Exploration of Covalent Inhibitors for Trypsin-Like Serine Proteases

Dissertation

zur Erlangung des Grades

„Doktor der Naturwissenschaften (Dr. rer. nat.)“

im Promotionsfach Pharmazeutische und Medizinische Chemie

am Fachbereich Chemie, Pharmazie, Geographie und Geowissenschaften

der Johannes Gutenberg-Universität Mainz

Patrick Müller, Master of Science

geb. in Mainz

Submitted to the Faculty of Chemistry, Pharmacy, Geography, and Geoscience.

Dean: [redacted]

Name of the 1st reviewer: [redacted]

Name of the 2nd reviewer: [redacted]

Date of the doctoral examination: 18.06.2024

Declaration of Authorship

I, Patrick Müller, declare that this thesis entitled, “Exploration of Covalent Inhibitors for Trypsin-Like Serine Proteases”, and the work presented in it are my own. I confirm that:

- I did this work completely while in candidature for a research degree at this university.
- I have clearly stated which parts of this thesis have already been submitted for another degree or qualification at this university or other institutions.
- I always attributed clearly where I have consulted the published work of others.
- I always gave the source when I quoted from other works. Except for such quotations, this thesis is entirely my work.
- I have acknowledged all main sources of help.
- I have made clear exactly what was done by others and what I have contributed myself.

Place, Date: Mainz, 23.04.2024 Signature: 

Acknowledgments

Table of Contents

List of Abbreviations

[*I*]	Inhibitor concentration
[*S*]	Substrate concentration
ACE2	Angiotensin-converting enzyme 2
AMC	7-Amino-4-methylcoumarin
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
ATF	Amino-terminal fragment
AU	Absorbtion units
Cys	Cysteine
Clp	Caseinolytic protease
CP	Cytoplasmic component
Da	Dalton
DENV	Dengue virus
DESC	Differentially expressed in squamous cell carcinoma gene
E	Envelope protein
EC	Extracellular component
ECM	Extracellular matrix
EGFR	Epidermal growth factor receptor
Em	Emission
ER	Endoplasmic reticulum
[EI]	(Non-)covalent reversible enzyme-inhibitor complex
[EI]*	Covalent irreversible enzyme-inhibitor complex
F	Fluorophore
FAK	Focal adhesion kinase
Fl	Fluorescence
FDA	Food and Drug Administration
FP	Fusion protein
FRET	Förster resonance energy transfer
GFD	Growth-factor domain
Gln	Glutamine
Gly	Glycine
GPI	Glycosyl phosphatidylinositol
GRB2	Growth factor receptor-bound protein 2

HAT	Human airway trypsin-like protease
His	Histidine
HMW	High-molecular weight
IC_{50}	Inhibitor concentration by half-maximal enzyme inhibition
Ile	Isoleucine
KD	Kringle domain
K_i	Inhibition constant
K_M	Michaelis-Menten constant
LC-MS	Liquid chromatography-mass spectrometry
Lys	Lysine
LDLRA	Low-density lipoprotein receptor class A domain
M	Membrane protein
Met	Hepatocyte growth factor receptor
Metallo	Metalloproteinase
N	Nucleocapsid protein
NBD	7-Nitrobenz-2-oxa-1,3-diazol-4-amine
NMR	Nuclear magnetic resonance
NS3	Non-structural protein 3 (serine protease)
NS2B-NS3	Non-structural protein 2B-Non-structural protein 3 (serine protease)
Nsp	Non-structural protein
NTD	N-terminal domain
PDB	Protein data bank
Phe	Phenylalanine
pNA	*p*-Nitroanilide
Pro-HGF	Pro-hepatocyte growth factor
Q	Quencher
Ras	Rat sarcoma
RBD	Receptor-binding domain
RNA	Ribonucleic acid
S	Spike protein
SAR	Structure-activity relationship
SARS-CoV-2	Severe acute respiratory syndrome-coronavirus type 2
Sec	Seconds
Ser	Serine
SRCR	Scavenger receptor cysteine-rich domain
STAT	Signal transducers and activators of transcription

Thr	Threonine
TM	Transmembrane
TMPRSS2	Transmembrane protease serine subtype 2
Trp	Tryptophan
TTSP	Type II transmembrane serine protease
Tyr	Tyrosine
uPA	Urokinase-type plasminogen activator
uPAR	Urokinase-type plasminogen activator receptor
uPAS	Urokinase-type plasminogen activator system
v	Reaction rate (velocity)
Val	Valine

Abstract

Eliminating past concerns about toxicity issues, covalent inhibitors have successfully overcome their bad reputation and ushered a new era in the development of protease inhibitors. This resulted in the approval of several drugs by the U.S. Food and Drug Administration. A peptidomimetic recognition sequence, combined with an electrophilic moiety (warhead), which is responsible for the covalent bond to the target enzyme, has been recognized as an effective inhibitor scaffold. The present dissertation is about the investigation and characterization of covalent peptidomimetic protease inhibitors, regarding their pharmacodynamic properties and improvement of the associated fluorometric inhibition assay (chapter A) as well as the evaluation of covalent inhibitors for trypsin-like serine proteases (chapter B).

Chapter A: To reduce possible side effects of covalent inhibitors, due to non-selective binding to physiological nucleophiles and other enzymes, the right choice of an appropriate warhead in combination with a suitable peptidomimetic recognition sequence is of utmost importance during drug development. Within the first project, an assortment of peptidomimetic scaffolds targeting different proteases, combined with various warheads were synthesized. Fluorometric inhibition studies of the inhibitors revealed the impact of both structure parts concerning their affinity to its target protease and selectivity profile. Additional NMR- and LC-MS-based reactivity studies were performed, investigating the chemoselectivity of the used warheads against model nucleophiles of serine/threonine and cysteine proteases. The observed results in both test systems were further explained by quantum mechanics and docking simulations.

Conventionally used fluorogenic substrates for inhibition assays display excitation and emission maxima in the lower wavelength area, which is vulnerable to assay interferences and bears the danger of leading to false positive or negative results. Here, the development and characterization of new 7-nitrobenz-2-oxa-1,3-diazol-4-amine (NBD)-based substrates presented an alternative strategy to mitigate common assay interferences, due to the shift of the excitation and emission wavelength of the NBD fluorophore to higher wavelengths and resulting in more reliable inhibition data.

Chapter B: Trypsin-like serine proteases are involved in many important physiological processes like blood coagulation and remodeling of the extracellular matrix. However, they are also associated with pathological events. Overexpression and dysregulation of urokinase-type plasminogen activator (uPA) and hepsin promotes tumor progression, proliferation and metastasis. Another member of this protease class represents the transmembrane protease serine subtype 2 (TMPRSS2), which enables the viral cell entry of coronaviruses and received a lot of attention during the SARS-CoV-2 pandemic. Within this chapter, peptidomimetic covalent protease inhibitors were designed to target each protease with high potencies and promising selectivity profiles.

Kurzfassung

Kovalente Inhibitoren konnten ihren schlechten Ruf bezüglich früherer Toxizitätsbedenken überwinden und läuteten eine neue Ära der Entwicklung von Proteaseinhibitoren ein, welche zu mehreren Arzneimittelzulassungen der U.S. Food and Drug Administration führte. Dabei kristallisierte sich eine peptidomimetische Erkennungssequenz, gepaart mit einer geeigneten elektrophilen Kopfgruppe (Warhead) als effektives Grundgerüst für die Entwicklung von Inhibitoren heraus. Die vorliegende Dissertation handelt von der Untersuchung kovalenter Proteaseinhibitoren bezüglich ihrer pharmakodynamischen Eigenschaften sowie Optimierung des fluoreszenzbasierten Inhibitionsassays (Kapitel A), und präsentiert das rationale Design kovalenter Inhibitoren gegenüber Trypsin-ähnlichen Serinproteasen (Kapitel B).

Kapitel A: Zur Vermeidung potenzieller Nebenwirkungen, verursacht durch unspezifische Bindung an physiologische Nukleophile und andere Enzyme, ist es von größter Wichtigkeit, die richtige Wahl des kovalenten Warheads und der peptidomimetischen Erkennungssequenz im Rahmen der Inhibitor-Entwicklung zu treffen. Das erste Projekt umfasst die Synthese kovalenter Inhibitoren mit Kombinationen verschiedener peptidomimetischer Grundgerüste und Warheads sowie die Charakterisierung ihrer Inhibition durch Fluoreszenzassays gegenüber unterschiedlichen Proteasen. Zusätzlich wurden NMR- und LC-MS-basierte Reaktivitätstestungen mit den ausgewählten Warheads gegenüber Modell-Nukleophilen der Serin-/Threonin- und Cysteinproteasen durchgeführt. Die erhaltenen Ergebnisse beider Testsysteme wurden mit quantenmechanischen und computergestützten Bindungsexperimenten untermauert.

Häufig verwendete Fluoreszenzsubstrate des oben erwähnten Inhibitionsassays besitzen Anregungs- und Emissionwellenlängen der jeweiligen freien Fluorophore im kurzwelligen Bereich. Dadurch besteht die Gefahr von typischen Assay-Interferenzen, welche zu falsch-positiven und -negativen Ergebnissen führen können. Aufgrund dessen wurden durch Nutzung von 7-Nitrobenz-2-oxa-1,3-diazol-4-amin (NBD) als Fluorophor Anregungs- und Emissionswellenlängen in den höheren Wellenlängenbereich verschoben, sodass typische Assay-Interferenzen vermieden und verlässliche Inhibitionsergebnisse erhalten werden können.

Kapitel B: Trypsin-ähnliche Serinproteasen stellen einen wichtigen Bestandteil von vielen physiologischen Prozessen wie der Blutgerinnung und der Umstrukturierung der extrazellulären Matrix dar, können jedoch bei Deregulierung pathologische Ereignisse fördern. Die Enzyme Urokinase-Typ Plasminogen-Aktivator (uPA) und Hepsin können beide bei Überexpression die Tumorprogression, -proliferation und -metastasierung unterstützen. Ein weiteres Beispiel stellt die Transmembran-Serinprotease 2 (TMPRSS2) dar, welche den viralen Zelleintritt von Coronaviren ermöglicht und erneut durch die SARS-CoV-2 Pandemie in den Fokus geriet. Kapitel B beschreibt das rationale Design von peptidomimetischen kovalenten Inhibitoren gegenüber den drei genannten Serinproteasen.

1. Introduction

1.1 Functional Role and Classification of Proteases

The hydrolysis of amide bonds is probably one of the oldest known mechanism of proteins. At the beginning of the 20th century, it was long considered that the process of proteolytic enzymes only served the purpose of receiving amino acid nutrients by nonspecific degradation of dietary proteins (protein catabolism).[1],[2] However, the pioneering work on trypsinogen activation and mechanism of blood clotting in the middle of the 20th century revealed, that precise cleavage of proteins by proteolytic enzymes, so called proteases, can regulate other proteins or even affect signal pathways.[3]–[5] This led to the concept of "proteolytic processing", in which inactive precursors or zymogens are converted to their physiologically active forms with selective enzymatic hydrolysis of peptide bonds.[6] According to this realization, many functional roles of proteases were discovered and characterized. This comprises intra- and extracellular processes like cell proliferation and cell death, tissue remodeling, immune response and hemostasis.[7]–[11] In contrast to the majority of other signaling pathways, protease signaling is mostly irreversible.[6] Thus, constant and strict regulation of the enzymatic activity, which includes gene expression, zymogen inactivation or endogenous inhibitors, plays a crucial role for maintaining key physiological roles.[6],[12],[13] Therefore, it is not difficult to understand that dysregulation of proteases can have severe pathological consequences such as cancer, cardiovascular or neurodegenerative diseases.[14]–[16]

The evolutionary complexity of living organisms presents a large repertoire of proteases, defined by different functionalities with various structural characteristics and specificities. The effort of the scientific community who seeks new applications for proteases under physiological or pathological conditions, led to the development of *degradomics* as an experimental field for identification and characterization of proteases, their substrates and inhibitors.[6],[17],[18] By using information retrieved from sequencing projects, combined with bioinformatic analysis, over 550 proteases (~2% of the genome) have been identified for the human degradome.[17],[19],[20] To provide an overview of the large number of different proteases, mainly two international systems for protease nomenclature are used. The Enzyme Commission Number (EC number), developed in the 1950s is a numerical classification for enzymes, based on the chemical reactions they catalyze.[21] Proteases are classified in class EC 3, as hydrolases, subclass EC 3.4 peptidases, due to the catalytic hydrolyzation of peptide bonds.[18] Furthermore, peptidases are divided into endopeptidases and exopeptidases, depending on their cleavage site. Exopeptidases cleave only one of the three amide bonds near the *C* (carboxypeptidase)- or *N* (aminopeptidase)-terminus. In contrast, endopeptidases can cleave bonds much further away from both termini but may also cleave peptide bonds near a *C*-or *N*-terminus.[22] Nowadays, peptidases can additionally be characterized by their catalytic mechanism, due to a better knowledge of protease

structures. The peptidase database MEROPS classifies proteases by their catalytic mechanism, whereby the names originate from the catalytic active amino acid or metal, present in the active site: (A) aspartic peptidases (EC 3.4.23), (C) cysteine peptidases (EC 3.4.22), (G) glutamic peptidases (EC 3.4.19), (M) metallo peptidases (EC 3.4.24), (S) serine peptidases (EC 3.4.21), (T) threonine peptidases (EC 3.4.25) and (U) unknown peptidases. A seventh group has been identified with the asparagine peptide lyases (N). However, due to a different catalytic mechanism without consumption of a water molecule, it is assigned to the lyases (EC 4.3.2).[23] Furthermore, the classes are categorized with respect to their tertiary structure and basis of common ancestry into clans and families (Figure 1).[22]

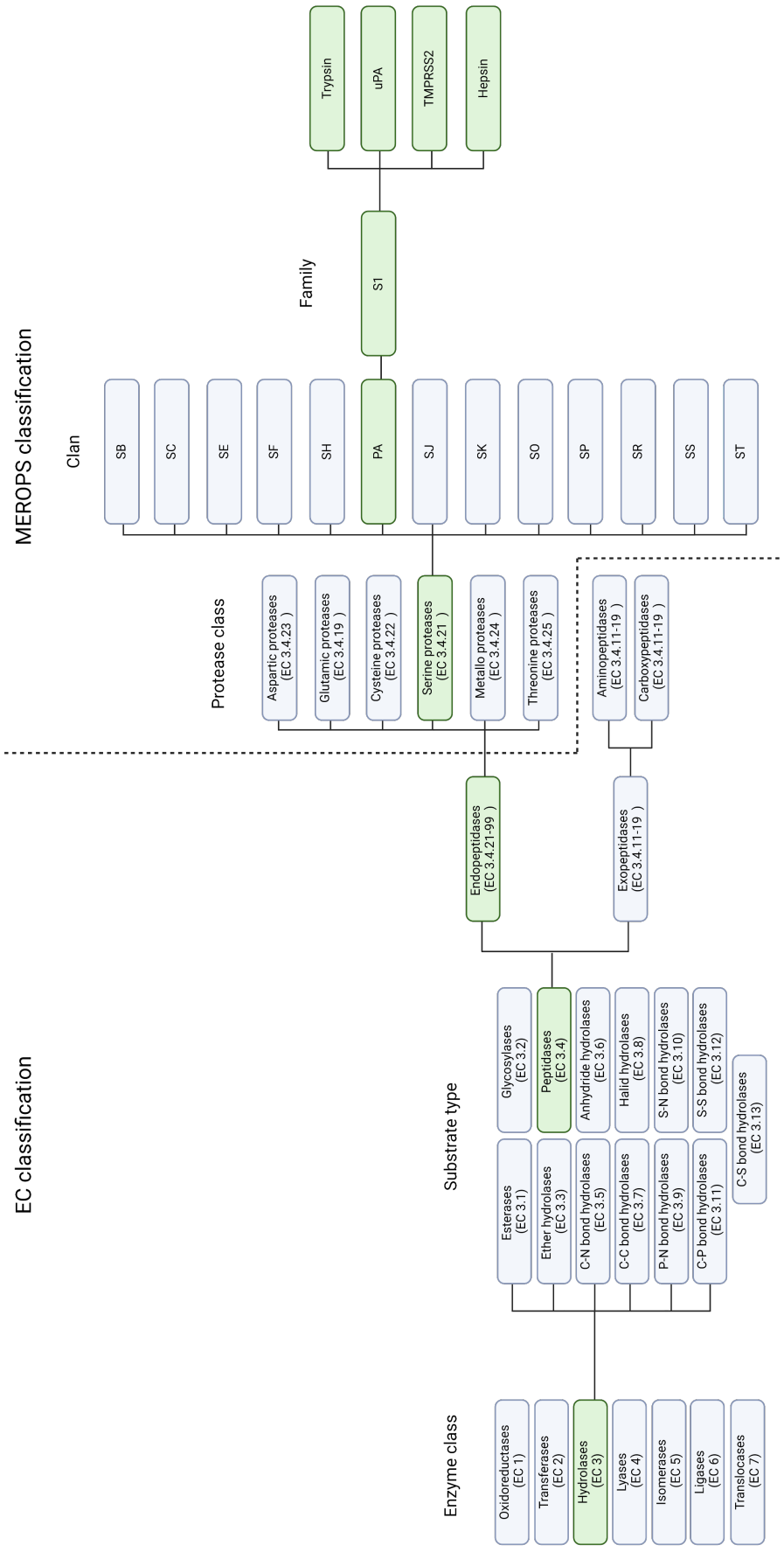


Figure 1: Classification of enzymes according to the enzyme commission number system (EC number; left) based on their catalyzed chemical reaction. More in depth classification by the MEROPS database of peptidases, which categorize proteolytic enzymes to their catalytic mechanism, structural properties (clan) and ancestral origin (family; right). The trypsin-like serine proteases trypsin, urokinase-type plasminogen activator (uPA), transmembrane protease serine subtype 2 (TMPRSS2) and hepsin (green highlighted) will be discussed in more detail, whereas the last three present a part of the scientific work in this doctoral thesis. Figure created with BioRender.com.

Proteases take on a broad range of physiological roles, which is mainly due to the large repertoire of specialized substrate pockets. This comprises binding pockets from highly promiscuous proteases that can cleave a big variety of substrates, to proteases with specificity for particular substrate sequences.[24],[25] Molecular interactions between the substrate and the binding pocket of the protease determines substrate specificity, which are mainly represented by electrostatic and hydrophobic interactions.[26],[27] The amino acid side chains of substrate peptides protrude into sub-pockets of the protease binding site. This results in a locked extended *β*-strand conformation of the substrate in the binding cleft. Therefore, folded proteins which consist of helices, turns/loops and sheets are protected from degradation.[28] Schechter and Berger developed a nomenclature, which describes the substrate-binding site of the protease and the residues of the substrate that occupy them.[28] The surface of the protease within the binding site, which can accommodate one side chain of a substrate residue is called subsite or subpocket. Beginning from the scissile bond, which is located between the residues P1 and P1′ of the substrate, subpockets of the binding sites are numbered S1–Sn towards the *N*-terminus (non-primed site) and S1′–Sn′ towards the *C*-terminus (primed site). In analogy, the side chains of the substrate are described as P1–Pn (non-primed site) and P1′–Pn′ (primed site; Figure 2).[1],[25]

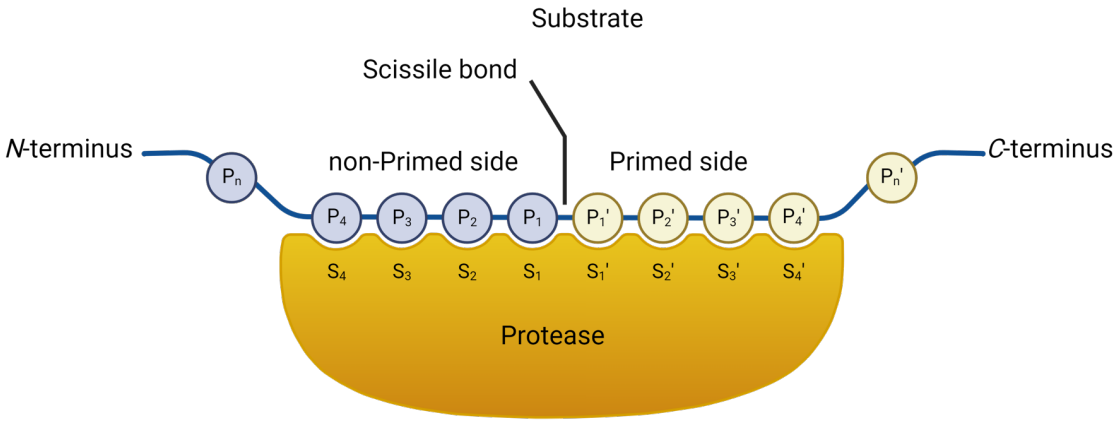


Figure 2: Schematic representation of a protease binding site with its substrate. Numbering of the substrate residues according to Schechter and Berger nomenclature, with P1–Pn towards the *N*-terminus (blue) and P1′–Pn′ towards the *C*-terminus (beige) and in similar fashion to the subpockets (S1–Sn, S1′–Sn′). Figure created with BioRender.com and modeled after Turk *et al.*[1]

The formation of a peptide bond occurs via a dehydration reaction, where water is released. Therefore, it is not surprising that the breakdown of a peptide is accomplished by hydrolysis. The six major protease classes can be divided into two classes, concerning their catalytic mechanism. Cysteine, serine and threonine proteases harbor a nucleophilic amino acid in their active site (covalent catalysis) and stabilize the tetrahedral transition-state by a covalent bond between the substrate and the catalytic active amino acid.

The aspartic, glutamic and metalloproteases instead use an activated water molecule (non-covalent catalysis) for formation of the tetrahedral transition-state during the peptide breakdown (Figure 3).[1],[22],[25]

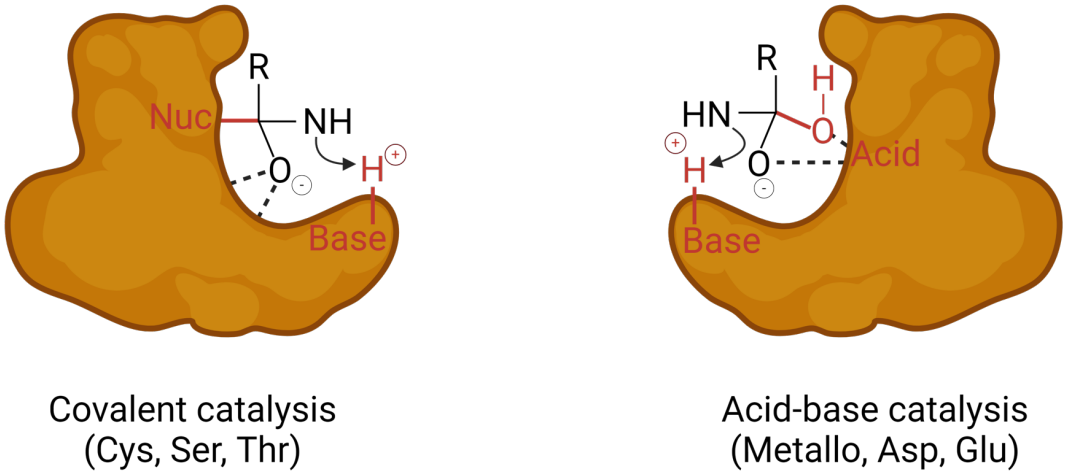


Figure 3: Catalytic mechanism of the six major protease classes. Cysteine, serine and threonine proteases harbor the respective amino acid residue that resembles the nucleophile of the catalytic site (covalent catalysis) and gets activated by a base in proximity (often histidine; left). Metallo-, aspartate and glutamate proteases harbor an activated water molecule in the active site as the nucleophile (acid-base catalysis), where aspartate, glutamate, or zinc (metalloproteases) work as acids and bases. Figure created with BioRender.com and modeled after Turk *et al*.[1]

1.2 Serine Proteases as Potential Therapeutic Targets

1.2.1 General Properties of Serine Proteases

Serine proteases (EC 3.4.21) represent almost one-third of all identified proteases and are considered as the largest proteolytic class.[24],[29],[30] The structure of the catalytic machinery of serine proteases was originally distinguished by the catalytic triad Asp-His-Ser of α-chymotrypsin in 1964.[31] X-ray crystallography revealed that the catalytic Ser195 was in proximity to His57 to form a hydrogen bond. Later, Asp102 was identified to also play an important catalytic role as part of the "charge relay" system due to polarization properties.[24],[31],[32] Before the structure elucidation of subtilisin, it was thought that all serine proteases would be homologous to α-chymotrypsin. However, 1969 X-ray crystallography showed that subtilisin harbors the same catalytic triad, with a totally different main-chain structure.[33] The presence of the identical amino acids in the active site in a different structural context of subtilisin demonstrated that there must be independent evolutionary paths to the triad.[32] Overall, there are four existing distinct protein folds, illustrated by chymotrypsin, subtilisin, carboxypeptidase and Clp protease that utilize the Asp-His-Ser triad in the same manner to catalyze hydrolysis of peptide bonds.[30] Moreover, further serine proteases mediate catalysis with different catalytic triads, such as a pair of His residues combined with the nucleophilic serine (His-Ser-His) or an exchange of the aspartate residue

with glutamate (Glu-His-Ser). Other proteases employ a simpler mechanism with dyads containing Lys or His paired with the nucleophilic serine.[30],[32] Figure 4 presents five examples of different serine proteases to highlight the structural diversity of the protease class. In total there are 13 clans and 51 families of serine proteases currently listed, according to the MEROPS database.[34] A summary of the different catalytic units in all serine protease clans is listed in Table 1.

Table 1: Illustration of all clans and families of serine proteases with their catalytic residues and characteristic protein fold.[35]

Clan	Family	Representative Member	Catalytic Residues	Fold
PA	14	Trypsin/Chymotrypsin	Asp-His-Ser	Greek-key β-barrels
SB	2	Subtilisin	Asp-His-Ser	3-layer sandwich
SC	7	Prolyl oligopeptidase	Asp-His-Ser	α/β
SE	3	D-Ala-D-Ala carboxypeptidase A	Lys-Ser	α-helical bundle
SF	2	LexA peptidase	His/Lys-Ser	all β
SH	5	Cytomegalovirus assemblin	His-His-Ser	α/β barrel
SJ	3	Lon peptidase	Ser-Lys	α/β
SK	3	Clp peptidase	Asp-His-Ser	α/β
SP	1	Nucleoporin	Lys-Ser	all β
SP	1	Aminopeptidase DmpA	His-Ser	4-layer sandwich
SR	1	Lactoferrin	Lys-Ser	3-layer sandwich
SS	1	L, D-Carboxypeptidase K	Glu-His-Ser	β-sheet + β-barrel
ST	8	Rhomboid	His-Ser	α-barrel

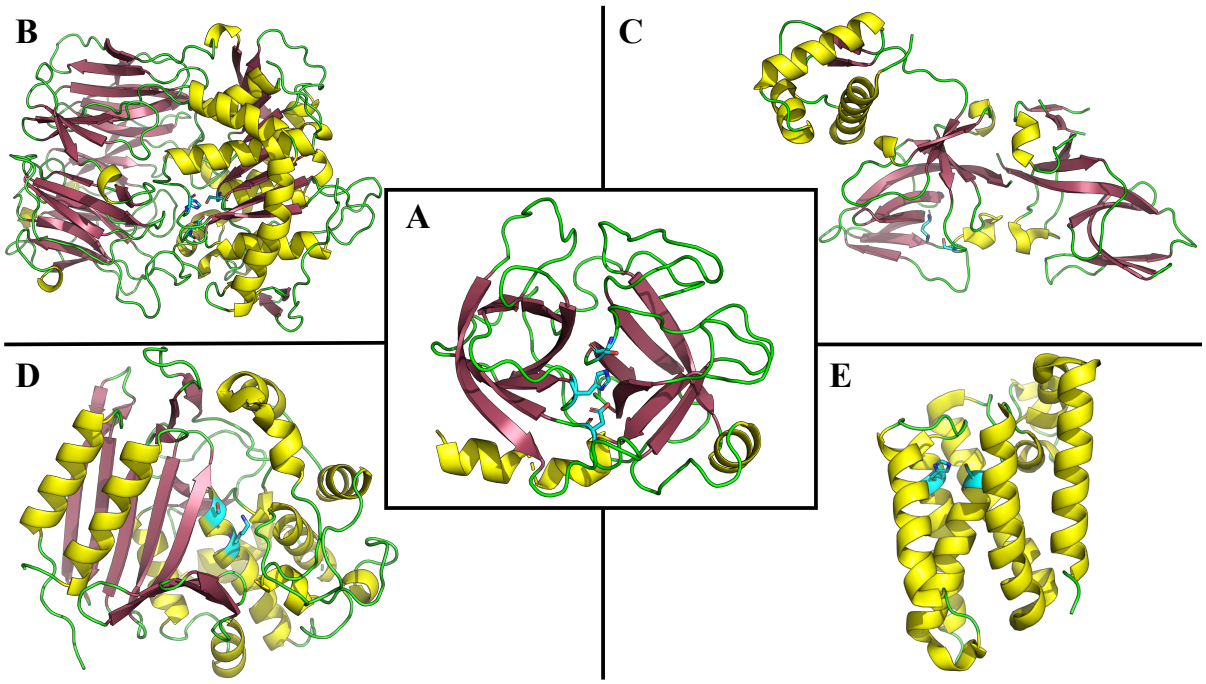


Figure 4: Overview of structurally diverse serine proteases from different clans, to highlight the various evolutionary paths of serine proteases. The tertiary structure of the proteases is shown, containing α-helices (yellow) and β-sheets (red). The catalytic residues are highlighted in sticks (cyan). (**A**) **Trypsin** (pdb: 1TRN) represents a prototypic example of serine proteases. The catalytic triad Asp-His-Ser is located between two similar β-barrels. (**B**) **Prolyl oligopeptidase** (pdb: 1QFS) presents a two-domain structure, with an *N*-terminal eight-bladed β-propeller and *C*-terminal α/β hydrolase domain, containing the Asp-His-Ser triad. (**C**) **LexxA repressor** (pdb: 1JHF) displays a scaffold that is mainly consisting of β-strands containing the catalytic dyad Lys-Ser. (**D**) **D-Ala-D-Ala-carboxypeptidase** (pdb: 3PTE) shows a catalytic dyad Lys-Ser on top of a α-helical bundle. (**E**) **Rhomboid protease** (pdb: 2IC8) presents an intramembrane protease with the dyad His-Ser located in a shallow surface within an α-barrel.[35] Figure created with PyMOL and PowerPoint.

Serine proteases are widely distributed in nature as well as found in most compartments of cellular life. But major differences in their appearance exist of each clan across species. For example, the PA clan proteases appears mostly in eukaryotes, but rarely in prokaryotic and plant genomes.[30],[36] Most serine proteases are endopeptidases, which catalyze the hydrolysis of an amide bond in the middle of a polypeptide. However, several exopeptidases were discovered, that can remove one or more amino acids from the termini of a peptide chain.[35],[37] This enables the proteases to cleave amino acids from large protein molecules, which expands the overall substrate-range and influence of physiological processes.[38]

Through evolutionary activities caused by gene duplication and divergence, four protease families cover around 40% of all proteolytic enzymes in human:[37] (1) the ubiquitin-specific proteases (cysteine proteases), which are responsible for the regulated protein turnover, (2) adamalysins (metalloproteases), controlling growth factors and integrin function, (3) prolyl oligopeptidases (serine proteases), cleaving mainly peptide hormones and neuropeptides, and (4) the trypsin-like serine proteases, which are

involved in many physiological mechanism like digestion, tissue remodeling, blood clotting etc.[39]–[46] Moreover, the trypsin-like protease family S1 presents the largest group of homologous proteases in human genome.[30]

1.2.2 Structural Biology and Catalytic Mechanism of Trypsin-Like Serine Proteases

The S1 family of the PA clan is composed of two subfamilies: S1A and S1B.[47] S1B proteases are found in all cellular life and regulate the intracellular protein turnover, while the S1A proteases are responsible for extracellular processes.[30],[35],[47] The S1A proteases are grouped into six functional categories for human: coagulation and immunity, tryptase, matriptase, kallikrein, granzymes, and digestion.[48]–[53] Trypsin, chymotrypsin and elastase belong to the S1A family and act as digestive enzymes.[54]–[56] They are endopeptidases and cleave polypeptides into shorter chains. Additional digestion can be mediated by different exopeptidases.[57],[58]

The trypsin-like architecture, here exemplarily demonstrated by trypsin, is characterized by two six-stranded β-barrels, where the β-strand topology presents a classic Greek-key architecture (Figure 5 A & C). The active site residues are placed in the cleft between the barrels. His57 and Asp102 are located on the *N*-terminal β-barrel, whereas Ser195 and the oxyanion hole (Ser195 and Gly193) originates from the *C*-terminal barrel.[59] Each barrel is partitioned with one end involved in catalysis, and the other in regulation (Figure 5 B). The entrance to the active site is decorated with eight surface exposed loops (Figure 5 C & D). Here, each β-barrel end donates three loops to the active site face, and two loops wrap around from the back of the protein.[30],[35],[47] Furthermore, the protein structure is stabilized by six conserved disulfide bridges against the reductive environment of the gastrointestinal tract.[60]

Proteolytic processing of inactivated zymogens is required for the activation of many trypsin-like serine proteases.[60] The initial cleavage of the proprotein precursor between residues 15 and 16 (trypsin numbering) occurs in most members of the S1A family. The following formation of an ion-pair between the *N*-terminus, which is originally unstructured, and Asp194 stabilizes the oxyanion hole and substrate-binding site. Simultaneously, residues near the active site change their flexible state in the zymogen to an ordered state in the active enzyme.[61],[62] Alternatively, the S1A serine protease tissue-type plasminogen activator (tPA) carries a Lys156, which builds an ion-pair with Asp194 and induces the conformational change without proteolytic cleavage at residue 15.[63]

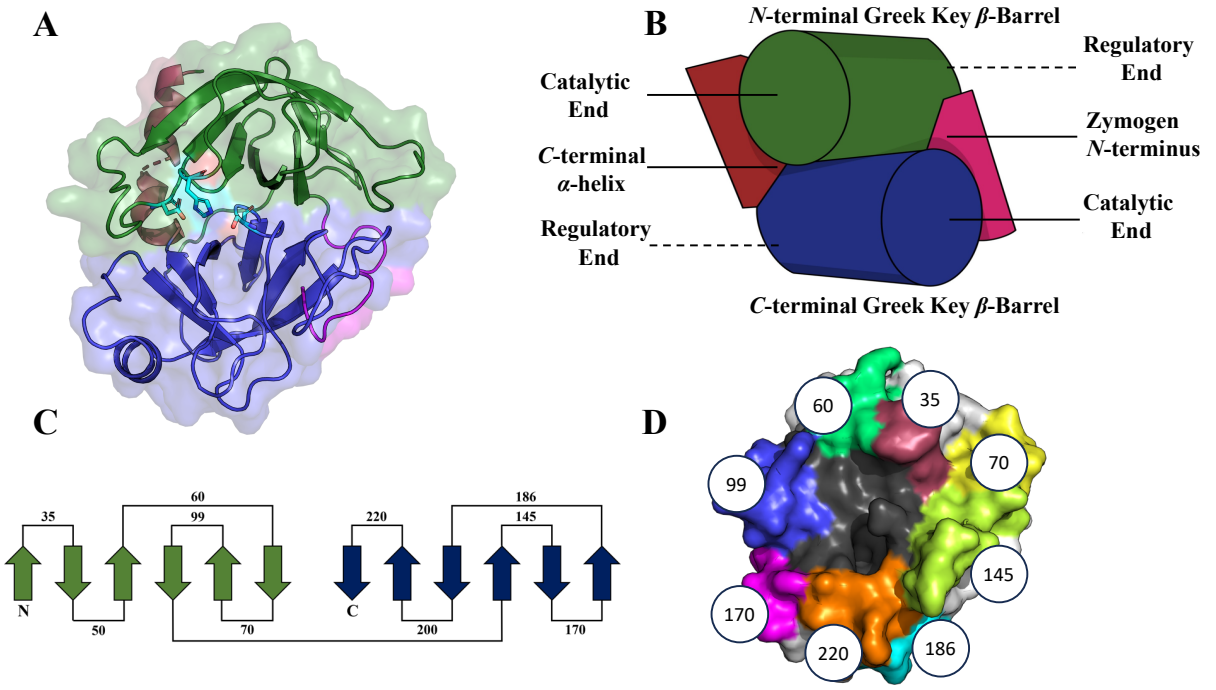


Figure 5: Overview of the trypsin fold. (**A**) X-ray crystal structure of trypsin (pdb:1TRN) showing the *N*-terminal *β*-barrel (green) and *C*-terminal *β*-barrel (blue) with the catalytic active site Asp-His-Ser (cyan) between them. The *C*-terminal *α*-helix (red) and zymogen-activation peptide (magenta) stabilize the overall structure. (**B**) Schematic representation of the trypsin fold. Catalytic and regulatory ends are partitioned and showing in opposing directions. (**C**) *β*-strand topology of trypsin displaying the structural similarity between both *β*-barrels and the Greek-key architecture. (**D**) Top view of trypsin with the catalytic active site in the middle (black), surrounded by eight loops. Numbers in (**C** & **D**) indicates the amino acid, which is included in the respective loop. Figure created with BioRender.com, PyMOL and PowerPoint. Modeled after Di cera *et al*..[35]

Although the characteristic protein fold of Clan PA proteases is presented in trypsin, chymotrypsin and elastase, they display different substrate specificities. Trypsin prefers basic amino acid side chains (Arg/Lys) at the P1 site of substrates, chymotrypsin large aromatic residues (Phe/Tyr/Trp) and elastase small aliphatic side chains (Ala).[35],[64],[65] The molecular reason for the differences is located in the S1 pocket of the proteases. The ca. 10 Å deep cavity harbors different amino acids at the bottom of the pocket. Trypsin displays a negatively charged Asp189, which allows formation of a salt bridge with Arg or Lys. Chymotrypsin harbors a Ser189 that eliminates the possibility of ionic interactions with the substrate and changes the favor to bulky aromatic residues and elastase´s specificity results due to the presence of Val216 and Thr226. Both residues block the S1 pocket for bulky structure elements from entering. In contrast, trypsin and chymotrypsin both harbor two glycine residues at that position.[30],[35],[47],[66]

The reaction mechanism of peptide hydrolysis by trypsin-like serine proteases with a catalytic triad (Asp-His-Ser) is executed in two steps. The hydrogen bond between the aspartate and histidine residue in the active site causes an increased electronegativity of the nitrogen atom (Figure 6 A). Therefore, histidine acts as a base, which is responsible for the deprotonation of the nucleophilic serine. The attack of the activated hydroxyl oxygen atom in the serine residue to the sp^2-hybridized carbonyl carbon atom of the amide bond results in the formation of a tetrahedral transition-state, which is stabilized by the

backbone nitrogen atoms of Gly and Ser (Figure 6 B). Both residues generate a positively charged pocket, called the oxyanion hole. The release of the *C*-terminal fragment takes place with donating the acquired proton from histidine to the leaving amine group (Figure 6 C). The second step starts with the deprotonation of a water molecule in the active site by histidine and following nucleophilic attack towards the acyl-enzyme intermediate (Figure 6 D). After the collapse of the second tetrahedral transition-state and the release of the *N*-terminal fragment, the return of the proton from histidine to the catalytic serine resets the enzyme for the next peptide bond hydrolysis cycle (Figure 6 E).[24],[30],[47]

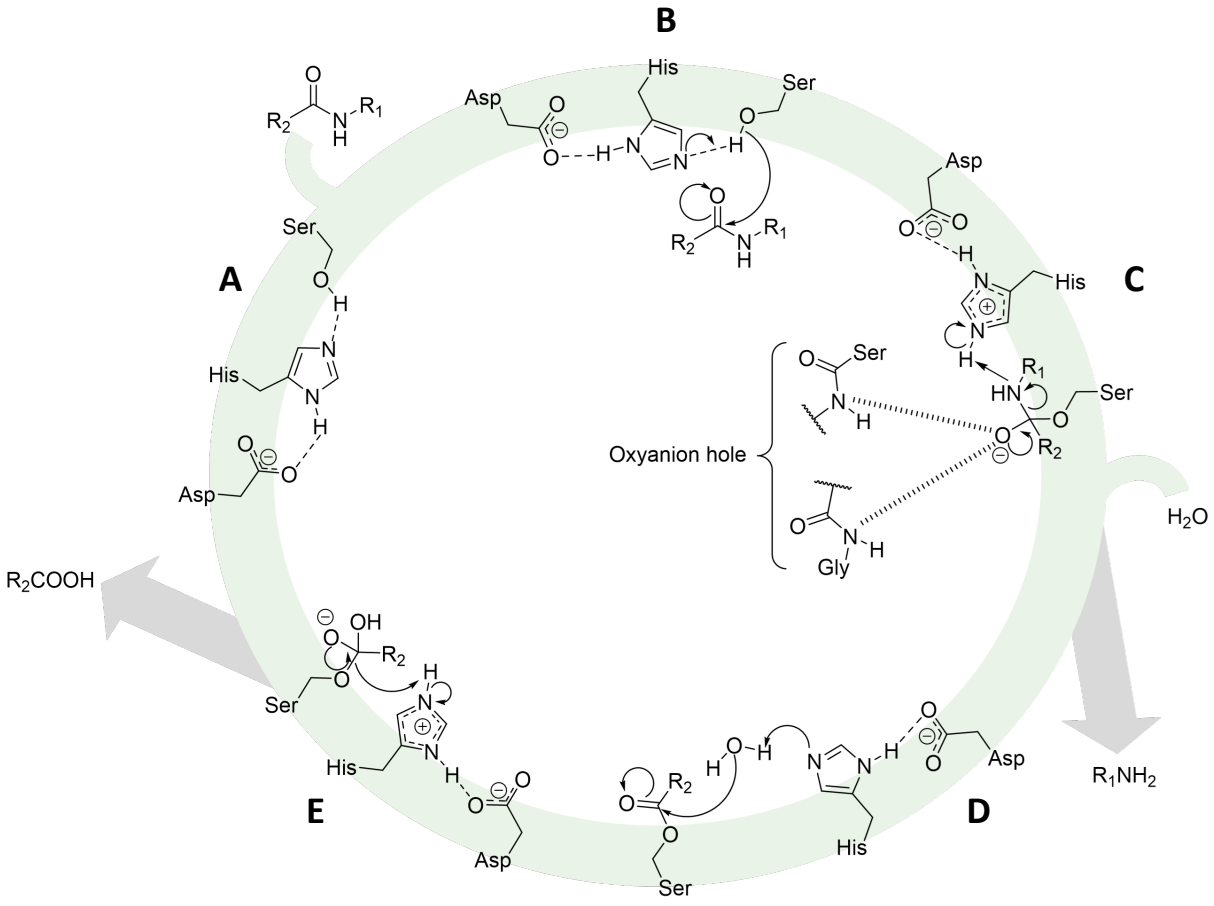


Figure 6: Catalytic reaction mechanism of trypsin-like serine proteases with the catalytic triad (Asp-His-Ser). (**A**) The charge relay system displaying the activation of the Ser residue by His and Asp. (**B**) Formation of the first tetrahedral transition-state by nucleophilic attack of Ser towards the amide bond. (**C**) Stabilization of the tetrahedral complex by the oxyanion hole and release of the *C*-terminal fragment. (**D**) Activation of one water molecule in the active site and following nucleophilic attack to the acyl-enzyme intermediate. (**E**) Collapse of the second tetrahedral transition-state by releasing the *N*-terminal fragment and recovery of the catalytic triad. Figure created with ChemDraw.

1.2.3 Physiological Roles of Serine Proteases

Out of 699 proteases in man, 178 are serine proteases and 138 of them belong to the S1 trypsin-like family.[30] Therefore, it's not surprising that this class of proteases is involved in pivotal physiological mechanisms.

The probably oldest known role of proteolytic enzymes presents the degradation of dietary proteins.[2] The protein catabolism displays the breakdown of proteins into monomers, which enables absorption for further degradation or reassembly.[67] Essential amino acids cannot be biosynthesized by humans but are important building blocks for the biosynthesis of vital proteins. Thus, their only source is polypeptide breakdown of dietary proteins through digestive enzymes. The exocrine pancreas secrets a variety of digestive enzymes, of which trypsin resembles the most known member. Trypsin is secreted in its zymogen form trypsinogen and gets activated either by enteropeptidase or other active trypsin proteins.[68] Additionally, trypsin continues to activate other proenzymes like chymotrypsinogen, proelastase and procarboxypeptidase.[69] Above all, trypsin hydrolyzes polypeptides at specific cleavage points (Arg/Lys, 1.2.2) into amino acids or oligopeptides, which can get further cleaved into amino acids by exopeptidases.[67] Then, active uptake mechanisms guide the amino acids into cells for biosynthesis of new proteins with vital intracellular functions.[70] They can also enter the citric acid cycle for energy generation or get modified and excreted mostly through the kidneys.[70],[71]

Another example of a crucial role of trypsin-like serine proteases can be found in the blood. The supply of nutrients and oxygen to the brain and muscles, which is dependent by an intact blood circulation, requires strategies to maintain this fragile system. Blood coagulation demonstrates an important physiological machinery and is always required by bleeding, which in the worst case could lead to death. The mechanism involves forming a tight mesh of cross-linked fibrin, which covers a platelet plug to stop the bleeding and stabilize blood circulation.[52] The coagulation cascade consists of multiple serine proteases that get activated one after the other, providing a biochemical amplification process.[3],[4] Upon damage of the vasculature, a tissue factor gets exposed and binds to the first serine protease coagulation factor VII.[72] The activation of this protease leads to cascade-like activation of several proteases. Key enzymes represent the trypsin-like serine proteases factor Xa and thrombin.[73] The latter is responsible for the cleavage of fibrinogen to fibrin, which gets cross-linked by the transglutaminase factor XIII and generates the abovementioned mesh to stabilize the platelet plug and stop bleeding.[74]

Protein catabolism and blood coagulation resembles just two important physiological roles of trypsin-like serine proteases. Further involvement of this protease class comprises cell proliferation and cell death, the immune response system and neural development, which can be researched in literature.[8]–[10],[75],[76]

1.2.4 Pathological Associations

On basis of the previously mentioned involvement of serine proteases in a variety of important physiological roles to enable life as we know it, it is not surprising that dysregulation of these enzymes could lead to severe pathological circumstances, represented for example by cancer. It is widely known that cancer arises from the transformation of normal cells into tumor cells in a multi-stage process and afterwards spreads via metastasis to other compartments of the human body.[77],[78] Dysregulated serine proteases have been associated with tumor development and metastasis for a variety of cancer-types.[79],[80] The trypsin-like serine proteases urokinase-type plasminogen activator (uPA) and kallikrein for example, play a crucial role in the remodeling of the extracellular matrix (ECM), angiogenesis and cell proliferation[81],[82]. Both proteases exert their proteolytic activity in the tumor microenvironment, mainly activating signaling networks and modulating the expression of genes to accelerate tumor growth and invasion. This includes pathways like the activation of other proteases like metalloproteases that promote the degradation of the ECM to enable metastasis, or activate signal pathways like the EGFR signal cascade, which result in cell proliferation and overall tumor progression.[83],[84]

Moreover, serine proteases are not only involved in pathological processes of cellular origin, but also a key element for the distribution of extrinsic pathological diseases like viruses. Two prominent examples are zika and dengue fever, which are relevant in tropical and subtropical countries. Spread by mosquitos, the zika and dengue virus can lead from mild flu-like symptoms such as fever to severe deadly symptoms with bleeding and cardiovascular collapse.[85] Dengue and zika viruses are small enveloped flaviviruses harboring a single-stranded (+) sense RNA genome, which get translated into a polyprotein after the viral cell entry.[86],[87] This polyprotein consists of the viral trypsin-like serine protease NS2B-NS3.[88] Co- and post translational processing by host proteases or the viral protease itself, leads to the formation of functional proteins.[89] The activated trypsin-like serine protease continues supporting the replication process of the virus by cleavage of the polyprotein and overall the spread of the virus infection in human.[88]

Due to the involvement of the trypsin-like serine proteases uPA, TMPRSS2 and hepsin within the work of this dissertation, their structures, physiological roles and pathological associations will be discussed in more detail later.

1.2.5 Serine Proteases as Clinical Therapeutics

Proteases are not only considered as key elements in important physiological or pathological pathways, but also used as potential therapeutics in medicine. Due to beneficial properties in recombinant production and a general high stability compared to other protease classes, serine proteases have attracted world-wide attention for biotechnological applications.[29],[90] The first approved protease-drug uPA by the Food and Drug Administration (FDA) in 1978 presented a new alternative therapeutical approach against thrombotic diseases, contrary to the conventional surgical removal of emboli.[91],[92] It

is used to dissolve blood clots in blood vessels, by proteolytic activation of the inactive zymogen plasminogen to plasmin. The increased plasmin activity leads to the degradation of fibrin clots to remove emboli.[93]

In contrast, thrombin presents another serine protease, which is a pivotal component of the coagulation cascade and responsible for the formation of stable blood clots by converting fibrinogen into fibrin monomers and initiating the following multimerization (1.2.3).[94] Since 2008, human thrombin, sometimes combined with fibrin or collagen and formulated into bandages, is used to treat bleeding from surgery or trauma.[95] Both examples highlight the potential therapeutic use of protease-drugs and present one additional opportunity for treatment of diseases in medicine.

1.3 Serine Protease Inhibitors

1.3.1 Endogenous Serine Protease Inhibitors

Because of the crucial physiological roles (1.2.3) and the accompanying severe diseases (1.2.4) by dysregulation, the proteolytic activity of serine proteases underlies a strict regimen by endogenous inhibitors to maintain the homeostatic equilibrium. Serpins (serine protease inhibitors) are the largest superfamily of endogenous protease inhibitors.[96],[97] Many disease-related proteases are regulated by serpins, e.g. α-1-antitrypsin inhibits trypsin to prevent pancreatitis due to chronic tissue breakdown or antithrombin inhibits thrombin to control coagulation.[56],[98] Additionally, it is often the case that one serpin can modulate multiple proteases, like α-1-antitrypsin inhibits trypsin, neutrophil elastase, tryptase etc.[99] Serpins inhibit the target serine protease in an irreversible manner. The active endogenous inhibitor inserts their reactive center loop into the catalytic binding site of the protease. The loop gets cleaved and binds to the nucleophilic serine residue of the enzyme. The covalent serpin-protease complex is stabilized by major conformational change of both proteins.[100]

1.3.2 Clinical Applications of Serine Protease Inhibitors

Recombinant endogenous inhibitors are of considerable interest for clinical applications, due to the severe pathology by deficiencies. For example, recombinant α-1-antitrypsin is approved by the FDA for delaying the progression of emphysema, which originates in the deficiency of α-1-antitrypsin in lungs.[101],[102] Another class of biological therapeutic inhibitors presents the Kunitz-based inhibitors and include for example the pancreatic trypsin inhibitor aprotinin, which display the typical irreversible inhibition mechanism of endogenous inhibitors (1.3.1).[103],[104] One additional famous member resembles ecallantide. The recombinant biological therapeutic is a kallikrein inhibitor that is FDA approved for treatment of hereditary angioedema, which is caused by a serpin inhibitor deficiency.[105] Nevertheless, due to the high molecular weight and presence of hydrophilic and hydrophobic structure elements, poor permeability and lacking oral bioavailability as well as vulnerability against protein degradation and excretion highlight general challenges for the use of recombinant endogenous inhibitors.[106]

The most frequently employed strategy to target serine proteases, and proteolytic enzymes in general, involves the synthesis of substrate mimics, that compete with the natural protease ligand in the active site of the enzyme. In the 1950s, a patent of a Japanese company mentioned the very first protease inhibitor. ε-Aminocaproic acid was described as a “compound for combatting diseases associated with plasmin activation and other proteolytic action”.[107],[108] Around 1960, *in vivo* and *in vitro* studies revealed positive therapeutic effects of the abovementioned compound to allergic diseases and bleeding caused by an increase of fibrinolytic activity, with almost no toxicity.[107],[109] Because of the promising clinical use, ε-aminocaproic acid was approved by the FDA as the first protease inhibitor in 1964 for perioperative hemostasis, even without knowing the exact mechanism of action.[110] Only seven years

later, it was discovered that ε-aminocaproic acid binds to the inactivated zymogen plasminogen and therefore prevents catalytic activation through other proteases.[111]

Compared with common drug development nowadays, the described fashion for discovery of ε-aminocaproic acid as a protease inhibitor by lucky coincidences remains exotic and represents an exception. Within time, principles and guidelines have been developed about "drug-like" properties and were expanded through new discoveries of protein-ligand interactions.[110],[112]–[114] New technologies and methods regarding structure elucidation, protein expression and the resulting improved characterization of the target proteins and their inhibitors built the foundation for the rational design of protease inhibitors.[110]

Beginning with ε-aminocaproic acid as the first (serine) protease inhibitor, between 1964 and 2023, at least 64 "small-molecule" protease targeting drugs with molecular weights ranging from 200 to 600 Da have been approved by the FDA.[110] The largest part of the approved drugs accounts for serine protease inhibitors with 29 representatives (Figure 7).

