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

Kinetic analysis of retroviral proteases

by

Helga Eizert (Rosmann)

Supervisor: Prof. Dr. József Tőzsér, D.Sc.

UNIVERSITY OF DEBRECEN
DOCTORAL SCHOOL OF MOLECULAR CELL AND IMMUNE BIOLOGY

DEBRECEN, 2013

KINETIC ANALYSIS OF RETROVIRAL PROTEASES

Short thesis for the degree of doctor of philosophy (Ph.D.) in theoretical medical sciences

by **Helga Eizert (Rosmann)**, M.Sc.

Supervisor: Prof. Dr. József Tőzsér, D.Sc.

Doctoral School of Molecular Cell and Immune Biology, University of Debrecen

*Head of the **Examination Committee***:	Prof. Dr. Lajos Gergely, M.D., D.Sc.
Members of the Examination Committee:	Dr. József Kónya, M.D., Ph.D. Dr. Mihály Kovács, D.Sc.

The examination takes place at the Department of Medical Microbiology, Medical and Health Science Center, University of Debrecen, 14th of October 2013.

*Head of the **Defense Committee***:	Prof. Dr. Lajos Gergely, M.D., D.Sc.
Reviewers:	Prof. Dr. Viktor Dombrádi, D.Sc. Dr. Péter Gál, Ph.D
Members of the Defense Committee:	Dr. József Kónya, M.D., Ph.D. Dr. Mihály Kovács, D.Sc.

The Ph.D. Defense takes place at the Lecture Hall of Building A, Department of Internal Medicine, Medical and Health Science Center, University of Debrecen
October 14, 2013, 13:00

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 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 and in zoonotically infected humans. Some inherent properties of FV vectors set them favorably apart from orthoretroviral vectors and ask for additional basic research on the viruses and on the application in gene therapy. The foamy virus PR is essential for viral infectivity, since mutation of the active site Asp residues resulted in non-infectious virions, as previously found for HIV-1 PR. Previously, HFV PR was

cloned in fusion with maltose binding protein (MBP) and characterized the fusion protein. 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. The obtained catalytic constants for HFV PR were much lower than those were previously determined for various retroviral proteases. The pH optimum of HFV PR was much higher than those values published for HIV-1 PR, depending on the used substrate, ionic strength and other experimental conditions. Furthermore, the dimer stability of HFV PR was much lower than that of HIV-1 PR. 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

Retroviruses in general

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. 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. 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 subdivided into seven groups defined by evolutionary relatedness (ICTV database).

All retroviruses contain three major coding regions with information for virion proteins: *gag*, which directs the synthesis of internal virion proteins matrix, capsid, and nucleoprotein; *pol*, which contains the information for the reverse transcriptase and integrase enzymes; and *env*, from which are derived the surface and transmembrane components of the viral envelope protein. An additional, smaller, coding region present in all retroviruses is *pro*, which encodes the virion protease.

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 (deltavirus), viral gene expression and replication are very poor in both tissue cultures and infected humans. HTLV causes highly malignant T-cell lymphomas in a small proportion of infected humans. 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 (ALV) 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 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. That gives a “plus” for studying the HFV and to the importance to know its function better. EIAV is a lentivirus with strong homology to the immunodeficiency viruses; it is an important model for HIV. 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. 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 leucosis, a disease characterized by the occurrence of clonal lymphoid tumors of B-cell origin. WDSV is a piscine retrovirus associated with skin tumors in walleyes.

Life cycle of Retroviruses

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. 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 double stranded DNA by reverse transcriptase, 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 us 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. 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.

All the retroviruses follow a similar life cycle including entry, reverse transcription, integration, translation, assembly and budding. 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. 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 forms with a condense core.

Foamy viruses are complex retroviruses whose replication cycle resembles that of hepadnaviruses, which exploits a late RT step and the utilization of different promoters for RNAs encoding virion-associated proteins and accessory proteins. In contrast to orthoretroviruses, the FV PR

cleaves the cognate Gag protein only once prior to or during budding. 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. Foamy virus genome contains the *pro* gene that codes for PR in the *pol* reading frame. Only the integrase domain is cleaved off from Pol resulting in a mature RT harboring the PR domain at the N-terminus (PR-RT). FV particles resemble to the immature forms of conventional retroviruses, suggesting the absence of Gag cleavage and consequent CA rearrangement in the extracellular phase of the life cycle. 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 PR-dependent disassembly pathway during entry into the target cell. 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). FV particles contain both viral RNA and DNA at ratio of about 5:1. 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.

Retroviral proteases

Retroviral proteases are key enzymes in viral propagation and are initially synthesized with other viral proteins as polyprotein precursors (Gag and Gag-Pro-Pol) that are subsequently cleaved by the viral protease at specific sites to produce mature, functional units. 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. 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. 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 PRs are significantly similar, particularly in the locations that are important in preserving both structure and function. 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 proteases, which have only a single flap, the active retroviral PRs possess 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.

