Short thesis for the Degree of Doctor of Philosophy (PhD)

The biochemical characterization of Human T-lymphotropic leukemia virus type 1, -2 and -3 proteases and investigation of cigarette smoke toxin and investigation of pancreatic lipase mutation as risk factors for chronic pancreatitis

by Norbert Kassay

Supervisor: Prof. Dr. József Tőzsér

UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF MOLECULAR CELL AND IMMUNE BIOLOGY

Debrecen, 2024.

THE BIOCHEMICAL CHARACTERIZATION OF HUMAN T-LYMPHOTROPIC LEUKEMIA VIRUS TYPE 1, -2 AND -3 PROTEASES AND INVESTIGATION OF CIGARETTE SMOKE TOXIN AND INVESTIGATION OF PANCREATIC LIPASE MUTATION AS RISK FACTORS FOR CHRONIC PANCREATITIS

By Norbert Kassay, biotechnologist

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

Supervisor: Dr. József Tőzsér, PhD, DSc

Examination Committee:

Head:	Dr. István Balogh, PhD, DSc
Members:	Dr. Viktor Dombrádi, PhD, DSc
	Dr. Attila Ambrus, PhD, DSc

The PhD examination takes place at the Library of the Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Debrecen on 2025.02.12., at 10 a-m.

Reviewers:	Dr. Andrea Kiss, PhD
	Dr. Judit Pongrácz, PhD, DSc

Defense Committee:

Head:	Dr. István Balogh, PhD, DSc
Members:	Dr. Viktor Dombrádi, PhD, DSc
	Dr. Attila Ambrus, PhD, DSc
	Dr. Andrea Kiss, PhD
	Dr. Judit Pongrácz, PhD, DSc

The PhD Defense takes place at the Lecture Hall of Bldg. A, Department of Internal Medicine, Faculty of Medicine, University of Debrecen on 2025.02.12., at 12 a-m.

1. INTRODUCTION

1.2. Retroviruses

1.2.1 Delta retroviruses

Human T-lymphotropic viruses (HTLV) belong to the delta genus of retroviruses, together with bovine leukaemia virus (BLV) and simian T-lymphotropic viruses (STLV), forming together the group of primate T-lymphotropic viruses (PTLV). Among human pathogenic retroviruses, HTLV-1 was the first to be discovered. This was followed shortly afterwards by the isolation of HTLV-2 and human immunodeficiency virus 1 (HIV-1) of the lentiviral genus. Later, other members of the T-lymphotropic viruses, HTLV-3 and HTLV-4, were identified.

1.2.2 Structure of retroviruses

There are simple and complex retroviruses. HTLV is one of the latter and encodes accessory and regulatory proteins other than the essential gag, pol and env proteins. The 100 nm diameter virus has a lipid envelope on the surface of where heterodimers of glycoproteins encoded by the *env* gene are expressed. The *gag* gene encodes the structural matrix (MA), capsid (CA) and nucleocapsid (NC) proteins, while the *pol* gene encodes non-structural proteins such as reverse transcriptase (RT), integrase (IN) and protease (PR) enzymes. The infectious virion also contains a positive-strand viral RNA genome.

The genome of HTLV-1 is bordered by two long terminal repeats (LTRs), the 5' LTR being the major promoter, from which the Gag, Gag-Pro and Gag-Pro-Pol polyproteins are transcribed as a function of the -1 frameshift. The Gag protein, encoded by the *gag* gene, is cleaved by the retroviral PR to form functional MA, CA and NC proteins. From the Gag-Pol precursor, the nucleocapsid is replaced by the transframe protein 1 (TF), the PR and two smaller proteins p1 and p2. In addition to MA, CA, TF, PR and p1, RT, RNase H and IN are cleaved from the Gag-Pro-Pol polyprotein. The *env* gene encodes elements of the envelope glycoprotein, which is cleaved by proteolytic cleavage into transmembrane (TM) and surface subunit (SU) proteins. The proteins

encoded by the pX region, p12, p13, p30, p8, which is derived from proteolytic cleavage of p12, and the transactivator protein (tax), rex or antisense HTLV-1 bZIP factor (HBZ) proteins are expressed near the 3' LTR. The genomic organisation of HTLV-2 and HTLV-3 is identical to HTLV-1 in its major elements, with differences in the pX region.

1.2.3. Life cycle of HTLV viruses

The life cycle of HTLV viruses is highly similar, but that of HTLV-3 it is poorly understood. They primarily infect T-cells but are also able to infect other immune cells. HTLV-1 prefers CD4 T-cells, HTLV-2 prefers CD8 T-cells, while HTLV-3 can infect both cell types to a similar extent. Infection is preceded by binding to a cell surface receptor. Following fusion with the membrane, the viral RNA genome is reverse transcribed into double-stranded DNA, followed by the formation of the so-called pre-integration complex from the remaining parts of the capsid core, which can enter the nucleus and integrate into the host cell genome. Subsequently, translation and expression of the viral polyproteins is carried out via the host cell's apparatus. The gag polyproteins translocate towards the plasma membrane where they begin their oligomerization and assembly. The polyproteins are cleaved by viral PR into functional units, allowing viral maturation then the virus enters another cells typically via cell-cell contact or budding.

In the later stages of infection, clonal expansion is more common. The relatively low variability of HTLV-1 (in a single individual) also implies that it replicates by mitotic division when incorporated into genomic DNA, *i.e.* replication of viral hereditary material is carried out by cellular DNA polymerases rather than by the more error-prone reverse transcriptase.

1.2.4. HTLV protease

HTLV-1 PR was first isolated in 1989, followed by several *in vitro* studies where the cleavage of the gag polyprotein was observed and MA, CA and NC proteins were identified. Subsequently, the cleavage sites

MA/CA, CA/NC, TF1/PR, PR/p1, p1/RT, RT/RH and RT/IN were also determined.

HTLV-1 PR consist of 125 amino acids, larger than HIV-1 PR, which contains 99 amino acids. Most of the "extra" amino acids are located at the C-terminal, mainly composed of hydrophobic amino acids and is unique to deltaretroviruses. The amino acids in the C-terminal have been found to require the presence of at least five amino acids for autoprocessing (G111-L115). HTLV-1 PR expressed *in vitro* is prone to autoproteolysis and most studies have therefore investigated the L40I stabilized protein or a mutant containing additional mutations (C90A, C109A). Other studies have found similar enzyme activity to the full-length protein when testing the 115 amino acid long HTLV-1 PR with the L40I stabilising mutation and His fusion tag at the N-terminal. Later, our group also tested a wild-type PR and a protein containing stabilizing mutations (L40I, C90A and C109A) of 125, 120 and 116 amino acids in length, but only observed activity in the full-length and five amino acid truncated PR and concluded that amino acids P116-L125 may play an important role in dimer stabilization. In addition to the N- and C-terminals forming the β-terminal dimerization surface, the H-bond system formed by the D-S/T-G-A consensus active site motif amino acids, the so-called "fireman's grip", also plays an important role in the dimerization of retroviral PRs. The HTLV-1 PR contains two extra amino acids in the flap region compared to the HIV-1 PR, thus forming a more extensive substrate-binding site and breaking the hairpin structure of the flap through the disruption of hydrogen bonds, resulting in a helical structure.

There is only ~28% identity between HIV-1 and HTLV-1 PRs in terms of the overall sequence, but X-ray crystallographic analysis shows a high degree of structural similarity, due to the conservation of the active site, where there is a higher degree of identity (45%), and low variability in the amino acids involved in dimerization. The conserved amino acids may also play an important role in maintaining the quaternary structure, in the proper folding of the protein and in forming its dynamic properties.

Similar to HIV-1 PR, HTLV-1 PR does not have a consensus cleavage site sequence either, but hydrophobic side chains predominate and the majority of autoproteolytic cleavage sites are leucine at the P1

site, as in other HTLV PRs. Seven substrate binding pockets (S4-S3') are distinguished in HTLV-1 PR. The nomenclature for naming substrate amino acids and substrate binding sites is that of Schechter and Berger (1967).