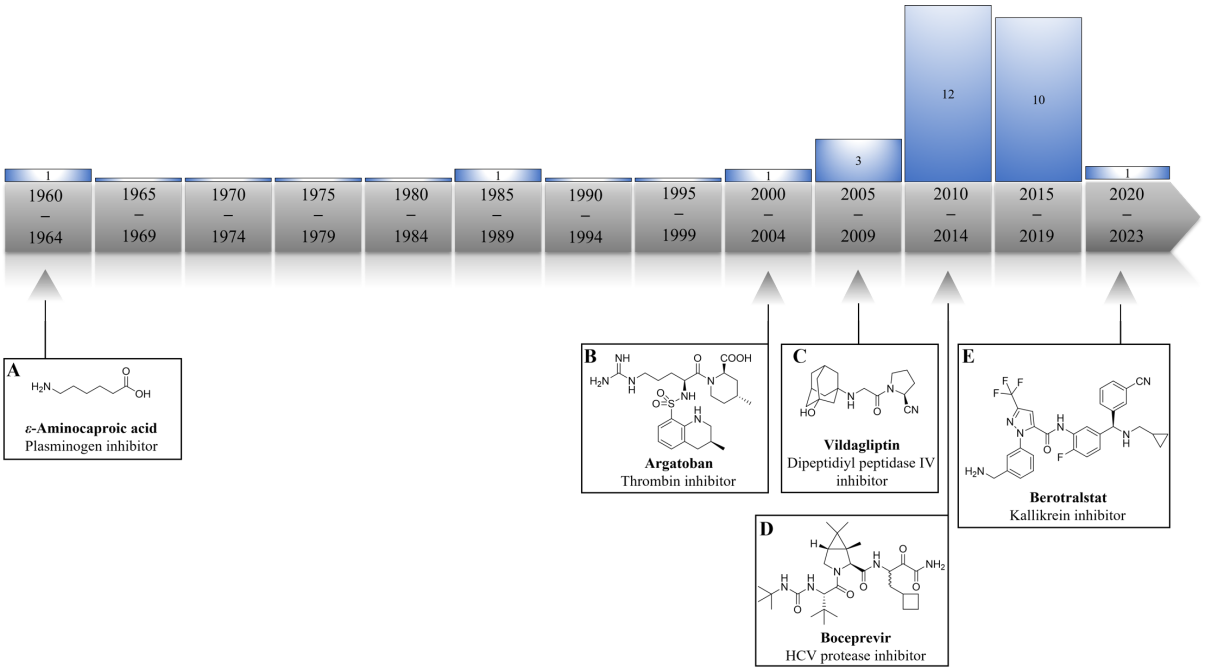


Figure 7: Timeline of all approved "small molecule" serine protease inhibitors by the FDA from 1964 until 2023. Examples of approved inhibitors are shown, which represent important derivates of different drug classes against their respective target. (**A**) ε-Aminocaproic acid as the first approved serine protease inhibitor by preventing the activation of plasminogen. (**B**) Argatoban as the first competitive serine protease inhibitor, targeting thrombin. (**C**) Vildagliptin represents the drug class of gliptines as oral hypoglycemics, targeting the dipeptidyl peptidase IV. (**D**) Boceprevir as the first hepatitis C virus inhibitor towards the viral serine protease. (**E**) Berotralstat as the latest approved serine protease inhibitor, targeting kallikrein for treatment of hereditary angioedema.[110] Figure created with PowerPoint and ChemDraw.

Besides ε-aminocaproic acid (Figure 7 A) and the lysine-analogue tranexamic acid, which belongs to the class of ε-amino acids and displays the same mechanism of action, all approved serine protease inhibitors can be assigned as competitive inhibitors, and compete with the natural substrate in the active site.[115] Argatoban (Figure 7 B) was the first FDA approved competitive serine protease inhibitor in 2000, which targets the trypsin-like serine protease thrombin.[116],[117] This was followed by five

approved direct oral anticoagulants, targeting thrombin or factor Xa, becoming the gold standard for treatment of thrombosis.[118]–[121] Besides the development of antithrombotic agents, the research field of oral hypoglycemics, so called gliptins represents another large drug class. The designed inhibitors block the serine protease dipeptidyl peptidase IV and are used for treatment of type 2 diabetes mellitus.[122] Starting with sitagliptin and vildagliptin (Figure 7 C) as the first representatives, nine more derivates were approved between 2006 and 2023.[39],[122]–[128] Another important drug class was the development of hepatitis C virus (HCV) inhibitors. In total nine HCV inhibitors were approved by the FDA, started with boceprevir in 2011 (Figure 7 D), which target the viral NS3 serine protease.[129],[130] The most recently approved serine protease inhibitor, represents berotralstat (Figure 7 E), which inhibits kallikrein and is used to prevent attacks of hereditary angioedema.[131]

Based on the presented FDA approved serine proteases it becomes apparent that due to the growing knowledge of protein-ligand interactions, rational design of protease inhibitors increased the tendency to the development of synthetic peptidomimetics as potent drugs (Figure 7).[40]–[43],[132] This inhibitor class combines simple peptides, which imitate the natural substrate of the targeted proteases, with nonpeptide structure elements. Peptidic drugs are often characterized by short half-times, due to proteolytic degradation, lack of oral bioavailability, rapid metabolism and following excretion. However, these limitations and other pharmacokinetic issues can get bypassed by introduction of non-peptide structure elements.[44]

This comprises for example the replacing of natural amino acids with unnatural derivates, side-chain substitutions or backbone modification to increase the metabolic stability.[45],[133],[134] Macrocyclization is another proven strategy to overcome proteolytic degradation and is often used to enhance the membrane permeability, due to hiding of polar structure elements by intramolecular interactions.[135] Covering polar groups is not only possible through conformational change induced by cyclisation, but through reversible drug derivatization. The prodrug approach describes the masking of structure elements, to overcome pharmacokinetic barriers like low oral drug absorption. For example, the thrombin inhibitor dabigatran showed very poor intestinal absorption after oral administration. In response, the highly basic benzamidine group was converted to the *n*-hexyl carbamate ester and the carboxylic acid into the ethyl ester moiety. This resulted in an improvement of the oral bioavailability and after enzymatic cleavage of the prodrug to an overall increase of dabigatran plasma concentrations.[136]

1.3.3 Targeted Covalent Inhibitors

In contrast to the traditional inhibition mechanism that displays the non-covalent reversible interaction between the inhibitor and the target enzyme, covalent inhibitors are compounds that by design are intended to form a covalent bond with their biological target and thereby suppress its biological function. This class of drugs are called “Targeted Covalent Inhibitor” (TCI).[137],[138]

The process of inhibiting their biological protein relies upon two steps. First, the inhibitor gets near to its target binding site and forms a non-covalent protein-inhibitor complex. This brings the electrophilic moiety of the inhibitor (warhead) in close proximity to a nucleophilic residue of the protein. The second step contains the spontaneous reaction between the functional groups of the inhibitor and protein, which results in a covalent bond and inactivation of the biological target. The covalent bond formation can be either reversible or irreversible, depending on the used warhead.[139]

Bortezomib was the first FDA approved protease inhibitor with the boronic acid moiety as warhead, targeting the catalytic threonine of the 26S proteasome for treatment against multiple myeloma.[140] The approval was a significant milestone for the development of covalent protease inhibitors. Previously, covalent inhibitors were avoided by the pharmaceutical industry, justified by high risks in terms of toxicity and side effects. The skepticism appears to stem largely from the possible activation of the immune system through haptenization of irreversibly covalently inhibited proteins and the random, covalent binding of highly reactive drugs to cellular macromolecules, that can result in tissue injury.[139],[141] Additionally, the introduced electrophilic trap can react with different physiologically important nucleophiles in the human body by insufficient selectivity, for example with glutathione.[142] Nevertheless, covalent inhibitors provide several advantages for clinical use over traditional non-covalent drugs. The improved efficacy, due to the prolonged inhibition time of proteins leads to lower dosage of drugs, which on the one hand minimizes off-target reactions and side effects and on the other hand reduces drug resistance mechanism, caused for example by efflux pumps.[143]–[145]

Since the approval of bortezomib, it has been recognized, that the rational design of covalent protease inhibitors, combining a peptidomimetic scaffold with a suitable warhead moiety, results in potent drug candidates with clinical use. The peptidomimetic structure element presents enhanced pharmacokinetic properties (stability, permeability etc.) and selectivity towards other proteases, while the warhead is responsible for the affinity to the target.

In this context, the aim of this dissertation is the investigation and characterization of both structure elements (peptidomimetic recognition sequence & warhead) and possible combinatorial effects, regarding their impact on the pharmacodynamic properties of protease inhibitors. The results are shown in the presentation of the scientific article in chapter A (5.1).

In 2009, the first covalent serine protease inhibitor vildagliptin was approved as an antidiabetic drug (Figure 7 C). It displays the nitrile moiety as the warhead and a peptidomimetic scaffold as the recognition sequence. Overall, 10 covalent serine protease inhibitors were approved by the FDA between 2009 and 2023, containing either a nitrile or a α-ketoamide moiety as the electrophilic trap.[146],[147]

Generally, covalent inhibitors can be classified into two major categories depending on the inhibition mechanism. Covalent reversible inhibitors are characterized by finite inhibition of the target protein, meaning the covalent bond breaks after a while and the enzyme regains its biological functionality. Examples of this inhibitory mechanism can be found in the HCV protease inhibitors and gliptins bearing the covalent reversible nitrile or α-ketoamide warheads. Further examples are aldehydes, boronic acids and α-ketoheterocycles that are usually used for serine and threonine protease inhibitors.[148]–[151] The nitrile, carbonyl derivates and boronic acid all provoke the reversible nucleophilic attack of the activated serine residue to the electrophilic carbon/boron atom, resulting in the respective imidate (Figure 8 A), hemiacetal/hemiketal (Figure 8 B) or boronate ester (Figure 8 C).[147]–[152]

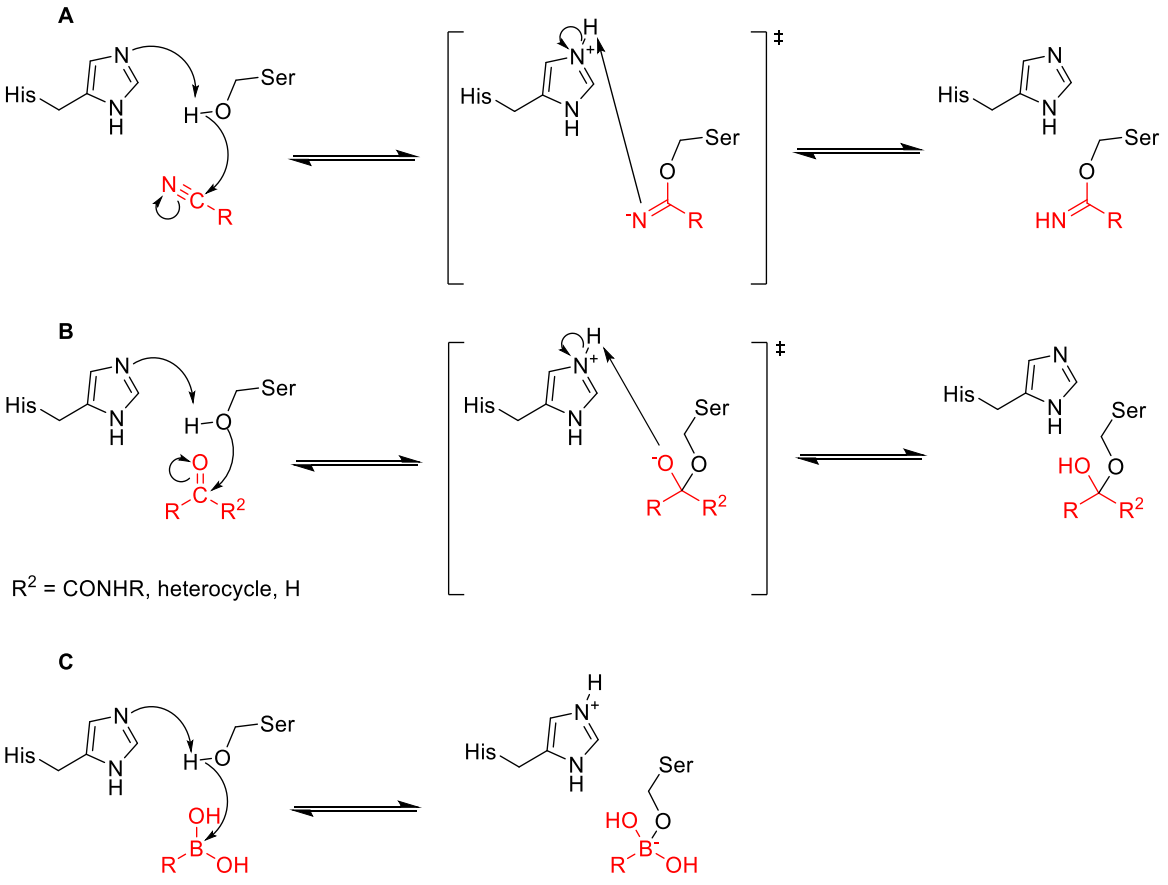


Figure 8: Overview of generally used covalent reversible warheads (red) with their mechanism of action. Reaction mechanism displayed using the example of a His-Ser dyad. (**A**) Nitrile inhibitors with their respective covalent reversible inhibition mechanism by formation of the imidate adduct. (**B**) Carbonyl-based inhibitors attacked by the activated serine residue at the electrophilic carbonyl carbon atom, leading to the reversible reacting hemiacetal/hemiketal. (**C**) Boronic acid as the electrophilic trap, which leads to the formation of a reversible boronate ester. Figure created with ChemDraw.

In contrast, the covalent adduct between an irreversible inhibitor and its target is permanent. This comprises for example electrophilic functionalities like α-halomethylketones and phosphonates, which are mainly developed for serine and threonine proteases.[153],[154] After the nucleophilic attack of the activated serine residue to the carbonyl carbon atom of the α-halomethylketone, there are two possible ways leading to the irreversibly inhibited enzyme. Either intramolecular cyclisation leads to the formation of an epoxid, which gets attacked by the histidine residue resulting in the irreversible product (Figure 9 A/1) or by direct substitution of the halogenide by histidine (Figure 9 A/2).[153] Cysteine proteases also get irreversibly inhibited by α-halomethylketones. However, the covalent bond between protease and inhibitor occurs with the nucleophilic cysteine residue, forming a thioether, instead of the histidine in the active site. Similar to the inhibition of the serine proteases, there are two possible mechanism, resulting to the thioether formation. First one is the direct substitution of the halide group by the thiol/thiolate anion (Figure 9 B/1). The second mechanism involves the formation of a thiohemiketal and a three-membered sulfonium intermediate, which presents the thioether adduct after rearrangement (Figure 9 B/2).[155] The phosphonate warhead shows a similar reaction mechanism to the boronic acid, but inhibits the enzyme irreversible, due to the strong P-O bond by ester cleavage (Figure 9 C).[156]

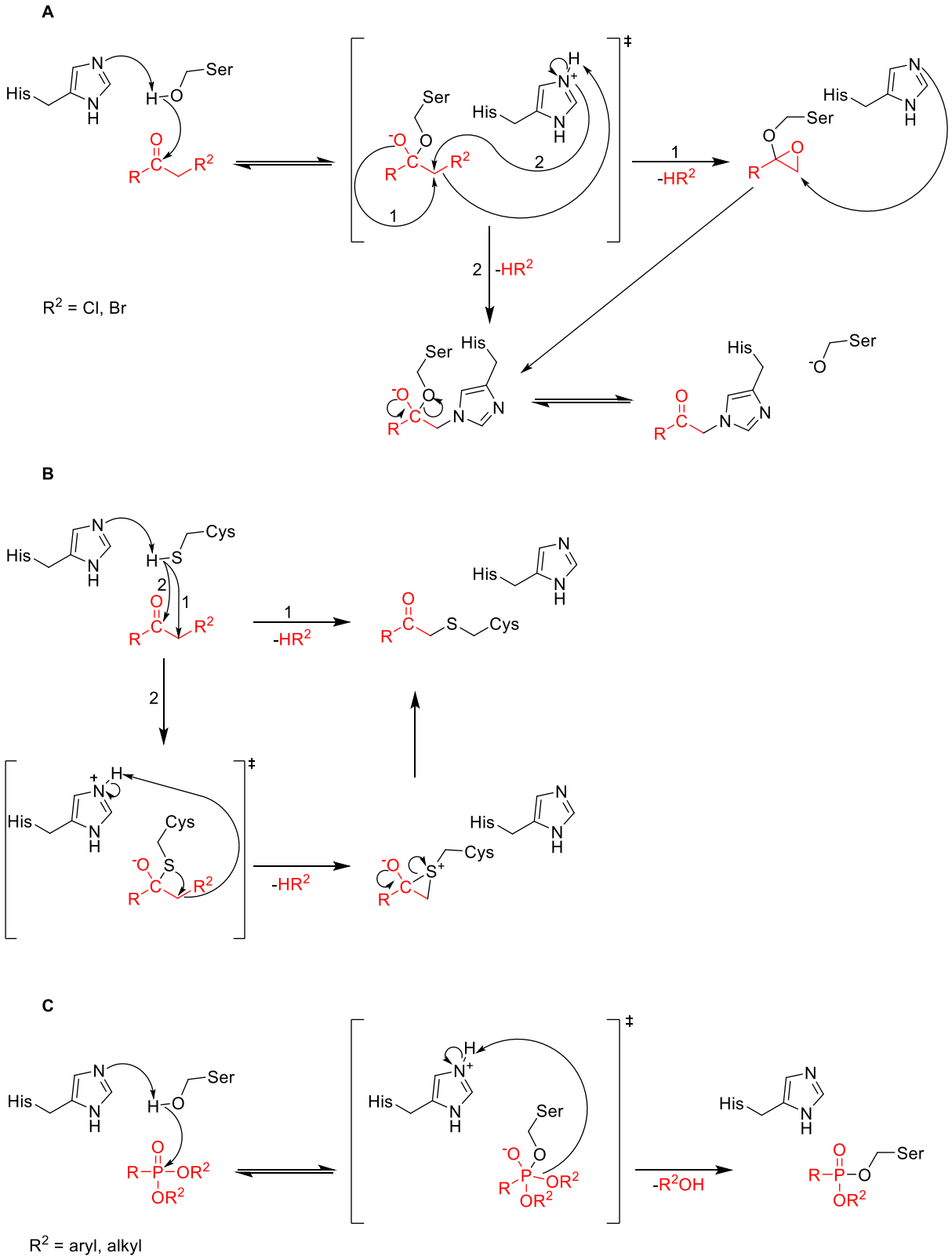


Figure 9: Overview of generally used covalent irreversible warheads (red) with their mechanism of action. Reaction mechanism displayed using the example of a His-Ser dyad. (**A**) Irreversible inhibition mechanism of α-halomethylketones displaying two possible paths for formation of the irreversibly inhibited enzyme. (**B**) Irreversible inhibition mechanism of cysteine proteases by α-halomethylketones displaying two possible routes for formation of the irreversibly inhibited enzyme. (**C**) Phosphonate-based inhibitors showing similar reaction mechanism to the boronic acid warhead but leading to an irreversible phosphonate ester derivate.[155] Figure created with ChemDraw.

Regarding the latest approved covalent serine protease inhibitors berotralstat (nitrile) and narlaprevir (α-ketoamide), it becomes clear that the development of this inhibitor-type is preferred over irreversible analogues, primarily, due to the abovementioned possible side effects of haptenization.[139],[141] Additionally, the increasing amount of approved covalent protease inhibitors clearly shows that the initial concerns regarding toxicity issues has been set aside by rational design of selective inhibitors and the concept of covalent drugs represents a promising possibility for developing potent drugs in medicinal chemistry.

In the following, the trypsin-like serine proteases urokinase-type plasminogen activator (uPA), transmembrane serine protease subtype 2 (TMPRSS2) and hepsin with their current state of inhibitor development will be presented. During the course of this work, new covalent-reversible inhibitors were designed and tested targeting the three abovementioned serine proteases with enhanced pharmacodynamic properties. The results are displayed in the presentation of the respective scientific articles in chapter B (5.1, 5.2, 5.3).

1.4 Urokinase-Type Plasminogen Activator (uPA)

The ubiquitous expressed serine protease uPA, together with its cell membrane-anchored uPA receptor (uPAR), forms the urokinase-type plasminogen activating system (uPAS). Along with the endogenous plasminogen activator inhibitors 1 & 2 (PAI-1, PAI-2), the complex is involved in numerous physiological and pathological processes, which comprise cell proliferation, adhesion and migration, ECM remodeling as well as tumor progression and metastasis.[83],[157]–[160]

1.4.1 Physiological Background of the uPA System (uPAS)

The urokinase-type plasminogen activator is secreted as a single chain zymogen (pro-uPA), consisting of a 53 kDa multidomain of 411 residues.[161] Once released into the extracellular environment, the zymogen gets cleaved between Lys158 and Ile159 by plasmin, thermolysin, trypsin or kallikrein for converting the proenzyme to the active disulfide bridge-linked high molecular weight uPA (HMW-uPA).[162] Alternatively, the zymogen can be processed at different positions by thrombin and elastase to enzymatically inactive HMW-uPA.[163] The active two-chain uPA is structured into the *N*-terminal A-chain, which consists of the kringle domain (KD) and the growth factor-like domain (GFD), and the *C*-terminal B-chain (Figure 10).[164] Studies have shown that the kringle domain is involved in chemotaxis, where the mechanism is not yet fully understood, while the GFD domain is responsible for the binding of the uPA to the uPAR.[165] The *C*-terminal B-chain harbors the catalytic serine protease domain.[164] Further processing of the two-chain uPA results in forming a shorter A-chain, the inactive amino-terminal fragment (ATF), and the low-molecular weight uPA (LMW-uPA, 33 kDa), which is still enzymatically active, but lost the possibility to bind to the uPAR.[164]

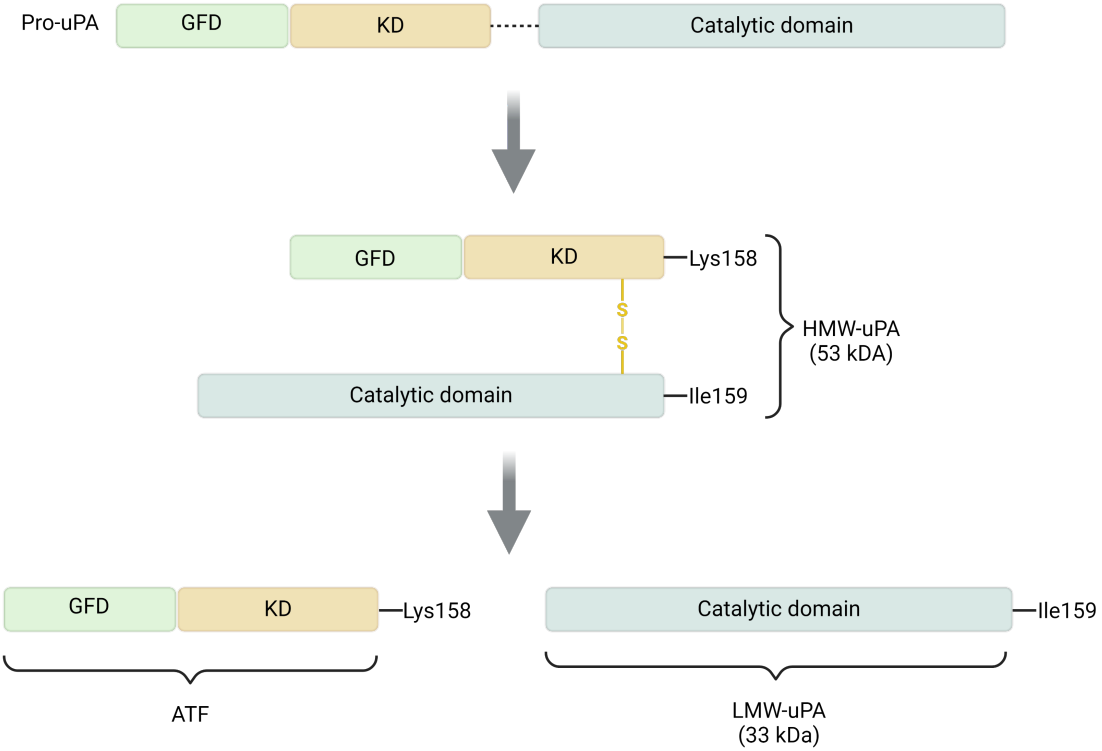


Figure 10: Proteolytic activation of the zymogen pro-uPA containing the *N*-terminal growth-factor domain (GFD), kringle domain (KD) and the *C*-terminal catalytic serine protease domain. After the first cleavage between Lys158 and Ile159 the HMW-uPA is formed and stabilized by a disulfide bond. Further processing leads to the LMW-uPA and amino-terminal fragment (ATF). S-S represents the disulfide bond. Figure created with BioRender.com and modeled after Mahmood *et al*..[164]

The uPA receptor is a 60 kDa glycoprotein consisting of three distinct domains (D1, D2 and D3) that are connect by flexible linker sequences.[166] During the maturation process, a glycosyl phosphatidylinositol (GPI) anchor is added at the *C*-terminus, which is responsible for the covalent connection between the uPAR and the cell membrane.[167] The receptor has no transmembrane nor cytoplasmatic domains. The binding of the pro-uPA or active enzyme involves all three domains of the receptor.[157] The role of the uPAR is not limited to uPA binding/activation. The receptor can also interact with different ligands like vitronectin or other transmembrane receptors, for example integrins to facilitate downstream cell signaling and enhance cell migration, survival and proliferation (Figure 11).[168]–[171] Due to the involvement of important physiological processes besides the binding and activation of the uPA, the uPAR also presents a valid target for cancer treatment, but will not be further discussed since the focus of this work is aimed at the uPA.[172],[173]

The main physiological process of uPA is the proteolytic activation of the zymogen plasminogen to its active form plasmin. Plasmin itself also is a trypsin-like serine protease which is directly involved in the degradation of the ECM by breakdown of different components like collagen, laminin, fibrin and fibronectin.[164],[174] Furthermore, plasmin can activate other proteases, like matrix metalloproteases (MMPs), that also catalyze the ECM degradation and cause the release of growth factors (Figure 11).

The *Fibroblast Growth Factor* 2 (FGF-2) and the *Hepatocyte Growth Factor* (HGF) can stimulate the expression of uPAS compartments, especially of the uPAR.[168],[175] In addition, plasmin can also convert pro-uPA to its active form, thus enhancing its own production. This phenomenon is termed “reciprocal zymogen activation”.[160],[176]

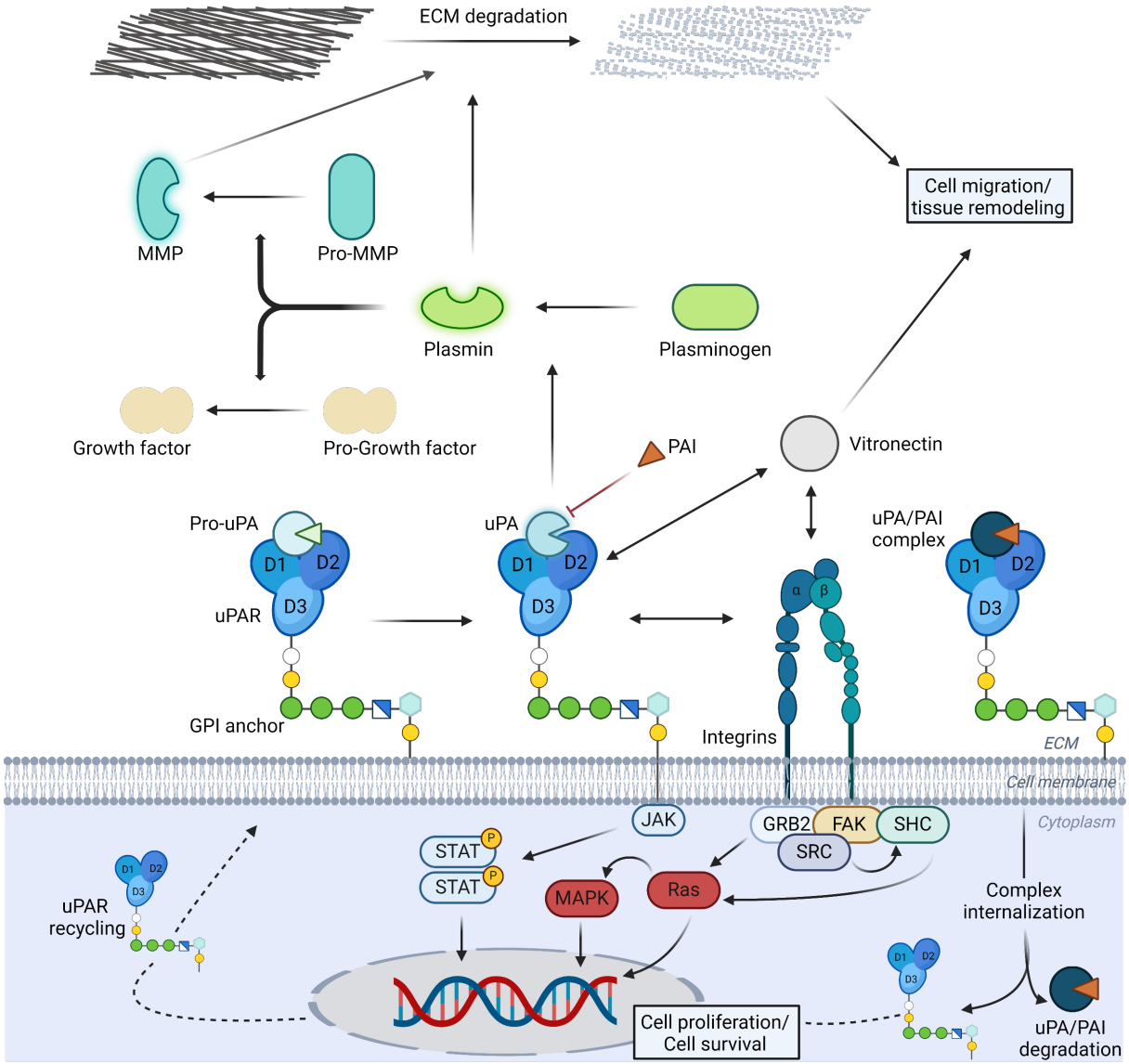


Figure 11:Schematic overview of the urokinase-type plasminogen activating system (uPAS) and mediated pathways. The GPI-anchored uPAR consisting of three domains (D1, D2, D3) binds the zymogen pro-uPA as well as the active enzyme. The activated serine protease converts plasminogen to plasmin, which induces ECM degradation either by itself or by activation of MMPs, that also lead to the release of growth factors. This results overall in cell migration and tissue remodeling. Furthermore, uPAR also binds to vitronectin, which enhances cell migration. Interaction of the uPAR with integrins activates different intracellular pathways, like JAK-STAT, focal adhesion kinase (FAK) and Ras/MAPK, which regulates cell proliferation and survival. Inhibited uPA by the endogenous inhibitor PAI leads to internalization of the uPA/uPAR, while the uPA/PAI complex gets degraded and the uPAR recycled (bottom right corner, 1.4.4). Figure created with BioRender.com and modulated after Mahmood.[164]

The uPA system is also involved in the regulation of the fibrinolytic system, due to the possible degradation of fibrin networks by plasmin.[177] This enhanced the intravascular blood circulation and highlights the uPA as an alternative therapeutic approach against thrombotic diseases (1.2.5)

1.4.2 Pathological Association of the uPA System

The uPA system presents a pivotal point for different physiological mechanism, which comprises ECM remodeling, wound healing and different cellular events like cell proliferation, adhesion and motility.[168],[177] However, these mechanisms can accelerate neoplastic evolution by dysregulation and facilitate metastasis progress immensely in a variety of solid tumors, including breast, ovarian, gastric, pancreatic, colorectal, liver cancer etc.[178]–[181] During cancer progression and metastasis malignant cells need to migrate through the ECM towards blood vessels for intravasation, extravasation and colonialization of target metastatic tissue. Studies have shown that the inhibition of uPA with antibodies reduced this mechanism. This proved the importance of the ECM degradation by the protease itself and further activated enzymes, especially plasmin, during the metastasis process (Figure 12).[182],[183] Before the invasion of peripheric tissue, the cancer cells need to detach from the primary tumor.[161] The release of growth factors (FGF-2, HGF) through the degradation of the ECM and the interaction of integrins and vitronectin with the uPAR accelerate cell migration and cell proliferation, which in turn improve tumor progression.[83],[161] Furthermore, growth factors stimulate angiogenesis, which provides the tumor tissue with oxygen and nutrients, and therefore supports tumor growth.[184]

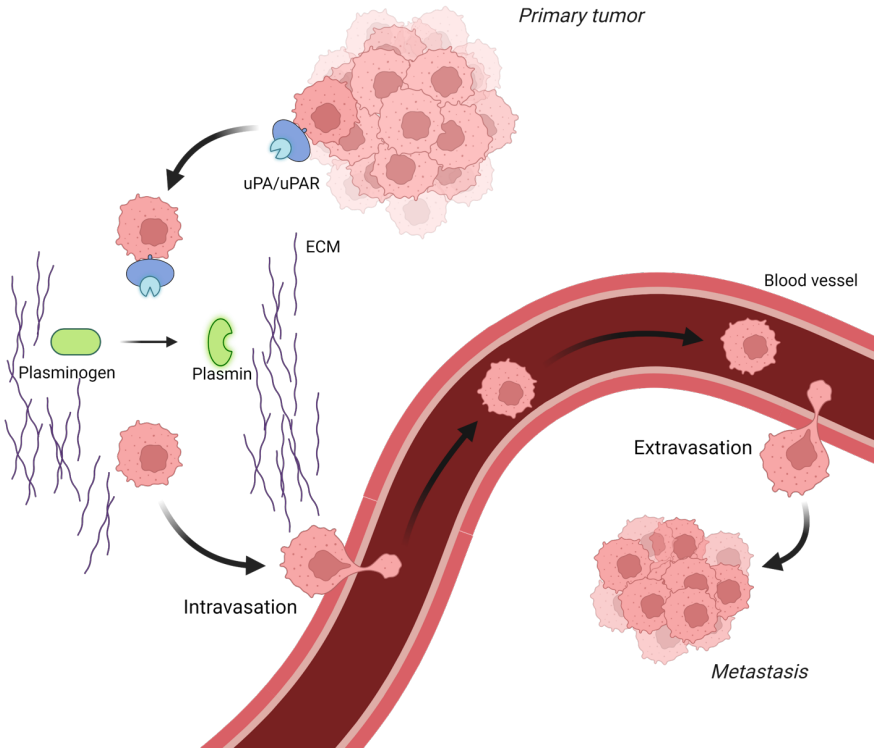


Figure 12: Involvement of the uPA/uPAR in tumor metastasis by activating plasminogen to plasmin. This leads to ECM degradation and acceleration of tumor progression through facilitated metastasis. Figure created with BioRender.com.

1.4.3 Structure of uPA

According to the MEROPS database, the urokinase-type plasminogen activator belongs to the trypsin-like serine proteases (EC 3.4.21.73, clan PA, family S1) harboring the catalytic triad Asp102-His57-Ser195. Thus, the uPA displays a high structure similarity to the characteristic protein fold of clan PA proteases with the corresponding catalytic mechanism (1.2.2). Typical for trypsin-like proteases, uPA consists of two opposed six-stranded β-barrels. The active site residues are located at the junction of the two barrels containing the recognition subsites S1–S4 with the catalytic triad (Figure 13, left).[180] The substrate binding groove contains a deep and well-defined S1 pocket, with a negatively charged Asp189 residue, responsible for the basic substrate specificity (Figure 13, right).[34],[179] The pocket is defined by two surrounding β-strands, consisting on one side of Gly216–Gly219 and Gly226–Tyr228 and on the other side of Ser190–Cys191–Gln192. One important difference over structurally related serine proteases displays Ser190 in the S1 pocket, which is sometimes exchanged by Ala190 (tPA, thombin) or Thr190 (kallikrein).[185],[186] It has been shown that a highly conserved water molecule at the S1 site can be displaced by inhibitors to enable polar interactions to Ser190, resulting in a stronger binding. In contrast, proteases with an Ala190 residue cannot be addressed in this manner. Furthermore, the S1 pocket is deeper and broader than other trypsin-like serine proteases. The S2 pocket is presented by a small depression on the surface of the uPA and characterized by His57, His99, and Trp215, allowing only small ligand residues and offers possible aromatic stacking interactions. The S3 pocket is on top of Gly216 and Trp215. Additionally, a new binding site S1β was identified from structure analysis, which is formed by the residues Lys143, Ser146, Gln192 and Gly218. Another characteristic structure element presents the loop region Asp97-Thr97a-Leu-97b that forms the S4 pocket in uPA. The “97” loop can adopt different conformations, depending on the bounded ligand structure.[180],[187]

Together, the overall bigger S1 pocket containing the displaceable conserved water molecule for enabling additional hydrogen bonding with Ser190, plus the unique subpocket S1β and the specific subpockets S2–S4 highlight possibilities to design potent and selective inhibitors targeting the uPA, which will be discussed in the following section.

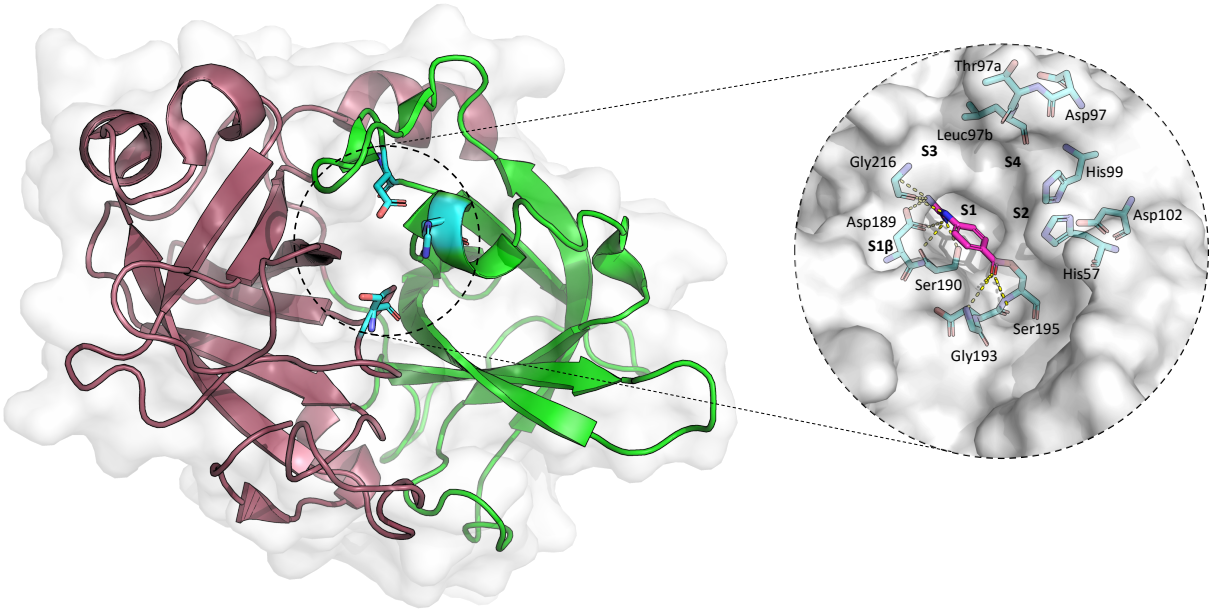


Figure 13: X-ray crystal structure of uPA (pdb: 7VM7) showing the characteristic trypsin-like fold with two opposed six-stranded *β*-barrels (green and red), harboring in the middle the catalytic active site (cyan). On the right side, visualization of the binding site (co-crystallized with nafamostat, magenta) formed by the **catalytic triad** (Asp102-His57-Ser195), deep **S1** pocket (Asp189, Ser190, Gly216), **S1β**, **S2** (His57, His99), **S3** (Gly216) and **S4** (Asp97-Thr97a-Leu-97b).[179] Some residues are not shown for clarity and are described in 1.4.3. Yellow dashed line presents polar interactions between the inhibitor and the residues of the binding site. Figure created with PyMOL and PowerPoint.

1.4.4 Inhibitors of uPA

The regulation of the proteolytic activity of uPA plays an important role, due to the abovementioned physiological influences regarding tissue remodeling, cell proliferation, adhesion and motility, and possible pathological association by dysregulation (1.4.1 & 1.4.2).

The activity of the uPA system is physiologically regulated by the serpin (serine protease inhibitor) family members.[188]–[190] The inhibitors PAI-1 and PAI-2 inhibit specifically the uPA, as described in 1.3.1, by presenting a surface-exposed loop, which will be recognized by the uPA as a pseudo-substrate. After cleavage, the newly generated *C*-terminal residue is bounded covalently to the active site of the uPA and inhibits the protease activity irreversibly. The serpins are considered as suicide inhibitors since they induced endocytosis of the uPA-uPAR-PAI complex after inhibition and get degraded together with the uPA, while the uPAR is recycled and transported back to the cell surface (Figure 11).[161],[191],[192]

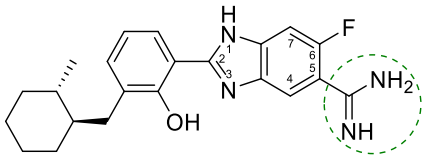
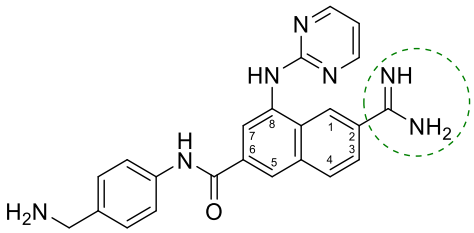
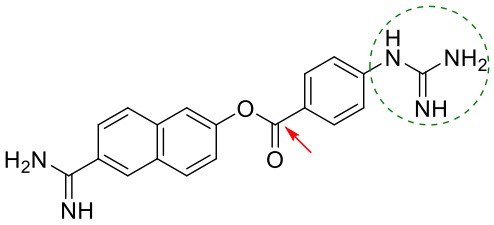
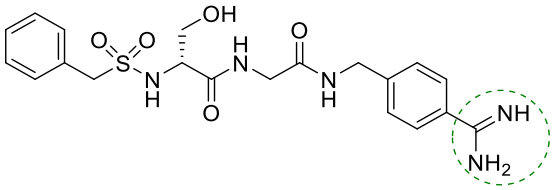
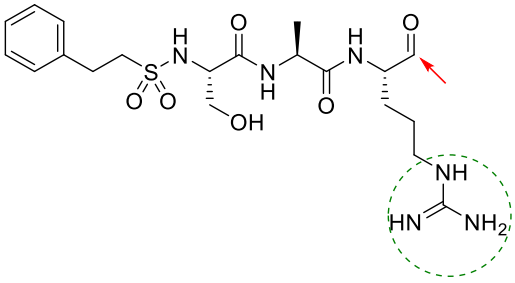
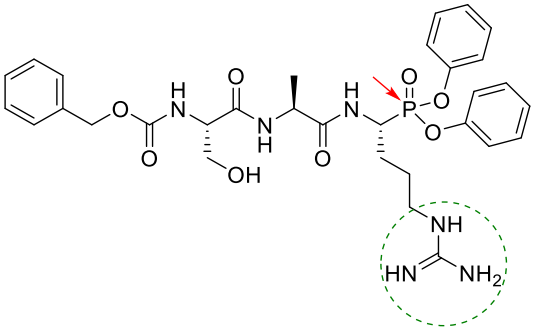
The first synthetic compounds which were designed to inhibit uPA represent small molecules with an aromatic amidine moiety, that addresses the deep S1 pocket with the negatively charged Asp189. The addition of a halogen atom in position 6 of the 5-amidinobenzimidazole inhibitor led to a significant increase of selectivity toward uPA against other proteases with an Ala190, due to the abovementioned displacement of the conserved water molecule and interaction of the Ser190 residue in the S1 pocket (1.4.3, Table 2 entry a).[187] Another class of aromatic amidines represent the 2-naphthamidines, which fit well in the S1 pocket and by substitution at position 8 utilize the unique S1β subpocket to improve potency and selectivity (Table 2 entry b).[193],[194] In recent years, the small molecule nafamostat received

increased attention due to the excellent inhibitory activity against serine proteases. The inhibitor presents an aryl guanidine group, which addresses the S1 pocket (Table 2 entry c). Furthermore, the uPA gets inhibited by covalent reversible modification (acylation) of the Ser195 residue.[195]

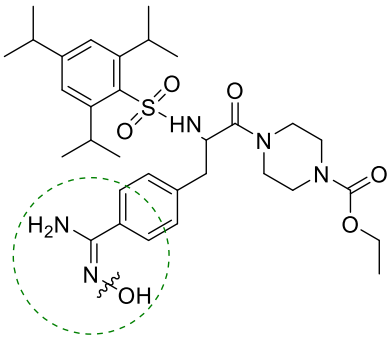
Structure-based design led to the development of peptidomimetic inhibitors. One of the first synthesized compounds are the 4-amidinobenzylamide-based inhibitors, which target almost every pocket of the binding groove (S1–S4), in contrast to the small molecule inhibitors, which primarily address the S1 pocket. By combining the aromatic amidine element of the previously mentioned small molecule inhibitors with the peptidomimetic backbone, promising drug candidates could be obtained with affinities in the single digit nanomolar range and moderate selectivity towards other trypsin-like serine proteases (Table 2 entry d).[196]

Introducing electrophilic moieties at the *C*-terminus, the originally non-covalent inhibitors were converted into covalent reversible aldehyde or irreversible phosphonate inhibitors (Table 2 e, f). Both inhibitors exhibit decent affinities in the nanomolar range with moderate selectivity.[180],[197] Because of the irreversible binding mechanism of the phosphonate warhead and the metabolic instability of the aldehyde group, which can lead to toxicity issues and deletion of the inhibitory activity towards the target protease, these inhibitor classes were not suitable for further drug development.[198]–[200]

Table 2: Selected uPA-targeting inhibitors representing typical structures of (non-)covalent small molecule/peptidomimetic drugs. Basic moiety which interacts with the negatively charged Asp189 residue in S1 is circled in green. Electrophilic moieties of covalent inhibitors are indicated with a red arrow.

	Inhibitor	K_i [nM]	Literature	Entry
Small Molecule Inhibitors		11	[187]	a
		0.6	[193]	b
		1	[195]	c
Peptidomimetic Inhibitors		36	[201]	d
		3	[180]	e
		1	[197]	f

Despite their high inhibitory potency towards uPA, most of the presented drugs are limited in their oral bioavailabilities and pharmacokinetic behaviors, due to the presence of the positively charged P1 moiety. To improve the efficacy of these inhibitors, the prodrug-approach can be used to cover the polar amidine/guanidine group of the respective inhibitor for better pharmacokinetic properties.[202] The amidine compound WX-UK1 (Mesupron©, upamostat, Wilex-AG) was converted into its amidoxime prodrug for better oral bioavailability. Although the comparatively low potency to the uPA, WX-UK1 is currently the most promising candidate yet, which is in clinical phase II trials as an antimetastatic agent in breast cancer, due to the enhanced pharmacokinetic behavior.[203],[204]



K_i = 410 nM (active drug)

Figure 14: Upamostat as the most promising uPA inhibitor with an K_i value of 410 nM of the active drug.[205] Shown in its prodrug form with an amidoxime moiety. Wavy line indicates cleavable bond of the prodrug. Figure created with ChemDraw.

1.5 Transmembrane Serine Protease Subtype 2 (TMPRSS2)

So far, all mentioned serine proteases describe the archetype of proteases like trypsin, thrombin and uPA, which are secreted from cells.[56],[161],[206] In fact, many proteases are membrane-associated.[207] Type II transmembrane serine proteases (TTSPs) are multidomain enzymes, which consist of an extracellular, a transmembrane, and an intracellular domain. Based on their phylogeny and architecture, the TTSP family contains over twenty members, which are grouped into four subfamilies: Hepsin/TMPRSS, Matriptase, HAT/DESC and Corin.[208] Here, I present the physiological background of the transmembrane serine protease subtype 2 (TMPRSS2), its role in viral infection and possible treatment targeting TMPRSS2.

1.5.1 Physiological Background of TMPRSS2

TMPRSS2 is expressed as a 492 amino acid protein, consisting of an *N*-terminal cytoplasmatic domain, a transmembrane domain, the LDL receptor class A domain (LDLRA) that binds calcium ions, a scavenger receptor cysteine-rich domain (SRCR) important for mediating non-opsonic phagocytosis upon ligand binding and the *C*-terminal serine protease domain (Figure 15).[178],[209]–[212] Autoactivation of the 54 kDa zymogen by proteolytic cleavage between Arg255 and Ile256 results in the catalytically active serine protease domain. Thereby, the protease domain can remain bounded to the pro-domain by

an interdomain disulfide bond or can be secreted into extracellular space.[213],[214] Multiple human organs express TMPRSS2, including the prostate, respiratory tract, gastrointestinal tract as well as the liver.[215],[216] The transmembrane protease is predominantly found in plasma membranes of epithelial cells, for example on the apical membrane of luminar cells of the prostate, on epithelial cells of nasal and tracheal tissues, and in the distal airways.[217],[218]

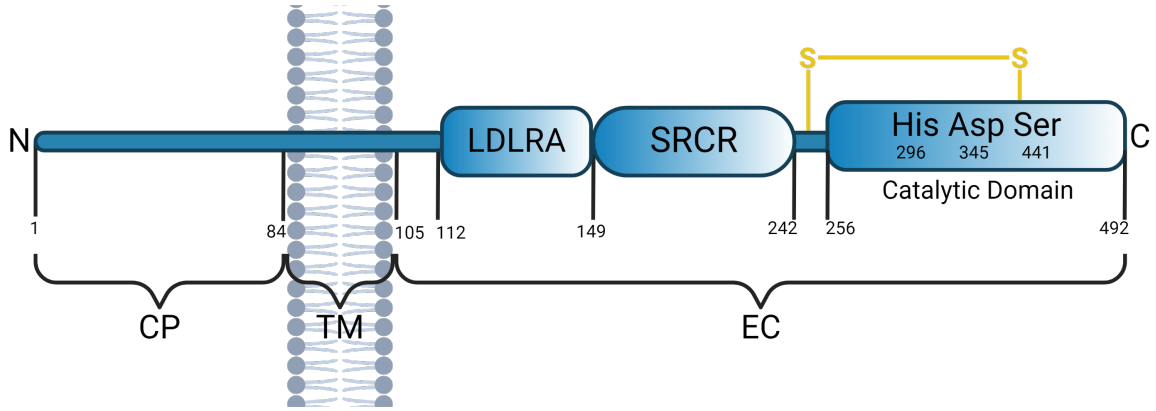


Figure 15: Architecture of (inactive) TMPRSS2 consisting of the cytoplasmic domain (CP), transmembrane domain (TM) and extracellular domain (EC), divided into the LDL receptor class A domain (LDLRA), scavenger receptor cysteine-rich domain (SRCR) and catalytic serine protease domain with the triad (His296-Asp345-Ser441). Numbers indicate amino acid positions and S-S represents the disulfide bond. Figure created with BioRender.com and modified from Wettstein *et al*..[212]

Although the biological function of mammalian TMPRSS2 remains largely elusive, several potential substrates have been described. *In vitro* studies using androgen-dependent prostate cancer cell lines revealed that the protease-activated receptor-2 (PAR-2) might be an endogenous substrate of TMPRSS2.[219] The release of calcium ions from intracellular stores after activation of PAR-2, indicate also initiation of downstream signaling.[220] Additional studies revealed that overexpression of TMPRSS2 activates the structurally related transmembrane protease matriptase, which in turn activates PAR-2.[212],[221],[222] Despite the fact that PAR-2 is found in several TMPRSS2 associated tissues and mediates protective functions in the respiratory and gastrointestinal tract, deeper investigations have to be performed to elucidate the connection between TMPRSS2 and the matriptase-PAR-2 axis.[223],[224] Due to the observed decreased amiloride-sensitive epithelial sodium channel (ENaC) activity and reduced protein levels in presence of TMPRSS2, the serine protease might also influence the airway surface liquid volume and mucociliary clearance.[217],[225] However, it remains unclear whether it is a TMPRSS2-mediated phenomenon or not.

Furthermore, studies of TMPRSS2-knockout mice did not reveal phenotypic abnormalities in terms of embryonic development, fertility and organ function, which points out the functional redundancy of one or more TTSPs family members.[226] In contrast to the physiological background, the pathological role of TMPRSS2 concerning viral cell entry has been investigated extensively and will be discussed in the following.

