér, J. (2005) Amino acid preferences for a critical substrate binding subsites of retroviral proteases in type 1 cleavage sites. *J. Virol.* **79**, 4213-4218.

Baldwin, D. N., and Linial, M. L. (1998) The role of Pol and Env in the assembly pathway of human foamy virus. *J. Virol.* **72**, 3658-3665.

Beck, Z. Q., Morris, G. M., and Elder, J. H. (2002) Defining HIV-1 protease substrate selectivity. *Curr. Drug targets Infect. Disord.* **2** (1), 37-50.

Billich, S., Knoop, M. T., Hansen, J., Strop, P., Sedlacek, J., Mertz, R., and Moelling, K. (1988) Synthetic peptides as substrates and inhibitors of human immunodeficiency virus-1 protease. *J. Biol. Chem.* **263** (34), 17905-17908.

Blaha, I., Nemec, J., Tőzsér, J., and Oroszlán, S. (1991) Synthesis of homologous peptides using fragment condensation: analogs of an HIV proteinase substrate. *Int. J. Pept. Protein Res.* **38** (5), 453-458.

Blaha, I., Tőzsér, J., Kim, Y., Copeland, T. D., and Oroszlán, S. (1992) Solid phase synthesis of the proteinase of bovine leukemia virus. Comparison of its specificity to that of HIV-2 proteinase. *FEBS Lett.* **309**, 389-393.

Bodem, J., Lochelt, M., Winkler, I., Flower, R. P., Delius, H., and Flügel, R. M. (1996) Characterization of the spliced *pol* transcript of feline foamy virus: the splice acceptor site of the *pol* transcript is located in *gag* of foamy viruses. *J. Virol.* **70**, 9024-9027.

Boross, P., Tőzsér, J., and Bagossi, P. (2006) Improved purification protocol for wild-type and mutant foamy virus proteases. *Protein Expr. Purif.* **46,** 343-347.

Bowser, P. R., Wolfe, M. J., Forney, J. L., and Wooster, G. A. (1988) Seasonal prevalence of skin tumors from walleye (*Stizostedion vitreum*) from Oneida Lake, New York. *J. Wildl. Dis.* **24**, 292-298.

Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72,** 248-254.

Brik, A., and Wong, C. H. (2003) HIV-1 protease: mechanism and drug discovery. *Org. Biomol. Chem.* **1**(1), 5-14.

Cameron, C. E., Grinde, B., Jacque, P., Jentoft, J., Leis, J., Wlodawer, A., and Weber, I. T. (1993) Comparison of the substrate-binding pockets of the Rous sarcoma virus and human immunodeficiency virus type 1 proteases. *J. Biol. Chem.* **268** (16), 11711-11720.

Cann, A. J., and Chen, I. S. Y. (1996) Human T-cell leukemia virus types 1 and 2. *Fields*, B. N., Knipe, D. M. (eds) *Virology*, 2nd ed., Raven Press, Ltd., New York, 1501-1519.

Cartelierri, M., Rudolph, W., Herchenrőder, O., Lendemann, D., and Rethwilm, A. (2005) determination of the relative amounts of gag and Pol polyproteins in foamy virus particles. *Retrovirology* **2**, 44.

Chiu, Y. L., and Greene, W. C. (2006) APOBEC3 cytidine deaminases: distinct antiviral actions along the retroviral life cycle. *J. Biol. Chem.* **281** (13), 8309-8312.

Chiu, Y. L, and Greene, W. C. (2006) Multifaceted antiviral actions of APOBEC3 cytidine deaminases. *Trends Immunol.* **27** (6), 291-297.

Coffin, J. M. (1992) Genetic diversity and evolution of retroviruses. *Curr. Top. Microbiol. Immunol.* **176**, 143-164.

Coffin, J. M. (1996) Retrovirus restriction revealed. *Nature* **382**, 762-763.

Coffin, J. M., Hughes, S. H., and Varmus, H. E. (1997) Retroviruses. *ColdSpring Harbor Laboratory Press*, NY, USA.

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., et al. (1995) *In vivo* emergence of HIV-1 variants resistant to multiple protease inhibitors. *Nature* **374** (6522), 569-571.

Daenke, S., Schramm, H. J., and Bangham, C. R. M. (1994) Analysis of substrate cleavage by recombinant protease of human T cell leukemia virus type 1 reveals preferences and specificity of binding. *J. Gen. Virol*. **75**, 2233-2239.

Darke, P. L., Nutt, R.F. , Brady, S. F., Garsky, V. M., Ciccarone, T. M., Leu, C. T., Lumma, P. K., Freidinger, R. M., Veber, D. F., and Sigal, I. S. (1988) HIV-1 protease specificity of peptide cleavage is sufficient for processing of gag and pol polyproteins. *Biochem. Biophys. Res. Commun.* **156** (1), 297-303.

Darke, P. L., Jordan, S. P., Hall, D. L., Zugay, J. A., Shafer, J. A., and Kuo, L. C. (1994) Dissociation and association of the HIV-1 protease dimer subunits: equilibria and rates. *Biochemistry* **33** (1), 98-105.

Dauber, D. S., Ziermann, R., Parkin, N., Maly, D. J., Mahrus, S., Harris, J. L., Ellman, J. A., Petropoulus, C., and Craik, C. S. (2002) Altered substrate specificity of drug-resistant human HIV-1 PR. *J. Virol*. **76**, 1359-1368.

De Clercq, E. (2004) Non-nucleoside reverse transcriptase inhibitors (NNRTIs): past, present, and future. *Chem. Biodivers.* **1** (1), 44-46.

Deng, H., Liu, R., Ellmeier, W., Choe, S., Unutmaz, D., Burkhart, M., Di Marzio, P., Marmon, S., Sutton, R. E., Hill, C. M., Davis, C. B., Peiper, S. C., Schall, T. J., Littman, D. R., and Landau, N. R. (1996) Identification of a major co-receptor for primary isolates of HIV-1. *Nature* **381** (6584), 647-648.

Doyon, L., Croteau, G., Thibeault, D., Poulin, F., Pilote, L., and Mamarre, D. (1996) Second locus involved in human immunodeficiency virus type 1resistance to protease inhibitors. *J. Virol.* **70,** 3763-3769**.**

Drake J. W. (1993) Rates of spontaneous mutation among RNA viruses. *Proc. Natl. Acad. Sci. USA* **90** (9), 4171-4175.

Dunn, B. M., Gutschina, A., Wlodaver, A., and Kay, J. (1994) Subsite preferences of retroviral proteinases. *Methods Enzymol.* **241**, 254-278.

Dunn, B. M., Goodenow, M. M., Gutschina, A., and Wlodawer, A. (2002) Retroviral proteases. *Genome Biol.* **3**(4), 3006.1-3006.7.

Eizert, H., Bander, P., Bagossi, P., Sperka, T., Miklóssy, G., Boross, P., Weber, I. T., and Tőzsér, J. (2008) Amino acid preferences of retroviral proteases for amino-terminal positions in a type 1 cleavage site. *J. Virol.* **82**(20), 10111-10117.

Enssle, J., Jordan, I., Mauer, B., and Rethwilm, A. (1996) Foamy virus reverse transcriptase is expressed independently from the Gag protein. *Proc. Natl. Acad. Sci. USA* **93**, 4137-4141.

Fehér , A., Weber, I. T., Bagossi, P., Boross, P., Mahalingam, B., Louis, J. M., Copeland, T. D., Torshin, I. Y., Harrison, R. W., and Tőzsér, J. (2002) Effect of sequence polymorphism and drug-resistance on two HIV-1 Gag processing sites. *Eur. J. Biochem.* **269**, 4114-4120 (1-7).

