Investigation of the biochemistry and function of neutrophil serine protease 4 (NSP4)

Dissertation zur Erlangung
des Doktorgrades der Naturwissenschaften
an der Fakultät für Biologie
der Ludwig-Maximilians-Universität München

angefertigt am
Max-Planck-Institut für Neurobiologie,
Abteilung Neuroimmunologie
und
Helmholtz Zentrum München,
Comprehensive Pneumology Center

vorgelegt von
Natascha Cynthia Perera
aus Sigmaringen

München, September 2013

Erstgutachterin:	Prof. Dr. Elisabeth Weiß
Zweitgutachterin:	PD Dr. Barbara Lösch
Mitgutachter:	Prof. Dr. Heinrich Jung
	Prof. Dr. Angelika Böttger
Sondervotum:	PD Dr. Dieter Jenne
Einreichung der Dissertation:	12. September 2013
Datum der mündlichen Prüfung:	31. Januar 2014

Summary

Serine proteases in cytoplasmic granules of neutrophils (NSPs), namely neutrophil elastase, cathepsin G (CG) and proteinase 3, have been under intense investigation for several decades. They are mainly known for their role in intracellular killing of pathogens and are also increasingly recognized as key regulators of innate immune responses. In 2009, I identified a fourth serine protease in neutrophils that has been completely overlooked and neglected so far. The aim of this thesis was an in-depth biochemical and functional characterization of this novel serine protease 4 (NSP4) of human neutrophils.

Using monoclonal antibodies to NSP4, the distribution of NSP4 in normal human tissues was studied. NSP4 was observed only in neutrophils and neutrophil precursors of the bone marrow. The content of NSP4 in neutrophil lysates was about 20-fold lower compared to CG. Nevertheless, NSP4 was found to be released into the supernatant upon neutrophil activation. NSP4 could be further identified as a novel azurophil granule protein of neutrophils by Western blot analyses of subcellular fractions. For the functional analysis, the production and yield of recombinant NSP4 was clearly improved using different expression systems and DNA construct modifications. The proteolytic specificity was analyzed using *E. coli* peptide libraries, mass spectrometry and several synthetic peptide libraries. All these analyses clearly revealed an arginine specificity for NSP4. Consistent with this, NSP4 was strongly inhibited by heparin-accelerated antithrombin and C1 inhibitor and, with lower efficacy, by α_1-proteinase inhibitor (α1PI). The data allowed me to generate an NSP4-specific α1PI variant that was shown to form covalent complexes with all NSP4 of neutrophil lysates and supernatants of activated neutrophils. This finding strongly indicated that NSP4 is fully processed and stored as an already activated enzyme in azurophil granules. In addition, dipeptidyl peptidase I (DPPI) was identified as the activator of NSP4 *in vivo*, as DPPI deficiency resulted in complete absence of NSP4 in a Papillon-Lefèvre patient. Analysis of cell-based calcium assays revealed that proteinase-activated receptor-2 may represent a potential natural substrate of NSP4. So far, NSP4-deficient mice did not show an abnormal phenotype under clean housing conditions. Activation of isolated neutrophils by phorbol esters or immune complexes was also not impaired. This study establishes NSP4 as the only arginine-specific pre-activated serine protease stored in azurophil granules of neutrophils that may fullfil a quite distinct, supportive role in neutrophil responses to tissue damage and bacterial infections.

Zusammenfassung

Die Granula-assoziierten Serinproteasen der Neutrophilen, genannt Neutrophilen-Elastase, Cathepsin G (CG) und Proteinase 3, werden bereits seit über 30 Jahren intensiv erforscht. Sie sind hauptsächlich für ihre Funktion bei der intrazellulären Degradation von pathogenen Keimen bekannt. Außerdem werden sie zunehmend als wichtige Regulatoren der angeborenen Immunantwort angesehen. Im Jahr 2009 identifizierte ich eine vierte Serinprotease in Neutrophilen, die bis dahin komplett übersehen wurde. Das Ziel dieser Arbeit war eine ausführliche Untersuchung der Biochemie und Funktion dieser neuen Serinprotease 4 (NSP4) in humanen Neutrophilen.

NSP4 konnte in immunhistochemischen Analysen normaler Gewebe nur in Neutrophilen und deren Vorläuferzellen im Knochenmark nachgewiesen werden. Die Menge von NSP4 im Gesamtzelllysat von Neutrophilen war 20-fach geringer als die von CG. Dennoch konnte ich nachweisen, dass NSP4 von aktivierten Neutrophilen sezerniert wird. In subzellulären Neutrophilen-Fraktionen war NSP4 in erster Linie mit azurophilen Granula assoziiert. Die Produktion und Ausbeute rekombinanter, aktiver NSP4 wurde mithilfe verschiedener DNA-Konstrukte und Expressionssysteme deutlich verbessert. Für die Bestimmung der proteolytischen Spezifität wurden *E. coli* Peptidbibliotheken, Massenspektrometrie und synthetische Peptidbibliotheken eingesetzt. Alle Ergebnisse ergaben eindeutig eine Arginin-Spezifität für NSP4. NSP4 wurde sehr gut von Heparin-Antithrombin, C1 Inhibitor und mit geringerer Effizienz von α_1-Proteinase-Inhibitor (α1PI) blockiert. Die Ergebnisse ermöglichten die Herstellung einer NSP4-spezifischen α1PI Variante, die nachweislich mit der gesamten NSP4 aus Neutrophilen-Zelllysat und aus Überständen aktivierter Neutrophilen einen kovalenten Komplex bildete. Dies war ein klarer Hinweis für die Speicherung und Sekretion der NSP4 als aktive Endoprotease. Außerdem konnte ich zeigen, dass NSP4 *in vivo* von Dipeptidylpeptidase I (DPPI) aktiviert wird, da in einem Papillon-Lefèvre Patienten ohne funktionelle DPPI auch NSP4 nicht nachweisbar war. Intrazelluläre Kalziummessungen in HaCaT-Zellen ergaben, dass der Proteinase-aktivierte Rezeptor-2 ein mögliches natürliches Substrat für NSP4 sein könnte. In einer Pathogen-freien Reinraumumgebung zeigten NSP4-defiziente Mäuse bislang noch keinen Phänotyp. Auch die Aktivierung isolierter Neutrophilen durch Phorbolester oder Immunkomplexe war nicht beeinträchtigt. In dieser Arbeit gelang es mir, die erste aktive Arginin-spezifische Serinprotease (NSP4) in azurophilen Granula von Neutrophilen zu identifizieren. Wegen ihres Vorkommens in allen sequenzierten Genomen höherer Vertebraten könnte NSP4 eine wichtige Rolle bei der Steuerung und Verstärkung Neutrophilen-abhängiger Immunantworten zukommen.

Acknowledgments

I would like to thank, primarily, my supervisor **PD Dr. Dieter Jenne** for teaching me how good research is done, for encouraging me to have my own ideas and for always being open for discussions and brainstorming of any kind. I highly appreciated his incredibly huge knowledge and broad interest in science and beyond.

I am very grateful to **Prof. Dr. Elisabeth Weiß**, my doctoral thesis supervisor at the faculty of Biology at the Ludwig-Maximilians-Universität München, for her time, valuable suggestions and support during the past years. I also want to thank all members of my thesis examination committee, namely **PD Dr. Barbara Lösch**, **Prof. Dr. Heinrich Jung**, **Prof. Dr. Angelika Böttger**, **Prof. Dr. Michael Schleicher** and **Prof. Dr. Heinrich Leonhardt** for their time and contribution to this work.

I want to thank **Prof. Dr. Hartmut Wekerle** for giving me the opportunity to work on this thesis in the Max-Planck-Institute of Neurobiology (MPIN), for his helpful feedback during progress meetings and his continuous interest in the project.

Prof. Dr. Oliver Eickelberg is acknowledged for hosting our group in the Comprehensive Pneumology Center (CPC) of the Helmholtz-Zentrum München. I am very grateful for his support during and after our relocation and for the warm welcome that we received in the CPC.

Prof. Dr. Wolfram Bode and **Prof. Dr. Achim Krüger** supported me as members of my thesis advisory committee and I am thankful for their contribution of ideas, opinions and time towards my research.

Further, I acknowledge the financial support by the **European Union Seventh Framework Programme** (FP7/2007-2013) under Grant Agreement 261382 (INTRICATE).

The whole Jenne Group contributed by providing a fantastic place to work and I want to thank my great colleagues **Heike Kittel**, **Lisa Stegmann**, **Thérèse Dau** and **Lisa Hinkofer**. I could always rely on them for support, scientific brainstorming and a lot of fun.

I have also benefitted greatly from my fellow students and colleagues, I want to thank especially **Marsilius Mues**, **Michael Koutrolos**, **Johannes Plett**, **Sarah Laurent**, **Marija Pesic**, **Latika Bhonsle**, **Markus Krumbholz** and **Franziska Uhl** for making my time in the lab and outside a great and very memorable experience. I would further like to extend my gratitude to all members of the MPIN and the CPC.

I dearly thank my boyfriend **Robert** for constantly believing in me, encouraging me and because he can always make me laugh.

I would like to dedicate this work to my loving family. My sister **Nadine** has encouraged and supported me for as long as I can remember. And I could always rely on my parents **Numan** and **Christine** for constant interest, support and understanding and I want to thank them for always being there for me.

Table of contents

1 Introduction

1.1 Neutrophil granulocytes

Differential blood counts show the percentage of different blood cell populations in the peripheral blood. Normally, neutrophil granulocytes, in the following briefly called neutrophils, are most abundant and make up 50 - 70% of all white blood cells in the circulation. The abundance of these cells already points to the importance of these cells for immune defense responses. Neutrophils are relatively short-lived cells that develop from myeloid progenitor cells in the bone marrow to mature band and segmented neutrophils. Neutrophils are constantly released into the blood stream. In response to local inflammatory danger signals, neutrophils are among the first immune cells that extravasate from the blood stream and migrate into infected or damaged tissues to form the early immune response (Nathan, 2006).

One hallmark of fully differentiated neutrophils is their lobulated nucleus. Hence, together with the other granulocytes, they were also specified as polymorphonuclear cells (PMN) of which they constitute the major part. Eosinophil and basophil granulocytes constitute only 1 - 5% and 0.5 - 1% of total leukocytes, respectively.

1.1.1 Formation of granules and of granule proteins

As the name granulocyte suggests, neutrophils are densely packed with a multitude of granules that are filled with bactericidal peptides and proteins like alpha-defensins, bactericidal/ permeability-increasing protein (BPI), myeloperoxidase (MPO) and several proteases (Faurschou and Borregaard, 2003). The granules are formed consecutively during neutrophil maturation in the bone marrow. Primary or azurophil granules are formed in myeloblasts and promyelocytes, secondary or specific granules appear during the myelocyte and metamyelocyte

stage and tertiary or gelatinase granules are made in band and segmented cells. The three distinct subsets were classified based on the presence of characteristic proteins: MPO, neutrophil gelatinase associated lipocalin (NGAL) and gelatinase B (matrix metalloproteinase 9, MMP-9) are markers for azurophil, specific and gelatinase granules, respectively. Secretory vesicles constitute the fourth type of storage organelles in neutrophils. These are created by endocytosis during the late stages of neutrophil development (Borregaard, 2010).

Le Cabec et al. (1996) have demonstrated that the storage process in granules is regulated by timing and not by sorting. The primary granules contain proteins that are expressed at the very early stage of neutrophil differentiation, the proteins in the secondary granules are expressed later and the content of the tertiary granules is made during the late maturation phase of band and segmented cells. Therefore, the time of expression of a gene can indicate in which granules the corresponding protein is stored.

1.2 Neutrophil immune response

The human body is firstly protected from microbe infections by the physical barriers provided by the skin and mucus membranes. If pathogenic microorganisms have overcome this barrier and gained access to tissues, local endothelial cells are activated by signals from microbes and macrophages in the tissues. Neutrophils are then recruited by activated endothelium and enter the tissue after extravasation. This transmigration through the blood vessel wall involves a multi-step cascade of tethering, rolling, firm attachment and diapedesis (Zarbock and Ley, 2008). In the course of this process, neutrophils are activated by integrin signaling and G-protein coupled receptor activation. At the site of infection, the central function of neutrophils is the elimination of pathogens. This is mainly achieved by phagocytosis and fusion of the endosome with the neutrophil cytoplasmic granules. Engulfed microorganisms can be destroyed in several ways, e.g. by reactive oxygen species (ROS), generated by NADPH-oxidase and MPO, bactericidal peptides and proteins or proteases like the neutrophil serine proteases. Activation of neutrophils also results in granule exocytosis. Secretory vesicles are released after priming of neutrophils and with increasing levels of neutrophil activation, also the tertiary, secondary and primary granules are sequentially released in this order.

While neutrophils are indespensable in the innate immune defense against intruding microorganisms (Newburger, 2006), prolonged neutrophil accumulation can also lead to pathological tissue damage as seen in several inflammatory diseases, like rheumatoid arthritis, inflammatory bowel disease and chronic lung diseases (Weiss, 1989).

1.3 Neutrophil serine proteases (NSPs)

1.3.1 Biosynthesis

Three NSPs have been known for more than 30 years (Baggiolini et al., 1978): neutrophil elastase (NE), cathepsin G (CG) and proteinase 3 (PR3). The genes for NE (ELA2, ELANE) and PR3 (PRTN3) are tightly linked in a cluster found on the short arm of human chromosome 19 (19p13.3) (Zimmer et al., 1992), whereas CTSG encodes CG and is found in a separate cluster on human chromosome 14 (14q11.2) with genes encoding granzymes and mast cell chymases (Caughey et al., 1993). All three are expressed as inactive precursors (zymogens) featuring a signal peptide and an N-terminal propeptide. After signal peptide processing, the two residues of the propeptide, actually a dipeptide, are removed by dipeptidylpeptidase I (DPPI, cathepsin C). Like all serine proteases of the chymotrypsin/trypsin family, the NSPs can only take up their mature form if the conserved N-terminal Ile16 is free to interact with Asp194 (numbering according to bovine chymotrypsinogen). This process leads to a conformational change rendering the active site accessible for potential substrates (Freer et al., 1970). After processing by DPPI, the NSPs are stored as mature enzymes in the azurophil granules of neutrophils (Adkison et al., 2002; Korkmaz et al., 2010).

1.3.2 Role of NSPs

The NSPs contribute to the bactericidal activity of neutrophils. Investigation of NSP-deficient mouse models has shown that CG and NE are important in killing certain Gram-negative and Gram-positive bacteria (Belaaouaj et al., 1998, 2000; Reeves et al., 2002) and clearing certain fungi (Tkalcevic et al., 2000; Reeves et al., 2002).

In addition, NSPs are also attributed a potential key function in the regulation of inflammation (Wiedow and Meyer-Hoffert, 2005; Pham, 2008; Kessenbrock et al., 2011). As the NSPs are also partly released by activated neutrophils, they can proteolytically modify chemokines, cytokines or cell surface receptors and thereby shape the immune response at a very early stage. For instance, Sambrano et al. (2000) have demonstrated that CG can activate the proteinase-activated receptor 4 (PAR-4) on platelets. In addition, Raptis et al. (2005) have revealed a role for CG in integrin clustering on neutrophils after interaction with immobilized immune complexes, leading to cytoskeletal rearrangements and increased ROS production. And PR3 and NE were both shown to degrade anti-inflammatory progranulin after immune complex-mediated inflammation *in vivo*, resulting in enhanced neutrophil dependent inflammation (Kessenbrock et al., 2008).

1.3.3 NSPs in human diseases

Papillon-Lefèvre syndrome (PLS)

All three NSPs are affected in PLS (Gorlin et al., 1964), a rare autosomal recessive disease that is caused by loss-of-function mutations in the gene encoding DPPI, the natural convertase of NE, CG and PR3. PLS is characterized by severe early-onset periodontitis and palmoplantar hyperkeratosis. Pham et al. (2004) have shown that the levels of NSPs are severely reduced in neutrophils of PLS patients. The unconverted zymogens of NE, CG and PR3 are thought to be lost by degradation, continuous secretion or other unknown mechanisms. However, neutrophils of PLS patients are not uniformly defective in their antimicrobial activity against common bacteria. This illustrates that the neutrophil bactericidal activity is not purely dependent on NSPs (Pham et al., 2004).

Hereditary neutropenias

Cyclic neutropenia and severe congenital neutropenia are caused by maturation arrest of neutrophil precursors during myelopoiesis. Mutations in the ELA2 gene (encoding NE) were linked to hereditary neutropenias (Horwitz et al., 2004) and are thought to disrupt normal intracellular

trafficking of NE. Intracellular accumulation of mutant NE could then result in apoptosis after induction of the unfolded protein response (Köllner et al., 2006).

Non-infectious inflammatory diseases

Animal model analysis has implicated NSPs in several non-infectious inflammatory diseases. Both a mouse model deficient in DPPI (convertase of NE, CG and PR3) and a mouse model deficient in NE and CG were protected against acute arthritis induced by anti-collagen antibodies (Adkison et al., 2002). Mouse models deficient in NE were found to be partially protected from smoke-induced emphysema (Shapiro, 2003) as well as from bleomycin-induced pulmonary fibrosis (Chua et al., 2007). Furthermore, Shimoda et al. (2007) reported that CG accelerates pathogenic inflammation in mice after ischemia/reperfusion injury.

Based on their anticipated broad impact on inflammatory processes, NSPs as a whole appeared to represent a very attractive therapeutic target. Indeed, both industry and academia have attempted to suppress the activity of NE, CG and PR3 by therapeutic inhibitors (Groutas et al., 2011; Korkmaz et al., 2010). However, besides α1-proteinase inhibitor (α1PI, α1-antitrypsin), a natural inhibitor of NE, small-molecule inhibitors have yielded unsatisfactory therapeutic results and did not receive clinical registration so far (Lucas et al., 2013).

1.4 Class 6 serine proteases

Based on their exon-intron organization, NE, CG and PR3 can be grouped to the sixth separate class of serine proteases (Jenne, 1994). The exon-intron organization refers to the number of exons and introns, and the phases of the introns. Introns can interrupt exons between two codons, after the first or after the second codon base, termed phase 0, 1 or 2, respectively. Intron phases are highly conserved and relatively stable in evolution, as mutations changing the intron phase requires concerted mutations both at the donor and acceptor site of an intron.

This sixth class in humans further encompasses the five lymphocyte granzymes A, B, H, K and M, the catalytically inactive serine protease homolog azurocidin, the mast cell chymase and complement factor D (CFD). All class members are implicated in immune defense reactions.

The granzymes and NSPs are important for the elimination of virally infected cells and bacteria, respectively, while CFD and mast cell chymase amplify the innate immune response by different mediator-dependant mechanisms.

1.5 Neutrophil serine protease 4 (NSP4)

A total of 178 serine protease homologs were identified in humans (Page and Di Cera, 2008). One of these could be identified as an additional member of the class 6 serine proteases, based on its exon-intron structure. PRSS57, the gene encoding NSP4, was annotated by computer algorithms as "protease, serine, 57". It exhibited 39% sequence identity to NE and PR3 and also featured four introns with the same phases, 1-2-0-0, and at identical positions as in granzymes and NSPs.

In my diploma work, 14 monoclonal antibodies (mAbs) to NSP4 were generated in cooperation with Dr. Elisabeth Kremmer (Helmholtz Zentrum München) and characterized by myself in much detail. In these studies, I showed that three of these mAbs were specific for NSP4 and did not cross-react with NE, CG, PR3, azurocidin or the five granzymes. Applying them in Western blot analyses of total cell lysates of different human blood cell populations, I demonstrated that NSP4 is present in neutrophils, but not in natural killer cells or other peripheral blood mononuclear cells (PBMCs). Furthermore, NSP4 showed enzymatic activity by cleaving the synthetic substrate Boc-Ala-Pro-Nva-Sbzl after norvaline, thus displaying a similar substrate specificity as NE and PR3. Based on the presence of this novel active serine protease in neutrophils, I designated it neutrophil serine protease 4, abbreviated as NSP4 in the following.

1.6 Protease specificity

1.6.1 The Schechter and Berger nomenclature

The specificity of a protease denotes the amino acid sequence that is preferentially cleaved by this protease. Schechter and Berger (1967) have introduced a widely accepted nomenclature for describing the interaction of a protease and its substrate. A prerequisite for successful peptide

bond hydrolyzation is the proper interaction between amino acid residues close to the peptide bond that is to be cleaved (scissile bond) and the active site of the protease. The active site in serine proteases accommodates the catalytic triad which is a highly coordinated structure consisting of His57, Asp102 and Ser195 (numbering according to bovine chymotrypsinogen). The cooperative action of the catalytic triad enables a nucleophilic attack of the hydroxyl group of Ser195 on the carbonyl carbon of the scissile bond of the substrate.

The substrate amino acids close to the scissile bond were designated P-residues (peptide). Thereby the residues N-terminal of the scissile bond are referred to as P1, P2, P3, P4, etc. and the residues on the C-terminal side P1', P2', P3', P4', etc. The scissile bond is located between P1 and P1', as illustrated in Figure 1.1. Accordingly, the protease subsites close to the active site that take up the respective substrate amino acids are called S1 and S1', S2 and S2', S3 and S3', etc.

1.6.2 Proteolytic specificities of NE, CG and PR3

NE preferentially cleaves substrates with medium to small aliphatic residues in P1, like Val, Ile, Leu or Ala (Bode et al., 1989). The specificity of PR3 is similar to NE, small aliphatic residues like Ala and Val are preferred in P1 (Rao et al., 1991). CG features chymotrypsin-like activity and strongly prefers Phe or Tyr at the P1 substrate position, but also cleaves substrates after Lys or Arg with much lower efficacy (Wysocka et al., 2007). This additional low Lys/Arg specificity of CG was only reported for human CG and was not seen in analyses of mouse or other mammalian CG. The slight change in specificity is thought to result from the Ala226Glu change that occured late in primate evolution and is only present in human CG (Raymond et al., 2010).

1.7 Inhibitors of neutrophil serine proteases

Under normal circumstances, excessive extracellular activity of NE, CG and PR3 is antagonized by distinct protease inhibitors. NE is inhibited by α1-proteinase inhibitor (α1PI, α1-antitrypsin), elafin (also known as skin-derived antileukoprotease), monocyte neutrophil elas-

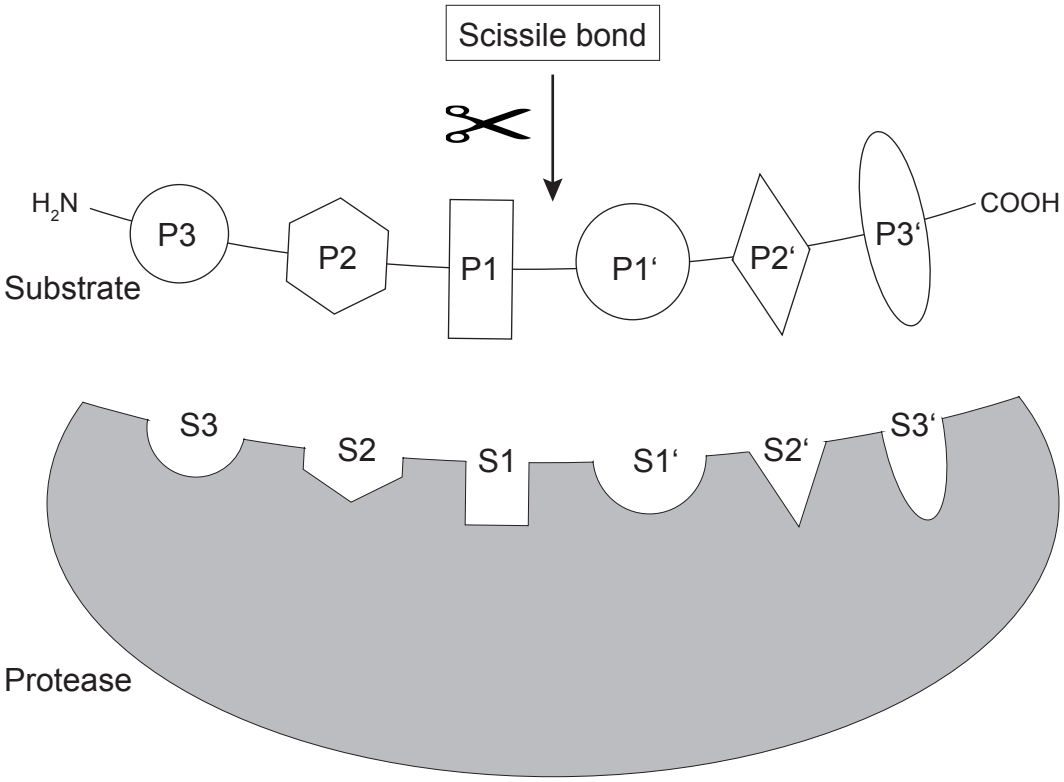


Figure 1.1: **The Schechter and Berger nomenclature.** Schematic representation of the interaction of the amino acids of the substrate next to the cleavage site (scissile bond) (P3-P3') and the protease subsites close to the active site (S3-S3'). Hydrolyzation of the substrate can only occur if the amino acids of the substrate fit well into the subsites of the protease.

tase inhibitor (MNEI), secretory leukocyte protease inhibitor (SLPI) and α2-makroglobulin. α1PI, elafin, MNEI and α2makroglobulin are also inhibitors of PR3. CG is inactivated by antichymotrypsin (ACT), SLPI, MNEI and α1PI (Korkmaz et al., 2008).

There are, however, several mechanisms by which neutrophil serine proteases can escape endogenous inhibition. As described in Owen and Campbell (1999), proteinase inhibitors can be inactivated by oxidative or proteolytic mechanisms and overwhelmed by the high local concentration of proteases at inflammation sites.

1.7.1 Serpins

α1PI, MNEI and ACT belong to the serpin family of *ser*ine *p*rotease *in*hibitors. These high-molecular weight protease inhibitors represent a suicide trap for serine proteases. The so-called reactive center loop (RCL) is an extended loop that is initally recognized as a substrate by the protease. In this way, the rate of inhibition is also dependent on how well the amino acids on the RCL fit to the proteolytic specificity of the protease.

Hydrolysis by serine proteases normally consists of five steps. (i) The formation of an initial non-covalent Michaelis complex is followed by (ii) the Ser195 attack on the scissile bond to form a tetrahedral intermediate. Next, (iii) the peptide bond is cleaved which results in a covalent acyl ester intermediate of the protease and the N-terminal part of the substrate while the C-terminal part of the substrate is released. After a water attack, (iv) a second tetrahedral intermediate is formed, which is followed by (v) separation of protease and cleaved substrate. In the case of the serpins, this process is arrested after (iii). Cleavage of the RCL allows the serpin to undergo a substantial conformational change in which the covalently attached protease is translocated to the distal end of the serpin and the active site of the protease is distorted. In this way, the acyl-enzyme intermediate is kinetically trapped representing the covalently linked serpin-protease complex (Olson and Gettins, 2011).

1.8 Proteinase-activated receptor-2 (PAR-2)

The family of proteinase-activated receptors (PAR) so far includes members 1-4. These receptors are activated after a proteolytic cleavage in their extracellular N-terminal part. This leads to exposure of a previously cryptic sequence that acts as activating ligand (tethered ligand) of the receptor (Steinhoff et al., 2005). Proteases can also cleave downstream of the activating sequence, thereby inactivating the respective receptor. The PAR can be activated by several proteases. Thrombin is classically reported as activator of PAR-1, PAR-3 and PAR-4 (Vu et al., 1991; Ishihara et al., 1997; Kahn et al., 1998) and trypsin can activate PAR-2 (Nystedt et al., 1995). Trypsin can also activate PAR-1, PAR-3 and PAR-4 to a lower extent, but thrombin does not activate PAR-2 (Shpacovitch et al., 2008).

Considering NSP4, PAR-2 appeared to be a very interesting natural substrate target, based on its sequence around the activation site, SKGR-/-SLIG. Like the other three, PAR-2 is expressed in various cells and organs and the functions reported so far reflect this variety (reviewed in Steinhoff et al. (2005)). It was already reported that proteases of neutrophils can activate PAR-2 on epithelium resulting in increased epithelial permeability (Chin et al., 2008). On the other hand, Ramachandran et al. (2011) revealed that none of NE, CG or PR3 can activate G-coupled PAR-2 calcium signaling, but all three can disarm and inactivate PAR-2. Hence, the PAR-2 activating protease in neutrophils could be NSP4.

1.9 NSP4-deficient mouse model

I was planning to generate an NSP4-deficient (NSP4-/-) mouse line. A homologous gene indeed exists in the mouse genome and has been designated Prss57. While cloning the mouse DNA constructs, I learned that Lexicon Pharmaceuticals Inc. (The Woodlands, TX, USA) in cooperation with Genentech Inc. (San Francisco, CA, USA) had already developed an NSP4 deficient mouse model (project PRT354N1) by targeting exons 2 and 3 with a lacZ-neomycin cassette (Figure 2.1). This mouse model was part of a large collection of knockout mouse lines for 472 genes encoding secreted and membrane proteins in mice (Tang et al., 2010). All mouse models

were subjected to a phenotypical screen to uncover alterations in embryonic development, metabolism, the immune system, the nervous system and the cardiovascular system. The results of the screens can be accessed and studied online (*http://mmrrc.mousebiology.org/phenotype/index.php*).

Four to six Prss57-deficient mice on a mixed 129S5/SvEvBrd $\times$ C57BL/6J-Tyr$^{c\text{-}Brd}$ hybrid background were investigated by Lexicon Pharmaceuticals and Genentech. No notable phenotype could be observed other than a marginal decrease in pheripheral blood CD8 and NK cells, a slight increase in B cells and slightly higher levels of tumor necrosis factor α (TNF-α) and monocyte chemotactic protein-1 after intraperitoneal challenge with lipopolysaccharides (LPS).

Expression of NSP4 in their study was investigated by lacZ encoded β-galactosidase and RT-PCR. LacZ activity was reported to be missing in all tissues they had studied (panel size and tissue range not mentioned). Expression of mRNA in wildtype (WT) mice was restricted to thymus and spleen. The cell type transcribing the gene, however, has not been identified. Bone marrow tissue was not analyzed by RT-PCR.

Hence, neutrophil functions and inflammatory models other than LPS stimulation were not specifically explored as NSP4 was not suspected to be expressed in neutrophils.

1.10 Objectives

Despite numerous studies on neutrophils, only NE, CG and PR3 were identified as active serine proteases in these short-lived cells. In my diploma thesis, I identified the fourth neutrophil serine protease, NSP4, which was completely overlooked so far. Nothing was known about this novel serine protease despite its expression in human neutrophils. The aim of this thesis was an in-depth characterization of NSP4. This new knowledge should then facilitate a functional comparison with the other NSPs and help to uncover the physiological functions of NSP4.

At first, the production and purification of recombinantly expressed NSP4 had to be improved. Before, NSP4 was expressed in a stably transfected HEK 293 cell line. Due to the low expression levels and many other proteins in the cell culture supernatant, this previous method was not satisfactory for the production and purification of NSP4. To this end, different expression systems and improvement steps for the expression of NSP4 in the stably transfected HEK 293 line were tested.

So far, I had only tested the presence of NSP4 in different human blood cell populations. For a better classification of the role of NSP4, protein expression in other tissues was studied. Moreover, with regard to NSP4 in neutrophils, a characterization of its biosynthesis, storage, activation and release by neutrophils was required. This basic knowledge helps to narrow down the broad range of potential natural substrates of NSP4 and enables us to assess its role in the correct biological context and in relationship to the long-known other three members, NE, CG and PR3.

To uncover the function of NSP4, it is indispensable to know which substrate sequences are preferentially hydrolyzed and when and where NSP4 is enzymatically active. In order to answer these questions, the proteolytic specificity of NSP4 was analyzed and possible inhibitors were tested. In addition, efforts were undertaken to determine the activity status of NSP4 in biological samples. To investigate the role of NSP4, I also assessed potential natural substrates of NSP4. Since I was interested in analyzing the evolutionary conserved function of NSP4, I also addressed NSP4 in the mouse regarding proteolytic specificity, activity and expression in comparison to human NSP4. Finally, isolated neutrophils of NSP4-/- and WT mice were examined in cell based *in vitro* assays.

2 Materials and Methods

2.1 Materials

2.1.1 Chemicals and consumables

Chemicals were obtained from the companies BD Biosciences (Franklin Lakes, NJ, USA), Bio-Rad Laboratories (Hercules, CA, USA), Biozym (Hessisch Oldendorf, Germany), Life Technologies (Carlsbad, CA, USA), Merck (Darmstadt, Germany), PAA Laboratories GmbH (Pasching, Austria), GE Healthcare (Little Chalfont, Buckinghamshire, UK), Qiagen (Hilden, Deutschland), Roth (Karlsruhe, Germany), and Sigma-Aldrich (St Louis, MO, USA), unless mentioned otherwise. Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (Ipswich, MA, USA). Consumables such as pipette tips, reaction tubes, centrifuge tubes and microwell plates were procured from Kisker (Steinfurt, Germany), Biozym, Eppendorf (Hamburg, Germany), BD Biosciences and Nunc/Thermo Fisher Scientific (Waltham, MA, USA). Cell culture flasks and dishes were from Corning (Corning, NY, USA) and BD Biosciences.

2.1.2 Oligonucleotides

Oligonucleotides were synthesized by Metabion (Martinsried, Germany). Oligonucleotide sequences can be found in the Appendix (7.2).

2.1.3 Plasmids

Plasmid	Construct	Origin
pcDNA5	FRT-V5-H_6-TOPO	Invitrogen, Life Technologies
S-NSP4(human)-pcDNA5	SigIgκ-S-EK-NSP4-H_6	Heike Kittel, Arbeitsgruppe PD Dr. Jenne (AG Jenne)
pTT5	-	NRC Biotechnology Research Institute (Ottawa, Ontario, Canada)
NSP4_precursor(human)-pcDNA3.1	SigNSP4-NSP4_precursor-cMyc-H_6	Heike Kittel, AG Jenne
MOG(human)-pTT5	SigIgκ-MOG-H_6	Heike Kittel, AG Jenne
Lamp2a(human)-pTT5	SigIgκ-Lamp2a-H_6	Heike Kittel, AG Jenne
pFASTBAC1	-	Invitrogen
Sumo3(human)-NSP4(human)-pET28M	Sumo3-EK-NSP4-H_6	Johannes Trambauer, AG Jenne
α1PI(human)-pTT5	SigIgκ-α1PI-H_6	Therese Dau, AG Jenne
SUMOstar-pTT5	SigIgκ-H_6-SUMOstar	Therese Dau, AG Jenne
Ubiquitin(human)-pCR2.1	Ubiquitin	Eurofins MWG

The vector maps can be found in the Appendix (7.1).

2.1.4 Cell lines

Cell line	ATCC no.	Origin
Flp-In^{TM}-293 [S-NSP4]	CRL-1573	Heike Kittel, AG Jenne; Invitrogen
HEK 293 EBNA1	CRL-10852	NRC Biotechnology Research Institute
HT1080	CCL-121	Luisa-Astrid Fratila, AG Sommerhoff
HaCaT		Luisa-Astrid Fratila, AG Sommerhoff

2.1.5 Recombinant and purified proteins

Protein	Origin
α1PI, human	Athens Research & Technology (Athens, GA, USA)
Antichymotrypsin, human	Athens Research & Technology
Antithrombin, human	Athens Research & Technology
Cathepsin G, human	Calbiochem (Merck, Darmstadt, Germany)
C1 inhibitor, human	Athens Research & Technology
Elafin, human	Proteo Biotech AG (Kiel, Germany)
Heparin, porcine	Sigma-Aldrich
MNEI, human	Creative Biomart (Shirley, NY, USA)
Neutrophil elastase, human	Elastin Products (Owensville, MO, USA)
Proteinase 3, human	Athens Research & Technology
SLPI, human	R&D Systems (Minneapolis, MN, USA)
Thrombin, bovine	BioCentrum, Krakow, Poland
Trypsin, bovine	Sigma-Aldrich

2.1.6 Antibodies

Primary antibodies

Antigen	Species	Clone	Provider	Cat. No.
Human NSP4	Rat	2F6	AG Jenne/AG Kremmer, IMI	-
Human α1PI†	Rabbit	polyclonal	Dako	A0012
Human CG	Mouse	19C3	Abcam	ab75572
Ovalbumin	Rabbit	serum	Sigma	C6534
Human haptoglobin	Rabbit	polyclonal	Dako	A0030

†Anti-α1PI antibodies were biotinylated by Bert Utecht, Utecht & Lüdemann GmbH, Klausdorf, Germany.

Secondary antibodies

Antigen	Species	Label	Provider	Cat. No.
Rat IgG+IgM	Goat	Peroxidase	Jackson ImmunoResearch	112-035-068
Mouse IgG+IgM	Goat	Peroxidase	Pierce (Thermo Scientific)	31444
Rabbit IgG	Goat	Peroxidase	Jackson ImmunoResearch	111-035-006

Other

Protein	Label	Provider	Cat. No.
Streptavidin	Peroxidase	Millipore	OR03L

2.1.7 Synthetic substrates

The list of substrates that were used for this thesis can be found in the Appendix (7.4).

2.1.8 Mouse strain

The NSP4 knockout mouse model was established by Lexicon Pharmaceuticals Inc. in a collaboration with Genentech Inc. The mouse model was generated on a mixed 129S5/SvEvBrd $\times$

C57BL/6J-Tyr$^{c\text{-}Brd}$ background. NSP4-deficient mice were purchased as cryo-preserved heterozygous embryos from the Mutant Mouse Regional Resource Center (MMRRC) at the University of California in Davis (UC Davis, Davis CA, USA) (stock no. 032536-UCD, strain name B6;129S5-Prss57tm1Lex/Mmucd).