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 orthoretrovirus proteases. In avian retroviruses and retrotransposons, the Thr is also replaced by a Ser residue.

Understanding the substrate specificity of HIV PRs is important for studying the molecular basis of drug resistance and development of new drugs. For the optimal catalysis, the minimal length of substrates is 7 amino acids and that is determined by the ability of substrate amino acid 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. To understand the similarities and differences among the specificities of retroviral proteases, to better understand the structure-activity relationship molecular models were built for all of the studied enzymes.

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. The mutations in PR alter single or multiple residues which can cause multidrug resistance and cross-resistance. Mutations in 45 out of 99 residues of the PR have been associated with the treatment with protease inhibitors. In some cases, the observed structural changes in mutations are in agreement with kinetic and stability changes.

3. RATIONALE AND SPECIFIC AIMS

Foamy virus is a model system for gene transfer. In comparison with HIV-1 and MLV-based vectors they are capable of similar or higher gene transfer efficiency. Our aim was to analyze the structure-function relationship in the foamy protease because knowing it better would help in understanding the unique replication cycle of the virus. 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 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. 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 rapidly development of drug resistance variants. 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. Previously, a large series of peptides containing single amino acid substitutions in the P4-P3' 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*. 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 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.

4. MATERIALS AND METHODS

Retroviral proteases

Purified retroviral proteases were prepared as described previously: HIV-2 PR, MPMV PR, MMTV PR, MMLV PR, HTLV-1 PR, BLV PR, WDSV PR and HFV PR (wild-type and mutants MBP fusion proteins) (Q8R, H22L, S25T, T28D, Q8R/T28D, H22L/T28D).

Oligopeptides

Oligopeptides were synthesized by solid-phase peptide synthesis on a model 430A automated peptide synthesizer or a semiautomatic Vega peptide synthesizer. Prior to use, all the oligopeptides were purified by RP-HPLC (reversed-phase high performance liquid chromatography). 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 (DTT)). 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), and a HFV oligopeptide substrate (SRAVN↓TVTQS).

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 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. Substrates and the cleavage products were separated using acetonitrile gradient (0-100%) in water in the presence of 0.05% TFA. Cleavage of peptides was monitored and detected by reversed-phase HPLC at 206 nm, the peak areas were integrated, and than the kinetic parameters were determined. Relative activities were calculated from the molar amounts 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 (Val-Ser-Gln-Val-Tyr↓Pro-Ile-Val-Gln). The relative activities for the HIV-1 PR, EIAV PR, and AMV PR have been reported previously. Measurements were performed in duplicate and the average values were calculated. The

standard error was less than 20%. Previous studies also indicated a strong correlation between relative activities and the specificity constants; therefore, the determined activity values can be considered as a measure of the k_{cat}/K_m values.

Preparation of the HFV PR mutants

The pMBP-HFV PR clone was used as a template for mutagenesis. The plasmid DNA (p13HFV) encoding PR was used with oligonucleotide primers with the base corresponding to the mutation to generate the mutant construct. Mutants were generated by the Quick-Change mutagenesis protocol with the appropiate oligonucleotide pairs.

Expression of the wild-type and mutant human foamy virus proteases

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. 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 and the supernatant was loaded on a column containing amylose resin. 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 using bovine serum albumin (BSA) as a standard protein. Purity of the protease was assesed by SDS-PAGE, using 10-20% gradient gels. The proteases were estimated to be 90% pure based on Coomassie staining.

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% TFA, and an aliquot was analyzed by reversed-phase high-performance liquid chromatography. Cleavage products of PR-catalyzed hydrolysis were previously identified by amino

acid analysis for wild type HFV PR and mutant enzymes produced the same cleavage fragments as indicated by identical retention times. The Fig. P program was used for determination of the kinetic parameters (K_m and k_{cat}).

Determination of pH optimum of HFV PRs

The pH optimum of the enzymes was determined in META buffer with pH in the range 3-9. Reactions 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.

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.

Sequence alignment and molecular modeling of retroviral PRs

Structure-based alignment of the HIV-1, HIV-2, SIV (simian immunodeficiency virus), 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. The initial multiple sequence alignment was made by ClustalW, structural alignment was made by Whatif, followed by manual corrections based on the structural alignment. In the building of the molecular models we have used crystal structures of known proteases: HIV-1, HIV-2, EIAV, FIV and RSV. Homologous model of HFV was built by Modeller from the crystal structure of HIV-1 PR complexed with an inhibitor. A model of VSQNY↓PIVQ oligopeptide was docked into the substrate binding site of each retroviral protease, the minimization and analysis procedure were applied as described previously with the help of Sybyl program package run on Silicon Graphics Fuel computer graphics system. Volume of the amino acid residues was retrieved from the literature.

5. RESULTS AND DISCUSSION

5.1. CHARACTERIZATION OF HFV PROTEASES

Site-directed mutagenesis and molecular modeling of HFV PR

Mutant forms of human foamy virus protease (Q8R, H22L, S25T, T28D, Q8R-T28D, H22L-T28D) were designed 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 PR. The mutants were chosen conform sequence alignment and verified using the homologous model. We replaced selected amino acid residues being close to the catalytic aspartates, which may form the protease subsites.

