SHORT THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

THE DEVELOPMENT AND APPLICATIONS OF A RECOMBINANT FLUORESCENT FUSION PROTEIN-BASED *IN VITRO* PROTEASE ASSAY PLATFORM FOR KINETIC AND SPECIFICITY STUDIES

by Beáta Bozóki

Supervisor: Dr. József Tőzsér

UNIVERSITY OF DEBRECEN
DOCTORAL SCHOOL OF MOLECULAR CELL AND IMMUNE BIOLOGY

DEBRECEN, 2020

THE DEVELOPMENT AND APPLICATIONS OF A RECOMBINANT FLUORESCENT FUSION PROTEIN-BASED *IN VITRO* PROTEASE ASSAY PLATFORM FOR KINETIC AND SPECIFICITY STUDIES

By Beáta Bozóki, MSc

Supervisor: Prof. József Tőzsér, PhD, DSc

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

Head of the **Examination Committee**: Prof. Ferenc Erdődi, PhD, DSc
Members of the Examination Committee: Prof. Péter Gál, PhD, DSc
Judit Szabó, PhD

The Examination took place at Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Debrecen
Debrecen, 16th February, 2015

Head of the **Defense Committee**: Prof. Gábor Szabó, PhD, DSc
Reviewers: Prof. Judit Erzsébet Pongrácz, PhD, DSc
Tamás Emri, PhD

Members of the Defense Committee: Krisztián Kvell, MD, PhD
Endre Kókai, PhD

The PhD Defense takes place at the Lecture Hall of Bldg. A, Department of internal Medicine, Faculty of Medicine, University of Debrecen at 13:00 on 11th September 2020.

1. INTRODUCTION

1.1. *In vitro* protease assays

Proteases are essential components of all living organisms from viruses to humans. Proteases account for 2 % of the human genome and 1-5 % of the genome of infectious organisms. The complete set of proteases in a given organism is called the 'degradome' and plays a role in almost all biological pathways, thus, alterations in the proteolytic system account for several pathological conditions, for example cancer, neurodegenerative disorders, and inflammatory and cardiovascular diseases. Proteases are key enzymes in the life-cycles of numerous infectious microorganisms and viruses, therefore, became potential therapeutical drug targets and diagnostic and prognostic biomarkers in the pharmaceutical industry. Consistent with this, there is a great demand for the development of both *in vitro* and *in vivo* technologies aiming either to profile the so called 'orphan proteases', the function and specificity of which are unexplored, or to further investigate the already described ones *e.g.* by compound profiling.

Quantification of proteolytic activity depends on the substrate type, the activity rate of the enzyme to be investigated, and on the sensitivity and precision requirements of the utilized detection system. Generally, *in vitro* protease assays detect protease activity on either whole proteins or short synthetic peptide molecules.

Due to their versatile nature, synthetic peptide substrate-based methods are the most popular protease assay platforms. Synthetic peptide sequences are usually modified by different reporter groups including fluorogenic, chromogenic, or luminogenic moieties. Class-specific substrates can be developed for the differentiation of endopeptidases or exopeptidases. Several oligopeptide-based assay formats have already been developed, but the time- and cost efficient high-throughput screening (HTS)-compatible platforms are the most preferred. *In vitro* protease assays using synthetic peptide substrates can be classified into (i) homogenous, (ii) separation-based and (iii) heterogeneous assay formats. In case of homogenous and separation-based assays both the substrate(s) and the protease are in an aqueous phase. In the former format the substrate turnover can be detected as a signal change directly from the reaction mixture without further sample processing, while in case of separation-based formats products and/or the remaining substrates are isolated

from the reaction mixture after the termination of the reaction at a pre-determined linear range, where the amount of the generated products give a robust signal. In contrast, in heterogeneous assay formats, the substrates are immobilized on solid surfaces, while the protease of interest is in the aqueous phase.

Recombinant protein engineering technology has introduced new possibilities in the field of protease assays by the utilization of recombinant proteins as substrates. These substrates are designed specifically to the experimental aims. They usually contain only a single cleavable site to eliminate the limitations hold for native protein substrates due to the multiple cleavage sites present in their sequences. Microbial systems offer cost-efficient, convenient, and powerful tools for recombinant protein production although protein insolubility, conformational, structural and stability issues, low purification yields, and host cell toxicity are potential challenges that are to be overcome. In order to solve the listed difficulties, several different fusion tags have been developed, that can be applied to improve the characteristics of recombinant protein substrates. Generally, a single fusion tag does not provide overall solution to all issues, therefore, dual- or multiple-tagging of the protein of interest is widely applied to satisfy the different needs. Similar to the oligopeptide-based methods, the vast majority of *in vitro* recombinant protein-based systems rely on fluorescent intensity readout. In a few methods a fluorescent moiety is attached to the recombinant substrate molecule either chemically or *via* a fluorescently labelled antibody, however, in most of the cases the source of the detectable fluorescence is a genetically encoded fluorescent protein. By the recent developments in fluorescent protein engineering, wide variety of the fluorescent proteins have become available with several different advantages and properties to best suit the experimental demands.

1.2. Human immunodeficiency virus type 1 (HIV-1) retropepsin

Inevitably, Human immunodeficiency virus (HIV), the causative agent in acquired immunodeficiency syndrome (AIDS) is one of the most well-known and investigated member of retroviruses. Due to the epidemic significance of HIV, great efforts have been made for investigating the role of the different viral elements and for the development of effective antiretroviral inhibitors, targeting the critical steps in the life-cycle of the virus especially for HIV type-1 (HIV-1). Owing to its central role in the maturation and assembly of the virus, HIV-1

retropepsin, also called as HIV-1 PR, has become one of the most frequently investigated and targeted viral components. The HIV-1 PR is composed of 99 amino acid residues and its unbounded form it exhibits a 2-fold rotational (C_2) symmetry. It acts as a dimer of two identical subunits with only one active site, which lies across the dimer interface. Due to its central role in the maturation process of HIV-1, the PR has been a major target for developing inhibitors. HIV-1 protease inhibitors (PIs) are essential part of the highly active antiretroviral therapy (HAART). Rapid development of resistant viral mutants is the consequence of the heterogeneity and the rapid turnover of the virus coupled with the error-prone nature of the viral RT and the selective pressure forced by the therapeutic PIs. Although innumerable attempts have been made to design selective inhibitors with improved effectiveness against resistant viral forms, the development of new inhibitors are still on demand.

1.3. Tobacco etch virus NIa protease (TEV PR)

TEV is the member of *Potyviridae* family, regarded as the most economically important class of plant viruses, as they infect numerous host plants causing enormous agricultural, economical, and biological losses worldwide. It has a 9.5 kilobase long (+) RNA genome. The translated polyprotein is processed by three proteases, including P1, HC-Pro and Nuclear Inclusion protein a (NIa). The NIa protease is also referred as TEV PR or TEVp. It is released from the polyprotein by proteolysis and is critically important for viral maturation, as it has the largest contribution to polyprotein processing. The primary structure of TEV PR is highly similar to that of picornavirus 3C protease and His46, Asp81 and Cys151 together are considered as the catalytic triad. Studies of Bazan *et al.* (1988) have confirmed that these type of viral cysteine proteases are structurally and functionally homologous to the trypsin-like family of serine protease but utilize a cysteine thiol instead of a serine hydroxyl as the active-site nucleophile. Regarding P1' position, the proteolytic activity is retained at the presence of any amino acids except Pro. This practically means that TEV PR can release almost every protein fused downstream to ENLYFQ↓S/G, leaving the desired amino acid at the N-terminus.