The S4 substrate-binding pocket is relatively large and consists mainly of hydrophobic amino acids, in contrast to the hydrophilic subunit that is characteristic of HIV-1 PR. A common feature of retroviral PRs is that the S4 pocket is located close to the protein surface, is open and less well defined than the other pockets. The S3 pocket of HTLV-1 PR is also significantly different from HIV-1 PR. Like the S4 pocket, it is relatively open and can interact more easily with both hydrophilic and hydrophobic side chains. The substrate binding pocket of HIV-1 PR S3 is more closed and deeper, with more amino acids involved in substrate binding. In contrast, the S2 pocket has hydrophobic characteristics and is relatively small, similar to most retroviral PRs. The main difference between HIV-1 PR and HTLV-1 PR is that HTLV-1 PR contains a methionine (M37) at the position corresponding to the amino acid D30 (aspartic acid) in HIV-1 PR, a feature only found in immunodeficiency viruses (HIV, SIV), whereas most retroviral PRs have a hydrophobic side chain at this position. The S1 pocket is large in both PRs. The S1 pocket of HTLV-1 has two more amino acids, and there are two differences in the amino acids forming the pocket. In general, retroviral PRs are characterized by relatively variable substrate binding sites away from the cleavage site, both in the number and type of residues of the binding substrate. Pockets are more conserved closer to the cleavage site of HTLV-1 and HIV-1 PRs.

Most studies on retroviral proteases use oligopeptides representing natural cleavage sites as substrates. HIV-1 PR has been studied in the most detail, while among delta retroviruses only the activity of HTLV-1 and BLV PRs has been investigated so far. Current literature data show that HTLV-1 PR has a significantly narrower specificity than HIV-1 PR. In a series of assays with substrates containing retroviral PR cleavage sites, in addition to HIV-1 and HTLV-1, HIV-2, equine infectious anaemia virus (EIAV), rous sarcoma virus (RSV), Mason-Pfizer monkey virus (MMTV), BLV and murine leukaemia virus (MuLV) cleavage sites

were also investigated. HIV-1 PR cleaved most substrates, except certain MuLV and RSV cleavage sites, while HTLV-1 PR cleaved only about 40% of the substrates tested. In contrast, data from a similar substrate series showed that BLV PR has similar specificity as compared to HIV-1 PR. This finding was confirmed by experiments with the HTLV-1 CA/NC cleavage site, where the P4-P1' cleavage site contained single point mutations. Although HTLV-1 PR exhibited catalytic activity for most substrates, it was less tolerant to amino acid substitutions at the P2-P1' sites, in contrast to BLV PR. In terms of specificity, BLV PR showed higher similarity to HTLV-1 PR, cleaved HTLV-1 CA/NC oligopeptides with greater efficiency and broader specificity than HIV-1, and showed a greater preference for large hydrophobic amino acid side chains at P2 and P1 sites, similar to HTLV-1 PR.

Depending on the subtype of the virus and the quality of disease progression, diseases caused by HTLV-1 are typically treated with conventional chemotherapy, a combination of zidovudine (AZT) and interferon-alpha (IFN), or hematopoietic stem cell transplantation. Successes in HIV-1 therapy have largely been achieved using specific drugs. Such possibilities also regularly appear in research related to HTLV-1, so far with low therapeutic efficiency. On the one hand, the strategies try to prioritize the targets used successfully in anti-HIV-1 therapy also in the case of HTLV-1, i.e. to prevent the entry of the virus into the cell by targeting the envelope glycoproteins or their receptors, as well as they plan to inhibit the activity of the enzymes that play a role in the life cycle of the virus, the reverse transcriptase, integrase and protease by developing inhibitors. An essential part of anti-HIV-1 therapy is the use of protease inhibitors, which so far have not been successfully used in anti-HTLV-1 therapy, which can also be explained by the high inhibition constant (K_i) values observed in *in vitro* experiments. Based on the experience gained so far, it is necessary to develop HTLV-1 specific inhibitors.

1.3. Chronic pancreatitis and its risk factors

1.3.1. Chronic pancreatitis

Chronic pancreatitis (CP) is a chronic inflammation of the pancreas that can develop as a result of long-term inflammation. It is characterised by parenchymal or intraductal calcification, pancreatic fibrosis, exocrine and endocrine insufficiency, and is therefore associated with a deficiency or altered function of certain digestive enzymes and may also cause loss of endocrine glands. As a result, digestive difficulties, diabetes and pain may develop, causing a significant deterioration in quality of life and a reduction in life expectancy. The disease cannot be reversed, but its progression and the severity of symptoms can be reduced.

1.3.2. Hereditary pancreatitis

CP can also be triggered by genetic factors. Hereditary pancreatitis is caused by mutations in certain digestive enzymes secreted by acinar cells or in closely related proteins. The gene most affected is *PRSS1*, which encodes cationic trypsin.

Trypsin plays a central role in the function of digestive enzymes encoded by the pancreas. Activation of these enzymes occurs in the duodenum. The process begins with cleavage by enterokinase produced in the small intestine, which converts trypsinogen to trypsin.

Like trypsin, other digestive enzymes are secreted in an inactive, zymogenic form, containing an activation peptide at their N-terminal as well. Subsequently, the trypsin is also capable of activating trypsinogen. Enterokinase has a narrow specificity, activating only trypsinogen among the digestive enzymes, the other zymogenic proteins (chymotrypsinogen, carboxypeptidase, proelastase, prophospholipase, kallikreinogen, procolipase) are activated by trypsin and are not capable of autoactivation. In addition to activation, trypsin function is also regulated by the SPINK1 inhibitor and chymotrypsinogen C (CTRC), also produced by acinar cells.

1.3.3. Endoplasmic reticulum stress

Mutations in the digestive enzymes produced by pancreatic acinar cells can also cause inappropriate protein folding, leading to a secretion defect, which may be associated with the accumulation of proteins within the cell and the resulting proteins may be involved in proteasomal or autophagy pathways. In most cases, the production of these proteins is also associated with an increase in ER stress. The resulting cellular stress response can be described as the unfolded protein response (UPR), a signalling pathway that aims to restore cellular homeostasis, but prolonged unresolved stress can lead to cell death.

There are three main pathways of UPR, "inositol requiring enzyme-1" (IRE1), "protein kinase RNA-like ER kinase" (PERK) and "activating transcription factor 6" (ATF6). They have in common that they are in the ER membrane as a complex and bind a "binding immunoglobulin protein" (BiP), which becomes free upon activation of the pathways and can act as a chaperone to bind to misfolded proteins.

ER stress can also be activated by reactive oxygen species (ROS). ROS are normally also produced in cells and accumulate mainly in the mitochondria and ER. Protein unfolding occurs in the ER, as does disulfide bond formation, which is associated with ROS formation. The process starts with the oxidation of the protein disulfide isomerase (PDI) by ERO1, followed by the reaction of the thiol group of the unfolded protein with the oxidized PDI. ERO1 uses molecular oxygen as an acceptor, resulting in the formation of hydrogen peroxide. Reduced PDI can be oxidized not only by ERO1 but also by GSSG, depending on the GSH:GSSG ratio. PDI does not only cover the PDI or PDI1 disulfide isomerase, it is a family of enzymes with mechanisms other than the function described here, a significant proportion of which also act as chaperones, which may explain their elevated levels during ER stress, and some members (ERP57, ERP72, P5) are presented in quantities close to those of PDI.

UPR and oxidative stress are closely related. Persistently high levels of ROS can cause oxidation of proteins and thus poor folding,

inducing a stronger response. Inappropriate folding of certain mutant proteins can trigger the UPR response, re-engaging in oxidative folding processes, thus depleting the GSH pool and leading to ROS accumulation. Their close relationship is further demonstrated by the fact that PDI is an essential activator of the PERK pathway and ERP72 is required for the ATF6 UPR response.

Nuclear factor erythroid 2-related factor 2 (Nrf2) plays a central role in the regulation of ER stress and oxidative stress. It can be activated in the UPR pathway by phosphorylation of PERK via ATF4, JNK proteins, and also in response to oxidative stress. As a transcription factor, it induces transcription of genes involved in redox homeostasis. One of these is NAD(P)H quinone dehydrogenase 1 (NQO1), which functions as a reductase in response to stress by using NAD(P)H as a donor via FAD in its active site. In addition, it is known to play a role in the stabilisation of certain proteins and is also involved in their regulation by binding the mRNA of certain proteins.

