

SHORT THESIS FOR DEGREE OF DOCTOR OF PHILOSOPHY(PhD)

by
Krisztina Matúz

Biochemical characterization of the xenotropic murine
leukemia virus-related virus (XMRV)

Supervisor: József Tőzsér PhD, DSc



**UNIVERSITY OF DEBRECEN
DOCTORAL SCHOOL OF MOLECULAR CELL AND IMMUNE BIOLOGY
DEBRECEN, 2013**

**Biochemical characterization of the xenotropic murine leukemia virus-related virus
(XMRV)**

Dissertation for doctoral (PhD) degree
in the discipline of theoretical medicine

by **Krisztina Matúz**
chemist

Supervisor: **József Tózsér PhD, DSc**

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

Head of the Examination Committee: Lajos Gergely PhD, DSc

Members of the Examination Committee: László Nyitray PhD, DSc
József Kónya PhD

The Examination takes place at Library of Department of Medical Microbiology, Faculty of Medicine, University of Debrecen on May 7th 2015 at 11 AM.

Head of the Defense Committee: Lajos Gergely PhD, DSc

Reviewers: Éva Bakó PhD
István Venekei PhD

Members of the Defense Committee: Lajos Gergely PhD, DSc
László Nyitray PhD, DSc
József Kónya PhD
Éva Bakó PhD
István Venekei PhD

The PhD defense takes place at Lecture Hall of Department of Pathology, Faculty of Medicine, University of Debrecen on May 7th 2015 at 1 PM.

1. INTRODUCTION

The XMRV virus isolated from human samples in 2006 was related to the chronic fatigue syndrome and prostate cancer thus it was in the lime-light of scientific interest. However, it has been proved recently that these assumptions are not correct. However, according to the latest findings, XMRV may be a cause of human tumor in mice, thus it can be assumed that the virus might as well cause human diseases. Besides, being a virus related to the xenotropic murine leukemia virus, XMRV is able to function as the 'helper virus' of vectors based on the widespread murine leukemia virus (MLV) thus there is a danger that the gene therapeutic vector, getting out of the target cells, will enter other cells or, in extreme cases, the human population as well. Thus detection of the potential therapeutic targets of XMRV seems to be justified. As in the case of every retrovirus, the viral protease (PR) enzyme which performs cleavage of viral precursor proteins into functional parts also plays a key role in the life cycle of XMRV therefore the protease is an important chemotherapeutic target. As opposed to inhibitors used in the clinical practice, a resistance is developed in a short time in the case of the HIV-1 protease, namely mutations occur which correspond to amino acid parts found in other retroviral proteases in similar sites.

Because of the use of XMRV as a gene therapeutic vector and its role played in the animal model of the leukemia virus, characterization of the XMRV protease became the focus of our examinations.

As in the case of every retrovirus, the viral protease (PR) plays a key role in the life-cycle of the XMRV as well, which performs the cleavage of viral precursor proteins into functional parts, thus the protease is an important target of chemotherapy. With the help of inhibitors designed against proteases, the unwanted physiological effects of those can be prevented or reduced.

Our task force is focused on the examination of retroviral proteases: in the case of several proteases, the kinetic parameters of the enzymes, the inhibition profile of the enzyme and the substrate specificity of the proteases were examined by the liquid chromatographic method. The experimental results were supported both with bioinformatic tools and structural biological data.

The murine leukemia virus (MLV) protease can be well used as an animal model of the retroviruses and its role played in gene therapy is also significant as among retroviral vectors used in clinical practice, together with those based on the HIV-1, vectors developed by the modification of the murine leukemia virus are the most significant. Therefore our task force studied the MLV protease in detail. We had the homology model of the proteinase which was built upon the chrystal structure of the HIV-1 proteinase with the help of the Modeler 3 program.

I started examining the protease of the XMRV as a newly identified virus in 2011, which was later identified as an artefact created in a laboratory. The chrystal structure of the XMRV protease formed with an inhibitor was also unknown. As a collaboration, we examined the biochemical characteristics of the XMRV protease produced and purified by the staff of the Macromolecular Chrystallographic Laboratory (National Cancer Institute – Frederick National Laboratory of Cancer Research, USA) lead by the American scientist, Alexander Wlodawer. Examination of the chrystal structure of the inhibitor complexes of the XMRV protease was performed by our collaboration partner.

As the XMRV sequence shows a very high level of correspondence with the sequence of the MLV proteinase, the chrystal structure of which protease has not been defined yet, thus definition of the chrystal structure of the XMRV proteinase is also significant with respect to structural information related to the MLV protease. We wished to contribute to the development of much more specific inhibitors for the XMRV and gammaretrovirus subfamily by a detailed knowledge of the chrystal structure of the enzyme-inhibitor, the mapping of the amino-acids forming the substrate-binding pockets and the examination of the interactions.

During my doctoral studies, I examined the kinetic parameters of the XMRV protease and the inhibition profile of the enzyme. As a support for experimental work, we used a bioinformatic method, in which we were assisted by Dr. János Mótyán, and we tried to interpret the binding method of the inhibitor in the case of two classical aspartile protease inhibitors (pepstatine A and acety-pepstatine) together with the X-ray chrystallographic results, and compare it with the binding method of the HIV-1 protease inhibitor.