2.1.9 Laboratory equipment

Autoklave	Varioklav, H+P Labortechnik, Oberschleissheim, Germany
Balances	PM 4800 Delta Range, Mettler-Toledo, Columbus, OH, USA
	2001 MP2, Sartorius, Göttingen, Deutschland
Centrifuges	5417R, Eppendorf
	5417C, Eppendorf
	Rotanta 460R, Hettich, Tuttlingen, Germany
	Rotanta/R, Hettich
Film developer	X-Omat M35, Kodak
FPLC	ÄKTAprime^{TM}, Amersham-Pharmacia
Gel electrophoresis chambers	
and power supplies	Bio-Rad, Amersham-Pharmacia, MPI workshop
Incubation shaker	HT Multitron, Infors, Bottmingen, Switzerland
Incubator	BBD6220, Heraeus, Thermo Scientific, Waltham, MA, USA
Incubator	Jouan, Germany
Icemachine	Ziegra, Isernhagen, Germany
Laminar flow	LaminAir HB 2472S, Heraeus, Hanau, Germany
Magnetic stirrer	MR3003, Heidolph, Kelheim, Germany
	KMO2 basic, IKA, Staufen, Germany
Microplate reader	FLUOStar Optima, BMG Labtech, Offenburg, Germany
Microscope	Leica DM IL, Leica Microsystems, Wetzlar, Germany
pH-meter	Inolab pH Level 1, WTW, Weilheim, Germany
PCR cycler	T3 Thermocycler, Biometra, Göttingen, Germany
Semidry blotter	Biometra, Göttingen, Germany
Shakers	130 basic, IKA KS
	260 basic, IKA KS
Spectrophotometer	ND-1000, peqlab, Erlangen, Germany
	Biophotometer, Eppendorf
Stirring cell	MPI workshop
Thermoblock	Thermomixer 5436, Eppendorf
Ultrasonic bath	Sonorex digital 10P, Bandelin, Berlin, Germany
Water bath incubator	MA6, Lauda, Lauda-Königshofen, Germany
Water preparation	Milli Q Advantage, Millipore

2.2 Molecular biological methods

2.2.1 Preparation of competent bacteria with $CaCl_2$

A single colony of *E. coli* DH5α was inoculated into 50 ml LB medium (Luria-Bertani-Broth) without antibiotics and grown in an Erlenmeyer flask at 37 °C with shaking (250 rpm) overnight. Then 4 ml of the preculture were added to 400 ml fresh LB medium. The cells were grown until the OD_{600nm} reached 0.375-0.6. The culture was kept on ice for 5-10 min in pre-cooled 50 ml tubes and pelleted by centrifugation at 3000 rpm and 4 °C for 7 min without brake. The cell pellets from 50 ml culture volume were resuspended in 10 ml ice-cold $CaCl_2$ solution and kept on ice for 30 min. For the remaining steps the cells were kept on ice and all flasks, pipette tips and tubes were pre-cooled at 4 °C. After centrifugation at 3000 rpm and 4 °C for 5 min, the pellets were taken in 2 ml ice-cold $CaCl_2$ solution and aliquots of 250 µl were stored at -80 °C.

$CaCl_2$-Solution	60 mM	$CaCl_2$
	15% (w/v)	glycerol
	10 mM	PIPES [piperazine-N,N'-bis(2-ethanesulfonic acid)]
		pH 7
LB medium	1% (w/v)	bacto tryptone
	0.5% (w/v)	yeast extract
	1% (w/v)	NaCl
		pH 7.5
		autoclaved
	(100 µg/ml	ampicillin)
LB agar (plates)	1 ×	LB medium
	1.5% (w/v)	bacto agar
		autoclaved

2.2.2 Transformation of competent bacteria by heat shock

Competent *E. coli* DH5α were thawed on ice for 10 min. I used 100 µl bacterial cells and 3-10 µl ligation preparation or 10 ng plasmid DNA in 1-10 µl per transformation. The DNA was added to the cells and left on ice for 30 min. This should allow the DNA to attach to the cell

walls. The heat shock was performed at 42 °C for 30 s. This permeabilizes the cell membrane for a short duration and the plasmid DNA can enter the cell. Directly after the heat shock, 900 µl pre-warmed LB medium without antibiotics was added and the cells were grown for 60 min at 37 °C and 250 rpm. Both the pTT5 and the pFASTBAC1 vector included the beta-lactamase gene which renders the bacterial clones resistant to the antibiotic ampicillin. I plated 100-200 µl of the cell suspension onto LB agar plates with 100 µg/ml ampicillin and they were incubated at 37 °C overnight.

2.2.3 Plasmid DNA purification from *E. coli*

For purification of plasmid DNA from 2-5 ml overnight culture of *E. coli*, the QIAprep Spin Miniprep Kit (Qiagen) and a microcentrifuge were used. The bacteria were lysed under alkaline conditions followed by adsorption of plasmid DNA onto silica in high salt. The EndoFree Plasmid Maxi Kit (Qiagen) was employed for purification of plasmid DNA from larger cultures (200-300 ml). Here, the bacteria were disrupted by alkaline lysis and the plasmid DNA was bound to QIAGEN anion-exchange resin. Both kits were used according to the manufacturer's instructions. The EndoFree kit includes an endotoxin removal buffer which prevents endotoxins (lipopolysaccharides, LPS) from binding to the anion-exchange resin. Purified plasmid DNA was then used for transfection of mammalian cells (HEK 293E) in cell culture. Endotoxins were removed because increased levels can lead to reduced transfection efficiencies.

2.2.4 Polymerase chain reaction (PCR)

The PCR enables the amplification of short DNA segments (50 bp - 10 kbp) from larger DNA, e.g. plasmids or genomic DNA. One forward and one reverse primer are required, which flank the desired DNA segment. The DNA polymerase can then add and link nucleotides to the primers, thereby extending the amplicon. The PCR is done in cycles, beginning with a denaturation step in which the strands of the DNA double helix separate, followed by an annealing step in which the primers bind to the ends of the desired DNA segment. The cycle is completed by the extension step in which the DNA polymerase extends the primers. Two different polymerases

were employed. The HotStarTaq DNA polymerase from Qiagen was used for mouse genotyping analyses and RT-PCR. The Phusion High-Fidelity DNA polymerase from Finnzymes (New England Biolabs) was applied for cloning of expression constructs. The latter one is a pyrococcus-like enzyme fused with a processivity enhancing domain increasing fidelity and speed.

HotStarTaq PCR reaction (25 µl)		
	1 µl	DNA template (mouse tail DNA, ∼100-200 ng)
	0.2 mM	dNTP-mix
	1 ×	PCR buffer (Qiagen)
	0.4 µM	forward primer (10 pmol absolute)
	0.4 µM	reverse primer (10 pmol absolute)
	0.2 µl	HotStarTaq DNA polymerase

HotStarTaq PCR programme "Touch down"			
15 min	95 °C	Denaturation	
10 s	95 °C	Denaturation	3 cycles
30 s	60 °C	Annealing	
30 s	72 °C	Extension	
10 s	95 °C	Denaturation	3 cycles
30 s	57 °C	Annealing	
30 s	72 °C	Extension	
10 s	95 °C	Denaturation	21 cycles
30 s	54 °C	Annealing	
30 s	72 °C	Extension	
10 min	72 °C	Annealing	

Phusion PCR reaction (25 µl)		
	2-10 ng	plasmid DNA
	or 50-100 ng	genomic DNA
	0.2 mM	dNTP-mix
	1 ×	HF Buffer (Finnzymes)
	0.4 µM	forward primer (10 pmol absolute)
	0.4 µM	reverse primer (10 pmol absolute)
	0.2 µl	Phusion DNA polymerase

Phusion PCR programme "Touch down"				
	30 s	98 °C	Denaturation	
	10 s	98 °C	Denaturation	3 cycles
	30 s	*	Annealing	
	30 s	72 °C	Extension	
	10 s	98 °C	Denaturation	3 cycles
	30 s	**	Annealing	
	30 s	72 °C	Extension	
	10 s	98 °C	Denaturation	18 cycles
	30 s	***	Annealing	
	30 s	72 °C	Extension	
	10 min	72 °C	Annealing	

* primer specific annealing temperature + 6 °C
** primer specific annealing temperature + 3 °C
*** primer specific annealing temperature

2.2.5 Formation of oligoduplexes

For the exchange of short sequences, two complimentary oligonucleotides were ordered and mixed at equimolar concentrations (e.g. 10 µM). The incubation at 95 °C for 5 min was followed by 2 min at room temperature and 5 min on ice.

2.2.6 Restriction digests

Restriction endonucleases are enzymes which cleave DNA, mostly at a specific recognition sequence. One unit is defined as the amount of restriction enzyme that completely digests 1 µg of substrate DNA (often Lambda DNA) in one hour at the optimal temperature. Restriction enzymes were purchased at New England Biolabs or SibEnzyme (*Abs*I; Novosibirsk, Russia) and used according to the manufacturer's instructions with the provided buffers. The substrate DNA (500 ng - 5 µg) was digested in 15-50 µl volume for 60 min.

2.2.7 Dephosphorylation of vector DNA

To avoid self-ligation of the vector plasmid without insert, the vectors were dephosphorylated after restriction digest. After restriction digest, 2-5 µg of the vector DNA was incubated with 0.5 µl Calf Intestinal Alkaline Phosphatase (New England Biolabs) in a volume of 30-50 µl at 37 °C.

2.2.8 Agarose gel electrophoresis

As DNA fragments are negatively charged, they can be separated in a gel matrix in an electric field on the basis of their size. Agarose gels were prepared in TAE buffer with SYBR Safe (Life Technologies) 1:10 000 or 1 µg/ml ethidium bromide for visualization of DNA under UV (λ= 254-366 nm). Depending on the size of expected DNA fragments, the percentage of agarose in the gels ranged from 0.7% to 3.5% (w/v). The samples were loaded into the gel slots after addition of loading buffer and run in TAE buffer at 90 V for gels of 40 ml. The size of the migrating DNA fragments was determined by comparison with molecular weight standards (Gene Ruler 1 kb, Fermentas, Thermo Fisher Scientific and 50 bp ladder, Roth).

TAE buffer (50 ×)	242 g	Tris
	100 ml	0.5 M EDTA pH 8.0
	57.1 ml	acetic acid
		H_2O to 1 L
DNA loading buffer (10 ×)	0.25% (w/v)	bromphenol blue
	0.25% (w/v)	xylenecyanol
	50% (v/v)	glycerol

2.2.9 DNA purification

After restriction digest or PCR, DNA fragments were purified by gel extraction (QIAquick Gel Extraction Kit, Qiagen) or with the QIAquick PCR Purification Kit (Qiagen) in accordance with the manufacturer's instructions.

2.2.10 Determination of DNA concentration

At the wavelength of 260 nm, the extinction coefficient for double stranded DNA in water is 0.020 $(\mu g/ml)^{-1}$ cm^{-1}. With this, the concentration of DNA was measured using a spectrophotometer and the Beer-Lambert law (Absorption = (extinction coefficient) $\times$ (length of solution) $\times$ (concentration of solution)).

2.2.11 Ligation of DNA

T4 DNA Ligase or the Quick Ligation Kit (both from New England Biolabs) were used for the ligation of insert and vector DNA. The manufacturer's instructions were followed. I used 50 ng vector DNA and a molar ration of 1 to 10 of vector and insert in a total volume of 20 µl. Ligation reactions employing T4 DNA Ligase were carried out for 2-4 h at room temperature. The Quick Ligation Kit enabled ligation within 5-15 min at room temperature.

2.2.12 Sequence analysis of DNA

DNA samples were sent to Eurofins MWG GmbH (Ebersberg, Germany) for sequence analysis.

2.2.13 Expression constructs

Expression in mammalian cells (HEK 293E)

Name	Construct	Insert	Vector	Primer	AT [°C]
S-NSP4(human)-pTT5	SigIgκ-S-EK-NSP4-H_6	*Nhe*I/*Pme*I digest of S-EK-NSP4-pcDNA5	*Nhe*I/*Nae*I digest of pTT5	-	-
SigPR3-NSP4(human)-pTT5	SigPR3-EK-NSP4-H_6	Oligoduplex coding for PR3 signal peptide	*Nhe*I/*Pml*I digest of S-NSP4-pTT5	DJ3305/ DJ3306	-
NSP4_precursor(human)-pTT5	SigNSP4-AQ-NSP4-cMyc-H_6	*Eco*RI/*Bsr*GI digest of NSP4 precursor-pcDNA3.1	*Eco*RI/*Bsr*GI digest of S-NSP4-pTT5	-	-
SUMOstar-NSP4(human)-pTT5	SigIgκ-SUMOstar-EK-NSP4-H_6	PCR with SUMOstar vector LIC7121; *Kpn*I/*Dra*III	*Kpn*I/*Dra*III digest of S-NSP4-pTT5	DJ3324/ DJ3325	50
MOG(human)-NSP4(human)-pTT5	SigIgκ-MOG-EK-NSP4-H6	*Nhe*I/*Hinc*II digest of MOG-pTT5	*Nhe*I/*Pml*I digest of S-NSP4-pTT5	-	-
Lamp2a(human)-NSP4(human)-pTT5	SigIgκ-Lamp2a-EK-NSP4-H_6	PCR with Lamp2a in pTT5; *Nhe*I/*Pml*I	*Nhe*I/*Pml*I digest of S-NSP4-pTT5	DJ3312/ DJ3314	60
NGS-NSP4(human)-pTT5	SigIgκ-NGS-EK-NSP4-H_6	Oligoduplex coding for NGS-peptide	*Kpn*I/*Pml*I digest of S-NSP4-pTT5	DJ3385/ DJ3386	-
α1PI*-pTT5	SigIgκ-α1PI(RCL=IKPR)-H_6	Oligoduplex coding for IKPR	*Abs*I/*Acc*I digest of α1PI(human)-pTT5	DJ3559/ DJ3560	-
mNSP4-pTT5	SigIgκ-H_6-EK-mNSP4	PCR with cDNA clone (imaGenes, Berlin, Germany); *Pst*I/*Age*I	*Sbf*I/*Age*I digest of H_6-SUMOstar-pTT5	DJ3352/ DJ3353	55
S-mNSP4-pTT5	SigIgκ-S-EK-mNSP4-H_6	PCR with mNSP4-pTT5; *Pml*I/*Age*I	*Pml*I/*Age*I digest of S-NSP4-pTT5	DJ3368/ DJ3369	55

AT = Annealing temperature; Sig = Signal peptide; S = S-tag peptide; H_6 = 6× His tag; EK = enterokinase cleavage site; AQ = natural propeptide of NSP4, Ala-Glu; cMyc = cMyc tag; MOG = myelin oligodendrocyte glycoprotein; Lamp = lysosomal-associated membrane protein; NGS = peptide sequence with glycosylation site, SRMHRNGSH; RCL = reactive center loop of a serpin; IKPR = RCL sequence of α1PI*; mNSP4 = mouse NSP4

Expression in insect cells (High Five, Sf9)

Name	Construct	Insert	Vector
SigGP67-pFASTBAC1	SigGP67	Gene synthesis of SigGP67 (Eurofins, Ebersberg); *Eco*RI/*Bgl*III	*Eco*RI/*Bam*HI digest of pFASTBAC1
Sumo3(human)-NSP4(human)-pFASTBAC1	SigGP67-Sumo3-EK-NSP4-H_6	*Nde*I digest of Sumo3-NSP4-pET28M	*Nde*I digest of SigGP67-pFASTBAC1
Ubiquitin(human)-pFASTBAC1	SigGP67-Ubiquitin	Gene synthesis of Ubiquitin (Eurofins); *Nde*I/*Stu*I	*Nde*I/*Stu*I digest of SigGP67-pFASTBAC1
Ubiquitin(human)-NSP4(human)-pFASTBAC1	SigGP67-Ubiquitin-EK-NSP4-H_6	*Pml*I/*Bam*HI digest of S-NSP4 pTT5	*Pml*I/*Bam*HI digest of Ubiquitin pFASTBAC1

AT = Annealing temperature; Sig = Signal peptide; GP67 = acidic glycoprotein of AcNPV baculovirus; H_6 = 6× His tag; EK = enterokinase cleavage site; S = S-tag peptide;

2.3 Recombinant protein expression

All cell culture methods described were carried out in a sterile laminar flow hood. Cell culture media and other solutions were autoclaved or sterile filtered using a filter with pore size 0.22 µm.

2.3.1 Expression of S-NSP4 in Flp-In™-293

The Flp-In™ system from Invitrogen (Life Technologies) enables expression of a protein of interest after stable integration of the gene of interest. Integration is achieved employing Flp-recombinase and FRT sites in host cell and expression vector. I was working with Flp-In™-293 cells that were already stably transfected with S-NSP4 in pcDNA 5 (Heike Kittel, AG Jenne).

Flp-In™-293 cell culture in cell stacks

For protein expression, Flp-In™-293 cells were grown in adherent cell culture in 5-level cell stacks (Invitrogen) at 37 °C and 5% CO_2 in a humidified atmosphere. The supernatant was harvested after 10-14 days. The cells were then detached from the vessel surface with trypsin preceded by a wash in 1 × phosphate buffered saline (PBS). One fifth of the cells was then again seeded in the cell stacks in 500 ml fresh medium.

Flp-In medium	1 ×	DMEM
	10% (v/v)	FCS
	75 µg/ml	hygromycin B (Invitrogen)
PBS	137 mM	NaCl
	2.7 mM	KCl
	10 mM	$Na_2HPO_4 \times H_2O$
	1.8 mM	KH_2PO_4
		pH 7.4

Flp-In™-293 single cell culture

The cells were cultured as single cell clones to enable a comparison of expression on this level. To ameliorate the growth of single cells, I used a 1:1 mixture of fresh Flp-In medium and conditioned Flp-In medium. Conditioned medium is medium in which the cells have been growing in for one passage and that was filtered through 0.22 µm. The cells were diluted to 7 cells/ml and 100 µl were seeded per well into 96-well cell culture plates. This should result in a statistic number of 0.7 cells per well. The growth at 37 °C and 5% CO_2 in a humidified atmosphere was monitored daily under a microscope. When single cell colonies of similar sizes were formed, the culture medium was refreshed and harvested after 9 days. The supernatants were then assayed for NSP4 by Western blot analysis. Well expressing clones were chosen and used for adaptation to low-serum and suspension growth.

Adaptation to low serum and suspension growth

The cells were adapted stepwise from Flp-In medium to low-serum medium with each split. After reaching confluency, the cells were split 1 to 10 and taken in medium with 7.5%, 5%, 2.5%, 1% and then 0.3% serum. The cells cultured in medium with 0.3% serum were transferred into Erlenmeyer flasks for suspension culture and grown in the 37 °C incubation shaker containing a humidified atmosphere of 8% CO_2 in air at 110 rpm. For suspension growth, the non-ionic surfactant Pluronic® F-68 (Invitrogen) was added to protect the cells from shear stress. I used 125 ml, 250 ml, 1 L and 3 L Erlenmeyer shaker flasks for cell cultures of 10-50 ml, 50-100 ml, 150-400 ml and 600-1200 ml, respectively. Supernatants were harvested after 7-9 days.

Flp-In low-serum medium	1 ×	Freestyle™ 293 Expression Medium (Invitrogen)
	0.3% (v/v)	FCS
	50 µg/ml	hygromycin B (Invitrogen)
	(0.1% (v/v)	Pluronic® F-68)

Freezing and thawing of Flp-In™-293

Flp-In™-293 were frozen at a concentration of 1×10^7 cells/ml in freezing medium. For this the cells were pelleted for 6 min at 330 × g and 4 °C, taken into freezing medium and aliquoted

into cryotubes. The cryotubes were then put into an ice-cold freezing container (Nalgene, Thermo Fisher Scientific) filled with isopropanol and at -80 °C. The freezing container provided the cooling rate of 1 °C/min. For long time storage, the cells were transferred into liquid nitrogen. Flp-In^{TM}-293 aliquots were carefully thawed in a 37 °C water bath. Just before complete thawing, the cells were taken into 10 ml pre-warmed low-serum Flp-In medium and centrifuged for 6 min at 330 × g. The supernatant was removed to reduce the DMSO concentration and the cells were taken into 30 ml pre-warmed low-serum Flp-In medium and grown in the incubation shaker at 37 °C in a humidified atmosphere with 8% CO_2 and 110 rpm.

Freezing medium	1 ×	Freestyle^{TM} 293 Expression medium (Invitrogen)
	1%	FCS
	10% (v/v)	DMSO

2.3.2 Expression of α1PI*, S-mNSP4 and all NSP4 constructs in pTT5 in HEK 293E

All constructs that were cloned into the pTT5-plasmid and listed in section 2.2.13 (Expression in mammalian cells) were expressed in the HEK 293 EBNA 1 (HEK 293E) cell line. The expression of S-NSP4, NGS-NSP4 and S-mNSP4 could be enhanced by using additional filler DNA for transfection of HEK 293E (please see below).

HEK 293E is a modified human embryonic kidney cell line stably transfected with a plasmid encoding the nuclear EBV-antigen 1 (EBNA-1). The plasmid also carries a geneticin resistance marker (G418) for maintenance of the plasmid within the cell. EBNA-1 transactivates the episomal replication of plasmids carrying the EBV specific replication of origin 'oriP' (Reisman and Sugden, 1986; Durocher et al., 2002), e.g. pTT5. Cells were cultured in suspension in 293 EBNA medium in Erlenmeyer flasks in the incubation shaker at 37 °C in a humidified atmosphere with 8% CO_2 and 110 rpm.

293 EBNA medium	1 ×	Freestyle^{TM} 293 Expression Medium (Invitrogen)
	0.1% (v/v)	Pluronic® F-68
	25 µg/ml	Geneticin® G418

Transfection of HEK 293E

Prior to transfection of HEK 293E with the expression constructs, cells were brought to a density of 1×10^6 cells/ml. Polyethyleneimine (PEI)/DNA complexes were prepared by adding PEI to DNA, both prediluted in Optipro serum free medium (Invitrogen), incubated at room temperature for 15 min and added to the cells. I used 2 µg PEI and 1 µg DNA for 1 ml of cell culture suspension. After 1 day, Bacto TC Lactalbumin Hydrolysate (BD Biosciences), which is the enzymatically hydrolyzed portion of milk whey, was added as an amino acid supplement to 0.5% final concentration. Supernatants were harvested after 4 days.

The use of filler DNA for improved transfection and expression of S-NSP4, NGS-NSP4 and S-mNSP4 in HEK 293E

Transfection was carried out as described above, but I mixed, or 'diluted' the plasmid of interest in ratios of 1:10 to 1:1000 with non-specific filler DNA. As filler DNA I used a bacterial plasmid without mammalian expression elements (Ubiquitin-pCR2.1).

2.4 Cell biological methods

2.4.1 Determination of cell count and viability

The number of viable cells/ml was determined in a hemacytometer (Neubauer improved) after mixing trypan blue solution and cell suspension in a 1:1 ratio. Trypan blue can enter the perforated membranes of dead cells and allows for discrimination from living cells. The number of cells in each of the four large quadrants was counted and the average of this was multiplied by the dilution factor. As the volume over one of the large quadrants is 100 nl, the result was multiplied by 10^4 to get the number of cells/ml.

Trypan blue solution	0.1% (w/v)	trypan blue
	1 ×	PBS pH 7.4

2.4.2 Analysis of human PMN

Isolation of human PMN from blood

Fresh blood from volunteering healthy donors or from a PLS patient was taken into 9 ml EDTA tubes to prevent blood clotting. I mixed 18 ml EDTA blood with 17 ml PBS, transferred it into a 50 ml tube and carefully underlaid it with 10 ml Pancoll human (Pan Biotech, Aidenbach, Germany). The tubes were centrifuged for 20 min at 500 $\times$ g at 20 °C without brake. After successful gradient centrifugation, erythrocytes and polymorphonuclear cells (PMN) can be found at the bottom of the tube whereas plasma, thrombocytes and peripheral blood mononuclear cells (PBMC, lymphocytes and monocytes) are still on top of the Pancoll layer. Plasma, thrombocytes, PBMC and Pancoll were removed and the rest was filled to 40 ml with PBS and mixed with 8 ml of 6% dextran in PBS. After 45 min dextran sedimentation of erythrocytes at room temperature, the PMN in the supernatant were taken and pelleted for 5 min at 500 $\times$ g. Residual erythrocytes were lysed in 5 ml 0.2% NaCl. After 1 min, 5 ml of 1.6% NaCl was added followed by 20 ml PBS after another min. The cells were centrifuged again and taken in PBS or RPMI 1640 (Gibco, Life Technologies). Neutrophil purity was determined by forward and sideward scatter flow cytometry and was routinely found $\sim$95%.

Isolation of PMN from a patient with Papillon-Lefèvre syndrome (PLS)

The PLS patient was a 18-year-old man that presented with mild skin findings. The patient carried a homozygous 854 C > T nucleotide exchange in the cathepsin C gene leading to a substitution of the normal proline residue 285 by leucine (P285L mutation). This mutation has already been reported and characterized and was associated with an almost complete loss of protein function (Noack et al., 2008). PMN were isolated from heparinized peripheral blood as described above.

Preparation of total cell lysates

For Western blot analysis, purified PMN were totally lysed. Cells were taken in cold RIPA buffer at 3 $\times$ 10^4 cells/µl and left on ice for 15 min. In thin-walled PCR tubes (0.5 ml, New England

Biolabs), cells were incubated 4 × 3 min in the ultrasonic bath, with intermittent incubations on ice for 5 min. Cell debris and DNA were pelleted by centrifugation for 10 min at 20 000 × g and 4 °C. For subsequent measurements of enzymatic activity of serine proteases, RIPA buffer without PMSF was used.

RIPA buffer	50 mM	Tris
	150 mM	NaCl
	0.5 M	EDTA
	(1 mM	PMSF)
	0.5% (w/v)	deoxycholic acid
	0.1% (w/v)	SDS
	0.5% (v/v)	Nonidet P-40
		pH 8.0

Degranulation of PMN

To analyze the granule content of neutrophils, PMN were activated with cytochalasin B (cytB) and phorbol-12-myristate-13-acetate (PMA). To this end, isolated PMN were taken in RPMI 1640 at 1.7 × 10^5 cells/µl and incubated with cytB (5 µg/ml) at 37 °C for 5 min. PMA was then added to a final concentration of 200 ng/ml. CytB and PMA were both solved in DMSO, therefore, DMSO alone was added to the control sample. After another 30 min at 37 °C, samples were centrifuged for 10 min at 500 × g and for 5 min at 20 000 × g. Resulting supernatants were analyzed by immunoblotting.

Detection of NSP4 in subcellular fractions of human neutrophils

The fractionation of human neutrophils was done in the laboratory of Prof. Dr. Niels Borregaard (The Phagocyte Laboratory, University of Copenhagen, Copenhagen, Denmark). Samples of 1-5 µl of each of the 5 fractions representing azurophil granules, specific granules, gelatinase granules, secretory vesicles with plasma membranes or cytoplasm were assayed by Western blotting with rat anti-NSP4 mAbs. The bands appearing after immunoblotting were quantitated in Image Lab 4.1 (Bio-Rad).

2.4.3 Calcium signaling in HaCaT and HT1080

Measurements of calcium signaling were done together with Luisa-Astrid Fratila in the laboratory of Prof. Dr. Christian Sommerhoff, Institut für Laboratoriumsmedizin, LMU, München. HaCaT (immortalized human keratinocytes) or HT1080 (fibrosarcoma cell line) were cultured in 24-well tissue culture plates until they were confluent. DMEM with 10% FCS was exchanged to HEPES buffer with 1 mM calcium chloride (0.5 ml per well). The calcium indicator Fluo-4 AM (Molecular Probes, Life Technologies) was added to a final concentration of 4 µM and Pluronic® F-127 (Life technologies) was added to 0.04%. Pluronic® F-127 is a non-ionic surfactant that has been found to facilitate the solubilization and loading of water-insoluble dyes into cells. Fluo-4AM features an ester group which enables crossing of the cell membrane and esterolysis in the cytosol. Although the charged dye molecules leak out of cells far more slowly than the uncharged precursor molecules, anion transporters promote the leakage of the fluorescent dye over time. Probenecid was used to inhibit anion transporters and block the efflux of intracellular Fluo-4. Intracellular calcium bound to Fluo-4 increases the fluorescence 100-fold. The cells were left in the dark incubator for 1 h and washed twice with HEPES buffer containing calcium chloride. For baseline measurements, the cells were first measured without any stimulators and then with stimulators at λEx = 494 nm and λEm = 516 nm. Negative controls were treated with HEPES buffer plus calcium alone or containing proNSP4 (S-NSP4 before EK-digest). Positive controls included bovine trypsin and thrombin.

HEPES buffer with Calcium	10 mM	HEPES
	12.4 mM	$NaHCO_3$
	137 mM	NaCl
	2.7 mM	KCl
	5 mM	D-Glucose × H_2O
	1 mM	$CaCl_2$ (H_2O)
	4 mM	Probenecid, freshly added

Desensitization of PAR-1 and PAR-2

The cells were pretreated with agonist peptides for PAR-1 (TFLLR-amide, H5848) and PAR-2 (SLIGKV-amide, H4624) at final concentrations of 200 µM and 50 µM, respectively. Both were

purchased from Bachem (Bubendorf, Switzerland).

2.5 Protein analysis

2.5.1 Purification of recombinant proteins

After centrifugation at 500 × g for 15 min, the cell culture supernatant was filtered through a 0.22 µm membrane (Millipore). Employing a stirring chamber under nitrogen pressure (2.5 × 10^5 Pa) and an ultrafiltration membrane with 10 kDa cut-off, the supernatant was concentrated fivefold, and then dialyzed against binding buffer at 4 °C. Nickel-affinity chromatography was used to purify 6 × His-tagged S-NSP4, NSP4 precursor, NGS-NSP4, S-mNSP4 or α1PI* after expression in Flp-In^{TM}-293 cells (S-NSP4, see 2.3.1) or HEK 293E cells (S-NSP4, NSP4 precursor, NGS-NSP4, S-mNSP4 and α1PI*, see 2.3.2). The protein solution was applied to a HisTrap HP column (GE Healthcare) previously equilibrated in binding buffer. The column was washed with binding buffer, and bound proteins were eluted applying a linear imidazole gradient from 50 mM to 1 M imidazole in binding buffer. Fractions were collected and analyzed by SDS PAGE and silver staining. The expected size is 31 kDa. The fractions that contained the protein of interest were pooled and concentrated.

Binding buffer	20 mM	Na_2HPO_4
	500 mM	NaCl
	50 mM	imidazole
		pH 7.5 (adjusted using HCl)

2.5.2 Processing of S-, NGS-NSP4 and S-mNSP4 by enterokinase

Purified proteins featuring an enterokinase cleavage site were concentrated about 20-fold by ultrafiltration using Amicon Ultra-4 centrifugal filter units with a cut-off of 10 kDa. With the same centrifugal devices, the residual buffer was exchanged 3 times to EK-buffer. $CaCl_2$ was added to a final concentration of 2 mM. The amino-terminal extension, the S-tag peptide or the NGS-peptide and the enterokinase (EK) recognition sequence was cleaved off by Tag- Off High

Activity Enterokinase (Novagen, Merck) to generate mature 6 × His-tagged NSP4. EK cleavage was accomplished by 1 U EK per 50 µg NSP4 for 16 h at room temperature. Conversion of NSP4 was verified by SDS PAGE and subsequent silver staining. There is a molecular weight shift of ∼3 kDa after removal of these synthetic propeptides.

EK-buffer	20 mM	Tris-HCl
	500 mM	NaCl
	(2 mM	$CaCl_2$)
		pH 7.4

2.5.3 Conversion of NSP4 precursor by dipeptidyl peptidase I

Dipeptidyl peptidase I (DPPI, cathepsin C) was first activated with 1 mM dithioerythritol (DTE) at 37 °C in DPPI buffer for 60 min. After purification, NSP4 precursor was incubated with 0.5 U/mg activated DPPI in DPPI buffer for 4 h at room temperature. The sample was analyzed by Edman sequencing before and after incubation with DPPI.

DPPI-buffer	50 mM	CH_3COONa
	100 mM	NaCl
		pH 5.5

2.5.4 Determination of protein concentration

Bicinchoninic acid (BCA) assay

Protein concentration was determined with BCA using the Uptima MicroBC Assay kit (Interchim, Montluçon Cedex, France) according to the manufacturer's instructions.

Spectrophotometric determination of protein concentration

Protein concentration was determined spectrophotometrically using the molar absorption coefficient as described by Grimsley and Pace (2004). For this, I measured the absorbance at 280 nm and applied the following equation: $c = (A_{280} \times MW) / (\epsilon_{280} \times l)$ (c, concentration; A,

absorption; MW, molecular weight; ϵ, extinction coefficient; l, length of solution). Extinction coefficients of each protein were determined based on the amino acid sequence employing the programme ProtParam (http://us.expasy.org/tools/ protparam.htm).

2.5.5 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE)

SDS PAGE enables separation of proteins on the basis of their size. Gel, sample and running buffer contained the anionic detergent SDS which binds to proteins and thereby unfolds their secondary structure. SDS treatment also confers a negative charge to all proteins in accordance with the length of their polypeptide chain. This enables a discrimination of proteins only by size and irrespective of their inherent charges. β-mercaptoethanol was added to the sample buffer for reduction of disulfide bridges. Prior to loading, protein samples were taken in sample buffer and incubated at 95 °C for 5 min. I used a discontinuous gel system that included a stacking gel and a separating gel. Low acrylamidic concentration and slightly acidic pH in the stacking gel permitted the formation of thin, sharply defined bands before entering the separating gel. The higher polyacrylamide content and the more basic pH in the separating matrix enabled the separation according to protein size. The narrower pores in the separation gel allowed smaller proteins to travel more rapidly through the gel matrix. For comparison of protein sizes, molecular weight standards were used (Prestained protein marker broad range, New England Biolabs; Roti-Mark prestained, Roth).

8 SDS polyacrylamide gels (15%)

Stacking gel	**Separating gel**	
13.6 ml	9.4 ml	H_2O
3.32 ml	20 ml	30% Acrylamid
2.56 ml		1 M Tris HCl, pH 6.8
	10 ml	1,5 M Tris HCl, pH 8.8
100 µl	200 µl	20% SDS
200 µl	400 µl	10% APS
20 µl		bromphenolblue
20 µl	16µl	TEMED

Running buffer (10 ×)	250 mM	Tris HCl
	1.92 M	glycine
	1% (w/v)	SDS
Sample buffer (4 ×), reducing	200 mM	Tris HCl, pH 6.8
	40% (v/v)	glycerol
	10% (w/v)	SDS
	30% (w/v)	β-mercaptoethanol
	0,2% (w/v)	bromphenolblue

2.5.6 Protein detection

Coomassie Blue staining

After SDS PAGE, the gel was incubated in Coomassie Blue staining solution for at least 30 min and then in Coomassie Blue destaining solution until the protein bands were visualized (1-3 hours).

Coomassie Blue staining solution	0.25% (w/v)	Coomassie Brilliant Blue
	45% (v/v)	methanol, p.a.
	10%	acetic acid
Coomassie Blue destaining solution	45% (v/v)	methanol, p.a.
	10%	acetic acid

Silver staining

After SDS PAGE, the gel was incubated in fixation solution for 60 min on a shaker. The gel was washed 3 × 20 s in dH_2O and sensitized in thiosulfate solution for 1 min. After rinsing with water, the gel was submerged into silver nitrate solution, incubated for 20 min and washed again in water. In developer solution, the protein bands became visible and the reaction was stopped in 50% methanol and 12% acetic acid after rinsing in water. The final fixation was achieved in 50% methanol.

Fixation solution	50% (v/v)	methanol
	12% (v/v)	acetic acid
	0,5 ml/l	37% formaldehyde
Thiosulfate solution	0.2 g/l	$Na_2S_2O_3 \times 5\ H_2O$
Silver nitrate solution	2 g/l	$AgNO_3$
	750 µl/l	37% formaldehyde
Developer solution	60 g/l	Na_2CO_3
	4 mg/l	$Na_2S_2O_3 \times 5\ H_2O$
	0.5 ml/l	37% formaldehyde

2.5.7 Edman sequencing

Ten micrograms of NSP4 precursor before and after incubation with DPPI was separated by reducing SDS PAGE and transferred in Edman transfer buffer to a polyvinylidene fluoride (PVDF) membrane. The membrane was stained with Coomassie Blue and the bands representing NSP4 were analyzed by N-terminal Edman sequencing (Reinhard Mentele, MPI Biochemistry).

Edman transfer buffer	90 mM	Tris
	90 mM	boric acid
	1 mM	EDTA
	10% (v/v)	methanol, p.a.

2.5.8 Measurement of enzymatic activity

2.5.9 Proteomic identification of protease cleavage sites (PICS)

For the initial specificity profiling, active recombinant NSP4 was sent to Dr. Oliver Schilling (University of Freiburg, Freiburg, Germany). As described in Schilling and Overall (2008), the cleavage sites of NSP4 were determined in peptide libraries from *E. coli*.

Thiobenzyl ester (Sbzl) substrates

Enzymatic activity of NE, PR3 and CG (100 nM) was determined using thiobenzyl ester substrates and Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoic acid) or DTNB). The free thiobenzyl ester group can react with DTNB which results in yellow 2-nitro-5-thiobenzoate ions. The rate of hydrolysis was read out at 405 nm. I used Boc-Ala-Pro-Nva-4-chloro-Sbzl for NE and PR3 and Suc-Phe-Leu-Phe-Sbzl for CG, both at 1 mM final concentration and with 0.5 mM DTNB in activity buffer.