Fehér, A., Boross, P., Sperka, T., Oroszlán, S., and Tőzsér, J. (2004) Expression of the murine leukemia virus protease in fusion with maltose-binding protein in *Escherichia coli. Prot. Expres. Purif.* **35**, 62-68.

Fehér, A., Boross, P., Sperka, T., Miklóssy, G., Kádas, J., Bagossi, P., Oroszlán, S., Weber, I. T., and Tőzsér, J. (2006) Characterization of the murine leukemia virus protease and its comparison with the HIV-1 protease. *J. Gen. Virol.* **87**, 1321-1330.

Felsentein, J. (1989) PHYLIP - Phylogeny Inference Package. Cladistics 5: 164-166.

Fenyőfalvi, G., Bagossi, P., Copeland, T. D., Oroszlán, S., Boross, P., and Tőzsér, J. (1999) Expression and characterization of human foamy virus proteinase. *FEBS Lett.* **462**, 397-401.

Fischer, N., Heinkelein, M., Lindemann, D., Enssle, J., Baum, C., Werder, E., Zentgraf, H., Muller, J. G., and Rethwilm, A. (1998) Foamy virus particle formation. *J. Virol*. **72**, 1610-1615.

Flügel, R. M., and Pfrepper, K. I. (2003) Proteolytic processing of foamy virus Gag and Pol proteins. *Curr. Top. Microbiol. Immunol.* **277**, 63-88.

Fodor, S. K., and Vogt, V. M. (2002) Characterization of the protease of a fish retrovirus, walley dermal sarcoma virus. *J. Virol.* **76**, 4341-4349.

Freed, E. O. (2001) HIV-1 replication. *Somat. Cell. Mol. Genet.* **26** (1-6), 13-33.

Gallo, R. C. and Montagnier, L**.** (1988) AIDS in 1988. *Sci. Am*. **259** (4), 41-48.

Gheysen, D., Jacobs, E., De Foresta, H., Thiriart, C., Francotte, M., Thines, D., and De Wilde, M. (1989) Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. *Cell* **59** (1), 103-112.

Ghysdael, J., Bruck, C., Kettmann, R., and Burny, A. (1984) Bovine leukemia virus. *Curr. Topic Microbiol. Immunol.* **112**, 1-19.

Griffiths, J. T., Phylip, L. H., Konvalinka, J., Strop, P., Gutschina, A., Wlodawer, A., Davenport, R. J. , Briggs, R., Dunn, B. M., and Kay, J. (1992) Different requirements for productive interaction between the active site of HIV-1 proteinase and substrate containing -hydrophobic* hydrophobic- or -aromatic*Pro- cleavage sites. *Biochemistry* **31**, 5193-5200.

Grinde, B., Cameron, C. E., Leis, J., Weber, I. T., Wlodawer, A., Burstein, H., and Skalka, A. M. (1992) Analysis of substrate interaction of the Rous sarcoma virus wild type and mutant proteases and human immunodeficiency virus-1 protease using a set of systematically altered peptide substrates. *J. Biol. Chem.* **267** (14), 9491-9498.

Gulnik, S. V., Suvorov, L. I., Liu, B., Yu, B., Anderson, B., Mitsuya, H., and Erickon, J. W. (1995) Kinetic characterization and cross-resistance patterns of HIV-1 protease mutants selected under drug pressure. *Biochemistry* **34** (29), 9282-9287.

Gutschina, A., and Weber, I. T. (1990) Comparison of inhibitor binding in HIV-1 protease and in non-viral aspartic proteases: the role of the flap. *FEBS Lett.* **269** (1), 269-272.

Gutschina, A., and Weber, I. T. (1991) Comparative analysis of the sequences and structures of HIV-1 and HIV-2 proteases. *Proteins* **10**, 325-339.

Gutschina, A. (1995). Molecular modeling of the structure of FIV protease. *Adv. Exp. Med. Biol.* **362**, 479-484.

Gutschina, A., Kervinen, J., Powell, D. J., Zdanov, A., Kay, J., and Wlodawer, A. (1996) Structure of equine infectious anemia virus proteinase complexed with an inhibitor, *Protein Sci.* **5**, 1453-1465.

Harrison, R. W., and Weber, I. T. (1994) Molecular dynamics simulations of HIV-1 protease with peptide substrate. *Prot. Eng.* **7**(11), 1353-1363.

Hartl, M. J., Wohrl, B. M., Rosch, P., and Schweimer, K. (2008) The solution structure of the simian foamy virus protease reveals a monomreric protein. *J. Mol. Bio.* **381**, 141-149.

Hartl, M. J., Mayr, F., Rethwilm, A, and Wohrl, B. M. (2010a) Biophysical and enzymatic properties of the simian and prototype foamy virus reverse transcriptases. *Retrovirology* **7** (5), 1-10.

Hartl, M. J., Schweimer, K., Reger, M. H., Schwarzinger, S., Bodem, J., Rosch, P., and Wohrl, B. M. (2010b) Formation of transient dimers by a retroviral protease. *Biochem. J.* **427**, 197-203

Hartl, M. J., Bodem, J., Jochheim, F., Rethwilm, A., Rosch, P., and Wohrl, B. M. (2011) Regulation of foamy virus protease activity by viral RNA - A novel and unique mechanism among retroviruses. *J. Virol.* **85**, 4462-4469.

Heinkelein, M., Pietschmann, T., Jármy, G., Dressler, M., Imrich, H., Thurow, J., Lindemann, D., Bock, M., Moebes, A., Roy, J., Herchenröder, O., and Rethwilm, A. (2000) Efficient intracellular retrotransposition of an exogenous primate retrovirus genome. *EMBO J.* **19**(13), 3436-3445.

Heinkelein, M., Rammling, M., Juretzek, T., Lindemann, D., and Rethwilm, A. (2003) Retrotransposition and cell-to-cell transfer of foamy viruses. *J. Virol.* **77**, 11855-11858.

Herchenroder, O., Renne, R., Loncar, D., Cobb, E. K., Murthy, K. K., Schneider, J., Mergia, A., and Luciw, P. A. (1994) Isolation, cloning, and sequencing of simian foamy viruses from chimpanzees (SFV cpz): high homology to human foamy virus (HFV). *Virology* **201**, 187-199.

Hertogs, K., Bloo, S., Kemp, S. D., Van den Eynde, C., Alcorn, T. M., Pauwels R., Van Houtte, M., Staszewski, S., Miller, V., and Larder, B. A. (2000) Phenotypic and genotypic analysis of clinical HIV-1 isolates reveal extensive protease inhibitor cross-resistance: a survey of over 6000 samples. *AIDS* **14** (9), 1203-1210.

Hirsch, V. M., Dapolito G., Hahn A., Lifson J., Montefiori D., Brown C. R., and Goeken R. (1998) Viral genetic evolution in macaques infected with molecularly cloned simian immunodeficiency virus correlates with the extent of persistent viremia. *J. Virol.* **72** (8), 6482-6489.

Hirsch, M. S., Conway, B., D'Aquila, R. T., Johnson, V. A., Brun-Vézinet, F., Clotet, B., Demeter, L. M., Hammer, S. M., Jacobsen, D. M., Kuritzkes, D. R., Loveday, C., Mellors, J. W., Vella, S., and Richman, D. D. (1998) Antiretroviral drug resistance testing in adults with HIV infection: implications for clinical management. *International AIDS Society - USA Panel. JAMA* **279** (24), 1984-1991.

