Development of Peptidomimetic Substrates and Inhibitors for Transmembrane Serine Proteases

Dissertation

zur Erlangung des Grades

'Doktor rerum naturalium (Dr. rer. nat.)' im Promotionsfach Chemie

am Fachbereich Chemie, Pharmazie und Geowissenschaften

der Johannes Gutenberg-Universität

in Mainz (D77)

vorgelegt von

Philip Maximilian Knaff

geboren in Esch-sur-Alzette (Luxemburg)

Mainz, 2021

The research studies for the present work were carried out from August 1st, 2018 until May 30th, 2021 at the Max Planck Institute for Polymer Research in Mainz in the department of Prof. Dr. Katharina Landfester under supervision of Prof. Dr. Volker Mailänder.

I, Philip Maximilian Knaff, hereby certify that I have prepared this dissertation entitled "Rational development of peptidomimetic substrates and inhibitors for transmembrane serine proteases" by myself and I have indicated all support used (literature, apparatus, materials) in the thesis. I have not submitted the work presented for any government or scientific examination and I have not submitted the work or any part of it to any other faculty or department as a dissertation.

....
Philip Maximilian Knaff

Mainz, July 14th, 2021

Dekan: Univ.-Prof. Dr. Tobias Reich

Prodekanin: Univ.-Prof. Dr. Tanja Schirmeister

Gutachter/in: 1. Prof. Dr. Katharina Landfester

2. Prof. Dr. Volker Mailänder

Datum der Einreichung: 20.07.2021

Content

Abstract

A new strategy for the treatment of cancer is the concept of hollow nanocarriers which function as a delivery vehicle for drugs and allow increased drug efficiency. For targeted release of the drug to the site of action, nanocarriers can be functionalized with a smart moiety. In the scope of the "HEPSIN-NANOLYSIS" project, this moiety comprises of a peptide which act as a proteolytically degradable substrate for the serine protease hepsin. As a pericellular-active protease, hepsin is easily accessible and its high overexpression in androgen-independent prostate cancer makes it of particular interest. The peptide substrate known in literature with the amino acid sequence RQLR↓VVGG lacks specificity towards hepsin, which limits the applicability of this therapeutic concept due to the large number of proteases present in blood. To tackle this issue, a methodology for finding and optimizing novel cleavable substrates is developed in the first study of chapter A. This methodology included the *in silico* preparation of combinatorial peptide libraries targeting the P- and P′-site and molecular docking of the respective library in the catalytic active sub site cavity of hepsin. Calculation of binding scores and analysis of binding modes allowed the selection of peptide combinations which were used to improve the reported substrate RQLR↓VVGG. Selected substrate candidates were synthesized as IQF substrates using the FRET system Mca/Dnp and enzymatically tested. *In vitro* screening against hepsin revealed catalytic efficiencies up to 2.3-fold higher than the literature known substrate. The stability in body fluids could be further improved through introduction of D-amino acids on a selected substrate. With this strategy, two substrate combinations were identified with resistance to degradation by serum or plasma while maintaining cleavability to hepsin.

The enzymatic activity of hepsin may not only be used as stimulus for triggered drug release, but it also presents an attractive drug target through its ability to promote the formation of cancer metastasis through cell migration. Therefore, the above-mentioned approach was used to design covalent peptidomimetic inhibitors for this serine protease. In this regard, eight top scoring peptide combinations targeting the P-site were connected with electrophilic serine traps. It was shown that the peptide combinations result in hepsin inhibitors having high potency and good off-target selectivity against coagulation proteases. Lead candidates were stable for up to 10 days in serum, plasma and medium. Further, the influence of the serine trap and truncated versions on a selected candidate were investigated which allowed deeper insides into structure-activity relationships. With this strategy, a variety of candidates with high antiproteolytic activity could be identified with new structures which may offer a promising scaffold for further drug development.

Abstract

Our strategy to prevent viral infections synergizes with potential anticancer therapeutics, too. The transmembrane serine proteases 2 (TMPRSS2) primes the SARS-CoV-2 spike protein and drives viral entry into human cells. Therefore, TMPRSS2 represents an interesting target for COVID-19 therapy and in analogy to hepsin, peptidomimetic TMPRSS2 inhibitors were developed. Molecular docking studies were used to identify tripeptidomimetic binders of the TMPRSS2 catalytic site, which were then connected to an electrophilic serine trap to create covalent and stable inhibitors. The synthesized compounds inhibited TMPRSS2 and selected coagulation proteases with good off-target selectivity and lead candidates were stable in cell culture medium, blood serum and plasma for at least ten days. Selected compounds were shown to prevent SARS-CoV-2 pseudoparticle and genuine SARS-CoV-2 wildtype infection in cell culture with comparable efficiency as camostat mesylate. Notable, the compounds also inhibited infection of SARS-CoV-2 variants of concern carrying spike mutations D614G, B.1.1.7 and B.1.315. The results highly motivate further development of the best antiviral TMPRSS2 inhibitors with regard to pharmacokinetic analysis and the potential route of application e.g. in form of an inhalative spray.

Publication list

[1] <u>**Philip Maximilian Knaff**</u>*, Christian Kersten, Ramona Willbold, Carole Champanhac, Daniel Crespy, Rainer Wittig, Katharina Landfester, Volker Mailänder. **From in-silico to Experimental Validation: Tailoring Peptide Substrates for a Serine Protease.** *Biomacromolecules* **2020** *21* (4), 1636-1643.

[2] <u>**Philip Maximilian Knaff**</u>*, Christian Kersten, Lukas Wettstein, Patrick Müller, Tatjana Weil, Carina Conzelmann, Janis Müller, Stefan Pöhlmann, Markus Hoffmann, Tanja Schirmeister, Jan Münch, Volker Mailänder. **Tailored Inhibitors for the Serine Protease Hepsin**. To be submitted.

[3] Patrick Müller*, Hannah Maus, Stefan Hammerschmidt, <u>**Philip Maximilian Knaff**</u>, Volker Mailänder, Tanja Schirmeister, Christian Kersten. **Interfering with Host Proteases in SARS-CoV-2 Entry as a Promising Therapeutic Strategy.** *Current Medicinal Chemistry.* **2021**.

[4] <u>**Philip Maximilian Knaff**</u>*, Christian Kersten*, Lukas Wettstein*, Patrick Müller, Tatjana Weil, Carina Conzelmann, Janis Müller, Maximillian Brückner, Stefan Pöhlmann, Markus Hoffmann Tanja Schirmeister, Jan Münch, Volker Mailänder. **Development of Potent Peptidomimetic Inhibitors for the Serine Protease TMPRSS2**. Nature Communications. **2021** (Submitted).

[5] Lukas Wettstein*, Tatjana Weil, Carina Conzelmann, Janis A. Müller, Rüdiger Groß, Maximilian Hirschenberger, Alina Seidel, Susanne Klute, Fabian Zech, Caterina Prelli Bozzo, Nico Preising, Giorgio Fois, Robin Lochbaum, <u>**Philip Maximilian Knaff**</u>, Volker Mailänder, Ludger Ständker, Dietmar Rudolf Thal, Christian Schumann, Steffen Stenger, Alexander Kleger, Günter Lochnit, Benjamin Mayer, Yasser B. Ruiz-Blanco, Markus Hoffmann, Konstantin M. J. Sparrer, Stefan Pöhlmann, Elsa Sanchez-Garcia, Frank Kirchhoff, Manfred Frick, & Jan Münch. **Alpha-1 antitrypsin inhibits TMPRSS2 protease activity and SARS-CoV-2 infection.** *Nature Communications* **2020**.

Introduction

Serine proteases are a class of enzymes which regulate a high variety of physiological processes and are therefore of high importance. Those processes are diverse and range for example from blood coagulation to fibrinolysis, cell growth, apoptosis and digestion. Consequently, balanced activation of serine proteases is a prerequisite for well-functioning homeostasis. If unbalanced, as it occurs for example in untreated cancer, an overactivation of serine proteases may support life threatening malignancies. Besides in cancer, serine proteases play also an important role in viral infections which according to the world health organization in 2021 pose the greatest public health risk due to their inherent pandemic potential.[1]

Chapter A is part of the project 'Hepsin cleavable nanocapsules for precision medicine in prostate carcinoma' (HEPSIN-NANOLYSIS) which is supported by the Deutsche Forschungsgemeinschaft and a collaborative project with Max Planck Institute for Polymer Research, Johannes Gutenberg University Mainz and the Institute for Laser Technologies in Medicine and Metrology at the University of Ulm. This project intends to use the enzymatic activity of hepsin for the selective release of antitumor drugs to prostate cancer cells.

Previous studies worked on a hepsin responsive nanocapsulate system which aimed to work as a delivery vehicle for the selective opening and release of bioactive substances on the surface of prostate cancer cells. With this strategy we expected to increase substance bioavailability and to reduce drug-induced side effects. This system was developed by using peptide-based triblock copolymers with the structure polymer-peptide-polymer and the technique of nanoprecipitation was used to form capsules. Those nanocapsules were first shown to be applicable as a system to optically detect the enzyme and for this purpose, a literature known hepsin-cleavable peptide substrate was used for the preparation of the copolymer and a hydrophilic fluorescent dye was encapsulated in high, self-quenching concentrations. After degradation of this peptide substrate by hepsin, the dye was released and the recovery of fluorescence intensity over time could be monitored.

Despite the successful preparation of hepsin sensitive nanocapsules, this concept is not yet functioning for potential drug delivery application. High susceptibility towards hepsin and low degradation by body fluids are necessarily required for site-specific drug release and long circulation times. For this reason, the development of highly selective peptide substrates is urgently needed. This aspect was addressed in the first study of chapter A, where a combinatorial strategy was used to identify peptide binders for the active binding site of hepsin. With this, the development of novel substrates for hepsin was investigated and the enzymatic activity in terms of hepsin degradability and stability in body fluids was further analyzed. The

second study describes the development of peptidomimetic inhibitors to downregulate the activity of hepsin. Overall, this chapter shows the development of novel peptide substrates which can be used for the preparation of hepsin cleavable nanocarriers, as well as the development of inhibitors for hepsin.

Chapter B is part of the project 'Fight-nCoV' which is supported by the EU's Horizon 2020 research and innovation program. The goal of this project is to develop potent therapeutic interventions for COVID-19 and potential newly emerging viral diseases. COVID-19 is the disease of SARS-CoV-2 infection which causes a mild to moderate respiratory illness. However, people with pre-existing health conditions and older people are more prone to develop a more serious disease with strong symptoms like respiratory failure and shock.

SARS-CoV-2 carries a surface spike (S) protein on the surface which can be primed by host proteases present in the respiratory tract, primarily TMPRSS2, and drive infection of human cells. The recent success on the development of efficient vaccines raises hope to limit further spread of infection. However, newly emerging variants of concerns (VOCs) show increased transmissibility and are capable of escaping the immune system. In this regard, a new route for potential COVID-19 treatment is investigated which targets the host cell serine protease TMPRSS2 to prevent viral cell entry. Recent studies showed that blocking of TMPRSS2 by an inhibitor can efficiently prevent SARS-CoV-2 infection without causing toxicity.[2, 3] Until now, only repurposed serine protease inhibitors of TMPRSS2 have been reported and there is a clear gap for new pharmaceutical compounds. Chapter B comprises of two studies which describe antiviral inhibitors for the surface-active protease TMPRSS2. The first study describes the development of a new class of TMPRSS2 inhibitors, based on peptidomimetic structures. The second study aimed to identify endogenous TMPRSS2 inhibitors based on a proteomic approach. Further, the potential antiviral activity of the inhibitors to prevent viral SARS-CoV-2 wildtype and VOC infection was analyzed by collaboration partners.

Chapter A - Development of substrates and inhibitors for the serine protease hepsin

Chapter A is divided into three subchapters. The first subchapter gives an overview over the research area and theoretical aspects of the underlying principles which were used in this dissertation. The second subchapter presents the first study **[1]**, which aimed to develop new peptide substrates for the serine protease hepsin. Those results were published in a peer-reviewed journal (Biomacromolecules) and the text below is therefore in nearly all instances a word-by-word reproduction of the peer-reviewed and published paper. The third subchapter presents the second study **[2]** and describes the development of peptidomimetic inhibitors for hepsin. This study is to be published in a peer-reviewed journal.

[1] <u>**Philip Maximilian Knaff***</u>, Christian Kersten, Ramona Willbold, Carole Champanhac, Daniel Crespy, Rainer Wittig, Katharina Landfester, Volker Mailänder. **From in-silico to Experimental Validation: Tailoring Peptide Substrates for a Serine Protease.** *Biomacromolecules* **2020** *21* (4), 1636-1643.

[2] <u>**Philip Maximilian Knaff***</u>, Christian Kersten, Lukas Wettstein, Patrick Müller, Tatjana Weil, Carina Conzelmann, Janis Müller, Stefan Pöhlmann, Markus Hoffmann, Tanja Schirmeister, Jan Münch, Katharina Landfester, Volker Mailänder. **Tailored Inhibitors for the Serine Protease Hepsin**. *To be submitted.*

1. Theoretical background

1.1 Prostate cancer as potential therapeutic target

Among new cancer cases, prostate cancer (PCa) shows the second highest incidence rate (1 414 259 cases) after lung cancer and the fifth highest mortality rate (375 304 cases) for men worldwide in 2020 (Figure 1.1).[4] The incidence rate increases significantly with age and reaches nearly 60% for men over the age of 65 years.[4, 5] The disease proceeds often asymptomatic and requires minimal treatment however common symptoms include frequent urination, nocturia and urination hesitancy. In a more advanced disease stage PCa may become metastatic. Ethnical data indicate a higher incidence rate among African-American men with a mortality rate twice as high compared to Caucasians.[6] Asian men are less likely of getting PCa.[7] The risk of prostate cancer correlates with family history and men with first degree relatives affected had a twice fold increase risk to develop PCa.[8] Several lifestyle factors were also shown to negatively affect PCa, above all smoking increased mortality rate by 61%. The high incidence rates make prostate cancer a considerable public health concern and patients may benefit from new diagnostic and therapeutic concepts.

According to the worlds health organization prostate cancer can be classified either as benign or malignant tumors. The benign tumors comprises of prostatic intraepithelial neoplasis (PIN), atypical small acinar proliferation (ASAP) or benign prostatic hyperplasia (BPH) which phenotypically show the precancerous cell formation or prostate enlargement.[9] Malignant tumors are classified based on their origin as epithelial which is the most redundant form of prostate cancer originated from epithelial cells or sarcoma which is rare and formed from connective tissue of the prostate. As a measure for the progression and aggressiveness of the tumor, a biopsy is examined immunohistochemically and graded by two complimentary scales (Gleason-Score and TNM-classification). The Gleason-Score was created in 1966 by Donald F. Gleason and updated by the International Society of Urological Pathology (ISUP) in 2005 and 2015.[10] The architectural pattern of the tumor is classified into 5 grades and calculated by addition of the primary grade of the predominant pattern with the secondary grade of the second most predominant pattern (Figure 1.1, Table 1.1).[11] TNM-classification gives information about the disease stage through 3 criteria. 1) The size and expansion of the tumor, 2) lymph node metastases and 3) distant metastasis.[12] Prostate cancer screening uses the prostate specific antigen (PSA) which is organ-specific and widely used for diagnostic but it is a unspecific biomarker for prostate cancer and thus controversial. Increased levels of serum PSA can be found in prostate cancer, benign prostatic hyperplasia, prostatitis or through sport-related activities like cycling.[13,14,15] The low positive predictive value ranges between 25-40% and the sensitivity and specificity of the testing ranges from 70 to 90% and 20 to 40%. [16,17] This

results in unnecessary prostate biopsies and wrong positive screening results. Thus, PSA testing as precautionary measures is not unambiguously recommended and causes unnecessary measures for men with prostate cancer that does not require treatment and men with a false-positive screening results.[18, 19] Besides PSA, the serine protease hepsin is discussed as new potential prostate cancer biomarker due to the significant upregulation in tumor tissue versus benign tissue. A recent study reported serum hepsin level specificity of 89%.[20]
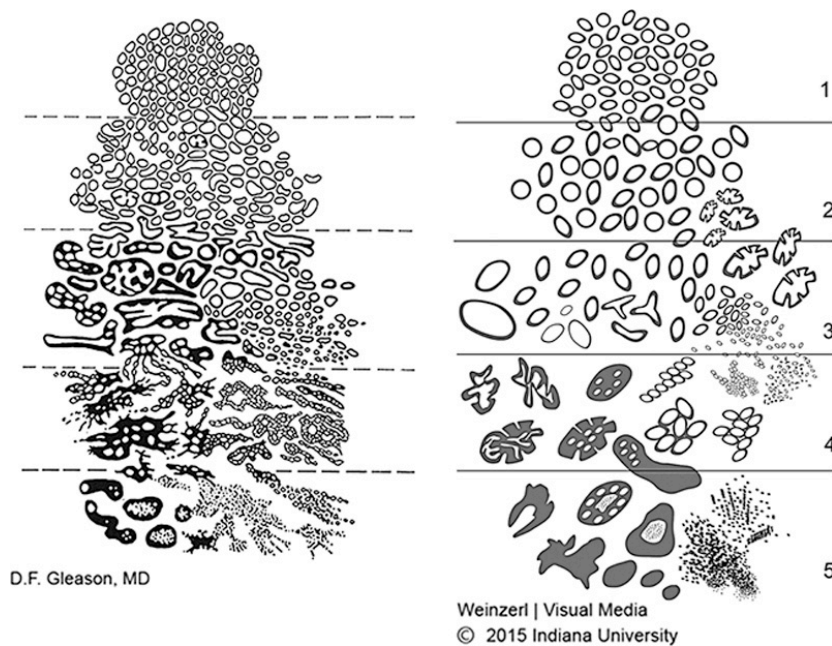
A
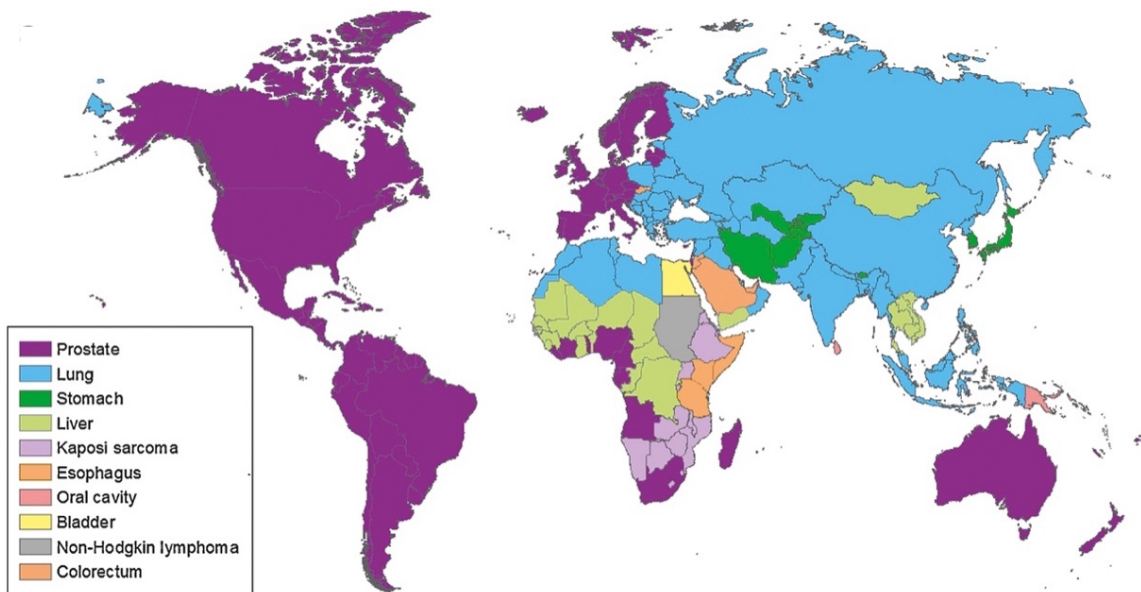
B

Figure 1.1. **A** Most commonly diagnosed cancer among men worldwide in 2008. The figure was adapted from Center M. M. et al.[41, 42] **B** Histologic patterns of prostate adenocarcinoma: original from Gleason F. D. (left) and 2015 modified ISUP Gleason schematic (right).[10, 21, 22]

Table 1.1. 2014 ISUP Gleason score and gleason grade groups. The table was adapted from Shah B. M. et al.[23]

Gleason Grade Group	Score	Definition
1	3+3 = 6	Only individual discrete well-formed glands
2	3+4 = 7	Predominantly well-formed glands with a lesser component ofpoorly/fused/cribriform glands
3	4+3 = 7	Predominantly poorly formed/fused/cribriform glands with a lesser component of well-formed glands
4	5+3, 3+5, 4+4 (Gleason score 8)	Only poorly formed/fused/cribriform glands (>95%) or Predominantly well-formed glands and lesser component lacking glands or predominantly lacking glands and a lesser component of well-formed glands
5	Gleason scores 9 and 10	Lack of gland formation (or with necrosis) (>95%) with or without formed/fused/cribriform glands

1.2 Therapy of prostate cancer

After diagnosis of prostate cancer, therapy options should be applied dependent on the disease progress and patients' health condition. Since most invasive therapy options like surgery, radiotherapy or chemotherapy have severe side effects, if applicable, non-invasive measures are often preferred and termed as expectant management. This includes active surveillance, which intends to intervene in case of disease progression or watchful waiting, which is meant for patients with severe health conditions, who will not benefit from an invasive therapy. For low-risk prostate cancer patients, active surveillance was shown to not compromise the cancer-specific survival at 10 years and is therefore in many cases a suitable option to prevent side-effects of invasive, aggressive treatments.[24] In case of disease progression of a localized prostate tumor or if active surveillance is not feasible active treatments are initiated which entails the use of radical prostatectomy, radiation therapy or a combination thereof (Figure 1.2). For metastatic tumor stages, a hormonal androgen deprivation therapy (ADT) in combination with chemotherapy is often reserved. ADT was observed by Huggins in 1941 and it aims to deprive the prostate tumor of androgens which reduces androgen receptor signaling that drive cell grow and proliferation.[25] However, within 1-2 years in most cases the disease emerges and progresses to the androgen-independent form of metastatic castration-resistant prostate cancer (mCRPC) which is incurable with median survival rates of 15 months.[25-27] The use of the drug docetaxel, a chemotherapeutic binder for microtubules preventing disassembly, increases the quality of life and survival for

nearly two months in men with mCRPC and was approved in 2004 by the FDA.[28, 29 30, 31] A recent clinical study showed that chemohormonal therapy by combination of ADT with the cytostatic docetaxel improved significantly the overall survival over ADT alone (Clinical Trials. gov number, NCT00309985).[32, 33] The therapeutic options of ADT, radical prostatectomy, radiation and chemotherapy have severe side effects and the development of new and more effective therapeutic or diagnostic approaches which may utilize the serine protease hepsin, as biomarker and metastasis supporting protein is of importance and subject of chapter A in this dissertation.[34]
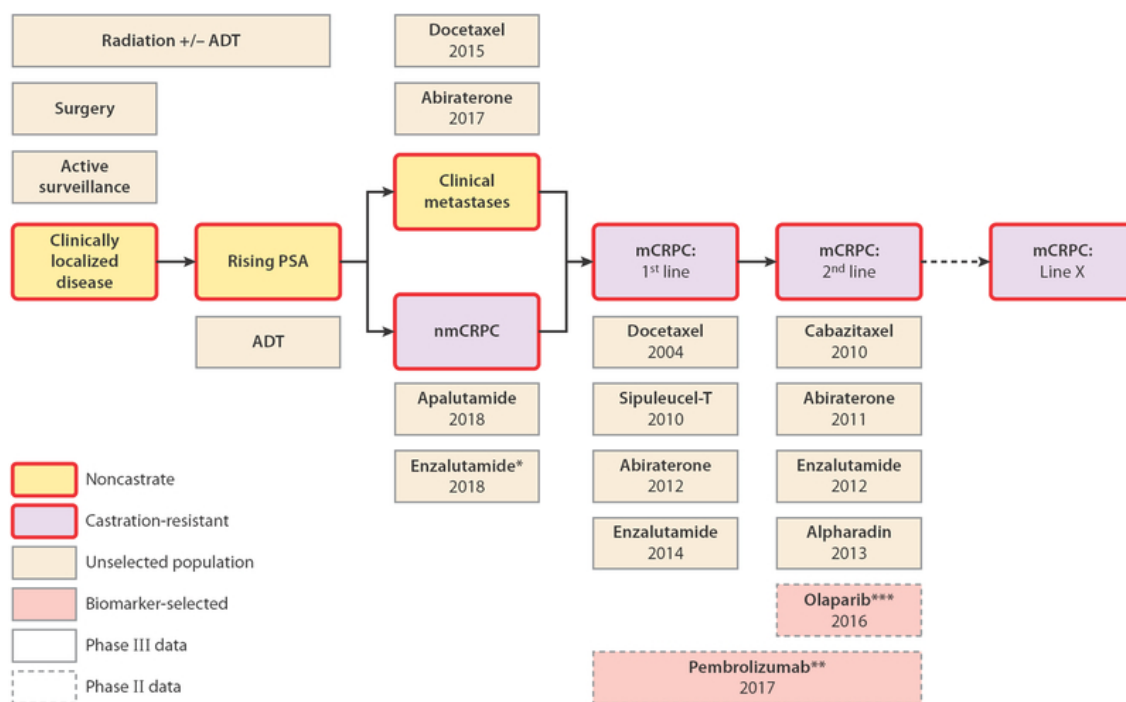

Figure 1.2. Clinical states of prostate cancer with management options in different clinical stages. mCPRC, metastatic castration-resistant prostate cancer; nmCPRC, nonmetastatic castration-resistant prostate cancer; PSA prostate-specific antigen. This Figure was adapted from Prostate Working Group 3 and Teo Y. M. et al.[31, 35]

1.3 Hepsin

Hepsin was first described by Davie and coworkers in 1988, who identified a novel trypsin-like protease in human liver by cDNA library screening using the colony hybridization technique. They named the protease hepsin after the greek word *hepar* - liver.[36] The gene of hepsin (*HPN*) is located on human chromosome 19 at q11-13.2 and result in the transcription of an mRNA with a size of 1.85 kilobasis.[37] The mRNA is translated into a 417 amino acid long single polypeptide chain with a calculated mass of ~45 kDa (due to glycosylation, the full-length

protein migrates at 51 kDa in SDS-PAGE) and activated through proteolytic cleavage between Arg162 and Ile163, thus separating the protease domain from the rest of the protein.[36, 38] The two resulting domains of hepsin are connected through a covalent disulfide bonds between Cys277 and Cys153.[39] Previous studies showed a concentration depended autoactivation of hepsin to the active, disulfide-linked two chain form, but it is still unclear if other proteases are involved in the activation process.[40]

Due to evolutionary relationships, some proteases share a similar folding of the catalytic active site and are therefore assigned to clans and families according to the MEROPS database.[41] The hepsin protease belongs to the serine and cysteine endoprotease clan (clan PA) among the family of trypsin-like proteases (family S1, subfamily A) and can be found under the MEROPS ID of S01.224.[41] This protease contains a characteristic type II transmembrane serine protease domain structure with a cytosolic N-terminus, a central hydrophobic domain which is involved in anchoring the protease in the cell membrane, a scavenger receptor cysteine rich domain (SCRD) and an extracellular C-terminus bearing the catalytic active domain (SPD) enabling localized cell surface proteolysis. Thus, hepsin is considered to be a member of Type II Transmembrane Serine Proteases and is further subgrouped in the hepsin/TMPRSS2 subfamily together with TMPRSS2-4, MSPL, Spinesin, and Enteropeptidase (Table 1.2).

Table 1.2. The family of transmembrane serine proteases. TTSP are divided into four subfamilies based on the highly variable stem region.

TTSP subfamily	Members
HAT/DESC	HAT, DESC1 HAT-like 4 TMPRSS11A HAT-like 5
Hepsin/TMPRSS2	Hepsin TMPRSS2-4 MSPL Spinesin Enteropeptidase
Matriptase	Matriptase 1-3 Polyserase-1
Corin	Corin

The catalytic active site of this serine protease is formed by the residues His203, Asp257 and Ser353 (His57, Asp102, Ser195 in standard chymotrypsin numbering system) which act as a coordinated network to induce covalent catalysis to hydrolyze the peptide bond of a substrate. As show in Figure 1.3, the proton of the hydroxyl group from Ser353 forms a hydrogen bond with the imidazole ring of His203 which, in turn, forms a hydrogen bond with the carboxyl group of Asp257. This arrangement of the catalytic triad ensures correct alignment and increases the nucleophilicity of the hydroxyl group through polarization. After binding of a substrate in the active site of the enzyme, the oxygen atom of the serine residue is now capable of performing a nucleophilic attack on the carbonyl atom of the peptide bond and the proton of the hydroxyl group can be directly transferred to the imidazole ring. A covalent tetrahedral transition state of the substrate is formed and the negative charge of the alkoxide is stabilized using the enzyme's backbone amides.[42] This stabilization is mainly driven through coordination with NH-groups and the respective pocket referred to as oxyanion pocket. The transition state collapses, and the attacked peptide bond of the substrate disintegrate into two fragments, whereas one fragment is covalently bound to the enzyme (acyl-enzyme intermediate). In analogy to the above-mentioned nucleophilic attack, a water molecule can then be coordinated by the imidazole ring of the histidine residue and hydrolyze the acyl-enzyme intermediate. The second fragment is released, and the catalytic triad is regenerated for a new cycle.
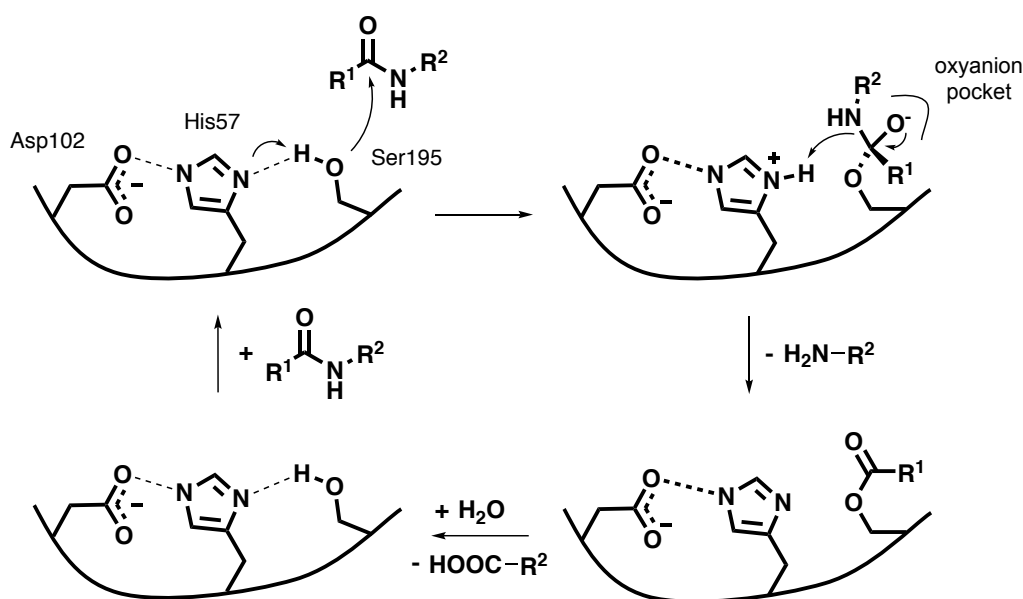
Figure 1.3. Catalytic triad. Asp102, His57 increase the nucleophilicity of Ser195 and act in concert to induce hydrolysis of the peptide bond. This figure was adapted from Berg et al. and Millies et al.[43, 44]

According to the Schechter & Berger nomenclature the specificity of proteases can be divided into subsites. "S" for subsite of the active site of the protease and "P" for the corresponding subsite of the substrate.[45] Protease substrate specificity can be determined by using a database like MEROPS, PS-SCL, or testing of individual substrates.[41,46,47] The MEROPS database collect experimental data from publications and creates a specificity matrix by listing the mentioned amino acid in the respective position. Positional scanning of substrate combinatorial libraries (PS-SCL) is a popular screening method to analyze the S4-S1 pocket preference. This strategy utilizes different libraries in which a fluorescent molecule like ACC is fixed at P1′ position and an amino acid of interest is fixed while the others are randomized by using isokinetic mixtures of amino acids.[48] Upon cleavage at the scissile bond P1↓P1′ the fluorophore is released which enables optical detection. This strategy serves as an experimental tool to identify substrate specificity for the P-site and thus, a high-resolution crystal structure of the extracellular region of hepsin bound with the tetrapeptide Lys-Gln-Leu-Arg was identified (PDB-ID: 1Z8G, resolution 1.55 Å).[46] This crystal structure allowed deeper insights in structure-activity relationships. As commonly observed with trypsin-like serine proteases, hepsin was shown to have a clear preference for Arg in P1 position which interacts with a strong salt bridge with Asp347 in the S1 pocket. The Leu in P2 position, binds to a small, S2 pocket (formed by residues His203, Tyr243, Asn254, Asp257, Ser376) with van der Waals interactions. The Glu in P3 position forms a hydrogen bond with Tyr301 and the Lys in P4 position binds the S4 pocket (formed by Trp377, Glu252, Glu 253, Asn254, Gln331) which reveals overall a negative electrostatic potential.[46] Yet the PS-SCL screening strategy has

some substantial drawbacks. Only S-site preferences can be investigated and the close proximity of the fluorophore to the cleavable bond in P1′ position may influence hydrolysis behavior. Furthermore, MEROPS database and PS-SCL screening strategy ignore cooperativity effects of adjacent amino acids. Such effects can be of intramolecular nature, where neighboring amino acids interact within the protease substrate and change accessibility of the cleavage site. They may also be of an intermolecular nature, where binding of a substrate residue in a protease subsite affects the binding affinity of neighboring amino acids. This phenomenon has been termed subsite cooperativity and is observed in many proteases.[49, 50]

The hepsin gene is ubiquitously expressed at low expression levels but shows a high expression in the liver, spleen and pancreas.[51-53] Hepsin plays a particular role through activation of growth factors. The growth zymogens pro-hepatocyte growth factor (pro-HGF) and pro-macrophage stimulating protein (pro-MSP) were identified as preferred endogenous substrates for the transmembrane serine proteases hepsin, matriptase and for the plasma protease hepatocyte growth factor activator (HGFA).[46, 53, 54, 55] The similar substrates of the proteases indicate overlapping functions. After proteolytic processing of pro-HGF and pro-MSP the active proteins HGF and MSP can bind to the tyrosine kinase receptors c-MET (hepatocyte growth factor receptor; HGFR) and RON (macrophage stimulating protein receptor; MST1R), respectively, and subsequent activate downstream signaling pathways. Activation of those receptors triggers events like cellular migration, proliferation and survival.[56, 53] Hepsin was also shown to directly degrade laminin, a protein of the basement membrane, and activate pro-matrix metalloproteinases thus promotes ECM degradation.[57, 58] Furthermore, it also amplifies plasmin through activation of pro-urokinase-type plasminogen activator (pro-uPA) which triggers the conversion of plasminogen to plasmin and the plasmin-mediated disruption of basement membranes.[59] Doxycycline-induced hepsin expression was shown to downmodulate desmosomal structures in epithelial cell membranes which are involved in cell cohesion and hemidesmosomes which connect epithelial cells with the basal lamina.[60]

Hepsin is not an essential gene and hepsin deficient mice were viable, fertile and grow normally.[61, 62] Abnormalities could be observed in increased levels of pro-HGF and reduced levels of active HGF which interfered with liver metabolism by reduced glycogen, lipid and protein synthesis, with adipocyte differentiation by increased energy metabolism, with inner ear by incomplete cochlear formation and profound hearing loss and with kidney by decreased uromodulin levels for urinary tract defense.[63, 64, 65, 66]

The TTSP-mediated activation of growth factor is mainly regulated in pericellular environment by two endogenous inhibitors named HGFA inhibitor type 1 and type 2 (HAI-1, HAI-2). In cancers like prostate, ovarian, renal and breast cancer increased overexpression of the hepsin gene was found in epithelial tissue.[67, 68, 69, 70] *HPN* was found to be upregulated in 90% of

prostate tumors and is often one of the highest expressed genes with up to 10-fold higher expression level compared to benign prostate tissue. [52, 71, 68, 72] Increased levels of hepsin are maintained throughout prostate cancer progression from primary tumors to metastasis and correlated with the Gleason score.[68, 73, 74] The oncogenic overexpression of hepsin leads to proteolytically shedding of HAI-1 which further reinforces the hepsin-to-HAI-1 ratio and thus increases proteolytic activity.[60] Such activity was proven *in vivo* to promote cancer progression and the formation of metastasis of a primary tumor to liver, lung and bone through disruption of the basement membrane.[75-77] Long term oral application of a small-molecule inhibitor *in vivo* was shown to block the cell-surface active protease hepsin and inhibited liver, lung and bone metastasis with no visually inspectable deficiencies in mice.[71, 78]

1.4 Solid phase peptide synthesis

The beginning of the area of peptide synthesis was marked by E. Fischer who synthesized the first dipeptide, glycyl-glycine, in 1901.[79] Since then several small biologically active peptides could be prepared like the synthesis of the hormone oxytocin by V. Vigneaud (Nobel Prize in Chemistry, 1955).[80] With increasing numbers of amino acids, the synthesis and purification of larger peptides is more challenging. To overcome those difficulties R. Merryfield presented in 1963 a new concept for the synthesis of a peptide (Nobel Prize in Chemistry, 1984).[81, 82] This concept included the use of insoluble solid resin particles to synthesize the peptide chain on. The amino acids can be coupled in a stepwise manner and the growing peptide chain can easily be handled and tuned. This solid support allows the removal of uncoupled excessive reagents and by-products through washing steps which tremendously facilitates the purification process and efficiency of the synthesis. In the following years more research was carried out and several strategies were developed to improve the solid-phase peptide synthesis (SPPS) such as automatization of the synthesis, application of microwave systems, novel solid supports and coupling reagents. Nowadays in SPPS, the transient protection groups for the amino groups that form the peptide bond determines the strategy of the chemistry. Widely used are N-terminal protection groups like *tert*-butyoxycarbonyl (Boc) group which are sensitive to acids or fluoren-9-ylmethyloxycarbonyl (Fmoc) group which is sensitive to bases.[81, 83] The chemistry of the Fmoc-group is often preferred due to the orthogonal protection scheme and milder reaction conditions. In the scope of this dissertation the strategy utilizing the Fmoc-group was used.
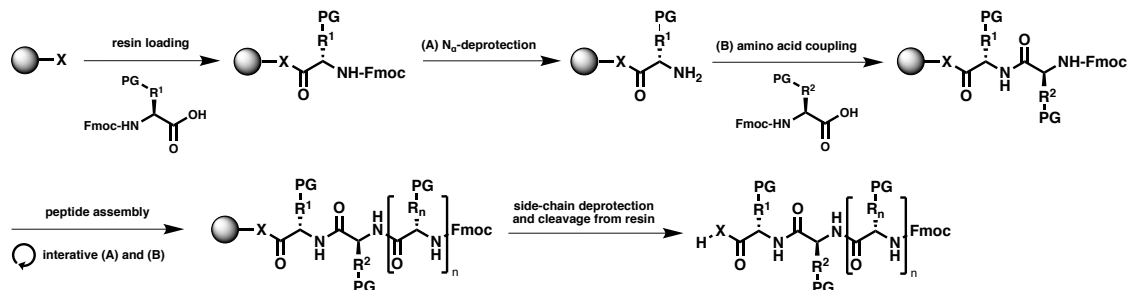
1.5 The Fmoc strategy

In the following section the principle of the Fmoc strategy and the most important underlying mechanisms will be presented further in detail. As shown in Figure 1.4, the synthesis starts by loading the resin with the first Fmoc protected amino acid. In this process, the carboxylic group of the amino acid is attached on a functionalized resin which allow N-terminal coupling of the subsequent amino acid. The coupling process consists of two iterative steps for each amino acid. In the first step, the amino protection group Fmoc is cleaved of with a mild base to yield the free amine (step (A) N_α-deprotection). After removal of the protection group, the free amine can then be coupled with the carboxylic group of the next Fmoc protected amino acid (step (B) amino acid coupling). To this point, the peptide chain was lengthened by one amino acid under formation of a peptide bond. Elongation of the chain with further amino acids can be achieved through alternating step (A) and step (B) until the desired sequence has been synthesized. In a final step, the acid labile side chains are deprotected and the immobilized peptide on the resin can be cleaved to yield the crude peptide.

Fmoc is a base-labile amine protection group which can readily be cleaved with a secondary amine e.g. piperidine. A detailed mechanism of the deprotection is shown in Figure 1.4. The electron-withdrawing fluorene makes the proton at the 9-position acidic which can be abstracted by using a 20% piperidine solution in DMF. The stabilized anion disintegrates in an elimination mechanism (*E*1*cB)* which triggers the release of carbon dioxide, dibenzofulvene and the fully deprotected amine.[84] The latter is highly reactive and will be immediately trapped by the secondary amine to form the dibenzofulvene-piperidine adduct.
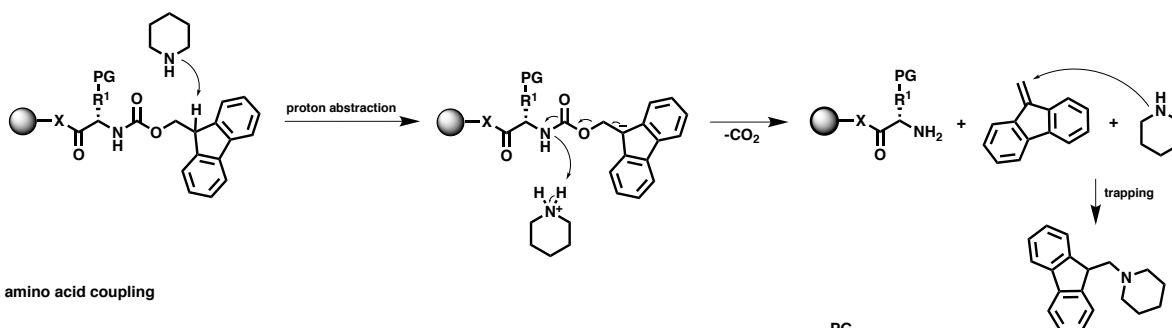
The solid phase peptide synthesis relies on high amino acid coupling yields for the successful synthesis of a peptide. Mixing of an amine with a carboxylic acid results first in a acid-base reaction and the formation of a salt. A condensation reaction yielding the amide can be either achieved at high temperatures or with reagents which accelerate the amide bond formation. Therefore, a panel of highly effectively peptide-coupling reagents have been developed which activate the carboxyl group through introduction of an electron withdrawing moiety which facilitates nucleophilic attack by an amino group.[85] The detailed mechanism of this reaction using the coupling reagents PyBOP/DIEA is presented in Figure 1.4. To activate the carboxylic acid, it is first deprotonated by a base e.g. diisopropylethylamine and the nucleophilic carboxylate anion is then coupled with PyBOP to create an activated acylphosphonium species.[86] This species represents a good leaving group and makes the nucleophilic attack of the amine from the growing peptide chain feasible. A new peptide bond is then formed under aminolysis. The driving force of this reaction is the formation of the phosphoric amide. Besides PyBOP a broad range of coupling reagents are nowadays commercially available. Popular are dicyclohexylcarbodiimide (DCC), diisopropylcarbodiimide (DIC), PyBOP (Benzotriazol-1-

yloxy-tripyrrolidino-phos-phonium hexafluorophosphate) and HCTU (2-(6-chloro-1H-benzotriazol-1-yl)-N,N,N',N'-tetramethylaminium hexafluoro-phosphate).[87]

General scheme of the Fmoc-strategy

(A) N$_\alpha$-deprotection
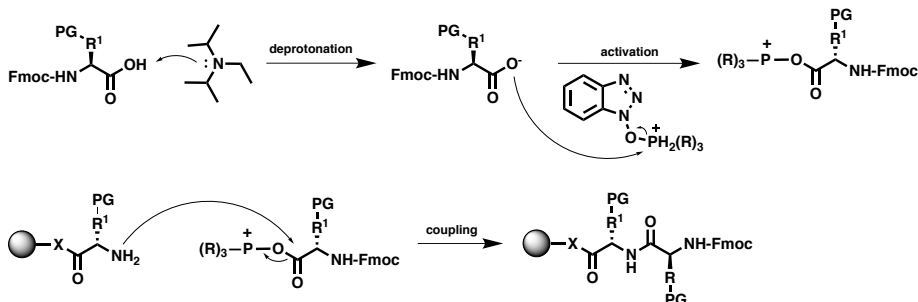

(B) amino acid coupling

Figure 1.4. Scheme for the solid-phase peptide synthesis of peptides using the Fmoc strategy. (A) Mechanism for the removal of the Fmoc group with piperidine. (B) PyBOP-mediated amide bond formation mechanism over the acyloxyphosphonium pathway. X = O or NH. PG = protection group. This figure was adapted from Biron et al. .[88]

1.6 Förster resonance energy transfer

The phenomenon of energy transfer between two molecules was theoretically described by Theodore Förster in 1946 and it describes the radiationless energy transfer from a donor molecule to an acceptor molecule.[89] This energy transfer is strongly depended on the distance between these molecules and rapidly falls off with increased distance.[90, 91] Most moieties used for traditional FRET are efficient in a distance less than 10 nm and the energy transfer was

shown to be suitable as a "molecular ruler" to investigate biological interactions.[92] A typical situation for FRET is shown in Figure 1.5. FRET is frequently used with donor and acceptor molecules as fluorophores and thus also called fluorescence resonance energy transfer but chromophores can also act as quenchers and absorb the transferred energy. The FRET efficiency is the proportion of excited states that transferred the energy to the acceptor moiety and can be calculated with equation 1.

$$E = \frac{R_0^6}{r^6 + R_0^6} \tag{1}$$

E is the FRET efficiency which depends on the distance r of the fluorophores to the inverse sixth power and R_0^6 is the distance at which FRET efficiency is $E = 50\%$.[92] The FRET efficiency can also be calculated using the fluorescence intensity of the donor in presence of the acceptor (I_{DA}) or in absence of the acceptor (I_D) using equation (2).

$$E = 1 - \frac{I_{DA}}{I_D} \tag{2}$$

Besides the distance, the transfer rate is dependent on the spectral overlap integral $J(\lambda)$ of the donor emission and the acceptor excitation spectra and the dipole orientation factor κ^2 of each fluorophore. κ^2 is dependent on the angle between the donor and acceptor dipole and can have values between 0 (perpendicular) and 4 (aligned). For freely rotatable fluorophores the averaged value is assumed to be 2/3.[93] The Förster distance R_0 can be calculated using equation 3.[94] The character n represents the refractive index of the solution and Q_D represents the fluorescence quantum yield of the donor in the absence of a quencher and describes the efficiency of the excited state being emitted as fluorescence.

$$R_0 = 0.211 \times \sqrt[6]{\kappa^2 n^{-4} Q_D J(\lambda)} \tag{3}$$

The Jablonski Diagram illustrates the FRET effect (Figure 1.5). Excitation of the donor molecule by absorption of a photon promotes electrons from the ground state S_0 to an excited energy level S_1. This electron rises to different vibrational levels and a part of this energy dissipates through vibrational energy relaxation. In absence of an acceptor molecule, the electron in the excited state relaxes to the ground state and emits a photon (fluorescence). The emitted photon has a higher wavelength which is described by the phenomena called Stokes shift.[95] In presence of a suitable acceptor the energy is transferred to the acceptor molecule in a non-radiative manner through dipole interactions and the excited electron in the donor returns to the ground state without light emission.[96]

1.7 Internally quenched fluorescent substrates

Since the discovery of FRET it became an important tool in the field of biochemistry with numerus applications e.g. interactions between proteins, in membranes, tracking of cellular signaling pathways or in small-molecule fluorescent probes.[97-100] FRET-based sensor for biomacromolecules include the type of internally quenched fluorescent substrates wherein donor and acceptor molecules are incorporated into a single probe. Those molecules are connected through a peptide which act as a recognition sequence for an enzyme to afford an enzyme-specific FRET sensor. An exemplary structure of an IQF substrate used in this dissertation is shown in Figure 1.5. The peptide comprises of eight amino acids and act as a substrate for the serine protease hepsin. The fluorophores (7-methoxycoumarin-4-yl)acetyl (Mca) and the chromophore 2,4-dinitrophenol (Dnp) are bound to the ε-amino group of lysine residues adjacent to the N-terminal arginine and C-terminal glycine residue. The stretched IQF substrate shows a maximum length of 3.9 nm between the lysine residues and lies within the R_0 value for the Mca/Dnp pair which was calculated to be 3.65 nm (Table 1.3).[101]

Table 1.3. Spectral properties of the FRET pair Mca/Dnp.[101, 102]

FRET pair	Donor Excitation Maximum	Donor Emission Maximum	Donor Quantum Yield	Acceptor Molar Extinction Coefficient	Förster Distance
Mca/Dnp	325 nm	392 nm	0.72	17300 M^{-1}cm^{-1}	3.65 nm

As prepared, the fluorophore Mca can be excited at 325 nm and the emitted fluorescence at 392 nm can efficiently be quenched by Dnp (Figure 1.5). Upon proteolytic cleavage of the substrate by hepsin, Mca loses the proximity to Dnp which prevent efficient energy transfer and the emitted fluorescence of Mca can be detected.[101]
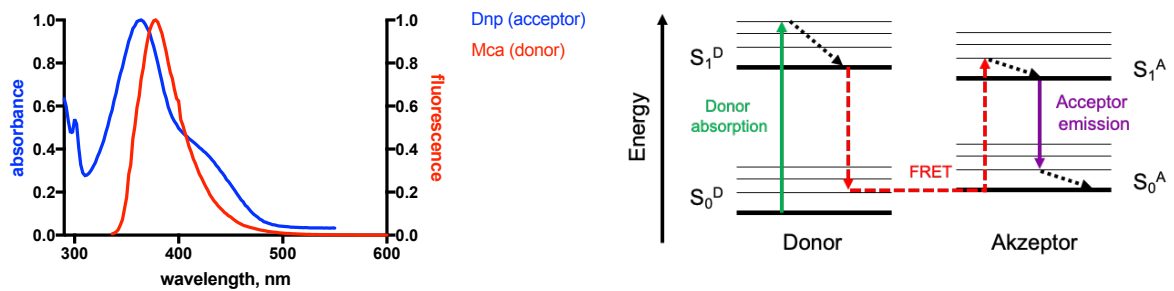
A
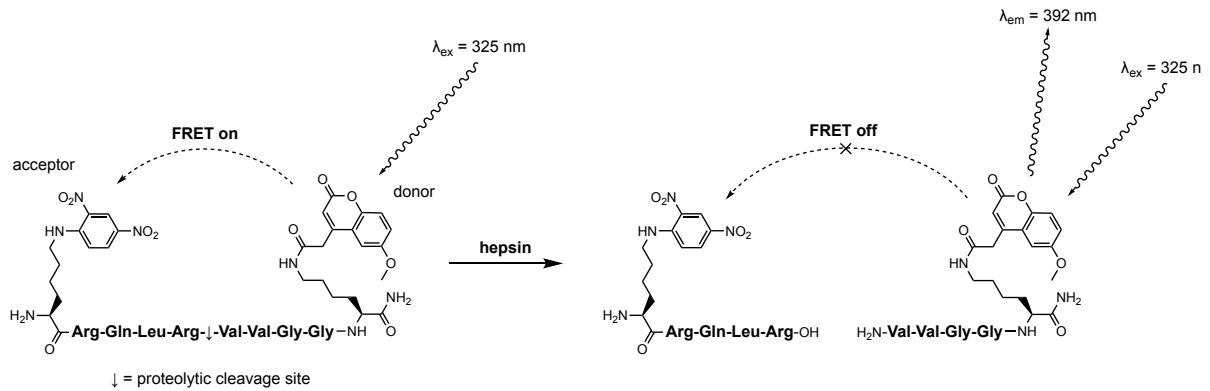
B

Figure 1.5. A The basic principles of IQF substrates in protease assays with Mca (C-terminal) as donor moiety and Dnp (N-terminal) as acceptor moiety. **B** Spectral overlap $J(\lambda)$ of the donor emission and the acceptor excitation spectra (left panel). Jablonski Diagram explaining the effect of FRET (right panel). This figure was adapted from Hochreiter et al.[96]

1.8 Enzyme kinetics

The chemical reaction of a reactant to a product in the absence of a catalyst proceeds usually very slow, especially in physiological surroundings.[103] The reactant and the product have different energy levels which are separated through an energetic barrier. This barrier is known as the activation energy $\Delta G_{no}^{\ddagger}$ which is necessary to transform the reactant in a high-energetic transition state shortly before reacting to the product. In this process, the activation energy $\Delta G_{no}^{\ddagger}$ is released together with the free Gibbs energy ΔG during the reaction. A catalyst can provide an alternative mechanism through molecular interaction with the reactant which lowers the activation energy $\Delta G_{enz}^{\ddagger}$ (Figure 1.6). As a result, more reactants are leveled up in the transition-state which fastens the reaction but do not change reaction equilibria.

Enzymes are specialized catalysts for biochemical reactions. In an enzyme-mediated catalytic process, the substrate binds to the so-called 'active site' of the enzyme which is formed through a three-dimensional arrangement of its amino acids. Substrate specificity can be generated due to the chemical properties of its amino acids and the driving force of substrate binding significantly depends on ionic- or hydrophobic interactions, hydrogens bonds or Van der Waals forces.

The catalytic activity of the enzyme can be influenced by several factors e.g., pH, temperature, presence of inhibitors or substrate concentration. The best-known model which describes enzyme kinetics was proposed by Leonor Michaelis and Maud Menten in 1913 (Scheme 1.1).[104] An enzyme E bind a substrate S through the key and lock interaction and forms in the first step an enzyme-substrate complex ES. This reaction is an equilibrium, and the association and dissociation rate constants of this complex are described with k_1 and k_{-1}, respectively. If no dissociation occurs, ES can be converted with the rate constants k_2 and k_{-2} to the enzyme-product complex EP. From this complex, the enzyme may dissociate and release the product P (Scheme 1.1). Since the formation of product is at the beginning of such an enzymatic reaction less, the reverse reaction to the EP complex is neglectable and a simplified reaction model explaining the Michaelis-Menten model is used (Scheme 1.1).
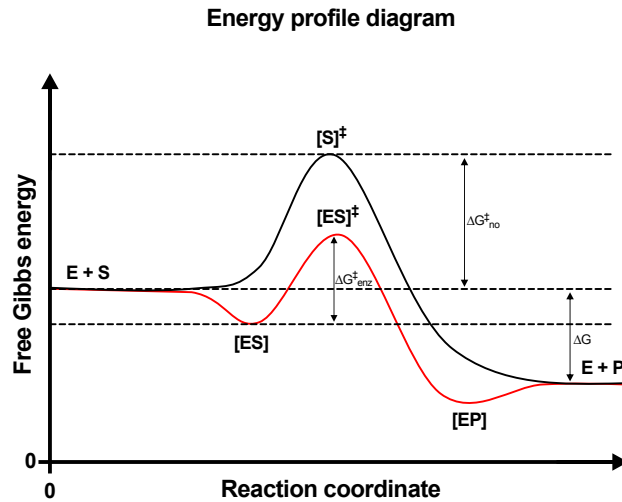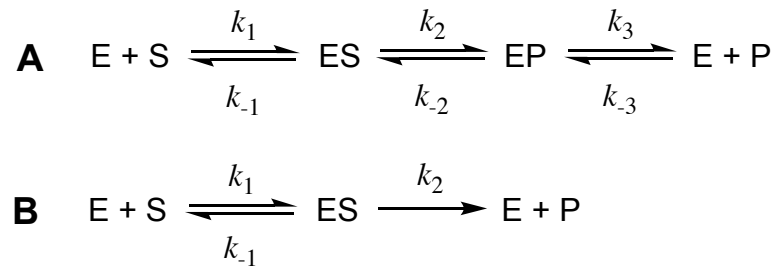
Energy profile diagram

Figure 1.6. Energy profile diagram of an uncatalyzed reaction (black line) or enzyme catalyzed reaction (red line) along a reaction coordinate.

A $\quad E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} EP \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} E + P$

B $\quad E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \overset{k_2}{\longrightarrow} E + P$

Scheme 1.1. A Enzymatic reaction pathway of a substrate to its product. **B** Simplified reaction equation explaining the Michaelis-Menten model.

The reaction rate of the product formation is depended on the substrate concentration and its dependency under steady-state of ES concentration can be described in a typical Michaelis-Menten plot (Figure 1.7). This hyperbolic curve shows at low substrate concentrations a proportional increase of the reaction rate which at some point reaches saturation. Here, all enzymes are saturated, and addition of more substrate does not increase the reaction velocity. The substrate concentration, at which half of the maximum velocity is reached is the K_M value and this represents a good measure for the affinity of the enzyme to its substrate. With programs providing nonlinear regression analysis, the asymptotic curve can be plotted and the values for V_{max} and K_M determined. The progress of the curve can be described with the Michaelis-Menten equation (4).

$$V = \frac{V_{max} \times [S]}{K_M + [S]} \tag{4}$$

Another possibility which uses linear regression to determine enzyme kinetics is called Lineweaver-Burk diagram. This diagram uses a straight-line equation with the form $f(x) = m \times x + n$ which can be plotted through the data points by double-reciprocal

representation of the reaction rate against the substrate concentration (Figure 1.7, equation 5).[105]

$$\frac{1}{V} = \frac{K_M}{V_{max}\,[S]} + \frac{1}{V_{max}} \tag{5}$$

The slope of the curve corresponds to the term K_M/V_{max} and a flat progress indicates high enzymatic activity. From the intersection of the ordinate, the reciprocal reaction rate $1/V_{max}$ can be retrieved while the intersection with the abscissa discloses the reciprocal value of the Michaelis-Menten constant $1/K_M$ which can be determined through extrapolation of the data points. An exemplary Michaelis-Menten and Lineweaver-Burk plot of the enzyme TMPRSS2 towards the fluorogenic substrate Boc-Gln-Ala-Arg-AMC is shown in Figure 1.7.
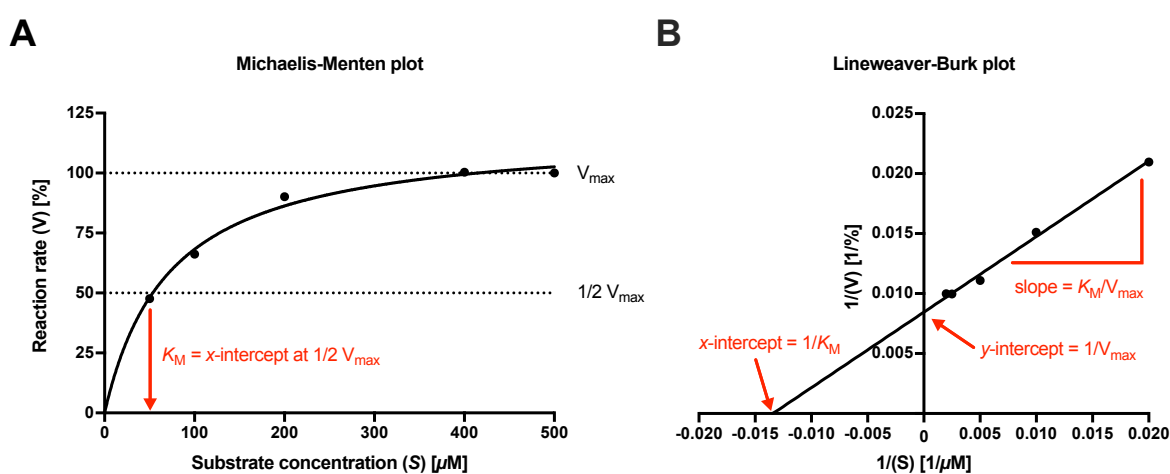

Figure 1.7. A Michaelis-Menten plot. Shown is the plot of the reaction rate in dependance of the substrate concentration using nonlinear regression. **B** Lineweaver-Burk plot. Double-reciprocal transformation of the Michaelis-Menten equation allows determination of kinetic parameters using linear regression. V_{max} = maximum velocity, K_M = Michaelis-Menten constant.

As previously discussed, the Michaelis-Menten constant is a measure for the enzyme's affinity to its substrate by formation of the enzyme-substrate complex ES. Once formed, the rate of product formation depends also on the capability of the enzyme to convert the substrate into product and is expressed as the rate constant k_2. This rate constant is also known as the turnover number $\underline{k}_{cat}$ which represents the maximum number of converted substrates per second into products at V_{max}. The fraction of both parameters k_{cat}/K_M describes the catalytic efficiency and can be used to compare the relative rates of reaction of different substrates to the enzyme.

The regulation of enzymes is mainly driven by inhibitors which can be either produced by the organism (endogenous) itself or are derived synthetically (exogenous) in the laboratory. There

are several types of inhibition like e.g. reversible or irreversible. In this dissertation, the type of reversible inhibitors with a substrate-like competitive binding mode is addressed. Those inhibitors resemble a substrate and compete with them for the catalytic active site of the enzyme. Upon binding of the inhibitor, the active site is blocked which prevents conversion of the substrate. Therefore, a higher substrate concentration is needed to reach the maximum reaction rate V_{max} which also increases Michaelis-Menten constant K_M. The characteristic changes on enzyme kinetics can be displayed using the Lineweaver-Burk diagram as shown in Figure 1.7. The y-intercept does not change which indicates that the maximum reaction rate V_{max} stays unaffected, instead the slope of the curve get steeper which shows the increase of the K_M value.

The influence of an inhibitor can be analyzed in assays where the enzyme degrades a substrate and upon degradation e.g. a fluorophore is liberated which can be monitored. In presence of an inhibitor, the active binding site is blocked which lowers the velocity of fluorophore liberation. For the comparison of several inhibitors, the potencies are determined which are usually analyzed as the half maximal inhibitory concentration (IC_{50}). The IC_{50} value is the concentration where half of the enzyme's reaction rate is reached, and it can be derived through non-linear regression of the velocity against the inhibitor concentration (Figure 1.8). Due to the strong substrate dependency of the IC_{50} value, a measure for the absolute inhibitory affinity is usually calculated using the Cheng-Prusoff equation 6.[106]
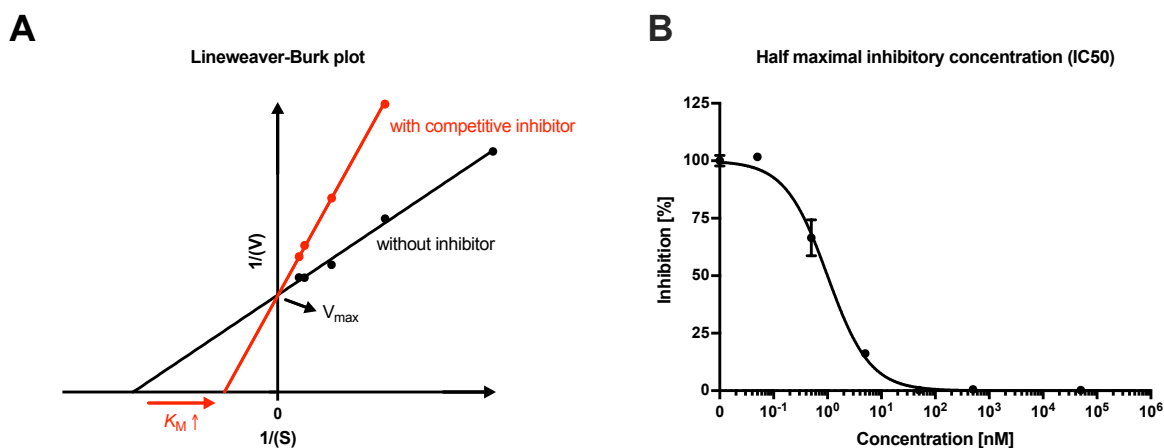
$$K_i = \frac{IC_{50}}{1 + \frac{[S]}{K_M}} \tag{6}$$

Figure 1.8. A Lineweaver-Burk plot showing the influence of competitive reversible inhibitors on enzyme kinetics. **B** Calculation of the half maximal inhibitory concentration (IC_{50}) of CM against the enzyme TMPRSS2. V_{max} = maximum velocity, K_M = Michaelis-Menten constant.

2. From *in silico* to experimental validation: tailoring peptide substrates for a serine protease

Copyright:

This sub-chapter 2 was published as a peer-reviewed journal **[1]**. Therefore, the text below is in nearly all instances a word-by-word reproduction of the peer-reviewed and published paper. Presented results are reprinted with permission from American Chemical Society, Biomacromolecules. Copyright © 2020 American Chemical Society.

[1] <u>**Philip Maximilian Knaff***</u>, Christian Kersten, Ramona Willbold, Carole Champanhac, Daniel Crespy, Rainer Wittig, Katharina Landfester, Volker Mailänder. **From in-silico to Experimental Validation: Tailoring Peptide Substrates for a Serine Protease.** *Biomacromolecules* **2020** *21* (4), 1636-1643.

Aim:

Hepsin, a pericellular transmembrane serine protease, is selectively overexpressed in prostate cancer and promotes tumor progression and metastasis. Hence, hepsin represents an attractive target for the development of peptide-based enzyme cleavable nanocarriers for efficient diagnostic or therapeutic purposes. In this context, the peptides serve as substrates for hepsin which are used as a proteolytic responsive trigger to open the nanocarriers in the proximity of prostate cancer cells. The aim of this study is to identify a peptide sequence with high specificity for hepsin and high stability in body fluids which can be used to selectively target hepsin expressing cancer cells.

Contribution:

I developed the methodology and carried out the molecular docking studies, synthesis and characterization of IQF substrates and stability assays in body fluids. Dr. Christian Kersten supported me with the molecular docking studies. Ramona Willbold performed testing of the IQF substrates against hepsin. The project was supervised by Dr. Carole Champanhac, Prof. Dr. Volker Mailänder and Prof. Dr. Katharina Landfester.

2.1 Abstract

Smart nanocarriers for the transport of drugs to tumor cells are nowadays of great interest for treating cancer. The use of enzymatic stimuli to cleave peptide-based drug nanocapsules for the selective release of nanocapsules cargo in close proximity to tumor cells opens new possibilities in cancer research. In the present work, we demonstrate a methodology for finding and optimizing cleavable substrate sequences by the type II transmembrane serine protease hepsin, which is highly overexpressed in androgen-independent prostate cancer. The design and screening of combinatorial libraries *in silico* against the binding cavity of hepsin allows the identification of a panel of promising substrates with high-calculated docking scores. *In vitro* screening verifies the predictions and showed that all substrates are cleaved by hepsin with higher efficiency than the literature known hepsin substrate RQLR↓VVGG. The introduction of D-amino acids on a selected peptide with the highest catalytic efficiency (k_{cat}/K_M) renders it resistant to cleavage by plasma or serum while maintaining their susceptibility to hepsin.

2.2 Introduction

Prostate carcinoma is the most common form of cancer in men, accounting for 20% of all incident cases of cancer.[107] The lack of visible symptoms can lead to undetected progression of the tumor. In the early stage, prostate cancer can be however efficiently cured by surgery.[108] In the later stages, prostate cancer can be treated with hormone therapy. Lowering androgen level decrease the tumor size but does not cure it completely.[109] If no regression occurs, these cancer cells eventually become androgen-independent and the therapy loses its efficiency.[110] The standard therapy for androgen independent prostate cancer is a combination of Paclitaxel and corticosteroid derivatives.[111, 112] This treatment leads to life extension and improves the quality of life in men. Nevertheless, median survival lies between 16 and 18 months.[113] Therefore, the development of a new effective therapy approach is highly desirable.

Our strategy is to develop a therapy based on a novel nanocarrier that releases cancer drugs to tumor cells via external stimuli. Nanostructures that respond to different stimuli such as light, pH value or temperature changes have already been presented, also enzymatic driven stimuli based on peptide-polymer conjugates have been reported as promising candidates.[114, 115, 116, 117, 118] Andrieu et al. synthesized peptide-polymeric hybrid capsules, produced by inverse miniemulsion, with incorporated enzymatic cleavable peptide substrates in the outer shell.[119]

In prostate cancer cells, type II transmembrane serine proteases (TTSP) like matriptase, matriptase-2, hepsin and DESC-1 are highly expressed.[120] Among those, hepsin was found to be the most overexpressed gene, especially involved in prostate cancer progression and metastasis.[71] The overexpression of this tumor-associated serine protease makes it an attractive target for enzymatic sensitive nanocarriers. These nanocarriers require a release mechanism based on peptide substrates to trigger the selective release of the cargo in close proximity to the tumor cells. Since serine proteases are present in blood and due to the similar substrate spectrum of type II transmembrane serine proteases, it is of essential importance to identify substrates with a preferential cleavage by hepsin. Béliveau et al. reported a peptide substrate (RQLR↓VVGG) which showed a higher specificity for hepsin among TTSP.[47]

We developed a methodology that can be used to rapidly identify and optimize enzyme cleavable substrates. The pre-requirements are a known crystal structure of the serine protease and the use of an "anchor"-residue (arginine) at the cleavable peptide bond for the definition of the peptide database. Based on the formerly identified high affinity hepsin substrate RQLR↓VVGG, we used molecular docking tools to predict 3D coordinates of peptides, which were then docked against the binding cavity of hepsin to obtain substrate-enzyme complexes.[13] Studies of their docking interactions helped us to select potential candidates for the development of hepsin-specific cleavable substrates. Following this

structure-based design, peptides with the highest potential were selected, based on their binding score. Next, the selected peptides were synthesized using a fluorophore/quencher system for live monitoring of enzymatic cleavage in solution (IQF substrates). Finally, we validated our predicted substrates by screening them against hepsin *in vitro,* showing that all selected substrates have a higher catalytic efficiency by hepsin. Moreover, all identified IQF substrates show an increased stability in blood serum and citrate plasma. In this study, we have selected one such peptide with the highest catalytic efficiency and introduced D-amino acids around the cleavage site. Our results suggest a panel of peptides that can be used for peptide-polymer nanocarrier systems.

2.3 Material and methods

Peptide Database Preparation. For the identification of novel high affinity hepsin substrates, virtual libraries were generated from 20 proteinogenic L-amino acids using CycloPs.[121] The libraries were composed of 8,000 tetramers for P1-P4-sites with a sequence motif of X-X-X-R and a 8,000 pentamers for P1-P4'-sites with a sequence motif of R-V-X-X-X. To avoid unwanted terminal charges, N-terminal acetylation (ace) and C-terminal *N*-methyl amidation (nme) caps were added. All structures were subsequently protonated (resulting in protonated Arg and Lys, deprotonated Asp and Glu and His carrying a hydrogen at the ε-nitrogen) and energy-minimized using the AMBER10EHT force field within MOE2015.[122,123,124,125]

Molecular Docking. A crystal structure of a hepsin-inhibitor complex is available for free in the protein data bank (PDB) with accession code 1Z8G.[126, 46] The docking was conducted using LeadIT-2.3.2.[127] For the P-site screening the binding-site was defined using the reference ligand, *N*-acetyl-6-ammonio-L-norleucyl-L-glutaminyl-*N*-[(1S)-4-1-(chloroacetyl)butyl]-L-leucinamide from the crystal structure, and including all protein residues within 6.5 Å of the ligand. The final binding site included His203, Tyr243, Ser251, Glu252, Glu253, Asn254, Asp257, Tyr301, Tyr328, Gln331, Lys333, Asp347, Gln350, Asp352, Ser353, Ser376, Thr379, Lys386 and Tyr390. The residues were protonated using the Protoss module within LeadIT resulting in deprotonated Asp and Glu residues and protonated Lys. His203 was protonated in δ-position to interact with Asp257. Both residues are part of the catalytic triad. Water molecules forming at least three interactions with the protein were kept as part of the receptor, namely water 188 and 302 of chain L and 409, 425, 438, 448, 476, 524 and 856 of chain A. To ensure the correct orientation within the binding site, pharmacophoric restraints were defined for C_α (xyz-coordinates 38.7/-2.4/29.4, radius 1 Å) and backbone nitrogen (xyz-coordinates 38.0/-3.5/30.3, radius 1 Å) of the arginine-moiety of the inhibitor and any atom to be close to C_α of lysine (xyz-coordinates 39.2/-11.3/33.0, radius 2 Å) (Figure 2.1). The receptor setup was validated by re-docking of the reference ligand in its intermediate alcohol-state for which the covalent bond to Ser353 detached, but no chlorine was added. The re-docking for the top pose had a score of -50.5 kJ/mol and a root mean square deviation (RMSD) of 2.6 Å, which was sufficient for the large number of rotating bonds. For the re-docking, the hybrid approach for enthalpy and entropy was used with 200 solutions for iteration and fragmentation. The docking of the 8,000 molecule P1-4-site peptide library was conducted under the same conditions.

For the P'-site docking, no reference ligand was available, therefore the binding site was manually defined using the arginine from P1-site as an anchor point. The final receptor included His186, His203, Asn209, Asn254, Asp257, Asn298, Thr299, Gln300, Tyr301, Tyr302, Gln304, Gln305, Asp347, Gln350, Asp352, Ser353, Ser376 and Tyr390 as well as water

molecules 448 and 524 of chain A being close to P1-site and as to be displaceable defined water molecules 302 and 409 of chain L and 682 and 713 of chain A. All further water molecules with two or less interactions with the protein were removed. Protonation was performed as for the P1-4-site docking resulting in deprotonated aspartates, His186 being protonated at ε-position, His203 in δ-position. To ensure the correct orientation within the binding site, pharmacophore constraints were added including an essential donor functionality on the arginine side-chain (xyz-coordinates 41.3/-7.0/25.7, radius 2 Å) and any atom being present 2 Å around xyz-coordinates 46.5/3.6/32.0 (Figure 2.1). The parameters for P1,1′-P4′-site, for base placement and scoring during docking were identical with the P1-4-site docking.
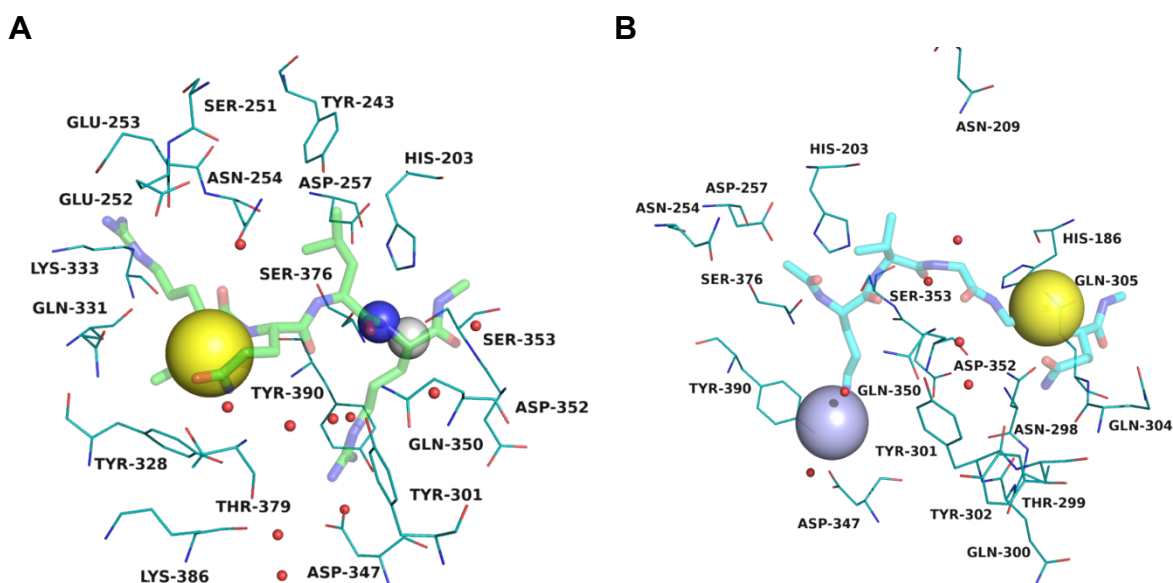

Figure 2.1. Binding-site definition for molecular docking with hepsin residues depicted as lines with teal carbon atoms and water molecules as red spheres (PDB-ID: 1Z8G). **A** Definition for P-site including pharmacophoric constraints as blue sphere for nitrogen atom, grey sphere for carbon atom and yellow sphere for any atom. Docking pose of ace-R-Q-L-R-nme with transparent sticks and green carbon atoms shown for orientation. **B** Definition of P′-site including pharmacophoric constraints for H-bond donor functionality (light blue sphere) and any atom (yellow sphere). Docking pose of ace-R-V-G-G-Q-nme depicted as transparent sticks and cyan carbon atoms shown for orientation.

IQF substrate synthesis. Peptides were synthesized as C-terminal amides on Rink amide resin using 9-fluorenylmethoxycarbonyl (Fmoc) strategy as described before.[128,129] All Fmoc-protected amino acids were purchased from Novabiochem, as well as, PyBOP, DIEA, Fmoc-Rink amide (loading 0.36 mmol/g), N,N-dimethylformamide (DMF, p.a., Sigma Aldrich), DCM (>99.9%, Carl Roth), and were used as received. All Fmoc-protected amino acids were dissolved in DMF at a concentration of 0.2 M, except for Fmoc-K(Mca)-OH which was used with a concentration of 0.077 M due to poor solubility. The coupling times for the peptides were

increased from 5 to 10 minutes and the fluorophore Fmoc-K(Mca)-OH and quencher Fmoc-K(Dnp)-OH were coupled twice on the same step to ensure high purities and yield. The peptides were cleaved from the resin with trifluoroacetic acid (TFA) containing a scavenger (TFA:TIPS:H$_2$O 94:3:3) for 3 h. The peptides were precipitated and washed three times with cold diethyl ether, freeze dried overnight and stored at -20 °C in the dark.