Activity buffer	50 mM	Tris
	150 mM	NaCl
	0.01%	Triton X-100
		pH 8.0

7-Amino-4-methylcoumarin substrates (AMC substrates)

Enzymatic activity of NSP4 towards substrates with Arg in P1 compared to Lys in P1 was determined with 1 mM H-Tyr-Arg-Phe-Arg-AMC or H-Tyr-Arg-Phe-Lys-AMC. Substrate hydrolysis was recorded as fluorescent signal after incubation with 250 nM NSP4 at λEx = 370 nm and λEm = 450 nm.

FRETS peptide libraries (Peptanova)

NSP4 (0.5 μM) or NE (0.1 μM) were incubated with 100 μM of each pool designated as FRETS 25-Ala, FRETS 25-Arg, FRETS 25-Ile, FRETS 25-Leu, FRETS 25-Met or FRETS 25-Val (PeptaNova GmbH, Sandhausen, Germany). Each library contained 25 peptides of the general structure D-A2pr(Nma)-Gly-[Phe/Ala/Val/Glu/Arg]-[Pro/Tyr/Lys/Ile/Asp]-P1-Ala-Phe-Pro-Lys(Dnp)-D-Arg-D-Arg, with different residues in P1 position. The proteolytic activity was monitored at λEx = 320 nm and λEm = 450 nm. The initial linear portion of the progress curve was used to calculate the enzymatic activity.

The peptides in the FRETS 25-Arg library were further analyzed by liquid chromatography - mass spectrometry (LC-MS). The library was incubated at a concentration of 0.5 mM with

6 μM NSP4, NE, PR3 or CG and given to Elisabeth Weyher-Stingl (MPI Biochemistry) for LC-MS analysis (ESI-TOF).

FRET peptide libraries (EMC microcollections)

I used 49, 49 and 19 individual FRET peptides for screening of positions P1'P2', P4P3 and P2, respectively. The FRET peptides were provided by Prof. Dr. Karl-Heinz Wiesmüller (EMC microcollections). All FRET peptides featured the structure Mca-Gly-P4-P3-P2-Arg-P1'-P2'-Pro-Lys (Dnp)-rr (r = D-Arg). The concentration of these substrates was determined by using the molar extinction coefficient of 2,4-Dinitrophenyl (17 300 $M^{-1}cm^{-1}$ at λ = 365 nm) and the Beer-Lambert law (see 2.2.10). The hydrolysis of the FRET peptides by NSP4, NE, PR3, CG or mNSP4 was followed at λEx = 320 nm and λEm = 405 nm in activity buffer. The proteases were used at concentrations from 25 to 100 nM and the FRET peptides were present at 1-20 μM. Specificity was assessed by recording the increase of fluorescence over time for each FRET substrate. Only the initial linear portion of the progress curve was used to calculate the enzymatic activity in RFU/min (increase of relative fluorescence units per minute).

PAR-2 FRET peptide

Cleavage of the PAR-2 derived FRET substrate Abz-GSKGR-SLIG-Y (NO_2)-D (EMC microcollections) was determined at λEx = 320 nm and λEm = 405 nm in activity buffer. The FRET peptide was present at 200 μM and NSP4 was added to a final concentration of 2.4 μM. For controls, I used 2.4 μM proNSP4 (unconverted S-NSP4) and enterokinase at the same concentrations as present in the NSP4 sample (1.2×10^{-3} U/μl). NSP4 was converted by enterokinase into its mature form.

2.5.10 Inhibition tests

Potential chemical inhibitors

PMSF. Inhibition of NSP4 by phenylmethylsulfonyl fluoride (PMSF) was assayed in comparison with NE. Both proteases were used at 50 nM and pretreated with different amounts

of PMSF (0.1-10 mM). After 40 min incubation at 37 °C, the best FRET-substrate of NSP4 (Mca-GFKVR-SRP-Lys (Dnp)-rr) and the α1PI FRET substrate (Mca-GEAIPM-SIPPEV-Lys (Dnp)-rr) were added to NSP4 and NE, respectively, at 13 μM. The residual activities of both enzymes after inhibitor treatment were used to estimate the concentration of inhibitor yielding 50% inhibition (IC50).

FPR-cmk. Inhibition of NSP4 by H-D-Phe-Pro-Arg-chloromethylketone (FPR-cmk, Bachem) was tested after preincubation of 100 nM NSP4 with 10, 50 or 100 μM FPR-cmk and subsequent activity assay with the FRET substrate Mca-GFPR-SRP-Lys (Dnp)-rr at a final concentration of 10 μM.

Potential natural inhibitors

Enzymatic inhibition of NSP4 (0.25 μM) was assayed after 45 min of incubation at 37 °C using the following inhibitors and concentrations, elafin, 10 μM; secretory leukocyte peptidase inhibitor (SLPI), 10 μM; α1-proteinase inhibitor (α1PI), 25 μM; monocyte/neutrophil elastase inhibitor (MNEI) with 1mM DTT, 3.3 μM; antithrombin with 300 U/mL heparin, 5.6 μM; C1 Inhibitor, 2.5 μM; or antichymotrypsin (ACT), 5.6 μM. The final volume was 50 μl. The residual enzymatic activity was determined as described for AMC substrates using H-Tyr-Arg-Phe-Arg-AMC. NSP4 without inhibitor was treated accordingly.

Complexation with inhibitors

To test whether NSP4 forms covalent complexes with serpins, NSP4 (0.8 μM) was incubated at 37 °C with 40 μM α1PI, 5 μM antithrombin and 300 U/mL heparin or 4 μM C1 inhibitor in PBS. Samples of 2 μl were taken after 0.5, 5, 30, and 100 min and analyzed by SDS PAGE and Western blot with anti-NSP4 mAbs.

2.6 Immunological methods

2.6.1 Western blot

Analysis by Western blotting or immunoblotting facilitates detection of specific proteins of interest after SDS PAGE applying specific antibodies. After separation by SDS PAGE, the proteins were transferred from the gel matrix onto PVDF membranes by semi-dry electroblotting. The PVDF membrane was first activated in 100% methanol, then gel and membrane were equilibrated in transfer buffer for 5-10 min. The transfer was carried out between 4 sheets of blotting paper soaked in transfer buffer. The proteins were transferred for 30 min at 5 mA/cm^2 in the Fastblot semi-dry blotter (Biometra). The membrane was washed in PBS-T and shaken in blocking milk for 60 min to avoid unspecific binding. The incubation with the first antibody in PBS-T was done overnight at 4 °C. After washing in PBS-T, the membrane was incubated with the secondary antibody in PBS-T for 60 min at room temperature while shaking. All secondary antibodies used for Western blotting were conjugated to horseradish peroxidase (HRP). The membrane was washed and bound HRP-conjugates were visualized using a chemiluminescent substrate (SuperSignal West Pico Chemiluminescent Substrate, Pierce). Membranes were developed with an ECL-Film (CL-XPosure, Thermo Fisher Scientific) or using the Chemidoc imaging system (Bio-Rad).

Dilutions	Rat anti-NSP4 2F6 (rat hybridoma SN)	1:25
	Rabbit anti-α1PI-biotin	1:13000
	Mouse anti-CG	1:50
	Rabbit anti-haptoglobin	1:1000
	Goat anti-rat-HRP	1:5000
	Goat anti-mouse-HRP	1:2500
	Goat anti-rabbit-HRP	1:5000
	Streptavidin-HRP	1:1000

Further information on antibodies is provided in 2.1.6.

Western transfer buffer	10% (v/v)	methanol, p.a.
	150 mM	glycine
	35 mM	Tris
		pH 8,3
PBS-T	140 mM	NaCl
	2,7 mM	KCl
	3,2 mM	$Na_2HPO_4 \times 12\ H_2O$
	1,5 mM	KH_2PO_4
	0,05% (v/v)	Tween 20
		pH 7,4
Blocking milk	1 ×	PBS-T
	5% (w/v)	Non-Fat Dry Milk Blocker, Bio-Rad
		pH 7,4

2.6.2 Immunohistochemistry

Immunological staining of formalin-fixed tissue samples was carried out by Dr. Walter Back (Pathological Institute Bremerhaven, Bremerhaven, Germany) using our rat anti-NSP4 antibodies. The rat hybridoma supernatants were diluted 1:30 and incubated with the sections overnight at 4 °C. Slides were developed with the Ultravision LP staining kit (Thermo Fisher Scientific) and counterstained with hematoxylin.

2.7 Mouse model analysis

The NSP4 knockout mice were purchased as cryo-preserved heterozygous embryos from the Mutant Mouse Regional Resource Center (MMRRC) at the University of California in Davis (see section 2.1.8). The gene of NSP4, Prss57, consists of 5 exons. By homologous recombination, the coding exons 2 and 3 had been targeted as shown in Figure 2.1.

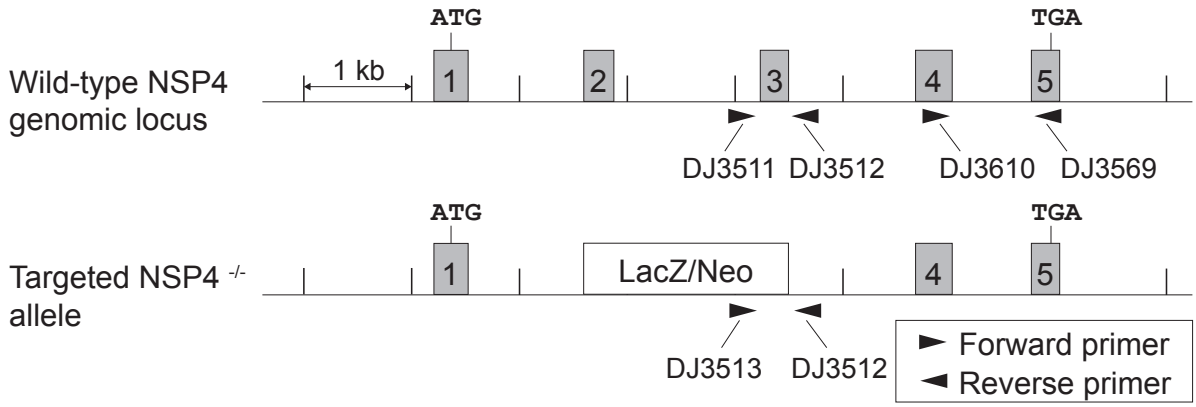


Figure 2.1: **Strategy for genotyping and analysis of NSP4 mRNA in the NSP4 knockout mouse model.** Schematic representation of the WT NSP4 locus and the targeted allele. Gene targeting resulted in deletion of exons 2 and 3. Heterozygous frozen embryos were supplied by MMRRC at UC Davis. Primers used for determination of genotype by PCR and for RT-PCR analysis of NSP4 are shown as arrow heads.

2.7.1 Isolation of genomic DNA from mouse tails

Mouse distal tail pieces of 2-5 mm were immersed in 500 µl tail buffer with 1 mg/ml proteinase K (Roth) and incubated overnight at 56 °C while shaking. After 5 min centrifugation at 20 000 × g, 400 µl of the supernatant was added to 500 µl isopropanol and briefly mixed. Precipitated DNA was pelleted for 10 min at 20 000 × g at 4 °C and the supernatant was removed. The precipitate was taken in 1 ml 70% ethanol and centrifuged for 5 min at 20 000 × g. The supernatant was discarded and the pellet was left to dry for 10 min. The DNA was solved in 100 µl Tris buffer shaking at 55 °C.

Tris buffer	20 mM	Tris
	150 mM	NaCl
		pH 7.4

2.7.2 Genotyping of mice

After isolation of genomic DNA from mouse tails, 1-2 µl were analyzed by PCR with HotStarTaq polymerase (see 2.2.4). The PCR with primers DJ3511 and DJ3512 amplifies the WT-allele represented by a band of 291 bp whereas the PCR with primers DJ3513 and DJ3512 amplifies

the knockout allele corresponding to a band of 237 bp after agarose gel electrophoresis (Figure 2.1). If only one of the PCR results in a positive band, the corresponding mouse is homozygous for the NSP4 gene and is either a WT mouse or an NSP4-/- mouse. If both PCR yield a positive band, the corresponding mouse is heterozygous for the NSP4 locus (NSP4+/-). As WT controls, I used non-transgenic littermates.

2.7.3 Isolation of bone marrow cells

Mice were sacrificed by CO_2 inhalation and the femurs of both hind legs were taken out and freed of soft tissue attachments. After cutting off the ends at both sides, the bone marrow was flushed out by forcing 5 ml RPMI 1640 through the bone using a syringe with a 21-gauge needle. To disperse the cells, the bone marrow solution was sent through a cell strainer.

2.7.4 Reverse transcriptase (RT) PCR

Isolation of mRNA

After isolation of bone marrow cells, the cells were pelleted for 5 min at 500 × g and resuspended in erythrocyte lysis buffer for 1-3 min. The cells were again pelleted and washed in RPMI 1640, pelleted and taken in 1 ml TRIzol ® (Life technologies) and left for 5 min at room temperature. After addition of 200 µl chloroform, it was mixed for 15 s and left for 15 min at room temperature. The solution was centrifuged for 10 min at 13 000 × g at 4 °C and the upper aqueous phase was added to 500 µl isopropanol and left for 15 min at room temperature before centrifuging for 10 min at 13 000 × g at 4 °C. The pellet was taken in 1 ml 75% ethanol and pelleted for 5 min at 5000 × g at 4 °C. The pellet was left to dry for 10 min and resuspended in 20 µl RNase free water. RNA concentration was determined by measuring the absorbance at 260 nm. According to the Beer-Lambert law and using the extinction coefficient of RNA (0.025 $(\mu g/ml)^{-1}$ cm^{-1}) a reading of 1.0 (absorption at 260 nm) is equivalent to 40 µg/ml single stranded RNA.

Erythrocyte lysis buffer (10 ml)	9 ml	0.8% NH_4Cl
	1 ml	0.2 M Tris HCl pH 7.5

Reverse transcription of mouse bone marrow mRNA

First-strand cDNA was generated employing the Verso cDNA synthesis kit (Thermo Fisher Scientific). One µg RNA was added to 0.5 µM reverse primer (Oligo dT) in a total volume of 12 µl and denatured at 65 °C for 10 min. The mix was kept on ice and 4 µl 5 x synthesis buffer, 2 µl 20 mM dNTPs, 1 µl RT-Enhancer and 1 µl Verso enzyme mix was added. The conditions for reverse transcription included 30 min at 42 °C and 2 min at 95 °C.

PCR of mouse bone marrow cDNA

Detection of the mouse NSP4 transcript in bone marrow cDNA was carried out by PCR as described in section 2.2.4 for HotStarTaq PCR with 1 µl cDNA and primers DJ3610 and DJ3569 ((Figure 2.1). I used a touchdown programme including 3, 3 and 24 cycles with annealing temperatures of 61, 58 and 55 °C, respectively. As a positive control, I used primers for mouse β-actin. The touchdown PCR programme for the β-actin transcript comprised 3, 3 and 24 cycles with annealing temperatures of 65, 52 and 59 °C, respectively. The primer sequences are given in the Appendix (7.2).

2.7.5 Isolation of neutrophils

After isolating the bone marrow cells from 3-4 mice, the cells were taken in 20 ml RPMI 1640 in a 50 ml tube. The suspension was carefully underlaid with 10 ml 62.1% Percoll (GE Healthcare) in 2 × HBSS. After centrifugation at 500 × g at 20 °C for 30 min without brake, the cell pellet containing PMN and erythrocytes is taken in erythrocyte lysis buffer (see section 2.7.4). After 1-3 min, the suspension is centrifuged for 5 min at 500 × g and the cells were washed in RPMI. The purity of isolated mouse neutrophils was determined by forward and sideward scatter in flow cytometry. The percentage of PMN in bone marrow cells could be routinely enriched from 32% to 85%.

62.1% Percoll (15 ml)	3 ml	10 × HBSS (Gibco, Life Technologies)
	2.69 ml	H_2O
	9.32 ml	Percoll (GE Healthcare)

2.7.6 Determination of neutrophil oxidative burst

Activation by PMA

After isolation of mouse neutrophils, the cells were resuspended in RPMI 1640 with 1% BSA at a density of 2×10^6 cells/ml and incubated with or without 200 ng/ml PMA for 15 min. Dihydrorhodamine 6 G (indicator for reactive oxygen species, Molecular Probes, Life Technologies) was added to a final concentration of 100 µM. Oxidative burst was followed at 37 °C by measuring fluorescence at λEx = 485 nm and λEm = 520 nm. The oxidative burst signals were compared between WT and NSP4-/- neutrophils using the linear portion of the enzymatic progress curve.

Activation by immune complexes

Immune complexes (IC) were prepared by coating 5 mg/ml ovalbumin (Sigma A5503) on Immulon 4 HBX plates (Thermo Fisher Scientific), at 4 °C overnight. After washing with PBS-T, rabbit anti-ovalbumin at 25 µg/ml was added to the plates and incubated for 60 min at room temperature. Residual antibodies were removed by three wash steps in PBS-T. After isolation of mouse neutrophils, the cells were dispensed in RPMI with 1% BSA at 2×10^6 cells/ml and incubated with 10 ng/ml mouse TNF-α. After 15 min preincubation at 37 °C, the cells were transferred to the preformed immune complexes or into control wells with only ovalbumin coating. Generation of reactive oxygen species was detected with Dihydrorhodamine 6 G as described above for PMA activated neutrophils.

3 Results

3.1 Production of NSP4

Production of natively folded NSP4 was not feasible in bacterial cells (Heike Kittel, Dieter Jenne, unpublished observations). After recombinant protein expression in *E. coli*, NSP4 could be detected only in inclusion bodies. Several refolding protocols were applied by Heike Kittel, but none of these yielded enzymatically active NSP4. The only possibility to produce active NSP4 was in Flp-In^{TM}-HEK 293 cells (Invitrogen, human embryonic kidney cell line), as reported previously (Natascha Perera, diploma thesis). In view of the low yields and many impurities in the cell culture medium, the method described previously was not satisfactory for the production and purification of NSP4. After nickel-chelate affinity chromatography, the elution fractions still contained other proteins. The main contaminant was suspected to be albumin originating from the cell culture medium which contained 5 to 10% FCS (Figure 3.1 A). In the presence of this major albumin impurity, it was impossible to measure the protein concentration and the yield of the NSP4 production. I estimated by SDS PAGE analysis that 1 L cell culture medium contained about 100-200 µg NSP4. As the Flp-In^{TM}-293 is growing adherently, it was difficult to scale up the cell culture. The maximum cell volume that I could cultivate was 500 ml in cell stacks with five levels and 3180 cm^2. The supernatant was routinely harvested after 10-14 days.

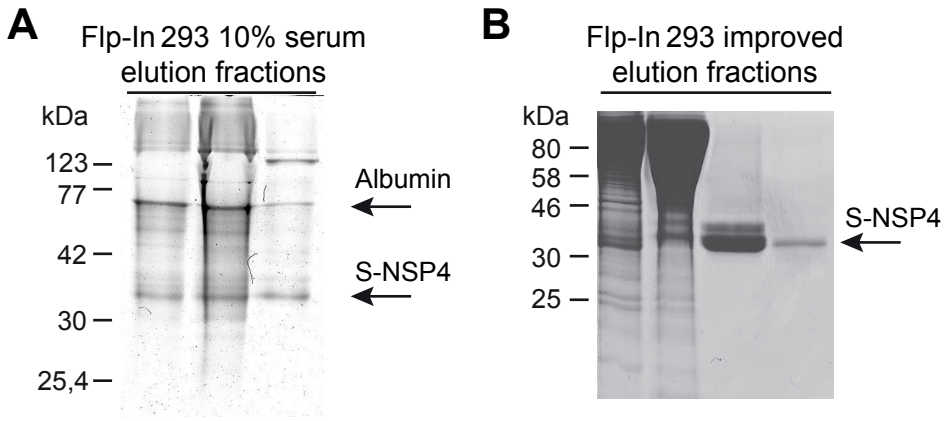


Figure 3.1: **Production of NSP4 in Flp-In^{TM}-293.** Elution fractions after purification of S-NSP4 from Flp-In^{TM}-293 cells stably transfected with SigIgκ-S-EK-NSP4-H_6-pcDNA5. A, Production in adherent FlpIn cell line (cell stacks). Flp-In 293 cells were cultivated for 14 days in cell stacks. The supernatant was harvested and 820 ml were purified using nickel-chelate chromatography. The volume of each elution fraction was 2 ml. Of each fraction a sample of 12 µl was analysed by SDS PAGE and Coomassie Blue staining. Albumin was suspected to be the major contaminant (top arrow). B, Improved production in Flp-In^{TM}-293 cell line, after serum reduction and adaptation to suspension culture. Flp-In^{TM}-293 cells were cultivated for nine days in suspension culture with low-serum medium. By nickel-chelate chromatography, 600 ml culture supernatant was purified. Each elution fraction contained 4 ml and samples of 21 µl of each fraction were analyzed by SDS PAGE and Coomassie Blue staining. NSP4 is represented by two bands which are probably the result of differential glycosylation, as NSP4 features two N-glycosylation sites.

3.1.1 Improved production in Flp-In^{TM}-293

The Flp-In^{TM}-293 were stably transfected with an NSP4 construct featuring an Igκ-signal peptide, an N-terminal S-tag peptide, an enterokinase cleavage site at the mature aminoterminus of NSP4 and a C-terminal 6×His-tag for purification (S-NSP4-pcDNA5). To improve the yields, I firstly established single cell clones from the cell line to compare expression levels of NSP4. The supernatants of the single cell clones were analyzed for NSP4 by Western blotting with rat anti-NSP4 mAb and the four best expressing clones were chosen for further culture. For additional improvement of yield and purity, the FCS in the cell culture medium was reduced. The serum content of the cell culture medium was stepwise reduced from 10% to 1% and at the same time, the cells were also adapted to grow in suspension. The concentration of 0.3%

FCS was found to enable both a sufficient growth rate of the cells and a satisfactory purification protocol of NSP4. Figure 3.1 B shows the purification of NSP4 from Flp-In^{TM}-293 growing in low-serum medium. With these improvements, the yields were 327 $\pm$ 61 µg (n = 12, $\pm$ SD) of pure NSP4 per 1 L of cell culture. The maximum volume for suspension cultures was 1200 ml in 3 L Erlenmeyer cell culture flasks. The culture was harvested after seven to nine days.

NSP4 migrates as two distinct bands, a dominant and a minor band. This cannot be the result of C- or N-terminal truncation. The S-NSP4 construct features a C-terminal 6$\times$ His-tag. If this was removed, purification by nickel-chelate chromatography would not have been feasible. The construct also features an N-terminal S-tag followed by an enterokinase cleavage site. After enterokinase digest (see Figure 3.3), both forms of NSP4 underwent a comparable molecular weight shift. Hence, the N-terminus was preserved in both forms. The difference in size of the two NSP4 forms is probably the result of heterogenous glycosylation. NSP4 features two potential N-glycosylation sites.

3.1.2 Production in HEK 293E

The 'HEK 293E Large Scale Transient Expression System' was developed by Yves Durocher (NRC Biotechnological Research Institute). The system was chosen for NSP4 production because it uses serum-free cell culture conditions and a convenient transient transfection protocol. The latter advantage permitted me to test different NSP4 cDNA constructs in a short period of time. Transfection of HEK 293E with the S-NSP4 construct (as described in 3.1.1) subcloned into the pTT5 vector resulted in very low yields of S-NSP4 in the cell culture supernatant. To improve the production, different signal peptides and fusions with solubility enhancing carriers such as Sumostar (a version of yeast SUMO, Lifesensors Inc.) were tested. Proteins which were secreted from HEK 293E cells at very high yields, such as MOG (myelin oligodendrocyte glycoprotein) and Lamp2A (lysosome associated membrane protein 2A) were also tried as fusion partners for NSP4. Improvement could be achieved by exchanging the N-terminal S-tag peptide to an N-terminal presequence including an N-glycosylation site (SRMHRNGSH; NGS-NSP4) as described in Liu et al. (2009) for MIC1.

The production could be further enhanced by partially replacing or 'diluting' the NSP4 plasmid

with filler DNA for the transfection. To this end, the NSP4 plasmid of interest was mixed with a bacterial plasmid devoid of any mammalian transcription elements (filler DNA). Figure 3.2 A depicts a Western blot analysis of S-NSP4 and NGS-NSP4 expressed after transfection with plasmid of interest in different ratios with filler DNA. For both constructs a plasmid mixture of 1 to 100 lead to the highest yield of NSP4 in culture supernatants. The yield of NGS-NSP4 was higher than that of S-NSP4 in HEK 293E. NGS-NSP4 is also represented by two to three bands which is probably the result of heterogenous glycosylation. Transfection of NGS-NSP4 partially replaced with filler DNA resulted in 373 $\pm$ 28 µg (n = 2, $\pm$ SD) of pure NSP4 from 1 L cell culture volume (Figure 3.2 B).

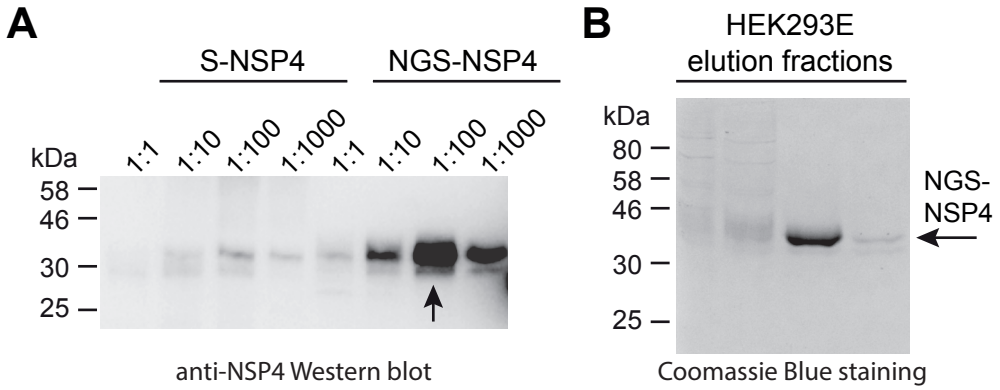


Figure 3.2: **Production of NSP4 in HEK 293E.** A, Cell culture supernatants after transfection of HEK 293E with S-NSP4 or NGS-NSP4 were analyzed by SDS PAGE and Western blotting with rat anti-NSP4 mAb. The expression was optimized by partially replacing or 'diluting' the plasmid of interest (S-NSP4 or NGS-NSP4) with filler DNA (bacterial plasmid without mammalian elements) for the transfection of HEK 293E. Samples of 24 µl of each supernatant (10 ml in total) were analyzed. The highest expression yields were obtained with NGS-NSP4 and a ratio of 1:100 of plasmid of interest and filler DNA (arrow). S-NSP4 is represented by two bands and NGS-NSP4 is represented by two to three bands which are probably the result of differential glycosylation. S-NSP4 features two N-glycosylation sites and NGS-NSP4 three N-glycosylation sites. B, Elution fractions after nickel-chelate chromatography of NGS-NSP4 after production in HEK 293E cells in a volume of 350 ml. For the transfection of cells, the amount of NGS-NSP4 plasmid DNA was reduced to 0.2% and filler DNA was added to 100% (1:500). A sample of 18 µl of each fraction, containing 4 ml in total, was analyzed by SDS PAGE and Coomassie Blue staining.

3.1.3 Other recombinant expression systems

Production of NSP4 was also attempted in insect and yeast cells. Insect cell protein expression was carried out both by Judith Scholl at the MPI Biochemistry (High Five and Sf9) and by Novartis. The production yields were too low for purification. The expression in yeast cells was done in pichia pastoris by U-Protein Express (Utrecht, The Netherlands), but NSP4 protein was undetectable in Western blot analysis.

3.1.4 Conversion of S- or NGS-NSP4 to mature NSP4 by enterokinase (EK)

Serine proteases of the trypsin/chymotrypsin family like NSP4 are only active after the generation of a free mature N-terminus. All NSP4 constructs carried N-terminal extensions of the mature N-terminus to prevent proteolytic activity during recombinant protein expression. The N-terminal extension consisted of an EK cleavage site and either the S-tag peptide or the NGS-peptide. Both N-terminal extensions could be removed by EK, resulting in mature NSP4. Figure 3.3 displays S-NSP4 before and after EK digestion. As incubation with EK resulted in a molecular weight shift of all S-NSP4, I assumed that S-NSP4 was completely converted to mature NSP4.

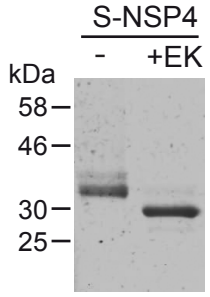


Figure 3.3: **Conversion of S-NSP4 to mature NSP4 by enterokinase (EK) digestion.** Purified S-NSP4 was concentrated 21-fold and is shown before and after incubation with EK. Three µl of both samples was analyzed by SDS PAGE and Coomassie Blue staining

3.2 Characterization of NSP4 in human neutrophils

3.2.1 NSP4 expression is restricted to the myelomonocyte cell lineage

NSP4 is expressed in polymorphonuclear leukocytes (PMNs), but not in other subsets of white blood cells. This was already shown in my diploma work. Other tissues, however, were not tested. To investigate the protein expression of NSP4 in different human tissues, immunohistochemistry was performed by PD Dr. Walter Back (Pathological Institute Bremerhaven, Bremerhaven, Germany) in cooperation with us. NSP4 was only detected in neutrophils and neutrophil precursors of bone marrow tissue, but not in lymph node, spleen, neural network, pancreas, prostate, or arteries (Figure 3.4). Hence, NSP4 exhibited a tissue-restricted expression pattern very similar to NE, CG, and PR3.

3.2.2 The amount of NSP4 is 20-fold lower than that of CG in neutrophils

NE, CG and PR3 are stored in large amounts in neutrophil granules (Campbell et al., 1989) and they are commonly prepared by purification from human neutrophils. In contrast, the content of NSP4 in neutrophils was suspected to be low. This would explain why NSP4 was always overseen in the past. Using purified NSP4 and CG as a protein standard, semiquantitative Western blotting with PMN lysates was performed. An exemplary Western blot comparison is shown in Figure 3.5 (A and B). To compare the expression of CG with NSP4 in human neutrophils, the content of NSP4 was calculated as a percentage of CG. The results are summarized in Figure 3.5 C. NSP4 in PMNs amounts to only 5% of the CG content.

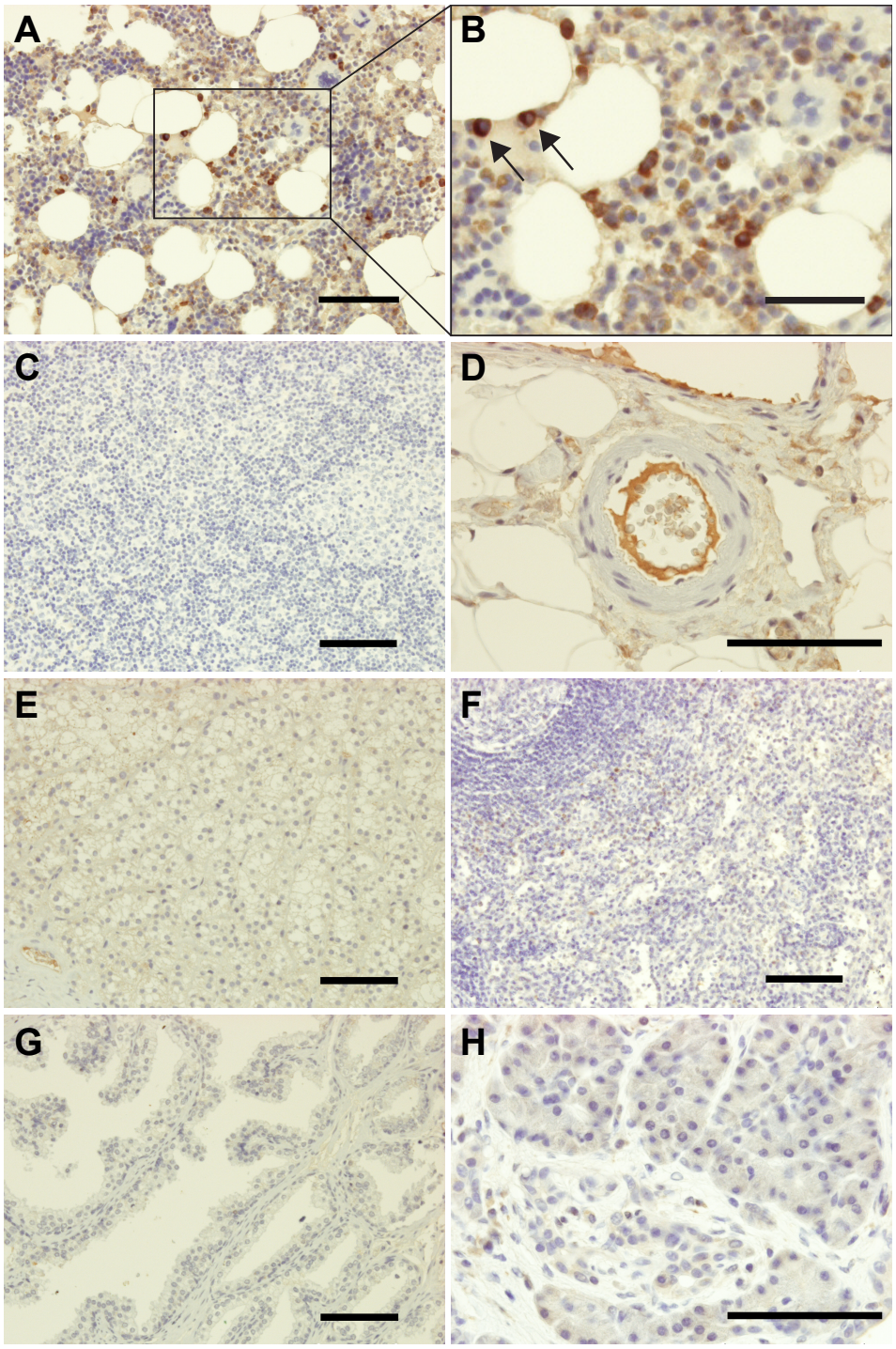


Figure 3.4: **NSP4 staining of different human tissue samples.** NSP4 was detected in neutrophil granulocytes and precursors in bone marrow tissue (A and B, at higher magnification in B with arrows). NSP4 was not detected in lymph nodes (C), arteries (D), neural network (E), spleen tissues (F), prostate (G) or pancreas (H). Scale bars: A and C-H, 100 µm; B, 50 µm. Staining was carried out by PD Dr. Walter Back (Pathological Institute Bremerhaven, Bremerhaven, Germany) using our rat anti-NSP4 mAb (2F6).

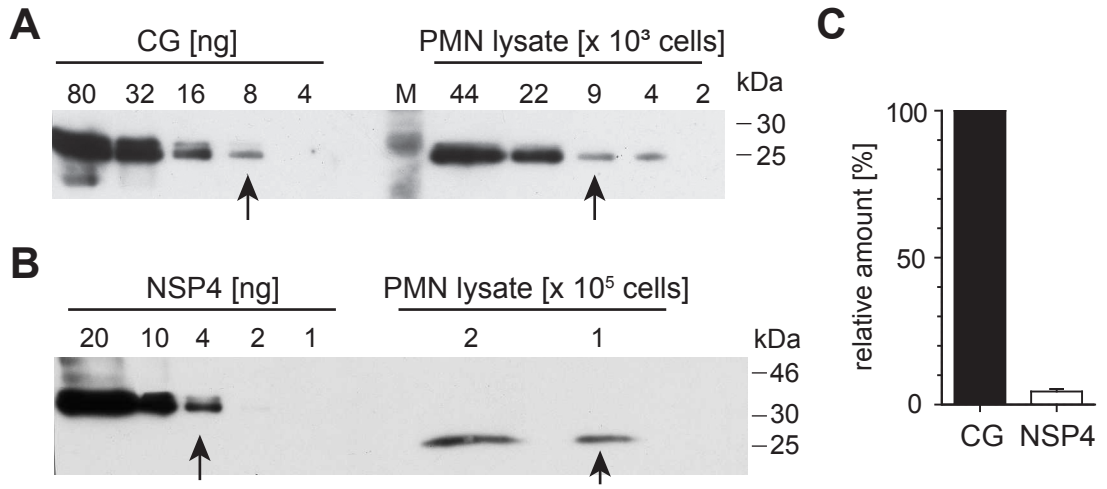


Figure 3.5: **The content of NSP4 was 20-fold lower compared to CG.** A and B, Western blot comparisons of different amounts of commercially available CG (purified from human neutrophils) and PMN lysate (A) and different amounts of recombinant S-NSP4 and PMN lysate (B). CG was detected with mouse anti-CG and NSP4 with rat anti-NSP4. In this exemplary Western blot, 8 ng of CG was detected in 9×10^3 PMNs (A, arrows) and 4 ng NSP4 was found in 1×10^5 PMNs (B, arrows). Hence, 1×10^6 PMNs contain 890 ng CG and 40 ng NSP4. Recombinant S-NSP4 features a higher molecular weight than natural NSP4 in human neutrophils because of additional N- and C-terminal sequences (S-tag, EK-site and 6× His-tag, approx. 4.2 kDa in total). Additionally, in contrast to recombinant NSP4 from HEK 293 cells, natural NSP4 presumably features truncated carbohydrate chains as was already shown for NE and CG (Watorek et al., 1993). M denotes the molecular weight standard. C, Data were pooled from three Western blot comparisons of CG and NSP4 in PMNs (n = 3, ± SD). For this comparison, the content of NSP4 in PMNs is given as a percentage of the content of CG in PMNs.