Ho, D. D. (1995) HIV-1 dynamics in vivo. *J. Biol. Regul. Homeost. Agents.* **9** (3), 76-77.

Hooks, J. J., and Gibbs, C. J. (1975) The foamy viruses. *Bacteriol. Rev.* **39** (3), 169-185.

Hruskova-Heidingsfeldova, O., Andreansky, M., Fabry, M., Blaha, I., Strop, P., and Hunter, E. (1995) Cloning, bacterial expression, and characterization of the Mason-Pfizer monkey virus proteinase. *J. Biol. Chem.* **270** (25), 15053-15058.

Hyland, L. J., Tomaszek, T. A. Jr., Meek, D. T. (1991) Human immunodeficiency virus-1 protease. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism. *Biochemistry* **30**, 8454-8463.

Ingr, M., Uhlikova, T., Strisovky, K., Majerova, E., and Konvalinka, J. (2003) Kinetics of the dimerization of retroviral proteases: The "fireman's grip" and dimerization. *Protein Sci.* **12**, 951-953.

Jaskolski, M., Miller, M., Rao, J. K., Leis, J., and Wlodawer, A. (1990) Structure of the aspartic protease from Rous sarcoma retrovirus refined at 2-Å resolution. *Biochemistry* **29** (25), 5889-5898.

Jandu, S. K., Ray, S., Brooks, L., and Leatherbarrow, R. J. (1990) Role of arginine 67 in the stabilization of chymotrypsin inhibitor 2: examination of amide proton exchange rates and denaturation thermodynamics of an engineered protein. *Biochemistry* **29** (26), 6264-6269.

Johnson, V. A., Brun-Vezinet, F., Clotet, B., Gunthard, H. F., Kuritzkes, D., Pillay, D., Shapiro, A., and Richman, D. D. (2009) Update of the Drug Resistance Mutations in HIV-1: December 2009. *Top. HIV Med.* **17** (5), 138-145.

Jordan, I., Enssle, J., Guttler, E., Mauer, B., and Rethwilm, A. (1996) Expression of human foamy virus reverse transcripatse involves a spliced pol mRNA. *Virology* **224**, 314-319.

Kádas, J., Weber, I. T., Bagossi, P., Miklóssy, G., Boross, P., Oroszlán, S., and Tőzsér, J. (2004) Narrow substrate specificity and sensitivity toward ligand binding site mutations of human T-cell leukemia virus type-1 protease. *J. Biol. Chem.* **279,** 27148-27157.

Kapust, R. B. and Waugh, D. S. (1999) Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* **8**, 1668-1674.

Kay, J., and Dunn, B. M. (1990) Viral proteinases: weakness in strength. *Biochim. Biophys. Acta* **1048** (1), 1-18.

Keller, A., Partin, K. M., Löchelt, M., Bannert, H., Flügel, R. M., and Cullen, B. R. (1991) Characterization of the transcriptional transactivator of human foamy retrovirus. *J. Virol.* **65,** 2589-2594.

Kempf, D. J. (1994) Design of symmetry-based peptidomimetic inhibitors of human immunodeficiency virus protease. *Methods Enzymol.* **241,** 334-354.

Kervinen, J., Lubkowski, J., Zdanov, A., Bhatt, D., Dunn, B. M., Hui, K. Y., Powell, D. J., Kay, J., Wlodawer, A., and Gutschina, A. (1998) Toward a universal inhibitor of retroviral proteases; comparative analysis of the interactions of LP-130 complexed with proteases from HIV-1, FIV, and EIAV. *Protein Sci*. **7,** 2314-2323.

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

Konvalinka, J., Strop, P., Velek, J., Cerna, V., Kostka, V., Phylip, L. H., Richards, A. D., Dunn, B. M., B. M., and Kay, J. (1990) Sub-site preferences of the aspartic proteinase from HIV-1. *FEBS Letters* **268,** 35-38.

Konvalinka, J., Litterst, M. A., Welker, R., Kottler, H., Rippmann, F., Heuser, A. M., and Krausslich, H.G. (1995a) An active-site mutation in the human immunodeficiency virus type 1 proteinase (PR) causes reduced PR activity and loss of PR-mediated cytotoxicity without apparent effect on virus maturation and infectivity. *J. Virol.* **69**, 7180-7186.

Konvalinka, J., Löchelt, M., Zentgraf, H., Flügel, R. M., and Krausslich, H. G. (1995b) Active foamy virus proteinase is essential for virus infectivity but not for formation of a Pol polyprotein. *J. Virol.* **69** (11), 7264-7268.

Kögel, D., Aboud, M., and Flügel, R. M. (1995) Molecular biological characterization of the human foamy virus reverse transcriptase and ribonuclease H domains. *Virology* **213**(1), 97-108.

Kuwata, T., Miyazaki, Y., Igarashi, T., Takehisa, J., and Hayami, M. (1997) The rapid spread of recombinants during a natural in vitro infection with two human immunodeficiency virus type 1 strains. *J. Virol.* **71** (9), 7088-7091.

Lapatto, R., Blundell, T., Hemmings, A., Overington, J., Wilderspin, A., Wood, S., Merson, J. R., Whittle, P. J., Danley, D. E., Geoghegan, K. F., et al. (1989) X-ray analysis of HIV-1 proteinase at 2.7 A resolution confirms structural homology among retroviral enzymes. *Nature* **342** (6247), 299-302.

Lee, E. G., and Linial, M. L. (2008) The C terminus of foamy retrovirus Gag contains determinants for encapsidation of Pol protein into virions. *J. Virol.* **82**, 10803-10810.

Lee, E. G., Roy, J., Jackson, D., Clark, P., Boyer, P. L., Hughes, S. H., and Linial, M. (2011) Foamy retrovirus integrase contains a Pol dimerization domain required for protease activation. *J. Virol.* **85**, 1655-1661.

Lehmann-Che, J., Giron, M. L., Delelis, O., Löchelt, M., Bittoun, P., Tobaly-Tapiero, J., de The, H., and Saib, A. (2005) Protease-dependent uncoating of a complex retrovirus. *J. Virol.* **79**, 9244-9253.

Leis, J., Baltimore, D., Bishop, J. M., Coffin, J., Fleissner E., Goff, S. P., Oroszlán, S., Robinson, H., Skalka, A. M., Temin, H. M., and Vogt, V. (1988) Standardized and simplified nomenclature for proteins common to all retroviruses. *J. Virol*. **62** (5), 1808-1809.

Leurs, C., Jansen, M., Pollok, K. E., Heinkelein, M., Schmidt, M., Wissler, M., Lindemann, D., Von Kalle, C., Rethwilm, A., Williams, D. A., et al. (2003) Comparison of three retroviral vector systems for transduction of nonobese diabetic /severe combined immunodeficiency mice repopulating human CD34+ cord blood cells. *Hum. Gene. Ther.* **14**, 509-519.

Li, M., Lako, G. S., Jaskolski, M., Rozycki, J., Alexandratos, J., Wlodawer, A., and Gutschina, A. (2005) Crystal structure of human T-cell leukemia virus protease, a novel target for anticancer drug design. *Proc. Natl. Acad. Sci. USA* **102** (51), 18332-18337.

Lindemann, D., and Rethwilm, A. (2011) Foamy Virus Biology and Its Application for Vector Development. *Viruses* **3**, 561-585.

Linial, M. (1999) Foamy viruses are unconventional retroviruses. *J. Virol.* **73**, 1747-1755.

Linial, M. (2000) Why aren't foamy viruses pathogenic? *Trends in Microb.* **8**(6), 284-289.