Human blood plasma and serum. Human blood was obtained from the Department of Transfusion Medicine Mainz from seemingly healthy donors in accordance with the Declaration of Helsinki. Blood was clotted overnight according to the standard protocol to generate human serum. A serum pool from seven volunteers was used and stored at -80 °C. To remove any protein aggregates after thawing, human serum and plasma were centrifuged for 15 minutes at 20,000 g before usage.

Enzymes. Recombinant human hepsin protein (HPN) was purchased from Bio-Techne GmbH (Wiesbaden, Germany).

IQF substrate kinetics. Hydrolysis of IQF peptide substrates by hepsin were measured by recording the fluorescence using a FLUOstar Omega BMG Labtech plate reader (Ex_λ 320 nm; Em_λ 400 nm) at 37 °C. The measurements were performed in a total volume of 100 µL in 50 mM ammonium bicarbonate buffer (pH 7.8). The enzyme concentration used was 0.2 ng/µL and the peptide concentrations varied between (0-200 µM). The linear part of the progress curve was used to determine the reaction velocity. The measurements were repeated, and the data presented are the average of at least three measurements. The catalytic parameters k_{cat} and K_M were calculated using Omega Data Analysis Software to directly fit the initial velocities to the Michaelis-Menten equation using non-linear regression analysis.

Serum and Plasma stability tests. The stability of the peptide were measured according a procedure from Jensen et al. .[130] In short, peptide stock solutions 0.1-5 mg/mL were prepared by dissolving the peptide in pure dimethyl sulfoxide. RPMI medium 1640 was supplemented with 12.5% (v/v) human serum or citrate plasma and aliquoted into 2 mL low protein binding Eppendorf tubes. All kinetic measurements were performed using a Tecan M1100 at 37 °C (Ex_λ 325 nm; Em_λ 392 nm). The substrate concentrations varied between (0-166 µM) and the total reaction volume was set to 100 µL per well. Measurements were conducted for 5 h. Non-linear regression analysis were performed according to IQF substrate kinetics method and the enzyme concentration was set to 1.0 µM by definition.

2.4 Results

Experimental Approach. The experimental workflow is schematically shown in Figure 2.2. Briefly, the known substrate RQLR↓VVGG (reference peptide), for the serine protease hepsin, was used to generate peptide combinations *in silico*. The arginine around the cleavage site R↓V was kept as an "anchor" while amino acids on each side of the cleavage site were permutated. Through this approach, two peptide libraries with the pattern X-X-X-R (for P-site) and R-V-X-X-X (for P′-site) were created with a total of 16,000 peptide combinations (Step 1). In the next step, these two peptide libraries were virtually screened with the software Molecular Operating Environment (MOE) against the binding pocket of hepsin and the peptides were ranked based on their calculated binding score. The peptides with the highest binding score were selected and peptide combinations were created in such a way, that one side of the reference peptide was changed while the other one was kept constant, respectively (Step 2). These peptide candidates were synthesized as internally quenched fluorescent peptides (IQF peptides) by solid phase peptide synthesis, using the Fmoc strategy. In the last step, the peptides were screened against the serine protease hepsin and the stabilities of the peptides were tested in serum and citrate plasma (Step 3).
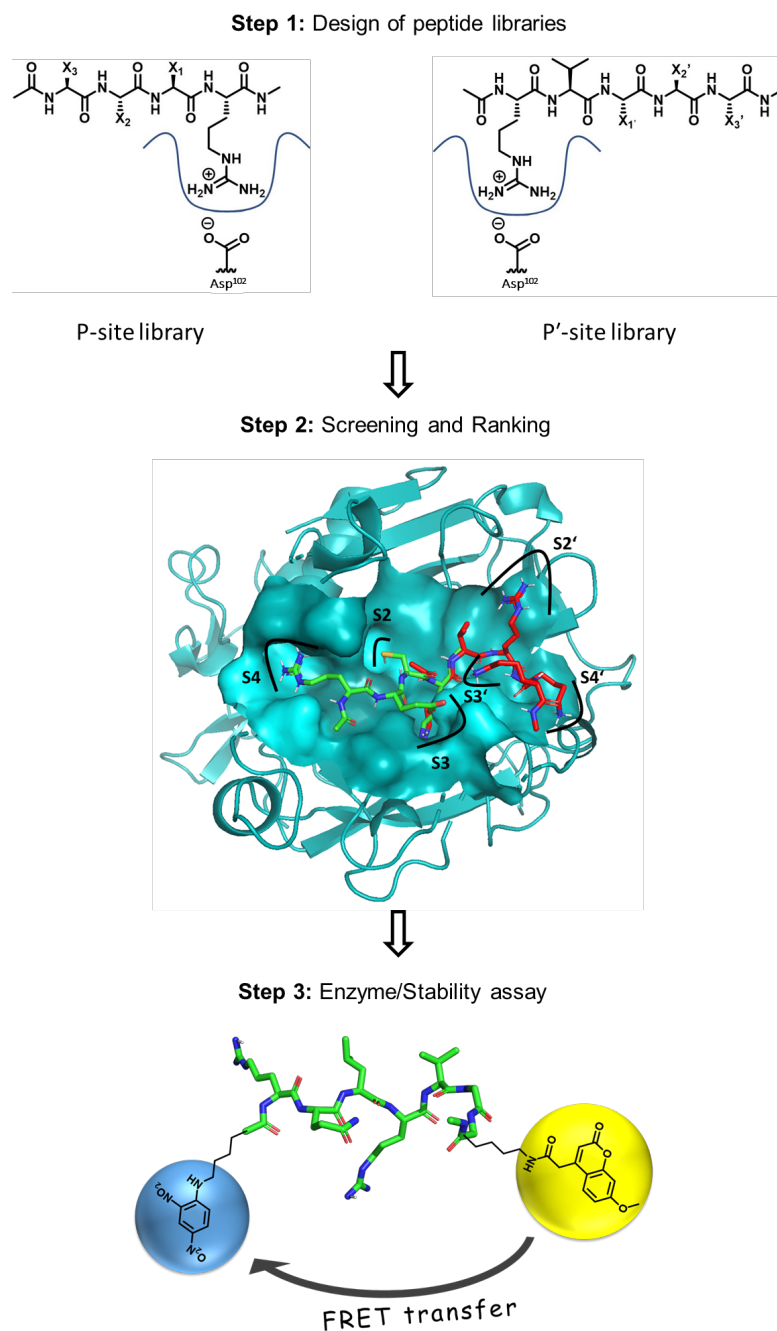
Step 1: Design of peptide libraries

P-site library   P'-site library

Step 2: Screening and Ranking

Step 3: Enzyme/Stability assay

FRET transfer

Figure 2.2. Workflow of the optimization process of hepsin cleavable IQF peptide substrates. Two different peptide libraries were designed (P-side library and P'-side library) and screened against the binding pocket of hepsin (protein coordinates taken from PDB-ID: 1Z8G, 1.55 Å resolution).[126] Afterwards, the peptides were ranked based on their calculated score and the peptides with the highest scores were synthesized by solid phase peptide synthesis.

IQF peptides. In the present work, internally quenched fluorescent peptides were prepared where donor and acceptor fluorophores were conjoined via ε-amino group of lysine in the peptide, thus being in close proximity. If the emission spectrum of the donor overlaps with the absorption spectrum of the quencher the energy transfer happens within the molecule and no

fluorescence can be recorded. The peptide between the FRET pair is a substrate for the serine protease which can be cleaved during an interaction. Upon cleavage of the IQF peptide, fluorophore and quencher are not close anymore and fluorescence can be detected. To confirm that cleavage occurs at the predicted position between suggested P1 (arginine) and P1′ position (valine), an IQF peptide with cleavage pattern R↓V was incubated with hepsin and the cleavage products were analyzed by MS (Figure 2.3).
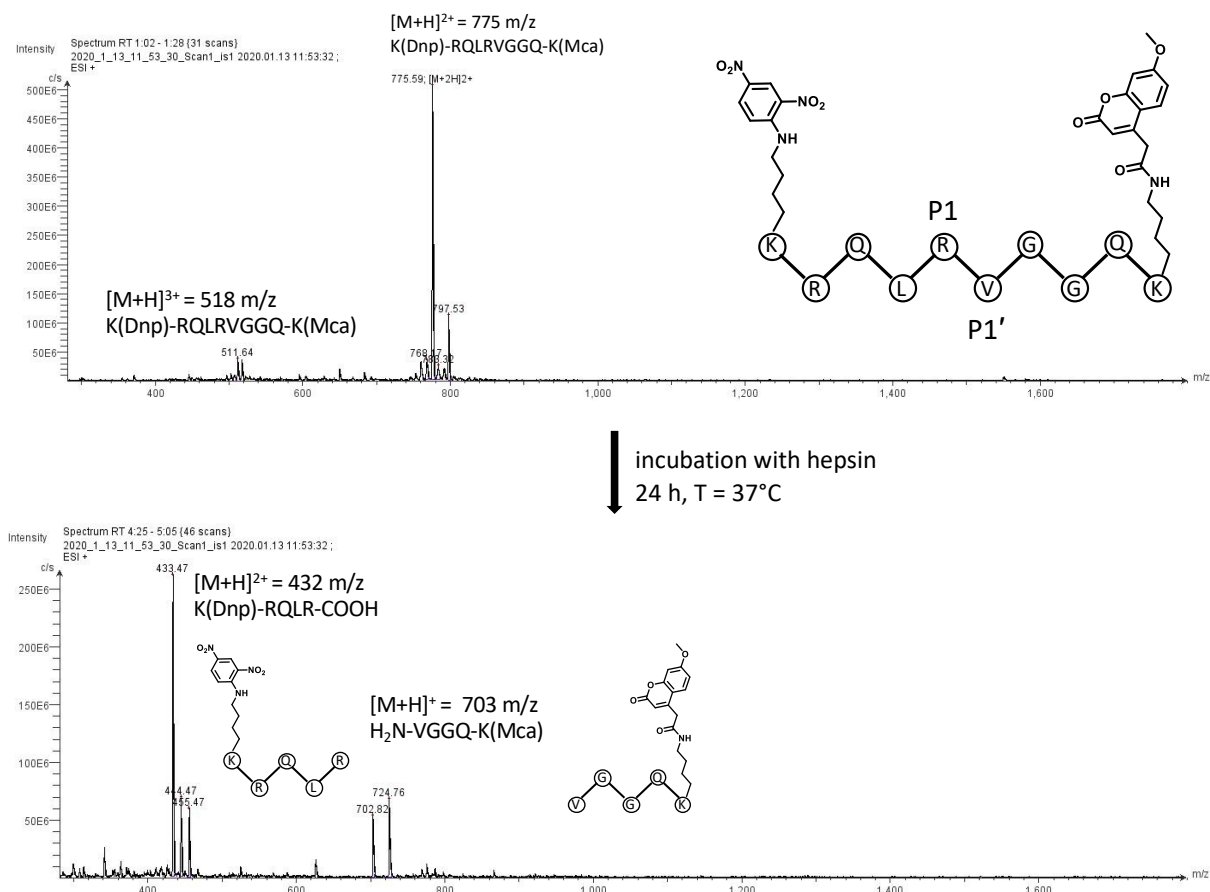

Figure 2.3. ESI mass spectra of the reference peptide with Dnp/Mca FRET pair. After treatment with hepsin, the peptide is cleaved between the scissile peptide bond R↓V.

Structure-based design of hepsin substrates. Prior knowledge and previous results identified the octapeptide (P4-4′) RQLR↓VVGG to be a high-affinity hepsin substrate.[41, 47] However, its identification was based on the similarity to known TTSP substrates such as the auto-activation sequences of matriptase (RQAR↓VVGG) or of the hepatocyte growth factor (KQLR↓VVNG), without a special focus on hepsin specificity or bloodstream stability, which must be optimized for the desired targeting. Therefore, a virtual screening to identify novel hepsin substrates with improved bloodstream stability was conducted. The high number of rotatable bonds and the connected high degree of conformational freedom for octameric

peptides is a general challenge in molecular docking.[131, 132] The process of docking was conducted in two parts: Identification of favored residues for P2-P4 sites of the binding site, keeping P1 residue constant as arginine; followed by identification of P2′-P4′ keeping P1′ constant as valine and also adding P1-residue arginine as an "anchor" residue due to its directional interactions within the sharply defined S1 pocket.[131] The 20 proteinogenic L-amino acids were included in the peptide design and 8,000 tetramers of the structure acetyl(ace)-X-X-X-R-amidation(nme) for P1-4 -sites and 8,000 pentamers ace-R-V-X-X-X-nme for the P1-4′ design were virtually generated (Figure 2.4).
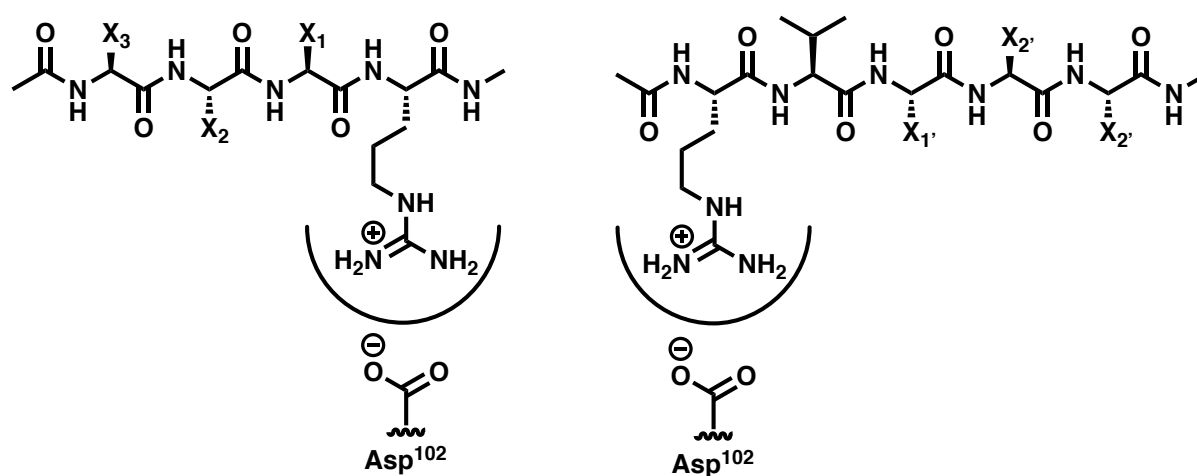

Figure 2.4. Architecture of the virtually generated peptide libraries. P-side library (left side) and P′-side library (right side) with X being natural amino acids. The arginine at P1 position serves as an anchor residue.

The molecules were subsequently docked into the prepared P-site and P′-site binding pockets, respectively, and ranked according to their score (Figure 2.5).
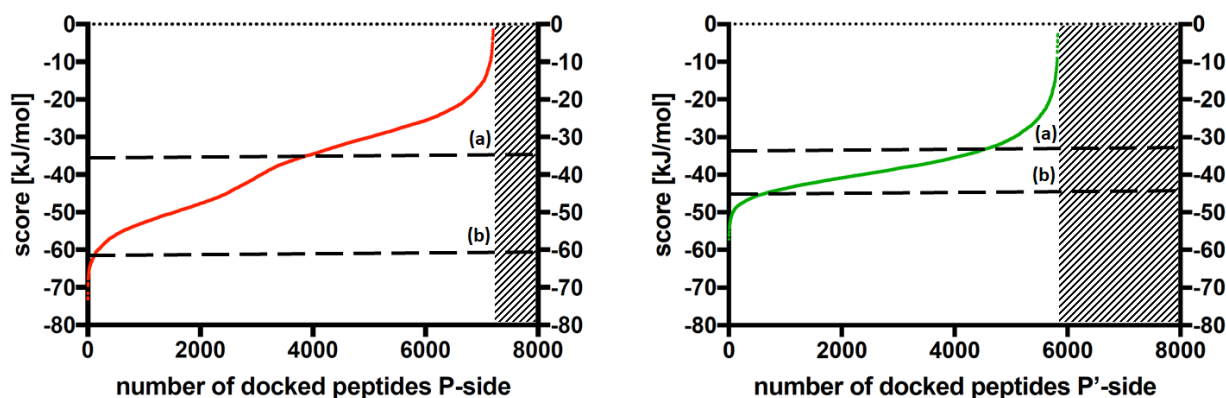

Figure 2.5. Ranked peptides of P-side and P′-side library based on their calculated score. (a) negative control, (b) reference peptide.

The highest scoring peptides were verified by visual inspection of the binding mode and the most promising sequences were created by combining either the high scoring P-site sequences with the P′-site of the reference peptide, or the high scoring P′-site sequences with the P-site of the reference peptide (Table 2.1). Those peptides were synthesized as IQF peptides for testing, as well as two low scoring peptides DLSR↓VSEV and GVTR↓VSWC serving as negative controls (Table 2.1).

Table 2.1. A Docking result for P-site (P1-P4) designed molecules with structure being acetyl-X-X-X-R-methylamide and **B** docking result for P′-site (P1-P4′) designed molecules with acetyl-R-V-X-X-X-methylamide structure and X being proteinogenic amino acids. **C** Designed IQF type peptide libraries for P-site and P′-site. (a) reference sequence of known substrate, (b) negative control.

A

P4	P3	P2	P1	↓	score [kJ/mol] (Rank)
R	Q	L	R		-60.1 (175)[a]
A	E	Q	R		-72.7 (2)
G	Q	R	R		-72.7 (3)
R	N	H	R		-71.6 (4)
Q	Q	V	R		-69.5 (5)
D	L	S	R		-33.3 (4256)[b]
G	V	T	R		-39.8 (3108)[b]

B

P1	↓	P1′	P2′	P3′	P4′	score [kJ/mol] (Rank)
R		V	V	G	G	-44.2 (836)[a]
R		V	H	S	L	-55.2 (4)
R		V	G	R	H	-55.1 (5)
R		V	G	G	Q	-53.9 (9)
R		V	S	E	V	-33.1 (4562)[b]
R		V	S	W	C	-33.1 (4552)[b]

C

			P4	P3	P2	P1	↓	P1′	P2′	P3′	P4′	
reference		K(Dnp)	R	Q	L	R		V	V	G	G	K(Mca)
P-site	**P1**	K(Dnp)	A	E	Q	R		V	V	G	G	K(Mca)
	P2	K(Dnp)	G	Q	R	R		V	V	G	G	K(Mca)
	P3	K(Dnp)	R	N	H	R		V	V	G	G	K(Mca)
	P4	K(Dnp)	Q	Q	V	R		V	V	G	G	K(Mca)
P′-site	**P′1**	K(Dnp)	R	Q	L	R		V	H	S	L	K(Mca)
	P′2	K(Dnp)	R	Q	L	R		V	G	R	H	K(Mca)
	P′3	K(Dnp)	R	Q	L	R		V	G	G	Q	K(Mca)
negative 1		K(Dnp)	D	L	S	R		V	S	E	V	K(Mca)
negative 2		K(Dnp)	G	V	T	R		V	S	W	C	K(Mca)

The binding modes of potential substrates (Figure 2.6) indicated that for the P4-position a polar residue was favored, potentially carrying a positive charge like Arg or Lys to interact with Glu253 or Glu252 *via* ionic interactions.
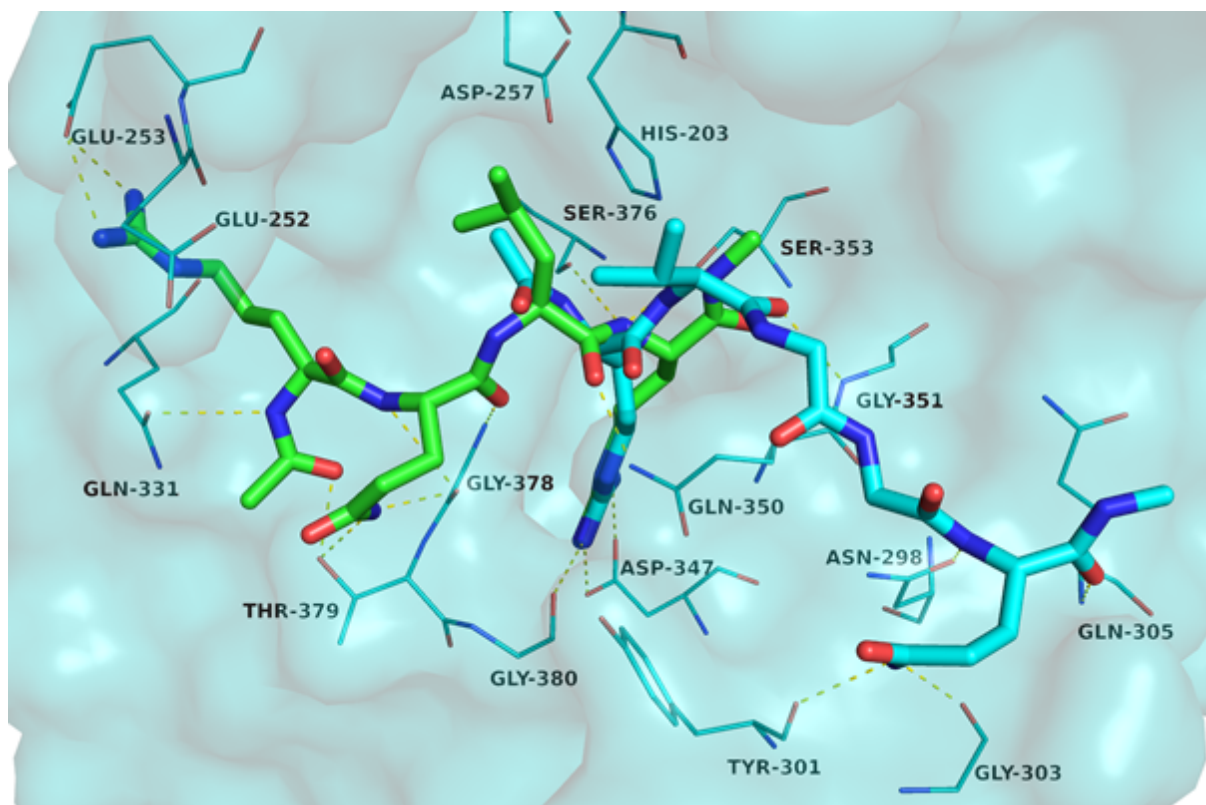
Figure 2.6. Superposition of predicted binding modes for P-site peptide ace-R-Q-L-R-nme (green carbon atoms) and P′-site peptide ace-R-V-G-G-Q-nme (cyan carbon atoms) in complex with hepsin (teal carbon atoms and surface, PDB-ID: 1Z8G). For a clear view only binding site residues that form polar interactions with the peptides or are part of the catalytic triad (Ser353, His203, Asp257) are labeled and depicted as lines.

For P3 also, a polar group, preferably Gln, interacting with Thr379 and the backbone of Gly378 is found. At P2′ polar and non-polar residues seemed to be tolerated in the docking, which is also supported by previous reports.[41] In case of Leu, a hydrophobic contact area is built with His203 of the catalytic triad and Asn254. Arg or Lys are crucial at the P1-site because the positively charged moiety interacts with Asp347, Gly380 backbone oxygen and an additional water molecule (wat302) in a deeply buried sub-pocket. Therefore, Arg was preliminary defined as the anchor residue. Differently from P-sites, the P′-sites are less clearly defined and more on the protein surface. From literature, no clearly favored residues are known, which is also reflected in the docking. Even though pentapeptides were docked during the P′-site screening, the overall scores were lower compared to the P-site tetrapeptide docking (Table 2.1, Figure 2.5) indicating that only weaker interactions can be formed. Small non-polar side chains like pre-defined Val seemed to fit well into P1′ interacting with His203, Cys188 and Leu187. For subsequent P2′-P4′ residues Gly, Gly and Gln polar interactions were only predicted for amides with Gly351 and Asn298 and the side-chain amide for P4′ Gln with Tyr301 and Gly303 of the protein backbone. As no deep cavities are present, smaller residues might be preferred.

Validation of predicted hepsin cleavable substrates. To study if the substrates predicted by molecular docking lead to an optimization of the known substrate (RQLR↓VVGG), we synthesized the selected candidates as IQF peptides with the FRET pair Mca/Dnp and assayed for enzymatic activity. Table 2.2 and Figure A2.1 present the kinetic values (k_{cat}, K_M, V_{max}, and k_{cat}/K_M) for all substrates. The term k_{cat}/K_M was used as catalytic efficiency to compare reaction rates of substrates transformed by enzymes relative to the reference peptide. Strikingly, all selected peptides, except negative controls, showed a higher catalytic efficiency during cleavage by hepsin. P3 peptide (319% rel. catalytic efficiency) showed no Michaelis-Menten behavior. Excluding this peptide from the rest showed that relative catalytic efficiencies of peptides with novel P-site ranges between 124-151%. Hepsin was more permissive with novel P′-site peptides, showing relative catalytic efficiencies between 120-232% indicating a better fit of the substrates in the hepsin binding cavity, thus leading to a higher cleavage rate (Table 2.2). P′3 K(Mca)-RQLR↓VGGQ-K(Dnp) led to a 2.3-fold increase of catalytic efficiency compared to the reference peptide K(Mca)-RQLR↓VVGG-K(Dnp). Substitutions of valine at P2′ position by a sterically more favorable glycine and glycine at P4′ position by a large polar glutamine suggest that these positions play an important role in substrate recognition by hepsin.

Table 2.2. Hydrolysis of IQF peptides by hepsin, serum and plasma. Constants were calculated from nonlinear regressions of hyperbolic Michaelis–Menten rate equations. Relative catalytic efficiencies (%) are the *catalytic efficiency* values of the IQF peptides relative to that of the reference peptide. Enzymatic measurements were performed in triplicates. N.c = not cleaved.

IQF substrate		Hepsin		Serum		Plasma	
		k_{cat}/K_M [sec^{-1} µM^{-1}]	rel. catalytic efficiency	k_{cat}/K_M [sec^{-1} µM^{-1}]	rel. catalytic efficiency	k_{cat}/K_M [sec^{-1} µM^{-1}]	rel. catalytic efficiency
reference		18.5×10^{-2}	100%	6.44×10^{-4}	100%	7.71×10^{-4}	100%
A)	P1	24.2×10^{-2}	131%	2.95×10^{-4}	46%	1.65×10^{-4}	37%
	P2	23.0×10^{-2}	124%	3.00×10^{-4}	47%	1.50×10^{-4}	21%
	P3	59.1×10^{-2}	319%	4.48×10^{-4}	70%	3.78×10^{-4}	49%
	P4	28.0×10^{-2}	151%	5.63×10^{-4}	87%	3.14×10^{-4}	41%
B)	P'1	22.2×10^{-2}	120%	5.22×10^{-4}	81%	3.57×10^{-4}	46%
	P'2	34.6×10^{-2}	187%	4.68×10^{-4}	73%	5.35×10^{-4}	69%
	P'3	24.9×10^{-2}	232%	4.51×10^{-4}	70%	5.64×10^{-4}	73%
C)	P'3.1	17.1×10^{-2}	95%	n.c.		n.c.	
	P'3.2	31.6×10^{-2}	176%	n.c.		n.c.	
	P'3.3	n.c.		2.70×10^{-4}	41%	2.01×10^{-4}	26%
	P'3.4	26.7×10^{-2}	148%	5.53×10^{-4}	86%	5.17×10^{-4}	67%
D)	P''1	14.9×10^{-2}	80%	4.77×10^{-4}	74%	2.62×10^{-4}	34%
	P''2	19.5×10^{-2}	105%	5.26×10^{-4}	82%	2.45×10^{-4}	32%
E)	negative 1	n.c.		2.58×10^{-4}	40%	2.62×10^{-4}	31%
	negative 2	n.c.		n.c.		n.c.	
	negative control	n.c.		n.c.		n.c.	

***In vitro* IQF peptide stability in blood serum and plasma.** An effective peptide polymer conjugate nanocapsule needs to be stable in biological fluids for an adequate time to reach the desired target. To test the *in vitro* stability, the identified IQF peptides (Table 2.1) were synthesized in a free form (with capped C-terminus, ace) and stability was tested in blood serum and citrate plasma. To gain an overall picture of the serum/plasma stability of the fluorescent peptide substrates, all peptides were first incubated at a fixed concentration (33.3 µM) with 12.5% serum and 12.5% plasma in RPMI cell culture medium (Figure 2.7).
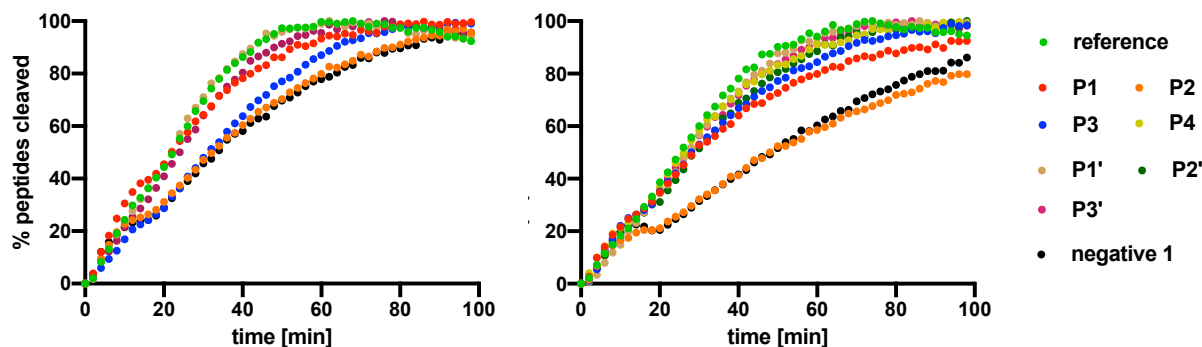
Figure 2.7. Exemplary presentation of the degradation of IQF peptides in serum (left panel) and plasma (right panel). Shown are the percentage of cleaved IQF peptides (33.3 µM) in 12.5% serum/plasma for a single experiment.

As shown in Figure 2.7, the reference peptide K(Mca)-RQLR↓VVGG-K(Dnp) is fully degraded after 60 min in serum. All identified IQF peptides with modified P- and P′-side were stable in serum and plasma over a longer period of time (Table 2.2, Figure A2.2 and Figure A2.3). Of these, IQF peptides with modified P-side show a higher stability than IQF peptides with modified P′-side. To better evaluate specificity of the IQF peptides against serum and plasma, we determined kinetic parameters using standard Michaelis-Menten kinetics. Of all the IQF peptides with modified P-site, the P2 peptide showed the highest stability in serum and plasma (Table 2.2). This IQF peptide yielded a k_{cat}/K_M value (3.00×10^{-4} s^{-1} µM^{-1}) in serum, 47% of the catalytic efficiency relative to the reference peptide, while showing only 21% relative catalytic efficiency in plasma. Changes in the P′-site of the reference peptide lead to an increase of stability in biological fluids. Relative catalytic efficiencies vary between 70-81% in serum and 46-73% in plasma but are not as pronounced as changes in the P-site, which vary between 46-87% in serum and 21-49% in plasma, suggesting that amino acids in the P-site play a superordinate role in recognizing substrates by proteases occurring in serum and plasma. The two docked peptides, which showed low binding affinities to hepsin in the *in silico* approach, showed the highest serum/plasma stabilities among all synthesized IQF peptides. Negative peptide 1 K(Mca)-DLSR↓VSEV-K(Dnp) showed a relative catalytic efficiency of 40% in serum

and 31% in plasma. Negative peptide 2 K(Mca)-GVTR↓VSWC-K(Dnp) and the negative control K(Mca)-AAAA-AAAA-K(Dnp) lacking the cleavage motif R↓V showed no degradation.

Optimization of P′3 peptide using D-amino acids. Successful validation of predicted hepsin cleavable substrates based on the reference peptide sequence RQLR-VVGG has shown the applicability of our approach. A series of peptides with higher hepsin cleavage efficiency and stability in biological fluids could be identified. The results observed are summarized in Table 2.2. To increase the specificity for hepsin and stability in blood, a strategy in analogy to Kumaresan et al. was applied.[133] Starting from the identified P′3 peptide K(Mca)-RQLR-VGGQ-K(Dnp), we introduced subsequently D-amino acids on both sides around the cleavage site R↓V to finally produce a panel of sequences (P′3.1 – P′3.4) in the form shown in Table 2.3.

Table 2.3. Pattern for the introduction of D-amino acids to increase stability in biological fluids.

substrate	P4	P3	P2	P1	↓	P1′	P2′	P3′	P4′
P′3	R	Q	L	R		V	G	G	Q
P′3.1	**r**	Q	L	R		V	G	G	Q
P′3.2	R	**q**	L	R		V	G	G	Q
P′3.3	R	Q	**l**	R		V	G	G	Q
P′3.4	R	Q	L	R		V	G	G	**q**

In biological fluids. Introduction of D-amino acids can heavily influence the degradation of peptides. Inverting the stereogenic center from L- to D-configuration for the P-site of the peptide could increase the stability in biological fluids. P4 position (arginine) and P3 position (glutamine) of P′3 peptide K(Mca)-RQLR-VGGQ-K(Dnp) was of great help and made the peptides nondegradable in blood serum and citrate plasma. In Figure 2.8, the fluorescence intensity over time is shown, as an example, at one concentration ($c_{peptide}$ = 166 µM). Incorporation of D-leucine (P′3.3) approximately doubled the stability of the peptide while modifying the P′-site with D-glutamine led to no significant benefit in stability. This further strengthens our observation that the P-site of the substrate plays a major role in substrate recognition by blood enzymes and should be especially taken into account for preventing degradation.
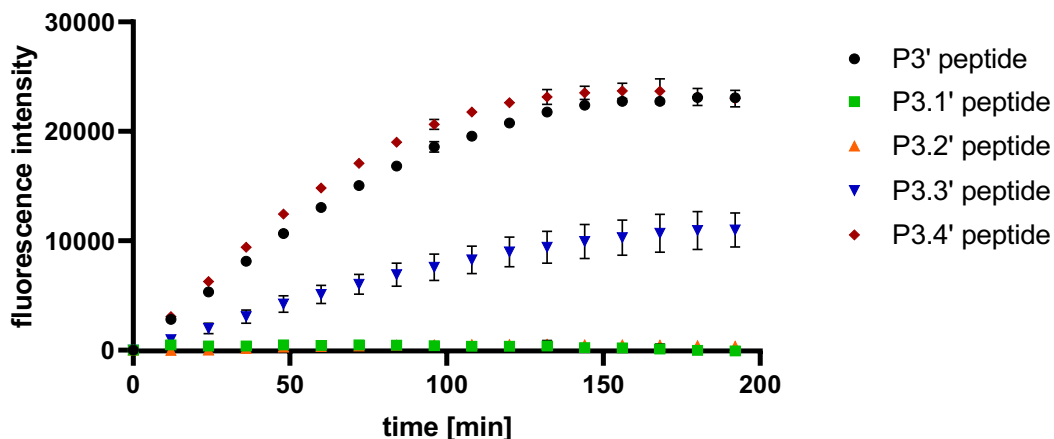
Figure 2.8. The fluorescence intensity of P3′ peptide and its D-amino acids variations P3.1′ – P3.4′. Exemplary representation of IQF substrate degradation (c = 166 μM) incubated with 12.5% blood serum at 37 °C for 5 h.

For hepsin. In order to investigate the effect of D-amino acids on substrate recognition and cleavage, the IQF substrates P′3.1-P′3.4 were screened against hepsin. Introduction of D-amino acids influence the enzymatic cleavage notably. The peptides P′3.2 and P′3.4 showed with 163% and 145% relative catalytic efficiency a higher cleavage rate than the reference peptide (but lower than P′3), while P′3.1 shows no significant difference relative to the reference peptide. Interestingly, the introduction of D-leucine in position P3′ (P′3.3 peptide) hinders enzymatic degradation completely. The close proximity of this position to the highly conserved cleavage site P1↓P1′ seems to play an important role in substrate recognition by hepsin.

Validation of substrates containing a combination top prime and top non-prime site sequence. As shown in Table 2.2 a variety of peptides with a higher catalytic efficiency against hepsin and stability in biological fluids have been identified. Until now, our approach aimed to improve the reference substrate RQLR-VVGG by substitution of either the P-site or the P′-site. In another set, we created two combinations with top prime (P3 and P4 peptide) and top non-prime site (P′3 peptide) as shown in Table 2.4.

Table 2.4. Pattern for the design of substrates containing top prime and top non-prime site sequence.

substrate	P4	P3	P2	P1	↓	P1′	P2′	P3′	P4′
P″1	Q	Q	V	R		V	G	G	Q
P″2	R	N	H	R		V	G	G	Q

Enzymatic degradation of P″1 peptide and P″2 peptide by hepsin revealed relative catalytic efficiencies of 80% and 105%, respectively. Both peptides showed a higher stability in serum (rel. catalytic efficiencies varied between 74% - 82%) and in plasma (rel. catalytic efficiencies varied between 32% - 34%) than the reference peptide. Substrates containing top prime site sequences and a top non prime site sequence showed no additive improvement in catalytic efficiency through proteolytic degradation by hepsin.

2.5 Discussion

We demonstrated a computational docking method to identify novel peptide substrates for the serine protease hepsin. For this purpose, we separated the peptide binding site into two smaller sub-pockets to reduce the conformational degree of freedom. In the next step, we built peptide libraries with a specific pattern, which were then screened *in silico* against the binding cavity of hepsin. Ranking of these libraries based on their calculated binding score and analysis of predicted binding modes allowed us to select substrate candidates for hepsin, which were synthesized and validated in an enzymatic assay using FRET system. With this method, we could identify a panel of peptides, specially tailored for hepsin binding pocket, which showed a higher cleavage efficiency and a higher stability in serum and plasma than the known reference peptide RQLR↓VVGG. Furthermore, a clear discrimination between predicted favored and predicted not favored substrates was observed *in vitro*. Our results suggest that changing the P-site of substrates leads to an increase of stability in biological fluids while changing P′-site of substrates increases the substrate preference for hepsin. P′3 peptide (RQLR↓VGGQ) was most promising and showed an increased relative catalytic efficiency up to 232% compared to the known hepsin substrate. Degradation assay in biological fluids showed for P′3 an increase of stability of 30% in blood serum and 27% in citrate plasma. Implementation of D-amino acids in P′3 could strongly increase stability in serum and plasma. P′3.1 and P′3.2 could resist degradation while keeping high cleavage efficiency by hepsin (P′3.1 95%, P′3.2 176%). Those peptides are promising candidates for the development of hepsin cleavable peptide-polymer nanocapsules. An additive improvement by combination of top prime and non prime peptides could not be observed. The method presented in this paper can be used as a guideline to find protease substrates - or with slight modifications - for peptidomimetic inhibitor design, if some pre-requirements are met. I) The crystal structure or a high-quality homology model of the target protease is available. II) At least one specific and deep subpocket of an "anchor" residue (like P1 Arg for hepsin) is present for peptide library definition. III) Pharmacophoric constraints can enhance the desired orientation within the binding-site. Additionally, non-proteinogenic amino acids and D-amino acids can be included in library design to increase the range of possibilities with regards to specificity and stability. The serine protease hepsin, as an attractive target for the development of novel prostate cancer therapies, could prove the successful applicability of our method.

3. Structure-based design of high affinity and selective peptidomimetic hepsin inhibitors

Aim:

The previously described study **[1]** in this dissertation reveals a panel of highly susceptible substrates for hepsin which were successfully developed using computation, solid-phase peptide synthesis and enzymatic testing. As a follow-up project, the study **[2]** intends to use a part of the obtained results to design with slight modifications highly active peptidomimetic inhibitors for hepsin. Those hepsin inhibitors may be used as an alternative strategy for potential anticancer application since conventional receptor tyrosine kinase inhibitors, which are commonly used are facing major issues concerning drug resistance and toxicity.

[2] <u>**Philip Maximilian Knaff**</u>*, Christian Kersten, Lukas Wettstein, Patrick Müller, Tatjana Weil, Carina Conzelmann, Janis Müller, Stefan Pöhlmann, Markus Hoffmann, Tanja Schirmeister, Jan Münch, Volker Mailänder. **Tailored Inhibitors for the Serine Protease Hepsin**. To be submitted.

Contribution:

I carried out the synthesis and characterization of the peptidomimetic inhibitors and performed the enzymatic activity measurements and stability studies. Dr. Christian Kersten supported with additional molecular docking studies and Patrick Müller supported with synthesis of multiple precursor compounds. The project was supervised by Prof. Dr. Volker Mailänder and Prof. Dr. Katharina Landfester.

3.1 Abstract

In many solid tumors, increased upregulation of transmembrane serine proteases (TTSP) leads to an overactivation of growth factors which promotes tumor progression. Here, we are using a combinatorial methodology to develop high affinity tetrapeptidic inhibitors. Virtual screening of 8.000 peptide combinations against the crystal structure of the TTSP hepsin identified a series of recognition sequences, customized for the non-prime substrate binding (P)-sites of this serine protease. Combination of the top recognition sequences with an electrophilic warhead resulted in highly potent inhibitors with good selectivity against coagulation proteases factor Xa and thrombin. Structure-activity relationships of two selected compounds were further elucidated by investigation of the stability in biological fluids as well as the influence of the warhead and truncated inhibitors on inhibitory potency. Overall, this methodology yielded compounds as selective inhibitors for potential cancer drug development where hepsin is overexpressed.

3.2 Introduction

Cancer emerges from a complex variety of gene mutations. Most frequently, the genes encoding for the protein class of the kinases and proteases are highly upregulated.[134, 135] The use of kinase inhibitors is nowadays a common treatment, but patients can rapidly acquire resistance. To overcome this obstacle, type II transmembrane serine proteases (TTSP) as an alternative drug target come in focus of research.[136] Hepsin is a TTSP which shows in healthy prostate tissue low expression levels while high expression levels have been shown to play critical roles in cancer like the androgen-independent prostate cancer.[37, 71] This protease activates oncogenic proteins such as pro-hepatocyte growth factor (pro-HGF) to HGF which then triggers downstream signaling pathways through c-MET activation. High proteolytic activity of hepsin was shown to evoke HGF related biological effects such as the loss of epithelial cell membrane integrity, survival and cellular growth. In healthy tissue, the activity of TTSP is regulated by the endogenous inhibitors HAI-1 and HAI-2 proteins. *In vivo* hepsin transgenic mice showed in primary prostate tumor models to progress cancer by promoting metastasis. Contrary, hepsin knockout mice have shown to develop normally with no apparent abnormalities. To specifically address hepsin-overexpressing tumors, the design of enzymatically cleavable peptide-based nanocapsules for a selective release came into focus. First investigations in this direction have been previously published by our group.[137, 138] The design of such peptide-based nanocapsules requires high enzymatic specificity to ensure targeted drug delivery. To meet these high demands, we improved the literature known substrate R-Q-L-R-V-V-G-G and developed novel substrates with a higher affinity and serum/plasma stability using a combinatorial *in silico* approach, paired with organic synthesis and biological testing.[138] The *in silico* approach included the design of a P- and a P'-site peptide library, consisting of 8,000 combinations each, which target the respective S- and S'-site of the binding pocket of hepsin. Peptides with the highest predicted affinity for hepsin were subsequently synthesized and tested for their suitability as a hepsin substrate as well as for their serum and plasma stability. In this paper, we aimed to validate the possibility of developing selective and high-affinity substrate-based inhibitors, tailored for the serine protease hepsin based on the substrate sequences identified previously (Figure 3.1). For this purpose the top binders of the S-site library were combined with an electrophilic ketobenzothiazole moiety to yield covalent reversible inhibitors.[139] Peptide-based inhibitors customized for the binding cavity of the target protease offer multiple interaction points and a larger contact area which increase selectivity in natural manner simply by the sequential choice of amino acids. Still, previous inhibitors showed activity among similar proteases like matriptase.[140-142] Even though matriptase is also upregulated in prostate cancer, unlike hepsin, matriptase-deficient mice showed low-life expectancy and should be considered as a potential

off-target.[143] Toxicity effects of hepsin and matriptase short-term inhibition still needs to be examined. *In vitro* screening revealed a panel of tetra-peptidomimetic inhibitors with activity in the low nanomolar (nM) range for hepsin with good off-target selectivity over matriptase, thrombin and factor Xa. Furthermore, we selected a compound to influence different warhead moieties as well as truncated recognition sequences to tri- and dipeptides. Our results show that we can use the screening approach as a fast and cost-efficient pipeline for the identification of possible binders which can further be used for the development of first-generation hepsin inhibitors. Further, the combinatorial nature of this approach allows new insights into structure-activity relationships.

Combinations of P- & P'-site candidates as IQF peptides
- **high affinity substrates**
- **stable in serum and plasma**

Introduction of electrophilic warhead on P-site candidates
- **potent, selective inhibitors**
- **stable in serum and plasma**
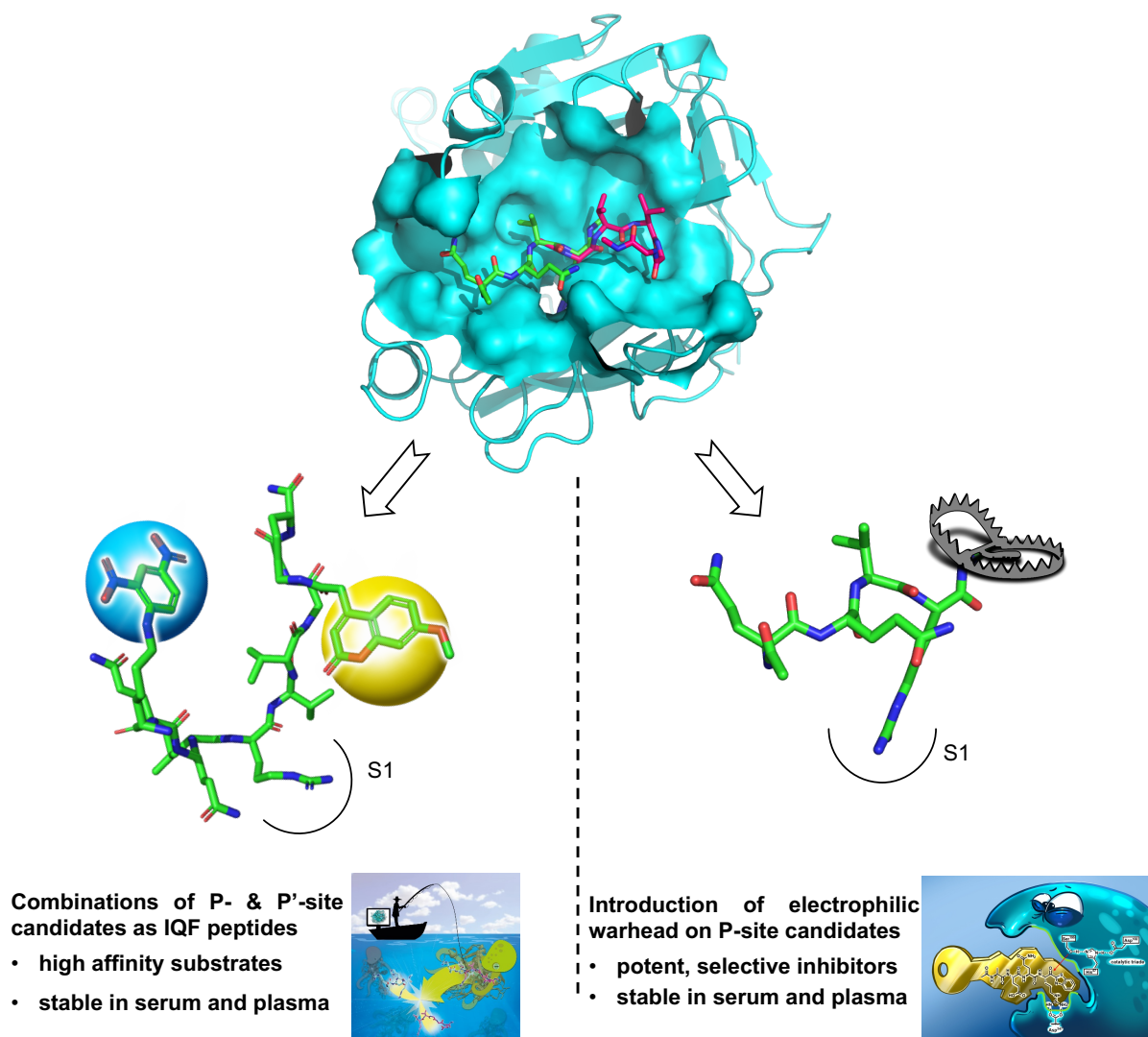
Figure 3.1. Scheme for the development of tailored peptide inhibitors. Computational docking results from the P-site library of a previous publication were used as recognition sequences, modified with an electrophilic warhead and tested for inhibition potency, selectivity and *in vitro* stability.

3.3 Material and methods

Molecular Docking. Molecular docking studies were performed as described previously with slight modifications.[138] Briefly, docking was performed with LeadIT-2.3.2 using the hepsin-inhibitor complex crystal structure (PDB-ID: 1Z8G), freely available in the protein data bank (PDB).[46, 144-146] The binding site was defined by all residues within 6.5 Å of the crystallographic reference ligand acetyl-Lys-Gln-Leu-Arg-chloromethylketone and further included highly coordinated water molecules 188, 302, 409, 425, 438, 448, 476, 524 and 856. Protonation states were determined with the Protoss module within LeadIT.[147] The docking setup was validated by re-docking of the reference ligand (FlexX-score -50.7 kJ/mol, RMSD 2.7 Å). Structures of potential inhibitors were generated using CycloPs and modified to contain an N-terminal acetyl cap and a C-terminal ketobenzothiazole (kbt) warhead.[121] Energetically minimized structures for docking were generated using the AMBER14:EHT force field within MOE2018.[123, 148] Differently from the previous approach, the docking was performed without any pharmacophoric constraints.[138] For selectivity studies, docking against matriptase was performed using PDB-ID: 4O9V as the receptor with residues 6.5 Å around the reference ligand including water molecule 1004 (re-docking FlexX-score -55.3 kJ/mol, RMSD 1.7 Å).[149] Figures were made with PyMOL.[150]

Chlorination of tritylhydroxide resin (Trt-OH).[151] The chlorination of trtiylhydroxide resin was performed based on a modified method described elsewhere.[151] In short, a 250 mL round bottom flask was rinsed with dry dichloromethane (DCM) and 25 g of Trt-OH (0.8 mmol/g, mesh 100-200, company: Iris Biotech) was added. The resin was suspended in a mixture of 50% DCM and 50% toluene, just enough to double the resin volume and 10 mL acetyl chloride was added. The glass vial was sealed and agitated for 24 h. The next day, the resin was dried and thoroughly washed with DCM (4 × 5 mL). Chlorinated resin (Trt-Cl) was stored in freezer.

Addition of first amino acid to Trt-Cl.[151] 15 mL of dry DCM was added to 0.8 g of Trt-Cl resin (0.8 mmol/g) and shaken for 10 min. Next, 2.5 equiv. (2.5 mmol) of amino acid and 5 equiv. (5 mmol) of activator base was added and the mixture was shaken overnight at room temperature. The next day, the resin was filtered and washed with *N,N*-dimethylformamide (DMF) (2 × 10 mL) and DCM (2 × 10 mL). Prior to use for solid phase peptide synthesis the resin was swollen with 10 mL DMF.

Solid phase peptide synthesis. Peptides were synthesized as C-terminal amides on a loaded 2-chlorotrityl chloride resin using 9-fluorenylmethoxycarbonyl (Fmoc) strategy. All Fmoc-protected amino acids were purchased from Novabiochem, as well as benzotriazol-1-

yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP), *N,N*-diisopropylethylamine (DIEA) , DMF, DCM and were used as received. All Fmoc-protected amino acids were dissolved in DMF at a concentration of 0.2 M. The coupling was done with the standard procedure of solid phase peptide synthesis.[152] In short, Fmoc deprotection was performed by treating the peptidyl resin with 20% piperidine in DMF for 10 min at 70 °C. After the reaction, the resin was washed with DMF and DCM, each (2 x 10 mL), then filtrated. The coupling was done by addition of 7.5 mL Fmoc-protected amino acid (3 equiv.), 3 mL of 1 M PyBOP (3 equiv.) and 1.5 mL of 0.5 M DIEA (3 equiv.). The reaction mixture was shaken at 75 °C for 30 minutes. Afterwards, the resin was washed as described above and used for the next step.

N-Cap modification and cleavage from the resin. The N-terminus of the peptides with protected side chains were acetyl capped using 20 mL of 0.5 M Ac_2O/DMF and 1 M DIEA/DMF. The reaction was shaken for 2 h at room temperature. The resin was filtered and washed with DMF and DCM, each (2 × 10 mL). The cleavage of the peptide with protected side chains was accomplished using 20 mL of 20% hexafluoroisopropanol (HFIP) in DCM. The mixture was put on an orbital shaker for 3 h at room temperature and then filtered. The solvent was taken off *in vacuo* and the peptide was precipitated using 50 mL of diethylether.

Synthesis of Boc-Arg(Mtr) weinreb amide (S1) *((S)-tert*-Butyl-(1-(methoxy(methyl)amino)-5-(3-((4-methoxy-2,3,6-trimethylphenyl)sulfonyl)guanidino)-1-oxopentan-2-yl)carbamate).[153] To a solution of *N*-(*tert*-butoxycarbonyl)-*N*-((4-methoxy-2,3,6-trimethylphenyl)sulfonyl)-L-arginine (1.56 g, 3.21 mmol, 1.0 equiv.) in 34 mL THF, *N,O*-dimethylhydroxyl-amine hydrochloride (0.64 g, 6.42 mmol, 2.0 equiv) and 1-hydroxybenzotriazole hydrate (0.54 g, 3.53 mmol, 1.1 equiv.) were dissolved at room temperature. DIEA (1.63 mL, 9.63 mmol, 3 equiv.) and EDCl (0.65 g, 3.37 mmol, 1.05 equiv.) were added and the solution was stirred for 4.5 h. The mixture was concentrated *in vacuo* and the reaction mixture extracted with ethyl acetate (150 mL), washed with 5% aqueous acetic acid (75 mL), sat. aqueous $NaHCO_3$ (75 mL), water (75 mL), and brine (75 mL). The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo* to yield the Weinreb amide **S1** (0.86 g, 1.6 mmol, yield: 69%) as a white powder (Figure A4). Purity (HPLC, 220 nm) > 80%. ^{1}H NMR (300 MHz, CD3OD): δ=1.43 (s, 9 H) 1.49–1.60 (m, 4 H) 1.87 (s, 1 H) 2.13 (s, 3 H) 2.61 (s, 3 H) 2.67 (s, 3 H) 3.17 (m, 3 H) 3.74 (s, 3 H) 3.90 (s, 3 H) 6.67 ppm (s, 1 H). MS (ESI): *m/z*: calcd for $C_{23}H_{39}N_5O_7S_2$ $[M+H]^+$ 530.3, $[2M+H]^+$ 1059.6, found $[M+H]^+$ 530.2, $[2M+H]^+$ 1059.4

Synthesis of Boc-Arg(Mtr) ketobenzothiazole (S2) ((*S*)-*tert*-Butyl-(1-(benzo[*d*]thiazol-2-yl)-5-(3-((4-methoxy-2,3,6-trimethylphenyl)sulfonyl)guanidino)-1-oxopentan-2-yl)carbamate).[142] To a solution of benzothiazole (0.47 g, 3.53 mmol, 1.1 equiv.), *n*-BuLi (1.6 M, 4 mL, 6.42 mmol,

2 equiv.) was added dropwise to THF (50 mL). After the mixture was stirred for an additional 30 minutes, Boc-Arg(Mtr) Weinreb amide (1.94 g, 3.21 mmol, 1 equiv.) was dissolved in THF (15 mL) and added slowly over 50 minutes. The mixture was stirred at -78.8 °C for 3 h. The reaction was quenched with sat. aqueous NH$_4$Cl (30 mL) and the aqueous layer was extracted with EtOAc (40 mL). The organic phase was collected, dried with Na$_2$SO$_4$, and then concentrated. The resulting residue was purified by semi-preparative RP-HPLC to yield the compound **S2** (0.32 g, 0.5 mmol, yield: 35%) as a yellow powder (Figure A4). Purity (HPLC, 220 nm) > 95%. ^{1}H NMR (300 MHz, CD3OD): δ=1.43 (s, 9 H) 1.66 (m, 4 H) 2.05 (s, 3 H) 2.55 (s, 3 H) 2.70 (s, 3 H) 3.23 (m, 2 H) 3.81 (s, 3 H) 5.32 (m, 1 H) 6.55 (s, 1 H) 7.69 (m, 2 H) 8.27 ppm (m, 2 H). MS (ESI): *m/z*: calcd for C$_{28}$H$_{37}$N$_5$O$_6$S$_2$ [M+H]$^+$ 604.2, [2M+H]$^+$ 1207.4, found [M+H]$^+$ 603.9, [2M+H]$^+$ 1206.6.

Synthesis of HCl•H-Arg(Mtr) ketobenzothiazole (S3) (*N*-(*N*-(4-Amino-5-(benzo[*d*]thiazol-2-yl)-5-oxopentyl)carbamimidoyl)-4-methoxy-2,3,6-trimethylbenzenesulfonamide).[141]
Compound **S2** (0.200 g, 0.40 mmol) was stirred in 1.5 M HCl/dioxane (10 mL) at room temperature for 18 h. The solvent was removed the resulting residue was dried *in vacuo* and purified by RP-HPLC (Figure A4). The received compound was used for the following peptide couplings.

Synthesis of Boc-Arg(Mtr) ketothiazole (**S4**) ((*S*)-*tert*-Butyl-(5-(3-((4-methoxy-2,3,6-trimethylphenyl)sulfonyl)guanidino)-1-oxo-1-(thiazol-2-yl)pentan-2-yl)carbamate).[154] To a solution of 2-bromothiazol (0.211 g, 1.29 mmol, 3.3 equiv.) in dry THF (10 mL) *n*-BuLi (2.5 M, 0.52 mL, 1.29 mmol, 3.3 equiv.) was added dropwise under inert atmosphere at -78 °C. The reaction mixture stirred for 1.5 h at -78 °C, followed by dropwise addition of compound **S1** (0.205 g, 0.39 mmol, 1 equiv.) at the same temperature. The resulting solution was stirred 2 h at -78 °C, after which sat. aqueous NH$_4$Cl (10 mL) was added. The organic phase was separated and the aqueous phase was extracted three times with EtOAc. The combined organic extracts were washed with brine (30 mL), dried over Na$_2$SO$_4$, filtered and concentrated *in vacuo*. The residue was purified on a silica column eluting with EtOAC/cyclohexane (4:1 v/v), to afford the compound S4 (0.12 g, 0.22 mmol, yield: 56%) as a white foam (Figure A4). Purity (LC, 254 nm) > 95%. ^{1}H NMR (300 MHz, CDCl$_3$): δ = 8.04 (d, 1 H), 7.72 (d, 1 H), 6.53 (s, 1 H), 5.64 (d, 1 H), 5.41 (s, 1 H), 3.83 (s, 3 H), 3.26 (m, 2 H), 2.67 (s, 3 H), 2.59 (s, 3 H), 2.12 (s, 3 H), 1.76 – 1.57 (m, 4 H), 1.41 (s, 9 H). ppm. MS (ESI): *m/z*: calcd for C$_{24}$H$_{35}$N$_5$O$_6$S$_2$ [M+H]$^+$ 554.2, found [M+H]$^+$ 554.2.

Synthesis of TFA•H-Arg(Mtr) ketothiazole (S5) ((*S*)-*N*-(*N*-(4-Amino-5-oxo-5-(thiazol-2-yl)pentyl)carbamimidoyl)-4-methoxy-2,3,6-trimethylbenzenesulfonamide). Compound **S4**

(0.256 g, 0.46 mmol) was stirred in DCM (3 mL) at 0 °C and TFA (1 mL) was added. The reaction mixture stirred for 1 h at ambient temperature, then isopropyl alcohol (0.5 mL) was added. The solution was concentrated *in vacuo* and triturated with diethyl ether. The supernatant was decanted and the residue was purified by RP-HPLC (Figure A4). The obtained compound **S5** was used for the following peptide couplings.

Synthesis of Boc-Arg(Mtr) alcohole (S6) ((*S*)-*tert*-Butyl-(1-hydroxy-5-(3-((4-methoxy-2,3,6-trimethylphenyl)sulfonyl)guanidino)pentan-2-yl)carbamate). To a solution of Boc-Arg(Mtr)-OH (0.3 g, 0.62 mmol, 1 equiv.) in dry THF (5 mL) were added NMM (0.063 g, 0.62 mmol, 1 equiv.) and EtOCOCl (0.067 g, 0.62 mmol, 1 equiv.) at -15 °C under argon. The reaction mixture stirred for 1 h at -15 °C, then transferred dropwise via canula into a stirred solution of $NaBH_4$ (0.047 g, 1.24 mmol, 2 equiv.) in water (15 mL). The resulting solution was stirred 5 minutes at 0 °C and then diluted with water (15 mL). The aqueous phase was extracted twice with EtOAc (10 mL). The combined organic extracts were dried over Na_2SO_4 and concentrated *in vacuo* to obtain the compound **S6** (0.24 g, 0.5 mmol, yield: 81%) as a colourless oil (Figure A4). Purity (LC, 254 nm) 98%. ^{1}H NMR (300 MHz, $CDCl_3$) δ = 6.52 (s, 1 H), 6.33 (s, 2 H) 5.15 (d, 1 H), 3.82 (s, 3 H), 3.55 (s, 2 H), 3.21 (s, 1 H), 2.69–2.66 (m, 5 H), 2.59 (s, 3 H), 2.12 (s, 3 H), 1.55 (s, 4 H), 1.40 (s, 9 H) ppm. LC-MS: *m/z*: calcd for $C_{21}H_{36}N_4O_6S$ [M+H]$^+$ 473.2, found [M+H]$^+$ 473.2.

Synthesis of TFA•H-Arg(Mtr) alcohol (S7) ((*S*)-*N*-(*N*-(4-Amino-5-hydroxypentyl)carbamimidoyl)-4-methoxy-2,3,6-trimethylbenzenesulfonamide). Compound **S6** (0.22 g, 0.47 mmol) was stirred in DCM (3 mL) at 0 °C and TFA (1 mL) was added. The reaction mixture stirred for 2 h at ambient temperature, then isopropyl alcohol (0.5 mL) was added. The solution was concentrated *in vacuo* and triturated with diethyl ether. The supernatant was decanted and the residue was purified by RP-HPLC. The received compound **S7** was used for the following peptide couplings (Figure A4).

Preparation of inhibitors. Respective serine traps (1.5 equiv.) were coupled with tripeptides (1.0 equiv.) bearing standard protection groups using PyBOP (1.5 equiv.) and DIEA (3 equiv.) in DMF. After the reaction was agitated for 4 h at room temperature, 3 mL of a deprotection solution was added (93% TFA, 3.5%TIPS, 3.5% H_2O) and further agitated for 8 h at room temperature. After concentrating, the crude inhibitor was precipitated in 50 mL cold diethylether and afterwards purified using RP-HPLC (Figure A5).

Purification, lyophilization and analysis of peptide- and warhead precursors and inhibitors. All precursor compounds were purified with a semi-preparative RP-HPLC. The

following gradient was applied: 95% H_2O / 5% ACN to 5% H_2O / 95% ACN in 30 min. Trifluoracetic acid was dissolved in the water to a concentration of 0.1%. Zorbax Eclipse XDB column C-18 9.4x250 mm 5 μm, Agilent Technologies. UV Vis detector model S-3702, Soma. For the detection a wavelength of 220 nm was used. After chromatographic purification, the fractions were collected and freeze dried overnight. The purified and lyophilized precursor compounds were stored in the freezer at -20 °C. The mass of the purified compounds was determined with MS-ESI. Expression-L compact mass spectrometer, Advion. Peptides were dissolved to a concentration of c = 0.01 mg/mL in MeOH + 0.1% formic acid. Injection was done by a syringe pump with a flow rate of 10 μL/min.

Enzymes. Recombinant human hepsin protein and factor Xa protein were purchased from Bio-Techne GmbH (Wiesbaden, Germany). Recombinant human thrombin protein and matriptase protein were purchased from R&D Systems (Minneapolis, MN, USA).

Determination of inhibitory constant K_i. The activity of the compounds against the recombinant human enzymes was determined in enzyme inhibition assays. Here, a ten-point dilution series for the inhibitors was prepared and incubated for 30 minutes with the enzyme in TNC buffer (25 mM Tris, 150 mM NaCl, 5 mM $CaCl_2$, 0.01% Triton X-100, pH = 8) prior to adding a fluorogenic reference substrate Boc-Gln-Ala-Arg-AMC for hepsin and matriptase or a chromogenic substrate D-Phe-Homopro-Arg-pNA, Bz-Ile-Glu-Gly-Arg-pNA for thrombin and factor Xa, respectively. The preincubation time did not significantly influence activity of the compounds (Table 3.1). However, due to consistency with published literature, the 30 minutes preincubation time was maintained. The measurements were performed on a Tecan infinite® M1000 and the fluorescence intensity was measured by exciting the AMC fluorophore at 380 nm wavelength and recording emission at 460 nm wavelength. The absorption of pNA was measured at 405 nm. Fluorescence intensities and absorption were measured every 2 min for 2 h. The end concentrations of the enzymes were 0.3 nM (hepsin) and 0.2 nM (matriptase) in 20 μL total volume and 0.6 nM (thrombin), 0.35 nM (factor Xa) in 100 μL total volume. The end concentration of the reference substrate was 100 μM (matriptase), 157 μM (hepsin), 200 μM (factor Xa) and 100 μM (thrombin). To determine the IC_{50} values, the concentration-response data were plotted with the program GraphPad prism version 8.4.2 (San Diego, California) and a nonlinear regression fit with the equation [Inhibitor] vs. normalized response was applied. The inhibitory constant K_i was calculated from the IC_{50} values using the Cheng-Prusoff equation (K_i = IC_{50}/[S]/K_M) for competitive reversible inhibitors.[106]

Table 3.1. Influence of preincubation on inhibitory activity of compound 3 and 4 against hepsin. Peptidomimetic inhibitors were added to hepsin with 30 minutes preincubation time or without preincubation prior addition of the fluorogenic reference substrate Boc-Gln-Ala-Arg-AMC. The velocity of substrate degradation was assessed by recording the fluorescence intensity at 460 nm. Shown are the means ± SD of triplicate measurements.

Compound	K_i [nM]	
	w/ preincubation	w/o preincubation
3	9,4 ± 3,2	6,2 ± 2,2
4	5,6 ± 0,6	6,3 ± 0,5

Stability of inhibitors in serum, plasma and cell culture medium. The stability of the inhibitors was measured according to a modified procedure from Jensen et al.[130] In short, hepsin inhibitor solutions of 1 mM were prepared by dissolving the peptide in pure dimethyl sulfoxide. 10 µL of the peptide solution were added to 1 mL RPMI medium 1640 supplemented with 25% (v/v) human serum, citrate plasma or only RPMI medium and incubated at 37 °C. At indicated intervals 100 µL samples were taken and mixed with 200 µL ethanol for precipitation of proteins. The cloudy solution was cooled at 4 °C for 15 minutes and centrifuged at 14,800 rpm for 2 minutes. The supernatant was aspirated and analyzed using analytical HPLC. The residual inhibitory constants K_i after 1 day incubation in biological fluids was measured as described above.

3.4 Results

Design and synthesis of a peptidomimetic kbt inhibitor library. In a previous publication, we reported the computer-aided development of peptide substrates for the serine protease hepsin.[138] In this former publication, we used molecular docking to screen two peptide libraries targeting the P- and P′-site of the hepsin's catalytic active cleft and ranked the peptides based on their calculated binding affinity. For the identification of novel peptide based hepsin inhibitors with specific antiproteolytic properties, we utilized the above-mentioned highest scoring peptide combinations targeting the P-site (Table A4.) and designed an inhibitor library by coupling the tripeptides with a highly reactive, C-terminal electrophilic group, the serine trap, which enables interaction with the protease's catalytic triad. Further, the P-site recognition sequence of the endogenous substrate HGF (Lys-Gln-Leu-Arg) and of the previously identified high affinity substrate (Arg-Gln-Leu-Arg) was also included.[46, 47] The common structure is as follows: N-terminal acetyl cap (ace), proteinogenic amino acids at P4-P3, arginine at P1 position and a ketobenzothiazole (kbt) warhead (Table 3.2).

Table 3.2. Assembled peptidomimetic inhibitor library for synthesis.[a] recognition sequence of previously identified high affinity substrate (RS).[b] recognition sequence of natural substrate (NS), PDB-ID: 1Z8G. For RS rank 2 docking pose was evaluated due to an inverted binding mode of the rank 1 pose. Kbt = ketobenzothiazole, kt = ketothiazole, ol = alcohol, carb = carboxyl, pvs = phenylvinylsulfon.

Compound	N-cap	P4	P3	P2	P1	Serine trap
RS[a]	ace	Arg	Gln	Leu	Arg	kbt
NS[b]	ace	Lys	Gln	Leu	Arg	kbt
1	ace	Arg	Glu	Cys	Arg	kbt
2	ace	Ala	Glu	Gln	Arg	kbt
3	ace	Gly	Gln	Arg	Arg	kbt
4	ace	Gln	Gln	Val	Arg	kbt
5	ace	Arg	His	Gln	Arg	kbt
6	ace	Glu	Gln	Pro	Arg	kbt
7	ace	Asn	Glu	Gln	Arg	kbt
3.1	ace	Gly	Gln	Arg	Arg	kt
3.2	ace	Gly	Gln	Arg	Arg	pvs
3.3	ace	Gly	Gln	Arg	Arg	carb
3.4	ace	Gly	Gln	Arg	Arg	ol
8	ace		Gln	Arg	Arg	kbt
9	ace			Arg	Arg	kbt
10	ace				Arg	kbt

The designed inhibitor library was virtually prepared and docked in the active binding cleft of hepsin. All compounds revealed plausible binding modes with likeliness for reaction and high calculated binding affinities (Table 3.3).

Table 3.3. Docking studies of assembled peptidomimetic inhibitor library. Shown are the determined docking FlexX-scores for rank 1 poses with distance between electrophilic carbon of the kbt warhead and nucleophilic Ser353 oxygen to estimate reaction likeliness.[a] P-site recognition sequence of a literature known preferred substrate RQLR↓VVGG.[b] P-site recognition sequence of the endogenous substrate HGF.[138]

Compound	FlexX-score [kJ/mol]	Distance electrophilic carbon to nucleophilic oxygen [Å]
RS[a]	-54.86	2.8
NS[b]	-53.99	2.6
1	-59.37	2.5
2	-64.25	2.6
3	-56.37	2.5
4	-61.44	3.2
5	-56.45	3.0
6	-52.12	2.7
7	-62.05	2.7

In the next step, we synthesized the designed inhibitor library using a two-step approach, previously reported in literature (Scheme 3.1). In the first step, precursor compounds were prepared. Using a solid phase peptide synthesis approach, N-acetylated side chain protected tripeptides with a free C-terminus (I) were synthesized. Next, the side-chain protected Arg-kbt warhead was prepared (II) and coupled with the tripeptidic compounds using PyBOP/DIEA as coupling reagents. Following side-chain deprotection, the compounds were subjected to RP-HPLC purification and then lyophilized (Figure A5).
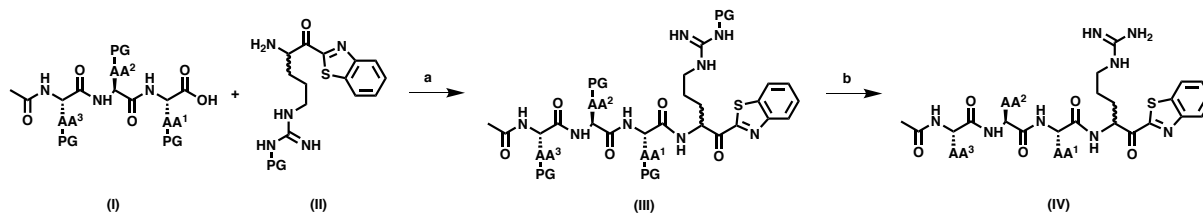
Scheme 3.1. Synthesis of tripeptidic inhibitors. Reagents and conditions: (a) PyBOP/DIEA, RT, 2 h; (b) TFA/TIPS/H_2O (93/3/4), 8 h. PG = protection group; AA = amino acid.

Evaluation of tetrapeptidomimetic kbt inhibitors. For the analysis of antiprotease activity, the inhibitors were incubated with the target protease hepsin or off-target proteases matriptase, thrombin and factor Xa upon addition of a reporter substrate. Overall, multiple members of the P-site library showed a high potency in the low nanomolar concentration range against hepsin and the closely related protease, matriptase, in the fluorogenic inhibition assay (Table 3.4, Figure A6). Compound **RS** with the P-site recognition sequence of the high hepsin susceptible substrate and compound **NS** with the P-site recognition sequence of the structural similar endogenous substrate HGF, revealed the highest activity against hepsin with K_i = 0.8 nM and K_i = 0.5 nM, respectively. The compounds **1-5** and **7** suppressed hepsin activity in the range of K_i = 1.7 - 60 nM. Among the most potent candidates, **RS**, **NS**, **1, 4** and **7** revealed a 3.9-26.6-fold selectivity for hepsin. In contrast, we found that the compounds **2** and **3** are more selective for matriptase (3.5- and 5.2-fold, respectively). The compounds **5** and **6** showed poor discrimination between the two proteases.

For systemic administration of hepsin inhibitors, high selectivity against the large amounts of coagulation proteases present in blood are necessarily required. Therefore, the inhibitors were also tested against thrombin and factor Xa. All compounds did not inhibit thrombin measurable in the used concentration range. The compounds **RS, NS, 1** and **2** inhibit factor Xa in the low nanomolar range with selectivity ratios varying between 12-73.6-fold. The highest factor Xa selectivity reveal compound **3** with 1194-fold and compound **4** with 248.4-fold.

To gain a better understanding of the structure-activity relationships and selectivity-determining features between hepsin and its off-target TTSPs, the protease binding sites and predicted binding modes of the inhibitors were elucidated in detail. Matriptase, thrombin and factor Xa share an overall sequence similarity of 55%, 46% and 49% with hepsin. Within the binding site (defined as residues within 6 Å of the reference ligand ace-Lys-Gln-Leu-Arg-chloromethylketone, PDB-ID: 1Z8G), similarity even increases to 65%, 58% and 60%, respectively. With the common preference of theses TTSPs for basic residues within the S1 pocket, special attention has to be paid in the design of selective inhibitors.[41] Based on the molecular docking studies, the arginine in P1 position is deeply buried in the S1 pocket of hepsin forming polar interactions with Asp347 and Gly380. Thereby, the electrophilic carbon is placed in proximity of the catalytic Ser353 which facilitates covalent-reversible binding (Figure 3.2). The S2 pocket of hepsin is characterized by a small hydrophobic sub-pocket which accommodates a methyl group of small hydrophobic amino acids like Val or Leu as found in **RS, NS** and **4** and is likely to be a selectivity-determining feature. Differently from hepsin, the S2-pocket of matriptase offers no small sub-pocket for a methyl group as it is blocked by Phe99 (corresponding to Asn254 in hepsin, Figure 3.2). Larger, polar residues like Gln in compound **2** or especially Arg in compound **3** might enable favorable contacts in the remote S2-pocket with Asn254 and Ser251 in hepsin or Asp96 in matriptase.[140] Gln at P3

further improved the affinity by additional polar interactions with Gln350 and Gly378 in hepsin which also formed H-bonds with the peptidic backbone. Lastly, P4 tolerates polar and or basic residues, while acidic residues as Glu (compound **6**) or small non-polar as Ala (compound **2**) do not make significant contribution. Similar binding modes were reported for multi-basic high affinity matriptase inhibitors before.[149, 155]
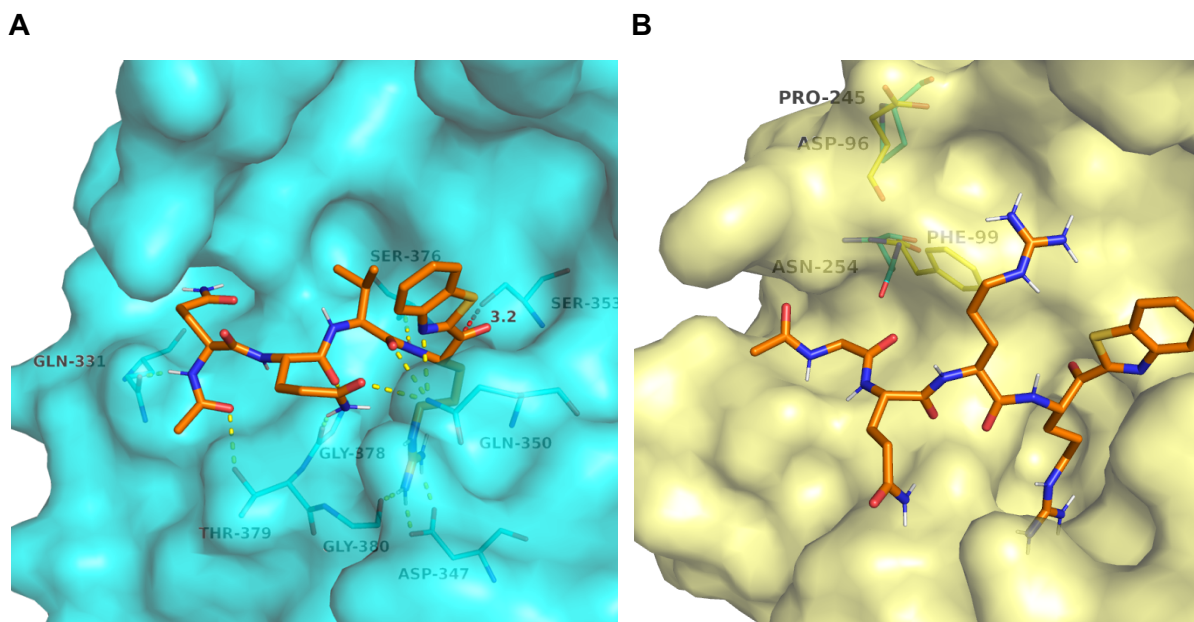
Figure 3.2. A Predicted binding mode of compound 4 in complex with hepsin (PDB-ID: 1Z8G). **B** Predicted binding mode of compound 3 in complex with matriptase (PDB-ID: 4O9V) with superposed residues of hepsin. Ligand carbon atoms are depicted in orange, hepsin carbon atoms and surface in cyan, matriptase carbon atoms and surface in yellow. Polar interactions are shown as yellow dashed lines, the distance between electrophilic carbon of kbt and nucleophilic oxygen of Ser as red dashed line. For clear view only residues described in the text are depicted and labeled.