TEV PR is one of the best studied and most extensively utilized proteases for biotechnological applications, due to its advantageous properties. It is widely used for affinity tag removal.

1.4. Venezuelan equine encephalitis virus non-structural protein 2 protease

Venezuelan equine encephalitis virus (VEEV) is the member of *Alphavirus* genus of the *Togaviridae* family and causes human and livestock disease mainly in Central and South America.

Alphaviruses including VEEV, are small spherical enveloped viruses, with ~70 nm diameter. The ~11 kilobase long viral mRNA consists of two cistrons: one which corresponds to the 5'-two-third of the genome is translated into a single polyprotein nsP123 or nsP1234, comprising the non-structural proteins, while the second open reading frame (ORF) is coding for the structural proteins that form the viral particles translated in the late stage of the infection. Non-structural proteins play a pivotal role in the formation of the viral replication complex and are released from the nsP123 and nsP1234 polyprotein upon the proteolytic activity of C-terminal region of nsP2, frequently referred to as the nsP2pro. In alphaviruses the nsP2pro processes the non-structural polyprotein at three different sites located between nsP1/nsP2; nsP2/nsP3 and nsP3/nsP4. Additionally, nsP2 protease is supposed to play an essential function in neutralizing early innate immune system, by mechanisms similar to other Group IV viruses comprising *Picornaviridae*, *Flaviviridae* and *Coronaviridae* including novel pathogens contributing to recent global pandemia e.g. middle east respiratory syndrome- and severe acute respiratory syndrome-associated coronavirus (MERS- and SARS-CoV) and SARS-CoV-2. Previously structural similarities were described between VEEV nsP2 and papain-like proteases/deubiquitinases of the SARS- and MERS-CoV. Based on these, VEEV nsP2pro is regarded as potential target for antiviral drug development. Furthermore, as these enzymes pose some degree of sequence specificity, they have been also considered as potential biotechnological tools for the *in vitro* removal of affinity tags from recombinant proteins.

Previously our laboratory investigated the activity of three alphaviral proteases for their utilization as tools for affinity tag removal in collaboration with Macromolecular Crystallography Laboratory (MCL), Center for Cancer Research, National Cancer Institute (NCI) (Frederick, MD, USA). The results of the study were published by Zhang *et al.* (2009). The nsP2pro domains of sindbis virus (SIN), semliki forest virus (SFV) and VEEV were generated by using bacterial expression system, and the activity of these proteases were investigated on different fusion protein substrates and on artificial oligopeptides substrates designed based on the P6-P6' residues of the three corresponding natural recognition sites. As the recognition site of the three enzyme is highly

similar, beside investigating the activity of the proteases on their cognate substrates, their cross-reactivity was also tested. Kinetic parameters on oligopeptide substrates could be determined for only three enzyme-substrate pairs including VEEV nsP2pro on SFV nsP1/nsP2 (SFV-1) substrate and SFV nsP2pro on its cognate SFV nsP1/nsP2 and nsP3/nsP4 cleavage site. The corresponding k_{cat}/K_M values were determined as 0.028 ± 0.005 mM^{-1} s^{-1}, 0.036 ± 0.008 mM^{-1} s^{-1} and 0.277 ± 0.050 mM^{-1} s^{-1}, respectively, and were considered as substantially lower than the catalytic efficiency of TEV and tobacco vein mottling virus (TVMV) proteases. Furthermore, their ability to process fusion protein substrates was also relatively poor, therefore, it was concluded that despite of the fact that alphavirus proteases theoretically exhibit sufficient sequence specificity to be useful reagents for affinity tags removal, in practice their cleavage properties, at least in their investigated form, could not offer any advantage over the already successfully utilized potyviral proteases.

2. OBJECTIVES

Several excellent biological and chemically-based methods have been designed to investigate the different proteolytic activities and to find/optimize substrates for the newly discovered proteases. However, many of them are relatively insensitive, expensive, time-consuming, and labor-intensive, some of them are applicable only for a group of proteases or provide incomplete coverage and no or only limited data on reaction kinetics. Accordingly, the development of HTS-compatible and quantitative assay formats offering new advantages and/or integrating the already existing ones of other methods is still on demand.

In this work our aims were the followings:

1) Establishment of an interchangeable recombinant fusion protein substrate system by the generation of the expression vectors and plasmids coding for substrates that consist of N-terminal His_6 and MBP fusion tags, a TEV PR recognition site, a cloning cassette for the insertion of the cleavage site of interest, and a C-terminal fluorescent protein (FP).

2) Optimization of the expression and the purification of the substrates and their application for the development of a quantitative, HTS-compatible, versatile fluorescence-based proteolytic assay platform. The assay was aimed to be optimized by using HIV-1 and TEV PRs, and to be utilized in specificity studies of VEEV PR

3) A VEEV PR (nsP2pro-1) was previously expressed by Zhang *et al.* (2009) and showed unexpectedly low activity *in vitro*. Thereby new recombinant VEEV nsP2 constructs were designed, in which the protease domain would be extended by (i) the N-terminal domain (VEEV nsP2) and (ii) Ala436-Met457 region (VEEV nsP2pro-2). The *in vitro* activity of the new construct was planned to be studied on His_6-MBP-mEYFP substrates comprising wild-type SFV-1 cleavage site or its P5, P4, P2, P1, P1', and P2' variants.

3. MATERIALS AND METHODS

3.1. Generation of the 'empty' expression vector of the substrates.

pDest-His_6-MBP-FP plasmids were prepared by Gateway Cloning Technology based on Tropea *et al.* (**2007**). The linear DNA sequences to be transferred *via* pDON221 donor vector into pDEST-His_6- were amplified by a two-step polymerase chain reaction (PCR). Gateway cloning steps were carried out according to the manufacturer's protocol using customized N1 and C primers.

3.2. Generation of substrate coding dsDNA by random mutagenesis

Annealed oligonucleotide primers, coding for SFV-1, flanked by *BamHI* and *NheI* ends were ligated into pT7-Blue-3 plasmid previously linearized by *BamHI* and *NheI* restriction endonucleases. Random mutagenesis of P1' residues SFV-1 was performed by QuickChange Lightning Multi-Site Directed Mutagenesis Kit (Stratagene) according to the manufacturer's manual but using oligonucleotide primer degenerated at the triplet coding for P1' site. Plasmids were purified using Qiagen Plasmid Midi Kit. The mutagenized DNA sequences of SFV-1 cleavage sites were cut out of the purified plasmids using *NheI* and *PacI* and separated on polyacrylamide gels containing 15 % urea and purified from the gel by Qiaex II Gel Extraction Kit.

3.3. Generation of substrate coding-expression vectors by ligation

Short dsDNAs coding for the protease cleavage site of interest were generated either by random mutagenesis or by annealing of complement oligonucleotides primers flanked by *NheI* and *PacI* sticky ends, and were ligated into pDest-His_6-MBP-FP plasmid linearized by *NheI* and *PacI*. Next day 100 µL of *E. coli* BL21(DE3) competent cells were transformed by the ligation reaction mixture (15 µL) using heat shock, followed by plasmid preparation from the cultures of grown colonies. Purified plasmids were verified by capillary DNA sequencing.