3.2.3 NSP4 is released after neutrophil activation

In response to neutrophil activation, NE, CG, and PR3 are released into the pericellular environment. *In vitro*, this effect can be induced by stimulation of purified human PMNs with the phorbol ester PMA. Pretreatment with cytochalasin B (cytB), a fungal toxin, was shown to further increase the release of azurophil granules (Yurewicz and Zimmerman, 1977). To test whether NSP4 is also released after activation of neutrophils, PMNs were stimulated with cytB and PMA. After this stimulation, NSP4 was detected in the supernatants by Western blot analysis with anti-NSP4 mAb (Figure 3.6). NSP4 could not be detected in supernatants of PMNs that

were treated likewise but without cytB and PMA. NSP4, therefore, most likely acts in concert with the other NSPs.

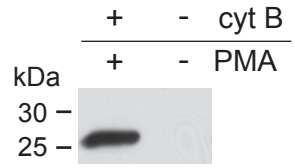


Figure 3.6: **NSP4 was released by activated neutrophils.** PMNs were incubated with or without cyt B and PMA for 30 min at 37 °C. The cell free supernatants of 4.5 $\times$ 10^6 PMN were analyzed by immunoblotting with rat anti-NSP4 mAb. NSP4 was detected in supernatants of activated PMNs, but not in cells treated without cytB and PMA.

3.2.4 NSP4 localizes to the azurophil granules of neutrophils

NE, CG and PR3 are found in the primary or azurophil granules of neutrophils. Targeting of proteins to different subsets of neutrophil storage granules is determined by the time of their expression (Le Cabec et al., 1996). Expression of a gene during a certain stage of granulocyte development in the bone marrow is correlated with the intracellular localization of its product. This question was addressed with the help of Prof. Dr. Niels Borregaard (University of Copenhagen, Copenhagen, Denmark). In collaboration with his laboratory, he detected the highest mRNA levels of NSP4 in the very early stage of myeloblasts/promyelocytes. This finding already suggested a targeting and sequestration of NSP4 into azurophil granules.

To confirm that NSP4 is stored in the azurophil granules, different subcellular fractions of neutrophils were assayed (provided by Prof. Dr. Niels Borregaard, Udby and Borregaard (2007)). After subcellular fractionation of neutrophils, the portions were assayed for specific marker proteins using ELISA by Prof. Dr. Niels Borregaard (Figure 3.7 A). In this way, different fractions were defined that corresponded (from the bottom) to the α-band, containing primarily the azurophil granules, the β1-band, which comprises the specific granules, the β2-band enriched for the gelatinase granules, the γ-band including secretory vesicles and plasma membranes, and the cytosolic fraction (c). All five portions were analyzed by Western blotting using anti-NSP4 mAb. Immunoreactivity was found in the α- and to a minor extent in the β1-band (Figure 3.7

B,C), which fully agrees with the distribution of MPO, a marker of azurophil granules (Faurschou and Borregaard, 2003). This comparison with MPO confirmed that the vast majority of NSP4 is targeted to and stored in azurophil granules of neutrophils alongside with NE, CG and PR3.

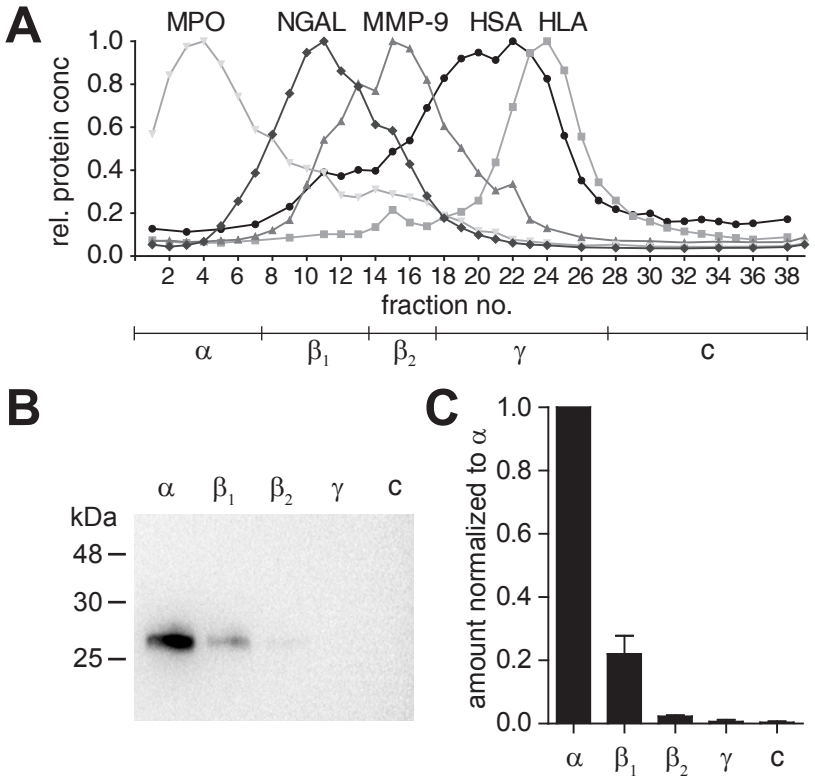


Figure 3.7: **NSP4 is stored in azurophil granules of neutrophils.** A, ELISA analysis showing the relative concentration of MPO (a marker of azurophil granules), NGAL (a marker of specific granules), MMP-9 (gelatinase granule marker), HSA (human serum albumin, marker of secretory vesicles) and HLA (HLA class I antigen, plasma membrane marker) in the subcellular fractions. The fractions were classified into α-, β1-, β2-, γ-band and the cytosolic fraction (c) as indicated below the graph. This data was provided by Prof. Dr. Niels Borregaard. B, 2 µl of each subcellular fraction, representing the amount of sample from 1×10^6 neutrophils, was analyzed by SDS PAGE and Western blotting with anti-NSP4 mAb. C, The immune-reactive bands in (B) were quantified using Image Lab 4.1 (Bio-Rad). The results were normalized to the signal of the α-fraction (n = 3, ± SD).

3.2.5 Processing of NSP4 precursor by DPPI

The human gene of NSP4 (PRSS57) encodes both a potential signal peptide and a prodipeptide (Ala-Gln) at the N-terminus of NSP4. This N-terminal organisation was already reported for the other NSPs (Korkmaz et al., 2010). During translocation into the endoplasmic reticulum, the signal peptide is assumedly cleaved off by signal peptidase. The remaining NSP4 precursor is inactive until its presumed propeptide is removed from the mature N-terminus (Ile-Ile-Gly-Gly) (see section 1.3.1). To investigate biosynthetic conversion of NSP4 into its mature form, I examined the processing of the natural NSP4 (NSP4 precursor) by dipeptidyl peptidase I (DPPI, cathepsin C) *in-vitro*. DPPI also activates the other NSPs during storage into azurophil granules. To this end, the full-length natural cDNA sequence of NSP4 with a C-terminal 6×His-tag was expressed in HEK 293E and purified. Edman sequencing of the purified precursor identified Ala-Gln-Ile-Ile as the N-terminal residues. As expected, signal peptidase cleavage occured right before the prodipeptide. The NSP4 precursor was then treated with DPPI and sequenced again. The sequence obtained was indeed that of the predicted mature N-terminus Ile-Ile-Gly-Gly. I therefore hypothesized that NSP4 is processed by DPPI to its mature form in neutrophil precursor cells like all other NSPs.

3.2.6 Patients with Papillon-Lefèvre syndrome (PLS) lack NSP4 in their neutrophils

The absence of DPPI in PLS patients is associated with a severe reduction of all three NSP activities. Moreover, NE, CG and PR3 are barely detectable in these patients at the protein level, neither as zymogens nor as mature enzymes (Pham et al., 2004). The effect of DPPI deficiency on NSP4 in humans was analyzed using PMNs from a PLS patient. The patient was an 18-year-old man who presented with mild skin findings. He carried a homozygous 854 C > T mutation in the DPPI gene leading to a P285L amino acid exchange. This mutation has been reported and characterized and was associated with an almost complete loss of enzyme activity (Noack et al., 2008). NSP4 was completely undetectable in total cell lysates or supernatants of activated PMNs in this patient (Figure 3.8). Haptoglobin was used as a control and this neutrophil granule protein (Theilgaard-Mönch et al., 2006) was detected at comparable levels in

PMNs from the PLS patient and the healthy control. This finding proves that proNSP4 is indeed another physiological substrate of cathepsin C and that N-terminally processing is required for its accumulation in neutrophil granules *in vivo*.

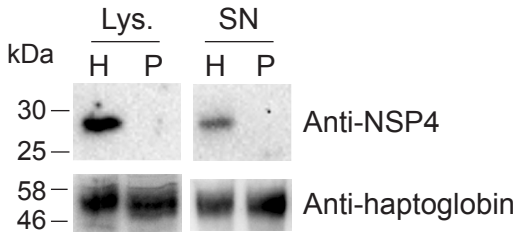


Figure 3.8: **NSP4 was not detected in PMNs of a PLS patient.** PMNs were isolated from a PLS patient (P) and compared with PMNs of a healthy control subject (H). Total cell lysates of 1×10^5 PMNs (Lys.) and supernatants of 3×10^5 PMNs after activation with cytB and PMA (SN) were probed with rat anti-NSP4 mAb and rabbit anti-haptoglobin antibody in immunoblotting.

3.3 Functional analysis of NSP4

3.3.1 Proteolytic specificity of NSP4

It is absolutely necessary to determine the proteolytic activity and cleavage specificity of an unknown protease to elucidate its function. Detailed information about cleavable peptide sequences facilitates the discovery of potential protein substrates, which can then point to a biological role for this protease in a specific biological context. The determination of the specificity of a protease is furthermore useful to construct specific reporter substrates or inhibitors. In this way, the activity of a specific protease can be monitored or inhibited, sometimes even *in vivo*.

Identification of NSP4 cleavage sites in an *E. coli* proteome

For the initial characterization of the cleavage specificity of NSP4, I collaborated with Dr. Oliver Schilling (University of Freiburg). Purified active NSP4 was analyzed by a recently developed method called proteomic identification of protease cleavage sites (PICS) (Schilling and Overall, 2008). Dr. Oliver Schilling used three proteome-derived peptide libraries, generated with either chymotrypsin, GluC (*Staphylococcus aureus* Protease V8) or trypsin digestion of

total *E. coli* proteins. Chymotrypsin generates peptides by cleaving after aromatic amino acids, GluC cleaves C-terminal of glutamic acid and trypsin hydrolyzes peptide bonds after basic amino acids, namely lysine or arginine. Some cleavage sites are always eliminated during library generation due to the cleavages generated by the protease. These cleavage sites are lost in the peptide library and cannot be analyzed by adding the protease of interest. Three peptide libraries were generated using three different proteases, as the cleavage specificity of NSP4 was completely unknown so far. Based on sequence comparisons between NSP4 and the other NSPs, I expected a specificity similar to NE or PR3. NE and PR3 cleave peptide bonds after aliphatic residues like Ile, Leu or Val. However, there was also evidence for a trypsin-like specificity and hydrolysis after Arg.

All peptides that were cleaved by NSP4 can be found in the Appendix (7.5). The results are summarized in Figure 3.9 A. Indeed, NSP4 showed a strong preference for arginine in P1 with both GluC and chymotrypsin PICS libraries. The peptides in the trypsin library were already cleaved by trypsin after arginine residues. Hence, analysis of the trypsin library showed additional P1 preferences for Gln and Ala. However, without Arg in P1, the preference for Gly in P1' is stronger than any preference in P1. The preferred cleavage before Gly in P1' was shared in all three libraries.

Comparison of Arg and Lys in P1

Considering size and charge of the side chain, arginine is very similar to lysine. In the preparation process of peptide libraries for PICS, lysines have to be dimethylated. A specificity towards non-methylated lysines cannot be revealed using the PICS technique. To overcome this, I compared Arg and Lys in two custom-made fluorescent substrates that were at hand in our laboratory. These fluorescent AMC (7-Amino-4-methylcoumarin) substrates were alike except for Arg/Lys in P1 (H-Tyr-Arg-Phe-[Arg/Lys]-AMC). NSP4 displayed a strong selectivity for Arg in P1, as only cleavage of the Arg-substrate could be observed (Figure 3.9 B).

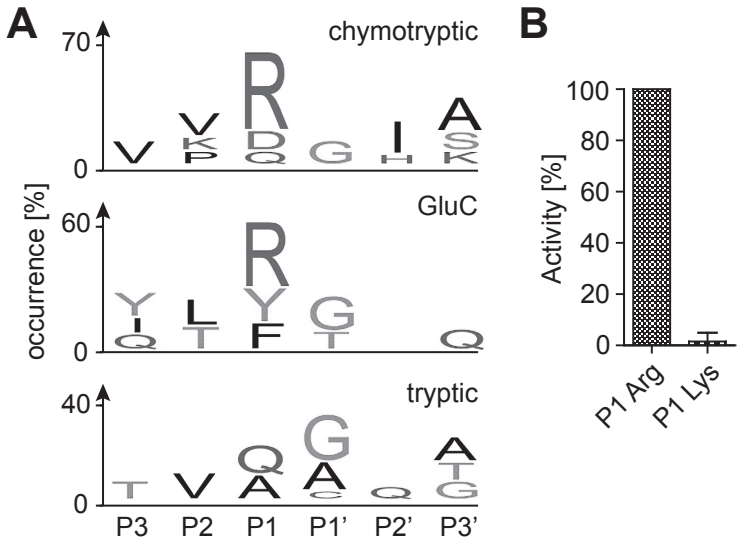


Figure 3.9: **NSP4 cleaves after Arg in P1 position.** A, PICS (proteomic identification of protease cleavage sites) specificity profile for NSP4 with chymotrypsin, GluC (*Staphylococcus aureus* Protease V8), and trypsin libraries (from top to bottom) generated from *E. coli*. The order and size of the letters for each position corresponds to the frequency of occurence of these amino acids in the cleaved peptides. Sequence logos were generated with IceLogo (Colaert et al., 2009). PICS analysis was carried out by Dr. Oliver Schilling. B, Proteolytic activity of 0.25 µM NSP4 and 1 mM AMC (7-Amino-4-methylcoumarin) substrates with either Arg or Lys in P1 position. Fluorescent signal intensity was determined with excitation at λEx = 370 nm and light emission at λEm = 450 nm (n = 3, ± SD).

Comparison of NE-substrates and Arg-substrates

To ensure that Arg in P1 is indeed preferred over NE-substrates with aliphatic residues in P1, six FRET (fluorescence resonance energy transfer) libraries were purchased from Peptanova. Each library contained 25 different FRET substrates of the structure D-A2pr(Nma)-Gly-[Phe/Ala/Val/Glu/Arg]- [Pro/Tyr/Lys/Ile/Asp] -P1- / -Ala-Phe-Pro-Lys(Dnp)-D-Arg-D-Arg. For each library, only the P1 position was different. I used libraries with Arg, Ala, Ile, Leu, Met or Val in P1. The libraries were chosen because too many fixed amino acids in a substrate library can influence the P1 specificity of a protease. The use of five different amino acids in P3 and P2 reduces the influence of these positions and enables a better screening of the preferred P1 position. Figure 3.10 shows the activity of NSP4 towards the six libraries in comparison with NE. Whereas NE showed hydrolyzing activity in all libraries, NSP4 only cleaved substrates in

the P1-arginine library.

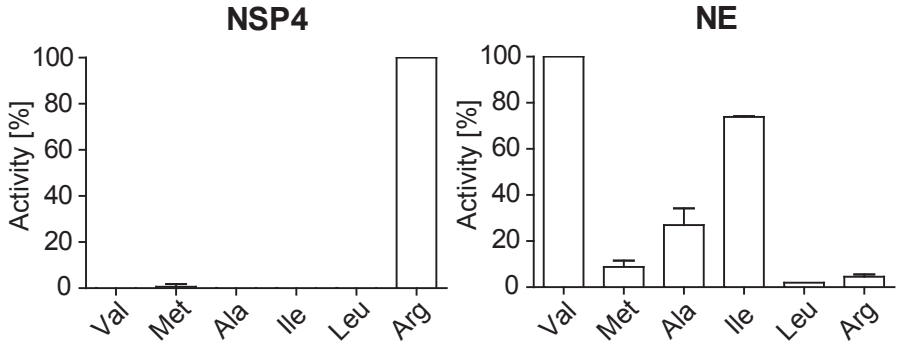


Figure 3.10: **NSP4 prefers Arg in P1 over aliphatic residues.** The proteolytic activities of recombinant NSP4 and NE (Elastin products) were assayed with six different FRET libraries (25 peptides in each) with different amino acids in P1 position. The activity was determined as increasing fluorescence over time for each library and each enzyme. The initial linear portion of the progress curve was used to calculate the enzymatic activity in RFU/min (increase of relative fluorescence units per minute). Average values with standard deviations (SD) are given as a percentage of the value of the library that was cleaved best by the tested protease. The maximum value was set to 100%. (NSP4: n=3, NE: n=2; ± SD).

3.3.2 Chemical inhibitors of NSP4

PMSF

A standard irreversible inhibitor for serine proteases is PMSF. The inhibition of NSP4 by PMSF was tested with different concentrations of PMSF and compared to the inhibition of NE by PMSF (Figure 3.11). The initial linear portion of the graph was used to evaluate the concentration of PMSF needed for 50% inhibition (IC 50). The IC 50 for NE was estimated to be 80 μM ± 4.4 μM and the IC 50 for NSP4 was estimated 705 μM ± 51 μM. PMSF inhibited NE 9-fold better than NSP4.

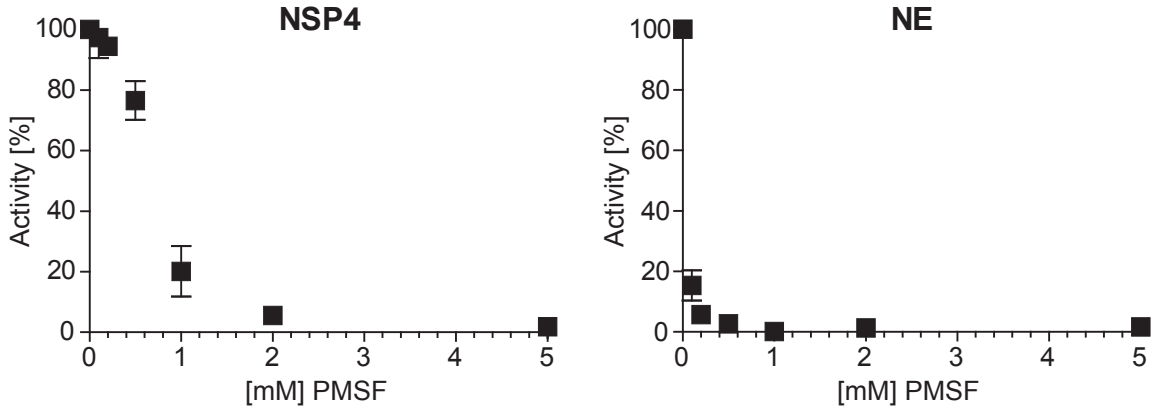


Figure 3.11: **PMSF inhibits NE better than NSP4.** Inhibition of 50 nM NSP4 or NE by increasing amounts of PMSF. NSP4 and NE were preincubated with PMSF and the residual activity was measured using FRET substrates (the α1PI FRET substrate Mca-GEAIPM-SIPPEV-Lys [Dnp]-rr for NE and Mca-GFKVR-SRP-Lys [Dnp]-rr for NSP4) at 13 μM. The activity of protease without inhibitor was set to 100% (n=3 ± SD).

FPR-cmk

In search for more efficient inhibition of NSP4, a chloromethylketone was tested with Phe-Pro-Arg in the P3-P1 positions (FPR-cmk), a specific inhibitor for serine proteases with arginine-specificity. NSP4 was preincubated with this inhibitor and Figure 3.12 illustrates that NSP4 was completely inhibited by 50 μM FPR-cmk.

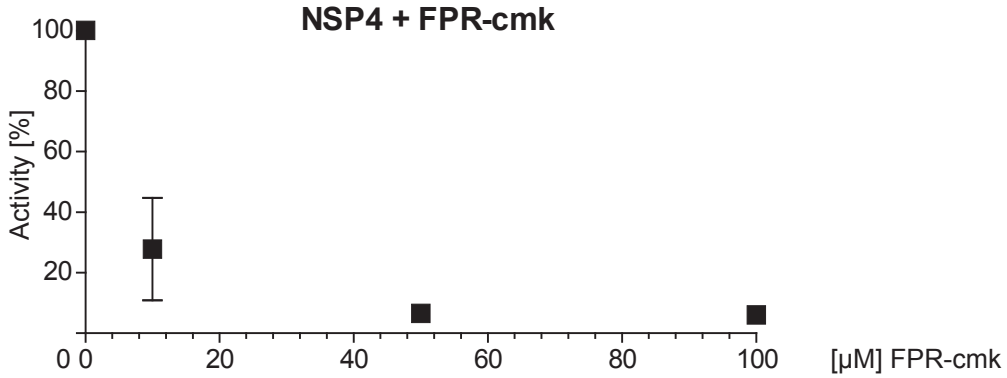


Figure 3.12: **NSP4 is strongly inhibited by FPR-cmk.** Inhibition of 100 nM NSP4 by different amounts of FPR-cmk. The residual activity was measured using a FRET substrate (Mca-GFPR-SRP-Lys [Dnp]-rr) at 10 μM final concentration. The activity of protease without inhibitor was set to 100% (n = 3 ± SD).

3.3.3 Natural inhibitors of NSP4

The inhibitory capacity of natural human serine protease inhibitors from blood or neutrophils on NSP4 was assayed by cleavage of H-Tyr-Arg-Phe-Arg-AMC by NSP4 after preincubation with inhibitors. I tested the inhibitors in different concentrations. I tried to use high molar excess of inhibitor over protease but of some inhibitors I had to use less than of others because of the amounts that were at hand in our laboratory at that time. However, I tested all inhibitors in at least 10-fold molar excess over NSP4. Many proteases are not inhibited at equimolar concentrations of inhibitor and protease, but only at molar excess of inhibitor over protease. Addition of 100-fold excess of α1PI, 22-fold excess of heparin accelerated antithrombin and 10-fold excess of C1 inhibitor completely prevented NSP4 activity. NSP4 activity was not inhibited by 40-fold excess of elafin, 40-fold excess of SLPI, 13-fold excess of MNEI or 22-fold excess of ACT (Figure 3.13 A).

Complexation with Serpins

α1PI, antithrombin and C1 inhibitor are serpins that, in most cases, form covalent complexes with the inhibited protease. The covalent complexes can only form after the protease has hydrolyzed a specific bond in the reactive center loop (RCL) of the serpin. Complex formation between NSP4 and these three serpins was studied by immunoblotting after different incubation times with these inhibitors using anti-NSP4 mAb. As I intended to analyze which of these potential natural inhibitors would be the fastest to inhibit NSP4 in *in vivo* situations, α1PI, heparin-activated antithrombin and C1 inhibitor were used at concentrations found in human plasma (Hortin et al., 2008), resulting in a 50-, 6.5-, or 5-fold molar excess of inhibitors, respectively. NSP4 formed covalent complexes with all three serpins as shown by the immunoreactive bands at approximately 80-90 kDa (Figure 3.13 B). Complexes with α1PI could be detected only after 30 min of incubation, whereas the complexation with heparin-activated antithrombin and C1 inhibitor could be observed already after 1 min and 5 min, respectively. The cleavage site in the RCL (reactive center loop) of α1PI is after Met, while antithrombin and C1 inhibitor feature an arginine residue in the P1 position of their RCL. This observation again confirms that NSP4 preferably hydrolyzes peptide bonds C-terminal of arginine. After prolonged incubation

times with α1PI, antithrombin or C1 inhibitor, there was still some uncomplexed NSP4 seen by Western blotting. As the proteolytic activity of NSP4 was completely abolished by these three serpins (Figure 3.13 A), a proportion of NSP4 is most likely inactive due to incorrect folding or aggregation.

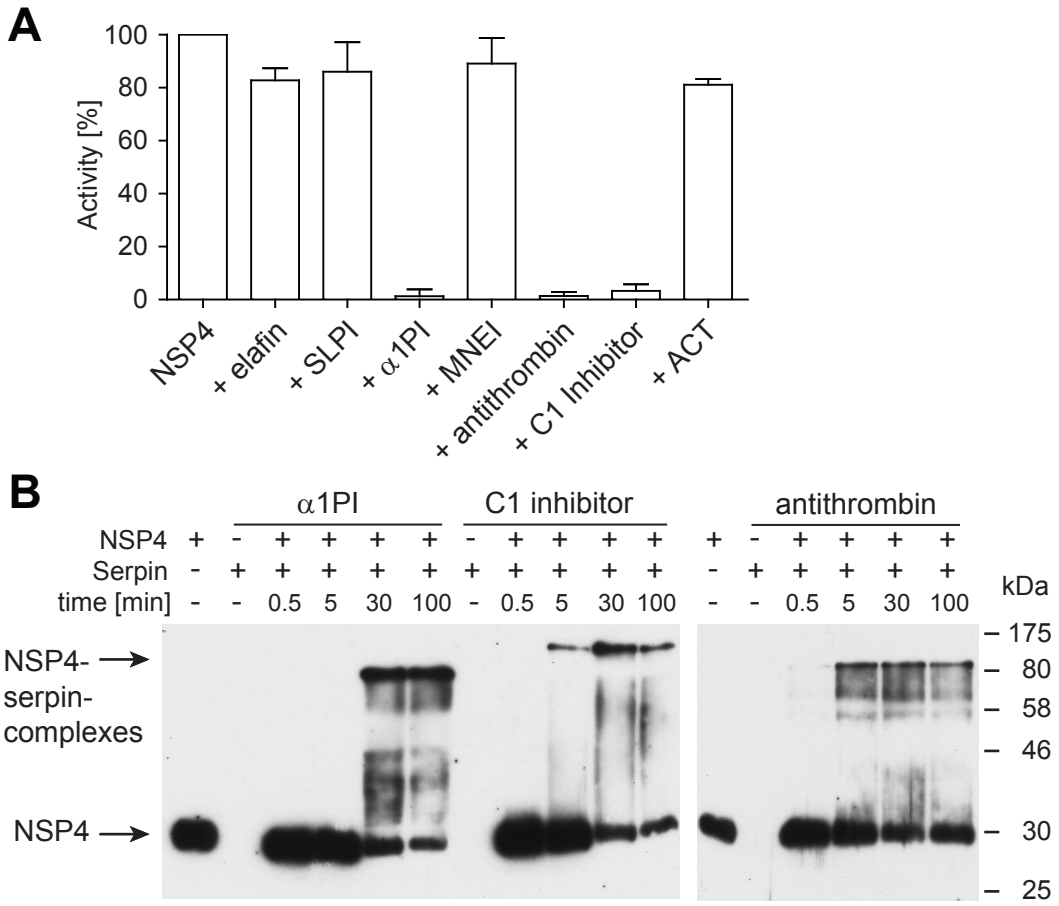


Figure 3.13: **NSP4 is inhibited by α1PI, antithrombin and C1 inhibitor and forms covalent complexes with these serpins.** A, 0.25 μM NSP4 was incubated with 10 μM elafin, 10 μM SLPI, 25 μM α1PI, 3.3 μM MNEI with 1 mM DTT, 5.6 μM antithrombin with 300 U/mL heparin, 2.5 μM C1 inhibitor, or 5.6 μM ACT in 50 μl. The residual activity of NSP4 was determined by measuring the hydrolysis rate of H-Tyr-Arg-Phe-Arg-AMC. The activity of NSP4 without inhibitor was set to 100%. Shown data are the mean from three independent experiments (n = 3, ± SD) with exception of MNEI and ACT (n = 2, ± SD). B, 0.8 μM NSP4 were kept at 37 °C with 40 μM α1PI, 4 μM C1 inhibitor or 5 μM antithrombin with 300 U/mL heparin and samples of 2 μl were analyzed after different timepoints by immunoblotting with anti-NSP4 mAb. NSP4 or the serpins alone were treated similarly for comparisons.

3.3.4 Positional specificity profiling of NSP4

Specificity analysis has unequivocally revealed a preference of NSP4 for arginine in P1. With this in mind, I wanted to study the impact of residues in the adjacent positions in more detail. Additional amino acids in close proximity to the scissile bond can also potently influence the cleavage preferences of NSP4. Analysis of the extended specificity of NSP4 for P4-P2' should facilitate the construction of specific substrates and inhibitors. Furthermore, information on amino acids that were preferred, accepted or excluded in certain positions by NSP4 will assist in selecting potential natural substrates of NSP4. For the detailed analysis of the positions P4-P2', intramolecularly quenched fluorogenic (förster resonance energy transfer, FRET) substrates with arginine in P1 were used. FRET substrates can by synthesized with extended peptide sequences before and after the scissile bond. Extended peptide sequences around the scissile bond are known to increase both sensitivity and specificity of proteases towards a substrate.

P3P2 preference of NSP4 (FRETS25Arg)

The starting point was the FRETS-25Arg library from Peptanova. The library was analyzed by LC-MS after incubations with NSP4. The company provides the user with information on accurate molecular weights of the resulting peptide fragments after all possible cleavage events. Using this information, the LC-MS data generated by Elisabeth Weyher-Stingl (MPI Biochemistry) permitted me to identify those substrates that were cleaved by NSP4. As early as 5 min of incubation with NSP4, cleavages of substrates with the P3-P2 combinations Phe-Pro, Phe-Tyr, Phe-Ile, Val-Pro, Arg-Pro, Val-Ile, Val-Lys and Arg-Ile could be detected. To identify unique substrates for NSP4, the library was also incubated with NE, CG and PR3. None of the peptide substrates was cleaved by CG and PR3. NE, however, was capable of digesting all substrates carrying an Ile in the P2 position. On account of this elastase activity, a Phe and a Pro was chosen for the positions P3 and P2, respectively, in further profiling experiments. This substrate was cut by NSP4 very fast. Pro confers a high degree of peptide resistance to other serine proteases as serine proteases do not cleave before and after proline residues. Placing a Pro in P2 prevents the cleavage after P3 as well as before P1.

P1'P2' preference of NSP4

A small collection of FRET peptide substrates with varying amino acids at the P1' and P2' positions was kindly provided by Prof. Dr. Karl-Heinz Wiesmüller (EMC microcollections, Tübingen, Germany). This library of 49 individual FRET substrates featured the sequences Mca-Gly-Phe-Pro-Arg-[Phe/Val/Ser/Asn/Glu/Arg/Gly]-[Phe/Val/Ser/Asn/Glu/Arg/Gly]-Pro-Lys(Dnp)-D-Arg-D-Arg. All substrate sequences are given in the Appendix (7.4). Seven characteristically different amino acids were used for the P1' and P2' positions. The enzymatic activity of NSP4 with each substrate is shown in Figure 3.14 A. For the enzymatic activity on FRET substrates only 25-40 nM NSP4 and 1-2 μM substrate were used. As expected, the efficacy of cleavage was already much better than with AMC substrates. The four best FRET substrates from this P1'P2' library (P1'-P2': S-F, S-S, S-N, S-R) were also examined for potential susceptibility to NE, CG or PR3-mediated hydrolysis (Figure 3.15 A). None of the three NSPs showed enzymatic activity towards these substrates. For further profiling studies, positions P1' and P2' were fixed to Ser-Arg as this combination was most efficiently cleaved by NSP4 and specific for NSP4.

P4P3 preferences of NSP4

The amino acid residues in P4 and P3 were varied in a similar manner yielding 49 individual FRET substrates (a kind gift from Prof. Dr. Karl-Heinz Wiesmüller) with the following sequences, Mca-Gly-[Phe/Ala/Ser/Asn/Glu/Arg/Gly]-[Tyr/Ala/Ser/Asn/Asp/Lys/Gly]-Pro-Arg-Ser-Arg-Pro-Lys(Dnp)-D-Arg-D-Arg. The enzymatic turnover of these individual substrates by NSP4 is shown in Figure 3.14 B. To demonstrate specificity, the three best substrates (P4-P3: F-K, A-K, A-N) were assayed for cleavage by the other NSPs. Figure 3.15 B shows that all three are specific for NSP4 as they were not hydrolyzed by NE, CG or PR3. For the final step of improving the P2 residue, the best combination Phe-Lys was fixed at the P4 and P3 position, respectively.

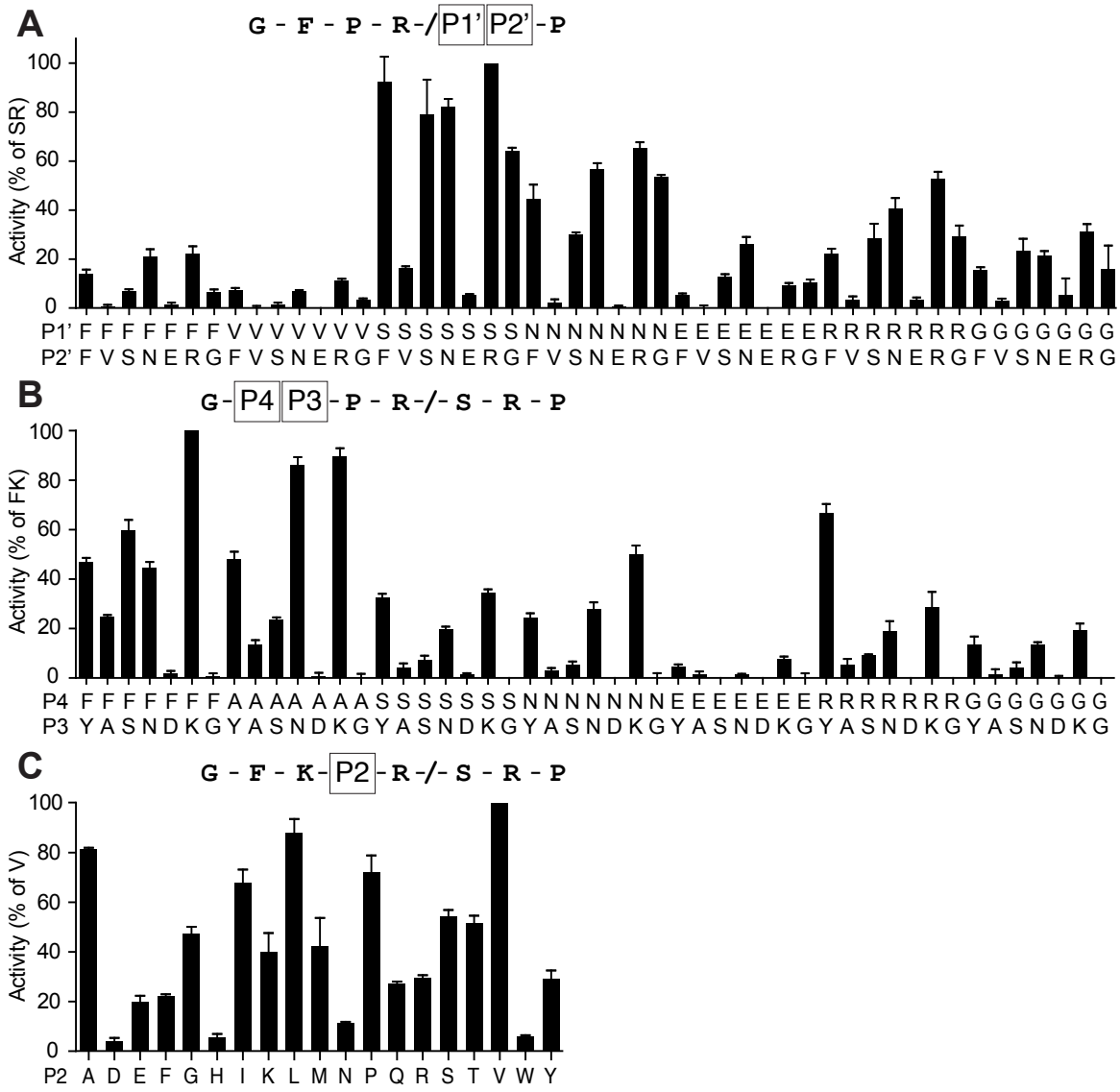


Figure 3.14: **NSP4 specificity analysis with FRET substrates.** Fluorescent signals were monitored at λEx = 320 nm and λEm = 405 nm. Results were normalized to the signal of the best substrate and are given in percent (n = 3, ± SD). A, Enzymatic activity of NSP4 cleaving FRET substrates with 49 different amino acid combinations at positions P1' and P2'. NSP4 was added to a final concentration of 40 nM, each FRET substrate was present at 2 μM. B, Profiling of positions P4 and P3. Proteolytic activity of NSP4 towards 49 FRET substrates differing in positions P4 and P3. NSP4 was added to a final concentration of 40 nM, each FRET substrate was present at 1 μM. C, Profiling of the P2 position. Enzymatic activity of NSP4 cleaving 19 FRET substrates with different amino acids in P2. NSP4 was added to a final concentration of 25 nM, each FRET substrate was present at 1 μM.