In the next step, we investigated the change in activity through progressive, N-terminal truncations on compound **3**. Overall, most of the above-mentioned rationales for contributing activity against hepsin could be confirmed and removal of the S2 and S3 residues led to a decrease in inhibitory potency (Table 3.4, Figure A7). The S4 Gly residue of compound **3** is negligible as it does not increase inhibition compared to compound **8** (ace-Gln-Arg-Arg-kbt). The smallest possible structural element still revealing inhibitory activity was shown to be compound **10** (ace-Arg-kbt), however, no broad selectivity can be expected among similar proteases. Further, truncation of the ketobenzothiazole (**3**) moiety to a ketothiazole (**3.1**) led to a ~3-fold decrease in activity against hepsin and matriptase while introduction of a phenylvinylsulfone (**3.2**), which is widely used for the development of potent cysteine protease inhibitors, strongly decreased activity. Replacement of the serine trap by a carboxylic acid (**3.3**) or alcohol (**3.4**) abolished any antiprotease activity (Table 3.4, Figure A8).

Table 3.4. Inhibition of hepsin and off-target proteases matriptase, thrombin and factor Xa in fluorogenic inhibition assay. Selectivity indices represent the quotient of K_i values of matriptase, thrombin, and factor Xa by the K_i value of hepsin. n.d = not determined. Kbt = ketobenzothiazole, kt = ketothiazole, ol = alcohol, carb = carboxyl, pvs = phenylvinylsulfon.

Compound	K_i [nM]				selectivity indices		
	Hepsin	Matriptase	Thrombin	Factor Xa	Matriptase	Thrombin	Factor Xa
RS	0.8	3.8	> 50,000	290.7	4.8	n.d	73.6
NS	0.5	12	> 50,000	196.1	26.6	n.d	15.8
1	1.7	6.5	> 50,000	93.1	3.9	n.d	14.4
2	60	17.1	> 50,000	205.1	0.3	n.d	12
3	7.3	1.4	> 50,000	1,668	0.2	n.d	1,194
4	5.1	79.1	> 50,000	19,643	15.5	n.d	248.4
5	21.7	55.8	> 50,000	1,167	2.6	n.d	20.9
6	467.2	374	> 50,000	2,647	0.8	n.d	7.1
7	20.2	157.6	> 50,000	3,081	7.8	n.d	19.6
3.1	19.8	3.9	> 50,000	34,529	0.2	n.d	8,847
3.2	14,364	4,971	> 50,000	> 50,000	0.3	n.d	n.d
3.3	> 50,000	> 50,000	> 50,000	> 50,000	n.d	n.d	n.d
3.4	> 50,000	> 50,000	> 50,000	> 50,000	n.d	n.d	n.d
8	4.3	0.4	> 50,000	442.2	0.1	n.d	1,147
9	72.4	5	> 50,000	12,626	0.1	n.d	2,544
10	4,623	144.6	> 50,000	> 50,000	0.03	n.d	n.d

Inhibitor stability of selected inhibitors in body fluids. Stability of peptidomimetic inhibitors is a main challenge in peptidic drug development. Peptidases might degrade the inhibitor before reaching the desired target in an adequate time. For this reason, the N-terminus of the P-site recognition sequence was modified by acetylation which potentially resist enzymatic degradation by aminopeptidases, endopeptidases and synthetases.[156] The compounds **3** and **4** were selected due to their high activity against hepsin and good selectivity over coagulation proteases for *in vitro* stability studies. Both compounds were spiked in 25% blood serum and the inhibitory activity was tested at different time points from 0 min to 10 days. As shown in

Figure 3.3, even after 1 day incubation, both compounds still have the potential to fully prevent degradation of the fluorogenic reference substrate at a concentration of 1 µM. After 10 - 20 days, the peptidomimetic inhibitors start to lose their potency.
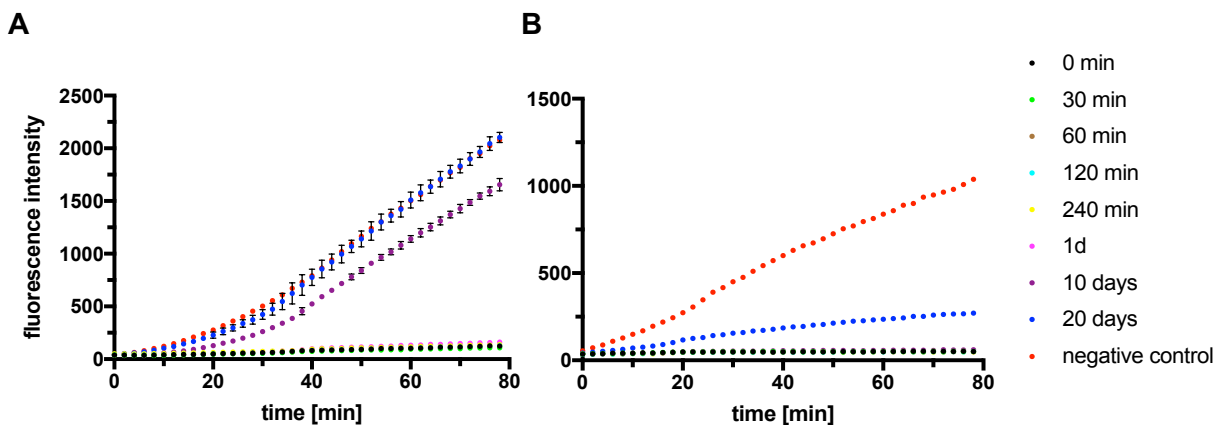
Figure 3.3. Inhibition of reference substrate degradation by hepsin with **A** compound 3 and **B** compound 4 at c = 1000 nM after spiking in 25% blood serum at various time points.

Analysis of the residual activity revealed only slight loss of initial potency after 1 day incubation in body fluids (Figure 3.4). The most prominent loss can be seen for compound **3** which lost 5-fold activity after incubation with plasma for 1 day incubation.
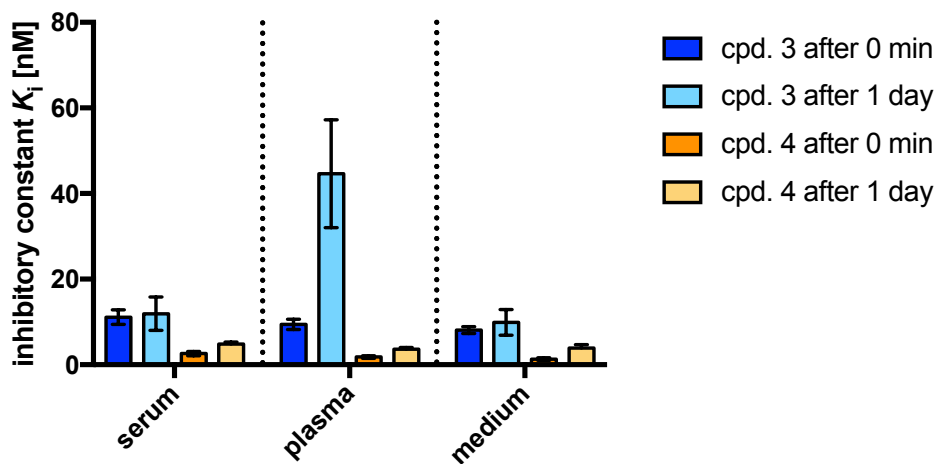
Figure 3.4. Residual inhibitory activity for compound 3 and compound 4 against hepsin after 0 minutes and 1 day incubation in 25% serum, 25% plasma and pure RPMI cell medium.

3.5 Discussion

In this study we describe the development of novel potent and stable peptidomimetic hepsin inhibitors. Hepsin represents a promising target through its ability to proteolytically activate growth factor proteins which lead to cell migration and eventually metastasis in cancerous tissue. As a transmembrane serine protease, its accessible localization on the cell-surface and its high upregulation in cancerous diseases, especially in prostate cancer compared to benign tissue makes it of particular interest.

We used a previously published strategy by our group which allowed the identification of suitable peptide "binders" which were calculated with high scoring binding affinities for the S- and S′-site of the binding cavity of hepsin. Assembling of those "binders" led to efficient substrates for hepsin. This study uses modifications of this strategy for the development of novel peptidomimetic inhibitors. For this purpose, the top scoring peptide combinations targeting the S-site of the binding pocket were used as recognition sequences for the design of an inhibitor library. The recognition sequences of tetrapeptides were N-acetyl capped and C-terminally connected with a reactive kbt warhead moiety to yield covalent, reversible inhibitors. The inhibitory potential was tested for isolated enzymes which revealed for multiple compounds high activity against hepsin and matriptase with good off-target selectivity over the coagulation proteases factor Xa and overall high selectivity against thrombin. The highest potency against hepsin was demonstrated for the compounds RS and NS with K_i < 1 nM, however, the compounds **3** and **4** with K_i ≈ 5-7 nM against hepsin revealed a significantly higher selectivity (250-1200-fold) against factor Xa. Like hepsin, the closely related TTSP matriptase was also shown be upregulated in cancer and to activate the endogenous ligand HGF. However, inhibition of this protease might lead to side-effects due to the broad expression in many epithelial cells. The highest matriptase selectivity revealed compound **4** (~16-fold) and our data suggest the small S2 pocket to be the selectivity-determining feature for hepsin.

We subjected our top compounds **3** and **4** stability tests in body fluids which showed high remaining potency even after 1 day incubation. The kbt moiety of the serine trap proved to be most active which is consistent with previous findings in literature.[141, 142, 157] Abandoning the electrophilic serine trap by replacing with hydroxy or carboxy groups resulted in a loss of inhibitory potency. However, substitutions on the heterocycle to additionally target the prime site of the binding pocket could offer a strategy to further increase hepsin selectivity. Our results might warrant deeper analysis of pharmacokinetics, broader protease selectivity or *in vivo* stability studies. Of note, structural similar serine proteases like TMPRSS2 were shown to play an important role in viral infections like COVID-19. The results in this study might be used for the design and development of new anticancer or antiviral drugs.

Chapter B - Inhibitors for TMPRSS2 to block SARS-CoV-2 infection

Chapter B comprises the subchapters 4-6. The fourth subchapter gives an overview over general aspects and a theoretical background about coronaviruses, in particular SARS-CoV-2, and the utilization of inhibitors for the host protease TMPRSS2 as potential therapeutic strategy. This subchapter is a selected excerpt from a published review **[3]**. The fifth subchapter presents the development of peptidomimetic TMPRSS2 inhibitors. It includes the computational design, synthesis and testing of the inhibitors. The results were submitted in a peer-reviewed journal **[4]**. The sixth subchapter presents a study which showed the protein alpha-1 antitrypsin as an endogenous inhibitor of TMPRSS2 and its implication of preventing viral infections **[5]**. These results were published in a peer-reviewed journal and the text below is therefore in nearly all instances a word-by-word reproduction of the peer-reviewed and published paper.

[3] Patrick Müller*, Hannah Maus*, Stefan Hammerschmidt*, **Philip Maximilian Knaff**, Volker Mailänder, Tanja Schirmeister and Christian Kersten. **Interfering with Host Proteases in SARS-CoV-2 Entry as a Promising Therapeutic Strategy.** *Current Medicinal Chemistry* **2021**.

[4] Philip Maximilian Knaff*, Christian Kersten*, Lukas Wettstein, Patrick Müller,Tatjana Weil, Carina Conzelmann, Janis Müller, Maximillian Brückner, Stefan Pöhlmann, Markus Hoffmann Tanja Schirmeister, Jan Münch, Volker Mailänder. **Development of Potent Peptidomimetic Inhibitors for the Serine Protease TMPRSS2**. *Nature Communications* **2021** (Submitted).

[5] Lukas Wettstein*, Tatjana Weil, Carina Conzelmann, Janis A. Müller, Rüdiger Groß, Maximilian Hirschenberger, Alina Seidel, Susanne Klute, Fabian Zech, Caterina Prelli Bozzo, Nico Preising, Giorgio Fois, Robin Lochbaum, **Philip Maximilian Knaff**, Volker Mailänder, Ludger Ständker, Dietmar Rudolf Thal, Christian Schumann, Steffen Stenger, Alexander Kleger, Günter Lochnit, Benjamin Mayer, Yasser B. Ruiz-Blanco, Markus Hoffmann, Konstantin M. J. Sparrer, Stefan Pöhlmann, Elsa Sanchez-Garcia, Frank Kirchhoff, Manfred Frick and Jan Münch. **Alpha-1 antitrypsin inhibits TMPRSS2 protease activity and SARS-CoV-2 infection.** *Nature Communications* **2021**.

4. Theoretical background

Copyright:

This sub-chapter 4 was published in a peer-reviewed journal **[3]**. Therefore, the text below is in nearly all instances a word-by-word reproduction of the peer-reviewed and published paper. Presented results are reprinted with permission from Bentham Science Publishers, Current Medicinal Chemistry. Copyright © 2021 Bentham Science Publishers.

[3] Patrick Müller*, Hannah Maus*, Stefan Hammerschmidt*, <u>**Philip Maximilian Knaff**</u>, Volker Mailänder, Tanja Schirmeister and Christian Kersten. **Interfering with Host Proteases in SARS-CoV-2 Entry as a Promising Therapeutic Strategy.** *Current Medicinal Chemistry* **2021** (Accepted). DOI: https://doi.org/10.2174/0929867328666210526111318.

Aim:

Studies have shown that cell entry of coronaviruses, including SARS-CoV-2, takes place through binding of a viral surface protein called spike onto host cell receptors. Subsequently, host proteases proteolytically process the spike protein which triggers a mechanism leading to fusion of viral with the host cell membrane. This review with the title "Interfering with Host Proteases in SARS-CoV-2 Entry as a Promising Therapeutic Strategy" aims to discuss the current state of protease inhibitor development which impair processing of the spike protein, thus preventing viral infection.

Contribution:

I contributed for this collaborative review the chapter "The coronavirus spike protein". This project was supervised by Dr. Christian Kersten and Prof. Dr. Tanja Schirmeister.

4.1 Viral infections as potential therapeutic target

In December 2019, a new respiratory disease has been reported in Wuhan (China). Later, the severe pneumonia with uncertain etiopathology was assigned to the family of coronaviruses. Due to genomic sequencing that found high similarity to severe acute respiratory syndrome-coronavirus (SARS-CoV) (79% sequence identity), the pathogen was called SARS-CoV-2.[158-161] The infection has rapidly spread out of Wuhan and evolved from a local disease to the pandemic CoV 2019 disease (COVID-19). Until September 27th 2020, the World Health Organization (WHO) has confirmed over 32.7 million COVID-19 infections and 991 000 deaths worldwide.[162] Since there are neither U.S. Food and Drug Administration (FDA) or World Health Organization (WHO) approved vaccines nor specific therapies, social distancing still constitutes the most effective way to control the continuous increasing numbers of infections.[163, 164] Therefore, the worldwide scientific community strives to get further in development of possibilities to cure the pandemic.[165, 166] Remdesivir is one of the first drugs which is used to counter disease progression of COVID-19. It constitutes an intravenous applicated monophosphoramidate nucleoside prodrug, which belongs to the RNA-polymerase inhibitors and was originally developed to treat the Ebola virus.[167, 168] Remdesivir received emergency use authorization from the FDA on May 1st, 2020 for the treatment of suspected or laboratory-confirmed COVID-19 infections in adults and children which were hospitalized with severe disease.[169, 170] Despite the use of remdesivir, it appeared that the treatment with the drug has little to no effect on hospitalized COVID-19 patients in overall mortality and duration of hospitalization.[171] Academic institutions and science funders are trying to discover potential treatments by drug repurposing or repositioning. This process involves massive experimental and virtual screening of compounds, that already demonstrated their beneficial safety profiles in humans. Thereby, medicines can readily enter clinical trials and reach approval more cost-effectively. In the past 20 years, not a single drug was approved based on drug repurposing approaches, for new indications. Remdesivir and other drugs like imatinib used for gastrointestinal stromal tumors (GIST) were not discovered with the above mentioned method, but rather from rational testing a reasonable mechanistic hypothesis.[172, 173] In this context, the research in biochemistry and structural biology is an important step for the progression in strategies to effectively treat SARS-CoV-2 infections. The worldwide barrier free sharing of this data, especially in correlation with the existing knowledge about the parent SARS-CoV and middle east respiratory syndrome-coronavirus (MERS-CoV), is of significant importance and provides the foundation of further development.

Coronaviruses belong to the family of *Coronaviridae*. They are classified into four genera, the *α*-CoV, *β*-CoV, *γ*-CoV and *δ*-CoV.[174, 175] Prior to December 2019, six members of CoVs were identified that can be transferred from animals to humans. Among them are the two *α*-CoVs

(HCoV-229E and HKU-NL63) and four β-CoVs (HCoV-OC43, HCoV-HKU1, SARS-CoV and MERS-CoV).[175] The β-CoVs (HCoV-229E, HKU-NL63) are characterized by self-limiting upper respiratory infections. In contrast, SARS-CoV, MERS-CoV and the current SARS-CoV-2 cause severe lower respiratory tract infections with acute respiratory distress syndrome additionally affecting extrapulmonary regions. Severe disease cases can lead to multiorgan dysfunction like diarrhea or renal- and liver-function impairment.[175] Despite its high similarity to SARS-CoV, the SARS-CoV-2 nucleotide sequence shares an even higher nucleotide similarity (96%) with the SARS-like (SL) CoV, RaTG13 (96%), discovered in a cave of Yunnan, China, in 2013.[174, 176] This data suggests its origin from a bat SL-CoV rather than a mutated form of SARS-CoV. SARS-CoV-2 consists, like the other CoVs, of a positive-sense single stranded (ss-(+))-RNA. Coronaviruses have the largest genome of all RNA-viruses known to date, with a length of 30 kb.[177, 178] The (+)-RNA genome encodes at least 29 proteins. Four of them are structural proteins: the spike (S)-, membrane (M)-, envelope (E)- and nucleocapsid (N)- protein.[179, 180] The membrane, envelope and spike proteins are located on the surface of the virus particle, while the nucleocapsid protein binds and protects the genomic RNA inside. Coronavirus virions are spherical particles with the prominent crown-shaped spike proteins sticking out of the particle, as depicted with cryo-electron microscopy (structures are freely available in the protein data bank, PDB).[145, 178, 181, 182] The S-protein of SARS-CoV-2 facilitates, in an analogous manner to SARS-CoV, virus-cell entry through the receptor binding domain (RBD) at the S1 subsite of the spike protein by interacting with the angiotensin-converting-enzyme receptor-2 (ACE2).[177, 183-185] For replication of the viral genome and the production of virus particles, the formation of a functional replicase complex is required. Therefore, the replicase gene of the RNA gets translated and the polyproteins pp1a and pp1ab are processed by autocleavage of the viral main protease (M^{pro}) and the papain-like protease (PL^{pro}) to non-structural proteins.[186-189] The before mentioned structural proteins (S-, M-, E, N-proteins) are responsible for the assembly of the new virus particles. Mature virions are transported to the cell surface and released by exocytosis.[178]
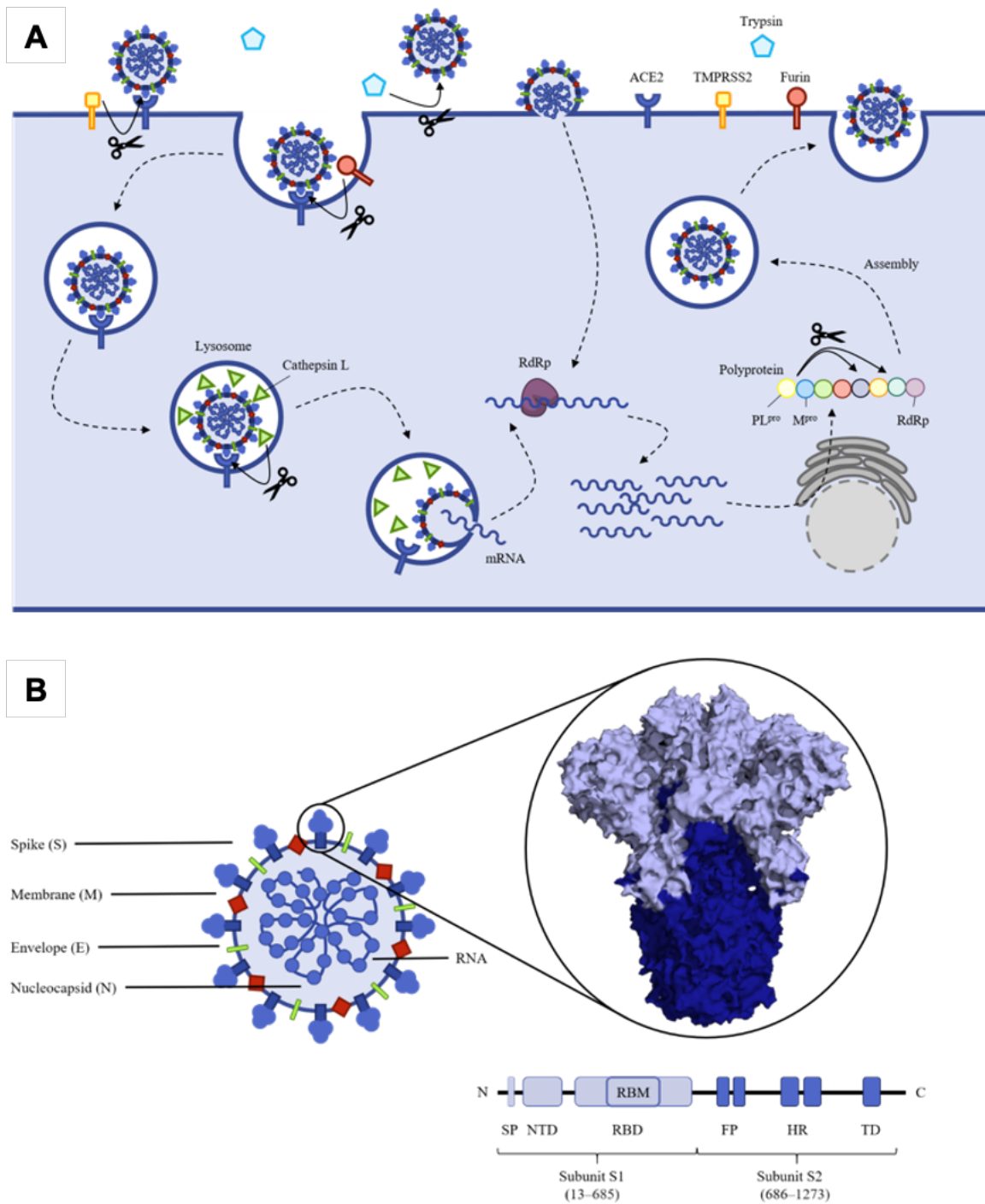
Figure 4.1. A Schematic life cycle of coronavirus and involved host cell proteases known to proteolytically activate coronavirus S-protein. Virus uptake to the cell can occur via endolysosomal pathway after ACE2 binding or by fusion of the virus with the plasma membrane. **B** Schematic structure of the SARS-CoV-2 virion (left) and S protein domains (right). The surface of the crystal structure (PDB- ID: 6ZB4) of the spike protein is colored in light blue and deep blue for subunit S1 and subunit S2, respectively.[190]

Due to the current global situation, the scientific community strives to develop methods and drugs to fight the disease utilizing different strategies. One of them is the inhibition of proteases to prevent viral spread. The differing roles of these proteases in the virus' life cycle allow to further divide them into two classes. The first class includes the viral cysteine proteases M[pro] and PL[pro] which are implicated in the processing of viral proteins.[191, 192] The second class of proteases is involved in the viral cell-entry mechanism to enable the infection of host cells with the viral RNA. This category of proteases comprises the transmembrane serine protease type-II (TMPRSS2), cathepsin L (CatL), furin and trypsin (Figure 4.1).[2, 193-195] This review will present the biochemical influence and the roles of host proteases in the viral cell entry of SARS-CoV-2 and additionally present the current state of the development of inhibitors, addressing the four mentioned cell-entry proteases.

4.2 The coronavirus spike protein

The 600-kDa trimeric spike (S) protein belongs to the class 1 viral fusion proteins and represents the main mediator for viral cell entry and, besides that, is responsible for tissue tropism. It is highly glycosylated with 66 N-linked glycans and located on the surface of the virus.[177, 196, 197] The spike protein can be found among all 7 coronaviruses infecting humans including SARS-CoV, MERS-CoV and SARS-CoV-2 and shows a high similarity to that of other viral glycoproteins like influenza virus hemagglutinin (HA) and the envelope glycoprotein GP120 from human immunodeficiency virus (HIV).[198] Those proteins are collectively referred to as class I membrane fusion proteins and cover the surface of the viruses in a crown-like shape (Figure 4.1).[199] Spike proteins exert crucial functions in receptor recognition, cell-virus fusion and release of the viral RNA inside the cell by binding the host cell receptor ACE2.[200] The spike protein (220 kDa, 1273 amino acids (aa) long) contains a high proportion of α-helical coiled-coil structures and is anchored in the viral membrane through a transmembrane domain (TD) with a short intraviral C-terminal segment and an N-terminal extracellular segment harboring the S1 and S2 subunits.[174, 196] The S1 subunit (aa 14–685) is involved in binding cellular receptors and cell-virus attachment and the S2 subunit (aa 686–1273) facilitates cell-virus membrane fusion.

The globular S1 subunit of the spike protein mediates attachment of the virus to the ACE2 receptor *via* the RBD.[201] The RBD of the S1 subunit is spanned by 222 aa (residues Arg319–Phe541) and binds SARS-CoV-2 S efficiently with an equilibrium dissociation constant (K_D) of 4.7-14.7 nM showing that RBD is a key functional component for ACE2 binding.[184, 196] A previous study revealed high resolution X-ray crystal structures of the RBD bound with ACE2 (PDB-ID: 6M0J) and showed that the ACE2 binding is nearly identical to that of SARS-CoV S

RBD.[184] Further, it has been elucidated that the RBD has five-stranded antiparallel β sheets (β1, β2, β3, β4, β7) and between β4 and β7 is an extended insertion forming the receptor binding motif (RBM) which harbors most of the contacting residues. The S2 subunit has several segments which play a key role in cell-virus membrane fusion and entry of the virus into the cell (Figure 4.1). Among them are fusion peptide (FP), heptad region 1 and 2 (HR1 and HR2) and transmembrane domain (TD).

Hoffmann *et al.* showed that SARS-CoV-2 S utilizes the receptor ACE2 for internalization.[2] ACE2 is abundantly detected in alveolar type II cells in the lung and in epithelial cells in the small intestines.[202] Other organs expressing cells for ACE2 include the heart and kidney. Binding of SARS-CoV-2 S with ACE2 over the RBD in the S1 subunit involves a total of 17 residues of S protein in contact with 20 residues of ACE2.[184] Outside of the RBD, an essential salt bridge is formed through the amino acid lysine in position 417 (Lys417) in SARS-CoV-2 S with Asp30 in ACE2. The RBD region of the S1 subunit is a less conserved region and mutations might change the nature of the virus due to varying binding interactions of RBD and ACE2 receptor leading to higher or lesser pathogenicity and transmissibility between hosts.[203, 204] The authors reported that mutation of aspartate (Asp) at position 614 to Gly (Asp614Gly) in the RBD gene resulted in a more pathogenic strain of SARS-CoV-2. After binding of the RBD region to the ACE2 receptor, a two-step hydrolysis process between the S1↓S2 and ↓S2' cleavage sites by host proteases induces dissociation of the S1 from the S2 subunit and activation of the spike protein.[205] Previous studies revealed involvement of the serine proteases TMPRSS2, furin and the endosomal cysteine protease CatL. Since these proteases are present in different sites of the cell, their localization also determines the cellular location of viral-cell membrane fusion. After priming, viral fusion is commenced by insertion of FPs into the target cell membrane and triggering of α-helical HR1 and HR2 to form a six-helical coiled-coil bundle (6-HB) which brings the viral envelope and cell membrane into proximity, thereby initializing membrane fusion and release of viral RNA inside the cell.[206, 207]

4.3 TMPRSS2

Virus entry of *coronaviridae* is mediated by the connection of S-proteins to cell membranes via cell receptor binding. To fuse with membranes, resulting in the release of genetic information, fusion proteins need to gain flexibility through several regulating events such as protonation, disulfide reduction or proteolytic cleavage events. Therefore, cell receptors as well as extracellular proteases resemble interesting targets for potential entry inhibitors. Amongst others, the proteolytic activity of human type II transmembrane serine proteinase 2 (TMPRSS2) was identified to be relevant for virus entry of influenza viruses SARS-CoV as well

as for SARS-CoV-2, prompting us to focus part of this review on structure, function and potential inhibitors of TMPRSS2.[2, 208-210]

The gene coding for the human type II transmembrane serine proteinase 2 (TMPRSS2), also known as epitheliasin, was first described by Paoloni-Giacobino *et al.* after they identified it by systematic exon-trapping experiments.[211, 212] It is located at 21q22.3.[212] The TMPRSS2 protease is assigned to the clan of proteases of mixed nucleophile, superfamily A (clan PA) and the family of trypsine like proteases (family S1) (MEROPS database).[41] Since it features membrane anchoring *via* an *N*-terminal proximal transmembrane domain, it is considered as one of the 17 known members of type II transmembrane serine proteases (TTSPs).[213]

The 44 kb spanning TMPRSS gene comprises 14 exons and 13 introns.[214, 215] It is transcribed into a 3.8 kb mRNA, that translates into a 492 aa protein, consisting of a cytoplasmic domain (exons 2 and 3, aa 1-78), a transmembrane domain (exon 4, aa 84-106), a LDL receptor class A domain (LDLRA, exon 5, aa 113-148) that binds calcium ions, a scavenger receptor cysteine-rich domain (SRCR, exons 6-8, aa 149-242), relevant for mediating non-opsonic phagocytosis upon ligand binding and immune response, and the serine protease (exons 9-13, aa 255-492).[212, 215-219] Autoactivated proteolytic cleavage of the full-length 54 kDa zymogen (due to glycosylation, the full-length protein migrates at 70 kDa in SDS-PAGE) between Arg255 and Ile256 leads to the 32 kDa free and catalytically active serine-protease domain.[209, 219, 220] Whereas the TMPRSS2 protease domain was initially reported to be secreted into extracellular space, other studies indicated that the protease domain mainly remains anchored to the pro-domain via a disulfide bond.[219-221] Interestingly, TMPRSS2 was also detected in extracellular vesicles, indicating functions beyond its expression sites.[222] Six additional Cys residues are predicted to form three intra-protease disulfide bonds.[212]

Due to a high sequence identity of type II serine proteases, similar folding of related enzymes can be expected.[220] Crystal structures of highly related chymotrypsin-like folded proteases (S1) e.g. enteropeptidase and matriptase/MT-SP1 revealed that these enzymes feature two six-stranded *β*-barrels connected by three *trans*-domain segments.[223, 224] The aa of the catalytic triad His296, Asp345 and Ser441 (His57, Asp102, Ser195 in standard chymotrypsin numbering system) are located in the aforementioned junction.[214] The active-site cleft runs perpendicular to this junction.[224]

According to the nomenclature for interactions with proteases and their respective substrates by Schechter & Berger, based on knowledge of related TTSP, the bottom of the S1 pocket of TMPRSS2 is probably formed by Asp435, Gly462 and Gly472, determining the selectivity for the basic residues Lys or Arg.[45, 212, 214] The substrates' scissile bond is properly oriented by interactions of Ser460, Trp461 and Gly462 with the P1 and P3 residues of the substrate.[214] Since TMPRRS2 harbors a Ser436, Arg and Lys are likely to be similarly preferred in P1

position. Ser436 is expected to form hydrogen bonds to the side chain of Lys, but not to Arg. Arg, however, can form a direct salt bridge with Asp435, whereas interactions of a P1-Lys residue with Asp435 are likely to be mediated by a conserved water molecule in the S1 pocket.[214, 225] Altogether, there is yet little known about the substrate specificity of TMPRSS2. For the P2 position, the small aa glycine and proline are well tolerated.[226] The P3 position seems to be less specific. Basic residues are favored, hydrophobic aa, however, are also tolerated, even in D-configuration.[225]

The TMPRSS2 protein is located in plasma membranes of epithelial cells and expressed in several human tissues that comprise large populations of epithelial cells, predominantly in prostate, colon and pancreas, but also on kidney and lung, with increased transcription levels in prostate gland.[208, 214, 219, 227, 228] The C-terminal TMPRSS2 protease domain faces extracellular or luminal spaces.[214]

Although located at the small human chromosome 21, there seems to be no evidence, that TMPRSS2 plays a role in development of the Down syndrome. Indeed, the biological functions of mammalian TMPRSS2 still remain unknown.[229] In this context, monitored TMPRSS2 knockout mice developed normally and survived to adulthood with no obvious phenotype in terms of organ histology, behavior or reproducibility.[229] Based on these findings, it was concluded that there is a functional redundancy involving one or more of the TTSPs family members or other serine proteases, which suggests, that targeting TRMPSS2 selectively with inhibitors could be well tolerated. It was proposed that TMPRSS2 is possibly involved in other specialized, but non-vital functions, such as stress coping or disease response.[229] The TMPRSS2 protease is expressed in human airways and modulates the amiloride-sensitive epithelial sodium channel (ENaC).[228] In contrast to prostasin, TMPRSS3, TMPRSS4 and PRSS14, the presence of TMPRSS2 decreases sodium channel activity and protein levels probably due to degradation.[228] Therefore, TMPRRS2 may influence airway surface liquid volume and mucociliary clearance, although the physiological relevance of these findings is strongly questioned.[228, 230]

However, the extracellular domains, expressed in the TMPRSS2 gene, namely the LDLRA and the SRCR domains together with an intracellular N-terminus suggest a role in intracellular signal transduction or a receptor function, binding low-density lipoproteins as well as calcium ions followed by non-opsonic internalization.[120, 219] Additionally, in other proteins, these calcium-binding domains can bind macromolecules (poly-anionic molecules such as modified lipoproteins, bacterial cell surface lipids and sulfated polysaccharides) and therefore be involved in immune response.[216-218, 231-234] A more recent study connected TMPRSS2 to the development of cancer pain.[235]

Since TMPRSS2 has a high expression in prostate and is found to be inducible by androgenic growth hormones, it probably is involved in gland development and tissue-specific functions.[214, 229] Pathophysiologically, there are increased TMPRSS2 levels in prostate and colon cancer.[219, 227, 236] In a patient with an aggressive form of prostate cancer, a mutated form of TMPRSS2 was found, leading to a nonsense and probably inactive protein.[236] However, studies designed to determine a correlation between TMPRSS2 germ line polymorphisms and the development of prostate cancer could not identify any associations.[237] Furthermore, TMPRSS2 regulates signaling through protease-activated receptor 2 (PAR2) and therefore is connected to tumor-cell migration and metastasis *via* increased production of the matrix metalloproteases (MMP) MMP-2 and MMP-9.[238]

Given its expression in airway epithelia cells in nose, trachea, and alveoli, TMPRSS2 has the potential to be involved in viral infections of respiratory organs.[214, 228] For long, trypsin has been commonly utilized in *in vitro* studies to proteolytically activate influenza A virus, thus enabling infection of cell cultures.[239, 240] Later on, *in vitro* assays exposed TMPRSS2 as well as human airway trypsin-like protease (HAT) to mediate influenza A virus entry by the essential step of cleavage of the precursor hemagglutinin protein (HA0) at a monobasic site into the two subunits HA1 and HA2, linked only via disulfide bonds. This was shown for the three human pathogenic HA subtypes H1, H2 and H3, all of them featuring a monobasic (Arg) cleavage site, suitable as a TMPRSS2 substrate.[208] This cleavage was found to be host-cell associated and protease inhibitor-sensitive, thus further supporting the role of membrane-anchored proteases in virus entry steps.[221] Another human pathogenic virus, whose propagation is mediated by TMPRSS2, is the human metapneumovirus, where the fusion protein (F) is processed at a site containing two Arg residues, separated by two uncharged aa.[241, 242] Other studies found TMPRSS2 to also mediate virus entry of MERS-CoV and porcine epidemic diarrhea virus (PEDV)[243, 244]

In the case of coronaviruses (CoVs), including severe acute respiratory syndrome (SARS), S-proteins, localized in the corona, mediate the essential step of cell entry by connecting viruses to host-cell plasma membrane receptors and catalyze subsequent membrane fusions.[209] For that, conformational flexibility of SARS-CoV S, provided by proteolytic cleavage at an individual cleavage site after Arg667[246] (S1↓S2 cleavage site) and Arg707[247] (↓S2' cleavage site) is necessary.[209, 241, 245-247] The highly conserved character of these Arg residues in all *coronaviridae* indicated a similar mechanism of activation.[246] These Arg residues resemble suitable cleavage sites for TMPRSS2.[248] The findings that *cis*-cleavage occurred between TMPRSS2 positive and S-protein positive cells provided the first evidence that TMPRSS2 cleaves and activates SARS-CoV S-protein. The shed cleavage fragments of S functioned as decoys for anti-S antibodies, thus probably affecting humoral immune response.[249] After the COVID-19 outbreak, TMPRSS2 was found to prime S-proteins of SARS-

CoV-2 as well.[2, 210] Investigations of the sequential and conformational differences in S1↓S2 cleavage sites of SARS-CoV and SARS-CoV-2 revealed that the site of SARS-CoV-2 consists of two additional Arg residues, preceding the former cleavage site. Based on findings, that the avian influenza became more virulent upon duplication of the cleavage site of HA, they suggested the multiple basic aa in S1↓S2 to account for the increased virulence of SARS-CoV-2.[250] The ↓S2' cleavage site of SARS-CoV-2 is similar to that of SARS-CoV, in respect of basic aa.[210]

The carboxypeptidase ACE2, also expressed in alveolar cells, constitutes the primary receptor for SARS-CoV as well as SARS-CoV-2 and binds trimeric S-protein at an area distinct from its catalytic site with high affinity (SARS-CoV K_D = 1.7 nM; SARS-CoV-2 K_D = 15.2 nM), underlining, that other proteases are likely to facilitate S-protein priming.[2, 245, 251-253] In an immunoprecipation assay, TMPRSS2 was captured by ACE2, thus suggesting heterodimerization in alveolar epithelia cell surfaces.[209] Coexpressed by type II pneumocytes, both proteins enhance cell entry of SARS-CoV.[209, 249] For influenza virus, however, besides ACE2, 2,6-linked sialic acids represent the main receptor determinants in the respiratory tract for triggering endosomal internalization and were found to be co-expressed with TMPRSS2 and HAT.[254, 255]

A recent study highlighted the relevance of TMPRSS2 in virus spread and infection's severity. T helper 1 cell (Th1)-prone C57BL/6 TMPRSS2 knockout mice showed significantly reduced body-weight loss, less virus spread, and less severe immunopathology compared to their SARS-CoV or MERS-CoV receptor expressing counterparts. Interestingly, inflammatory chemokine and/or cytokine responses upon nasal application of polyinosinic:polycytidylic acid (poly I:C – a toll-like receptor-3 agonist) were weakened in the knockout mice, further supporting the immunorelevance of TMPRSS2.[229, 256] Keeping the androgen inducible TMPRSS2 expression in mind, it seems plausible that patients receiving androgen depletion therapy as prostate cancer treatment, exhibited a 4-fold lower risk of SARS-CoV-2 infection compared with patients who did not undergo this treatment.[257]

For research purposes, the promiscuous irreversible serine protease inhibitor **13** (Figure 4.2), previously shown to inhibit trypsin-like serine proteases,[262] can be utilized to inhibit TMPRSS2 effectively at 10 µM.[238, 258] Another promiscuous serine protease inhibitor widely used for research purposes is aprotinin, that was shown to inhibit TMPRSS2-mediated virus replication in a multicycle replication assay of influenza virus moderately at 50 µM, whereas there was strong inhibition in presence of the commercial irreversible trypsin inhibitor pefabloc® SC (4-(2-aminoethyl)benzolsulfonylfluoride hydro-chloride, **14**) and ovomucoid.[259]

Although these inhibitors possess only limited potential for drug development due to low specificity and hence, pronounced toxicities or low bioavailability of the poly- or oligopeptides

aprotinin, leupeptin (**15**) and ovomucoid, an aprotinin containing aerosol formulation has been approved in Russia for local respiratory application in mild to moderate influenza infections.[260] Other polypeptidic agents with TMPRSS2 inhibitory activity that could be administered by aerosols, are plasminogen activator inhibitor 1 (PAI-1), and hepatocyte growth factor activator inhibitor 2 (HAI-2), according to previous reports of inhibitory activity against several related serine proteases.[261-263]

Meyer *et al.* identified the first synthetic inhibitors for TMPRSS2 by screening substrate analogues, previously reported as inhibitors for TTSP (HAT, matriptase and thrombin), all harboring a 4-amidinobenzylamide in the P1 pocket, featuring a competitive reversible binding mode.[264-266] Besides some structure-activity relationships, they identified **17** and some of its close analogues to inhibit TMPRSS2 with a K_i of 18-20 nM.

In a second approach*,* synthetic inhibitors for matriptase and thrombin were screened and revealed. the matriptase inhibitor **18**, previously published by Steinmetzer *et al.* to inhibit TMPRSS2 with an K_i of 8 nM.[265, 267] A major drawback of this highly potent compound is its tribasic character, diminishing its membrane permeability. To overcome this issue, they assayed another set of published matriptase inhibitors, harboring only one or two basic residues.[268, 269] In this set, the dibasic compound **19** was most potent (K_i = 0.9 nM). In line with matriptase, monobasic inhibitors possess only reduced potency. In a recent study, the antiviral activity of **18** and a related monobasic inhibitor against SARS-CoV-2 in Calu-3 cells was demonstrated.[210]

Bromhexine hydrochloride (**16**) is an approved drug without significant side effects and used in expectorant cough syrups. **16** prevented pancreas cancer metastasis in mice by inhibiting TMPRSS2 with an IC$_{50}$ of 0.75 µM.[270] Conversely, an *in-vitro* assay with isolated TMPRSS2 could not observe inhibition of TMPRSS2, leaving its mechanism of action unclear.[271] Nevertheless, bromhexine hydrochloride, has already entered clinical trials for the treatment of COVID-19.

In an approach to identify a selective TMPRSS2 inhibitor, a single-stranded DNA-like antisense fragment called peptide-conjugated phosphorodiamidate morpholino oligomer (PPMO) was designed. This fragment can readily enter cells and alter splicing patterns of mRNA, leading to incomplete and, hence, inactive forms of TMPRSS2. Consequently, HA cleavage could be reduced 100-1000-fold in Calu-3 lung cells.

The probably best characterized TMPRSS2 inhibitor was identified in an attempt to overcome the issue, that SARS-CoV exploits at least two distinct pathways to enter host cells. A TMPRSS2-dependent entry, directly through the plasma membrane and an endosomal pathway, mediated by CatL.[209, 241, 272, 273] In this respect, the combined administration of inhibitors of TMPRSS2 (**17**) and CatL were able to effectively suppress pseudotyped viruses,

bearing either SARS-CoV or SARS-CoV-2 S-protein, entry in several cells (HeLa, BHK-21 and resected cells from human airway epithelia).[2, 3] Of all tested serine protease inhibitors, only the approved drug camostat (**20**) mesilate, showed inhibition of TMPRSS2 mediated entry. **20** is successfully used for the treatment of chronic forms of pancreatitis, cancer and liver fibrosis with only minimal side effects.[274-276] In TMPRSS2 and CatL expressing cells up to 80% inhibition of virus entry could be achieved by treatment with **20**, indicating 20-30% endosomal entry. In combination with a cathepsin inhibitor, up to 95% of virus entry could be avoided.[3] These inhibition values are in line with more recent findings that human coronaviruses prefer TMPRSS2-mediated cell entry over endosomal uptake.[277] In various other studies, **20** was found to reduce influenza virus titers in infected mice and influenza virus replication and HA1 titers in the supernatant of human tracheal surface epithelial (HTE) cells without affecting RNA titers.[260, 278-280] It also protected Calu-3 lung cells from SARS-CoV-2 infections and even improved survival rates of SARS-CoV-infected mice.[278] Compound **20** is currently undergoing phase II clinical trials.[281] Additionally, in Vero-TMPRSS2 cells, 10 µM of **20** reduced MERS-CoV entry and viral RNA titers in cells 15-fold. In Calu-3 cells, 100 µM of **20** reduced viral RNA titers 270-fold.[243] Shrimp *et al.* employed an assay on isolated TMPRSS2 and found **20** to inhibit TMPRSS2 activity with an IC_{50} of 6.2 nM.[271] After administration, **20** undergoes rapid biotransformation ($t_{1/2}$ <1 min). Carboxyesterase hydrolyses the side-chain ester to yield its active metabolite 4-(4-guanidinobenzoyloxy)phenylacetic acid.[282]

The related drug nafamostat (**21**) was found to be even more effective in preventing MERS-CoV entry, as it reduced viral entry 100-fold at 1 nM.[283] Besides the same application as **20** in cancer and pancreatitis treatment,[289] **21** mesilate is used to prevent blood coagulation during hemodialysis by inhibiting the serine protease thrombin, and thus proteolysis of fibrinogen to fibrin.[284, 285] Later studies provided the IC_{50} of **21** to be 0.27 nM and assessed its antiviral effects against SARS-CoV-2 infections of Vero E6 cells.[271] The EC_{50} was found to be 22.50 µM with a relatively low SI of 4.44, narrowing the therapeutic window of **21**.[286] Nafamostat is currently in clinical trials for SARS-CoV-2 treatment as well. Another analogue of **20**, gebexate (**22**), was found to be a less potent inhibitor of TMPRSS2 (IC_{50} = 130 nM).[44, 271]

Whereas currently, no crystal structures of TMPRSS2 are available in the PDB, a complex of **20** with prostasin is reported (PDB-ID: 3FVF).[292] Prostasin and TMPRSS2 share 67% sequence identity within the active site (defined as residues 6 Å around **20**, sequences aligned with MOE2018). In the reported prostatin-**20** complex, **20** is covalently bound to the catalytical Ser195 (corresponding to Ser441 in TMPRSS2). The oxoanion hole is formed by the backbone N-H donor functionalities of Gly193, Asp194 and the active-site Ser195 residue (in TMPRSS2 Gly439, Asp440 and Ser441, respectively). The phenylguanidine moiety of **20** is deeply buried in the S1 pocket, forming ionic interactions with Asp189 (Asp435 in TMPRSS2) and multiple

H-bonds with the acceptors in the backbone of Asp217, Arg224, Ala-90 (corresponding to Ser463, Arg470 and Ser436) and a water molecule (Figure 4.3).
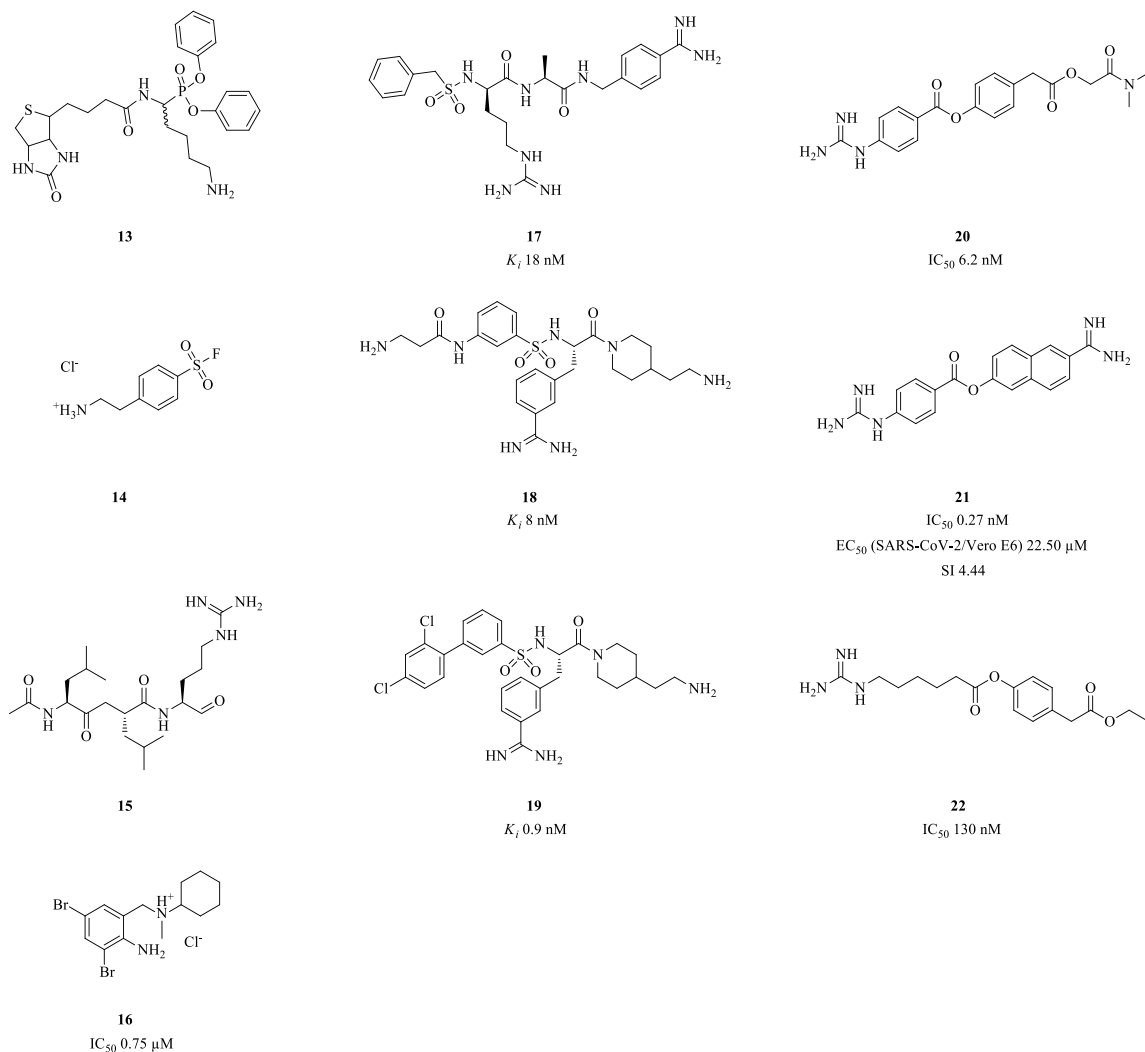
Figure 4.2. Inhibitors for TMPRSS2. If available, inhibitory constants are depicted beneath the respective structure.
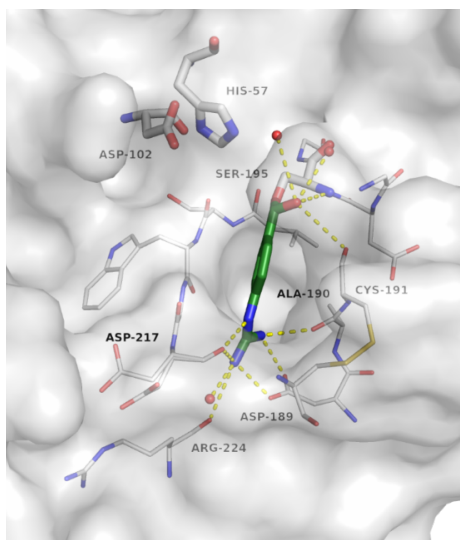
Figure 4.3. Camostat (20, green carbon atoms) –prostasin (grey carbon atoms and surface) complex (PDB-ID: 3FVF). For a clear view, only residues forming polar interactions with the ligand or being part of the catalytic triad Ser195-His57-Asp102 are labelled.

5. Peptidomimetic inhibitors of TMPRSS2 to block SARS-CoV-2 infection

Aim:

Chapter A showed the involvement of the transmembrane serine protease hepsin in prostate cancer and presented strategies to control this disease. The closely related transmembrane serine protease TMPRSS2 shares a high similarity with hepsin and was recently shown to proteolytically activates multiple viruses, including SARS-CoV-2, which initiates viral uptake and infection of host cells. In analogy to study **[2]**, the aim of this study **[4]** intends to develop potent peptidomimetic inhibitors for the host protease TMPRSS2 to suppress SARS-CoV-2 infection and spreading.

[4] <u>**Philip Maximilian Knaff**</u>*, Christian Kersten*, Lukas Wettstein*, Patrick Müller, Tatjana Weil, Carina Conzelmann, Janis Müller, Maximillian Brückner, Stefan Pöhlmann, Markus Hoffmann Tanja Schirmeister, *Katharina Landfester,* Jan Münch, Volker Mailänder. **Development of Potent Peptidomimetic Inhibitors for the Serine Protease TMPRSS2**. Nature Communications. **2021** (Submitted).

Contribution:

I carried out the synthesis and characterization of the peptidomimetic inhibitors and performed the enzymatic activity measurements and stability studies. I further supported the molecular docking studies which were mainly performed by Dr. Christian Kersten. Patrick Müller supported with synthesis of multiple precursor compounds. Lukas Wettstein, Carina Conzelmann, Tatjana Weil and Janis Müller tested the inhibitors against viruses. The project was supervised by Prof. Dr. Volker Mailänder, Prof. Dr. Katharina Landfester and Prof. Dr. Jan Münch.

5.1 Abstract

The transmembrane serine protease 2 (TMPRSS2) primes the SARS-CoV-2 spike (S) protein for host cell entry and represents a promising target for COVID-19 therapy. Here, we describe the *in silico* development and *in vitro* characterization of peptidomimetic TMPRSS2 inhibitors. Molecular docking studies identified peptidomimetic binders of the TMPRSS2 catalytic site, which were synthesized and coupled to an electrophilic serine trap. The compounds inhibited TMPRSS2 with good off-target selectivity on selected coagulation proteases, and lead candidates were shown to be stable in blood serum and plasma for at least 10 days. Finally, we show that selected peptidomimetics inhibit SARS-CoV-2 S-driven pseudovirus entry and authentic SARS-CoV-2 infection with comparable efficacy as camostat mesylate. Notably, the peptidomimetic TMPRSS2 inhibitors also prevented infection of SARS-CoV-2 variants of concern with spike mutations D614G, B.1.1.7 and B.1.153. In sum, our study reports antivirally active and stable TMPRSS2 inhibitors with prospects for further preclinical and clinical development as antiviral agents against SARS-CoV-2 and other TMPRSS2-dependent viruses.

5.2 Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the causative agent of coronavirus disease 2019 (COVID-19). By the end of April 2021, the WHO reported more than 147 million confirmed SARS-CoV-2 infections worldwide, resulting in more than 3.1 million deaths since its first occurrence in late 2019.[287] COVID-19 is characterized by a mild to moderate respiratory illness and infected individuals usually recover without requiring special treatment. Older people, and those with underlying medical problems like cardiovascular disease, diabetes, chronic respiratory disorders, and cancer are more likely to develop serious illness, characterized by respiratory failure, shock, and multiorgan dysfunction.[288-291] SARS-CoV-2 is primarily transmitted through aerosols and droplets of saliva. Inhaled virus may then establish infection in epithelial cells of the upper respiratory tract from which it may further disseminate to lower airway epithelial and alveolar cells, and other organs such as the gastrointestinal tract or heart.[292-295] Measures to prevent or mitigate SARS-CoV-2 spread include lock-down strategies, social distancing, quarantining, use of face masks and hygiene concepts. The recent implementation of effective SARS-CoV-2 vaccination programs are the best defense against COVID-19 and have raised hopes that the pandemic is nearing an end. However, the emergence of viral variants of concern that escape pre-existing immunity and are associated with increased transmissibility and higher case fatality rates, as well as the slow vaccine rollout in most countries, may compromise efforts to control the pandemic.[296-300] Currently, no specific drug or therapeutic vaccine are available and approved for the treatment of SARS-CoV-2 infected patients. Thus, it remains imperative to develop potent therapeutic interventions for COVID-19 therapy and numerous compounds, mostly repurposed drugs are currently evaluated in clinical trials worldwide.

SARS-CoV-2 is an enveloped, positive-sense single stranded RNA virus. Infection is mediated by the viral spike (S) protein, a homotrimeric transmembrane glycoprotein. The spike glycoprotein is composed of S1 and S2 subdomains. The S1 subdomain encodes for the receptor binding domain and is responsible for binding to angiotensin-converting enzyme 2 (ACE2), the primary receptor for SARS-CoV-2.[2, 174, 301] Subsequently, the transmembrane protease serine subtype 2 (TMPRSS2) primes the S protein which triggers conformational changes in S2 leading to fusion of the viral with the cellular membrane and delivery of the nucleocapsid into the cytoplasm.[2] Of note, TMPRSS2 not only cleaves and primes SARS-CoV-2 spike, but also surface proteins of several other viruses, including the hemagglutinin (HA) of certain influenza A virus strains, the fusion protein (F) of human metapneumovirus and the spike proteins of human coronavirus 229E (HCoV-229E), Middle East respiratory syndrome coronavirus (MERS-CoV), and SARS-CoV.[2, 241, 243, 272, 302, 303] TMPRSS2 priming is essential for triggering fusion of these virions with target cell membranes and disruption of TMPRSS2

expression was found to markedly reduce influenza A virus, SARS-CoV and MERS-CoV infection and pathogenesis in mice. Importantly, TMPRSS2 knockout mice are phenotypically similar to wild-type animals, suggesting that the protease is not essential, rendering TMPRSS2 a very promising target for broad-spectrum antiviral agents with presumably low side effects.[304, 256, 229]

TMPRSS2 belongs to the family of type II transmembrane serine proteases (TTSP) which control a variety of physiological processes, including epithelial differentiation, homeostasis, iron metabolism, hearing and blood pressure regulation.[305] The family of TTSP comprises a total of 17 members in four subfamilies (matriptase, corin, hepsin/TMPRSS and HAT/DSC subfamily) with a similar domain structure including an intracellular N-terminus, a transmembrane domain which anchors the protease in the cell membrane, and an extracellular C-terminus harboring a serine protease domain.[306] TMPRSS2 belongs to the hepsin/TMPRSS subfamily with a total of seven serine proteases including TMPRSS2 - 5, MSPL, hepsin and enteropeptidase. Development of TMPRSS2 inhibitors is hampered by the fact that no crystal structure is available. Until now, only few substrate analogue inhibitors of TMPRSS2 have been described.[226, 264] Approved drugs that are known to inhibit TMPRSS2 may be suitable for off-label use and repurposing in COVID-19 prevention and therapy. Noteworthy are camostat mesylate (CM) and nafamostat, which are used for treatment of chronic pancreatitis in Japan, as well as the endogenous protease inhibitor alpha-1 antitrypsin (α_1AT), and the mucolytic cough suppressant bromhexine.[2, 307-309] However, the selectivity of some of these inhibitors is low, little is known about structure-activity relationships and the therapeutic effect of CM in COVID-19 has not yet been reported.

Herein, we describe our efforts to develop novel peptidomimetic inhibitors for TMPRSS2 and SARS-CoV-2 infection. Using computational modeling and docking of combinatorial peptide libraries we identified high scoring binders which were synthesized and characterized with respect to protease inhibition, selectivity and antiviral activity. This approach allowed to identify lead candidates that efficiently inhibit TMPRSS2 enzyme activity and block SARS-CoV-2 spike driven entry into target cells. Furthermore, we demonstrate that these peptidomimetic TMPRSS2 inhibitors prevent authentic SARS-CoV-2 infection, including the variants of concern B.1.1.7 and B.1.351. The tested peptidomimetics are stable in human plasma and serum for at least 10 days, suggesting that these novel TMPRSS2 inhibitors are promising leads for further development as antiviral drugs in COVID-19 therapy.

5.3 Material and methods

Molecular modelling. For TMPRSS2, a homology model was built using the Swiss-Model web server.[310] The template structure was selected based on the serine protease hepsin in complex with *N*-acetyl-6-ammonio-L-norleucyl-L-glutaminyl-*N*-[(1S)-4-w-1-(chloroacetyl)butyl]-L-leucinamide (PDB-ID: 1Z8G)[46] with a sequence identity to TMPRSS2 of 42.49% and a 1.5 Å resolution. For subsequent docking studies, both the homology model and the crystal structure of matriptase in complex with *N*-(3-phenylpropanoyl)-3-(1,3-thiazol-4-yl)-L-alanyl-*N*-[(1S,2S)-1-(1,3-benzothiazol-2-yl)-5-carbamimidamido-1-hydroxypentan-2-yl]-L-valinamide (PDB-ID: 6N4T)[311] as a surrogate model were used. The focused serine-protease inhibitor library was derived from the ZINC15 database.[312] Commercially available compounds with reported affinity for factor Xa, hepsin, kalikrenin, plasminogen, thrombin, matriptase-1, matriptase-2, trypsin and urokinase were pooled and after duplicate removal a total of 315 trypsin-like serine protease inhibitors were obtained. Sequences of tripeptides for docking studies on TMPRSS2 homology/surrogate were generated using CycloPs 8.5.9 and included proteinogenic and non-proteinogenic amino acids (aa).[121] The generated SMILES were modified to carry a N-terminal acetyl cap (ace) and a C-terminal aldehyde warhead. Prior to docking, all molecules were protonated and energetically minimized using MOE2019.[148] Hereby, the MMFF94x[313] forcefield was used for small molecules and AMBER14:EHT[123] for peptidic molecules. For molecular docking with LeadIT-2.3.2, the binding site was defined to include all residues within 6 Å around the reference ligand of the hepsin homology model (PDB-ID: (1Z8G)).[144] For matriptase, all residues within 6.5 Å around the crystallographic reference ligand (PDB-ID: 6N4T) and water molecules forming at least three interactions with the target and ligand were included. Structures were protonated with the Protoss module[147] within LeadIT-2.3.2. All dockings were performed using standard settings and the enthalpy-entropy hybrid approach. The docking strategy was validated for matriptase surrogate model by redocking of the ligand (PDB-ID: 6N4T), and for the TMPRSS2 homology model and matriptase surrogate model by docking of the substrate ace-D-Arg-Pro/Gly-Arg-nme and by a binder *vs* non-binder discrimination using 56 published TMPRSS2 inhibitors and 314 decoys generated for four inhibitors present at ZINC using the database of useful decoys enhanced (DUD-E; Figure 5.1).[226, 264,314] Results were analyzed by FlexX score and visual pose inspection to select molecules for purchase and synthesis. Figures were made with PyMOL.[150]
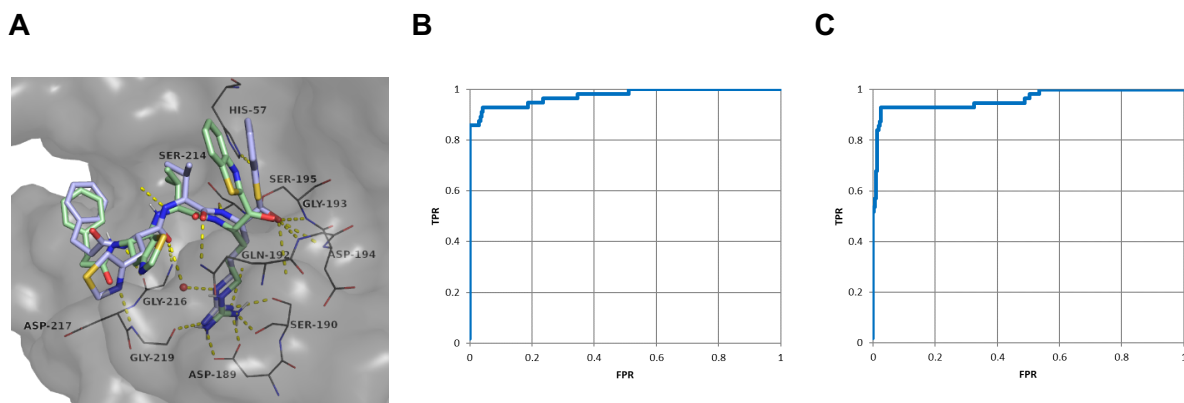
Figure 5.1. A Re-docking result for reference ligand-matriptase complex (PDB-ID: 6N4T). Docking pose with light green carbon atoms, crystallographic binding mode with light blue carbon atoms, FlexX-score = 67.2 kJ/mol, RMSD = 1.8 Å. **B** ROC curve for binder vs non-binder discrimination of molecular docking using a TMPRSS2 homology model (56 binder and 314 non-binder successfully docked, ROC AUC = 0.98). **C** ROC curve for binder vs non-binder discrimination of molecular docking using matriptase-1 surrogate model (PDB-ID: 6N4T, 56 binder and 314 non-binder successfully docked, ROC AUC = 0.96). FPR: false positive rate, TPR: true positive rate.

Chlorination of tritylhydroxide resin (Trt-OH).[151] The chlorination of trtiylhydroxide resin was performed based on a modified method described elsewhere.[151] In short, a 250 mL round bottom flask was rinsed with dry dichloromethane (DCM) and 25 g of Trt-OH (0.8 mmol/g, mesh 100-200, company: Iris Biotech) was added. The resin was suspended in a mixture of 50% DCM and 50% toluene, just enough to double the resin volume and 10 mL acetyl chloride was added. The glass vial was sealed and agitated for 24 h. The next day, the resin was dried and thoroughly washed with DCM (4 × 5 mL). Chlorinated resin (Trt-Cl) was stored in a freezer at -20 °C.

Addition of first amino acid to Trt-Cl.[151] 15 mL of dry DCM was added to 0.8 g of Trt-Cl resin (0.8 mmol/g) and shaken for 10 minutes. Next, 2.5 equiv. (2.5 mmol) of amino acid and 5 equiv. (5 mmol) of activator base was added and the mixture was shaken overnight at room temperature. The next day, the resin was filtered and washed with *N*,*N*-dimethylformamide (DMF) (2 × 10 mL) and DCM (2 × 10 mL). Prior to use for solid phase peptide synthesis the resin was swollen with 10 mL DMF.

Solid phase peptide synthesis. Peptides were synthesized as C-terminal amides on a loaded 2-chlorotrityl chloride resin using 9-fluorenylmethoxycarbonyl (Fmoc) strategy. All Fmoc-protected amino acids were purchased from Novabiochem, as well as benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP), *N*,*N*-diisopropylethylamine

(DIEA), DMF, DCM and were used as received. All Fmoc-protected amino acids were dissolved in DMF at a concentration of 0.2 M. The coupling was done with the standard procedure of solid phase peptide synthesis.[152] In short, Fmoc deprotection was performed by treating the peptidyl resin with 20% piperidine in DMF for 10 min at 70 °C. After the reaction, the resin was washed with DMF and DCM, each (2 x 10 mL), then filtrated. The coupling was done by addition of 7.5 mL Fmoc-protected amino acid (3 equiv.), 3 mL of 1 M PyBOP (3 equiv.) and 1.5 mL of 0.5 M DIEA (3 equiv.). The reaction mixture was shaken at 75 °C for 30 minutes. Afterwards, the resin was washed as described above and used for the next step.

N-Cap modification and cleavage from the resin. The N-terminus of the peptides with protected side chains were acetyl capped using 20 mL of 0.5 M Ac_2O/DMF and 1 M DIEA/DMF. The reaction was shaken for 2 h at room temperature. The resin was filtered and washed with DMF and DCM, each (2 × 10 mL). The cleavage of the peptide with protected side chains was accomplished using 20 mL of 20% hexafluoroisopropanol (HFIP) in DCM. The mixture was put on an orbital shaker for 3 h at room temperature and then filtered. The solvent was taken off *in vacuo* and the peptide was precipitated using 50 mL of diethylether.

Synthesis of Boc-Arg(Mtr) weinreb amide (S1) *((S)-tert*-Butyl-(1-(methoxy(methyl)amino)-5-(3-((4-methoxy-2,3,6-trimethylphenyl)sulfonyl)guanidino)-1-oxopentan-2-yl)carbamate).[153]
To a solution of *N*-(*tert*-butoxycarbonyl)-*N*-((4-methoxy-2,3,6-trimethylphenyl)sulfonyl)-L-arginine (1.56 g, 3.21 mmol, 1.0 equiv.) in 34 mL THF, *N,O*-dimethylhydroxyl-amine hydrochloride (0.64 g, 6.42 mmol, 2.0 equiv) and 1-hydroxybenzotriazole hydrate (0.54 g, 3.53 mmol, 1.1 equiv.) were dissolved at room temperature. DIEA (1.63 mL, 9.63 mmol, 3 equiv.) and EDCI (0.65 g, 3.37 mmol, 1.05 equiv.) were added and the solution was stirred for 4.5 h. The mixture was concentrated *in vacuo* and the reaction mixture extracted with ethyl acetate (150 mL), washed with 5% aqueous acetic acid (75 mL), sat. aqueous $NaHCO_3$ (75 mL), water (75 mL), and brine (75 mL). The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo* to yield the Weinreb amide **S1** (0.86 g, 1.6 mmol, yield: 69%) as a white powder (Figure A4). Purity (HPLC, 220 nm) > 80%. ^{1}H NMR (300 MHz, CD3OD): δ=1.43 (s, 9 H) 1.49–1.60 (m, 4 H) 1.87 (s, 1 H) 2.13 (s, 3 H) 2.61 (s, 3 H) 2.67 (s, 3 H) 3.17 (m, 3 H) 3.74 (s, 3 H) 3.90 (s, 3 H) 6.67 ppm (s, 1 H). MS (ESI): *m/z*: calcd for $C_{23}H_{39}N_5O_7S_2$ [M+H]+ 530.3, [2M+H]+ 1059.6, found [M+H]+ 530.2, [2M+H]+ 1059.4

Synthesis of Boc-Arg(Mtr) ketobenzothiazole (S2) ((*S*)-*tert*-Butyl-(1-(benzo[*d*]thiazol-2-yl)-5-(3-((4-methoxy-2,3,6-trimethylphenyl)sulfonyl)guanidino)-1-oxopentan-2-yl)carbamate).[142]
To a solution of benzothiazole (0.47 g, 3.53 mmol, 1.1 equiv.), *n*-BuLi (1.6 M, 4 mL, 6.42 mmol, 2 equiv.) was added dropwise to THF (50 mL). After the mixture was stirred for an additional

30 minutes, Boc-Arg(Mtr) Weinreb amide (1.94 g, 3.21 mmol, 1 equiv.) was dissolved in THF (15 mL) and added slowly over 50 minutes. The mixture was stirred at -78.8 °C for 3 h. The reaction was quenched with sat. aqueous NH$_4$Cl (30 mL) and the aqueous layer was extracted with EtOAc (40 mL). The organic phase was collected, dried with Na$_2$SO$_4$, and then concentrated. The resulting residue was purified by semi-preparative RP-HPLC to yield the compound **S2** (0.32 g, 0.5 mmol, yield: 35%) as a yellow powder (Figure A4). Purity (HPLC, 220 nm) > 95%. ^{1}H NMR (300 MHz, CD3OD): δ=1.43 (s, 9 H) 1.66 (m, 4 H) 2.05 (s, 3 H) 2.55 (s, 3 H) 2.70 (s, 3 H) 3.23 (m, 2 H) 3.81 (s, 3 H) 5.32 (m, 1 H) 6.55 (s, 1 H) 7.69 (m, 2 H) 8.27 ppm (m, 2 H). MS (ESI): *m/z*: calcd for C$_{28}$H$_{37}$N$_5$O$_6$S$_2$ [M+H]$^+$ 604.2, [2M+H]$^+$ 1207.4, found [M+H]$^+$ 603.9, [2M+H]$^+$ 1206.6.

Synthesis of HCl•H-Arg(Mtr) ketobenzothiazole (S3) (*N*-(*N*-(4-Amino-5-(benzo[*d*]thiazol-2-yl)-5-oxopentyl)carbamimidoyl)-4-methoxy-2,3,6-trimethylbenzenesulfonamide).[141]

Compound **S2** (0.200 g, 0.40 mmol) was stirred in 1.5 M HCl/dioxane (10 mL) at room temperature for 18 h. The solvent was removed the resulting residue was dried *in vacuo* and purified by RP-HPLC (Figure A4). The received compound was used for the following peptide couplings.

Synthesis of Boc-Arg(Mtr) ketothiazole (S4) ((*S*)-*tert*-Butyl-(5-(3-((4-methoxy-2,3,6-trimethylphenyl)sulfonyl)guanidino)-1-oxo-1-(thiazol-2-yl)pentan-2-yl)carbamate).[154] To a solution of 2-bromothiazol (0.211 g, 1.29 mmol, 3.3 equiv.) in dry THF (10 mL) *n*-BuLi (2.5 M, 0.52 mL, 1.29 mmol, 3.3 equiv.) was added dropwise under inert atmosphere at -78 °C. The reaction mixture stirred for 1.5 h at -78 °C, followed by dropwise addition of compound **S1** (0.205 g, 0.39 mmol, 1 equiv.) at the same temperature. The resulting solution was stirred 2 h at -78 °C, after which sat. aqueous NH$_4$Cl (10 mL) was added. The organic phase was separated and the aqueous phase was extracted three times with EtOAc. The combined organic extracts were washed with brine (30 mL), dried over Na$_2$SO$_4$, filtered and concentrated *in vacuo*. The residue was purified on a silica column eluting with EtOAC/cyclohexane (4:1 v/v), to afford the compound S4 (0.12 g, 0.22 mmol, yield: 56%) as a white foam (Figure A4). Purity (LC, 254 nm) > 95%. ^{1}H NMR (300 MHz, CDCl$_3$): δ = 8.04 (d, 1 H), 7.72 (d, 1 H), 6.53 (s, 1 H), 5.64 (d, 1 H), 5.41 (s, 1 H), 3.83 (s, 3 H), 3.26 (m, 2 H), 2.67 (s, 3 H), 2.59 (s, 3 H), 2.12 (s, 3 H), 1.76 – 1.57 (m, 4 H), 1.41 (s, 9 H). ppm. MS (ESI): *m/z*: calcd for C$_{24}$H$_{35}$N$_5$O$_6$S$_2$ [M+H]$^+$ 554.2, found [M+H]$^+$ 554.2.

Synthesis of TFA•H-Arg(Mtr) ketothiazole (S5) ((*S*)-*N*-(*N*-(4-Amino-5-oxo-5-(thiazol-2-yl)pentyl)carbamimidoyl)-4-methoxy-2,3,6-trimethylbenzenesulfonamide). Compound **S4** (0.256 g, 0.46 mmol) was stirred in DCM (3 mL) at 0 °C and TFA (1 mL) was added. The

reaction mixture stirred for 1 h at ambient temperature, then isopropyl alcohol (0.5 mL) was added. The solution was concentrated *in vacuo* and triturated with diethyl ether. The supernatant was decanted and the residue was purified by RP-HPLC (Figure A4). The obtained compound **S5** was used for the following peptide couplings.

Synthesis of Boc-Arg(Mtr) alcohole (S6) ((*S*)-*tert*-Butyl-(1-hydroxy-5-(3-((4-methoxy-2,3,6-trimethylphenyl)sulfonyl)guanidino)pentan-2-yl)carbamate). To a solution of Boc-Arg(Mtr)-OH (0.3 g, 0.62 mmol, 1 equiv.) in dry THF (5 mL) were added NMM (0.063 g, 0.62 mmol, 1 equiv.) and EtOCOCl (0.067 g, 0.62 mmol, 1 equiv.) at -15 °C under argon. The reaction mixture stirred for 1 h at -15 °C, then transferred dropwise via canula into a stirred solution of NaBH$_4$ (0.047 g, 1.24 mmol, 2 equiv.) in water (15 mL). The resulting solution was stirred 5 minutes at 0 °C and then diluted with water (15 mL). The aqueous phase was extracted twice with EtOAc (10 mL). The combined organic extracts were dried over Na$_2$SO$_4$ and concentrated *in vacuo* to obtain the compound **S6** (0.24 g, 0.5 mmol, yield: 81%) as a colourless oil (Figure A4). Purity (LC, 254 nm) 98%. ^{1}H NMR (300 MHz, CDCl$_3$) δ = 6.52 (s, 1 H), 6.33 (s, 2 H) 5.15 (d, 1 H), 3.82 (s, 3 H), 3.55 (s, 2 H), 3.21 (s, 1 H), 2.69–2.66 (m, 5 H), 2.59 (s, 3 H), 2.12 (s, 3 H), 1.55 (s, 4 H), 1.40 (s, 9 H) ppm. LC-MS: *m/z*: calcd for C$_{21}$H$_{36}$N$_4$O$_6$S [M+H]$^+$ 473.2, found [M+H]$^+$ 473.2.

Synthesis of TFA•H-Arg(Mtr) alcohol (S7) ((*S*)-*N*-(*N*-(4-Amino-5-hydroxypentyl)carbamimidoyl)-4-methoxy-2,3,6-trimethylbenzenesulfonamide). Compound **S6** (0.22 g, 0.47 mmol) was stirred in DCM (3 mL) at 0 °C and TFA (1 mL) was added. The reaction mixture stirred for 2 h at ambient temperature, then isopropyl alcohol (0.5 mL) was added. The solution was concentrated *in vacuo* and triturated with diethyl ether. The supernatant was decanted and the residue was purified by RP-HPLC. The received compound **S7** was used for the following peptide couplings (Figure A4).

Preparation of inhibitors. Respective serine traps (1.5 equiv.) were coupled with dipeptides (1.0 equiv.) bearing standard protection groups using PyBOP (1.5 equiv.) and DIEA (3 equiv.) in DMF. After the reaction was agitated for 4 h at room temperature, 3 mL of a deprotection solution was added (93% TFA, 3.5% TIPS, 3.5% H$_2$O) and further agitated for 8 h at room temperature. After concentrating, the crude inhibitor was precipitated in 50 mL cold diethylether and afterwards purified using RP-HPLC (Figure A9).