3.4. Expression of substrates and cell lysis

Small-scale overnight starter cultures were initiated by adding 10 µL freezer stocks of *E. coli* BL21(DE3) cells containing the expression plasmids, to 5 mL LB medium (containing 100 µg/mL ampicillin) in a 50 mL centrifuge tube and incubated overnight at 37 °C. Next day 5 mL of the starter culture was

added to 50 mL LB containing 100 µg/mL ampicillin in a 500 mL Erlenmeyer flask. (For small scale expression 2.5 mL of the starter culture was added to 15 mL LB containing 100 µg/mL ampicillin in a 50 mL centrifuge tube.) Cells were grown at 37 °C up to an absorbance of 0.6-0.8 at 600 nm. Protein expression was induced by the addition of 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) followed by incubation for 3 hours at 37 °C (or 4 hours at 37 °C for small scale expression). Cells were harvested by centrifugation at 4000 g for 15 minutes at 4 °C (Jouan CR 412) and were stored at least for 1 hour at -70 °C before lysis. For cell lysis, pellets were thawed on ice for 15 minutes and suspended in 2 mL lysis buffer (50 mM sodium-acetate, 300 mM NaCl, 10 mM imidazole, 0.05 % Tween 20, pH 8.0) (1 mL for small scale expression) containing PMSF at a final concentration of 25 µg/mL. Lysozyme and DNase were added in 1 mg/mL and 10 U/mL final concentrations, respectively. Cell suspensions were vortexed and occasionally mixed during incubation on ice for 10 minutes, then were aliquoted into microcentrifuge tubes and sonicated for 3 minutes. Tubes were centrifuged at 10000 g for 15-20 minutes (Eppendorf 5415D) at room temperature, and cleared bacterial lysate were collected.

3.5. Substrate expression for VEEV nsP2pro-2 kinetic study

The expression of the wild-type, P4-Glu, P4-Thr, P4-Arg, P4-Gly, P1-Gly, P1'-Thr, and P2'-Ser SFV-1 cleavage site sequences were performed as described in Section 3.4, but the cells were grown at 37 °C up to an absorbance of 0.5–0.6 at 600 nm, and the suspension was incubated for 2 hours at 37 °C after induction by IPTG. Protein translation was arrested by addition of tetracycline in 200 µg/mL final concentration, and the culture was further incubated at 37 °C for 2 h.

3.6. Purification of the substrates

Cleared bacterial lysates containing the desired recombinant fusion substrate were added to Ni-NTA magnetic agarose beads (Qiagen or Cube Biotech). The suspensions were incubated for 20-60 min at room temperature while continuously rotating. Substrate-attached magnetic beads (SAMBs) were washed three times by 1 % Tween 20 (pH 7.0), three times by washing buffer (50 mM NaH_2PO_4, 300 mM NaCl, 5 mM imidazole, 0.05 % Tween 20, pH 7.0) and three times by cleavage buffer (50 mM NaH_2PO_4, 300 mM NaCl, 0.05 % Tween 20, pH 7.0) using Dynamag^{TM}-2 magnetic particle concentrator (MPC).

Hereafter, substrates were eluted from the beads using either elution buffer A (100 mM EDTA, 0.05 % Tween 20, pH 8.0) used in Section 3.12 – 3.15, or elution buffer B (50 mM NaH_2PO_4, 300 mM NaCl, 500 mM imidazole, 0.05 % Tween 20, pH 8.0) used in Section 3.16-3.18. Buffer was exchanged to cleavage buffer by using 10K Amicon tubes. Total protein concentrations were determined by BCA protein assay and/or by measuring absorbance at 280 nm (NanoDrop 2000), the theoretical extinction coefficients and molecular weights were calculated by ProtParam tool of ExPASy.

3.7. Calibration of the substrates

The purified substrates dissolved in either elution buffer or cleavage buffer, were serially diluted and each dilution points were transferred into black half-area plates. Relative fluorescent intensities were measured by using Biotek Synergy2, H1 or Victor Wallac 1420 devices according to the excitation and emission wavelength of mApple, mTurquoise2, mEYFP and mCherry. Blank-corrected relative fluorescent intensity values (RFU) were plotted against the substrate concentration (mM). Linear regression was performed, and the parameters of the fitted lines were determined by using Microsoft Excel 2010 (Microsoft). Linear correlation of the blank-corrected fluorescence intensity values to the corresponding concentrations was a prerequisite for evaluation, at least in the concentration range of the assay.

3.8. Cloning of VEEV nsP2 and VEEV nsP2pro-2 expression plasmids

Plasmid encoding the cDNA of nonstructural proteins of VEEV were a gift from Dr. Christine L. Pugh (United States Army Research Institute of Infectious Diseases). Plasmids coding for the different VEEV PR constructs were prepared by Gateway Cloning Technology based on Tropea *et al*. (2007). The linear DNA sequences coding for VEEV nsP2 and VEEV nsP2pro-2 had been previously amplified in a two-step PCR and were transferred via pDON221 donor vector into pDEST™ 17 (further referred to as pDEST-His_6) and pDEST-His_6-MBP and according to manufacturer's instruction. Expression vectors pBB2546 (pDEST-His_6-VEEVnsP2), pBB2547 (pDEST-His_6-VEEVnsP2pro-2), pBB2549 (pDEST-His_6-MBP-VEEVnsP2), and pBB2550 (pDEST-His_6-MBP-VEEVnsP2pro-2) from the selected and enriched clones were isolated by Qiagen Miniprep Kit.

3.9. Expression of VEEV nsP2pro-2 and VEEV nsP2

Small-scale overnight starter cultures were initiated from freezer stocks of pBB2546, pBB2547, pBB2549 and pBB2550 plasmids in *E. coli* Rozetta cells. Next day 30 mL of the starter culture was added to 1000 mL LB medium containing 125 μg/mL ampicillin, 30 μg/mL chloramphenicol and 0.2 % glucose. Cells were incubated at 37 °C up to an absorbance of 0.5–0.6 at 600 nm, the protein expression was induced by the addition of 5 mL of 200 mM IPTG, followed by incubation for 4 hours at 30 °C while shaking at 250 rpm. Cells were harvested by centrifugation (5000 rpm, 10 min), the cell pellets were stored at -80 °C. To improve the solubility of His_6-MBP-VEEV nsP2, in case of pBB2549 plasmid, expression at 18 °C was also tested.