1.3.4. Aberrant protein pathway

A defect in the folding of normally secreted proteins produced by the pancreas is most often associated with activation of the UPR response, which can lead to an increase in CHOP transcription and apoptosis.

Although most cationic trypsin variants are properly secreted, they are also capable of causing pancreatitis by this pathway. Some non-secreted mutants, in addition to inducing ER stress, are autoactivated within the cell, inducing programmed cell death.

In addition to trypsin, other digestive enzymes are also affected by mutations resulting in inappropriately folded protein. However, the effects of the resulting proteins are independent of trypsin and are primarily attributable to the UPR. These include CTRC, CPA1, CEL, but the folding defect may also affect other secretory proteins, and the severity of the stress response may be related to the amount of protein produced. The SPINK-1 mutants are an exception to this group, as they are non-functional inhibitors unable to inhibit trypsin activation,

however, the G48E, D50E, Y54H and R67C mutants have not yet been associated with pancreatitis.

1.3.5. Human pancreatic lipase

Human pancreatic lipase (PNLIP) is secreted in large amounts in the pancreas, similar to cationic trypsin, and catalyses the hydrolysis of triglycerides to fatty acids in the small intestine, efficiently cleaving both short and long chain triglycerides. It is assisted by colipase, but has the opposite effect to bile acids, which can inhibit lipase activity. Although colipase and bile acids are not essential, they are required for the efficient digestion and absorption of triglycerides.

There are several known mutations in PNLIP. Some variants cause lipase deficiency, which is associated with impaired absorption of lipids from food. It has not yet been clearly linked to CP, but several mutations carry a high risk. Some variants have higher activity, others are not secreted properly. The latter mutations, with the exception of T221M, have only been detected in heterozygous form and have not been associated with CP. Some mutations may make the PNLIP protein more sensitive to degradation by PRSS1 and CTRC, and such protein variants are found in some patients with non-alcoholic CP, but the link between CP and mutations has not yet been explored. Non-secreted PNLIP variants (A174P, T221M, G233E, C254R and V454F) have also been shown to increase ER stress. Several of these mutations are located near the active site, Gly 233 being part of the ß9 loop. Previous experiments have investigated their effects on ER *in vitro* in HEK and AR42J cell lines. For all five protein variants, an increase in XBP-1 mRNA maturation was observed, as well as an increase in BiP mRNA levels. These effects were more expressed in the AR42J rat acinar cell line, which may be explained by the relatively extensive ER of the cell.

2. OBJECTIVES

To date, several studies have investigated the differences between HTLV-1 and HTLV-2 viruses. These have mainly focused on differences in regulatory proteins, signalling pathways and pathogenic outcomes. Only limited information is available on HTLV-3 and the differences between HTLV proteases have not been discussed in previous publications. Although there is a considerable amount of experimental data available for HTLV-1 PR, the prevalence of HTLV-2 and the health risk of HTLV-3 has made it necessary to investigate HTLV-2 and HTLV-3 PR. The main objectives identified in our work were:

- to optimise the expression and purification of HTLV-2 and HTLV-3 PRs

- investigate the substrate specificity and susceptibility for inhibitors of HTLV-1, HTLV-2 and HTLV-3 proteases

- investigating the autoprocessing ability of HTLV-2 and HTLV-3 proteases

Inherited chronic pancreatitis can be caused by mutations in proteins that are not properly folded and secreted. Smoking may be present as an independent risk factor. It was hypothesised that the co-existence of smoking and genetic risk factors may reinforce each other. A poorly studied human pancreatic lipase (PNLIP) variant (G233E) was used as a model for genetic factors. We aimed to certify our hypothesis using a human (HEK293AD) and a rat pancreatic cell line (AR42J). Our aim was:

- to study the effect of major cigarette smoke components and PNLIP mutation on cell viability

- to examine endoplasmic reticulum stress markers

- to investigate the association between HQ and the G233E PNLIP mutation

3. MATERIALS AND METHODS

3.1 Characterisation of HTLV proteases

3.1.1 Expression of HTLV-1, HTLV-2 and HTLV-3 proteases

For the expression of HTLV-1 PR, an expression plasmid (pET11a) coding for a stabilized version of the enzyme (that contains L40I, C90A and C109A mutations) was used. For HTLV-2 PR expression, a bacterial expression plasmid containing the codon-optimized sequence of the protease was designed (Genscript Biotech). For HTLV-3 PR, a codon-optimized sequence of the Pyl43 strain was used. The sequence of the protease was cloned from pCR2.1-TOP0 plasmid into pET11a expression vector using NdeI and BamHI restriction endonucleases.

The expression plasmids were transformed into BL21(DE3) *E. coli* cells by heat shock. The transformed cells were grown in ampicillin-supplemented Luria-Bertani (LB) medium, and the protein expression was induced by the addition of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG).

3.1.2 Purification of HTLV-1, HTLV-2 and HTLV-3 proteases

The cells expressing the recombinant proteins were harvested by centrifugation and then dissolved in buffer A (50 mM Tris-HCl, 1 mM dithiothreitol (DTT), 1 mM ethylenediaminetetraacetic acid (EDTA), pH 8.2) and lysed by sonication, followed by ultracentrifugation. The pellet containing the inclusion bodies was then dissolved in buffer B (50 mM Tris-HCl, 1 mM DTT, 1 mM EDTA, 1 v/v % Triton X-100, pH 8.2). This purification step was not included in the method described previously for HTLV-1 PR. After two centrifugation steps, the pellet was dissolved in buffer C (50 mM Tris-HCl, 1 mM DTT, 1 mM EDTA, 1 v/v % Triton X-100, 1 M urea, pH 8.2) and finally solubilized in buffer D (50 mM Tris-HCl, 7.5 M guanidine-HCl, 5 mM DTT, 5 mM EDTA, pH 8.2). Proteins were separated by SDS-PAGE.

After solubilisation, proteins were purified by reversed-phase HPLC on a POROS R2 column in an Äkta Purifier. Separation was

performed using an increasing water/acetonitrile (0-100%) gradient in the presence of 0.05% trifluoroacetate (TFA). The purity of the eluted fractions was checked by SDS-PAGE. The fractions with the highest purity (>90%) were further processed.

Proper folding of the enzymes was ensured by dialysis against buffer E (20 mM piperazine-N,N'-bis(2-ethanesulfonic acid) (PIPES), 1 mM EDTA, 100 mM NaCl, 10 v/v % glycerol, 0.5 v/v % NP-40, 5 mM DTT, pH 7.0).

3.1.3. Oligopeptides and inhibitors

Synthetic oligopeptides representing HTLV-1, HTLV-2 and HTLV-3 PR cleavage sites were used as substrates, such as MA/CA, CA/NC, TF1/PR and PR/p1 cleavage sites, as well as wild-type and P4, P3, P2, P1 or P1′ site modified HTLV-1 MA/CA (KTKVL*VVQPK) cleavage sites.

IB-268 (KTKVL-r-VVQPK) and IB-269 (APQVL-r-PVMHP) inhibitors containing reduced peptide bonds were synthesized and provided by Dr Ivo Blaha (Ferring Leciva, Prague, Czech Republic). DMP-323 (a tight-binding inhibitor of HIV-1 PR) and inhibitors currently or previously used in anti-HIV-1 therapy (atazanavir, darunavir, indinavir, ritonavir and saquinavir) were in-house stocks.

3.1.4. HTLV-1, HTLV-2 and HTLV-3 protease activity assays

Kinetic studies were performed using 10 µl of 2x incubation buffer (0.5 M K_3PO_4, 4 M NaCl, 10 v/v % glycerol, 10 mM DTT, pH 5.6), 5 µl of HTLV-1, HTLV-2 or HTLV-3 protease and 0.5-5 µl of oligopeptide substrate representing a natural HTLV PR cleavage site. The final volume of the reaction was 20 µl. The reactions were initiated by addition of the enzyme and incubated at 37°C for 0.5-4 h and stopped by addition of TFA.