2. THEORETICAL BACKGROUND

2.1. General description of retroviruses

Retroviruses are viruses having a positive diploid RNA genome which belong to the *Retroviridae* family. It is proved scientifically that they are capable of infecting any class of vertebrata, and beyond passive viremia, the consequences of infection may be the occurrence malignancy, neurological disorders, anemia or immune deficiency.

The diameter of retroviral virions varies between 80 and 100 nm, and viral glycoproteins can be found in their outer lipid envelope (Env). Envelope proteins contain a heterodimeric triplet which consists of a surface receptor binding subunit (SU) and a trans-membrane subunit (TM) being in a non-covalent interaction with this. The inner protein nucleus is built of matrix (MA) and capsid (CA) proteins, their shapes and positioning is different in various species of the family. The nucleus contains a viral genome associated with the nucleocapsid (NC) protein which also encodes 3 enzymes necessary for the life cycle: protease (PR), reverse transcriptase (RT) and integrase (IN) (Coffin, 1992, 1997).

According to the organization of the viral genome, we can differentiate between simple and complex retroviruses. In the case of the virus having the simplest viral genome, the murine leukemia virus (MLV), only three genes coded by the virus are necessary for retroviral replication (Varmus and Brown, 1989). *Gag*, which encodes structural proteins of the virion, *pol* which encodes viral enzymes, and *env* which encodes the envelope protein. These three genes can always be found in the same order in the integrated provirus ($5'$ -*gag-pol-env*- $3'$) and repeating terminal sequences (LTR, long terminal repeat) of specific length occurring during reverse transcription can be found at both ends. LTR contains promoter and enhancer elements necessary for the efficient transcription of the retroviral genome and sequences necessary for the polyadenylation of the mRNA at the $3'$ -terminal. In the case of simple retroviruses, two different transcripts are formed from the integrated provirus: genomic RNA, which also serves as an mRNA for the synthesis of the Gag and Pol proteins, and a smaller inserted mRNA which encodes the Env protein. Viral gene expression is jointly regulated by *cis*-activating viral DNA or RNA sequences and *trans*-activating factors encoded by the host cell.

In the case of the virus related to the xenotropic murine leukemia virus (XMRV), which was isolated in 2006, the genomic structure is very similar to that of the MLV.

2.2. The life-cycle of retroviruses

The life-cycles of retroviruses can be divided into two phases which we call early and late phases. As the first step of the early phase the virus enters the host cell via membrane fusion or receptor-mediated endocytosis. Reverse transcription takes place in the entering capsid structure where the positive single-stranded RNA genome is transcribed into a double-stranded DNA genome. A preintegration complex (PIC) is formed from the genomic DNA and some protein components of the entering capsid, which enters the nucleus. This is a passive step in the case of most retroviruses, thus

they are only capable of infecting dividing cells the nuclear membrane of which is not unharmed; however in the case of HIV-1 and other retroviruses belonging to the *Lentivirus* subfamily active transport of PIC enables the infection of non-dividing cells as well. Viral DNA is built in the genome of the host cell with the help of IN being an essential part of PIC. The first step of the late phase is transcription of the viral DNA which takes place with the mediation of cellular RNA-polimerase II. A portion of the mRNA molecules formed in this way leaves the nucleus unmodified and will be a template of the Gag and Gag-Pro-Pol poliproteins during the translation, another portion gets into the virus envelope and serves as the genetic material of the genome of the successor viruses. *Env*-coded proteins are translated from a smaller inserted mRNA which first become glycosylated then, during transportation into the plasma membrane, they will split into a surface (SU) and a transmembrane (TM) protein under the influence of a cellular protease. Several multiply inserted mRNAs serve as a template for the synthesis of supplementary proteins, such as Tat, Rev, Vif, Nef and Vpr. Gag proteins accumulate in the Env protein rich parts of the host cell membrane, on the inner surface of the membrane, then the 'immature' virus particle having a 'doughnut-shaped' capsid structure leaves the cell via 'unbinding' (budding). The virus will become 'mature' and virulent after the activation the protease cleaving the poliproteins at fixed sites. In this case the virus particle has a solid conical inner structure in the case of HIV-1, while it has an icosahedral structure in the case of MLV. As only 'mature' virus particles are virulent, functioning of the PR is essential during virus replication.

2.3. Characterization of the HIV-1 protease

Retroviral proteases are proteins consisting of 99-138 amino acid parts with a molecular weight of 11-15 kDa which show several features characteristic of aspartic proteases, such as inhibition by pepstatin and enzyme inactivation induced by the mutation of catalytic aspartil. However, as opposed to the single-chain cellular aspartic proteases carrying two domains topologically similar but not completely identical at the same time, retroviral proteases are enzymes built of two identical subsites functioning in a dimeric form. The primary and secondary structures of retroviral proteases are analogous to one of the domains of the cellular aspartic proteases: they contain several β -folds and one or two short α -helices per subsite depending on the enzyme. The N- and C-terminal chains of the two subsites form a four-layered anti-parallel β -fold spliced.