Linial, M. (2007) Foamy viruses, *Fields Virol*. Lippincot Williams & Wilkins, PhiladelphiaKnipe DM, Howley PM, 2245-2262.

Loeb, D. D., Hutchison, C. A. 3rd, Edgell, M. H., Farmerie, W. G., and Swanstrom, R. (1989) Mutational analysis of human immunodeficiency virus type 1 protease suggests functional homology with aspartic proteinases. *J. Virol.* **63** (1), 111-121.

Löchelt, M., and Flügel, R. M. (1995) The molecular biology of primate spumaviruses. In: Levy JA (ed) Plenum Press, New York, *The Retroviridae* **4**, 239-292.

Löchelt, M., and Flügel, R. M. (1996) The human foamy virus *pol* gene is expressed as a Pro-Pol polyprotein and not as a Gag-Pol fusion protein. *J. Virol.* **70**, 1033-1040.

Louis, J. M., Nashed, M. T., Parris, K. D., Kimmel, A. R., and Jerina, D. M. (1994) Kinetics and mechanism of autoprocessing of human immunodeficiency virus type 1 protease from an analog of the Gag-Pol polyprotein. *Proc. Natl. Acad. Sci.* **91**(17), 7970-7974.

Louis, J. M., Oroszlán, S., and Tőzsér, J. (1999) Stabilization from autoproteolysis and kinetic characterization of the human T-cell leukemia virus type 1 proteinase. *J. Biol. Chem.* **274**, 6660-6666.

Louis, J. M., Weber, I. T., Tőzsér, J., Clore, G. M., and Gronenborn, A. M. (2000) HIV-1 protease: maturation, enzyme specificity and drug-resistance. *Adv. Pharmacol.* **49**, 111-146.

Luban, J., and Goff, S. P. (1994) Mutational analysis of *cis*-acting packaging signals in human immunodeficiency virus type 1 RNA. *J. Virol*. **68** (6), 3784-3793.

MacArthur, M. W., and Thornton, J. M. (1991) Influence of proline residues on protein conformation. *J. Mol. Biol.* **218**(2), 397-412.

McCornack, M. A., Kakalis, L. T., Caserta, C., Handschumacher, R. E., and Armitage, I. M. (1997) HIV protease substrate conformation: modulation by cyclophilin A. *FEBS Lett.* **414**(1), 84-88.

Maddon, P. J., Dalgleish, A. G., McDougal, J. S., Clapham, P. R., Weiss, R. A., and Axel, R. (1986) The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* **47**(3), 333-348.

Maguire, M., Shortino, D., Klein, A., Harris, W., Manohitharajah, V., Tisdale, M., Elston, R., Yeo, J., Randall, S., Xu, F., Parker, H., May, J., and Snwden, W. (2002) Emergence of resistance to protease inhibitor amprenvir in human immunodeficiency virus type 1-infected patients: selection of four alternative viral protease genotypes and influence of viral susceptibility to coadministered reverse transcriptase nucleoside inhibitors. *Antimicrob. Agents Chemother*. **46** (3), 731-738.

Mahalingam, B., Louis, J. M., Reed, C. C., Adomat, J. M., Krouse, J., Wang, Y. F., Harrison, R. W., and Weber, I. T. (1999) Structural and kinetic analysis of drug resistant mutants of HIV-1 protease. *Eur. J. Biochem.* **263,** 238-245.

Mahalingam, B., Boross, P., Wang, Y. F., Louis, J. M., Fischer, C. C., Tőzsér, J., Harrison, R. W., and Weber, I. T. (2002) Combining mutations in HIV-1 protease to understand mechanisms of resistance. *Proteins* **48,** 107-116.

Mammano F., Pettit, C., and Clavel, F**.** (1998) Resistance-associated loss of viral fitness in human immunodeficiency virus type 1: phenotypic analysis of protease and Gag coevolution in protease inhibitor-treated patients. *J. Virol.* **72** (9), 7632-7637.

Mansky, L. M., and Temin, H. M. (1995) Lower *in vivo* mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J. Virol.* **69** (8), 5087-5094.

Maurer, B., Bannert, H., Darai, G., and Flügel, R. M. (1988) Analysis of the primary structure of the long terminal repeat and *gag* and *pol* genes of the human spumaretrovirus. *J. Virol.* **62**, 1590-1597.

Meiering, C. D., and Linial, M. L. (2001) Historical perspective of foamy virus epidemiology and infection. *Clin. Microbiol. Rev.* **14** (1), 165-176.

Menéndez-Arias, L., Young, M., and Oroszlán, S. (1992) Purification and characterization of the mouse mammary tumor virus protease expressed *in Escherichia coli*. *J. Biol. Chem.* **267**, 24134-24139.

Menéndez-Arias, L., Gotte, D., and Oroszlán, S. (1993) Moloney murine leukemia virus protease: bacterial expression and characterization of the purified enzyme. *Virology* **196**, 557-563.

Mergia, A., and Heinkelein, M. (2003) Foamy virus vectors. *Curr. Top. Microbiol. Immunol.* **277**, 131-159.

Merkulov, G. V., Swiderek, K. M., Brachmann, B., and Boeke, J. D. (1996) A critical proteolytic cleavage site near the C-terminus of the yeast retrotransposon Ty1 Gag protein. *J. Virol.* **70**, 5548-5556.

Miller, M., Jaskolski, M., Rao, J. K. M., Leis, J., and Wlodawer, A. (1989) Crystal structure of a retroviral protease proves relationship to aspartic protease family. *Nature* **337**, 576-579.

Moebes, A., Enssle, J., Bieniasz, P. D., Heinkelein, M., Lindemann, D., Bock, M., McClure, M. O., and Rethwilm, A. (1997) Human foamy virus reverse transcription that occurs late in the viral replication cycle. *J. Virol*. **71**, 7305-7311.

Montagnier, L. (2009) 25 Years after HIV Discovery: Prospects for Cure and Vaccine (Nobel Lecture). *Angew. Chem. Int. Ed*. **48**, 5815-5826.

Morozov, V. A., Copeland, T. D., Nagashima, K., Gonda, M. A., and Oroszlán, S. (1997) Protein composition and morphology of human foamy virus intracellular cores and extracellular particles. *Virology* **228**, 307-317.

Mulichak, A. M., Hui, J. O., Tomasselli, A. G., Heinrikson, R. L., Curry, K. A., Tomich, C. S., Thaisrivongs, S., Sawyer, T. K., and Watenpaugh, K. D. (1993) The crystallographic structure

of the protease from human immunodeficiency virus type 2 with two synthetic peptidic transition state analog inhibitors. *J. Biol. Chem.* **268**(18), 13103-13109.

Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A.W., Martelli, G. P., Mayo, M. A., and Summers, M. D. (eds) (1994) Virus taxonomy: The classification and nomenclature of viruses. *Retroviridae*. Springer-Verlag, Vienna, Austria.

Murray, P. R., Rosenthal, K. S., Kobayashi, G. S, and Pfaller, M. A. (1997) *Medical Microbiology* (3rd edition), 507.

Murray, S. M. and Linial, M. L (2006) Foamy virus infection in primates. *J. Med. primatol.* **35**, 225-235.

Murray, S. M., Picker L. J., Axthelm, M. K., Hudkins, K., Alpers, C. E., and Linial, M. L. (2008) Replication in a superficial epithelial cell niche explains the lack of pathogenicity of primate foamy virus infections. *J. Virol.* **82**, 5981-5985.