1.5.2 Role of TMPRSS2 in Viral Cell Entry

Enveloped viruses like influenza and coronaviruses use their envelope glycoproteins to interact with cell surface receptors to initiate cell entry. Before the membrane fusion of the virus and the host cell occurs, proteolytic activation is sometimes necessary. Depending on the virus, some glycoproteins either get cleaved after viral egress from producer cells, immediately before the membrane fusion, or experience multiple independent cleavage events.[227],[228] The involvement of specific proteases for these events is dependent on the cleavage site motifs on the glycoproteins. Because of the high expression levels of TMPRSS2 in the human airways, the TTSP has an important role in viral cell entry of relevant respiratory diseases.[229]

Since the outbreak of a new respiratory disease in Wuhan (China) 2019 and the assignment to the family of the coronavirus, these viruses regained attention.[230],[231] Due to the high (RNA-)genome similarity to severe acute respiratory syndrome-coronavirus (SARS-CoV; 79% sequence identity), it was called SARS-CoV-2.[232] The virus rapidly spread and became a pandemic, which caused over 770 million infection and 7 million deaths worldwide, according to the WHO so far.[233]

Coronaviruses belong to the family of *Coronaviridae* and can be divided into four classes, which comprise seven coronaviruses that infect humans (hCoVs): the α-CoV (hCoV-229E and hCoVNL63), β-CoV (hCoV-OC43, hCoV-HKU1, SARS-CoV, MERS-CoV, SARS-CoV-2), γ-CoV and δ-CoV.[234] The α-CoVs and β-CoVs (hCoV-OC43, hCoV-HKU1) usually cause mild respiratory illness, whereas SARS-CoV, MERS-CoV and SARS-CoV-2 are characterized by severe courses of disease.[227] SARS-CoV-2 belongs to the positive-sense single-stranded RNA viruses (ss-(+))-RNA, which encodes at least 29 proteins in total, consisting of four structural proteins: the spike (S)-, membrane (M)-, envelope (E)-, and nucleocapsid (N)-protein.[235] The virions display a spherical structure, with the prominent crown-shaped spike- , membrane- and envelope-proteins located on the surface, while the nucleocapsid protein binds and protects the genomic RNA inside the particle.[236]

Coronavirions bind to the angiotensin-converting enzyme 2 receptor (ACE2) on the cell surface which promotes the viral cell entry together with other host factors (TMPRSS2). After the release and uncoating of the viral RNA, the intracellular replication process of the SARS-CoV-2 virus immediately starts with translation of non-structural proteins and the following formation of a replication complex (Figure 16). Afterwards, the polyproteins pp1a and pp1ab are expressed by translation of the replicase gene of the RNA. Pp1a and pp1ab get processed through autocleavage by the main viral protease (M^{pro}) and papain-like protease (PL^{pro}). This results in the formation of non-structural proteins (nsps), which form the viral replication/transcription complex, and structural proteins (S-, M-, E-proteins). These translocate into endoplasmic reticulum (ER) membranes and transit through the ER to the Golgi intermediate compartment, where interaction with N-protein encapsulated, freshly produced genomic

RNA results in virus assembly.[237],[238] The finished virus particles are transported to the cell surface and release by exocytosis.[227]

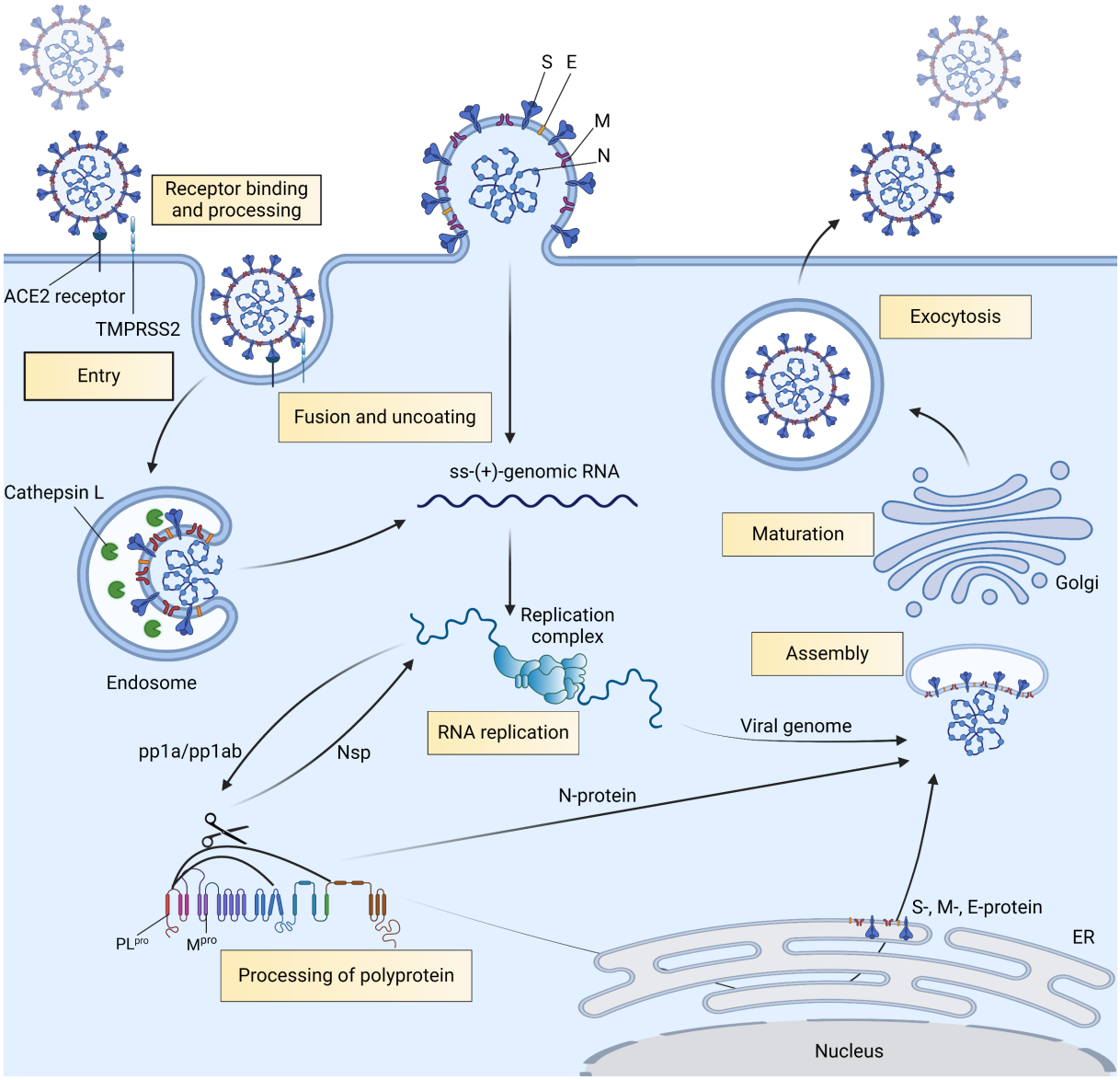


Figure 16: Schematic replication cycle of the SARS-CoV-2 coronavirus. The virus particle binds with the S-protein to its specific receptor (ACE2), gets processed by TMPRSS2 or cathepsin L, and fusions with the cellular or endosomal membrane. Immediately after uptake, the uncoated RNA gets translated into non-structural proteins leading to the formation of the replication/transcription complex. Replicated viral RNA is translated into polyproteins pp1a and pp1ab, which get processed by viral proteases (M^{pro} & PL^{pro}) to nsps and structural proteins (S-, M-, E-, N-proteins). These result in formation of assembled virus particles and replication complexes. The virions get maturated in the Golgi apparatus and released via exocytosis. Figure created with BioRender.com and modulated after V´kovski *et al*. [227]

The 600 kDa homotrimeric S-protein on the viral surface is the main mediator for the viral cell entry and plays a crucial role for the receptor recognition of the ACE2 receptor, cell-virus fusion and intracellular release of the viral RNA.[239] The glycoprotein is divided into the functional parts S1 and S2. During the intracellular biosynthesis of the virions, the S1 and S2 subunits are cleaved by the serine protease furin in the golgi-apparatus (maturation, Figure 16), but are still non-covalently associated with

each other.[240] The surface-exposed S1 subunit contains the receptor-binding domain (RBD), which is responsible for the ACE2 recognition and binding. The transmembrane S2 domain contains the fusion protein (FP) that is crucial for the fusion of viral and cellular membranes.[239],[241] After receptor binding of the virus particle, a second cleavage at the S2 subunit, termed the S2´ site occurs, mainly by TMPRSS2 at the cell surface or by cathepsin L in the endosomal compartment.[242]–[244] This induces the release of the fusion peptide, initiating pore formation and enables membrane fusion (Figure 17).[245]

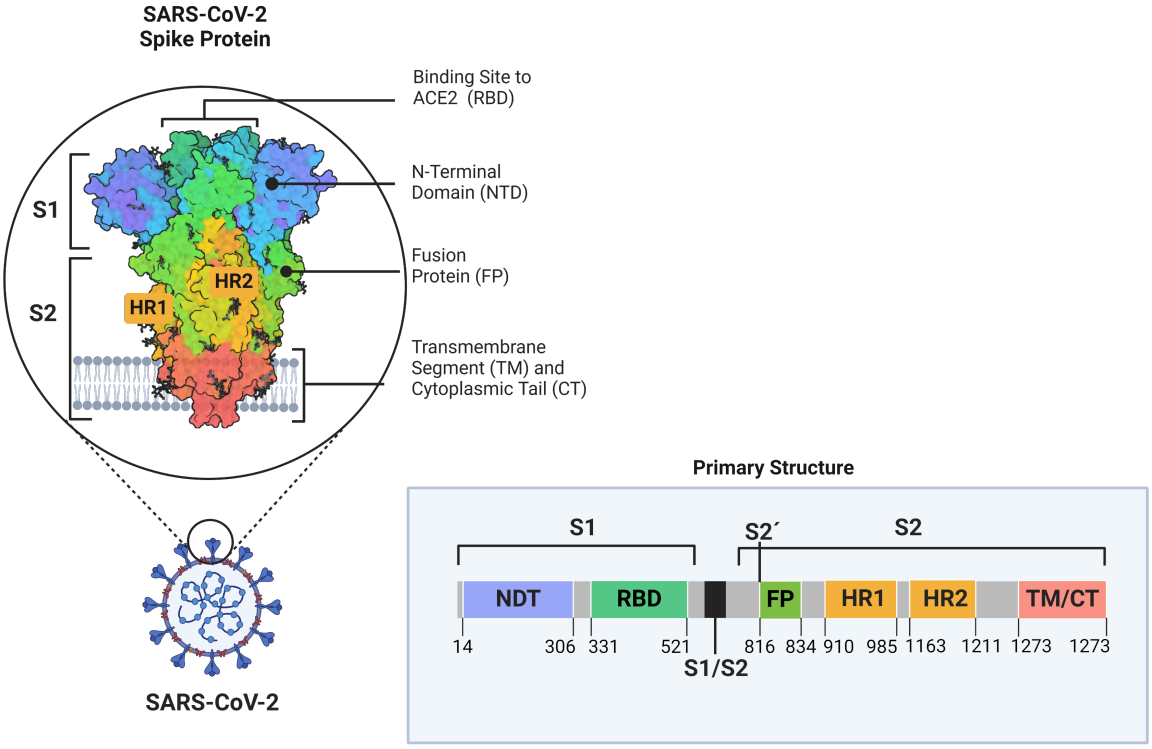


Figure 17: SARS-CoV-2 virus particle with structural insight of the S-protein (pdb: 6VXX) and its primary structure (numbers indicate amino acid positions), consisting of the S1 und S2 subunit. S1 contains the receptor binding site to ACE2 (RBD, mint green) and *N*-terminal domain (NTD, blue). S2 contains the fusion protein (FP, green), heptad repeat 1 & 2 (HR1/HR2, orange), which is involved in the formation of the fusion pore, as well as the transmembrane segment, cytoplasmic tail (TM, CT, red) and S2´ cleavage site. Figure created with BioRender.com and modified after Jackson *et al*. and Müller *et al*..[211],[246]

Although both proteases (cathepsin L & TMPRSS2) can mediate the virus-cell fusion in cell culture, several *in vivo* studies suggest a preference of TMPRSS2-dependent viral cell entry. Firstly, coronavirus infect and replicate primarily in epithelial cells of the respiratory tract while the expression levels of cathepsins in lungs are low. Furthermore, TMPRSS2-knockout mice showed a reduced pathogenicity and spread of *β*-CoVs (MERS-CoV, SARS-CoV and SARS-CoV-2).[247],[248] Hence, targeting TMPRSS2 to prevent the viral cell entry of SARS-CoV-2 represents a promising strategy for treatment of the virus infection.

1.5.3 Structure of the TMPRSS2 Protease Domain

The extracellular catalytic serine protease domain at the *C*-terminus belongs to the PA clan, S1 family of trypsin-like serine proteases according to the MEROPS database.[34] Therefore, it consists of the typical catalytic triad of Asp345-His296-Ser441 and displays the canonical trypsin fold with two six-stranded *β*-barrels harboring the active site cleft in between (1.2.2, Figure 18, left).[249] In general, the serine protease domain is highly conserved with all TTSPs, but divergent substrate specificity is conferred by eight highly variable surface-exposed loops. The characteristic negativley charged Asp435, together with Ser436 and Gly464 form the S1 subpocket and lead to the trpysin-like substrate specificity of an positively charged Arg or Lys residue (Figure 18, right). The S2 pocket harbors Lys342 that likely prefers small and/or electronegative substrates. S3 and S4 appears suitable for various amino acids with the respective Gln438 and Thr341 residues. Overall, the active site of TMPRSS2 seems to be flexible and therefore capable of binding different substrate sequences with the biggest selectivity behavior caused by S1 and S2 pocket.[249]

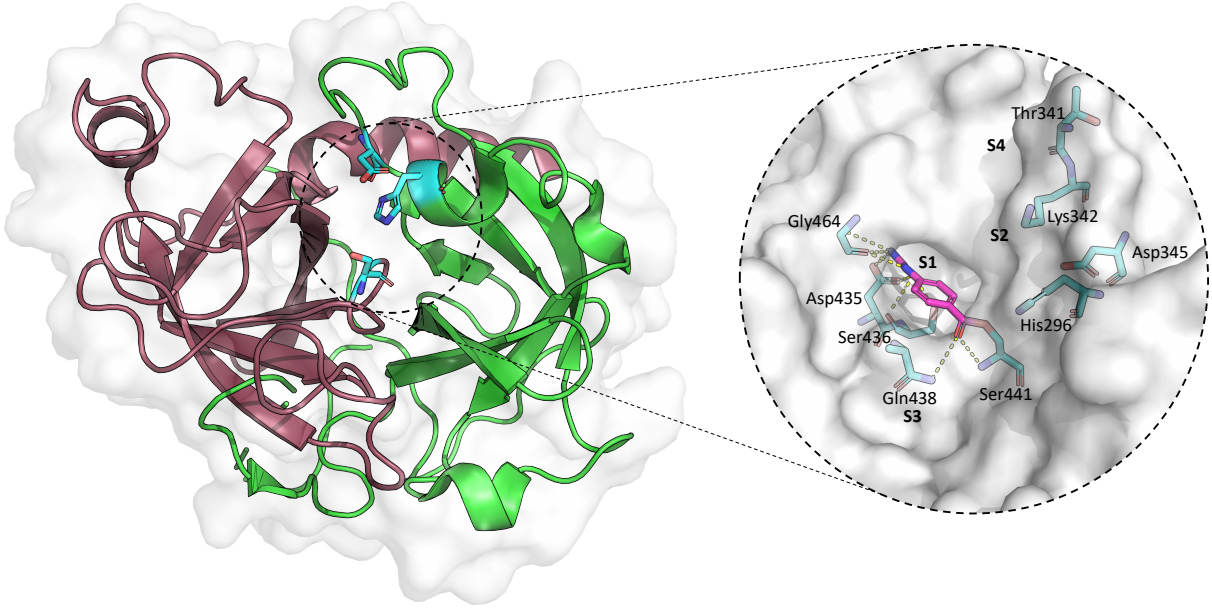


Figure 18: X-ray crystal structure of TMPRSS2 protease domain (pdb: 7MEQ) showing the characteristic trypsin-like fold with two opposed six-stranded *β*-barrels (green and red), harboring in the middle the catalytic active site (cyan). On the right side, visualization of the binding site (co-crystallized with nafamostat, magenta) formed by the **catalytic triad** (Asp435-His296-Ser441), deep **S1** pocket (Asp435, Ser436, Gly464), **S2** (Lys342), **S3** (Gln438) and **S4** (Thr341).[249] Yellow dashed line presents polar interactions between the inhibitor and the residues of the binding site. Figure created with PyMOL and PowerPoint.

1.5.4 Inhibitors of TMPRSS2

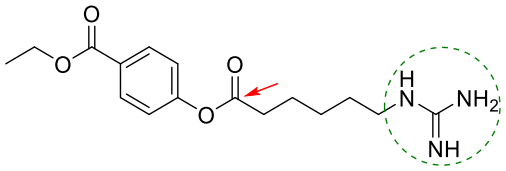
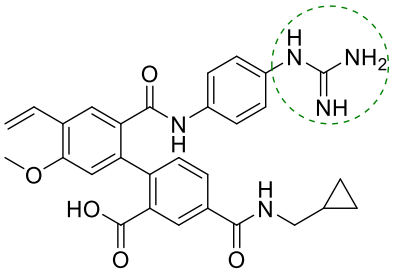
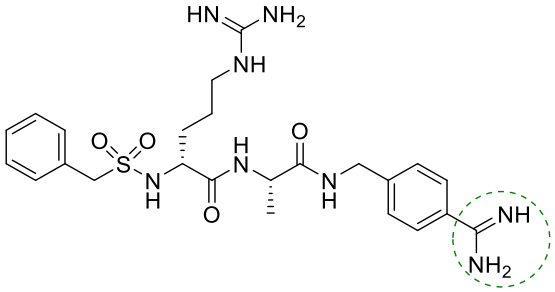
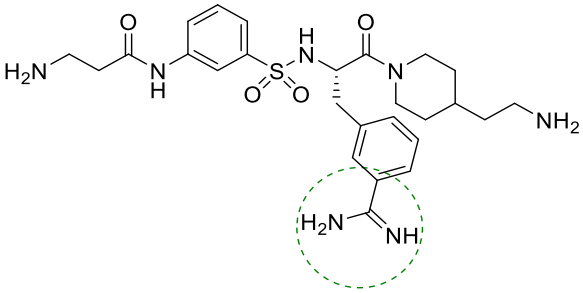
The characterization of the serine protease domain of TMPRSS2 by X-ray crystallography in 2022 simplified and accelerated the development of potential drugs against TMPRSS2. Since the outbreak of COVID-19, several known serine protease inhibitors or new designed drugs targeting the catalytic domain of TMPRSS2 have been described and will be discussed in the following.

The small molecule inhibitor gabexate was predicted to inhibit the proteolytic activity of TMPRSS2 by virtual screening.[250] Afterwards, reduced enzymatic activity could be shown with nanomolar affinities in *in vitro* assays (Table 3 entry a). The inhibition mechanism is based on the reversible acylation of the catalytically active Ser441, comparable with nafamostat (1.4.4). However, in further studies no reduction of SARS-CoV-2 mediated entry into lung cells could be detected nor an effect to the respective virus-cell membrane fusion.[251]

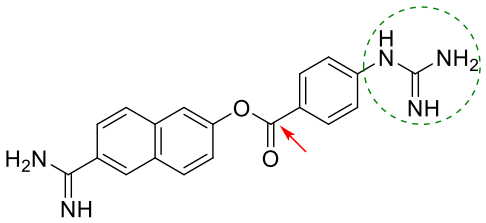
The small molecule inhibitor nafamostat, approved in Japan and Korea for treatment of pancreatitis and used as an anticoagulant, was detected to inhibit recombinant TMPRSS2. The inhibitor displays affinities in the subnanomolar range and blocks the catalytical serine in a covalent-reversible mechanism (Figure 19).[252] In contrast to gabexate, the drug also inhibits SARS-CoV-2 infection in human cells.[253] Another example of drug repurposing represents the non-covalent kallikrein inhibitor avoralstat, which is in clinical trials for treatment against hereditary angioedema and showed positive results in blocking SARS-CoV-2 cell entry and infection (Table 3 entry) b.[254]

Based on previously designed peptidomimetic inhibitors against other TTSPs and trypsin-like serine proteases, several TMPRSS2 substrate analogues were developed. One of the most potent non-covalent inhibitors belong to the groups of 4-amidinobenzylamide and 3-amidinophenylalanyl inhibitors with affinities in the low nanomolar range (Table 3 entry c, d).[255] Similar to the small molecules, the positively charged amidine moiety of both inhibitor groups protrudes into the S1 pocket and interacts with the negatively charged Asp435.

Table 3: Selected TMPRSS2 targeting inhibitors representing typical structures of (non-)covalent small molecule/peptidomimetic drugs. Basic moiety, which interacts with the negatively charged Asp435 residue in S1 in green circled. Electrophilic moiety of covalent inhibitors indicated with a red arrow.

	Inhibitor	K_i [nM]	Literature	Entry
Small Molecule Inhibitors		96	[252]	a
		1	[254]	b
Peptidomimetic Inhibitors		18	[255]	c
		8	[255]	d

Due to the positive effect of nafamostat in blocking the viral cell entry, first randomized clinical trials were performed to investigate the efficacy and safety in hospitalized patients with severe COVID-19 patients. The convincing results regarding safety and efficacy qualify nafamostat as a possible drug for treatment of SARS-CoV-2 infections. This progress of TMPRSS2 inhibitor-development highlights the opportunity for an alternative or additional therapeutical approach in comparison to currently treatment by targeting the viral host protease M^{pro}.[256],[257]



K_i = 0.2 nM

Figure 19: Nafamostat as the most promising TMPRSS2 inhibitor with an K_i value of 0.2 nM.[252] Figure created with ChemDraw.

1.6 Hepsin

Hepsin represents another type II transmembrane serine protease and belongs to the same subfamily (Hepsin/TMPRSS) as the above described TMPRSS2 (1.5).[208],[258] Therefore, the architecture of hepsin as well as the biosynthesis and zymogen activation is similar to the presented structure-related TMPRSS2 (1.5) and will not be discussed further.[71],[259]

1.6.1 Physiological Function and Pathological Association of Hepsin

In 1988 hepsin was first identified in human hepatocytes and named to the greek word *hepar* – liver.[260] In addition to the liver, the serine protease is expressed in a wide range of tissues, including prostate, lung, kidney, thyroid tissues etc. and is associated with multiple physiological functions. [261] *In vivo* experiments revealed that hepsin-deficient mice have low levels of liver glycogen, triglyceride, blood glucose and cholesterol. Furthermore, the levels of hepatic triglyceride and cholesterol remained low, when hepsin-knockout mice were treated with a high fat diet, compared to similarly treated wild-type mice.[259] Responsible for these observations is the regulatory mechanism of the glycogen, lipid and protein metabolism. Hepsin cleaves the pro-hepatocyte growth factor (pro-HGF) to the active two-chain HGF, which represents a ligand for the receptor tyrosine kinase Met. Pro-HGF is synthesized in the liver and binds, but not activates, Met. After the activation, downstream signaling increases the glycogen, lipid and protein synthesis in hepatocytes. Hepsin-mediated pro-HGF cleavage and following Met activation have been shown in *in vivo* and *in vitro* experiments.[262]–[264]

Overexpression of hepsin is reported in different cancer-types, including prostate, lung, breast cancer etc. and displays different mechanism enhancing tumor progression.[79],[265] These comprise the

upregulation of the HGF/Met signaling, induced by high expression levels of hepsin in prostate cancer.[262] Furthermore, *ex vivo* studies revealed the activation of pro-matrix metalloproteases by hepsin leading to the degradation of the extracellular matrix.[266] Both pathological-associated mechanism lead to an aggressive metastatic behavior as well as enhance the tumor progression and survival.[259],[267],[268] Thus, the transmembrane serine protease hepsin represents a promising target for treatment against numerous cancer-types.

1.6.2 Structure of the Hepsin Protease Domain

The protease domain of hepsin is represented by a trypsin-like serine protease (clan PA, S1 family), harboring the catalytic triad Asp257-His203-Ser353 in the active site between the two six-stranded *β*-barrels of the canonical trypsin fold (1.2.2, Figure 20, left). A negatively charged Asp347 combined with Gly380 form the deep S1 pocket and lead to the characteristic trypsin-like substrate specificity for a positively charged Arg or Lys residue (Figure 20, right). The S2 site is defined by a shallow binding pocket, consisting of Ser376 and His203, which accommodate mostly smaller hydrophobic or polar groups. S3 contains Asn254 and Gly378 that enable hydrogen bonding with different polar residues. The oxygen atom of Gln331 point towards the S4 pocket, which gives an overall negative electrostatic potential and explains why a lysine is favored in this position.[262] Similar to TMPRSS2, hepsin presents a flexible active site able for binding various substrates, but partly displaying different amino acid residues in the subpockets, which offer the opportunity for development of potent and selective inhibitors.

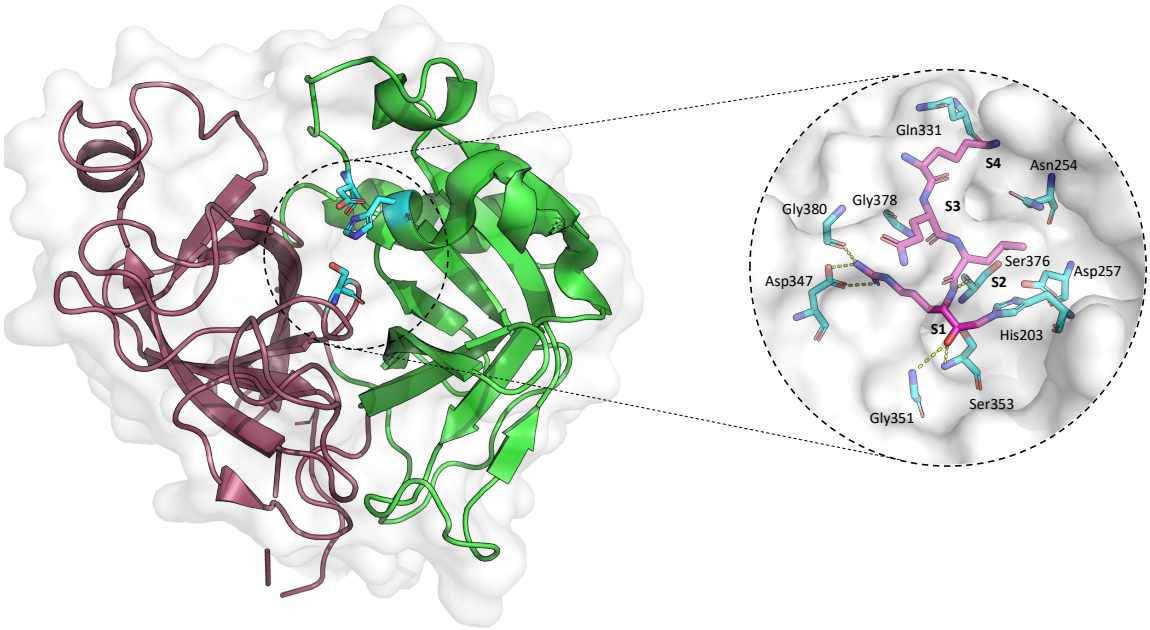


Figure 20: X-ray crystal structure of the hepsin protease domain (pdb: 1Z8G) showing the characteristic trypsin-like fold with two opposed six-stranded *β*-barrels (green and red), harboring in the middle the catalytic active site (cyan). On the right side, visualization of the binding site (co-crystallized with the peptidomimetic inhibitor acetyl-Lys-Gln-Leu-Arg-chloromethylketone, magenta) formed by the **catalytic triad** (Asp257-His203-Ser353), deep **S1** pocket (Asp347, Gly380), **S2** (Ser376, His203), **S3** (Asn254 and Gly378) and **S4** (Gln331). Yellow dashed line presents polar interactions between the inhibitor and the residues of the binding site.[262] Figure created with PyMOL and PowerPoint.

1.6.3 Inhibitors of Hepsin

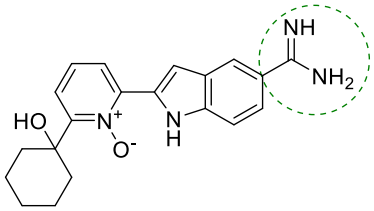
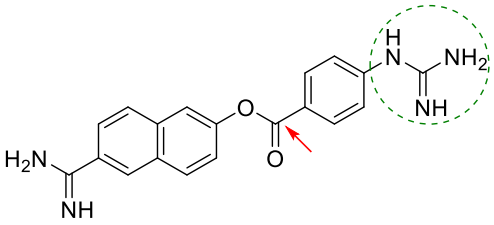
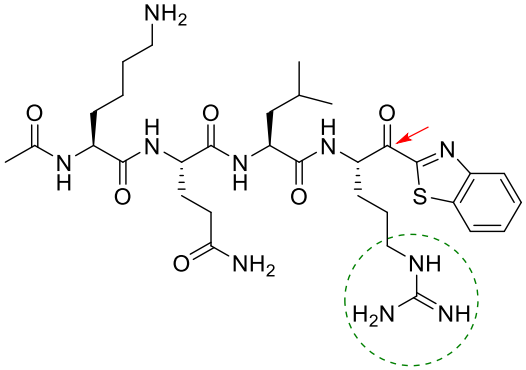
The involvement of hepsin overexpression in several cancer-types has attracted considering interests in developing hepsin inhibitors. Numerous low molecular weight inhibitors have been designed and can be categorized into two classes. First ones present heterocyclic amidine/guanidine-based compounds.[269]

1*H*-Indole-5-carboxamidine derivates were found as potent hepsin inhibitors showing affinities in the nanomolar range (Table 4 entry a), which protrude the basic amidine moiety into the S1 pocket for electrostatic interactions with the negatively charged Asp347. Substitution at C-2 position with a pyridine *N*-oxide derivate resulted in an increase of affinity, due to additional interactions between the *tert*-hydroxyl and *N*-oxide group with the catalytic Ser353 and His203.[270]

Other members of this inhibitor class are the benzamidine/guanidine derivates, listed with the small molecule nafamostat as the most potent compound. The covalent-reversible inhibitor displays inhibition rates in the subnanomolar range (Table 4 entry b).[271]

The second class represents peptide-derived analogs. Tetrapeptides were developed to mimic the natural substrate of hepsin, which were attached at the *C*-terminus with an α-ketobenzothiazole warhead. This resulted in very potent covalent-reversible inhibitors with affinities in the subnanomolar range (Table 4 entry c).[272]

Table 4: Selected hepsin targeting inhibitors representing typical structures of (non-)covalent small molecule/peptidomimetic drugs. Basic moiety, which interacts with the negatively charged Asp347 residue in S1 in green circled. Electrophilic moiety of covalent inhibitors indicated with a red arrow.

	Inhibitor	K_i [nM]	Literature	Entry
Small Molecule Inhibitors		100	[270]	a
		0.5	[271]	b
Peptidomimetic Inhibitors		0.1	[272]	c

1.7 Challenges in the Development of Trypsin-Like Serine Proteases

The structural comparison between the presented uPA, TMPRSS2 and hepsin inhibitors as well as binding sites suggest cross-reactivity of the drugs towards all three proteases. Indeed, for example nafamostat (K_i (uPA): 1 nM, K_i (TMPRSS2): 0.2 nM, K_i (hepsin): 0.5 nM) and gabexate (K_i (uPA): 191 nM, K_i (TMPRSS2): 96 nM, K_i (hepsin): not available) display similar literature known inhibitory effects on the proteases.[195],[252],[273] This highlights one major problem in the development of potential trypsin-like serine protease inhibitors for clinical use. The high similarity of the binding pockets, caused by the conserved trypsin-fold, results frequently in off-target interactions of synthesized inhibitors. Due to the involvement of this protease class in important processes, like hemostasis and general supply of nutrients and therefore maintenance of the physiological system (1.2.3), off-target inhibition can lead to severe side effects.[274]–[276]

Therefore, it is of utmost importance to synthesize not only potent inhibitors towards the target protease, but selective ones. This can be achieved for example by structure-based design, which includes identification and characterization of potential binding sites as well as discovering unique properties like single amino acid mutation and unusual shapes of subpockets, that could present opportunities for designing highly selective inhibitors (1.4.3).[185],[277]–[280] Additionally, different computational tools to simulate possible binding events can help to get a deeper insight of the binding cleft and enhance the drug design.[281]

Another challenge is the insufficient oral bioavailability and general problematic pharmacokinetic properties of the highly basic and polar inhibitors described in 1.4.4 .[282] One strategy to overcome this issue presents the aforementioned prodrug approach (1.4.4) to mask the polar groups, for example by conversion to their respective amidoxime, carbamate or *N*-hydroxyguanidine derivates, until their target compartment is reached.[202],[283] Another possibility displays a more technological approach, by nanoparticle formulation of polar drugs to enhance the pharmacokinetic properties and lead to sufficient drug concentration at their target.[284],[285]

1.8 Investigation of Covalent Protease Inhibitors

The goal of rational drug design of protease inhibitors is the development of selective and active compounds — generally a small peptidomimetic molecule that inhibits the action of the enzyme (1.3.2). The evaluation of the pharmacodynamic properties in terms of affinity and specificity towards the target protease as well as the inhibition mechanism of the inhibitor present the next steps to gain important information for identifying active compounds and further optimization.[286],[287]

1.8.1 Fluorometric Protease Inhibition Assays

The gold standard for initial potency investigations of protease inhibitors in medicinal chemistry represent the fluorometric based *in vitro* assays. This method is widely used, due to the easily accessible instrumentation, time efficiency and high sensitivity.[288],[289] Additionally, there is a large supply of different, commercially available fluorometric protease substrates, which cover the majority of proteases. The first fluorometric substrates resembles the so-called Förster energy transfer (FRET) substrates in 1973, which includes a quencher (Q) molecule and a fluorophore (F).[290] Due to the enhanced detection sensitivity, the newly invented fluorometric method replaced the previously used colorimetric *p*-nitroanilide (*p*NA) substrates.[291] Three years later the implementation of the 7-amino-4-methyl coumarin (AMC) derivates led to new fluorescent protease substrates, which were less vulnerable to assay interferences compared to the FRET and *p*NA derivates.[292]

The process of fluorometric protease inhibition assays is described as follows: The fluorescent peptide gets cleaved by the target protease via amide bond hydrolyzation. This leads either in an increased distance between the quencher molecule and the fluorophore regarding the FRET substrates or in the separation of the fluorescent AMC group from the peptide (Figure 21).[293] In both cases, the fluorophore is now able to get excited for example by a microplate reader at a specific excitation wavelength and following fluorescence can be detected at a corresponding emission wavelength. The resulting increase of fluorescence over time is a direct indicator for the proteolytic activity of the enzyme. In presence of an inhibitor, the impaired protease is hindered for substrate cleavage. Thus, the resulting decrease in fluorescence over time is proportional to the inhibitory potency and can be evaluated by different methods.[288],[294]

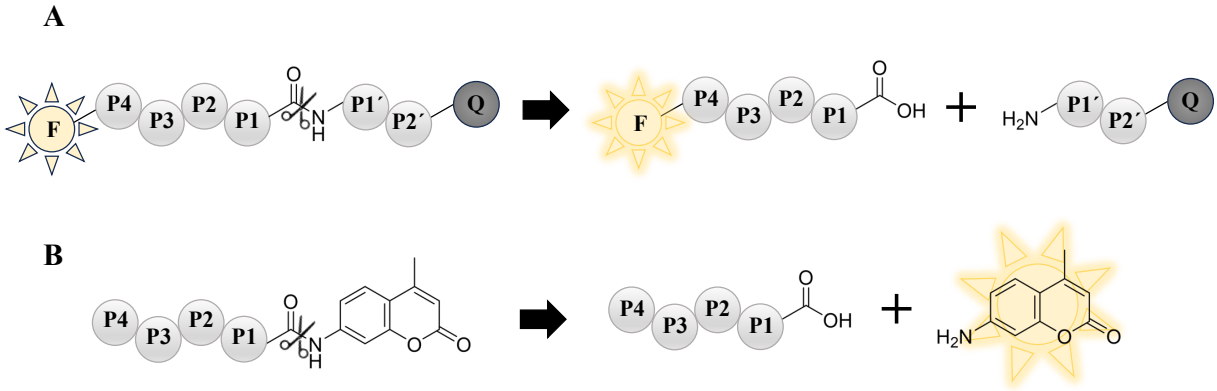


Figure 21: Principle of fluorometric substrates. (**A**) After the enzymatic cleavage of the peptide bond, the distance between fluorophore (F) and quencher (Q) increases. This allows fluorescence measurement, which was suppressed by the quencher. (**B**) Enzymatic cleavage results in release of 7-amino-4-methyl coumarin (AMC) and following detection of fluorescence. Figure created with ChemDraw and PowerPoint.

It is important to check the inhibitory mechanism of the designed compounds to determine the appropriate kinetic data, since the evaluation of reversible and irreversible binding inhibitors differs, and the use of the same kinetic constants would be meaningless and lead to false information.[288] Usually, covalent interactions (reversible and irreversible) can be predicted due to structure elements in most cases by the appearance of electrophilic traps (warheads).[295],[296] However, the exact covalent binding mode is only evident through binding experiments like LC/MS analysis with the target protein.[297]

Within the scope of this dissertation, the evaluation of covalent reversible and two-step covalent irreversible inhibitors were performed. Therefore, both analysis methods will be presented in the following.

1.8.2 Evaluation of (Non-)Covalent Reversible Protease Inhibitors by Fluorometric Assays

Reversible inhibitors regardless of covalent or non-covalent nature, are characterized by the formation and dissociation of an enzyme-inhibitor complex [EI] (Figure 22).

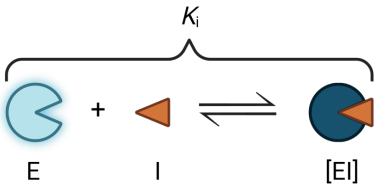


Figure 22: Schematic presentation of (non-)covalent reversible inhibition. E = unbound enzyme, I = unbound inhibitor, [EI] = (non-)covalent reversible enzyme-inhibitor complex, K_i = inhibition constant, which describes the equilibrium for formation and dissociation of the enzyme-inhibitor complex. Figure created with BioRender.com.

To characterize the potency of this inhibitor type, the IC_{50} value represents an adequate parameter. It indicates the inhibitor concentration necessary for 50% enzyme inhibition. At the beginning, fluorescence progress curves for different inhibitor concentrations are measured, which follow a linear

relationship (Figure 23, left). Afterwards, the relative slopes between a negative control without inhibitor and the measurements with different inhibitor concentrations are calculated, which represents the residual enzyme activity. Then, the IC_{50} value can be calculated with the following 4-parameter equation **1**.

$$y = \frac{y_{\max} - y_{\min}}{1 + \left(\frac{[I]}{IC_{50}}\right)^{s}} + y_{\min} \qquad \mathbf{1}$$

Here, the inhibitor concentration $[I]$ is plotted semilogarithmically against the remaining enzymatic activity $[y]$. y_{max} displays the maximum enzymatic activity at small inhibitor concentration and y_{min} the minimum values at high inhibitor concentrations. The Hill factor s describes the steepness of the middle part of the sigmoidal dose-response curve (Figure 23, right).[288]

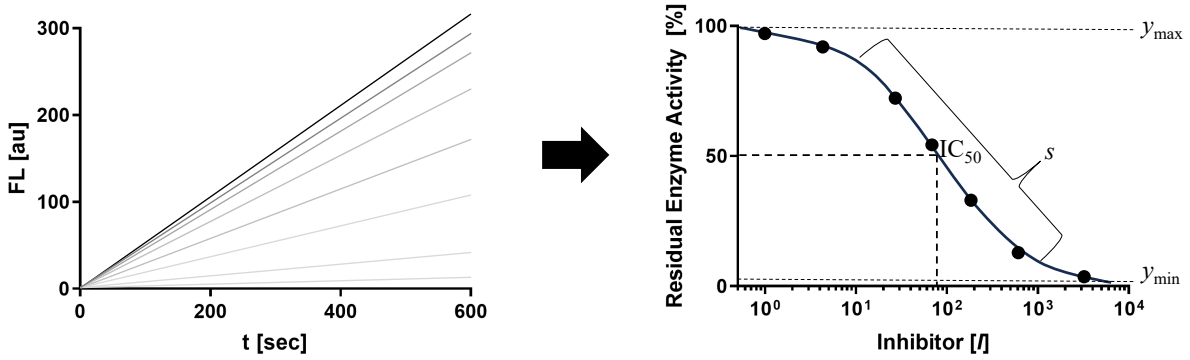


Figure 23: Evaluation of (non-)covalent reversible protease inhibitors by fluorometric assays. Resulting linear fluorescence-progress curves after 10 min measurement without addition of inhibitor (black line) and in presence of different inhibitor concentration (grey lines, left). Typical sigmoidal dose-response curve, after plotting the inhibitor concentration $[I]$ against the residual enzyme activity [%], showing the IC_{50} value, y_{max} and y_{min} as well as the hill factor s. Figure created with GraphPad and PowerPoint.

For competitive inhibitors, the IC_{50} is dependent on the substrate affinity of the protease, characterized by the Michaelis Menten constant (K_M), and its concentration. In contrast, the inhibitory constant K_i is independent of the abovementioned values and therefore recommended to use instead, for better comparison of the inhibition activity towards different proteases. It describes the equilibrium constant for formation and dissociation of the enzyme-inhibitor complex and can be calculated with the mathematical Cheng-Prussoff equation **2**.[288],[298] It presents the relationship between the IC_{50} value with the inhibition constant K_i, the substrate concentration $[S]$ and the Michaelis Menten constant K_M. Regarding non-competitive inhibitors, the IC_{50} value is independent of the substrate affinity and concentration, therefore equals the K_i value.[288]

$$IC_{50} = K_i\left(1 + \frac{[S]}{K_M}\right) \qquad \mathbf{2}$$

1.8.3 Evaluation of Two-Step Covalent Irreversible Protease Inhibitors

In case of irreversible inhibitors, the characterization with the IC_{50} value would be meaningless, because the kinetic parameter is not only dependent of the substrate concentration and affinity towards the protease (1.8.2), but also on the incubation time of the inhibitor with enzyme prior to the substrate addition. This is explained by the two-step inhibition mechanism, which added the irreversible inactivation step of the covalent-modified enzyme [EI]*, after formation of the reversible enzyme-inhibitor complex [EI] (Figure 24).[288]

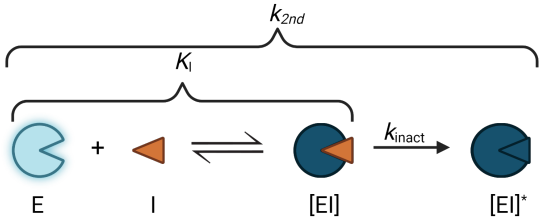


Figure 24: Schematic presentation of two-step covalent irreversible inhibition. E = unbound enzyme, I = unbound inhibitor, [EI] = (non-)covalent reversible enzyme-inhibitor complex, [EI]* covalent irreversible enzyme-inhibitor complex, K_I = inhibition constant, which describes the equilibrium for formation and dissociation of the reversible enzyme-inhibitor complex, k_{2nd} = inactivation efficiency, that comprises the inactivation rate (k_{inact}) over the inhibition constant (K_I). Figure created with BioRender.com.

Due to the irreversible covalent bond, the IC_{50} would decrease exponentially with time. This concludes, that if the incubation time is long enough, the IC_{50} value would always equal half of the total enzyme concentration, which can lead to false information regarding the inhibitor potency.[288],[299] One hint for a time-dependent irreversible reacting inhibitor displays the fluorescence-progress curves over time, due to their non-linear behavior following the exponential equation **3**, compared to the linear shaped progress curves of reversible inhibitors or in absence of inhibitor (Figure 26, left). However, non-linear fluorescence progress curves are not sufficient, to determine the irreversible inhibition mechanism of compounds. Additional experiments like dilution or dialysis assays should be conducted for final proof.[300],[301]

$$F = A\left(1 - e^{-k_{obs}t}\right) + B + offset \qquad \mathbf{3}$$

$$F = \frac{v_i - v_s}{k_{obs}}\left(1 - e^{-k_{obs}t}\right) + v_s\, t + offset \qquad \mathbf{4}$$

The exponential equation **3** expresses the behavior of the fluorescence intensity F, described by the background fluorescence (*offset*), A ($\frac{v_i - v_s}{k_{obs}}$) and B ($v_s\, t$), which are related to the initial v_i and steady-state velocities v_s of the reaction between the inhibitor and the enzyme **4**. Due to the irreversible inhibition of

the enzyme activity after the formation of the covalent bond by the inhibitor, v_s and therefore B equals 0, which leads to the simplified equation **5**.

$$F = \frac{v_i}{k_{obs}} \left(1 - e^{-k_{obs}t}\right) + offset \qquad \mathbf{5}$$

The apparent first-order rate constant k_{obs} describes the conversion between the initial rate from (non-)covalent inhibition v_i to the inactivation steady-state v_s velocity of the reaction and can be calculated by fitting the fluorescence intensity against the incubation time t using equation **5** (Figure 25).

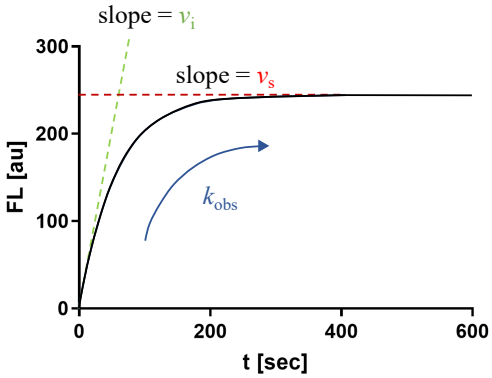


Figure 25: Non-linear fluorescence-progress curve of a two-step irreversible inhibition mechanism, showing the initial velocity v_i ((non-)covalent inhibition) and final velocity v_s (inactivation at reaction completion). k_{obs} describes the rate from (non-)covalent inhibition v_i to the inactivation of the enzyme v_s.

The apparent inhibition constant K_I^{app} (represents the equilibrium constant for formation and dissociation of the reversible enzyme-inhibitor complex, 1.8.2) and k_{inact}, which describes the inactivation rate of the enzyme by irreversible inhibition, can be determined by plotting the inhibitor concentration $[I]$ to the obtained k_{obs} values and fitted to equation **6** (Figure 26, right).

$$k_{obs} = \frac{k_{inact}[I]}{K_I^{app} + [I]} \qquad \mathbf{6}$$

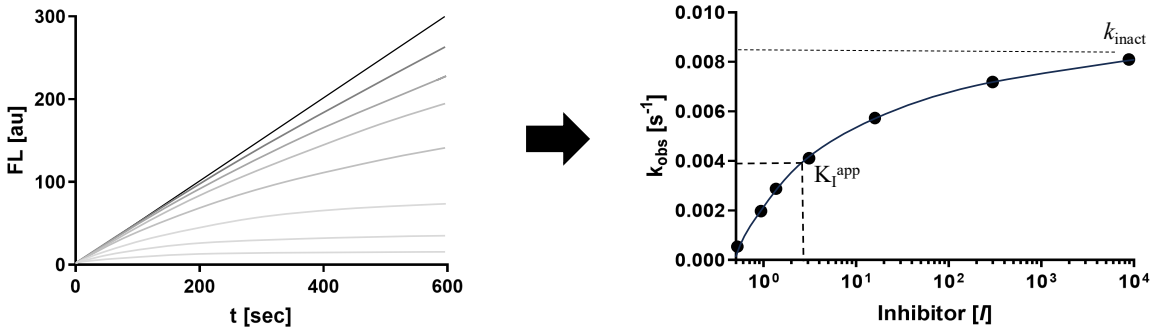


Figure 26: Evaluation of two-step covalent-irreversible protease inhibitors by fluorometric assays. Resulted non-linear fluorescence-progress curves after 10 min measurement without addition of inhibitor (black line) and in presence of different inhibitor concentration (grey lines, left). Hyperbolic dose-response curve, after plotting the inhibitor concentration $[I]$ against the calculated k_{obs} values [s^{-1}], showing the K_I^{app} and k_{inact} value. Figure created with GraphPad and PowerPoint.

The resulting K_I^{app} value can be corrected of the substrate concentration [*S*] and affinity by equation **2**, obtaining the K_I value for an adequate comparison of the inhibition data between different proteases. Furthermore, using equation **7** leads to the calculation of k_{2nd}, which describes the inactivation efficiency of the irreversible inhibitor (Figure24).[302]

$$k_{2\mathrm{nd}} = \frac{k_{\mathrm{inact}}}{K_{\mathrm{I}}} \qquad \mathbf{7}$$

1.8.4 Interferences of Fluorometric Protease Inhibition Assays

At least equally important as the correct evaluation of the respective inhibitor type is the vulnerability of a fluorometric assay system regarding non-specific interferences. Assay robustness describes the ability of the measured signal to avoid perturbation by numerous effects, that originate from various sources like assay components, including the inhibitors themselves.[303]–[305] One common interference during the fluorometric protease inhibition assay is the inner-filter effect. It defines the absorption of the excitation UV light by the analyzed compound, which is caused by similar absorption maxima between the fluorophore and the drug. Due to the spectral overlap of the excitation light and absorption maximum of the inhibitor, the fluorescence is impaired and leads to the false assumption of a positive inhibition result (Figure 27 A). Similar to the inner-filter effect and frequently wrongly confused with each other represents fluorophore quenching by the investigated compound. In contrast to the first mentioned interference, quenching describes the absorption of the fluorescence emitted by the fluorophore, which leads again to false-positive inhibition results (Figure 27 B).[303],[304] Autofluorescence decay represents another phenomenon, which appears when the excitation wavelength of the analyzed compound and fluorophore overlaps as well as the emission wavelengths of both are located in the same detection wavelength. This leads to a fluorescence increase in the beginning, but might change to a decrease of the fluorescence intensity, due to degradation of the compound (Figure 27 C).[305]

The abovementioned examples represent typical fluorescence-based assay interferences, using AMC- and FRET-based fluorogenic substrates, during the evaluation of protease inhibitors. To avoid false results, it is necessary to validate the potential inhibitory activity of the analyzed compounds with additional orthogonal experiments.[303] This leads to more time-consuming and expensive experiments, which affect the overall drug development. Therefore, one project within this dissertation presents the synthesis and characterization of alternative fluorogenic substrates, which are designed to mitigate common assay interferences (chapter A, 4.2).

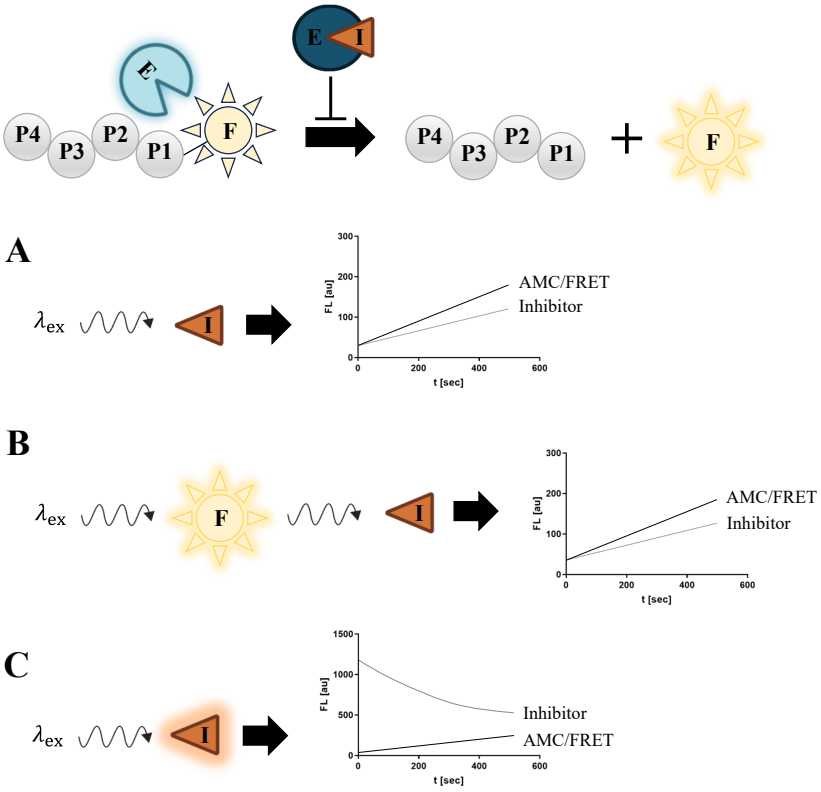


Figure 27: Overview of common fluorometric protease assay interferences caused by the analyzed inhibitor. E = enzyme, I = inhibitor, F = fluorophore, λ_{ex} = excitation wavelength. Showing fluorescence-progress curves of AMC- or FRET-based fluorogenic substrates without addition of inhibitor (black line) and in presence of an inhibitor inducing the respective interference (grey line). (**A**) The inner-filter effect describes the absorption of the excitation UV light by the inhibitor. (**B**) Fluorophore quenching presents the absorption of the fluorescence emitted by the fluorophore. (**C**) Autofluorescence decay is represented by an increase of the fluorescence signal due to the spectral overlap of the excitation and emission wavelength of the analyzed inhibitor and fluorophore. The overall fluorescence decreases after time, because of the degradation of the inhibitor. Figure created with BioRender.com and modified after Maus et al.[306]

2 Projects and Objectives

The scientific work in this doctoral thesis can be categorized into two chapters focused on covalent protease inhibitors. The first one represents the investigation of peptidomimetic covalent inhibitors by different experiments, resulting in a deeper insight and better understanding of this inhibitor class, and improvement of the associated fluorometric inhibition assay.

2.1 Chapter A: Investigation and Characterization of Warheads and Substrates for Proteases and their Inhibitors

The positive outcome regarding the pharmacodynamic properties of covalent protease inhibitors, suggests promising results by selecting a suitable peptidomimetic scaffold for the target protein, which prevents possible off-target interactions with other structure-related enzymes. Furthermore, it is of great importance to choose a matching warhead towards the catalytically active amino acid, not just for securing high affinity properties, but to reduce possible side effects like haptenization or off-target binding with physiological nucleophiles.[141] Therefore, the first project represents an extensive systematic study regarding the influence of both structure elements and possible combinatorial effects in terms of selectivity and reactivity.

An assortment of various peptidomimetic scaffolds, in combination with different reversible- and irreversible-reacting warheads were synthesized, targeting different proteases. Based on fluorometric inhibition studies (1.8), the impact of the peptidomimetic scaffold and respective warhead, regarding the affinity and selectivity profiles were evaluated. Additionally, reactivity studies with tool compounds containing the used warheads in the *in vitro* study, reacting with model nucleophiles in solution as well as docking studies and quantum mechanics simulations allowed further characterization of both structure motives.

Fluorometric inhibition studies represent an important investigation tool for evaluation of protease inhibitors regarding affinity and selectivity profiles (1.8). Hence, it is of great interest to minimize possible interferences, that originate from different sources and lead to false information and overall failing development of potential drug candidates (1.8.4).[305] The second part of this project highlights possible errors during fluorometric inhibition studies and presents suggestions for improving the vulnerability against assay interferences.

The following publications are part of the results obtained in this project:

(1) **Patrick Müller**[†], Mergim Meta[†], Jan Laurenz Meidner[†], Marvin Schwickert, Jessica Meyr, Kevin Schwickert, Christian Kersten, Collin Zimmer, Stefan Josef Hammerschmidt, Ariane Frey, Albin Lahu, Sergio de la Hoz-Rodriguez, Laura Agost-Beltrán, Santiago Rodriguez, Kira Diemer, Wilhelm Neumann, Florenci V. Gonzàlez, Bernd Engels and Tanja Schirmeister[*]. "Investigation of the Compatibility between Warheads and Peptidomimetic Sequences of Protease Inhibitors—A Comprehensive Reactivity and Selectivity Study" *International Journal of Molecular Sciences* **2023**, *24*(8), 7226; https://doi.org/10.3390/ijms24087226.