For the substrates containing HIV-1, EIAV, RSV, MMTV, MPMV, MuLV or BLV protease cleavage sites, a longer incubation time (24 h) was used.

The cleavage products and substrate were separated on a Nova-Pack C18 RP-HPLC column on a LaChrom HPLC instrument. Separation was performed using a 0-100% water-acetonitrile gradient in the presence of 0.05% TFA.

Kinetic parameters were determined based on the peak area values calculated for the substrate and the resulting products by fitting the data to the Michaelis-Menten equation. A Gauss equation was used to determine the optimum pH, while linear regression was used to determine the optimum NaCl concentration and temperature. To calculate the catalytic constants k_{cat}/K_M, the concentration of the active enzyme was determined by active centre titration using a previously described method.

3.1.5 Determination of optimal pH, temperature and ionic strength for HTLV-1, HTLV-2, and HTLV-3 proteases

To investigate the effects of pH, temperature and ionic strength, enzyme reactions were carried out in 2x META buffer (100 mM MES [2-(N-morpholino)-ethanesulfonic acid], 200 mM Tris, 100 mM sodium acetate). The reaction mixtures contained 10 µl buffer, 5 µl enzyme and 5 µl substrate. Reactions were initiated by addition of the enzyme, cleavage reactions were carried out as described in the HTLV-1, HTLV-2 and HTLV-3 protease activity assays section.

Oligopeptides representing HTLV-1 PR/P1, HTLV-2 PR/P1 and HTLV-3 TF1/PR cleavage sites were used as substrates for HTLV-1, HTLV-2, and HTLV-3 PRs. The effect of NaCl was investigated in the 0-2 M concentration range, buffers with pH 4.5-8.0 were used to determine the pH optimum, while the determination of the temperature optimum was investigated in the 20-40°C temperature range.

3.1.6 Determination of amino acid preferences of HTLV-2 and HTLV-3 proteases

Substrates representing the wild-type HTLV-1 PR CA/NC cleavage site (KTKVL*VVQPK) and its P4, P3, P2, P1 and P1′-modified

variants were used to compare the amino acid preferences of HTLV-2 and HTLV-3 PR to those of HTLV-1 PR. To study whether the S5 and S4 pockets are involved in substrate recognition, besides the previously described substrates (P5-P5'), shorter substrate variants (P4-P5′ and P3-P5′) were also tested. Enzyme reactions were performed as described in section HTLV-1, HTLV-2 and HTLV-3 protease activity assays. The 20 µl reaction mixture contained 10 µl of 2x incubation buffer, 5 µl of HTLV-1, HTLV-2 or HTLV-3 PR (final concentration of active enzymes varied between 0.1 and 34.4 nM) and 5 µl of substrate (final concentration 0.4-0.5 mM). The reactions were initiated by addition of the enzyme followed by incubation for 0.5-4 h at 37°C. The reactions were stopped by the addition of TFA. Relative activity values were determined for each substrate. In each case, the value obtained for the KTKVL*VVQPK substrate was considered to be 100%.

3.1.7. Inhibition of HTLV-1, HTLV-2 and HTLV-3 proteases

The reaction mixtures consisted of 4.8 µl substrate, 0.2 µl inhibitor, 5 µl enzyme (HTLV-1, HTLV-2 and HTLV-3 PR) and 10 µl 2x incubation buffer. The substrates represented HTLV-2 and HTLV-3 MA/CA cleavage sites. Concentrations of the sock solutions were 100 µM and 0.5-100 µM for HIV-1 PR as well as IB-268 and IB-269 inhibitors, respectively. The inhibitors were dissolved in DMSO and 0.2 µl of DMSO was added to the control samples. IB-269 inhibitor was used to determine the active enzyme concentration.

3.1.8 Site-directed mutagenesis of HTLV-2 and HTLV-3 proteases

The coding sequences of HTLV-2 and HTLV-3 PRs were complemented with an 8 residue-long N-terminal linker, corresponding to the P8-P1 residues of the TF1/PR natural cleavage site. The resulting sequence was cloned into a pMALc2x vector, which allowed to express the recombinant protease fused to an N-terminal maltose-binding protein (MBP). Cloning was performed by overlap extension PCR, the PCR product was digested with EcoRI and BamHI restriction endonucleases

and ligated into the pMALc2x vector. This was followed by site-directed mutagenesis of HTLV-2 (L37D, L37N, L57G, A59I and F67Q) and HTLV-3 PRs (I37D, I37N, L57G, A59I and F67Q) using QuikChange II mutagenesis kits. The success of mutagenesis was verified by sequencing.

3.1.9. Study of the autoprocessing ability of HTLV-2 and -3 proteases

Expression constructs were transformed into BL21(DE3) *E. coli* cells with heat shocks. The recombinant HTLV-2 and HTLV-3 PRs were expressed fused with an N-terminal MBP fusion tag (MBP-HTLV-2 PR and MBP-HTLV-3 PR). Expression was induced by the addition of 1 mM IPTG and cells were lysed by sonication in buffer A. Cell lysates were separated on 16% polyacrylamide gel. After SDS-PAGE, proteins were transferred onto nitrocellulose membrane. Western-blot was performed according to a protocol previously optimized by the research group. The membrane was blocked for 1 h in TTBS solution containing 5% milk powder (Tris-buffered saline, pH 7.5, 0.01% Tween-20). An anti-MBP monoclonal antibody at a dilution of 1:4000 (4°C, 16 h) was used as primary antibody, and an anti-rabbit HRP-conjugated antibody at a dilution of 1:10000 (room temperature, 1 h) as secondary antibody. The proteins were detected by chemiluminescent substrate using X-ray film using a KODAK Medical X-ray processor 102 developer.

3.2. Analysis of the main cigarette smoke components and the G233E PNLIP mutation

3.2.1 Expression plasmids and adenovirus vectors

For PNLIP expression, a previously described construct was used that coded for PNLIP fused to a C-terminal decahistidine (His_{10}) fusion tag, cloned into a pcDNA3.1(-) vector.

Adenoviruses containing the lipase coding sequence were amplified in HEK293AD human embryonic kidney cells. Adenoviruses were released from cells by freeze-thaw cycles and purified on anion-exchange columns. The columns were washed with 0.4 ml of 0.1 M sodium hydroxide and centrifuged (1 min, 2000 g), followed by another washing step with 0.4 ml of 0.1 M sodium acetate buffer (pH 5.0) (1 min, 2000 g). The columns were equilibrated with sample buffer (50 mM HEPES, 2 mM $MgCl_2$, 0.1% Tween-20, pH 8.0) and centrifuged. After re-equilibration, cell lysates containing adenoviruses were diluted 10× with sample buffer, pipetted onto the column and centrifuged (3 min, 2000 g). Columns were washed with sample buffer supplemented with 0.2 M NaCl and eluted with sample buffer containing 1 M NaCl. The solutions containing the purified adenovirus were supplemented with 0,1 ml of 50 v/v % glycerol and stored at -70°C until use. Viral titres were determined in IFU/ml (infectious unit/ml) (AdEasy Viral Titer Kit).

3.2.2 Cell culture cultures and gene transfer

HEK293AD cells were maintained in DMEM (high glucose DMEM culture medium supplemented with 10 v/v % FBS, 50 U/ml penicillin and 50 µg/ml streptomycin solutions) in cell culture medium at 37°C in a cell culture incubator. The day before the experiments, the cells were plated on a 12-well cell culture plate (180,000/well) in 0.5 ml medium. Branched-chain polyethyleneimine (PEI) solution was prepared as described previously. For transfection, 12 µl of PEI solution was mixed with 2 µg of plasmid DNA in Opti-MEM medium. After 20 min incubation (room temperature), the mixture was added to the cells and placed in a cell culture incubator. After 6 hours, the medium was

removed, the cells were washed with phosphate buffered saline (PBS) and the supernatant was replaced with Opti-MEM solution supplemented with 50 µg/ml streptomycin 50 µg/ml penicillin.

AR42J rat acinus cells were maintained in DMEM medium supplemented with high glucose 20 v/v % FBS, 50 U/ml penicillin and 50 µg/ml streptomycin at 37°C in a cell culture incubator.