Naldini, L., Blomer, U., Gage, F. H., Trono, D., and Verma, I. M. (1996) Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. *Proc. Natl. Acad. Sci. USA* **93**(21), 11382-11388.

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., and Springer, J. P. (1989).Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* **337**, 615-620.

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., et al. (1995) Flexibility and function in HIV-1 protease. *Nat. Struct. Biol.* **2** (4), 274-280.

Ohtaka, H., and Freire, E. (2005) Adaptive inhibitors of the HIV-1 protease. *Prog. Biophys. Mol. Biol*. **88** (2), 193-208.

Oroszlán, S. and Luftig, R. B. (1990) Retroviral proteinases. *Curr. Top. Microbiol.* **157,** 153-185.

Oroszlán, S. and Tőzsér, J. (1990) The retroviral proteinases. *Virolog* **1,** 369-378.

Ozer, N., Haliloglu, T., and Schiffer, C. A. (2006) Substrate specificity in HIV-1 protease by a biased sequence search method. *Proteins* **64** (2), 444-456.

Pahl, A., and Flügel, R. M. (1993) Endonucleolytic cleavages and DNA-joining activities of the integration protein of human foamy virus. *J. Virol.* **67** (9), 5426-5434.

Pauwels, R. (2006) Aspects of successful drug discovery and development. *Antiviral research* **71,** 77-89.

Perelson, A. S., Neumann, A. U., Markowitz, M., Leonard, J. M., and Ho, D. D. (1996) HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* **271** (5255), 1582-1586.

Pettit, S. C., Simsic, J., Loeb, D. D., Everitt, L., Hutchinson III, C. A., and Swanstrom, R. (1991) Analysis of retroviral protease cleavage sites reveals two types of cleavage sites and the structural requirements of the P1 amino acid. *J. of Biol. Chemistry* **266**(22), 145539-14547.

Pettit, S. C., Sheng, N., Tritch, V., Erickson-Viitanen, S., and Swanstrom, R. (1998) The regulation of sequential processing of HIV-1 Gag by the viral protease. *Adv. Exp. Med. Biol.* **436**, 15-25.

Phylip, L. H., Richards, A. D., Kay, J., Kovalinka, J., Strop, P., Blaha, I., Velek, J., Kostka, V., Ritchie, A. J., Broadhurst, A. V., et al. (1990) Hydrolysis of synthetic chromogenic substrates by HIV-1 and HIV-2 proteinases. *Biochem. Biophys. Res. Commun.* **171** (1), 439-444.

Pfrepper, K. I., Löchelt, M., Schnolzer, M., and Flügel, R. M. (1997) Expression and molecular characterization of an enzymatically active recombinant human spumaretrovirus protease. *Biochem. Biophys. Res. Commun.* **237**, 548-553.

Pfrepper, K. I., Löchelt, M., Rackwitz, H. R., Schnölzer, M., Heid, H., and Flügel, R. M. (1999) Molecular characterization of proteolytic processing of the Gag proteins of human spumavirus. *J. Virol*. **73**, 7907-7911.

Pietschmann, T., Heinkelein, M., Heldmann, M., Zentgraf, A., Rethwilm, A., and Lindemann, D. (1999) Foamy virus capsids require the cognate envelope protein for particle export. *J. Virol*. **73**, 2613-2621.

Polgár, L., Szeltner, Z., and Boros, I. (1994) Substrate-dependent mechanism in the catalysis of human immunodeficiency protease.*Biochemistry* **33**, 9351-9357.

Poorman, R. A., Tomaselli, A. G., Heinrikson, R. L., and Kézdy, F. J. (1991) A cumulative specificity model for proteases from human immunodeficiency virus types 1 and 2, inferred from statistical analysis of an extended substrate data base. *J. Biol. Chem*. **266** (22), 14554-14561.

Powell, P. D. (1996) Can evolutionary principles help resolve the aquried immune deficiency syndrome crisis*? Med. Hypotheses* **46** (2), 130-134.

Prabu-Jeyabalan, M., Nalvaika, E., and Schifffer, C. A. (2002) Substrate shapes determine specificity of recognition for HIV-1 PR: analysis of crystal structures of six substrate complexes. *Structure* **10,** 369-381.

Prabu-Jeyabalan, M., Nalivaika, E. A., King, N. M., and Schiffer, C. A. (2003) Viability of a drug-resistant human immunodeficiency virus type 1 protease variant: structural insights for better antiviral therapy. *J. Virol*. **77** (2), 1306-1315.

Preston, B. D., Poiesz, B. J., and Loeb, L. A. (1988) Fidelity of HIV-1 reverse transcriptase. *Science* **242** (4882), 1168-1171.

Rao, C., Scarborough, P. E., Lowther, W. T., Kay, J., Batley, B., Rapundalo, S., Klutchko, S., Taylor, M. D., and Dunn, B. M. (1991) Structure-function database for active site binding to the aspartic proteinases. *Adv. Exp. Med. Biol*. **306,** 143-147.

Rethwilm, A., Erlwein, O., Baunach, G., Maurer, B., and ter Meulen, V. (1991) The transcriptional transactivator of human foamy virus maps to the *bel 1* genomic region. *Proc. Natl. Acad. Sci. USA* **88**, 941-945.

Rethwilm, A. (2003) The replication strategy of foamy viruses. *Curr. Top. Microbiol. Immunol.* **277**, 1-26.

Rethwilm, A. (2005) Foamy viruses. In: Mahy BWJ, ter Leulen V (eds) Topley and Wilson's microbiology and microbial infections-virology, Hodder Arnold, London, 1304-1321.

Ridky, T. and Leis, J. (1995) Development of drug resistance to HIV-1 protease inhibitors. *J. Biol. Chem*. **270** (50), 29621-29623.

Robertson, D. L., Sharp, P. M., McCutchan, F. E., and Hahn, B.H. (1995) Recombination in HIV-1. *Nature* **374** (6518), 124-126.

Rose, R. B., Rose, J. R., Salto, R., Craik, C. S., and Stroud, R. M. (1993) Structure of the protease from simian immunodeficiency virus: complex with an irreversible nonpeptide inhibitor. *Biochemistry* **32** (46), 12498-12507.

Rosin, C. D., Belew, R. K., Walker, W. L., Morris, G. M., Olson, A. J., and Goodsell, D. S. (1999) Coevolution and subsite decomposition for the design of resistance-evading HIV-1 protease inhibitors. *J. Mol. Biol.* **287**, 77-92.

Rothenaigner, I., Kramer, S., Meggendorfer, M., Rethwilm, A., and Brack-Werner, R. (2009) Transduction of human neural progenitor cel with foamy virus vectors for differentitation-depenedent gene expression. *Gen. Ther*. **16**, 349-358.

Roy, J., Rudolph, W., Juretzek, T., Gartner, K., Bock, M., Herchenroder, O., Lindemann, D., Heinkelein, M., and Rethwilm, A. (2003) Feline foamy virus genome and replication strategy. *J. Virol.* **77**, 11324-11331.

Rustigian, R., Johnston, P., and Reihart, H. (1955) Infection of monkey kidney tissue cultures with virus-like agents. *Proc. Soc. Exp. Biol.Med.* **88** (1), 8-16.

Sali, A., and Blundell, T. L. (1993) Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol*. **234**, 779-815.

Schechter, I., and Berger, A**. (**1967) On the size of the active site in proteases. I.papain. *Biochem. Biophys. Res. Commun*. **27,** 157-162.

Schliephake, A. W., and Rethwilm, A. (1994) A nuclear localization of foamy virus Gag precursor protein. *J. Virol*. **68**, 4946-4954.