3.10. Purification of VEEV nsP2pro-2

Bacterial cells containing the His_6-MBP-VEEV nsP2pro-2 were suspended in 50 mL of ice-cold buffer A (50 mM sodium phosphate, 150 mM NaCl, 25 mM imidazole, pH 7.5). The cells were lysed with an APV Gaulin Model G homogenizer (Invensys, Albertslund, Denmark) at 10,000 psi and centrifuged at 30,000 g for 30 min at 4 °C. The supernatant was filtered through a 0.22 μm polyethersulfone membrane, applied to a 15 mL HisTrap FF crude affinity column (GE Healthcare, Piscataway, NJ) equilibrated in buffer A, and then eluted with a linear gradient from 5-50 % buffer B (50 mM sodium phosphate, 150 mM NaCl, 500 mM imidazole, pH 7.5). Fractions containing high amount the fusion protein were pooled and concentrated using an Amicon YM30 membrane (Millipore, Billerica, MA), then diluted 6-fold with 50 mM sodium phosphate buffer (containing 150 mM NaCl, pH 7.5) to reduce the imidazole concentration to approximately 25 mM. The fusion protein was then digested with a 5 mg/mL stock solution of the His-tagged TEV PR (70:1 v/v) overnight at 4 °C. Products of the digest were applied to a 20 mL HisTrap FF crude affinity column equilibrated with buffer A. Flow-through fractions containing high amount of VEEV nsP2pro-2 were pooled. The sample was concentrated to about 10 mg/mL using an Amicon YM30 membrane and applied to a HiPrep 26/60 Sephacryl S200 column (GE Healthcare) equilibrated with 50 mM Tris-HCl buffer (containing 150 mM NaCl, 2 mM Tris (2-carboxyethyl) phosphine hydrochloride (TCEP), pH 7.5). The peak fractions (5 mL each) containing the VEEV nsP2pro-2 protease were pooled and concentrated. The concentration of the purified and pooled VEEV nsp2pro-2 was determined as 2.9 mg/mL by A=280 method. The theoretical extinction coefficient and molecular weight were

calculated by ProtParam tool of ExPASy. Aliquots were flash-frozen with liquid nitrogen and stored at -80 °C until further use.

3.11. Ni-NTA magnetic bead-based assays

SAMBs were dissolved in cleavage buffer in order to generate the SAMB stock solution that is used for generation of the assay samples. The volume of the cleavage buffer was dependent on the individual experimental design and was calculated based on the density of magnetic bead, on the number and volume of samples to be assayed. For time course, inhibitory, and pH-dependence experiments, equal amounts of homogenous SAMB suspensions, while for substrate-dependent kinetic measurements, increasing amounts of the homogenous SAMB suspensions were measured into 2.0 mL Protein Lobind Micro-centrifuge tubes (Eppendorf). The tubes were applied to an MPC, the supernatant was removed, and the beads were suspended in equal volume of cleavage buffer. After the addition of the enzyme/enzyme buffer or elution buffer, the final volume of the reaction buffer was 70 µL. Substrate control (C) samples were also prepared in the same way as the reaction (R) samples but SAMBs were suspended in elution buffer, instead of the cleavage buffer. Parallel to the reaction samples, substrate blank (B) samples, where enzyme buffer was added to the SAMBs instead of the enzyme were prepared. Reactions were terminated by separating the magnetic beads on MPC and the fluorescence of the supernatants were measured. The amount of C-terminal fluorescent cleavage product in the reaction samples were calculated by dividing the blank corrected RFU values by the slope of the cleavage-buffer-based calibration curve. The concentration of the eluted substrate (mM) in the supernatants of the C sample was calculated by dividing their corrected RFU values by the slope of the elution buffer-based calibration curve. For calculating the molar concentration of the substrates in each R sample, the substrate concentration (mM) of the SAMB stock solution used for creating the assay samples was be determined according to the followings $c_{SAMB} = (c_C \times V_r)/V_{SAMB}$, where c_{SAMB} is the molar concentration of the SAMB stock solution; c_C is the molar concentration of the eluted substrate in the C sample; V_r is the volume of the reaction mixture after the addition of the reaction buffer and the enzyme buffer; and V_{SAMB} is the volume of the SAMB stock solution in the C sample.

3.12. Time course kinetic studies of HIV-1 PR

Cleavage of His_6-MBP-VSQNY↓PIVQ-mTurquoise2 and His_6-MBP-VSQNY↓PIVQ-mApple substrates (at 1.6 μM and 2.4 μM final concentrations, respectively) by HIV-1 PR (at 36.4 nM final concentration) was followed by measuring fluorescence after 0, 20, 40, and 60 minutes of incubation at 37 °C, while continuously shaking at 600 rpm. Molar concentration of the C-terminal fluorescent cleavage product in the reaction samples were plotted against time (min). Linear regression was performed, and the parameters of the fitted lines were determined by GraphPad Prism Version 5.00.

3.13. Inhibition of HIV-1 PR

His_6-MBP-VSQNY↓PIVQ-mTurquoise2 recombinant substrate (0.028 mM final concentration) was used to study inhibitory effect of amprenavir (ranging from 1 nM to 1 μM final total concentrations) on HIV-1 PR (at 36.4 nM final concentration). The assay samples were incubated for 60 minutes at 37 °C, while continuously shaking at 600 rpm. Initial velocity values (nMs^{-1}) were plotted against the logarithms of amprenavir concentrations (nM). Inhibitory effect at 50 % initial velocity value (IC_{50}) was determined by fitting five parameter logistic curve on the data using GraphPad Prism Version 5.00. The active enzyme concentration (6.05 nM) was determined at y=0 of the line fitted on the velocity values plotted against the amprenavir concentration (nM) in the low nanomolar range, and further used for assessing kinetic parameters and inhibitory constant.

3.14. Substrate-dependent kinetic studies of HIV-1 PR

Cleavage reactions were performed by using 36.4 nM HIV-1 PR (final, total enzyme concentration) on increasing amount of His_6-MBP-VSQNY↓PIVQ-mTurquoise2 and His_6-MBP-VSQNY↓PIVQ-mApple substrates. The assay samples were incubated for 60 minutes at 37 °C while continuously shaking at 600 rpm. Initial velocity values (nMs^{-1}) were plotted against the initial substrate concentration (mM). Kinetic parameters were determined at <20 % substrate turnover by Michaelis-Menten non-linear regression analysis using GraphPad Prism Version 5.00.

3.15. Measurements with TEV PR

The substrate-dependent kinetic measurements were performed as described in Section 3.14, by incubating the reactions with TEV PR (45.7 nM

final concentration) at 30 °C. To study dependence of enzyme activity on pH, TEV PR (91.4 nM final concentration) was incubated with His_6-MBP-VSQNY↓PIVQ-mTurquoise2 recombinant substrate (at 0.03 mM final concentration) in cleavage buffer. The pH of the cleavage buffer was set to 6.0, 6.5, 7.0, 7.5, 8.0, and 8.5. The obtained initial velocity values (nMs^{-1}) were plotted against pH. Significance of the differences between the initial velocity values of the enzyme at pH 6.5-8.5 range and the highest one measured at pH 7.0 were determined by using unpaired t-test.

3.16. Microplate-based specificity studies of VEEV nsp2pro-2

The specificity of VEEV nsP2pro-2 was studied by mEYFP-fused recombinant substrates coding either for the wild type or for the modified SFV-1 cleavage site sequences. For screening, a 96-well plate–based adaptation of the Ni-NTA magnetic bead-based assay platform was used. For screening of substrates, VEEV nsP2pro-2 (2.1-6.0 μM final concentration) was incubated with His_6-MBP-mEYFP substrates (1-5 μM final concentration) at 30 °C while continuously shaking at 600 rpm. Incubation time was set to be 40 minutes in case of P5, P4, P1, and P2' variants, while 20-22 hours for P1'-modified substrates. The final volume of the assay samples was 70 μL/well. The pH of the applied buffers was 7.5. During the experiments flat-bottom 96 well plates (Greiner BioONE or Qiagen) were applied in combination with 96-Well Magnet Type A magnetic particle concentrator (Qiagen). During the incubation and washing of the beads, plates were shaken by a digital shaker (IKA MS3). The determination of the concentration of the products and substrates were carried out as it is described in Section 3.11. Cleavage preferences of VEEV nsp2pro-2 on the different substrates was compared based on the substrate conversion (%) rate. Significance of the differences between the conversion rates of the P5, P4, P2, P1, and P2'variants and the wild-type substrate were analyzed (i) by unpaired t-test and (ii) by one-way analysis of variances (ANOVA).