Own contributions: Inhibitor synthesis (compounds 3, 4, 6, 7, 8, 9, 25, 51, 52, 65, 99, 101–107, 109–111), uPA and rhodesain inhibition assays, NMR-reactivity tests, LC-MS analysis, molecular docking and writing of the original draft plus editing of the manuscript.

Contribution from other authors: Inhibitor synthesis (compounds 11–22, 26–96, 112–117), cathepsin S, proteasome β5-subunit, and M^{pro} inhibition assays, LC-MS reactivity tests, molecular docking, quantum mechanics and kinetic simulations and writing of the original draft plus editing of the manuscript.

(2) Hannah Maus[†], **Patrick Müller**[†], Mergim Meta, Sabrina N. Hoba, Stefan Josef Hammerschmidt, Robert A. Zimmermann, Collin Zimmer, Natalie Fuchs, Tanja Schirmeister and Fabian Barthels[*]. "Next Generation of Fluorometric Protease Assays: 7-Nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-Amides) as Class-Spanning Protease Substrates" *Chemistry − A European Journal* **2023**, *29*(50), 1−11; doi.org/10.1002/chem.202301855.

Own contributions: Substrate synthesis (2a–j, 4, 6a–d, 7a–j, 18–20) and writing of the original draft plus editing of the manuscript.

Contribution from other authors: Inhibitor synthesis, protein expression and purification, substrate stability assays, molecular docking, kinetic characterization of substrates, investigations on assay interferences and writing of the original draft plus editing of the manuscript.

2.2 Chapter B: Evaluation of Covalent Inhibitors for Trypsin-Like Serine Proteases

This part of the doctoral thesis describes the development of covalent inhibitors targeting the trypsin-like serine proteases TMPRSS2, uPA and hepsin.

The important physiological role (1.2.2) and their pathological association (1.2.3) highlight the possibility for treatment of various diseases by targeting trypsin-like serine proteases.[80],[161],[244] The transmembrane serine protease subtype 2 (TMPRSS2) cleaves the SARS-CoV-2 spike protein and enables viral cell entry (1.5.2). Hepsin presents another type of transmembrane serine protease and leads if upregulated to an increase of growth factor release, which supports tumor progression in numerous cancer-types (1.6.1).[307] The first two projects aim to develop potent peptidomimetic covalent inhibitors targeting TMPRSS2 and hepsin. *In silico* studies by molecular docking and virtual screenings were executed to identify suitable peptidomimetic recognition sequences, which were coupled with different warheads and synthesized. Afterwards, the inhibitors were evaluated with fluorometric inhibition assays (1.8) regarding their affinity and selectivity. Additionally, stability and cytotoxicity tests as well as specific cell tests concerning the viral cell entry of the coronavirus were performed.

Based on the results of the previously described work, new peptidomimetic covalent uPA inhibitors were designed and evaluated in the last project. Furthermore, due to a high structure similarity between uPA and TMPRSS2, structure-activity relationship (SAR-) studies were executed with the abovementioned designed TMPRSS2 inhibitors for further optimization.

The following publications are part of the results obtained in this project:

(3) Lukas Wettstein[†], Philip Maximilian Knaff[†], Christian Kersten[†], **Patrick Müller**, Tatjana Weil, Carina Conzelmann, Janis A. Müller, Maximilian Brückner, Markus Hoffmann, Stefan Pöhlmann, Tanja Schirmeister, Katharina Landfester, Jan Münch and Volker Mailänder[*]. "Peptidomimetic inhibitors of TMPRSS2 block SARS-CoV-2 infection in cell culture" *Communications Biology* **2022**, *5*, 681; https://doi.org/10.1038/s42003-022-03613-4.

Own contributions: Inhibitor synthesis (S1−S5) and writing of the original draft plus editing of the manuscript.

Contributions from other authors: Inhibitor synthesis (1−8), inhibition-, cytotoxicity- and stability assays, molecular docking, pseudovirus inhibition assay, SARS-CoV-2 inhibition assay and writing of the original draft plus editing of the manuscript.

(4) Philip Maximilian Knaff, **Patrick Müller**, Christian Kersten, Lukas Wettstein, Jan Münch, Katharina Landfester and Volker Mailänder*. "Structure-Based Design of High-Affinity and Selective Peptidomimetic Hepsin Inhibitors" *Biomacromolecules* **2022**, *23*, 2236−2242; https://doi.org/10.1021/acs.biomac.1c01011.

Own contributions: Inhibitor synthesis (S1−S5) and writing of the original draft plus editing of the manuscript.

Contributions from other authors: Inhibitor synthesis (1−10), inhibition-, cytotoxicity- and stability-assays, molecular docking and writing of the original draft plus editing of the manuscript.

(5) **Patrick Müller**, Collin Zimmer, Ariane Frey, Gideon Holzmann, Annabelle Carolin Weldert and Tanja Schirmeister*. „Ligand-Based Design of Selective Peptidomimetic uPA and TMPRSS2 Inhibitors with Arg Bioisosteres" *International Journal of Molecular Sciences* **2024**, *25*(3), 1375; https://doi.org/10.3390/ijms25031375.

Own contributions: Inhibitor synthesis (3, 5, 9, 11, 12a, 13, 14a, 15, 21−45), inhibition assay and writing of the original draft plus editing of the manuscript.

Contributions of others: Inhibitor synthesis (6b−e, 10b−f, 12b−f, 14b−f, 17−20), parallel artificial membrane permeation assay (PAMPA), absorption spectra, calculation of physicochemical properties, calculation of protein similarity, stability studies and writing of the original draft plus editing of the manuscript.

3 List of Publications

3.1 Publications as Part of this Doctoral Thesis

3.1.1 Chapter A: Investigation and Characterization of Warheads and Substrates for Proteases and their Inhibitors

(1) **Patrick Müller**[†], Mergim Meta[†], Jan Laurenz Meidner[†], Marvin Schwickert, Jessica Meyr, Kevin Schwickert, Christian Kersten, Collin Zimmer, Stefan Josef Hammerschmidt, Ariane Frey, Albin Lahu, Sergio de la Hoz-Rodriguez, Laura Agost-Beltrán, Santiago Rodriguez, Kira Diemer, Wilhelm Neumann, Florenci V. Gonzàlez, Bernd Engels and Tanja Schirmeister[*]. "Investigation of the Compatibility between Warheads and Peptidomimetic Sequences of Protease Inhibitors—A Comprehensive Reactivity and Selectivity Study" *International Journal of Molecular. Sciences* **2023**, *24* (8), 7226; https://doi.org/10.3390/ijms24087226.

(2) Hannah Maus[†], **Patrick Müller**[†], Mergim Meta, Sabrina N. Hoba, Stefan Josef Hammerschmidt, Robert A. Zimmermann, Collin Zimmer, Natalie Fuchs, Tanja Schirmeister and Fabian Barthels[*]. "Next Generation of Fluorometric Protease Assays: 7-Nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-Amides) as Class-Spanning Protease Substrates" *Chemistry − A European Journal* **2023**, *29* (50), 1−11; doi.org/10.1002/chem.202301855.

3.1.2 Chapter B: Evaluation of Covalent Inhibitors for Trypsin-Like Serine Proteases

(3) Lukas Wettstein[†], Philip Maximilian Knaff[†], Christian Kersten[†], **Patrick Müller**, Tatjana Weil, Carina Conzelmann, Janis A Müller, Maximilian Brückner, Markus Hoffmann, Stefan Pöhlmann, Tanja Schirmeister, Katharina Landfester, Jan Münch and Volker Mailänder[*]. "Peptidomimetic inhibitors of TMPRSS2 block SARS-CoV-2 infection in cell culture" *Communications Biology* **2022**, *5*, 681; https://doi.org/10.1038/s42003-022-03613-4.

(4) Philip Maximilian Knaff, **Patrick Müller**, Christian Kersten, Lukas Wettstein, Jan Münch, Katharina Landfester and Volker Mailänder[*]. "Structure-Based Design of High-Affinity and Selective Peptidomimetic Hepsin Inhibitors" *Biomacromolecules* **2022**, *23*, 2236−2242; https://doi.org/10.1021/acs.biomac.1c01011.

(5) **Patrick Müller**, Collin Zimmer, Ariane Frey, Gideon Holzmann, Annabelle Carolin Weldert and Tanja Schirmeister[*]. „Ligand-Based Design of Selective Peptidomimetic uPA and TMPRSS2 Inhibitors with Arg Bioisosteres" *International Journal of Molecular Sciences* **2024**, *25*(3), 1375; https://doi.org/10.3390/ijms25031375.

3.2 Publications beyond this Doctoral Thesis

3.2.1 Research Article

(6) Armin Welker, Christian Kersten, Chrsitin Müller, Ramakanth Madhugiri, Collin Zimmer, **Patrick Müller**, Robert A. Zimmermann, Stefan J. Hammerschmidt, Hannah Mus, John Ziebuhr, Christoph Sotriffer, Tanja Schirmeister. "Structure-Activity Relationships of Benzamides and Isoindolines Designed as SARS-CoV Protease Inhibitors Effective against SARS-CoV-2" *ChemMedChem* **2021**, 16 (2): 340–354; doi: 10.1002/cmdc.202000548.

3.2.2 Review Article

(7) **Patrick Müller**‡, Hannah Maus‡, Stefan J. Hammerschmidt‡, Philip M. Knaff, Volker Mailänder, Tanja Schirmeister, Christian Kersten. "Interfering with Host Proteases in SARS-CoV-2 Entry as a Promising Therapeutic Strategy" *Current Medicinal Chemistry* **2022**, 29 (4): 635–665, doi: 10.2174/0929867328666210526111318.

3.2.3 Other Publication

(8) Stefan Josef Hammerschmidt, **Patrick Müller**, Tanja Schirmeister. "SARS-CoV-PLpro-Inhibitoren als mögliche Breitspektrum-Virostatika" *BIOspektrum* **2021**; 27(3):254–256. https://doi.org/10.1007/s12268-021-1576-6.

4 Chapter A: Investigation and Characterization of Warheads and Substrates for Proteases and their Inhibitors

4.1 Investigation of the Compatibility between Warheads and Peptidomimetic Sequences of Protease Inhibitors—A Comprehensive Reactivity and Selectivity Study

4.1.1 Context, Project Summary, and own Contributions

Numerous SAR-campaigns aim for discovery of new potential lead structures and clinical candidates, by designing and synthesizing new protease inhibitors. Covalent peptidomimetic inhibitors gained a lot of attention in drug development in recent years (1.3.2, 1.3.3). Selecting a suitable peptidomimetic recognition sequence, paired with an appropriate electrophilic warhead led to promising drug candidates in the past with enhanced pharmacodynamic properties, but can also display toxicity risks, due to non-selective off-target protein binding.[141]

Therefore, we investigated in an extensive systematic study the influence of peptidomimetic scaffolds and warheads as well as possible combinatorial effects in different experiments regarding their pharmacodynamic properties.

We synthesized covalent reversible and irreversible inhibitors with peptidomimetic recognition sequences suited for five different proteases that comprises serine-, threonine- and cysteine catalytic mechanisms. Seven warheads were chosen with different specificities for cysteine, and serine/threonine proteases. Afterwards, the synthesized compounds were tested against all five proteases to get a deeper insight into their affinity and selectivity profile. The *in vitro* results highlighted the crucial role of the peptidomimetic backbone regarding the selectivity properties, displaying mostly affinity of inhibitors to their targeted enzyme, with exception of structurally similar proteases, that showed cross-reactivity of the respective inhibitors. Choosing a matching warhead for the target nucleophile of the proteases resulted in high affinity inhibitors against the target enzyme. Furthermore, for some proteases it was demonstrated, that only the right combination of a specific warhead and peptidomimetic recognition sequence displayed inhibitory activity at all. Molecular docking with the synthesized inhibitors and tested proteases could confirm the observed activity of the compounds, by finding the warheads in close proximity to the catalytic amino acid and the peptidomimetic recognition sequences protruding in the expected subpockets of the enzymes.

Within the reactivity test system, containing model nucleophiles and tool compounds with the used electrophilic warheads of the inhibitors, the chemoselectivity were investigated by NMR and LC-MS measurements and additionally explained with quantum mechanics simulations. The results could

partially confirm the observed enzyme inhibition assay data with the detected reactivity between model nucleophile and warhead in solution. Nevertheless, some differences could be observed, which may have occurred by using different conditions in the chemical test system and biochemical *in vitro* studies. Overall, the extensive study highlights the crucial role of the peptidomimetic sequence and warheads to the selectivity and affinity profile, as well as characterized the properties of the used electrophilic traps, which help the scientific community for developing new lead structures and potential drug candidates (Figure 28).

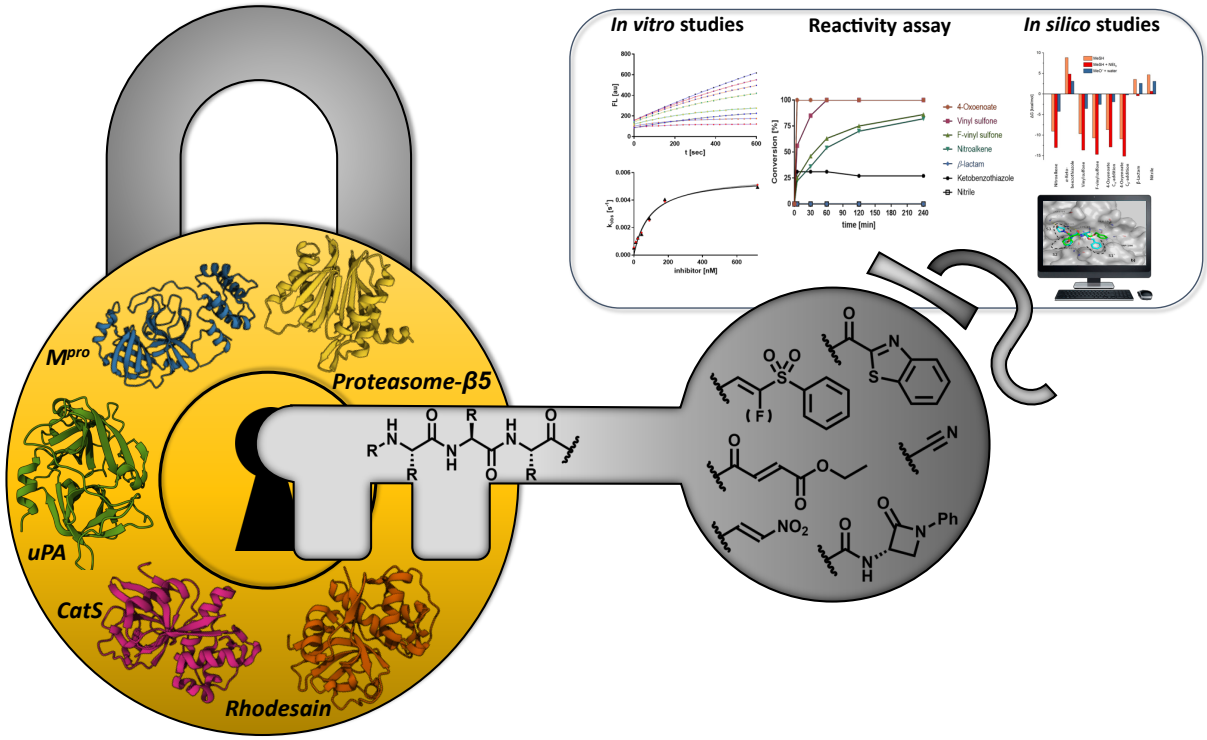


Figure 28: Investigation and characterization of covalent protease inhibitors regarding their pharmacodynamic properties. *In vitro* studies towards five different proteases were performed of the synthesized inhibitors, consisting of different peptidomimetic sequences and warheads (*α*-ketobenzothiazole, (F-)vinylsulfones, *α*-oxoenoate, nitrile, nitroalkene and *β*-lactam). Additional reactivity assays and *in silico* studies for a deeper insight of the inhibitor properties were executed. Figure prepared with PowerPoint.

Own contributions: Inhibitor synthesis (compounds 3, 4, 6–9, 25, 51, 52, 99, 101–107, 109–111), uPA, rhodesain inhibition assay, NMR-reactivity tests, molecular docking and writing of the original draft plus editing of the manuscript.

Contribution from other authors: Inhibitor synthesis (compounds 11–22, 26–50, 53–98, 112–117), cathepsin S, proteasome *β*5-subunit, M^{pro} inhibition assay, LC-MS reactivity tests, molecular docking, quantum mechanics and kinetic simulations and writing of the original draft plus editing of the manuscript.

This work has been published in International Journal of Molecular Sciences (impact factor 5.60).

The appended Supporting Information represents an abridged version. The full version can be accessed online at https://www.mdpi.com/1422-0067/24/8/7226.

4.1.2 Publication

The following publication quoted within “” from page 58 to page 123 is the same as the manuscript cited on page 53.

“

International Journal of
Molecular Sciences

MDPI

Article

Investigation of the Compatibility between Warheads and Peptidomimetic Sequences of Protease Inhibitors—A Comprehensive Reactivity and Selectivity Study

Patrick Müller [1,†], Mergim Meta [1,†], Jan Laurenz Meidner [1,†], Marvin Schwickert [1], Jessica Meyr [2], Kevin Schwickert [1], Christian Kersten [1], Collin Zimmer [1], Stefan Josef Hammerschmidt [1], Ariane Frey [1], Albin Lahu [1], Sergio de la Hoz-Rodríguez [3], Laura Agost-Beltrán [3], Santiago Rodríguez [3], Kira Diemer [2], Wilhelm Neumann [2], Florenci V. Gonzàlez [3], Bernd Engels [2] and Tanja Schirmeister [1,*]

[1] Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Staudinger Weg 5, D-55128 Mainz, Germany
[2] Institute of Physical and Theoretical Chemistry, Julius-Maximilians-University of Wuerzburg, Emil-Fischer-Straße 42 Süd, D-97074 Wuerzburg, Germany
[3] Departament de Química Inorgànica i Orgànica, Universitat Jaume I, 12080 Castelló de la Pana, Spain
* Correspondence: schirmei@uni-mainz.de; Tel.: +49-6131-39-25742
† These authors contributed equally to this work.

Abstract: Covalent peptidomimetic protease inhibitors have gained a lot of attention in drug development in recent years. They are designed to covalently bind the catalytically active amino acids through electrophilic groups called warheads. Covalent inhibition has an advantage in terms of pharmacodynamic properties but can also bear toxicity risks due to non-selective off-target protein binding. Therefore, the right combination of a reactive warhead with a well-suited peptidomimetic sequence is of great importance. Herein, the selectivities of well-known warheads combined with peptidomimetic sequences suited for five different proteases were investigated, highlighting the impact of both structure parts (warhead and peptidomimetic sequence) for affinity and selectivity. Molecular docking gave insights into the predicted binding modes of the inhibitors inside the binding pockets of the different enzymes. Moreover, the warheads were investigated by NMR and LC-MS reactivity assays against serine/threonine and cysteine nucleophile models, as well as by quantum mechanics simulations.

Keywords: covalent inhibitors; in vitro study; protease inhibitors; peptidomimetic sequence; warhead; reactivity and selectivity study

Citation: Müller, P.; Meta, M.; Meidner, J.L.; Schwickert, M.; Meyr, J.; Schwickert, K.; Kersten, C.; Zimmer, C.; Hammerschmidt, S.J.; Frey, A.; et al. Investigation of the Compatibility between Warheads and Peptidomimetic Sequences of Protease Inhibitors—A Comprehensive Reactivity and Selectivity Study. *Int. J. Mol. Sci.* **2023**, *24*, 7226. https://doi.org/10.3390/ijms24087226

Academic Editor: Raffaele Capasso

Received: 24 March 2023
Revised: 6 April 2023
Accepted: 10 April 2023
Published: 13 April 2023

1. Introduction

The human organism expresses about 600 different proteases falling into five different catalytic classes: aspartic, cysteine, metallo, serine and threonine proteases [1,2]. With their ability to catalyze irreversible protein hydrolysis, these members of the degradome manage the functions of many proteins through various mechanisms, such as activating or inactivating, e.g., growth factors, cytokines and other enzymes. As a result, they play an important role in physiological and developmental processes. These include DNA replication, cell proliferation and differentiation, but also tissue remodeling and neuronal outgrowth [3,4]. Due to their essential roles in such vital processes, dysregulation of these proteins causes severe pathologic conditions, such as cancer and neurodegenerative or cardiovascular disorders [5,6]. Furthermore, proteases play a key role in infectious diseases of, for example, parasitic or viral origin. African trypanosomiasis, also called sleeping sickness, and Chagas disease are caused by parasites and are classified as neglected tropical diseases and constitute important health issues in Latin American and Sub-Saharan African countries. For both diseases, proteases have been identified, which are essential for the

development of the parasites and the progression of the disease [7,8]. The 2019–2020 coronavirus (SARS-CoV-2) outbreak is the most recent example of a viral disease with global impact and burden. The viral replication and spreading is associated with proteases playing crucial roles in the viral life cycle, turning them into valid targets for the design of new anti-infectives [9,10].

Over the course of time, various protease inhibitors have been discovered either by targeted design or serendipity. Depending on the target binding site and inhibition mechanism, the molecular structures vary significantly. These range from small molecules to macrocyclic drugs and from non-covalent to covalent inhibition types [11–13]. Until recently, covalent modifiers which consist of an electrophilic trap (warhead) were controversially discussed as therapeutics due to the possibility of unselective reactions with off-target proteins and associated immunogenicity and toxicity. These compounds are emerging as potential drugs due to various inherent advantages, such as longer residence times and an accompanying lower drug dosage necessary for effective therapy [14]. There are many covalent drugs that have been approved, including some protease inhibitors, such as the proteasome inhibitors bortezomib or carfilzomib, for treatment of multiple myeloma, which inhibit the proteasome's β5-subunit in an irreversible manner, due to the permanent covalent bond to the catalytically active Thr-1. On the other hand, the nitriles saxagliptin and vildagliptin for treatment of type 2 diabetes and the recent first-approved cysteine protease inhibitor nirmatrelvir for treatment of COVID-19 bind covalent-reversibly to their target proteases, due to the decomposition of the (thio)-imidate adduct formed between the inhibitor and the amino acid of the protease, which is preferable since covalent-reversible inhibition leads to a lower risk of haptenization and binding to off-targets [15–17].

The binding of such covalent protease inhibitors proceeds in two stages. A peptidic or peptidomimetic recognition sequence is mainly responsible for the non-covalent interactions (first step) with the substrate binding pockets. It mainly determines the selectivity profile of the inhibitor towards the protease of interest, due to polar and non-polar interactions between the peptidic residues and the enzyme sub pockets. In the second step, the reaction between the warhead and an active site amino acid residue leads to the formation of a covalent bond, either reversibly or irreversibly, between the drug and the enzyme. This step mainly determines the affinity of the inhibitor to the target protease [14,18]. However, the warhead must be suitable for the respective nucleophilic amino acid residue in the active site. Depending on the type of nucleophile, different warheads can be used to target thiol or hydroxy groups of amino acid residues. Functional groups, such as β-lactams, but also boronic acids, which are all considered hard electrophiles with regard to the HSAB theory, are warheads targeting mainly serine and threonine-based proteases. Unsaturated, vinylogous Michael-acceptor-like structures, which are considered soft electrophiles, preferably react with cysteine proteases [18–21]. There are also warheads, e.g., ketones, aldehydes and nitriles, that are similarly suitable for serine-, threonine- and cysteine-based proteases [22–25]. Thus, exchanging the warhead can lead to different reactivity and affinity profiles, and alterations to the peptidomimetic/peptidic sequence may affect the selectivity of an inhibitor.

Within this extensive systematic study, we selected peptidomimetic sequences specifically to ensure a high affinity towards the protease of interest, which will be discussed below. We collected information about different kinds of warheads regarding their electrophilic properties and inhibition mechanisms to obtain a well-balanced assortment to potentially target cysteine and serine-/threonine proteases and combined them with the sequences (Figure 1) [18,21,22,24]. In vitro testing of all inhibitors on every target, first with the suited peptidomimetic sequence with differing warheads for their on-target and afterwards towards the off-target proteases, revealed the impact of the peptidomimetic sequences and the warheads on affinity and selectivity. The results indicate that, depending on the protease, every tested warhead behaved differently. The experimental results were compared with molecular docking results, visualizing putative binding modes in order to achieve a better understanding of the characteristics of the tested compounds.

Additionally, a reactivity study was carried out using model compounds containing the seven different warhead types, which were reacted with hydroxy and thiol model nucleophiles representing serine, threonine and cysteine proteases. Quantum mechanical computations of the reactions between the warheads and model nucleophiles were used to explain the experimental reactivity test data. These data highlight the preference of the warheads for specific active site residues.

To our knowledge, this is the first systematic study of this extent to evaluate the inhibition properties of peptidomimetic inhibitors with different warheads described in the literature, including in vitro testing towards a series of selected proteases, reactivity tests of the warheads in solution with model nucleophiles and in silico studies (docking and quantum mechanics and kinetic simulations) to explain the experimentally obtained data.

For our studies, the urokinase-type plasminogen activator (uPA) was chosen as a serine, the β5-subunit of the proteasome as a threonine and human cathepsin S (CatS), SARS-CoV-2 main protease (M^{pro}) and *T. brucei* rhodesain (TbCatL) as representatives of cysteine proteases.

The uPA belongs to the trypsin-like serine protease superfamily and contains a catalytic triad consisting of Ser195, His57 and Asp102 [26]. The enzyme is involved in several physiological functions, such as the degradation of the extracellular matrix (ECM), cell migration and thrombolysis [27,28]. Dysregulation of the uPA is involved in the metastasis of several cancer species [29]. We chose Ac–(L)Gly–(L)Thr–(L)Ala–(L)Arg–(warhead) as the specific peptidomimetic sequence for the uPA-inhibitors because of its high selectivity, which has been reported in the literature [30].

The 20S proteasome is responsible for most of the protein degradation in cells but can also lead to cancer by dysfunction [31]. It consists of three β-subunits (β1, β2 and β5), each containing a catalytic threonine. Here, we focus on the β5-subunit with the catalytic triad Thr1, Lys33 and Asp17, as it has the greatest impact on the proteolytic activity of the 20S proteasome. We selected the peptidomimetic sequence of bortezomib Pyz–(L)Phe–(L)Leu–(warhead) because of its clinically proven properties as a potent drug [32].

As cysteine proteases, we chose CatS, M^{pro} and rhodesain. Since CatS and rhodesain are both members of the papain family, they would allow a closer examination of the selectivity of the tested inhibitors towards related proteases [33]. CatS contains a catalytic dyad consisting of Cys25 and His164 [34]. It is partly tethered at the cell surface and involved in tissue remodeling, which can lead to cancer cell growth and spreading [35]. We utilized the peptidomimetic sequence morpholine–(L)cyAla–(L)Ser(OBn)–(warhead) which has been reported in the literature because of its described affinity and selectivity properties [36].

In contrast to the aforementioned proteases, rhodesain and M^{pro} do not originate from the human organism but play significant roles in the progression of infectious diseases. Rhodesain is essential for the development of the parasite *Trypanosoma brucei rhodesiense*, which is responsible for the sleeping sickness "Human African Trypanosomiasis". Analogously to CatS, it contains a catalytic dyad consisting of Cys25 and His159 [37]. There are various peptidomimetic sequences that have been published for rhodesain inhibitors. We decided to utilize Cbz–(L)Phe–(L)*h*Phe–(warhead), as it is a commonly used sequence with great affinity and selectivity [38]. M^{pro} originates from SARS-CoV-2 and plays a key role in the virus replication. The active site contains Cys145 and His164 as a catalytic dyad [39]. Similar to the newly published M^{pro} inhibitors, we chose 4-(OMe)-1*H*-indole–(L)Leu–3-[(3*S*)-2-oxopyrrolidin-3-yl]-(L)Ala–(warhead) as the general structure [40]. All peptidomimetic sequences and warheads are illustrated in Figure 1.

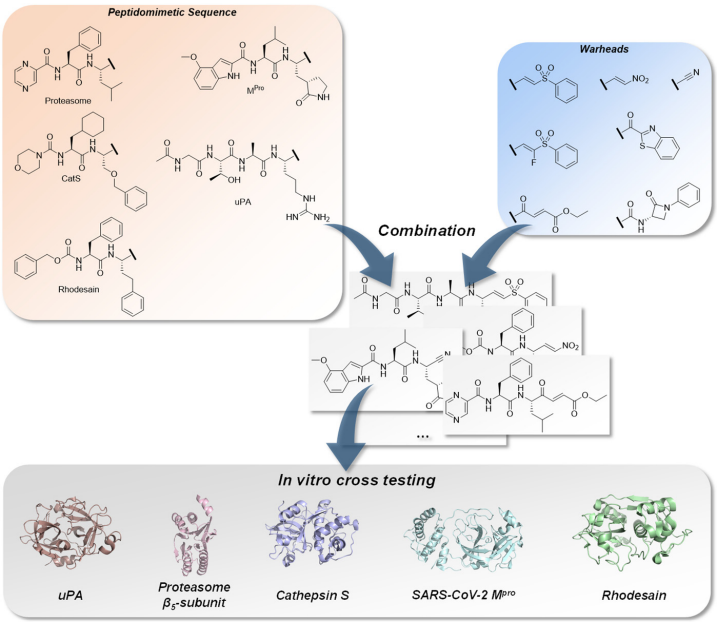


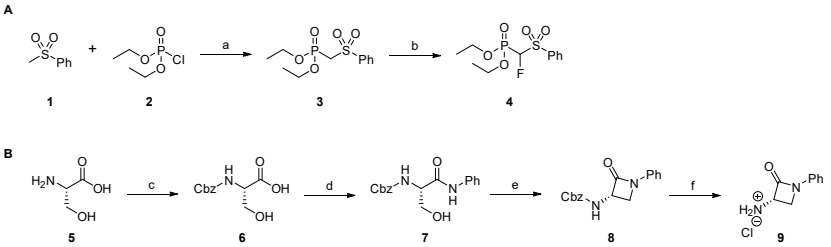
Figure 1. Combination of characteristic peptidomimetic inhibitor sequences for the targets: urokinase-type plasminogen activator (uPA), PDB-ID: 1W10 [41], proteasome β5-subunit, PDB-ID: 5LF3 [42], cathepsin S, PDB-ID: 1MS6 [43], SARS-CoV-2 main protease (M^{pro}), PDB-ID: 6XR3 [44] and rhodesain, PDB-ID: 2P7U [45], with selected warheads (vinyl sulfone, F-vinyl sulfone, nitroalkene, α-ketobenzothiazole, 4-oxoenoate, nitrile and β-lactam). The resulting compounds were tested on each target to determine affinity and selectivity.

2. Results

2.1. Chemistry

2.1.1. Synthesis of Precursors

All tested substances were synthesized in multi-step reactions [19,21]. Regarding the synthesis of the (F-)vinyl sulfone and β-lactam compounds, the same precursor molecules were used repeatedly. The preparation of these precursors is shown in Scheme 1.



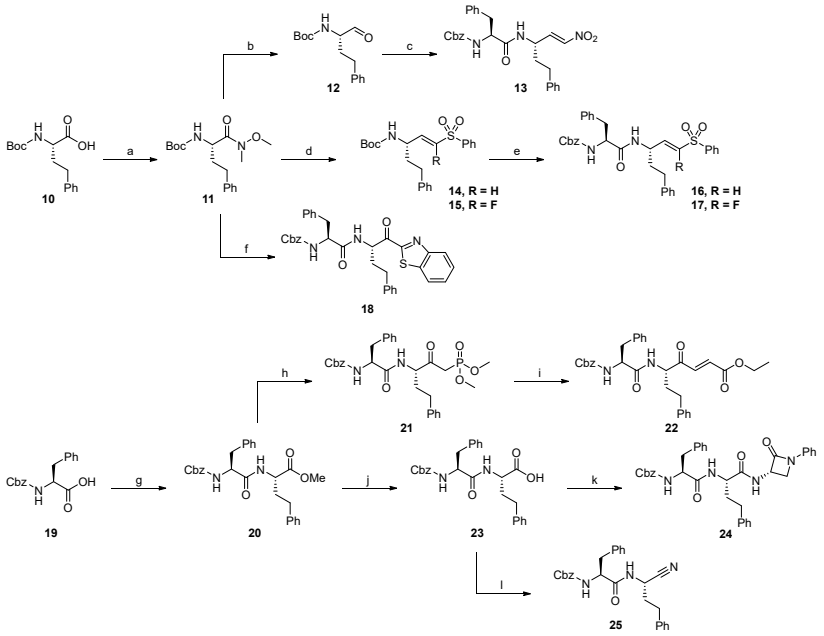
Scheme 1. Synthesis of precursor molecules. (**A**) Synthesis of phosphonate building block **4**. (**B**) Synthesis of β-Lactam building block **9**. Reaction conditions: (a) *n*-BuLi, DECP, THF, −78 °C; (b) **3**, LHMDS, Selectfluor®, THF, DMF, −78 °C, 3 h, 49%; (c) Cbz-Cl, $NaHCO_3$, H_2O, 12 h, rt, 90%; (d) aniline, TBTU, HOBt · H_2O, EtOAc, 12 h, rt, 74%; (e) $ImSO_2$, NaH, DMF, F20 °C, 1.5 h, 77%; (f) Pd/C, H_2, THF, 88%.

In a substitution reaction on diethyl chlorophosphate (DECP) **2** using methyl phenyl sulfone **1** and *n*-butyllithium (*n*-BuLi), the phosphonate **3** was prepared. Subsequent fluorination of **3** with Selectfluor® led to phosphonate **4**. These precursors were used for the synthesis of vinyl sulfone warheads.

The synthesis of the *β*-lactam precursor **9** was conducted from L-serine. Benzyloxycarbonyl (Cbz) protection followed by amide coupling of the free carboxylic acid moiety with aniline led to the intermediate **7**. The following cyclisation was performed using 1,1′-sulfonyldiimidazol ($ImSO_2$) and sodium hydride (NaH). Cbz deprotection with hydrogen and palladium on carbon (Pd/C) yielded precursor **9**.

2.1.2. Rhodesain Inhibitors

The synthesis of substances with the targeting structure designed for rhodesain was conducted according to Scheme 2.



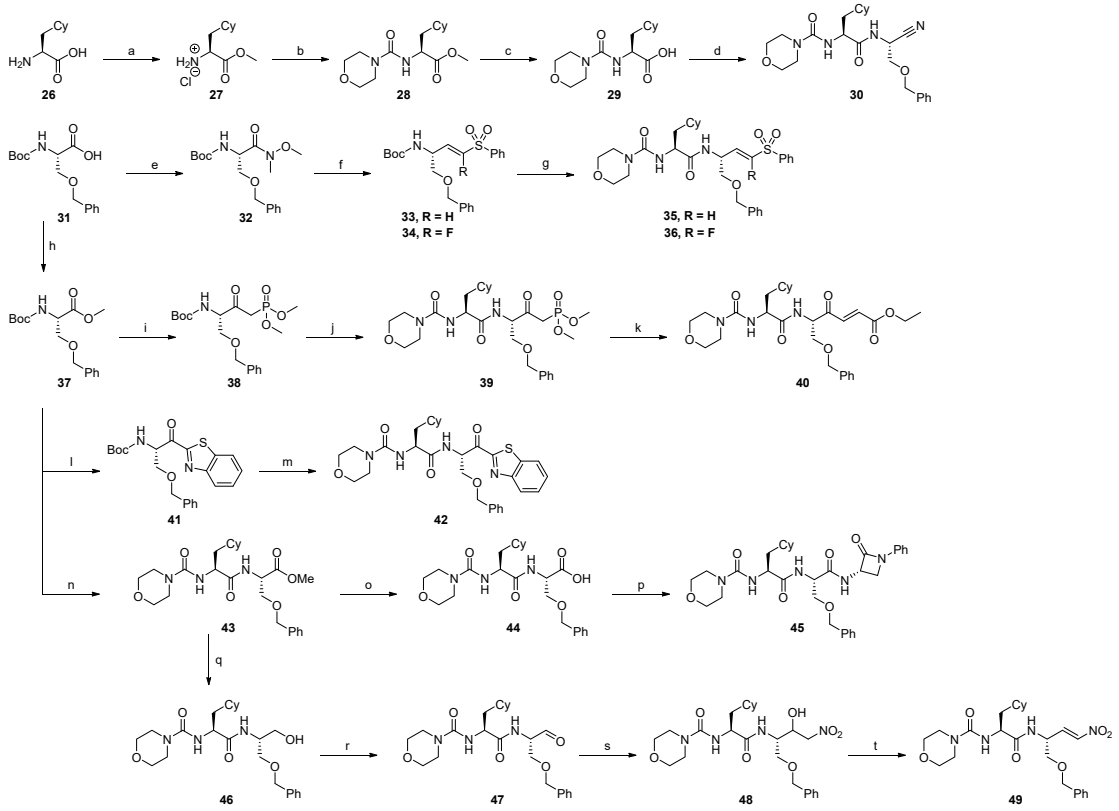
Scheme 2. Synthesis of rhodesain compounds. Reaction conditions: (a) *N*,*O*-dimethylhydroxylamine · HCl, DCC, HOBt · H_2O, DIPEA, THF, rt, 12 h, 46%; (b) LAH, THF, 0 °C, 2 h, 67%; (c) 1. $MeNO_2$, Et_3N, DCM, rt, 8 h, 2. TFA, DCM, rt, 0.5 h, 3. Cbz-(L)Phe-OH, EDC · HCl, TEA, DCM, rt, 12 h, 4. MsCl, DIPEA, DCM, rt, 2 h, 75%; (d) **3**/**4**, LiHMDS, THF, −80 °C, 12 h, 59% (**14**), 57% (**15**); (e) TFA, DCM, Cbz-(L)Phe-OH, T3P, DIPEA, DMF, rt, 12 h, 65% (**16**), 48% (**17**); (f) 1. benzothiazole, *n*-BuLi, THF, −78 °C, 3 h, 2. TFA, DCM, rt, 2 h, 3. Cbz-(L)Phe-OH, EDC · HCl, HOBt · H_2O, Et_3N, DCM, rt, 8 h, 56%; (g) H-(L)*h*Phe-OMe, HATU, 2,4,6-collidine, DCM/DMF, rt, 16 h, quant.; (h) DMMP, *n*-BuLi, THF, –70 °C, 2 h, 98%; (i) ethyl glyoxylate, K_2CO_3, EtOH, rt, 2 h, 76%; (j) LiOH, THF/H_2O, rt, 16 h, quant.; (k) **9**, HATU, 2,4,6-collidine, DCM/DMF, rt, 16 h, 67%; (l) 1. EDC · HCl, HOBt · H_2O, NH_4OH, DMF, rt, 12 h, 2. TFAA, pyridine, DMF, 0 °C, 0.1 h, 40%.

The first step of the synthesis of rhodesain inhibitors was the conversion of Boc-(L)*h*Phe-OH **10** into Weinreb amide **11**. From this intermediate, the nitroalkene inhibitor **13** was accessible by reduction to aldehyde **12** and subsequent Henry reaction followed by standard deprotection and amide coupling to connect the P2-P3 residues. In a similar way, the vinyl sulfone **16** and F-vinyl sulfone **17** were obtained, whereby the aldehyde **12** was used in a Horner–Wadsworth–Emmons (HWE) reaction with the precursors **3** and **4**

followed by the attachment of the P2-P3 residues. The α-ketobenzothiazole inhibitor **18** was prepared by alkylation of the Weinreb amide **11** with benzothiazole and subsequent attachment of the P2-P3 residues. Starting from Boc-(L)*h*Phe-OH, the methyl ester **20** was prepared by amide coupling. A following alkylation with dimethyl methylphosphonate (DMMP) and HWE reaction with ethyl glyoxylate led to the 4-oxoenoate **22**. For the synthesis of β-lactam **24**, hydrolysis of methyl ester **20** and amide coupling with precursor **9** yielded the desired product. Nitrile **25** was prepared from carboxylic acid **23** via amide coupling with ammonia followed by dehydration.

2.1.3. Cathepsin S Inhibitors

Compounds designed for the inhibition of cathepsin S were synthesized according to Scheme 3.



Scheme 3. Synthesis of cathepsin S compounds. Reaction conditions: (a) $SOCl_2$, MeOH, –10 °C, 16 h, 91%; (b) morpholine, triphosgene, $NaHCO_3$, $CHCl_2$, 0 °C; 16 h, 98%; (c) LiOH, THF/H_2O, 3 h, 97%; (d) 1. NaCN, NH_4Cl, NH_3, 2-(benzyloxy)acetaldehyde, Et_2O, 2. HATU, 2,4,6-collidine, DCM/DMF, rt, 16 h, 53%; (e) *N*,*O*-dimethylhydroxylamine · HCl, DCC, HOBt · H_2O, DIPEA, THF, –15–0°C, 16 h, 80%; (f) 1. LAH, Et_2O, 0 °C, 2 h, 2. **3**/**4**, KHMDS/ LHMDS, THF, –78 °C, 3 h; 75% (**33**), 44% (**34**); (g) 1. 4 N HCl in 1,4-dioxane, 2. **29**, HATU, collidine, DCM/DMF, rt, 16 h, 63% (**35**), 60% (**36**); (h) MeI, DMF,

0 °C, 16 h, 97%; (i) *n*-BuLi, DMMP, THF, –78 °C, 3 h, 79%; (j) 1. 4 N HCl in 1,4-dioxane, 2. **29**, HATU, 2,4,6-collidine, DCM/DMF, rt, 16 h, 52%; (k) LiCl, ethyl glyoxylate, DIPEA, MeCN, 0 °C, 2 h, 39%; (l) benzothiazole, *n*-BuLi, THF, –78 °C, 3 h, 38%; (m) 1. 4 N HCl in 1,4-dioxane, 2. **29** HATU, 2,4,6-collidine, DCM/DMF, rt, 16 h, 43%; (n) 1.4 N HCl in 1,4-dioxane, 2. **29**, HATU, collidine, DCM/DMF, rt, 16 h, 90%; (o) LiOH, THF/H_2O, rt, 3 h, 99%; (p) **9**, HATU, 2,4,6-collidine, DCM/DMF, rt, 16 h, 81%; (q) $NaBH_4$, MeOH, THF, 0 °C, 16 h, 91%; (r) Dess–Martin–Periodinan, DCM, rt, 16 h, 70%; (s) NaH, $MeNO_2$, THF, 0 °C, 1 h, 58%; (t) MsCl, Et_3N, DCM, 0 °C, 3 h, 45%.

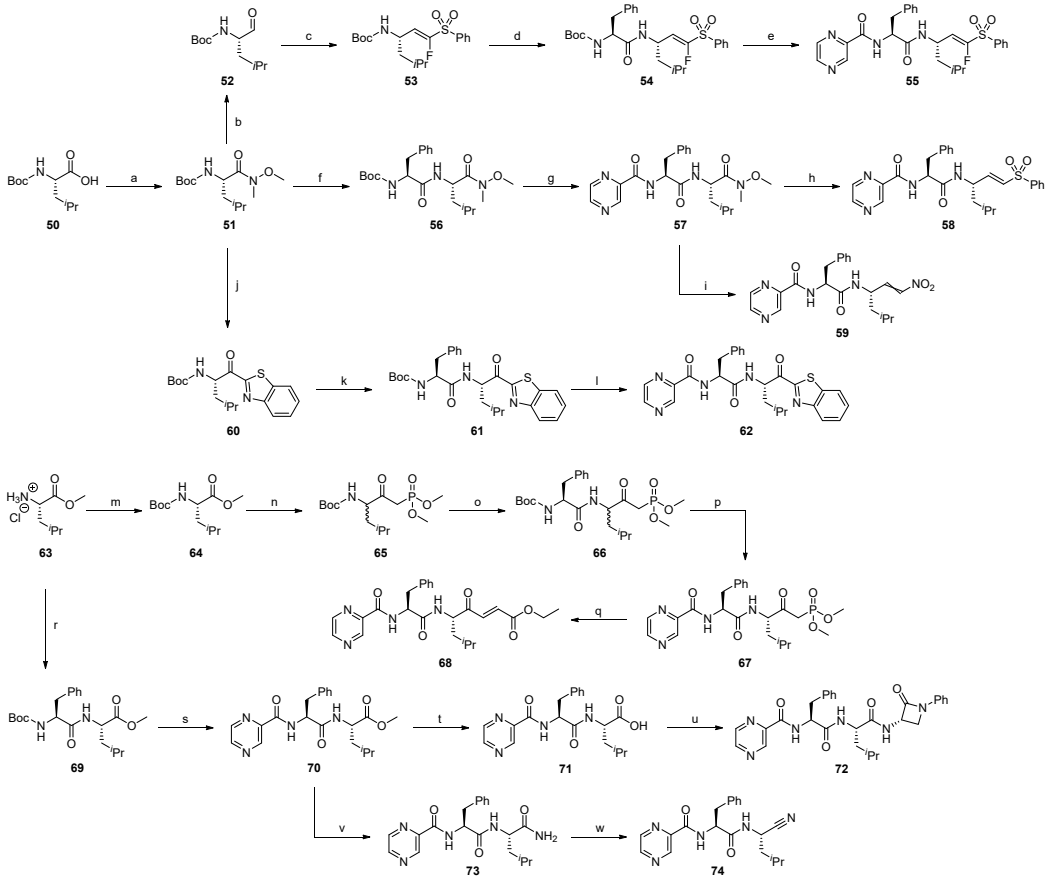
For the synthesis of cathepsin S inhibitors, the P2–P3 intermediate **29** was used repeatedly. It was prepared by attaching a morpholino-urea residue to cyclohexyl alanine **26** followed by hydrolysis of the methyl ester. In a direct conversion from **29**, the nitrile inhibitor **30** was prepared by amide coupling with ammonia and dehydration. From Boc-(L)Ser(OBn)-OH **31**, the vinyl sulfone **35** and F-vinyl sulfone **36** were obtained by conversion into Weinreb amide **32** followed by reduction, HWE reaction with the precursors **3** and **4** and subsequent standard deprotection and amide coupling with intermediate **29**. Boc-(L)Ser(OBn)-OH **31** was also converted into the methyl ester **37**, which was used for the synthesis of the 4-oxoenoate **40**. Therefore, an alkylation with DMMP and subsequent introduction of the P2 and P3 residues by deprotection and amide coupling led to the phosphonate intermediate **39**, which was converted into the desired product by HWE reaction with ethyl glyoxylate. The α-ketobenzothiazole **42** was prepared from methyl ester **37** in an alkylation reaction with benzothiazole and attachment of the P2-P3 residues by deprotection and amide coupling with intermediate **29**. Starting from methyl ester **37**, deprotection and amide coupling with intermediate **29** led to the methyl ester intermediate **43**, which was converted into the β-lactam **45** by hydrolysis and amide coupling with precursor **9**. The nitroalkene **49** also was prepared from methyl ester **43** by firstly converting it to the alcohol **46** and then to aldehyde **47**, which was used in a Henry reaction with nitromethane and subsequent dehydration.

2.1.4. Proteasome β5-Subunit Inhibitors

Compounds designed for the inhibition of the proteasome β5-subunit were synthesized according to Scheme 4.

The synthesis of proteasome β5-subunit targeting compounds started from Boc-(L)Leu-OH **50**, which was converted into the Weinreb amide **51**. From this, the F-vinyl sulfone **55** was prepared by reduction to aldehyde **52** and subsequent HWE reaction followed by a standard deprotection and amide coupling procedure connecting the P2 and P3 residues. For the vinyl sulfone inhibitor **58**, a different route was taken. First, the Weinreb amide intermediate **56** containing the P2 and P3 residues was prepared by standard amide coupling. Subsequent reduction and HWE reaction led to the desired inhibitor. The Weinreb amide **57** was also the intermediate for nitroalkene **59**, which was prepared by reduction and Henry reaction with subsequent dehydration. From Weinreb amide **51**, the α-ketobenzothiazole moiety was introduced by alkylation. The attachment of the P2 and P3 residues by standard deprotection and amide coupling yielded the α-ketobenzothiazole **62**.

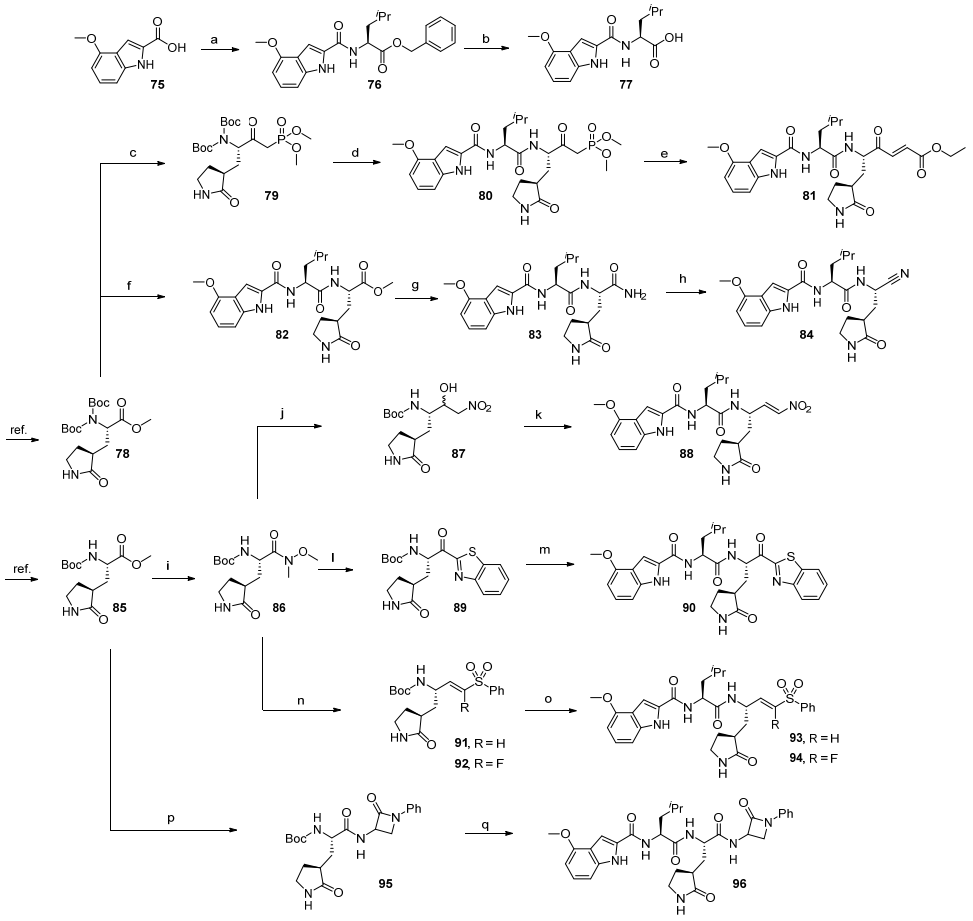
The 4-oxoenoate inhibitor **68** was prepared by HWE reaction of ethyl glyoxylate with the phosphonate intermediate **67**. The latter was synthesized by starting with the Boc protection of H-(L)Leu-OMe · HCl **63**, followed by alkylation of the methyl ester with DMMP and successive deprotection/amide coupling to introduce the P2 and P3 residues. In the same way, the introduction of the P2 and P3 residues to H-(L)Leu-OMe · HCl **63** led to the methyl ester intermediate **70**, from which the β-lactam **72** was prepared by hydrolysis and subsequent amide coupling with precursor **9**. Methyl ester **70** was also converted into the nitrile **74** by ammonolysis and dehydration.



Scheme 4. Synthesis of proteasome β5-subunit compounds. Reaction conditions: (a) *N*,*O*-dimethylhydroxylamine · HCl, TBTU, HOBt · H_2O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 86%; (b) LAH, THF, 0 °C, 30 min, 29%; (c) **4**, LiCl, DBU, MeCN, 0 °C, 1 h, 80%; (d) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. Boc-Phe-OH, TBTU, HOBt · H_2O, 2,4,6-collidine, DCM, 0 °C, 16 h, 97%; (e) 1. 4 N HCl in 1,4-dioxane, rt, 1.5 h, 2. pyrazinoic acid, TBTU, HOBt · H_2O, 2,4,6-collidine, DCM, 0 °C, 16 h, 67%; (f) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. Boc-(L)Phe-OH, TBTU, HOBt · H_2O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 97%; (g) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. pyrazinoic acid, TBTU, HOBt · H_2O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 88%; (h) 1. LAH, THF, 0 °C, 1 h, 2. **3**, LiCl, DBU, MeCN, 0 °C, 1.5 h, 11%; (i) 1. LAH, THF, 0 °C, 1 h, 2. $MeNO_2$ Et_3N, DCM, 0–20 °C, 16h, 3. MsCl, DIPEA, DCM, rt, 3 h, 19%; (j) benzothiazole, *n*-BuLi, THF, –78 °C, 6 h, 65%; (k) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. Boc-(L)Phe-OH, TBTU, HOBt · H_2O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 68%; (l) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. pyrazinoic acid, TBTU, HOBt · H_2O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 53%; (m) Boc_2O, $NaHCO_3$, water, 1,4-dioxane, 3 h, rt, 99%; (n) *n*-BuLi, THF, –78 °C, 6 h, 88%; (o) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. Boc-(L)Phe-OH, TBTU, HOBt · H_2O, 2,4,6-collidine, DCM, DMF, 0–20 °C, 16 h, 72%; (p) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. pyrazinoic acid, TBTU, HOBt · H_2O, 2,4,6-collidine, DCM, DMF, 0–20 °C, 16 h, 66%; (q) ethyl glyoxylate, LiCl, DIPEA, MeCN, 1 h, 0 °C, 79%; (r) Boc-(L)Phe-OH, TBTU, HOBt · H_2O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 82%; (s) 1. 4 N HCl in 1,4-dioxane, rt, 1 h, 2. pyrazinoic acid, TBTU, HOBt · H_2O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 82%; (t) LiOH, water, THF, rt, 17 h, quant.; (u) **9**, TBTU, HOBt · H_2O, 2,4,6-collidine, DCM, 0–20 °C, 16 h, 65%; (v) 7 N NH_3 in MeOH, rt, 48 h, 89%; (w) cyanuric chloride, DMF, 0 °C, 48 h, 47%.