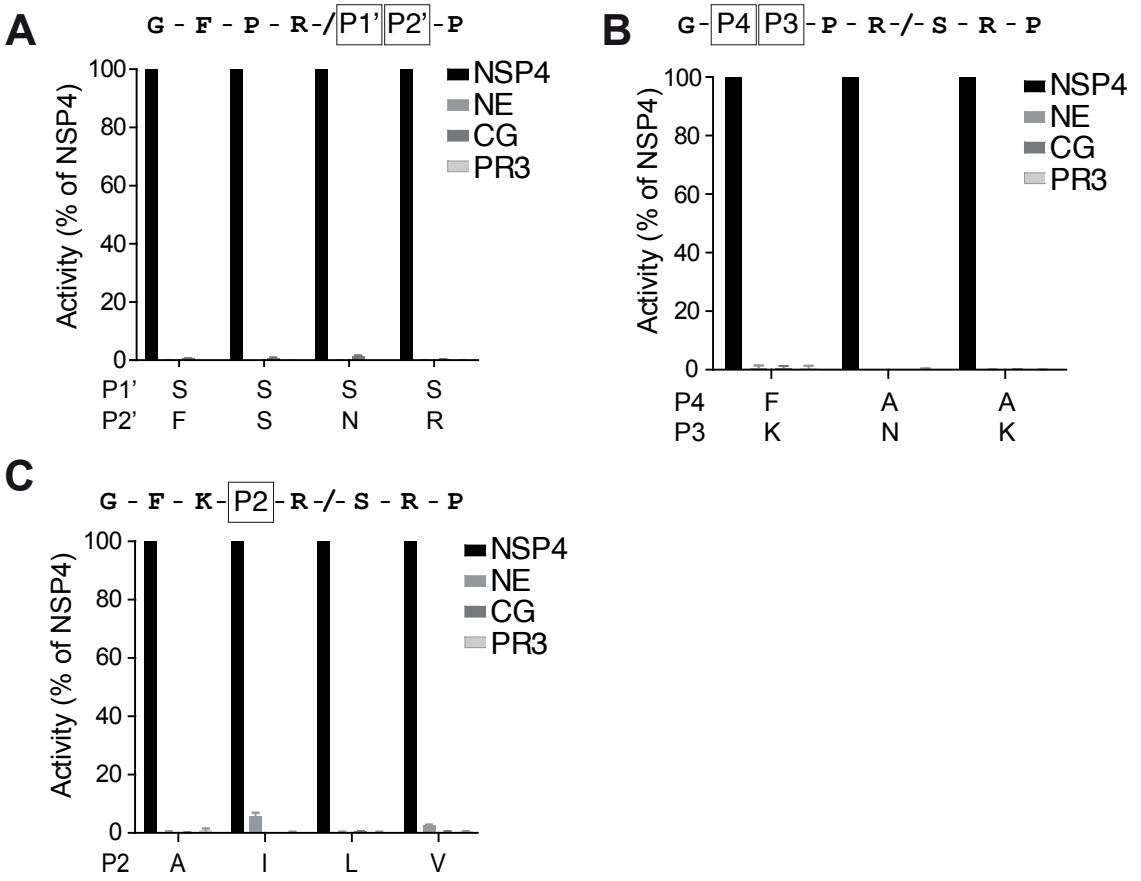


Figure 3.15: **Specificity of FRET substrates towards NSP4.** Fluorescent signals were monitored at λEx = 320 nm and λEm = 405 nm. Results were normalized to the signal initiated by NSP4 and are given in percent (n = 3, ± SD).A, Enzymatic activity of NSP4 in comparison to NE, CG and PR3 using the best substrates from Figure 3.14 A. All proteases were used at 100 nM and the FRET substrates at 20 μM. B, Enzymatic activity of NSP4 in comparison to NE, CG and PR3 using the best substrates from Figure 3.14 B. All proteases were used at 100 nM and the FRET substrates at 10 μM. C, Enzymatic activity of NSP4 in comparison to NE, CG and PR3 using the best substrates from Figure 3.14 C. All proteases were used at 50 nM and the FRET substrates at 5 μM.

P2 preferences of NSP4

The P2 preferences of NSP4 were determined in a library of 19 individual FRET substrates (kindly synthesized by Prof. Dr. Karl-Heinz Wiesmüller). The substrates of the structure Mca-Gly-Phe-Lys-P2-Arg-Ser-Arg-Pro-Lys (Dnp)-D-Arg-D-Arg comprised all natural amino acids in position P2 except Cys. Cysteine was omitted due to its potential to form disulfide bonds by oxidation. The five substrates cleaved best by NSP4 featured Val, Leu, Ala, Pro or Ile in P2 (Figure 3.14 C). All in all, these data agree with our initial screening of the FRETS-25Arg library yielding Pro, Ile, Tyr and Lys in P2. Cleavage of the four best substrates (the substrate with Pro in P2 was already given in Figure 3.15 B) by CG or PR3 was not detectable. There was, however, a very low activity of NE towards the substrates with Ile or Val in P2. Even if Val in P2 increases the cleavage efficacy by NSP4, Pro in P2 increases the specificity of the substrates by preventing cleavages N- or C-terminal of Pro by all NSPs.

With these highly improved and specific substrates of NSP4, detection of NSP4 activity in PMN lysate and after PMN degranulation was examined. Unfortunately, it was not possible to detect NSP4 activity in these biological samples. Considering the low levels of NSP4 in neutrophils (see 3.2.2), I suspected that our FRET substrates were still not sensitive enough.

3.3.5 Production of an NSP4 reactive serpin forming covalent complexes (α1PI*)

The proteolytic activity of a protease can be demonstrated by the formation of a covalent complex with a serpin, as shown in section 3.3.3. To test whether NSP4 is active in neutrophil lysates and after neutrophil activation, a specific serpin for NSP4 was required. α1PI could not be used because it reacts much faster with NE than with NSP4. Antithrombin could not be used because it is degraded very rapidly by NE (Jochum et al., 1981; Jordan et al., 1987). To overcome these problems, I constructed a tailor-made, optimized α1PI variant with targeted specificity for NSP4. The P4 to P1 residues of the RCL of α1PI were changed from Ala-Ile-Pro-Met to Ile-Lys-Pro-Arg (α1PI*). Ile in P4 was adopted from antithrombin. Lys-Pro-Arg in P3-P2-P1 were taken from the specificity profiling in section 3.3.4. I chose Pro in P2 over Val, Ile and

Ala, because it prevented the cleavage by other serine proteases before and after the P2 position. This Pro was expected to increase the stability of α1PI*. The sequence of the optimized RCL was evaluated as a FRET peptide substrate (Mca-Gly-Ile-Lys-Pro-Arg-Ser-Arg-Pro-Lys (Dnp)-D-Arg-D-Arg). This FRET substrate perfomed comparably well against the best FRET derived from the specificity profiling (Mca-Gly-Phe-Lys-Val-Arg-Ser-Arg-Pro-Lys (Dnp)-D-Arg-D-Arg) (Figure 3.16 A). Moreover, the other NSPs did not cleave this FRET substrate (Figure 3.16 B).

α1PI* was produced in HEK 293E after transient transfection and was represented by two bands after SDS PAGE and silver staining (Figure 3.16 C). Figure 3.16 D depicts α1PI* analyzed by Western blotting with anti-α1PI antibody after incubation with NSP4, NE, PR3 or CG. α1PI* formed a covalent complex with NSP4 that is visible after 5 min and 30 min of incubation. Without NSP4, the upper α1PI* band was the major α1PI*-band. After addition of NSP4, only α1PI* of the upper band appears to have formed a complex with NSP4, as only the density of this band was reduced. α1PI* that is represented by the lower band seems to be inactive. PR3 did not form a covalent complex with α1PI* or degrade it. After 5 min of incubation with NE, the upper band of α1PI* was not detected anymore and after 5 min of incubation with CG, α1PI* was probably completely degraded. Formation of the covalent complex of α1PI* and NSP4 is specific for NSP4 as it did not occur with PR3, NE or CG.

The two bands of α1PI* could be explained either by proteolytic truncation or differential glycosylation patterns. α1PI* features three potential N-glycosylation sites. Proteolytic truncation of α1PI* could have occurred during expression in HEK 293E. Several matrix metalloproteinases including collagenase (MMP-1) (Mast et al., 1991) or gelatinase B (MMP-9) (Liu et al., 2000) can cleave α1PI and destroy its inhibitory activity. Both of these proteases were shown to be expressed in HEK cells (Liu and Wu, 2006). This would also explain, why the lower band of α1PI* seems inactive. α1PI* of the upper band was probably cleaved and inactivated by NE which results in a merge of the two bands in a single lower band.

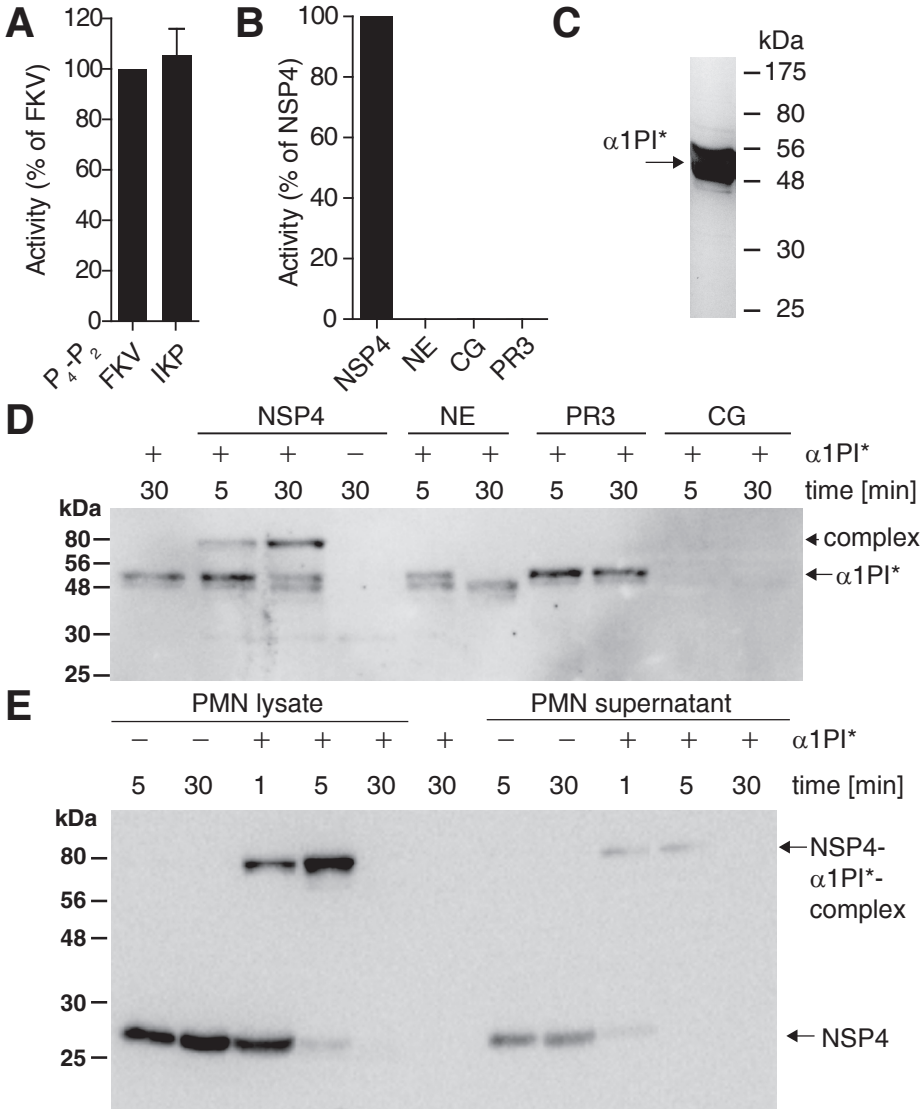


Figure 3.16: **Production of an NSP4 specific serpin and detection of active NSP4 in neutrophil lysate and released by activated neutrophils.** A, Enzymatic activity of 25 nM NSP4 cleaving the best substrate from Figure 3.14 (FKV, Mca-Gly-Phe-Lys-Val-Arg-Ser-Arg-Pro-Lys [Dnp]-D-Arg-D-Arg) in comparison to the FRET peptide sequence designed for α1PI* (IKP, Mca-Gly-Ile-Lys-Pro-Arg-Ser-Arg-Pro-Lys [Dnp]-D-Arg-D-Arg), both at 1 μM (n = 3, ± SD). B, Enzymatic activity of NSP4 cleaving the α1PI* FRET (5 μM) in comparison to the other NSPs. All proteases were used at final concentrations of 50 nM (n = 2, ± SD). C, α1PI* was recombinantly expressed in HEK 293E and is shown after SDS PAGE and silver staining. α1PI* is represented by two bands that either result from proteolytic truncation during expression in HEK 293E or from heterogenous glycosylation. D, α1PI* was incubated with NSP4, NE, PR3 or CG and analyzed after different incubation times by immunoblotting with anti-α1PI antibody. E, Total cell lysates of human neutrophils or supernatants of cytB and PMA stimulated neutrophils were incubated with or without α1PI* and analyzed by Western blotting with anti-NSP4 mAb.

3.3.6 Detection of active NSP4 in neutrophil lysate and released by neutrophils

The NSP4-optimized serpin α1PI* was employed to detect active NSP4 in total cell lysates of human neutrophils. Furthermore, human neutrophils were activated by cytB and PMA and the resulting supernatants were also incubated with α1PI*. The samples were analyzed after different incubation times by Western blotting with anti-NSP4 mAb. Figure 3.16 E shows that NSP4 in neutrophil lysates and after degranulation is completely active and forms a covalent complex with α1PI*. As early as 1 min, the heterodimeric band of α1PI* and NSP4 could be observed. After 30 min, the complex was most likely degraded by other proteases in the sample, while NSP4 in the sample without α1PI* is still visible. *In vitro*, in the absence of additional proteases, degradation of the α1PI*-NSP4 complex did not occur (Figure 3.16 D). Proteases of the trypsin/chymotrypsin family were reported to become susceptible to degradation by several proteases after serpin-complexation (Stavridi et al., 1996; Huntington et al., 2000).

This finding strongly indicates that NSP4 is stored and released by PMNs as an active endoproteinase. This result was expected, as NSP4 was not detected and most likely degraded in PMNs of a PLS-patient (3.2.6) without DPPI convertase activity.

3.4 Is PAR-2 a potential natural substrate of NSP4?

Potential natural substrates of NSP4 were chosen on the basis of the specificity of NSP4 and accessibility of the substrate to NSP4 during neutrophil responses. Only pericellular and membrane proteins that come into close contact with active NSP4 can be cleaved by NSP4. Several possible natural substrates of NSP4 were tested for this thesis, like complement C3, prorenin, hepatocyte growth factor, haptoglobin or SIRP-alpha. So far, cleavage of these substrates by NSP4 could not be shown *in-vitro*. Here, I will only present the data on PAR-2 as a potential natural substrate of NSP4.

3.4.1 PAR-2 FRET

PAR-2 is activated by proteolytic cleavage and the amino acids at the cleavage site fit very well to the specificity profile of NSP4. Initially, I tested the cleavage of PAR-2 by NSP4 using a FRET substrate derived from the activation cleavage site of PAR-2 (Abz-GSKGR-SLIG-Y (NO_2)-D). Figure 3.17 depicts the enzymatic activity of NSP4 towards the PAR-2 peptide substrate. S-NSP4 (unconverted pro-form) and enterokinase did not cleave the PAR-2 FRET.

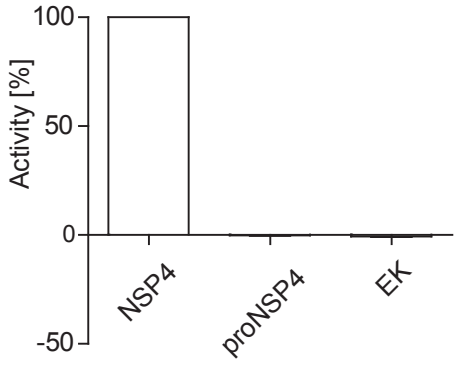


Figure 3.17: **Enzymatic activity of NSP4 hydrolyzing the PAR-2 peptide substrate.** The FRET substrate Abz-GSKGR-SLIG-Y (NO_2)-D is derived from the activation cleavage site of PAR-2 and was efficiently cleaved by NSP4, but not by its proform (S-NSP4) or its convertase (EK, enterokinase). The FRET substrate was present at 200 μM, NSP4 or proNSP4 at 2.4 μM and EK at 1.2×10^{-3} U/μl (n = 2, ± SD).

3.4.2 NSP4 triggered calcium signaling in HaCaT, but not in HT1080

Activation of PAR receptors results in intracellular calcium signaling. The effect of NSP4 on calcium signaling in HaCaT and HT1080 cells was investigated together with Luisa-Astrid Fratila in the laboratory of Prof. Dr. Christian Sommerhoff (Institut für Laboratoriumsmedizin, LMU, München). Calcium signaling was monitored using Fluo-4 AM, a calcium sensitive dye. HaCaT cells are derived from keratinocytes which predominantly express PAR-2 and lower levels of PAR-1, -3 and -4 (Santulli et al., 1995), whereas HT1080 is a fibrosarcoma cell line that expresses higher levels of PAR-1 and lower levels of PAR-2, -3 and -4. As positive controls we used bovine trypsin, which activates mainly PAR-2 but also PAR-1, -3 and -4 (Steinhoff et al., 2005), and bovine thrombin, which activates PAR-1, -3 and -4, but not PAR-2 (Shpacovitch

et al., 2008). NSP4 addition lead to an increase in calcium levels in HaCaT cells (Figure 3.18 A), but not in HT1080 cells (Figure 3.19 A). If this process was triggered by PAR receptors, it would suggest signaling through PAR-2, as HaCaT are high and HT1080 are low in PAR-2 receptor expression. As expected, the signal generated by trypsin was higher in HaCaT (Figure 3.18 B) than in HT1080 (Figure 3.19 B), while the effect of thrombin was stronger in HT1080 (Figure 3.19 C) than in HaCaT (Figure 3.18 C). The unprocessed zymogen of NSP4 (S-NSP4) was used as a negative control and had no effect on the calcium levels in HaCaT (Figure 3.18 D).

We tested 2 μM, 1 μM, 100 nM and 10 nM NSP4. Only with 2 μM and 1 μM NSP4, an increase in calcium levels could be detected. So the effect of NSP4 on HaCat was far lower compared to the effect of trypsin, which was still detectable at 1 nM concentration. As noted in section 3.3.3, the recombinantly expressed NSP4 is partly inactive. But even if we assume that only 40% of NSP4 molecules are active, the concentration of NSP4 required for calcium increase in HaCaT cells is still very high.

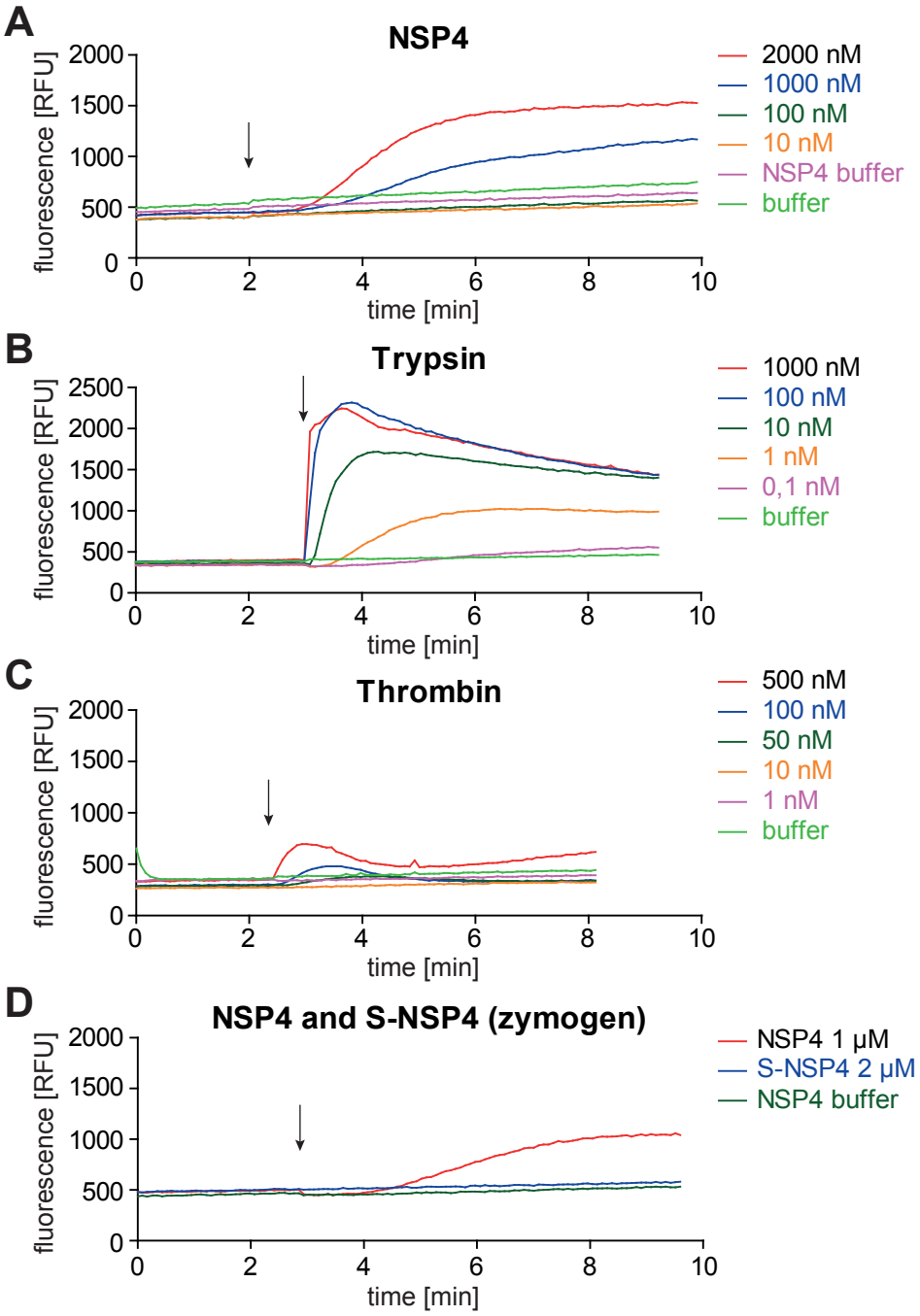


Figure 3.18: **NSP4 triggered calcium signaling in HaCaT cells.** HaCaT cells were pretreated with the calcium sensitive dye Fluo4 AM. In this way, intracellular increase in calcium levels could be monitored at λE = 494 nm and λEm = 516 nm. The cells were treated with different amounts of NSP4 (A), trypsin (B) or thrombin (C). Stimulation by the NSP4 zymogen (S-NSP4) or NSP4 buffer (EK buffer) served as controls and are shown in D. The time point of stimulation is indicated by a vertical arrow. Data are representative of n = 2 experiments that were carried out together with Luisa-Astrid Fratila (Institut für Laboratoriumsmedizin, LMU, München).

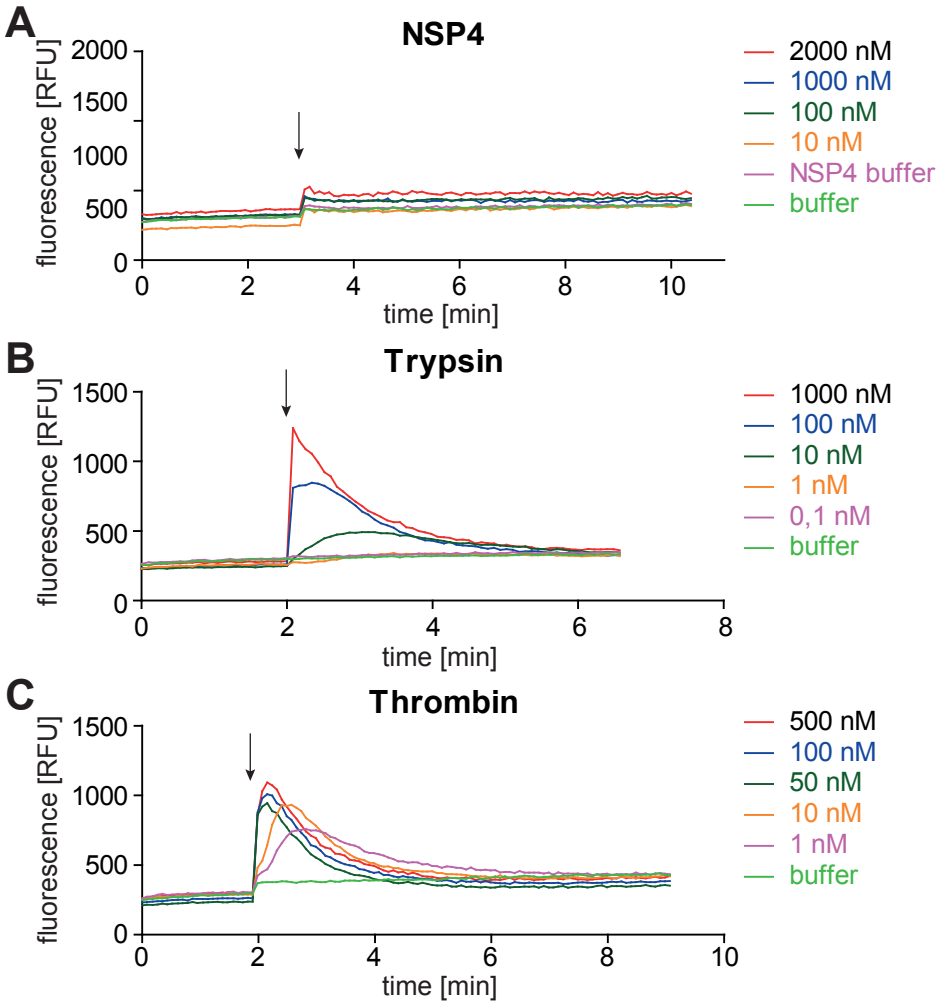


Figure 3.19: **Calcium levels in HT1080 cells after treatment with NSP4, trypsin or thrombin.** HT1080 cells were pretreated with the calcium sensitive dye Fluo4 AM. In this way, intracellular increase in calcium levels could be monitored at λE = 494 nm and λEm = 516 nm. The cells were treated with different amounts of NSP4 (A), trypsin (B) and thrombin (C). The time point of stimulation is indicated by an arrow. Data are representative of n = 2 experiments that were carried out together with Luisa-Astrid Fratila (Institut für Laboratoriumsmedizin, LMU, München).

3.4.3 PAR-2 but not PAR-1 agonists prevented NSP4 induced calcium signaling in HaCaT cells

Although the NSP4 concentration required for calcium increase in HaCaT cells was very high compared to trypsin, I was interested whether the calcium increase was triggered by PAR-2 activation. Many external stimuli can lead to calcium fluxes across cellular membranes. To this

end, Luisa and I analyzed the effect of NSP4 on calcium signaling in HaCaT after desensitization with PAR-1 and PAR-2 specific agonist peptides. These peptides were derived from the activating sequence of the respective PAR. The cells were first treated with the agonist peptides. This should result in PAR activation and removal of the receptor from the cell surface. Subsequent stimulation by a PAR activating protease would then result in a reduced calcium signal. In control experiments, successful desensitization was examined using trypsin. Pretreatment of HaCaT with the PAR-1 agonist did not impair the calcium increase after trypsin treatment (Figure 3.20 C). Treatment with the PAR-2 agonist before trypsin lead to 50% decrease of the calcium signal that could be generated by trypsin (Figure 3.20 D). This is in complete accordance with trypsin cleaving predominantly PAR-2, but also PAR-1, -3 and -4 with lower efficacy. Prestimulation of HaCaT by the PAR-1 agonist peptide before NSP4 exposure did not reduce the calcium signal generated by NSP4 (Figure 3.20 A). However, desensitization with the PAR-2 agonist completely prevented calcium signaling in response to stimulation with NSP4 (Figure 3.20 B).

This data strongly indicates that the calcium increase in HaCaT cells in response to NSP4 was due to PAR-2 activation by this novel neutrophil serine protease. Nevertheless, the relevance of NSP4 activating PAR-2 *in vivo* is questionable because the calcium increase in HaCaT cells was only triggered by NSP4 at rather high concentrations (1 μM).

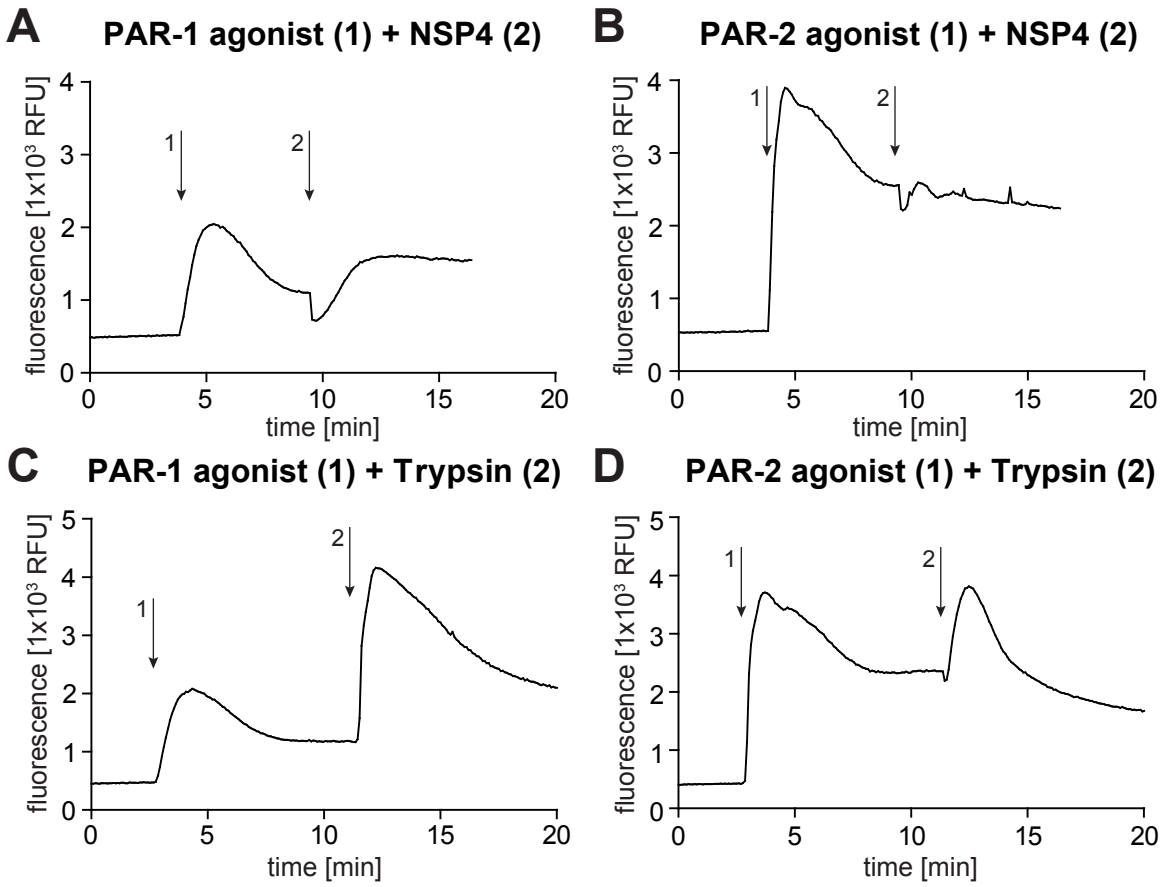


Figure 3.20: **Calcium signaling in HaCaT in response to NSP4 or trypsin after pretreatment with PAR-1 or 2 agonist peptides.** After pretreatment of HaCaT with Fluo4 AM, intracellular calcium signaling was followed at λEx = 494 nm and λEm = 516 nm. First, cells were desensitized with PAR-1 activating peptides (H-Thr-Phe-Leu-Leu-Arg-NH2) (A, C) or Par-2 activating peptides (H-Ser-Leu-Ile-Gly-Lys-Val-NH2) (B, D) at concentrations of 200 μM and 50 μM, respectively (arrow 1) and then treated with 1 μM NSP4 (A, B) or 0.8 μM trypsin (C, D) for comparison (arrow 2). Data are representative of n = 2 experiments that were carried out together with Luisa-Astrid Fratila (Institut für Laboratoriumsmedizin, LMU, München).

3.5 NSP4 in the mouse

The mouse homolog of NSP4 (mNSP4) (gene symbol Prss57) was studied at the protein level by expressing the mouse cDNA and its potential role *in vivo* was evaluated using knockout embryos that were purchased from MMRRC (UC Davis, Davis, California, USA). Firstly, I wanted to see whether mNSP4 featured a similar cleavage specificity to that of human NSP4.

The two species homologs share 77.4% sequence identity. Hence, a conserved specificity and function in human and mouse was anticipated.

3.5.1 Production of mNSP4

A precursor of mNSP4, S-mNSP4, was expressed in HEK 293E. Transfection was carried out with 90% of S-mNSP4 plasmid replaced by filler DNA (a bacterial plasmid devoid of any mammalian transcription elements) (1:10). The yield of S-mNSP4 after nickel affinity chromatography was 180 µg per 1 L expression culture. The elution fractions are shown in Figure 3.21 A. Complete conversion of S-mNSP4 to mature mNSP4 was achieved after enterokinase digestion (Figure 3.21 B).

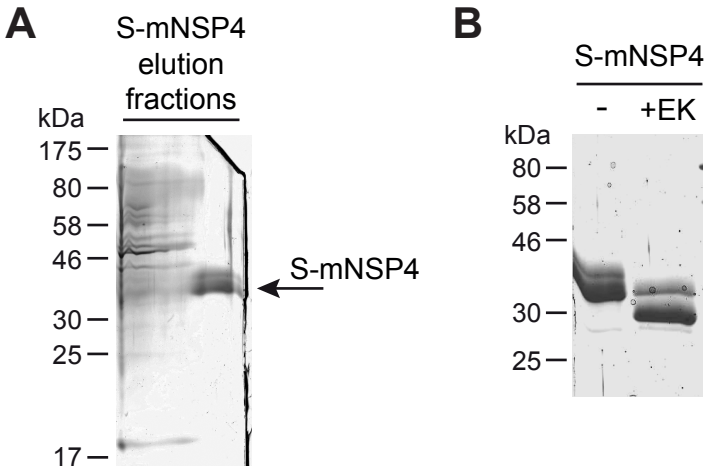


Figure 3.21: **Production of mNSP4 in HEK 293E.** A, Elution fractions after nickel-chelate chromatography of S-mNSP4 after expression in HEK 293E in a volume of 600 ml. For the transfection of cells, the plasmid encoding S-mNSP4 was reduced to 10% and filler DNA was added to 100% (1:10). A sample of 18 µl of each fraction, containing 4 ml in total, was analyzed by SDS PAGE and silver staining. S-mNSP4 is represented by two bands which are probably the result of heterogenous glycosylation, as it features two N-glycosylation sites. B, S-mNSP4 was concentrated 25-fold and is shown before and after incubation with EK. Three µl of each sample was analyzed by SDS PAGE and Coomassie Blue staining.

3.5.2 Specificity comparison between human and mouse NSP4

The specificity of mNSP4 was analyzed in comparison to human NSP4. I used three characteristic FRET substrates that were used for profiling the P2 position of NSP4. Figure 3.22 indicates that mNSP4 also hydrolyzed substrates with arginine in the P1 position. Furthermore, mNSP4 displayed the same preference for Val over Ser and Asp in P2 like human NSP4.

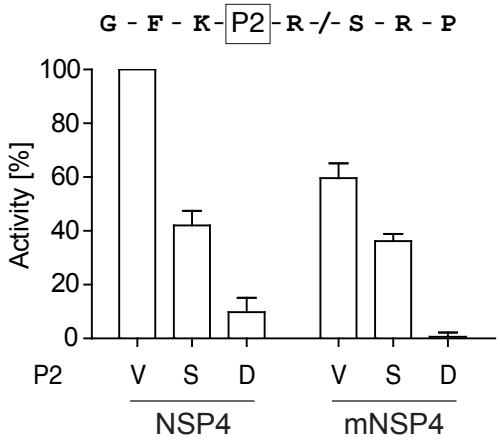


Figure 3.22: **Specificity comparison of human and mouse NSP4.** Enzymatic activity of NSP4 and mNSP4 cleaving three individual FRET substrates with Val, Ser or Asp in P2 position. NSP4 and mNSP4 were used at concentrations of 25 nM and the FRET substrates were present at 1 μM. Fluorescent signals were monitored at λEx = 320 nm and λEm = 405 nm. Results were normalized to the signal of the substrate cleaved best by NSP4 and are given in percent (n = 3, ± SD).

3.5.3 Analysis of mNSP4 mRNA in WT and NSP4-/- mice

The NSP4 knockout mouse model was obtained as cryopreserved embryos that were either heterozygous for the NSP4 knockout or homozygous for the wildtype (WT) gene. The mice were genotyped by PCR and bred to homozygosity in order to establish an NSP4-deficient mouse colony. Non-transgenic littermates were used as WT control mice.

The rat mAb for human NSP4 did not cross-react with mNSP4. Therefore, the knockout of NSP4 in the mouse line was confirmed by RT-PCR. The transcript of mNSP4 was detected in bone marrow cells of WT animals but not in NSP4-/- mice. β-actin served as a control and

its transcript could be verified in both WT and NSP4-/- mice (Figure 3.23). The expression of NSP4 in mouse bone marrow once again suggests functional conservation in mouse and man.

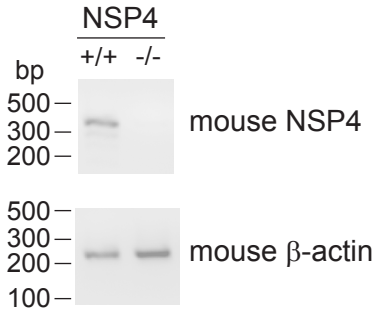


Figure 3.23: **RT-PCR analysis of mNSP4 in WT and knockout mice.** Bone marrow of WT (+/+) or NSP4-/- mice was probed by RT-PCR for mRNA transcripts of mNSP4 or mouse β-actin. Samples are shown after 3% agarose gel electrophoresis.

3.5.4 Production of reactive oxygen species (ROS) was not impaired in NSP4-/- mice

Neutrophil function in NSP4-/- and WT mice was tested by stimulation of neutrophils with PMA and measurements of ROS (the oxidative burst). This is a frequently used readout to determine cellular activation of neutrophils. Activation by the phorbol ester PMA is receptor-independent and was shown to directly activate protein kinase C (Castagna et al., 1982), resulting in the production of ROS. To this end, mouse neutrophils were isolated from bone marrow and treated with PMA. ROS production was monitored after addition of dihydrorhodamine 6G, a dye that becomes fluorescent upon oxidation. Figure 3.24 illustrates that there was no significant difference in PMA triggered oxidative burst in NSP4-/- compared to WT neutrophils. Thus, the production of ROS by neutrophils was generally functional in NSP4-deficient mice.