3.17. Time course kinetic studies of VEEV nsP2pro-2

Cleavage of the wild-type His_6-MBP-SFV-1-mEYFP substrate (at 0.02 mM final concentration) by VEEV nsP2pro-2 (at 3.0 μM final total concentration) was followed by measuring fluorescence after 0, 5, 10, 15, 20, 30, 50, 70, and 90 minutes at 30 °C, while continuously shaking at 600 rpm. Molar concentration of the C-terminal fluorescent cleavage product in the reaction samples were plotted against time (min). Linear regression was

performed, and the parameters of the fitted lines were determined by GraphPad Prism Version 5.00.

3.18. Substrate-dependent kinetic studies of VEEV nsP2pro-2

Kinetic measurements were performed on the wild type and P4-Gln, P4-Thr, P4-Ala, P4-Gly, P1-Gly, P1'-Thr, and P2'-Ser variants of His_6-MBP-SFV-1-mEYFP. The samples were incubated for 10 minutes at 30 °C, while continuously shaking at 600 rpm. Evaluation of the enzyme kinetic parameters was performed as described in Section 3.14. When calculating the catalytic constants 100 % activity was assumed for the VEEV PR nsP2pro-2. Significance of the differences between the k_{cat}/K_M values of the tested variants and the wild-type substrate were determined by unpaired t-test.

3.19. In-solution digestion of the recombinant substrates

In-solution digestion reactions were initialized by the addition of 2-5 µL purified HIV-1 PR (4.1 µM) or 5 µL TEV PR (21 µM) or 5 µL VEEV nsP2pro-2 (73 µM) to their corresponding purified recombinant fluorescent substrates (0.01-0.03 mM) dissolved in cleavage buffer (70 µL final volume). The reaction mixtures were incubated for 1-22 hours at 37 °C in case of HIV-1 PR or at 30 °C in case of TEV PR and VEEV nsP2pro-2 and were stopped by the addition of 6X denaturing or non-denaturing sample loading buffer. Samples were analyzed by polyacrylamide gel electrophoresis (PAGE).

3.20. Kinetic measurements on HIV-1 PR by RP-HPLC

Enzyme reactions were initiated by the mixing of 5 µL purified HIV-1 PR (dialyzed against 2X cleavage puffer, 430 nM final concentration), 10 µL 2X cleavage buffer, and 5 µL synthetic oligopeptide substrate (in 0.47-2.35 mM final concentrations) representing the naturally occurring cleavage site between the matrix and the capsid protein (MA/CA) of HIV-1 PR (VSQNY↓PIVQ). The mixtures were incubated at 37 °C for 10 minutes and stopped by the addition of 9 volumes of 1 % trifluoroacetic acid (TFA). The samples were injected onto Nova-Pak C18 reversed-phase chromatography column (Waters Associates, Inc.) using an automatic injector. Separation of substrates and cleavage products was performed by using acetonitrile gradient (0 to 100 %) in water, in the presence of 0.05 % TFA. Peptides were detected at 206 nm, followed by integration of the peak areas. Reactions were monitored at < 20 % substrate hydrolysis. Kinetic parameters were determined as described in Section 3.14.

Active site titration of the enzyme was performed at the same reaction setting as above at 0.47 mM final oligopeptide substrate concentration but using 2X cleavage buffer contained protease inhibitor amprenavir in final concentration ranging from 1 nM to 80 nM. Active enzyme concentration (22.05 nM) has been determined as described in Section 3.13.

3.21. PAGE and in-gel renaturation of the fluorescent proteins

Purified recombinant protein substrates and cleavage products were analyzed by either native or sodium-dodecyl-sulfate (SDS)-PAGE using 14 % separating and 4 % stacking gel. In case of denaturing PAGE, the electrophoresis was followed by rinsing the gels in distilled water for 30 minutes at room temperature in order to wash the SDS out from the gel and thereby to renature the separated proteins. In case of non-denaturing PAGE analysis, the separated fluorescent proteins were in their native state, therefore the washing step could be omitted. The fluorescent proteins in the unstained gels were visualized by using Dark Reader Blue transilluminator and by using UV imaging function of an AlphaImager gel documentation system (ProteinSimple). For staining the gels, PageBlue Protein Staining Solution was applied, and then band intensities were determined by Dr. János András Mótyán using ImageJ 1.43 software and by GelAnalyzer 2010a program (www.gelanalyzer.com).

4. RESULTS AND DISCUSSION

4.1. Expression vector system

The empty pDEST-His_6-MBP-FP plasmids - that carry the coding sequence of a His_6 affinity tag and an MBP fusion protein, followed by a TEV PR cleavage site (ENLYFQ↓G), the 'cloning cassette', and a C-terminal fluorescent protein (FP) including ECFP, mTurquoise2, EGFP, mEYFP, mApple, or mCherry - were prepared by using Gateway cloning technology. Due to their common origin, the terminal sequences of the different FPs are often identical, therefore, the designed N1 and C primers are suitable for the amplification of those FPs whose terminal nucleotide sequence is identical to that of EGFP. The 'cloning cassette' contains cleavage sites of *PacI* and *NheI* restriction endonucleases next to each other and allows the insertion of the nucleotide sequence of the proteolytic cleavage site of interest by ligation. The dsDNA sequences coding for the cleavage site sequence of interest can be generated either by (i) random mutagenesis or (ii) by the hybridization of complementary, *E. coli*-optimized oligonucleotide primers. The flexibility of the 'cloning cassette' allows the insertion of wide variety of the sequences by a simple one-step PCR reaction into the 'empty' expression vector. The inexpensive and efficient generation of the fluorescent fusion substrates harboring the cleavage site of interest, makes the system especially attractive for the generation of substrate libraries. Notably, the coding sequences of the fluorescent proteins are not in the ORF without the insertion of the properly designed linear dsDNA fragments, therefore, the fluorescent proteins are translated only after a successful ligation.

4.2. Recombinant fluorescent fusion protein substrates

Recombinant fusion protein substrates contain an N-terminal hexahistidine affinity tag (His_6) fused to a maltose binding protein (MBP), followed by cleavage sites of TEV and the inserted cleavage site sequence of the protease of interest, while on the C-terminal end a fluorescent protein variant is fused to the substrate construct. The role of the His_6 affinity tag is to enable the recombinant proteins to be immobilized on metal-chelate surfaces, and to facilitate the purification of the proteins by immobilized metal affinity chromatography (IMAC). The N-terminal MBP fusion protein enhances the water-solubility of the recombinant protein substrate and improves its folding, while TEV protease cleavage site serves as an internal control cleavage site and makes the construct suitable for investigation TEV PR activity. In order to avoid substrate

aggregation, substrates possessing monomeric fluorescent protein forms including mTurquoise2, mEYFP, mApple, and mCherry were preferred. The flexibility of the platform allows the selection of the fluorescent proteins that most suit the experimental purposes and the given instrumentation.