2.1.5. SARS-CoV-2 M^{pro} Inhibitors

Compounds designed for the inhibition of SARS-CoV-2 M^{pro} were synthesized according to Scheme 5.

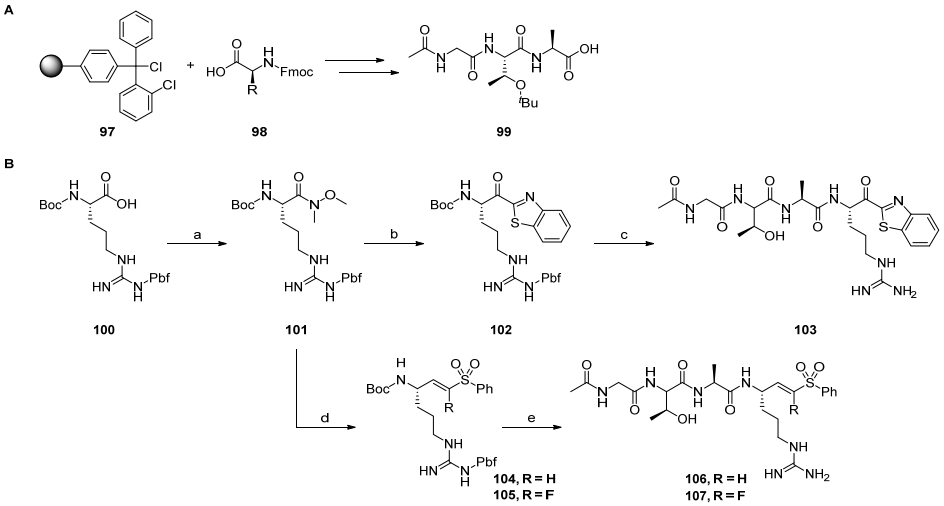


Scheme 5. Synthesis of SARS-CoV-2 M^{pro} compounds. Reaction conditions: (a) H-(L)Leu-OBn · *p*TsOH, TBTU, DIPEA, DCM, 0–20 °C, 16 h, 93%; (b) H_2, Pd/C (10%), EtOH, quant.; (c) DMMP, *n*-BuLi, THF, –78 °C, 5 h, 24%; (d) 1. TFA, DCM, 0 °C, 3 h, 2. **77**, EtOCOCl, NMM, THF, –20 °C, 2 h, 40%; (e) ethyl glyoxylate, LiCl, DIPEA, MeCN, 0 °C, 2 h, 47%; (f) 1. TFA, DCM, 0 °C, 3 h, 2. **77**, HATU, 2,4,6-collidine, DMF, 0–20 °C, 16 h, 83%; (g) 1. LiOH, THF/H_2O, 0–4 °C, 16 h, 2. NH_4OH, HATU, OxymaPure®, 2,4,6-colllidine, DMF, 0–20 °C, 16 h, 65%; (h) burgess reagent, DCM, rt, 2 h, 67%; (i) 1. LiOH, THF/MeOH/H_2O, 0–4 °C, 16 h, 2. *N*,*O*-dimethylhydroxylamine · HCl, HATU, 2,4,6-colllidine, DMF, 0–20 °C, 16 h, 76%; (j) 1. LAH, THF, −20 °C, 2 h, 2. $MeNO_2$, EtN_3, DCM, rt, 15 h, 71%; (k) 1. TFA, DCM, 0 °C, 3 h, 2. **77**, EDC · HCl, HOBt · H_2O, DIPEA, DCM, 0–20 °C, 16 h, 3. MsCl, DIPEA, DCM, 0–20 °C, 16 h, 50%; (l) benzothiazole, *n*-BuLi, THF, –78 °C, 5 h, 59%; (m) 1. TFA, DCM, 0 °C, 3 h, 2. **77**, EtOCOCl, NMM, THF, –20 °C, 2 h, 54%; (n) 1. LAH, THF, −20 °C, 2 h, 2. **3**/**4**, LiCl, DBU, MeCN, 0 °C, 2 h, 65% (**89**), 35% (**90**); (o) 1. TFA, DCM, 0 °C, 3 h, 2. **77**, EtOCOCl, NMM, THF, –20 °C, 2 h, 18% (**91**), 15% (**92**); (p) 1. LiOH, water, THF, 16 h, 2. **9**, TBTU, HOBt · H_2O, DIPEA, DCM, 0–20 °C, 48 h, 39%; (q) TFA, DCM, 0 °C, 3 h, 2. **77**, EtOCOCl, NMM, THF, –20 °C, 2 h, 80%.

Potential SARS-CoV-2 M^{pro} inhibitors were synthesized, starting from the rigidized glutamine analogs **78** and **85**, which had been prepared according to methods reported in the literature [46,47]. The P2–P3 residues fragment of the potential inhibitors was prepared by standard amide coupling with **75** and subsequent deprotection, yielding the intermediate **77**. From glutamine analog **78**, the 4-oxoenoate **81** was prepared by alkylation with DMMP and subsequent deprotection and amide coupling with **77** followed by HWE reaction with ethyl glyoxylate. Also starting from **78**, deprotection and amide coupling with **77** followed by hydrolysis and coupling with ammonia and subsequent dehydration yielded the nitrile inhibitor **84**. Starting with the preparation of Weinreb amide **86** from glutamine analog **85**, the nitroalkene **88** was accessible through reduction, a subsequent Henry reaction with nitromethane followed by dehydration and final deprotection and amide coupling with **77**. Introduction of the α-ketobenzothiazole moiety to **86** and connection of the P2–P3 residues by deprotection and coupling with **77** led to α-ketobenzothiazole **90**. Similarly, the reduction of **86** and HWE reaction with the precursors **3** and **4** and subsequent attachment of the P2–P3 residues yielded the vinyl sulfone **93** and F-vinyl sulfone **94**. For the β-lactam **96**, hydrolysis of **85** and amide coupling with precursor **9** followed by attachment of the P2–P3 residues yielded the desired product.

2.1.6. uPA Inhibitors

Compounds designed for the inhibition of the uPA were synthesized according to Scheme 6.



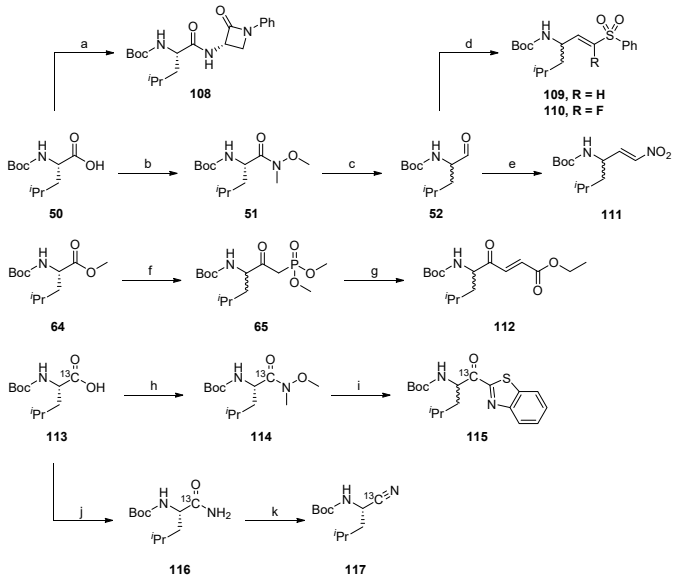
Scheme 6. Synthesis of the uPA compounds. (**A**) solid phase peptide synthesis of building block **99**. (**B**) combined synthesis of the final uPA compounds. Reaction conditions: (a) *N*,*O*-dimethylhydroxylamine· HCl, TBTU, DIPEA, DCM, rt, 12 h, 95%; (b) benzothiazole, *n*-BuLi, THF, –78 °C, 2 h, 76%; (c) 1. TFA, DCM, rt, 0.5 h, 2. HATU, DIPEA, DMF, DCM, rt, 12 h, 3. TFA, DCM, rt, 2 h, 10%; (d) 1. LAH, THF, 0 °C, 2. **3**/**4**, LiCl, DBU, MeCN, 0 °C, 1 h, 72% (**104**), 31% (**105**); (e) 1. TFA, DCM, rt, 0.5 h, 2. **99,** HATU, DIPEA, DMF, DCM, rt, 12 h, 3. TFA, DCM, rt, 2 h, 16% (**106**), 11% (**107**).

The potential uPa inhibitors are based on a peptide sequence which was synthesized via a standard Fmoc solid-phase peptide synthesis (SPPS) protocol. The obtained peptide **99** was coupled to the α-ketobenzothiazole intermediate **102**, which had been prepared from Boc- L)Arg(Pbf)-OH **100** by alkylation of its Weinreb amide with benzothiazole to yield the α-ketobenzothiazole **103**, after deprotection. The vinyl sulfone **106** and F-vinyl sulfone **107** were prepared by reduction of Weinreb amide **101**, followed by a subsequent

HWE reaction with the precursors **104** and **105**, which were then coupled with **99** and finally deprotected. The inhibitors with the β-lactam, nitrile and 4-oxoeonoate moiety were not synthetically accessible due to the acidic conditions for the Pbf-deprotection to obtain the final inhibitors.

2.1.7. Synthesis of Reactivity Probes

Substances designed for reactivity assay were synthesized according to Scheme 7.



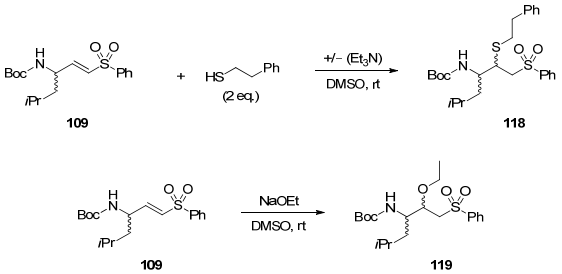
Scheme 7. Synthesis of the reactivity probes. Reaction conditions: (a) **9**, HATU, collidine, DCM, DMF, rt, 16 h, 97%; (b) TBTU, HOBt · H_2O, *N*,*O*-dimethylhydroxylamine · HCl, 2,4,6-collidine, rt, 12 h, quant.; (c) 1. LAH, Et_2O, 0 °C, 1 h, 89%; (d) **3**/**4**, LiCl, DBU, MeCN, 0 °C, 1 h, 50% (**109**), 30% (**110**); (e) 1. NaH, $MeNO_2$, THF, 0 °C, 1 h, 2. MsCl, Et_3N, DCM, 0 °C, 1 h, 15%; (f) DMMP, *n*-BuLi, THF, –78 °C, 1.5 h, 70%; (g) ethyl glyoxylate, LiCl, DBU, MeCN, 0 °C, 1 h, 45%; (h) TBTU, HOBt · H_2O, *N*,*O*-dimethylhydroxylamine · HCl, 2,4,6-collidine, 0–20 °C, 16 h, 95%; (i) benzothiazole, *n*-BuLi, THF, –78 °C, 3.5 h, 56%; (j) EDC · HCl, HOBt · H_2O, NH_4OH, DMF, rt, 16 h, 14%; (k) TFAA, pyridine, THF, –10 °C, 2 h, 74%.

For the synthesis of the reactivity probes, leucine was chosen as the model amino acid due to availability and to avoid side-chain reactivity. The different warheads were synthesized in the same way as described above for the full peptidic/peptidomimetic inhibitors. The β-lactam **108**, (F-)vinyl sulfone **109**, **110** and nitroalkene **111** reactivity probes were synthesized starting from Boc-(L)Leu-OH **50**, whereas the 4-oxoenoate **112** was prepared from Boc-(L)Leu-OMe **64**. Boc-(L)Leu-1-^{13}C-OH **113** was the starting material for the ^{13}C-labelled α-ketobenzothiazole **115** and nitrile **117** reactivity probes.

2.2. Reactivity Tests

To investigate the reactivity between the different warheads towards the three classes of proteases (serine, threonine and cysteine proteases), their behavior in model systems under the same reaction conditions (solvent, nucleophile and base) using either NMR or LC-MS analysis was investigated. We used reactivity probes with a Boc-L-Leu-(warhead) sequence. Leucin was chosen as a P1 amino acid to minimize influences of the side chain and due to synthetic accessibility. 2-Phenylethanethiol was used as a model nucleophile to mimic the thiol moiety of cysteine proteases, and sodium ethoxide was used as a ser-

ine/threonine replacement. DMSO-d_6 was used as solvent. Under these conditions, the nucleophile is deprotonated, simulating the activated serine or threonine in the catalytic triad of serine and threonine proteases, while ethanol as protonated alcohol species turned out to be unreactive in preliminary test reactions. The reactivity tests using 2-phenylethanthiol were carried out in the presence and absence of triethylamine as a base. This allowed for a reactivity comparison of the warheads towards protonated and deprotonated nucleophilic thiol species. Generating a deprotonated thiol species in the presence of triethylamine simulates the deprotonated cysteine in the catalytic dyad of cysteine proteases. Scheme 8 illustrates the reaction of the reactivity assay with both model nucleophiles and the vinyl sulfone moiety **109** as an example.



Scheme 8. Reaction scheme of the reactivity assay with both model nucleophiles under equal reaction conditions and the vinyl sulfone moiety **109**.

The reactivity tests of all Michael acceptors, **109**, **110**, **111** and **112**, the α-ketobenzothiazole **115** and nitrile **117** were investigated using an NMR-based analysis method, while the β-lactam **108** reactivity was investigated via LC-MS, due to its lack of proton signals, which could be used for evaluation of the reactivity in the ^{1}H-NMR studies, and the irreversible reaction mechanism, which allowed the LC-MS analysis. Additionally, LC-MS analyses of all reactions were performed in order to prove the formation of the expected reaction products. Formation of the expected adducts with the nitrile **117** (PhEtSH/PhEtS$^-$/EtONa), the α-ketobenzothiazole **115** (PhEtSH/PhEtS$^-$) and the nitroalkene **111** (EtONa) could not be observed. This may have been due to the covalent reversible reaction mechanism of the nitrile and α-ketobenzothiazole and the overall difficult ionization of the specific compounds by an electron spray ionization mass spectrometer.

Method A (NMR): ^{1}H-NMR spectra were recorded for the respective warhead and nucleophile mixture, before the addition of the nucleophile (0 min) and after 5, 30, 60, 120 and 240 min reaction time. For quantification, the double bond-signals (doublet/doublet of doublets, around 7.4–6.7 ppm) of the Michael acceptors were integrated relative to 1,3-dioxolane as an internal standard. The acetal CH_2 signal of the internal standard at 5.3 ppm was used as a reference.

The α-ketobenzothiazole **115** and nitrile **117** were similarly analyzed by ^{13}C-NMR. Therefore, the corresponding ^{13}C-leucin derivates were synthesized (Scheme 7). Quantifications of the reactions were carried out by using the integral of the carbonyl carbon atom signal at 195 ppm for the α-ketobenzothiazole and 120 ppm for the nitrile moiety. The reference signal of DMSO-d_6 was set to 39.52 ppm.

In Figure 2, the ^{1}H-NMR spectra of the test reaction of the 4-oxoenoate **112** with 2-phenylethanthiol **118** are shown exemplarily. After four hours, 92% conversion of the inhibitor to the product **120** was observed. LC-MS analysis confirmed the diastereomeric formation of the expected product **120**.

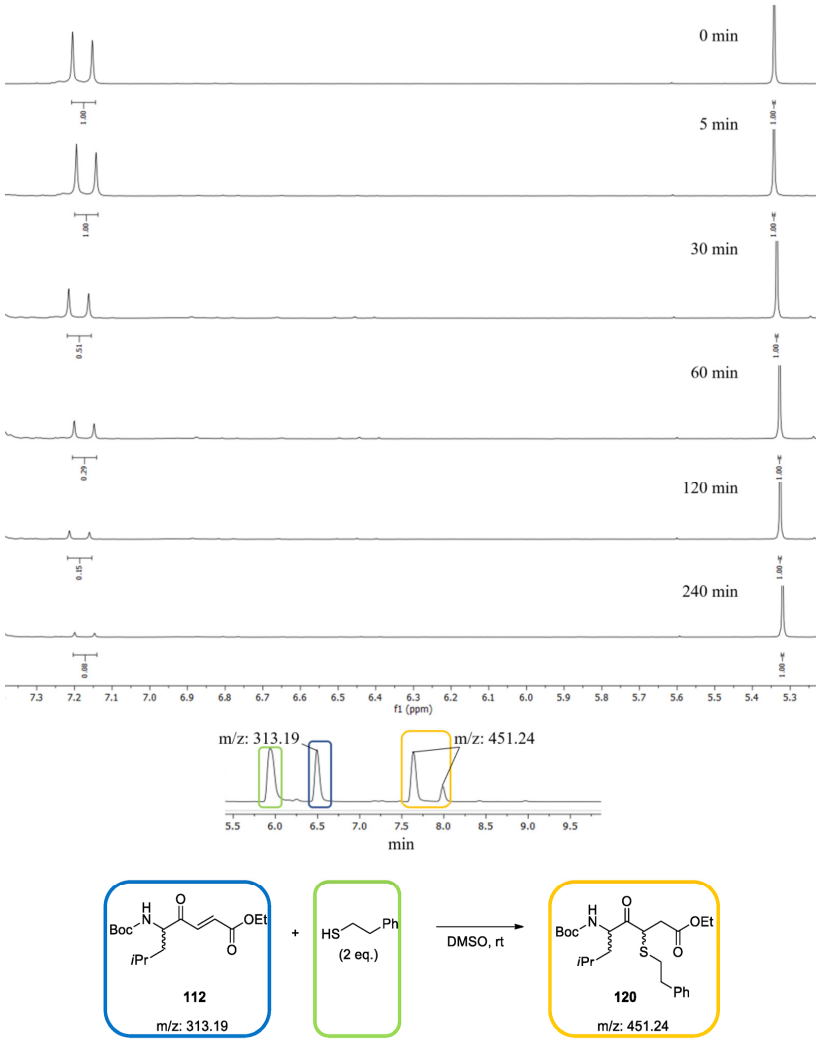


Figure 2. ^{1}H-NMR spectra of the 4-oxoenoate **112** before the addition (0 min) and after 5, 30, 60, 120 and 240 min reaction time with 2-phenylethanthiol in the absence of triethylamine. The integrals of the β-proton of the double bond at 7.17 ppm in relation to the 2-CH_2 signal of the internal standard 1,3-dioxolane at 5.3 ppm are given. LC-MS analysis of the same reaction at 30 min.

Method B (LC-MS): The reactivity of the β-lactam test compound **108** with the nucleophiles was investigated using an LC-MS-based method. To quantify the conversion, the AUCs were determined at 254 nm. In Figure 3, the UV spectra of the test reaction of the β-lactam **108** with EtONa are shown exemplarily. After one hour, complete conversion of the inhibitor to the adduct **122** was observed.

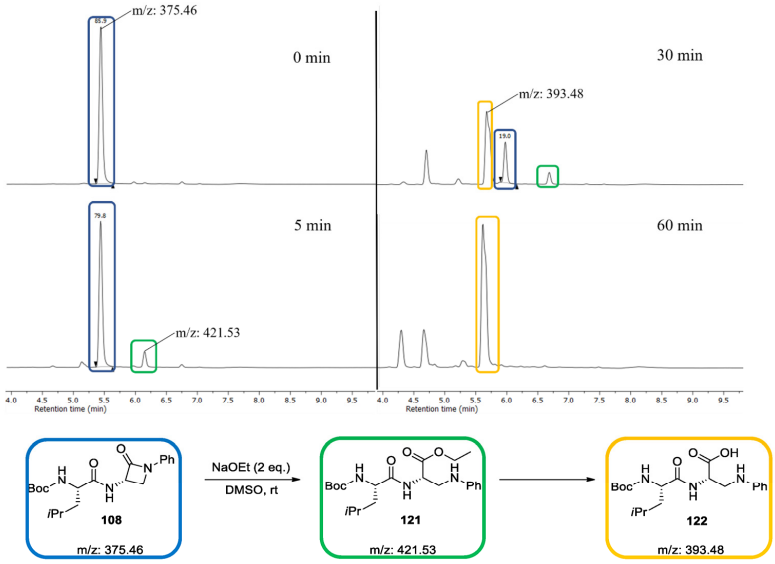


Figure 3. LC-MS spectra of the *β*-lactam **108** before (0 min) the addition and after 5, 30 and 60 min reaction time with sodium ethoxide.

All ^{1}H-NMR/^{13}C-NMR spectra and chromatograms of the reactivity tests are presented in the Supporting Information (Figures S1–S18). The reactivity test results of all warhead compounds with PhEtSH, PhEtSH + Et_3N and EtONa are shown in Figure 4.

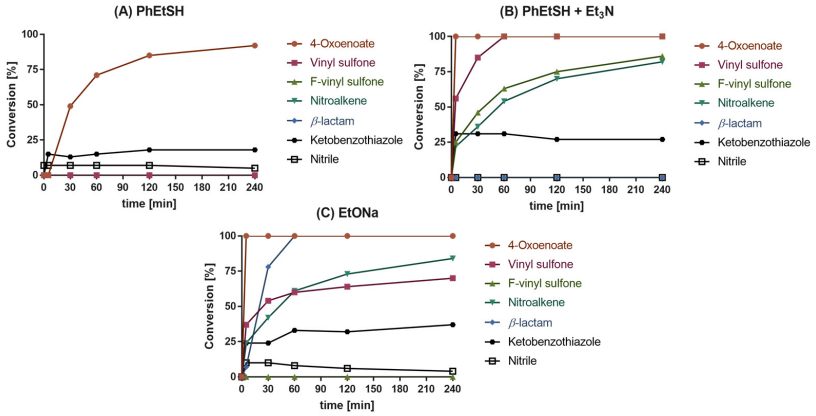


Figure 4. Progress curves of the reactions of the different warhead compounds with the model nucleophiles PhEtSH, PhEtSH + Et_3N and EtONa as measured by NMR and LC-MS analysis.

As depicted in Figure 4A, the 4-oxoenoate **112**, the *α*-ketobenzothiazole **115** and the nitrile **117** moiety did indeed react with PhEtSH under non-basic conditions. In contrast, conversion was not observed with the vinyl sulfone **109**, F-vinyl sulfone **110**, *β*-lactam **108** and nitroalkene **111** warheads. After 240 min, the 4-oxoenoate **112** had nearly completely (92%) reacted with PhEtSH, while only 18% conversion of the *α*-ketobenzothiazole **115** was observed. The equilibrium of the *α*-ketobenzothiazole **116** was reached after 5 min. Similarly, with the nitrile moiety **117**, only 7% conversion was detected, indicating that the

formed thioimidate adduct is relatively unstable (Figure 4A). With the addition of Et_3N (Figure 4B) the overall reactivity increased. Every warhead except the β-lactam **108** and the nitrile **117** reacted with the deprotonated thiol species. Full conversion of the 4-oxoenoate **112** could be observed after 5 min, followed by the vinyl sulfone **109**, which took 60 min for complete reaction. The F-vinyl sulfone **110** and nitroalkene **111** both showed similar reactivity with the thiolate species, with a maximum conversion of 86% and 82% after 240 min, respectively. The α-ketobenzothiazole **115** also showed an increased reactivity, with around 30% conversion. The reactivity tests with EtONa as nucleophile revealed the 4-oxoenoate **112** moiety as the most reactive warhead, which was completely consumed after 5 min (Figure 4C). However, LC-MS analysis did not prove the formation of the expected product but rather unspecific conversion of **112** (see Supplementary Materials, Figure S8). The nitroalkene **111** showed a similar behavior in comparison to the reactivity test with the deprotonated thiolate species, with a conversion of 84% after 240 min. The β-lactam **108** compound showed full conversion after 60 min. In contrast to the deprotonated PhEtSH species, the results indicated a much slower reactivity of the vinyl sulfone **109** with a conversion of 70% after 240 min. No conversion with EtONa was observed for the F-vinyl sulfone **110**. The α-ketobenzothiazole **115** showed a higher conversion in the presence of EtONa (37%) than with PhEtSH, but reached this maximum only after 60 min, showing a slower reaction rate compared to the deprotonated thiol species at 5 min. The equilibrium between the α-ketobenzothiazole **115** and hemiacetal shifted to 37% conversion and was higher compared to the reactivity test with the thiol nucleophiles. The nitrile **117** showed a similar conversion at 5 min with EtONa compared to the protonated thiol species, with 10% conversion, but again decreased after a period of time, which again indicates the instability of the imidate adduct under basic conditions.

The high reactivity of the 4-oxoenoate **112** warhead with the thiolate is in accordance with the high inhibitory potency of dipeptidyl 4-oxoenoate-based compounds against cysteine protease [48]. The missing reactivity of both vinyl sulfones **109** and **110** toward protonated thiol species and the high reactivity with deprotonated thiols are also in agreement with the high activity of vinyl sulfone inhibitors against cysteine proteases with a thiolate residue in the catalytic center, as reported in the literature.

Nitroalkenes are classified as cysteine targeting warheads, which is also confirmed by the observed reactivity with the model thiolate nucleophile [49].

β-lactams are commonly known as warheads in antibacterial agents with transpeptidase-inhibiting properties but have also been used in the development of serine protease inhibitors [21,50,51]. The reactivity tests demonstrate the preference for alcoholate-based nucleophiles, since they only reacted with EtONa and not with PhEtSH/$PhEtS^-$.

α-Ketobenzothiazole derivatives are used as potent serine and cysteine protease inhibitors [52,53]. Therefore, the reactivity of the α-ketobenzothiazole **115** moiety towards all three model nucleophiles was expected. In accordance with the HSAB concept, the stability of the tetrahedral (thio)hemiacetal decreased from the hard sodium ethoxide to the soft thiol/thiolate nucleophiles (EtONa>$PhEtS^-$PhEtSH) after 240 min.

The observed reaction of the nitrile **117** with both nucleophiles (PhEtSH/EtONa) is in accordance with the well-known reactivity of nitrile-based drugs. The observed instability of the (thio)-imidate adduct might have been due to the neutral or basic reaction conditions in solution [54,55]. In contrast, the (thio)imidate adduct is stabilized by interaction with amino acid residues of the enzyme pocket [56].

2.3. Quantum Mechanics Simulations

As model nucleophiles for the QM simulations, methanethiol/ate and methanolate were used. While the formed products were identical, the warheads vinyl sulfone, F-vinyl sulfone and nitroalkene exhibited varying reactivities for PhEtSH and PhEtSH in the presence of triethylamine. Only for 4-oxoenoate and α-ketobenzothiazole was significant reactivity towards PhEtSH observed, whereas for PhEtSH + Et_3N, all warheads except the nitrile and the β-lactam showed reactivity (Figure 4). Since most of the reaction energies

with both MeSH and MeSH + Et_3N were computed to be exergonic (Figure 5A), this cannot be explained merely by thermodynamics. For instance, the experimental results do not show reactivity of the warheads vinyl sulfone, F-vinyl sulfone and nitroalkene with PhEtSH, despite a computed negative free energy of reaction. Thus, to determine whether a reaction can be expected to take place, it is important to consider the whole reaction path, including the activation barriers, which determine the kinetics. Previous calculations have revealed that MeSH is often insufficiently nucleophilic to allow a reaction to occur at room temperature [57]. A base, such as triethylamine, serves as interim storage for the thiol proton before it is transferred to the warhead. By deprotonating the thiol prior to the nucleophilic attack, the nucleophilicity of MeSH is strongly increased, thereby decreasing the associated activation barriers considerably (Figures S22C, S23B and S24A–C). Following the addition of the nucleophile, the proton is transferred back from the base to the anionic intermediate. Unlike the 4-oxoenoate, vinyl sulfone, F-vinyl sulfone and nitroalkene warheads, the β-lactam warhead does not show any reactivity with PhEtSH + Et_3N. The computed reaction mechanism revealed three consecutive steps to obtain the product (Figure 5C). First, the nucleophile attacks the amide carbonyl group (TS1), resulting in a tetrahedral anionic intermediate (Int1). The rate-determining step is the opening of the lactam ring in the second step (TS2). This was computed to be about 33 kcal mol^{-1} for MeSH + NEt_3, which is in excellent agreement with the experimental data. In the last step, the former amide nitrogen is protonated by the base to yield the final product (TS3). For the nitrile warhead, a weak reaction with PhEtSH but none with PhEtSH + triethylamine was observed experimentally, which cannot be explained by reference to the computational data. As described in the reactivity tests, this might have been due to the instability of thioimidates in basic solution.

As a result of our calculations, the difference in reactivity between PhEtSH and PhEtSH + Et_3N for the vinyl sulfone, F-vinyl sulfone and nitroalkene warheads was attributed to a significant reduction in activation barriers caused by proton transfer from the nucleophile to the base prior to the nucleophilic attack. We therefore investigated the reason for the reactivity of 4-oxoenoate and α-ketobenzothiazole warheads with PhEtSH in DMSO in the absence of a base. For 4-oxoenoate, a conversion of 92% was observed experimentally, which corresponds to a computed free energy of reaction of about –9 kcal mol^{-1} for the nucleophilic attack at C_α and about –11 kcal mol^{-1} for the addition at C_β. The reaction of PhEtSH with α-ketobenzothiazole, however, showed only about 18% conversion, and the corresponding product was computed to be 8 kcal mol^{-1} (Figure 5B) and 9 kcal mol^{-1} for the thermodynamic calculation with a bigger basis set (Figure 5A). The solvent used in the experiments was not completely free of water, and, as a result, water molecules were able to catalyze the nucleophilic attack for α-ketobenzothiazole and 4-oxoenoate, as well as the keto-enol tautomerization for the latter (Figure 5B and Figure S22A,B) [58]. Our calculations demonstrate that traces of water in the solvent can function as a base to catalyze the reaction of MeSH with the warhead. The activation barrier for α-ketobenzothiazole is reduced from more than 40 kcal mol^{-1} to roughly 25 kcal mol^{-1}, and the product energy is lowered to 1 kcal mol^{-1} (Figure 5B). Similarly, water catalyzes both the nucleophilic attack of MeSH at the 4-oxoenoate warhead and the subsequent keto-enol tautomerization, leading to a decreased activation barrier of 26 kcal mol^{-1} for the first step (TS1) and one of 20–25 kcal mol^{-1} for the second step (TS2). Additionally, the product energy is even more exergonic at –17-(–18) kcal mol^{-1} (Figure S22B). Contrary to the reaction without water, the proton does not have to be transferred directly from the thiol to the atom to be protonated. Instead, it is shuffled along a chain of water molecules. The keto-enol tautomerization is favored for the C_β-addition, but the barrier associated with the rate-determining nucleophilic attack is nearly identical (Figure S22B). Thus, it is expected that both reactions should occur in solution. For the reaction with an enzyme, the conformation of the binding pocket will likely determine at which carbon atom the nucleophilic attack will occur.

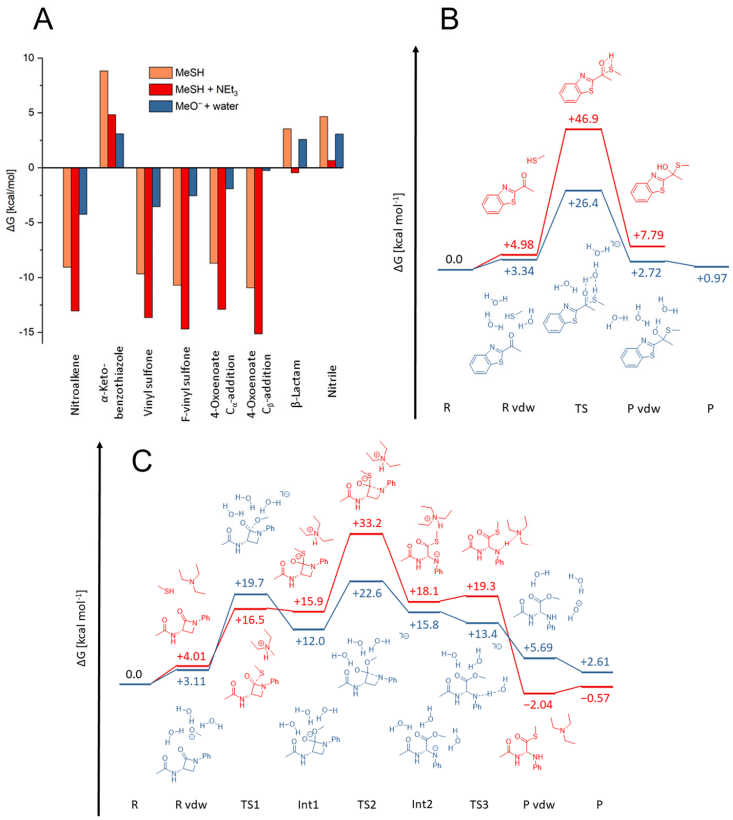


Figure 5. (**A**) Free energies of the reactions for all inhibitor warheads with MeSH, MeSH + Et_3N and MeO^- + $3H_2O$, computed as described in the Supplementary Materials section and Figures S24 by ωB97X−D/6−311++G**//ωB97X−D/6-31+G* calculations. (**B**) Free energy reaction paths of the α-ketobenzothiazole warhead with MeSH (red) and with MeSH in the presence of three water molecules (blue). For MeSH (red), the van der Waals complex (P vdw) and separated product molecules (P) are identical since the reaction yields only a single product molecule. For MeSH + water (blue), the energies are referenced on MeSH + $2H_2O$ and the α-ketobenzothiazole warhead + H_2O (R). (**C**) Free energy reaction paths of the β-lactam warhead with MeSH + NEt_3 (red) and with MeO^- in the presence of three water molecules (blue). The reaction proceeds in three consecutive steps: first, nucleophilic attack at the amide carbonyl group (TS1); second, the opening of the lactam ring (TS2); and third, the proton transfer from the base (NEt_3 or H_2O) to the former amide nitrogen (TS3).

To mimic the reaction of the warheads with NaOEt, we calculated the reaction path with MeO^- and included three water molecules to allow for protonation of the intermediates to obtain the final products and stabilize the reactive anionic species (Figures 5C, S22D, S23D and S24A–C). The reaction can either terminate at the anionic intermediate or proceed to the neutral adduct by transferring one proton from a water molecule, depending on the basicity of the intermediate, i.e., the intermediate carbanion is poorly stabilized for the vinyl sulfone, hence the reaction progresses to form the neutral addition product (Figure S24A). The nitroalkene carbanion, however, is strongly stabilized, and our calculations suggest that the reaction might stop at the intermediate (Figure S24C). Analogously, the α-ketobenzothiazole forms a deprotonated hemiacetal (Figure S23B). Experimentally, no reactivity of the F-vinyl sulfone warhead with NaOEt was observed, which was not supported by our calculations

and is contradictory to chemical intuition (Figure S24B). As previously stated, the barrier for the β-lactam ring opening in reaction with MeSH + Et_3N was computed to be over 30 kcal mol^{-1}, explaining the lack of reactivity in the experiments. Since the anionic species and ring opening are better stabilized in the reaction with MeO^- + $3H_2O$, only 23 kcal mol^{-1} is required in this step, which is consistent with the experimental data (Figure 5C).

2.4. In Vitro Evaluation of the Synthesized Compounds

Inhibition of the target enzymes was tested via fluorometric assays. Therefore, fluorogenic AMC- or FRET-based substrates with appropriate peptide sequences for the different proteases were used (see Supplementary Materials, Figures S19–S20).

The potential inhibitors were initially screened against all five target enzymes at 20 µM. A cut-off value of 80% inhibition at this concentration was set to differentiate the non-active (n.a.) compounds from active ones.

For the reversible inhibitors (α-ketobenzothiazole, nitroalkene, F-vinyl sulfone and nitrile), the IC_{50} values were determined and converted to corresponding K_i values using the Cheng–Prusoff equation [56]. Regarding the irreversible inhibitors (vinyl sulfone, 4-oxoenoate and β-lactam) the K_i, k_{inact} and k_{2nd} values were determined (see Table S1) [56]. For a better overview, the pK_i values were calculated and are presented in Figure 6.

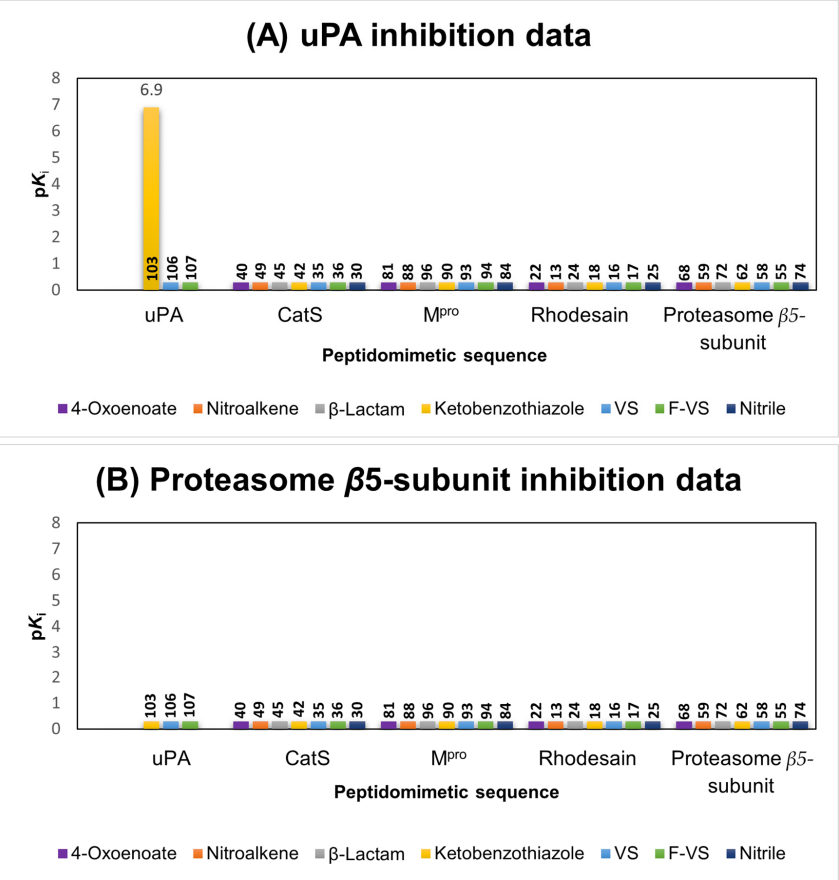


Figure 6. *Cont.*

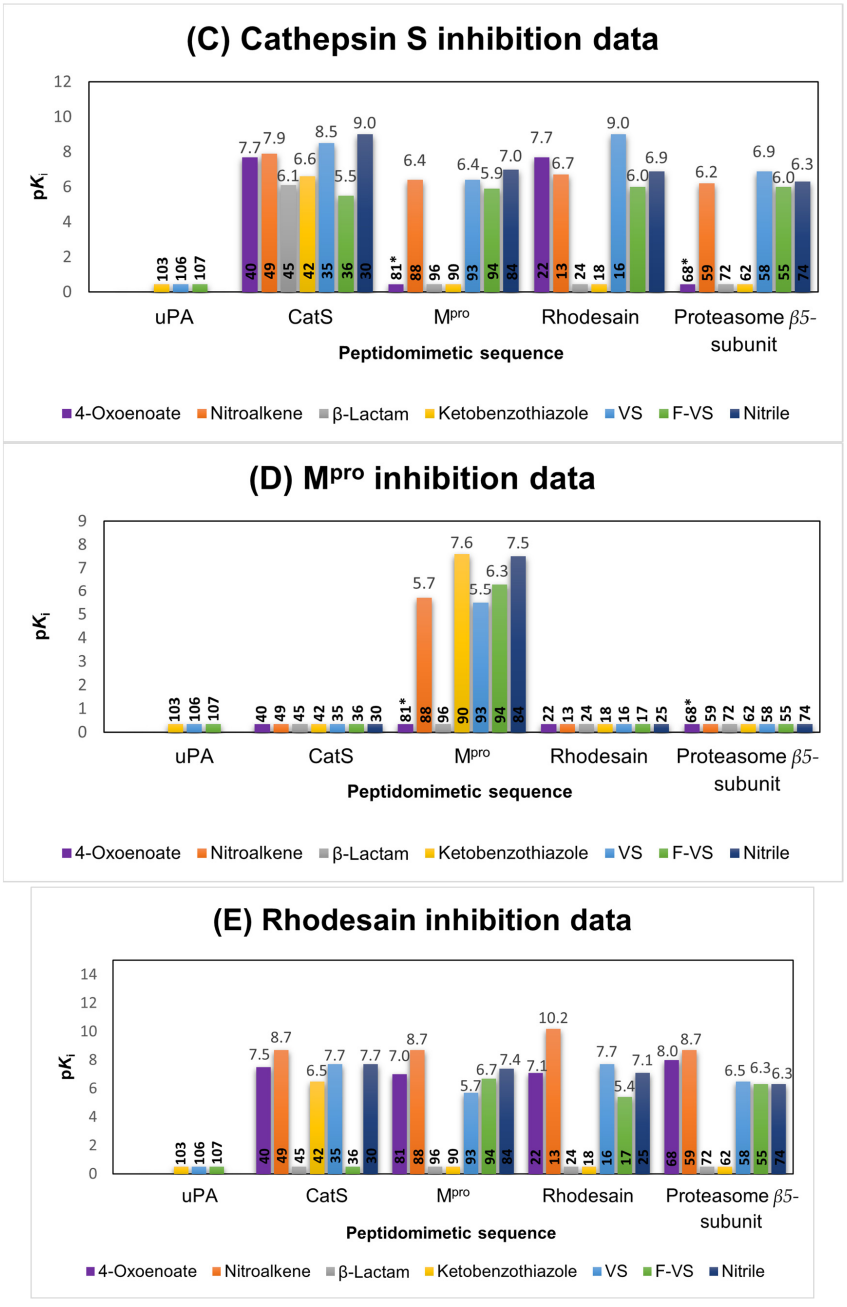


Figure 6. Inhibition data for the assays with uPA (**A**), proteasome β5-subunit (**B**), cathepsine S (**C**), SARS-CoV-2 M^{pro} (**D**) and rhodesain (**E**). pK_i values were calculated from the K_i values ($-\log_{10}(K_i/M)$) [59]. The height of a bar indicates the inhibitory potency of an inhibitor towards the target enzyme, and the color of a bar indicates the warhead of the inhibitor; the peptidomimetic sequence is indicated by the enzyme name under the bars, e.g., the purple bar with the height value

of 7.5 for rhodesain inhibition (Figure 6E) by inhibitor **40** with the oxoenoate warhead and the CatS sequence. * Respective compounds were inactive in the in vitro assay due to instability towards DTT in the buffer.

In the following, the inhibition data will be analyzed for each enzyme, first with their suited peptidomimetic sequences (Figure 6, parts A, B, C, D and E), followed by cross testing against the other enzymes.

uPA. Only the α-ketobenzothiazole inhibitor **103** was found to be active. The combination of the appropriate sequence for uPA with the α-ketobenzothiazole warhead resulted in a potent inhibitor with a pK_i value of 6.9. Other enzymes were not inhibited (Figure 6A).

Proteasome β5-subunit. None of the compounds with the Pyz–(L)Phe–(L)Leu sequence (**55**, **85**, **59**, **62**, **68**, **72** and **74**), which is well-known from the potent boronic acid-based inhibitor bortezomib, showed inhibition of the proteasome at 20 μM, independently of the warhead used (Figure 6B). Moreover, none of the other compounds with any of the other peptidomimetic sequences showed any inhibition. This highlights the general difficulty of addressing this protease with peptidomimetic inhibitors [56]. An alternative warhead which reacts preferably with Ser or Thr proteases is the epoxide functionality, which is also present in the approved proteasome inhibitor carfilzomib. Although very potent, due to its unpredictable reaction mechanism, this warhead was not included in this study [60].

CatS. Regarding the in vitro testing of the cysteine protease CatS, a total of 20 hits were detected (Figure 6C). The most potent inhibitors with the fitting CatS sequence were the nitrile **30** (pK_i = 9) and the vinyl sulfone **35** (pK_i = 8.5). The nitroalkene **49** (pK_i = 7.9) also showed high affinity towards CatS, followed by the 4-oxoenoate **40** (pK_i = 7.7), the α-ketobenzothiazole **42** (pK_i = 6.5), β-lactam **45** (pK_i = 6.1) and F-vinyl sulfone **36** (pK_i = 5.5). Since CatS and rhodesain are both papain-like cysteine proteases with similar active sites, cross reactivity between these two series was expected and has been well described in the literature [61]. The vinyl sulfone with the rhodesain-targeting sequence **16** (pK_i = 9) showed the same inhibition constant as the corresponding inhibitor with the CatS sequence. The vinyl sulfones with the proteasome and the M^{Pro} sequences inhibited CatS to lower degrees (pK_i = 6.9 and 5.5). A comparison of the 4-oxoenoates of the CatS and rhodesain series yielded the same results, since both exhibited the same pK_i value of 7.7 for inhibition of CatS. The 4-oxoenoates designed for targeting the proteasome **68** and the M^{Pro} **81** were essentially inactive against CatS (no inhibition in the initial screening at 20 μM). This can be explained by the instability of these compounds in the CatS assay buffer containing dithiothreitol (DTT).

The F-vinyl sulfones, which are reversibly reacting counterparts of the vinyl sulfones, inhibited CatS to a lower degree, and exchange of the peptidomimetic sequence (**36**, pK_i = 5.5 vs. **17**, pK_i = 6 vs. **55**, pK_i = 6 vs. **94**, pK_i = 5.9) had little to no effect, except for the compounds with the uPA sequence (**103**, **106** and **107**), which was not active at 20 μM against CatS.

The nitroalkene inhibitor which contains the CatS sequence showed a high on-target affinity but changing the sequence to any of the other targeting sequences led to less potent inhibitors (**13**, pK_i = 6.7 vs. **88**, pK_i = 6.4 μM vs. **59**, pK_i = 6.2). Interestingly, the α-ketobenzothiazole- (**42**, pK_i = 6.6) and β-lactam- (**45**, pK_i = 6.1) based inhibitors showed only significant inhibition of CatS if connected to the respective CatS sequence, indicating the strong dependency of a suitable peptidomimetic sequence combined with one of these warheads.

M^{pro}. In comparison to the M^{pro} inhibitors α-ketobenzothiazole **90** (pK_i = 7.6) and nitrile **84** (K_i = 7.5) described in the literature, vinyl sulfone **93** (pK_i = 5.5), F-vinyl sulfone **94** (pK_i = 6.3) and nitroalkene **88** (pK_i = 5.7), all of which contain the appropriate M^{pro} peptidic sequence, showed weaker inhibition (Figure 6D) [40,62]. A clear preference of the protease for specific warheads could be observed. The vinylogous warheads (vinyl

sulfone, F-vinyl sulfone and nitroalkene) showed significantly weaker inhibition than the α-ketobenzothiazole- and nitrile-based compounds. As also observed in the model reactivity studies, the in vitro studies with CatS and both 4-oxoenoate inhibitors with the M^{pro} **81** and the proteasome β5-subunit **68** sequences revealed instability in the buffer with DTT. The β-lactam **96** as well as all compounds containing a targeting structure designed for other proteases were inactive at 20 μM. This indicates a high specificity of the M^{pro} towards its peptidomimetic sequence.

Rhodesain. The results showed similar trends to those found for CatS, with 20 compounds active in the assays (Figure 6E). The most potent was the nitroalkene **13** which contains the corresponding rhodesain peptidic sequence (pK_i = 10.2), followed by the vinyl sulfone **16**, 4-oxoenoate **22** and nitrile **25**, which showed similar inhibition constants (pK_i-= 7.1–7.7). The F-vinyl sulfone **17** showed moderate inhibition (pK_i = 5.4), and the β-lactam **24** and α-ketobenzothiazole **18** were inactive at 20 μM, indicating a preference of rhodesain for vinylogous warheads. Comparable to the CatS study, inhibitors lacked selectivity between rhodesain and CatS due to the structural similarity of the proteases. This is evident through the high pK_i values of the synthesized CatS inhibitors with 4-oxoenoate- **40** (pK_i = 7.5), nitroalkene- **49** (pK_i = 8.7), vinyl sulfone- **35** (pK_i = 7.7) and nitrile- **30** (pK_i = 7.7) moieties. Surprisingly, the α-ketobenzothiazole **42** designed for targeting CatS showed significant inhibition (pK_i = 6.5), whereas the α-ketobenzothiazole with the rhodesain peptidic sequence **18** was inactive. Among the compounds designed for M^{pro} and the proteasome, the nitroalkene derivatives **88** and **59** showed the same potency as the CatS analogue **49**, both with a pK_i value of 8.7. Interestingly, the vinyl sulfones with the M^{pro} **93** (pK_i = 5.7) and the proteasome sequence **58** (pK_i = 6.5) showed significantly lower affinity compared to the vinyl sulfone designed for rhodesain **16** (pK_i = 4.7). Differently, the F-vinyl sulfones **55** (pK_i = 6.3) and **94** (pK_i = 7.7) showed higher affinities than the analogue with the rhodesain sequence **17** (pK_i = 5.4). The 4-oxoenoate inhibitor **68** (pK_i-= 8.0) designed for the proteasome and the one designed for the M^{pro} **81** (pK_i = 7.0) also showed strong inhibition. All other inhibitors with the β-lactam and α-ketobenzothiazole moiety were inactive, as well as the compounds containing the uPA sequence (**103**, **106** and **107**).

2.5. Molecular Docking

To further elucidate the impact of the different warhead types on the binding modes of the inhibitors, protease–inhibitor complexes were investigated with non-covalent and covalent docking [63,64]. For the non-covalent docking, special emphasis was laid on the distances between the reactive nucleophilic carbon atoms of the corresponding warheads to the thiol(ate) or hydroxyl(ate) side chains of the cysteine/serine(threonine) active site amino acids, respectively, as estimates for covalent-bond-formation likeliness. Additionally, the impact of the different warhead moieties on the binding conformation of the inhibitors with otherwise identical peptidomimetic recognition sequences was analyzed. The covalent docking setup was used to investigate whether realistic poses for the covalent complexes could be generated and whether larger conformational rearrangements of the ligand may occur after the covalent reaction.

Conventional non-covalent docking yielded generally reasonable binding modes for all complexes resembling interactions of the crystallographic reference ligands and peptidomimetic recognition sequences in their expected subpockets. Additionally, electrophilic warheads were regularly found in close proximity to the nucleophilic catalytic amino acids (Figure 7, Tables S2 and S3).

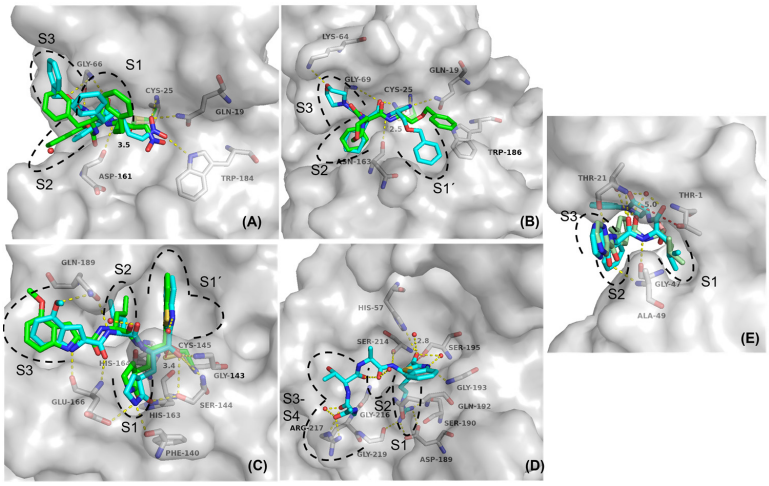


Figure 7. Predicted binding modes and polar interactions (yellow dashed lines) of different inhibitor classes with different enzymes (white carbon atoms and surface). For non-covalent docking poses, the distance between electrophilic carbon and nucleophilic sulfur or oxygen is shown as a red dashed line, with the distance measured in Å. For a clear view, only amino acids that form polar interactions with the ligands are shown as sticks and labelled. Black dashed lines indicate subpocket locations. Non-covalent docking poses are shown with cyan C-atoms and the covalent docking poses with green C-atoms. (**A**) Superposition of the non-covalent and the covalent docking pose of **13** with rhodesain, PDB-ID (2P7U). (**B**) Superposition of the non-covalent and the covalent docking pose of **30** with CatS, PDB-ID (1MS6). (**C**) Superposition of the non-covalent and the covalent docking pose of **90** with SARS-CoV-2-M^{pro}, PDB-ID (6XR3). (**D**) Predicted binding mode of non-covalently docked **103** with uPA, PDB-ID (1W10). (**E**) Superposition of the non-covalent docking pose of **72** and bortezomib (palegreen C-atoms) with the β-5 subunit of human 20S-proteasome, PDB-ID (5LF3).

The docking with rhodesain (pdb entry 2P7U) indicated that the introduction of a Michael-acceptor system as the warhead led to binding poses similar to the co-crystallized reference ligand, with all the essential interactions between inhibitor and enzyme being nearly identical, as exemplified for the docking poses of the nitroalkene inhibitor **13** (Figure 7a). The poses of the covalent and the non-covalent docking showed that the overall orientation of the inhibitor inside the active site should not change much after the covalent reaction, since the final covalent enzyme–inhibitor complex is very similar to the non-covalent complex (Figure 7a). The corresponding electrophilic C-atoms of all warheads were predicted to be in close proximity to the sulfur atom of Cys25 (2.54–3.50 Å), suggesting a high probability for a nucleophilic attack. High docking scores were also found for the nitroalkene inhibitor **13** (FlexX score: –24.03 kJ/mol; MOE score: –2.66), indicating that it should form very favorable non-covalent interactions while correctly placing the electrophilic warhead (distance to Cys25 sulfur: 3.50 Å). This is consistent with the in vitro data, showing that the nitroalkene moiety represents the most potent inhibitor class for rhodesain. Since the α-ketobenzothiazole designed for CatS (**42**) surprisingly inhibited rhodesain with a submicromolar affinity, we compared the non-covalent docking poses between **42** and the ketobenzothiazole with the rhodesain sequence (**18**) (Figure S21). Superposition of the non-covalent docking poses showed that both inhibitors have almost the same positioning with the warhead close to Cys-25 (2.5 Å) inside the active site of rhodesain, indicating that both compounds should have similar affinities towards rhodesain. This makes it hard to explain why inhibitor **42** had a significantly higher affinity for rhodesain in the in vitro testing. Since molecular docking is an inaccurate method, flawed docking

poses are no rarity. The non-covalent docking method used in this case might not be suited to explaining this in vitro result. The results of the docking with CatS (1MS6) showed similar trends, since the distances between the electrophilic C-atoms of the warheads and the sulfur atom of Cys25 were again in close proximity in all cases (2.69–3.37 Å). The vinyl sulfone **35** and the nitrile **25** had high scores (FlexX score: –27.35/–26.22 kJ/mol; MOE score: –5.32/–3.00 kcal/mol) combined with similar binding geometries for the covalent and the non-covalent docking poses (shown for nitrile inhibitor **25**, Figure 7b). These data are in accordance with the in vitro data showing that the nitrile warhead was the most potent one, but other warheads also led to productive enzyme inhibition.