Shafer, R.W. (2002) Genotypic testing for human immunodeficiency virus type 1 drug resistance. *Clin. Microbiol. Rev.* **15** (2), 247-277.

Sperka, T., Boross, P., Eizert, H., Tőzsér, J., and Bagossi, P. (2006) Effect of mutation on the dimer stability and the pH optimum on the human foamy virus protease. *Prot. Engin. Design & Selec.* **8** (19), 369-375.

Sperka, T., Miklóssy, G., Tie, Y., Bagossi, P., Zahuczky, G., Boross, P., Weber, I. T., Harrison, R. W., and Tőzsér, J. (2007) Bovine leukemia virus protease: comparison with human T-lymphotropic virus and human immunodeficiency virus proteases. *J. Gen. Virol.* **88**, 2052-2063.

Szeltner, Z., and Polgár, L. (1996a) Conformational stability and catalytic activity of HIV-1 protease are both enhanced at high salt concentration. *J. Biol. Chem.* **271** (10), 5458-5463.

Szeltner, Z., and Polgár, L. (1996b) Rate-determining steps in HIV-1 protease catalysis. The hydrolysis of the most specific substrate. *J. Biol. Chem.* **271** (50), 32180-32184.

Stephens, R. M., Casey, J. W., and Rice, N. R. (1986) Equine infectious anemia virus gag and pol genes: relatedness to visna and AIDS virus. *Science* **231** (4738), 589-594.

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

Swain, A. L., Miller, M. M., Green, J., Rich, D. H., Schneider, J., Kent, S. B., and Wlodaver, A. (1990) X-gray crystallographic structure of a complex between a synthetic protease of human immunodeficiency virus 1 and a substrate-based hydroxyethylamine inhibitor. *Proc. Natl. Acad. Sci. USA* **87** (22), 8805-8809.

Swanstrom, R., and Wills, J. W. (1997) Synthesis, assembly, and processing of viral proteins. In *Retroviruses*, Coffin, J. M., Hughes, S. H., and Varmus, H. E., Eds.; Cold Spring Harbour Laboratory Press: Cold Spring Harbour, NY, USA, 1997.

Tamalet, C., Pasquier, C., Yahi, N., Colson, P., Poizot-Martin, I., Lepeu, G., Gallais, H., Massip, P., Puel, J., and Izopet, J. (2000) Prevalence of drug resistant mutants and virological response to combination therapy in patients with primary HIV-1 infection. *J. Med. Virol.* **61** (2), 181-186.

Tandon, V. K., and Chhor, R. B. (2005) Current status of anti-HIV agents. *Curr. Med. Chem. Anti-Infective Agents* **4**, 3-28.

Tessmer, U., and Krausslich, H. G. (1998) Cleavage of human immunodeficiency virus type 1 protease from the N-terminally adjacent p6* protein is essential for efficient Gag polyprotein processing and viral infectivity. *J. Virol.* **72** (4), 3459-3463.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680.

Tobaly-Tapiero, J., Bittoun, P., Neves, M., Guillemin, M. C., Lecellier, C-H., Puvion-Dutilleul, F., Gicquel, B., Zientara, S., Giron, M-L., de The, H., and Saib, A. (2000) Isolation and characterization of an equine foamy virus. *J. Virol*. **74**, 4064-4077.

Tomasselli, A. G., and Heinrikson, R. L. (1994) Specificity of retroviral proteases: an analysis of viral and nonviral protein substrates. *Methods Enzymol*. **241**, 279-301.

Tomasselli, A. G., and Heinrikson, R. L. (2000) Targeting the HIV-protease in AIDS therapy: a current clinical perspective. *Biochim. Biophys. Acta* **1477**(1-2), 189-214.

Tong, L., Pav, S., Mui, S., Lamarre, D., Yoakim, C., Beaulieu, P., and Anderson, P. C. (1995) Crystal structure of HIV-2 protease in complex with inhibitors containing the hydroxyethylamine dipeptide isostere. *Structure* **3**, 33-40.

Tőzsér, J., Gustchina, A., Weber, I. T., Blaha, I., Wondrak, E. M., and Oroszlán, S. (1991a) Studies on the role of the S4 substrate binding site of HIV proteinases. *FEBS Lett.* **279,** 356-360.

Tőzsér, J., Blaha, I., Copeland, T.D., Wondrak, E.M. Oroszlán, S. (1991b) Comparison of the HIV-1 and HIV-2 proteinases using oligopeptide substrates representing cleavage sites in Gag and Gag-Pol polyproteins. *FEBS Lett.* **281,** 77–80.

Tőzsér J., Weber I. T., Gustchina A., Blaha I., Copeland T. D., Louis J. M., and Oroszlán S. (1992) Kinetic and modeling studies of S3-S3' subsites of HIV proteinases. *Biochemistry* **31**, 4793-800.

Tőzsér J., Friedman, D., Weber I. T., Blaha, I., and Oroszlán, S. (1993) Studies on the substrate specificity of the proteinase of equine infectious anemia virus using oligopeptide substrates. *Biochemistry* **32**, 3347-3353.

Tőzsér, J., Bagossi, P., Weber, I. T., Copeland, T. D., and Oroszlán, S. (1996) Comparative studies on the substrate specificity of avian myeloblastosis virus proteinase and lentiviral proteinases. *J. Biol. Chem*. **271**, 6781-6788.

Tőzsér, J., Yin, F. H., Cheng, Y. S. E., Bagossi, P., Weber, I. T., Harrison, R. W., and Oroszlán, S. (1997) Activity of tethered HIV-1 PR containing mutations in the flap region of one subunit. *Eur. J. Biochem*. **244**, 235-241.

Tőzsér, J. (2003) Stages of HIV replication and targets for therapeutic intervention. *Curr. Topics in Med. Chem.* **3**, 1447-1457.

Tőzsér, J., and Oroszlán, S. (2003) Proteolytic events of HIV-1 replication as targets for therapeutic intervention.*Curr. Pharm. Des.* **9**, 1803-1815.

Tőzsér, J. (2010) Comparative Studies on Retroviral Proteases: Substrate Specificity. *Viruses* **2,** 147-165.

Towbin, H., Staehelin, T., and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* **76**, 4350-4354.

Trobridge, G., and Rusell, D. W. (2004) Cell cycle requirements for transduction by foamy virus vectors compared to those of oncovirus and lentivirus vectors. *J. Virol.* **78** (5), 2327-2335.

Turner, B. G., and Summers, M. F. (1999) Structural biology of HIV. *J. Mol. Biol*. **285**(1), 1-32.

http://www.unaids.org (report 2010.pdf)

UNAIDS (2011). "AIDS epidemic update: December 2011." http://www.unaids.org.

Vance, J. E., LeBlanc, D. A., Wingfield, P., and London, R. E. (1997) Conformational selectivity of HIV-1 protease cleavage of X-Pro peptide bonds and its implications. *J. Biol. Chem.* **272**, 15603-15606.

Veverka, V., Bauerova, H., Zabransky, A., Pichova, I., and Hrabal, R. (2001) Backbone resonance assignment of protease from Mason-Pfizer monkey virus. *J. Biomol. NMR* **20**, 291-292.

Veverka, V., Bauerova, H., Zabransky, A., Lang, J., Ruml, T., Pichova, I., and Hrabal, R. (2003) Three-dimensional structure of a monomeric form of a retroviral protease. *J. Mol. Biol.* **333**, 771-780.