HIV-1 protease can be characterized by three specific regions: active center, the so-called flap and the conserved region situated near the C-terminal. The catalytic triade (Asp-Thr-Gly) encoding the

active center is situated near the N-terminal. The catalytic triades of the subsites form a loop which is analogous to the ψ -structure described with cellular aspartic proteases. Catalytic triplets are bound to each other through a network of hydrogen bonds. The flexible flap region is more or less conservative which moves at the binding of the substrate and the inhibitor and bends on the ligand, thus stabilizing the complex. The third conserved region (Gly-Arg-Asn) is situated near the C-terminal and plays an important role in dimerization by forming ionic pairs.

2.4. Endogén retroviruses

Endogenous retroviruses are endogenous viral elements that are derived from retroviruses, and they are abundant in the genomes of jawed vertebrates. The replication cycle of a retrovirus entails the insertion ("integration") of a DNA copy of the viral genome into the nuclear genome of the host cell. Most retroviruses infect somatic cells, but occasional infection of germline cells (cells that produce eggs and sperm) can also occur. Rarely, retroviral integration may occur in a germline cell that goes on to develop into a viable organism. This organism will carry the inserted retroviral genome as an integral part of its own genome - an "endogenous" retrovirus (ERV) that may be inherited by its offspring as a novel allele. Many ERVs have persisted in the genome of their hosts for millions of years. However, most of these have acquired inactivating mutations during host DNA replication and are no longer capable of producing virus. ERVs can also be partially excised from the genome by a process known as recombinational deletion, in which recombination between the identical sequences that flank newly integrated retroviruses results in deletion of the internal, protein-coding regions of the viral genome. Endogenous retroviruses can play an active role in shaping genomes. Most studies in this area have focused on the genomes of humans and higher primates, but other vertebrates, such as mice and sheep, have also been studied in depth. The long terminal repeat (LTR) sequences that flank ERV genomes frequently act as alternate promoters and enhancers, often contributing to the transcriptome by producing tissue-specific variants. In addition, the retroviral proteins themselves have been co-opted to serve novel host functions, particularly in reproduction and development. Recombination between homologous retroviral sequences has also contributed to gene shuffling and the generation of genetic variation. The majority of ERVs that occur in vertebrate genomes are ancient, inactivated by mutation, and have reached genetic fixation in their host species.

For these reasons, they are extremely unlikely to have negative effects on their hosts except under unusual circumstances. Nevertheless, it is clear from studies in birds and non-human mammal species

including mice, cats and koalas, that younger (i.e., more recently integrated) ERVs can be associated with disease. This has led researchers to propose a role for ERVs in several forms of human cancer and autoimmune disease, although conclusive evidence is lacking. In humans, ERVs have been proposed to be involved in multiple sclerosis (MS). In 2004 it was reported that antibodies to HERVs were found in greater frequency in the sera of people with schizophrenia. Human endogenous retrovirus (HERV) proviruses comprise a significant part of the human genome, with approximately 98,000 ERV elements and fragments making up nearly 8%.

This is because most HERVs are merely traces of original viruses, having first integrated millions of years ago. However, one family of viruses has been active since the divergence of humans and chimpanzees. Immunological studies have shown some evidence for T cell immune responses against HERVs in HIV-infected individuals.

The hypothesis that HIV induces HERV expression in HIV-infected cells led to the proposal that a vaccine targeting HERV antigens could specifically eliminate HIV-infected cells.

2.5. Characterization of the virus related to the xenotropic murine leukemia virus (XMRV)

Xenotropic murine leukemia virus-related virus (XMRV) is a laboratory-derived gammaretrovirus, that arose from the recombination of two endogenous mouse retroviruses during the middle of 1990s. The chimeric virus was first described in 2006 as an apparently novel retrovirus and potential human pathogen. Initial reports linked the virus to prostate cancer and later to chronic fatigue syndrome (CFS), leading to considerable interest in the scientific and patient communities, investigation of XMRV as a potential cause of multiple medical conditions, and public health concerns about the safety of the donated blood supply. XMRV has now been established as a laboratory contaminant. False positive detection of XMRV may also occur because of contamination of clinical specimens and laboratory reagents with other mouse retroviruses or related nucleic acids. There is no evidence that XMRV can infect humans, nor has it been demonstrated that XMRV is associated with or causes any human disease.

XMRV is a murine leukemia virus (MLV) that formed through the recombination of the genomes of two parent MLVs known as preXMRV-1 and preXMRV-2. MLVs belong to the virus family Retroviridae and the genus gammaretrovirus and have a single-stranded RNA genome that replicates through a DNA intermediate. The name XMRV was given because the discoverers of the virus initially thought that it was a novel potential human pathogen that was related to but distinct from MLVs.