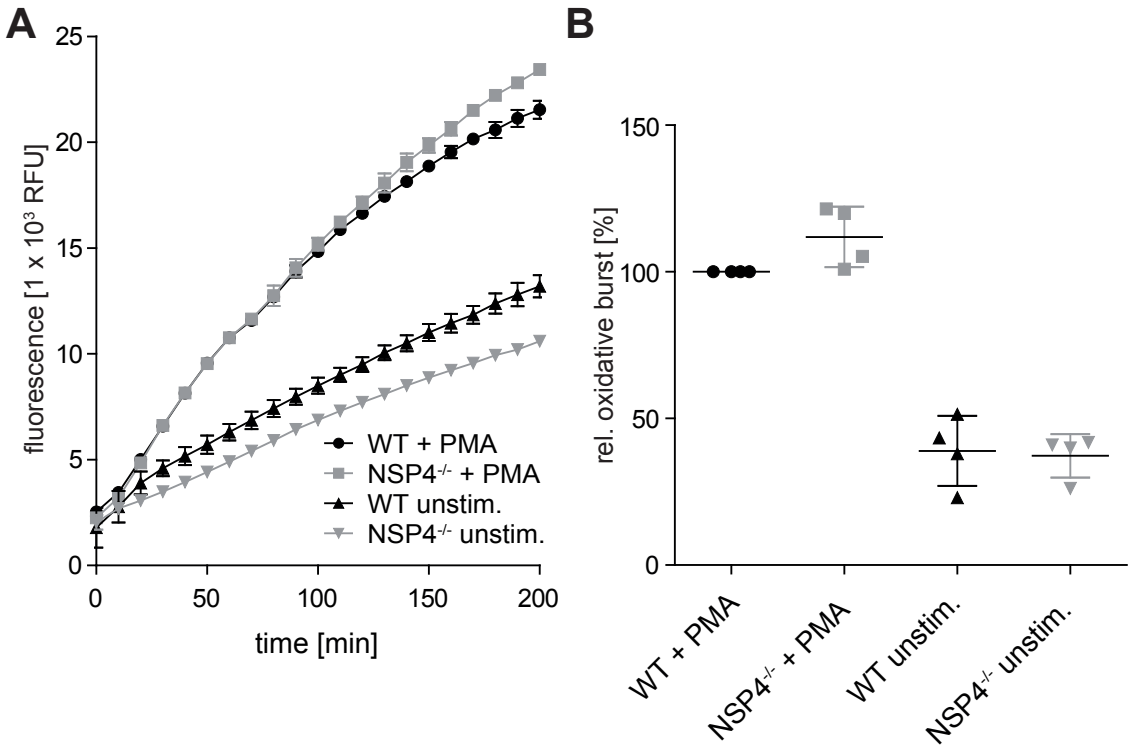


Figure 3.24: **Oxidative burst in NSP4-/- neutrophils compared to WT neutrophils in response to PMA.** Mouse PMNs were isolated from NSP4-/- or WT-mice and stimulated with PMA. Unstimulated cells were used as a control for PMA activation. As a readout for neutrophil activation, I measured neutrophil oxidative burst over time. For this, the cells were treated with dihydrorhodamine 6G, a dye that becomes fluorescent upon oxidation. The fluorescent signal was recorded at λEm = 520 nm with excitation at λEx = 485 nm. Representative Data and SD of triplicates are shown in A. Results of independent experiments were normalized to the oxidative burst in WT neutrophils (n = 4, ± SD) and are shown in B.

3.5.5 Neutrophil activation by immune complexes was not impaired in NSP4-/- mice

In many human diseases, the formation of immune complexes (ICs) is an important trigger of neutrophil activation and neutrophil-dependent inflammation (Jancar and Sánchez Crespo, 2005). Analysis of neutrophil function in double knockout mouse models of CG/NE (Raptis et al., 2005) or NE/PR3 (Kessenbrock et al., 2008) revealed impaired production of ROS upon stimulation with ICs. Activation by ICs is different from PMA activation in that it is mediated via Fcγ receptors and β_2 integrins in PMNs (Zhou and Brown, 1994). To test activation of

NSP4-/- neutrophils by ICs, ovalbumin (OVA) was immobilized on microtiter plates and treated with anti-OVA rabbit serum. For unstimulated controls, treatment with anti-OVA was omitted. Isolated neutrophils were primed with TNF-α and added to the immobilized ICs or to immobilized OVA. Treatment with TNF-α increases the abundance and activity of receptors on the plasma membrane, such as the complement receptors or Fcγ receptors. Determination of the reactive burst in neutrophils from NSP4-/- mice in comparison to WT control mice revealed no difference in IC-mediated neutrophil activation (Figure 3.25).

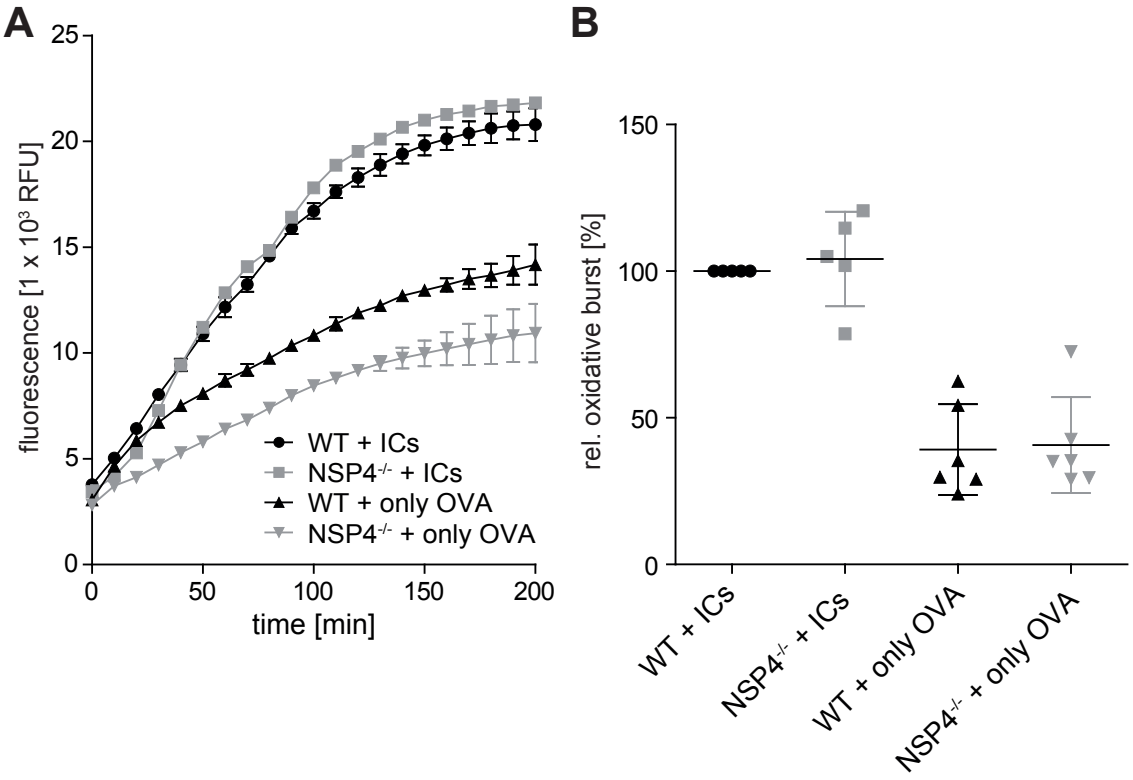


Figure 3.25: **Oxidative burst in NSP4-/- neutrophils compared to WT neutrophils in response to immune complexes (ICs).** Mouse PMNs were isolated from NSP4-/- or WT-mice and stimulated with TNF-α and ICs (ovalbumin and anti-ovalbumin). TNF-α primed PMNs were placed on immobilized ovalbumin without anti-ovalbumin antibodies to serve as a control. As a readout for neutrophil activation, I measured neutrophil oxidative burst over time. For this, the cells were treated with dihydrorhodamine 6G, a dye that becomes fluorescent upon oxidation. The fluorescent signal was recorded at λEm = 520 nm with excitation at λEx = 485 nm. Representative Data and SD of triplicates are shown in A. Results of independent experiments were normalized to the oxidative burst in WT neutrophils (n = 5, $\pm$ SD) and are shown in B.

4 Discussion

4.1 NSP4, a novel neutrophil protease highly conserved in evolution

In the present study, I identified a fourth active serine protease which is stored in azurophil granules of neutrophils, and therefore gave it the name of NSP4. NSP4 has so far been overlooked and not considered in neutrophil research. The reason for NSP4 escaping discovery for so long, despite intense investigations of NE, CG and PR3, may be found in the lower content of NSP4 in neutrophils (Figure 3.5).

The gene of NSP4 (PRSS57) was identified in the telomeric region of the short arm of human chromosome 19, in a distance of about 200 kb from the AZU1-PRTN3-ELANE-CFD gene cluster. Homologs of NSP4 are also found in all known mammals that have been sequenced. Even in birds (chicken and zebra finch) and bony fishes (Tetraodon and zebrafish), NSP4-like genes can be identified and are physically linked to the same unrelated genes as in humans (namely FGF22 and POLRMT). While the zebra finch genome contains only one copy, NSP4 was expanded in fishes and is represented by a cluster of several paralogous genes. The presence of NSP4 homologs in birds and bony fishes indicates that NSP4 preceded the emergence of NE, CG and PR3 and developed as a class 6 serine protease very early during vertebrate evolution.

4.2 NSP4, an unusual serine protease with arginine specificity

4.2.1 S1 pocket specificity

NSP4 is an arginine-specific serine protease. This was clearly demonstrated by PICS analysis (Figure 3.9A), FRET libraries (Figure 3.10) and by its inhibitor profile. NSP4 was inhibited best by classical arginine-specific protease inhibitors, like FPR-cmk (Figure 3.12), antithrombin or C1 inhibitor (Figure 3.13).

Arginine-specific serine proteases mostly carry an aspartate residue at the bottom of their S1-pocket (position 189, chymotrypsinogen numbering), as seen in trypsin or complement factor D. NSP4, however, features a glycine residue at position 189. This is also highly conserved in all NSP4 homologs from vertebrates. There is yet another critical residue determining the binding properties of the S1 pocket. As Perona et al. (1993) have demonstrated in an anionic rat trypsin D189G/G226D mutant, the combination of G189 and D226 can also cause a preference for basic amino acids in the P1 position. Not only NSP4, but also NE and PR3 feature this residue combination. In NE and PR3, however, the acidic side chain of Asp226 is shielded by two hydrophobic side chains of Val190 (NE) / Ile190 (PR3) and Val 216 (Figure 4.1). This may explain why the cleavage specificity of NSP4 is so unexpectedly unique and differs from all other NSPs. Although not very common, two other natural serine proteases with arginine/lysine specificity were described that do not occupy an aspartate residue at position 189, but at position 226. Of these two, complement factor B exhibits Asn in 189 and complement factor C2 displays a Ser189 at the bottom of the S1 pocket (Xu, 2000).

4.2.2 The extended specificity of NSP4

Positional specificity profiling of NSP4 revealed clear preferences also for P4-P2, P1' and P2' positions. While small residues like Ser or Gly were favoured in P1', FRET substrates with Val in P1' or P2' were not well accepted by NSP4 (Figure 3.14A). This suggests that NSP4 does not activate other serine protease precursors, e. g. those of the complement and coagulation

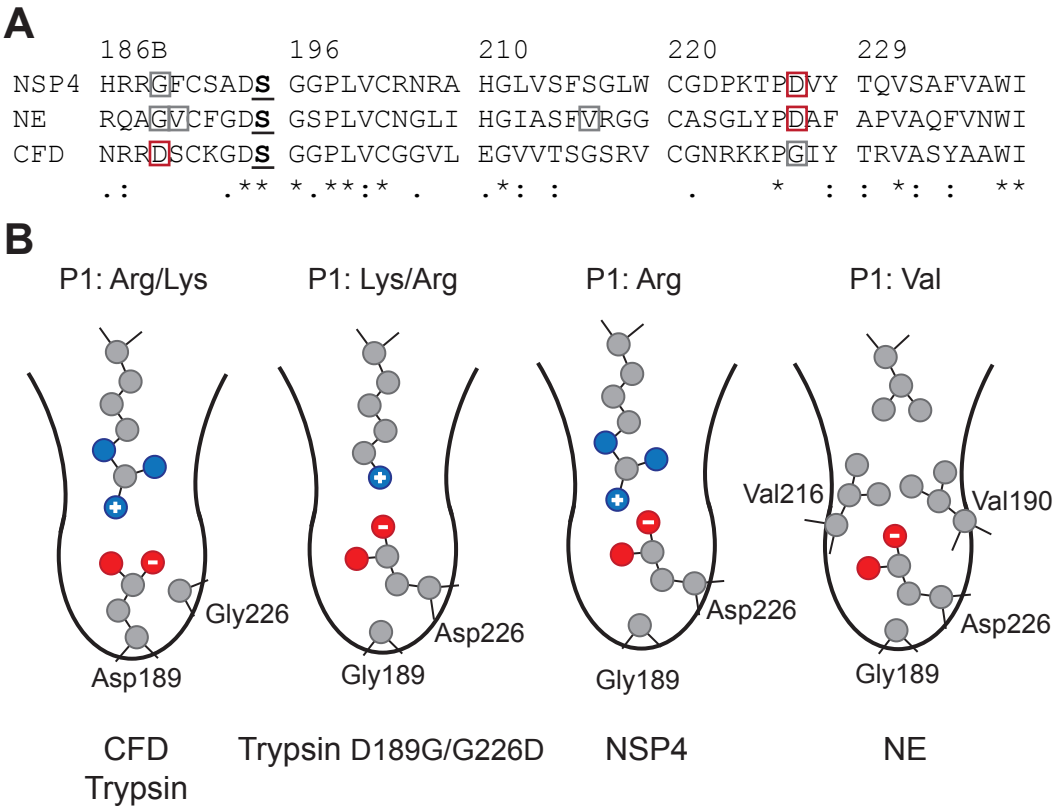


Figure 4.1: **Sequence alignment of residues around Ser195 in NSP4, NE and CFD, and schematic models of the S1 specificity pockets.** A, topologically aligned amino acid residues (single letter code) are numbered according to chymotrypsinogen. Ser195 is highlighted in bold and underlined. Residues that affect the substrate specificity in the S1 pocket are boxed and shown in the schematic model in B. An asterisk (*) below the alignments indicates positions which agree completely; a colon (:) or a period (.) indicate conservation of residues with strong or weak similarities (scoring > 0.5 or = < 0.5 in the Gonnet PAM 250 matrix), respectively. B, specificity for arginine in P1 is mostly caused by an aspartate residue at the bottom of the S1 pocket (position 189). However, as was demonstrated for the rat anionic trypsin variant R189G/G226R, also a combination of Gly189 and Asp226 can result in a preference for arginine or lysine in P1. NSP4 and NE both feature the Gly189/Asp226 combination. Nevertheless, only NSP4 can cleave after arginine residues. This is most likely caused by restriction of Asp226 by Val190 and Val216 in NE.

systems, between an Arg at the C-terminal end of a prodomain and the highly conserved N-terminus of the serine protease domain (Arg-/-(Ile/Val)-(Ile/Val)), as seen during the activation of serine protease cascades. This finding was also supported by the PICS analysis that revealed a preference for small residues in P1' like Gly or Ala.

Acidic residues close to the scissile bond in the substrate sequence seemed to prevent cleavage by NSP4, as demonstrated for Glu in P1', P2', P4 and P2 and for Asp in the P3 and P2 positions (Figure 3.14). FRET substrates with Gly in P4 or P3 were also not well cleaved by NSP4. This knowledge about preferences is highly valuable for the identification of potential natural substrates of NSP4.

4.3 Detection of active NSP4

When searching for the function of a protease, it is necessary to know when and where the respective protease is active. Storage of NSP4 as mature, active endoprotease was already suspected as it carried a two residue long propeptide. DPPI, which cleaves off dipeptides from the N-terminus and also activates NE, CG and PR3 during their biosynthesis in neutrophil precursors, was identified as the natural convertase of NSP4 (Figure 3.8). To confirm this conclusion, a highly specific and, in particular, a very sensitive detection method was developed. The detection of active NSP4 by complexation with α1PI* in Western blotting was based on both unique immunological properties and catalytic activities of NSP4. With this method, I could ascertain that NSP4 is indeed stored and released as active endoprotease by activated neutrophils (Figure 3.16E).

α1PI* can also be used to study the functional consequence of NSP4 inhibition in human neutrophils *in vitro*. Systemic applications of the serpin variant, however, may result in unpredictable off-target effects. It is anticipated that α1PI* inhibits other arginine-specific proteases, e.g. those of the coagulation and complement system. There is a naturally occurring Met358Arg α1PI variant (Pittsburgh mutation) that is known to inhibit thrombin, complement C1s, plasma kallikrein, factor XIa and Hageman factor fragment XIIf (Scott et al., 1986). Already heterozygous individuals suffer from episodic bleedings. The potential problems of a systemically

acting α1PI variant with arginine in P1 were hereby already demonstrated. Inhibition of NSP4 by the Pittsburgh mutant is conceivable and might explain its attenuating effect on NE release from activated neutrophils (Wachtfogel et al., 1994).

Concerning the function of NSP4, the storage and release by neutrophils as active endoprotease alludes to a role in the context of neutrophil immune responses. NSP4 could cleave substrates that are also stored in neutrophils or expressed by other cells that will come into close contact with neutrophils during immune responses.

4.4 PAR-2 as a potential NSP4 substrate

The effect of PAR-2 activation *in vivo* can hardly be predicted. PAR-2 is ubiquitously expressed on many different cell types. Therefore, the functions ascribed to PAR-2 are highly diverse and complex (Steinhoff et al., 2005; Chignard and Pidard, 2006). Likewise the role of PAR-2 in inflammation is controversially discussed. There is evidence for both a protective role (Cocks et al., 1999), but also for a proinflammatory role in two different mouse models of colitis (Hyun et al., 2010).

Several other proteases were already reported as activators of PAR-2, such as mast cell tryptase (Steinhoff et al., 1999), trypsin IV (Cottrell et al., 2004) or kallikrein-related peptidase 14 (Gratio et al., 2011). Matriptase was even reported as PAR-2 activator in an *in vivo* setting (Camerer et al., 2010). The authors suggested that activation of PAR-2 by matriptase triggered neural tube closure in mouse embryos. Matriptase, tryptase, trypsin IV and kallikrein related peptidase 14 could activate PAR-2 at lower concentrations (0.1 - 90 nM) than NSP4 (1 μM). In most situations it is therefore rather unlikely that PAR-2 is activated by NSP4 *in vivo*.

PAR-2 is also expressed in neutrophils and stored in azurophil granules like NSP4. In this special situation, the concentration of NSP4 could be sufficient to activate PAR-2 directly after activation of PMNs and release of the azurophil granule content. This might then lead to an enhanced motility of neutrophils, as was demonstrated after activation of PAR-2 by agonist peptides or trypsin in three-dimensional collagen lattices (Shpacovitch et al., 2004). In this

case, NSP4 would also assume an immunomodulatory function, as was proposed for NE, CG and PR3 (Meyer-Hoffert, 2009).

Activation of PAR-2 by the other NSPs has become a rather controversial issue (Chignard and Pidard, 2006). Based on the substrate specificities of NE, CG and PR3, the activating cleavage site SKGR-/-SLIGKV seems to be a very unlikely target for these three NSPs. Consistent with this, Loew et al. (2000) revealed that in a recombinant peptide sequence derived from the N-terminal, extracellular domain of PAR-2 (Arg^{31} to Pro^{79}), NE, CG and PR3 did not hydrolyze PAR-2 at the activation cleavage site, but downstream of it. The latter cleavages that occur C-terminally between the activation site and the first transmembrane domain are believed to inactivate PAR-2. Accordingly, it was also reported that NE, CG and PR3 do not activate PAR-2 but disarm the receptor, so that it cannot be activated by trypsin anymore (Dulon et al., 2003). Contrary to this, there were two reports demonstrating that NE, CG and PR3 can also activate PAR-2 (Uehara et al., 2002, 2003). Both articles were, however, retracted in 2010 (Sugawara and Muramoto, 2010; Muramoto and Sugawara, 2010). Of the four NSPs, NSP4 is the most likely protease to activate PAR-2 on neutrophils.

4.5 The NSP4-deficient mouse model

So far, no notable phenotype was discovered in the NSP4-deficient mouse model. I investigated basic neutrophil functions by determining the activation by PMA or immune complexes, but only on isolated neutrophils. It is quite possible that more complex inflammatory mechanisms are disrupted in the absence of NSP4. Further investigation of the NSP4-deficient mouse model will help in elucidating the function of this novel serine protease. Since NSP4 features a substrate specificity which is highly conserved in vertebrate evolution and differs sharply from NE, CG or PR3, a significant and relevant function for NSP4 may be postulated.

The most striking roles of NSPs were revealed using mouse models simultaneously deficient in at least two members of the NE, CG or PR3 group (Raptis et al., 2005; Kessenbrock et al., 2008). Analysis of a mouse deficient for NSP4 and another NSP should therefore also be very informative. Unfortunately, a combination of NE/PR3 and NSP4 is not feasible by just crossing

mouse lines, as the corresponding genes are located too close together (approx. 200 kb). A combination of an NSP4 knockout mice with CG-deficiency would be feasible.

4.6 The function of NSP4

Neutrophils are highly efficient killers and eaters of invading microbes. To fulfill this task, they are equipped with a large arsenal of microbicidal and digestive proteins. After microorganisms are engulfed, antimicrobial peptides and proteins are released into the resulting phagolysosomes. At the same time, the membrane bound NADPH oxidase system produces large amounts of oxygen radicals and hydrogen peroxide in a respiratory burst. These are then used by MPO to generate the antibacterial hypochlorous acid (HOCl).

The relative importance of antibacterial biologicals like NSPs, azurocidin, defensins, lysozyme and BPI, in comparison to chemical weapons produced by phagocytes is still in debate (Roos and Winterbourn, 2002; Williams, 2006). There are large species differences in the composition of granules. For example, defensins and azurocidin, an antimicrobial catalytically inactive NE homolog, are completely absent in mice. Hence, the reported importance of NSPs for bacterial defence might be species-dependent (Belaaouaj et al., 1998; Weinrauch et al., 2002).

Indeed, serine proteases with direct bactericidal activity do not appear as smart effective weapons against bacteria. This is especially true for NSP4 with its highly restricted substrate specificity. By mutating the respective target sequences, bacteria could quickly escape the attack from proteases. Also, many bacteria are well adjusted to the protease-rich milieu of the gastrointestinal tract. Moreover, microbes also produce protease inhibitors like ecotin, serpins, alpha-2-macroglobulins or trap the active site region of proteases by peptidoglycans on their surface (Mcgrath et al., 1995).

The adaptation of host-derived specificities in comparison to rapidly changing infectious challenges is rather slow. This suggests that NSP4, NE, CG and PR3 rather contribute to host defense on a regulatory indirect level, by accelerating the recruitment of phagocytes, enhancing endocytosis and amplifying the respiratory burst (Bank and Ansorge, 2001; Pham, 2006; Kessenbrock

et al., 2011). For a successful clearence of bacteria, these pathogens have to be controlled before they can multiply and spread into the body. This is achieved by a rapid mobilization of the phagocytes and clustering in high densities at infected tissue sites. Lowering the activation threshold of neutrophils is therefore more important than adding another weapon with target-restricted specialty to the neutrophil arsenal. However, this life-saving mechanism comes at a high price in special instances, such as multiple organ failure in severe sepsis and local tissue damage by dense phagocyte infiltrations due to an excess release of NSPs (Weiss, 1989).

The low abundance in neutrophils and the strict conservation of this gene in vertebrates strengthens the assumption that NSP4 is not just digesting bacterial proteins in the phagolysosome, but rather regulating neutrophil responses and innate immune reactions.

4.7 Conclusions

In summary, NSP4 can be classified as a genuine member of granule-associated, pre-activated serine proteases of neutrophils with restricted sequence specificity. NSP4 is the only arginine-specific serine protease in neutrophils. Consequently, a non-redundant role for NSP4 may be suspected, as the other NSPs cannot compete for NSP4 substrate cleavage sites. Nevertheless, a synergistic role for NSP4 in combination with one or more of the other NSPs is highly plausible. For a better understanding of the biological interplay of all four NSPs, NSP4 should be considered in future studies of neutrophil immune responses.

Under clean housing conditions, the NSP4-knockout mice did not show an abnormal phenotype. PLS patients not only lack the vast majority of NE, PR3 and CG, but also the arginine-specific NSP4 in their neutrophils. The phenotype of PLS is relatively mild. PLS patients carry only a marginally increased risk for bacterial infections. NSP4 might therefore appear rather dispensable in general host defense. However, NSP4 clearly has specific non-redundant proteolytic capacities and could fulfill a life-saving function under exceptional emergency conditions.

5 Bibliography

Please note: The number(s) after each reference stand for the number(s) of the page(s) on which this reference was cited.

Adkison, A. M., Raptis, S. Z., Kelley, D. G., and Pham, C. T. N. Dipeptidyl peptidase I activates neutrophil-derived serine proteases and regulates the development of acute experimental arthritis. *J. Clin. Invest*. 109(3):363–371. (2002).

Baggiolini, M., Bretz, U., Dewald, B., and Feigenson, M. E. The polymorphonuclear leukocyte. *Agents Actions*. 8(1-2):3–10. (1978).

Bank, U. and Ansorge, S. More than destructive: neutrophil-derived serine proteases in cytokine bioactivity control. *J. Leukoc. Biol*. 69(2):197–206. (2001).

Belaaouaj, A., McCarthy, R., Baumann, M., Gao, Z. M., Ley, T. J., Abraham, S. N., and Shapiro, S. D. Mice lacking neutrophil elastase reveal impaired host defense against gram negative bacterial sepsis. *Nat. Med*. 4(5):615–618. (1998).

Belaaouaj, A., Kim, K. S., and Shapiro, S. D. Degradation of outer membrane protein A in Escherichia coli killing by neutrophil elastase. *Science*. 289(5482):1185–1187. (2000).

Bode, W., Meyer, E., and Powers, J. C. Human leukocyte and porcine pancreatic elastase: X-ray crystal structures, mechanism, substrate specificity, and mechanism-based inhibitors. *Biochemistry*. 28(5):1951–1963. (1989).

Borregaard, N. Neutrophils, from marrow to microbes. *Immunity*. 33(5):657–670. (2010).

Camerer, E., Barker, A., Duong, D. N., Ganesan, R., Kataoka, H., Cornelissen, I., Darragh, M. R., Hussain, A., Zheng, Y.-W., Srinivasan, Y., Brown, C., Xu, S.-M., Regard, J. B., Lin, C.-Y., Craik, C. S., Kirchhofer, D., and Coughlin, S. R. Local protease signaling contributes to neural tube closure in the mouse embryo. *Dev. Cell*. 18(1):25–38. (2010).

Campbell, E. J., Silverman, E. K., and Campbell, M. A. Elastase and cathepsin G of human monocytes. Quantification of cellular content, release in response to stimuli, and heterogeneity in elastase-mediated proteolytic activity. *J. Immunol*. 143(9):2961–2968. (1989).

Castagna, M., Takai, Y., Kaibuchi, K., Sano, K., Kikkawa, U., and Nishizuka, Y. Direct activation of calcium-activated, phospholipid-dependent protein kinase by tumor-promoting phorbol esters. *J. Biol. Chem*. 257(13):7847–7851. (1982).

Caughey, G. H., Schaumberg, T. H., Zerweck, E. H., Butterfield, J. H., Hanson, R. D., Silverman, G. A., and Ley, T. J. The human mast cell chymase gene (CMA1): mapping to the cathepsin G/granzyme gene cluster and lineage-restricted expression. *Genomics*. 15(3): 614–620. (1993).

Chignard, M. and Pidard, D. Neutrophil and pathogen proteinases versus proteinase-activated receptor-2 lung epithelial cells: more terminators than activators. *Am. J. Respir. Cell Mol. Biol*. 34(4):394–398. (2006).

Chin, A. C., Lee, W. Y., Nusrat, A., Vergnolle, N., and Parkos, C. Neutrophil-mediated activation of epithelial protease-activated receptors-1 and -2 regulates barrier function and transepithelial migration. *J. Immunol*. 181(8):5702–5710. (2008).

Chua, F., Dunsmore, S. E., Clingen, P. H., Mutsaers, S. E., Shapiro, S. D., Segal, A. W., Roes, J., and Laurent, G. J. Mice lacking neutrophil elastase are resistant to bleomycin-induced pulmonary fibrosis. *Am. J. Pathol*. 170(1):65–74. (2007).

Cocks, T. M., Fong, B., Chow, J. M., Anderson, G. P., Frauman, A. G., Goldie, R. G., Henry, P. J., Carr, M. J., Hamilton, J. R., and Moffatt, J. D. A protective role for protease-activated receptors in the airways. *Nature*. 398(6723):156–160. (1999).

Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., and Gevaert, K. Improved visualization of protein consensus sequences by iceLogo. *Nat. Methods*. 6(11):786–787. (2009).

Cottrell, G. S., Amadesi, S., Grady, E. F., and Bunnett, N. W. Trypsin IV, a novel agonist of protease-activated receptors 2 and 4. *J. Biol. Chem*. 279(14):13532–13539. (2004).

Dulon, S., Candé, C., Bunnett, N. W., Hollenberg, M. D., Chignard, M., and Pidard, D. Proteinase-activated receptor-2 and human lung epithelial cells: disarming by neutrophil serine proteinases. *Am. J. Respir. Cell Mol. Biol*. 28(3):339–346. (2003).

Durocher, Y., Perret, S., and Kamen, A. High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. *Nucleic Acids Res*. 30(2):E9. (2002).

Faurschou, M. and Borregaard, N. Neutrophil granules and secretory vesicles in inflammation. *Microbes Infect*. 5(14):1317–1327. (2003).

Freer, S. T., Kraut, J., Robertus, J. D., Wright, H. T., and Xuong, N. H. Chymotrypsinogen: 2.5-angstrom crystal structure, comparison with alpha-chymotrypsin, and implications for zymogen activation. *Biochemistry*. 9(9):1997–2009. (1970).

Gorlin, R. J., Sedano, H., and Anderson, V. E. The syndrome of palmar-plantar hyperkeratosis and premature periodontal destruction of the teeth. A clinical and genetic analysis of the Papillon-Lefèvre syndrome. *J. Pediatr*. 65:895–908. (1964).

Gratio, V., Loriot, C., Virca, G. D., Oikonomopoulou, K., Walker, F., Diamandis, E. P., Hollenberg, M. D., and Darmoul, D. Kallikrein-related peptidase 14 acts on proteinase-activated receptor 2 to induce signaling pathway in colon cancer cells. *Am. J. Pathol*. 179(5):2625–2636. (2011).

Grimsley, G. R. and Pace, C. N. Spectrophotometric determination of protein concentration. *Curr. Protoc. Protein Sci*. Chapter 3:Unit 3.1. (2004).

Groutas, W. C., Dou, D., and Alliston, K. R. Neutrophil elastase inhibitors. *Expert Opin. Ther. Pat*. 21(3):339–354. (2011).

Hortin, G. L., Sviridov, D., and Anderson, N. L. High-abundance polypeptides of the human plasma proteome comprising the top 4 logs of polypeptide abundance. *Clin. Chem*. 54(10): 1608–1616. (2008).

Horwitz, M., Benson, K. F., Duan, Z., Li, F.-Q., and Person, R. E. Hereditary neutropenia: dogs explain human neutrophil elastase mutations. *Trends Mol. Med*. 10(4):163–170. (2004).

Huntington, J. A., Read, R. J., and Carrell, R. W. Structure of a serpin-protease complex shows inhibition by deformation. *Nature*. 407(6806):923–926. (2000).

Hyun, E., Andrade-Gordon, P., Steinhoff, M., Beck, P. L., and Vergnolle, N. Contribution of bone marrow-derived cells to the pro-inflammatory effects of protease-activated receptor-2 in colitis. *Inflamm. Res*. 59(9):699–709. (2010).

Ishihara, H., Connolly, A. J., Zeng, D., Kahn, M. L., Zheng, Y. W., Timmons, C., Tram, T., and Coughlin, S. R. Protease-activated receptor 3 is a second thrombin receptor in humans. *Nature*. 386(6624):502–506. (1997).

Jancar, S. and Sánchez Crespo, M. Immune complex-mediated tissue injury: a multistep paradigm. *Trends Immunol*. 26(1):48–55. (2005).

Jenne, D. E. Structure of the azurocidin, proteinase-3, and neutrophil elastase genes - implications for inflammation and vasculitis. *Am. J. Respir. Crit. Care Med*. 150(6):S147–S154. (1994).

Jochum, M., Lander, S., Heimburger, N., and Fritz, H. Effect of human granulocytic elastase on isolated human antithrombin III. *Hoppe Seylers Z. Physiol. Chem*. 362(2):103–112. (1981).

Jordan, R. E., Kilpatrick, J., and Nelson, R. M. Heparin promotes the inactivation of antithrombin by neutrophil elastase. *Science*. 237(4816):777–779. (1987).

Kahn, M. L., Zheng, Y. W., Huang, W., Bigornia, V., Zeng, D., Moff, S., Farese, R. V., Tam, C., and Coughlin, S. R. A dual thrombin receptor system for platelet activation. *Nature*. 394 (6694):690–694. (1998).

Kessenbrock, K., Fröhlich, L., Sixt, M., Lämmermann, T., Pfister, H., Bateman, A., Belaaouaj, A., Ring, J., Ollert, M., Fässler, R., and Jenne, D. E. Proteinase 3 and neutrophil elastase enhance inflammation in mice by inactivating antiinflammatory progranulin. *J. Clin. Invest*. 118(7):2438–2447. (2008).

Kessenbrock, K., Dau, T., and Jenne, D. E. Tailor-made inflammation: how neutrophil serine proteases modulate the inflammatory response. *J. Mol. Med*. 89(1):23–28. (2011).

Köllner, I., Sodeik, B., Schreek, S., Heyn, H., von Neuhoff, N., Germeshausen, M., Zeidler, C., Krüger, M., Schlegelberger, B., Welte, K., and Beger, C. Mutations in neutrophil elastase causing congenital neutropenia lead to cytoplasmic protein accumulation and induction of the unfolded protein response. *Blood*. 108(2):493–500. (2006).

Korkmaz, B., Moreau, T., and Gauthier, F. Neutrophil elastase, proteinase 3 and cathepsin G: physicochemical properties, activity and physiopathological functions. *Biochimie*. 90(2): 227–242. (2008).

Korkmaz, B., Horwitz, M. S., Jenne, D. E., and Gauthier, F. Neutrophil elastase, proteinase 3, and cathepsin G as therapeutic targets in human diseases. *Pharmacol. Rev*. 62(4):726–759. (2010).

Le Cabec, V., Cowland, J. B., Calafat, J., and Borregaard, N. Targeting of proteins to granule subsets is determined by timing and not by sorting: the specific granule protein NGAL is localized to azurophil granules when expressed in HL-60 cells. *Proc. Natl. Acad. Sci. U S A*. 93(13):6454–6457. (1996).

Liu, C.-H. and Wu, P.-S. Characterization of matrix metalloproteinase expressed by human embryonic kidney cells. *Biotechnol. Lett*. 28(21):1725–30. (2006).

Liu, Y., Nguyen, A., Wolfert, R. L., and Zhuo, S. Enhancing the secretion of recombinant proteins by engineering N-glycosylation sites. *Biotechnol. Prog*. 25(5):1468–1475. (2009).

Liu, Z., Zhou, X., Shapiro, S. D., Shipley, J. M., Twining, S. S., Diaz, L. A., Senior, R. M., and Werb, Z. The serpin alpha1-proteinase inhibitor is a critical substrate for gelatinase B/MMP-9 in vivo. *Cell*. 102(5):647–655. (2000).

Loew, D., Perrault, C., Morales, M., Moog, S., Ravanat, C., Schuhler, S., Arcone, R., Pietropaolo, C., Cazenave, J. P., van Dorsselaer, A., and Lanza, F. Proteolysis of the exodomain of recombinant protease-activated receptors: prediction of receptor activation or inactivation by MALDI mass spectrometry. *Biochemistry*. 39(35):10812–10822. (2000).

Lucas, S. D., Costa, E., Guedes, R. C., and Moreira, R. Targeting COPD: Advances on low-molecular-weight inhibitors of human neutrophil elastase. *Med. Res. Rev*. 33 Suppl 1:E73–E101. (2013).

Mast, A. E., Enghild, J. J., Nagase, H., Suzuki, K., Pizzo, S. V., and Salvesen, G. Kinetics and physiologic relevance of the inactivation of alpha 1-proteinase inhibitor, alpha 1-antichymotrypsin, and antithrombin III by matrix metalloproteinases-1 (tissue collagenase), -2 (72-kDa gelatinase/type IV collagenase), and -3 (stromelysin). *J. Biol. Chem*. 266(24): 15810–15816. (1991).

Mcgrath, M. E., Gillmor, S. A., and Fletterick, R. J. Ecotin: lessons on survival in a protease-filled world. *Protein Sci*. 4(2):141–148. (1995).

Meyer-Hoffert, U. Neutrophil-derived serine proteases modulate innate immune responses. *Front Biosci (Landmark Ed.)*. 14(5):3409–3418. (2009).

Muramoto, K. and Sugawara, S. Retraction: Neutrophil serine proteinases activate human nonepithelial cells to produce inflammatory cytokines through protease-activated receptor 2. *J. Immunol*. 184(7):4043. (2010).