The crystal structure of HIV-1 PR complexed with an inhibitor was the basis for the model of HFV PR. The Modeller program was used to build the initial model of all studied retroviral 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. The HFV PR molecular modeling revealed 47% of sequence similarity with the HIV-1 PR in the substrate binding region. The HFV and HIV-1 proteases have different lengths; the HIV-1 PR is with 99 residues, while HFV PR has 125 residues. In spite of the differences in length, we predicted that the foamy PR model share the conserved core region of HIV-1 PR. Sequence comparison of HIV-1 PR with HFV PR revealed a sequence identity of about 23% and similarity of about 30%. The overall structure was expected to be similar, but specific structural features of HFV PR remain unpredictable until a crystal structure of the foamy PR will be solved.

Purification of HFV PRs

With the use of protein expression and affinity chromatography purification protocol we obtained pure and active enzymes. The specific activity of the fusion protein was similar to that of the processed enzyme. So the more convenient purified fusion proteins were used in all further experiments. The fusion proteins were purified by affinity chromatography on amylose resin. The fractions eluted from the amylose column showing high absorbance were collected and the fusion foamy virus protein was precipitated by ammonium-sulfate (4 M final concentration), on ice. The precipitate was dissolved in lysis buffer containing 0.1% β-mercaptoethanol, on ice. The activities of the foamy proteases were measured

immediately after incubation on ice because staying at -20°C made them less active. The purification protocol was successfully applied to the mutants and wild-type unprocessed enzymes, either.

Variation of catalytic constants between the mutants and wild-type HFV PRs

The activity of mutant MBP-HFV PRs was compared with that obtained for the wild-type fusion protein. The specificity constant of S25T mutant was the same as the wild-type, but the values of the other single mutants were lower than that of the wild-type enzyme, respectively. The specificity constant of the double mutants: H22L-T28D mutant was also lower, but the corresponding value of Q8R-T28D mutant was higher than that of wild-type fusion protein. 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. 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. The small variation of the apparent 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 correctly folded enzyme.

pH optimum of HFV PRs

We have determined the pH profile of the wild-type and mutant fusion enzymes. 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. 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. The increased pH optimum might be a consequence of increased dimer stability. The dimerization of HIV-1 PR is strongly dependent on the pH, forming less stable dimers at higher pH.

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). 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 S25T mutation, in good agreement with the suggested role of this residue in

dimerization. The urea stability of the H22L mutant was increased at both pH, as expected. The T28D and the Q8R single mutants showed higher stability against urea than the wild-type enzyme. T28D mutant was more stable than the double mutant Q8R-T28D enzyme, which suggests that other structural features may also play a role. 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.

5.2 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 sets based on the naturally occurring type 1 cleavage site between MA and CA proteins of HIV-1. 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 one exception of P3-Val; therefore, this enzyme was omitted from further analysis. This work together with a previous one 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 constant; therefore the relative activities determined was considered as measure of the specificity constants in this study.

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 residues may depend on the residues forming S1 subsite and may also be a function of the residues at P3. 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. 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, which showed the highest preference for Leu side chain at this position.

Test of P3 specificity

S3 subsite 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. Unlike the specificity of the S1 site, various residues were observed as the preferred ones when the S3 binding sites were mapped. 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. The specificity at S3 also appears to correlate with the phylogenetic tree: 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. 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.

Crystal structures and molecular models of PRs suggested that S3 subsites are generally large. Consequently, the P3 side chain may be positioned to interact with the hydrophobic internal residues of the enzyme. These pockets can accept various side chains, and the variation of activity is relatively small as compared to neighborhood subsites, S4 and S2.

Test of P4 specificity

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.

Similarly to S3, again, various residues were found to fit preferably to the S4 sites of the PRs. In some cases small, even polar residues were preferred, like in case of primate lentiviral PRs, as HIV-1 PR, preferring Gly and Ser. In contrast, hydrophobic residues at P4 form better substrates for EIAV than for HIV PRs, due to the presence of the additional flap residues 50-52 that contribute to the S4 subsite. Other enzymes were not very selective, able to accept various types of hydrophobic or polar residues, suggesting that they have a more enclosed pocket. 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 an 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. A side chain at P4 may either interact with the solvent molecules or may form interactions with residues of the binding pockets.

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 a broad spectrum inhibitors, from which multiple viral strains and even different retroviruses can not escape by mutations.

7. Keywords: retroviral protease, oligopeptide substrate, enzyme kinetics, substrate specificity, pH optimum, dimer stability, molecular model, foamy virus

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

DEBRECENI EGYETEM KENÉZY ÉLETTUDOMÁNYI KÖNYVTÁRA EGYETEMI ÉS NEMZETI KÖNYVTÁR

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. 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" - Foundatioan 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 for 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. students 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