The day before the experiments, cells were plated in a 12-well cell culture plate (400,000/well) in the presence of 100 nM dexamethasone. Cells were incubated for 48 h, washed with PBS solution, then Opti-MEM supplemented with 50 µg/ml of 50 U/ml penicillin and 50 µg/ml streptomycin was applied to the cells and transduced with adenovirus vectors ($5x10^7$ IFU/ml final concentration).

3.2.3. Chemical treatments

Stock solutions of acrolein and crotonaldehyde were freshly prepared before the experiments. Hydroquinone was dissolved in water and stored at -20°C until use. HEK293AD and AR42J cells were treated with solutions of 10-200 µM final concentrations.

3.2.4. Viability tests

For viability studies, MTT reagent was dissolved in PBS solution at a concentration of 5 mg/ml. The conditioned medium was removed from the cells and 1 ml of Opti-MEM culture medium supplemented with 0.1 ml of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) was added to the cells. To form formazan crystals, the cell culture plate was incubated at 37°C (10-60 min). The solution was then replaced with DMSO (0.2 ml) and suspended to dissolve the crystals. The resulting formazan-containing solution was transferred to a 96-well cell culture plate and the absorbance was measured at 544 nm.

3.2.5. Flow cytometry

For the measurements, medium was removed from HEK293AD and AR42J cells, and the cells were harvested from the cell culture plate with 0.5x trypsin. The cells were then transferred to microcentrifuge tubes, washed and centrifuged twice (100 g, 5 min) in 1 ml PBS. Cell death was measured by using FITC Annexin V/Dead Cell Apoptosis Kit. Cells were filtered after incubation with FITC-conjugated Annexin V and propidium iodide, placed into FACS tubes, and then measured by using a BD FACSAria III flow cytometer (BD Biosciences) (DE-ÁOK, BMBI Stem Cell Research Laboratory) with a 488 nm laser 530/30 BP filter for FITC and a 561 nm laser 610/20 BP filter for propidium iodide measurements.

3.2.6 SDS-polyacrylamide gel electrophoresis

48 hours post-transfection, medium was collected from HEK293AD (200 µl) and AR42J (40 µl) cells and the proteins were precipitated with 10% trichloroacetic acid, followed by centrifugation (17000 g, 10 min). The pellets were dissolved in Laemmli buffer supplemented with 100 mM DTT and heat-denatured (95°C, 5 min). Samples were analysed by SDS-PAGE gel and proteins were visualized with Coomassie Brilliant Blue R-250 stain.

3.2.7 RT-PCR and qPCR analyses

RNA was isolated by NucleoSpin RNA Plus kit (Macherey-Nagel) and 2 µg was used for reverse transcription. From the resulting cDNA, PCR was used to determine the amount of mature XBP-1 mRNA using human and rat XBP-1 primer pairs. PCR products were separated on 2% agarose gel, the band intensities were determined by densitometry. The expression levels of BiP, CHOP, GAPDH and NQO1 genes were determined based on the relative amounts of the transcribed cDNA. The listed TaqMan assays were used for the experiments: BIP/HSPA5 (Hs00607129_gH, Rn00565250_m1), CHOP/DDIT3 (Hs00358796_g1,

Rn00492098_g1), ERP72 (Rn00587766_m1), GAPDH (Hs02758991_g1, Rn01775763_g1), Nqo1 (Hs00168547_m1, Rn00566528_m1) and PDI (Rn00564459_m1).

3.2.8. Protein immunoblot

Cells were washed twice with PBS, scraped from the cell culture plate, and dissolved in 1 ml PBS, followed by centrifugation (850 g, 10 min). Cells were lysed by repeated freeze-thaw cycles in 200 µl of lysis buffer supplemented with inhibitors (Reporter Lysis Buffer, Promega), the mixture was centrifuged (2400 g, 5 min) and the supernatant was collected. For Western-blot, 2.5 µg of AR42J lysate and 5 µg of HEK293AD lysate were used, dissolved in Laemmli buffer and denatured at 95°C. This was followed by SDS-PAGE and transfer onto nitrocellulose membrane (100V, 1 h). The membrane was blocked in TTBS solution containing 5% milk powder for 1 h, then anti-GAPDH antibody was added to the mixture at 1:15000 dilution (1 h) and HRP-conjugated anti-rabbit antibody was used as a secondary antibody (1 h) at 1:10000 dilution. To detect the presence of PNLIP, HRP-conjugated anti-His antibody was used at a dilution of 1:2000 (1 hour incubation). Proteins were detected by chemiluminescent substrate using an Azure 600 gel detection system.

3.2.9 Statistical analysis

Values were analyzed using MedCalc software, subjected to a T-test, where P-values less than <0.05 were considered significant. Data were plotted using GraphPad 5 software.

4. RESULTS

4.1 Characterisation of HTLV proteases

4.1.1 Expression and purification of HTLV proteases

The HTLV-1, -2 and -3 proteases were expressed in BL21(DE3) *E. coli* cells. The extraction of proteins from the inclusion bodies was the first step of purification, using the same conditions for each enzyme. HTLV-2 and -3 proteases had not been previously studied by others, so it was necessary to optimize their solubilization and purification. For this purpose, we used a protocol previously described for HTLV-1 PR. An additional washing step was applied during solubilization (details in Materials and Methods). After solubilisation, the proteins were purified by a reversed-phase HPLC-based method. The eluted fractions were collected and analysed by SDS-PAGE gels. The fractions with the highest purity (>95%) were dialysed against buffer E and used for cleavage reactions.

4.1.2 Determination of optimal reaction conditions and enzyme stability studies

The effects of different reaction conditions (temperature, pH, ionic strength) on the activities of HTLV-2 and HTLV-3 PRs were investigated and compared to the optimal conditions of HTLV-1 PR. After the enzyme reactions, the substrate and the cleavage products were separated by using a reversed-phase HPLC-based method. Oligopeptide substrates representing the natural cleavage sites of HTLV-1, HTLV-2 and HTLV-3 PR were used for activity measurements.

The increase of the temperature caused elevation of HTLV-2 and HTLV-3 PRs' activity, the highest activity was observed at ~40°C while a significantly lower activity (~40%) was determined at lower temperature (35°C). In agreement with the previous observations, the temperature optimal for the enzyme activity was close to 37°C. In addition to the HTLV-1 protease, this temperature was used in previous studies for BLV, as well as HIV-1, HIV-2 PR, EIAV, MMTV and human foamy virus (HFV) PRs.

The activities of the HTLV-2 and HTLV-3 PRs showed similar dependence on the NaCl concentration. The highest activity was observed at 2 M NaCl concentration. A similar trend was previously observed for HTLV-1 PR, as well. This effect of NaCl on the enzyme activity is a common feature of retroviral proteases, such as HIV-1, HFV and BLV, as well as retrovirus-like proteases, *e.g.* Ty1 retrotransposon PR and human retroviral-like aspartic protease 1 (ASPRV1).

The pH dependence was investigated in the range 4.5-8.0. The pH optimum of HTLV-1 (pH 6.11 ± 0.03) and HTLV-2 PR (pH 6.14 ± 0.06) were nearly the same, whereas that of the HTLV-3 PR was found to be lower (pH 5.56 ± 0.04). The pH optimum for HTLV proteases was found to be similar to those of other retroviral proteases and a slightly acidic pH is optimal for the protease activity, such as in the case of HIV-1 PR (pH 4.0-6.0) (151), BLV PR (pH 4.0-6.5), HFV PR (pH 6.6), and ASPRV1 PR (6.27 ± 0.02).

4.1.3. Comparison of catalytic properties

Oligopeptides representing the autoproteolytic cleavage sites of each HTLV protease (MA/CA, CA/NC, TF1/PR, PR/P1) were used to determine the catalytic efficiencies of HTLV proteases. For the measurements, the conditions that were previously optimized for HTLV-1 protease (0.25 M K_3PO_4, 2 M NaCl, 5 v/v % glycerol, 5 mM DTT, pH 5.6, 0.5-4 h, 37 °C) were used rather than the most optimal conditions that were determined in our study. This enabled us to obtain data that enable direct comparison of the studied proteases.