Nathan, C. Neutrophils and immunity: challenges and opportunities. *Nat. Rev. Immunol*. 6(3): 173–182. (2006).

Newburger, P. E. Disorders of neutrophil number and function. *Hematology Am. Soc. Hematol. Educ. Program*. pages 104–110. (2006).

Noack, B., Görgens, H., Schacher, B., Puklo, M., Eickholz, P., Hoffmann, T., and Schackert, H. K. Functional Cathepsin C mutations cause different Papillon-Lefèvre syndrome phenotypes. *J. Clin. Periodontol*. 35(4):311–316. (2008).

Nystedt, S., Emilsson, K., Larsson, A. K., Strömbeck, B., and Sundelin, J. Molecular cloning and functional expression of the gene encoding the human proteinase-activated receptor 2. *Eur. J. Biochem*. 232(1):84–89. (1995).

Olson, S. T. and Gettins, P. G. W. Regulation of proteases by protein inhibitors of the serpin superfamily. *Prog. Mol. Biol. Transl. Sci*. 99:185–240. (2011).

Owen, C. A. and Campbell, E. J. The cell biology of leukocyte-mediated proteolysis. *J. Leukoc. Biol*. 65(2):137–150. (1999).

Page, M. J. and Di Cera, E. Serine peptidases: classification, structure and function. *Cell. Mol. Life Sci*. 65(7-8):1220–1236. (2008).

Perona, J. J., Tsu, C. A., Mcgrath, M. E., Craik, C. S., and Fletterick, R. J. Relocating a negative charge in the binding pocket of trypsin. *J. Mol. Biol*. 230(3):934–949. (1993).

Pham, C. T. N. Neutrophil serine proteases: specific regulators of inflammation. *Nat. Rev. Immunol*. 6(7):541–550. (2006).

Pham, C. T. N. Neutrophil serine proteases fine-tune the inflammatory response. *Int. J. Biochem. Cell Biol*. 40(6-7):1317–1333. (Jan. 2008).

Pham, C. T. N., Ivanovich, J. L., Raptis, S. Z., Zehnbauer, B., and Ley, T. J. Papillon-Lefèvre syndrome: correlating the molecular, cellular, and clinical consequences of cathepsin C/dipeptidyl peptidase I deficiency in humans. *J. Immunol*. 173(12):7277–7281. (2004).

Ramachandran, R., Mihara, K., Chung, H., Renaux, B., Lau, C. S., Muruve, D. A., DeFea, K. A., Bouvier, M., and Hollenberg, M. D. Neutrophil elastase acts as a biased agonist for proteinase-activated receptor-2 (PAR2). *J. Biol. Chem*. 286(28):24638–24348. (2011).

Rao, N. V., Wehner, N. G., Marshall, B. C., Gray, W. R., Gray, B. H., and Hoidal, J. R. Characterization of proteinase-3 (PR-3), a neutrophil serine proteinase. Structural and functional properties. *J. Biol. Chem*. 266(15):9540–9548. (1991).

Raptis, S. Z., Shapiro, S. D., Simmons, P. M., Cheng, A. M., and Pham, C. T. N. Serine protease cathepsin G regulates adhesion-dependent neutrophil effector functions by modulating integrin clustering. *Immunity*. 22(6):679–691. (2005).

Raymond, W. W., Trivedi, N. N., Makarova, A., Ray, M., Craik, C. S., and Caughey, G. H. How immune peptidases change specificity: cathepsin G gained tryptic function but lost efficiency during primate evolution. *J. Immunol*. 185(9):5360–5368. (2010).

Reeves, E. P., Lu, H., Lortat-Jacob, H., Messina, C. G. M., Bolsover, S., Gabella, G., Potma, E. O., Warley, A., Roes, J., and Segal, A. W. Killing activity of neutrophils is mediated through activation of proteases by K(+) flux. *Nature*. 416(6878):291–297. (2002).

Reisman, D. and Sugden, B. trans activation of an Epstein-Barr viral transcriptional enhancer by the Epstein-Barr viral nuclear antigen 1. *Mol. Cel. Biol*. 6(11):3838–3846. (1986).

Roos, D. and Winterbourn, C. C. Immunology. Lethal weapons. *Science*. 296(5568):669–671. (2002).

Sambrano, G. R., Huang, W., Faruqi, T., Mahrus, S., Craik, C., and Coughlin, S. R. Cathepsin G activates protease-activated receptor-4 in human platelets. *J. Biol. Chem*. 275(10):6819–6823. (2000).

Santulli, R. J., Derian, C. K., Darrow, A. L., Tomko, K. A., Eckardt, A. J., Seiberg, M., Scarborough, R. M., and Andrade-Gordon, P. Evidence for the presence of a protease-activated receptor distinct from the thrombin receptor in human keratinocytes. *Proc. Natl. Acad. Sci. U S A*. 92(20):9151–9155. (1995).

Schechter, I. and Berger, A. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun*. 27(2):157–162. (1967).

Schilling, O. and Overall, C. M. Proteome-derived, database-searchable peptide libraries for identifying protease cleavage sites. *Nat. Biotechnol*. 26(6):685–694. (2008).

Scott, C. F., Carrell, R. W., Glaser, C. B., Kueppers, F., Lewis, J. H., and Colman, R. W. Alpha-1-antitrypsin-Pittsburgh. A potent inhibitor of human plasma factor XIa, kallikrein, and factor XIIf. *J. Clin. Invest*. 77(2):631–634. (1986).

Shapiro, S. D. Immunology: Mobilizing the army. *Nature*. 421(6920):223–224. (2003).

Shimoda, N., Fukazawa, N., Nonomura, K., and Fairchild, R. L. Cathepsin g is required for sustained inflammation and tissue injury after reperfusion of ischemic kidneys. *Am. J. Pathol*. 170(3):930–940. (2007).

Shpacovitch, V., Feld, M., Hollenberg, M. D., Luger, T., and Steinhoff, M. Role of protease-activated receptors in inflammatory responses, innate and adaptive immunity. *J. Leukoc. Biol*. 83(6):1309–1322. (2008).

Shpacovitch, V. M., Varga, G., Strey, A., Gunzer, M., Mooren, F., Buddenkotte, J., Vergnolle, N., Sommerhoff, C. P., Grabbe, S., Gerke, V., Homey, B., Hollenberg, M., Luger, T. A., and Steinhoff, M. Agonists of proteinase-activated receptor-2 modulate human neutrophil cytokine secretion, expression of cell adhesion molecules, and migration within 3-D collagen lattices. *J. Leukoc. Biol*. 76(2):388–398. (2004).

Stavridi, E. S., O'Malley, K., Lukacs, C. M., Moore, W. T., Lambris, J. D., Christianson, D. W., Rubin, H., and Cooperman, B. S. Structural change in alpha-chymotrypsin induced by complexation with alpha 1-antichymotrypsin as seen by enhanced sensitivity to proteolysis. *Biochemistry*. 35(33):10608–10615. (1996).

Steinhoff, M., Corvera, C. U., Thoma, M. S., Kong, W., McAlpine, B. E., Caughey, G. H., Ansel, J. C., and Bunnett, N. W. Proteinase-activated receptor-2 in human skin: tissue distribution and activation of keratinocytes by mast cell tryptase. *Exp. Dermatol*. 8(4):282–294. (1999).

Steinhoff, M., Buddenkotte, J., Shpacovitch, V., Rattenholl, A., Moormann, C., Vergnolle, N., Luger, T. A., and Hollenberg, M. D. Proteinase-activated receptors: transducers of proteinase-mediated signaling in inflammation and immune response. *Endocr. Rev*. 26(1): 1–43. (2005).

Sugawara, S. and Muramoto, K. Retraction: Activation of human oral epithelial cells by neutrophil proteinase 3 through protease-activated receptor-2. *J. Immunol*. 184(7):4042. (2010).

Tang, T., Li, L., Tang, J., Li, Y., Lin, W. Y., Martin, F., Grant, D., Solloway, M., Parker, L., Ye, W. L., Forrest, W., Ghilardi, N., Oravecz, T., Platt, K. A., Rice, D. S., Hansen, G. M., Abuin, A., Eberhart, D. E., Godowski, P., Holt, K. H., Peterson, A., Zambrowicz, B. P., and de Sauvage, F. J. A mouse knockout library for secreted and transmembrane proteins. *Nat. Biotechnol*. 28(7):749–755. (2010).

Theilgaard-Mönch, K., Jacobsen, L. C., Nielsen, M. J., Rasmussen, T., Udby, L., Gharib, M., Arkwright, P. D., Gombart, A. F., Calafat, J., Moestrup, S. r. K., Porse, B. T., and Borregaard, N. Haptoglobin is synthesized during granulocyte differentiation, stored in specific granules, and released by neutrophils in response to activation. *Blood*. 108(1):353–361. (2006).

Tkalcevic, J., Novelli, M., Phylactides, M., Iredale, J. P., Segal, A. W., and Roes, J. Impaired immunity and enhanced resistance to endotoxin in the absence of neutrophil elastase and cathepsin G. *Immunity*. 12(2):201–210. (2000).

Udby, L. and Borregaard, N. Subcellular fractionation of human neutrophils and analysis of subcellular markers. *Methods Mol. Biol*. 412:35–56. (2007).

Uehara, A., Sugawara, S., Muramoto, K., and Takada, H. Activation of human oral epithelial cells by neutrophil proteinase 3 through protease-activated receptor-2. *J. Immunol. (RETRACTED)*. 169(8):4594–603. (2002).

Uehara, A., Muramoto, K., Takada, H., and Sugawara, S. Neutrophil serine proteinases activate human nonepithelial cells to produce inflammatory cytokines through protease-activated receptor 2. *J. Immunol. (RETRACTED)*. 170(11):5690–6. (2003).

Vu, T. K., Hung, D. T., Wheaton, V. I., and Coughlin, S. R. Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation. *Cell*. 64(6): 1057–1068. (1991).

Wachtfogel, Y. T., Bischoff, R., Bauer, R., Hack, C. E., Nuijens, J. H., Kucich, U., Niewiarowski, S., Edmunds, L. H., and Colman, R. W. Alpha 1-antitrypsin Pittsburgh (Met358–>Arg) inhibits the contact pathway of intrinsic coagulation and alters the release of human neutrophil elastase during simulated extracorporeal circulation. *Thromb. Haemost*. 72(6):843–847. (1994).

Watorek, W., van Halbeek, H., and Travis, J. The isoforms of human neutrophil elastase and cathepsin G differ in their carbohydrate side chain structures. *Biol. Chem. Hoppe Seyler*. 374 (6):385–393. (1993).

Weinrauch, Y., Drujan, D., Shapiro, S. D., Weiss, J., and Zychlinsky, A. Neutrophil elastase targets virulence factors of enterobacteria. *Nature*. 417(6884):91–94. (2002).

Weiss, S. J. Tissue destruction by neutrophils. *N. Engl. J. Med*. 320(6):365–376. (1989).

Wiedow, O. and Meyer-Hoffert, U. Neutrophil serine proteases: potential key regulators of cell signalling during inflammation. *J. Intern. Med.* 257(4):319–328. (2005).

Williams, R. Killing controversy. *J. Exp. Med.* 203(11):2404. (2006).

Wysocka, M., Logowska, A., Bulak, E., Jaskiewicz, A., Miecznikowska, H., Lesner, A., and Rolka, K. New chromogenic substrates of human neutrophil cathepsin G containing non-natural aromatic amino acid residues in position P1 selected by combinatorial chemistry methods. *Mol. Divers.* 11(2):93–99. (2007).

Xu, Y. Mutational Analysis of the Primary Substrate Specificity Pocket of Complement Factor B. Asp(226) is a major structural determinant for p(1)-Arg binding. *J. Biol. Chem.* 275(1): 378–385. (2000).

Yurewicz, E. C. and Zimmerman, M. Cytochalasin B-dependent release of azurophil granule enzymes from human polymorphonuclear leukocytes. *Inflammation.* 2(4):259–264. (1977).

Zarbock, A. and Ley, K. Mechanisms and consequences of neutrophil interaction with the endothelium. *Am. J. Pathol.* 172(1):1–7. (2008).

Zhou, M. J. and Brown, E. J. CR3 (Mac-1, alpha M beta 2, CD11b/CD18) and Fc gamma RIII cooperate in generation of a neutrophil respiratory burst: requirement for Fc gamma RIII and tyrosine phosphorylation. *J. Cell Biol.* 125(6):1407–1416. (1994).

Zimmer, M., Medcalf, R. L., Fink, T. M., Mattmann, C., Lichter, P., and Jenne, D. E. Three human elastase-like genes coordinately expressed in the myelomonocyte lineage are organized as a single genetic locus on 19pter. *Proc. Natl. Acad. Sci. U S A.* 89(17):8215–8219. (1992).

6 Abbreviations

α1PI	α_1-proteinase inhibitor, α_1-antitrypsin
ACT	α_1-antichymotrypsin
AEBSF	4-(2-Aminoethyl) benzenesulfonyl fluoride
AMC	7-Amino-4-methylcoumarin
APS	Ammonium persulfate
AT	Annealing temperature
BCA	Bicinchoninic acid
BPI	Bactericidal/permeability-increasing protein
Boc	Tert-butoxycarbonyl
BSA	Bovine serum albumin
CG	Cathepsin G
CFD	Complement factor D
cmk	Chloromethyl ketone
cMyc	c-Myc epitope tag
CytB	Cytochalasin B
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
Dnp	2,4-Dinitrophenyl
dNTP	Deoxyribonucleotide triphosphate
DPPI	Dipeptidyl peptidase I
DTE	Dithioerythritol
DTT	Dithiothreitol
DTNB	5,5'-Dithiobis-(2-nitrobenzoic acid)
EBNA1	Epstein-Barr nuclear antigen 1
EBV	Epstein-Barr virus
ECL	Electrochemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EK	Enterokinase
ESI-TOF	Electrospray ionization - time-of-flight
FCS	Fetal calf serum
FPLC	Fast protein liquid chromatography
FRET	Fluorescence/Förster resonance energy transfer

FRT	Flp recombination target
GP67	acidic glycoprotein of AcNPV baculovirus
H_6	6x His tag
HBSS	Hank's Balanced Salt Solution
HEK	Human embryonic kidney cell line
HLA	HLA class I antigen
HRP	Horse radish peroxidase
HSA	Human serum albumin
IC	Immune complex
lacZ	Gene encoding β-galactosidase
LAMP2A	Lysosomal-associated membrane protein 2A
LB	Luria-Bertani
LC-MS	Liquid chromatography–mass spectrometry
λEm	Emission wavelength
λEx	Excitation wavelength
LPS	Lipopolysaccharides
mAb	monoclonal antibody
Mca	7-Methoxycoumarin-4-acetic acid
MMP-9	Matrix metalloproteinase 9
MNEI	Monocyte/neutrophil elastase inhibitor
mNSP4	Mouse NSP4
MOG	Myelin oligodendrocyte glycoprotein
MPO	Myeloperoxidase
NE	Neutrophil elastase
NGAL	Neutrophil gelatinase-associated lipocalin
NGS	peptide sequence with N-glycosylation site, SRMHRNGSH
NK	Natural killer cells
Nma	2-(N-methylamino)benzoyl
NSP	Neutrophil serine protease
NSP4	Neutrophil serine protease 4
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PAR	Proteinase-activated receptor
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEI	Polyethylenimine
PICS	Proteomic identification of protease cleavage sites
PLS	Papillon-Lefèvre syndrome

PMA	Phorbol-12-myristate-13-acetate
PMN	Polymorphonuclear cells
PMSF	Phenylmethylsulfonyl fluoride
PR3	Proteinase 3
PVDF	Polyvinylidene difluoride
RCL	Reactive center loop
RFU	Relative fluorescence unit
rH	Relative humidity
ROS	Reactive oxygen species
RPMI	Roswell Park Memorial Institute medium
RT-PCR	Reverse transcription polymerase chain reaction
S	S-tag peptide
Sbzl	Thiobenzyl ester
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SigIgκ	Signal peptide derived from Igκ
siRNA	Small interfering RNAs
SLPI	Secretory leukocyte peptidase inhibitor
SN	Supernatant
TNF-α	Tumor necrosis factor α
WT	Wildtype

7 Appendix

7.1 Vector maps

7.1.1 pTT5

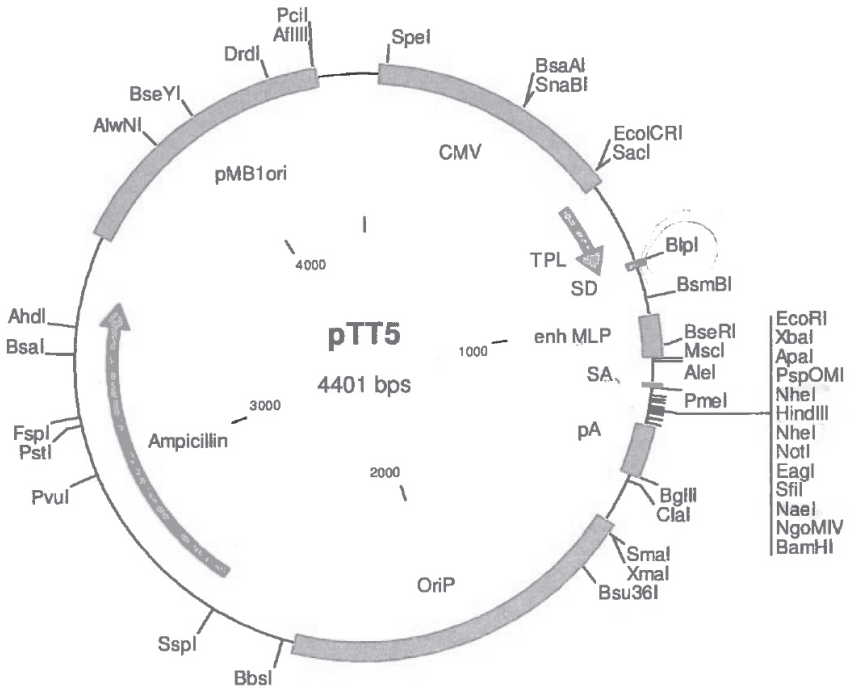
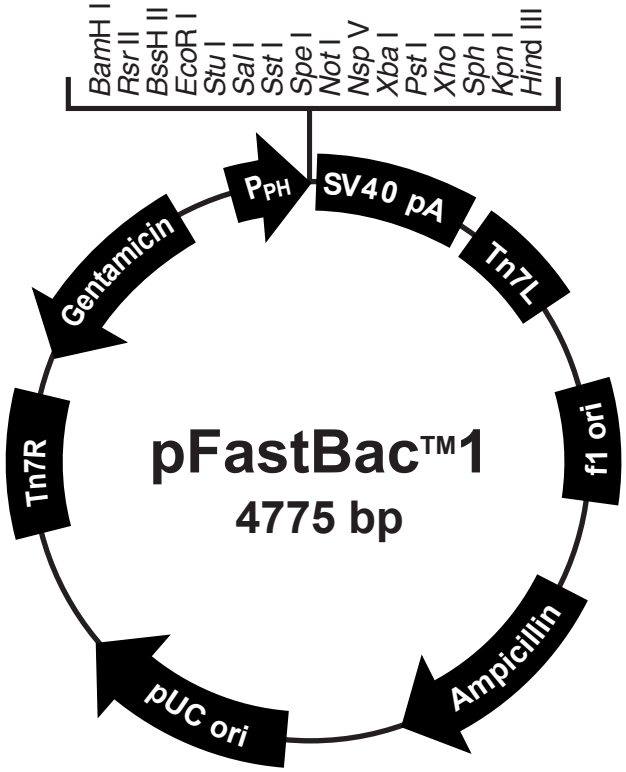


Figure 7.1: **Vector map of the pTT5 vector.** The pTT5 vector contains an EBV specific origin of replication (OriP) as well as a promoter of the cytomegalovirus (CMV), an *E. coli* origin of replication (pMB1ori), an ampicillin resistance gene, an adenovirus tripartite leader (TPL) and major late promotor (enh. MLP), splice donor (SD) and splice acceptor (SA) sites and a rabbit beta-globin polyadenylation site (pA). All the relevant recognition sites of the restriction enzymes are depicted. Plasmid provider: NRC Biotechnology Research Institute.

7.1.2 pFASTBAC1



Comments for pFastBac™1
4775 nucleotides

f1 origin: bases 2-457
Ampicillin resistance gene: bases 589-1449
pUC origin: bases 1594-2267
Tn7R: bases 2511-2735
Gentamicin resistance gene: bases 2802-3335 (complementary strand)
Polyhedrin promoter (P_{PH}): bases 3904-4032
Multiple cloning site: bases 4037-4142
SV40 polyadenylation signal: bases 4160-4400
Tn7L: bases 4429-4594

Figure 7.2: **Vector map of the pFASTBAC1 vector.** P_{PH}, Polyhedrin promoter; SV40 pA, SV40 polyadenylation signal; Tn7L and Tn7R, Mini Tn7 elements that permit site-specific transposition of the gene of interest into the baculovirus genome; f1 origin, origin of replication of f1 phage, allows rescue of single stranded DNA; Ampicillin and Gentamycin, resistence genes for selection in *E. coli*; pUC ori, origin of replication, permits high-copy replication and maintenance in *E. coli*. Source: Invitrogen.

7.2 List of primers

DJ3305	5'-Pho-CTA GCC CCA CCA TGG CTC ACC GGC CCC CCA GCC CTG CCC TGG CGT CCG TGC TGC TGG CCT TGC TGC TGA GCG GTG CTG CCC GAG CTC AC-3'
DJ3306	5'-GTG AGC TCG GGC AGC ACC GCT CAG CAG CAA GGC CAG CAG CAC GGA CGC CAG GGC AGG GCT GGG GGG CCG GTG AGC CAT GGT GGG G-3'
DJ3312	5'-CGG CGG CTA GCA AAC ATG GAG ACA GAC AGA-3'
DJ3314	5'-TGG ACA CGT GGT TGT CGT CAT CTG CAC TGC-3'
DJ3324	5'-CGG GGT ACC AGG TTC CAC TGG TGA CTC CCT GCA GGA CTC AGA A-3'
DJ3325	5'-GGG GGT CAC CTC GTG GCC CCC GAT GAT ACC TCC AAT CTG TTC TCT-3'
DJ3352	5'-TCC CTG CAG GGC GAC GAC GAC GAC AAG ATC GTG GGC GGC CAT GAG-3'
DJ3353	5'-CCC ACC GGT CAC CGA ACC ACA TCC CAG AT-3'
DJ3368	5'-TGG ACA CGT GGA TGA CGA CGA CAA GAT CGT G-3'
DJ3369	5'-CCC CAC CGG TCC GAA CCA CAT CCC AGA TC-3'
DJ3385	5'-Pho-CAG GTT CCA CTG GTA GCC GCA TGC ACC GCA ACG GCA AGG AC-3'
DJ3386	5'-Pho-GTG GCT GCC GTT GCG GTG CAT GCG GCT ACC AGT GGA ACC TGG TAC-3'
DJ3511	5‘-TGTATGACCTCATGACTTCAGG-3‘
DJ3512	5‘-CCCAGCCCCACAAGTAAATTAA-3‘
DJ3513	5‘-GCAGCGCATCGCCTTCTATC -3‘
DJ3559	5'-TCG AGA TCA AGC CTA GAA GT-3'
DJ3560	5'-Pho-ATA CTT CTA GGC TTG ATC-3'
DJ3569	5'-ATCCCAGATCCAGGTCACAA-3'
DJ3610	5'-ATT GCC TCG GAG AAA TGC TA-3'

7.3 DNA and protein sequences

7.3.1 S-NSP4 (human)

The amino acid sequence is shown in single letter code. The sequence begins with the Igκ signal peptide. The N-terminal S-tag is marked in bold and underlined. The enterokinase cleavage site was underlined. The sequence of mature NSP4 is marked in bold. The C-terminal 6x His-tag is marked in italics. The stop codon is represented by a star (*).

```
1    ATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTACCAGGTTCCACTGGT  60
1    M  E  T  D  T  L  L  L  W  V  L  L  L  W  V  P  G  S  T  G   20

61   GACGTGAAAGAAACCGCTGCTGCTAAATTCGAACGCCAGCACATGGACAGCGGAAGCGGT  120
21   D  V  K  E  T  A  A  A  K  F  E  R  Q  H  M  D  S  G  S  G   40

121  GGACACGTGGATGACGACGACAAGATCATCGGGGGCCACGAGGTGACCCCCCACTCCAGG  180
41   G  H  V  D  D  D  D  K  I  I  G  G  H  E  V  T  P  H  S  R   60

181  CCCTACATGGCATCCGTGCGCTTCGGGGGCCAACATCACTGCGGAGGCTTCCTGCTGCGA  240
61   P  Y  M  A  S  V  R  F  G  G  Q  H  H  C  G  G  F  L  L  R   80

241  GCCCGCTGGGTGGTCTCGGCCGCCCACTGCTTCAGCCACAGAGACCTCCGCACTGGCCTG  300
81   A  R  W  V  V  S  A  A  H  C  F  S  H  R  D  L  R  T  G  L   100

301  GTGGTGCTGGGCGCCCACGTCCTGAGTACTGCGGAGCCCACCCAGCAGGTGTTTGGCATC  360
101  V  V  L  G  A  H  V  L  S  T  A  E  P  T  Q  Q  V  F  G  I   120

361  GATGCTCTCACCACGCACCCCGACTACCACCCCATGACCCACGCCAACGACATCTGCCTG  420
121  D  A  L  T  T  H  P  D  Y  H  P  M  T  H  A  N  D  I  C  L   140

421  CTGCGGCTGAACGGCTCTGCTGTCCTGGGCCCTGCAGTGGGGCTGCTGAGGCTGCCAGGG  480
141  L  R  L  N  G  S  A  V  L  G  P  A  V  G  L  L  R  L  P  G   160

481  AGAAGGGCCAGGCCCCCCACAGTGGGGACACGGTGCCGGGTGGCTGGCTGGGGCTTTGTG  540
161  R  R  A  R  P  P  T  V  G  T  R  C  R  V  A  G  W  G  F  V   180

541  TCTGACTTTGAGGAGCTGCCGCCTGGACTGATGGAGGCCAAGGTCCGAGTGCTGGACCCG  600
181  S  D  F  E  E  L  P  P  G  L  M  E  A  K  V  R  V  L  D  P   200

601  GACGTCTGCAACAGCTCCTGGAAGGGCCACCTGACACTTACCATGCTCTGCACCCGCAGT  660
201  D  V  C  N  S  S  W  K  G  H  L  T  L  T  M  L  C  T  R  S   220

661  GGGGACAGCCACAGACGGGGCTTCTGCTCGGCCGACTCCGGAGGGCCCCTGGTGTGCAGG  720
221  G  D  S  H  R  R  G  F  C  S  A  D  S  G  G  P  L  V  C  R   240

721  AACCGGGCTCACGGCCTCGTTTCCTTCTCGGGCCTCTGGTGCGGCGACCCCAAGACCCCC  780
241  N  R  A  H  G  L  V  S  F  S  G  L  W  C  G  D  P  K  T  P   260

781  GACGTGTACACGCAGGTGTCCGCCTTTGTGGCCTGGATCTGGGACGTGGTTCGGAAGACC  840
261  D  V  Y  T  Q  V  S  A  F  V  A  W  I  W  D  V  V  R  K  T   280

841  GGTCATCATCACCATCACCATTGA                                      864
281  G  H  H  H  H  H  H  *                                        287
```

7.3.2 NGS-NSP4 (human)

The amino acid sequence is shown in single letter code. The sequence begins with the Igκ signal peptide. The N-terminal NGS-peptide is marked in bold and underlined. The enterokinase cleavage site was underlined. The sequence of mature NSP4 is marked in bold. The C-terminal 6x His-tag is marked in italics. The stop codon is represented by a star (*).

```
1    ATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTACCAGGTTCCACTGGT  60
1    M  E  T  D  T  L  L  L  W  V  L  L  L  W  V  P  G  S  T  G   20

61   AGCCGCATGCACCGCAACGGCAGCCACGTGGATGACGACGACAAGATCATCGGGGGCCAC  120
21   S  R  M  H  R  N  G  S  H  V  D  D  D  D  K  I  I  G  G  H   40

121  GAGGTGACCCCCCACTCCAGGCCCTACATGGCATCCGTGCGCTTCGGGGGCCAACATCAC  180
41   E  V  T  P  H  S  R  P  Y  M  A  S  V  R  F  G  G  Q  H  H   60

181  TGCGGAGGCTTCCTGCTGCGAGCCCGCTGGGTGGTCTCGGCCGCCCACTGCTTCAGCCAC  240
61   C  G  G  F  L  L  R  A  R  W  V  V  S  A  A  H  C  F  S  H   80

241  AGAGACCTCCGCACTGGCCTGGTGGTGCTGGGCGCCCACGTCCTGAGTACTGCGGAGCCC  300
81   R  D  L  R  T  G  L  V  V  L  G  A  H  V  L  S  T  A  E  P   100

301  ACCCAGCAGGTGTTTGGCATCGATGCTCTCACCACGCACCCCGACTACCACCCCATGACC  360
101  T  Q  Q  V  F  G  I  D  A  L  T  T  H  P  D  Y  H  P  M  T   120

361  CACGCCAACGACATCTGCCTGCTGCGGCTGAACGGCTCTGCTGTCCTGGGCCCTGCAGTG  420
121  H  A  N  D  I  C  L  L  R  L  N  G  S  A  V  L  G  P  A  V   140

421  GGGCTGCTGAGGCTGCCAGGGAGAAGGGCCAGGCCCCCCACAGTGGGGACACGGTGCCGG  480
141  G  L  L  R  L  P  G  R  R  A  R  P  P  T  V  G  T  R  C  R   160

481  GTGGCTGGCTGGGGCTTTGTGTCTGACTTTGAGGAGCTGCCGCCTGGACTGATGGAGGCC  540
161  V  A  G  W  G  F  V  S  D  F  E  E  L  P  P  G  L  M  E  A   180

541  AAGGTCCGAGTGCTGGACCCGGACGTCTGCAACAGCTCCTGGAAGGGCCACCTGACACTT  600
181  K  V  R  V  L  D  P  D  V  C  N  S  S  W  K  G  H  L  T  L   200

601  ACCATGCTCTGCACCCGCAGTGGGGACAGCCACAGACGGGGCTTCTGCTCGGCCGACTCC  660
201  T  M  L  C  T  R  S  G  D  S  H  R  R  G  F  C  S  A  D  S   220

661  GGAGGGCCCCTGGTGTGCAGGAACCGGGCTCACGGCCTCGTTTCCTTCTCGGGCCTCTGG  720
221  G  G  P  L  V  C  R  N  R  A  H  G  L  V  S  F  S  G  L  W   240

721  TGCGGCGACCCCAAGACCCCCGACGTGTACACGCAGGTGTCCGCCTTTGTGGCCTGGATC  780
241  C  G  D  P  K  T  P  D  V  Y  T  Q  V  S  A  F  V  A  W  I   260

781  TGGGACGTGGTTCGGAAGACCGGTCATCATCACCATCACCATTGA                 825
261  W  D  V  V  R  K  T  G  H  H  H  H  H  H  *                  274
```

7.3.3 NSP4 precursor (human)

The amino acid sequence is shown in single letter code. The sequence begins with the signal peptide. The N-terminal prodipeptide is underlined. The sequence of mature NSP4 is marked in bold. The C-terminal 6x His-tag is marked in italics. The stop codon is represented by a star (*).

```
1    ATGGGGCTCGGGTTGAGGGGCTGGGGACGTCCTCTGCTGACTGTGGCCACCGCCCTGATG  60
1    M  G  L  G  L  R  G  W  G  R  P  L  L  T  V  A  T  A  L  M   20

61   CTGCCCGTGAAGCCCCCCGCAGGCTCCTGGGGGGCCCAGATCATCGGGGGCCACGAGGTG 120
21   L  P  V  K  P  P  A  G  S  W  G  A  Q  I  I  G  G  H  E  V   40

121  ACCCCCCACTCCAGGCCCTACATGGCATCCGTGCGCTTCGGGGGCCAACATCACTGCGGA 180
41   T  P  H  S  R  P  Y  M  A  S  V  R  F  G  G  Q  H  H  C  G   60

181  GGCTTCCTGCTGCGAGCCCGCTGGGTGGTCTCGGCCGCCCACTGCTTCAGCCACAGAGAC 240
61   G  F  L  L  R  A  R  W  V  V  S  A  A  H  C  F  S  H  R  D   80

241  CTCCGCACTGGCCTGGTGGTGCTGGGCGCCCACGTCCTGAGTACTGCGGAGCCCACCCAG 300
81   L  R  T  G  L  V  V  L  G  A  H  V  L  S  T  A  E  P  T  Q   100

301  CAGGTGTTTGGCATCGATGCTCTCACCACGCACCCCGACTACCACCCCATGACCCACGCC 360
101  Q  V  F  G  I  D  A  L  T  T  H  P  D  Y  H  P  M  T  H  A   120

361  AACGACATCTGCCTGCTGCGGCTGAACGGCTCTGCTGTCCTGGGCCCTGCAGTGGGGCTG 420
121  N  D  I  C  L  L  R  L  N  G  S  A  V  L  G  P  A  V  G  L   140

421  CTGAGGCTGCCAGGGAGAAGGGCCAGGCCCCCCACAGCGGGGACACGGTGCCGGGTGGCT 480
141  L  R  L  P  G  R  R  A  R  P  P  T  A  G  T  R  C  R  V  A   160

481  GGCTGGGGCTTCGTGTCTGACTTTGAGGAGCTGCCGCCTGGACTGATGGAGGCCAAGGTC 540
161  G  W  G  F  V  S  D  F  E  E  L  P  P  G  L  M  E  A  K  V   180

541  CGAGTGCTGGACCCGGACGTCTGCAACAGCTCCTGGAAGGGCCACCTGACACTTACCATG 600
181  R  V  L  D  P  D  V  C  N  S  S  W  K  G  H  L  T  L  T  M   200

601  CTCTGCACCCGCAGTGGGGACAGCCACAGACGGGGCTTCTGCTCGGCCGACTCCGGAGGG 660
201  L  C  T  R  S  G  D  S  H  R  R  G  F  C  S  A  D  S  G  G   220

661  CCCCTGGTGTGCAGGAACCGGGCTCACGGCCTCGTTTCCTTCTCGGGCCTCTGGTGCGGC 720
221  P  L  V  C  R  N  R  A  H  G  L  V  S  F  S  G  L  W  C  G   240

721  GACCCCAAGACCCCCGACGTGTACACGCAGGTGTCCGCCTTTGTGGCCTGGATCTGGGAC 780
241  D  P  K  T  P  D  V  Y  T  Q  V  S  A  F  V  A  W  I  W  D   260

781  GTGGTTCGGAAGACCGGTCATCATCACCATCACCATTGA                       819
261  V  V  R  K  T  G  H  H  H  H  H  H  *                        272
```

7.3.4 S-mNSP4 (mouse)

The amino acid sequence is shown in single letter code. The sequence is begins with the Igκ signal peptide. The N-terminal S-tag is marked in bold (line 2). The enterokinase cleavage site was underlined. The sequence of mature mNSP4 is marked in bold (lines 3-14). The C-terminal 6x His-tag is marked in italics. The stop codon is represented by a star (*).

```
1    ATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTACCAGGTTCCACTGGT  60
1    M  E  T  D  T  L  L  L  W  V  L  L  L  W  V  P  G  S  T  G   20

61   GACGTGAAAGAAACCGCTGCTGCTAAATTCGAACGCCAGCACATGGACAGCGGAAGCGGT  120
21   D  V  K  E  T  A  A  A  K  F  E  R  Q  H  M  D  S  G  S  G   40

121  GGACACGTGGATGACGACGACAAGATCGTGGGCGGCCATGAGGTAACGCCACATTCCCGG  180
41   G  H  V  D  D  D  D  K  I  V  G  G  H  E  V  T  P  H  S  R   60

181  CCCTATATGGCCTCTGTGAGCTTCGAGGGCCATCACTACTGTGGAGGCTTCCTGATTCAT  240
61   P  Y  M  A  S  V  S  F  E  G  H  H  Y  C  G  G  F  L  I  H    80

241  ACCCATTGGGTGGTGTCAGCTGCCCATTGCTTCAGTGACAGGGACCCCTCCATGGGGCTG  300
81   T  H  W  V  V  S  A  A  H  C  F  S  D  R  D  P  S  M  G  L   100

301  GTGGTGCTGGGGGCCCACGCACTGCTCGCCCCAGAGCCCACACAGCAGACATTTAGCATT  360
101  V  V  L  G  A  H  A  L  L  A  P  E  P  T  Q  Q  T  F  S  I   120

361  GCAGCTGCTGTCAGCCATCCTGATTTCCAGCCTGCCACACAGGCCAACGATATCTGCCTG  420
121  A  A  A  V  S  H  P  D  F  Q  P  A  T  Q  A  N  D  I  C  L   140

421  CTAAGGTTGAACGGCTCTGCTGTCCTGGGCCCCGCTGTGAGGCTATTACGATTGCCTCGG  480
141  L  R  L  N  G  S  A  V  L  G  P  A  V  R  L  L  R  L  P  R   160

481  AGAAATGCTAAGCCACCTGCTGCGGGCACGCGTTGCCATGTATCCGGCTGGGGCTTCGTG  540
161  R  N  A  K  P  P  A  A  G  T  R  C  H  V  S  G  W  G  F  V   180

541  TCTGACTTTGAGGAGCCTCCGCCTGGACTGATGGAAGTCGAGGTGCGCATACTGGATCTG  600
181  S  D  F  E  E  P  P  P  G  L  M  E  V  E  V  R  I  L  D  L   200

601  AGTGTCTGCAACAGCTCCTGGCAGGGCCAGCTGAATCCTGCCATGCTCTGCACTCACAGT  660
201  S  V  C  N  S  S  W  Q  G  Q  L  N  P  A  M  L  C  T  H  S   220

661  GGGGACCGCCGGCGGCGTGGTTTCTGCTCGGCAGATTCTGGGGGACCCTTGGTCTGTGGA  720
221  G  D  R  R  R  R  G  F  C  S  A  D  S  G  G  P  L  V  C  G   240

721  AGACGGGCCCATGGTCTTGTATCTTTCTCAGGCCTCTGGTGTGGTGACCCCAAGACTCCG  780
241  R  R  A  H  G  L  V  S  F  S  G  L  W  C  G  D  P  K  T  P   260

781  GATGTCTACACTCAAGTGTCAGCGTTTGTGACCTGGATCTGGGATGTTGTTCGGACCGGT  840
261  D  V  Y  T  Q  V  S  A  F  V  T  W  I  W  D  V  V  R  T  G   280

841  CATCATCACCATCACCATTGA                                        861
281  H  H  H  H  H  H  *                                          286
```