For SARS-CoV-2 M^{pro} (6XR3), the distances between the electrophilic C-atoms and the Cys145 sulfur atom were overall slightly higher (2.90–4.91 Å) compared to the papain-like cysteine proteases. The α-ketobenzothiazole warhead seems to have a very favorable positioning in the binding pocket, as illustrated by the close proximity (3.41 Å) of the electrophilic C-atom to the thiol of the enzyme Cys145 (Figure 7c). Superposition of the covalent and non-covalent docking poses of **90** showed almost identical positioning of the inhibitor inside the enzyme, with most of the polar interactions retained.

Out of all the investigated warheads in this series, only the nitrile, the α-ketobenzothiazole and the β-lactam warheads are known to react with oxygen containing amino acid residues in serine (uPA) or threonine (proteasome) proteases.

For the uPA, which was the only target with only one hit in the enzymatic assay, non-covalent docking revealed a large distance between the electrophilic C-atom and the hydroxy-group in the active site for the β-lactam (5.07Å) as a known serine warhead. Only the α-ketobenzothiazole inhibitor **103**, which had one of the highest scores out of all the inhibitors (FlexX score: –51.59 kJ/mol), was in close proximity to the oxygen of Ser195 (2.84 Å distance to the electrophilic C-atom). This inhibitor also showed a high potency in the in vitro study (Figure 7d). Finally, docking of the β-lactam containing inhibitor **72** designed for the proteasome revealed that the warhead position was, again, too far away from the threonine oxygen (4.95 Å), possibly preventing a covalent reaction (Figure 7e). This could be explained by the shifted positioning of the lactam moiety compared to the other warheads. Although the docking of the nitrile and α-ketobenzothiazole inhibitors **74** and **62** might suggest that these compounds should inhibit their target sufficiently since the warheads are positioned correctly and in close proximity (2.27 Å/3.16 Å) to the Thr-1 oxygen atom, there was still no inhibition with these warheads in the in vitro study. This might have been due to wrongly generated binding poses, since docking approaches are not always reliable and cannot be considered flawless in all cases. A possible explanation why none of the compounds designed to address the β5-subunit of the proteasome showed any inhibition might be the catalytic dyad in the active site consisting of Lys33 and Thr1 compared to the catalytic dyads or even triads in the other enzymes, where the deprotonation of the active site residue is assisted by histidine and/or asparagine. The lysine residue might not always be able to deprotonate the threonine in the active site, depending on the inhibitor, and thus facilitate the covalent reaction step with a warhead [65].

2.6. Comparison of the Reactivity Assay Results with the In Vitro Study

Based on the reactivity assay, all Michael acceptors (4-oxoenoate **112**, (F-) vinyl sulfone **109**/**110** and nitroalkene **111**) showed high reactivity toward the deprotonated cysteine model nucleophile, which is congruent with the observed behavior of the synthesized compounds designed for CatS and rhodesain inhibition in the in vitro studies. Furthermore, the α-ketobenzothiazole warhead **115** showed a strong reactivity for both model nucleophiles ($PhEtS^-$/EtONa), which is consistent with the correspondent uPA and M^{pro} inhibitors **103** and **90** in the protease assays. However, the nitrile **117** showed no reaction with the deprotonated cysteine but with the serine model nucleophile, which contradicts the high inhibitory activity against the cysteine proteases and the missing inhibition by the proteasome β5-subunit inhibitor **74**. This might have been due to the aforementioned

instability of the thioimidate adduct in basic conditions compared to the stabilized adduct in the enzyme pocket and the overall difficulty of addressing the proteasome β5-subunit. The β-lactam **108** showed only a strong reactivity towards the serine model nucleophile, but the corresponding bortezomib derivative **72** did not inhibit the proteasome β5-subunit, which might have been due to the shift of the electrophilic center of the β-lactam moiety into the S1′ pocket and the resulting increase in distance. The 4-oxoenoate moiety **112** was the only warhead that showed high reactivity toward the protonated cysteine model nucleophile, which might hint at non-selective reactivity behavior toward thiol species under physiological conditions. This could also be observed in the in vitro studies. The 4-oxoenoate compounds designed for the M^{pro} **81** and proteasome-β5-subunit **68** both reacted quickly with DTT in the respective buffer solutions and appeared to be inactive.

3. Discussion

Covalent targeting has become a popular and powerful concept in drug discovery, and great efforts have been devoted to developing and repurposing different warheads [66]. In this first extensive systematic study, we aimed to achieve a deeper insight into the reactivities and selectivities of a selection of electrophilic traps combined with established peptidomimetic sequences for the uPA, CatS, β5-subunit of the proteasome, SARS-CoV-2 M^{pro} and rhodesain, which represent cysteine, serine and threonine proteases. Based on these peptidomimetic sequences, we synthesized compounds decorated with warheads of different specificities. We chose the Michael acceptors ((F-)vinyl sulfone, nitroalkene and 4-oxoenoate) as cysteine-targeting and β-lactam as serine/threonine-targeting representatives. Furthermore, nitriles and α-ketobenzothiazoles were used, as they are applicable for both hydroxy- and thiol-containing nucleophiles. The compounds were tested on each target to analyze their affinities as well as their selectivity profiles.

Based on the in vitro studies, it is evident that the peptidomimetic sequences of the synthesized compounds play a crucial role in the selectivity towards the tested on-target and off-target proteases. This could be observed by the selectivity profile towards the cysteine protease M^{pro} and serine protease uPA. Only the inhibitors with the suited peptidomimetic sequence for M^{pro} (**84**, **88**, **90**, **93** and **94**) and for uPA (**103**) displayed inhibitory activity towards their targeted protease. Furthermore, the selection of a suitable warhead for the specific type of protease nucleophile ensures high affinity to the target or even activity in the first place, as demonstrated with the bortezomib congeners and the α-ketobenzothiazole inhibitor **103** as the only affine compound towards the uPA. The structurally similar papain-like proteases CatS and rhodesain showed that cross reactivity can occur, despite the design of well-defined peptidomimetic sequences. Therefore, the combination of both a highly reactive warhead towards the target protease, for example, the nitrile **30** group for CatS or the nitroalkene **13** for rhodesain, with a suitable peptidomimetic sequence can lead to potent inhibitors with promising pharmacodynamic properties.

Non-covalent docking yielded reasonable binding modes for all compounds resembling interactions of the crystallographic reference ligands and peptide recognition sequences in their expected subpockets. Additionally, electrophilic warheads were regularly found in close proximity to the nucleophilic catalytic amino acids, except for the β-lactams.

A reactivity test system with tool compounds of the used warheads and model nucleophiles was established to evaluate chemoselectivity. The findings confirmed the high reactivity of the 4-oxoenoate, the (F-)vinyl sulfones and the nitroalkene moieties towards the deprotonated thiol nucleophile/cysteine model, and high affinity of the Michael acceptor inhibitors towards the cysteine proteases was observed. Analogously to the in vitro studies of the uPA and M^{pro} target, the α-ketobenzothiazole warhead was found to be a potent electrophilic trap for both cysteine and serine proteases. Nevertheless, some major differences in reactivity could be observed, which might have been due to different conditions used in the chemical test system and the biochemical in vitro studies. To the best of our knowledge, this is the first extensive study in which different warhead types were combined with different peptidic recognition units and in which the resulting compounds

were cross tested against different protease types. Similar published studies limited their focus to testing different warheads on one target or exchanging the peptidic backbone while retaining the same warhead [37,59,67,68].

4. Material and Methods

The material as well as the methods used for this study are described in the Supporting Information. The authors have cited additional references within the Supporting Information [21,42,43,46,47,63,64,69–89] (Supplementary Figures of the reactivity study (Figures S1–S18), of the fluorometric inhibition assays (Figures S19 and S20), of molecular docking (Figure S21), of quantum mechanics simulation (Figures S22–S25) and of the NMR-spectra and HPLC-chromatograms of the final inhibitors (Figures S26–S137) can be accessed in the supporting information).

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms24087226/s1.

Author Contributions: P.M.: synthesis, reactivity assay and in vitro investigation, writing—original draft; M.M.: synthesis, molecular docking and in vitro investigation, writing—original draft; J.L.M.: synthesis and in vitro investigation, writing—original draft; M.S.: synthesis, review and editing; J.M.: quantum mechanics simulations, review and editing; K.S.: synthesis; C.K.: review and editing; C.Z.: rhodesain expression; S.J.H.: SARS-CoV-2 M^{pro} expression; A.F.: synthesis, A.L.: synthesis, S.d.l.H.-R.: synthesis, L.A.-B.: synthesis, S.R.: synthesis, K.D.: quantum mechanics simulations, W.N.: quantum mechanics simulations, F.V.G.: validation, review and editing; B.E.: validation, review and editing; T.S.: validation, review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: Financial support from the DFG (Deutsche Forschungsgemeinschaft) in the framework of the CRC 1066 (Nanodimensional Polymeric Therapeutics for Tumor Therapy) project Q5 (Targeting and Immunomodulator Structures and their Coupling to Therapeutic Nanosystems for Oncological Application) is gratefully acknowledged. This research was also funded by the Generalitat Valenciana (PROMETEO with ref. CIPROM/2021/079) and Universitat Jaume I (UJI-B2021-71 and SomUJIcontracovid crowdfunding campaign). L.A-B. thanks the Ministerio de Universidades for funding a PhD fellowship (ref. FPU19/04913).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We thank Sabine Maehrlein for the in vitro testing of all synthesized compounds against the proteasome β5-subunit and Katrina Schorstein for help in the synthesis. The authors wish to thank the Serveis Centrals d'Instrumentació Científica of Universitat Jaume I for technical support. The authors gratefully acknowledge the computing time provided to them on the high-performance computer Noctua2 at the NHR Center PC2. This system is funded by the Federal Ministry of Education and Research and the state governments participating on the basis of the resolutions of the GWK for national high-performance computing at universities (www.nhr-verein.de/unsere-partner, accessed on 23 March 2023). The authors gratefully acknowledge the computational and data resources provided by the Leibniz Supercomputing Centre (www.lrz.de, accessed on 9 April 2023).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. López-Otín, C.; Overall, C.M. Protease degradomics: A new challenge for proteomics. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 509–519. [CrossRef] [PubMed]
2. Grozdanić, M.; Vidmar, R.; Vizovišek, M.; Fonović, M. Degradomics in Biomarker Discovery. *Proteom. Clin. Appl.* **2019**, *13*, 1800138. [CrossRef] [PubMed]
3. Ruggiano, A.; Ramadan, K. DNA–protein crosslink proteases in genome stability. *Commun. Biol.* **2021**, *4*, 11. [CrossRef]
4. Lee, C.W.; Stankowski, J.N.; Chew, J.; Cook, C.N.; Lam, Y.W.; Almeida, S.; Carlomagno, Y.; Lau, K.F.; Prudencio, M.; Gao, F.B.; et al. The lysosomal protein cathepsin L is a progranulin protease. *Mol. Neurodegener.* **2017**, *12*, 55. [CrossRef] [PubMed]

5. Eatemadi, A.; Aiyelabegan, H.T.; Negahdari, B.; Mazlomi, M.A.; Daraee, H.; Daraee, N.; Eatemadi, R.; Sadroddiny, E. Role of protease and protease inhibitors in cancer pathogenesis and treatment. *Biomed. Pharmacother.* **2017**, *86*, 221–231. [CrossRef] [PubMed]
6. Liu, C.L.; Guo, J.; Zhang, X.; Sukhova, G.K.; Libby, P.; Shi, G.P. Cysteine protease cathepsins in cardiovascular disease: From basic research to clinical trials. *Nat. Rev. Cardiol.* **2018**, *15*, 351–370. [CrossRef]
7. Previti, S.; Ettari, R.; Calcaterra, E.; Di Chio, C.; Ravichandran, R.; Zimmer, C.; Hammerschmidt, S.; Wagner, A.; Bogacz, M.; Cosconati, S.; et al. Development of Urea-Bond-Containing Michael Acceptors as Antitrypanosomal Agents Targeting Rhodesain. *ACS Med. Chem. Lett.* **2022**, *13*, 1083–1090. [CrossRef]
8. Rocha, D.A.; Silva, E.B.; Fortes, I.S.; Lopes, M.S.; Ferreira, R.S.; Andrade, S.F. Synthesis and structure-activity relationship studies of cruzain and rhodesain inhibitors. *Eur. J. Med. Chem.* **2018**, *157*, 1426–1459. [CrossRef]
9. Kincaid, J.R.A.; Caravez, J.C.; Iyer, K.S.; Kavthe, R.D.; Fleck, N.; Aue, D.H.; Lipshutz, B.H. A sustainable synthesis of the SARS-CoV-2 Mpro inhibitor nirmatrelvir, the active ingredient in Paxlovid. *Commun. Chem.* **2022**, *5*, 156. [CrossRef]
10. Müller, P.; Maus, H.; Hammerschmidt, S.J.; Knaff, P.M.; Mailänder, V.; Schirmeister, T.; Kersten, C. Interfering with Host Proteases in SARS-CoV-2 Entry as a Promising Therapeutic Strategy. *Curr. Med. Chem.* **2022**, *29*, 635–665. [CrossRef]
11. Knaff, P.M.; Müller, P.; Kersten, C.; Wettstein, L.; Münch, J.; Landfester, K.; Mailänder, V. Structure-Based Design of High-Affinity and Selective Peptidomimetic Hepsin Inhibitors. *Biomacromolecules* **2022**, *23*, 2236–2242. [CrossRef] [PubMed]
12. Tsantrizos, Y.S.; Bolger, G.; Bonneau, P.; Cameron, D.R.; Goudreau, N.; Kukolj, G.; LaPlante, S.R.; Llinàs-Brunet, M.; Nar, H.; Lamarre, D. Macrocyclic inhibitors of the NS3 protease as potential therapeutic agents of hepatitis C virus infection. *Angew. Chemie Int. Ed.* **2003**, *42*, 1356–1360. [CrossRef] [PubMed]
13. Maus, H.; Barthels, F.; Hammerschmidt, S.J.; Kopp, K.; Millies, B.; Gellert, A.; Ruggieri, A.; Schirmeister, T. SAR of novel benzothiazoles targeting an allosteric pocket of DENV and ZIKV NS2B/NS3 proteases. *Bioorg. Med. Chem.* **2021**, *47*, 116392. [CrossRef] [PubMed]
14. Martin, J.S.; MacKenzie, C.J.; Fletcher, D.; Gilbert, I.H. Characterising covalent warhead reactivity. *Bioorg. Med. Chem.* **2019**, *27*, 2066–2074. [CrossRef] [PubMed]
15. Wang, Y.-H.; Zhang, F.; Diao, H.; Wu, R. Covalent Inhibition Mechanism of Antidiabetic Drugs—Vildagliptin vs Saxagliptin. *ACS Catal.* **2019**, *9*, 2292–2302. [CrossRef]
16. Lamb, Y.N. Nirmatrelvir Plus Ritonavir: First Approval. *Drugs* **2022**, *82*, 585–591. [CrossRef]
17. Robak, P.; Robak, T. Bortezomib for the Treatment of Hematologic Malignancies: 15 Years Later. *Drugs R D* **2019**, *19*, 73–92. [CrossRef]
18. Johe, P.; Jung, S.; Endres, E.; Kersten, C.; Zimmer, C.; Ye, W.; Sönnichsen, C.; Hellmich, U.A.; Sotriffer, C.; Schirmeister, T.; et al. Warhead Reactivity Limits the Speed of Inhibition of the Cysteine Protease Rhodesain. *ACS Chem. Biol.* **2021**, *16*, 661–670. [CrossRef]
19. Santos, M.; Moreira, R. Mini-Reviews. *Med. Chem.* **2007**, *7*, 1040–1050.
20. Adams, J.; Kauffman, M. Development of the Proteasome Inhibitor Velcade^{TM} (Bortezomib). *Cancer Investig.* **2004**, *22*, 304–311. [CrossRef]
21. Dražić, T.; Kopf, S.; Corridan, J.; Leuthold, M.M.; Bertoša, B.; Klein, C.D. Peptide-β-lactam Inhibitors of Dengue and West Nile Virus NS2B-NS3 Protease Display Two Distinct Binding Modes. *J. Med. Chem.* **2020**, *63*, 140–156. [CrossRef] [PubMed]
22. Fleming, F.F.; Yao, L.; Ravikumar, P.C.; Funk, L.; Shook, B.C. Nitrile-Containing Pharmaceuticals: Efficacious Roles of the Nitrile Pharmacophore. *J. Med. Chem.* **2010**, *53*, 7902–7917. [CrossRef] [PubMed]
23. Bullock, T.L.; Breddam, K.; Remington, J.S. Peptide Aldehyde Complexes with Wheat Serine Carboxypeptidase II: Implications for the Catalytic Mechanism and Substrate Specificity. *J. Mol. Biol.* **1996**, *255*, 714–725. [CrossRef] [PubMed]
24. Hu, X.; Lin, C.; Xu, Q.; Zhou, X.; Zeng, P.; McCormick, P.J.; Jiang, H.; Li, J.; Zhang, J. Structural Basis for the Inhibition of Coronaviral Main Proteases by a Benzothiazole-Based Inhibitor. *Viruses* **2022**, *14*, 2075. [CrossRef] [PubMed]
25. Akiyama, Y.; Tsutsumi, S.; Hatsushiba, E.; Ohuchi, S.; Okonogi, T. Peptidyl α-keto thiazole as potent thrombin inhibitors. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 533–538. [CrossRef]
26. Vincenza Carriero, M.; Patrizia Stoppelli, M. The Urokinase-type Plasminogen Activator and the Generation of Inhibitors of Urokinase Activity and Signaling. *Curr. Pharm. Des.* **2011**, *17*, 1944–1961. [CrossRef] [PubMed]
27. Ismail, A.A.; Shaker, B.T.; Bajou, K. The plasminogen–activator plasmin system in physiological and pathophysiological angiogenesis. *Int. J. Mol. Sci.* **2021**, *23*, 337. [CrossRef] [PubMed]
28. Smith, H.W.; Marshall, C.J. Regulation of cell signalling by uPAR. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 23–36. [CrossRef]
29. Kumar, A.A.; Buckley, B.J.; Ranson, M. The Urokinase Plasminogen Activation System in Pancreatic Cancer: Prospective Diagnostic and Therapeutic Targets. *Biomolecules* **2022**, *12*, 152. [CrossRef]
30. Li, C.Y.; de Veer, S.J.; Law, R.H.P.; Whisstock, J.C.; Craik, D.J.; Swedberg, J.E. Characterising the Subsite Specificity of Urokinase-Type Plasminogen Activator and Tissue-Type Plasminogen Activator using a Sequence-Defined Peptide Aldehyde Library. *ChemBioChem* **2019**, *20*, 46–50. [CrossRef]
31. Collins, G.A.; Goldberg, A.L. The Logic of the 26S Proteasome. *Cell* **2017**, *169*, 792–806. [CrossRef] [PubMed]
32. Kisselev, A.F.; van der Linden, W.A.; Overkleeft, H.S. Proteasome Inhibitors: An Expanding Army Attacking a Unique Target. *Chem. Biol.* **2012**, *19*, 99–115. [CrossRef] [PubMed]

33. Kerr, I.D.; Wu, P.; Marion-Tsukamaki, R.; Mackey, Z.B.; Brinen, L.S. Crystal Structures of TbCatB and Rhodesain, Potential Chemotherapeutic Targets and Major Cysteine Proteases of Trypanosoma brucei. *PLoS Negl. Trop. Dis.* **2010**, *4*, e701. [CrossRef] [PubMed]
34. Pauly, T.A.; Sulea, T.; Ammirati, M.; Sivaraman, J.; Danley, D.E.; Griffor, M.C.; Kamath, A.V.; Wang, I.K.; Laird, E.R.; Seddon, A.P.; et al. Specificity determinants of human cathepsin S revealed by crystal structures of complexes. *Biochemistry* **2003**, *42*, 3203–3213. [CrossRef] [PubMed]
35. Wilkinson, R.D.A.; Williams, R.; Scott, C.J.; Burden, R.E. Cathepsin S: Therapeutic, diagnostic, and prognostic potential. *Biol. Chem.* **2015**, *396*, 867–882. [CrossRef] [PubMed]
36. Chen, J.-C.; Uang, B.-J.; Lyu, P.-C.; Chang, J.-Y.; Liu, K.-J.; Kuo, C.-C.; Hsieh, H.-P.; Wang, H.-C.; Cheng, C.-S.; Chang, Y.-H.; et al. Design and Synthesis of α-Ketoamides as Cathepsin S Inhibitors with Potential Applications against Tumor Invasion and Angiogenesis. *J. Med. Chem.* **2010**, *53*, 4545–4549. [CrossRef] [PubMed]
37. Jung, S.; Fuchs, N.; Johe, P.; Wagner, A.; Diehl, E.; Yuliani, T.; Zimmer, C.; Barthels, F.; Zimmermann, R.A.; Klein, P.; et al. Fluorovinylsulfones and -Sulfonates as Potent Covalent Reversible Inhibitors of the Trypanosomal Cysteine Protease Rhodesain: Structure-Activity Relationship, Inhibition Mechanism, Metabolism, and in Vivo Studies. *J. Med. Chem.* **2021**, *64*, 12322–12358. [CrossRef]
38. Previti, S.; Ettari, R.; Cosconati, S.; Schirmeister, G.; Chouchene, K.; Wagner, A.; Hellmich, U.A.; Ulrich, K.; Krauth-Siegel, R.L.; Wich, P.R.; et al. Development of Novel Peptide-Based Michael Acceptors Targeting Rhodesain and Falcipain-2 for the Treatment of Neglected Tropical Diseases (NTDs). *J. Med. Chem.* **2017**, *60*, 6911–6923. [CrossRef]
39. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; et al. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020**, *582*, 289–293. [CrossRef]
40. Owen, D.R.; Allerton, C.M.N.; Anderson, A.S.; Aschenbrenner, L.; Avery, M.; Berritt, S.; Boras, B.; Cardin, R.D.; Carlo, A.; Coffman, K.J.; et al. An oral SARS-CoV-2 M pro inhibitor clinical candidate for the treatment of COVID-19. *Science* **2021**, *374*, 1586–1593. [CrossRef]
41. Zeslawska, E.; Jacob, U.; Schweinitz, A.; Coombs, G.; Bode, W.; Madison, E. Crystals of urokinase type plasminogen activator complexes reveal the binding mode of peptidomimetic inhibitors. *J. Mol. Biol.* **2003**, *328*, 109–118. [CrossRef] [PubMed]
42. Schrader, J.; Henneberg, F.; Mata, R.A.; Tittmann, K.; Schneider, T.R.; Stark, H.; Bourenkov, G.; Chari, A. The inhibition mechanism of human 20S proteasomes enables next-generation inhibitor design. *Science* **2016**, *353*, 594–598. [CrossRef] [PubMed]
43. Ward, Y.D.; Thomson, D.S.; Frye, L.L.; Cywin, C.L.; Morwick, T.; Emmanuel, M.J.; Zindell, R.; McNeil, D.; Bekkali, Y.; Marc Girardot, M.; et al. Design and synthesis of dipeptide nitriles as reversible and potent Cathepsin S inhibitors. *J. Med. Chem.* **2002**, *45*, 5471–5482. [CrossRef] [PubMed]
44. Hattori, S.-I.; Higashi-Kuwata, N.; Hayashi, H.; Allu, S.R.; Raghavaiah, J.; Bulut, H.; Das, D.; Anson, B.J.; Lendy, E.K.; Takamatsu, Y.; et al. A small molecule compound with an indole moiety inhibits the main protease of SARS-CoV-2 and blocks virus replication. *Nat. Commun.* **2021**, *12*, 668. [CrossRef] [PubMed]
45. Kerr, I.D.; Lee, J.H.; Farady, C.J.; Marion, R.; Rickert, M.; Sajid, M.; Pandey, K.C.; Caffrey, C.R.; Legac, J.; Hansell, E.; et al. Vinyl sulfones as antiparasitic agents and a structural basis for drug design. *J. Biol. Chem.* **2009**, *284*, 25697–25703. [CrossRef] [PubMed]
46. Vuong, W.; Vederas, J.C. Improved Synthesis of a Cyclic Glutamine Analogue Used in Antiviral Agents Targeting 3C and 3CL Proteases Including SARS-CoV-2 M pro. *J. Org. Chem.* **2021**, *86*, 13104–13110. [CrossRef]
47. Tian, Q.; Nayyar, N.K.; Babu, S.; Chen, L.; Tao, J.; Lee, S.; Tibbetts, A.; Moran, T.; Liou, J.; Guo, M.; et al. An efficient synthesis of a key intermediate for the preparation of the rhinovirus protease inhibitor AG7088 via asymmetric dianionic cyanomethylation of N-Boc-l-(+)-glutamic acid dimethyl ester. *Tetrahedron Lett.* **2001**, *42*, 6807–6809. [CrossRef]
48. Royo, S.; Rodríguez, S.; Schirmeister, T.; Kesselring, J.; Kaiser, M.; González, F.V. Dipeptidyl Enoates As Potent Rhodesain Inhibitors That Display a Dual Mode of Action. *ChemMedChem* **2015**, *10*, 1484–1487. [CrossRef]
49. Latorre, A.; Schirmeister, T.; Kesselring, J.; Jung, S.; Johé, P.; Hellmich, U.A.; Heilos, A.; Engels, B.; Krauth-Siegel, R.L.; Dirdjaja, N.; et al. Dipeptidyl Nitroalkenes as Potent Reversible Inhibitors of Cysteine Proteases Rhodesain and Cruzain. *ACS Med. Chem. Lett.* **2016**, *7*, 1073–1076. [CrossRef]
50. Shah, S.K.; Finke, P.E.; Brause, K.A.; Chandler, G.O.; Ashe, B.M.; Weston, H.; Maycock, A.L.; Mumford, R.A.; Doherty, J.B. Monocyclic β-lactam inhibitors of human leukocyte elastase. Stereospecific synthesis and activity of 3,4-disubstituted-2-azetidinones. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 2295–2298. [CrossRef]
51. Han, W.T.; Trehan, A.K.; Kim Wright, J.J.; Federici, M.E.; Seiler, S.M.; Meanwell, N.A. Azetidin-2-one derivatives as inhibitors of thrombin. *Bioorg. Med. Chem.* **1995**, *3*, 1123–1143. [CrossRef] [PubMed]
52. Steert, K.; Berg, M.; Mottram, J.C.; Westrop, G.D.; Coombs, G.H.; Cos, P.; Maes, L.; Joossens, J.; Van der Veken, P.; Haemers, A.; et al. α-Ketoheterocycles as Inhibitors of Leishmania mexicana Cysteine Protease CPB. *ChemMedChem* **2010**, *5*, 1734–1748. [CrossRef] [PubMed]
53. 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.; et al. In-Depth Study of Tripeptide-Based α-Ketoheterocycles as Inhibitors of Thrombin. Effective Utilization of the S1 'Subsite and Its Implications to Structure-Based Drug Design. *J. Med. Chem.* **2005**, *48*, 1984–2008. [CrossRef] [PubMed]
54. DiNinno, F.; Ernest, V.L. Facile Synthesis of β-Thioxo Esters from β-Enamino Esters. *J. Org. Chem.* **1979**, *44*, 3271–3273. [CrossRef]

55. Delprino, L.; Giacomotti, M.; Dosio, F.; Brusa, P.; Ceruti, M.; Grosa, G.; Cattel, L. Toxin-Targeted Design for Anticancer Therapy. I: Synthesis and Biological Evaluation of New Thioimidate Heterobifunctional Reagents. *J. Pharm. Sci.* **1993**, *82*, 506–512. [CrossRef] [PubMed]
56. Brogi, S.; Ibba, R.; Rossi, S.; Butini, S.; Calderone, V.; Gemma, S.; Campiani, G. Covalent Reversible Inhibitors of Cysteine Proteases Containing the Nitrile Warhead: Recent Advancement in the Field of Viral and Parasitic Diseases. *Molecules* **2022**, *27*, 2561. [CrossRef]
57. Barthels, F.; Meyr, J.; Hammerschmidt, S.J.; Marciniak, T.; Räder, H.-J.; Ziebuhr, W.; Engels, B.; Schirmeister, T. 2-Sulfonylpyrimidines as Privileged Warheads for the Development of S. aureus Sortase A Inhibitors. *Front. Mol. Biosci.* **2022**, *8*, 804970. [CrossRef]
58. Paasche, A.; Schiller, M.; Schirmeister, T.; Engels, B. Mechanistic Study of the Reaction of Thiol-Containing Enzymes with α,β-Unsaturated Carbonyl Substrates by Computation and Chemoassays. *ChemMedChem* **2010**, *5*, 869–880. [CrossRef]
59. Silva, D.G.; Ribeiro, J.F.R.; De Vita, D.; Cianni, L.; Franco, C.H.; Freitas-Junior, L.H.; Moraes, C.B.; Rocha, J.R.; Burtoloso, A.C.B.; Kenny, P.W.; et al. A comparative study of warheads for design of cysteine protease inhibitors. *Bioorg. Med. Chem. Lett.* **2017**, *27*, 5031–5035. [CrossRef]
60. Kim, K.B.; Crews, C.M. From epoxomicin to carfilzomib: Chemistry, biology, and medical outcomes. *Nat. Prod. Rep.* **2013**, *30*, 600. [CrossRef]
61. Cianni, L.; Feldmann, C.W.; Gilberg, E.; Gütschow, M.; Juliano, L.; Leitão, A.; Bajorath, J.; Montanari, C.A. Can Cysteine Protease Cross-Class Inhibitors Achieve Selectivity? *J. Med. Chem.* **2019**, *62*, 10497–10525. [CrossRef] [PubMed]
62. Konno, S.; Kobayashi, K.; Senda, M.; Funai, Y.; Seki, Y.; Tamai, I.; Schäkel, L.; Sakata, K.; Pillaiyar, T.; Taguchi, A.; et al. 3CL Protease Inhibitors with an Electrophilic Arylketone Moiety as Anti-SARS-CoV-2 Agents. *J. Med. Chem.* **2022**, *65*, 2926–2939. [CrossRef] [PubMed]
63. Chemical Computing Group ULC. *Molecular Operating Environment (MOE)*; Chemical Computing Group ULC: Montreal, QC, Canada, 2020; Available online: https://www.chemcomp.com/. (accessed on 23 March 2022).
64. *LeadIT/FlexX*, Version 2.3.2; GmbH, BioSolveIT: Sankt Augustin, Germany, 2017.
65. Serrano-Aparicio, N.; Moliner, V.; Świderek, K. Nature of Irreversible Inhibition of Human 20S Proteasome by Salinosporamide A. The Critical Role of Lys–Asp Dyad Revealed from Electrostatic Effects Analysis. *ACS Catal.* **2021**, *11*, 3575–3589. [CrossRef]
66. Gehringer, M.; Laufer, S.A. Emerging and Re-Emerging Warheads for Targeted Covalent Inhibitors: Applications in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.* **2019**, *62*, 5673–5724. [CrossRef]
67. Chenna, B.C.; Li, L.; Mellott, D.M.; Zhai, X.; Siqueira-Neto, J.L.; Calvet Alvarez, C.; Bernatchez, J.A.; Desormeaux, E.; Alvarez Hernandez, E.; Gomez, J.; et al. Peptidomimetic Vinyl Heterocyclic Inhibitors of Cruzain Effect Antitrypanosomal Activity. *J. Med. Chem.* **2020**, *63*, 3298–3316. [CrossRef]
68. Vankadara, S.; Dawson, M.D.; Fong, J.Y.; Oh, Q.Y.; Ang, Q.A.; Liu, B.; Chang, H.Y.; Koh, J.; Koh, X.; Tan, Q.W.; et al. A Warhead Substitution Study on the Coronavirus Main Protease Inhibitor Nirmatrelvir. *ACS Med. Chem. Lett.* **2022**, *13*, 1345–1350. [CrossRef] [PubMed]
69. Ludewig, S.; Kossner, M.; Schiller, M.; Baumann, K.; Schirmeister, T. Enzyme Kinetics and Hit Validation in Fluorimetric Protease Assays. *Curr. Top. Med.Chem.* **2010**, *10*, 368–382. [CrossRef] [PubMed]
70. Barthels, F.; Marincola, G.; Marciniak, T.; Konhäuser, M.; Hammerschmidt, S.; Bierlmeier, J.; Distler, U.; Wich, P.R.; Tenzer, S.; Schwarzer, D.; et al. Irreversible and Selective Inhibitors of Staphylococcus aureus Sortase A. *ChemMedChem* **2020**, *15*, 839–850. [CrossRef]
71. Amendola, G.; Ettari, R.; Previti, S.; Di Chio, C.; Messere, A.; Di Maro, S.S.; Hammerschmidt, J.; Zimmer, C.; Zimmermann, R.A.; Schirmeister, T.; et al. Lead Discovery of SARS-CoV-2 Main Protease Inhibitors through Covalent Docking-Based Virtual Screening. *J. Chem. Inf. Model.* **2021**, *61*, 2062–2073. [CrossRef]
72. Schirmeister, T.; Kesselring, J.; Jung, S.; Schneider, T.H.; Weickert, A.; Becker, J.; Lee, W.; Bamberger, D.; Wich, P.R.; Distler, U.; et al. Engels, Quantum Chemical-Based Protocol for the Rational Design of Covalent Inhibitors. *J. Am. Chem. Soc.* **2016**, *138*, 8332–8335. [CrossRef]
73. Caffrey, C.R.; Hansell, E.; Lucas, K.D.; Brinen, L.S.; Hernandez, A.A.; Cheng, J.; Roush, W.R.; Stierhof, Y.-D.; Bogyo, M.; Steverding, D.; et al. Active site mapping, biochemical properties and subcellular localization of rhodesain, the major cysteine protease of Trypanosoma brucei rhodesiense. *Mol. Biochem. Parasitol.* **2001**, *118*, 61–73. [CrossRef]
74. Berman, H.M. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [CrossRef]
75. Berman, H.; Henrick, K.; Nakamura, H. Announcing the worldwide Protein Data Bank. *Nat. Struct. Mol. Biol.* **2003**, *10*, 980. [CrossRef] [PubMed]
76. Halgren, T.A. MMFF94s option for energy minimization studies. *J. Comput. Chem.* **1999**, *20*, 720–729. [CrossRef]
77. *The PyMOL Molecular Graphics System*, version 2.5.2; Schrödinger, LLC: New York, NY, USA, 2021.
78. Reulecke, I.; Lange, G.; Albrecht, J.; Klein, R.; Rarey, M. Towards an Integrated Description of Hydrogen Bonding and Dehydration: Decreasing False Positives in Virtual Screening with the HYDE Scoring Function. *ChemMedChem* **2008**, *3*, 885–897. [CrossRef] [PubMed]
79. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Peterson, G.A.; Nakatsuji, H.; et al. *Gaussian 16 (Revision A.03)*; Gaussian Inc.: Wallingfort, CT, USA, 2016.

80. Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615. [CrossRef]
81. Marenich, A.V.; Cramer, C.J.; Truhlar, D.G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396. [CrossRef]
82. Pliego, J.R., Jr.; Riveros, J.M. Gibbs energy of solvation of organic ions in aqueous and dimethyl sulfoxide solutions. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1622–1627. [CrossRef]
83. Ben-Naim, A. Standard Thermodynamics of Transfer. Uses and Misuses. *J. Phys. Chem.* **1978**, *82*, 792–803. [CrossRef]
84. Spina, R.; Colacino, E.; Martinez, J.; Lamaty, F. Poly(ethylene glycol) as a Reaction Matrix in Platinum- or Gold-Catalyzed Cycloisomerization: A Mechanistic Investigation. *Chem.-A Eur. J.* **2013**, *19*, 3817–3821. [CrossRef]
85. Ho, A.; Cyrus, K.; Kim, K.-B. Towards Immunoproteasome-Specific Inhibitors: An Improved Synthesis of Dihydroeponemycin. *Eur. J. Org. Chem.* **2005**, *2005*, 4829–4834. [CrossRef]
86. St-Georges, C.; Désilets, A.; Béliveau, F.; Ghinet, M.; Dion, S.P.; Colombo, É.; Boudreault, P.-L.; Najmanovich, R.J.; Leduc, R.; Marsault, É. Modulating the selectivity of matriptase-2 inhibitors with unnatural amino acids. *Eur. J. Med. Chem.* **2017**, *129*, 110–123. [CrossRef] [PubMed]
87. Costanzo, M.J.; Yabut, S.C.; Almond, H.R.; Andrade-Gordon, P.; Corcoran, T.W.; de Garavilla, L.; Kauffman, J.A.; Abraham, W.M.; Recacha, R.; Chattopadhyay, D.; et al. Potent, Small-Molecule Inhibitors of Human Mast Cell Tryptase. Antiasthmatic Action of a Dipeptide-Based Transition-State Analogue Containing a Benzothiazole Ketone. *J. Med. Chem.* **2003**, *46*, 3865–3876. [CrossRef]
88. Engel-Andreasen, J.; Wellhöfer, I.; Wich, K.; Olsen, C.A. Backbone-Fluorinated 1,2,3-Triazole-Containing Dipeptide Surrogates. *J. Org. Chem.* **2017**, *82*, 11613–11619. [CrossRef] [PubMed]
89. Dutton, F.E.; Lee, B.H.; Johnson, S.S.; Coscarelli, E.M.; Lee, P.H. Restricted Conformation Analogues of an Anthelmintic Cyclodepsipeptide. *J. Med. Chem.* **2003**, *46*, 2057–2073. [CrossRef]

Supporting information

Investigation of the compatibility between warheads and peptidomimetic sequences of protease inhibitors - a comprehensive reactivity and selectivity study

Patrick Müller[1†], Mergim Meta[1†], J. Laurenz Meidner[1†], Marvin Schwickert[1], Jessica Meyr[2], Kevin Schwickert[1], Christian Kersten[1], Collin Zimmer[1], Stefan Hammerschmidt[1], Ariane Frey[1], Albin Lahu[1], Sergio de la Hoz-Rodríguez[3], Laura Agost-Beltrán[3], Santiago Rodríguez[3], Kira Diemer[2], Wilhelm Neumann[2], Florenci V. Gonzàlez[3], Bernd Engels[2] and Tanja Schirmeister[1*]

[1] Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Mainz, Germany

[2] Institute of Physical and Theoretical Chemistry, Julius-Maximilians-University of Wuerzburg, Wuerzburg, Germany

[3] Departament de Química Inorgànica I Orgànica, Universitat Jaume I, 12080 Castello, Spain

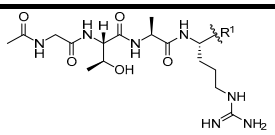
[*] Author to whom correspondence should be addressed.

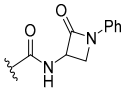
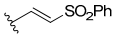
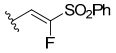
[†] These authors contributed equally to this work.

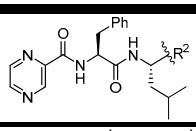
Content

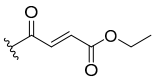
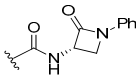
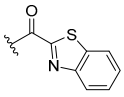
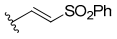
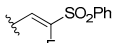
In vitro results of the synthesized compounds

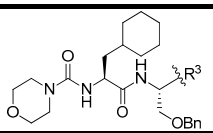
Table S1. *In vitro* results of the tested compounds against the corresponding protease. The peptidic sequence targeting the protease is shown on top.



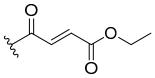
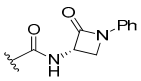
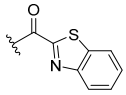
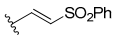
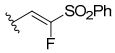
	K_i (µM)/ k_{inact} (s^{-1})/ k_{2nd} (M^{-1} s^{-1})				
R^1	**uPA**	**Proteasome β5-subunit**	**CatS**	**M^{Pro}**	**Rhodesain**
 **103**	0.141 ± 0.027 [a]	n.a.	n.a.	n.a.	n.a.
 **106**	n.a.	n.a.	n.a.	n.a.	n.a.
 **107**	n.a.	n.a.	n.a.	n.a.	n.a.

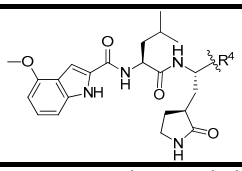


	K_i (µM)/ k_{inact} (s^{-1})/ k_{2nd} (M^{-1} s^{-1})				
R^2	**uPA**	**Proteasome β5-subunit**	**CatS**	**M^{Pro}**	**Rhodesain**
 **68**	n.a.	n.a.	n.a.*	n.a.	0.009 ± 0.001/ 0.006 ± 0.001/ 6.65 x 10^5 ± 1.57 x10^4
 **59**	n.a.	n.a.	0.619 ± 0.054 [a]	n.a.	0.002 ± 0.005 [a]
 **72**	n.a.	n.a.	n.a.	n.a.	n.a.
 **62**	n.a.	n.a.	n.a.	n.a.	n.a.
 **58**	n.a.	n.a.	0.126 ± 0.071/ 0.013 ± 0.002/ 1.00 x 10^5 ± 1.25 x10^4	n.a.	0.331 ± 0.041/ 0.019 ± 0.001/ 5.85 x 10^4 ± 1.66 x10^3
 **55**	n.a.	n.a.	0.934 ± 0.138 [a]	n.a.	0.462 ± 0.039 [a]
 **74**	n.a.	n.a.	0.238 ± 0.023[a]	n.a.	0.560 ± 0.267

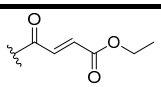
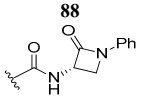
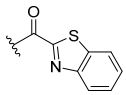
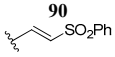


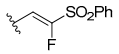
K_i (µM)/ k_{inact} (s^{-1})/ k_{2nd} (M^{-1} s^{-1})

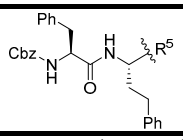
R^3	uPA	Proteasome β5-subunit	CatS	M^{Pro}	Rhodesain
 **40**	n.a.	n.a.	0.022 ± 0.004/ 0.0023 ± 0.00007/ 1.04 x 10^5 ± 1.04 x10^4	n.a.	0.035 ± 0.009/ 0.003 ± 0.0001/ 1.03 x 10^5 ± 1.09 x10^4
 **49**	n.a.	n.a.	0.012 ± 0.001 [a]	n.a.	0.002 ± 0.0002 [a]
 **45**	n.a.	n.a.	0.86 ± 0.13/ 0.00046 ± 0.00002/ 5.34 x 10^5 ± 8.29 x10^4	n.a.	n.a.
 **42**	n.a.	n.a.	0.233 ± 0.024 [a]	n.a.	0.324 ± 0.089 [a]
 **35**	n.a.	n.a.	0.003 ± 0.0006/ 0.02022 ± 0.0006/ 7.24 x 10^6 ± 2.36 x10^5	n.a.	0.022 ± 0.002/ 0.006 ± 0.0004/ 2.57 x 10^5 ± 1.04 x10^3
 **36**	n.a.	n.a.	3.215 ± 0.612 [a]	n.a.	n.a.
 **30**	n.a.	n.a.	0.001 ± 0.0001 [a]	n.a.	0.021 ± 0.002



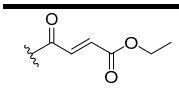
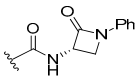
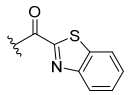
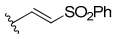
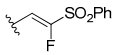
K_i (µM)/ k_{inact} (s^{-1})/ k_{2nd} (M^{-1} s^{-1})

R^4	uPA	Proteasome β5-subunit	CatS	M^{pro}	Rhodesain
 **81**	n.a.	-	n.a*	n.a*	0.099 ± 0.024/ 0.006 ± 0.0009/ 6.41 x 10^4 ± 3.00 x10^3-
 **88**	n.a.	n.a.	0.388 ± 0.277 [a]	2.197 ± 0.184 a)	0.002 ± 0.0001 [a]
 **96**	n.a.	n.a.	n.a.	n.a.	n.a.
 **90**	n.a.	n.a.	n.a.	0.026 ± 0.003 a)	n.a.
 **93**	n.a.	n.a.	0.330 ± 0.065/ 0.0087 ± 0.0001/ 2.62 x 10^4 ± 2.41 x10^3	3.531 ± 1.112/ 0.0137 ± 0.0061/ 2.79 x 10^3 ± 3.12 x10^2	2.002 ± 0.604/ 0.029 ± 0.008/ 1.49 x 10^4 ± 1.09 x10^4

 **94**	n.a.	n.a.	1.351 ± 0.09 [a]	0.517 ± 0.033 [a]	0.217 ± 0.031 [a]
 **84**	n.a.	n.a.	0.104 ± 0.009	0.031 ± 0,005 [a])	0.042 ± 0.005



K_i (µM)/ k_{inact} (s^{-1})/ k_{2nd} (M^{-1} s^{-1})

R^5	uPA	Proteasome β-5-subunit	CatS	M^{Pro}	Rhodesain
 **22**	n.a.	n.a.	0.018 ± 0.003/ 0.0022 ± 0.0001/ 1.24 x 10^5 ± 1.86 x10^4	n.a.	0.079 ± 0.007/ 0.008 ± 0.001/ 1.06 x 10^5 ± 4.57 x10^2
 **13**	n.a.	n.a.	0.201± 0.093 [a]	n.a.	0.060 ± 0.006 nM [a]
 **24**	n.a.	n.a.	n.a.	n.a.	n.a.
 **18**	n.a.	n.a.	n.a.	n.a.	n.a.
 **16**	n.a.	n.a.	0.001 ± 0.0006/ 0.0037 ± 0.0006/ 2.88 x 10^6 ± 8.56 x10^5	n.a.	0.020 ± 0.002/ 0.005 ± 0.0001/ 2.90 x 10^5 ± 3.17 x10^3
 **17**	n.a.	n.a.	0.966 ± 0.104)	n.a.	3.565 ± 0.437 [a]
 **25**	n.a.	n.a.	0.136 ± 0.012 [a]	n.a.	0.088 ± 0.0174 [a]

a) Calculated from IC_{50} value with the Cheng–Prusoff equation.
n.d. not determined.
n.a. not active.

Reactivity assay

Method A (^{1}H-NMR):

The corresponding Michael-acceptor (0.07 mmol, 1 eq) and the reference compound 1,3-dioxolan (0.5 eq) were dissolved in DMSO-d_6 (0.5 mL). The reaction was measured in time intervals before (0 min) and 5, 30, 60, 120 and 240 mins after the addition of the corresponding nucleophile (PhEtSH/EtONa, 2 eq) in presence or absence of Et_3N (2 eq). The CH_2-signal of 1,3-dioxlan was set at 5.3 ppm to get uniform shifts of the olefin signals in the ^{1}H-NMR spectra.

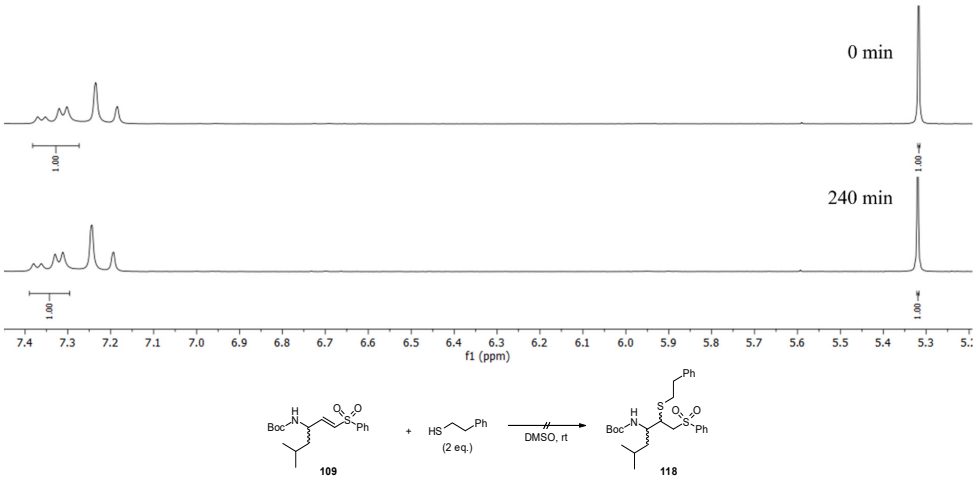


Figure S1. ^{1}H-NMR spectra of **109** after certain time intervals. Reaction of 2-phenylethanthiol with the vinyl sulfone moiety.

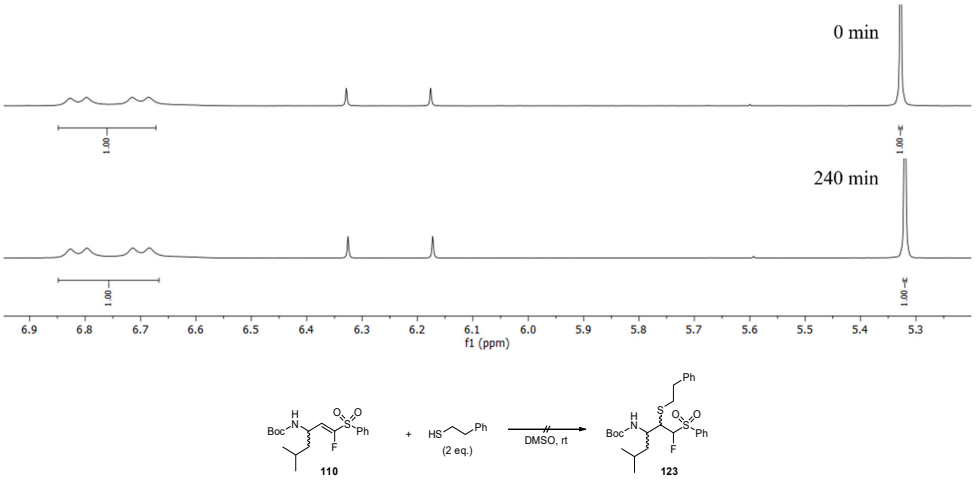


Figure S2. ^{1}H-NMR spectra of **110** after certain time intervals. Reaction of 2-phenylethanthiol with the F-vinyl sulfone moiety.

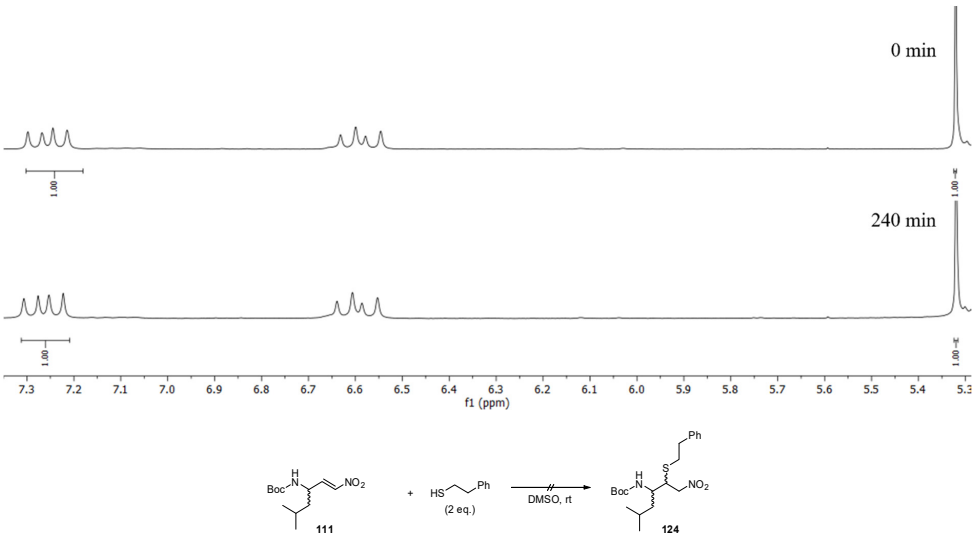


Figure S3. ^{1}H-NMR spectra of **111** after certain time intervals. Reaction of 2-phenylethanthiol with the nitroalkene moiety.

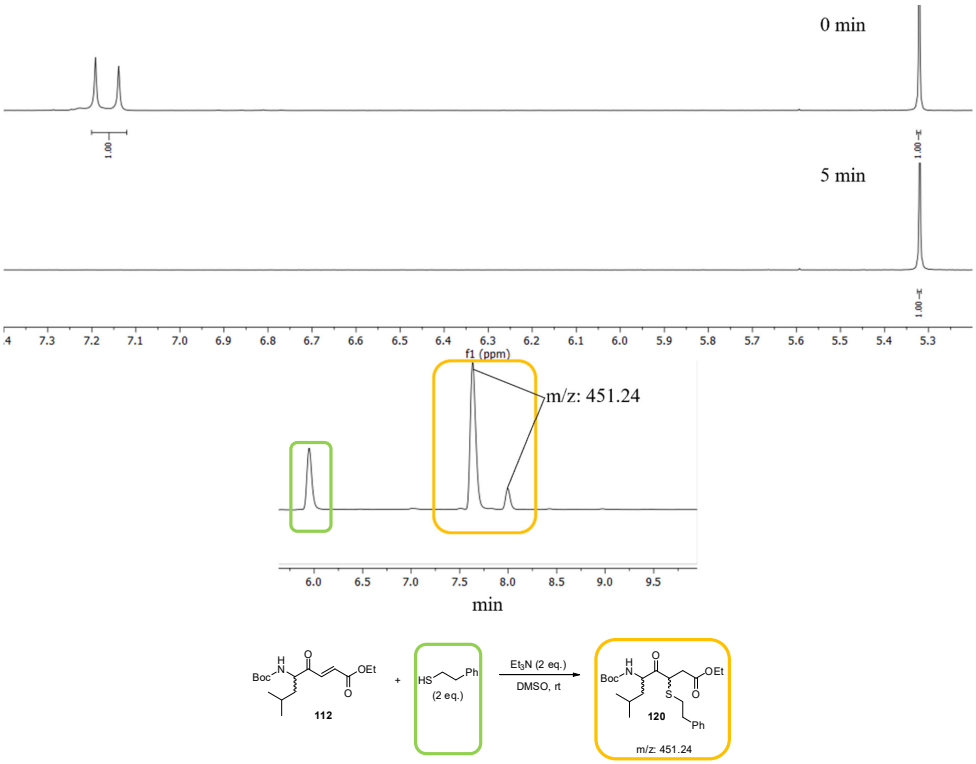


Figure S4. ^{1}H-NMR spectra of **112** after certain time intervals. Reaction of 2-phenylethanthiol with the 4-oxoenoate moiety in presence of Et_3N. LC/MS analysis of the same reaction at 10 min.