Some measurements required longer incubation times, which were not associated with a decrease or inactivation of enzyme activity. The stabilities of HTLV-1, HTLV-2 and HTLV-3 PRs were determined by using HTLV-3 PR/P1 substrate.

HTLV-1 PR showed the highest k_{cat}/K_M values of all substrates tested on substrates representing HTLV-1 PR cleavage sites, while lower values were observed for HTLV-2 and HTLV-3 PR cleavage sites. All substrates excepting HTLV-3 CA/NC were cleaved by the protease, the same peptide was cleaved with low efficiency by HTLV-2 and HTLV-3

PRs. For HTLV-1 PR, the lowest k_{cat}/K_M was observed for the HTLV-1 TF1/PR cleavage site, which was cleaved with similar efficiency by HTLV-2 and HTLV-3 PRs.

HTLV-2 PR cleaved all the tested substrates. The highest catalytic efficiency was observed for HTLV-2 TF1/PR and PR/P1 substrates. The cleavage sites of HTLV-1 and HTLV-2 CA/NC differ only at the P5' side chain, where lysine is present for HTLV-1 and arginine for HTLV-2, so only the HTLV-1 CA/NC substrate was used, the HTLV-2 CA/NC substrate was not tested in the experiments.

HTLV-3 also cleaved all the oligopeptides tested, but interestingly the highest k_{cat}/K_M was observed for the HTLV-1 PR/P1 cleavage site, which was the most preferred substrate of HTLV-1 PR, as well.

We also examined the activity of the enzymes on substrates representing HIV-1, EIAV, RSV, MMTV, MPMV, BLV, and MuLV PR cleavage sites. Although, HTLV-1 PR was able to cleave certain substrates, HTLV-2 and HTLV-3 PRs did not cleave the tested oligopeptides, indicating that HTLV-1 PR has a broader specificity as compared to HTLV-2 and HTLV-3 PRs. Some of the tested peptides inhibited HTLV-1 or HIV-1 PR at >0.1 mM concentration, hence some substrates were supposed to have similar effect on HTLV-2 and HTLV-3 PRs.

4.1.4. Amino acid preferences

To determine the amino acid preferences of HTLV proteases, oligopeptide substrates representing the wild-type HTLV-1 CA/NC (KTKVL*VVQPK) cleavage site and its N-terminally truncated (P4-P5′ and P3-P5′) and P4, P3, P2, P1 or P1′-modified variants were used.

As compared to the wild-type substrate (KTKVL*VVQPK), all of the three HTLV proteases showed lower activity on the P4-P5′ truncated substrates. The shortest P3-P5' substrate was not cleaved or was cleaved only to a negligible extent. The obtained relative activities were consistent with the lower catalytic constants previously observed for HTLV-1 PR. Lower values were also observed when shortened substrates

were used, indicating the importance of S4 and S5 sites in substrate recognition for both HTLV-2 and HTLV-3 PR.

Modification of the substrate at the P4 position was well tolerated by all three proteases. The results obtained are consistent with the finding that HTLV-1 PR has a relatively large hydrophobic S4 substrate-binding pocket. We found that the studied HTLV proteases prefer similar P4 amino acids. The lowest relative activity was observed for the P4-Gly and P4-Asp substrates, while the P4-Arg variant was not cleaved by the enzymes.

The S3 pockets were found to accept different amino acid side-chains, the preferences observed for the three proteases were found to be similar, but the P3-Phe and P3-Val variants were accepted by HTLV-1 PR only. The P3-Ser side-chain was more preferred than P3-Asp and the wild-type lysine residue.

Of the tested P2 variants, all three proteases preferred the P2-Ile mutant most. HTLV-2 and HTLV-3 PR showed higher relative activity values for P2-Ile than for the wild-type substrate. In agreement with the results of the former study on P2 amino acid preferences, we also found that isoleucine was the most preferred in this position, while P2-Asn and P2-lys were not acceptable for HTLV1 PR. P2-Asn was a relatively good substrate for HTLV-2 PR, although, the composition of the substrate-binding pocket does not underlie this.

The P1-Tyr and P1-Phe substrate variants were cleaved efficiently only by HTLV-1 PR. In a previous publication of our working group these variants were found to be good substrates for HTLV-1 PR. HTLV-3 did not cleave the P1-Ala mutant, it processed the P1-Leu and P1-Gly variants with low efficiency, while HTLV-2 did not cleave any of the substrate variants with at least 1% efficiency.

In agreement with the results of former specificity studies, we did not observe cleavage of the P1'-Ala, -Gly or -Lys substrate variants for HTLV-1 PR, and HTLV-2 and HTLV-3 PRs also had poor substrate cleavage except for P1'-Leu in the case of HTLV-3 PR.

4.1.5. Inhibition studies

The inhibition of HTLV-2 and HTLV-3 PRs was tested with several inhibitors used in HIV-1 therapy (atazanavir, darunavir, indinavir, ritonavir and saquinavir) and with DMP-323, which is also an effective HIV-1 PR inhibitor. The inhibitors were applied at a final concentration of 1 μM and showed a moderate inhibition, they caused less than 50% decrease of the activity. The obtained data are in agreement with the K_i values previously determined for BLV, HTLV-1 and MuLV PRs, for which low inhibition was also observed. Darunavir showed the highest inhibitory effect (25-30%) in the inhibition studies of the three HTLV PRs. In contrast, only indinavir showed a significant inhibitory effect on HTLV-1 PR.

In addition to the HIV-1 PR inhibitors, two experimental HTLV-1 PR inhibitors (IB-268, IB-269) were also tested. K_i values for HTLV-1 PR have been previously determined, which were 298 nM for IB-268 and 465 nM for IB-269, but were not shown to be effective against HIV-1 PR. The inhibition constant obtained for IB-268 on HTLV-2 PR was almost one tenth (37 nM) of that found for HTLV-1 PR, while HTLV-3 PR showed a similar value (214 nM) to HTLV-1 PR. In IB-269, the activity of HTLV-2 and HTLV-3 PRs was inhibited to a similar extent, significantly more efficiently than HTLV-1 PR.

4.1.6 Sequence analysis and substrate binding site analysis

The sequences of HTLV PRs show only low identity. The sequence identity between HTLV-1 and -2 PR is 50% and between HTLV-1 and HTLV-2 PR is 49%. The HTLV-2 and -3 PRs show 57% identity, but there is a higher conservation in the case of the substrate binding regions (HTLV-1/HTLV-2 PR: 82%, HTLV-1/HTLV-3 PR: 56%, HTLV-2/HTLV-3 PR: 76%). For substrate-binding pockets, the highest variability was found at subsite S4, where there was 57% identity between HTLV-1 and HTLV-2 proteases, 57% for HTLV-3 PR, and only 43% between HTLV-2 and HTLV-3 PR. There is a higher degree of conservation closer to subsite S1. HTLV-2 and HTLV-3 PRs have the same S3 pocket composition, HTLV-1 differs from them in one residue. The S2 pocket of HTLV-1 and HTLV-3 PRs differ only at position 37,

and there are only 2 residue differences between HTLV-1/HTLV-2 and HTLV-2/HTLV-3 PRs. The S1 pockets of HTLV-2 and HTLV-3 PRs have the same composition, HTLV-1 PR differs from the other HTLV PRs at the 10th position. The composition of the S4-S1 substrate-binding pockets of HTLV-1 PR has been previously determined by members of our research group, and in our work the composition of the substrate-binding pockets of HTLV-2 and HTLV-3 PRs were determined based on the comparison of the protein sequences.

4.1.7. Testing autoprocessing ability

The mutations already studied for HTLV-1 PR were also used to investigate the autoproteolysis of HTLV-2 and HTLV-3 PR. The sequence identity between HIV-1 and HTLV PRs is lower than between HTLV-1, -2 and -3 PRs. To study the effect of the mutations on the autoprocessing ability, the sequences of the HTLV PRs were modified by replacing their residues with those of the HIV-1 PR in the studied positions. The mutations were designed to test the sensitivity of HTLV proteases to mutations affecting the flap (positions 57 and 59) and the active site (positions 37 and 67). Each of the modified residues is involved in the assembly of a substrate binding pocket. The residue at position 37 is part of the S2 and S4 pockets, while the one at position 57 is part of the S2-S4 pockets, and the 59th and 67th residues are involved in the formation of the S1-S2 and S4 pockets. L57 is conserved among HTLV PRs, and mutation of the corresponding position (E48) in the HIV-1 PR to leucine is responsible for the resistance against indinavir and saquinavir. The I50 residue of HIV-1 and HIV-2 PRs is thought to play a role in the flexibility of the flap region, while in HTLV-1 PR, the A59I mutation at the corresponding position caused inactivation of the enzyme. In HTLV-1 PR the F67 is part of the S4 pocket and the enzyme is intolerant of substitution for glutamine at this position (F67Q mutation).