The recombinant substrates were successfully expressed in *E. coli* BL21(DE3) strain in 50 mL final volume, and the small-scale expressions in 15 mL final volume were also optimized and used for the 96-well-based specificity screening of VEEV PR nsP2pro-2. As some FPs produced by *E. coli* cells may have a longer maturation time; in order to increase the fluorescent yield of the substrate solution, the protein translation can be optionally arrested by the tetracycline treatment, as it was applied in case of mEYFP-fused substrates expressed for VEEV nsP2pro-2 kinetic measurements.

Cells were lysed under native conditions, and the cleared lysates containing the soluble substrates were used directly for sample preparation in a Ni-NTA magnetic agarose bead-based protease assay system or for substrate purification. Purified substrates can be further used for calibration, in-solution digestion and/or for PAGE analysis.

4.3. Ni-NTA magnetic-bead-based protease assay

The substrates attached to Ni-chelate coated beads *via* their His_6 tag are accessible for cleavage by the protease of interest. Strong nature of this interaction results in low level of impurities and high signal-to-noise ratio of the assay. After processing, the N-terminal cleavage products (and the uncleaved substrates) remain attached to the beads, while the fluorescent C-terminal cleavage products are released into the supernatant and after separation can be detected with high sensitivity.

The workflow of the developed protease assay includes six steps. (i) The recombinant substrates, either purified or dissolved in the cleared cell lysate, are incubated with the affinity magnetic beads to prepare SAMBs. After several washing steps, the SAMBs are aliquoted for the preparation of assay samples. (ii) The initiation of reactions by the addition of the protease. Upon cleavage, the proteolytic fragments are released into the supernatant. (iii) The reactions are terminated by the separation of the magnetic beads from the reaction mixture containing the fluorescent cleavage products and the enzyme. (iv) The supernatants are applied to the wells of a microtiter plate and the fluorescence is determined by fluorimetry. (v) Calibration curves are generated using purified fluorescent substrates solved in each assay buffers. (vi) The concentrations of C-

terminal fluorescent cleavage products and the applied substrate are determined based on the calibration curves. There are three different sample types applied in the developed protease assay: (i) reaction (R) sample contains used for assessing proteolytic cleavage properties; (ii) substrate blank (B) sample to assess spontaneous substrate dissociation during the reaction; and (iii) substrate control (C) sample to determine the initial substrate concentration in the reaction.

Calibration procedures are performed to determine the quantities of fluorescent substrates and cleavage products. Based on our experiences, relative fluorescence intensities are directly proportional to the substrate concentration in the range applied in our experiments. We found that coefficient of variance (CV %) values of the slopes were <15 % among substrates possessing the same C-terminal FP and treated identically.

The principle of the assay has already been utilized earlier in similar protease assay, however our system may provide a novel tool for the investigation of proteases by combining and improving the advantages of some already existing methods by offering a true separation-based assay, supporting the determination of enzyme kinetic parameters and providing a detailed substrate quantification procedure. Our assay offers a low volume (down to 70 µL) and HTS-compatible format.

4.4. PAGE analysis and in-gel renaturation

The substrates and cleavage products, both in the case of the Ni-NTA magnetic bead-based assay and in-solution digestion, can be analyzed by native or reducing PAGE. In our PAGE experiments, cleavage reaction of HIV-1, TEV, and VEEV nsP2pro-2 proteases were analyzed. Samples were prepared at either denaturing or non-denaturing conditions. In contrast to non-denaturing conditions, denaturing sample preparation included application of SDS and β-ME, and heat-treatment prior to PAGE. Fluorescence of the protein bands in the case of non-denatured samples was readily detected in the gel after PAGE, under visible light (even by naked eye), by using a dark-reader blue light transilluminator or UV imaging using a gel documentation system. In case of denatured samples, where SDS distorts protein structure, fluorescence of the analyzed proteins could not be detected by any of the above listed tools. In order to regenerate protein structure and fluorescence of the separated proteins, the gels were washed with distilled water in order to remove SDS. We found that removal of the SDS made the detection of fluorescence possible even by naked eye or by using either blue light transilluminator or UV imaging. There were

differences observed in the renaturation abilities of mApple and mTurquoise2. mTurquoise2 showed substantially better renaturation ability compared to mApple, when illuminated by blue or UV light. Based on the densitometry, band intensity of the renatured mTurquoise2 was at least 50 % compared to the non-denatured. In contrast with this, mApple showed only ~10 % fluorescence intensity after renaturation, however, it was still detectable in the gel after UV illumination. The acylimine bond, being responsible in part for the red shift of mApple (and he vast majority of dsRed derived FPs) may be disrupted upon denaturation and this change is supposed to be responsible for the lower renaturation ability of mApple and mCherry.

We have observed that in contrast to denaturing conditions, the cleavage products were not separated fully from each other in the case of non-denaturing sample preparation. The advantage of protein denaturation/renaturation in this protease assay is that the cleavage products can be separated from each other based on their molecular weight, and their native charge or shape does not affect their migration during electrophoresis, however after renaturation the appropriate bands can be identified based on their fluorescence. This can be highly beneficial if the molecular weights of other cleavage products (or contaminants) that have no fluorescence closely resemble those of the fluorescent products. Relatively small (~2 kDa) differences can also be detected by separation of denatured cleavage products, which may be indicative if cleavages occur at alternative sites, as well.

4.5. Standardized manual and video protocol

In order to support the acquisition of the method and minimize initial assay failure rate, a tutorial video protocol has been prepared. The video protocol is and the corresponding detailed manual can be found here: https://www.jove.com/video/58824/use-recombinant-fusion-proteins-fluorescent-protease-assay-platform.

4.6. Kinetic and inhibition studies of HIV-1 PR

For the demonstration, that the developed assay system is suitable for the determination of enzyme kinetic parameters, we performed both time- and substrate-dependent kinetic measurements by HIV-1 PR, using mTurquoise2- and mApple-fused substrates. Kinetic time course studies revealed that the formation of the fluorescent cleavage products was satisfactorily linear in the examined timeframe at below the lowest substrate concentration of the

substrate-dependent studies. For the characterization and comparison of the proteolytic reactions, kinetic parameters were calculated based on the data of substrate-dependent kinetic measurements. Results demonstrated that k_{cat}/K_M values determined for HIV-1 PR using mApple and mTurquoise2-fused substrates closely resemble each other (18.5 ± 1.40 mM^{-1} s^{-1} and 19.69 ± 0.14 mM^{-1} s^{-1}), while the individual k_{cat} and K_M values differed substantially.

Comparison of the kinetic parameters of HIV-1 PR obtained by the magnetic bead-based protease assay with the values determined by using synthetic oligopeptide substrates (15.94 ± 4.82 mM^{-1} s^{-1}) revealed that k_{cat}/K_M values determined by the different methods closely resemble each other. Notably, the individual kinetic parameters were substantially higher on synthetic oligopeptide substrate, which can be interpreted by the conformational differences of the substrates, as oligopeptide substrates represents the binding sites in an extended form, in contrast to recombinant fusion protein substrate, where the cleavage site is likely to be folded in a different conformation and its flexibility can be affected by the different protein domains.