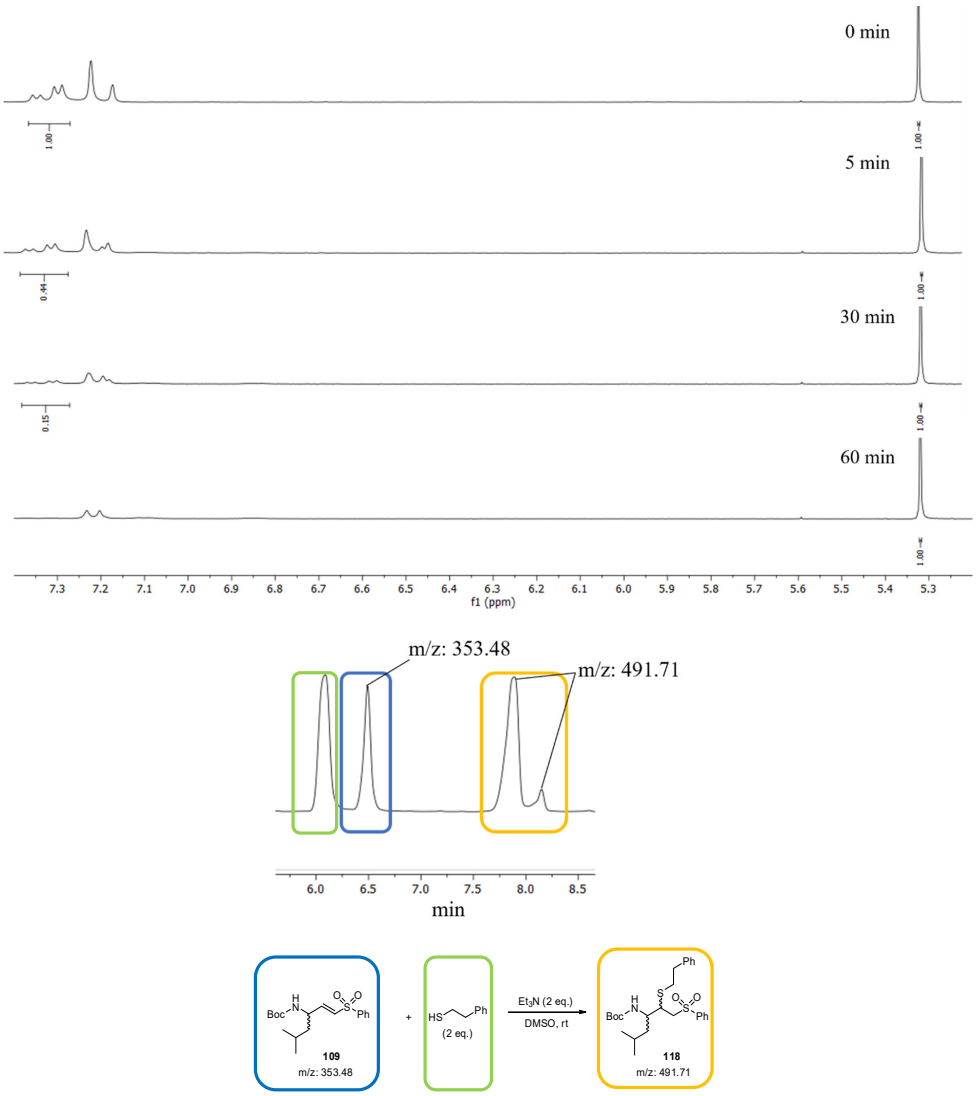


Figure S5. ^{1}H-NMR spectra of **119** after certain time intervals. Reaction of 2-phenylethanthiol with the vinyl sulfone moiety in presence of Et_3N. LC/MS analysis of the same reaction at 10 min.

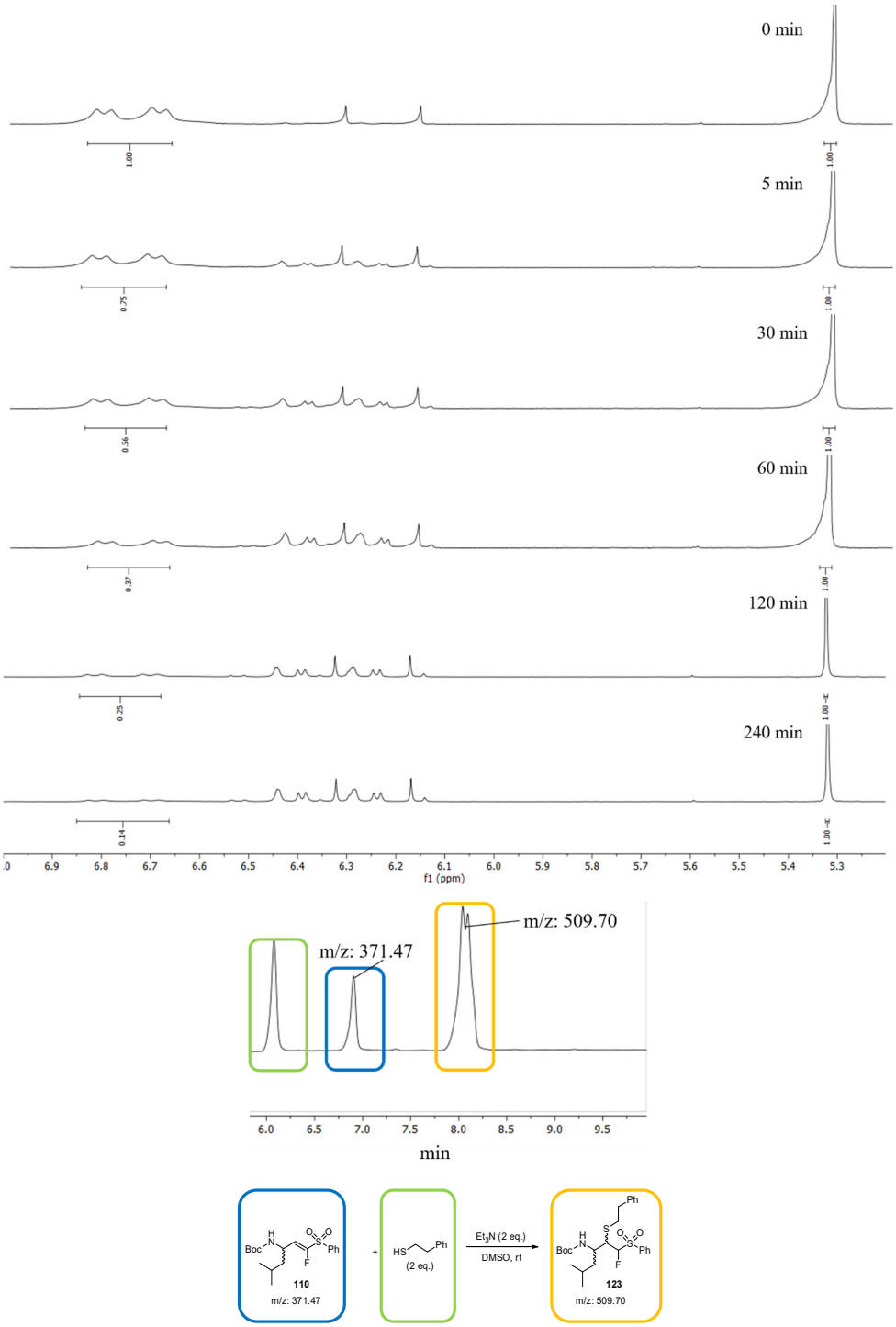


Figure S6. ^{1}H-NMR spectra of **110** after certain time intervals. Reaction of 2-phenylethanthiol with the F-vinyl sulfone moiety in presence of Et_3N. LC/MS analysis of the same reaction at 120 min.

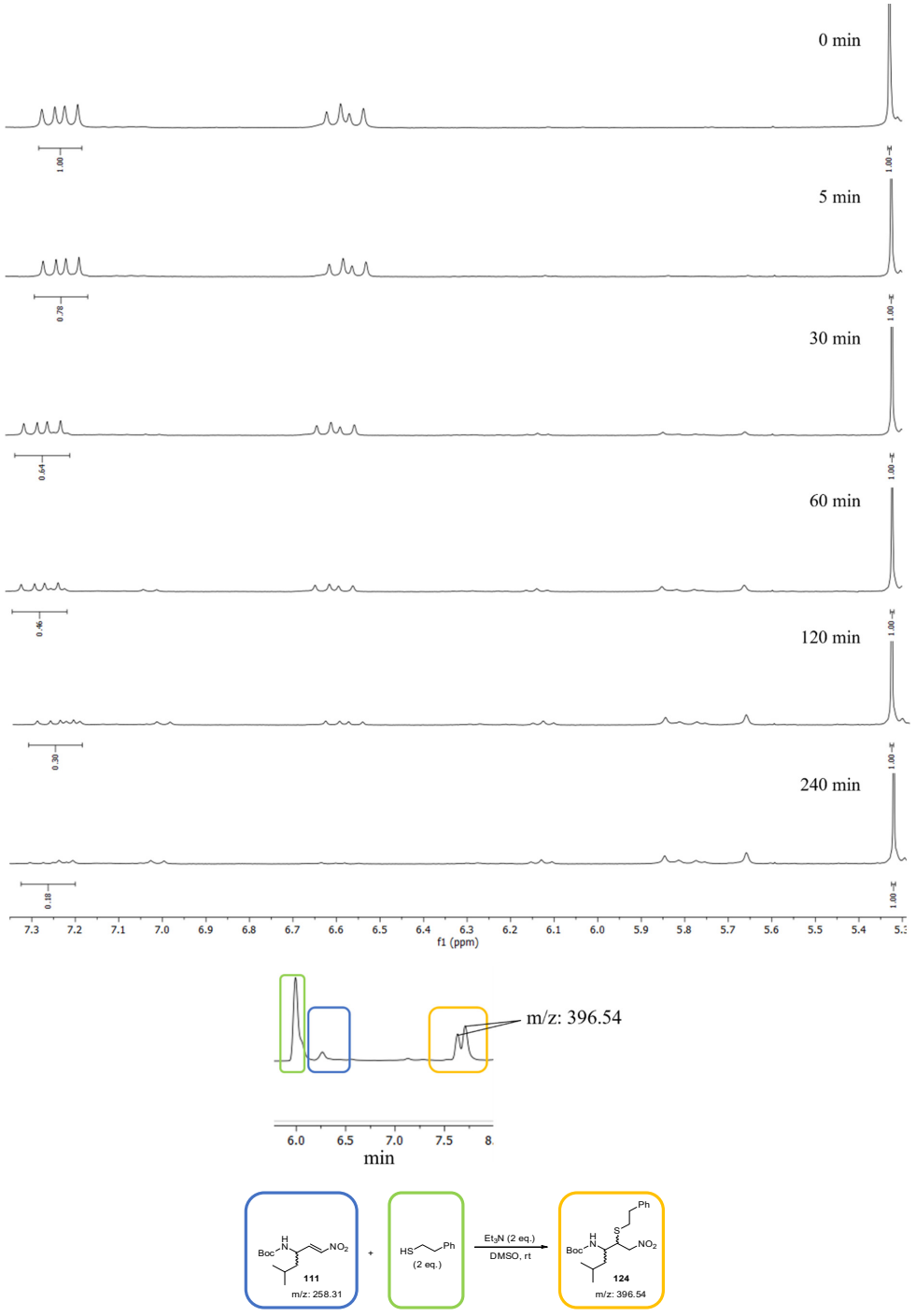


Figure S7. ^{1}H-NMR spectra of **111** after certain time intervals. Reaction of 2-phenylethanthiol with the nitroalkene moiety in presence of Et_3N. LC-MS analysis of the same reaction at 240 min. The blue marked peak was identified by NMR analysis as **111**.

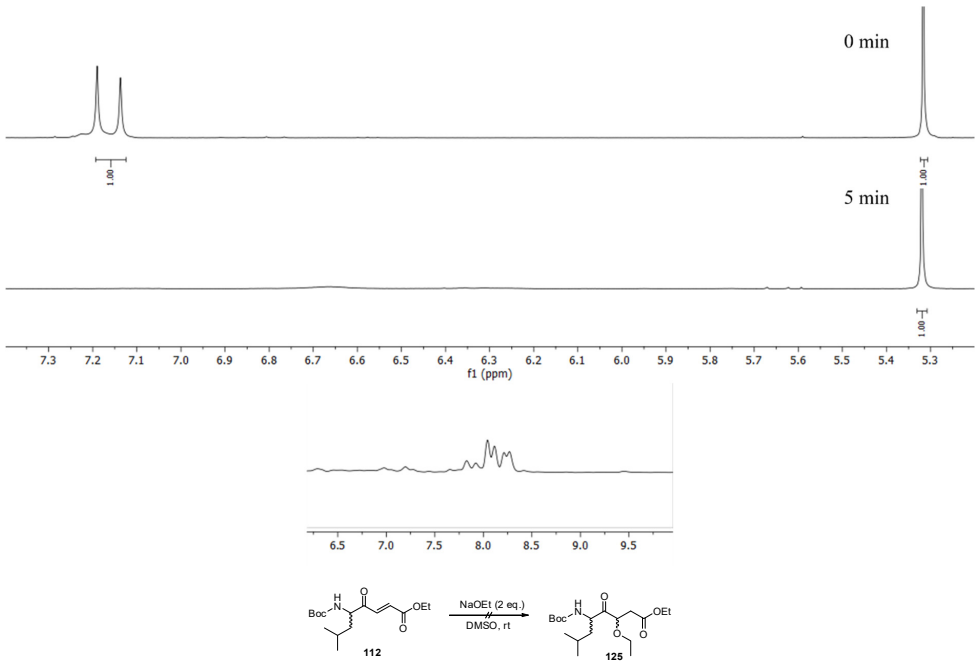


Figure S8. ^{1}H-NMR spectra of **112** after certain time intervals. Reaction of sodium ethanolate with the oxoenoate moiety. LC/MS analysis of the same reaction at 10 min.

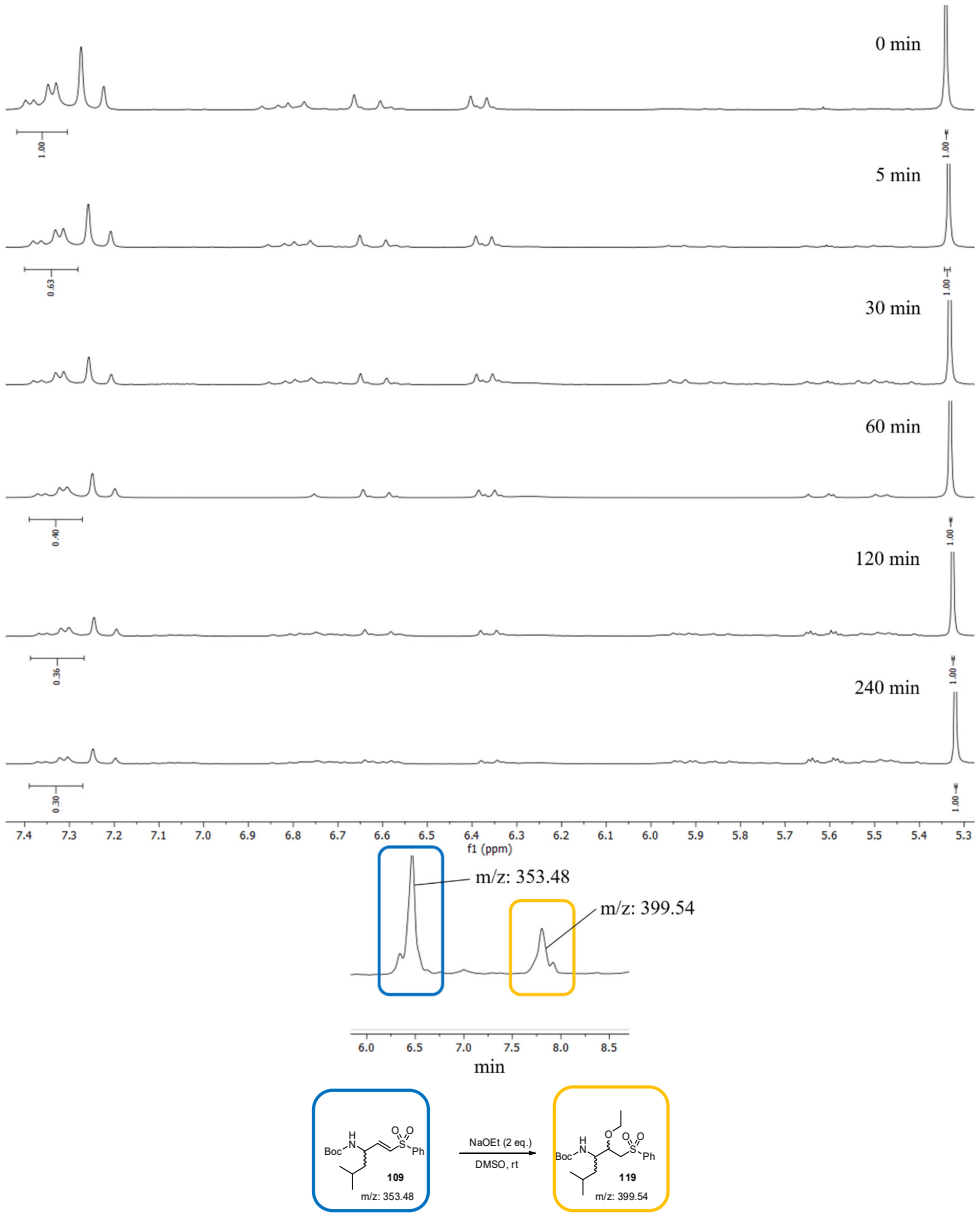


Figure S9. ^{1}H-NMR spectra of **109** after certain time intervals. Reaction of sodium ethanolate with the vinyl sulfone moiety. LC/MS analysis of the same reaction at 10 min.

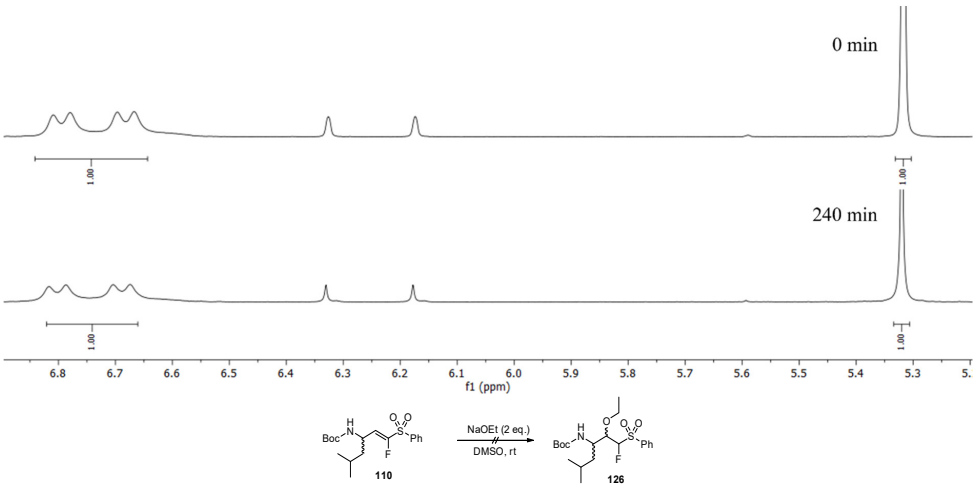


Figure S10. ^{1}H-NMR spectra of **110** after certain time intervals. Reaction of sodium ethanolate with the F-vinyl sulfone moiety.

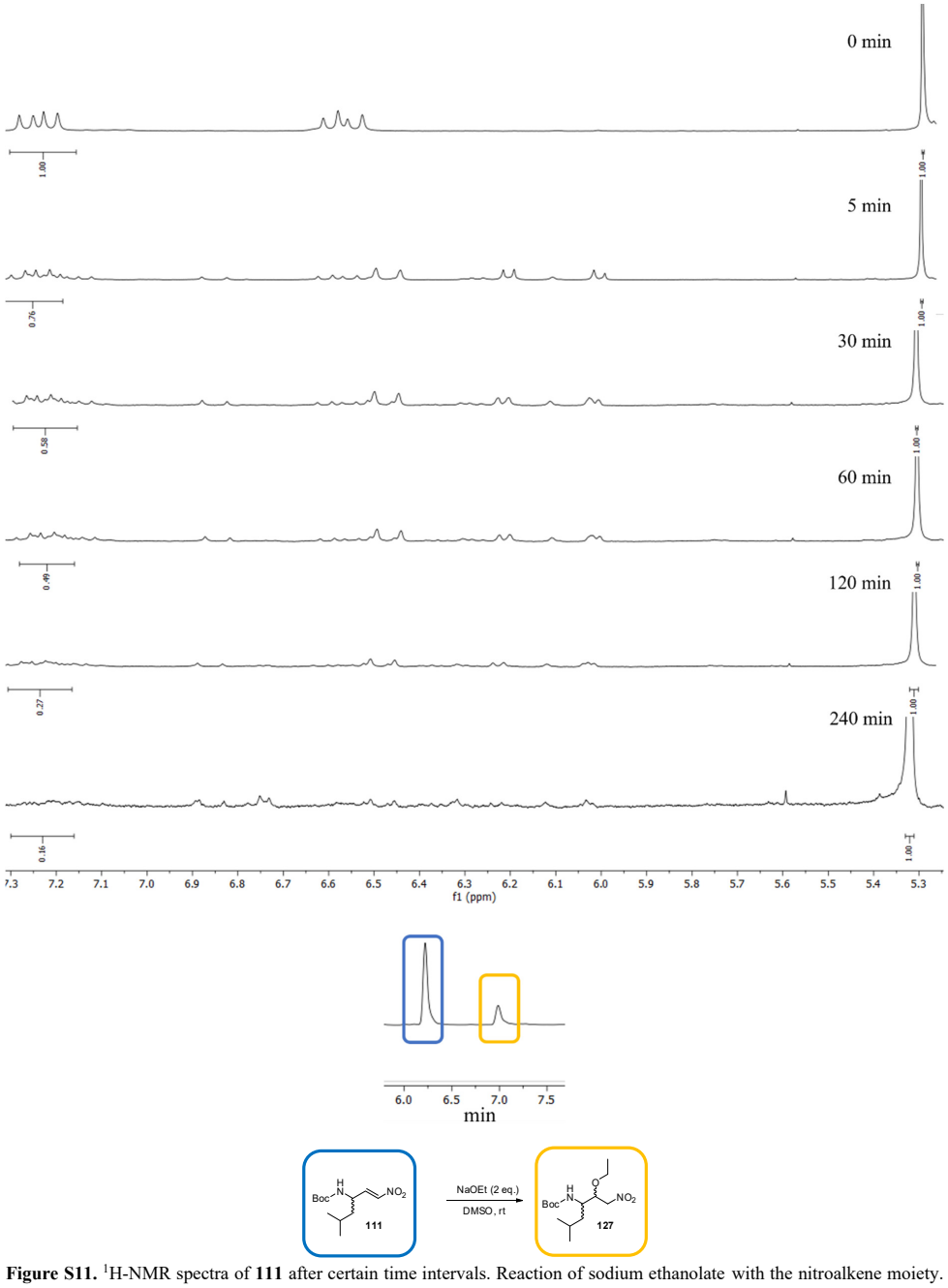


Figure S11. ^{1}H-NMR spectra of **111** after certain time intervals. Reaction of sodium ethanolate with the nitroalkene moiety. LC/MS analysis of the same reaction at 10 min. The blue marked peak was identified by NMR analysis as **111**. The yellow marked peak corresponds to the supposed product **127**, identification with LC-MS was not possible.

Method A (^{13}C-NMR):

The α-ketobenzothiazole or nitrile **115**, **117** (0.07 mmol, 1 eq) was dissolved in DMSO-d_6 (0.5 mL). The reaction was measured in time intervals before (0 min), 5, 30, 60, 120 and 240 mins after the addition of the corresponding nucleophile (PhEtSH/EtONa, 2 eq) in presence or absence of Et_3N (2 eq). The reference signal of DMSO-d_6 was set at 39.52 ppm.

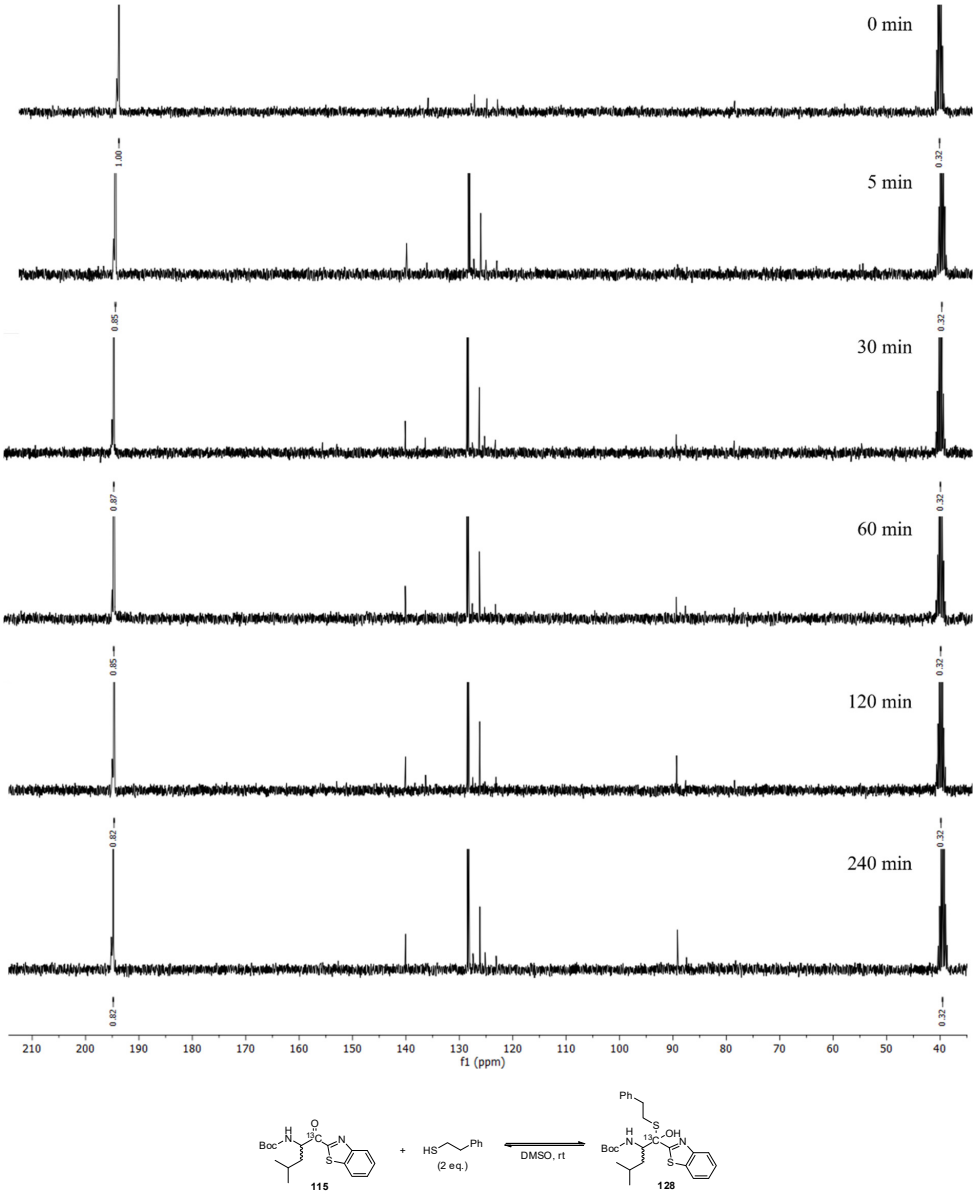


Figure S12. ^{13}C-NMR spectra of **115** after certain time intervals. Reaction of 2-phenylethanthiol with the ketobenzothiazole moiety.

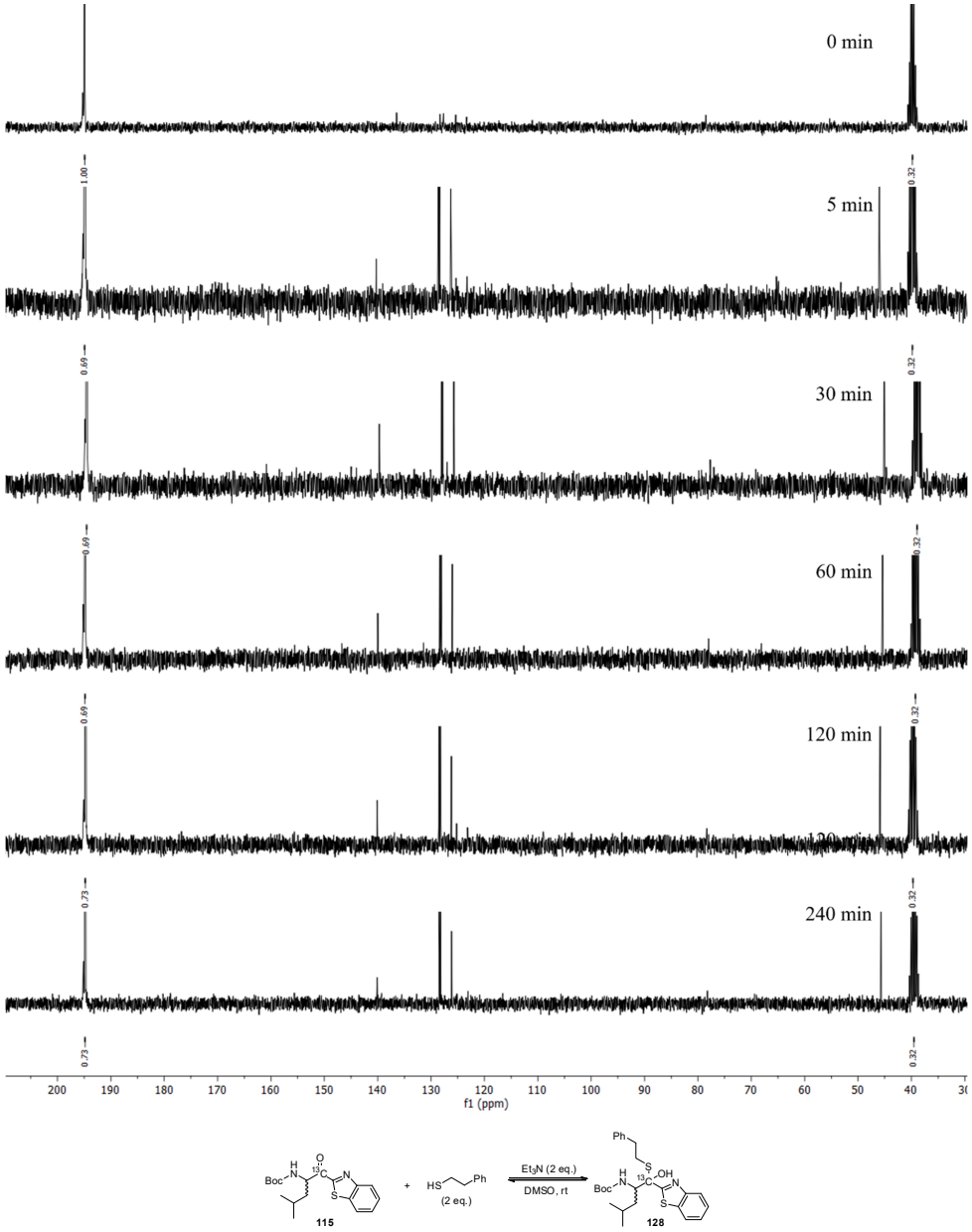
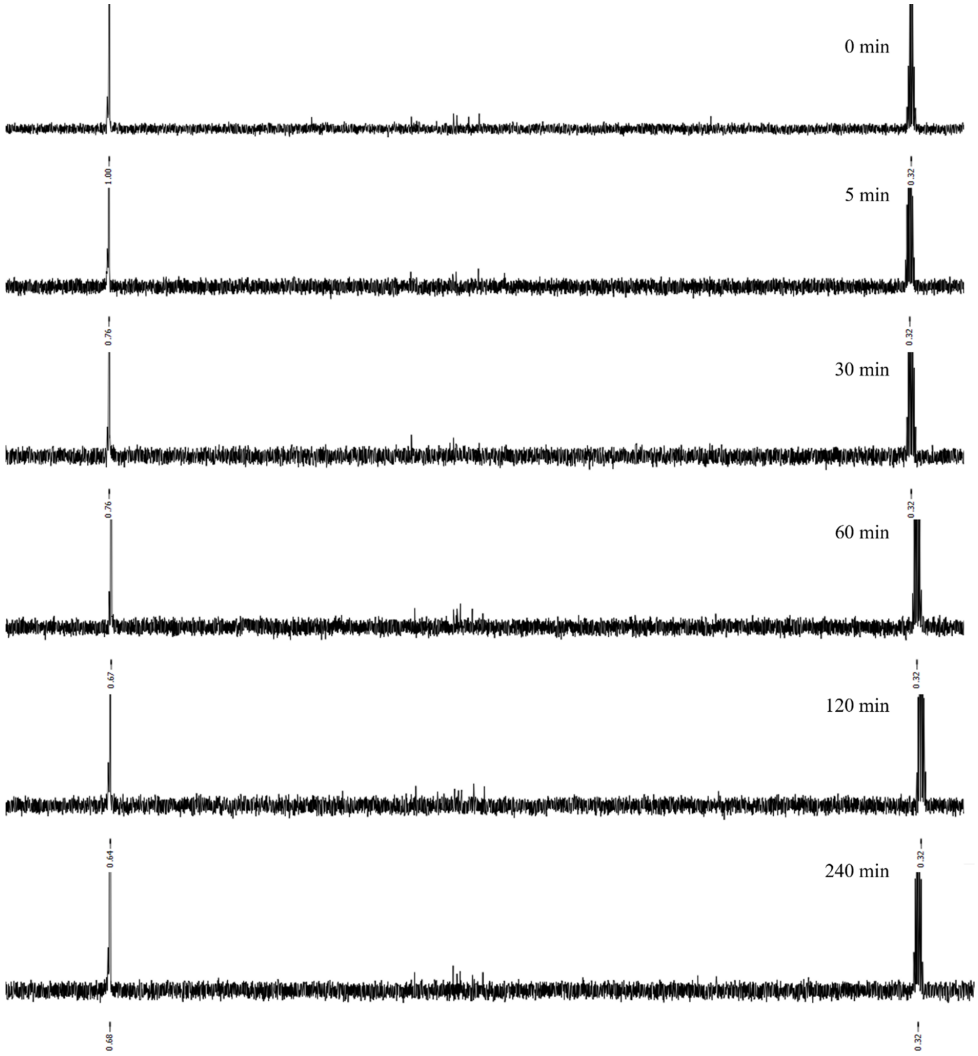


Figure S13. ^{13}C-NMR spectra of **115** after certain time intervals. Reaction of 2-phenylethanthiol with the ketobenzothiazole moiety in presence of Et_3N.


0 min
5 min
30 min
60 min
120 min
240 min

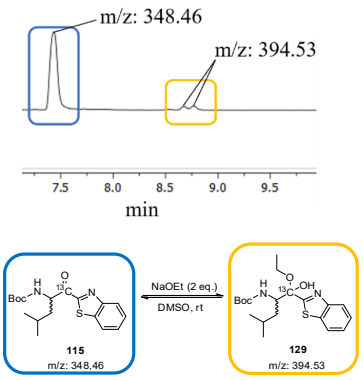

m/z: 348.46
m/z: 394.53
7.5
8.0
8.5
9.0
9.5
min
NaOEt (2 eq.)
DMSO, rt
115
m/z: 348,46
129
m/z: 394.53

Figure S14. ^{13}C-NMR spectra of **115** after certain time intervals. Reaction of sodium ethanolate with the *α*-ketobenzothiazole moiety. LC/MS analysis of the same reaction at 10 min.

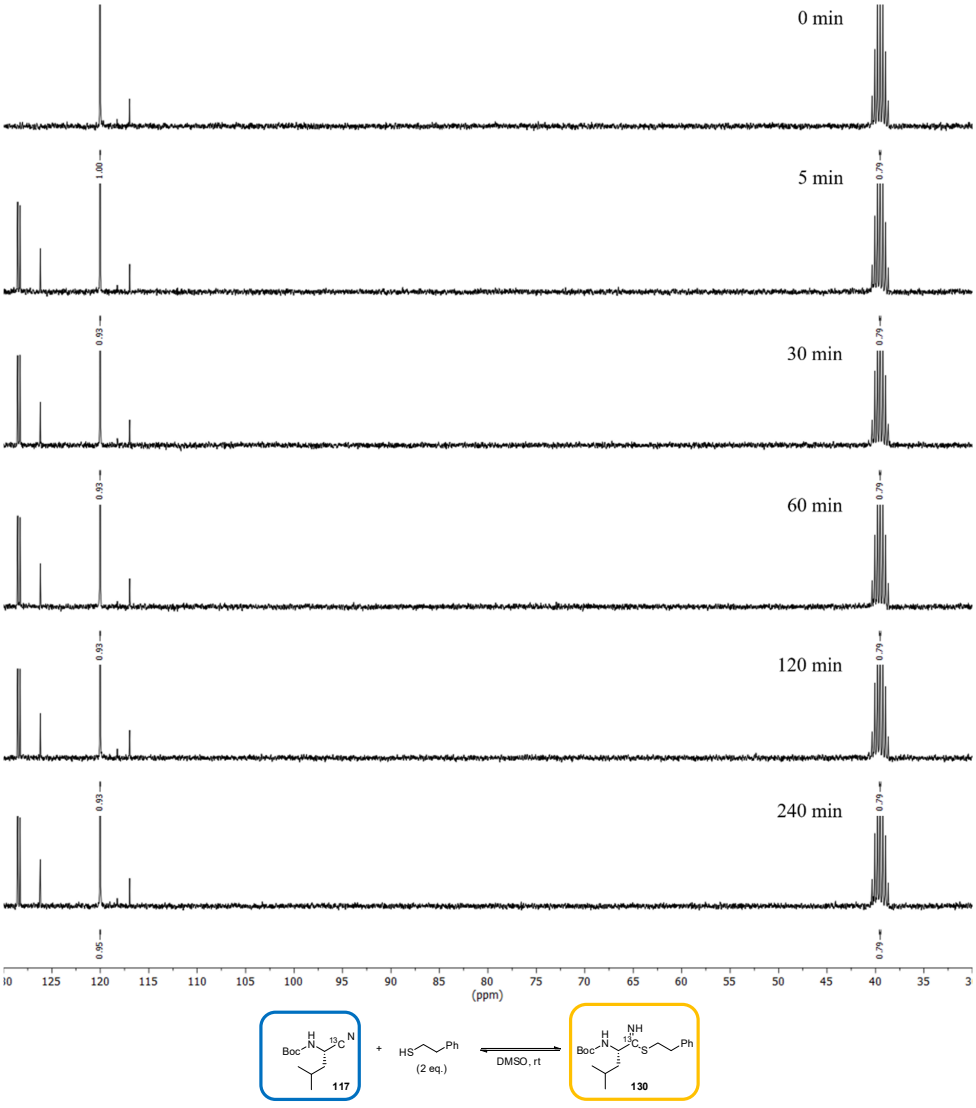


Figure S15. ^{13}C-NMR spectra of **117** after certain time intervals. Reaction of 2-phenylethanthiol with the nitrile moiety. The blue marked peak was identified by NMR analysis as **117**. The yellow marked peak corresponds to the supposed product **130**, identification with LC-MS was not possible.

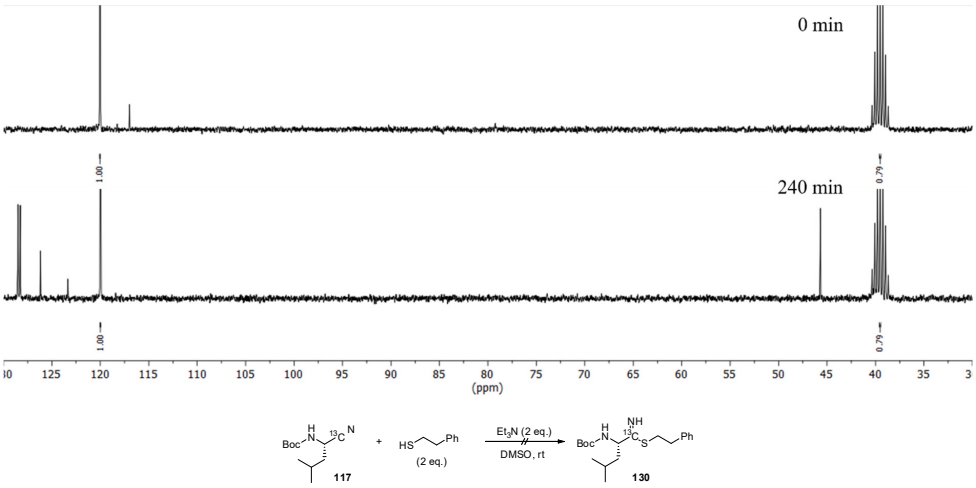


Figure S16. ^{13}C-NMR spectra of **117** after certain time intervals. Reaction of 2-phenylethanthiol with the nitrile moiety in presence of Et_3N.

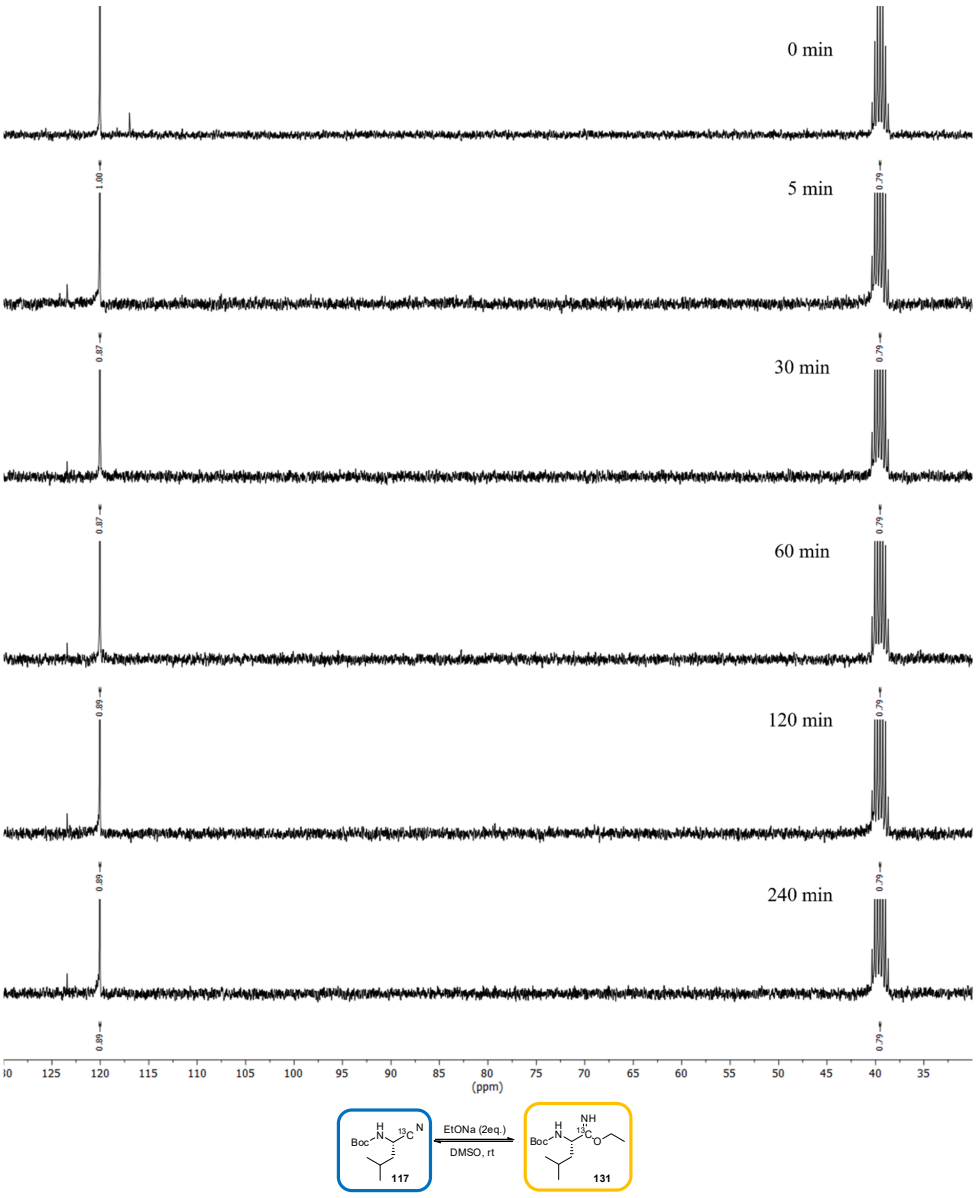


Figure S17. ^{13}C-NMR spectra of **117** after certain time intervals. Reaction of sodium ethanolate with the nitrile moiety. The blue marked peak was identified by NMR analysis as **117**. The yellow marked peak corresponds to the supposed product **131**, identification with LC-MS was not possible.

Fluorometric inhibition assays

The inhibitory activities of the compounds against the proteases were determined with assays based on fluorogenic or FRET-based substrates. The fluorescence was measured in a white flat-bottom 96-well plate from Greiner Bio-One using a Tecan Infinite F2000 Pro plate reader. Measurements were performed as triplicates. The substrate and the compounds were prepared as stock solutions in DMSO. Each well contained a total volume of 200 µL, consisting of 185 µL buffer, 5 µL inhibitor in DMSO or pure DMSO as negative control, 5 µL substrate in DMSO and 5 µL enzyme solution in buffer. Dilution series between 200 µM and 50 pM were prepared for the determination of the inhibition constants. The fluorescence signal was measured every 30 s for 10 min at 25 °C with the corresponding excitation and emission wavelengths. IC_{50} values for the reversible inhibitors were calculated by fitting the remaining enzymatic activity to the four parameter IC_{50} equation with Y [ΔF/min] as the substrate hydrolysis rate, Y_{max} as the maximum value of the dose response curve at inhibitor concentrations [I] = 0 µM, Y_{min} as the minimum value at high inhibitor concentrations and *s* as the hill coefficient.[1] The sigmoidal dose-response curves of the compounds **30** and **94** against rhodesain are shown exemplary in Figure **S19**.

$$y = \frac{y_{\max} - y_{\min}}{1 + \left(\frac{[\mathrm{I}]}{\mathrm{IC}_{50}}\right)^{s}} + y_{\min}$$

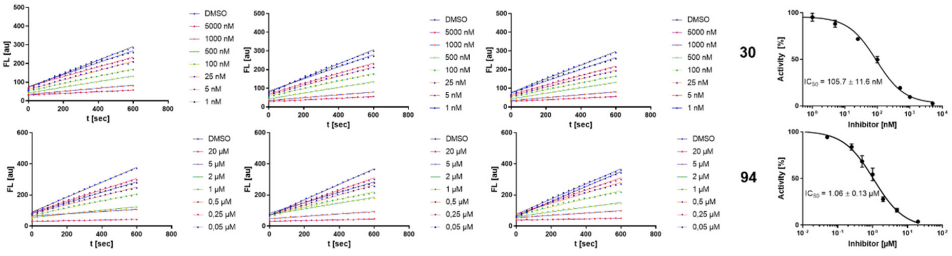


Figure S19. (left) Fluorescence progress curves for **30** and **94** against rhodesain. (right) Plots showing the respective IC_{50} values from sigmoidal fits.

Due to the dependence of the IC_{50} value on the substrate affinity and concentration, the K_i values were calculated with the Cheng-Prusoff equation (1) for appropriate comparison of the inhibitory activities to the other enzymes and inhibitors.[1]

$$K_I = \frac{\mathrm{IC}_{50}}{1 + \frac{[S]}{K_\mathrm{M}}} \qquad 1$$

To evaluate the inhibitory activities of the irreversible inhibitors, reporting IC_{50} values is less suitable, because the IC_{50} values are heavily depending on the incubation time of the enzyme and inhibitor. Therefore, the dissociation constant of the reversible enzyme-inhibitor complex K_I, the maximum inactivation rate k_{inact} and the second-order rate constants of inhibition k_{2nd} were determined. The substrate-conversion plots in presence of the compounds **16**, and **35** against rhodesain are indicating a time-dependent inhibition as shown exemplary in Figure **S20**. It is shown, that the apparent first-order rate constant k_{obs} changed hyperbolically with the inhibitor concentration. A limiting value was approached at higher inhibitor concentrations, indicating a two-step inhibition-mechanism.[1,2]

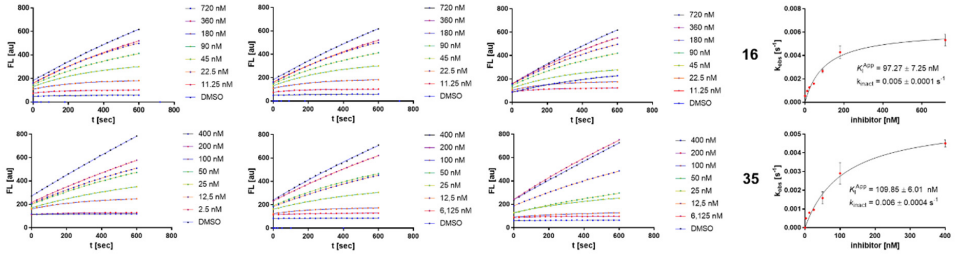


Figure S20. (left) Fluorescence progress curves for **16** and **35** against rhodesain. (right) k_{obs} against [I] plots showing the respective K_I^{App} and k_{inact} values from hyperbolic fits.

For determination of the k_{obs} values, progress curves were analyzed by non-linear regression analysis up to 10 min with the exponential equation (2). F represents the fluorescence intensity, which is proportional to the product concentration [P] and offset represents the background fluorescence.

$$F = [P]^{\propto}(1 - e^{-k_{obs} \cdot t}) + \text{offset} \qquad 2$$

The received k_{obs} values were fitted against the inhibitor concentrations using the hyperbolic equation (3) to get the k_{inact} and K_I^{app} values. For compound **93** no k_{obs} saturation was observed. A Lineweaver-Burk linearization (4) was used to receive K_I^{app} and $k_{inact.}$

$$k_{obs} = \frac{k_{inact}[I]}{K_I^{App} + [I]} \qquad 3$$

$$\frac{1}{k_{obs}} = \frac{K_I^{App}}{k_{inact}[I]}\frac{1}{[I]} + \frac{1}{k_{inact}} \qquad 4$$

The K_I^{app} values were corrected to the K_I values by the Cheng-Prussoff equation (5), for better comparison of the inhibitory constants to the other enzymes and compounds.

$$K_{\mathrm{I}} = \frac{K_{\mathrm{I}}^{\mathrm{App}}}{1 + \frac{[S]}{K_{\mathrm{M}}}} \qquad 5$$

The second-order rate constant of inhibition k_{2nd} were calculated from the equation (6).[1]

$$K_{2\mathrm{nd}} = \frac{k_{\mathrm{inact}}}{K_{\mathrm{I}}} \qquad 6$$

Buffers and substrates

The following buffers and substrates were used for the respective assays: Rhodesain (50 mM Na-acetate pH 5.5, 5 mM EDTA, 200 mM NaCl, 50 µM DTT, 0.005% Brij 10 µM Cbz-Phe-Arg-AMC, 1 nM rhodesain), Cathepsin S (50 mM KH_2PO_4, pH 6.5, 50 mM K_2HPO_4, 2.5 mM DTT, 2.5 mM EDTA, Z-Val-Arg-Arg-AMC, 10 nM CatS), β-5-subunit of the proteasome (50 mM Tris HCl, 25 mM NaCl, 10 mM NaCl, $MgCl_2 \cdot 6\ H_2O$, 100 µM Succ-Leu-Leu-Val-Tyr-AMC, 0.02 mg/mL proteasome β5-subunit), Sars-Cov-2 M^{Pro} (20 mM TRIS, pH 7.5, 0.1 mM, NaCl 200 mM, DTT 1 mM, 50 nm M^{Pro}), an internally quenched 14-mer fluorogenic (FRET) peptide (DABCYL-KTSAVLQSGFRKME-EDANS)), uPA (50 mM Tris HCl, 150 mM pH 7.5, NaCl, 10 mM $CaCl_2$, 0.005% TX-100, 240µM Cbz-Gly-Gly-Arg-AMC, 2.5 U uPA). As positive controls, several well-known inhibitors have been used for the respective enzymes and their biological activity was reproduced.

The human uPa and recombinant CatS were purchased from Sigma Aldrich, and the proteasome β5-subunit from Enzo Life sciences. Rhodesain and SARS-CoV-2 M^{pro} were expressed under the conditions described below.

Protein Expression and purification

SARS-CoV-2 M^{pro}.