However, numerous studies have failed to find evidence of XMRV in prostate cancer. Researchers in Germany found no XMRV-specific sequences in the DNA or RNA of samples from prostate cancer patients, and no XMRV-specific antibodies were detected in blood serum samples. Another German study found no XMRV association with non-familial (sporadic) prostate cancer, and no XMRV link was found in Irish prostate cancer patients with the R462Q mutation. Similar results were reported in The Netherlands, Japan, and Mexico. Geographical differences were initially suspected as the culprit in these discrepancies, because the positive results were obtained in the United States. However, U.S. studies have also found no evidence of XMRV. Researchers at the National Cancer Institute, Johns Hopkins and the Mayo Clinic tested over 1,000 samples from prostate cancer patients but found no evidence of the virus. A causal role of XMRV in cancer has not been established, and XMRV does not appear to be capable of transforming cells directly. In prostate cancer, XMRV protein has been found in tumour-associated but nonmalignant stromal cells, but in one study it was not found in the actual prostate cancer cells, raising the possibility that the virus may indirectly support tumorigenesis. However, in another study, XMRV proteins and nucleic acids were found in malignant cells. XMRV has been proposed as a cause for conditions including autism, fibromyalgia, multiple sclerosis, amyotrophic lateral sclerosis, and Parkinson's disease. However, there is currently no evidence to support these hypotheses. No XMRV was found in 230 autistic children or in 204 controls, nor was XMRV detected in fibromyalgia or MS patients or in children with idiopathic diseases. The results of several studies support the contamination hypothesis. Scientists reported that they found ample sources of contamination in laboratory reagents, tissue samples and blood. Samples that were positive for XMRV were also positive for mouse DNA contaminants.

2.6. OBJECTIVES

We had the following objectives in my doctoral work:

Characterization of the XMRV protease enzyme expressed by *E.coli* cells, definition of the kinetic parameters of the enzyme with the help of the liquid chromatographic method, examination of dimerization and dissociation ability of the enzyme and comparison of that with the parameters of the HIV-1 protease.

Examination of the inhibition profile of the enzyme with various anti-protease inhibitors using two different methods: with the help of gel electrophoresis in a medium with a low ionic strength and with the help of a liquid chromatographic method in a medium with a high ionic strength.

We wished to examine the enzyme inhibition of the XMRV protease in the case of pepstatine A and acety-pepstatine inhibitors. We complemented our experimental work with molecular modeling and we wished to interpret the enzyme-inhibitor interactions together with the data derived from the X-ray crystallographic examinations of XMRV PR-inhibitor complexes submitted to us by our American collaboration partner.

We wished to compare the inhibition method of the XMRV and the HIV-1 enzyme in the case of the two classical aspartate protease inhibitors (pepstatine A and acety-pepstatine) and interpret the interactions.

3. MATERIALS AND METHODS

3.1 Purification of the XMRV and HIV-1 protease enzymes and the MLV Gag fragment substrate

The purified XMRV protease was submitted to us by Alexander Wlodawer and his colleagues. We produced the HIV-1 proteases with the help of *E.coli* cells described above. Expression of the stabilized HIV-1 proteases was carried out in *E.coli* BL21(DE3) cultures carrying the appropriate plasmids, growing them at 37 °C in a Luria-Bertani nutrient solution containing 100 µg/ml ampicillin then we induced expression with IPTG for 3.5 hours. We accumulated the cells by centrifuging, we digested the pellet and by sonication on ice. We centrifuged the lysatum, then we suspended it with a lysis puffer containing 3 M urea and sonicated it again. We solved the pellet gained after centrifuging in a denaturing buffer containing 8M urea then, after another centrifuging, we filtered the materials floating at the top of the buffer through a filter with a pore size of 0.22 µm. We purified the proteins on a POROS 20 R2 chromatographic column with HPLC in the presence of 0.05v/v% TFA applying a linear acetonitrile-water gradient (0-100%). We checked the purity of the fractions containing protease on 16% polyacryamide gels.

We produced the MLV Gag fragment substrate (MLV Gag Δ 2) in the expression system and corresponding to the purification protocol published by our task force. We transformed the MLV Gag Δ 2 construction cloned in a pET23b vector provided with His₆ at the C-terminal end in *E.coli* BL32(DE3) competent cells with the help of a thermal shock. We made the cells grow at 37 °C in a Luria-Bertani nutrient solution containing 100 µg/ml ampicillin then we induced expression with IPTG for 2 hours. We accumulated the cells by centrifuging, then after removing the materials floating at the top of the buffer, we picked up the pellet in a lysis puffer and we digested it by sonication on ice. We centrifuged

the lysatum, then we filtered the materials floating at the top of the buffer through a filter with a pore size of 0.22 μm and we purified them with nickel-chelate affinity chromatography applying 250 mM imidazol gradient. We separated the proteins by gel filtering with the help of a Superdex TM 30 column. We determined the efficiency of expression with the Bradford spectrophotometric measuring method.

3.2 Examination of activity

For measuring protease activity of XMRV we used RSLLY↓PALTP synthetic oligopeptide substrate. The substrate is a P3 Leu-substituted decapeptide which contained the sequence (RSLLY↓PALTP) of the MLV MA/p12 cleavage site. We determined the quantity of active enzymes by active center titration. In these experiments we measured with different concentrations inhibitor solutions and in the reaction compound. We used amprenavir inhibitor for the titration of the HIV-1 and XMRV PR active centers. We selected the applied substrate concentration range depending on the K_M value. We incubated the reaction compounds for 1 hour at 37 °C, then we stopped the enzyme reaction with 180 μl 1v/v% TFA-solution and we mounted the solutions gained in this way on a Nova-Pak C₁₈ reverse phase chromatographic column with the help of an automatic injector. We separated the substrate from the products and the puffer components with a linear acetonitrile-water gradient (0-100%) in the presence of 0.05v/v% TFA. We followed up separation on 206 nm and we calculated the extent of hydrolysis by integrating the chromatographic curve. We set enzyme concentrations in a way that hydrolysis of the substrates should remain under 20%. We defined kinetic parameters by adjusting the reaction rate and the substrate concentration data to the Michaelis-Menten equation, with the method of non-linear regression, with the help of the SigmaPlot 8.0 program. The standard deviations of the kinetic constants were under 20%. We calculated the catalytic constant by using the active enzyme concentrations gained by active center titration.