Velazquez-Campoy, A., Muzammil, S., Ohtaka, H., Schon, A, Vega, S., and Freire, E. (2003) Structural and thermodynamic basis for resistance to HIV-1 protease inhibition: implications for inhibitor design. *Curr. Drug Targets Infect. Disord.* **3** (4), 311-328.

Vickrey, J. F., Logsdon, B. C., Proteasa, G., Palmer, S., Winters, M. A., Merigan, T. C., and Kovari, L. C. (2003) HIV-1 protease variants from 100-fold drug resistant clinical isolates: expression, purification, and crystallization. *Prot. Expr. Purif.* **28** (1), 165-172.

Vriend, G. (1990) WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* **8** (1), 52-56.

Wainberg, M. A., and Friedland, G. (1998) Public health implications of antiretroviral therapy and HIV drug resistance. *JAMA* **279** (24), 1977-83.

Wang, C., Castro, A. F., Wilkes, D. M., and Altenberg, G. A. (1999) Expression and purification of the first nucleotide-binding domain and linker region of human multidrug resistance gene product: comparison of fusions to glutathione S-transferase, thioredoxin and maltose-binding protein. *Biochem. J.* **338** (Pt 1), 77-81.

Weber, I. T. (1989) Structural alignment of retroviral protease sequences. *Gene* **85** (2), 565-566.

Weber, I. T. (1990) Comparison of the crystal structures and intersubunit interactions of human immunodeficiency and Rous sarcoma virus proteases. *Proteins* **7**, 172-184.

Weber, I. T., Tőzsér, J., Wu, J., Friedman, D., and Oroszlán, S. (1993) Molecular model of equine infectious anemia virus proteinase and kinetic measurements for peptide substrates with single amino acid substitutions. *Biochemistry* **32**, 3354-3362.

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., et al. (1995) Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* **373** (6510), 117-122.

Wiegers, K., Rutter, G., Kottler, H, Tessmer, U., Hohenberg, H., and Krausslich, H. G. (1998) Sequential steps in human immunodeficiency virus particle maturation revealed the alterations of individual Gag polyprotein processing. *J. Virol*. **72**, 2846-2854.

Wlodawer, A., Miller, M., Jaskolski, M., Satyanarayana, B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J., and Kent, S. B. H. (1989) Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. *Science* **245**, 616-621.

Wlodawer, A., and Erickson, J. W. (1993) Structure-based inhibitors of HIV-1 protease. *Annu. Rev. Biochem.* **62,** 543-585.

Wlodawer, A., Gutschina, A., Reshetnikova, L., Lubkowski, J., Zdanov, A., Hui, K. Y., Angleton, E. L., Farmerie, W. G., Goodnow, M. M., et al. (1995) Structure of an inhibitor complex of the proteinase from feline immunodeficiency virus. *Nat. Struct. Biol.* **2**, 480-488.

Wlodawer, A., and Vondrasek, J. (1998) Inhibitors of HIV-1 protease: a major succes of structure-assisted drug design. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 249-284.

Wlodawer, A., and Gutschina, A. (2000) Structural and biochemical studies of retroviral proteases. *Biochim. Biophys. Acta* **1477**, 16-34.

Wondrak, E. M., Nashed, N. T., Haber, M. T., Jerina, D. M., and Louis, J. M. (1996) A transient precursors of the HIV-1 protease. Isolation, characterization, and kinetics of maturation. *J. Mol. Biol*. **271** (8), 4477-4481.

Wu, H., Myszka, D. G., Tendian, S. W., Brouillette, C. G., Sweet, R.W., Chaiken, I. M., and Hendrickson, W. A. (1996) *Proc. Natl. Acad. Sci. USA* **93** (26), 15030-15035.

Wu, J., Adomat, J. M., Ridky, T. W., Louis, J. M., Leis, J., Harrison, R. W., and Weber, I. T. (1998) Structural basis for specificity of retroviral proteases. *Biochemistry* **37**, 4518-4526.

Wu, J., Schiffer, C. A., Gonzales, M. J., Taylor, J., Kantor, R., Chou, S., Israelski, D., Zolopa, A. R., Fessel, W. J., and Schafer, R. W. (2003) Mutation patters and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments. *J. Virol.* **77** (8), 4836-4847.

Wyatt, R., and Sodroski, J. (1998) The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science* **280**, 1884-1888.

Yamazaki, T., Hinck, A. P., Wang, Y. X., Nicholson, L. K., Torchia, D. A., Wingfield, P., Stahl, S. J., Kaufman, J. D., Chang, C. H., Domaille, P. J., and Lam, P. Y. (1996) Three-dimensional solution structure of the HIV-1 protease complexed with DMP323, a novel cyclic urea-type inhibitor, determined by nuclear magnetic resonance spectroscopy. *Protein Sci.* **5**(3), 495-506.

Yap, M. W., Lindeman, D., Stanke, N., Reh, J., Westphal, D., Hanenberg, H., Ohkura, S., and Stoye, J. P. (2008) Restriction of Foamy Viruses by Primate Trim5α. *J.Virol.* **82**(11), 5429-5439.

Yoshinaka, Y., Katoh, I., Coopeland, T. D., and Oroszlán, S. J. (1985) Murine leukemia virus protease is encoded by the gag-pol gene and is synthesized through suppression of an amber termination codon. *Proc. Natl. Acad. Sci. USA* **82**(6), 1618-1622.

Yu, S. F., Baldwin, D. N., Gwynn, S. R., Yendapalli, S., and Linial, M. L. (1996) Human foamy virus replication: a pathway distinct from that of retroviruses and hepadnaviruses. *Science* **271**, 1579-1582.

Yu, S. F., Sullivan, M. D., and Linial, M. L. (1999) Evidence that the human foamy virus genome is DNA. *J. of Virol*. **73**, 1565-1572.

Yu, L., Garwicz, D., and Rögnvaldsson, T. (2005) Comprehensive Bioinformatic Analysis of the Specificity of Human Immunodeficiency Virus Type 1 Protease. *J. Virol.* **79** (19): 12477-12486.

Zabranska, H., Tuma, R., Kluh, I., Svatos, A., Ruml T., Hrabal, R., and Pichova, I. (2007) The role of the S-S bridge in retroviral protease function and virion maturation. *J. Mol. Biol.* **365**, 1493-1504.

Zabransky, A., Andreansky, A. M., Hruskova-Heidingsfeldova, O., Havlicek, V., Hunter, E., Ruml, T., and Pichova, I. (1998) Three active forms of aspartic proteinase from Mason-Pfizer monkey virus. *Virology* **245**, 250-256.

Zahuczky, G., Boross, P., Bagossi, P., Emri, G., Copeland, T. D., Oroszlán, S., Louis, J. M., and Tőzsér, J. (2000) Cloning of the bovine leukemia virus proteinase in *Escherichia coli* and comparison of its specificity to that of human T-cell leukemia virus proteinase. *Biochem. Biophys. Acta* **1478**, 1-8.

Zamyatin, A. A. (1972) Protein volume in solution. *Prog. Biophys. Mol. Biol.* **24**, 107-123.

Zemba, M., Wilk, T., Rutten, T., Wagner, A., Flugel, R. M., and Lochelt, M. (1998) The carboxy-terminal p3-Gag domain of the human foamy virus gag precursor is required for efficient virus infectivity. *Virology* **247**, 7-13.

Zennou, V., Mammano, F., Paulous, S., Mathez, D., and Clavel, F. (1998) Loss of viral fitness associated with multiple Gag and Gag-Pol processing defects in human deficiency virus type 1 variants selected for resistance to protease inhibitors in vivo. *J. Virol.* **72**, 3300-3306.