To investigate the autoproteolysis of HTLV-2 and HTLV-3 PR, the enzymes were expressed as fusion proteins containing an N-terminal MBP fusion tag and an 8 residue-long linker prior to the protease. This

sequence corresponds to the autoproteolytic cleavage site of the Gag-Pro polyprotein which precedes the protease and allows the protease to process from the fusion protein. The autoproteolyis was followed by Western-blot and cleavage efficiencies were compared to those previously determined for HTLV-1 PR.

The effects of the mutations on the autoproteolytic activities were similar. Substitution of the 37th residue (MBP-HTLV-1: methionine, MBP-HTLV-2: leucine, MBP-HTLV-3: isoleucine) with asparagine prevented processing in MBP-HTLV-1 and MBP-HTLV-2 PRs, but did not cause any change in MBP-HTLV-3 PR. Mutation of the same residue to aspartate abolished self-proteolysis. The L57G mutation prevented autoproteolysis of MBP-HTLV-3 PR but not that of MBP-HTLV-2 PR.

4.2. Investigation of the association between hydroquinone and the G233E PNLIP mutation

4.2.1 Effect of major components of cigarette smoke on the viability of HEK293AD cells

To characterize the effect of the main components of cigarette smoke, HEK293AD cells were tested in the presence of elevating concentrations of acrolein, crotonaldehyde or hydroquinone (HQ). The approximate viability of the cells was determined by examining the metabolic activity of the cells.

Acrolein at concentrations up to 20 μM did not affect cell metabolism, but at higher concentrations it progressively decreased cell viability. Conversely, HQ and crotonaldehyde caused a significant reduction in viability even at low concentrations, with a concentration of 100 μM being associated with a reduction in viability of more than 80%. Because of its significant effect on metabolic activity and its good storability, HQ was investigated in more detail. Flow cytometry was used to investigate whether reduced cellular metabolic activity was associated with programmed cell death or necrosis by monitoring the amount of cells labelled with FITC annexin V and propidium iodide stains in response to increasing HQ concentrations. HQ at 25 μM had no significant effect, but programmed cell death was observed at higher HQ levels of 50 and 75 μM, as determined by an increase in the proportion of cells labelled with FITC annexin V staining, followed by an increase in the number of cells stained with propidium iodide.

4.2.2 Examination of secretion and its effect on the viability of hydroquinone and inappropriately folded PNLIP mutant HEK293AD cells

Previous studies have already investigated the secretion of the G233E mutant PNLIP, it was found that in contrast to the wild-type lipase, the misfolded G233E variant was not secreted in HEK293T and AR42J cells, however the expression of the G233E mutant PNLIP in HEK293AD cells has not been investigated to date. To verify secretion,

plasmids expressing wild-type and mutant lipase were introduced by transient transfection, conditioned medium was collected and analyzed by SDS-PAGE. As with the cell lines mentioned above, HEK293AD cells secreted only wild-type PNLIP and were therefore suitable for the desired experiments. We also monitored the effect of HQ on secretion, adding 25 μM HQ to the cells reduced the secretion of wild-type lipase, but had no effect on intracellular PNLIP levels.

To investigate the additive effect of HQ and mutant PNLIP, HEK293AD cells were transfected with plasmid expressing wild-type and mutant lipase and incubated with HQ (0, 25 μM). The concentration of HQ used was not yet associated with cell death, and thus proved suitable for studying early cellular responses. Cells transfected with empty vector and not treated with HQ were considered as controls. The presence of HQ was responsible for a 30% decrease in metabolic activity. By itself, wild-type PNLIP reduced cell viability by 18%, whereas the mutant reduced it more significantly by 31%. The presence of HQ and wild-type lipase together were responsible for 48% and HQ and mutant lipase for 57% reduction.

4.2.3 Effects of HQ and the misfolded PNLIP variant on ER stress in HEK293AD cells

The ER stress-inducing effect of some non-secreted PNLIP variants, such as G233E mutant, which we studied in more detail, has been previously demonstrated in HEK293T and AR42J cells. The effect of HQ ER on stress has also been investigated in other mammalian cells in several studies. To investigate the effect of HQ and G233E PNLIP mutation on ER stress, HEK293AD cells were transfected with (transient) plasmids containing wild-type and mutant lipase. Cells were incubated with HQ (0 and 25 μM) and XBP-1 maturation and mRNA levels of additional ER stress markers BiP, CHOP and NQO1 were determined. Samples transfected with empty plasmid were considered as control. The percentage of mature XBP-1 mRNA in cells expressing wild-type lipase was only slightly increased to 16% compared to the control (10%). Compared to the wild type, we observed a nearly twofold increase in

mutants, which further increased in the presence of HQ to 42%, higher than the effect of HQ alone.

Expression of wild-type PNLIP caused no change in BiP mRNA levels, in contrast, the mutant induced a 2.5x increase. HQ caused a 1.9-2.2x increase in cells expressing control and wild type. When HQ was added to the mutants, a higher change of 3.6x was observed than in the mutants. Similarly, to BiP mRNA, wild-type PNLIP did not cause a significant change in CHOP mRNA levels compared to the control. The presence of the mutant increased CHOP levels by 1.6x. In control cells, the addition of HQ was associated with a slightly lower change and, as before, the combined presence of the mutant and HQ caused the largest 2.6x increase. Neither did wild-type PNLIP increase transcription of NQO1 mRNA, with the mutant responsible for a 1.6x change. The addition of HQ also caused an increase in mRNA levels, the highest when used in combination with mutants (2.5x).

4.2.4. Effect of HQ on AR42J cell viability

The effect of HQ was also investigated on AR42J rat pancreatic cells differentiated with dexamethasone prior to the experiments. The cells were incubated with HQ (0 and 40 μM, 24 h) and then monitored for cell viability by measuring their metabolic activity. The concentration of HQ was increased up to 100 μM, which was associated with a relatively small decrease in viability of 20%, with cells showing lower sensitivity compared to HEK293AD cells.

We also tested the presence of apoptosis at different HQ concentrations (0, 40, 100 and 200 μM). For this purpose, annexin V and propidium iodide staining were used and cells were observed using a flow cytometer. Up to 40 μM HQ concentration, cells remained stable, annexin V positive cells appeared only at 100 μM. At higher concentrations (200 μM), cells also started to stain with propidium iodide. Our results show that programmed cell death is observed in AR42J cells in the presence of higher HQ concentrations compared to HEK293AD cells.

4.2.5 Examination of the secretion and viability of AR42J cells in the presence of HQ and the misfolded PNLIP mutant

The effect of HQ (40 μM) on secretion was examined in AR42J cells expressing wild-type and mutant lipase. The amount of PNLIP was also monitored in the media and cell lysates collected from the cells. Consistent with our previously described observations, the G233E variant was not secreted, and HQ significantly reduced the secretion of wild-type PNLIP, while intracellular lipase levels were unchanged.

We investigated whether the G233E PNLIP mutation and HQ have a concomitant effect by examining cellular metabolic activity. To this end, AR42J cells were infected with adenovirus encoding wild-type and mutant PNLIP and incubated in the presence of HQ (0, 40 μM). The addition of HQ reduced viability by 28%, which is comparable to the data obtained for HEK293AD cells where 25 μM HQ was applied. A 9% reduction in wild-type lipase was observed when HQ was also added to the cells, with a slightly stronger effect than when HQ alone was applied. The G233E PNLIP mutation was responsible for a 45% reduction, which increased to 60% when HQ was added.