Purification, lyophilization and analysis of peptide- and warhead precursors and inhibitors. All precursor compounds were purified with a semi-preparative RP-HPLC. The following gradient was applied: 95% H$_2$O / 5% ACN to 5% H$_2$O / 95% ACN in 30 minutes.

Trifluoracetic acid was dissolved in the water to a concentration of 0.1%. Zorbax Eclipse XDB column C-18 9.4x250 mm 5 µm, Agilent Technologies, California, USA. UV Vis detector model S-3702, Ever Seiko Corporation, Tokyo, Japan. For the detection a wavelength of 220 nm was used. After chromatographic purification, the fractions were collected and freeze dried overnight. The purified and lyophilized precursor compounds were stored in the freezer at -20 °C. The mass of the purified compounds was determined with MS-ESI. Expression-L compact mass spectrometer, Advion, New York, USA. Peptides were dissolved to a concentration of c = 0.01 mg/mL in MeOH + 0.1% formic acid. Injection was done by a syringe pump with a flow rate of 10 µL/min.

Enzymes. Recombinant human hepsin protein and factor Xa protein were purchased from Bio-Techne GmbH (Wiesbaden, Germany). Recombinant human thrombin protein and matriptase protein were purchased from R&D Systems (Minneapolis, MN, USA).

Determination of inhibitory constant K_i. The activity of the compounds against the recombinant human enzymes was determined in enzyme inhibition assays. Here, a ten-point dilution series for the inhibitors was prepared and incubated for 30 minutes with the enzyme in TNC buffer (25 mM Tris, 150 mM NaCl, 5 mM $CaCl_2$, 0.01% Triton X-100, pH = 8) prior to adding a fluorogenic reference substrate Boc-Gln-Ala-Arg-AMC for matriptase and TMPRSS2 or a chromogenic substrate D-Phe-Homopro-Arg-pNA, Bz-Ile-Glu-Gly-Arg-pNA for thrombin and factor Xa, respectively. The measurements were performed on a Tecan infinite® M1000 and the fluorescence intensity was measured by exciting the AMC fluorophore at 380 nm wavelength and recording emission at 460 nm wavelength. The absorption of pNA was measured at 405 nm. Fluorescence intensities and absorption were measured every 2 minutes for 2 h. The end concentrations of the enzymes were 0.2 nM (matriptase) and 0.2 nM (TMPRSS2) in 20 µL total volume and 0.6 nM (thrombin), 0.35 nM (factor Xa) in 100 µL total volume. The end concentration of the reference substrate was 100 µM (matriptase), 100 µM (TMPRSS2), 200 µM (factor Xa) and 100 µM (thrombin). To determine the IC_{50} values, the concentration-response data were plotted with the program GraphPad prism version 8.4.2 (San Diego, California) and a nonlinear regression fit with the equation [Inhibitor] vs. normalized response was applied. The inhibitory constant K_i was calculated from the IC_{50} values using the Cheng-Prusoff equation (K_i = $IC_{50}/[S]/K_M$) for competitive reversible inhibitors.[106] The K_M value was determined to be 77 µM for TMPRSS2 (Figure 5.2). Fluorescence intensities of the lowest inhibitor concentration were set to 100%.
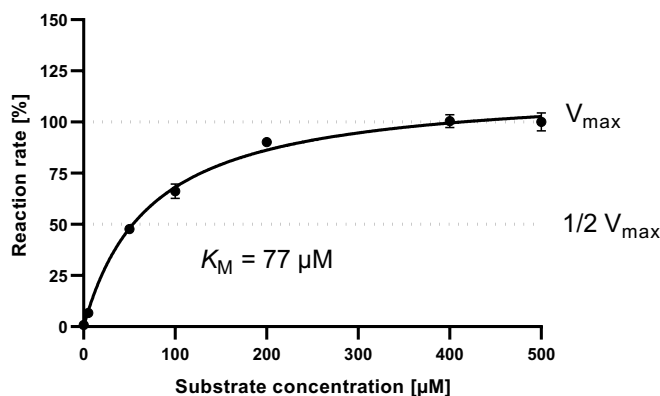
Figure 5.2. Michaelis-Menten constant (K_M) of the fluorogenic reference substrate Boc-Gln-Ala-Arg-AMC for TMPRSS2. The data points were plotted using the Michaelis-Menten equation in GraphPad Prism prism version 8.4.2 (San Diego, California). Shown are the means ± SD of triplicate measurements.

Analysis of cellular TMPRSS2 expression. The human colon adenocarcinoma cell line Caco-2 was purchased from the German Collection of Microorganisms and Cell Cultures (DSMZ, Germany) was maintained in Eagle's Minimum Essential Medium (EMEM) supplemented with 10% FBS, 100 U /mL penicillin, 100 mg/mL streptomycin, and 2 mM glutamine (all Invitrogen, Germany). For the validation of the expression of the transmembrane serine protease TMPRSS2, 100,000 Caco-2 cells were resuspended in 100 µL of Dulbecco's phosphate buffered saline (DPBS, Sigma-Aldrich) and incubated with TMPRSS2 antibody (rabbit anti-human IgG from ThermoFisher Scientific, PA5-14264) at final concentrations of 100 µg/mL, 40 µg/mL 20 µg/mL, and 10 µg/mL)] for 30 minutes at 4 °C. After the separation from unbound antibody molecules by centrifugation (200 × g for 3 min) and resuspension of the cells in 100 µL DPBS, 1 µL of a FITC-labeled secondary donkey anti-rabbit IgG (ThermoFisher, A16024) was added and incubated for 30 minutes at 4 °C. Following a final centrifugation (200 × g for 3 minutes), the cells were resuspended in 1 mL DPBS and analyzed by flow cytometry. The measurements were perfomed on an Attune™ NxT cytometer (ThermoFisher, Waltham, USA) with a 488 nm laser for excitation of bound secondary antibody molecules (FITC) and a 530/30 nm band pass filter for emission detection. Using the Attune™ NxT software (ThermoFisher), Caco-2 cells were selected by the FSC/SSC plot, thereby excluding cell debris. From this dot plot gating of Caco-2 cells, a histogram plot of the BL1-H emission filter signal was generated. The signal of untreated Caco-2 cells (autofluorescence) was gated to one percent, whereby all other samples refer to the percentage of events within this gate. For the data analysis, GraphPad prism version 8.4.2 was applied.

Analysis of inhibition of cellular TMPRSS2 activity. 10,000 Caco-2 cells in 100 µL RPMI medium supplemented with 10% FBS, 100 U/mL penicillin, 100 mg/mL streptomycin, and 2 mM glutamine (all Invitrogen, Germany) were seeded per well in a 96-well plate and incubated at 37 °C for four days until full confluency of the cells. The cells were washed two times with PBS and RPMI medium without FBS was added. For the determination of IC_{50} values, 1 µL of inhibitor was incubated for 30 minutes at room temperature prior to adding 2 µL of 10 mM reference substrate (Boc-Gln-Ala-Arg-AMC). The fluorescence intensity was measured as described above. Fluorescence intensities of the lowest inhibitor concentration were set to 100%.

SARS-CoV-2 pseudoparticles. To generate replication deficient lentiviral pseudoparticles carrying the SARS-CoV-2 spike protein (LV(Luc)-CoV-2), 900,000 HEK293T cells were seeded in 2 mL DMEM supplemented with 10% FBS, 100 U/mL penicillin, 100 mg/mL streptomycin, and 2 mM glutamine. The next day, the medium was refreshed and cells were transfected with a total of 1 µg DNA using polyethyleneimine (PEI). To this end, 2% of SARS-2 spike plasmid (encoding the spike protein of SARS-CoV-2 isolate Wuhan-Hu-1, NCBI reference sequence YP_009724390.1, SARS-CoV-2 variant B.1.1.7 or B.1.351) were mixed with pCMVdR8_91 (encoding HIV structural proteins gag and pol) and pSEW-Luc2 (crippled lentiviral vector encoding the luciferase reporter gene) in a 1:1 ratio in serum-free medium. Plasmid DNA was mixed with PEI at a DNA:PEI ratio of 1:3 (3 µg PEI per 1 µg DNA), incubated for 20 minutes at room temperature and added to cells dropwise. At 8 h post transfection, medium was removed, cells were washed with 2 mL of PBS and 2 mL of HEK293T medium with 2.5% FCS were added. At 48 h post transfection, pseudoparticles containing supernatants were harvested and clarified by centrifugation for 5 minutes at 1500 rpm.

SARS-CoV-2 strains and propagation. Viral isolates BetaCoV/Netherlands/01/NL/2020 of the pandemic D614G variant (#010V-03903), BetaCoV/France/IDF0372/2020 (#014V-03890) and virus lineage B.1.1.7 hCoV-19/Netherlands/NH-RIVM-20432/2020 (#014V-04031) were obtained from the European Virus Archive global. The virus lineage B.1.351 2102-cov-IM-r1-164 was isolated, sequenced and kindly provided by Michael Schindler. All strains were propagated on Vero E6 or Caco-2 cells. To this end, 70-90% confluent cells in 75 cm² cell culture flasks were inoculated with SARS-CoV-2 isolate (multiplicity of infection (MOI) of 0.03-0.1) in 3.5 mL serum-free medium. Cells were incubated for 2 h at 37 °C, before adding 20 mL medium containing 15 mM HEPES. Cells were incubated at 37 °C and supernatant harvested when a strong cytopathic effect (CPE) was visible. Supernatants were centrifuged for 5 minutes at 1,000 × g to remove cellular debris, and then aliquoted and stored at -80 °C as virus stocks.

Infectious virus titer was determined as plaque forming units (PFU) on Vero E6 cells, which was used to calculate MOI.

Pseudovirus inhibition assay. 10,000 Caco-2 cells were seeded in 100 µL DMEM supplemented with 10% FBS, 100 U/mL penicillin, 100 mg/mL streptomycin, 2 mM glutamine, 1× non-essential amino acids and 1mM sodium pyruvate. The next day, medium was replaced by 60 µL of fresh medium and cells were treated with 20 µL of serial dilutions of TMPRSS2 inhibitors or small molecule protease inhibitors for 2 h at 37 °C, followed by transduction with 20 µL of infectivity normalized LV(Luc)-CoV-2 pseudoparticles. Transduction rates were assessed after 48 h by measuring luciferase activity in cell lysates with a commercially available kit (Promega). Briefly, cells were washed with PBS and incubated with 40 µL cell culture lysis reagent for 10 minutes at room temperature. 30 µL of lysates were transferred to opaque 96-well plates and mixed with 50 µL of Luciferase assay substrate. Luminescence was recorded immediately for 0.1 s/well in an Orion II Microplate luminometer (Berthold) with simplicity 4.2 software. Luciferase activities in absence of inhibitors were set to 100% and IC_{50} were determined by linear regression using GraphPad Prism version 8.4.2.

SARS-CoV-2 inhibition assay. 25,000 Caco-2 cells were seeded in 100 µL respective medium. The next day 40 µL of medium were removed and cells were treated with 20 µL of serial dilutions of TMPRSS2 inhibitors or small molecule protease inhibitors for 2 h at 37 °C, followed by infection with 20 µL SARS-CoV-2 of the respective virus strain or at a multiplicity of infection (MOI) of 5×10^{-4}. Infection rates were assessed at 2 days post infection by in-cell ELISA for SARS-CoV-2 nucleocapsid or spike. Briefly, cells were fixed by adding 180 µL 8% paraformaldehyde (PFA) for 30 minutes at room temperature and permeabilized by incubation with 100 µL 0.1% Triton-X for 5 minutes. After washing once with PBS, cells were stained with 1:5,000 diluted anti-spike protein antibody 1A9 (Biozol GTX-GTX632604) or anti-nucleocapsid antibody (Sinobiological 40143-MM05) in antibody buffer (10% FCS and 0.3% Tween 20 in PBS) for 1 h at 37 °C. After 2 washes with 0.3% Tween 20 in PBS, the secondary HRP-conjugated antibody (Thermo Fisher #A16066) (1:15,000) was incubated for 1 h at 37 °C. Cells were washed 3 times with 0.3% Tween 20 in PBS, TMB peroxidase substrate (Medac #52-00-04) was added for 5 minutes and the reaction stopped using 0.5 M H_2SO_4. The optical density (OD) was recorded at 450 nm - 620 nm using the Asys Expert 96 UV microplate reader (Biochrom) with DigiRead 1.26 software. Values were corrected for the background signal derived from uninfected cells and untreated controls were set to 100% infection.

Cytotoxicity assay 10,000 Caco-2 cells were seeded in 100 µL respective medium. The next day, medium was replaced by 80 µL of fresh medium and cells were treated with 20 µL of

serial dilutions of peptidomimetic TMPRSS2 inhibitors or small molecule protease inhibitors. Cell viability was assessed after 48 h with a commercially available kit (Promega). Briefly, medium was removed and cells were lysed with 100 µL CellTiter-Glo reagent for 10 minutes at room temperature. 50 µL of lysates were transferred to opaque 96-well plates and luminescence was recorded immediately for 0.1 s/well in an Orion II Microplate luminometer (Berthold) with simplicity 4.2 software. Luciferase activities in absence of inhibitors were set to 100%.

Stability of inhibitors in serum, plasma and cell culture medium. The stability of the inhibitors was measured according to a modified procedure from Jensen et al..[130] In short, TMPRSS2 inhibitor solutions of 1 mM were prepared by dissolving the peptide in pure dimethyl sulfoxide. 10 µL of the peptide solution were added to 1 mL RPMI medium 1640 supplemented with 25% (v/v) human serum, citrate plasma or only RPMI medium and incubated at 37 °C. At indicated intervals, 100 µL samples were taken and mixed with 200 µL ethanol for precipitation of proteins. The cloudy solution was cooled at 4 °C for 15 minutes and centrifuged at 14,800 rpm for 2 minutes. The supernatant was aspirated and analyzed using analytical HPLC. The residual inhibitory constants K_i after 10 days incubation in biological fluids was measured as described above. For the data analysis, GraphPad prism version 7.0a was applied.

5.4 Results

Structure-based design of TMPRSS2 inhibitors. For the identification of novel peptide based TMPRSS2 inhibitors as a potential treatment of SARS-CoV-2 infection, molecular docking studies were performed. As no crystal structure of TMPRSS2 is available in the protein data bank (PDB) we used matriptase-1 as a surrogate model.[145,311] Matripase-1 shares 41% sequence identity with TMPRSS2 and previously described substrate analogue TMPRSS2 inhibitors showed no selectivity over matriptase.[226] Additionally, a homology model of TMPRSS2 was built using hepsin (43% sequence identity) as a template. The docking procedure of those crystal structures was tested by redocking of reference ligands for matriptase and TMPRSS2 and receiver operator characteristic (ROC) curve analysis which indicated a very strong ability to discriminate between known binders (Figure 5.1, Table 5.1).

Table 5.1. FlexX-scores of tripeptidic substrate-analogue ligands for docking receptor validation. Scores are in kJ/mol, ace: N-terminal acetyl-cap, nme: C-terminal N-methylamide cap. [a]Predicted binding modes is shown in Figure 5.3.

Ligand sequence	FlexX-score TMPRSS2 homology model	FlexX-score matriptase surrogate
ace-D-Arg-Gly-Arg-nme	-48.8	-54.9
ace- D-Arg-Pro-Arg-nme[a]	-53.8	-64.2
ace-Arg-Gly-Arg-nme	-51.1	-56.3
ace-Arg-Pro-Arg-nme	-48.0	-57.7

Previous studies identified the substrate requirements of TMPRSS2, revealing a preference for arginine in P1 position, a glycine or proline residue in P2 position, and a D-configured arginine in P3 position.[226] Based on this substrate preference, a reference binder comprising a N-terminal acetyl-cap and a C-terminal aldehyde serine trap with the sequence ace-D-Arg-Gly/Pro-Arg-aldehyde was designed and docked to the matriptase-based TMPRSS2 surrogate model (Figure 5.3) and TMPRSS2 homology model (Figure A10). Our dockings show that both the reference binder and the published TMPRSS2 inhibitor CM bind to the reactive center of the surrogate and the homology model (Figure 5.3; Figure A10).

A

B
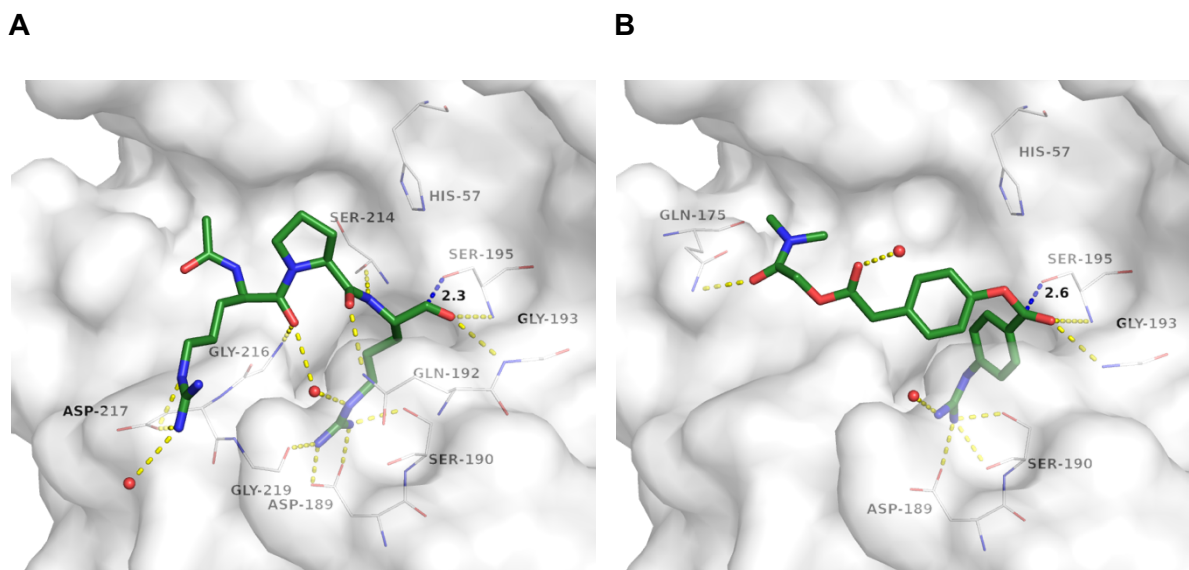
Figure 5.3. Predicted binding mode of **A** reference binder ace-D-Arg-Pro-Arg-aldehyde and **B** camostat mesylate (CM) to matriptase as surrogate for TMPRSS2 (white carbon atoms and surface). For clear view only residues forming polar interactions (yellow dashed lines), and the catalytic residues Ser-195 and His-57 are depicted as lines. Distance between nucleophilic serine oxygen and electrophilic carbon atom of the serine trap is illustrated by a dashed blue line. Carbon atoms are shown in green, oxygens in red and nitrogens in blue.

To optimize the binding affinity of the reference binder, we altered the residues at P1-P3 position using both proteinogenic and non-proteinogenic amino acids and docked the resulting structures to both, the matriptase surrogate and TMPRSS2 homology model. Compounds were then ranked based on their binding score (Table 5.2-5.4), the plausibility of their binding mode to the S1-S3 sub-pockets of the TMPRSS2 active site and proximity of the aldehyde serine trap to the catalytic Ser186/195 (TMPRSS2 homology model/matriptase enumeration), as well as commercial availability of their building blocks.

Table 5.2. Docking results of P1-sidechain screening of peptidomimetic inhibitors with ace-Pro-X-aldehyde sequence for 7 potential basic amino acid-derivatives. X: position of altered amino acid, ace: N-terminal acetyl cap, mPG: meta-Phenylguanidine, pPG: para-Phenylguanidine, mBA: meta-Benzamidine, pBA: para-Benzamidine, Orn: Ornithin. Scores are in kJ/mol.

Ligand ace-Pro-X-aldehyde	FlexX-score TMPRSS2 homology model	FlexX-score matriptase surrogate
pBA	-51.2	-42.1
Arg	-45.1	-50.8
mPG	-40.1	-41.9
mBA	-33.9	-31.0
pPG	-28.3	-34.4
Lys	-22.9	-32.9
Orn	-19.9	-29.0

Table 5.3. Excerpt of docking results of P2-sidechain screening of peptidomimetic inhibitors with ace-D-Arg-X-Arg-aldehyde sequence. A total of 369 molecules were docked. Scores are in kJ/mol. AZc: Azetidine-2-carboxylic acid, Pip: Pipecolinic acid, Pgl: 3,4-Dichlorophenylglycine, Cyc: Cyclobutylalanine.

Ligand sequence ace-D-Arg-X-Arg-aldehyde	FlexX-score (rank) TMPRSS2 homology model	FlexX-score (rank) matriptase surrogate
Arg	-57.4 (1)	-57.8 (59)
Orn	-51.0 (2)	-59.7 (25)
Pro	-49.9 (8)	-60.5 (16)
Pgl	-49.2 (9)	-58.8 (41)
Pro	-48.4 (15)	-61.2 (9)
Lys	-47.5 (23)	-58.6 (45)
Thr	-46.9 (29)	-62.2 (4)
Azc	-46.9 (39)	-60.1 (22)
Val	-45.9 (54)	-62.1 (6)
Pip	-44.1 (101)	-55.9 (110)
Leu	-43.5 (127)	-56.6 (96)
Ile	-41.9 (183)	-61.5 (8)
Cyc	-39.8 (269)	-52.9 (227)

Table 5.4. Excerpt of docking results of P3-sidechain screening of peptidomimetic inhibitors with ace-X-Pro/Gly-Arg-aldehyde sequence. For both X-Pro-Arg and X-Gly-Arg sequences 388 molecules were docked and ranked separately. Scores are in kJ/mol.

Ligand sequence ace-X-Y-Arg-aldehyde	FlexX-score (rank) TMPRSS2 homology model		FlexX-score (rank) Matriptase surrogate	
X	Y = Pro	Y = Gly	Y = Pro	Y = Gly
Arg	-52.0 (1)	-48.3 (3)	-55.1 (298)	-56.4 (28)
D-Arg	-44.2 (135)	-45.5 (10)	-64.9 (1)	-56.4 (27)
His	-39.7 (348)	-40.5 (144)	-54.5 (364)	-54.3 (90)
D-His	-49.3 (6)	-41.8 (92)	-62.2 (16)	-57.4 (6)
Trp	-45.1 (97)	-44.2 (20)	-63.5 (7)	-54,7 (71)
D-Trp	-45.1 (104)	-42.8 (50)	-59.7 (70)	53.5 (122)
Asn	-40.5 (329)	-40.2 (160)	-52.9 (364)	-56.5 (23)
D-Asn	-46.7 (47)	-42.1 (78)	-57.1 (190)	-51.8 (201)
Met	-42.3 (253)	-38.4 (261)	-57.8 (158)	-47.9 (369)
D-Met	-38.7 (370)	-35.8 (368)	-59.4 (77)	-50.1 (281)
Glu	-39.1 (362)	-40.8 (127)	-54.3 (323)	-50.9 (241)
D-Glu	-40.6 (325)	-38.6 (252)	-50.6 (384)	-51.9 (193)

Overall, D-configuration for P3 residue was favored to improve metabolic stability. The most promising compounds were chosen for solid phase synthesis, where the aldehyde serine trap used for *in silico* modelling was exchanged by a well-established ketobenzothiazole, yielding a library of peptidomimetic inhibitors (Table 5.5).

Table 5.5. Assembled peptidomimetic inhibitor library selected for synthesis.

Cpd.	N-cap	P3	P2	P1	Serine trap
1	ace	D-Arg	Pro	Arg	kbt
2	ace	Arg	Pro	Arg	kbt
3	ace	D-His	Pro	Arg	kbt
4	ace	His	Pro	Arg	kbt
5	ace	Asn	Pro	Arg	kbt
6	ace	D-Arg	Pip	Arg	kbt
7	ace	D-Arg	Cyc	Arg	kbt
8	ace	D-Arg	Thr	Arg	kbt

Abbreviations: Cpd, compound; ace: N-terminal acetyl cap; Cyc: Cyclobutylalanine; Pip: Pipecolinic acid; kbt: ketobenzothiazole. The bond between the P1 position and serine trap is the site of nucleophilic attack by the protease catalytic triad.

To identify literature-known inhibitors for TMPRSS2, a focused library consisting of 315 commercially available inhibitors reported for the closely related trypsin-like serine proteases factor Xa, hepsin, kallikrein, plasminogen, thrombin, matriptase-1, matriptase-2, trypsin and urokinase was virtually screened. The most promising virtual screening (VS) hits based on score, binding pose and availability were subsequently purchased and tested (Table 5.6).

Table 5.6. Virtual screening hits selected for testing. [a]Potential covalent inhibiton, [b]control compound as low affinity reference. Scores are in kJ/mol.

ZINC ID	Synonym / international nonproprietary name (INN)	FlexX-score (rank) TMPRSS2 homology model	FlexX-score (rank) Matriptase surrogate
3809827	Melagatran	-50.2 (6)	-56.3 (2)
1910616	Dabigatran	-48.7 (8)	-48.2 (15)
13826594	Leupeptin	-44.9 (17)	-51.2 (14)
26376363	PPack[a]	-36.8 (88)	-54.1 (6)
3871842	Camostat[a]	-16.5 (281)	-25.1 (207)
36634	Benzamidine[b]	-33.1 (117)	-33.2 (119)

Designed peptidomimetic inhibitors block TMPRSS2 and matriptase activity. We next investigated the impact of our peptidomimetic inhibitors on the activity of closely related matriptase and TMPRSS2 enzymes. To this end, the respective purified proteases were incubated with the compounds **1-8**, followed by adding a protease specific reporter substrate that allowed monitoring of protease activity over time. Overall, the compounds suppressed TMPRSS2 activity in the low nanomolar range (K_i = 2.5 - 215.9 nM) and inhibit matriptase with comparable activity (Table 5.2; Figure A11). Compound **1** which contains the peptide sequence of the reference binder showed an activity of K_i = 86.7 nM, while the compounds **2**, **3**, **4**, **5** and **7** were most active against isolated TMPRSS2, with inhibitory constants of 2.5 - 57.5 nM. The compounds **2** (K_i = 3.8 nM) and **5** (K_i = 2.5 nM) were 2-3-fold more active than the active metabolite of CM, FOY-251 (K_i = 9.7 nM). Compounds **6** and **8** were the least active with K_i values of 71.5 and 215.9 nM, respectively.

For systemic administration of the protease inhibitors, high selectivity over off-target proteases is required to reduce side effects. To investigate potential interference of the peptidomimetic inhibitors with serine proteases involved in coagulation, we assessed their activity against thrombin and factor Xa (Table 5.7; Figure A11). All compounds excluding compound **1** display a > 100-fold selectivity against thrombin (Table 5.7). Compound **1** shows no selectivity over factor Xa while compounds **2**, **3**, **4**, **5, 6** and **7** reveal 1.6- to 38.9-fold selectivity compared to TMPRSS2.

Table 5.7. Overview of inhibitory activity (K_i) of synthesized TMPRSS2 inhibitors 1-8 against TMPRSS2, matriptase, thrombin and factor Xa. Selectivity indices represent the quotient of K_i values of matriptase, thrombin, and factor Xa by the K_i value of TMPRSS2.

Compound	K_i [nM]				selectivity indices		
	TMPRSS2	Matriptase	Thrombin	Factor Xa	Matriptase	Thrombin	Factor Xa
1	86.7	1.7	2,077	4.1	0.02	24	0.05
2	3.8	3.3	599	106.7	0.9	158	28.1
3	9.1	14.0	4,088	271.3	1.5	449	29.8
4	8.5	11.5	7,217	331	1.4	849	38.9
5	2.5	5.2	1,046	41.1	2.1	418	16.4
6	71.5	75.6	> 50,000	472.4	1.1	>699	6.6
7	57.7	30	> 50,000	94.1	0.5	>867	1.6
8	215.9	159.8	> 50,000	965.5	0.7	232	4.5
CM	0.4	0.6	> 50,000	1,785.5	1.5	>125,000	4,464
FOY-251	9.7	173.4	> 50,000	697	17.9	>5,155	71.0

The activity of the inhibitors against TMPRSS2 correlated with their activity against matriptase and revealed a linear correlation (Figure 5.4). Yet, compound **1** showed the highest selectivity (~51-fold) for matriptase while FOY-251 showed the highest selectivity (~18-fold) for TMPRSS2.
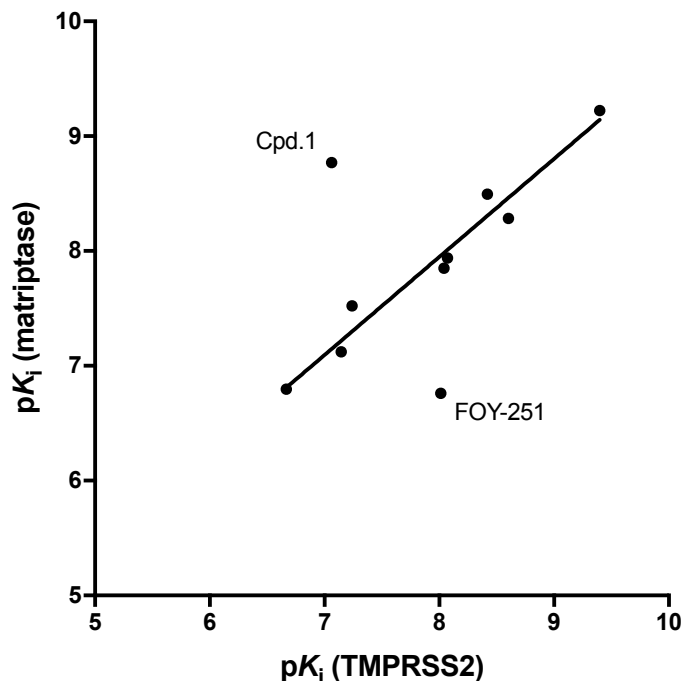
Figure 5.4. Correlation of the pK_i values from compounds 1-8, CM and FOY-251 against TMPRSS2 and matriptase. A linear regression fit (R^2 = 0.96) was plotted through the data points excluding cpd. 1 and FOY-251.

Truncation of the ketobenzothiazole serine trap moiety to a ketothiazole did not improve activity against matriptase, nor thrombin/factor Xa selectivity and further reduction to an alcohol abolished antiprotease activity (Table 5.8, Figure A12). Considering the inhibitory constants and selectivity over potential off-target coagulation proteases, the compounds **2, 4**, **5** and **7** were further analyzed for inhibition of cellular TMPRSS2 activity.

Table 5.8. Influence of serine trap on biological activity of compound 7 against matriptase, thrombin and factor Xa. The ketobenzothiazole (7) serine trap moiety was truncated to ketothiazole (7-2) and further reduced to the alcohol (7-3) for complete abolishment of electrophilicity. N.i = no inhibition within concentration range.

Compound	K_i [nM]		
	Matriptase	**Thrombin**	**Factor Xa**
7	30	n.i	94.1
7-2	141.2	n.i	98.7
7-3	n.i	n.i	n.i

Among the selected, high scoring literature known inhibitors, only Leupeptin and PPack showed activity against matriptase with K_i = 1452 nM and K_i = 12.4 nM, respectively (Table 5.9). The anticoagulant drugs PPack, Dabigatran and Melagatran show inhibitory activities comparable to literature known values.[315-317] Due to low activity against matriptase and poor selectivity against the coagulation proteases, those inhibitors were not further investigated.

Table 5.9. K_i values of reference inhibitors against matriptase, TMPRSS2, thrombin and factor Xa. N.i = no inhibition within concentration range.

	K_i [nM]		
Compound	**Matriptase**	**Thrombin**	**Factor Xa**
CM	0.6	n.i	n.i
FOY-251	173.4	n.i	n.i
Melagatran	n.i	21	110.6
Dabigatran	n.i	15.3	267.5
Leupeptin	1,452	n.i	n.i
Ppack	12.4	0.2	38.6
Benzamidine	n.i	n.i	n.i

Having demonstrated that the designed peptidomimetics inhibit cell free TMPRSS2 activity, we next analyzed inhibition of cell associated enzyme activity. For this we used SARS-CoV-2 permissive Caco-2 cells, which show upregulation of the TMPRSS2 gene and express TMPRSS2 on the cell surface (Figure 5.5).[318]
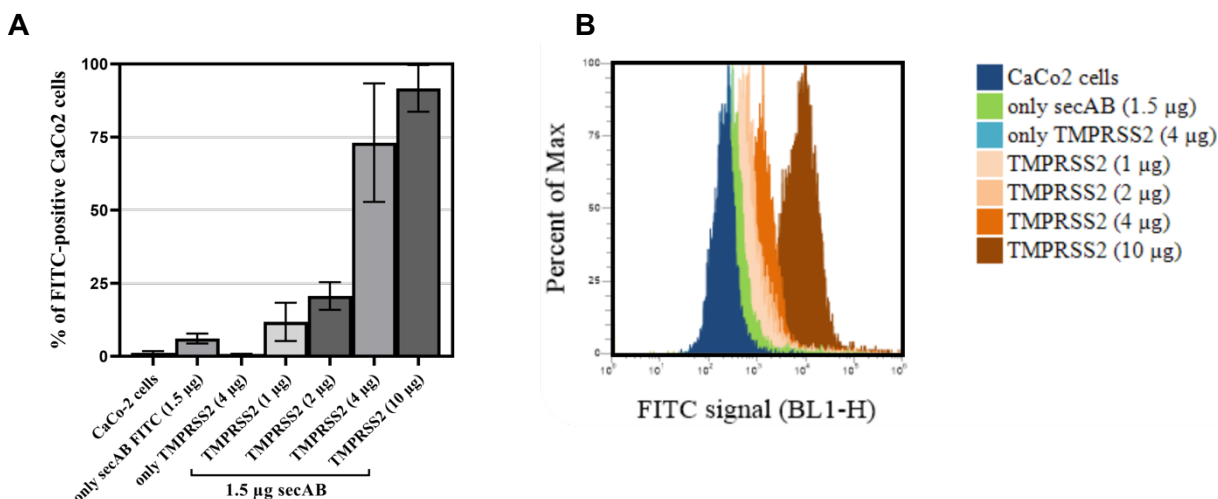
Figure 5.5. The transmembrane serine protease TMPRSS2 is expressed on the surface of Caco-2 cells. The Figure shows the detection of the transmembrane serine protease TMPRSS2 on human Caco-2 cells, which was performed by first incubating the cells with different amounts of a TMPRSS2 antibody (rabbit anti-human) followed by the detection of bound TMPRSS2 antibodies on the cell surface via an anti-rabbit FITC-labeled secondary antibody. Only secAB (secondary antibody) FITC and only TMPRSS2 were used as negative control groups. **A** shows the percentage of FITC-positive Caco-2 cells, where the bars represent the percentage of gated cell measurements generated by flow cytometry (n = 3, ± standard deviation). **B** shows an overlay of the normalized FITC signal histograms for each group (one out of three histograms is exemplary shown for each group).

Cells were incubated with the respective inhibitors and treated with fluorogenic protease substrate. The most potent inhibitors against matriptase and TMPRSS2 also efficiently prevented cell-mediated proteolysis of the fluorogenic substrate with half maximum inhibitory concentrations (IC_{50}) of 12.7 - 234.2 nM, with compounds **2** (IC_{50} = 32 nM) and **5** (IC_{50} = 12.7 nM) being most active (Table 5.10, Figure A14). Taken together, our results demonstrate that the synthesized peptidomimetic inhibitors potently reduce the activity of purified matriptase and TMPRSS2 while showing no activity against thrombin. Further, the selected most potent inhibitors display selectivity over factor Xa and reduce cellular protease activity.

Table 5.10. IC_{50} values of peptidomimetic TMPRSS2 inhibitors, camostat mesylate (CM) and FOY-251 measured on Caco-2 cells.

Compound	IC_{50} [nM]
2	32
4	45.5
5	12.7
7	234.2
CM	42.2
FOY-251	377.2

TMPRSS2-specific peptidomimetic inhibitors block SARS-CoV-2 infection. We next analyzed whether compounds **1-8** may inhibit SARS-CoV-2 spike driven viral entry. For this, Caco-2 cells treated with serial dilutions of the compounds (and CM as control) were inoculated with luciferase encoding lentiviral pseudoparticles carrying the wildtype SARS-CoV-2 spike protein. Transduction rates were determined 2 days later by measuring cell-associated luciferase activity and showed a concentration dependent inhibition of viral entry for all analyzed compounds (Figure A15). Compound **5** was most efficient with an IC_{50} value of 467.2 nM and was even more potent than CM ($IC_{50} \sim 747.5$ nM). Compounds **1-4** and **7** suppressed spike driven entry with IC_{50} values between 1,200 and 2,068 nM while compounds **6** and **8** were the least antivirally active with IC_{50} values between 5,604-12,085 nM, respectively (Table 5.11, Figure A15).

As the wildtype virus has largely been replaced by other SARS-CoV-2 clades, and variants of concern (VOC) with increased transmissibility and virulence emerged, we also determined the activity of selected compounds against the B.1.1.7 (emerged in the UK) and B.1.351 (emerged in South Africa) spikes. TMPRSS2 inhibitors **2**, **4**, **5** and **7** suppressed cell entry mediated by both spike variants in a dose-dependent manner with IC_{50} values ranging between 260.7-1,597 nM. (Table 5.11, Figure A15). Collectively these data show that the designed TMPRSS2 inhibitors suppress SARS-CoV-2 spike driven viral transduction.

Table 5.11. IC$_{50}$ values of peptidomimetic TMPRSS2 inhibitors and camostat mesylate (CM) against pseudotype lentivirus.

	IC$_{50}$ [nM]		
Compound	**Wildtype**	**B.1.1.7**	**B.1.351**
1	1,201	n.d.	n.d.
2	1,617	390.5	433.9
3	3,243	n.d.	n.d.
4	1,322	485.7	141.2
5	467.2	260.7	153.3
6	12085	n.d.	n.d.
7	2068	1597	920.8
8	5604	n.d.	n.d.
CM	747.5	98.73	156.4

We next analyzed whether the inhibitors may also block authentic SARS-CoV-2 infection. To this end, Caco-2 cells were supplemented with serial dilutions of the compounds **1-8** or CM and were then infected with a wildtype SARS-CoV-2 isolate obtained from France. Infection rates were determined 2 days later by quantifying intracellular spike protein expression by ELISA.[319] All compounds including CM suppressed SARS-CoV-2 infection in a concentration dependent manner (Figure A16). The IC$_{50}$ values were, however, generally higher as compared to the pseudotype experiment (Table 5.12). Compounds **2**, **4** and **5** were the most potent inhibitors with IC$_{50}$ values of 4.6, 5.7 and 4,7 µM, respectively, similar to CM (3.6 µM). The remaining compounds were less active with IC$_{50}$ values > 10 µM.

Finally, we determined the antiviral activity of compounds **2**, **4**, **5** and **7** against a SARS-CoV-2 isolate harboring the D614G mutation, which increases viral infectivity, and the VOCs B.1.1.7 or B.1.351 (Figure A16).[320] The four selected compounds as well as CM inhibited all three tested SARS-CoV-2 isolates. Compounds **2** and **5** suppressed "D614G" infection with IC$_{50}$ values of 17.1 and 11.9 µM, respectively and were even more active than CM (26.7 µM) (Table 5.12). SARS-CoV-2 variants B.1.1.7 and B.1.351 were most efficiently inhibited by compound **5** (IC$_{50}$ of 6.8 and 6.3 µM, respectively) and CM (IC$_{50}$ of 16.4 and 9.3 µM, respectively). Compounds **2**, **4** and **7** showed moderately higher IC$_{50}$ as compared to CM and compound **5**, (Table 5.12), but were still capable of blocking infection entirely. Thus, the designed

peptidomimetic TMPRSS2 inhibitors prevent SARS-CoV-2 infection with comparable antiviral activity as CM, which is currently evaluated in clinical trials as COVID-19 therapeutic.

Table 5.12. IC_{50} values of peptidomimetic TMPRSS2 inhibitors and camostat mesylate (CM) against SARS-CoV-2 wildtype and variants of concerns.

Compound	IC_{50} [nM]			
	Wildtype	**D614G**	**B.1.1.7**	**B.1.351**
1	18,889	n.d	n.d	n.d
2	4,567	17,126	~20,000	6,880
3	12,569	n.d	n.d	n.d
4	5,667	15,405	~20,000	6,985
5	4,654	11,867	6,782	6,260
6	>100,000	n.d	n.d	n.d
7	20,427	51,170	21,649	24,011
8	24,460	n.d	n.d	n.d
CM	3,625	26,672	16,407	9,307

The DMSO solvent control neither affected transduction of cells with SARS-CoV-2 pseudoparticles nor the infection with wildtype virus (Figure 5.6) and we did not observe cytotoxic effects from the compounds tested that exceed toxicity of the solvent control (Figure 5.7).
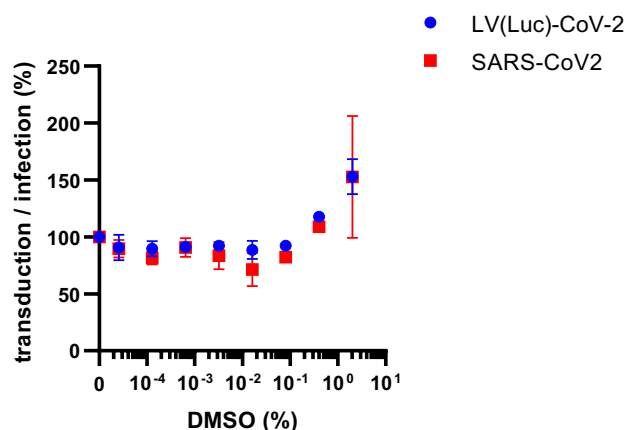
Figure 5.6. Impact of DMSO on SARS-CoV-2 spike mediated entry and infection. DMSO was added to Caco-2 cells at concentrations corresponding to the maximum DMSO concentrations in transduction/infection experiments. After 1 h cells were either transduced with lentiviral SARS-CoV-2 pseudoparticles or infected with SARS-CoV-2 wildtype. Transduction/infection rates were assessed 2 days later by measuring luciferase activity in cell lysates or by ELISA, respectively. Shown are the mean ± SEM of two experiments, performed in triplicates.
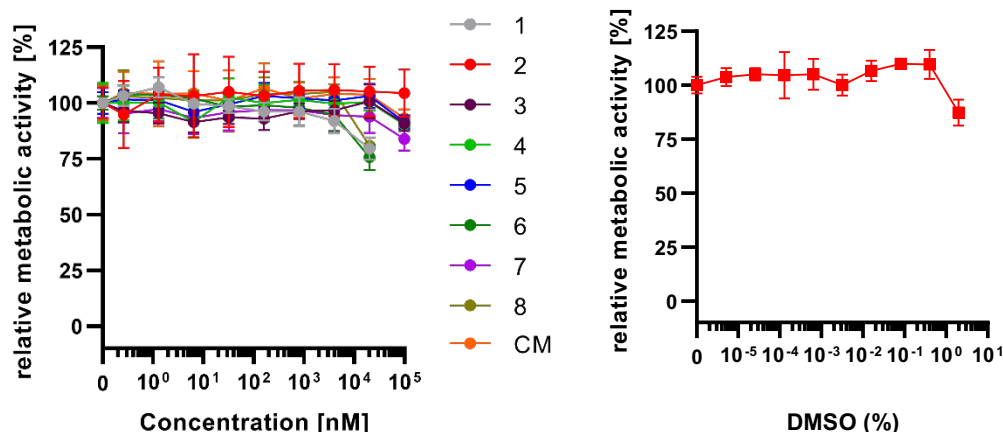
Figure 5.7. Cytotoxicity of TMPRSS2 inhibitors. TMRPSS2 inhibitors and the small molecule camostat mesylate (CM) were added to Caco-2 cells. Cell viability was assessed 2 days post addition by measuring ATP content in cell lysates. Due to low stock concentration compounds 1, 6 and 8 were tested at a maximum concentration of 20,000 nM. As solvent control, DMSO concentrations corresponding to the maximum DMSO concentration of cells were assessed. Shown is the mean ± SD of one experiment, performed in triplicates.

In vitro **stability of selected leads in body fluids and epimerization studies.** Stability of peptidomimetic inhibitors is a main challenge in peptide drug development. Peptidases in blood might degrade the inhibitor before reaching the desired target in an adequate time. Therefore, the stability of two compounds in body fluids was assessed. Compound **2** was selected due to

the high potency, selectivity as well as potency in suppressing genuine SARS-CoV-2 infection, and compound **7** was selected since the structure comprises exogenous amino acids with potential resistance to proteolysis. In a first analysis, compound **7** was spiked into 25% human serum and chromatographic analysis of peptide content after up to 10 days confirmed the presence of residual inhibitor (Figure 5.8). Due to the strong electrophilicity of the ketobenzothiazole serine trap, the P1 Arg C_α atom is prone to epimerization.[139] After introduction to blood serum, compound **7** rapidly epimerized within 30 minutes (Figure 5.8).
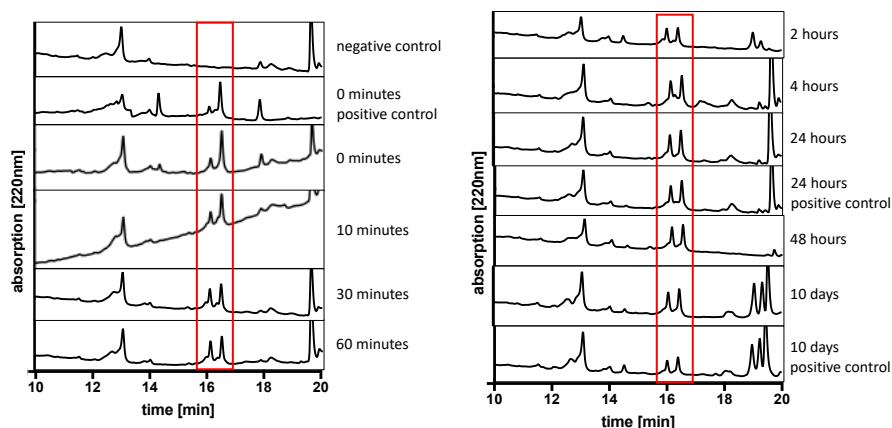
Figure 5.8. UV absorbance profile (220 nm) of RP-HPLC chromatograms of 7 after various time points of incubation with 25% serum. Two epimers can be observed D-Arg (t_R = 16:00 minutes) and L-Arg (t_R = 16:30 minutes). Negative control = 25% blood serum in RPMI; positive control = inhibitor in pure RPMI medium.

Activity measurements on purified enzyme show that the compounds **2** and **7** retained their inhibition towards matriptase (Figure 5.6). To further investigate the residual activity of the compounds, the inhibitory constant K_i were determined and compared (Figure 5.6). For the compounds **2** and **7**, the loss of activity was 8-fold in serum (**2**, K_i = 25.3 nM; **7**, K_i = 71.7 nM) and 6-7-fold in plasma (**2**, K_i = 17.7 nM; **7**, K_i = 78.3 nM).

A
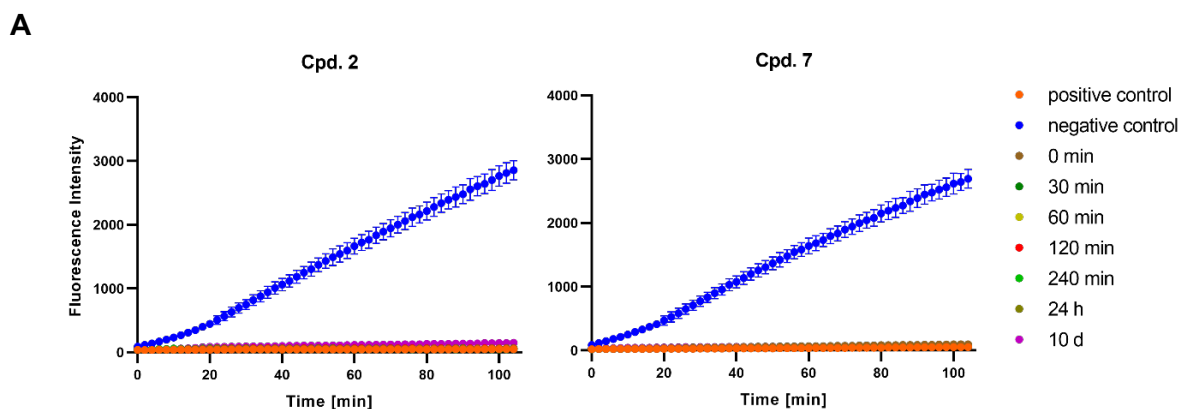
B
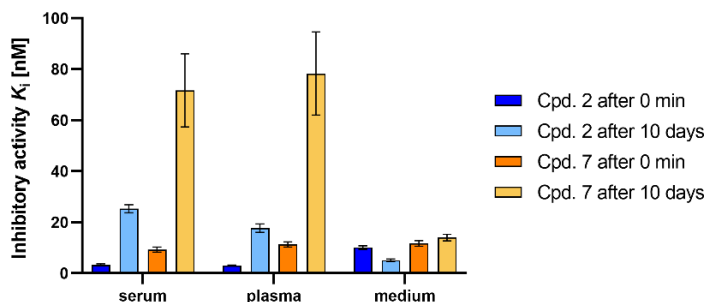
Figure 5.9. A Inhibition of reference substrate degradation by matriptase with compound 2 (left panel) and 7 (right panel) at c = 1,665 nM after incubation in blood serum at various time points. Positive control = inhibitor without incubation, negative control = no inhibitor. **B** Residual inhibitory activity K_i for compound 2 and compound 7 against matriptase after 0 minutes and 10 days incubation in 25% serum or 25% plasma in RPMI medium or pure RPMI medium.

For a better understanding of the structural influence of P1 Arg C_α, the epimeric mixture of compound **2**, due to higher chromatographic resolution, was separated and the enzymatic activity of each epimer determined (Figure 5.10).

A
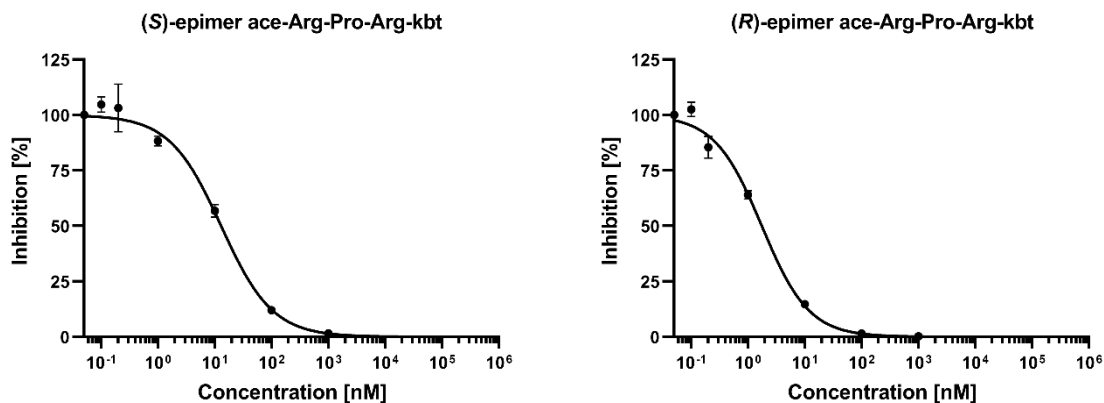
B

Compound	P1 Arg chirality	matriptase K_i [nM]
2S,R	S,R	3.3
2S	S	6.2
2R	R	0.9

Figure 5.10. A Dose-response curves of (S)- and (R)-epimers of compound 2 (ace-Arg-Pro-Arg-kbt) and **B** calculated inhibitory activity against matriptase. Shown is the mean ± SD of one experiment, performed in triplicates.

The assignment of the (*S*)- and (*R*)-epimer in the chromatogram was determined not by crystallographic analysis but by elution order and comparison with similar compounds in literature.[139] Both epimers showed high enzymatic activity with the (*R*)-epimer being more active than the (*S*)-epimer (**2*S***, K_i = 6.2 nM; **2*R***, K_i = 0.9 nM). Incubation of an epimer in TNC buffer at pH = 8, which was used as medium for the determination of enzymatic activity, did not significantly alter the ratio of (*S*)- and (*R*)-epimer (Figure 5.11).

A
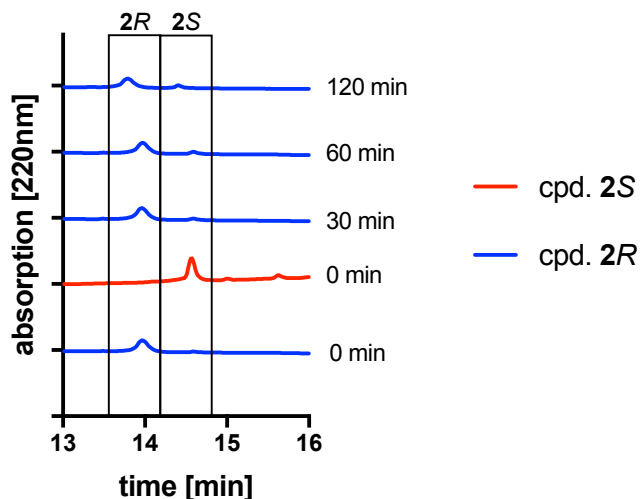
B

	ratio **2*R***:**2*S*** epimer
0 min incub. **2*R***	96:4
30 min incub. **2*R***	93:7
60 min incub. **2*R***	89:11
120 min incub. **2*R***	87:13

Figure 5.11. A HPLC chromatogram of compound 2*S* (red) and 2*R* (blue). The chemical stability of compound 2*R* in TNC buffer at pH = 8 with regard to racemization was analyzed at different time points (0, 30, 60 and 120 minutes). **B** The ratio of 2*R* and 2*S* epimer was determined through integration of the area under the curves.

5.5 Discussion

Here we describe novel, potent and stable peptidomimetic inhibitors of TMPRSS2 that block SARS-CoV-2 infection. Targeting TMPRSS2 is a promising antiviral strategy because the protease is not only essential for SARS-CoV-2 entry, but also primes glycoproteins of various other viruses for subsequent fusion and infection. Since TMPRSS2 is a host and not a viral protein TMPRSS2-targeting therapeutics that block its enzymatic activity should be less likely to induce resistance mutations.

To develop TMPRSS2 inhibitors, we used literature data on TMPRSS2 substrate preferences and designed a reference binder, which was used as a template for the preparation of peptidomimetic libraries, which then were screened *in silico* against the binding cavity of matriptase as a surrogate model for TMPRSS2 and against a TMPRSS2 homology model. A library of recognition sequences was compiled by incorporation of the identified top scoring amino acids in the template. The recognition sequences were connected with an electrophilic ketobenzothiazole warhead moiety as a reactive functional group to yield a panel of inhibitors.[140, 141, 321] We tested the inhibitors against isolated enzymes and our data identified the four compounds **2**, **4**, **5** and **7** as potential hits with high activity against TMPRSS2 and matriptase, and good off-target selectivity against coagulation proteins thrombin and factor Xa. CM and its rapidly forming active metabolite FOY-251 in blood were used for comparison since they have been shown previously to efficiently inhibit TMPRSS2 proteolytic activity and CM is currently evaluated in clinical trials for COVID-19.[281, 322] Our best candidates show inhibitory activities in the same range as CM, and the compounds **2** and **5** even show a 2-3-fold higher activity against isolated TMPRSS2 than FOY-251. The ketobenzothiazole serine trap moiety revealed the highest activity which may be attributed to a preferential fit in the hydrophobic S1′ pocket in contrast to the less hydrophobic and smaller ketothiazole. Substitutions on the benzothiazole with amino acids to interact more specifically with the S'-site might be of interest in further studies.[141]

Furthermore, the high activity of the top compounds was confirmed through cleavage of a fluorogenic substrate on Caco-2 epithelial cells, which serve as an intestinal model carrying the TMPRSS2 protease on the surface. Finally, the inhibitors blocked SARS-CoV-2 spike driven viral entry into Caco-2 cells and infection of Caco-2 cells by authentic SARS-CoV-2 wildtype and variants of concern in a concentration dependent manner. Thus, the designed inhibitors likely block TMPRSS2 mediated proteolytic priming of the viral spike protein, thereby preventing subsequent receptor binding and fusion. These data also show that SARS-CoV-2 VOCs are still dependent on TMPRSS2 as essential co-factor for cell entry and demonstrate that VOCs that escape from preexisting immunity are equally sensitive to entry inhibitors, as previously shown for soluble ACE2 or fusion-inhibiting peptide EK1 and EK1C4.[297, 323]

Two compounds (**2** and **7**) were incubated in body fluids for up to 10 days and still showed residual inhibitory activity in the sub nanomolar range which is remarkable considering the literature known stability issues of peptide therapeutics.[324, 325] The rapid epimerization of the compounds in blood serum did not alter the activity significantly, suggesting that structurally simplified inhibitors may be developed. The high stability in body fluids and potent anti-TMPRSS2 and anti-SARS-CoV-2 activity warrant further preclinical development of selected compounds. Of note, the TMPRSS2 inhibitors will not only act against SARS-CoV-2 but potentially also block other TMPRSS2-dependent coronaviruses such as SARS-CoV and MERS-CoV, and likely also future novel emerging coronaviruses, and also TMPRSS2-dependent viruses from other viral families. In sum, TMPRSS2 represents an attractive drug target in COVID-19 and downregulation of its enzymatic activity with active and selective inhibitors should significantly improve health rehabilitation. Here we showed a new direction for the fast development of peptidomimetic inhibitors and our results offer potential candidates with comparable activities to CM whose efficacy may be further elucidated in *in vivo* studies.

6. Alpha-1 antitrypsin inhibits TMPRSS2 protease activity and SARS-CoV-2 infection

Copyright:

This sub-chapter 6 was published as a peer-reviewed journal **[5]**. Therefore, the text below is in nearly all instances a word-by-word reproduction of the peer-reviewed and published paper. Presented results are reprinted with permission from American Chemical Society, Biomacromolecules. Copyright © 2021 American Chemical Society.

[5] Lukas Wettstein*, Tatjana Weil, Carina Conzelmann, Janis A. Müller, Rüdiger Groß, Maximilian Hirschenberger, Alina Seidel, Susanne Klute, Fabian Zech, Caterina Prelli Bozzo, Nico Preising, Giorgio Fois, Robin Lochbaum, **Philip Maximilian Knaff**, Volker Mailänder, Ludger Ständker, Dietmar Rudolf Thal, Christian Schumann, Steffen Stenger, Alexander Kleger, Günter Lochnit, Benjamin Mayer, Yasser B. Ruiz-Blanco, Markus Hoffmann, Konstantin M. J. Sparrer, Stefan Pöhlmann, Elsa Sanchez-Garcia, Frank Kirchhoff, Manfred Frick, & Jan Münch. **Alpha-1 antitrypsin inhibits TMPRSS2 protease activity and SARS-CoV-2 infection.** *Nature Communications* **2021**.

Aim:

The previous study showed the development of potent peptidomimetic inhibitors for TMPRSS2 which successfully blocked SARS-CoV-2 infection. Since the airway epithelium of the respiratory tract system is the frontline of defense against pathogens, this study aims to screen peptide/protein libraries derived from the respiratory tract to identify endogenous factors that block SARS-CoV-2 infection with prospects for potential COVID-19 therapy.

Contribution:

I carried out the surface plasmon resonance analysis studies, to verify the interaction of the identified protein α_1AT with the protease TMPRSS2. The project was supervised by Prof. Dr. Jan Münch and Prof. Dr. Volker Mailänder.

6.1 Abstract

SARS-CoV-2 is a respiratory pathogen and primarily infects the airway epithelium. As our knowledge about innate immune factors of the respiratory tract against SARS-CoV-2 is limited, we generated and screened a peptide/protein library derived from bronchoalveolar lavage for inhibitors of SARS-CoV-2 spike-driven entry. Analysis of antiviral fractions revealed the presence of α_1-antitrypsin (α_1AT), a highly abundant circulating serine protease inhibitor. Here, we report that α_1AT inhibits SARS-CoV-2 entry at physiological concentrations and suppresses viral replication in cell lines and primary cells including human airway epithelial cultures. We further demonstrate that α_1AT binds and inactivates the serine protease TMPRSS2, which enzymatically primes the SARS-CoV-2 spike protein for membrane fusion. Thus, the acute phase protein α_1AT is an inhibitor of TMPRSS2 and SARS-CoV-2 entry, and may play an important role in the innate immune defense against the novel coronavirus. Our findings suggest that repurposing of α_1AT-containing drugs has prospects for the therapy of COVID-19.

6.2 Introduction

SARS-CoV-2 is mainly transmitted through inhalation of droplets and aerosols and subsequent infection of cells of the respiratory tract.[326] In many cases, infection is limited to the upper airways resulting in no or mild symptoms. Severe disease is caused by viral dissemination to the lungs ultimately resulting in acute respiratory distress syndrome, cytokine storm, multi-organ failure, septic shock, and death.[290, 291] The airway epithelium acts as a frontline defense against respiratory pathogens via the mucociliary clearance and its immunological functions.[327] The epithelial lining fluid is rich in innate immunity peptides as well as proteins with antibacterial and antiviral activity, such as lysozyme, lactoferrin or defensins.[328] Currently, our knowledge about innate immune defense mechanisms against SARS-CoV-2 in the respiratory tract is limited.

To identify endogenous antiviral peptides and proteins, we previously generated peptide/protein libraries from body fluids and tissues and screened the resulting fractions for antiviral factors.[329] This approach allowed to identify novel modulators of HIV-1, CMV, and HSV-2 infection, with prospects for clinical development as antiviral drugs.[329-332] In this work, we set out to identify factors of the respiratory tract that block SARS-CoV-2 infection. Screening a peptide/protein library derived from bronchoalveolar lavage allowed to identify α_1-antitrypsin (α_1AT), a highly abundant circulating serine protease inhibitor, as SARS-CoV-2 entry inhibitor. We show that α_1AT suppresses viral replication in cell lines and human airway epithelial

cultures and that it binds and inactivates the serine protease TMPRSS2, which enzymatically primes the SARS-CoV-2 spike protein for membrane fusion. Hence, α_1AT blocks SARS-CoV-2 entry by inhibiting TMPRSS2 and may play an important role in the innate immune defense against the novel coronavirus. Our findings suggest that repurposing of α_1AT-containing drugs has prospects for treatment of COVID-19.

6.3 Material and methods

Protein–protein docking. As starting structures for the docking simulation we used a homology model of the extracellular fragment of TMPRSS2, based on the structure with PBD-ID: 1z8g which is composed by a SRCR and a Peptidase S1 domain.[46] This model was downloaded from the SWISSMODEL repository under the access link: https://swissmodel.expasy.org/ repository/uniprot/O15393?csm= C05B5531C8A311C7.[333] In addition, we employed the structure of the α_1AT protein with PDB-ID: 3cwm[3]. HADDOCK (2.4) was used to dock the model of TMPRSS2 and the α_1AT inhibitor.[334, 335] This program uses a set of active residues as restraints to the first step of the docking algorithm, where the rotational and translational degrees of freedom of both interacting proteins are explored. The formation of contacts between the active residues is favored during this step of the algorithm. Hence, the residues forming the catalytic triad of TMPRSS2 and their closest neighbors (< 5 Å) were selected as active, while the residues in the serpin loop of α_1AT were also declared as active. During the docking simulation, 10,000 structures were analyzed. After internal clustering and refinement steps, seven models were obtained and the one with the most favorable binding score was selected. Next, we further refined the extracted model by means of a simulated annealing procedure in an explicit water box. The annealing involved 20 ns of heating and cooling ramps exploring a temperature range of 275 –350 K with steps of 5 K, each cycle of the annealing lasted 3 ns. From all cycles, the structures equilibrated at 300 K were sampled and those with minimum potential energy content were used for further analyses.

Molecular dynamics simulations. All simulations were performed using NAMD2.13 and the CHARMM36m force field.[336-339] VMD 1.9.3 was employed for structural analysis and visualization.[340] Molecular dynamics simulations were done with explicit TIP3P water molecules.[341] The simulations were performed at 1 atm and 300 K, with the pressure and temperature controlled via Langevin dynamics simulations.[342, 343] An electrostatic cut-off of 14 Å was used, together with the Particle Mesh Ewald method for the treatment of long-range interactions.[344] The annealing cycles were carried out by gradually changing the system temperature, and that of the Langevin piston for the pressure control, in steps of 5 K within the

ranges mentioned above. The convergence analyses for all the checkpoints are summarized in Tables A5 and A6.

Generation of a peptide/protein library from lungs. 6.5 kg of human lung were obtained from deceased individuals without known diseases from pathology of Ulm University. The organs were frozen at -20 °C and freeze-dried. Lung material was homogenized and peptides and small proteins were extracted using ice-cold acetic acid. Then the mixture was centrifuged at 3500 × g and the supernatant was applied to a 0.45 µm filter. Thereafter, the obtained peptides and proteins were separated by a 30 kDa molecular weight cut-off ultrafiltration step. The filtrate was separated by reversed-phase (RP) chromatography with a Sepax Poly RP300 50 × 300 mm column (Sepax Technologies, Newark DE, USA 260300-30025) at a flow rate of 100 mL/min and a gradient ranging from 0.1% TFA (Merck, 1082621000) in ultrapure water (buffer A) to 0.1% TFA in acetonitrile (J.T.Baker, JT9012-3, buffer B). Reversed-phase chromatographic fractions of 50 mL were collected to constitute the lung peptide bank, from which 1 mL-aliquots (2%) were lyophilized and used for antiviral testing.

Generation of a peptide/protein library from BAL. Clinical samples of bronchoalveolar lavage (BAL) comprising a total of 20 L were collected and immediately frozen for further processing. Peptide/protein extraction was done by acidification with acetic acid to pH 3, followed by centrifugation at 3500 × g and filtration (0.45 µm) of the supernatant. Further, the filtered BAL was subjected to a 30 kDa molecular weight cut-off. Chromatographic fractionation of the filtrate was performed by using a reversed-phase (PS/DVB) Sepax Poly RP300 HPLC 30 × 250 mm column (Sepax Technologies, Newark DE, USA 260300-30025), at a flow rate of 55 mL/min and a gradient ranging from 0.1% TFA in ultrapure water (buffer A) to 0.1% TFA in acetonitrile. Seventy-three reversed-phase chromatographic fractions of 50 mL were collected to constitute the BAL peptide bank, from which 1 mL aliquots (2%) were lyophilized and used for antiviral testing. For further purification of active fractions, a second reversed-phase C18 HPLC 4.6 × 250 mm column (Phenomenex 00G-4605-EO) was used at a flow rate of 0.8 mL/min and a gradient ranging from 0.1% TFA in ultrapure water (buffer A) to 0.1% TFA in acetonitrile, followed again by lyophilization prior to antiviral testing.

Cell culture. Unless stated otherwise, HEK293T cells (ATCC, CRL-3216) were cultivated in DMEM supplemented with 10% fetal calf serum (FCS), 2 mM L-glutamine, 100 U/mL penicillin, and 100 mg/mL streptomycin. Caco-2 cells (ATCC, HTB-37) were cultivated in DMEM supplemented with 10% FCS, 2 mM glutamine, 100 U/mL penicillin and 100 mg/µL streptomycin, 1× non-essential amino acids (NEAA) and 1 mM sodium pyruvate. TMPRSS2-expressing Vero E6 cells (kindly provided by the National Institute for Biological Standards and

Control (NIBSC), #100978) were cultivated in DMEM supplemented with 10% fetal calf serum (FCS), 2 mM L-glutamine, 100 U/mL penicillin, 100 mg/mL streptomycin and 1 mg/mL geneticin. Human Small Airway Epithelial Cells (Lonza, CC-2547, batch: 18TL082942, donor: 68 years, female) were cultivated in SAGM™ Small Airway Epithelial Cell Growth Medium (Lonza, CC-3118).

Generation of lentiviral pseudotypes. For generation of lentiviral SARS-CoV-2 pseudoparticles (LV(Luc)-CoV-2-S) 900,000 HEK293T cells were seeded in 2 mL HEK293T medium. The next day, medium was replaced and cells were transfected with a total of 1 µg DNA using polyethyleneimine (PEI). To this end, 2% of pCG1-SARS-2-S (encoding the spike protein of SARS-CoV-2 isolate Wuhan-Hu-1, NCBI reference Sequence YP_009724390.1) were mixed with 98% of pCMVdR8_91 (encoding a replication-deficient lentivirus) and pSEW-Luc2 (encoding a luciferase reporter gene, both kindly provided by Christian Buchholz) in a 1:1 ratio in OptiMEM. Plasmid DNA was mixed with PEI at a DNA:PEI ratio of 1:3 (3 µg PEI per 1 µg DNA), incubated for 20 min at room temperature and added to cells dropwise. At 8 h post transfection, medium was removed, cells were washed with 2 mL of PBS and 2 mL of HEK293T medium with 2.5% FCS were added. At 48 h post transfection, pseudoparticles containing supernatants were harvested and clarified by centrifugation for 5 minutes at 450 × g.

Generation of VSV-based pseudotypes. For generation of VSV-based SARS-CoV-2 pseudoparticles (VSV(Luc_eGFP)-CoV-2-S), HEK293T were seeded in 30 mL HEK293T medium in a T175 cell culture flask. The next day, cells were transfected with a total 44 µg pCG1-SARS-2-S using PEI. Plasmid DNA and PEI were mixed in 4.5 mL of OptiMEM at a 2:1 ratio (2 µg PEI per 1 µg DNA), incubated for 20 min at room temperature and added to cells dropwise. 24 h post transfection, medium was replaced and cells were transduced with VSV-G-protein pseudotyped VSV encoding luciferase and GFP reporter gene (kindly provided by Gert Zimmer, Institute of Virology and Immunology, Mittelhäusern/Switzerland).[345] At 2 h post transduction, cells were washed three times with PBS and cultivated for 16 h in HEPES-buffered HEK293T medium. Virus containing supernatants were then harvested and clarified by centrifugation for 5 minutes at 450 × g, residual pseudoparticles harboring VSV-G-protein were blocked by addition of anti-VSV-G hybridoma supernatant at 1/10 volume ratio (I1, mouse hybridoma supernatant from CRL-2700; ATCC). Virus stocks were concentrated 10-fold using a 100 kDa Amicon molecular weight cutoff and stored at -80 °C until use.

SARS-CoV-2 strains and propagation. Viral isolate BetaCoV/France/IDF0372/2020 (#014V-03890) and BetaCoV/Netherlands/01/ NL/2020 (#010V-03903) were obtained from the European Virus Archive global and propagated on Vero E6 or Caco-2 cells. To this end, 70–

90% confluent cells in $75\,cm^2$ cell culture flasks were inoculated with SARS-CoV-2 isolate (multiplicity of infection (MOI) of 0.03–0.1) in 3.5 mL serum-free medium. Cells were incubated for 2 h at 37 °C, before adding 20 mL medium containing 15 mM HEPES. Cells were incubated at 37 °C and supernatant harvested when a cytopathic effect (CPE) was visible. Supernatants were centrifuged for 5 minutes at $1000 \times g$ to remove cellular debris, and then aliquoted and stored at -80 °C as virus stocks. Infectious virus titer was determined as plaque-forming units (PFU) on Vero E6 cells, which was used to calculate MOI.

Screening lung and BAL library for inhibitors of SARS-CoV-2 pseudoparticle entry. 10,000 Caco-2 cells were seeded in 100 µL respective medium in a 96-well flat-bottom plate. The next day, medium was replaced by 40 µL of serum-free medium. For screening peptide containing fractions, 10 µL of the solubilized fraction (in dH_2O) were added to cells. Cells were inoculated with 50 µL of infectivity normalized LV(Luc)-CoV2 (or LV(Luc)-no GP control). Transduction rates were assessed by measuring luciferase activity in cell lysates at 48 hours post transduction with a commercially available kit (Promega). Values for untreated controls were set to 100% transduction.

Gel electrophoresis and western blotting. Gel electrophoresis of active fractions was performed on a 4–12% Bis–Tris protein gel (NuPAGE™). Prior to electrophoresis, samples were reduced with 50 mM β-mercaptoethanol and heated for 10 minutes at 90 °C. The gel was either directly stained with Coomassie G-250 (GelCode™ Blue Stain) or blotted on PVDF membranes. Blotted membranes were stained with a polyclonal anti-$α_1$AT antibody (1:1000; Proteintech 16382-1-AP). After incubation with IRDye anti-rabbit secondary antibodies (1:20,000; LiCor), the staining was visualized using an Odyssey Infrared Imager (Licor).

Tryptic in-gel digestion of proteins. Bands of interest were excised and the proteins were digested with trypsin. Tryptic peptides were eluted from the gel slices with 1% trifluoric acid.

Matrix-assisted laser-desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS). MALDI-TOF-MS was performed on an Ultraflex TOF/TOF mass spectrometer (Bruker Daltonics, Bremen) equipped with a nitrogen laser and a LIFT-MS/MS facility. The instrument was operated in the positive-ion reflectron mode using 2.5-dihydroxybenzoic acid and methylendiphosphonic acid as matrix. Sum spectra consisting of 200–400 single spectra were acquired. For data processing and instrument control the Compass 1.4 software package consisting of FlexControl 4.4, FlexAnalysis 3.4 4, Sequence Editor and BioTools 3.2 and ProteinScape 3.1. were used. External calibration was performed with a peptide standard (Bruker Daltonics).

Database search. Proteins were identified by MASCOT peptide mass fingerprint search (http://www.matrixscience.com) using the Uniprot Human database (version 20200226, 210438 sequence entries; $p < 0.05$). For the search, a mass tolerance of 75 ppm was allowed and oxidation of methionine as variable modification was used.

Pseudoparticle inhibition experiments. One day prior to transduction, 10,000 Caco-2 cells were seeded in 100 µL respective medium in a 96-well plate. For the addition of α_1AT prior to infection, medium was replaced by 80 µL of serum-free medium and cells were incubated with serial dilutions of Prolastin (α_1AT) for 0, 1, 2, and 4 h at 37 °C followed by infection with 20 µL of infectivity normalized VSV(Luc)-CoV-2-S pseudoparticles. To investigate whether α_1AT acts post viral entry, cells were inoculated with 20 µL of infectivity normalized VSV(Luc)-CoV-2 pseudoparticles. After 2 h, cells were washed with 100 µL PBS and 100 µL serum-free medium as well as 20 µL of serially diluted α_1AT were added. Transduction rates were assessed by measuring luciferase activity in cell lysates at 16 h post transduction with a commercially available kit (Promega). Values for untreated controls were set to 100% transduction.

TMPRSS2 activity measurement. 20,000 HEK 293T cells were seeded in 100 µL of the respective medium in a 96-well flat-bottom plate. The next day, cells were transfected with 100 ng of TMPRSS2 (addgene 53887, kindly provided by Roger Reeves, Johns Hopkins University, Baltimore, United States) per well using PEI transfection reagent. Briefly, plasmid DNA was mixed with PEI in a 3:1 ratio in serum-free medium, incubated for 20 minutes at room temperature and added to cells dropwise. At 12 h post transfection, medium was removed and 60 µL of PBS were added, followed by serial dilutions of Prolastin or inhibitor control. After incubation for 15 min at 37 °C, 20 µL of protease substrate Boc-Gln-Ala-Arg-AMC were added. Fluorescence intensity was recorded at an excitation wavelength of 380 nm and emission wavelength of 460 nm in 1 minute intervals for 2 h at 37 °C in a Synergy™ H1 microplate reader (BioTek, USA) with Gen 5 3.04 software. For assessing the activity of recombinant human TMPRSS2, 25 µL of serially diluted Prolastin or inhibitor control in assay buffer (50 mM Tris-HCL, 0.154 mM NaCl pH 8.0) were incubated with 25 µL of recombinant TMPRSS2 enzyme for 10 minutes at 37 °C, followed by addition of 20 µM Boc-Gln-Ala-Arg-AMC protease substrate. Fluorescence intensity was measured after 3 h at an excitation wavelength of 360 nm and emission wavelength of 465 nm in a Tecan Genios with Magellan V6.4 software.

Binding free energy calculations. The Central Limit Free Energy Perturbation (CL-FEP) approach was employed to estimate the binding affinities of the complexes of the enzymes TMPRSS2 and trypsin with α_1AT.[346] The structure of the trypsin-α_1AT complex reported with PDB-ID: 1OPH was used for the calculations, while the complex TMPRSS2-α_1AT was modeled

via docking and subjected to further structural refinement, see computational details in supplementary materials.[347] CL-FEP allows an unbiased end-state calculation of free energy changes directly from explicit solvent simulations. The sampling was performed with molecular dynamics simulations of the individual proteins, the enzyme-α_1AT complexes and the bulk solvent. The proteins were sampled under wall-type restraints on their bound-state conformations, which permits to focus the sampling on the most relevant states to the binding energy. The force constant for these restraints was $100\,kcal\,mol^{-1}\,Å^{-2}$ and a restraint-free range of $5\,Å$ was allowed with respect to the initial conformation. $100\,ns$ of sampling simulations were collected for each simulation box, with the energy sampled every $5\,ps$. The simulation boxes and molecular dynamics setup were obtained using the CLFEP-GUI web server (https://clfep.zmb.uni-due.de/). The CL-FEP analyses were performed with ten checkpoints containing increasing fractions of the total energy samples, an oversampling ratio of $osr = 20$ was fixed to bring the free energy variance until the level of (kT), and the second-order cumulant estimator (C2) was evaluated to determine the free energy changes at each checkpoint.[290] The final estimate corresponds to the average among the converged checkpoints. The error is obtained as the standard deviation among the individual estimations. The convergence analyses for all the checkpoints are summarized in Tables S6.2 and S6.3.[346]

Surface plasmon resonance. SPR assay was carried out on a Reichert SR7000 system using SPRAutolink 1.1.9 software. To measure the binding affinity between TMPRSS2 and α_1AT, TMPRSS2 (Creative BioMart) was immobilized on a 11-mercaptoundecanoic acid (MUA) functionalized gold sensor chip through covalent coupling of amine groups. The surface was activated using 0.4 M N-hydroxysuccinimide (NHS) and 0.1 M 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) coupling reagents and the TMPRSS2 protein was coupled at a concentration of $12.5\,\mu g/mL$ in $10\,mM$ acetate buffer at $pH = 5.5$. Residual activated carboxylic groups were blocked using a 1 M ethanolamine solution at $pH = 8.5$. α_1AT was dissolved at different concentrations (0, 0.2, 0.5, 1, 2, 5, 10, and 20 µM) in HEPES buffer ($10\,mM$ HEPES, $150\,mM$ NaCl, 0.05% Tween-20; $pH = 7.5$) and introduced with a flow rate of $10\,\mu L/min$. The data were analyzed with Prism 7 using non-linear regression fit and the value for the equilibrium dissociation constant K_D was obtained by fitting the response at equilibrium against the concentration. Successful immobilization of TMPRSS2 was verified using a TMPRSS2 antibody (rabbit anti-human IgG, ThermoFischer Scientific, PA5-14264) dissolved in HEPES buffer to a concentration of $0.2\,mg/mL$ (1:10 of $2\,mg/mL$ stock). As a negative control, TMPRSS2 antibody solution was passed through a channel with no TMPRSS2 protein immobilization.