3.3. Examination of dimerization and urea dissociation

In the case of the XMRV and the HIV-1 proteases, we defined the virtual dimerization values (K_{dapp}) with the help of activity measurement using a KARVnL↓NphEAnL-NH₂ substrate. The reaction compound contained different concentration of enzyme dissection. We incubated the samples at 37 °C- on for 20 minutes, then we stopped the reaction with a 1v/v% TFA solution and we subjected the

reaction compounds to an HPLC analysis. We depicted relative specific activity with the help of the SigmaPlot 8.0 program geared to protease concentration, then we applied non-linear insertion.

We performed the stability examination of active dimer proteases with the help of urea denaturation. We also measured the UC_{50} value of the dissociation constant (urea concentration measured in the case of 50% dissociation) with the help of the HPLC method, applying a KARVNle↓F(NO₂)EAL-NH₂ substrate, the concentration of urea was 0-4 M. After the HPLC analysis of the products we depicted relative specific activity geared to urea concentration, then we applied a non-linear insertion with the help of the SigmaPlot 8.0 software.

3.4. Examinations of inhibition

3.4.1. Examinations of inhibitions with the HPLC method

For measuring protease activity of XMRV we used RSLLY↓PALTP synthetic oligopeptide substrate. In the case of HIV-1 protease we used VSQNY↓PIVQ substrate. During the inhibition experiments, we applied different inhibitor solutions, acetyl-pepstatin, pepstatin A, amprenavir, TL-3 and atazanavir, which solved with DMSO, and we put DMSO into the compound instead of the inhibitor as a control. We incubated the reactions at 37°C for 1 hour while shaking, then we stopped them by adding 1v/v% trifluoroacetic acid (TFA). We changed the concentration of the enzyme in a way that the conversion of the substrate should be under 20%. We separated the products occurring during the enzyme reaction with the help of reverse-phase chromatography. We identified the cleavage products on the basis of their retention times. We defined the K_i values with the help of the IC₅₀ values, on the basis of the $K_i = (IC_{50} - [E]/2) / (1 + [S]/K_m)$ relationship, where [E] and [S] are protease and substrate concentrations. We defined kinetic parameters by adjusting the reaction rate and the substrate concentration data to the Michaelis-Menten equation, with the method of non-linear regression, with the help of the SigmaPlot program.

3.4.2. Examinations of inhibition with the SDS PAGE method

We incubated the recombinant MLV Gag Δ2 fragment (Δp12-CA-NC-ΔPR) substrate for 1 hour at 37°C, the cleavage compound contained amprenavir or TL-3 solved in DMSO. During the inhibition experiments performed with acetyl-pepstatin, pepstatin A, acetyl-pepstatin and pepstatin A inhibitors. In the case of the control samples either the inhibitor solution or the enzyme dissection was missing

from the compound. We stopped the enzyme reactions by adding an SDS-PAGE sample loading buffer and we denatured the proteins at 95°C for 8 minutes, then we separated the products with the help of SDS-PAGE and identified them with the help of the Protein ladder protein molecule standard after Coomassie Brilliant Blue painting.

3.5. Molecular modeling

We performed structure-based multiple sequence insertion for the HIV-1, MLV and the XMRV PR sequences. We used the crystal structure data of the HIV-1 PR and acetyl-pepstatin complex (PDB ID: 5HVP) and the XMRV PR and pepstatin A complex (PDB ID: 3SM1) to build a 3D model of the complexes of the HIV-1 and the XMRV PR-s with one or two pepstatin A and acetyl-pepstatin inhibitor molecules with the help of the SYBYL program package. The starting structures of the enzyme-inhibitor complexes were produced by inserting the inhibitor molecules into active sites. At the binding of a single inhibitor molecule, we removed the catalytic water molecule and substituted it with the hydroxyl group of the central statin derivative. We left the catalytic water in the case of the binding of two inhibitor molecules. We refined the position of the inhibitors with a minimalization procedure with the help of the SYBYL program package. We calculated the interaction energy between the enzyme and the inhibitor molecules for each substrate binding site. Calculations and molecular displays were performed on the Silicon Graphics Fuel workstation.

3.6. X-ray crystallography

X-ray crystallographic examination of the XMRV PR was performed by Alexander Wlodawer and his colleagues. In the case of TL-3 and pepstatin, inhibitors were added in a 4:1 XMRV protease (monomer): inhibitor mole proportion; in the case of amprenavir this proportion was 1:1, acetyl-pepstatin was used in a 4:1 PR: inhibitor mole proportion. Each crystallization was performed with the hanging drop vapor diffusion method. The crystals of the XMRV PR/TL-3 complex were growing at pH 5.5 (PBD: 3SLZ) , while crystals of the XMRV PR/pepstatin A complex at pH 7.0 (PBD: 3SM1), and the crystals of the amprenavir complex at pH 4.75 (PBD kód: 3SM2). Calculations were performed with the help of the PyMol program.