4.2.6 Effects of HQ and inappropriately folded PNLIP variant on ER stress in AR42J cells

The relationship of HQ and the G233E PNLIP mutation with ER stress was also investigated by monitoring the immature/mature XBP-1 mRNA ratio and the transcript levels of BiP, CHOP and NQO1. Cells transduced with empty adenovirus were considered as controls. The mRNA ratio of mature XBP-1 in cells expressing control and wild-type lipase ranged from 12-16%, whereas it increased significantly to 37% in the G233E variant. The addition of HQ caused a 31% increase in both control and wild-type PNLIP-expressing cells, but as high as 45% in mutant-expressing cells, the highest value among the conditions tested. In cells expressing wild-type PNLIP, BiP mRNA levels were unchanged

compared to the control, whereas the mutant was responsible for an increase of nearly 2.1x. HQ significantly increased BiP transcription (2.7x-6.1x), reaching the highest value in the presence of the G233E PNLIP mutant. The level of CHOP did not change significantly in cells expressing wild-type PNLIP, whereas the G233E mutation was responsible for a 2.1x increase. HQ alone was associated with a 3x increase in CHOP levels, with a magnitude of change that increased nearly 6x in the presence of wild-type and mutant PNLIP. Addition of HQ to the cells caused a 1.5x increase in NQO1 transcription. Although the wild-type lipase only slightly increased transcription, the increase in mRNA levels was 2.3x when HQ was present, compared to the 1.5x change induced by HQ. Even more significantly, expression of the mutant increased the level of NQO1, which was further enhanced by the addition of HQ.

To further confirm the role of oxidative stress, we also examined the transcription of two protein disulfide isomerases, PDI and ERP72, in AR42J cells. A 2.1x increase in ERP72 mRNA levels was observed upon HQ, while the expression of the G233E variant caused an almost equivalent change, whereas wild-type lipase was responsible for a 1.4x increase. In lipase-producing cells, the addition of HQ was equally associated with a 4.5x increase in ERP72 mRNA levels. Expression of wild-type lipase was not associated with a change in PDI mRNA levels, and the mutant also caused only a small 1.3x increase. The addition of HQ caused a 1.5-1.9x increase and was not significantly affected by genetic factors.

5. SUMMARY

The work presented in the PhD thesis deals with the study of two proteins involved in human disease, providing insights into the biochemical background of HTLV retroviral protease function and the relationship between components in cigarette smoke and the normally secreted pancreatic enzymes that are not properly folded.

Our work involved the optimization of the expression and purification of HTLV-2 and -3 PR and the investigation of their biochemical properties. We found that the enzymatic properties of HTLV-2 and -3 PRs are highly similar and have a narrower substrate specificity than HTLV-1 PR, and their amino acid preferences are similar. Similar to HTLV-1, HTLV-2 and -3 PRs have a higher tolerance to amino acids P4 and P3, while S2-S1' pockets have a more stringent amino acid preference. The HIV-1 PR inhibitors tested were not able to inhibit HTLV-2 and -3 PR effectively, whereas the experimental HTLV-1 PR inhibitors IB-268 and IB-269 were significantly more effective. Autoprocessing assays with modified HTLV-2 and -3 PRs confirmed that HTLV PRs are more sensitive to active site mutations than HIV-1 PRs due to their different replication strategies.

In our work on risk factors for chronic pancreatitis, we used HQ to model the potential impact of smoking and the PNLIP protein misfolding was represented by the G233E mutant. We found that each of the effects alone reduced cell viability, that application of higher concentrations of HQ was able to induce programmed cell death, and that it also inhibited the secretion of wild-type PNLIP. Their combined use was associated with even more significant changes. We also observed an increase in the expression of ER stress markers, and changes in the mRNA level of a pro-apoptotic marker followed this trend. In our experiments, we observed a cumulative effect of the HQ and G233E mutant PNLIP protein, suggesting that although PNLIP mutants that do not fold properly may not always be clearly associated with pancreatic inflammatory disease, the cigarette smoke component hydroquinone may be able to significantly increase the risk of this, or even more severe symptoms.

6. PUBLICATIONS ON WHICH THE PHD THESIS IS BASED

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/425/2022.PL
Subject: PhD Publication List

Candidate: Norbert Kassay
Doctoral School: Doctoral School of Molecular Cellular and Immune Biology
MTMT ID: 10081734

List of publications related to the dissertation

1. **Kassay, N.**, Toldi, V., Tőzsér, J., Szabó, A.: Cigarette smoke toxin hydroquinone and misfolding pancreatic lipase variant cooperatively promote endoplasmic reticulum stress and cell death. *PLoS One. 17* (6), 1-19, 2022.
   DOI: http://dx.doi.org/10.1371/journal.pone.0269936
   IF: 3.752 (2021)
2. **Kassay, N.**, Mótyán, J. A., Matúz, K., Golda, M., Tőzsér, J.: Biochemical Characterization, Specificity and Inhibition Studies of HTLV-1, HTLV-2, and HTLV-3 Proteases. *Life (Basel). 11* (2), 1-21, 2021.
   DOI: http://dx.doi.org/10.3390/life11020127
   IF: 3.251

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

List of other publications

3. Mótyán, J. A., **Kassay, N.**, Matúz, K., Tőzsér, J.: Different Mutation Tolerance of Lentiviral (HIV-1) and Deltaretroviral (BLV and HTLV) Protease Precursors.
   Viruses-Basel. 14 (9), 1-17, 2022.
   DOI: http://dx.doi.org/10.3390/v14091888
   IF: 5.818 (2021)

4. Toldi, V., **Kassay, N.**, Szabó, A.: Missense PNLIP mutations impeding pancreatic lipase secretion cause protein misfolding and endoplasmic reticulum stress.
   Pancreatology. 21 (7), 1317-1325, 2021.
   DOI: http://dx.doi.org/10.1016/j.pan.2021.07.008
   IF: 3.977

Total IF of journals (all publications): 16,798
Total IF of journals (publications related to the dissertation): 7,003

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

09 September, 2022

7. ACKNOWLEDGEMENTS

I would like to thank **Dr. József Tőzsér** for his guidance and professional advice and for giving me the opportunity to do my PhD studies at the Department of Biochemistry and Molecular Biology of University of Debrecen!

I would like to thank **Dr. András Szabó** for giving me the opportunity to get involved in the research topics he was leading and for his professional advice!

I would like to thank **Dr. János András Mótyán** for his comments and professional advice which contributed to the success of my work!

I would like to thank **Dr. Krisztina Joóné Matúz**, my former MSc supervisor, for her support during my PhD studies. Thank you to all the former and current staff of the Laboratory of Retroviral Biochemistry: **Ambrus Viktor**, **Dr. Lívia Diána Gazda**, **Dr. Mária Golda**, **Gyula Hoffka**, **Szilvia Janics-Pető**, **Balázs Kunkli**, **Tamás Richárd Linkner**, **Dr. Márió Miczi**, **Noémi Miltner**, **Dr. Mohamed Mahdi Faisal**, **Katalin Nagy**, **Dr. Zsófia Ilona Szojka**, **Dr. Vanda Toldi**, **Dr. Ferenc Tóth** and **Ágota Nagyné Veres**, for their help and support.

Thank you to all the other staff members of the Department of Molecular Biology and Biochemistry and to all those who contributed to the PhD thesis!

Last but not least, thank you to my **family** without whom this dissertation would not have been possible!

The wild-type HTLV-2 PR-pET11a expression plasmid was provided by Dr. Mária Golda, the wild-type HTLV-3 PR-pET11a plasmid by Dr. Krisztina Joóné Matúz, while the wild-type and mutant PNLIP plasmids by Dr. Vanda Toldi. The composition of the substrate binding pockets was determined by Dr. János András Mótyán.

The works presented in this thesis were supported in part be the following projects: GINOP-2.3.2-15-2016-00044 "PHARMPROT teaming", TKP2020-IKA-04, TKP2021-EGA-20 (Biotechnology), TÁMOP 4.2.2.A-11/1/KONV-2012-0023 "VÉD-ELEM", TÁMOP-4.2.2.D-15/1/KONV-2015-0016, ÚNKP-21-5 New National Excellence Programme, FK127942 OTKA, K125238 OTKA.