Cytotoxicity assay. To assess cytotoxicity of α_1AT and camostat mesylate (CM), 10,000 Caco-2 cells were seeded in 100 µL medium in a 96-well flat-bottom plate. The next day, medium was replaced by 80 µL of serum-free Caco-2 medium and cells were treated with serial dilutions of Prolastin, CM or DMSO as solvent control for CM. After 48 h, cell viability was assessed by measuring ATP levels in cells lysates with a commercially available kit (CellTiter-Glo®, Promega).

Virus-induced cytopathic effect analysis. To quantify SARS-CoV-2 wildtype infection, virus-induced cell death was inferred from remaining cell viability determined by MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) assay. To this end, 20,000 TMPRSS2-expressing Vero E6 or 30,000 Caco-2 cells were seeded in 96-well plates in 100 µL respective medium. The next day, medium was replaced with serum-free medium and the respective compound of interest was added. After incubation for 1 h at 37 °C the cells were infected with a MOI of 0.001 of the SARS-CoV-2BetaCoV/France/IDF0372/2020 or BetaCoV/Netherlands/01/NL/2020, in a total volume of 180 µL. 2 (TMPRSS2-expresing Vero E6) or 3 (Caco-2) days post infection, infection was quantified by detecting remaining metabolic activity. To this end, 36 µL of CellTiter 96® AQueous One Solution Reagent (Promega G3580) were added to the medium and incubated for 3 h at 37 °C. Then, optical density (OD) was recorded at 620 nm using an Asys Expert 96 UV microplate reader (Biochrom)with DigiRead 1.26 software. To determine infection rates, sample values were subtracted from untreated control and untreated control set to 100%. For time of addition experiments, 18,000 TMPRSS2-expressing Vero E6 cells were seeded in 100 µL respective medium. The next day, respective compounds were added 1 h prior to, simultaneously with, or 2, 4 and 24 h post infection with SARS-CoV-2 (BetaCoV/France/IDF0372/2020) at a MOI of 0.01. Viral inoculum was removed 2 h post infection and cells were washed twice with PBS. Cell viability was assessed at 2 days post infection by MTS assay as described above.

Plaque assay. 700,000 TMPRSS2-expressing Vero E6 cells were seeded in 12-well plates in 1 mL of the respective medium. The next day, medium was refreshed and cells were treated with respective compound 1 h prior to, simultaneously with or 1.5 h post infection with SARS-CoV-2 (BetaCoV/Netherlands/01/NL/2020). After infection, cells were overlaid with respective medium supplemented with cellulose. At 2 days post infection, cells were fixed in 4% paraformaldehyde in PBS for 45 minutes, washed one with PBS and stained with 0.5% crystal violet in 0.1% triton in H_2O for 30 minutes. Staining solution was removed and cells were washed three times with water. Virus-induced plaques were quantified in ImageJ 1.53c by assessing number of pixels occupied by plaques.

Infection of small airway epithelial cells. 30,000 SAEC were seeded in 100 µL of respective medium in 96-well flat bottom plate. The next day, medium was refreshed and Prolastin was added 1 h prior to or 3 h and 24 h post infection with SARS-CoV-2 (BetaCoV/France/IDF0372/2020) at a MOI of 1 for 3 h. Cells were washed three times with PBS and fresh medium was added. Immediately after the wash step (0 dpi) and at 6 dpi, supernatant samples were collected for qPCR analysis. To analyze virus replication by RT-qPCR, RNA was isolated using the QIAamp Viral RNA Mini Kit (Qiagen) according to the manufacturer's instructions. For RT-qPCR, samples were thawed and 5 µL of lysate used as sample in a 20 µL reaction using Fast Virus 1-Step Mastermix (Thermo Fisher, # 4444436), 0.5 µM Taqman primers targeting SARS-CoV-2-*ORF1b-nsp14* and 0.25 µM probe (for primer sequences see Table 6.1).

Table 6.1. List of primers and their sequences.

Name	Sequence
SARS-CoV-2-*ORF1b-nsp14* forward primer	5'-TGGGGYTTTACRGGTAACCT-3'
SARS-CoV-2-*ORF1b-nsp14* reverse primer	5'-AACRCGCTTAACAAAGCACTC-3'
probe	5'-FAM-GCAAATTGTGCAATTTGCGG-TAMRA-3'

Cycling conditions in were as follows: 1 cycle of reverse transcription (50 °C, 300 s) and RT-inactivation (95 °C, 20 s); 40 cycles of denaturation (95 °C, 5 s) and extension (60 °C, 30 s) in a Step One Plus qPCR cycler (AppliedBiosystems) with Step One Software 2.3. RNA isolated from virus stocks of BetaCoV/France/IDF0372/2020 with copy numbers previously estimated using synthetic SARS-CoV-2 RNA standard (Twist Bioscience, #102024) was used to determine genome copies from Ct values.

Generation of human airway epithelial cells. Differentiated air–liquid interface cultures of human airway epithelial cells (HAECs) were generated from primary human basal cells isolated from airway epithelia. Cells were expanded in a T75 flask (Sarstedt) in Airway Epithelial Cell Basal Medium supplemented with Airway Epithelial Cell Growth Medium SupplementPack (both Promocell). Growth medium was replaced every two days. Upon reaching 90% confluence, HAECs were detached using DetachKIT (Promocell) and seeded into 6.5 mm Transwell filters (Corning Costar). Filters were precoated with Collagen Solution (StemCell Technologies) overnight and irradiated with UV light for 30 minutes before cell seeding for collagen crosslinking and sterilization. 3.5×10^4 cells in 200 µL growth medium

were added to the apical side of each filter, and an additional 600 µL of growth medium was added basolaterally. The apical medium was replaced after 48 h. After 72–96 h, when cells reached confluence, the apical medium was removed and basolateral medium was switched to differentiation medium. Differentiation medium consisted of a 1:1 mixture of DMEM-H and LHC Basal (Thermo Fisher) supplemented with Airway Epithelial Cell Growth Medium SupplementPack and was replaced every 2 days. Air-lifting (removal of apical medium) defined day 0 of air–liquid interface (ALI) culture, and cells were grown at ALI conditions until experiments were performed at day 25–28. To avoid mucus accumulation on the apical side, HAEC cultures were washed apically with PBS for 30 minutes every 3 days from day 14 onwards.

SARS-CoV-2 infection of HAECs. Immediately before infection, the apical surface of HAECs grown on Transwell filters were washed three times with 200 µL PBS to remove accumulated mucus. Then, 10 µM of α_1AT or 5 µM remdesivir were added into the basal medium and onto the apical surface. Cells were infected with 9.25×10^2 plaque-forming units (PFU) of SARS-CoV-2 (BetaCoV/France/IDF0372/2020). After incubation for 2 h at 37 °C, viral inoculum was removed and cells were washed three times with 200 µL PBS and again cultured at the air–liquid interface. At 1, 2, and 3 days post infection, cells were fixed for 30 minutes in 4% paraformaldehyde in PBS, permeabilized for 10 minutes with 0.2% saponin and 10% FCS in PBS, washed twice with PBS and stained with anti-SARS-CoV-2 spike (ab252690, Abcam) and anti-alpha-tubulin (MA1-8007, Thermo Scientific) diluted 1:300 to 1:500, respectively, in PBS, 0.2% saponin and 10% FCS over night at 4 °C. Subsequently, cells were washed twice with PBS and incubated for 1 h at room temperature in PBS, 0.2% saponin and 10% FCS containing AlexaFluor 488-labeled anti-rabbit and AlexaFluor 647-labeled anti-rat secondary antibody, respectively (all 1:500; Thermo Scientific) and DAPI + phalloidin AF 405 (1:5000; Thermo Scientific). Images were taken on an inverted confocal microscope (Leica TCS SP5, Leica Microsystems, Leica application suite version 2.7.3.9723) using a ×40 lens (Leica HC PL APO CS2 40 × 1.25 OIL). Images for the blue (DAPI), green (AlexaFluor 488) and far-red (AlexaFluor 647) channels were taken in sequential mode using appropriate excitation and emission settings that were kept constant for all the acquisitions. For quantification, randomly chosen locations in each filter were selected and z-stacks were acquired. A maximum z projection was performed and anti-SARS-CoV-2 positive cells per area (0.15 mm^2) were visually identified and counted.

6.4 Results

Identification of α$_1$AT as SARS-CoV-2 inhibitor. To discover endogenous antiviral peptides and proteins, we extracted polypeptides from 6.5 kg of homogenized human lung or 20 liters of pooled bronchoalveolar lavage (BAL), and separated them by chromatographic means. The corresponding fractions were added to human epithelial colorectal carcinoma (Caco-2) cells and the cells were inoculated with luciferase expressing lentiviral pseudoparticles carrying the SARS-CoV-2 spike protein.[2] None of the fractions of the lung library suppressed infection (Figure 6.1).
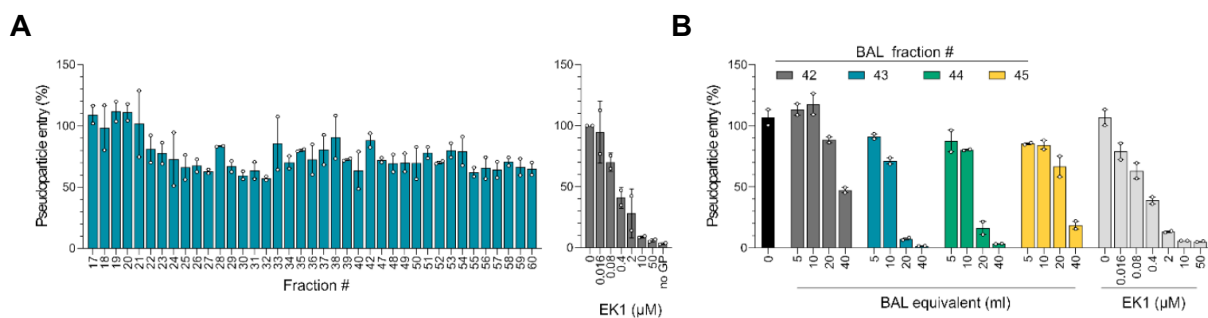
Figure 6.1. Screening peptide/protein libraries for anti-SARS-CoV-2 activity. **A** Caco-2 cells were treated with protein containing fractions of a lung library or EK1 as inhibitor control and transduced with lentiviral SARS-CoV-2 spike pseudoparticles. **B** Caco-2 cells were treated with serial dilutions of bronchoalveolar lavage library (BAL) fractions with anti-SARS-CoV-2 activity and transduced with lentiviral SARS-CoV-2 spike pseudoparticles. Transduction rates in a and b were determined 2 days post addition of pseudoparticles by measuring luciferase activities in cell lysates. The means ± SEM from n=2 independent experiments (**A**) or n=1 experiment (**B**) are shown, each experiment was performed in biological duplicates.

In contrast, fractions 42–45 of the BAL library prevented SARS-CoV-2 spike driven entry with an efficiency comparable to that of 10 µM EK1, a coronavirus spike-specific peptide fusion inhibitor (Figure 6.2).[348] Titration of BAL fractions 42–45 onto Caco-2 cells confirmed dose-dependent inhibition of SARS-CoV-2 spike pseudoparticles (Figure 6.1).

A

B
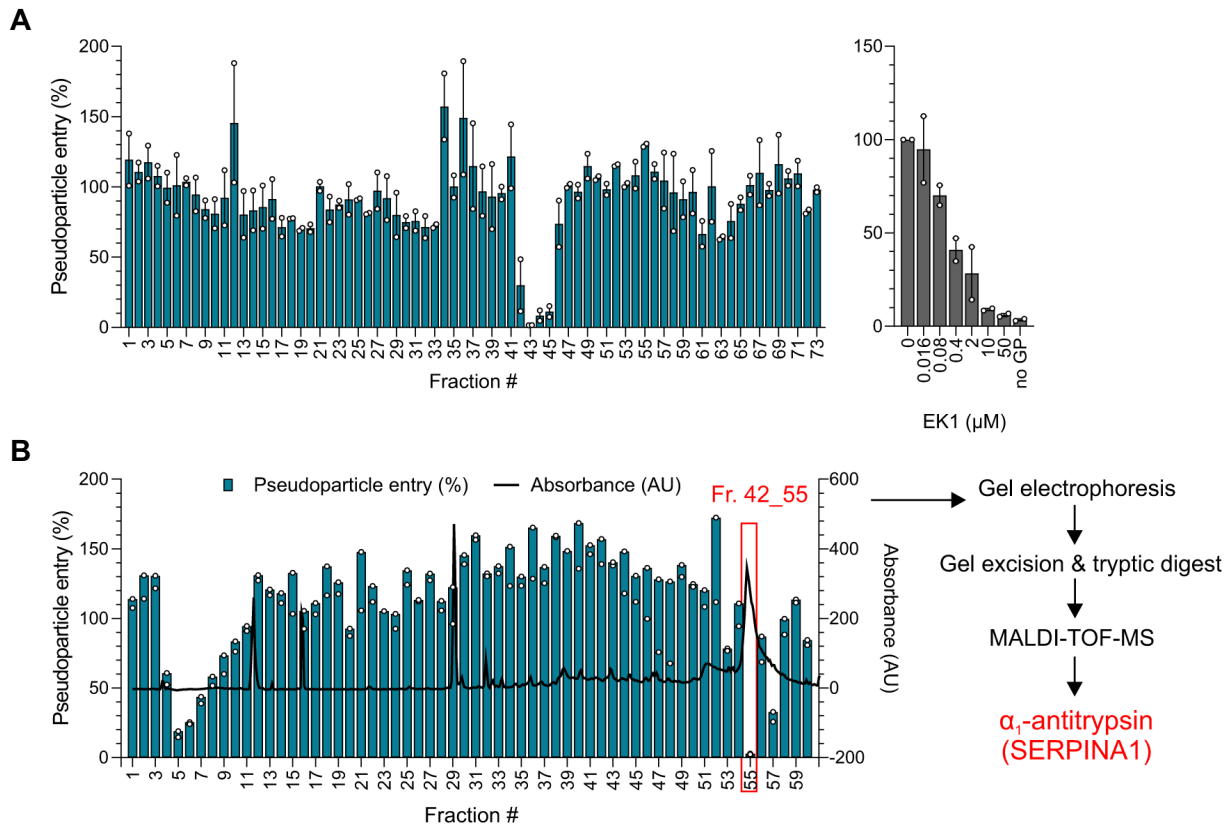
Figure 6.2. A Caco-2 cells were treated with peptide/protein containing fractions of the bronchoalveolar lavage library (or EK1 peptide as inhibitor control) and transduced with luciferase-encoding lentiviral SARS-CoV-2 spike pseudoparticles. **B** Caco-2 cells were treated with subfractions of mother fraction 42 (see **A**) of the bronchoalveolar lavage library and transduced with lentiviral SARS-CoV-2 spike pseudoparticles. Blue columns represent pseudoparticle entry and black line absorbance at 280 nm of the corresponding fraction. Transduction rates in **A** and **B** were determined 2 days after the addition of pseudoparticles by measuring luciferase activities in cell lysates. The means ± SEM from $n = 2$ (**A**) or individual datapoints from $n = 1$ (**B**) independent experiments are shown, each performed in biological duplicates. The active fraction 42_55 (red box) was analyzed via gel electrophoresis, gel excision, tryptic digest, and MALDI-TOF-MS to identify α_1AT, a serine protease inhibitor (serpin).

To isolate the antiviral factor responsible for blocking spike-driven entry, the BAL mother fraction 42 was further separated chromatographically and the resulting sub-fractions analyzed for antiviral activity. As shown in Figure 6.2, sub-fractions 42_3 to 42_8 and 42_57 reduced, and sub-fraction 42_55 almost completely prevented host cell entry of SARS-CoV-2 spike pseudoparticles. Analysis of these inhibitory fractions by gel electrophoresis revealed distinct protein bands in sub-fractions 42_55 and 42_57, whereas no protein was detectable in sub-fractions 42_5 to 42_7 (Figure 6.3).
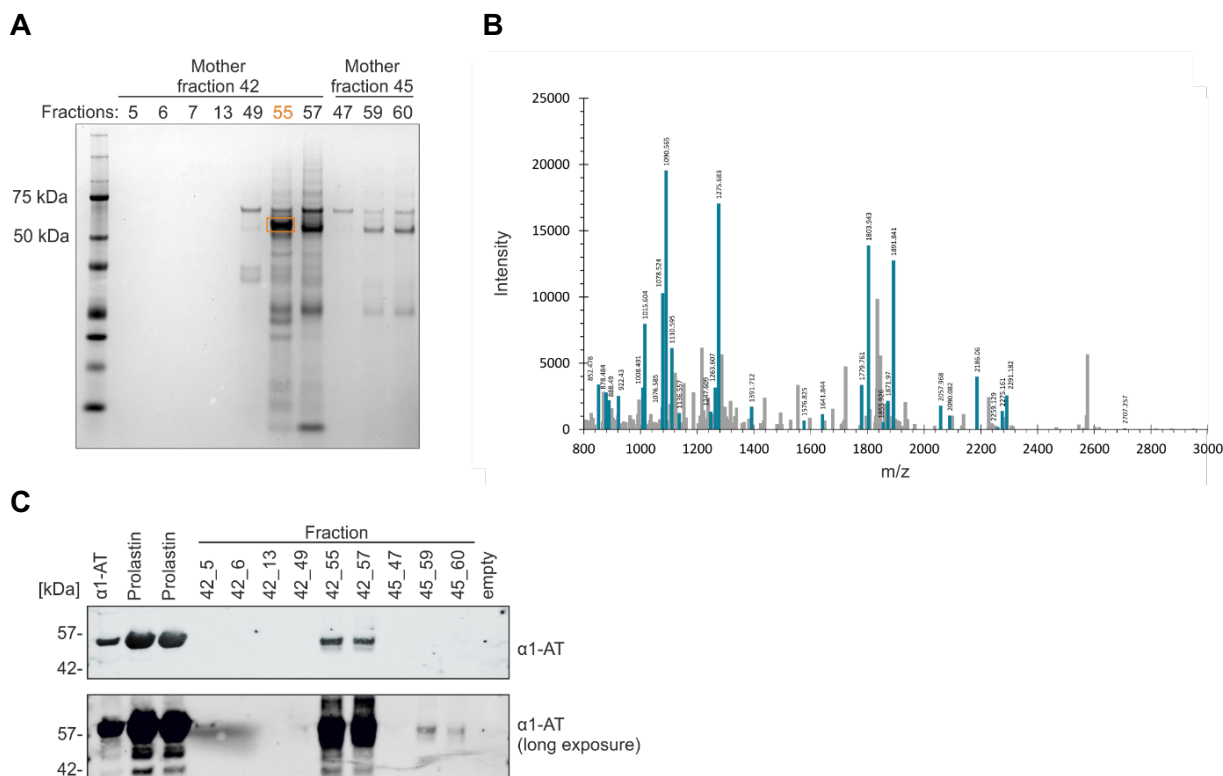
Figure 6.3. Analysis of BAL fractions with anti-SARS-CoV-2 activity. **A** Gel electrophoresis of active fractions was performed on a 4-12% Bis-Tris protein gel. Prior to electrophoresis, samples were reduced by addition of 50 mM β-mercaptoethanol and heated for 10 minutes at 90 °C. Protein gel was stained with Coomassie G-250. The boxed band in fraction 42_55 was cut out and subjected in to an in-gel tryptic digest. **B** MALDI-TOF-MS of the tryptic digest of boxed band from fraction 42_55. Mass signals assigned to the identified α_1AT are shown in blue. Also see Table A7. **C** For WB analysis, 1 µg of α_1AT purified from human serum obtained from Merck or from two Prolastin batches and 1 µL of active BAL fractions were blotted on PVDF membranes and stained with anti-α_1AT antibody. Analysis of fractions by SDS-PAGE, MALDI-TOF and Western Blot were conducted once as Bronchoalveolar lavage (BAL) fractions were limited.

The most active sub-fraction 42_55 contained a prominent band at ~ 52 kDa, which was also present in other active fractions but hardly found in neighboring fractions showing no antiviral activity (e.g., 42_49) (Figure 6.3). This band was excised from the gel, digested with trypsin and subjected to MALDI-TOF-MS revealing a 100% sequence identity to α_1-antitrypsin (SERPINA1) (Figure 6.3 and Table A7), a 52 kDa protease inhibitor.[349] The presence of α_1-antitrypsin (α_1AT) in inhibitory fractions 42_55 and 42_57 was confirmed by western blot analysis with an α_1AT-specific antibody (Figure 6.2). α_1AT belongs to the serine protease inhibitor (serpin) superfamily and protects lung tissue from digestive enzymes released by immune cells, in particular neutrophil elastase.[350] The serpin has a reference range in blood of 0.9–2 mg/mL (corresponding to ~ 17–38 µM), but the concentration can rise 4- to 5-fold upon

acute inflammation.[351] α_1AT purified from donor blood is also available as pharmaceutical product (e.g., Prolastin) for intravenous substitution therapy of α_1AT deficiency, a hereditary disorder that leads to chronic uncontrolled tissue breakdown in the lower respiratory tract.[352]

α_1AT inhibits SARS-CoV-2 spike mediated pseudovirus entry. To test whether the serpin indeed inhibits SARS-CoV-2, Caco-2 cells were exposed to Prolastin, a pharmaceutical preparation of α_1AT, or camostat mesylate (CM), a small molecule inhibitor of the SARS-CoV-2 spike priming protease TMPRSS2.[3, 183] α_1AT and CM both suppressed SARS-CoV-2 spike pseudoparticle entry with half-maximal inhibitory concentrations (IC$_{50}$) of ~38.5 µM for α_1AT, and ~0.05 µM for CM, respectively (Figure 6.5). Cell viability assays showed that α_1AT displayed no cytotoxic effects at concentrations of up to 160 µM (8.3 mg/mL), whereas CM reduced cell viability at concentrations of 200 µM, due to DMSO in the stock (Figure 6.4).
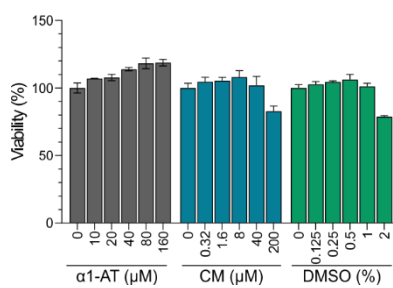
Figure 6.4. Cell viability assay. To assess cytotoxicity of Prolastin (α_1AT, grey) and camostat mesylate (CM, blue), Caco-2 cells were treated with serial dilutions of the compounds (and DMSO as solvent control for CM, green). After 48 h, cell viability was assessed by measuring ATP levels in cells lysates with a commercially available kit (CellTiter-Glo®, Promega). The mean ± SD of n=1 experiment performed in biological triplicates is shown.

The antiviral activity of α_1AT and CM was specific for the coronavirus spike because entry of pseudoparticles carrying the G-protein of VSV was not affected (Figure 6.4). In fact, time-of-addition experiments demonstrated that α_1AT prevented single-round SARS-CoV-2 spike pseudovirus entry only if added prior (1–4 h) to or during infection but not if added 2 h post infection (Figure 6.4). Taken together, these data show that α_1AT specifically targets SARS-CoV-2 spike-driven entry.
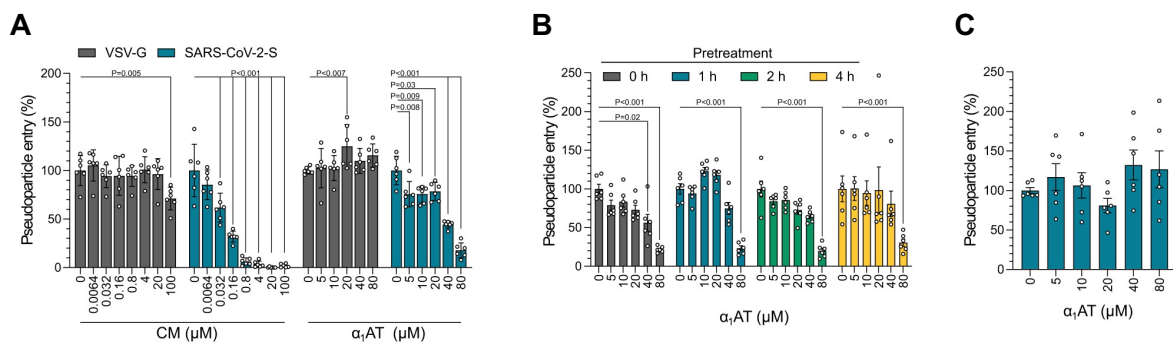
Figure 6.5. A Prolastin (α_1AT) and the control small molecule inhibitor camostat mesylate (CM) were added to Caco-2 cells 1 h prior to transduction of cells with rhabdoviral SARS-CoV-2 spike (blue) or VSV-G pseudoparticles (gray). **B** Caco-2 cells were treated for the indicated hours with Prolastin (α_1AT) and were then transduced with rhabdoviral SARS-CoV-2 spike pseudoparticles. **C** Prolastin (α_1AT) was added 2 h post transduction of Caco-2 cells with rhabdoviral SARS-CoV-2 spike pseudoparticles. Transduction rates in **A**–**C** were determined at 16 h after addition of pseudoparticles by measuring luciferase activities in cell lysates. The mean ± SEM from $n = 2$ experiments in biological triplicates are shown (2-way ANOVA with Dunett´s multiple comparison test).

α_1AT inhibits SARS-CoV-2 infection and replication. To determine whether α_1AT inhibits not only spike pseudoparticles but also wild-type SARS-CoV-2, we examined its activity against two SARS-CoV-2 isolates from France (bearing the spike variant D614) and the Netherlands (bearing the fitter spike variant G614).[353, 354] For this, we assessed survival rates of TMPRSS2-expressing Vero E6 cells infected in the absence or presence of EK1, CM or α_1AT by MTS assay. In the absence of drugs, infection by both SARS-CoV-2 isolates resulted in virus-induced cytopathic effects (CPE) and reduced cell viability by ~80% (Figure 6.6 and Figure 6.7). Microscopic evaluation revealed the absence of CPE in the presence of high concentrations of EK1, CM or α_1AT, and MTS assay confirmed a concentration-dependent inhibition of cell death and viral replication by EK1 and CM (Figure 6.6 and Figure 6.7) with average IC_{50} values against both SARS-CoV-2 isolates of 2.8 µM for EK1, and 3.6 µM for CM, respectively (Figure 6.6). α_1AT inhibited the French SARS-CoV-2 isolate with an IC_{50} of 21.2 µM (1.1 mg/mL), and the Dutch strain with an IC_{50} of 17.3 µM (0.9 mg/mL) (Figure 6.6). Almost complete rescue of cell viability was observed at α_1AT concentrations of 40–80 µM (Figure 6.6 and Figure 6.7). A similar antiviral activity of α_1AT against both SARS-CoV-2 isolates was determined in Caco-2 cells (Figure 6.7).
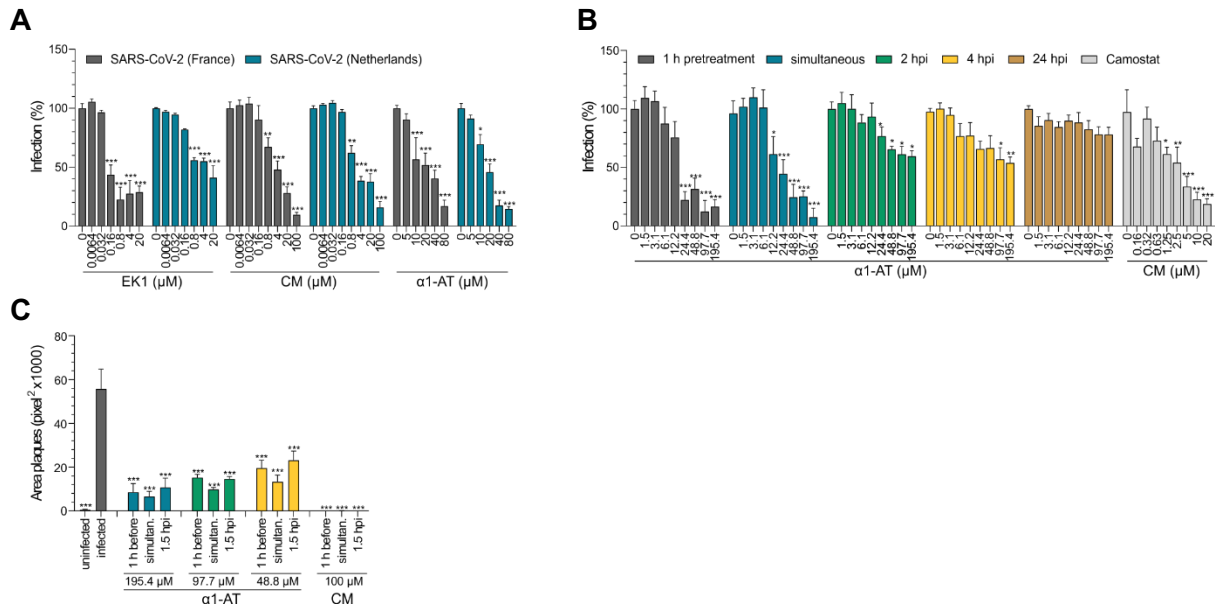
Figure 6.6. A TMPRSS2-expressing Vero E6 cells were treated with Prolastin (α₁AT), EK1 or camostat mesylate (CM) for 1 h, and infected with SARS-CoV-2 isolates either from France (gray) or the Netherlands (blue) at a MOI of 0.001. Virus-induced cytopathic effects were assessed at 2 days post infection by MTS assay (see Figure 6.7 A, B for raw and cell viability data). **B** TMPRSS2-expressing Vero E6 cells were treated with Prolastin (α₁AT) at indicated timepoints prior to, simultaneously with or post infection with SARS-CoV-2 at a MOI of 0.001. Camostat mesylate (CM) control was added 1 h prior to infection. Infection rates were assessed at 2 days post infection by MTS assay (see also Figure 6.8 for raw and cell viability data). **C** TMPRSS2-expressing Vero E6 cells were treated with Prolastin (α₁AT) or CM 1 h prior to, simultaneously with or 1.5 h post infection with SARS-CoV-2. At 1.5 h post infection cellulose overlay was performed. At 2 days post infection, cells were stained with crystal violet (see Figure 6.9) and plaque areas were quantified. The mean ± SEM from *n* = 1 (A, EK1 and CM) or *n* = 2 independent experiments in biological triplicates (A, α₁AT, B) or duplicates (C) and quadruplicates (C, infected) are shown. (2-way ANOVA with Dunett´s multiple comparison test (A, B), ordinary one-way ANOVA with Dunett´s multiple comparison test (C)).
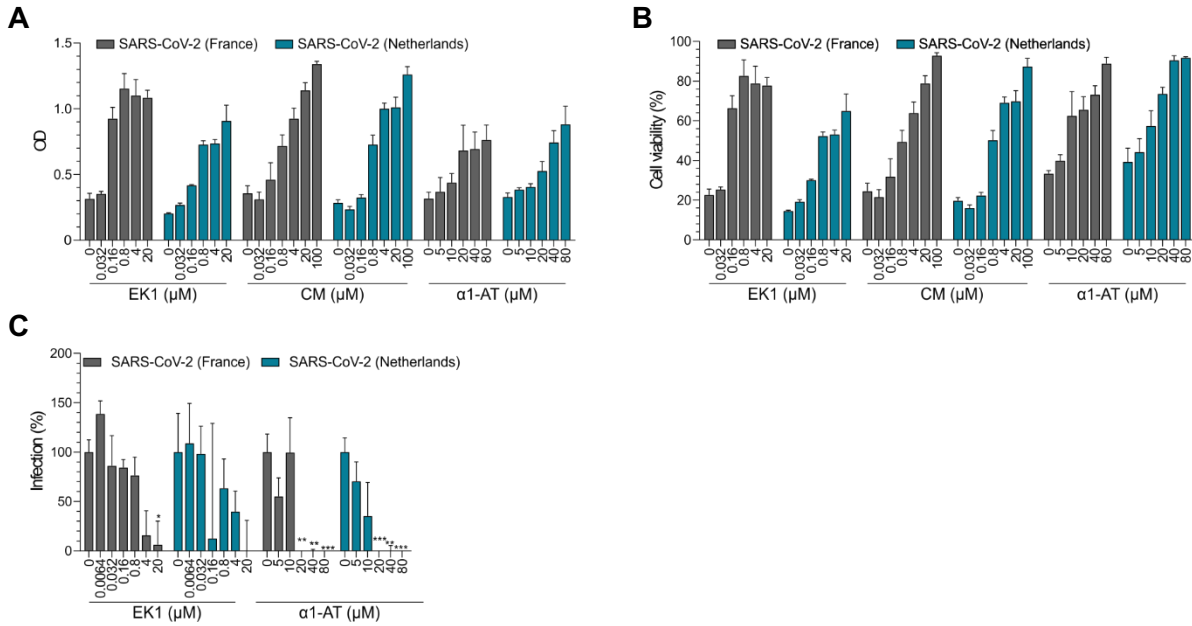
Figure 6.7. α_1AT inhibits SARS-CoV-2 infection. **A** TMPRSS2-expressing Vero E6 cells were exposed to indicated concentrations of EK1, camostat mesylate (CM)and α_1AT and then infected with a French (grey) and Dutch (blue) SARS-CoV-2 isolate. 2 days later, virus-induced cytopathic effect (CPE) was determined by MTS assay. Optical density (OD) was recorded at 620 nm using an Asys Expert 96 UV microplate reader (Biochrom). Graph shows raw data (OD_{620}). **B** Cell viability rates as calculated from A). To determine infection rates, sample values were subtracted from untreated control and untreated control set to 100% (see Figure 6.3a). **C** Caco-2 cells were exposed to EK1 and α_1AT and infected with two SARS-CoV-2 isolates. At day 2, viral CPE was quantified by MTS assay and data were processed as described above. Values shown in A are the means ± SEM from n=1(CM and EK1) or n=2 independent experiments (α_1AT) performed in biological triplicates, values in C show the means ± SEM from n=1 experiment in biological triplicates.

We next performed time-of-addition experiments and found that α_1AT most effectively inhibited SARS-CoV-2 replication when added 1 h prior to or simultaneously with infection (Figure 6.6, Figure 6.8), confirming pseudovirus data (Figure 6.5). α_1AT also suppressed SARS-CoV-2 if added 2 or 4 h post infection (Figure 6.3 and Figure 6.8). Considering the 10 h time span of the viral life cycle, these data suggest that the serpin inhibits spreading virus infection, i.e., infection of uninfected target cells by progeny virus.[355] Accordingly, no antiviral effect of α_1AT was observed when added 24 h post infection, when already two replication cycles have been completed (Figure 6.3 and Figure 6.8).
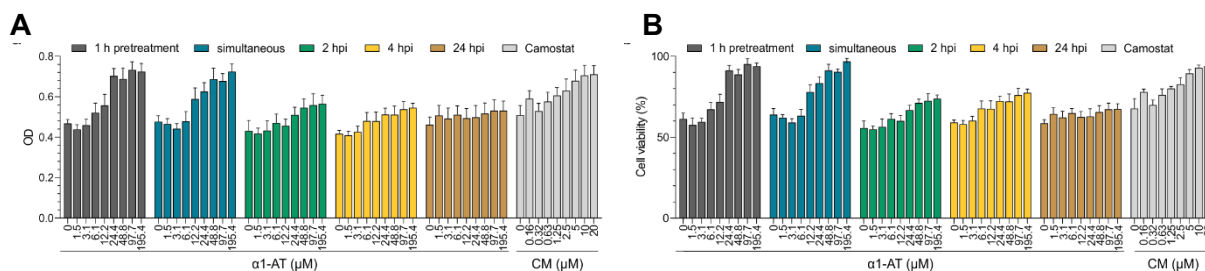
Figure 6.8. α_1AT inhibits SARS-CoV-2 infection and replication. **A** TMPRSS2-expressing Vero E6 cells were treated with Prolastin (α_1AT) at indicated time points prior to, simultaneously with or post infection with SARS-CoV-2. Camostat mesylate (CM) control was added 1 h prior to infection. 2 days post infection, virus-induced cytopathic effect (CPE) was determined by MTS assay. Optical density (OD) was recorded at 620 nm using an Asys Expert 96 UV microplate reader (Biochrom). Graph shows raw data (OD_{620}). **B** Cell viability as calculated from **A**. To determine infection rates, sample values were subtracted from untreated control and untreated control set to 100% (see Figure 6.6). The means ± SEM from n=2 independent experiments in biological triplicates are shown.

To investigate whether α_1AT may inhibit cell-to-cell spread of SARS-CoV-2, a plaque assay with cellulose containing medium (which prevents cell-free viral replication) was performed (Figure 6.9) and plaque sizes were quantified (Figure 6.6). α_1AT concentrations of 50 to 200 µM blocked viral spread most effectively when the serpin was added 1 h prior to infection (54–87% reduction) or simultaneously with infection (75–89% reduction). When adding α_1AT 1.5 h post infection, spread was still inhibited by 49–83%, demonstrating effective inhibition of SARS-CoV-2 infection and cell-to-cell viral spread by the serpin.

A

B
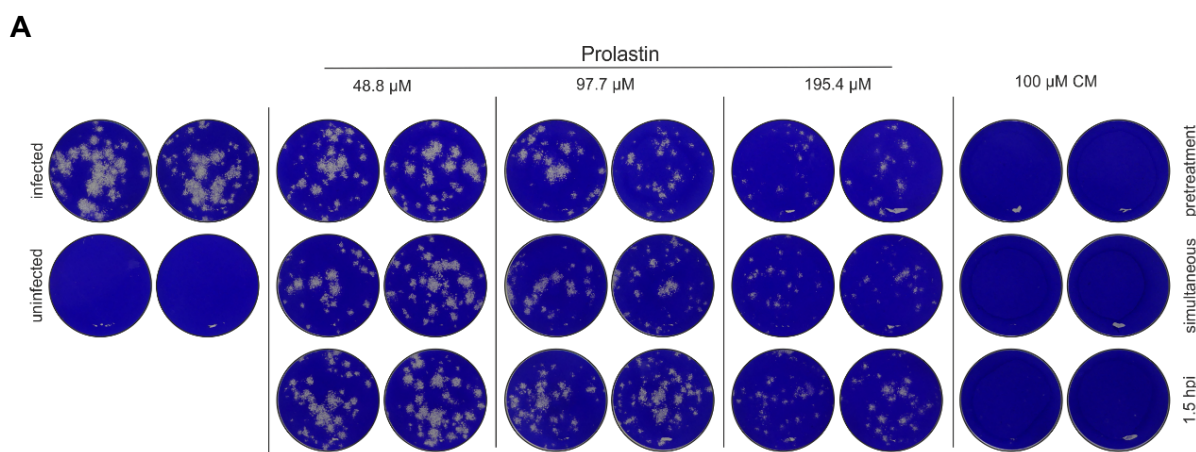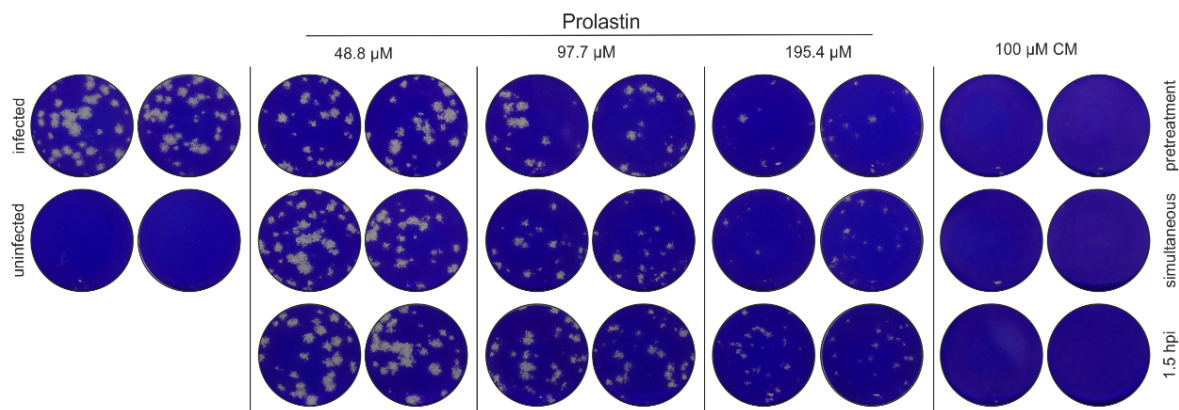
Figure 6.9. α_1AT reduces SARS-CoV-2 replication. **A/B** TMPRSS2-expressing Vero E6 cells were treated with Prolastin (α_1AT) or CM 1 h prior to, simultaneously with or 1.5 h post infection (hpi) with SARS-CoV-2. At 1.5 h post infection cellulose overlay was performed. Two days post infection, cells were stained with crystal violet to visualize virus-induced plaques. Images show plaques at 2 days post infection from n=2 independent experiments (A and B) each performed in duplicates or quadruplicates (infected). Two representative replicates of infected control are shown.

α_1AT inhibits SARS-CoV-2 replication in primary human airway cells. To corroborate the antiviral activity of α_1AT in human primary target cells, we used small airway epithelial cells (SAECs) that support low level SARS-CoV-2 replication. For this, SAECs pretreated with buffer only (PBS), 80 µM of α_1AT, or 100 µM of CM were infected with a high dose (MOI of 1) of SARS-CoV-2. Furthermore, α_1AT was added 3 and 24 h post infection to PBS pretreated cells. Cells were cultivated in the presence of the inhibitors for 6 days, and then viral genome copies in supernatants were quantified by RT-qPCR. As shown in Figure 6.10, CM and α_1AT present during infection reduced viral titers by ~92 and 83%, respectively. α_1AT that was added 3 and 24 h post infection also inhibited viral replication, albeit to a lesser extend (67 and 58% reduction, respectively). We next analyzed antiviral activity of α_1AT in fully differentiated primary human airway epithelial cells (HAECs) grown at the air–liquid interface. HAECs derived from two donors were treated with 10 µM (0.5 mg/mL) of α_1AT and then exposed to SARS-CoV-2. As control, we used 5 µM of remdesivir, which has previously been shown to suppress coronavirus replication in HAECs.[356] At days 1, 2, and 3 post infection cells were fixed and stained with antibodies against SARS-CoV-2 spike, and α-tubulin as marker for ciliated cells at the apical surface (Figure 6.10).[357, 358] In infected, PBS-treated HAECs, SARS-CoV-2 spike expression was readily detectable, mostly in neighboring ciliated cells, and increased between day 2 and 3, demonstrating productive infection and viral spread in the epithelia. Spike expression levels in α_1AT and remdesivir treated cultures were greatly reduced

at days 2 and 3 in both donors (Figure 6.10 and Figure 6.11). Thus, α_1AT suppresses SARS-CoV-2 infection of HAECs.
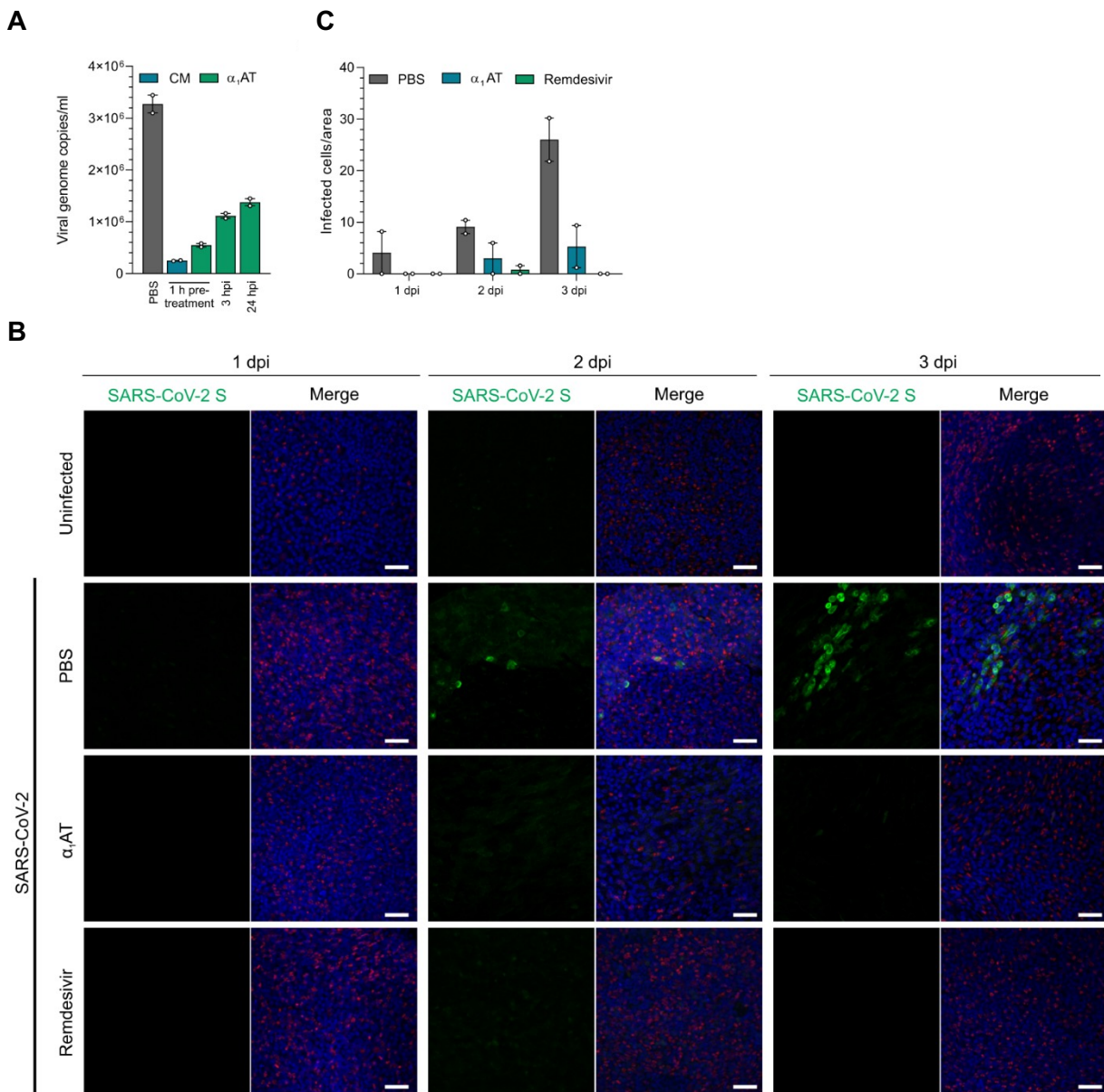
Figure 6.10. A Small airway epithelial cells (SAECs) were treated with 100 µM of Prolastin (α_1AT, green) 1 h prior to, or 3 and 24 h post infection (hpi) with SARS-CoV-2 at a MOI of 1. 100 µM CM (blue) added 1 h prior to infection served as control. Immediately after the inoculum was removed (0 dpi) and at day 6 post infection, supernatants were harvested and subjected to RT-qPCR specific for SARS-CoV-2 ORF1b nsp14. Virus titers from day 0 were subtracted from titers at day 6. The means of technical duplicates of one experiment performed in triplicates are shown. **B** The apical and basal site of human airway epithelial cells (HAEC) grown at air–liquid interface was exposed to PBS, α_1AT (10 µM or 0.5 mg/mL) and remdesivir (5 µM) and then inoculated with SARS-CoV-2 (9.25×10^2 PFU) for 2 h. Cells were fixed at day 1, 2, and 3 post infection, stained with DAPI (cell nuclei, blue), a SARS-CoV-2 specific spike antibody (SARS-CoV-2 S, green) and an α-tubulin-specific antibody (red). Images shown are derived from one donor and represent maximum projections of serial sections along the basolateral to apical cell axis. Scale bar: 50 µm. **C** Number of infected cells per area in mock- (PBS, gray), α_1AT- (blue)

or remdesivir-treated (green), SARS-CoV-2 infected HAECs. Values represent the mean number of infected HAECs from 2 donors at 5 random spots per culture, treatment and day ± SEM. For more images, see Figure 6.11.

A

B
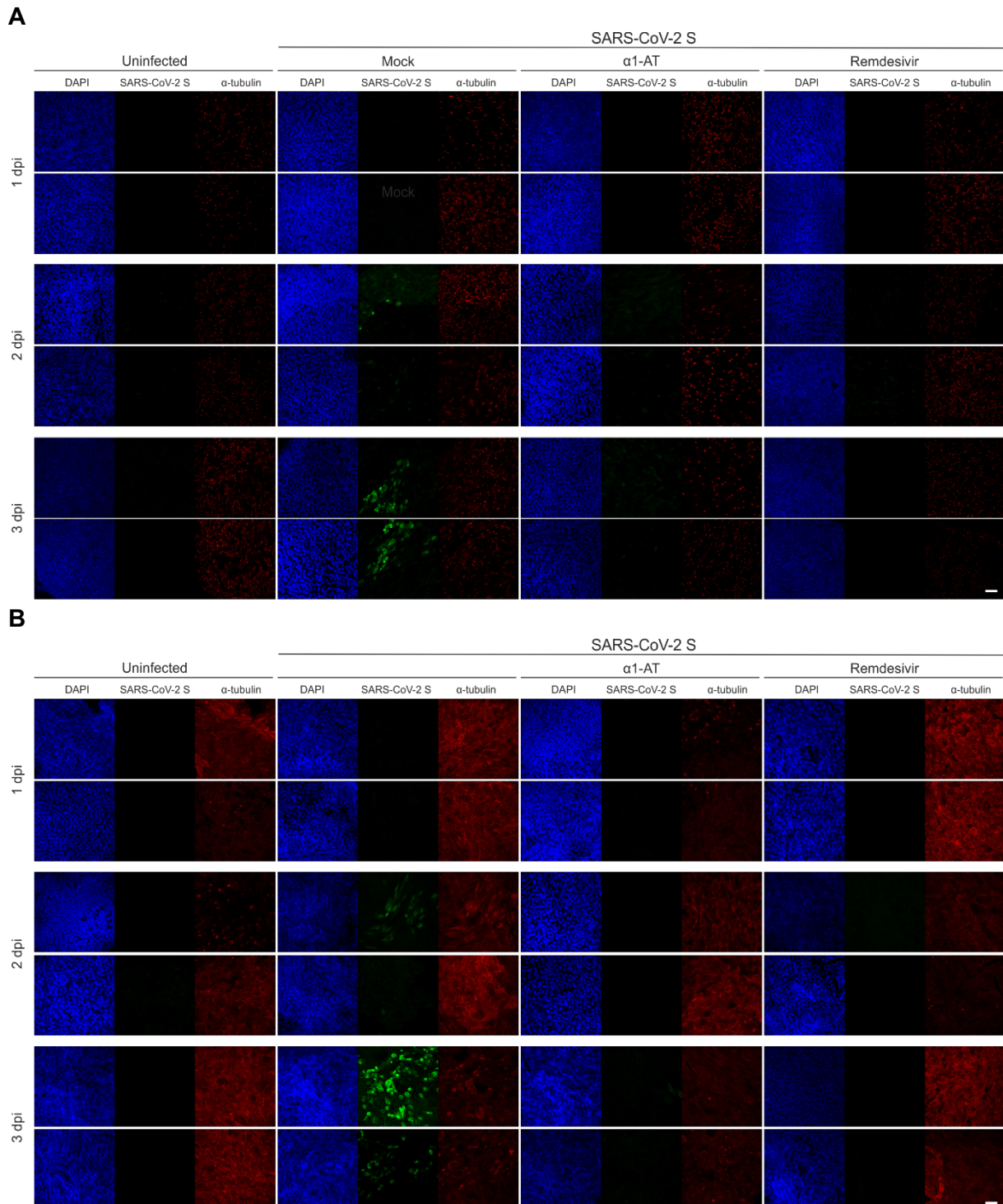
Figure 6.11. α1AT inhibits SARS-CoV-2 replication in primary human airway cells. **A/B** The apical and basal site of human airway epithelial cells (HAEC) derived from two donors (A and B) was exposed to PBS, α1AT (10 μM or 0.5 mg/mL) and remdesivir (5 μM) and then inoculated with SARS-CoV-2. Cells were fixed at days 1, 2, and 3 days post infection, stained with DAPI (cell nuclei, blue), a SARS-CoV-2

specific spike antibody (S, green) and an α-tubulin-specific antibody (red). Images shown here represent maximum projections of serial sections along the basolateral to apical cell axis. Scale bar: 50 μm. Analysis was conducted for two cultures per donor and condition.

α₁AT binds and inhibits TMPRSS2 protease activity. Finally, we set out to explore the mechanism underlying SARS-CoV-2 inhibition by α₁AT. A recent preprint publication suggests that α₁AT may suppress TMPRSS2, the spike priming protease, similar to camostat mesylate.[359] First, we established a computational model of the Michaelis complex of α₁AT and TMPRSS2 using protein–protein docking calculations complemented by structural refinement and observed a calculated binding free energy of −10.2 ± 2.2 kcal/mol (Figure 6.13). We applied this modeling to the complex of α₁AT (Pittsburgh variant, M358R) and trypsin (S195A) and observed a calculated binding free energy of −10 ± 1.6 kcal/mol (Figure 6.12).

$\Delta G_{exp} = -11.33$ kcal/mol
$\Delta G_{calc} = -10 \pm 1.6$ kcal/mol
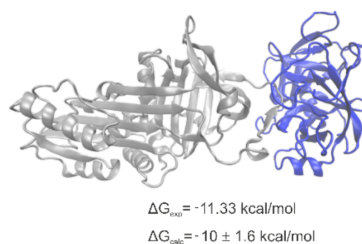
Figure 6.12. Michaelis complex of α₁AT (Pittsburgh variant, M358R, grey) and trypsin (S195A, blue). sBased on PDB-ID: 1OPH. ΔG_{exp} and ΔG_{calc} refer to the computationally calculated and experimentally determined binding free energies of the complex, respectively.

This calculated binding free energy is close to the experimentally determined binding free energy of the α₁AT (Pittsburgh variant, M358R) and trypsin (S195A) complex, therefore verifying our modeling approach.[360] The computed structure of the Michaelis complex suggests that this initial step of the reaction mechanism is favored by the interaction of Leu353 and Ala355 of α₁AT (numbers according to PDB-ID: 3cwm29), with a hydrophobic patch of TMPRSS2 formed by Tyr416, Leu419, and Trp461, which is located next to the catalytic triad (His296, Asp345, Ser441) (Figure 6.14). Thus, the anchoring site, defined by the hydrophobic patch, can effectively position and restrain the reactive center loop prior to the cleavage. These structural insights together with the calculated affinity of this complex, support a favorable interaction between TMPRSS2 and α₁AT. To verify a direct interaction of α₁AT with TMPRSS2, surface plasmon resonance analysis was performed. To this end, recombinant TMPRSS2 was immobilized on a metal surface and subjected to increasing doses (0–10 μM) of α₁AT (Figure 6.13).
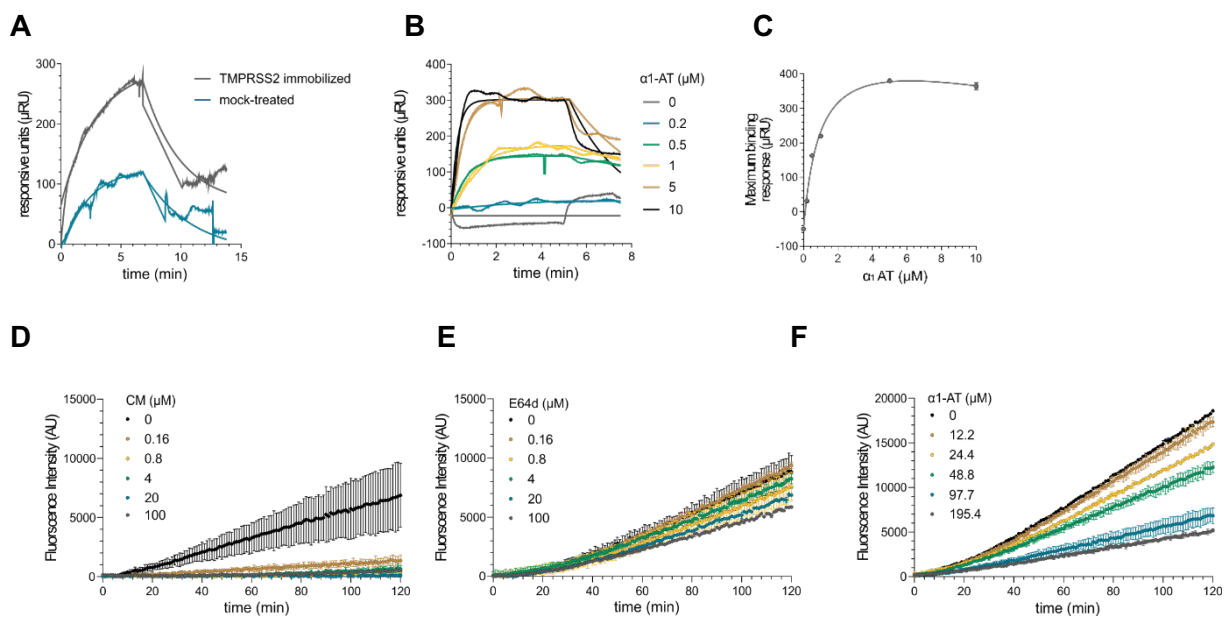
Figure 6.13. α1AT binds the extracellular region of TMPRSS2 and inhibits TMPRSS2 protease activity. **A** Validation of TMPRSS2 immobilization for surface plasmon resonance analysis. A gold sensor chip was either immobilized with recombinant TMPRSS2 (grey) or mock-treated (blue). Binding response of a TMPRSS2 specific antibody (rabbit anti-human IgG, Thermo Fischer Scientific, PA5-14264) was analyzed over time. **B** Surface plasmon resonance analysis of α₁AT-TMPRSS2 interaction. Immobilized recombinant TMPRRS2 was subjected to varying concentrations of α₁AT and binding response was measured over time. **C** Maximum binding response at increasing α₁AT concentrations (determined by nonlinear regression from values in b) allowed determination of dissociation constant K_D of 941 ± 297 nM for α₁AT-TMPRSS2 interaction. **D-F** HEK293T cells were transfected with a TMPRSS2 expression plasmid and treated with Prolastin (α₁AT), CM or E64d followed by incubation with the fluorogenic TMRPSS2 protease substrate Boc-Gln-Ala-Arg-AMC. Graphs shows fluorescence intensities over 2 h that were subtracted by values for mock-transfected HEK293T cells and represent one experiment.

A K_D of 941 ± 297 nM was measured by performing equilibrium analysis which was reached within 5 minutes interaction time of α₁AT with TMPRSS2 (Figure 6.13), proving a physical interaction of the protease inhibitor with the protease. To investigate whether α₁AT not only binds but also inhibits proteolytic activity of TMPRSS2, the protease was overexpressed in HEK293T cells and incubated with α₁AT, CM, or the cysteine protease inhibitor E-64d.[183] Enzymatic activity was assessed by adding a specific substrate that emits fluorescence after proteolytic cleavage. CM but not E-64d suppressed TMPRSS2 activity at sub-micromolar concentrations (Figure 6.13 and Figure 6.14). Interestingly, 100 and 200 μM of α₁AT effectively inhibited cell-associated TMPRSS2 activity (Figure 6.13 and Figure 6.14). Experiments performed with recombinant TMPRSS2 enzyme confirmed effective and dose-dependent inhibition of TMPRSS2 activity by α₁AT at physiologically relevant concentrations of 5–50 μM

(Figure 6.14). These data demonstrate that the abundant serpin α₁AT is an endogenous inhibitor of TMPRSS2 proteolytic activity.
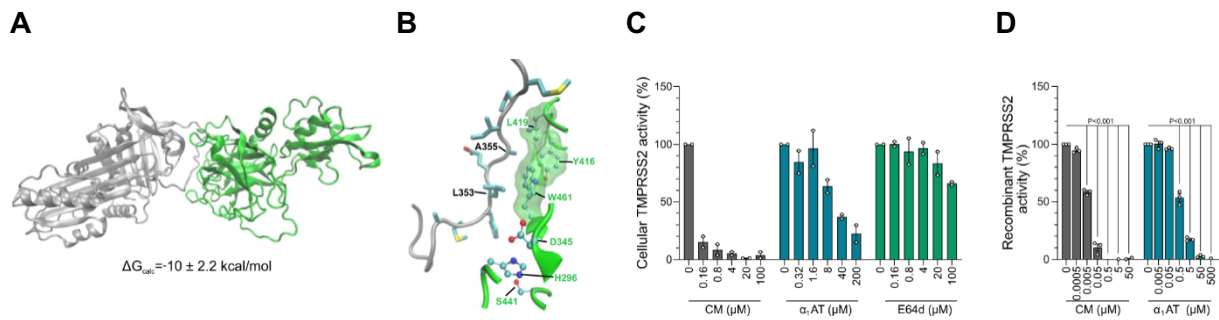
Figure 6.14. A Protein–protein docking analysis of a homology model for the TMPRSS2 extracellular fragment (green, PDB-ID: 1z8g) and α₁AT (gray, PDB-ID: 3cwm) and computationally calculated binding free energy (ΔG_{calc}) of the complex. **B** Detailed view on α₁AT-TMPRSS2 binding interface. The sidechains of α₁AT (gray) residues are represented with sticks, while sidechains of TMPRSS2 (green) are shown with balls and sticks. Hydrogen atoms are omitted for clarity, carbon, oxygen, nitrogen or sulfur atoms of amino acid side chains depicted in light blue, red, dark blue or yellow, respectively. The hydrophobic patch near the TMPRSS2 catalytic triad is highlighted with a green transparent surface. **C** α₁AT inhibits cell-associated TMPRSS2 activity. HEK293T cells were transfected with a TMPRSS2 expression plasmid and treated with Prolastin (α₁AT, blue), camostat mesylate (CM, gray) or E-64d (green) followed by incubation with the fluorogenic TMRPSS2 protease substrate Boc-Gln-Ala-Arg-AMC. Graph shows the relative area under the curve analysis of fluorescence intensities over 2 h that were corrected by values for mock-transfected HEK293T cells. **D** α₁AT inhibits recombinant TMPRSS2 enzyme activity. Recombinant human TMPRSS2 was mixed with Prolastin (α₁AT, blue) or CM (gray) prior to addition of fluorogenic TMPRSS2 protease substrate Boc-Gln-Ala-Arg-AMC, graph shows relative fluorescence intensities after 3 h of incubation. The mean ± SEM of $n = 2$ (C) or $n = 3$ (D) independent experiments in biological duplicates are shown (ordinary one-way ANOVA with Dunett´s multiple comparison test).

6.5 Discussion

This study demonstrates that α_1AT is a potent inhibitor of SARS-CoV-2 which blocks viral replication in cell lines, primary small airway epithelial cells and fully differentiated airway epithelium cultures. α_1AT binds and inhibits TMPRSS2, a transmembrane serine protease that cleaves the SARS-CoV-2 spike protein to enable viral fusion. α_1AT is the most abundant serine protease inhibitor (serpin) in the circulation (0.9–2 mg/mL, 17–38 µM), and levels of this acute phase protein further increase during acute inflammation.[351] We show that plasma-derived α_1AT blocks SARS-CoV-2 infection with IC_{50} values of 10–20 µM, which are well within the physiological range, suggesting that plasma itself may exert anti-SARS-CoV-2 activity. In fact, a recent preprint demonstrates that naïve serum exhibits inhibition of SARS-CoV-2 entry and suggested α_1AT and to a lesser degree α-2-macroglobulin as potential antiviral factors.[361] Thus, α_1AT may serve as natural inhibitor of the novel coronavirus, in particular during acute SARS-CoV-2 infection, when blood α_1AT concentrations increase, as recently shown in a cohort of 40 COVID-19 patients.[362] Further studies to clarify the physiological relevance of α_1AT in SARS-CoV-2 infection and whether inter-individual differences in α_1AT concentrations in blood, lungs, and other organs correlate with viral loads and disease progression are thus highly warranted.

α_1AT was isolated as inhibitor of SARS-CoV-2 from a complex peptide/protein library that was generated from pooled bronchoalveolar lavage (BAL). The corresponding fractions presumably contain all soluble peptides and small proteins present in lung fluid and should allow identification of those innate immune factors that are most relevant for controlling viral infection *in vivo*, at least in the absence of a specific antiviral host immune response. α_1AT-containing fractions 42–45 were more potent in inhibiting SARS-CoV-2 entry than all remaining BAL- and lung-derived fractions, suggesting a relevant role of the serpin *in vivo*. In normal individuals, α_1AT levels range between 10–40 µM in alveolar interstitial fluid, and 2–5 µM in alveolar extracellular lining fluid.[363-365] Assuming that α_1AT concentrations further increase during the acute phase response, the serpin may act as relevant innate immune factor against SARS-CoV-2 in the respiratory tract.

Biomolecular modeling, surface plasmon resonance, and biochemical analysis established α_1AT as inhibitor of TMPRSS2, a cell surface serine protease that is involved in cell-cell and cell-matrix interactions, and in prostate cancer metastasis.[238, 270, 366] TMPRSS2 not only primes the spike protein of SARS-CoV-2 but is also utilized as entry factor by other coronaviruses such as SARS-CoV, MERS-CoV or the common cold corona virus 229E.[272, 303, 367] Furthermore, TMPRSS2 is also the major hemagglutinin-activating protease of influenza A virus in human airways.[208, 368, 369] It is, therefore, not only of great interest to clarify whether α_1AT may act as

broad-spectrum inhibitor against respiratory viral pathogens, but also to evaluate its role in prostate cancer.

α_1AT is an approved drug for treatment of α_1AT deficiency implying repurposing for the therapy of COVID-19. α_1AT deficiency is a hereditary disorder that results in reduced circulating concentrations of α_1AT and consequently a chronic uninhibited breakdown of tissue in the lungs, mainly mediated by neutrophil elastase.[370] Several products containing α_1AT purified from human plasma (such as Prolastin used herein) are approved for decades for intravenous augmentation therapy in patients suffering from α_1AT deficiency. Of note, α_1AT may not only be beneficial in COVID-19 therapy because of its direct antiviral effect by targeting TMPRSS2-mediated SARS-CoV-2 entry, but also by inhibiting neutrophil elastase, which has been proposed to act as alternative spike priming protease and might contribute to pulmonary inflammation in COVID-19.[371-374] Moreover, a recent study revealed that the proinflammatory IL-6 to α_1AT ratio in patients with severe COVID-19 was more than two times higher compared to a pneumonia control cohort.[362] Of note, α_1AT can also be administered via inhalation and at substantially higher doses than those in routine α_1AT deficiency.[375] In some studies, α_1AT has been administered at doses of 120 mg/kg and even 250 mg/kg without causing side effects, resulting in a 5-fold increase of serpin concentration in lung epithelial lining fluid of α_1AT deficient patients.[365, 376, 377] However, whether α_1AT infusion or inhalation allows to reach local concentrations of the serpin that are sufficient to block SARS-CoV-2 in lungs or other organs without causing severe side effects remains to be addressed in clinical studies. Nevertheless, we and others suggest that α_1AT supplementation may have a therapeutic benefit because of its antiviral and anti-inflammatory properties. Thus, rapid evaluation of α_1AT for the treatment of severe COVID-19 disease is highly warranted and four clinical trials (NCT04495101, NCT04385836, NCT04547140, EudraCT: 2020-001391-15) have been initiated to evaluate the therapeutic potential of α_1AT in hospitalized COVID-19 patients.[361, 362, 378, 379]

Conclusion and outlook

In modern medicine novel strategies with therapeutic or diagnostic purposes for tackling diseases are constantly developed, aiming to improve preexisting treatment options with higher efficiencies, reduced side-effects and thereby to increase peoples' life quality.

Type II transmembrane serine proteases are a class of enzymes that are located on the cell-surface and mediate pivotal roles in tissue development and homeostasis. Studies showed that aberrant overactivation of those proteases is frequently observed in cancer and coincides with the degradation of extracellular matrix components, epithelial integrity and downregulation of endogenous serine protease inhibitors HAI-1. In prostate cancer, the type II transmembrane serine protease hepsin is a consistent biomarker indicating malignant tissue transformation and recent results proof hepsin as a promising target. In addition, new findings showed that the closely related type II transmembrane serine protease TMPRSS2 is exploited by respiratory viruses which utilize its proteolytic activity to activate viral surface proteins which mediate host cell infection and spreading of the virus. Within this thesis, new approaches for the development of peptidomimetic compounds which specifically interact with hepsin (Chapter A) and TMPRSS2 (Chapter B) are investigated.

In chapter A new peptidomimetic substrates and inhibitors were developed by using molecular docking studies on the binding cavity of hepsin. Therefore, the binding cavity of hepsin was separated into two subpockets (prime and non prime) and combinatorial peptide libraries were built with a specific pattern and screened against the subpockets. The *in silico* results revealed high scoring peptides combinations for each subpocket which were used as substitutes for a literature known hepsin substrate. A library of octapeptide substrates was assembled and synthesized using solid-phase peptide synthesis with a FRET pair which allow real-life monitoring of enzymatic peptide degradation. Enzymatic testing against hepsin revealed high catalytic efficiencies of the synthesized IQF substrates. To further improve the stability in serum and plasma, D-amino acids were introduced on a selected IQF substrate which revealed combinations with resistance against degradation by diluted serum and plasma while maintaining high cleavability by hepsin compared to the reference peptide substrate. Combination of top prime and top non-prime peptides did not lead to an additive improvement. This result strengthens the observation that neighboring amino acids in a peptide substrate exert strong cooperative effects and that conventional strategies to probe substrate specificities like PS-SCL may not fully uncover optional peptide combinations.

As a follow-up project, the highest scoring peptide combinations targeting the non-prime subsite pocket of hepsin were used as recognition sequences and connected with a reactive serine trap. Enzymatic measurements revealed two compounds having high potency against

hepsin and good selectivity against coagulation proteases. Of note, one compound showed a 15.5-fold selectivity against the closely related protease matriptase, presumably caused by small amino acids in the P2 pocket which could represent the matriptase selectivity determining position. Stability studies on two selected compounds revealed high remaining activity up to 1 day in diluted blood serum.

This chapter tackles the rational design of peptidomimetic substrates and inhibitors for hepsin from a new perspective through the usage of combinatorial peptide libraries. Further research in this area is conceivable as the identified, potentially improved peptide substrates could be used for the development of hepsin responsive peptide-based nanocarriers to selectively target prostate cancer cells. For this purpose, it is of importance to ensure high selectivity against other proteases and therefore, more enzymatic measurements against other, hepsin related proteases are necessarily required to gain broader knowledge about the peptide substrate specificity. Other measures to increase stability might include the use of poly(ethylene glycol) or poly(ethylene phosphate) on the surface of nanocarriers which is commonly used to reduce non-specific cellular uptake. In this regard, a thoroghful analysis of proteins, which adsorb on the surface of the nanocarriers after exposing into blood, as well as regulation of the adsorbed proteins might be helpful to adjust hepsin accessibility and cleavability.

The peptidomimetic inhibitors offer a great platform for further development of pharmaceutical compounds. Our studies prove high activity of selected compounds against hepsin and selectivity against off-target proteases. This opens promising opportunities for further optimization by using D-configurated or synthetic moieties. In addition, substitutions on the heterocycle of the ketobenzothiazole serine trap might enable more specific interactions with the prime recognition site of the catalytic active cleft potentially enabling increased potency or selectivity. For future drug development, attention should be drawn to potential route of application and required pharmacokinetic properties. The perspective is that this pipeline could be used for any other serine proteases as a guideline to improve the efficiency of the discovery of new substrates or inhibitors.

Chapter B describes the identification of novel inhibitors for the transmembrane serine protease TMPRSS2 to block viral infection. Here, a similar pipeline as described above was used to identify tailored peptides for TMPRSS2. At first, based on literature data, information about substrate preferences of TMPRSS2 were obtained and a tripeptide reference binder was created which was used as a starting point for the preparation of peptide libraries. Due to the lack of a protease crystal structure, the libraries were docked against two substitute structures and the top scoring combinations were assembled to a peptidomimetic inhibitor library. Enzymatic testing against TMPRSS2 and off-target proteases matriptase, thrombin and factor

Xa revealed four compounds as potentially promising with high potency and off-target protease selectivity. In further studies, the top compounds were shown to block SARS-CoV-2 spike driven entry into Caco-2 cells and to block infection of Caco-2 cells by SARS-CoV-2 wildtype virus and its variants of concern. Two selected compounds were incubated in blood serum and plasma which showed still high remaining activity after 10 days. The above-described further developments for inhibitors are also conceivable in this area. In addition, the lung being the most attacked organ in COVID-19 might favor a potential inhalative application of TMPRSS2 inhibitors.

Literature

1.   WHO, Prioritizing Diseases for Research and Development in Emergency Contexts. WHO: Genf, Switzerland, 2021.

2.   Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Kruger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N. H.; Nitsche, A.; Muller, M. A.; Drosten, C.; Pohlmann, S., SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020,** *181* (2), 271-280 e8.

3.   Kawase, M.; Shirato, K.; van der Hoek, L.; Taguchi, F.; Matsuyama, S., Simultaneous treatment of human bronchial epithelial cells with serine and cysteine protease inhibitors prevents severe acute respiratory syndrome coronavirus entry. *J Virol* **2012,** *86* (12), 6537-45.

4.   EM, F. J.; F, L.; M, C.; L, M.; M, P.; A, Z.; I, S., Global cancer observatory: cancer today. International Agency for Research on Cancer.

5.   AM, N.; N, H.; M, K.; D, M.; A, B.; M, Y.; J, R.; Z, T.; A, M.; DR, L.; HS, C.; EJ, F.; KA, C. SEER Cancer Statistics Review. https://seer.cancer.gov/csr/1975_2015/.

6.   DeSantis, C. E.; Siegel, R. L.; Sauer, A. G.; Miller, K. D.; Fedewa, S. A.; Alcaraz, K. I.; Jemal, A., Cancer statistics for African Americans, 2016: Progress and opportunities in reducing racial disparities. *CA Cancer J Clin* **2016,** *66* (4), 290-308.

7.   Sung, H.; Ferlay, J.; Siegel, R. L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **2021**.

8.   Steinberg, G. D.; Carter, B. S.; Beaty, T. H.; Childs, B.; Walsh, P. C., Family history and the risk of prostate cancer. *Prostate* **1990,** *17* (4), 337-47.

9.   Brawer, M. K., Prostatic intraepithelial neoplasia: an overview. *Rev Urol* **2005,** *7 Suppl 3*, S11-8.

10.   Epstein, J. I.; Egevad, L.; Amin, M. B.; Delahunt, B.; Srigley, J. R.; Humphrey, P. A.; Committee, G., The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System. *Am J Surg Pathol* **2016,** *40* (2), 244-52.

11.   Bailar, J. C.; Mellinger, G. T.; Gleason, D. F., Survival rates of patients with prostatic cancer, tumor stage, and differentiation--preliminary report. *Cancer Chemother Rep* **1966,** *50* (3), 129-36.

12.   Wittekind, C., [2010 TNM system: on the 7th edition of TNM classification of malignant tumors]. *Pathologe* **2010,** *31* (5), 331-2.

13.   Vickers, A. J.; Cronin, A. M.; Aus, G.; Pihl, C. G.; Becker, C.; Pettersson, K.; Scardino, P. T.; Hugosson, J.; Lilja, H., A panel of kallikrein markers can reduce unnecessary biopsy for

prostate cancer: data from the European Randomized Study of Prostate Cancer Screening in Göteborg, Sweden. *BMC Med* **2008,** *6*, 19.

14.	Bruyère, F.; Amine Lakmichi, M., [PSA interest and prostatitis: literature review]. *Prog Urol* **2013,** *23* (16), 1377-81.

15.	Kindermann, W.; Lehmann, V.; Herrmann, M.; Loch, T., [Influencing of the PSA concentration in serum by physical exercise (especially bicycle riding)]. *Urologe A* **2011,** *50* (2), 188-96.

16.	Schröder, F. H.; Carter, H. B.; Wolters, T.; van den Bergh, R. C.; Gosselaar, C.; Bangma, C. H.; Roobol, M. J., Early detection of prostate cancer in 2007. Part 1: PSA and PSA kinetics. *Eur Urol* **2008,** *53* (3), 468-77.

17.	Ferro, M.; Buonerba, C.; Terracciano, D.; Lucarelli, G.; Cosimato, V.; Bottero, D.; Deliu, V. M.; Ditonno, P.; Perdonà, S.; Autorino, R.; Coman, I.; De Placido, S.; Di Lorenzo, G.; De Cobelli, O., Biomarkers in localized prostate cancer. *Future Oncol* **2016,** *12* (3), 399-411.

18.	Früherkennung, Diagnose und Therapie der verschiedenen Stadien des Prostatakarzinoms. Deutsche Gesellschaft für Urologie e. V. (Hrsg.): 2019; Vol. Kurzversion 5 .1 – Mai 2 019.