4. RESULTS

4.1. Enzyme kinetic and stability examinations

4.1.1. Kinetic examination of the XMRV protease with the HPLC method

We defined the kinetic parameters of the XMRV protease. We followed up enzyme cleavage on a synthetic decapeptide substrate, which contained the sequence of the MLV MA/p12 cleavage site (RSLLY↓PALTP). Kinetic parameters of the XMRV protease are the following: $K_m = 0.216 \pm 0.027$ mM, $k_{cat} = 0.55 \pm 0.04$ s⁻¹, and $k_{cat}/K_m = 2.55 \pm 0.37$ mM⁻¹ s⁻¹.

4.1.2. Dimerization and urea dissociation examination of the XMRV and the HIV-1 proteases

We defined the value of the apparent dimerization constant (K_d) of the XMRV PR as 115 nM with the application of the KARVnL↓NphEAnL-NH₂ chromogenic substrate; we compared it with the value of the HIV-1 protease defined with the HPLC method and we found that in the case of the HIV-1 protease it dramatically decreased to the value of 1.0 nM, which is in accordance with the values defined earlier in literature.

When defining the stability of the dimers, we also examined the urea denaturation ability of the enzymes with the application of a chromogenic substrate and with the HPLC method. We found the value of the HIV-1 PR urea-dissociation constant (UD_{50}) similar to that defined photometrically earlier in literature; however the XMRV PR proved to be much more sensitive to urea concentration, which is in accordance with the high K_{dapp} value.

4.2. Examinations of inhibition

4.2.1. Examinations of inhibition with the HPLC method

On the basis of the K_i values, amprenavir proved to be the most efficient inhibitor among the tested inhibitors in the case of the XMRV PR, we also performed active center titration of the enzyme with this HIV-1 protease inhibitor. Pepstatin A proved to be the least efficient inhibitor, it did not function as an analogous XMRV PR inhibitor with a temporary state, which means that binding of the inhibitor to the enzyme takes place not in the traditional way.

Acetyl-pepstatin proved to be a more suitable inhibitor in the case of the HIV-1 protease than pepstatin A. Acetyl-pepstatin inhibited both the XMRV and the HIV-1 protease effectively, the

apparent K_i value was significantly lower in the case of the HIV-1 than in the case of the XMRV protease. The HIV-1 protease was inhibited by both inhibitors in a nanomolar concentration.

4.2.2. Examinations of inhibition with the SDS PAGE method

We examined the activity of the XMRV protease on a recombinant MLV Gag fragment substrate (MLVGag Δ 2, Δ p12-CA-NC- Δ PR) as well, which contains p12/CA, CA/NC and NC/PR cleavage sites. At the cleavage of the recombinant protein with the XMRV protease we stipulated the presence of the CA (31 kDa) and the Δ p12-CA (34 kDa) fragments. Amprenavir also proved to be a more efficient inhibitor at the cleavage of the MLV Gag fragment substrate than TL-3.

The XMRV protease is relatively strongly inhibited by acetyl-pepstatin as compared to pepstatin A in the case of applying a recombinant substrate. We compared the inhibition measured with the HPLC method in a buffer with a high ionic strength (2 M NaCl) with the inhibition concentration defined in a buffer with a low ionic strength (examined with an SDS-PAGE analysis) and we found that 3 μ M acetyl-pepstatin was enough for a nearly 50% inhibition of the protein cleavage of the enzyme, in the case of pepstatin A this value was μ M.

4.3. Molecular modeling

During calculations of interaction energy, we examine enzyme-inhibitor interactions and explain the binding of pepstatin A and the acetyl-pepstatin inhibitors to the HIV-1 PR and the XMRV PR in different ways. We studied two binding methods: in the first one we modeled the binding of one inhibitor, as we found out in the crystal structure of the HIV-1 PR/acetyl-pepstatin, while in the second binding method two molecules are simultaneously bound to the enzyme opposite each other, as it can be seen in the crystal structure of XMRV PR/pepstatin A. A single molecule of acetyl-pepstatin operates as a temporary state analogue in the case of the HIV-1 PR where the hydroxyl group of the central statin (Sta) derivate is situated among the catalytic aspartils, substituting the catalytic water molecule.

The crystal structure of the complex of the HIV-1 protease with the acetyl-pepstatin inhibitor is known, and it was also successfully crystallized with pepstatin A. Binding of the enzyme to pepstatin A takes place in a unique way, it may be accomplished in various ways but it differs from the binding method of other inhibitors of the XMRV protease.