Zhang, Y. M., Imamichi, H., Imamichi, T., Lane, H. C., Fallon, J., Vasudevachari, M. B., and Salzman, N. P. (1997) Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites. *J. Virol.* **71,** 6662-6670.

Zhao, B., Winborne, E., Minnich, M. D., Culp, J. S., Debouck, C., and Abdel-Meguid, S. S. (1993) Three-dimensional structure of a simian immunodeficiency virus protease/inhibitor complex. Implications for the design of human immunodeficiency virus type 1 and 2 protease inhibitors. *Biochemistry* **32** (48), 13054-13060.

8. LIST OF PUBLICATIONS

UNIVERSITY AND NATIONAL LIBRARY UNIVERSITY OF DEBRECEN
KENÉZY LIFE SCIENCES LIBRARY

Register Number: DEENKÉTK/270/2012.
Item Number:
Subject: Ph.D. List of Publications

Candidate: Enikő Helga Eizert

Neptun ID: H4C01T

Doctoral School: Doctoral School of Molecular Cell and Immune Biology

List of publications related to the dissertation

1. **Eizert, E.H.**, Bander, P., Bagossi, P., Sperka, T., Miklóssy, G., Boross, P., Weber, I.T., Tőzsér, J.: Amino acid preferences of retroviral proteases for amino-terminal positions in a type-1 cleavage site.
J. Virol. 82 (20), 10111-10117, 2008.
DOI: http://dx.doi.org/10.1128/j-vi.00418-08
IF:5.308

2. Sperka, T., Boross, P., **Eizert, E.H.**, Tőzsér, J., Bagossi, P.: Effect of mutations on the dimer stability and the pH optimum of the human foamy virus protease.
Protein Eng. Des. Sel. 19 (8), 369-375, 2006.
DOI: http://dx.doi.org/10.1093/protein/gzl021
IF:3

H-4032 Debrecen, Egyetem tér 1. e-mail: publikaciok@lib.unideb.hu

List of other publications

3. Tárkányi, I., Horváth, A., Szatmári, I., **Eizert, E.H.**, Vámosi, G., Damjanovich, S., Ségal-Bendirdjian, E., Aradi, J.: Inhibition of human telomerase by oligonucleotide chimeras, composed of an antisense moiety and a chemically modified homo-oligonucleotide.
FEBS Lett. 579 (6), 1411-1416, 2005.
DOI: http://dx.doi.org/10.1016/j.febslet.2005.01.041
IF:3.415

Total IF: 11.723
Total IF (publications related to the dissertation): 8.308

The Candidate's publication data submitted to the Publication Database of the University of Debrecen have been validated by Kenezy Life Sciences Library on the basis of Web of Science, Scopus and Journal Citation Report (Impact Factor) databases.

10 September, 2012

9. KEYWORDS: retroviral proteases, oligopeptide substrate, enzyme kinetics, substrate specificity, pH optimum, dimer stability, molecular model

10. ACKNOWLEDGEMENTS

I would like to first thank my superviser Prof. Dr. József Tőzsér for all of his patience, encouragement, critical remarks and guidance. He has been so supportive throughout my degree period and after it.

I am very thankful to Dr. János Aradi for my acceptance and accommodation in the Biochemistry Department and basic research teaching. He has encouraged me so much with his enthusiasm to science and brilliant research ideas. I have learned so much from both of you, which can be of benefit for my career as well as for my whole life.

I would also like to thank Prof. Dr. László Fésüs for admitting me into the Department of Biochemistry and Molecular Biology program and providing me the precious opportunity to complete this degree, especially for giving me the freedom to choose the research area I am interested in.

Thank you, all my labmates, Krisztina Matúz, Pálma Bander, Gabriella Miklóssy, Péter Bagossi, Tamás Sperka, Péter Boross and János Kádas. All of you made the lab such an exciting place and I have had a great time with all of you. Krisztina, you always listen to me, share with me all the happiness and sadness and have helped me so much with my accommodation in the retroviral lab. Tamás taught me all the basics of retroviral protease assay and retroviral protease purification. Peti kept all computers working properly and he made the molecular modeling. It really helped me saving a lot of time. Great thanks to Péter, he taught me everything about protein studies. Thanks for Gabriella for teaching me all the basic lab techniques and cell culturing.

The support from my family, especially my parents who have encouraged me all the time, is greatly appreciated. Robert, my husband, has helped me so much on everything; I am particularly thankful for the unconditioned love and understanding. I am so thankful to you for not only being the greatest life-partner, but also for being the best friend of mine. My little angels, Christian and Hanna, thanks for your innocent smiles which gave me strength to hold the line in writing the thesis.

Thanks for all the people, old and new, who were devoted to the Biochemistry Department by surrounding me with love and care, and much help in this period.

The Hungarian Ministry of Education is also acknowledged for granting me a Márton Áron scholarship and I am so thankful for choosing me in this Ph.D. program. I must thank the funding from Richter Gedeon pharmaceutical company for supporting my work.

11. LIST OF PRESENTATIONS

1. Presentations and posters related to the thesis:

J., Tőzsér, **H., Eizert**, T., Sperka, J., Kádas, G., Miklóssy, P., Boross, and P., Bagossi (2006) „Comparison of specificity of retroviral proteases" - Foundation of Hungarian Biochemistry. Annual Meeting, Pécs

H., Eizert (2006) „Comparison of substrate specificity of retroviral proteases on a critical substrate binding subsite in type 1 cleavage sites" - Scientific Meeting of Ph.D. and TDK students, Debrecen

H., Eizert, P., Bagossi, T., Sperka, A., Fehér, J., Kádas, G., Zahuczky, G., Miklóssy, P., Boross, and J., Tőzsér (2005) „Amino acid preferences for P1 and P4 sites of retroviral proteases in type 1 cleavage sites" - 30th FEBS Congress and 9th IUBMB Conference. The Protein World, Budapest

H., Eizert (2005) „Comparison of specificity of retroviral proteases" - Ph.D. Meeting of Hungarians over the border on Science Day at MÁSZ (Márton Áron Szakkollégium), Budapest

H., Eizert, P., Bagossi, T., Sperka, A., Fehér, J., Kádas, G., Zahuczky, G., Miklóssy, P., Boross, S., Oroszlán, and J., Tőzsér (2005) „Amino acid preferences for substrate binding subsites of retroviral proteases in type 1 cleavage sites" - Sixth DRP Symposium, Antiviral Drug Resistance, Chantilly, Virginia, USA

2. Other presentations and posters:

D., Popa, O., Micle, R., Iovan, **H., Rosmann**, and P., Marusca (2011) „Aspects of etiology and treatment of oropharyngeal candidiasis" - Romanian Meeting of Medicine Laboratories, Piatra Neamt, Romania

I., Tárkányi, A., Horváth, I., Szatmári, **H., Eizert.,** G., Vámosi, S., Damjanovics, E., Ségal-Bendirdjiian, and J., Aradi (2005) „Inhibition of human telomerase by oligonucleotide chimeras

composed of an antisense moiety and a chemically modified homo-oligonucleotide" - Cold Spring Harbor Meeting. Telomeres & Telomerase, New York, USA

H., Eizert (2005) „Cell surface tioredoxin's covalent interaction with chemically modified oligonucleotides" - Scientific Meeting for Ph.D. and TDK students, Debrecen

H., Eizert, A., Horváth, J., Szöllősi, and J., Aradi (2004) „Cell surface tioredoxin's covalent interaction with chemically modified oligonucleotides" - Scientific Days for Ph.D. students, Budapest