7.3.5 α1PI*

The amino acid sequence is shown in single letter code. The sequence is begins with the Igκ signal peptide. The sequence of α1PI* is marked in bold. The amino acids in the reactive center loop (RCL) that were optimized for specific interaction with NSP4 were underlined. The C-terminal 6x His-tag is marked in italics. The stop codon is represented by a star (*).

```
1    ATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTACCAGGTTCCACTGGT  60
1    M  E  T  D  T  L  L  L  W  V  L  L  L  W  V  P  G  S  T  G   20

61   GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 120
21   E  D  P  Q  G  D  A  A  Q  K  T  D  T  S  H  H  D  Q  D  H   40

121  CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 180
41   P  T  F  N  K  I  T  P  N  L  A  E  F  A  F  S  L  Y  R  Q   60

181  CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 240
61   L  A  H  Q  S  N  S  T  N  I  F  F  S  P  V  S  I  A  T  A   80

241  TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 300
81   F  A  M  L  S  L  G  T  K  A  D  T  H  D  E  I  L  E  G  L   100

301  AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 360
101  N  F  N  L  T  E  I  P  E  A  Q  I  H  E  G  F  Q  E  L  L   120

361  CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 420
121  R  T  L  N  Q  P  D  S  Q  L  Q  L  T  T  G  N  G  L  F  L   140

421  AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 480
141  S  E  G  L  K  L  V  D  K  F  L  E  D  V  K  K  L  Y  H  S   160

481  GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 540
161  E  A  F  T  V  N  F  G  D  T  E  E  A  K  K  Q  I  N  D  Y   180

541  GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 600
181  V  E  K  G  T  Q  G  K  I  V  D  L  V  K  E  L  D  R  D  T   200

601  GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 660
20   V  F  A  L  V  N  Y  I  F  F  K  G  K  W  E  R  P  F  E  V   220

661  AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 720
221  K  D  T  E  E  E  D  F  H  V  D  Q  V  T  T  V  K  V  P  M   240

721  ATGAAGCGTTTAGGCATGTTTAACATCCAGCACAGTAAGAAGCTGTCCAGCTGGGTGCTG 780
241  M  K  R  L  G  M  F  N  I  Q  H  S  K  K  L  S  S  W  V  L   260

781  CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 840
261  L  M  K  Y  L  G  N  A  T  A  I  F  F  L  P  D  E  G  K  L   280

841  CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 900
281  Q  H  L  E  N  E  L  T  H  D  I  I  T  K  F  L  E  N  E  D   300

901  AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 960
301  R  R  S  A  S  L  H  L  P  K  L  S  I  T  G  T  Y  D  L  K   320

961  AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1020
321  S  V  L  G  Q  L  G  I  T  K  V  F  S  N  G  A  D  L  S  G   340
```

```
1021 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1080
341  V  T  E  E  A  P  L  K  L  S  K  A  V  H  K  A  V  L  T  I   360

1081 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTCCTCGAGATCAAGCCTAGAAGTATA 1140
361  D  E  K  G  T  E  A  A  G  A  M  F  L  E  I  K  P  R  S  I   380

1141 CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1200
381  P  P  E  V  K  F  N  K  P  F  V  F  L  M  I  E  Q  N  T  K   400

1201 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAAAACCGGTCATCATCACCAT 1260
401  S  P  L  F  M  G  K  V  V  N  P  T  Q  K  T  G  H  H  H  H   420

1261 CACCATTGA                                                    1269
421  H  H  *                                                       422
```

7.4 List of synthetic protease substrates

Sbzl substrates

Sequence	Cleaving protease	Provider
Boc-Ala-Pro-Nva-4-chloro-Sbzl	NE, PR3	Bachem
Suc-Phe-Leu-Phe-Sbzl	CG	Bachem

AMC substrates

Sequence	Cleaving protease	Provider
H-Tyr-Arg-Phe-Arg-AMC	NSP4	Novabiochem (Merck)
H-Tyr-Arg-Phe-Lys-AMC	n/a	Novabiochem

Individual FRET substrates - P1'P2'

Sequence (single letter code)	Cleaving protease	Provider
Mca-GFPRFFP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRFVP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRFSP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRFNP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRFEP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRFRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRFGP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRVFP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRVVP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRVSP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRVNP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRVEP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRVRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRVGP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRSFP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRSVP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRSSP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRSNP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRSEP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRSGP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRNFP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRNVP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRNSP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRNNP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRNEP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRNRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRNGP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPREFP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPREVP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRESP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRENP-Lys(Dnp)-rr	NSP4	EMC microcollections

Individual FRET substrates - P1'P2' - Continued

Sequence (single letter code)	Cleaving protease	Provider
Mca-GFPREEP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRERP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPREGP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRRFP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRRVP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRRSP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRRNP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRREP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRRRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRRGP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRGFP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRGVP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRGSP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRGNP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRGEP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRGRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFPRGGP-Lys(Dnp)-rr	NSP4	EMC microcollections

r = D-arginine

Individual FRET substrates - P4P3

Sequence (single letter code)	Cleaving protease	Provider
Mca-GFYPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFAPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFSPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFNPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFDPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFGPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GAYPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GAAPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GASPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GANPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GADPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections

Individual FRET substrates - P4P3 - Continued

Sequence (single letter code)	Cleaving protease	Provider
Mca-GAKPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GAGPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GSYPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GSAPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GSSPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GSNPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GSDPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GSKPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GSGPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GNYPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GNAPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GNSPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GNNPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GNDPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GNKPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GNGPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GEYPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GEAPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GESPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GENPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GEDPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GEKPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GEGPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GRYPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GRAPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GRSPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GRNPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GREPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GRKPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GRGPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GGYPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GGAPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections

Individual FRET substrates - P4P3 - Continued

Sequence (single letter code)	Cleaving protease	Provider
Mca-GGSPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GGNPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GGDPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GGKPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GGGPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections

r = D-arginine

Individual FRET substrates - P2

Sequence (single letter code)	Cleaving protease	Provider
Mca-GFKARSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKDRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKERSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKFRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKGRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKHRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKIRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKKRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKLRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKMRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKNRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKPRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKQRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKRRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKSRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKTRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKVRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKWRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections
Mca-GFKYRSRP-Lys(Dnp)-rr	NSP4	EMC microcollections

r = D-arginine

Other FRET substrates

Sequence (single letter code)	Cleaving protease	Provider
Mca-GIKPRSRP-Lys(Dnp)-rr (α1PI* FRET)	NSP4	EMC microcollections
Abz-GSKGR-SLIG-Y (NO_2)-D (PAR-2 FRET)	NSP4	EMC microcollections

r = D-arginine

7.5 Lists of NSP4 cleavage sites in PICS libraries

Nonprime-side sequences are reported up to the position P10. X denotes ambiguous nonprime side residues. A small number of cleavage sites have P1 residues, which correspond to a potential cleavage site of the digestion protease used for library generation. These cases might be related to incomplete amine protection during library preparation. Hence, prime-side cleavage products with such amino-termini are omitted from the proteomic identification of protease cleavage sites (PICS) analysis, as well as prime-side cleavage products with amino termini that correspond to a protein amino terminus. One exemplary protein is stated for each identified peptide sequence. X!Tandem was applied to assign peptides to tandem mass spectra and the X!Tandem hyperscore is reported together with the PeptideProphet probability score. C signifies carboxyamidomethylated cysteine and K signifies dimethylated lysine. Some peptides were identified repeatedly by LC-MS/MS; however, the set of sequences was rendered nonredundant before analysis in the form of sequence logos.

7.5.1 Chymotrypsin

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
QFDAAFVPEV	AAQKAPASP	0.86	25.6	−0.7	955.5	17888	
YHGRVQALAD	AAREAGLQF	0.96	25.2	0.6	1049.5	17897	
AHEAGEFFMR	AGSATVRPTEGAGGTL	1.00	45.5	−0.3	1531.7	17875	
AHEAGEFFMR	AGSATVRPTEGAGGTL	1.00	45.6	−1.5	1531.7	17875	
GAVIAYEPVW	AIGTGKSATPAQAQAVH	1.00	52.6	−0.8	1722.9	17903	
GAVIAYEPVW	AIGTGKSATPAQAQAVH	1.00	49.6	−1.7	1722.9	17903	
NFNVPFVVVR	AISDVADQQSHL	1.00	50.3	−0.9	1370.6	17863	
NFNVPFVVVR	AISDVADQQSHL	1.00	53.9	−1.3	1370.6	17863	
MSFELPALPY	AKDALAPH	0.97	32.1	−0.7	937.5	17879	X
MSFELPALPY	AKDALAPH	0.99	29.3	−2.2	937.5	17879	X
KLDLERTVIR	APADGWVTNL	0.84	26.8	2.0	1130.5	17896	
LAVDESFQPT	AVGFAEALNNKDKPE	1.00	45.9	2.2	1745.9	17899	
LCRNCKIVKR	DGVIRVIC	0.85	31.6	−0.9	1018.5	17896	
SLXGTIIVGR	DIAHAKL	0.96	36.5	−0.5	882.5	17878	
SLXGTIIVGR	DIAHAKL	0.83	34.3	−0.2	882.5	17878	
KLTKRMRVIR	EKVDATKQY	1.00	35.7	−0.9	1224.6	17904	
KLTKRMRVIR	EKVDATKQY	0.99	32.8	−1.1	1224.6	17904	
NNPFFVSLKD	GAQKEADKLGY	0.95	38.2	−0.3	1322.7	17901	
NNPFFVSLKD	GAQKEADKLGY	0.97	38.1	−0.5	1322.7	17901	
GIANTFIAAQ	GHDVGKSLY	0.82	29.7	1.7	1090.5	17892	
KPEDAVLDVQ	GIATVTPAIVQ	0.97	41.8	−1.2	1156.6	17899	
KPEDAVLDVQ	GIATVTPAIVQ	0.97	38.1	−0.9	1156.6	17899	
KPEDAVLDVQ	GIATVTPAIVQAC	0.90	35	−3.0	1387.7	17899	

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
GQNVEFEIQD	GQKGPAAVNVTAI	0.90	27.1	−2.2	1340.7	17881	
HVIAGKAVAL	KEAMEPEFKTY	0.85	36.9	−2.1	1515.7	17889	X
SGVRAIDTKC	KIEQAPGQH	0.99	28.1	−0.9	1122.5	17896	
PDEGIPAVCF	KLKDGEDPGY	0.99	27.9	−0.8	1264.6	17877	X
PDEGIPAVCF	KLKDGEDPGY	0.94	31.1	−2.7	1264.6	17877	X
PDEGIPAVCF	KLKDGEDPGY	0.97	30.8	−2.6	1264.6	17877	X
KNPQKNLYTF	KNQASNDLPN	1.00	24.6	−1.1	1215.6	87082	X
KNPQKNLYTF	KNQASNDLPN	0.97	19.9	−1.1	1215.6	87082	X
SGGAGIQADL	KTFSALGAY	0.81	30.9	−7.4	1072.5	17884	X
MVEVFLERGY	KVVSGGTDNHLF	0.99	41.4	−0.6	1388.7	17889	X
AKSVKFKYPR	QRKTVVADGVGQGY	0.93	29.2	−3.0	1592.8	17863	
WRQKGTGRAR	SGSIKSPIW	0.89	30.7	0.0	1089.6	17897	
ARPPLVVISR	SHADAELKEY	1.00	42.2	−2.0	1277.6	17906	
ARPPLVVISR	SHADAELKEY	1.00	42.1	−2.0	1277.6	17906	
KEASAGKLVR	TLAAVRDAKEAA	1.00	42.3	−1.9	1330.7	17904	
KEASAGKLVR	TLAAVRDAKEAA	1.00	39.5	−1.0	1330.7	17904	
LSQYDFPGDD	TPIVRGSAL	0.97	31	−1.0	1000.5	17897	
RDXKGAVASL	TSVAKLPF	0.83	24.9	0.5	977.5	17871	X
VVNTIRGIVK	VAAVKAPGF	1.00	41.8	−1.2	974.5	17905	
KHLPEPFRIR	VIEPVKR	0.91	41.5	1.3	955.6	87082	
KHLPEPFRIR	VIEPVKR	0.98	41.7	−0.7	955.6	87082	
DIVWDFRLPR	VKSISASGH	0.98	33.5	−1.4	1000.5	17877	
MSTAKL	VKSKATNLLY	0.99	27.3	−0.6	1279.7	17870	X
CKPTSPGRRH	VVKVVNPELH	0.96	29.1	−1.9	1248.7	17897	
CKPTSPGRRH	VVKVVNPELH	0.94	29.4	−1.0	1248.7	17897	

7.5.2 GluC

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
GYTLNGRTIR	AAMVTVAKAKA	1.00	43.6	−2.3	1203.7	17889	
GYTLNGRTIR	AAMVTVAKAKA	1.00	50.5	−1.6	1203.7	17889	
KYYDPRVWLR	AGQTSMIARLE	0.93	37.1	−0.6	1263.6	17892	
KYYDPRVWLR	AGQTSMIARLE	1.00	48.2	−0.2	1263.6	17892	
KYYDPRVWLR	AGQTSMIARLE	0.97	40.8	−0.6	1263.6	17892	
SQASDSYYYR	AKQAVYR	0.78	25.4	−1.4	950.5	17886	
PFLEHDDANR	ALMGANMQR	0.99	38	−0.7	1078.5	17904	
PFLEHDDANR	ALMGANMQR	1.00	41.9	−2.1	1078.5	17904	
FQPTAVGFAE	ALNNKDKPE	0.80	37.3	−0.6	1171.6	17899	X
KLEHAVPMAK	ALVAGGVRVLE	1.00	38	−2.8	1170.6	17881	
KLEHAVPMAK	ALVAGGVRVLE	1.00	49.3	−0.4	1170.6	17881	
APDQTTTIVR	ANSSTTTAAEPLKM	0.93	30.3	−0.8	1536.7	17886	
APDQTTTIVR	ANSSTTTAAEPLKM	0.96	37	−1.0	1536.7	17886	
LAVDESFQPT	AVGFAEALNNKDKPE	1.00	53.5	0.0	1745.9	17899	
HVSPGALDAE	AYGVKSTIED	1.00	35.1	7.3	1197.6	17905	X
GCVELVQPGG	FDELVQIYE	0.91	27.5	0.5	1242.5	14569	
VALKEAMEPE	FKTYQQQVAK	1.00	33.1	0.4	1383.7	17889	X
LVVSTLNNPF	FVSLKDGAQKE	1.00	46.1	−4.4	1364.7	17901	
LVVSTLNNPF	FVSLKDGAQKE	1.00	42.3	−3.6	1364.7	17901	
LTASTVIQYR	GAIIPREAMPA	0.96	29.6	−0.5	1212.6	87082	
ESGVLNQQPY	GFNTRFE	0.76	23.4	−2.2	957.4	17877	
ESGVLNQQPY	GFNTRFE	1.00	27.3	0.3	957.4	17877	
GNGDCHIILR	GGKEPNYSAK	0.90	30.1	−1.5	1193.6	17869	
GNGDCHIILR	GGKEPNYSAK	0.97	33.3	0.3	1193.6	17869	

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
TTHKTLAGPR	GGLILAKGGSEE	1.00	37.2	−0.4	1245.6	17889	
TTHKTLAGPR	GGLILAKGGSEE	1.00	37.4	−0.5	1245.6	17889	
TEEYGNMIDM	GILDPTKVTR	0.97	43.4	−1.5	1214.7	17905	
TEEYGNMIDM	GILDPTKVTR	0.82	36.2	−2.5	1214.7	17905	
GAILVVAATD	GPMPQTRE	0.99	34.8	0.3	1002.4	17897	
YTSPRVMQAQ	GSQLTNKYAE	0.93	29.5	−0.9	1225.6	17889	
IEDVFSISGR	GTVVTGRVE	0.97	38	−1.3	1004.5	17897	
IEDVFSISGR	GTVVTGRVE	0.96	41.7	−2.4	1004.5	17897	
ARGAIFGLTR	GVNANHIIR	0.82	38.9	−2.3	1080.5	17903	
ARGAIFGLTR	GVNANHIIR	1.00	50.6	−1.4	1080.5	17903	
GDGVGMAIRA	GVPVQDM	0.80	22.6	−1.6	832.3	17869	
QLQQVLMMSR	HNLRAPLANNGSVLE	0.75	37.4	1.9	1691.8	17872	
QLQQVLMMSR	HNLRAPLANNGSVLE	1.00	52.6	1.0	1691.8	17872	
SNYFFDTTQG	HSQINGCTVR	0.77	25.4	−0.8	1258.6	87082	
IVGGGIANTF	IAAQGHDVGKSLYE	0.99	38.6	0.4	1602.8	17892	
IVGGGIANTF	IAAQGHDVGKSLYE	1.00	38.8	−0.1	1602.8	17892	
WWMLHEETVY	KGGDTVTLNE	0.75	28.5	−0.9	1148.5	87082	
GKSASAKSLF	KLQTLGLTQ	1.00	35.2	−1.1	1116.6	17887	
GKSASAKSLF	KLQTLGLTQ	1.00	31.7	−0.8	1116.6	17887	
KNPQKNLYTF	KNQASNDLPN	1.00	33.4	−1.6	1215.6	87082	
KNPQKNLYTF	KNQASNDLPN	0.99	21.7	−2.3	1215.6	87082	
KNPQKNLYTF	KNQASNDLPN	1.00	30	−1.4	1215.6	87082	
VYIEGQLRTR	KWTDQSGQDRYTTE	1.00	38.5	−0.3	1829.8	17904	
VYIEGQLRTR	KWTDQSGQDRYTTE	1.00	42	0.4	1829.8	17904	
ETGYSNKVLD	LIAHISK	0.91	34.5	−0.8	896.5	17880	
QKEADKLGYN	LVVLDSQNNPAKE	1.00	37.5	−0.7	1541.8	17901	

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
SNYQPSPMVR	MMKADAAPVSAQE	0.99	35.2	1.5	1463.6	17892	
SNYQPSPMVR	MMKADAAPVSAQE	1.00	35	0.0	1463.6	17892	
AQKEADKLGY	NLVVLDSQNNPAKE	0.95	37.6	0.9	1655.8	17901	
AQKEADKLGY	NLVVLDSQNNPAKE	1.00	37.9	−0.2	1655.8	17901	
AQKEADKLGY	NLVVLDSQNNPAKE	1.00	44.7	−0.2	1655.8	17901	
KDPLDNTYTR	NMYIVKH	0.99	32	−0.6	1019.5	17874	
KDPLDNTYTR	NMYIVKH	1.00	28.4	−0.1	1019.5	17874	
AESVKNIFGY	QYTIPTHQGRGAE	0.95	35.1	−0.3	1544.7	87082	
AESVKNIFGY	QYTIPTHQGRGAE	1.00	39.2	−0.9	1544.7	87082	
GINCNLTLLF	SFAQARACAE	1.00	31.2	−2.7	1197.5	17861	
YAKDGISYTF	SIVPNALGKDDE	0.91	36.7	−2.3	1372.7	17871	
YAKDGISYTF	SIVPNALGKDDE	0.99	40.1	0.6	1372.7	17871	
EMFGYATQLR	SLTKGRASYTME	0.78	32.3	−0.1	1458.7	17897	
KGLNVMQNLL	TAHPDVQAVFAQNDE	1.00	34.4	4.9	1728.8	17901	
KGLNVMQNLL	TAHPDVQAVFAQNDE	1.00	48.3	−2.1	1728.7	17901	
KGLNVMQNLL	TAHPDVQAVFAQNDE	1.00	41	−2.0	1728.7	17901	
CKKNPQKNLY	TFKNQASNDLPN	0.91	24.4	−0.7	1463.7	87082	
CKKNPQKNLY	TFKNQASNDLPN	1.00	30.8	1.3	1463.7	87082	
CKKNPQKNLY	TFKNQASNDLPN	0.98	24	−0.3	1463.7	87082	
SVKNIFGYQY	TIPTHQGR	0.93	33.5	−1.2	996.5	87082	
SVKNIFGYQY	TIPTHQGR	1.00	37.3	−0.3	996.5	87082	
SVKNIFGYQY	TIPTHQGRGAE	0.76	34.3	−1.4	1253.6	87082	
KEASAGKLVR	TLAAVRDAKEAA	0.95	34.8	−1.4	1330.7	17904	
VQVNEADALY	TNPAQARE	1.00	41.3	−2.3	973.4	87082	
VQVNEADALY	TNPAQARE	1.00	38.6	−2.2	973.4	87082	
EHAVPMAKAL	VAGGVRVLE	1.00	47.4	−1.0	986.5	17881	

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
MVMPGDNIKM	VVTLIHPI	1.00	28.6	3.2	978.6	17897	
GVKLINAVQD	VYLDSKITIA	0.91	25.8	−1.6	1237.7	87082	
KTTTTERILF	YTGVNHKIGE	0.96	34.8	−0.1	1232.6	17897	
KTTTTERILF	YTGVNHKIGE	0.95	32.9	0.2	1232.6	17897	

7.5.3 Trypsin

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
VHLEGGFVGM	AAAPSGASTGSR	1.00	58.1	1.0	1119.5	17891	
VHLEGGFVGM	AAAPSGASTGSR	1.00	55	−0.5	1119.5	17891	
EEAEKITTVQ	AAIDYINGHQA	1.00	41.7	0.5	1259.6	17873	
EEAEKITTVQ	AAIDYINGHQA	1.00	45.3	−1.5	1259.6	17873	
HLEGGFVGMA	AAPSGASTGSR	0.98	33	−0.5	1048.5	17891	
LEEIGVVCQQ	AGGHAAFVDAGK	1.00	57.2	−2.8	1215.6	87082	
LEEIGVVCQQ	AGGHAAFVDAGK	1.00	53.8	−0.5	1215.6	87082	
EWAKHNINVN	AIAPGYMATNNTQQLR	1.00	77.6	0.4	1835.9	17892	
EWAKHNINVN	AIAPGYMATNNTQQLR	1.00	49.6	−3.1	1835.9	17892	
VDHGKTTLTA	AITTVLAK	0.65	33	−1.4	931.5	17897	
VDHGKTTLTA	AITTVLAK	0.85	32.8	−7.5	931.5	17897	
HIPADQFPAQ	ALACELYK	0.74	39.1	−1.9	1082.5	87082	
HIPADQFPAQ	ALACELYK	0.99	37.9	−0.4	1082.5	87082	
HIPADQFPAQ	ALACELYK	1.00	41.8	0.2	1082.5	87082	
EVMGGLGGFG	ALCALPQK	0.81	30.8	−3.3	1015.5	17888	
GDGTTTATVL	AQAIITEGLK	0.90	45.5	−0.1	1158.6	17905	
GDGTTTATVL	AQAIITEGLK	0.83	49.5	4.0	1158.6	17905	
LAVDESFQPT	AVGFAEALNNK	1.00	49.6	2.3	1248.6	17899	
LAVDESFQPT	AVGFAEALNNK	0.99	45.4	0.6	1248.6	17899	
LAVDESFQPT	AVGFAEALNNKDKPE	1.00	53.4	0.4	1745.9	17899	
LAVDESFQPT	AVGFAEALNNKDKPE	1.00	57.1	0.4	1745.9	17899	
NPDEAVAIGA	AVQGGVLTGDVK	1.00	61.9	−1.8	1258.7	17861	
DLNPKAMTPV	AWWMLHEE	1.00	38	0.0	1188.5	87082	
KVQNASYQVA	AYLADEIAK	1.00	41.7	−1.1	1108.5	17877	
KVQNASYQVA	AYLADEIAK	0.84	35.2	0.1	1108.5	17877	

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
GTNTIGSSEA	CMLGGMAMK	1.00	44.8	0.4	1113.5	17877	
GTNTIGSSEA	CMLGGMAMK	1.00	37.9	−3.0	1113.5	17877	
TPTRHYAHVD	CPGHADYVK	1.00	42.4	−4.3	1161.5	17897	
TPTRHYAHVD	CPGHADYVK	1.00	42.5	−1.7	1161.5	17897	
GNARQNLATF	CQTWDDENVHK	1.00	45.2	−3.1	1546.6	17877	
GNARQNLATF	CQTWDDENVHK	1.00	51.6	−1.4	1546.6	17877	
DYDRITKLAR	EAVEGAKL	0.99	31	−0.5	931.5	17881	X
DYDRITKLAR	EAVEGAKL	1.00	37.8	−2.1	931.5	17881	X
ESKKGYINSL	GALTGGQALQQAK	1.00	65.8	−2.1	1357.7	17904	
ESKKGYINSL	GALTGGQALQQAK	1.00	65.8	−1.6	1357.7	17904	
VDESFQPTAV	GFAEALNNKDKPE	0.97	41.7	2.7	1575.8	17899	
VDESFQPTAV	GFAEALNNKDKPE	0.87	45.9	0.9	1575.8	17899	
APAELLFEEF	GFTVDNVVAK	1.00	38	−0.9	1164.6	48994	
APAELLFEEF	GFTVDNVVAK	1.00	41.1	−1.6	1164.6	48994	
KTDQKVVTLS	GFVESQAQAEEAVK	1.00	49.6	−0.8	1607.7	17908	
KTDQKVVTLS	GFVESQAQAEEAVK	1.00	57.3	2.7	1607.7	17908	
RGWQVPAFTL	GGEATDIVVMR	0.99	38	0.6	1234.6	17877	
RGWQVPAFTL	GGEATDIVVMR	1.00	41.9	−0.4	1234.6	17877	
EAVAIGAAVQ	GGVLTGDVK	0.90	38.3	−1.1	960.5	17861	
KYLSDHPKLQ	GIAQQNSFK	0.88	35.2	4.7	1107.5	17877	
KPEDAVLDVQ	GIATVTPAIVQACTQDK	1.00	45.7	−2.5	1888.0	17899	
KPEDAVLDVQ	GIATVTPAIVQACTQDK	1.00	57.1	−0.7	1887.9	17899	
KPEDAVLDVQ	GIATVTPAIVQACTQDK	1.00	53.3	−1.6	1887.9	17899	
VGVDVVAEAT	GLFLTDETAR	0.78	29.1	−1.7	1209.6	17880	
VGVDVVAEAT	GLFLTDETAR	1.00	42.7	3.7	1209.6	17880	
YLFVDMAHVA	GLVAAGVYPNPVPH	1.00	54.2	3.4	1477.7	17889	
YLFVDMAHVA	GLVAAGVYPNPVPH	1.00	50.1	0.0	1477.7	17889	
ITEGLKAVAA	GMNPMDLK	0.98	36.5	1.3	1020.4	17905	

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
ITEGLKAVAA	GMNPMDLK	0.95	36.5	−2.6	1020.4	17905	
CIAAGIASLW	GPAHGGANEAALK	1.00	72	−1.8	1307.6	17869	
CIAAGIASLW	GPAHGGANEAALK	1.00	75.9	−0.9	1307.6	17869	
NAAADLAAIS	GQKPLITK	0.77	23.1	−3.5	1027.6	17897	
KLGPYEFICT	GRPDEGIPAVCFK	0.83	36.6	−2.5	1560.7	17877	
KLGPYEFICT	GRPDEGIPAVCFK	1.00	39.7	1.1	1560.7	17877	
NYRNHFVTIL	GTIQGEQPGFINK	1.00	58.9	−0.3	1503.7	87082	
NYRNHFVTIL	GTIQGEQPGFINK	1.00	53.3	0.4	1503.7	87082	
FRNAEFLQAY	GVAIADGPLK	0.82	38.6	−0.2	1055.6	17875	
KARGITINTS	HVEYDTPTR	0.81	33.2	−2.9	1204.5	17897	
KARGITINTS	HVEYDTPTR	0.97	33.2	−1.6	1204.5	17897	
WDFRLPRVKS	ISASGHK	0.65	23.5	−1.6	814.4	17877	
WDFRLPRVKS	ISASGHK	0.70	26.8	−1.6	814.4	17877	
KNPQKNLYTF	KNQASNDLPN	1.00	26.9	−1.6	1215.6	87082	
KNPQKNLYTF	KNQASNDLPN	1.00	30.1	−2.0	1215.6	87082	
GEIEFWFAMI	KVTTIIVM	0.73	20	3.1	1019.6	48994	
NLKAMYSIAK	KYDIPVVMDSAR	1.00	38.9	1.7	1508.7	87082	X
NLKAMYSIAK	KYDIPVVMDSAR	0.97	33.7	−1.4	1508.7	87082	X
GETEDATIAD	LAVGTAAGQIK	0.89	37.1	−0.1	1143.6	17891	
GETEDATIAD	LAVGTAAGQIK	1.00	41.2	−0.5	1143.6	17891	
DLRCVNMVAD	LWHAPAPK	0.82	32.4	−1.5	1034.5	17877	
DLRCVNMVAD	LWHAPAPK	0.91	30	1.9	1034.5	17877	
TNTIGSSEAC	MLGGMAMK	1.00	42	−0.8	953.4	17877	
TNTIGSSEAC	MLGGMAMK	1.00	38.7	0.5	953.4	17877	
AELDDIFSVQ	NLMHPAYK	0.98	32.4	−1.6	1088.5	87082	
AELDDIFSVQ	NLMHPAYK	0.93	38.8	−1.7	1088.5	87082	
ALNMIDYGLD	NLPGGPL	0.68	26.4	−2.2	754.4	17890	
GLEEIGVVCQ	QAGGHAAFVDAGK	1.00	78.5	−1.8	1343.6	87082	

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
GLEEIGVVCQ	QAGGHAAFVDAGK	1.00	77.7	−1.1	1343.6	87082	
DGLEEIGVVC	QQAGGHAAFVDAGK	1.00	86.9	−0.8	1471.7	87082	
DGLEEIGVVC	QQAGGHAAFVDAGK	1.00	86.9	−0.5	1471.7	87082	
GLKQCKANPW	QQFAETHNK	1.00	44.5	−0.7	1217.5	17871	
GLKQCKANPW	QQFAETHNK	0.98	39.3	−0.8	1217.5	17871	
EHDPMEIWAT	QSSTLVEVLAK	0.93	35.3	0.2	1289.7	17903	
EHDPMEIWAT	QSSTLVEVLAK	0.97	39.3	−2.2	1289.7	17903	
AESVKNIFGY	QYTIPTHQGR	1.00	46.3	−1.7	1287.6	87082	
AESVKNIFGY	QYTIPTHQGR	1.00	46.3	−1.0	1287.6	87082	
PYIVATITSN	SAGGQPVSLANLK	0.94	35.1	0.3	1356.7	87082	
PYIVATITSN	SAGGQPVSLANLK	1.00	42	−2.4	1356.7	87082	
ARLMATMKEA	SAGKLVR	0.98	32.7	0.2	845.5	17904	
YAGQDIVSNA	SCTTNCLAPLAK	1.00	45.4	−6.1	1450.7	17880	
RKIKAAQYVA	SHPGEVCPAK	0.99	43.7	−0.6	1196.5	17868	
RKIKAAQYVA	SHPGEVCPAK	1.00	38.5	−0.8	1196.5	17868	
RKIKAAQYVA	SHPGEVCPAK	1.00	47.8	−0.8	1196.5	17868	
PFACIAAGIA	SLWGPAHGGANEAALK	0.96	39.7	1.3	1693.8	17869	
PFACIAAGIA	SLWGPAHGGANEAALK	0.97	43.9	−1.7	1693.8	17869	
HDPMEIWATQ	SSTLVEVLAK	0.95	35.2	−0.8	1161.6	17903	
HDPMEIWATQ	SSTLVEVLAK	0.99	38.2	0.0	1161.6	17903	
RTMACGIAGL	SVAADSLSAIK	1.00	56.7	−0.1	1176.6	17871	
RTMACGIAGL	SVAADSLSAIK	1.00	65.1	4.3	1176.6	17871	
RTMACGIAGL	SVAADSLSAIK	1.00	50.7	−5.6	1176.6	17871	
CKKNPQKNLY	TFKNQASNDLPN	1.00	29.6	−0.6	1463.7	87082	
CKKNPQKNLY	TFKNQASNDLPN	0.98	26.7	1.4	1463.7	87082	
KEASAGKLVR	TLAAVRDAKEAA	0.99	38.5	−0.6	1330.7	17904	X
KEASAGKLVR	TLAAVRDAKEAA	0.99	43.5	−1.0	1330.7	17904	X
KEASAGKLVR	TLAAVRDAKEAA	0.99	35.7	−0.8	1330.7	17904	X

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
KEASAGKLVR	TLAAVRDAKEAA	0.97	34.6	0.1	1330.7	17904	X
EEKARGITIN	TSHVEYDTPTR	1.00	46.4	0.9	1392.6	17897	
EEKARGITIN	TSHVEYDTPTR	1.00	46.3	−1.4	1392.6	17897	
TTGEHEVSFQ	VHSEVFAK	0.83	27.9	−0.7	1031.5	17906	
QQAVAAHKFN	VLASQPADFDR	0.81	25.5	−0.7	1305.6	17901	
PDEAVAIGAA	VQGGVLTGDVK	0.99	40.8	−0.2	1187.6	17861	
PDEAVAIGAA	VQGGVLTGDVK	0.99	46.6	−1.8	1187.6	17861	
MVMPGDNIKM	VVTLIHPI	0.83	26.1	0.3	978.6	17897	
MVMPGDNIKM	VVTLIHPI	0.99	29.7	−1.4	978.6	17897	

8 Publications and Meetings

8.1 Publications in peer-reviewed journals

- **Perera NC**, Schilling O, Kittel H, Back W, Kremmer E, Jenne DE. NSP4, an elastase-related protease in human neutrophils with arginine specificity. *Proc. Natl. Acad. Sci. U S A*. 2012 Apr 17;109(16):6229-34. Epub 2012 Apr 2.
- **Perera NC**, Jenne DE. Perspectives and potential roles for the newly discovered NSP4 in the immune system. *Expert Rev. Clin. Immunol*. 2012 Aug;8(6):501-3 (Editorial)
- **Perera NC**, Wiesmüller K, Torp Larsen M, Schacher B, Eickholz P, Borregaard N, Jenne DE. NSP4 is stored in azurophil granules and released by activated neutrophils as active endoprotease with restricted specificity. *J. Immunol*. 2013 Sep 1;191(5):2700-7. Epub 2013 Jul 31.
- O'Donoghue AJ, Jin Y, Knudsen GM, **Perera NC**, Jenne DE, Murphy JE, Craik CS and Hermiston T. Global Substrate profiling of proteases in human Neutrophil Extracellular Traps reveals consensus motif predominantly contributed by Elastase. Accepted by *PLoS One*. 2013 Aug 12.

8.1.1 Conference abstract

- **Perera NC**, Torp Larsen M, Wiesmüller K, Borregaard N, Jenne DE. 1.102 NSP4 is stored in azurophilic granules and released by activated neutrophils as an active endoprotease with restricted specificity. *Eur. J. Clin. Invest*. 2013; 43: 36.

8.2 Presentations at international conferences

8.2.1 Oral presentations

- A novel granule-associated serine protease in human neutrophils. *26th International Winter School on Proteinases and Their Inhibitors.* Italy, 2009. (Best talk award)
- A Novel Granule-Associated Serine Protease of Human Neutrophils. *Gordon Research Conference: Proteolytic Enzymes and Their Inhibitors.* Italy, 2010.
- NSP4 is a novel granule associated serine protease of human neutrophils. *28th International Winter School on Proteinases and Their Inhibitors.* Italy, 2011.
- Substrate specificity profiling of NSP4, the novel serine protease in human neutrophils. *29th International Winter School on Proteinases and Their Inhibitors.* Italy, 2012.
- NSP4 is stored in azurophil granules and released by activated neutrophils as an active endoprotease with restricted specificity. *30th International Winter School on Proteinases and Their Inhibitors.* Italy, 2013.

8.2.2 Poster presentations

- A Novel Granule-Associated Serine Protease of Human Neutrophils. *Gordon Research Conference: Proteolytic Enzymes and Their Inhibitors.* Italy, 2010. (Baxxter Poster Prize)
- NSP4 is stored in azurophil granules and released by activated neutrophils as active endoprotease. *47th Annual Scientific Meeting of the European Society for Clinical Investigation (ESCI).* Portugal, 2013. (ESCI Travel Grant)

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt ist.

München, den ....

(Unterschrift)

Erklärung

Hiermit erkläre ich, *

☐ dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist.

☐ dass ich mich anderweitig einer Doktorprüfung ohne Erfolg **nicht** unterzogen habe.

☐ dass ich mich mit Erfolg der Doktorprüfung im Hauptfach .....

und in den Nebenfächern .....

bei der Fakultät für ..... der .....

(Hochschule/Universität)

unterzogen habe.

☐ dass ich ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich der Doktorprüfung zu unterziehen.

München, den..... .....

(Unterschrift)

*) Nichtzutreffendes streichen