Previously, the specificity constant for HIV-1 PR on VSQNY↓PIVQ oligopeptide substrate have been determined as 45.3 mM^{-1} s^{-1} by our laboratory at highly different buffer conditions including pH 5.6 and high ionic strength (2 M NaCl), while in our current experiments the pH and the ionic strength was set to 7.0 and 300 mM NaCl, respectively, in order to optimize the performance of Ni-NTA magnetic beads in the assay. For the determination of the kinetic parameters, the active enzyme concentration of HIV-1 PR has been determined using amprenavir. The effect of amprenavir was tested on the cleavage of His_6-MBP-VSQNY↓PIVQ-mTurquoise2 recombinant fusion protein substrate and VSQNY↓PIVQ oligopeptide substrate, as well. Based on the dose-response curve of the Ni-NTA magnetic based inhibitory study the IC_{50} (21.43 nM) and K_i values (6.75 nM) were also determined.

4.7. Kinetic and pH dependence studies of TEV PR

Kinetic measurements of TEV PR were performed on His_6-MBP-VSQNY↓PIVQ-mTurquoise2 and -mApple substrates. The specificity constants determined for TEV PR by different substrates were highly similar and were comparable with previously published value by Kapust *et al.* (2002), however individual kinetic parameters differ substantially, which can be explained by the conformational differences between the applied substrate types, similarly to case of the HIV-1 PR kinetic results. Furthermore, differences of the reaction

conditions may also contribute to the difference in the kinetic parameters, as oligopeptide-based assay was performed at higher ionic strength (400 mM NaCl) and in the presence of reducing agents (*e.g.* dithiothreitol), compared to Ni-NTA magnetic bead-based assay, where 300 mM NaCl was applied in the absence of reducing agents. Experiments using TEV PR have given excellent opportunity to demonstrate the suitability of the Ni-NTA magnetic bead-based assay at different pH conditions. The activity of TEV PR was followed at pH between 6.0-8.5, and in accordance with the results of Parks *et al*., (1995), no significant differences were observed in the activities at pH 6.5-8.5. At pH 6.0, the interaction between Ni-NTA agarose beads and the His_6-tagged substrate weakens significantly, and we observed high rate of spontaneous dissociation of the intact substrate (>15 %), in agreement with the manufacturer's protocol.

4.8. Expression vectors encoding different VEEV PR construct

In this work, expression plasmids coding for two new VEEV PR constructs were prepared in an effort to generate a protease with improved activity. (i) One, representing the almost full-length VEEV nsP2 protein (Gly1-Thr785), as according to Das *et al.* (2014), Chikungunya virus nsP2 activity was supposed to be dependent on contributions from both the N- and C-terminal domains and this suggests the possibility of an interaction or some kind of functional crosstalk between them. (ii) The other construct called as VEEV nsP2pro-2, is a variant of VEEV nsP2pro-1 that contains a short N-terminal extension (Ala436-Met457). The rationale behind VEEV nsP2pro-2 design was that there is a strong sequence conservation in the middle of 436-457 region, and it is possible that this segment has structural importance for protease activity.

Beside engineering of the N-terminus, we decided to terminate the C-terminus of both constructs at Thr785. The ORFs of both VEEV nsP2 and VEEV nsP2pro-2 have been successfully cloned into both pDEST-His_6 and pDEST-His_6-MBP destination vectors.

4.9. Expression and purification of VEEV PR constructs

During bacterial expression in Rozetta cells, we have found that at 30 °C for 4 hours incubation after IPTG induction, both the His_6- and the His_6-MBP-fusioned VEEV nsP2pro-2 forms were overexpressed, however, only the His_6-MBP-fusioned form was confirmed to be soluble. These results are consistent with the previously described phenomena, that N-terminal MBP-fusion facilitates water-solubility of the recombinant proteins and improves

folding. In case of VEEV nsP2, both His_6- and His_6-MBP-attached forms were not pronouncedly overexpressed and none of the constructs were found to be soluble. Unfortunately, neither lower temperature (18 °C) nor longer incubation (overnight) improved the solubility. Based on this, we assume that poor solubility of the His_6-fused form may be caused by inherent structural properties and folding, which may have been improved by fusion with MBP, as it was seen in the case of His_6-MBP-VEEV nsP2pro-2. But, the relatively large size of His_6-MBP-VEEV nsP2 (~131 kDa) possibly overwhelms the advantageous effect of MBP and hindered successful expression at the applied conditions.

VEEV nsP2pro-2 have been successfully purified from its His_6-MBP attached form, at the very same purification conditions and laboratory instruments that were applied during the purification of nsP2pro-1 by Zhang *et al.* (**2009**). The final product was judged to be at least 95 % pure by SDS-PAGE.

4.10. Specificity and study of VEEV nsP2pro-2

We have studied VEEV nsP2pro-2 activity on His_6-MBP-mEYFP fusion protein substrates comprising cleavage sites representing either the wild type or modified SFV-1 sequence. SFV-1 variants were designed by the substitution of residues at P5, P4, P2, P1, P1', and P2' sites. Due to its proximity to the cleavage site, the importance of the Gly residue in P1' position was investigated extensively by testing all the possible amino acid substitution in this position. Furthermore, other SFV-1 variants have also been generated, based on natural cleavage site sequences of other related alphavirus strains. These variants include the following modifications: P5-Gln, P4-Glu, P4-Thr, P4-Arg, P4-Gly, P2-Ala, P2-Val, P1-Gly, P1-Val, P2'-Pro, and P2'-Ser.

We have compared the cleavability of the different variants to that of the wild type SFV-1 substrate. The results of the 96-well microplate-based screening revealed substantially lower substrate conversion rates for all P1' variants as compared to the wild type. For P4-Glu and P4-Thr, we observed significantly higher substrate turnover as compared to wild-type substrate, while in case of P4-Arg *versus* the wild-type the difference was not statistically significant. Moderate decrease in the substrate conversion rates has been observed for P4-Gly, P1-Gly, P2'-Pro and P2'-Ser variants compared to the wild type, respectively, while the activity of VEEV nsP2pro-2 was almost fully or completely inhibited by P5-Gln, P2-Ala, P2-Val and P1-Val substitutions. For all these cases, the statistical probe resulted in statistically significant

differences compared to the wild-type. The results of one-way ANOVA were in accordance with that of the unpaired t-test in each case.

The results are part of a manuscript that is under preparation for submission and are to be interpreted along with the *in silico* results of Dr. János András Mótyán.

4.11. Kinetic measurements of VEEV nsP2pro-2

We have successfully determined the k_{cat}, K_M and k_{cat}/K_M values for VEEV nsP2pro-2 on the wild type and P4-Glu, P4-Thr, P4-Ala, P4-Gly, P1-Gly, P1'-Thr, and P2'-Ser variants of His_6-MBP-SFV-1-mEYFP recombinant substrate. In agreement with results of the microplate-based specificity results, highest catalytic efficiency was determined for P4-Gln variant (5-fold higher as compared to the wild type substrate) and P4-Thr and P4-Arg mutants were also found to be processed with 3.7- and 1.7-fold higher cleavage efficiency, respectively. Although in case of P4-Arg the difference was not statistically significant. The k_{cat}/K_M values of variants including P4-Gly, P1-Gly, and P2-Ser were calculated to be within the lower error range of the k_{cat}/K_M value of the wild type cleavage site, and were statistically similar to the wild-type. In contrast, catalytic efficiency on P1'-Thr variant was found to be approximately 4-fold lower than that of the native sequence, this difference was statistically significant.