19.	Bundesausschuss, G., Prostatakrebsscreening mittels PSA -Test. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen: Köln, 2019; pp 1- 86.

20.	Beard, J.; Eddington, S.; Bowman, N.; Cole, A.; Brien, T.; Johnson, B. P., Abstract 6450: Circulating hepsin as a novel serum biomarker in prostate cancer patients. *Cancer Research* **2020,** *80* (16 Supplement), 6450.

21.	Ferlay, J.; Soerjomataram, I.; Dikshit, R.; Eser, S.; Mathers, C.; Rebelo, M.; Parkin, D. M.; Forman, D.; Bray, F., Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* **2015,** *136* (5), E359-86.

22.	Center, M. M.; Jemal, A.; Lortet-Tieulent, J.; Ward, E.; Ferlay, J.; Brawley, O.; Bray, F., International variation in prostate cancer incidence and mortality rates. *Eur Urol* **2012,** *61* (6), 1079-92.

23.	Shah, M.; Raju, K.; Kumar G, H., Revisiting Prostate Biopsy with 2014 ISUP Modified Gleason Score and Gleason Grade - A Cross Section Study. Biomedical Research and Therapy: 2018; Vol. 5, pp 2918-2925.

24.	Romero-Otero, J.; García-Gómez, B.; Duarte-Ojeda, J. M.; Rodríguez-Antolín, A.; Vilaseca, A.; Carlsson, S. V.; Touijer, K. A., Active surveillance for prostate cancer. *Int J Urol* **2016,** *23* (3), 211-8.

25.	Huggins, C.; Hodges, C. V., Studies on prostatic cancer. *Cancer res* **1941,** *1*, 297.

26.   Harris, W. P.; Mostaghel, E. A.; Nelson, P. S.; Montgomery, B., Androgen deprivation therapy: progress in understanding mechanisms of resistance and optimizing androgen depletion. *Nat Clin Pract Urol* **2009,** *6* (2), 76-85.

27.   Moreira, D. M.; Howard, L. E.; Sourbeer, K. N.; Amarasekara, H. S.; Chow, L. C.; Cockrell, D. C.; Pratson, C. L.; Hanyok, B. T.; Aronson, W. J.; Kane, C. J.; Terris, M. K.; Amling, C. L.; Cooperberg, M. R.; Freedland, S. J., Predicting Time From Metastasis to Overall Survival in Castration-Resistant Prostate Cancer: Results From SEARCH. *Clin Genitourin Cancer* **2017,** *15* (1), 60-66.e2.

28.   Yvon, A. M.; Wadsworth, P.; Jordan, M. A., Taxol suppresses dynamics of individual microtubules in living human tumor cells. *Mol Biol Cell* **1999,** *10* (4), 947-59.

29.   Petrylak, D. P.; Tangen, C. M.; Hussain, M. H.; Lara, P. N.; Jones, J. A.; Taplin, M. E.; Burch, P. A.; Berry, D.; Moinpour, C.; Kohli, M.; Benson, M. C.; Small, E. J.; Raghavan, D.; Crawford, E. D., Docetaxel and estramustine compared with mitoxantrone and prednisone for advanced refractory prostate cancer. *N Engl J Med* **2004,** *351* (15), 1513-20.

30.   Saad, F.; Hotte, S. J., Guidelines for the management of castrate-resistant prostate cancer. *Can Urol Assoc J* **2010,** *4* (6), 380-4.

31.   Teo, M. Y.; Rathkopf, D. E.; Kantoff, P., Treatment of Advanced Prostate Cancer. *Annu Rev Med* **2019,** *70*, 479-499.

32.   Sweeney, C. J.; Chen, Y. H.; Carducci, M.; Liu, G.; Jarrard, D. F.; Eisenberger, M.; Wong, Y. N.; Hahn, N.; Kohli, M.; Cooney, M. M.; Dreicer, R.; Vogelzang, N. J.; Picus, J.; Shevrin, D.; Hussain, M.; Garcia, J. A.; DiPaola, R. S., Chemohormonal Therapy in Metastatic Hormone-Sensitive Prostate Cancer. *N Engl J Med* **2015,** *373* (8), 737-46.

33.   Kyriakopoulos, C. E.; Chen, Y. H.; Carducci, M. A.; Liu, G.; Jarrard, D. F.; Hahn, N. M.; Shevrin, D. H.; Dreicer, R.; Hussain, M.; Eisenberger, M.; Kohli, M.; Plimack, E. R.; Vogelzang, N. J.; Picus, J.; Cooney, M. M.; Garcia, J. A.; DiPaola, R. S.; Sweeney, C. J., Chemohormonal Therapy in Metastatic Hormone-Sensitive Prostate Cancer: Long-Term Survival Analysis of the Randomized Phase III E3805 CHAARTED Trial. *J Clin Oncol* **2018,** *36* (11), 1080-1087.

34.   Tannock, I. F.; de Wit, R.; Berry, W. R.; Horti, J.; Pluzanska, A.; Chi, K. N.; Oudard, S.; Théodore, C.; James, N. D.; Turesson, I.; Rosenthal, M. A.; Eisenberger, M. A.; Investigators, T., Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer. *N Engl J Med* **2004,** *351* (15), 1502-12.

35.   Scher, H. I.; Morris, M. J.; Stadler, W. M.; Higano, C.; Basch, E.; Fizazi, K.; Antonarakis, E. S.; Beer, T. M.; Carducci, M. A.; Chi, K. N.; Corn, P. G.; de Bono, J. S.; Dreicer, R.; George, D. J.; Heath, E. I.; Hussain, M.; Kelly, W. K.; Liu, G.; Logothetis, C.; Nanus, D.; Stein, M. N.; Rathkopf, D. E.; Slovin, S. F.; Ryan, C. J.; Sartor, O.; Small, E. J.; Smith, M. R.; Sternberg, C. N.; Taplin, M. E.; Wilding, G.; Nelson, P. S.; Schwartz, L. H.; Halabi, S.; Kantoff, P. W.; Armstrong, A. J.; 3, P. C. C. T. W. G., Trial Design and Objectives for Castration-Resistant Prostate Cancer: Updated Recommendations From the Prostate Cancer Clinical Trials Working Group 3. *J Clin Oncol* **2016,** *34* (12), 1402-18.

36.   Leytus, S. P.; Loeb, K. R.; Hagen, F. S.; Kurachi, K.; Davie, E. W., A novel trypsin-like serine protease (hepsin) with a putative transmembrane domain expressed by human liver and hepatoma cells. *Biochemistry* **1988,** *27* (3), 1067-74.

37.   Tsuji, A.; Torres-Rosado, A.; Arai, T.; Le Beau, M. M.; Lemons, R. S.; Chou, S. H.; Kurachi, K., Hepsin, a cell membrane-associated protease. Characterization, tissue distribution, and gene localization. *J Biol Chem* **1991,** *266* (25), 16948-53.

38.   Zhukov, A.; Hellman, U.; Ingelman-Sundberg, M., Purification and characterization of hepsin from rat liver microsomes. *Biochim Biophys Acta* **1997,** *1337* (1), 85-95.

39.   Somoza, J. R.; Ho, J. D.; Luong, C.; Ghate, M.; Sprengeler, P. A.; Mortara, K.; Shrader, W. D.; Sperandio, D.; Chan, H.; McGrath, M. E.; Katz, B. A., The structure of the extracellular region of human hepsin reveals a serine protease domain and a novel scavenger receptor cysteine-rich (SRCR) domain. *Structure* **2003,** *11* (9), 1123-31.

40.   Vu, T. K.; Liu, R. W.; Haaksma, C. J.; Tomasek, J. J.; Howard, E. W., Identification and cloning of the membrane-associated serine protease, hepsin, from mouse preimplantation embryos. *J Biol Chem* **1997,** *272* (50), 31315-20.

41.   Rawlings, N. D.; Waller, M.; Barrett, A. J.; Bateman, A., MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* **2014,** *42* (Database issue), D503-9.

42.   Ménard, R.; Storer, A. C., Oxyanion hole interactions in serine and cysteine proteases. *Biol Chem Hoppe Seyler* **1992,** *373* (7), 393-400.

43.   Berg, J. M.; Tymoczko, J. L.; Stryer, L., Katalytische Strategien. In *Stryer Biochemie*, Berg, J. M.; Tymoczko, J. L.; Stryer, L., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2013; pp 257-291.

44.   Millies, B., *Entwicklung neuer nicht-kompetitiver Inhibitoren flaviviraler NS2B/NS3-Proteasen*. Johannes Gutenberg-Universität Mainz: 2019.

45.   Schechter, I.; Berger, A., On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **1967,** *27* (2), 157-62.

46.   Herter, S.; Piper, D. E.; Aaron, W.; Gabriele, T.; Cutler, G.; Cao, P.; Bhatt, A. S.; Choe, Y.; Craik, C. S.; Walker, N.; Meininger, D.; Hoey, T.; Austin, R. J., Hepatocyte growth factor is a preferred in vitro substrate for human hepsin, a membrane-anchored serine protease implicated in prostate and ovarian cancers. *Biochem J* **2005,** *390* (Pt 1), 125-36.

47.   Beliveau, F.; Desilets, A.; Leduc, R., Probing the substrate specificities of matriptase, matriptase-2, hepsin and DESC1 with internally quenched fluorescent peptides. *FEBS J* **2009,** *276* (8), 2213-26.

48.   Poręba, M.; Szalek, A.; Kasperkiewicz, P.; Drąg, M., Positional scanning substrate combinatorial library (PS-SCL) approach to define caspase substrate specificity. *Methods Mol Biol* **2014,** *1133*, 41-59.

49.   Ng, N. M.; Pike, R. N.; Boyd, S. E., Subsite cooperativity in protease specificity. *Biol Chem* **2009,** *390* (5-6), 401-7.

50.   Boyd, S. E.; Kerr, F. K.; Albrecht, D. W.; de la Banda, M. G.; Ng, N.; Pike, R. N., Cooperative effects in the substrate specificity of the complement protease C1s. *Biol Chem* **2009,** *390* (5-6), 503-7.

51.   Fagerberg, L.; Hallström, B. M.; Oksvold, P.; Kampf, C.; Djureinovic, D.; Odeberg, J.; Habuka, M.; Tahmasebpoor, S.; Danielsson, A.; Edlund, K.; Asplund, A.; Sjöstedt, E.; Lundberg, E.; Szigyarto, C. A.; Skogs, M.; Takanen, J. O.; Berling, H.; Tegel, H.; Mulder, J.; Nilsson, P.; Schwenk, J. M.; Lindskog, C.; Danielsson, F.; Mardinoglu, A.; Sivertsson, A.; von Feilitzen, K.; Forsberg, M.; Zwahlen, M.; Olsson, I.; Navani, S.; Huss, M.; Nielsen, J.; Ponten, F.; Uhlén, M., Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics* **2014,** *13* (2), 397-406.

52.   Dhanasekaran, S. M.; Barrette, T. R.; Ghosh, D.; Shah, R.; Varambally, S.; Kurachi, K.; Pienta, K. J.; Rubin, M. A.; Chinnaiyan, A. M., Delineation of prognostic biomarkers in prostate cancer. *Nature* **2001,** *412* (6849), 822-6.

53.   Ganesan, R.; Kolumam, G. A.; Lin, S. J.; Xie, M. H.; Santell, L.; Wu, T. D.; Lazarus, R. A.; Chaudhuri, A.; Kirchhofer, D., Proteolytic activation of pro-macrophage-stimulating protein by hepsin. *Mol Cancer Res* **2011,** *9* (9), 1175-86.

54.   Owen, K. A.; Qiu, D.; Alves, J.; Schumacher, A. M.; Kilpatrick, L. M.; Li, J.; Harris, J. L.; Ellis, V., Pericellular activation of hepatocyte growth factor by the transmembrane serine proteases matriptase and hepsin, but not by the membrane-associated protease uPA. *Biochem J* **2010,** *426* (2), 219-28.

55.   Shimomura, T.; Miyazawa, K.; Komiyama, Y.; Hiraoka, H.; Naka, D.; Morimoto, Y.; Kitamura, N., Activation of hepatocyte growth factor by two homologous proteases, blood-coagulation factor XIIa and hepatocyte growth factor activator. *Eur J Biochem* **1995,** *229* (1), 257-61.

56.   Comoglio, P. M.; Giordano, S.; Trusolino, L., Drug development of MET inhibitors: targeting oncogene addiction and expedience. *Nat Rev Drug Discov* **2008,** *7* (6), 504-16.

57.   Tripathi, M.; Nandana, S.; Yamashita, H.; Ganesan, R.; Kirchhofer, D.; Quaranta, V., Laminin-332 is a substrate for hepsin, a protease associated with prostate cancer progression. *J Biol Chem* **2008,** *283* (45), 30576-84.

58.   Wilkinson, D. J.; Desilets, A.; Lin, H.; Charlton, S.; Del Carmen Arques, M.; Falconer, A.; Bullock, C.; Hsu, Y. C.; Birchall, K.; Hawkins, A.; Thompson, P.; Ferrell, W. R.; Lockhart, J.; Plevin, R.; Zhang, Y.; Blain, E.; Lin, S. W.; Leduc, R.; Milner, J. M.; Rowan, A. D., The serine

proteinase hepsin is an activator of pro-matrix metalloproteinases: molecular mechanisms and implications for extracellular matrix turnover. *Sci Rep* **2017,** *7* (1), 16693.

59.		Moran, P.; Li, W.; Fan, B.; Vij, R.; Eigenbrot, C.; Kirchhofer, D., Pro-urokinase-type plasminogen activator is a substrate for hepsin. *J Biol Chem* **2006,** *281* (41), 30439-46.

60.		Tervonen, T. A.; Belitškin, D.; Pant, S. M.; Englund, J. I.; Marques, E.; Ala-Hongisto, H.; Nevalaita, L.; Sihto, H.; Heikkilä, P.; Leidenius, M.; Hewitson, K.; Ramachandra, M.; Moilanen, A.; Joensuu, H.; Kovanen, P. E.; Poso, A.; Klefström, J., Deregulated hepsin protease activity confers oncogenicity by concomitantly augmenting HGF/MET signalling and disrupting epithelial cohesion. *Oncogene* **2016,** *35* (14), 1832-46.

61.		Wu, Q.; Yu, D.; Post, J.; Halks-Miller, M.; Sadler, J. E.; Morser, J., Generation and characterization of mice deficient in hepsin, a hepatic transmembrane serine protease. *J Clin Invest* **1998,** *101* (2), 321-6.

62.		Yu, I. S.; Chen, H. J.; Lee, Y. S.; Huang, P. H.; Lin, S. R.; Tsai, T. W.; Lin, S. W., Mice deficient in hepsin, a serine protease, exhibit normal embryogenesis and unchanged hepatocyte regeneration ability. *Thromb Haemost* **2000,** *84* (5), 865-70.

63.		Li, S.; Peng, J.; Wang, H.; Zhang, W.; Brown, J. M.; Zhou, Y.; Wu, Q., Hepsin enhances liver metabolism and inhibits adipocyte browning in mice. *Proc Natl Acad Sci U S A* **2020,** *117* (22), 12359-12367.

64.		Guipponi, M.; Tan, J.; Cannon, P. Z.; Donley, L.; Crewther, P.; Clarke, M.; Wu, Q.; Shepherd, R. K.; Scott, H. S., Mice deficient for the type II transmembrane serine protease, TMPRSS1/hepsin, exhibit profound hearing loss. *Am J Pathol* **2007,** *171* (2), 608-16.

65.		Hsu, Y. C.; Huang, H. P.; Yu, I. S.; Su, K. Y.; Lin, S. R.; Lin, W. C.; Wu, H. L.; Shi, G. Y.; Tao, M. H.; Kao, C. H.; Wu, Y. M.; Martin, P. E.; Lin, S. Y.; Yang, P. C.; Lin, S. W., Serine protease hepsin regulates hepatocyte size and hemodynamic retention of tumor cells by hepatocyte growth factor signaling in mice. *Hepatology* **2012,** *56* (5), 1913-23.

66.		Brunati, M.; Perucca, S.; Han, L.; Cattaneo, A.; Consolato, F.; Andolfo, A.; Schaeffer, C.; Olinger, E.; Peng, J.; Santambrogio, S.; Perrier, R.; Li, S.; Bokhove, M.; Bachi, A.; Hummler, E.; Devuyst, O.; Wu, Q.; Jovine, L.; Rampoldi, L., The serine protease hepsin mediates urinary secretion and polymerisation of Zona Pellucida domain protein uromodulin. *Elife* **2015,** *4*, e08887.

67.		Chen, Z.; Fan, Z.; McNeal, J. E.; Nolley, R.; Caldwell, M. C.; Mahadevappa, M.; Zhang, Z.; Warrington, J. A.; Stamey, T. A., Hepsin and maspin are inversely expressed in laser capture microdissected prostate cancer. *J Urol* **2003,** *169* (4), 1316-9.

68.		Magee, J. A.; Araki, T.; Patil, S.; Ehrig, T.; True, L.; Humphrey, P. A.; Catalona, W. J.; Watson, M. A.; Milbrandt, J., Expression profiling reveals hepsin overexpression in prostate cancer. *Cancer Res* **2001,** *61* (15), 5692-6.

69.   Zacharski, L. R.; Ornstein, D. L.; Memoli, V. A.; Rousseau, S. M.; Kisiel, W., Expression of the factor VII activating protease, hepsin, in situ in renal cell carcinoma. *Thromb Haemost* **1998,** *79* (4), 876-7.

70.   Xing, P.; Li, J. G.; Jin, F.; Zhao, T. T.; Liu, Q.; Dong, H. T.; Wei, X. L., Clinical and biological significance of hepsin overexpression in breast cancer. *J Investig Med* **2011,** *59* (5), 803-10.

71.   Stephan, C.; Yousef, G. M.; Scorilas, A.; Jung, K.; Jung, M.; Kristiansen, G.; Hauptmann, S.; Kishi, T.; Nakamura, T.; Loening, S. A.; Diamandis, E. P., Hepsin is highly over expressed in and a new candidate for a prognostic indicator in prostate cancer. *J Urol* **2004,** *171* (1), 187-91.

72.   Landers, K. A.; Burger, M. J.; Tebay, M. A.; Purdie, D. M.; Scells, B.; Samaratunga, H.; Lavin, M. F.; Gardiner, R. A., Use of multiple biomarkers for a molecular diagnosis of prostate cancer. *Int J Cancer* **2005,** *114* (6), 950-6.

73.   Stamey, T. A.; Warrington, J. A.; Caldwell, M. C.; Chen, Z.; Fan, Z.; Mahadevappa, M.; McNeal, J. E.; Nolley, R.; Zhang, Z., Molecular genetic profiling of Gleason grade 4/5 prostate cancers compared to benign prostatic hyperplasia. *J Urol* **2001,** *166* (6), 2171-7.

74.   Goel, M. M.; Agrawal, D.; Natu, S. M.; Goel, A., Hepsin immunohistochemical expression in prostate cancer in relation to Gleason's grade and serum prostate specific antigen. *Indian J Pathol Microbiol* **2011,** *54* (3), 476-81.

75.   Wu, Q.; Parry, G., Hepsin and prostate cancer. *Front Biosci* **2007,** *12*, 5052-9.

76.   Valkenburg, K. C.; Hostetter, G.; Williams, B. O., Concurrent Hepsin overexpression and adenomatous polyposis coli deletion causes invasive prostate carcinoma in mice. *Prostate* **2015,** *75* (14), 1579-85.

77.   Klezovitch, O.; Chevillet, J.; Mirosevich, J.; Roberts, R. L.; Matusik, R. J.; Vasioukhin, V., Hepsin promotes prostate cancer progression and metastasis. *Cancer Cell* **2004,** *6* (2), 185-95.

78.   Tang, X.; Mahajan, S. S.; Nguyen, L. T.; Béliveau, F.; Leduc, R.; Simon, J. A.; Vasioukhin, V., Targeted inhibition of cell-surface serine protease Hepsin blocks prostate cancer bone metastasis. *Oncotarget* **2014,** *5* (5), 1352-62.

79.   Fourneau, E. F. U. E., Über einige Derivate des Gly kocolls. *Ber* **1901,** *34*, 2868-2877.

80.   Vigneaud, V. d.; Ressler, C.; Swan, C. J. M.; Roberts, C. W.; Katsoyannis, P. G.; Gordon, S., The synthesis of an octapeptide amide with the hormonal activity of oxytocin. *Journal of the American Chemical Society* **1953,** *75* (19), 4879-4880.

81.   Merrifield, R. B., Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *Journal of the American Chemical Society* **1963,** *85* (14), 2149-2154.

82.   The Nobel Prize in Chemistry 1984. https://www.nobelprize.org/prizes/chemistry/1984/summary/ (accessed March 22, 2021).

83.   Carpino, L. A.; Han, G. Y., 9-Fluorenylmethoxycarbonyl function, a new base-sensitive amino-protecting group. *Journal of the American Chemical Society* **1970,** *92* (19), 5748-5749.

84.   O'Ferrall, R. A. M.; Slae, S., β-Elimination of 9-fluorenylmethanol in aqueous solution: an E1cB mechanism. *Journal of the Chemical Society B: Physical Organic* **1970,** (0), 260-268.

85.   El-Faham, A.; Albericio, F., Peptide coupling reagents, more than a letter soup. *Chem Rev* **2011,** *111* (11), 6557-602.

86.   Frérot, E.; Coste, J.; Pantaloni, A.; Dufour, M.-N.; Jouin, P., PyBOP® and PyBroP: Two reagents for the difficult coupling of the α,α-dialkyl amino acid, Aib. *Tetrahedron* **1991,** *47* (2), 259-270.

87.   Montalbetti, C.; Falque, V., Amide Bond Formation and Peptide Coupling. *Tetrahedron* **2005,** *61*, 10827-10852.

88.   Bédard, F.; Biron, E., Recent Progress in the Chemical Synthesis of Class II and S-Glycosylated Bacteriocins. *Front Microbiol* **2018,** *9*, 1048.

89.   Forster, T., Energiewanderung und fluoreszenz. *Naturwissenschaften* **1946,** *33* (6), 166-175.

90.   Förster, T., Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen der Physik* **1948,** *437* (1-2), 55-75.

91.   Förster, T., Excitation transfer. *Comparative Effects of Radiation* **1960**, 300-319.

92.   Stryer, L.; Haugland, R. P., Energy transfer: a spectroscopic ruler. *Proc Natl Acad Sci U S A* **1967,** *58* (2), 719-26.

93.   Sahoo, H., Förster resonance energy transfer – A spectroscopic nanoruler: Principle and applications. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews* **2011,** *12* (1), 20-30.

94.   Lakowicz, J., *Principles of Fluorescence Spectroscopy*. 2006; Vol. 1.

95.   Stokes, G. G., XXX. On the change of refrangibility of light. *Philosophical Transactions of the Royal Society of London* **1852,** *142*, 463-562.

96.   Hochreiter, B.; Garcia, A. P.; Schmid, J. A., Fluorescent proteins as genetically encoded FRET biosensors in life sciences. *Sensors (Basel)* **2015,** *15* (10), 26281-314.

97.   Wu, L.; Huang, C.; Emery, B. P.; Sedgwick, A. C.; Bull, S. D.; He, X. P.; Tian, H.; Yoon, J.; Sessler, J. L.; James, T. D., Förster resonance energy transfer (FRET)-based small-molecule sensors and imaging agents. *Chem Soc Rev* **2020,** *49* (15), 5110-5139.

98.   Truong, K.; Ikura, M., The use of FRET imaging microscopy to detect protein-protein interactions and protein conformational changes in vivo. *Curr Opin Struct Biol* **2001,** *11* (5), 573-8.

99.   Loura, L. M.; Prieto, M., FRET in Membrane Biophysics: An Overview. *Front Physiol* **2011,** *2*, 82.

100.   Aoki, K.; Kiyokawa, E.; Nakamura, T.; Matsuda, M., Visualization of growth signal transduction cascades in living cells with genetically encoded probes based on Förster resonance energy transfer. *Philos Trans R Soc Lond B Biol Sci* **2008,** *363* (1500), 2143-51.

101.   Poreba, M.; Szalek, A.; Rut, W.; Kasperkiewicz, P.; Rutkowska-Wlodarczyk, I.; Snipas, S. J.; Itoh, Y.; Turk, D.; Turk, B.; Overall, C. M.; Kaczmarek, L.; Salvesen, G. S.; Drag, M., Highly sensitive and adaptable fluorescence-quenched pair discloses the substrate specificity profiles in diverse protease families. *Sci Rep* **2017,** *7*, 43135.

102.   Masui, Y.; Takemoto, T.; Sakakibara, S.; Hori, H.; Nagai, Y., Synthetic substrates for vertebrate collagenase. *Biochem Med* **1977,** *17* (2), 215-21.

103.   Berg, J. M.; Tymoczko, J. L.; Gatto jr, G. J.; Stryer, L., Enzyme: Grundlegende Konzepte und Kinetik. In *Stryer Biochemie*, Berg, J. M.; Tymoczko, J. L.; Gatto jr, G. J.; Stryer, L., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2018; pp 255-297.

104.   Michaelis, L.; Menten, M. L.; Johnson, K. A.; Goody, R. S., The original Michaelis constant: translation of the 1913 Michaelis-Menten paper. *Biochemistry* **2011,** *50* (39), 8264-9.

105.   Lineweaver, H.; Burk, D., The Determination of Enzyme Dissociation Constants. *Journal of the American Chemical Society* **1934,** *56* (3), 658-666.

106.   Cheng, Y.; Prusoff, W. H., Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem Pharmacol* **1973,** *22* (23), 3099-108.

107.   Ferlay, J.; Autier, P.; Boniol, M.; Heanue, M.; Colombet, M.; Boyle, P., Estimates of the cancer incidence and mortality in Europe in 2006. *Ann Oncol* **2007,** *18* (3), 581-92.

108.   Wallis, C. J. D.; Saskin, R.; Choo, R.; Herschorn, S.; Kodama, R. T.; Satkunasivam, R.; Shah, P. S.; Danjoux, C.; Nam, R. K., Surgery Versus Radiotherapy for Clinically-localized Prostate Cancer: A Systematic Review and Meta-analysis. *Eur Urol* **2016,** *70* (1), 21-30.

109.   Sharifi, N.; Gulley, J. L.; Dahut, W. L., Androgen deprivation therapy for prostate cancer. *JAMA* **2005,** *294* (2), 238-44.

110.   Attard, G.; Cooper, C. S.; de Bono, J. S., Steroid hormone receptors in prostate cancer: a hard habit to break? *Cancer Cell* **2009,** *16* (6), 458-62.

111.   Heidenreich, A.; Bastian, P. J.; Bellmunt, J.; Bolla, M.; Joniau, S.; van der Kwast, T.; Mason, M.; Matveev, V.; Wiegel, T.; Zattoni, F.; Mottet, N.; European Association of, U., EAU guidelines on prostate cancer. Part II: Treatment of advanced, relapsing, and castration-resistant prostate cancer. *Eur Urol* **2014,** *65* (2), 467-79.

112.   Ndibe, C.; Wang, C. G.; Sonpavde, G., Corticosteroids in the management of prostate cancer: a critical review. *Curr Treat Options Oncol* **2015,** *16* (2), 6.

113.   Tannock, I. F.; de Wit, R.; Berry, W. R.; Horti, J.; Pluzanska, A.; Chi, K. N.; Oudard, S.; Theodore, C.; James, N. D.; Turesson, I.; Rosenthal, M. A.; Eisenberger, M. A.; Investigators, T. A. X., Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer. *N Engl J Med* **2004,** *351* (15), 1502-12.

114.   Wang, Y.; Deng, Y.; Luo, H.; Zhu, A.; Ke, H.; Yang, H.; Chen, H., Light-Responsive Nanoparticles for Highly Efficient Cytoplasmic Delivery of Anticancer Agents. *ACS Nano* **2017,** *11* (12), 12134-12144.

115.   Sun, X.; Zhang, G.; Wu, Z., Nanostructures for pH-sensitive Drug Delivery and Magnetic Resonance Contrast Enhancement Systems. *Curr Med Chem* **2018,** *25* (25), 3036-3057.

116.   Giaquinto, M.; Ricciardi, A.; Aliberti, A.; Micco, A.; Bobeico, E.; Ruvo, M.; Cusano, A., Light-microgel interaction in resonant nanostructures. *Scientific Reports* **2018,** *8* (1), 9331.

117.   Dong, H.; Pang, L.; Cong, H.; Shen, Y.; Yu, B., Application and design of esterase-responsive nanoparticles for cancer therapy. *Drug Deliv* **2019,** *26* (1), 416-432.

118.   Klok, H.-A., Peptide/Protein–Synthetic Polymer Conjugates: Quo Vadis. *Macromolecules* **2009,** *42* (21), 7990-8000.

119.   Andrieu, J.; Kotman, N.; Maier, M.; Mailander, V.; Strauss, W. S.; Weiss, C. K.; Landfester, K., Live monitoring of cargo release from peptide-based hybrid nanocapsules induced by enzyme cleavage. *Macromol Rapid Commun* **2012,** *33* (3), 248-53.

120.   Bugge, T. H.; Antalis, T. M.; Wu, Q., Type II transmembrane serine proteases. *J Biol Chem* **2009,** *284* (35), 23177-81.

121.   Duffy, F. J.; Verniere, M.; Devocelle, M.; Bernard, E.; Shields, D. C.; Chubb, A. J., CycloPs: generating virtual libraries of cyclized and constrained peptides including nonnatural amino acids. *J Chem Inf Model* **2011,** *51* (4), 829-36.

122.   Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics* **2006,** *65* (3), 712-725.

123.   Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **2015,** *11* (8), 3696-713.

124.   Wang, J.; Cieplak, P.; Kollman, P. A., How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry* **2000,** *21* (12), 1049-1074.

125.   Inc., C. C. G. *Moleculat Operating Environment (MOE)*, 2015.

126.   Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *Nucleic Acids Res* **2000,** *28* (1), 235-42.

127.   Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G., A Fast Flexible Docking Method using an Incremental Construction Algorithm. *Journal of Molecular Biology* **1996,** *261* (3), 470-489.

128.   Chan, W.; White, P., *Fmoc solid phase peptide synthesis: a practical approach*. OUP Oxford: 1999; Vol. 222.

129.   Coin, I.; Beyermann, M.; Bienert, M., Solid-phase peptide synthesis: from standard procedures to the synthesis of difficult sequences. *Nature Protocols* **2007,** *2*, 3247.

130.   Jenssen, H.; Aspmo, S. I., Serum stability of peptides. *Methods Mol Biol* **2008,** *494*, 177-86.

131.   Pons, C.; Grosdidier, S.; Solernou, A.; Pérez-Cano, L.; Fernández-Recio, J., Present and future challenges and limitations in protein–protein docking. *Proteins: Structure, Function, and Bioinformatics* **2010,** *78* (1), 95-108.

132.   Andrusier, N.; Mashiach, E.; Nussinov, R.; Wolfson, H. J., Principles of flexible protein–protein docking. *Proteins: Structure, Function, and Bioinformatics* **2008,** *73* (2), 271-289.

133.   Kumaresan, P. R.; Natarajan, A.; Song, A.; Wang, X.; Liu, R.; DeNardo, G.; DeNardo, S.; Lam, K. S., Development of tissue plasminogen activator specific "on demand cleavable" (odc) linkers for radioimmunotherapy by screening one-bead-one-compound combinatorial peptide libraries. *Bioconjug Chem* **2007,** *18* (1), 175-82.

134.   Cicenas, J.; Zalyte, E.; Bairoch, A.; Gaudet, P., Kinases and Cancer. *Cancers (Basel)* **2018,** *10* (3).

135.   Martin, C. E.; List, K., Cell surface-anchored serine proteases in cancer progression and metastasis. *Cancer Metastasis Rev* **2019,** *38* (3), 357-387.

136.   Tanabe, L. M.; List, K., The role of type II transmembrane serine protease-mediated signaling in cancer. *FEBS J* **2017,** *284* (10), 1421-1436.

137.   Fuchs, A. V.; Kotman, N.; Andrieu, J.; Mailänder, V.; Weiss, C. K.; Landfester, K., Enzyme cleavable nanoparticles from peptide based triblock copolymers. *Nanoscale* **2013,** *5* (11), 4829-39.

138.   Knaff, P. M.; Kersten, C.; Willbold, R.; Champanhac, C.; Crespy, D.; Wittig, R.; Landfester, K.; Mailander, V., From In Silico to Experimental Validation: Tailoring Peptide Substrates for a Serine Protease. *Biomacromolecules* **2020,** *21* (4), 1636-1643.

139.   Costanzo, M. J.; Almond, H. R.; Hecker, L. R.; Schott, M. R.; Yabut, S. C.; Zhang, H. C.; Andrade-Gordon, P.; Corcoran, T. W.; Giardino, E. C.; Kauffman, J. A.; Lewis, J. M.; de Garavilla, L.; Haertlein, B. J.; Maryanoff, B. E., In-depth study of tripeptide-based alpha-ketoheterocycles as inhibitors of thrombin. Effective utilization of the S1' subsite and its implications to structure-based drug design. *J Med Chem* **2005,** *48* (6), 1984-2008.

140.   Damalanka, V. C.; Han, Z.; Karmakar, P.; O'Donoghue, A. J.; La Greca, F.; Kim, T.; Pant, S. M.; Helander, J.; Klefström, J.; Craik, C. S.; Janetka, J. W., Discovery of Selective Matriptase and Hepsin Serine Protease Inhibitors: Useful Chemical Tools for Cancer Cell Biology. *J Med Chem* **2019,** *62* (2), 480-490.

141.   Han, Z.; Harris, P. K.; Karmakar, P.; Kim, T.; Owusu, B. Y.; Wildman, S. A.; Klampfer, L.; Janetka, J. W., α-Ketobenzothiazole Serine Protease Inhibitors of Aberrant HGF/c-MET and MSP/RON Kinase Pathway Signaling in Cancer. *ChemMedChem* **2016,** *11* (6), 585-99.

142.   Damalanka, V. C.; Wildman, S. A.; Janetka, J. W., Piperidine carbamate peptidomimetic inhibitors of the serine proteases HGFA, matriptase and hepsin. *Medchemcomm* **2019,** *10* (9), 1646-1655.

143.   Oberst, M. D.; Singh, B.; Ozdemirli, M.; Dickson, R. B.; Johnson, M. D.; Lin, C. Y., Characterization of matriptase expression in normal human tissues. *J Histochem Cytochem* **2003,** *51* (8), 1017-25.

144.   Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G., A fast flexible docking method using an incremental construction algorithm. *J Mol Biol* **1996,** *261* (3), 470-89.

145.   Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *Nucleic Acids Research* **2000,** *28* (1), 235-242.

146.   *LeadIT*, 2.3.2; BioSolveIT GmbH: Sankt Augustin, Germany, 2017.

147.   Bietz, S.; Urbaczek, S.; Schulz, B.; Rarey, M., Protoss: a holistic approach to predict tautomers and protonation states in protein-ligand complexes. *J Cheminform* **2014,** *6*, 12.

148.   *Molecular Operating Environment (MOE)*, Chemical Computing Group, Inc.: Montreal, QC, Canada, 2019.

149.   Goswami, R.; Mukherjee, S.; Ghadiyaram, C.; Wohlfahrt, G.; Sistla, R. K.; Nagaraj, J.; Satyam, L. K.; Subbarao, K.; Palakurthy, R. K.; Gopinath, S.; Krishnamurthy, N. R.; Ikonen, T.; Moilanen, A.; Subramanya, H. S.; Kallio, P.; Ramachandra, M., Structure-guided discovery of 1,3,5 tri-substituted benzenes as potent and selective matriptase inhibitors exhibiting in vivo antitumor efficacy. *Bioorg Med Chem* **2014,** *22* (12), 3187-203.

150.   *The PyMOL Molecular Graphics System. Version 2.2.0*, Schrödinger LLC.

151.   Corporation, C. Trityl-Hydroxide (TRT-OH) SpheriTideResinLoading Procedure. http://cem.com/media/contenttype/media/literature/525_TechNote_SpheriTide_TRT-OHLoading.pdf.

152.   Coin, I.; Beyermann, M.; Bienert, M., Solid-phase peptide synthesis: from standard procedures to the synthesis of difficult sequences. *Nat Protoc* **2007,** *2* (12), 3247-56.

153.   Lin, J.; Deng, H.; Jin, L.; Pandey, P.; Quinn, J.; Cantin, S.; Rynkiewicz, M. J.; Gorga, J. C.; Bibbins, F.; Celatka, C. A.; Nagafuji, P.; Bannister, T. D.; Meyers, H. V.; Babine, R. E.; Hayward, N. J.; Weaver, D.; Benjamin, H.; Stassen, F.; Abdel-Meguid, S. S.; Strickler, J. E., Design, synthesis, and biological evaluation of peptidomimetic inhibitors of factor XIa as novel anticoagulants. *J Med Chem* **2006,** *49* (26), 7781-91.

154.   Adang, A. E.; de Man, A. P.; Vogel, G. M.; Grootenhuis, P. D.; Smit, M. J.; Peters, C. A.; Visser, A.; Rewinkel, J. B.; van Dinther, T.; Lucas, H.; Kelder, J.; van Aelst, S.; Meuleman, D. G.; van Boeckel, C. A., Unique overlap in the prerequisites for thrombin inhibition and oral bioavailability resulting in potent oral antithrombotics. *J Med Chem* **2002,** *45* (20), 4419-32.

155.   Furtmann, N.; Häußler, D.; Scheidt, T.; Stirnberg, M.; Steinmetzer, T.; Bajorath, J.; Gütschow, M., Limiting the Number of Potential Binding Modes by Introducing Symmetry into Ligands: Structure-Based Design of Inhibitors for Trypsin-Like Serine Proteases. *Chemistry* **2016,** *22* (2), 610-25.

156.   Powell, M. F.; Stewart, T.; Otvos, L.; Urge, L.; Gaeta, F. C.; Sette, A.; Arrhenius, T.; Thomson, D.; Soda, K.; Colon, S. M., Peptide stability in drug development. II. Effect of single amino acid substitution and glycosylation on peptide reactivity in human serum. *Pharm Res* **1993,** *10* (9), 1268-73.

157.   Han, Z.; Harris, P. K.; Jones, D. E.; Chugani, R.; Kim, T.; Agarwal, M.; Shen, W.; Wildman, S. A.; Janetka, J. W., Inhibitors of HGFA, Matriptase, and Hepsin Serine Proteases: A Nonkinase Strategy to Block Cell Signaling in Cancer. *ACS Med Chem Lett* **2014,** *5* (11), 1219-24.

158.   Wu, F.; Zhao, S.; Yu, B.; Chen, Y. M.; Wang, W.; Song, Z. G.; Hu, Y.; Tao, Z. W.; Tian, J. H.; Pei, Y. Y.; Yuan, M. L.; Zhang, Y. L.; Dai, F. H.; Liu, Y.; Wang, Q. M.; Zheng, J. J.; Xu, L.; Holmes, E. C.; Zhang, Y. Z., A new coronavirus associated with human respiratory disease in China. *Nature* **2020,** *579* (7798), 265-269.

159.   Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; Niu, P.; Zhan, F.; Ma, X.; Wang, D.; Xu, W.; Wu, G.; Gao, G. F.; Tan, W., A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine* **2020,** *382* (8), 727-733.

160.   Rossi, G. A.; Sacco, O.; Mancino, E.; Cristiani, L.; Midulla, F., Differences and similarities between SARS-CoV and SARS-CoV-2: spike receptor-binding domain recognition

and host cell infection with support of cellular serine proteases. *Infection* **2020,** *48* (5), 665-669.

161.   Viruses, C. S. G. o. t. I. C. o. T. o., The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* **2020,** *5* (4), 536-544.

162.   Dennison Himmelfarb, C. R.; Baptiste, D., Coronavirus Disease (COVID-19): Implications for Cardiovascular and Socially At-risk Populations. *Journal of Cardiovascular Nursing* **2020,** *35* (4).

163.   Remuzzi, A.; Remuzzi, G., COVID-19 and Italy: what next? *Lancet* **2020,** *395* (10231), 1225-1228.

164.   Saglietto, A.; D'Ascenzo, F.; Zoccai, G. B.; De Ferrari, G. M., COVID-19 in Europe: the Italian lesson. *Lancet* **2020,** *395* (10230), 1110-1111.

165.   Dhama, K.; Sharun, K.; Tiwari, R.; Dadar, M.; Malik, Y. S.; Singh, K. P.; Chaicumpa, W., COVID-19, an emerging coronavirus infection: advances and prospects in designing and developing vaccines, immunotherapeutics, and therapeutics. *Hum Vaccin Immunother* **2020,** *16* (6), 1232-1238.

166.   Shah, B.; Modi, P.; Sagar, S. R., In silico studies on therapeutic agents for COVID-19: Drug repurposing approach. *Life Sci* **2020,** *252*, 117652.

167.   Warren, T. K.; Jordan, R.; Lo, M. K.; Ray, A. S.; Mackman, R. L.; Soloveva, V.; Siegel, D.; Perron, M.; Bannister, R.; Hui, H. C.; Larson, N.; Strickley, R.; Wells, J.; Stuthman, K. S.; Van Tongeren, S. A.; Garza, N. L.; Donnelly, G.; Shurtleff, A. C.; Retterer, C. J.; Gharaibeh, D.; Zamani, R.; Kenny, T.; Eaton, B. P.; Grimes, E.; Welch, L. S.; Gomba, L.; Wilhelmsen, C. L.; Nichols, D. K.; Nuss, J. E.; Nagle, E. R.; Kugelman, J. R.; Palacios, G.; Doerffler, E.; Neville, S.; Carra, E.; Clarke, M. O.; Zhang, L.; Lew, W.; Ross, B.; Wang, Q.; Chun, K.; Wolfe, L.; Babusis, D.; Park, Y.; Stray, K. M.; Trancheva, I.; Feng, J. Y.; Barauskas, O.; Xu, Y.; Wong, P.; Braun, M. R.; Flint, M.; McMullan, L. K.; Chen, S. S.; Fearns, R.; Swaminathan, S.; Mayers, D. L.; Spiropoulou, C. F.; Lee, W. A.; Nichol, S. T.; Cihlar, T.; Bavari, S., Therapeutic efficacy of the small molecule GS-5734 against Ebola virus in rhesus monkeys. *Nature* **2016,** *531* (7594), 381-5.

168.   Tchesnokov, E. P.; Feng, J. Y.; Porter, D. P.; Götte, M., Mechanism of Inhibition of Ebola Virus RNA-Dependent RNA Polymerase by Remdesivir. *Viruses* **2019,** *11* (4).

169.   Goldman, J. D.; Lye, D. C. B.; Hui, D. S.; Marks, K. M.; Bruno, R.; Montejano, R.; Spinner, C. D.; Galli, M.; Ahn, M. Y.; Nahass, R. G.; Chen, Y. S.; SenGupta, D.; Hyland, R. H.; Osinusi, A. O.; Cao, H.; Blair, C.; Wei, X.; Gaggar, A.; Brainard, D. M.; Towner, W. J.; Muñoz, J.; Mullane, K. M.; Marty, F. M.; Tashima, K. T.; Diaz, G.; Subramanian, A.; Investigators, G.-U.-.-. Remdesivir for 5 or 10 Days in Patients with Severe Covid-19. *N Engl J Med* **2020,** *383* (19), 1827-1837.

170. McMahon, J. H.; Udy, A.; Peleg, A. Y., Remdesivir for the Treatment of Covid-19 - Preliminary Report. *N Engl J Med* **2020,** *383* (10), 992-993.

171. Pan, H.; Peto, R.; Henao-Restrepo, A. M.; Preziosi, M. P.; Sathiyamoorthy, V.; Abdool Karim, Q.; Alejandria, M. M.; Hernández García, C.; Kieny, M. P.; Malekzadeh, R.; Murthy, S.; Reddy, K. S.; Roses Periago, M.; Abi Hanna, P.; Ader, F.; Al-Bader, A. M.; Alhasawi, A.; Allum, E.; Alotaibi, A.; Alvarez-Moreno, C. A.; Appadoo, S.; Asiri, A.; Aukrust, P.; Barratt-Due, A.; Bellani, S.; Branca, M.; Cappel-Porter, H. B. C.; Cerrato, N.; Chow, T. S.; Como, N.; Eustace, J.; García, P. J.; Godbole, S.; Gotuzzo, E.; Griskevicius, L.; Hamra, R.; Hassan, M.; Hassany, M.; Hutton, D.; Irmansyah, I.; Jancoriene, L.; Kirwan, J.; Kumar, S.; Lennon, P.; Lopardo, G.; Lydon, P.; Magrini, N.; Maguire, T.; Manevska, S.; Manuel, O.; McGinty, S.; Medina, M. T.; Mesa Rubio, M. L.; Miranda-Montoya, M. C.; Nel, J.; Nunes, E. P.; Perola, M.; Portolés, A.; Rasmin, M. R.; Raza, A.; Rees, H.; Reges, P. P. S.; Rogers, C. A.; Salami, K.; Salvadori, M. I.; Sinani, N.; Sterne, J. A. C.; Stevanovikj, M.; Tacconelli, E.; Tikkinen, K. A. O.; Trelle, S.; Zaid, H.; Røttingen, J. A.; Swaminathan, S.; Consortium, W. S. T., Repurposed Antiviral Drugs for Covid-19 - Interim WHO Solidarity Trial Results. *N Engl J Med* **2021,** *384* (6), 497-511.

172. Edwards, A., What Are the Odds of Finding a COVID-19 Drug from a Lab Repurposing Screen? *J Chem Inf Model* **2020,** *60* (12), 5727-5729.

173. Din, O. S.; Woll, P. J., Treatment of gastrointestinal stromal tumor: focus on imatinib mesylate. *Ther Clin Risk Manag* **2008,** *4* (1), 149-62.

174. Ou, X.; Liu, Y.; Lei, X.; Li, P.; Mi, D.; Ren, L.; Guo, L.; Guo, R.; Chen, T.; Hu, J.; Xiang, Z.; Mu, Z.; Chen, X.; Chen, J.; Hu, K.; Jin, Q.; Wang, J.; Qian, Z., Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat Commun* **2020,** *11* (1), 1620.

175. Chan, J. F.; Kok, K. H.; Zhu, Z.; Chu, H.; To, K. K.; Yuan, S.; Yuen, K. Y., Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerg Microbes Infect* **2020,** *9* (1), 221-236.

176. Zhou, P.; Yang, X. L.; Wang, X. G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H. R.; Zhu, Y.; Li, B.; Huang, C. L.; Chen, H. D.; Chen, J.; Luo, Y.; Guo, H.; Jiang, R. D.; Liu, M. Q.; Chen, Y.; Shen, X. R.; Wang, X.; Zheng, X. S.; Zhao, K.; Chen, Q. J.; Deng, F.; Liu, L. L.; Yan, B.; Zhan, F. X.; Wang, Y. Y.; Xiao, G. F.; Shi, Z. L., A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020,** *579* (7798), 270-273.

177. Yao, H.; Song, Y.; Chen, Y.; Wu, N.; Xu, J.; Sun, C.; Zhang, J.; Weng, T.; Zhang, Z.; Wu, Z.; Cheng, L.; Shi, D.; Lu, X.; Lei, J.; Crispin, M.; Shi, Y.; Li, L.; Li, S., Molecular Architecture of the SARS-CoV-2 Virus. *Cell* **2020,** *183* (3), 730-738.e13.

178. Fehr, A. R.; Perlman, S., Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol* **2015,** *1282*, 1-23.

179. Kim, D.; Lee, J. Y.; Yang, J. S.; Kim, J. W.; Kim, V. N.; Chang, H., The Architecture of SARS-CoV-2 Transcriptome. *Cell* **2020,** *181* (4), 914-921.e10.

180.   Astuti, I.; Ysrafil, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): An overview of viral structure and host response. *Diabetes Metab Syndr* **2020,** *14* (4), 407-412.

181.   Bárcena, M.; Oostergetel, G. T.; Bartelink, W.; Faas, F. G.; Verkleij, A.; Rottier, P. J.; Koster, A. J.; Bosch, B. J., Cryo-electron tomography of mouse hepatitis virus: Insights into the structure of the coronavirion. *Proc Natl Acad Sci U S A* **2009,** *106* (2), 582-7.

182.   Neuman, B. W.; Adair, B. D.; Yoshioka, C.; Quispe, J. D.; Orca, G.; Kuhn, P.; Milligan, R. A.; Yeager, M.; Buchmeier, M. J., Supramolecular architecture of severe acute respiratory syndrome coronavirus revealed by electron cryomicroscopy. *J Virol* **2006,** *80* (16), 7918-28.

183.   Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N.-H.; Nitsche, A., SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *cell* **2020,** *181* (2), 271-280.

184.   Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X., Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020,** *581* (7807), 215-220.

185.   Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F., Structural basis of receptor recognition by SARS-CoV-2. *Nature* **2020,** *581* (7807), 221-224.

186.   Shin, D.; Mukherjee, R.; Grewe, D.; Bojkova, D.; Baek, K.; Bhattacharya, A.; Schulz, L.; Widera, M.; Mehdipour, A. R.; Tascher, G.; Geurink, P. P.; Wilhelm, A.; van der Heden van Noort, G. J.; Ovaa, H.; Müller, S.; Knobeloch, K. P.; Rajalingam, K.; Schulman, B. A.; Cinatl, J.; Hummer, G.; Ciesek, S.; Dikic, I., Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature* **2020,** *587* (7835), 657-662.

187.   Harcourt, B. H.; Jukneliene, D.; Kanjanahaluethai, A.; Bechill, J.; Severson, K. M.; Smith, C. M.; Rota, P. A.; Baker, S. C., Identification of severe acute respiratory syndrome coronavirus replicase products and characterization of papain-like protease activity. *J Virol* **2004,** *78* (24), 13600-12.

188.   Chen, S.; Jonas, F.; Shen, C.; Hilgenfeld, R.; Higenfeld, R., Liberation of SARS-CoV main protease from the viral polyprotein: N-terminal autocleavage does not depend on the mature dimerization mode. *Protein Cell* **2010,** *1* (1), 59-74.

189.   Satija, N.; Lal, S. K., The molecular biology of SARS coronavirus. *Ann N Y Acad Sci* **2007,** *1102*, 26-38.

190.   Toelzer, C.; Gupta, K.; Yadav, S. K. N.; Borucu, U.; Davidson, A. D.; Kavanagh Williamson, M.; Shoemark, D. K.; Garzoni, F.; Staufer, O.; Milligan, R.; Capin, J.; Mulholland, A. J.; Spatz, J.; Fitzgerald, D.; Berger, I.; Schaffitzel, C., Free fatty acid binding pocket in the locked structure of SARS-CoV-2 spike protein. *Science* **2020,** *370* (6517), 725-730.

191.   Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Bai, F.; Liu, H.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H., Structure of M. *Nature* **2020,** *582* (7811), 289-293.

192.   Klemm, T.; Ebert, G.; Calleja, D. J.; Allison, C. C.; Richardson, L. W.; Bernardini, J. P.; Lu, B. G.; Kuchel, N. W.; Grohmann, C.; Shibata, Y.; Gan, Z. Y.; Cooney, J. P.; Doerflinger, M.; Au, A. E.; Blackmore, T. R.; van der Heden van Noort, G. J.; Geurink, P. P.; Ovaa, H.; Newman, J.; Riboldi-Tunnicliffe, A.; Czabotar, P. E.; Mitchell, J. P.; Feltham, R.; Lechtenberg, B. C.; Lowes, K. N.; Dewson, G.; Pellegrini, M.; Lessene, G.; Komander, D., Mechanism and inhibition of the papain-like protease, PLpro, of SARS-CoV-2. *EMBO J* **2020,** *39* (18), e106275.

193.   Barile, E.; Baggio, C.; Gambini, L.; Shiryaev, S. A.; Strongin, A. Y.; Pellecchia, M., Potential Therapeutic Targeting of Coronavirus Spike Glycoprotein Priming. *Molecules* **2020,** *25* (10).

194.   Xia, S.; Lan, Q.; Su, S.; Wang, X.; Xu, W.; Liu, Z.; Zhu, Y.; Wang, Q.; Lu, L.; Jiang, S., The role of furin cleavage site in SARS-CoV-2 spike protein-mediated membrane fusion in the presence or absence of trypsin. *Signal Transduct Target Ther* **2020,** *5* (1), 92.

195.   Jaimes, J. A.; Millet, J. K.; Whittaker, G. R., Proteolytic Cleavage of the SARS-CoV-2 Spike Protein and the Role of the Novel S1/S2 Site. *iScience* **2020,** *23* (6), 101212.

196.   Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C. L.; Abiona, O.; Graham, B. S.; McLellan, J. S., Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020,** *367* (6483), 1260-1263.

197.   Watanabe, Y.; Allen, J. D.; Wrapp, D.; McLellan, J. S.; Crispin, M., Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **2020,** *369* (6501), 330-333.

198.   Tan, L.; Wang, Q.; Zhang, D.; Ding, J.; Huang, Q.; Tang, Y. Q.; Miao, H., Lymphopenia predicts disease severity of COVID-19: a descriptive and predictive study. *Signal Transduct Target Ther* **2020,** *5* (1), 33.

199.   White, J. M.; Delos, S. E.; Brecher, M.; Schornberg, K., Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. *Crit Rev Biochem Mol Biol* **2008,** *43* (3), 189-219.

200.   Tortorici, M. A.; Walls, A. C.; Lang, Y.; Wang, C.; Li, Z.; Koerhuis, D.; Boons, G. J.; Bosch, B. J.; Rey, F. A.; de Groot, R. J.; Veesler, D., Structural basis for human coronavirus attachment to sialic acid receptors. *Nat Struct Mol Biol* **2019,** *26* (6), 481-489.

201.   Kuhn, J. H.; Li, W.; Choe, H.; Farzan, M., Angiotensin-converting enzyme 2: a functional receptor for SARS coronavirus. *Cell Mol Life Sci* **2004,** *61* (21), 2738-43.

202.   Hamming, I.; Timens, W.; Bulthuis, M. L. C.; Lely, A. T.; Navis, G. J.; van Goor, H., Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *The Journal of pathology* **2004,** *203* (2), 631-637.

203.   Becerra-Flores, M.; Cardozo, T., SARS-CoV-2 viral spike G614 mutation exhibits higher case fatality rate. *Int J Clin Pract* **2020,** *74* (8), e13525.

204.   Wang, C.; Liu, Z.; Chen, Z.; Huang, X.; Xu, M.; He, T.; Zhang, Z., The establishment of reference sequence for SARS-CoV-2 and variation analysis. *J Med Virol* **2020,** *92* (6), 667-674.

205.   Hoffmann, M.; Kleine-Weber, H.; Pöhlmann, S., A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. *Molecular Cell* **2020,** *78* (4), 779-784.e5.

206.   Barrett, C. T.; Dutch, R. E., Viral Membrane Fusion and the Transmembrane Domain. *Viruses* **2020,** *12* (7).

207.   Jaimes, J. A.; Whittaker, G. R., Feline coronavirus: Insights into viral pathogenesis based on the spike protein structure and function. *Virology* **2018,** *517*, 108-121.

208.   Böttcher, E.; Matrosovich, T.; Beyerle, M.; Klenk, H. D.; Garten, W.; Matrosovich, M., Proteolytic activation of influenza viruses by serine proteases TMPRSS2 and HAT from human airway epithelium. *J Virol* **2006,** *80* (19), 9896-8.

209.   Shulla, A.; Heald-Sargent, T.; Subramanya, G.; Zhao, J.; Perlman, S.; Gallagher, T., A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. *J Virol* **2011,** *85* (2), 873-82.

210.   Bestle, D.; Heindl, M. R.; Limburg, H.; Van Lam van, T.; Pilgram, O.; Moulton, H.; Stein, D. A.; Hardes, K.; Eickmann, M.; Dolnik, O.; Rohde, C.; Klenk, H. D.; Garten, W.; Steinmetzer, T.; Böttcher-Friebertshäuser, E., TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Sci Alliance* **2020,** *3* (9).

211.   Konttinen, Y. T.; Porola, P.; Konttinen, L.; Laine, M.; Poduval, P., Immunohistopathology of Sjögren's syndrome. *Autoimmun Rev* **2006,** *6* (1), 16-20.

212.   Paoloni-Giacobino, A.; Chen, H.; Peitsch, M. C.; Rossier, C.; Antonarakis, S. E., Cloning of the TMPRSS2 gene, which encodes a novel serine protease with transmembrane, LDLRA, and SRCR domains and maps to 21q22.3. *Genomics* **1997,** *44* (3), 309-20.

213.   Hooper, J. D.; Clements, J. A.; Quigley, J. P.; Antalis, T. M., Type II transmembrane serine proteases. Insights into an emerging class of cell surface proteolytic enzymes. *J Biol Chem* **2001,** *276* (2), 857-60.

214.   Jacquinet, E.; Rao, N. V.; Rao, G. V.; Zhengming, W.; Albertine, K. H.; Hoidal, J. R., Cloning and characterization of the cDNA and gene for human epitheliasin. *Eur J Biochem* **2001,** *268* (9), 2687-99.

215.   Jacquinet, E.; Rao, N. V.; Rao, G. V.; Hoidal, J. R., Cloning, genomic organization, chromosomal assignment and expression of a novel mosaic serine proteinase: epitheliasin. *FEBS Lett* **2000,** *468* (1), 93-100.

216.	Krieger, M.; Herz, J., Structures and functions of multiligand lipoprotein receptors: macrophage scavenger receptors and LDL receptor-related protein (LRP). *Annu Rev Biochem* **1994,** *63*, 601-37.

217.	Resnick, D.; Chatterton, J. E.; Schwartz, K.; Slayter, H.; Krieger, M., Structures of class A macrophage scavenger receptors. Electron microscopic study of flexible, multidomain, fibrous proteins and determination of the disulfide bond pattern of the scavenger receptor cysteine-rich domain. *J Biol Chem* **1996,** *271* (43), 26924-30.

218.	Liu, L.; Yang, J.; Qiu, L.; Wang, L.; Zhang, H.; Wang, M.; Vinu, S. S.; Song, L., A novel scavenger receptor-cysteine-rich (SRCR) domain containing scavenger receptor identified from mollusk mediated PAMP recognition and binding. *Dev Comp Immunol* **2011,** *35* (2), 227-39.

219.	Afar, D. E.; Vivanco, I.; Hubert, R. S.; Kuo, J.; Chen, E.; Saffran, D. C.; Raitano, A. B.; Jakobovits, A., Catalytic cleavage of the androgen-regulated TMPRSS2 protease results in its secretion by prostate and prostate cancer epithelia. *Cancer Res* **2001,** *61* (4), 1686-92.

220.	Netzel-Arnett, S.; Hooper, J. D.; Szabo, R.; Madison, E. L.; Quigley, J. P.; Bugge, T. H.; Antalis, T. M., Membrane anchored serine proteases: a rapidly expanding group of cell surface proteolytic enzymes with potential roles in cancer. *Cancer Metastasis Rev* **2003,** *22* (2-3), 237-58.

221.	Zhirnov, O. P.; Ikizler, M. R.; Wright, P. F., Cleavage of influenza a virus hemagglutinin in human respiratory epithelium is cell associated and sensitive to exogenous antiproteases. *J Virol* **2002,** *76* (17), 8682-9.

222.	Chi, M.; Shi, X.; Huo, X.; Wu, X.; Zhang, P.; Wang, G., Dexmedetomidine promotes breast cancer cell migration through Rab11-mediated secretion of exosomal TMPRSS2. *Ann Transl Med* **2020,** *8* (8), 531.

223.	Lu, D.; Fütterer, K.; Korolev, S.; Zheng, X.; Tan, K.; Waksman, G.; Sadler, J. E., Crystal structure of enteropeptidase light chain complexed with an analog of the trypsinogen activation peptide. *J Mol Biol* **1999,** *292* (2), 361-73.

224.	Friedrich, R.; Fuentes-Prior, P.; Ong, E.; Coombs, G.; Hunter, M.; Oehler, R.; Pierson, D.; Gonzalez, R.; Huber, R.; Bode, W.; Madison, E. L., Catalytic domain structures of MT-SP1/matriptase, a matrix-degrading transmembrane serine proteinase. *J Biol Chem* **2002,** *277* (3), 2160-8.

225.	Perona, J. J.; Craik, C. S., Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J Biol Chem* **1997,** *272* (48), 29987-90.

226.	Meyer, D.; Sielaff, F.; Hammami, M.; Bottcher-Friebertshauser, E.; Garten, W.; Steinmetzer, T., Identification of the first synthetic inhibitors of the type II transmembrane serine protease TMPRSS2 suitable for inhibition of influenza virus activation. *Biochem J* **2013,** *452* (2), 331-43.

227.   Lin, B.; Ferguson, C.; White, J. T.; Wang, S.; Vessella, R.; True, L. D.; Hood, L.; Nelson, P. S., Prostate-localized and androgen-regulated expression of the membrane-bound serine protease TMPRSS2. *Cancer Res* **1999,** *59* (17), 4180-4.

228.   Donaldson, S. H.; Hirsh, A.; Li, D. C.; Holloway, G.; Chao, J.; Boucher, R. C.; Gabriel, S. E., Regulation of the epithelial sodium channel by serine proteases in human airways. *J Biol Chem* **2002,** *277* (10), 8338-45.

229.   Kim, T. S.; Heinlein, C.; Hackman, R. C.; Nelson, P. S., Phenotypic analysis of mice lacking the Tmprss2-encoded protease. *Mol Cell Biol* **2006,** *26* (3), 965-75.

230.   Garty, H.; Palmer, L. G., Epithelial sodium channels: function, structure, and regulation. *Physiol Rev* **1997,** *77* (2), 359-96.

231.   Hohenester, E.; Sasaki, T.; Timpl, R., Crystal structure of a scavenger receptor cysteine-rich domain sheds light on an ancient superfamily. *Nat Struct Biol* **1999,** *6* (3), 228-32.

232.   Brown, M. S.; Herz, J.; Goldstein, J. L., LDL-receptor structure. Calcium cages, acid baths and recycling receptors. *Nature* **1997,** *388* (6643), 629-30.

233.   Nykjaer, A.; Conese, M.; Christensen, E. I.; Olson, D.; Cremona, O.; Gliemann, J.; Blasi, F., Recycling of the urokinase receptor upon internalization of the uPA:serpin complexes. *EMBO J* **1997,** *16* (10), 2610-20.

234.   Kounnas, M. Z.; Church, F. C.; Argraves, W. S.; Strickland, D. K., Cellular internalization and degradation of antithrombin III-thrombin, heparin cofactor II-thrombin, and alpha 1-antitrypsin-trypsin complexes is mediated by the low density lipoprotein receptor-related protein. *J Biol Chem* **1996,** *271* (11), 6523-9.

235.   Lam, D. K.; Dang, D.; Flynn, A. N.; Hardt, M.; Schmidt, B. L., TMPRSS2, a novel membrane-anchored mediator in cancer pain. *Pain* **2015,** *156* (5), 923-930.

236.   Vaarala, M. H.; Porvari, K.; Kyllönen, A.; Lukkarinen, O.; Vihko, P., The TMPRSS2 gene encoding transmembrane serine protease is overexpressed in a majority of prostate cancer patients: detection of mutated TMPRSS2 form in a case of aggressive disease. *Int J Cancer* **2001,** *94* (5), 705-10.

237.   Lubieniecka, J. M.; Cheteri, M. K.; Stanford, J. L.; Ostrander, E. A., Met160Val polymorphism in the TRMPSS2 gene and risk of prostate cancer in a population-based case-control study. *Prostate* **2004,** *59* (4), 357-9.

238.   Wilson, S.; Greer, B.; Hooper, J.; Zijlstra, A.; Walker, B.; Quigley, J.; Hawthorne, S., The membrane-anchored serine protease, TMPRSS2, activates PAR-2 in prostate cancer cells. *Biochem J* **2005,** *388* (Pt 3), 967-72.

239.   Lazarowitz, S. G.; Choppin, P. W., Enhancement of the infectivity of influenza A and B viruses by proteolytic cleavage of the hemagglutinin polypeptide. *Virology* **1975,** *68* (2), 440-54.

240.   Klenk, H. D.; Rott, R.; Orlich, M.; Blödorn, J., Activation of influenza A viruses by trypsin treatment. *Virology* **1975,** *68* (2), 426-39.

241.   Shirogane, Y.; Takeda, M.; Iwasaki, M.; Ishiguro, N.; Takeuchi, H.; Nakatsu, Y.; Tahara, M.; Kikuta, H.; Yanagi, Y., Efficient multiplication of human metapneumovirus in Vero cells expressing the transmembrane serine protease TMPRSS2. *J Virol* **2008,** *82* (17), 8942-6.

242.   van den Hoogen, B. G.; Bestebroer, T. M.; Osterhaus, A. D.; Fouchier, R. A., Analysis of the genomic sequence of a human metapneumovirus. *Virology* **2002,** *295* (1), 119-32.

243.   Shirato, K.; Kawase, M.; Matsuyama, S., Middle East respiratory syndrome coronavirus infection mediated by the transmembrane serine protease TMPRSS2. *J Virol* **2013,** *87* (23), 12552-61.

244.   Shirato, K.; Matsuyama, S.; Ujike, M.; Taguchi, F., Role of proteases in the release of porcine epidemic diarrhea virus from infected cells. *J Virol* **2011,** *85* (15), 7872-80.

245.   Li, F.; Berardi, M.; Li, W.; Farzan, M.; Dormitzer, P. R.; Harrison, S. C., Conformational states of the severe acute respiratory syndrome coronavirus spike protein ectodomain. *J Virol* **2006,** *80* (14), 6794-800.

246.   Belouzard, S.; Chu, V. C.; Whittaker, G. R., Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proc Natl Acad Sci U S A* **2009,** *106* (14), 5871-6.

247.   Bosch, B. J.; van der Zee, R.; de Haan, C. A.; Rottier, P. J., The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J Virol* **2003,** *77* (16), 8801-11.

248.   Kam, Y. W.; Okumura, Y.; Kido, H.; Ng, L. F.; Bruzzone, R.; Altmeyer, R., Cleavage of the SARS coronavirus spike glycoprotein by airway proteases enhances virus entry into human bronchial epithelial cells in vitro. *PLoS One* **2009,** *4* (11), e7870.

249.   Glowacka, I.; Bertram, S.; Müller, M. A.; Allen, P.; Soilleux, E.; Pfefferle, S.; Steffen, I.; Tsegaye, T. S.; He, Y.; Gnirss, K.; Niemeyer, D.; Schneider, H.; Drosten, C.; Pöhlmann, S., Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. *J Virol* **2011,** *85* (9), 4122-34.

250.   Perdue, M. L.; García, M.; Senne, D.; Fraire, M., Virulence-associated sequence duplication at the hemagglutinin cleavage site of avian influenza viruses. *Virus Res* **1997,** *49* (2), 173-86.

251.   Tian, X.; Li, C.; Huang, A.; Xia, S.; Lu, S.; Shi, Z.; Lu, L.; Jiang, S.; Yang, Z.; Wu, Y.; Ying, T., Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. *Emerg Microbes Infect* **2020,** *9* (1), 382-385.

252.   Li, F.; Li, W.; Farzan, M.; Harrison, S. C., Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* **2005,** *309* (5742), 1864-8.

253.   Beniac, D. R.; deVarennes, S. L.; Andonov, A.; He, R.; Booth, T. F., Conformational reorganization of the SARS coronavirus spike following receptor binding: implications for membrane fusion. *PLoS One* **2007,** *2* (10), e1082.

254.   Nicholls, J. M.; Bourne, A. J.; Chen, H.; Guan, Y.; Peiris, J. S., Sialic acid receptor detection in the human respiratory tract: evidence for widespread distribution of potential binding sites for human and avian influenza viruses. *Respir Res* **2007,** *8*, 73.

255.   Bertram, S.; Heurich, A.; Lavender, H.; Gierer, S.; Danisch, S.; Perin, P.; Lucas, J. M.; Nelson, P. S.; Pohlmann, S.; Soilleux, E. J., Influenza and SARS-coronavirus activating proteases TMPRSS2 and HAT are expressed at multiple sites in human respiratory and gastrointestinal tracts. *PLoS One* **2012,** *7* (4), e35876.

256.   Iwata-Yoshikawa, N.; Okamura, T.; Shimizu, Y.; Hasegawa, H.; Takeda, M.; Nagata, N., TMPRSS2 Contributes to Virus Spread and Immunopathology in the Airways of Murine Models after Coronavirus Infection. *J Virol* **2019,** *93* (6).

257.   Montopoli, M.; Zumerle, S.; Vettor, R.; Rugge, M.; Zorzi, M.; Catapano, C. V.; Carbone, G. M.; Cavalli, A.; Pagano, F.; Ragazzi, E.; Prayer-Galetti, T.; Alimonti, A., Androgen-deprivation therapies for prostate cancer and risk of infection by SARS-CoV-2: a population-based study (N = 4532). *Ann Oncol* **2020,** *31* (8), 1040-1045.

258.   Hawthorne, S.; Hamilton, R.; Walker, B. J.; Walker, B., Utilization of biotinylated diphenyl phosphonates for disclosure of serine proteases. *Anal Biochem* **2004,** *326* (2), 273-5.

259.   Böttcher, E.; Freuer, C.; Steinmetzer, T.; Klenk, H. D.; Garten, W., MDCK cells that express proteases TMPRSS2 and HAT provide a cell system to propagate influenza viruses in the absence of trypsin and to study cleavage of HA and its inhibition. *Vaccine* **2009,** *27* (45), 6324-9.

260.   Zhirnov, O. P.; Klenk, H. D.; Wright, P. F., Aprotinin and similar protease inhibitors as drugs against influenza. *Antiviral Res* **2011,** *92* (1), 27-36.

261.   Dittmann, M.; Hoffmann, H. H.; Scull, M. A.; Gilmore, R. H.; Bell, K. L.; Ciancanelli, M.; Wilson, S. J.; Crotta, S.; Yu, Y.; Flatley, B.; Xiao, J. W.; Casanova, J. L.; Wack, A.; Bieniasz, P. D.; Rice, C. M., A serpin shapes the extracellular environment to prevent influenza A virus maturation. *Cell* **2015,** *160* (4), 631-643.

262.   Faller, N.; Gautschi, I.; Schild, L., Functional analysis of a missense mutation in the serine protease inhibitor SPINT2 associated with congenital sodium diarrhea. *PLoS One* **2014,** *9* (4), e94267.

263.   Szabo, R.; Hobson, J. P.; List, K.; Molinolo, A.; Lin, C. Y.; Bugge, T. H., Potent inhibition and global co-localization implicate the transmembrane Kunitz-type serine protease inhibitor

hepatocyte growth factor activator inhibitor-2 in the regulation of epithelial matriptase activity. *J Biol Chem* **2008,** *283* (43), 29495-504.

264.   Sielaff, F.; Bottcher-Friebertshauser, E.; Meyer, D.; Saupe, S. M.; Volk, I. M.; Garten, W.; Steinmetzer, T., Development of substrate analogue inhibitors for the human airway trypsin-like protease HAT. *Bioorg Med Chem Lett* **2011,** *21* (16), 4860-4.

265.   Sisay, M. T.; Steinmetzer, T.; Stirnberg, M.; Maurer, E.; Hammami, M.; Bajorath, J.; Gütschow, M., Identification of the first low-molecular-weight inhibitors of matriptase-2. *J Med Chem* **2010,** *53* (15), 5523-35.

266.   Biela, A.; Sielaff, F.; Terwesten, F.; Heine, A.; Steinmetzer, T.; Klebe, G., Ligand binding stepwise disrupts water network in thrombin: enthalpic and entropic changes reveal classical hydrophobic effect. *J Med Chem* **2012,** *55* (13), 6094-110.

267.   Steinmetzer, T.; Schweinitz, A.; Stürzebecher, A.; Dönnecke, D.; Uhland, K.; Schuster, O.; Steinmetzer, P.; Müller, F.; Friedrich, R.; Than, M. E.; Bode, W.; Stürzebecher, J., Secondary amides of sulfonylated 3-amidinophenylalanine. New potent and selective inhibitors of matriptase. *J Med Chem* **2006,** *49* (14), 4116-26.

268.   Hammami, M.; Rühmann, E.; Maurer, E.; Heine, A.; Gütschow, M.; Klebe, G.; Steinmetzer, T., New 3-amidinophenylalanine-derived inhibitors of matriptase. *MedChemComm* **2012,** *3* (7), 807-813.

269.   Steinmetzer, T.; Dönnecke, D.; Korsonewski, M.; Neuwirth, C.; Steinmetzer, P.; Schulze, A.; Saupe, S. M.; Schweinitz, A., Modification of the N-terminal sulfonyl residue in 3-amidinophenylalanine-based matriptase inhibitors. *Bioorg Med Chem Lett* **2009,** *19* (1), 67-73.

270.   Lucas, J. M.; Heinlein, C.; Kim, T.; Hernandez, S. A.; Malik, M. S.; True, L. D.; Morrissey, C.; Corey, E.; Montgomery, B.; Mostaghel, E.; Clegg, N.; Coleman, I.; Brown, C. M.; Schneider, E. L.; Craik, C.; Simon, J. A.; Bedalov, A.; Nelson, P. S., The androgen-regulated protease TMPRSS2 activates a proteolytic cascade involving components of the tumor microenvironment and promotes prostate cancer metastasis. *Cancer Discov* **2014,** *4* (11), 1310-25.

271.   Shrimp, J. H.; Kales, S. C.; Sanderson, P. E.; Simeonov, A.; Shen, M.; Hall, M. D., An Enzymatic TMPRSS2 Assay for Assessment of Clinical Candidates and Discovery of Inhibitors as Potential Treatment of COVID-19. *bioRxiv* **2020**.

272.   Matsuyama, S.; Nagata, N.; Shirato, K.; Kawase, M.; Takeda, M.; Taguchi, F., Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2. *J Virol* **2010,** *84* (24), 12658-64.

273.   Simmons, G.; Reeves, J. D.; Rennekamp, A. J.; Amberg, S. M.; Piefer, A. J.; Bates, P., Characterization of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry. *Proc Natl Acad Sci U S A* **2004,** *101* (12), 4240-5.

274.   Talukdar, R.; Tandon, R. K., Pancreatic stellate cells: new target in the treatment of chronic pancreatitis. *J Gastroenterol Hepatol* **2008,** *23* (1), 34-41.

275.   Sai, J. K.; Suyama, M.; Kubokawa, Y.; Matsumura, Y.; Inami, K.; Watanabe, S., Efficacy of camostat mesilate against dyspepsia associated with non-alcoholic mild pancreatic disease. *J Gastroenterol* **2010,** *45* (3), 335-41.

276.   Okuno, M.; Kojima, S.; Akita, K.; Matsushima-Nishiwaki, R.; Adachi, S.; Sano, T.; Takano, Y.; Takai, K.; Obora, A.; Yasuda, I.; Shiratori, Y.; Okano, Y.; Shimada, J.; Suzuki, Y.; Muto, Y.; Moriwaki, Y., Retinoids in liver fibrosis and cancer. *Front Biosci* **2002,** *7*, d204-18.

277.   Shirato, K.; Kawase, M.; Matsuyama, S., Wild-type human coronaviruses prefer cell-surface TMPRSS2 to endosomal cathepsins for cell entry. *Virology* **2018,** *517*, 9-15.

278.   Zhou, Y.; Vedantham, P.; Lu, K.; Agudelo, J.; Carrion, R.; Nunneley, J. W.; Barnard, D.; Pöhlmann, S.; McKerrow, J. H.; Renslo, A. R.; Simmons, G., Protease inhibitors targeting coronavirus and filovirus entry. *Antiviral Res* **2015,** *116*, 76-84.

279.   Lee, M. G.; Kim, K. H.; Park, K. Y.; Kim, J. S., Evaluation of anti-influenza effects of camostat in mice infected with non-adapted human influenza viruses. *Arch Virol* **1996,** *141* (10), 1979-89.

280.   Yamaya, M.; Shimotai, Y.; Hatachi, Y.; Lusamba Kalonji, N.; Tando, Y.; Kitajima, Y.; Matsuo, K.; Kubo, H.; Nagatomi, R.; Hongo, S.; Homma, M.; Nishimura, H., The serine protease inhibitor camostat inhibits influenza virus replication and cytokine production in primary cultures of human tracheal epithelial cells. *Pulm Pharmacol Ther* **2015,** *33*, 66-74.

281.   Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N. H.; Nitsche, A.; Müller, M. A.; Drosten, C.; Pöhlmann, S., SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. (1097-4172 (Electronic)).

282.   Midgley, I.; Hood, A. J.; Proctor, P.; Chasseaud, L. F.; Irons, S. R.; Cheng, K. N.; Brindley, C. J.; Bonn, R., Metabolic fate of 14C-camostat mesylate in man, rat and dog after intravenous administration. *Xenobiotica* **1994,** *24* (1), 79-92.

283.   Yamamoto, M.; Matsuyama, S.; Li, X.; Takeda, M.; Kawaguchi, Y.; Inoue, J. I.; Matsuda, Z., Identification of Nafamostat as a Potent Inhibitor of Middle East Respiratory Syndrome Coronavirus S Protein-Mediated Membrane Fusion Using the Split-Protein-Based Cell-Cell Fusion Assay. *Antimicrob Agents Chemother* **2016,** *60* (11), 6532-6539.

284.   Chen, X.; Xu, Z.; Zeng, S.; Wang, X.; Liu, W.; Qian, L.; Wei, J.; Yang, X.; Shen, Q.; Gong, Z.; Yan, Y., The Molecular Aspect of Antitumor Effects of Protease Inhibitor Nafamostat Mesylate and Its Role in Potential Clinical Applications. *Front Oncol* **2019,** *9*, 852.

285.   Sadahiro, T.; Yuzawa, H.; Kimura, T.; Oguchi, M.; Morito, T.; Mizushima, S.; Hirose, Y., Current Practices in Acute Blood Purification Therapy in Japan and Topics for Further Study. *Contrib Nephrol* **2018,** *196*, 209-214.