In the case of the binding of pepstatin A to the XMRV PR, none of the inhibitor molecules binding to an enzyme functions as a temporary state analogue. Pepstatin A and acetyl-pepstatin molecules only differ in their N-terminal endings. We calculated the interaction energies for both binding methods of the inhibitors to the enzymes. The calculated individual subsite interaction energies only showed a significant difference in the case of the two inhibitors where the acetyl (Ace) or isovaleryl (Iva) groups enter into an interaction with the enzyme. Our results show that binding of the larger Iva group of pepstatin A resulted in a greater interaction energy in all of the examined binding methods than binding of the smaller Ace group of acetyl-pepstatin, in accordance with the generally hydrophobic nature of the enzyme-ligand interactions. As opposed to this, enzyme kinetic examinations showed a greater inhibition effect for both kind of PR in the case of acetyl-pepstatin as compared to pepstatin A. Therefore better binding of acetyl-pepstatin to retroviral proteases as compared to pepstatin A cannot be explained only by taking interaction energies into consideration. Simulation of the binding of the two pepstatin A molecules to the XMRV PR resulted in a significantly greater enzyme-inhibitor interaction energy as compared to the binding of a single molecule, while preference for a single inhibitor binding could be observed in the case of the HIV-1 PR. In the case of the XMRV PR, preference for the binding of two inhibitors could only be seen in the case of the pepstatin A inhibitor, while simulation of the binding of acetyl-pepstatin resulted in approximately the same interaction energy in the different binding methods.

Results of the structural and biochemical examinations of the XMRV PR show that as compared to other retropepsins despite the significant differences in enzyme topology, interactions of the enzyme with the well-characterized inhibitors of this enzyme class are similar both in a structural and a kinetic sense, expect for the unusual binding method of pepstatin A. Detailed description of the substrate binding pockets for the subfamily including XMRV, MLV and similar retroviruses may contribute to the development of much more specific inhibitors.

Both the MLV (XMRV) and the HIV-1 PR-s preserve the mainly hydrophobic character of the substrate binding sites. However, differences were also described in substrate specificity, the reason for which is probably that other amino acid parts participate in the building of the substrate binding subsites. Bindings of the larger Sta4 groups to the S4/S4' sites are less favorable in the case of the HIV-1 PR, and this difference is also confirmed by the lower calculated interaction energies of the S4/S4' subsites as compared to the XMRV PR. While in the case of the XMRV PR the binding of two molecules of the pepstatin A inhibitor is favorable from an energetic point of view, in the case of the

XMRV PR, the binding of two molecules of acetyl-pepstatin provided very similar interaction energies due to the significantly lower S1-P1 and S1'-P1' interactions provided by the acetyl groups as compared to the isovaleryl group of pepstatin A.

4.4. X-ray crystallography

The crystal structure of the complexes of XMRV PR with aspartic protease inhibitors was examined by Wlodawer and his colleagues. The flap regions (which cover the active sites in retropepsins) which are partly disordered in the apo-enzyme, were completely ordered in the inhibitor complexes.

In the TL-3 complex of the symmetrical C2 inhibitor, the inhibitor molecule is bound to the protease dimer in a canonically extended conformation. Electron density is very square corresponding to the inhibitor and TL-3 apparently binds mainly in one direction and only a low binding rate (20%) can be observed in the other direction. Although the inhibitor is symmetrical, its method of binding to the enzyme is not symmetrical.

However the binding method of pepstatin A is very unusual. Instead of the binding of a single pepstatin A molecule to the enzyme, two molecules are bound with the N-terminal isovaleryl groups to both sides of the catalytic aspartil near the aspartils. Only 4 amino acid parts of pepstatin A are ordered in both molecules, and the C-terminal parts of the inhibitors are disordered, which cannot be seen in the electron density. The catalytic water molecule which can usually be found between the two aspartils in the structures of the non-inhibited proteases is also present, although a single molecule of TL-3 or amprenavir is bound to the enzyme, while pepstatin A – with two molecules, but the interactions between the inhibitors and the enzyme are remarkably similar.

In the case of the complex of the XMRV PR with acetyl-pepstatin, the very weak electron density present near the catalytic aspartils suggests that the binding method should be similar to that of pepstatin A, i.e. it binds to two opposing molecules with the outer parts in a disordered manner.

5. SUMMARY

Our objective was the kinetic examination of the viral protease related to the xenotropic murine leukemia virus (XMRV) and the examination of the stability and inhibition of the enzyme with the help of the HPLC as well as the SDS-PAGE methods. We examined the kinetic parameters of the enzyme (K_m , k_{cat} , k_{cat}/K_m) with the help of a synthetic oligopeptide substrate while the stability of the enzyme, i.e. its ability to dimerize (K_{dapp}) and the tendency of the enzyme towards dissociation (UD_{50}) - in a buffer containing an incrementing urea concentration by applying a chromogenic substrate. We examined the inhibition of the protease with the help of various protease inhibitors (amprenavir, TL-3, acetyl-pepstatin and pepstatin A), while applying the HPLC method in a high salt concentration buffer, and the SDS-PAGE method in the case of a low salt concentration to follow up the inhibition measurements. In the case of the SDS-PAGE method we applied an MLV Gag fragment as a substrate. We compared the stability values of the enzyme with the values of the HIV-1 protease. We modeled the enzyme-inhibitor complex formed in this way (in the case of acetyl-pepstatin and pepstatin A) and we examined the interaction energies and the potential binding of the inhibitors.

The XMRV protease proved to be much more sensitive to the concentration of urea than the HIV-1 protease. Among the inhibitors tested on the basis of K_i values amprenavir proved to be the most efficient inhibitor while pepstatin A – the least efficient inhibitor.