The kinetic parameters determined on the wild-type substrate were compared to the literature data obtained for VEEV nsP2pro-1 (Zhang *et al.*, 2009; Hu *et al.*, 2016). The difference between the previously reported and the herein determined K_M values is considered to be caused by the conformational differences of the applied peptide/protein substrates, as it has been observed in the case of TEV PR and HIV-1 PR kinetic measurements, as well. The k_{cat} and k_{cat}/K_M values are not fully comparable because no selective inhibitor has been identified for this enzyme. However, if enzyme activity is considered as 100 % at each reaction conditions, k_{cat}/K_M values are within the same order of magnitude, which suggests that the catalytic efficiency of the investigated VEEV nsP2pro constructs is similar on the different substrate-types comprising the wild-type cleavage site sequence of SFV-1. These results imply that extension of VEEV nsP2pro-1 by the highly conserved Ala436-Met457 segment does not cause substantial change in the activity of VEEV protease *in vitro*.

5. SUMMARY

We presented the development and the utility of a Ni-NTA magnetic bead-based fluorescent protease assay platform using recombinant protein substrates. The substrates consist of N-terminal His_6 and MBP fusion tags, a TEV PR cleavage site sequence, the recognition site for the protease of interest and a C-terminal fluorescent protein. The expression vector contains a cloning cassette between the TEV PR recognition site and FP domain, which cassette enables the one-step insertion of a short dsDNA that can be generated *via* random mutagenesis or oligonucleotide primers and can code for the protease cleavage site of interest. The fluorescent substrate expression was demonstrated at different scales. To test the on-bead application of the substrates, enzyme kinetic, inhibition, and pH optimum studies were performed on TEV and HIV-1 PRs. Besides fluorimetric studies, substrates and products were assayed by PAGE, as well. We found that the denatured proteins could be renatured after reducing SDS-PAGE, and the fluorescent proteins were detected in the gel upon blue light and/or UV illumination. Interestingly, different fluorescent proteins were found to have different ability for renaturation. The standardized working instructions and a corresponding tutorial video of on-bead application and the PAGE analysis were prepared to support the execution of the assay procedure.

The platform was applied in a microplate-optimized format to characterize the *in vitro* activity and specificity of the purified VEEV nsP2pro-2 construct on mEYFP-fused SFV-1 substrate variants modified at P5, P4, P2, P1, P1', and P2' positions. Kinetic parameters were determined for seven variants and for the wild type. The results indicated that VEEV nsP2pro-2 processed P4-Glu, P4-Thr variants with significantly higher, P4-Arg, P4-Gly, P1-Gly and P2'-Ser variants with not significantly different, while P1'-Thr variant with significantly lower efficiency compared to wild-type substrate. VEEV nsP2pro-2 did not show improved catalytic performance on the wild type SFV-1 cleavage site compared to VEEV nsP2pro-1, previously purified and examined by our collaborators.

The developed assay platform can be adaptable to HTS and automation-based pharmaceutical protease substrate discovery, inhibitor screening and/or drug development, whereas the results of VEEV PR specificity studies could contribute to the investigation of alphaviruses or other important Group IV viruses.

ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my advisor prof. Dr. József Tőzsér, for providing me with the opportunity to work in the Laboratory of Retroviral Biochemistry. Beside my advisor, I am greatly beholden to Dr. János András Mótyán, the 'best labmate ever', for his precious instructions, great ideas, and support from lab bench to scientific publications. Additionally, I would like to say special thanks to the co-authors of my publications including Dr. Ferenc Tóth for the expression and purification of HIV-1 PR and to Lívia Gazda and Márió Miczi for their precious support in the kinetic and PAGE measurements of the study. I would like to express my great thanks to Dr. David S. Waugh for his collaboration and for providing me with the opportunity for a short-term internship in MCL, Center for Cancer Research, NCI via Campus Hungary Program TÁMOP-4.2.4B/2-11/1-2012-0001. During this internship I have performed the cloning and purification work of VEEV nsP2 and VEEV nsP2pro-2 construct.

Furthermore, I would like to say special thanks to Prof. Dr. Attila Bácsi, for providing me with the access to Biotek Synergy2 device in the Department of Immunology, and to the co-worker of the Multimedia and E-learning Technical Center for the production of the audio and the video record of the JoVE publication

My sincere gratitude goes to my colleagues and fellow labmates at the Laboratory of Retroviral Biochemistry; great thanks to Szilvia Janics-Pető, Dr. Krisztina Matúz, Dr. János Kádas and Mária Golda for their invaluable brain-stormings, chats, and for their helping hands.

I would like to extend my gratitude to my former colleagues at Gedeon Richter Plc, especially to Dr. Zoltán Urbányi, Dr. Márton Megyeri and to Dr. Erik Bogsch for their constant encouragement.

Special thanks to Kornél, Gergő and Patrik Csontos for allowing me time away from my family. I am greatly beholden to my mother, my sister, my father and to my mother- and farther-in-law. And the last but not at least, I exclusively dedicate this thesis to my grandparents Sándor Papp and Julianna Sápi.

The financial support of the is work was provided by Hungarian Scientific Research Fund and PHARMPROT teaming project GINOP-2.3.2-15-2016-00044.

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/375/2019.PL
Subject: PhD Publikációs Lista

Candidate: Beáta Bozóki
Neptun ID: SDC8DN
Doctoral School: Doctoral School of Molecular Cellular and Immune Biology

List of publications related to the dissertation

1. **Bozóki, B.**, Gazda, L., Tóth, F., Miczi, M., Mótyán, J. A., Tőzsér, J.: A recombinant fusion protein-based, fluorescent protease assay for high throughput-compatible substrate screening.
   Anal. Biochem. 540-541, 52-63, 2018.
   DOI: http://dx.doi.org/10.1016/j.ab.2017.11.001
   IF: 2.507
2. **Bozóki, B.**, Mótyán, J. A., Miczi, M., Gazda, L., Tőzsér, J.: Use of Recombinant Fusion Proteins in a Fluorescent Protease Assay Platform and Their In-gel Renaturation.
   JoVE. 143, 1-15, 2018.
   DOI: http://dx.doi.org/10.3791/58824
   IF: 1.108

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

List of other publications

3. Mótyán, J. A., Miczi, M., **Bozóki, B.**, Tőzsér, J.: Data supporting Ni-NTA magnetic bead-based fluorescent protease assay using recombinant fusion protein substrates.
Data in Brief. 18, 203-208, 2018.
DOI: http://dx.doi.org/10.1016/j.dib.2018.03.031

4. Bagossi, P., Bander, P., **Bozóki, B.**, Tőzsér, J.: Discovery and significance of new Human T-lymphotropic viruses: HTVL-3 and HTLV-4.
Expert Rev. Anti. Infect. Ther. 7 (10), 1235-1249, 2009.
IF: 2.857

Total IF of journals (all publications): 6,472
Total IF of journals (publications related to the dissertation): 3,615

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.

20 November, 2019

DEBRECENI EGYETEM
Egyetemi és Nemzeti Könyvtár