286. Wang, M.; Cao, R.; Zhang, L.; Yang, X.; Liu, J.; Xu, M.; Shi, Z.; Hu, Z.; Zhong, W.; Xiao, G., Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res* **2020,** *30* (3), 269-271.

287. WHO, World Health Organization: 2021.

288. Ji, H. L.; Zhao, R.; Matalon, S.; Matthay, M. A., Elevated Plasmin(ogen) as a Common Risk Factor for COVID-19 Susceptibility. *Physiol Rev* **2020,** *100* (3), 1065-1075.

289. Li, G.; Liu, Y.; Jing, X.; Wang, Y.; Miao, M.; Tao, L.; Zhou, Z.; Xie, Y.; Huang, Y.; Lei, J.; Gong, G.; Jin, P.; Hao, Y.; Faria, N. R.; Clercq, E.; Zhang, M., Mortality risk of COVID-19 in elderly males with comorbidities: a multi-country study. *Aging (Albany NY)* **2020,** *13* (1), 27-60.

290. Tay, M. Z.; Poh, C. M.; Rénia, L.; MacAry, P. A.; Ng, L. F. P., The trinity of COVID-19: immunity, inflammation and intervention. *Nat Rev Immunol* **2020,** *20* (6), 363-374.

291. Gupta, A.; Madhavan, M. V.; Sehgal, K.; Nair, N.; Mahajan, S.; Sehrawat, T. S.; Bikdeli, B.; Ahluwalia, N.; Ausiello, J. C.; Wan, E. Y.; Freedberg, D. E.; Kirtane, A. J.; Parikh, S. A.; Maurer, M. S.; Nordvig, A. S.; Accili, D.; Bathon, J. M.; Mohan, S.; Bauer, K. A.; Leon, M. B.; Krumholz, H. M.; Uriel, N.; Mehra, M. R.; Elkind, M. S. V.; Stone, G. W.; Schwartz, A.; Ho, D. D.; Bilezikian, J. P.; Landry, D. W., Extrapulmonary manifestations of COVID-19. *Nat Med* **2020,** *26* (7), 1017-1032.

292. Hou, Y. J.; Okuda, K.; Edwards, C. E.; Martinez, D. R.; Asakura, T.; Dinnon, K. H.; Kato, T.; Lee, R. E.; Yount, B. L.; Mascenik, T. M.; Chen, G.; Olivier, K. N.; Ghio, A.; Tse, L. V.; Leist, S. R.; Gralinski, L. E.; Schäfer, A.; Dang, H.; Gilmore, R.; Nakano, S.; Sun, L.; Fulcher, M. L.; Livraghi-Butrico, A.; Nicely, N. I.; Cameron, M.; Cameron, C.; Kelvin, D. J.; de Silva, A.; Margolis, D. M.; Markmann, A.; Bartelt, L.; Zumwalt, R.; Martinez, F. J.; Salvatore, S. P.; Borczuk, A.; Tata, P. R.; Sontake, V.; Kimple, A.; Jaspers, I.; O'Neal, W. K.; Randell, S. H.; Boucher, R. C.; Baric, R. S., SARS-CoV-2 Reverse Genetics Reveals a Variable Infection Gradient in the Respiratory Tract. *Cell* **2020,** *182* (2), 429-446.e14.

293. Xiao, F.; Tang, M.; Zheng, X.; Liu, Y.; Li, X.; Shan, H., Evidence for Gastrointestinal Infection of SARS-CoV-2. *Gastroenterology* **2020,** *158* (6), 1831-1833.e3.

294. Lin, L.; Jiang, X.; Zhang, Z.; Huang, S.; Fang, Z.; Gu, Z.; Gao, L.; Shi, H.; Mai, L.; Liu, Y.; Lin, X.; Lai, R.; Yan, Z.; Li, X.; Shan, H., Gastrointestinal symptoms of 95 cases with SARS-CoV-2 infection. *Gut* **2020,** *69* (6), 997-1001.

295. Lindner, D.; Fitzek, A.; Bräuninger, H.; Aleshcheva, G.; Edler, C.; Meissner, K.; Scherschel, K.; Kirchhof, P.; Escher, F.; Schultheiss, H. P.; Blankenberg, S.; Püschel, K.; Westermann, D., Association of Cardiac Infection With SARS-CoV-2 in Confirmed COVID-19 Autopsy Cases. *JAMA Cardiol* **2020,** *5* (11), 1281-1285.

296. Sah, P.; Vilches, T. N.; Moghadas, S. M.; Fitzpatrick, M. C.; Singer, B. H.; Hotez, P. J.; Galvani, A. P., Accelerated vaccine rollout is imperative to mitigate highly transmissible COVID-19 variants. *EClinicalMedicine* **2021,** *35*.

297. Hoffmann, M.; Arora, P.; Groß, R.; Seidel, A.; Hörnich, B. F.; Hahn, A. S.; Krüger, N.; Graichen, L.; Hofmann-Winkler, H.; Kempf, A.; Winkler, M. S.; Schulz, S.; Jäck, H. M.; Jahrsdörfer, B.; Schrezenmeier, H.; Müller, M.; Kleger, A.; Münch, J.; Pöhlmann, S., SARS-CoV-2 variants B.1.351 and P.1 escape from neutralizing antibodies. *Cell* **2021,** *184* (9), 2384-2393.e12.

298. Wang, Z.; Schmidt, F.; Weisblum, Y.; Muecksch, F.; Barnes, C. O.; Finkin, S.; Schaefer-Babajew, D.; Cipolla, M.; Gaebler, C.; Lieberman, J. A.; Oliveira, T. Y.; Yang, Z.; Abernathy, M. E.; Huey-Tubman, K. E.; Hurley, A.; Turroja, M.; West, K. A.; Gordon, K.; Millard, K. G.; Ramos, V.; Da Silva, J.; Xu, J.; Colbert, R. A.; Patel, R.; Dizon, J.; Unson-O'Brien, C.; Shimeliovich, I.; Gazumyan, A.; Caskey, M.; Bjorkman, P. J.; Casellas, R.; Hatziioannou, T.; Bieniasz, P. D.; Nussenzweig, M. C., mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *bioRxiv* **2021**.

299. Frampton, D.; Rampling, T.; Cross, A.; Bailey, H.; Heaney, J.; Byott, M.; Scott, R.; Sconza, R.; Price, J.; Margaritis, M.; Bergstrom, M.; Spyer, M. J.; Miralhes, P. B.; Grant, P.; Kirk, S.; Valerio, C.; Mangera, Z.; Prabhahar, T.; Moreno-Cuesta, J.; Arulkumaran, N.; Singer, M.; Shin, G. Y.; Sanchez, E.; Paraskevopoulou, S. M.; Pillay, D.; McKendry, R. A.; Mirfenderesky, M.; Houlihan, C. F.; Nastouli, E., Genomic characteristics and clinical effect of the emergent SARS-CoV-2 B.1.1.7 lineage in London, UK: a whole-genome sequencing and hospital-based cohort study. *Lancet Infect Dis* **2021**.

300. Davies, N. G.; Abbott, S.; Barnard, R. C.; Jarvis, C. I.; Kucharski, A. J.; Munday, J. D.; Pearson, C. A. B.; Russell, T. W.; Tully, D. C.; Washburne, A. D.; Wenseleers, T.; Gimma, A.; Waites, W.; Wong, K. L. M.; van Zandvoort, K.; Silverman, J. D.; Diaz-Ordaz, K.; Keogh, R.; Eggo, R. M.; Funk, S.; Jit, M.; Atkins, K. E.; Edmunds, W. J.; Group, C. C.-W.; Consortium, C.-G. U. C.-U., Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* **2021,** *372* (6538).

301. Letko, M.; Marzi, A.; Munster, V., Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol* **2020,** *5* (4), 562-569.

302. Limburg, H.; Harbig, A.; Bestle, D.; Stein, D. A.; Moulton, H. M.; Jaeger, J.; Janga, H.; Hardes, K.; Koepke, J.; Schulte, L.; Koczulla, A. R.; Schmeck, B.; Klenk, H. D.; Bottcher-Friebertshauser, E., TMPRSS2 Is the Major Activating Protease of Influenza A Virus in Primary Human Airway Cells and Influenza B Virus in Human Type II Pneumocytes. *J Virol* **2019,** *93* (21).

303. Bertram, S.; Dijkman, R.; Habjan, M.; Heurich, A.; Gierer, S.; Glowacka, I.; Welsch, K.; Winkler, M.; Schneider, H.; Hofmann-Winkler, H.; Thiel, V.; Pöhlmann, S., TMPRSS2 activates the human coronavirus 229E for cathepsin-independent host cell entry and is expressed in viral target cells in the respiratory epithelium. *J Virol* **2013,** *87* (11), 6150-60.

304. Hatesuer, B.; Bertram, S.; Mehnert, N.; Bahgat, M. M.; Nelson, P. S.; Pohlmann, S.; Schughart, K., Tmprss2 is essential for influenza H1N1 virus pathogenesis in mice. *PLoS Pathog* **2013,** *9* (12), e1003774.

305.   Szabo, R.; Bugge, T. H., Membrane-anchored serine proteases in vertebrate cell and developmental biology. *Annu Rev Cell Dev Biol* **2011,** *27*, 213-35.

306.   Ramsay, A. J.; Hooper, J. D.; Folgueras, A. R.; Velasco, G.; Lopez-Otin, C., Matriptase-2 (TMPRSS6): a proteolytic regulator of iron homeostasis. *Haematologica* **2009,** *94* (6), 840-9.

307.   Wettstein, L.; Conzelmann, C.; Mueller, J. A.; Weil, T.; Gross, R.; Hirschenberger, M.; Seidel, A.; Klute, S.; Zech, F.; Preising, N., Alpha-1 antitrypsin inhibits SARS-CoV-2 infection. *BioRxiv* **2020**.

308.   Maggio, R.; Corsini, G. U., Repurposing the mucolytic cough suppressant and TMPRSS2 protease inhibitor bromhexine for the prevention and management of SARS-CoV-2 infection. *Pharmacol Res* **2020,** *157*, 104837.

309.   Depfenhart, M.; de Villiers, D.; Lemperle, G.; Meyer, M.; Di Somma, S., Potential new treatment strategies for COVID-19: is there a role for bromhexine as add-on therapy? *Intern Emerg Med* **2020,** *15* (5), 801-812.

310.   Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T., SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **2018,** *46* (W1), W296-W303.

311.   Beliveau, F.; Tarkar, A.; Dion, S. P.; Desilets, A.; Ghinet, M. G.; Boudreault, P. L.; St-Georges, C.; Marsault, E.; Paone, D.; Collins, J.; Macphee, C. H.; Campobasso, N.; Groy, A.; Cottom, J.; Ouellette, M.; Pope, A. J.; Leduc, R., Discovery and Development of TMPRSS6 Inhibitors Modulating Hepcidin Levels in Human Hepatocytes. *Cell Chem Biol* **2019,** *26* (11), 1559-1572 e9.

312.   Sterling, T.; Irwin, J. J., ZINC 15--Ligand Discovery for Everyone. *J Chem Inf Model* **2015,** *55* (11), 2324-37.

313.   Halgren, T. A., Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry* **1996,** *17* (5-6), 490-519.

314.   Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet, B. K., Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *J Med Chem* **2012,** *55* (14), 6582-94.

315.   Kovach, I. M.; Kelley, P.; Eddy, C.; Jordan, F.; Baykal, A., Proton bridging in the interactions of thrombin with small inhibitors. *Biochemistry* **2009,** *48* (30), 7296-304.

316.   Wienen, W.; Stassen, J. M.; Priepke, H.; Ries, U. J.; Hauel, N., In-vitro profile and ex-vivo anticoagulant activity of the direct thrombin inhibitor dabigatran and its orally active prodrug, dabigatran etexilate. *Thromb Haemost* **2007,** *98* (1), 155-62.

317.   Gustafsson, D.; Antonsson, T.; Bylund, R.; Eriksson, U.; Gyzander, E.; Nilsson, I.; Elg, M.; Mattsson, C.; Deinum, J.; Pehrsson, S.; Karlsson, O.; Nilsson, A.; Sörensen, H., Effects of

melagatran, a new low-molecular-weight thrombin inhibitor, on thrombin and fibrinolytic enzymes. *Thromb Haemost* **1998,** *79* (1), 110-8.

318.   Bertram, S.; Glowacka, I.; Blazejewska, P.; Soilleux, E.; Allen, P.; Danisch, S.; Steffen, I.; Choi, S. Y.; Park, Y.; Schneider, H.; Schughart, K.; Pöhlmann, S., TMPRSS2 and TMPRSS4 facilitate trypsin-independent spread of influenza virus in Caco-2 cells. *J Virol* **2010,** *84* (19), 10016-25.

319.   Conzelmann, C.; Gilg, A.; Groß, R.; Schütz, D.; Preising, N.; Ständker, L.; Jahrsdörfer, B.; Schrezenmeier, H.; Sparrer, K. M. J.; Stamminger, T.; Stenger, S.; Münch, J.; Müller, J. A., An enzyme-based immunodetection assay to quantify SARS-CoV-2 infection. *Antiviral Res* **2020,** *181*, 104882.

320.   Yurkovetskiy, L.; Wang, X.; Pascal, K. E.; Tomkins-Tinch, C.; Nyalile, T. P.; Wang, Y.; Baum, A.; Diehl, W. E.; Dauphin, A.; Carbone, C.; Veinotte, K.; Egri, S. B.; Schaffner, S. F.; Lemieux, J. E.; Munro, J. B.; Rafique, A.; Barve, A.; Sabeti, P. C.; Kyratsous, C. A.; Dudkina, N. V.; Shen, K.; Luban, J., Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. *Cell* **2020,** *183* (3), 739-751.e8.

321.   Colombo, E.; Désilets, A.; Duchêne, D.; Chagnon, F.; Najmanovich, R.; Leduc, R.; Marsault, E., Design and synthesis of potent, selective inhibitors of matriptase. *ACS Med Chem Lett* **2012,** *3* (7), 530-4.

322.   Hoffmann, M.; Hofmann-Winkler, H.; Smith, J. C.; Krüger, N.; Sørensen, L. K.; Søgaard, O. S.; Hasselstrøm, J. B.; Winkler, M.; Hempel, T.; Raich, L.; Olsson, S.; Yamazoe, T.; Yamatsuta, K.; Mizuno, H.; Ludwig, S.; Noé, F.; Sheltzer, J. M.; Kjolby, M.; Pöhlmann, S., Camostat mesylate inhibits SARS-CoV-2 activation by TMPRSS2-related proteases and its metabolite GBPA exerts antiviral activity. *bioRxiv* **2020**.

323.   Zhou, D.; Dejnirattisai, W.; Supasa, P.; Liu, C.; Mentzer, A. J.; Ginn, H. M.; Zhao, Y.; Duyvesteyn, H. M. E.; Tuekprakhon, A.; Nutalai, R.; Wang, B.; Paesen, G. C.; Lopez-Camacho, C.; Slon-Campos, J.; Hallis, B.; Coombes, N.; Bewley, K.; Charlton, S.; Walter, T. S.; Skelly, D.; Lumley, S. F.; Dold, C.; Levin, R.; Dong, T.; Pollard, A. J.; Knight, J. C.; Crook, D.; Lambe, T.; Clutterbuck, E.; Bibi, S.; Flaxman, A.; Bittaye, M.; Belij-Rammerstorfer, S.; Gilbert, S.; James, W.; Carroll, M. W.; Klenerman, P.; Barnes, E.; Dunachie, S. J.; Fry, E. E.; Mongkolsapaya, J.; Ren, J.; Stuart, D. I.; Screaton, G. R., Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine-induced sera. *Cell* **2021,** *184* (9), 2348-2361.e6.

324.   Nguyen, L. T.; Chau, J. K.; Perry, N. A.; de Boer, L.; Zaat, S. A.; Vogel, H. J., Serum stabilities of short tryptophan- and arginine-rich antimicrobial peptide analogs. *PLoS One* **2010,** *5* (9).

325.   Li, H.; Kem, D. C.; Zhang, L.; Huang, B.; Liles, C.; Benbrook, A.; Gali, H.; Veitla, V.; Scherlag, B. J.; Cunningham, M. W.; Yu, X., Novel retro-inverso peptide inhibitor reverses angiotensin receptor autoantibody-induced hypertension in the rabbit. *Hypertension* **2015,** *65* (4), 793-9.

326.   Hou, Y. J.; Okuda, K.; Edwards, C. E.; Martinez, D. R.; Asakura, T.; Dinnon Iii, K. H.; Kato, T.; Lee, R. E.; Yount, B. L.; Mascenik, T. M., SARS-CoV-2 reverse genetics reveals a variable infection gradient in the respiratory tract. *Cell* **2020,** *182* (2), 429-446.

327.   Vareille, M.; Kieninger, E.; Edwards, M. R.; Regamey, N., The airway epithelium: soldier in the fight against respiratory viruses. *Clinical microbiology reviews* **2011,** *24* (1), 210.

328.   Rogan, M. P.; Geraghty, P.; Greene, C. M.; O'Neill, S. J.; Taggart, C. C.; McElvaney, N. G., Antimicrobial proteins and polypeptides in pulmonary innate defence. *Respiratory research* **2006,** *7* (1), 1-11.

329.   Münch, J.; Ständker, L.; Adermann, K.; Schulz, A.; Schindler, M.; Chinnadurai, R.; Pöhlmann, S.; Chaipan, C.; Biet, T.; Peters, T., Discovery and optimization of a natural HIV-1 entry inhibitor targeting the gp41 fusion peptide. *Cell* **2007,** *129* (2), 263-275.

330.   Borst, E. M.; Ständker, L.; Wagner, K.; Schulz, T. F.; Forssmann, W.-G.; Messerle, M., A peptide inhibitor of cytomegalovirus infection from human hemofiltrate. *Antimicrobial agents and chemotherapy* **2013,** *57* (10), 4751.

331.   Groß, R.; Bauer, R.; Krüger, F.; Rücker-Braun, E.; Olari, L.-R.; Ständker, L.; Preising, N.; Rodríguez, A. A.; Conzelmann, C.; Gerbl, F., A placenta derived C-terminal fragment of β-hemoglobin with combined antibacterial and antiviral activity. *Frontiers in microbiology* **2020,** *11*, 508.

332.   Forssmann, W. G.; The, Y. H.; Stoll, M.; Adermann, K.; Albrecht, U.; Tillmann, H. C.; Barlos, K.; Busmann, A.; Canales-Mayordomo, A.; Giménez-Gallego, G.; Hirsch, J.; Jiménez-Barbero, J.; Meyer-Olson, D.; Münch, J.; Pérez-Castells, J.; Ständker, L.; Kirchhoff, F.; Schmidt, R. E., Short-term monotherapy in HIV-infected patients with a virus entry inhibitor against the gp41 fusion peptide. *Sci Transl Med* **2010,** *2* (63), 63re3.

333.   Bienert, S.; Waterhouse, A.; de Beer, T. A.; Tauriello, G.; Studer, G.; Bordoli, L.; Schwede, T., The SWISS-MODEL Repository-new features and functionality. *Nucleic Acids Res* **2017,** *45* (D1), D313-D319.

334.   Pearce, M. C.; Morton, C. J.; Feil, S. C.; Hansen, G.; Adams, J. J.; Parker, M. W.; Bottomley, S. P., Preventing serpin aggregation: the molecular mechanism of citrate action upon antitrypsin unfolding. *Protein Sci* **2008,** *17* (12), 2127-33.

335.   van Zundert, G. C. P.; Rodrigues, J. P. G. L.; Trellet, M.; Schmitz, C.; Kastritis, P. L.; Karaca, E.; Melquiond, A. S. J.; van Dijk, M.; de Vries, S. J.; Bonvin, A. M. J. J., The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol* **2016,** *428* (4), 720-725.

336.   Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E., Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys* **2020,** *153* (4), 044130.

337.   Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.;   Kalé, L.; Schulten, K., Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* **2005,** *26* (16), 1781-1802.

338.   Vanommeslaeghe, K.; MacKerell, A. D., CHARMM additive and polarizable force fields for biophysics and computer-aided drug design. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2015,** *1850* (5), 861-871.

339.   Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D.; Pastor, R. W., Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *The Journal of Physical Chemistry B* **2010,** *114* (23), 7830-7843.

340.   Humphrey, W.; Dalke, A.; Schulten, K., VMD: visual molecular dynamics. *J Mol Graph* **1996,** *14* (1), 33-8, 27-8.

341.   Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983,** *79* (2), 926-935.

342.   Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R., Constant pressure molecular dynamics simulation: The Langevin piston method. *The Journal of Chemical Physics* **1995,** *103* (11), 4613-4621.

343.   Martyna, G. J.; Tobias, D. J.; Klein, M. L., Constant pressure molecular dynamics algorithms. *Journal of Chemical Physics* **1994,** *101*, 4177-4189.

344.   Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993,** *98* (12), 10089-10092.

345.   Berger Rentsch, M.; Zimmer, G., A vesicular stomatitis virus replicon-based bioassay for the rapid and sensitive determination of multi-species type I interferon. *PLoS One* **2011,** *6* (10), e25858.

346.   Ruiz-Blanco, Y. B.; Sanchez-Garcia, E., CL-FEP: An End-State Free Energy Perturbation Approach. *J Chem Theory Comput* **2020,** *16* (3), 1396-1410.

347.   Dementiev, A.; Simonovic, M.; Volz, K.; Gettins, P. G., Canonical inhibitor-like interactions explain reactivity of alpha1-proteinase inhibitor Pittsburgh and antithrombin with proteinases. *J Biol Chem* **2003,** *278* (39), 37881-7.

348.   Xia, S.; Yan, L.; Xu, W.; Agrawal, A. S.; Algaissi, A.; Tseng, C.-T. K.; Wang, Q.; Du, L.; Tan, W.; Wilson, I. A., A pan-coronavirus fusion inhibitor targeting the HR1 domain of human coronavirus spike. *Science advances* **2019,** *5* (4), eaav4580.

349.   Carrell, R. W.; Jeppsson, J.-O.; Laurell, C.-B.; Brennan, S. O.; Owen, M. C.; Vaughan, L.; Boswell, D. R., Structure and variation of human α1–antitrypsin. *Nature* **1982,** *298* (5872), 329-334.

350. Travis, J.; Salvesen, G. S., HUMAN PLASMA PROTEINASE INHIBITORS. *Annual Review of Biochemistry* **1983,** *52* (1), 655-709.

351. Janciauskiene, S.; Welte, T., Well-Known and Less Well-Known Functions of Alpha-1 Antitrypsin. Its Role in Chronic Obstructive Pulmonary Disease and Other Disease Developments. *Ann Am Thorac Soc* **2016,** *13 Suppl 4*, S280-8.

352. Gadek, J. E.; Klein, H. G.; Holland, P. V.; Crystal, R. G., Replacement therapy of alpha 1-antitrypsin deficiency. Reversal of protease-antiprotease imbalance within the alveolar structures of PiZ subjects. *J Clin Invest* **1981,** *68* (5), 1158-65.

353. Baric, R. S., Emergence of a Highly Fit SARS-CoV-2 Variant. *N Engl J Med* **2020,** *383* (27), 2684-2686.

354. Korber, B.; Fischer, W. M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E. E.; Bhattacharya, T.; Foley, B.; Hastie, K. M.; Parker, M. D.; Partridge, D. G.; Evans, C. M.; Freeman, T. M.; de Silva, T. I.; McDanal, C.; Perez, L. G.; Tang, H.; Moon-Walker, A.; Whelan, S. P.; LaBranche, C. C.; Saphire, E. O.; Montefiori, D. C.; Group, S. C.-G., Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **2020,** *182* (4), 812-827.e19.

355. Bar-On, Y. M.; Flamholz, A.; Phillips, R.; Milo, R., SARS-CoV-2 (COVID-19) by the numbers. *Elife* **2020,** *9*.

356. Pizzorno, A.; Padey, B.; Julien, T.; Trouillet-Assant, S.; Traversier, A.; Errazuriz-Cerda, E.; Fouret, J.; Dubois, J.; Gaymard, A.; Lescure, F. X.; Dulière, V.; Brun, P.; Constant, S.; Poissy, J.; Lina, B.; Yazdanpanah, Y.; Terrier, O.; Rosa-Calatrava, M., Characterization and Treatment of SARS-CoV-2 in Nasal and Bronchial Human Airway Epithelia. *Cell Rep Med* **2020,** *1* (4), 100059.

357. Krüger, J.; Groß, R.; Conzelmann, C.; Müller, J. A.; Koepke, L.; Sparrer, K. M. J.; Weil, T.; Schütz, D.; Seufferlein, T.; Barth, T. F. E.; Stenger, S.; Heller, S.; Münch, J.; Kleger, A., Drug Inhibition of SARS-CoV-2 Replication in Human Pluripotent Stem Cell-Derived Intestinal Organoids. *Cell Mol Gastroenterol Hepatol* **2021,** *11* (4), 935-948.

358. Park, T. J.; Mitchell, B. J.; Abitua, P. B.; Kintner, C.; Wallingford, J. B., Dishevelled controls apical docking and planar polarization of basal bodies in ciliated epithelial cells. *Nat Genet* **2008,** *40* (7), 871-9.

359. Azouz, N. P.; Klingler, A. M.; Callahan, V.; Akhrymuk, I. V.; Elez, K.; Raich, L.; Henry, B. M.; Benoit, J. L.; Benoit, S. W.; Noé, F.; Kehn-Hall, K.; Rothenberg, M. E., Alpha 1 Antitrypsin is an Inhibitor of the SARS-CoV-2-Priming Protease TMPRSS2. *Pathog Immun* **2021,** *6* (1), 55-74.

360. Stratikos, E.; Gettins, P. G., Major proteinase movement upon stable serpin-proteinase complex formation. *Proc Natl Acad Sci U S A* **1997,** *94* (2), 453-8.

361.   Oguntuyo, K. Y.; Stevens, C. S.; Siddiquey, M. N.; Schilke, R. M.; Woolard, M. D.; Zhang, H.; Acklin, J. A.; Ikegame, S.; Hung, C. T.; Lim, J. K.; Cross, R. W.; Geisbert, T. W.; Ivanov, S. S.; Kamil, J. P.; Lee, B., In plain sight: the role of alpha-1-antitrypsin in COVID-19 pathogenesis and therapeutics. *bioRxiv* **2020**.

362.   McElvaney, O. J.; McEvoy, N. L.; McElvaney, O. F.; Carroll, T. P.; Murphy, M. P.; Dunlea, D. M.; Ní Choileáin, O.; Clarke, J.; O'Connor, E.; Hogan, G.; Ryan, D.; Sulaiman, I.; Gunaratnam, C.; Branagan, P.; O'Brien, M. E.; Morgan, R. K.; Costello, R. W.; Hurley, K.; Walsh, S.; de Barra, E.; McNally, C.; McConkey, S.; Boland, F.; Galvin, S.; Kiernan, F.; O'Rourke, J.; Dwyer, R.; Power, M.; Geoghegan, P.; Larkin, C.; O'Leary, R. A.; Freeman, J.; Gaffney, A.; Marsh, B.; Curley, G. F.; McElvaney, N. G., Characterization of the Inflammatory Response to Severe COVID-19 Illness. *Am J Respir Crit Care Med* **2020,** *202* (6), 812-821.

363.   Wewers, M. D.; Casolaro, M. A.; Crystal, R. G., Comparison of alpha-1-antitrypsin levels and antineutrophil elastase capacity of blood and lung in a patient with the alpha-1-antitrypsin phenotype null-null before and during alpha-1-antitrypsin augmentation therapy. *Am Rev Respir Dis* **1987,** *135* (3), 539-43.

364.   Crystal, R. G., Alpha 1-antitrypsin deficiency, emphysema, and liver disease. Genetic basis and strategies for therapy. *J Clin Invest* **1990,** *85* (5), 1343-52.

365.   Hubbard, R. C.; Sellers, S.; Czerski, D.; Stephens, L.; Crystal, R. G., Biochemical efficacy and safety of monthly augmentation therapy for alpha 1-antitrypsin deficiency. *JAMA* **1988,** *260* (9), 1259-64.

366.   Ko, C. J.; Huang, C. C.; Lin, H. Y.; Juan, C. P.; Lan, S. W.; Shyu, H. Y.; Wu, S. R.; Hsiao, P. W.; Huang, H. P.; Shun, C. T.; Lee, M. S., Androgen-Induced TMPRSS2 Activates Matriptase and Promotes Extracellular Matrix Degradation, Prostate Cancer Cell Invasion, Tumor Growth, and Metastasis. *Cancer Res* **2015,** *75* (14), 2949-60.

367.   Gierer, S.; Bertram, S.; Kaup, F.; Wrensch, F.; Heurich, A.; Krämer-Kühl, A.; Welsch, K.; Winkler, M.; Meyer, B.; Drosten, C.; Dittmer, U.; von Hahn, T.; Simmons, G.; Hofmann, H.; Pöhlmann, S., The spike protein of the emerging betacoronavirus EMC uses a novel coronavirus receptor for entry, can be activated by TMPRSS2, and is targeted by neutralizing antibodies. *J Virol* **2013,** *87* (10), 5502-11.

368.   Hatesuer, B.; Bertram, S.; Mehnert, N.; Bahgat, M. M.; Nelson, P. S.; Pöhlmann, S.; Pöhlman, S.; Schughart, K., Tmprss2 is essential for influenza H1N1 virus pathogenesis in mice. *PLoS Pathog* **2013,** *9* (12), e1003774.

369.   Limburg, H.; Harbig, A.; Bestle, D.; Stein, D. A.; Moulton, H. M.; Jaeger, J.; Janga, H.; Hardes, K.; Koepke, J.; Schulte, L.; Koczulla, A. R.; Schmeck, B.; Klenk, H. D.; Böttcher-Friebertshäuser, E., TMPRSS2 Is the Major Activating Protease of Influenza A Virus in Primary Human Airway Cells and Influenza B Virus in Human Type II Pneumocytes. *J Virol* **2019,** *93* (21).

370.   Greene, C. M.; Miller, S. D.; Carroll, T.; McLean, C.; O'Mahony, M.; Lawless, M. W.; O'Neill, S. J.; Taggart, C. C.; McElvaney, N. G., Alpha-1 antitrypsin deficiency: a conformational disease associated with lung and liver manifestations. *J Inherit Metab Dis* **2008,** *31* (1), 21-34.

371. Hu, J.; He, C.-L.; Gao, Q.-Z.; Zhang, G.-J.; Cao, X.-X.; Long, Q.-X.; Deng, H.-J.; Huang, L.-Y.; Chen, J.; Wang, K.; Tang, N.; Huang, A.-L., The D614G mutation of SARS-CoV-2 spike protein enhances viral infectivity and decreases neutralization sensitivity to individual convalescent sera. *bioRxiv* **2020**, 2020.06.20.161323.

372. Bhattacharyya, C.; Das, C.; Ghosh, A.; Singh, A. K.; Mukherjee, S.; Majumder, P. P.; Basu, A.; Biswas, N. K., Global Spread of SARS-CoV-2 Subtype with Spike Protein Mutation D614G is Shaped by Human Genomic Variations that Regulate Expression of TMPRSS2 and MX1 Genes. *bioRxiv* **2020**, 2020.05.04.075911.

373. Barnes, B. J.; Adrover, J. M.; Baxter-Stoltzfus, A.; Borczuk, A.; Cools-Lartigue, J.; Crawford, J. M.; Daßler-Plenker, J.; Guerci, P.; Huynh, C.; Knight, J. S.; Loda, M.; Looney, M. R.; McAllister, F.; Rayes, R.; Renaud, S.; Rousseau, S.; Salvatore, S.; Schwartz, R. E.; Spicer, J. D.; Yost, C. C.; Weber, A.; Zuo, Y.; Egeblad, M., Targeting potential drivers of COVID-19: Neutrophil extracellular traps. *J Exp Med* **2020,** *217* (6).

374. Leppkes, M.; Knopf, J.; Naschberger, E.; Lindemann, A.; Singh, J.; Herrmann, I.; Stürzl, M.; Staats, L.; Mahajan, A.; Schauer, C.; Kremer, A. N.; Völkl, S.; Amann, K.; Evert, K.; Falkeis, C.; Wehrfritz, A.; Rieker, R. J.; Hartmann, A.; Kremer, A. E.; Neurath, M. F.; Muñoz, L. E.; Schett, G.; Herrmann, M., Vascular occlusion by neutrophil extracellular traps in COVID-19. *EBioMedicine* **2020,** *58*, 102925.

375. Hubbard, R. C.; McElvaney, N. G.; Sellers, S. E.; Healy, J. T.; Czerski, D. B.; Crystal, R. G., Recombinant DNA-produced alpha 1-antitrypsin administered by aerosol augments lower respiratory tract antineutrophil elastase defenses in individuals with alpha 1-antitrypsin deficiency. *J Clin Invest* **1989,** *84* (4), 1349-54.

376. Campos, M. A.; Geraghty, P.; Holt, G.; Mendes, E.; Newby, P. R.; Ma, S.; Luna-Diaz, L. V.; Turino, G. M.; Stockley, R. A., The Biological Effects of Double-Dose Alpha-1 Antitrypsin Augmentation Therapy. A Pilot Clinical Trial. *Am J Respir Crit Care Med* **2019,** *200* (3), 318-326.

377. Campos, M. A.; Kueppers, F.; Stocks, J. M.; Strange, C.; Chen, J.; Griffin, R.; Wang-Smith, L.; Brantly, M. L., Safety and pharmacokinetics of 120 mg/kg versus 60 mg/kg weekly intravenous infusions of alpha-1 proteinase inhibitor in alpha-1 antitrypsin deficiency: a multicenter, randomized, double-blind, crossover study (SPARK). *COPD* **2013,** *10* (6), 687-95.

378. de Loyola, M. B.; Dos Reis, T. T. A.; de Oliveira, G. X. L. M.; da Fonseca Palmeira, J.; Argañaraz, G. A.; Argañaraz, E. R., Alpha-1-antitrypsin: A possible host protective factor against Covid-19. *Rev Med Virol* **2021,** *31* (2), e2157.

379. Bai, X.; Hippensteel, J.; Leavitt, A.; Maloney, J. P.; Beckham, D.; Garcia, C.; Li, Q.; Freed, B. M.; Ordway, D.; Sandhaus, R. A.; Chan, E. D., Hypothesis: Alpha-1-antitrypsin is a promising treatment option for COVID-19. *Med Hypotheses* **2021,** *146*, 110394.

Appendix

reference peptide RQLRVVGG

P1 peptide AEQRVVGG

P2 peptide GQRRVVGG

P3 peptide RNHRVVGG

P4 peptide QQVRVVGG

P1' peptide RQLRVHSL

P2' peptide RQLRVGRH

P3' peptide RQLRVGGQ

P3.1' peptide rQLRVVGG

P3.2' peptide RqLRVGGQ

P3.4' peptide RQLRVGGq
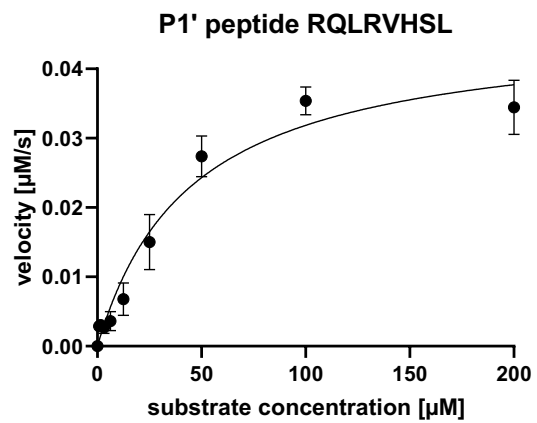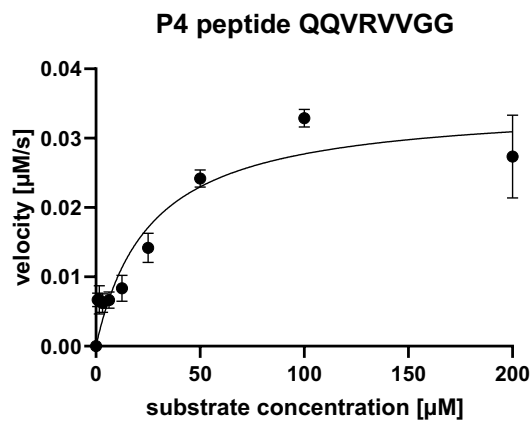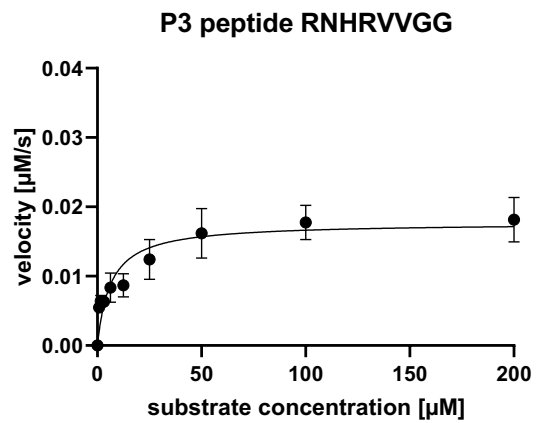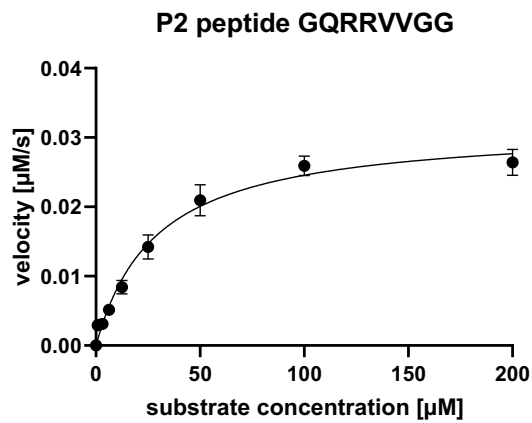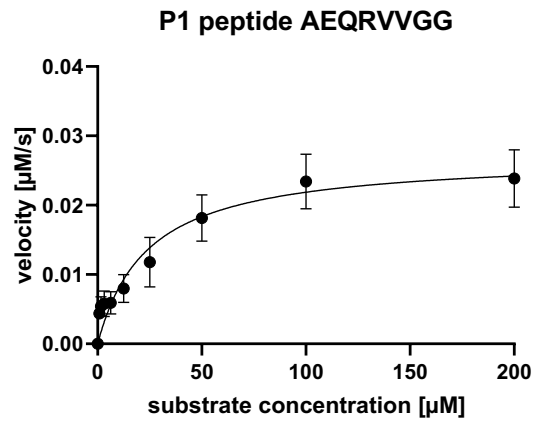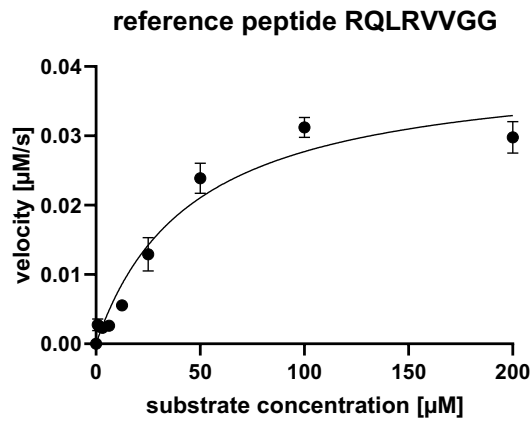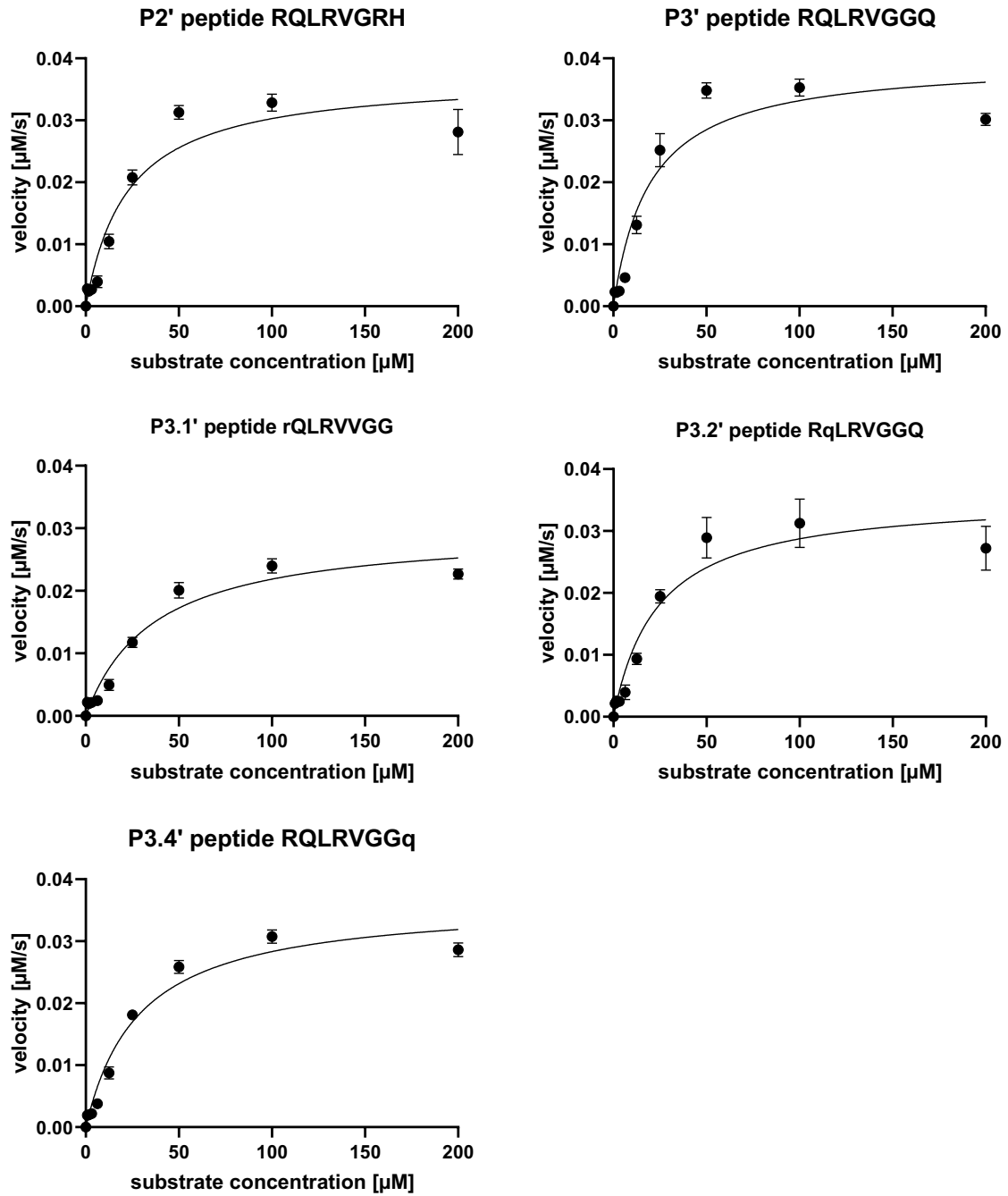
Figure A1. Michaelis Menten analysis of IQF substrate degradation by hepsin.
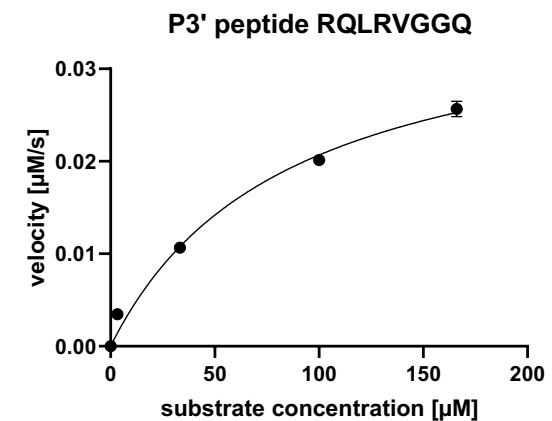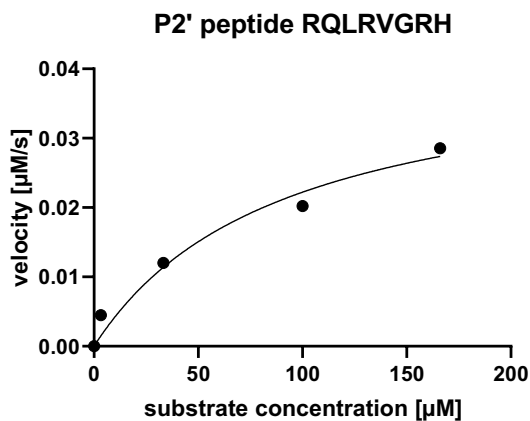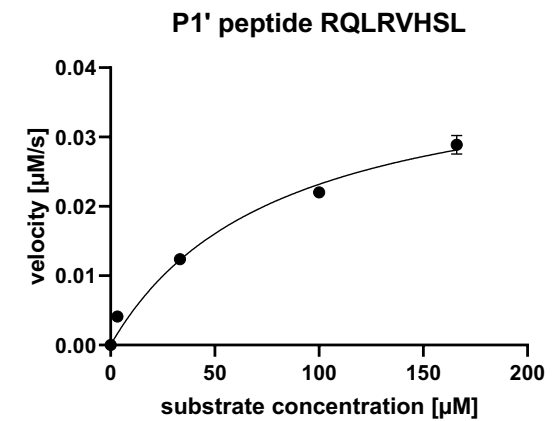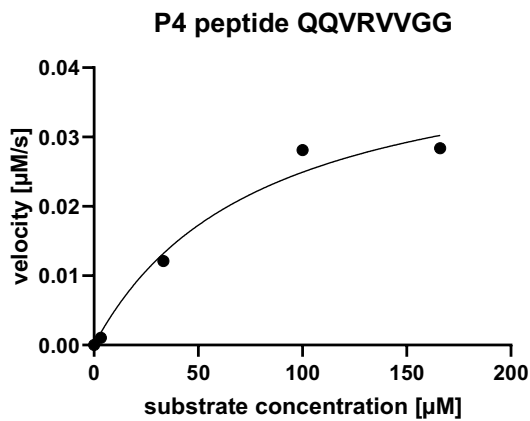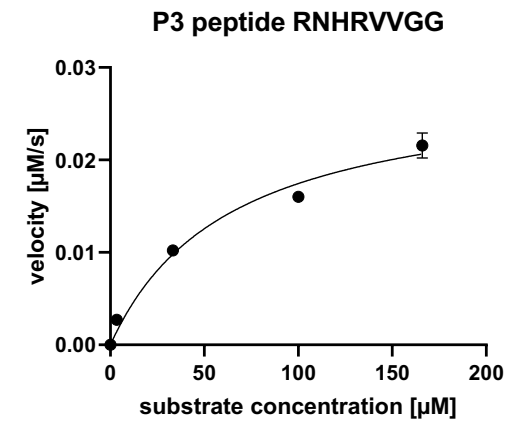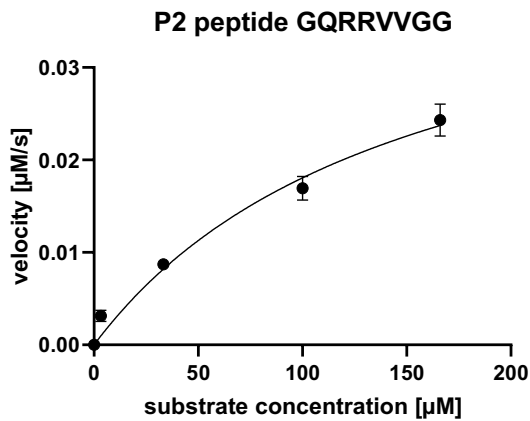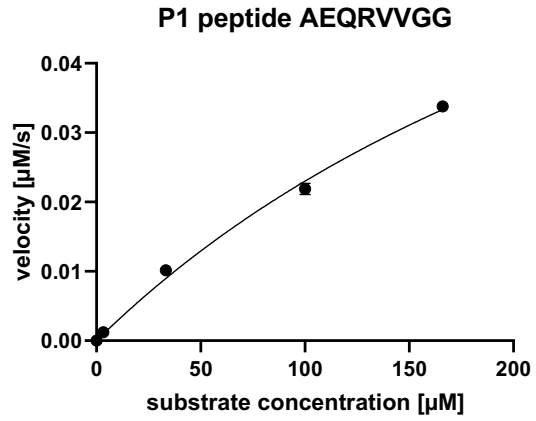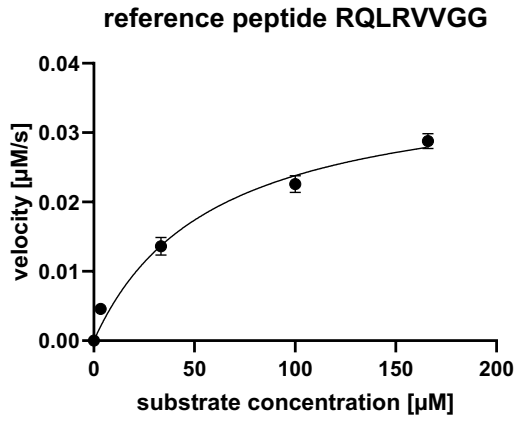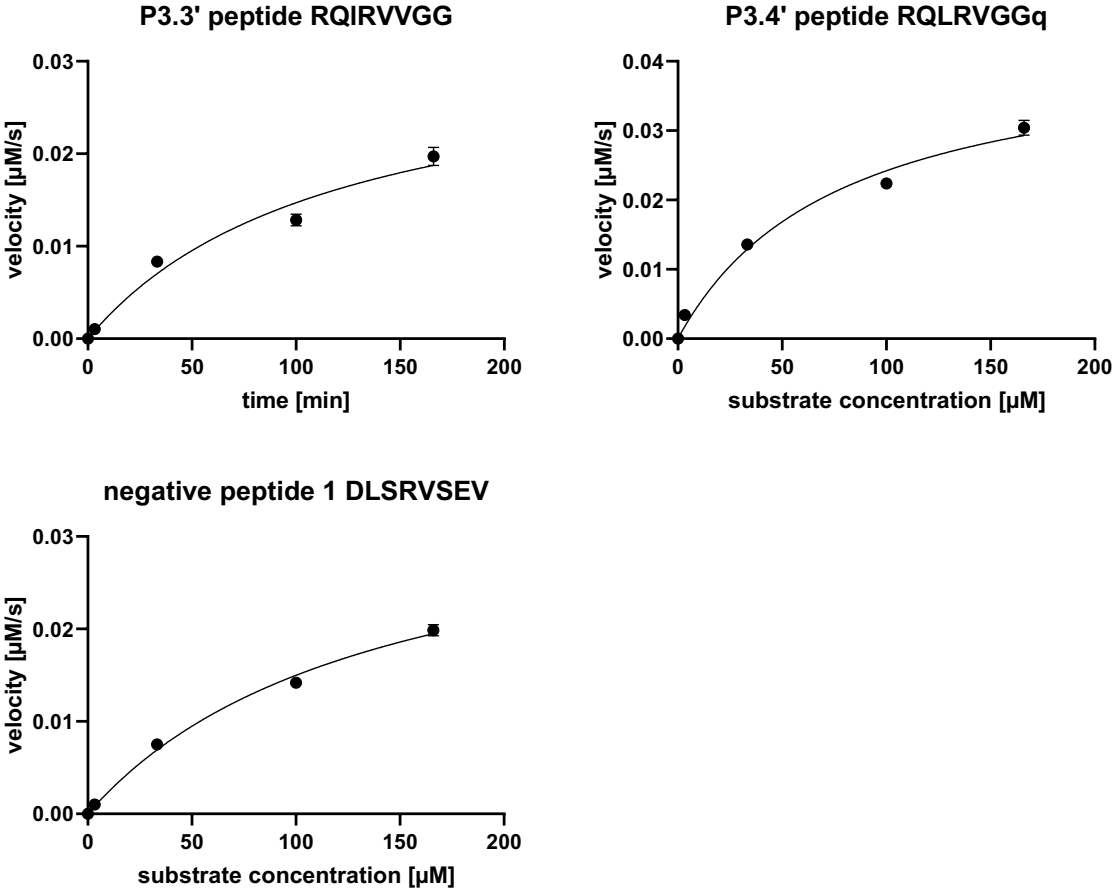
reference peptide RQLRVVGG

P1 peptide AEQRVVGG

P2 peptide GQRRVVGG

P3 peptide RNHRVVGG

P4 peptide QQVRVVGG

P1' peptide RQLRVHSL

P2' peptide RQLRVGRH

P3' peptide RQLRVGGQ

Figure A2. Michaelis Menten analysis of IQF substrates in 12.5% blood serum.

reference peptide RQLRVGGQ

P1 peptide AEQRVVGG

P2 peptide GQRRVVGG

P3 peptide RNHRVVGG

P4 peptide QQVRVVGG

P1' peptide RQLRVHSL

P2' peptide RQLRVGRH

P3' peptide RQLRVGGQ
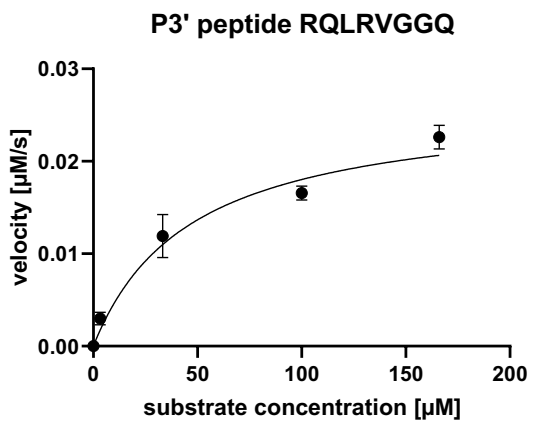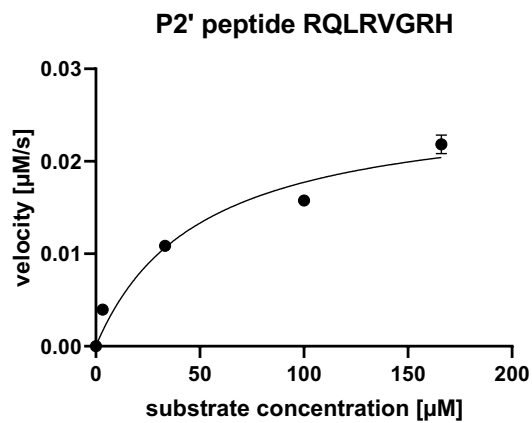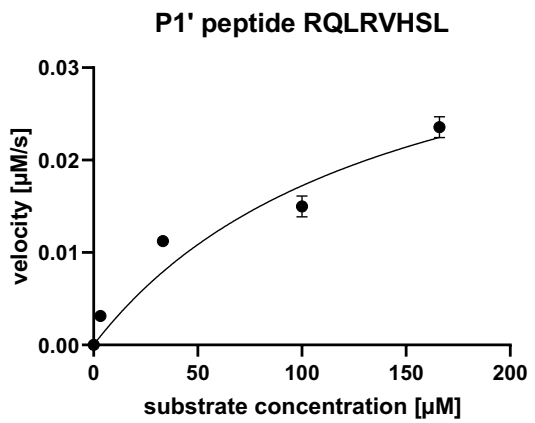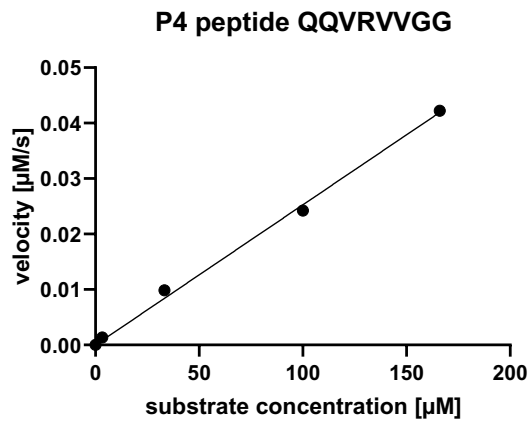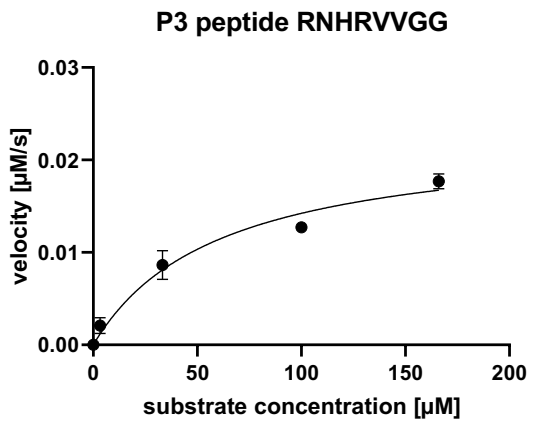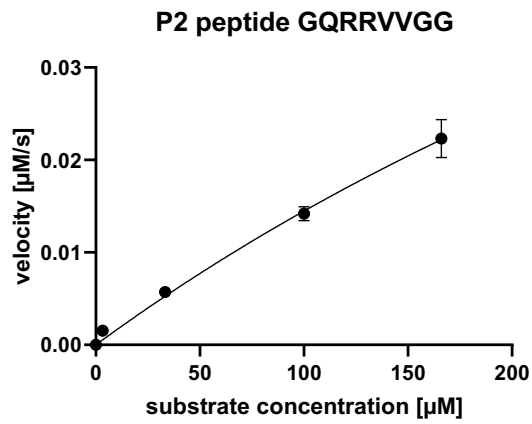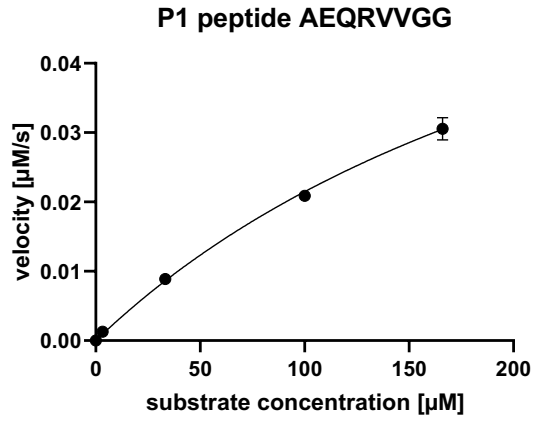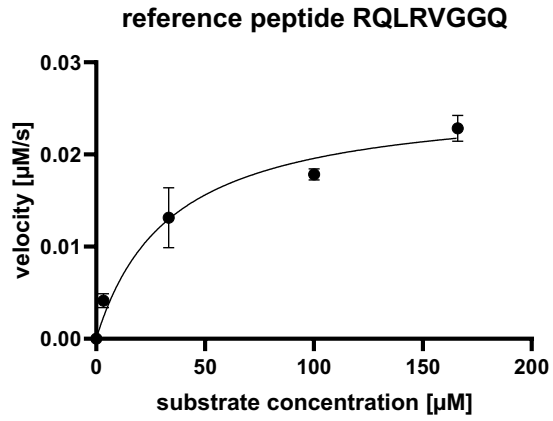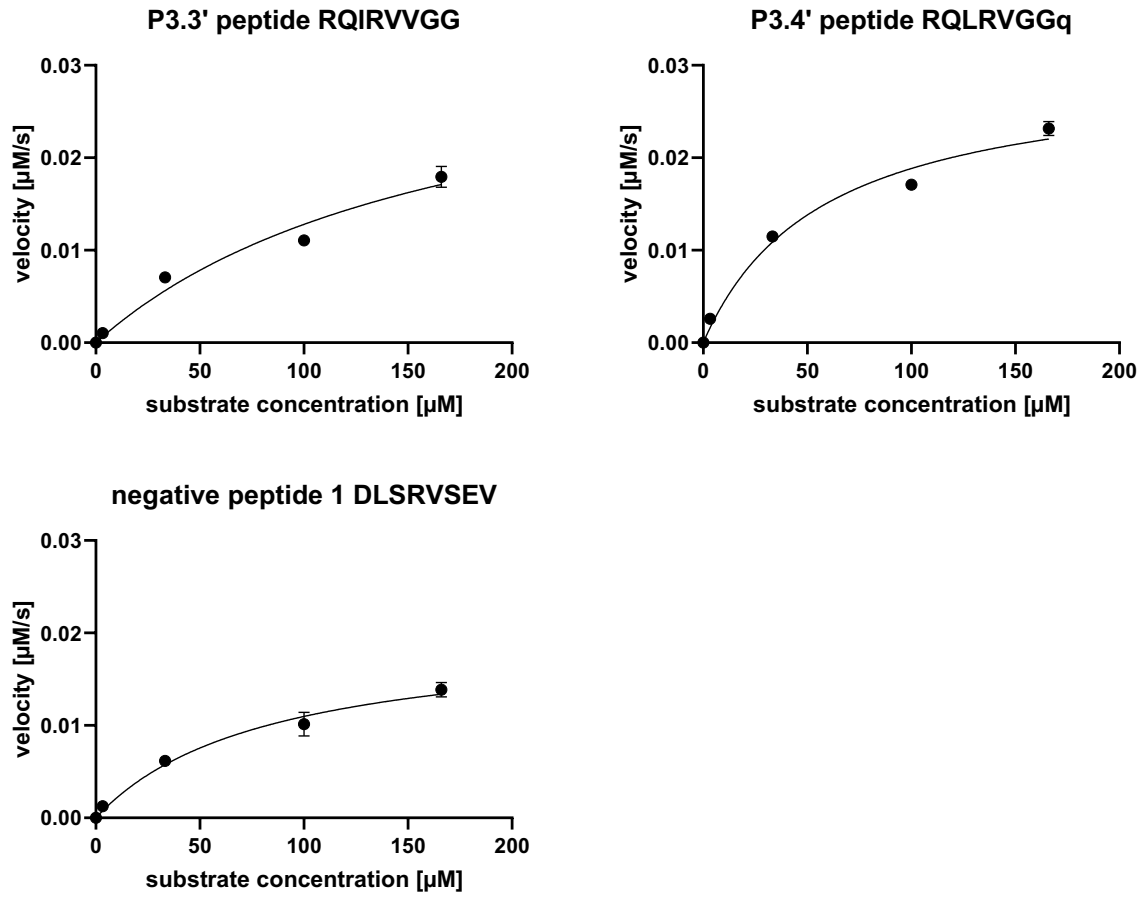
Figure A3. Michaelis Menten analysis of IQF substrates in 12.5% blood plasma.

Analytical data of precursors and serine traps

Boc-Arg(Mtr) Weinreb amide (S1)

((S)-tert-Butyl-(1-(methoxy(methyl)amino)-5-(3-((4-methoxy-2,3,6-trimethylphenyl)sulfonyl)guanidino)-1-oxopentan-2-yl)carbamate)

Structure:

Exact Mass: 529,26
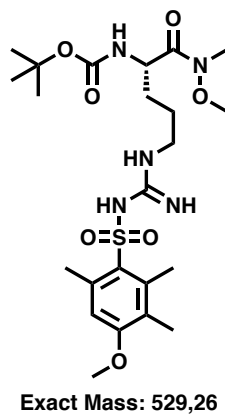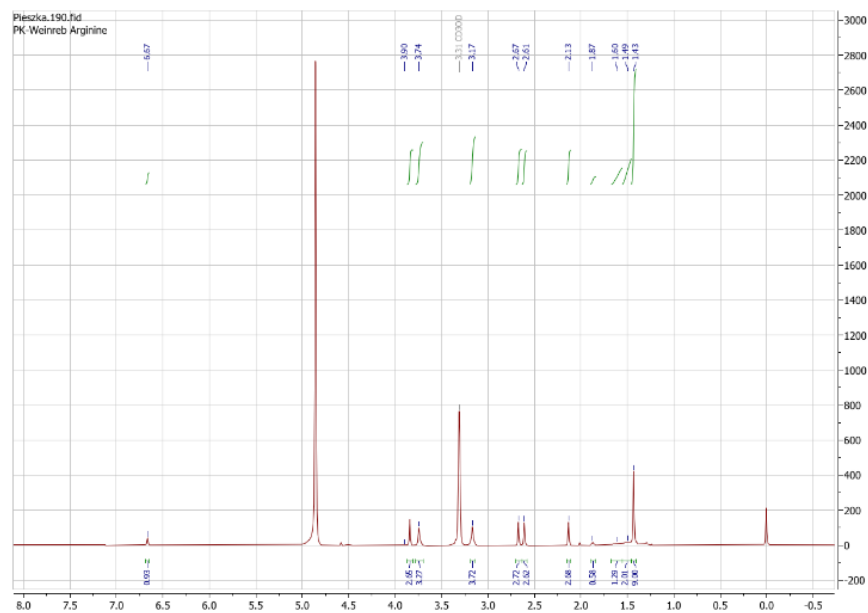
NMR (300 MHz, CD3OD):

^{1}H NMR (300 MHz, CD3OD): δ=1.43 (s, 9 H) 1.49–1.60 (m, 4 H) 1.87 (s, 1 H) 2.13 (s, 3 H) 2.61 (s, 3 H) 2.67 (s, 3 H) 3.17 (m, 3 H) 3.74 (s, 3 H) 3.90 (s, 3 H) 6.67 ppm (s, 1 H).

MS-ESI:

HPLC:

ESI (MS) *m/z*: calcd for $C_{23}H_{39}N_5O_7S_2$ [M+H]$^+$ 530.3, [2M+H]$^+$ 1059.6, found [M+H]$^+$ 530.2, [2M+H]$^+$ 1059.4

Retention time 24 min. Purity (HPLC, 220 nm) > 80%.

Boc-Arg(Mtr) ketobenzothiazole (S2)

tert-Butyl-(1-(benzo[*d*]thiazol-2-yl)-5-(3-((4-methoxy-2,3,6-trimethylphenyl)sulfonyl)guanidino)-1-oxopentan-2-yl)carbamate

Structure:

NMR (300 MHz, CD3OD):

Exact Mass: 603,22
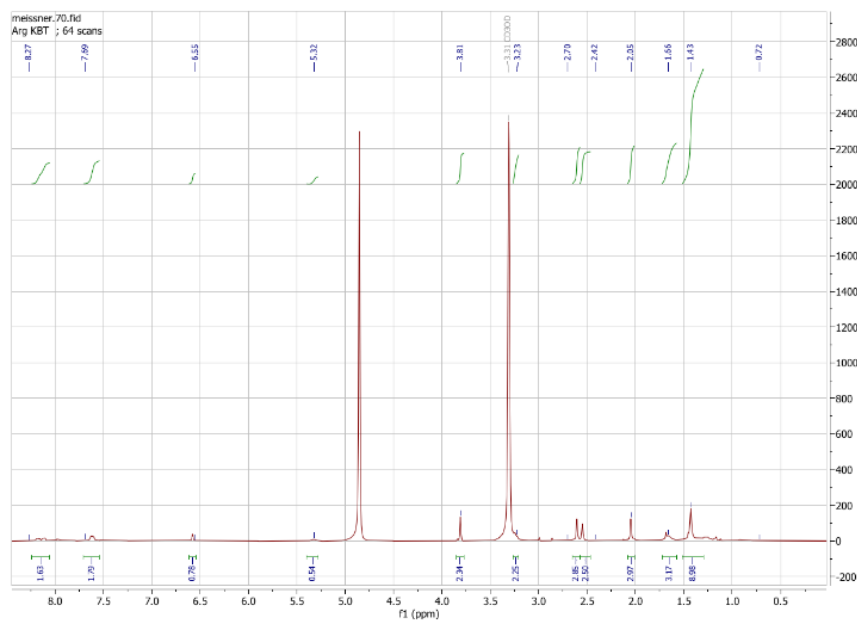
^{1}H NMR (300 MHz, CD3OD): δ=1.43 (s, 9 H) 1.66 (m, 4 H) 2.05 (s, H) 2.55 (s, 3 H) 2.70 (s, 3 H) 3.23 (m, 2 H) 3.81 (s, 3 H) 5.32 (m, 1 6.55 (s, 1 H) 7.69 (m, 2 H) 8.27 ppm (m, 2 H).

MS-ESI:
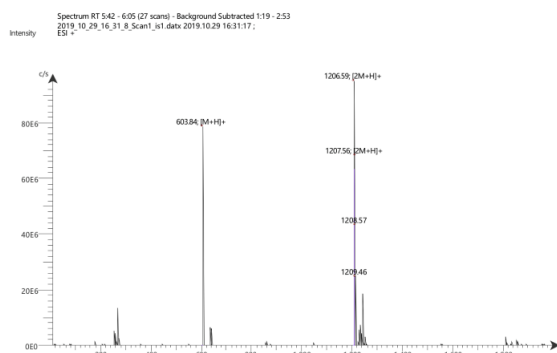
HPLC:
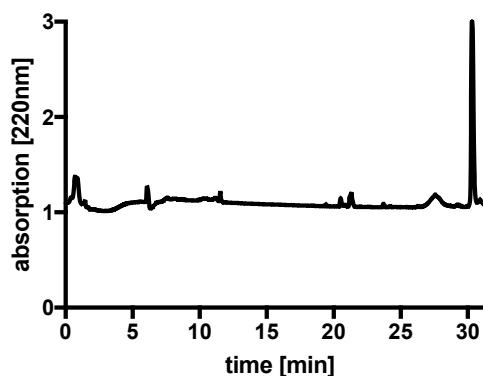
ESI (MS) *m/z*: calcd for $C_{28}H_{37}N_5O_6S_2$ [M+H]$^+$ 604.2, [2M+H]$^+$ 1207.4, found [M+H]$^+$ 603.9, [2M+H]$^+$ 1206.6.

Retention time 30-31 min. Purity (HPLC, 220 nm) > 90%.

H₂N-Arg(Mtr) ketobenzothiazole (S3)

(*N*-(*N*-(4-amino-5-(benzo[*d*]thiazol-2-yl)-5-oxopentyl)carbamimidoyl)-4-methoxy-2,3,6-trimethylbenzenesulfonamide)

Structure:

NMR (300 MHz, CD3OD):

Exact Mass: 503,17
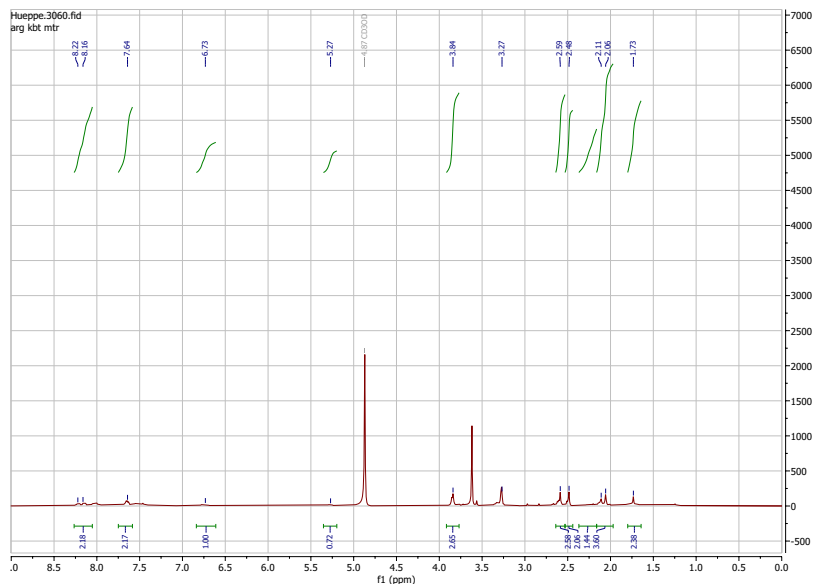
¹H NMR (300 MHz, CD3OD): δ = 1.73 (m, 2 H) 2.06–2.11 (m, 4 H) 2.24–2.36 (m, 1 H) 2.48 (s, 2 H) 2.59 (s, 3 H) 3.84 (s, 3 H) 5.27 (m, 1 H) 6.73 (s, 1 H) 7.64 (m, 2 H) 8.16–8.22 ppm (m, 2 H).

MS-ESI:

HPLC:

Retention time 19-20 min. Purity (HPLC, 220 nm) > 80%
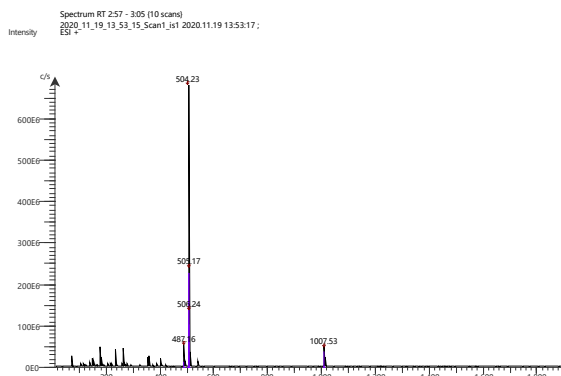
ESI (MS) *m/z*: calcd [M+H]⁺ 504.2, [2M+H]⁺ 1007.4 found [M+H]⁺ 504.2, [2M+H]⁺ 1007.5.

Boc-Arg(Mtr) ketothiazole (S4)

t*ert*-Butyl-(5-(3-((4-methoxy-2,3,6-trimethylphenyl)sulfonyl)guanidino)-1-oxo-1-(thiazol-2-yl)pentan-2-yl)carbamate

Structure: NMR (300 MHz, CDCl₃):
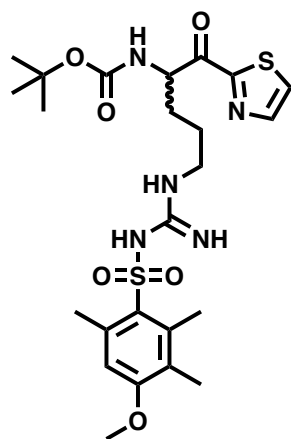
Exact Mass: 553,20
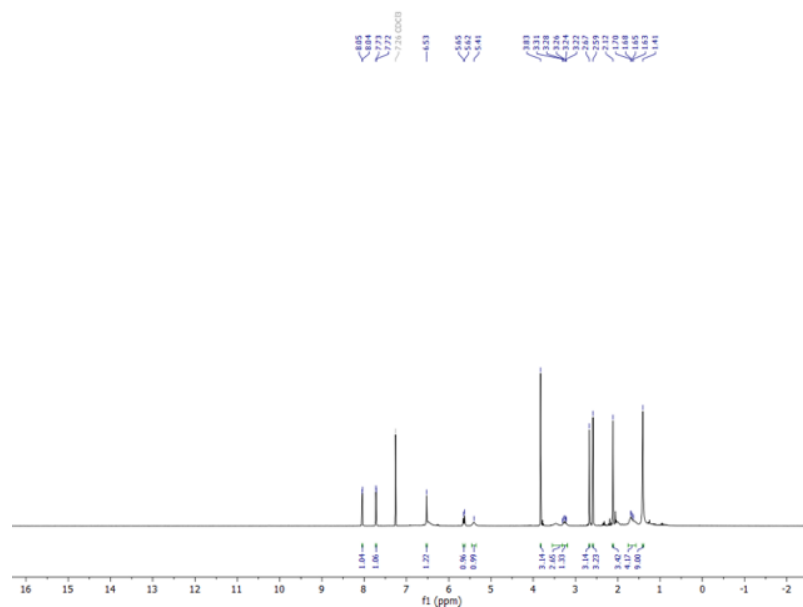
^{1}H NMR (300 MHz, CDCl₃): δ = 8.04 (d, 1 H), 7.72 (d, 1 H), 6.53 (s, 1 H), 5.64 (d, 1 H), 5.41 (s, 1 H), 3.83 (s, 3 H), 3.26 (m, 2 H), 2.67 (s, 3 H), 2.59 (s, 3 H), 2.12 (s, 3 H), 1.76 – 1.57 (m, 4 H), 1.41 (s, 9 H). ppm.

LC-MS data:
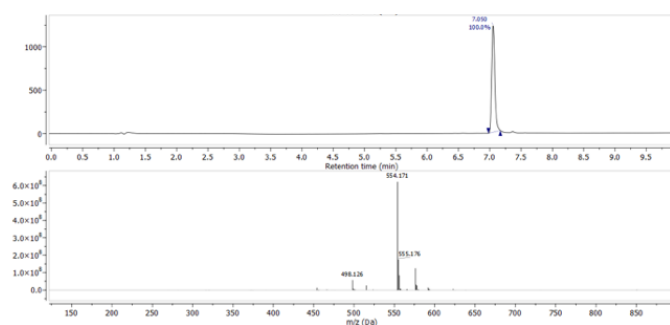
MS (ESI): *m/z*: calcd for C$_{24}$H$_{35}$N$_5$O$_6$S$_2$ [M+H]$^+$ 554.2, found [M+H]$^+$ 554.2.

Retention time 7 min. Purity (254 nm) > 95%.

H₂N-Arg(Mtr) ketothiazole (S5)

N-(*N*-(4-amino-5-oxo-5-(thiazol-2-yl)pentyl)carbamimidoyl)-4-methoxy-2,3,6-trimethylbenzenesulfonamide

Structure:
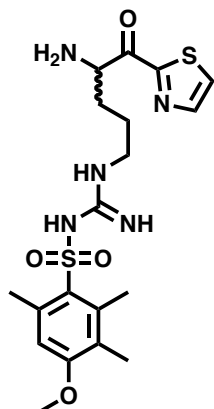
Exact Mass: 453,15

LC-MS data:
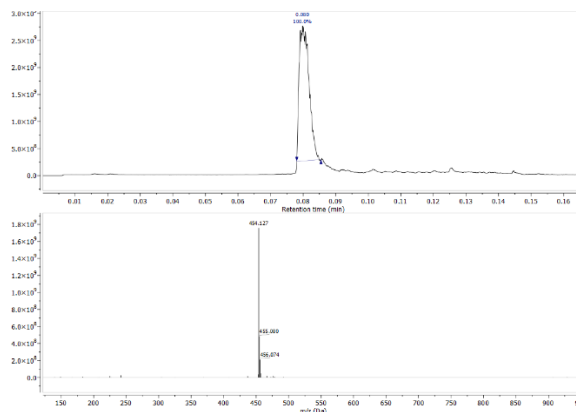
MS (ESI): *m/z*: calcd for $C_{20}H_{27}N_5O_4S_2$ [M+H]$^+$ 453.2, found [M+H]$^+$ 454.1.