On the basis of molecular modeling we established that the S4/S4' interaction also substantially contributes to the binding of the acetyl-pepstatin and pepstatin inhibitors to the HIV-1 and XMRV proteases. The two-inhibitor binding method of the XMRV protease-pepstatin A is more favorable from an energetic point of view, while in the case of acetyl-pepstatin, the interaction of the dual binding method provided similar interaction energies. The Sta6-S3' interaction is much more favorable in the case of the HIV-1 PR than in the case of the XMRV PR, thus HIV-1 prefers the single binding method.

X-ray crystallographic data were disclosed to us by our collaboration partner, on the basis of which we established that pepstatin A binds to the enzyme in a different manner as compared to other inhibitors, as two inhibitor molecules are bound to the PR dimer of XMRV while TL-3 or amprenavir with a single molecule, yet the interactions between the inhibitors and the enzyme are remarkably similar.

7. ACKNOWLEDGEMENTS

I would like to thank my supervisor, Prof. Dr. József Tőzsér, full professor for his professional directions and support given during my work.

I would like to thank Prof. Dr. László Fésüs, full professor, academic and Head of Department for enabling me to finish my studies at Debrecen University, Faculty of Medicine, Institute of Biochemistry and Molecular Biology and supporting my work throughout my studies.

I would like to thank Dr. Péter Boross, assistant lecturer for his useful advice during my practical work.

I would like to thank Dr. Tamás Sperka, Dr. János Kádas, Dr. Gabriella Miklóssy, Dr. Péter Bagossi† and Dr. János Mótyán for their professional help and Dr. Angelika Varga, Dr. Helga Eizert and Mária Golda, Dr. Mohamed Mahdi my colleagues and friends for that I could count on them any time. Besides, I would like to thank Pálma Bander, Dr. Bence Farkas, Bea Bozóki, Ferenc Tóth, Kata Nagy and Lívía Gazda for their professional and humane support.

I would like to thank Szilvia Pető and Katalin Szabó, assistants, for the numberless examples of practical help.

I would like to acknowledge the assistance of the employees of Debrecen University, Faculty of Medicine, Institute of Biochemistry and Molecular Biology and all those who contributed to the preparation of my Ph.D. theses in any form.

I would like to thank my mother, father and my husband for their support without whom all these could not have been achieved.

Register Number: DEENKÉTK/303/2012.

Item Number:

Subject: Ph.D. List of Publications

Candidate: Krisztina Matúz

Neptun ID: MUW3NK

Doctoral School: Doctoral School of Molecular Medicine

List of publications related to the dissertation

1. **Matúz, K.**, Mótyán, J.A., Li, M., Wlodawer, A., Tózsér, J.: Inhibition of XMRV and HIV-1 proteases by pepstatin A and acetyl-pepstatin.
FEBS J. 279 (17), 3276-3286, 2012.
DOI: <http://dx.doi.org/10.1111/j.1742-4658.2012.08714>
IF:3.79 (2011)
2. Li, M., Gustchina, A., **Matúz, K.**, Tózsér, J., Namwong, S., Goldfarb, N.E., Dunn, B.M., Wlodawer, A.: Structural and biochemical characterization of the inhibitor complexes of xenotropic murine leukemia virus-related virus protease.
FEBS J. 278 (22), 4413-4424, 2011.
DOI: <http://dx.doi.org/10.1111/j.1742-4658.2011.08364.x>
IF:3.79



List of other publications

3. Kádas, J., Boross, P., Weber, I.T., Bagossi, P., **Matúz, K.**, Tőzsér, J.: C-terminal residues of mature human T-lymphotropic virus type 1 protease are critical for dimerization and catalytic activity. *Biochem. J.* 416 (3), 357-364, 2008.
DOI: <http://dx.doi.org/10.1042/BJ20071132>
IF:4.371

4. Sperka, T., Miklóssy, G., Tie, Y., Bagossi, P., Zahuczky, G., Boross, P., **Matúz, K.**, Harrison, R.W., Weber, I.T., Tőzsér, J.: Bovine leukemia virus protease: Comparison with human T-lymphotropic virus and human immunodeficiency virus protease. *J. Gen. Virol.* 88 (7), 2052-2063, 2007.
DOI: <http://dx.doi.org/10.1099/vir.0.82704-0>
IF:3.12

Total IF: 15.071

Total IF (publications related to the dissertation): 7.58

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

17 October, 2012



8. APPENDIX

Posters:

Matúz K., Bagossi P., Kádas J., Tőzsér J.: Prediction of the structure of deltaretrovirus proteases with the help of mutagenesis examinations. *Traveling Session of the Hungarian Biochemical Association of the year 2006*. Pécs, 30 August-2 September, 2006.

Miklóssy G., **Matúz K.**, Kádas J., Tőzsér J., Bagossi P.: The protease inhibiting effect of trans-dominant negative HIV. *Traveling Session of the Hungarian Biochemical Association of the year 2006*. Pécs, 30 August-2 September, 2006.

Matúz K., Boross P., Bagossi P., Miklóssy G., Tőzsér J.: Examination of mutant deltaretrovirus proteases. *Traveling Session of the Hungarian Biochemical Association of the year 2007*. Debrecen, 26 August-29 August, 2007.

Miklóssy G., Kádas J., **Matúz K.**, Tőzsér J., Bagossi P.: The trans-dominant inhibition effect of mutant HIV-1 proteases. *Traveling Session of the Hungarian Biochemical Association of the year 2007*. Debrecen, 26 August-29 August, 2007.