Retention time 0.08 min. Purity (254 nm) > 95%.

Boc-Arg(Mtr) alcohol (S6)

(*S*)-tert-Butyl-(1-hydroxy-5-(3-((4-methoxy-2,3,6-trimethylphenyl)sulfonyl)guanidino)pentan-2-yl)carbamate

Structure:
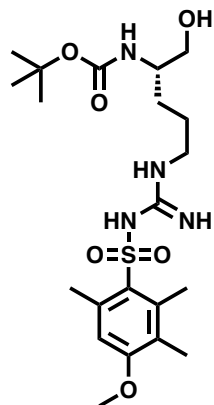
NMR (300 MHz, CDCl₃):

Exact Mass: 472,24
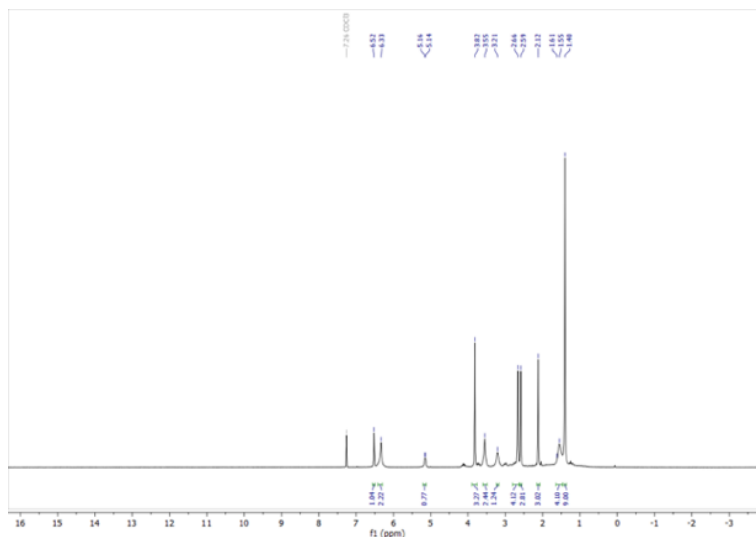
¹H NMR (300 MHz, CDCl₃) δ = 6.52 (s, 1 H), 6.33 (s, 2 H) 5.15 (d, 1 H), 3.82 (s, 3 H), 3.55 (s, 2 H), 3.21 (s, 1 H), 2.69–2.66 (m, 5 H), 2.59 (s, 3 H), 2.12 (s, 3 H), 1.55 (s, 4 H), 1.40 (s, 9 H) ppm.

LC-MS data:
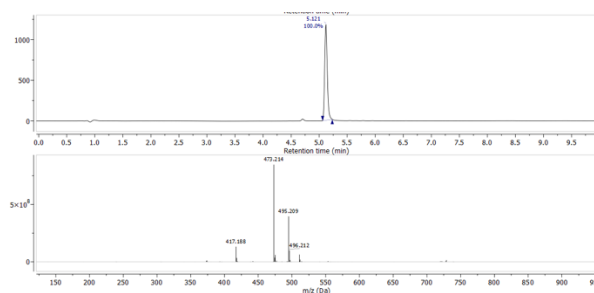
LC-MS *m/z*: calcd for C$_{21}$H$_{36}$N$_4$O$_6$S [M+H]$^+$ 473.2, found [M+H]$^+$ 473.2. Retention time 5 min. Purity (254 nm) > 98%.

H₂N-Arg(Mtr)-OH (S7)

((*S*)-N-(N-(4-amino-5-hydroxypentyl)carbamimidoyl)-4-methoxy-2,3,6-trimethylbenzenesulfonamide)
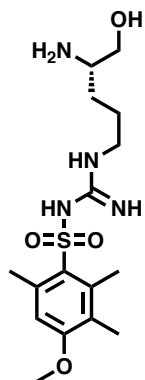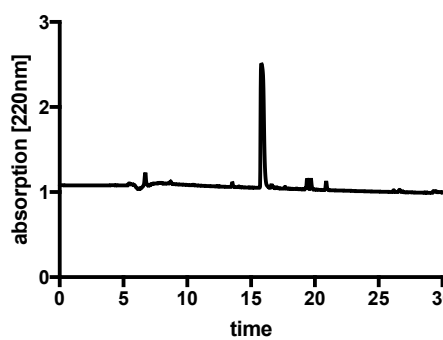
Structure   HPLC: MS-ESI:
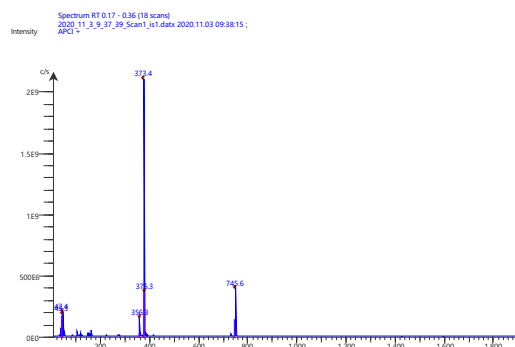
Exact Mass: 372,18

Retention time 16 min. Purity (HPLC, 220 nm) > 90%

ESI (MS) *m/z*: calcd [M+H]⁺ 373.3, [2M+H]⁺ 745.4 found [M+H]⁺ 373.4, [2M+H]⁺ 745.6.

Figure A4. Analytical data (structure, HPLC, MS, NMR data) of precursors and serine traps.

Analytical data of peptidomimetic hepsin inhibitors

Cpd. 1

Structure:

Exact Mass: 721,28
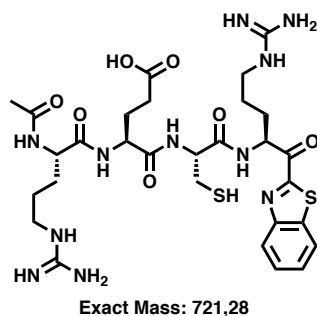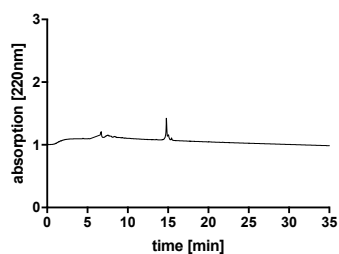
HPLC:

Retention time 15.0 min. Purity (HPLC, 220 nm) > 90%.

MS-ESI:
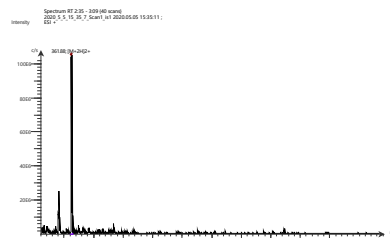
ESI (MS) *m/z*: calcd [M+H]$^+$ 722.3, [M+2H]$^{2+}$ 361.7 found [M+H]$^+$ 361.9.

Cpd. 2

Structure:

Exact Mass: 661,26
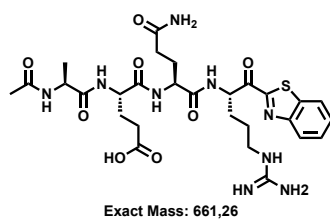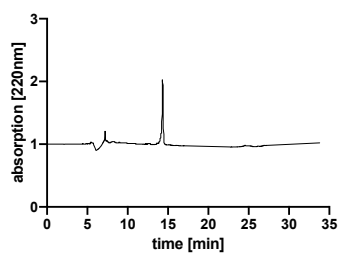
HPLC:

Retention time 14.5 min. Purity (HPLC, 220 nm) > 90%.
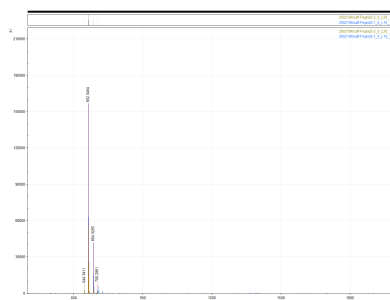
MALDI-TOF:

MALDI-TOF *m/z*: calcd [M+H]$^+$ 662.3, [M+Na]$^+$ 684.2 found [M+H]$^+$ 662.3, [M+Na]$^+$ 684.3

Cpd. 3

Structure:

Exact Mass: 674,31
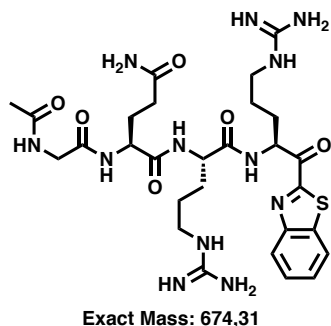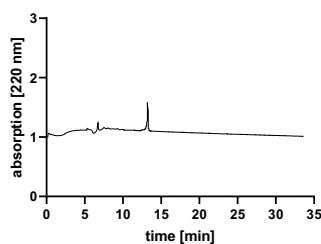
HPLC:

Retention time 13.5 min. Purity (HPLC, 220 nm) > 90%.

MS-ESI:
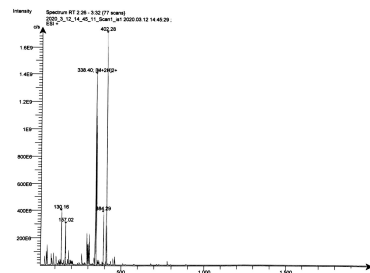
ESI (MS) *m/z*: calcd [M+H]⁺ 675.3, [M+2H]²⁺ 338.2 found [M+2H]²⁺ 338.4.

Cpd. 4

Structure:

Exact Mass: 688,31
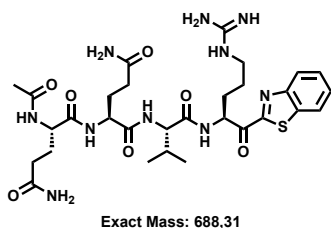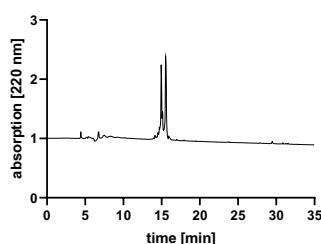
HPLC:

Retention time 14.5 - 15.5 min. Purity (HPLC, 220 nm) > 90%.
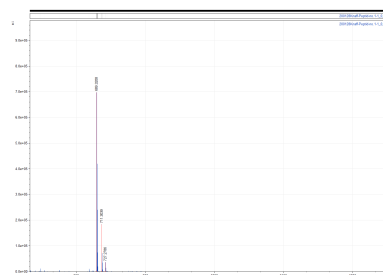
MALDI-TOF:

MALDI-TOF *m/z*: calcd [M+H]⁺ 689.3, [M+Na]⁺ 711.2 found [M+H]⁺ 689.3, [M+Na]⁺ 711.3

Cpd. 5

Structure:

Exact Mass: 754,34
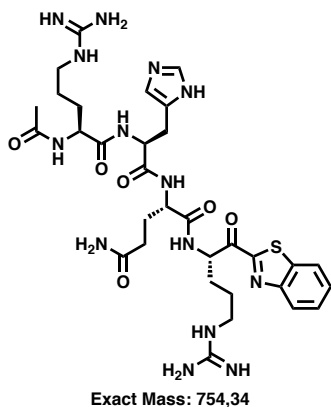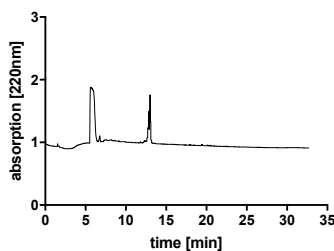
HPLC:

Retention time 13 min. Purity (HPLC, 220 nm) > 90%.
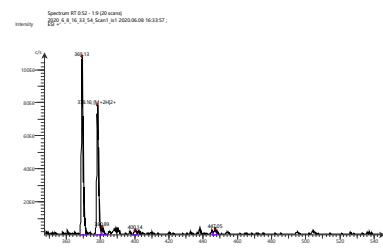
MS-ESI:

ESI (MS) *m/z*: calcd [M+H]⁺ 378.2 found [M+2H]²⁺ 378.2.

Cpd. 6

Structure:

Exact Mass: 687,28
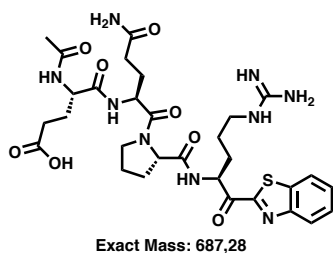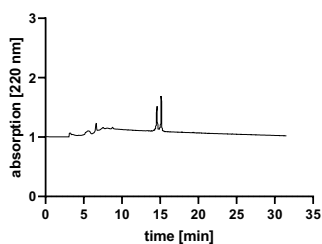
HPLC:

Retention time 14.5 - 15.5 min. Purity (HPLC, 220 nm) > 90%.
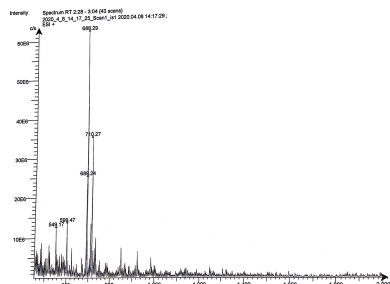
MS-ESI:

ESI (MS) *m/z*: calcd $[M+H]^+$ 688.3 found $[M+2H]^{2+}$ 688.3.

Cpd. 7

Structure:

Exact Mass: 704,27
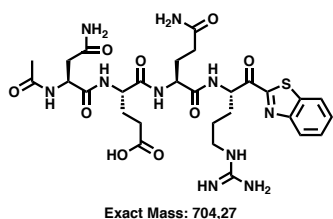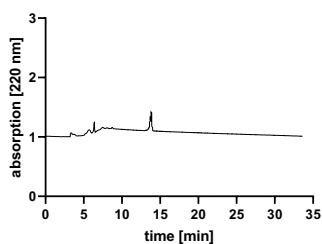
HPLC:

Retention time 14.0 min. Purity (HPLC, 220 nm) > 90%.
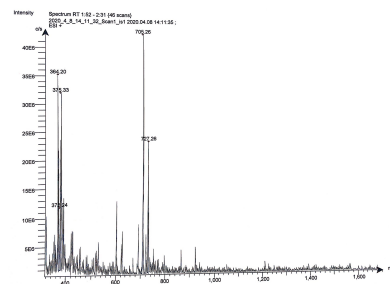
MS-ESI:

ESI (MS) *m/z*: calcd $[M+H]^+$ 705.3, $[M+Na]^+$ 727.2 found $[M+H]^+$ 705.3, $[M+Na]^+$ 727.3

Cpd. 8

Structure:

Exact Mass: 624,29
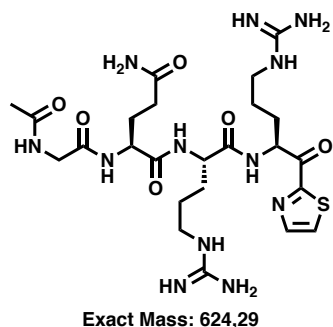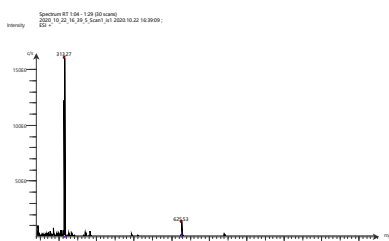
MS-ESI:

ESI (MS) *m/z*: calcd $[M+H]^+$ 625.3, $[M+2H]^{2+}$ 313.2 found $[M+H]^+$ 625.5, $[M+Na]^+$ 313.3

Cpd. 9

Structure:

Exact Mass: 679,32
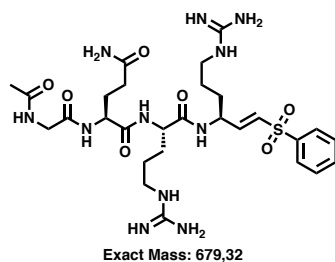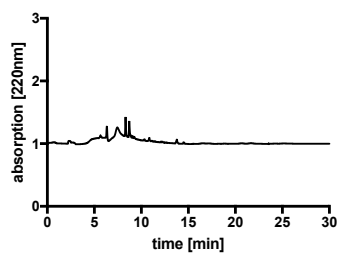
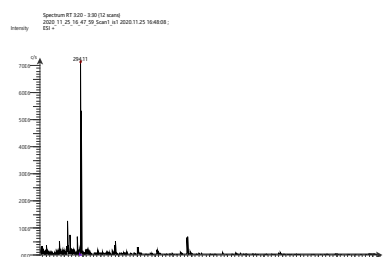**HPLC:**

Retention time 8-9 min. Purity (HPLC, 220 nm) > 90%.
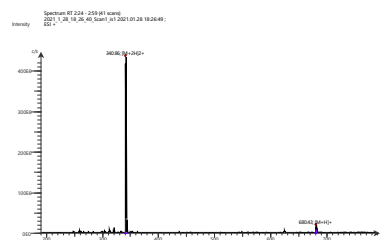
MS-ESI:

ESI (MS) *m/z*: calcd [M+H]$^+$ 679.3, [M+2H]$^{2+}$ 340.7 found [M+2H]$^{2+}$ 340.9

Cpd. 10

Structure:

Exact Mass: 557,30
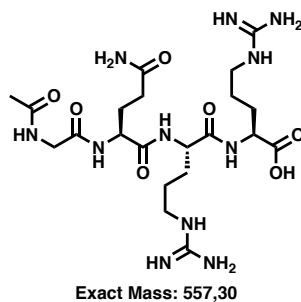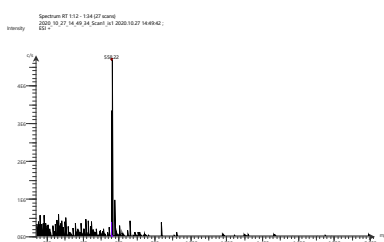
MS-ESI:

ESI (MS) *m/z*: calcd [M+H]$^+$ 558.3 found [M+H]$^+$ 558.2

Cpd. 11

Structure:

Exact Mass: 543,32
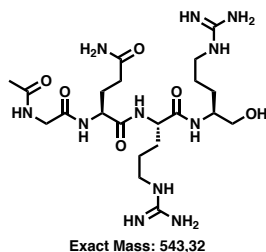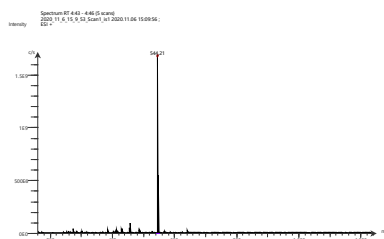
MS-ESI:

ESI (MS) *m/z*: calcd [M+H]$^+$ 544.3 found [M+H]$^+$ 544.2

Cpd. 12

Structure: **MS-ESI:**

Exact Mass: 617,29
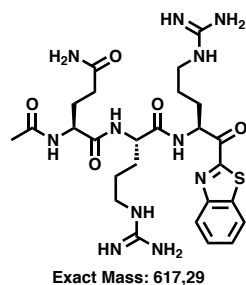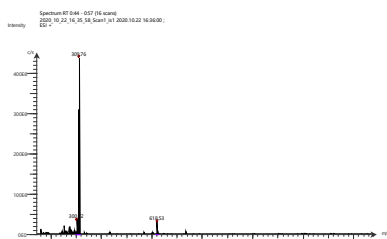
ESI (MS) *m/z*: calcd [M+H]⁺, 618.2 [M+2H]²⁺ 309.6 found [M+H]⁺ 618.5. [M+2H]²⁺309.8

Cpd. 13

Structure: **MS-ESI:**

Exact Mass: 489,23
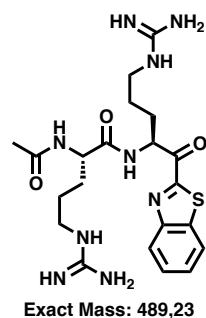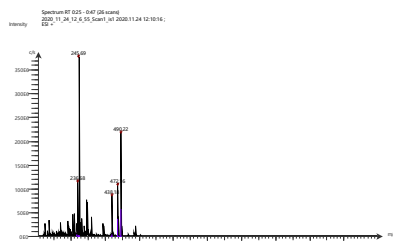
ESI (MS) *m/z*: calcd [M+H]⁺ 490.2 [M+2H]²⁺ 245.6 found [M+H]⁺ 490.2. [M+2H]²⁺245.7

Cpd. 14

Structure: **MS-ESI:**

Exact Mass: 333,13
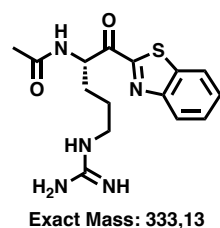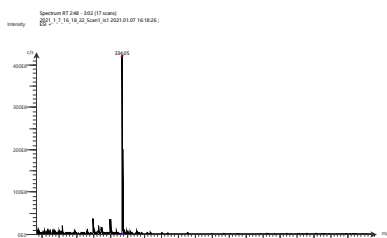
ESI (MS) *m/z*: calcd [M+H]⁺ 333.1 found [M+H]⁺ 334.1

Figure A5. Analytical data (structure, HPLC, MS) of peptidomimetic inhibitors.
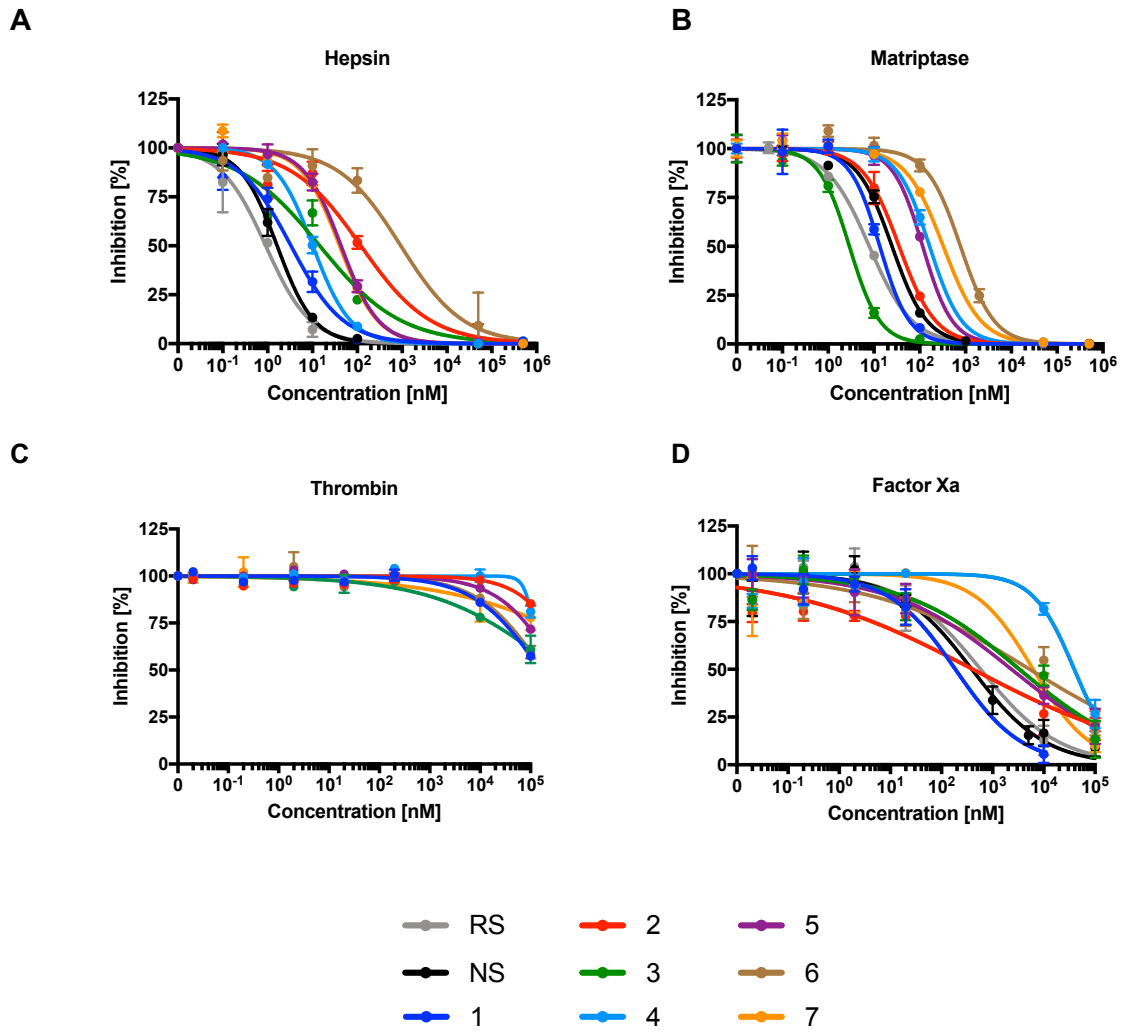
Figure A6. Peptidomimetic inhibitors block activity of purified hepsin (**A**), matriptase (**B**), thrombin (**C**) and factor Xa (**D**). Peptidomimetic inhibitors were added to isolated enzymes and after 30 minutes, the fluorogenic reference substrate Boc-Gln-Ala-Arg-AMC was added to hepsin and matriptase and the chromogenic reference substrates, D-Phe-Homopro-Arg-pNA or Bz-Ile-Glu-Gly-Arg-pNA, were added to thrombin or factor Xa, respectively. The velocity of substrate degradation was assessed by recording the fluorescence intensity at 460 nm or the absorbance at 405 nm within 2 h. Shown are the means ± SD of triplicate measurements.
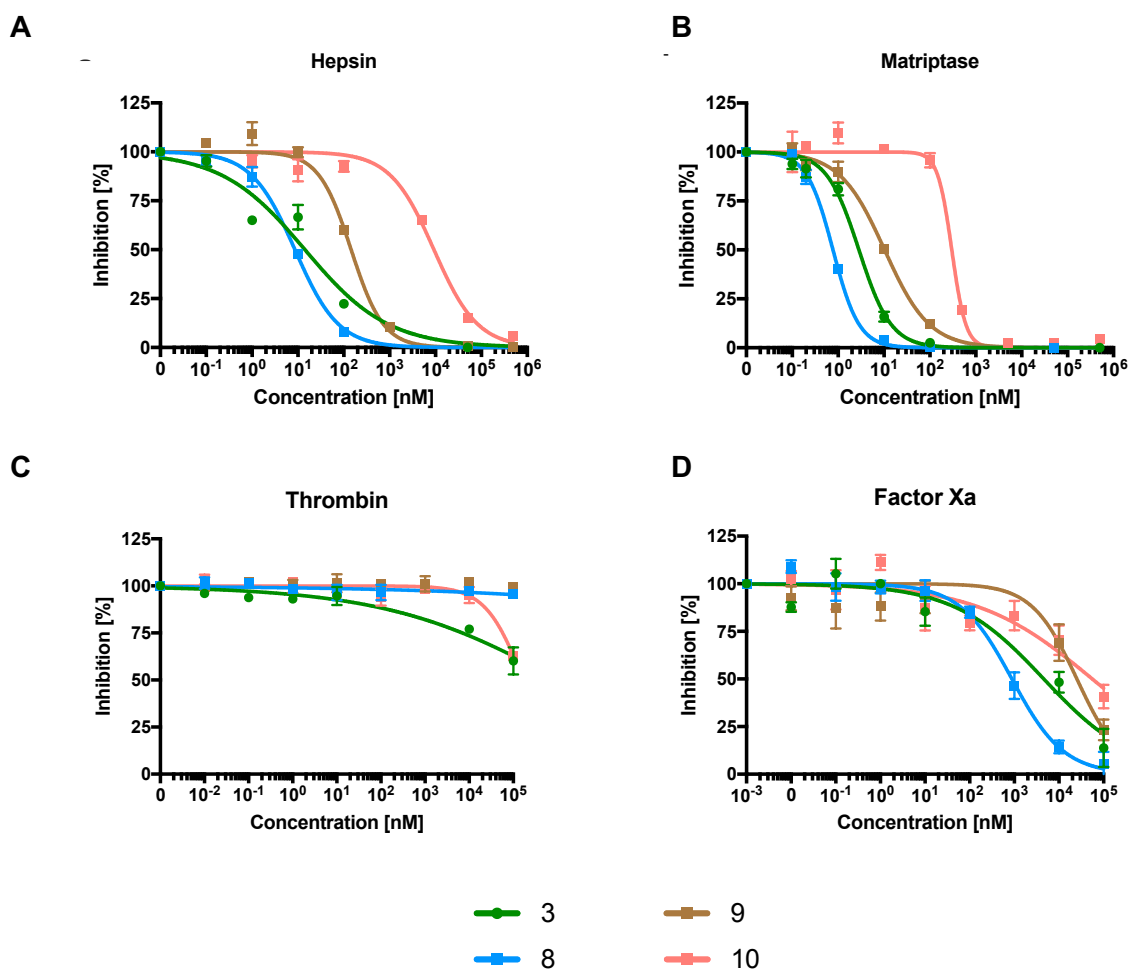
Figure A7. Influence of truncation on inhibitory activity of compound 3 against hepsin (**A**), matriptase (**B**), thrombin (**C**) and factor Xa (**D**). Peptidomimetic inhibitors were added to isolated enzymes and after 30 minutes, the fluorogenic reference substrate Boc-Gln-Ala-Arg-AMC was added to hepsin or matriptase and the chromogenic substrates, D-Phe-Homopro-Arg-pNA or Bz-Ile-Glu-Gly-Arg-pNA, were added to thrombin or factor Xa, respectively. The velocity of substrate degradation was assessed by recording the fluorescence intensity at 460 nm or the absorbance at 405 nm within 2 h. Shown are the means ± SD of triplicate measurements.
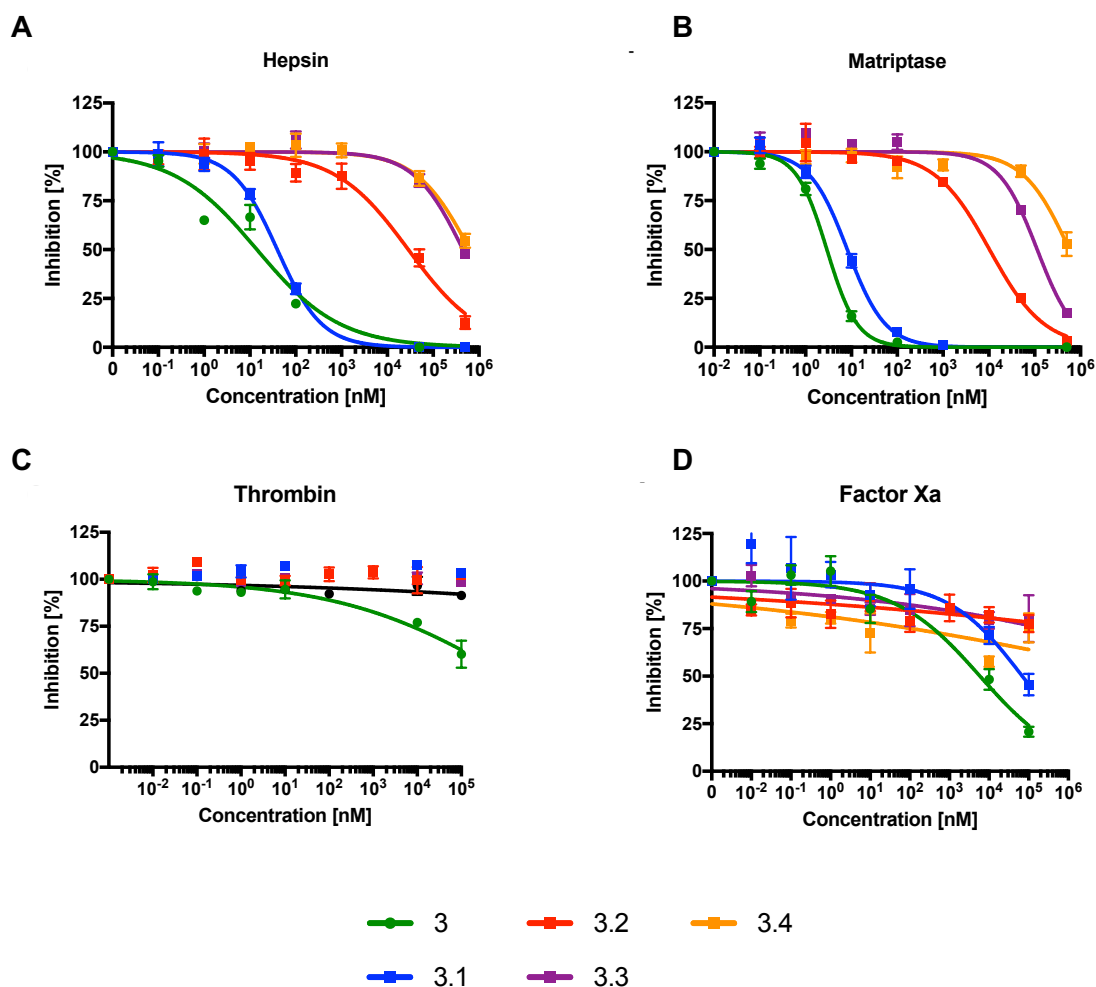
Figure A8. Influence of serine trap on biological activity of compound 3 against hepsin (**A**), matriptase (**B**), thrombin (**C**) and factor Xa (**D**). Peptidomimetic inhibitors were added to isolated enzymes and after 30 minutes, the fluorogenic reference substrate Boc-Gln-Ala-Arg-AMC was added to hepsin or matriptase and the chromogenic substrates, D-Phe-Homopro-Arg-pNA or Bz-Ile-Glu-Gly-Arg-pNA, were added to thrombin or factor Xa, respectively. The velocity of substrate degradation was assessed by recording the fluorescence intensity at 460 nm or the absorbance at 405 nm within 2 h. Shown are the means ± SD of triplicate measurements.

Analytical data of peptidomimetic TMPRSS2 inhibitors

Cpd. 1

Structure	HPLC:	MS-ESI:

Exact Mass: 586,28
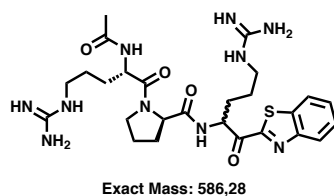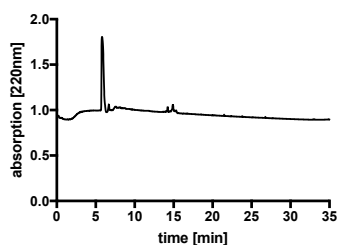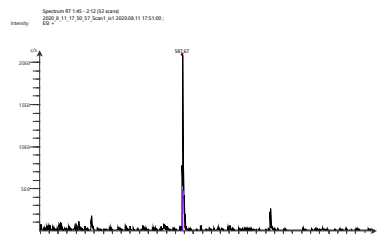
Retention time 15.5 min. Purity (HPLC, 220 nm) > 90%.

ESI (MS) *m/z*: calcd [M+H]⁺ 587.3, [2M+H]⁺ 1174.6 found [M+H]⁺ 587.7.

Cpd. 2

Structure	HPLC:	MS-ESI:

Exact Mass: 586,28
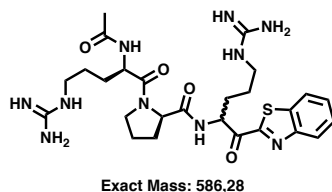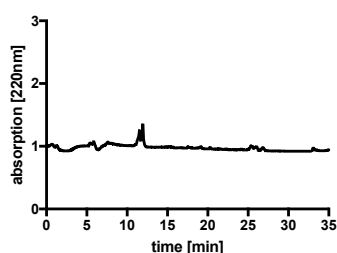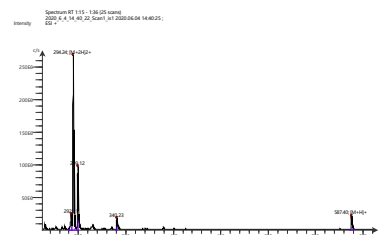
Retention time 12.5 min. Purity (HPLC, 220 nm) > 90%.

ESI (MS) *m/z*: calcd [M+H]⁺ 587.3, [M+2H]²⁺ 294.2 found [M+H]⁺ 587.4, [M+2H]²⁺ 294.2.

Cpd. 2*S*

Structure

HPLC:

MS-ESI:

Exact Mass: 586,28
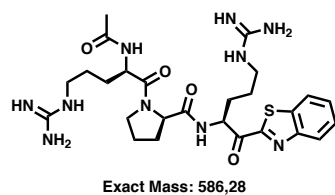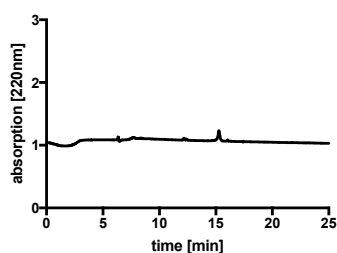
Retention time 15.2 min.
Purity (HPLC, 220 nm) >
90%.

ESI (MS) *m/z*: calcd [M+H]$^+$
587.3, [M+2H]$^{2+}$ 294.2 found
[M+H]$^+$ 587.4, [M+2H]$^{2+}$ 294.1.

Cpd. 2*R*

Structure
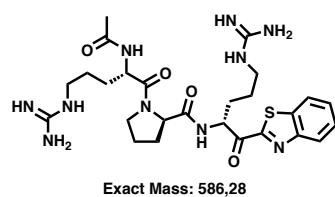
HPLC:
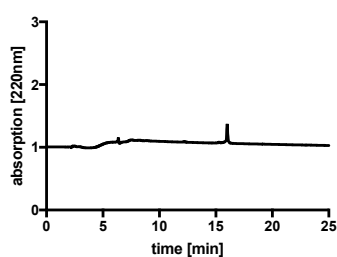
MS-ESI:

Exact Mass: 586,28

Retention time 16.0 min.
Purity (HPLC, 220 nm) >
90%.
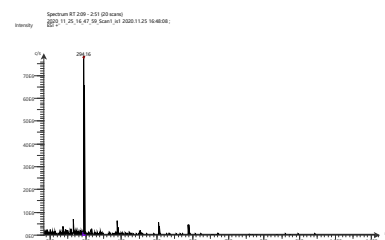
ESI (MS) *m/z*: calcd [M+H]$^+$
587.3, [M+2H]$^{2+}$ 294.2 found
[M+H]$^+$ 587.4, [M+2H]$^{2+}$ 294.1.

Cpd. 3

Structure

HPLC:

MS-ESI:

Exact Mass: 567,24
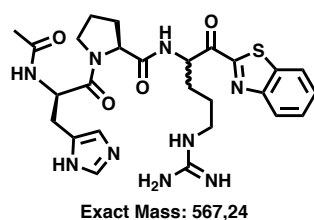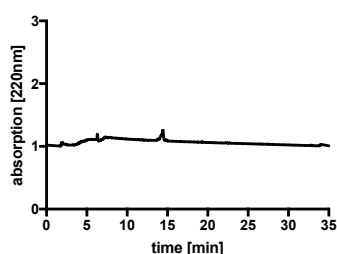
Retention time 14.5 min.
Purity (HPLC, 220 nm) >
90%.
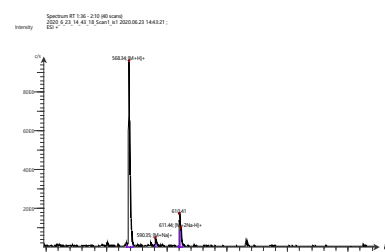
ESI (MS) *m/z*: calcd [M+H]$^+$
568.2, [M+Na]$^+$ 590.2 found
[M+H]$^+$ 568.3, [M+Na]$^+$ 590.3.

Cpd. 4

Structure

Exact Mass: 567,24
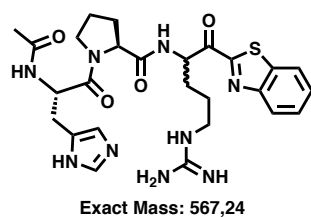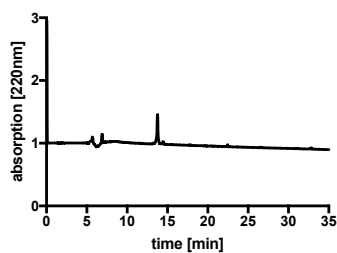
HPLC:

Retention time 14.0 min. Purity (HPLC, 220 nm) > 90%.
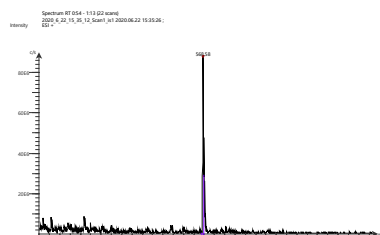
MS-ESI:

ESI (MS) *m/z*: calcd [M+H]+ 568.2 found [M+H]+ 568.6.

Cpd. 5

Structure

Exact Mass: 544,22
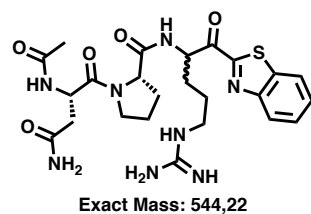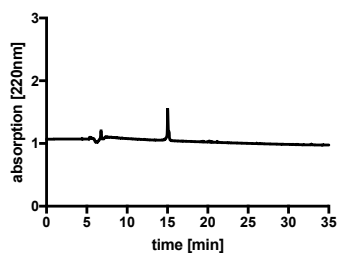
HPLC:

Retention time 14.0 min. Purity (HPLC, 220 nm) > 90%.
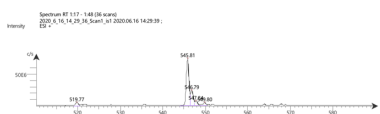
MS-ESI:

ESI (MS) *m/z*: calcd [M+H]+ 545.2 found [M+H]+ 545.8.

Cpd. 6

Structure
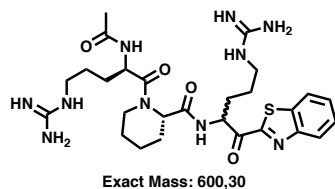
Exact Mass: 600,30
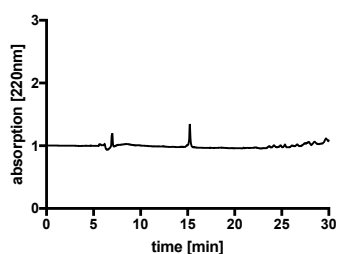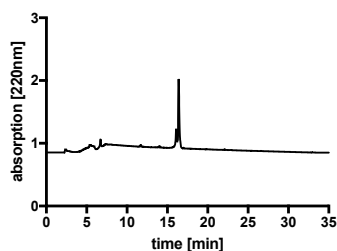
HPLC:

Retention time 15.0 min. Purity (HPLC, 220 nm) > 90%.
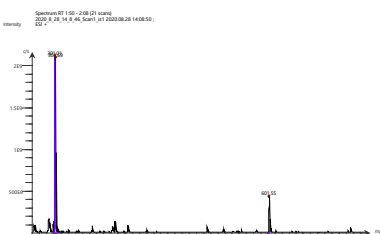
MS-ESI:

ESI (MS) *m/z*: calcd [M+H]+ 601.3, [M+2H]2+ 301.2, found [M+H]+ 601.6 [M+2H]2+ 301.7.

Cpd. 7

Structure   HPLC: MS-ESI:
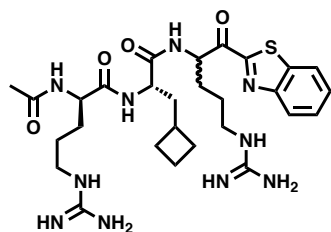
Exact Mass: 614,31
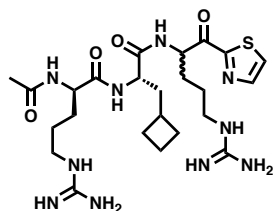
Retention time 16.0 min.
Purity (HPLC, 220 nm) >
90%.
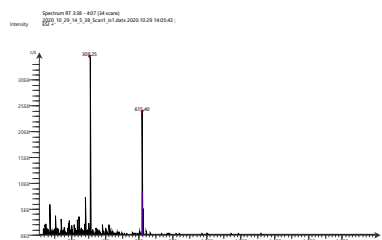
ESI (MS) *m/z*: calcd [M+H]$^+$
615.3, [M+2H]$^{2+}$ 308.2, found
[M+H]$^+$ 615.4 [M+2H]$^{2+}$ 308.3.

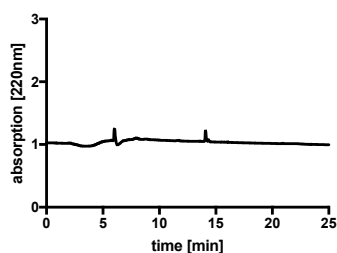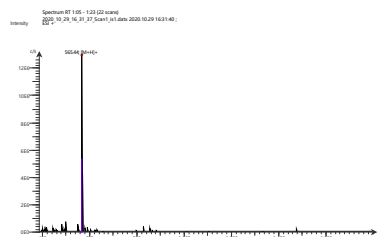## Cpd. 7-2

Structure   HPLC: MS-ESI:

Exact Mass: 564,30
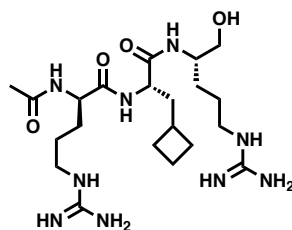
Retention time 14.0 min.
Purity (HPLC, 220 nm) >
90%.

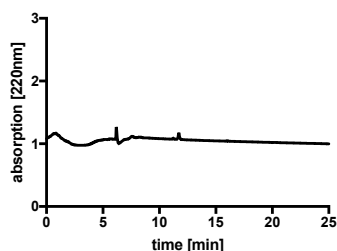ESI (MS) *m/z*: calcd [M+H]$^+$
565.3 found [M+H]$^+$ 565.4.

Cpd. 7-3
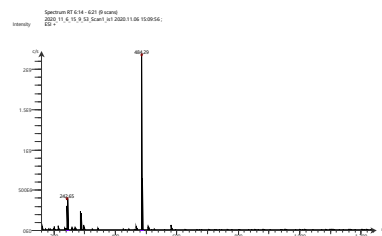
Structure   HPLC: MS-ESI:

Exact Mass: 483,33
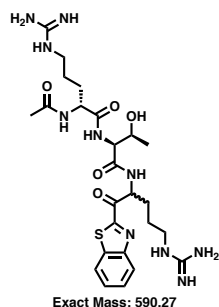
Retention time 12.0 min.
Purity (HPLC, 220 nm) >
90%
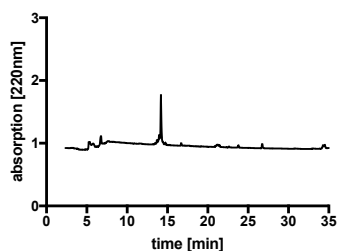
ESI (MS) *m/z*: calcd [M+H]$^+$
484.3, [M+2H]$^{2+}$ 242.7, found
[M+H]$^+$ 484.3 [M+2H]$^{2+}$ 242.7.

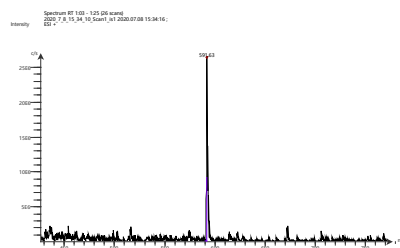## Cpd. 8

| Structure | HPLC: | MS-ESI: |

Exact Mass: 590,27

Retention time 14.0 min. Purity (HPLC, 220 nm) > 90%

ESI (MS) *m/z*: calcd $[M+H]^+$ 591.3 found $[M+H]^+$ 591.6.

Figure A9. Analytical data (structure, HPLC, MS-ESI, NMR data) of peptidomimetic TMPRSS2 inhibitors.
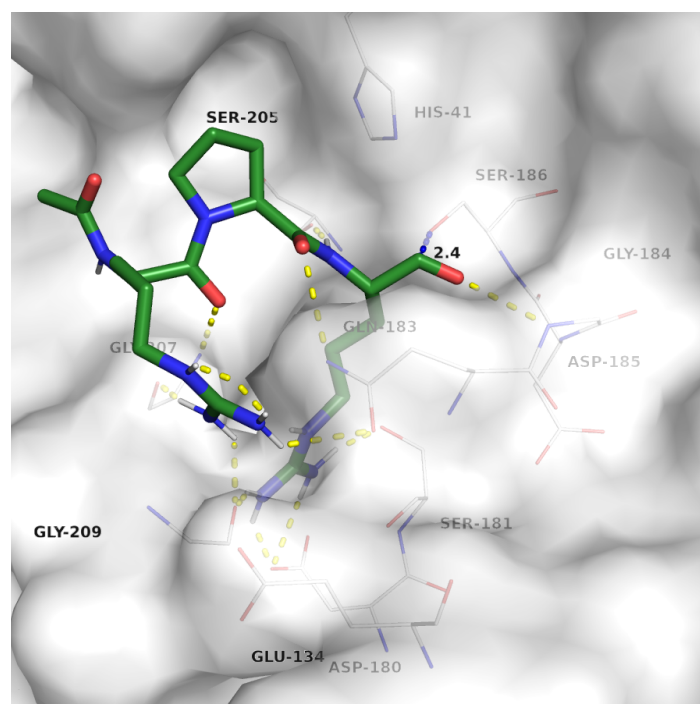
Figure A10. Predicted binding mode of substrate-like ligand ace-D-Arg-Pro-Arg-aldehyde in complex with TMPRSS2 homology model (white carbon atoms and surface). Yellow dashed lines indicate hydrogen bonds, the blue dashed line the distance between catalytic serine nucleophile oxygen and carbon of the P1|P1′ amide bond. For clear view, only residues forming polar contacts with the ligand and His14/57 of the catalytic triad are labeled and depicted as lines. Carbon atoms are shown in green, oxygens in red and nitrogens in blue.

A

TMPRSS2

B

matriptase

C

thrombin

D

factor Xa

1	3	5	7	CM	
2	4	6	8	FOY-251	
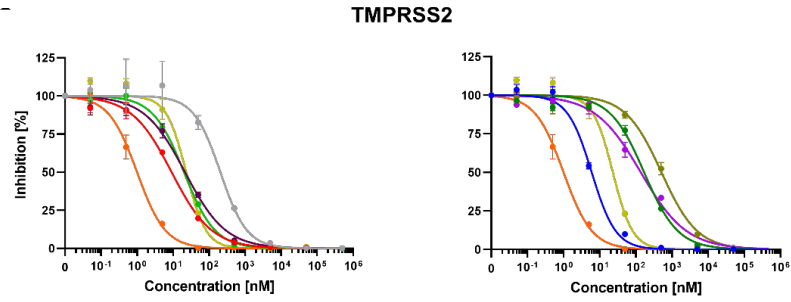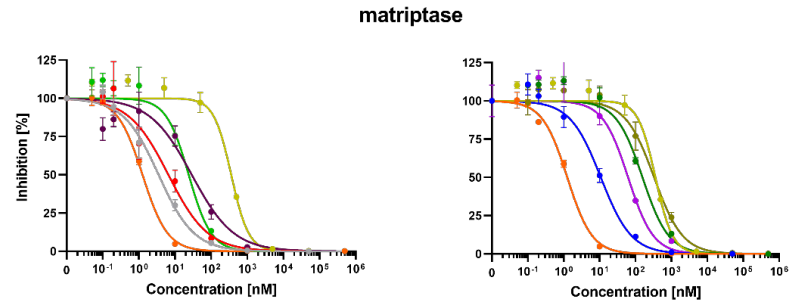
Figure A11. Peptidomimetic inhibitors block activity of purified TMPRSS2 (**A**), matriptase (**B**), thrombin (**C**) and factor Xa (**D**). Peptidomimetic inhibitors, camostat mesylate (CM) and FOY-251 were added to isolated enzymes. After 30 minutes, the fluorogenic reference substrate Boc-Gln-Ala-Arg-AMC was added to matriptase and TMPRSS2 and the chromogenic substrates, D-Phe-Homopro-Arg-pNA or Bz-Ile-Glu-Gly-Arg-pNA, were added to thrombin or fact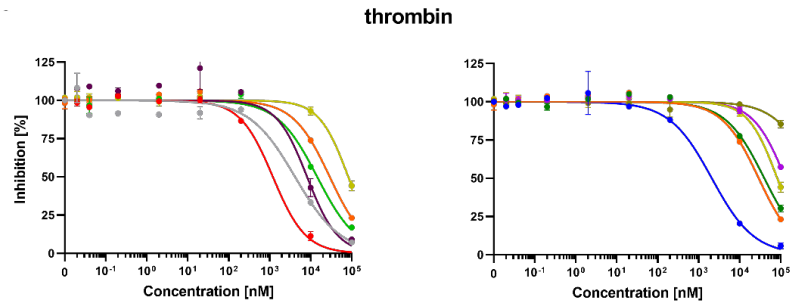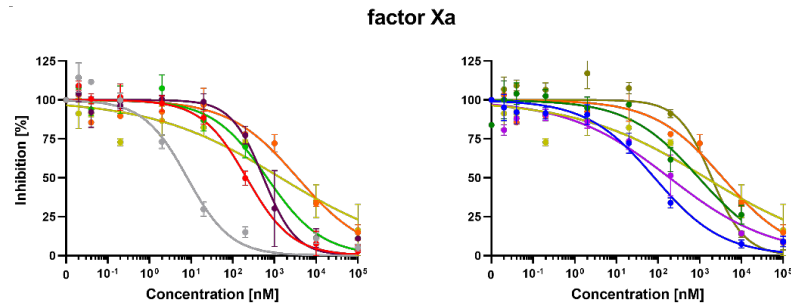or Xa, respectively. The velocity of substrate degradation was assessed by recording the fluorescence intensity at 460 nm or the absorbance at 405 nm within 2 h. Shown are the means ± SD of triplicate measurements.

Figure A12. Influence of serine trap on biological activity of compound 7 against **A** matriptase, **B** thrombin and **C** factor Xa. Peptidomimetic inhibitors with a ketobenzothiazole (7), ketothiazole (7-2) or alcohol (7-3) serine trap residue were added to isolated enzymes. After 30 minutes, the fluorogenic reference substrate Boc-Gln-Ala-Arg-AMC was added to matriptase and TMPRSS2 and the chromogenic substrates, D-Phe-Homopro-Arg-pNA or Bz-Ile-Glu-Gly-Arg-pNA, were added to thrombin or factor Xa, respectively. The velocity of substrate degradation was assessed by recording the fluorescence intensity at 460 nm or the absorbance at 405 nm within 2 h. Shown are the means ± SD of triplicate measurements.
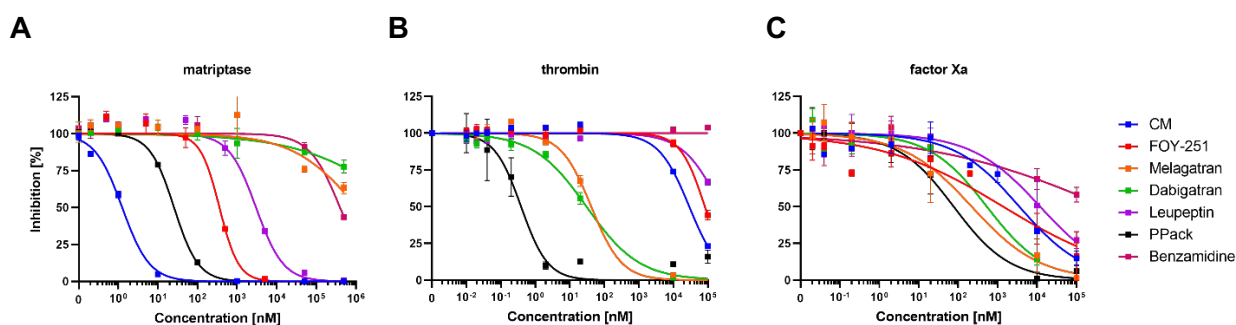
Figure A13. Biological activity of virtual screening hits against **A** matriptase, **B** thrombin and **C** factor Xa. After 30 minutes, the fluorogenic reference substrate Boc-Gln-Ala-Arg-AMC was added to matriptase and 8were added to thrombin or factor Xa, respectively. The velocity of substrate degradation was assessed by recording the fluorescence intensity at 460 nm or the absorbance at 405 nm within 2 h. Shown are the means ± SD of triplicate measurements.
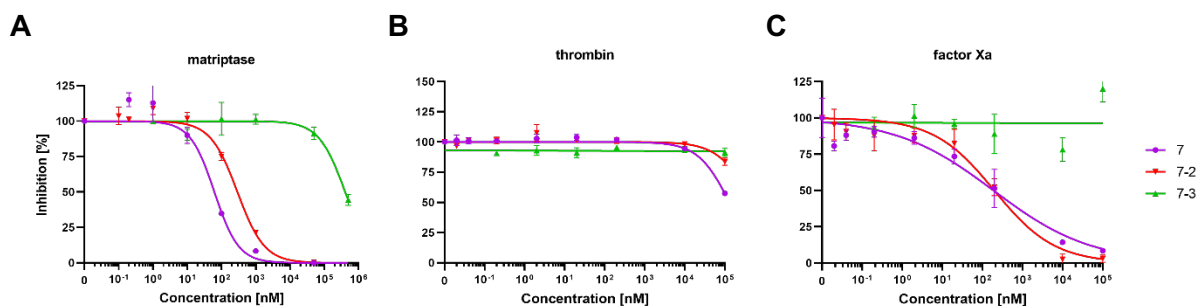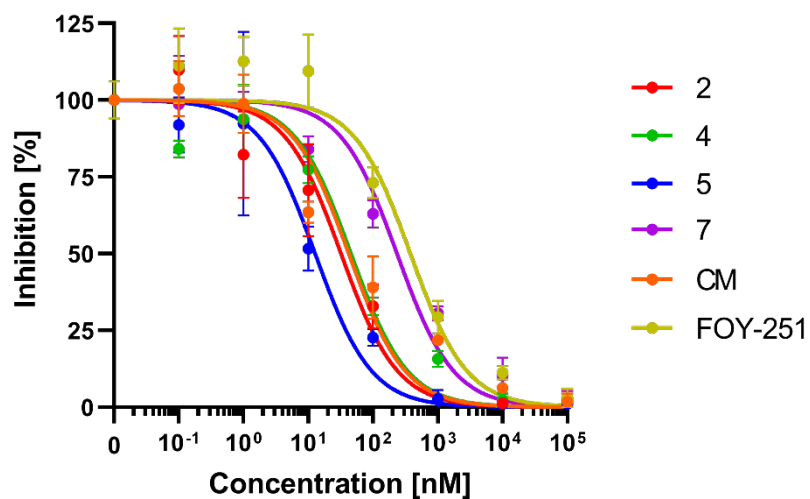
Figure A14. Peptidomimetic inhibitors block protease activity of Caco-2 cells. Peptidomimetic compounds 2, 4, 5 and 7 as well as camostat mesylate (CM) and FOY-251 were added to Caco-2 cells. After 30 min, the fluorogenic reference substrate Boc-Gln-Ala-Arg-AMC was added, and the reaction rate of substrate degradation was assessed by recording the fluorescence intensity within 2 h. Shown are the means ± SD of triplicate measurements. Calculated IC_{50} values for each compound are presented in Table 5.10.
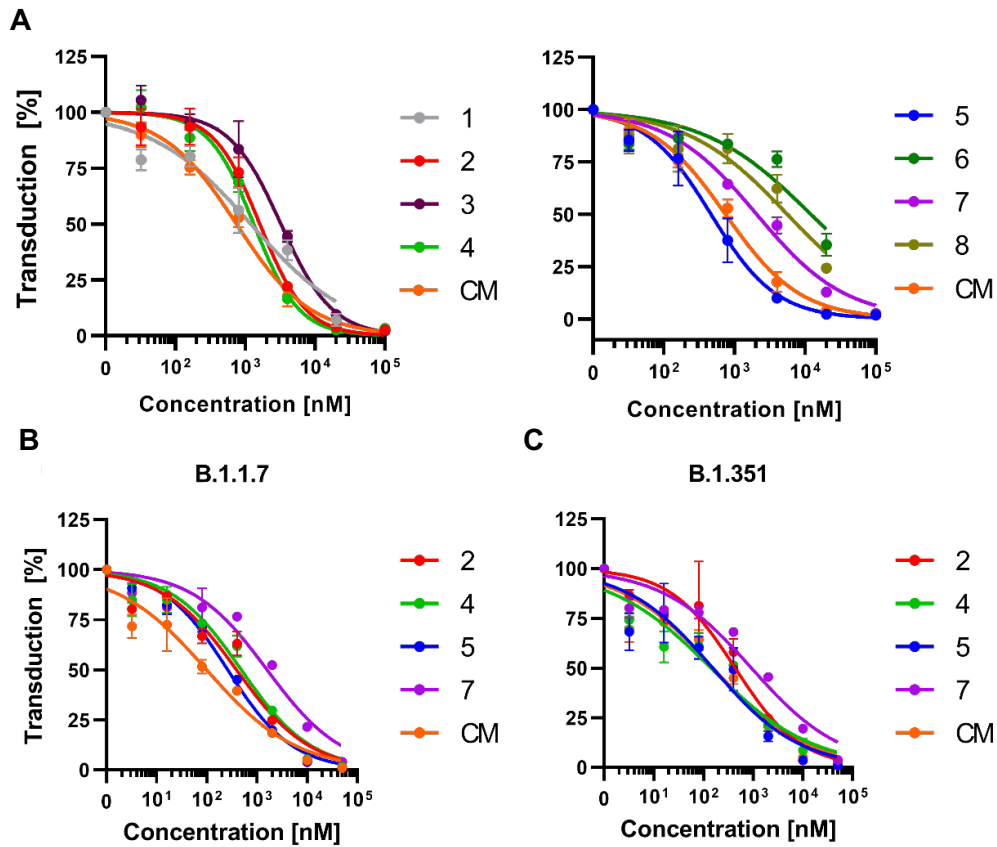
Figure A15. Peptidomimetic inhibitors reduce SARS-CoV-2 spike driven entry Peptidomimetic inhibitors and the small molecule camostat mesylate (CM) were added to Caco-2 cells. After 1 h, cells were transduced with lentiviral SARS-CoV-2 pseudoparticles carrying the spike protein of SARS-CoV-2 wildtype (**A**), the B1.1.7 (**B**) or B.1.351 (**C**) variant of concern. Transduction rates were assessed 2 days post transduction by measuring luciferase activity in cell lysates. Shown are the means ± SEM of two independent experiments, each performed in triplicates. Calculated IC_{50} values for each compound are presented in Table 5.11.

A

B D614G

C B.1.1.7

D B.1.351
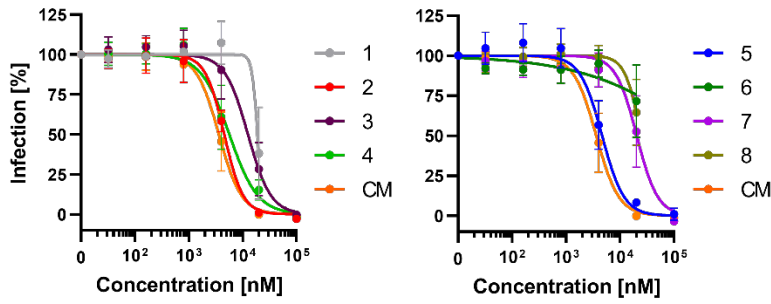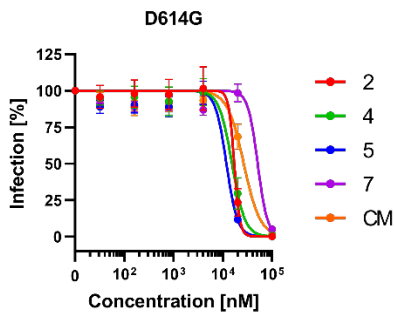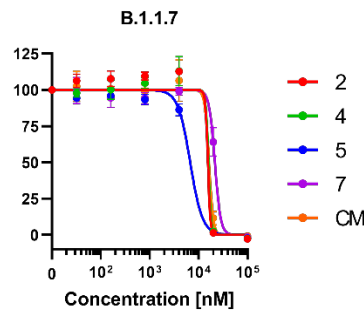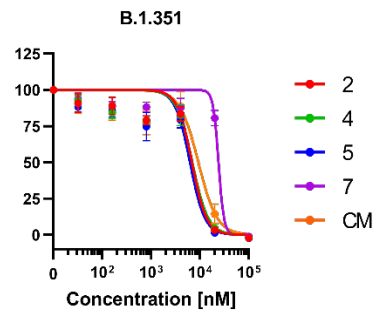
Figure A16. Peptidomimetic inhibitors reduce SARS-CoV-2 infection. Peptidomimetic inhibitors and the small molecule camostat mesylate (CM) were added to Caco-2 cells. After 1 h, cells were infected with SARS-CoV-2 WT (**A**), SARS-CoV-2 bearing the spike D614G mutation (**B**), or the variants of concern B.1.1.7 (**C**) and B.1.351 (**D**). Infection rates were determined 2 days post transduction by in cell ELISA for the viral N protein. Shown are the means ± SEM of three independent experiments, each performed in triplicates. Calculated IC_{50} values for each compound are listed in Table 5.12.

Table A1. Hydrolysis of IQF peptides with optimized substrate processing sites by hepsin. Hydrolysis of the IQF peptides (0–200 µM) was monitored containing 0.2 ng/µL hepsin and the constants were calculated from nonlinear regressions of hyperbolic Michaelis–Menten rate equations. Relative catalytic efficiencies (%) are the *catalytic efficiency* values of the IQF peptides relative to that of the reference peptide. Enzymatic measurements were performed in triplicates.

IQF substrate	K_M [µM]	k_{cat} [sec^{-1}]	k_{cat}/K_M [sec^{-1} µM^{-1}]	rel. catalytic efficiency
reference	47.7	8.6	18.5×10^{-2}	100%
P1	25.0	5.7	24.2×10^{-2}	131%
P2	30.4	6.7	23.0×10^{-2}	124%
P3	6.3	3.7	59.1×10^{-2}	319%
P4	26.2	7.4	28.0×10^{-2}	151%
P'1	48.1	9.9	22.2×10^{-2}	120%
P'2	22.6	7.8	34.6×10^{-2}	187%
P'3	19.6	8.3	42.9×10^{-2}	232%
P'3.1	36.9	6.3	17.1×10^{-2}	95%
P'3.2	23.7	7.5	31.6×10^{-2}	176%
P'3.3			not cleaved	
P'3.4	28.5	7.6	26.7×10^{-2}	148%
P''1	63.9	9.8	14.9×10^{-2}	800%
P''2	27.5	5.1	19.5×10^{-2}	105%
negative 1			not cleaved	
negative 2			not cleaved	
negative control			not cleaved	

Table A2. Michaelis Menten analysis of IQF substrates in 12.5% blood serum. Relative catalytic efficiencies (%) are the *catalytic efficiency* values of the IQF peptides relative to that of the reference peptide. Enzymatic measurements were performed in triplicates.

IQF substrate	K_M [µM]	k_{cat} [sec^{-1}]	k_{cat}/K_M [sec^{-1} µM^{-1}]	rel. catalytic efficiency
reference	58.5	3.77×10^{-2}	6.44×10^{-4}	100%
P1	353.5	1.04×10^{-2}	2.95×10^{-4}	46%
P2	150.3	4.52×10^{-2}	3.00×10^{-4}	47%
P3	63.7	2.86×10^{-2}	4.48×10^{-4}	70%
P4	79.3	4.47×10^{-2}	5.63×10^{-4}	87%
P'1	79.6	4.16×10^{-2}	5.22×10^{-4}	83%
P'2	90.2	4.23×10^{-2}	4.68×10^{-4}	73%
P'3	84.4	3.81×10^{-2}	4.51×10^{-4}	70%
P'3.1	not cleaved			
P'3.2	not cleaved			
P'3.3	118.8	3.22×10^{-2}	2.70×10^{-4}	41%
P'3.4	77.7	4.30×10^{-2}	5.53×10^{-4}	86%
P''1	159.3	7.60×10^{-2}	4.77×10^{-4}	74%
P''2	44.60	2.34×10^{-2}	5.26×10^{-4}	82%
negative 1	138.5	3.57×10^{-2}	2.58×10^{-4}	40%
negative 2	not cleaved			
negative control	not cleaved			

Table A3. Michaelis Menten analysis of IQF substrates in 12.5% blood plasma. Relative catalytic efficiencies (%) are the *catalytic efficiency* values of the IQF peptides relative to that of the reference peptide. Enzymatic measurements were performed in triplicates.

IQF substrate	K_M [µM]	k_{cat} [sec^{-1}]	k_{cat}/K_M [sec^{-1} µM^{-1}]	rel. catalytic efficiency
reference	34.0	2.62×10^{-2}	7.71×10^{-4}	100%
P1	285.4	8.23×10^{-2}	2.88×10^{-4}	37%
P2	700.5	1.12×10^{-2}	1.65×10^{-4}	21%
P3	60.3	2.28×10^{-2}	3.78×10^{-4}	49%
P4	526.4	1.66×10^{-2}	3.14×10^{-4}	41%
P′1	97.4	3.48×10^{-2}	3.58×10^{-4}	46%
P′2	49.5	2.65×10^{-2}	5.35×10^{-4}	69%
P′3	46.9	2.65×10^{-2}	5.65×10^{-4}	73%
P′3.1	not cleaved			
P′3.2	not cleaved			
P′3.3	173.5	3.50×10^{-2}	2.01×10^{-4}	26%
P′3.4	57.3	2.96×10^{-2}	5.17×10^{-4}	67%
P″1	136.5	3.57×10^{-2}	2.62×10^{-4}	32%
P″2	1213	29.67×10^{-2}	2.45×10^{-4}	34%
negative 1	82.6	2.00×10^{-2}	2.43×10^{-4}	31%
negative 2	not cleaved			
negative control	not cleaved			

Table A4. Docking result for P-site (P1−P4)-designed molecules with the structure being ace-X-X-X-R-methylamide. The letters under P1-P4 represent the one letter code for amino acids. [a] P-site recognition sequence of a literature known preferred substrate RQLR↓VVGG.[b] P-site recognition sequence of the endogenous substrate HGF.[138]

P4	P3	P2	P1	↓	score [kJ/mol] (Rank)
Arg	Gln	Leu	Arg		-60.1 (175)[a]
Lys	Gln	Leu	Arg		-33.9 (4139)[b]
Arg	Glu	Cys	Arg		-73.1 (1)
Ala	Glu	Gln	Arg		-72.7 (2)
Gly	Gln	Arg	Arg		-72.7 (3)
Gln	Gln	Val	Arg		-69.5 (4)
Arg	His	Gln	Arg		-69.1 (6)
Glu	Gln	Pro	Arg		-68.9 (7)
Asn	Glu	Gln	Arg		-67.6 (8)

Table A5. The convergence analyses for the system α1AT–Trypsin. Binding free energy, importance sampling ratio (ISR) and convergence parameters obtained for all the checkpoints in the system α1-AT – Trypsin.

Checkpoint (%)	ΔG_{C2} (kcal/mol)	ISR (%)	Convergence
1 (37)	-10.6 ± 0.94	12.77	WARNING3
2 (44)	-20 ± 1.2	23.59	WARNING3
3 (51)	-5.5 ± 0.71	19.19	WARNING3
4 (58)	-5.3 ± 0.83	43.68	WARNING3
5 (65)	-7.7 ± 0.53	19.85	OK
6 (72)	-8.4 ± 0.83	23.35	WARNING3
7 (79)	-12.5 ± 0.63	22.57	WARNING3
8 (86)	-13.5 ± 0.59	37.78	WARNING3
9 (93)	-10.7 ± 0.45	39.67	OK
10 (100)	-11.6 ± 0.48	35.12	OK
Average*	-10 ± 1.6		

*ISR-weighted average

Table A6. The convergence analyses for the system α1AT–TMPRSS2. Binding free energy, importance sampling ratio (ISR) and convergence parameters obtained for all the checkpoints in the system α1AT–TMPRSS2.

Checkpoint (%)	ΔG_{C2} (kcal/mol)	ISR (%)	Convergence
1 (37)	-13.4 ± 0.95	8.58	WARNING2
2 (44)	-9.2 ± 0.82	3.31	WARNING2
3 (51)	-10.0 ± 0.59	4.74	OK
4 (58)	-1.2 ± 0.66	11.10	WARNING3
5 (65)	-13.2 ± 0.58	15.98	OK
6 (72)	-15.2 ± 0.76	23.39	WARNING3
7 (79)	-11.9 ± 0.64	23.28	WARNING3
8 (86)	-10.4 ± 0.69	22.13	WARNING3
9 (93)	-7.9 ± 0.54	23.75	OK
10 (100)	-11.0 ± 0.68	48.13	WARNING3
Average*	-10 ± 2.2		

*ISR-weighted average

Table A7. Identified α1AT peptides in MALDI-TOF MS analysis from fraction 42_55. Predominant band in fraction 42_55 (Figure 6.3) was analyzed by MALDI-TOF mass spectrometry and identified mass-to-charge ratios with corresponding peptides are shown.

Mass-to-charge ratio (m/z) measured	Corresponding Peptide		AA position in α1-AT
	AA sequence	Calculated relative mass	
852.48	R.SASLHLPK.L	851.49	307 - 314
888.49	K.AVLTIDEK.G	887.50	360 - 367
922.43	K.FLENEDR.R	921.42	299 - 305
1,008.49	K.QINDYVEK.G	1,007.49	180 - 187
1,015.61	K.SVLGQLGITK.V	1,014.61	325 - 334
1,076.58	K.LSSWVLLMK.Y	1,075.61	259 - 267
1,090.57	K.WERPFEVK.D	1,089.56	218 - 225
1,110.60	K.LSITGTYDLK.S	1,109.60	315 - 324

1,247.61	R.LGMFNIQHCK.K	1,246.60	248 - 257
1,263.61	R.LGMFNIQHCK.K	1,262.59	248 - 257
1,576.83	R.DTVFALVNYIFFK.G	1,575.83	203 - 215
1,641.84	K.ITPNLAEFAFSLYR.Q	1,640.86	50 - 63
1,779.76	K.TDTSHHDQDHPTFNK.I	1,778.76	35 - 49
1,803.94	K.LQHLENELTHDIITK.F	1,802.95	284 - 298
1,855.93	K.FNKPFVFLMIEQNTK.S	1,854.97	390 - 404
1,871.97	K.FNKPFVFLMIEQNTK.S	1,870.97	390 - 404
1,891.84	K.DTEEEDFHVDQVTTVK.V	1,890.85	226 - 241
2,057.97	K.LYHSEAFTVNFGDTEEAK.K	2,056.94	161 - 178
2,259.13	K.GTEAAGAMFLEAIPMSIPPEVK.F	2,258.13	368 - 389
2,275.16	K.GTEAAGAMFLEAIPMSIPPEVK.F	2,274.13	368 - 389
2,291.18	K.GTEAAGAMFLEAIPMSIPPEVK.F	2,290.12	368 - 389
878.48	K.FLEDVKK.L	877.49	154 - 160
1,078.52	K.FLENEDRR.S	1,077.52	299 - 306
1,136.56	K.KQINDYVEK.G	1,135.59	179 - 187
1,275.68	K.GKWERPFEVK.D	1,274.68	216 - 225
1,391.71	R.LGMFNIQHCKK.L	1,390.69	248 - 258
2,090.08	K.ELDRDTVFALVNYIFFK.G	2,089.09	199 - 215
2,186.06	K.LYHSEAFTVNFGDTEEAKK.Q	2,185.03	161 - 179
2,707.26	K.LQHLENELTHDIITKFLENEDR.R	2,706.36	284 - 305

Curriculum Vitae

Personal Information

Name: Philip Maximilian Knaff

Birth: 30.05.1994

Place of birth: Esch-sur-Alzette (Luxembourg)

Nationality: German & Luxembourg

Education

08.2018 - present: Doctoral candidate in Chemistry at Max Planck Institute for Polymer Chemistry in Mainz, Germany

10.2016 - 06.2018: Master of Science in Biomedical Chemistry at Johannes Gutenberg University in Mainz, Germany

03.2016 - 10.2016: Research at Institute for Organic Chemistry at University of Alabama in Tuscaloosa, USA

04.2013 - 03.2016: Bachelor of Science in Biomedical Chemistry at Johannes Gutenberg University in Mainz, Germany

04.2000 - 03.2013: Graduation from Werner Heisenberg high school in Bad Dürkheim, Germany

Work Experience

06.2018 - 07.2018: Internship at Department of Forensic Toxicology Service at Laboratoire National de Santé in Dudelange, Luxembourg

08.2017 - 09.2017: Tutor for students at institute for Biochemistry

03.2016 - 04.2016: Internship at University Medical Center

08.2015 - 10.2015: Working student at BASF SE

08.2014 - 09.2014: Working student at BASF SE

Acknowledgement

First of all, I want to thank Prof. Dr. Katharina Landfester and Prof. Dr. Volker Mailänder for giving me the opportunity to work on those very interesting research projects under such good working conditions. I am very grateful for your supervision and support throughout my doctoral studies.

I further want to thank my cooperation partners for sharing knowledge, ideas and for reviewing the manuscripts. Special thanks for the great collaborative work to Dr. Christian Kersten, Patrick Müller and Prof. Dr. Tanja Schirmeister from the Pharmaceutical Institute in Mainz, who accompanied all of my three major projects.

Many thanks to my colleagues Lukas Wettstein and Prof. Dr. Jan Münch from the Institute of Molecular Virology in Ulm. A virus tremendously affects people's lives all around the globe and fighting at the scientific forefront for the search of a therapy is very encouraging. I hope our research will make a contribution with great benefit to society.

Further I want to thank the technicians Beate Müller, Christine Rosenauer, Gunnar Glaser, Katrin Kirchhoff and Angelika Manhart at AK Landfester. I appreciate very much their great expertise in analytical methods.

I want to thank the people from the working group "AK Landfester" and also from our small subgroup "AG Mailänder" for the great working atmosphere and friendships. Here at MPIP, people from all around the globe come together with different academic backgrounds bringing international spirit and new ideas. I will keep good memories about the discussions in chemistry-related fields but also from our recreation activities. Hopefully, we will see our self-built boat cross the rhine river!

Finally, I want to thank my family, especially my parents and my brother for their long-term support and encouragement during my studies. Thank you.