The expression of SARS-CoV-2 Mpro was performed as described previously.[3] Briefly, the pMal-c2 plasmid (New England Biolabs, Ipswich, MA, USA), harboring the DNA of the entire SARS-CoV-2 M^{pro} coding sequence framed by a short sequence specifying the 5 *C*-terminal residues of nonstructural protein 4 (nsp4/nsp5 cleavage site) at the 5' end and the sequence of a hexahistidine tag at the 3' end. The presence of the nsp4/nsp5 cleavage site between the plasmid's MBP and the M^{pro} sequence together with the native nsp5/nsp6 cleavage site between

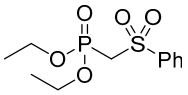
Synthesis

General Methods and Materials:

All reagents and solvents were purchased commercially and used as provided by the supplier without further purification. Solvents for synthesis, extraction, and chromatography were of analytical grade. Moisture-sensitive reactions were carried out under argon atmosphere, and anhydrous solvents were used as provided by the commercial supplier. Reaction progress was monitored by thin-layer chromatography using Alugram Xtra F254 silica plates from Macherey-Nagel and/or LC-MS. Therefore, an Agilent 1100 series HPLC system and an Agilent Poroshell 120 EC-C18, 150 x 2.10 mm, 4 μm column coupled to an Agilent 1100 series LC/MSD Trap with electron spray ionization (ESI), was used. The identities and purities of compounds were determined by the same LC-MS system with a gradient of acetonitrile and water (+0.1% formic acid). Signals were detected at 210/254 nm with quantitation by AUC and masses were determined in positive ionization mode (ESI). HPLC purification was performed with the Agilent 1290 II Infinity Preparative LC System using an InfinityLab Pursuit XRs C18, 30 x 250mm, 5 µm, preparative LC column. Flash chromatography was performed with the Biotage Isolera^{TM} One system using prepacked columns from Biotage. Silica gel (0.040 – 0.063 mm) from Macherey-Nagel was used for column chromatography. Optical rotations $[\alpha]_D^{22}$ were measured on an P3000 polarimeter from Krüss at 22 °C and are reported in ml·dm^{-1}·g^{-1} with the concentration c being g/100 ml. Fourier-transformed ATR-corrected IR spectra were measured on an Avatar 330 single crystal spectrometer from ThermoNicolet. Melting points (uncorrected) were measured with an MPM-H3 using semi-open capillaries. NMR spectra were recorded as stated individually on Bruker Fourier 300 MHz, Bruker Avance DSX 400 MHz and Bruker Avance III 600 MHz. Chemical shifts are indicated in parts per million (ppm), with the solvent resonance ($CDCl_3$, DMSO-d_6 or CD_3OD from Deutero GmbH) as internal standard. The purity of all compounds tested in biological assays was ≥95% as determined by LC-MS.

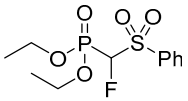
Synthesis of the vinyl sulfone precursors[4]

3, Diethyl ((phenylsulfonyl)methyl)phosphonate



Methyl phenyl sulfone (1.5 g, 9.9 mmol, 1 eq) was dissolved in dry THF (30 mL) under argon atmosphere and cooled to 0 °C. Subsequently, *n*-BuLi (2.5 M in hexanes, 9.6 mL, 24 mmol, 2.5 eq) was added dropwise and the reaction mixture stirred for one hour at 0 °C. Afterwards, diethyl chlorophosphate (2.4 mL, 16 mmol 1.2 eq) was added and stirring was continued for one hour. The reaction was quenched with the addition of saturated NH_4Cl (10 mL) and diluted with DCM (30 mL). The aqueous phase was extracted with DCM (3x 20 mL) and the combined organic layers were dried over anhydrous Na_2SO_4. The solvent was removed under reduced pressure and the crude product was purified by column chromatography (CH/EA 4:1) to yield the desired product as a colorless solid (1.47 g, 5 mmol, 51%). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 8.01 – 7.95 (m, 2H), 7.70 – 7.63 (m, 1H), 7.61 – 7.53 (m, 2H), 4.20 – 4.09 (m, 4H), 3.75 (d, J = 16.9 Hz, 2H), 1.28 (td, J = 7.1, 0.6 Hz, 6H). ^{13}C NMR (75 MHz, $CDCl_3$): δ/ppm = 140.0, 134.2, 129.2, 128.4, 63.5, 54.8, 52.9 (d, J = 6.5 Hz), 23.6 16.3. FT-IR: ν/cm^{-1} = 2988, 2898, 1324, 1257, 1155, 1018, 970, 796, 751, 683, MS (ESI) *m/z* calculated for $[C_{11}H_{18}O_5PS]^+$ ($[M+H]^+$): 293.1, found 293.0.

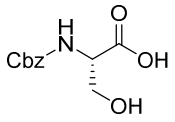
4, Diethyl (fluoro(phenylsulfonyl)methyl)phosphonate



3 (1 g, 4.45 mmol, 1 eq) was dissolved in dry THF (15 mL) and cooled to −78 °C. LiHMDS (1 M in THF, 5.6 mL, 5.56 mmol, 1.25 eq) was added and stirring was continued for 30 min at −78 °C. Afterwards, Selectfluor® (2.4 g, 6.66 mmol, 1.5 eq) in DMF (10 mL) was added dropwise and the mixture stirred for two hours. The reaction was quenched with saturated NH_4Cl solution (15 mL) and diluted with EtOAc (30 mL). The aqueous phase was extracted with EtOAc (3x 20 mL) and the combined organic phases were dried over anhydrous Na_2SO_4 and the solvent was removed under reduced pressure. The crude product was purified by column chromatography (CH/EA 4:1) to yield the desired product as a colorless solid (0.81 g, 2.6 mmol, 58%). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 8.03 – 7.96 (m, 1H), 7.75 – 7.68 (m, 1H), 7.60 (dd, J = 8.3, 6.9 Hz, 2H), 5.38 (dd, J = 45.5, 6.6 Hz, 1H), 4.35 – 4.19 (m, 4H), 1.34 (t, J = 7.1 Hz, 6H). ^{13}C NMR (75 MHz, $CDCl_3$): δ/ppm = 136.3, 135.0, 129.9, 129.3, 100.0, 97.9, 97.0, 94.9, 65.0 (t, J = 7.1 Hz), 16.41 (d, J = 5.9 Hz). mp: 61 – 65 °C. FT-IR: ν/cm^{-1} = 2990, 2907, 1449, 1331, 1263, 1221, 1159, 1061, 1006, 967, 852, 784, 755, 716, 699, 680. MS (ESI) *m/z* calculated for $[C_{11}H_{17}FO_5PS]^+$ ($[M+H]^+$): 311.0, found 311.0.

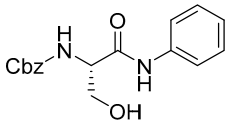
Synthesis of the *β*-lactam precursor[22]

6, (*S*)-2-(((Benzyloxy)carbonyl)amino)-3-hydroxypropanoic acid



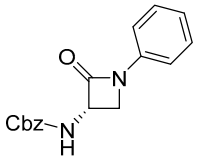
H-Ser-OH (1 g, 9.6 mol, 1 eq) and $NaHCO_3$ (2 g, 24 mmol, 2.5 eq) were dissolved in water (20 mL) and cooled to 0 °C. Benzyl chloroformate (2.01 mL, 14.4 mmol, 1.5 eq) was added dropwise. The reaction mixture was stirred for two hours at 0 °C and overnight at room temperature. The aqueous solution was acidified to pH = 2 with hydrochloric acid (1 M) and then extracted with EtOAc (3x, 10 mL). The combined organic layers were dried over Na_2SO_4, filtered and the solvent was removed under reduced pressure to yield the desired product as a colorless solid (2 g, 8.64 mmol, 90%). ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 7.40 – 7.25 (m, 5H), 5.04 (s, 2H), 4.06 (dt, J = 8.2, 5.1 Hz, 1H), 3.66 (d, J = 5.1 Hz, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ/ppm = 172.2, 156.1, 137.0, 128.4, 127.8, 65.5, 61.3, 56.7. mp: 115 – 116 °C. $[\alpha]_D^{20}$ = +9 (c 1.00, DMSO). FT-IR: ν/cm^{-1} = 3320, 2935, 1710, 1690, 1567, 1477, 1285, 996, 821, 755. MS (ESI) *m/z* calculated for $[C_{11}H_{14}NO_5]^+$ ($[M+H]^+$): 240.1, found 262.1.

7, (*S*)-Benzyl (3-hydroxy-1-oxo-1-(phenylamino)propan-2-yl)carbamate



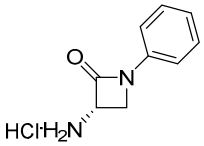
6 (3.26 g, 13.6 mmol, 1 eq) was dissolved in EtOAc (50 mL) under argon atmosphere and cooled to 0 °C. NMM (1.65 mL, 15 mmol, 1.1 eq) and TBTU (5.3 g, 16.3 mmol, 1.2 eq) were added and stirring was continued for 30 min at 0 °C. Afterwards, aniline (1.4 mL, 15 mmol, 1.1 eq) was added and the reaction mixture stirred for 12 h. The mixture was evaporated under reduced pressure and the crude product was purified by column chromatography (CH/EA 1:1) to yield the desired product as a colorless oil (3.32 g, 10.06 mmol, 74%). ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 7.65 – 7.58 (m, 2H), 7.39 – 7.25 (m, 7H), 7.09 – 7.01 (m, 1H), 5.07 – 5.03 (m, 2H), 4.94 (dt, J = 10.7, 5.8 Hz, 1H), 4.24 (dt, J = 7.9, 5.8 Hz, 1H), 3.66 (dq, J = 10.7, 5.5, 4.8 Hz, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ/ppm = 169.1, 155.9, 138.9, 136.9, 128.6, 128.3, 127.8, 127.7, 123.3, 119.2, 65.5, 61.7, 57.7. $[\alpha]_D^{20}$ = +11 (c 1.00, DMSO). FT-IR: ν/cm^{-1} = 2241, 2108, 1740, 1689, 1510, 1341, 1049, 1015, 832, 754, MS (ESI) *m/z* calculated for $[C_{17}H_{19}N_2O_4]^+$ ($[M+H]^+$): 315.1, found 315.1.

8, (S)-Benzyl (2-oxo-1-phenylazetidin-3-yl)carbamate



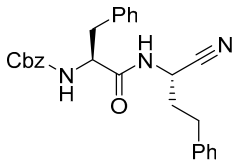
7 (1.5 g, 4.8 mmol, 1 eq) was dissolved in DMF (25 mL) under argon atmosphere and cooled to 0 °C. 1,1-sulfonyldiimidazole (1.42 g, 7.20 mmol, 1.5 eq) was added and stirred for 30 min at 0 °C. Afterwards, the reaction mixture was cooled to –20 °C and NaH (60%, 0.20 g, 7.20 mmol, 1.5 eq) was added portion wise. After additional 30 min of stirring, mixture of MeOH (0.04 mL) and water (20 mL) were added. The precipitate was filtered under reduced pressure and washed with cold water (15 mL) to yield the desired product as a white solid (1.1 g, 3.70 mmol, 77%). ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 8.09 (d, J = 8.4 Hz, 1H), 7.39 – 7.33 (m, 9H), 7.13 – 7.07 (m, 1H), 5.06 (s, 2H), 4.88 (ddd, J = 8.8, 5.8, 2.9 Hz, 1H), 3.94 (dt, J = 8.8, 5.8 Hz, 1H), 3.60 (dd, J = 5.8, 2.9 Hz, 2H). ^{13}C NMR (75 MHz, DMSO-d_6): δ/ppm = 164.7, 155, 138.0, 136.7, 129.2, 128.3, 127.9, 127.8, 123.6, 116.2, 65.7, 56.2, 46.4. $[\alpha]_D^{20}$ = –5 (c 1.00, DMSO). FT-IR: ν/cm^{-1} = 1754, 1512, 1501, 1380, 1341, 1254, 1023, 1004, 813, 756. MS (ESI) m/z calculated for $[C_{17}H_{17}N_2O_3]^+$ ($[M+H]^+$): 297.1, found 297.4.

9, (*S*)-3-Amino-1-phenylazetidin-2-one hydrochloride



9 (0.5 g, 1.7 mmol, 1 eq) was dissolved in THF (5 mL). Subsequently, Pd/C (50 mg, 10%) was added, and the reaction mixture was stirred for two hours under a H_2 atmosphere (3 bar). Afterwards, the solution was filtered through Celite® 545 and acidified with hydrochloric acid (1 M, 7 mL). The desired product was obtained after removing the solvent under reduced pressure as a white solid (0.3 g, 1.5 mmol, 88%). ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 9.14 (s, 1H), 7.44 – 7.36 (m, 4H), 7.16 (ddt, J = 6.3, 5.2, 2.9 Hz, 1H), 4.67 (dd, J = 5.6, 2.5 Hz, 1H), 4.0 (dd, J = 6.7, 5.6 Hz, 1H), 3.74 (dd, J = 6.7, 2.6 Hz, 1H). ^{13}C NMR (75 MHz, DMSO-d_6): δ/ppm = 160.3, 137.4, 129.4, 124.5, 116.5, 53.5, 44.0. $[\alpha]_D^{20}$ = +17 (c 1.00, MeOH). FT-IR: ν/cm^{-1} = 1749, 1532, 1509, 1368, 1339, 1244, 1025, 1000, 813, 756. MS (ESI) m/z calculated for $[C_9H_{11}N_2O]^+$ ($[M+H]^+$): 163.1, found 163.0.

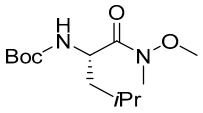
25, Benzyl ((*S*)-1-(((*S*)-1-cyano-3-phenylpropyl)amino)-1-oxo-3-phenylpropan-2-yl)carbamate



23 (0.15 g, 0.32 mmol, 1 eq) was dissolved in DMF (3 mL) and the solution was cooled to 0 °C. EDC · HCl (0.07 g, 0.39 mmol, 1.2 eq) and HOBt · H_2O (0.05 g, 0.39 mmol, 1.2 eq) were added and the mixture was stirred for 30 min before NH_4OH solution (25% in water, 3 mL) was added. The mixture was allowed to warm up to room temperature and was stirred overnight. The solvent was removed under reduced pressure and EtOAc (5 mL) was added. The organic phase was washed with saturated $NaHCO_3$ solution (5 mL) and hydrochloric acid (1 M, 5 mL), dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The crude product was used without further purification in the next step and solved in dry DMF (3 mL) under argon atmosphere. The solution cooled to 0 °C, pyridine (0.08 mL, 0.98 mmol, 3 eq) and TFA anhydride (0.14 mL, 0.98 mmol, 3 eq) were added and the reaction mixture stirred for 60 min. Subsequently, the reaction was quenched with saturated NH_4Cl solution (5 mL) and evaporated under reduced pressure. The crude product was purified by preparative HPLC to yield the desired product as a colorless solid (83.7 mg, 0.19 mmol, 60%). ^{1}H NMR (600 MHz, $CDCl_3$): δ/ppm = 7.35 – 7.03 (m, 15H), 6.35 (s, 1H), 5.32 – 5.22 (m, 1H), 5.10 – 5.01 (m, 2H), 4.77 – 4.66 (m, 1H), 4.41 – 4.29 (m, 1H), 3.10 – 2.98 (m, 2H), 2.66 (dt, J = 29.3, 7.1 Hz, 2H), 1.98 (h, J = 7.1, 6.6 Hz, 2H) ppm. ^{13}C NMR (600 MHz, $CDCl_3$): δ/ppm = 170.6, 156.2, 139.1, 135.8, 129.4, 129.1, 128.9, 128.7, 128.5, 128.2, 127.5, 126.8, 117.8, 67.5, 56.2, 40.1, 38.3, 34.3, 31.4. FT-IR: ν/cm^{-1} = 3294, 3024, 1667, 1529, 1445, 1254, 1035, 894, 748, 689. MS (ESI) *m/z* calculated. for $[C_{27}H_{28}N_3O_3]^+$ ($[M+H]^+$): 442.2, found: 442.0. Purity: 97%.

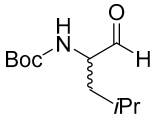
Synthesis of proteasome targeting compounds

51, *tert*-Butyl (*S*)-(1-(methoxy(methyl)amino)-4-methyl-1-oxopentan-2-yl)carbamate



To a 0 °C cold solution of Boc-Leu-OH **50** (2.04 g, 8.81 mmol, 1 eq) in DCM (80 mL) were added HOBt · H_2O (1.35 g, 8.81 mmol, 1 eq) and 2,4,6-collidine (2.34 mL, 17.62 mmol, 2 eq). After stirring at 0 °C for 30 min, TBTU (2.83 g, 8.81 mmol, 1 eq) was added. The solution was stirred for further 30 min at 0 °C and *N*,*O*-dimethylhydroxylamine · HCl (0.89 g, 8.81 mmol, 1 eq) was added. After stirring at room temperature overnight, the solvent was removed under reduced pressure and the residue was taken up in EtOAc (100 mL). The mixture was washed with saturated $NaHCO_3$ solution (3x 80 mL) and hydrochloric acid (1 M, 3x 80 mL) and filtered over a small silica column. The filtrate was concentrated under reduced pressure to yield the desired product as a colorless oil (2.24 g, 8.18 mmol, 93%). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 5.05 (d, J = 9.5 Hz, 1H), 4.70 (s, 1H), 3.77 (s, 3H), 3.18 (s, 3H), 1.78–1.62 (m, 1H), 1.47–1.38 (m, 11H), 0.99–0.88 (m, 6H). ^{13}C NMR (75.5 MHz, $CDCl_3$): δ/ppm = 174.0, 155.8, 79.6, 61.7, 49.1, 42.2, 32.3, 28.5, 24.8, 23.5, 21.7. $[\alpha]_D^{22}$ = –9 (*c* 1.00, $CHCl_3$). FT-IR: ν/cm^{-1} = 3324, 2958, 2936, 2870, 1709, 1658, 1501, 1389, 1366, 1250, 1165, 1045, 1016, 989, 876. MS (ESI) *m*/*z* calculated for $[C_{13}H_{26}N_2O_4+Na]^+$ ($[M+Na]^+$): 297.2, found: 297.1.

52, *tert*-Butyl (4-methyl-1-oxopentan-2-yl)carbamate



To a solution of **51** (1.51 g, 5.50 mmol, 1 eq) in dry THF (20 mL) was added $LiAlH_4$ (271 mg, 7.15 mmol, 1.3 eq) portion wise at 0 °C under argon atmosphere. After stirring at 0 °C for 30 min, diethyl ether (50 mL) and $KHSO_4$ solution (0.33 M, 80 mL,) were added. The suspension was filtered, and the filtrate was extracted with diethyl ether (2x 50 mL). The combined organic extracts were washed with hydrochloric acid (1 M, 2x 40 mL) and saturated $NaHCO_3$ solution (2x 40 mL) and dried over anhydrous Na_2SO_4. After removing the solvent under reduced pressure, the desired product was obtained as a colorless oil (345 mg, 1.60 mmol, 29%). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 9.57 (s, 1H), 4.95 (d, J = 7.3 Hz, 1H), 4.23 (s, 1H), 1.81–1.70 (m, 1H), 1.67–1.58 (m, 1H), 1.44 (s, 9H), 1.39–1.32 (m, 1H), 0.98–0.94 (m, 6H). ^{13}C NMR (75.5 MHz, $CDCl_3$): δ/ppm = 200.5, 155.8, 80.2, 58.5, 38.3, 28.4, 24.8, 23.2. $[\alpha]_D^{22}$ = –9 (*c* 1.00, $CHCl_3$). FT-IR: ν/cm^{-1} = 3352, 2958, 2932, 2871, 1689, 1507, 1455, 1391, 1366, 1249, 1164, 1045, 1010, 873, 779.

Synthesis of uPA targeting compounds

General procedure of Fmoc-Solid phase peptide synthesis (SPPS)

A resin Loading

Solid phase peptide synthesis was prepared in a fritted 12 mL polypropylene syringe. 1 g of 2-Chlorotritylchloride resin (2-CTC resin, 1.2 mmol loading capacity) was prewetted in 8 mL DCM for 15 min and drained. The first amino acid (3.6 mmol, 3 eq relative to resin loading capacity) was added in 1.8 M NMM/DCM (6 mL) and the mixture was mixed on a rocker for 12 h. After draining the solution, the resin was washed with DMF (3x 6 mL, 1 min each) and DCM (3 x 6 mL, 1 min each).

B resin capping

The remaining free 2-CTC resin linkers were capped with methanol in a solution of DCM/MeOH/DIPEA (9:2:1). The resin was mixed on a rocker for one hour, drained and washed with DCM (3x 6 mL).

C Fmoc deprotection

The resin was treated with a 20% piperidine in DMF solution (2x 6 mL, 10 min each) and subsequently washed with DMF (3x 6 mL).

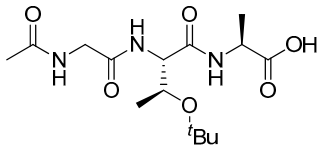
D peptide coupling with HATU

A coupling cocktail was prepared consisting of the specific Fmoc-protected amino acid (3.6 mmol, 3 eq) with HATU (3.6 mmol, 3 eq) and DIEA (10.8 mmol, 3 eq) in DMF (4.8 mL). The solution was added to the resin and mixed on a rocker for 3 h with exception of the amino acid Fmoc-Arg(Pbf)-OH which coupled overnight. After the reaction, the resin was drained and washed with DMF (3x 6 mL, 1 min each) and DCM (3x 6 mL, 1 min each).

E resin cleavage under preservation of protecting groups

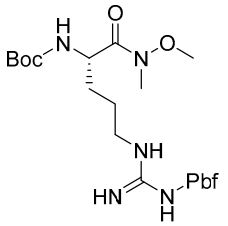
After the last coupling, the resin washed with DMF (3x 6 mL, 1 min each) and DCM (3x 6 mL, 1 min each). Then the cleavage cocktail containing AcOH/Trifluoroisopropanol/DCM (1:1:4, 6 mL) was added and mixed on a rocker for 1.5 h. The resin was drained and washed with DCM (5 mL). The combined organic solvents were concentrated under reduced pressure to yield the crude linear peptide. The purification of the crude product was done by reversed phase flash chromatography (H_2O/MeCN, gradient).

99, *N*-(Acetylglycyl)-*O*-(tert-butyl)-L-allothreonyl-L-alanine



The title compound was prepared according to general procedures A, B, C, D, and E on 1.2 mmol scale. The product was purified by reversed phase flash chromatography (H_2O/MeCN, gradient) and obtained as a colorless solid after lyophilization (414 mg, 0.6 mmol, 50%) ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 8.19 (t, J = 5.8 Hz, 1H), 7.95 (d, J = 7.0 Hz, 1H), 7.53 (d, J = 8.5 Hz, 1H), 4.26 – 4.16 (m, 2H), 3.92 (qd, J = 6.2, 3.9 Hz, 2H), 3.74 (dd, J = 7.2, 5.8 Hz, 2H), 1.86 (s, 3H), 1.28 (d, J = 7.2 Hz, 3H), 1.12 (s, 9H), 1.01 (d, J = 6.2 Hz, 3H). ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 173.8, 169.9, 169.3, 168.9, 73.8, 67.2, 57.3, 47.7, 42.3, 28.1, 22.4, 19.3, 17.5. mp: 87 – 95 °C. $[\alpha]_D^{20}$ = +12 (c 1.00, $CHCl_3$). FT-IR: ν/cm^{-1} = 668, 701, 1081, 1121, 1159, 1192, 1214, 1371, 1518, 1636. MS (ESI) *m/z* calculated for $[C_{15}H_{28}N_3O_6]^+$ ($[M+H]^+$): 346.2, found 368.1. Purity: 97.8%.

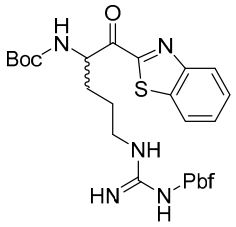
101, *tert*-Butyl (*S*)-(1-(methoxy(methyl)amino)-1-oxo-5-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)pentan-2-yl)carbamate



(Adapted from St-Georges *et al.*)[27]

Boc-Arg(Pbf)-OH **100** (6.00 g, 11.39 mmol, 1 eq) was dissolved in DCM (40 mL) under argon atmosphere. At 0 °C DIPEA (15.9 mL, 31.12 mmol, 8 eq) and TBTU (4.38 g, 13.67 mmol, 1.2 eq) were added. The reaction mixture stirred for 30 min at 0 °C, *N*,*O*-dimethyl hydroxylamine · HCl was added and stirred for 12 h. The solution was diluted with DCM (20 mL) and washed with saturated $NaHCO_3$ solution (3x 20 mL) and saturated NaCl solution (3x 20 mL). The organic solution dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The crude product was purified by column chromatography (EA) to yield the desired product as a colorless solid (6.16 g, 10.82 mmol, 95%). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 6.38 (s, 1H), 5.49 (d, J = 8.8 Hz, 1H), 4.61 (s, 1H), 3.73 (s, 3H), 3.18 (s, 3H), 2.95 (s, 3H), 2.56 (s, 3H), 2.50 (s, 3H), 2.08 (s, 3H), 1.74 – 1.52 (m, 4H), 1.45 (s, 6H), 1.41 (s, 9H) ppm, ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 156.4, 156.0, 138.8, 132.7, 124.8, 117.7, 86.6, 80.3, 61.7, 43.3, 41.0, 30.9, 28.7, 28.4, 24.9, 19.4, 18.0, 12.5. mp: 76 – 84 °C. $[\alpha]_D^{20}$ = −7 (c 1.00, $CHCl_3$). FT-IR: ν/cm^{-1} = 668, 756, 1106, 1165, 1456, 1556, 1652, 2362, 2974, 2981. MS (ESI) *m/z* calculated for $[C_{26}H_{44}N_5O_7S]^+$ ($[M+H]^+$): 570.3, found 570.3.

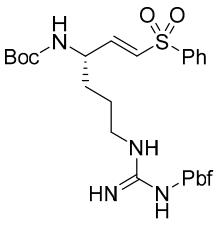
102, *tert*-Butyl (1-(benzo[*d*]thiazol-2-yl)-1-oxo-5-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)pentan-2-yl)carbamate



(Adapted from Costanzo *et al.*)[28]

Benzothiazole (3.47 g, 25.69 mmol, 10.2 eq) was dissolved in dry THF (70 mL) under argon atmosphere and cooled to −78 °C. *n*-BuLi (2.5 in hexanes, 2,97 mL, 25.19 mmol, 10 eq) was added dropwise and the reaction stirred for 1.5 h at the same temperature. Afterwards, **101** (1.43 g, 2.51 mmol, 1 eq) dissolved in dry THF (15 mL) was added slowly to the reaction mixture and stirred for two hours at −78 °C. The reaction mixture was quenched with saturated NH_4Cl solution (20 mL). The organic phase was separated and washed three times each with saturated $NaHCO_3$ solution (25 mL) and saturated NaCl solution (25 mL). The organic phase was dried over anhydrous $NaSO_4$ and concentrated under reduced pressure. The crude product was purified by column chromatography (CH/EA, 1:1 to 0:1) to yield the product as a colorless solid (1.11 g, 1.90 mmol, 76%). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 8.23 – 8.15 (m, 1H), 8.00 – 7.91 (m, 1H), 7.60 – 7.49 (m, 2H), 6.31 (s, 1H), 6.21 (s, 2H), 5.65 (d, J = 8.6 Hz, 1H), 5.61 – 5.49 (m, 1H), 3.57 – 3.18 (m, 2H), 2.92 (s, 3H), 2.54 (s, 3H), 2.49 (s, 3H), 2.06 (s, 3H), 1.72 (s, 4H), 1.45 (s, 6H), 1.41 (s, 9H). ppm, ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 158.8, 156.2, 153.5, 138.5, 137.3, 133.0, 132.5, 128.3, 127.4, 126.0, 124.7, 122.5, 117.5, 86.4, 80.6, 43.3, 40.8, 28.7, 28.4, 25.5, 19.3, 18.0, 12.5. mp: 105 – 115 °C. $[\alpha]_D^{20}$ = +2 (c 1.00, $CHCl_3$). FT-IR: ν/cm^{-1} = 668, 756, 1102, 1164, 1216, 1369, 1483, 1552, 1622, 1694. MS (ESI) *m/z* calculated for $[C_{31}H_{42}N_5O_6S_2]^+$ ($[M+H]^+$): 644.3, found 644.3. Purity: 98%.

104, *tert*-Butyl (*S*,*E*)-(6-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)-1-(phenylsulfonyl)hex-1-en-3-yl)carbamate

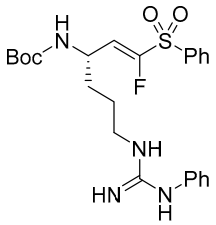


(Adapted from Andreasen *et al.*)[29]

101 (1 g, 1.76 mmol, 1 eq) was dissolved in dry THF (20 mL) under argon atmosphere and cooled to 0 °C. $LiAlH_4$ (1 M in THF, 2.2 mL, 1.25 eq) was added dropwise to the solution at 0 °C and stirred one hour at the same temperature. The reaction mixture was quenched with $KHSO_4$ solution (5%, 10 mL) and diluted with EtOAc (20 mL). The organic phase was washed with saturated $NaHCO_3$ solution (25 mL) and saturated NaCl solution (25 mL), dried over

anhydrous Na_2SO_4 and evaporated under reduced pressure, to yield the crude aldehyde as a colorless oil (0.67 g). The crude aldehyde (0.39 g, 0.76 mmol, 1 eq) and LiCl (0.04 g, 0.91 mmol, 1.2 eq) were dissolved in dry MeCN (15 mL) under argon atmosphere and cooled to 0 °C. Subsequently DBU (0.12 mL, 0.76 mmol, 1 eq) and 10 min later **3** (0.22 g, 0.76 mmol, 1 eq) in MeCN (2 mL) were added dropwise to the solution and stirred one hour at 0 °C and one hour at room temperature. The reaction mixture was quenched with citric acid solution (10%, 8 mL). The organic phase was separated and the aqueous phase extracted with EtOAc (3x 20 mL). The combined organic extracts were dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The crude product was purified by column chromatography (CH/EA 1:2) to yield the desired product colorless solid (0.36 g, 0.55 mmol, 72%). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 7.89 – 7.80 (m, 2H), 7.63 – 7.47 (m, 3H), 6.90 (dd, J = 15.0, 4.6 Hz, 2H), 6.51 – 6.33 (m, 3H), 5.26 (d, J = 8.7 Hz, 1H), 4.31 (s, 1H), 3.30 – 3.16 (m, 2H), 2.55 (s, 3H), 2.08 (s, 3H), 1.68 – 1.54 (m, 4H), 1.45 (s, 6H), 1.34 (s, 9H) ppm, ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 156.1, 155.6, 146.9, 140.1, 138.6, 133.6, 132.6, 130.3, 129.5, 127.7, 124.9, 117.7, 86.6, 43.3, 28.7, 28.3, 19.4, 18.1, 12.6. mp: 73 – 77 °C, $[\alpha]_D^{20}$ = –9 (c 1.00, $CHCl_3$), FT-IR: ν/cm^{-1} = 2957, 1706, 1512, 1358, 1302, 1141, 1091, 908, 748, 731. MS (ESI) *m/z* calculated for $[C_{31}H_{45}N_4O_7S_2]^+$ ($[M+H]^+$): 649.3, found 649.3.

105, *tert*-Butyl (*S*,*E*)-(1-fluoro-6-(3-phenylguanidino)-1-(phenylsulfonyl)hex-1-en-3-yl)carbamate



101 (1 g, 1.76 mmol, 1 eq) was dissolved in dry THF (20 mL) under argon atmosphere and cooled to 0 °C. $LiAlH_4$ (1 M in THF, 2.2 mL, 1.25 eq) was added dropwise to the solution at 0 °C and stirred one hour at the same temperature. The reaction mixture was quenched with $KHSO_4$ solution (5%, 10 mL) and diluted with EtOAc (20 mL). The organic phase was washed with saturated $NaHCO_3$ solution (25 mL) and saturated NaCl solution (25 mL), dried over anhydrous Na_2SO_4 and evaporated under reduced pressure, to yield the crude aldehyde as a colorless oil (0.67 g). The crude aldehyde (0.25 g, 0.49 mmol, 1 eq) and LiCl (0.03 g, 0.59 mmol, 1.2 eq) were dissolved in dry MeCN (8 mL) under argon atmosphere and cooled to 0 °C. Subsequently DBU (0.08 mL, 0.49 mmol, 1 eq) and 10 min later **4** (0.15 g, 0.49 mmol, 1 eq) in MeCN (1 mL) were added dropwise to the solution and stirred one hour at 0 °C and one hour at room temperature. The reaction mixture was quenched with citric acid solution (10%, 6 mL). The organic phase was separated, and the aqueous phase extracted with EA (3x 20 mL). The combined organic extracts were dried over Na_2SO_4 and evaporated under reduced pressure. The crude product was purified with column chromatography (CH/EA, 1:1) to yield the desired product as a colorless oil (0.1 g, 0.15 mmol, 31%). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 7.92 (dd, J = 7.2, 1.8 Hz, 2H), 7.72 – 7.64 (m, 1H), 7.57 (dd, J = 8.3, 6.9 Hz, 2H), 6.37 (s, 2H), 6.19 (d, J = 35.7 Hz, 2H), 5.21 (s, 1H), 4.41 (s, 1H), 3.20 (s, 2H), 2.96 (s, 2H), 2.56 (s, 3H), 2.50 (s, 3H), 2.09 (s, 3H), 1.67 – 1.50 (m, 4H), 1.46 (s, 6H), 1.30 (s, 9H) ppm, ^{13}C NMR

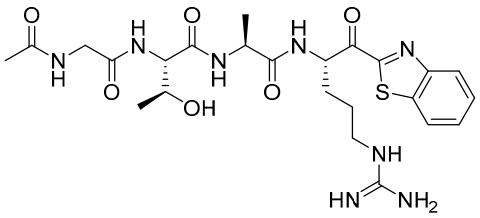
(75 MHz, CDCl3): δ/ppm = 156.1, 155.5, 152.4, 134.7, 129.6, 128.8, 124.9, 117.8, 86.7, 43.3, 40.9, 28.7, 28.3, 25.4, 19.4, 18.0, 12.6. $[\alpha]_D^{20}$ = +9 (*c* 1.00, $CHCl_3$). FT-IR: ν/cm^{-1} = 2974, 1692, 1628, 1563, 1369, 1245, 1161, 1094, 737, 658. MS (ESI) *m/z* calculated for $[C_{31}H_{44}FN_5O_7S_2]^+$ ($[M+H]^+$): 666.3, found 666.8.

General procedure for the amide bond formation between the H_2N-Arg-warhead moiety and 99

The respective H_2N-Arg-warhead moieties **102**, **104**, **105** were deprotected by using a deprotection mix of 16% TFA in DCM (4 ml) and stirred for 30 min at rt. The solvent was removed under reduced pressure to yield the deprotected amines as trifluoroacetate salts.

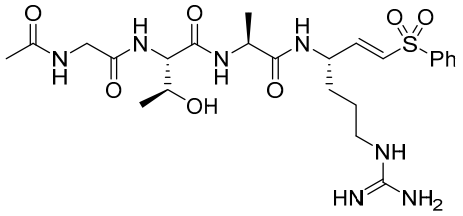
99 (1.0 eq) was dissolved in DCM and cooled to 0 °C with an ice-water bath. DIPEA (3 eq) and HATU (1.2 eq) were added and stirring was continued for 30 min at 0 °C. Then, the respective deprotected amine (1 eq) as a solution in DMF were added, stirred for 30 min at 0 °C and 16 h at room temperature. DCM and water were added, and the aqueous phase was extracted three times with DCM. The combined organic phases were dried over Na_2SO_4 and the solvent was removed under reduced pressure.

103, (2*S*,3*S*)-2-(2-Acetamidoacetamido)-*N*-((*S*)-1-(((*S*)-1-(benzo[*d*]thiazol-2-yl)-5-guanidino-1-oxopentan-2-yl)amino)-1-oxopropan-2-yl)-3-hydroxybutanamide



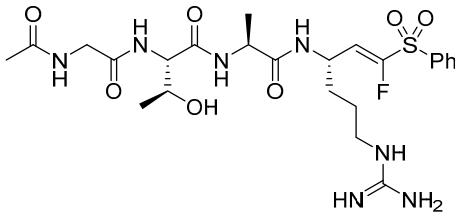
The crude product was used without further purification for the following deprotection of the Pbf- and O^{t}Bu-protecting group. It was dissolved in TFA/DCM (50%, 2 mL) and stirred for 2 h at room temperature. The solution was evaporated under reduced pressure and purified by preparative HPLC (12.8 mg, 0.017 mmol, 10%). ^{1}H NMR (400 MHz, DMSO-d_6): δ/ppm = 8.57 – 8.39 (m, 1H), 8.31 – 8.24 (m, 2H), 8.23 – 8.16 (m, 1H), 8.08 – 7.96 (m, 1H), 7.83 – 7.72 (m, 1H), 7.71 – 7.63 (m, 2H), 7.60 – 7.54 (m, 1H), 5.54 – 5.41 (m, 1H), 5.00 (d, *J* = 35.3 Hz, 1H), 4.36 (ddd, *J* = 10.0, 7.2, 2.7 Hz, 1H), 4.17 (tdd, *J* = 12.1, 8.0, 4.2 Hz, 1H), 3.97 (ddt, *J* = 16.9, 11.2, 5.2 Hz, 1H), 3.83 – 3.69 (m, 2H), 3.14 (p, *J* = 6.6 Hz, 2H), 1.97 (td, *J* = 13.7, 11.4, 6.7 Hz, 2H), 1.90 – 1.82 (m, 2H), 1.79 – 1.69 (m, 0H), 1.65 – 1.54 (m, 2H), 1.27 – 1.17 (m, 3H), 1.06 – 0.98 (m, 3H). ^{13}C NMR (400 MHz, DMSO-d_6): δ/ppm = 193.1, 172.6, 169.9, 169.7, 169.5, 164.5, 156.8, 153.0, 136.5, 128.3, 127.6, 125.3, 123.3, 66.7, 58.2, 54.3, 47.9, 42.2, 42.2, 27.8, 25.2, 22.4, 19.7, 18. $[\alpha]_D^{20}$ = +34 (*c* 1.00, DMSO). FT-IR: ν/cm^{-1} = 660, 668, 756, 1136, 1215, 1673, 1684, 2922, 2964, 3327, MS (ESI) *m/z* calculated for $[C_{24}H_{34}N_8O_6S+H]^+$ ($[M+H]^+$): 563.2, found 563.2. Purity: 98%.

106, (2*S*,3*S*)-2-(2-acetamidoacetamido)-N-((*S*)-1-(((*S*,*E*)-6-guanidino-1-(phenylsulfonyl)hex-1-en-3-yl)amino)-1-oxopropan-2-yl)-3-hydroxybutanamide



The crude product was used without further purification for the following deprotection of the Pbf- and O^{t}Bu-protecting group. It was solved in TFA/DCM (50%, 6 mL) and stirred for 2 h at room temperature. The solution was evaporated under reduced pressure and purified preparative HPLC (18.0 mg, 0.031 mmol, 16%). ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 8.27 – 8.15 (m, 1H), 8.00 (dd, J = 7.8, 3.3 Hz, 2H), 7.86 – 7.60 (m, 7H), 6.84 (d, J = 3.3 Hz, 1H), 6.81 – 6.65 (m, 1H), 4.57 – 4.45 (m, 1H), 4.23 – 3.66 (m, 5H), 3.13 – 2.99 (m, 2H), 1.64 – 1.37 (m, 4H), 1.22 (dd, J = 7.2, 4.9 Hz, 2H), 1.01 (dd, J = 6.3, 1.3 Hz, 3H). ^{13}C NMR (75 MHz, DMSO-d_6): δ/ppm = 172.1, 170.2, 170.1, 169.9, 169.6, 156.8, 146.5, 140.2, 133.8, 130.1, 129.7, 127.1, 66.6, 59.0, 58.3, 48.6, 42.4, 42.2, 30.1, 25.0, 22.4, 19.7, 17.8, 17.6. $[\alpha]_D^{20}$ = +13 (c 1.00, DMSO). FT-IR: ν/cm^{-1} = 675, 709, 734, 796, 1130, 1341, 1445, 1540, 1636, 3283, MS (ESI) *m/z* calculated for $[C_{24}H_{38}N_7O_7S]^+$ ([M+H]$^+$): 568.7, found 568.2. Purity: 100%.

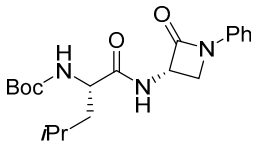
107, (2*S*,3*S*)-2-(2-acetamidoacetamido)-*N*-((*S*)-1-(((*S*,*E*)-1-fluoro-6-guanidino-1-(phenylsulfonyl)hex-1-en-3-yl)amino)-1-oxopropan-2-yl)-3-hydroxybutanamide



The crude product was used without further purification for the following deprotection of the Pbf- and O^{t}Bu-protecting group. It was solved in TFA/DCM (50%, 2 mL) and stirred for 2 h at rt. The solution was evaporated under reduced pressure and purified by preparative HPLC (5.6 mg, 0.021 mmol, 6%). ^{1}H NMR (600 MHz, DMSO-d_6): δ/ppm = 8.25 – 8.19 (m, 1H), 8.09 – 8.01 (m, 1H), 7.94 – 7.90 (m, 2H), 7.86 (td, J = 7.6, 3.8 Hz, 1H), 7.76 – 7.71 (m, 2H), 7.54 (t, J = 5.8 Hz, 1H), 7.34 (s, 1H), 6.83 (s, 2H), 6.32 (ddd, J = 36.4, 33.6, 9.1 Hz, 1H), 5.11 (s, 1H), 4.63 – 4.51 (m, 1H), 4.21 – 4.05 (m, 2H), 4.02 – 3.88 (m, 1H), 3.86 – 3.68 (m, 2H), 3.07 – 3.00 (m, 2H), 1.86 (d, J = 11.5 Hz, 2H), 1.58 – 1.29 (m, 4H), 1.23 – 1.17 (m, 3H), 1.05 (t, J = 6.3 Hz, 3H). ^{13}C NMR (151 MHz, DMSO-d_6): δ/ppm = 171.9, 171.7, 170.12, 170.0, 169.9, 169.8, 169.7, 169.6, 158.3, 158.1, 156.6, 136.4, 136.3, 135.4, 135.3, 130.3, 130.2, 128.2, 128.2, 118.7, 66.5, 66.4, 59.2, 58.1, 48.5, 43.8, 43.9, 43.8, 42.3, 42.0, 30.6, 30.5, 24.8, 24.8, 22.4, 22.4, 19.7, 19.6, 17.7, 17.6. $[\alpha]_D^{20}$ = +27 (c 1.00, DMSO). FT-IR: ν/cm^{-1} = 680, 713, 740, 781, 1145, 1335, 1444, 1561, 1640, 3260. MS (ESI) *m/z* calculated for $[C_{24}H_{36}FN_7O_7SNa]^+$ ([M+Na]$^+$): 586.7, found 586.2. Purity: 98%.

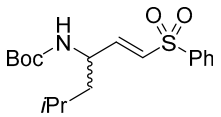
Synthesis of reactivity test compounds

108, *tert*-Butyl ((2S)-4-methyl-1-oxo-1-((2-oxo-1-phenylazetidin-3-yl) amino) pentan-2-yl) carbamate



Boc-Leu-OH **50** (0.13 g, 0.67 mmol, 1.0 eq) was dissolved in a mixture of DCM/DMF (9:1) at 0 °C. Under stirring HATU (0.31 g, 0.80 mmol, 1.2 eq) was added in portions. Afterwards 2,4,6- collidine (0.24 g, 2.0 mmol, 3.0 eq) was added and stirred for an additional 10 min at 0 °C. A solution of **9** (0.14 g, 0.67 mmol, 1 eq) in DCM (5 ml) was added dropwise and the reaction mixture was allowed to reach room temperature and was stirred overnight. H_2O (30 mL) was added, and the organic phase separated. The aqueous phase was extracted with EtOAc (3 × 20 mL). The combined organic extracts were washed with aq. sat. NaCl-solution (40 mL) and dried over Na_2SO_4. The solvent was evaporated under reduced pressure to give a crude product that was purified with column chromatography (EA:Cy = 7:4) yielding the title compound as an off-white solid (0.2 g, 0.54 mmol, 81%). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 7.65 – 7.45 (m, 1H), 7.38 – 7.20 (m, 4H), 7.06 (tdd, J = 5.9, 4.6, 3.0 Hz, 1H), 5.18 – 4.99 (m, 2H), 3.88 (t, J = 5.7 Hz, 1H), 3.51 (dd, J = 5.9, 2.7 Hz, 1H), 1.75 – 1.44 (m, 3H), 1.39 (s, 10H), 0.90 (dd, J = 6.3, 3.8 Hz, 6H). ^{13}C NMR (75 MHz, $CDCl_3$): δ/ppm = 173.6, 163.8, 156.0, 137.8, 129.3, 124.5, 116.8, 80.3, 65.4, 55.4, 53.1, 47.4, 41.6, 28.4, 24.8, 23.1, 21.9, 14.2. mp: 84 – 85 °C. $[\alpha]_D^{20} = -16$ (c 1.00, $CHCl_3$). MS (ESI) *m/z* calculated for $[C_{20}H_{28}N_3O_4Na]^+$ ($[M+Na]^+$): 389.2, found: 398.2. Purity: 99%.

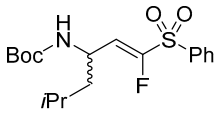
109, *tert*-Butyl (*S,E*)-(5-methyl-1-(phenylsulfonyl)hex-1-en-3-yl)carbamate



3 (0.52 g, 1.77 mmol, 1 eq) and LiCl (0.09 g, 2.21 mol, 1.25 eq) were dissolved in dry MeCN (18 mL) under argon atmosphere and cooled to 0 °C. Subsequently DBU (0.27 mL, 1.77 mmol, 1 eq) and 10 min later **52** (0.38 g, 1.77 mmol, 1 eq) in MeCN (5 mL) were added dropwise to the solution and stirred one hour at 0 °C and one hour at room temperature. The reaction mixture was quenched with citric acid solution (10%, 10 mL). The organic phase was separated, and the aqueous phase extracted with EtOAc (3x 20 mL). The combined organic extracts were dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The crude product was purified with column chromatography (CH/EA 1:6) to yield the product as colorless oil (0.18 g, 0.52 mmol, 30%). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 7.89 – 7.83 (m, 2H), 7.64 – 7.48 (m, 3H), 6.87 (dd, J = 15.0, 4.9 Hz, 1H), 6.42 (dd, J = 15.0, 1.3 Hz, 1H), 4.48 – 4.34 (m, 1H), 1.74 – 1.59 (m, 2H), 1.40 – 1.32 (m, 9H), 0.92 (d, J = 6.6 Hz, 6H). ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 155.0, 147.2, 140.4, 133.5, 130.1, 129.4, 127.7, 43.4, 28.3, 24.8, 22.8, 22.1. mp: 72 –

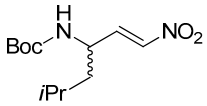
76 °C. $[\alpha]_D^{20}$ = −7 (*c* 1.00, $CHCl_3$). FT-IR: *ν*/cm^{-1} = 731, 748, 908, 1091, 1141, 1302, 1358, 1512, 1706, 2957. MS (ESI) *m/z* calculated for $[C_{13}H_{20}NO_2S]^+$ ($[M-Boc+H]^+$): 254.12, found 254.01. Purity: 99%.

110, *tert*-Butyl (*S,E*)-(1-fluoro-5-methyl-1-(phenylsulfonyl)hex-1-en-3-yl)carbamate



4 (0.22 g, 0.72 mmol, 1 eq) and LiCl (0.04 g, 0.90 mmol, 1.25 eq) were dissolved in dry MeCN (10 mL) under argon atmosphere and cooled to 0 °C. Subsequently DBU (0.11 mL, 0.72 mmol, 1 eq) and 10 min later **52** (0.16 g, 0.72 mmol, 1 eq) in MeCN (3 mL) were added dropwise to the solution and stirred one hour at 0 °C and one hour at room temperature. The reaction mixture was quenched with citric acid solution (10%, 10 mL). The organic phase was separated, and the aqueous phase extracted with EtOAc (3x 20 mL). The combined organic extracts were dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The crude product was purified with column chromatography (CH/EA 1:7) to desired product colorless oil (0.08 g, 0.21 mmol, 30%). ^{1}H NMR (300 MHz, $CDCl_3$): *δ*/ppm = 7.99 – 7.91 (m, 2H), 7.73 – 7.53 (m, 3H), 6.12 (d, *J* = 31.9 Hz, 1H), 4.53 (s, 1H), 1.65 – 1.45 (m, 2H), 1.35 (s, 9H), 0.90 (dd, *J* = 6.5, 2.5 Hz, 6H). ^{13}C NMR (75 MHz, $CDCl_3$): *δ*/ppm = 156.3, 154.9, 152.3, 137.3, 134.5, 129.5, 128.7, 119.1, 43.4, 28.3, 24.8, 22.5, 22.3. $[\alpha]_D^{20}$ = +26 (*c* 1.00, $CHCl_3$). FT-IR: *ν*/cm^{-1} = 683, 717, 756, 1080, 1153, 1335, 1448, 1507, 1703, 2965. MS (ESI) *m/z* calculated for $[C_{18}H_{26}FNO_4SNa]^+$ ($[M+Na]^+$): 394.1, found 394.1. Purity: 98%.

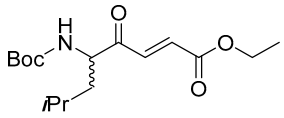
111, *tert*-Butyl ((3*S*)-2-hydroxy-5-methyl-1-nitrohexan-3-yl) carbamate



52 (0.39 g, 1.84 mmol, 1 eq) was dissolved in DCM (4 ml) and cooled to 0 °C. $MeNO_2$ (0.67 g, 11.06 mmol, 6 eq) and DBU (0.27 mL, 1.84 mmol, 1 eq) were added and the reaction mixture stirred for three hours at room temperature. The reaction got quenched with saturated NH_4Cl solution (6 mL) and extracted with DCM (15 mL). The organic solution was washed with brine (13 mL) and dried over anhydrous Na_2SO_4. The solvent was evaporated under reduced pressure and the crude product was used without further purification. The crude nitroaldol intermediate (0.44 g, 1.59 mmol, 1 eq) was dissolved in DCM (9 ml) and cooled to 0 °C. MsCl (0.36 g, 3.18 mmol, 2 eq) and DIPEA (1.11 mL, 6.36 mmol, 4 eq) were added and the reaction mixture stirred for 90 min at 0 °C. The reaction was quenched with saturated NH_4Cl solution (14 mL) and extracted with DCM (15 mL). The organic solution was washed with brine (20 mL) and dried over anhydrous Na_2SO_4. The solvent was evaporated under reduced pressure and the crude product was purified by reversed phase flash chromatography (H_2O/MeCN, gradient) to yield the desired product colorless solid (0.04 g, 0.16 mmol, 10%). ^{1}H NMR (300 MHz, $CDCl_3$): *δ*/ppm = 7.12 (dd, *J* = 13.3, 5.3 Hz, 1H), 7.04 (dd, *J* = 13.3, 0.8 Hz, 1H), 4.47 – 4.36

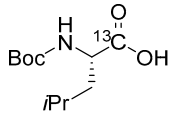
(m, 0H), 1.77 – 1.62 (m, 1H), 1.46 – 1.41 (m, 11H), 0.95 (dd, J = 6.6, 1.0 Hz, 6H). ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 155.0, 142.7, 139.7, 80.5, 43.4, 28.4, 24.8, 22.7, 22.1. $[\alpha]_D^{20}$ = +11 (c 1.00, $CHCl_3$). FT-IR: ν/cm^{-1} = 2963, 1698, 1560, 1507, 1358, 1231, 1161, 1049, 849, 756. Purity: 99%.

112, Ethyl (*S,E*)-5-((*tert*-butoxycarbonyl)amino)-7-methyl-4-oxooct-2-enoate



A solution of **65** (926 mg, 2.74 mmol, 1 eq) and LiCl (140 mg, 3.29 mmol, 1.2 eq) in dry MeCN (20 ml) was cooled to 0 °C. DBU (410 µL, 2.74 mmol, 1 eq) was added, followed by the dropwise addition of a solution of freshly distilled ethyl glyoxylate (559 µL, 5.49 mmol, 2 eq) in dry MeCN (5 ml), while stirring at 0 °C. After stirring for two hours at 0 °C, hydrochloric acid (0.1 M, 30 ml) was added and stirring was continued at room temperature for 15 min, followed by extraction with a mixture of chloroform an isopropyl alcohol (3:1, 4x 15ml). The combined organic layers were dried over anhydrous Na_2SO_4 and the solvent was removed under reduced pressure. The resulting residue was purified by reversed phase flash chromatography (H_2O/MeCN, gradient) yielding the desired product as a colorless resin (368 mg, 1.23 mmol, 45%). ^{1}H NMR (300 MHz, Chloroform-*d*): δ/ppm = 7.18 (d, J = 15.8 Hz, 1H), 6.78 (d, J = 15.8 Hz, 1H), 5.06 (d, J = 8.2 Hz, 1H), 4.65 – 4.47 (m, 1H), 4.24 (q, J = 7.1 Hz, 2H), 1.81 – 1.64 (m, 1H), 1.60 – 1.48 (m, 1H), 1.41 (s, 9H), 1.40 – 1.33 (m, 1H), 1.30 (t, J = 7.1 Hz, 3H), 0.94 (dd, J = 16.2, 6.5 Hz, 6H). ^{13}C NMR (75 MHz, $CDCl_3$): δ/ppm = 198.64, 165.20, 155.48, 136.31, 132.15, 80.00, 61.40, 57.22, 40.36, 28.23, 24.86, 23.17, 21.69, 14.06. $[\alpha]_D^{22}$ = +13 (c 1.00, $CHCl_3$). FT-IR: ν/cm^{-1} = 3367, 2960, 2872, 1700, 1507, 1470, 1391, 1367, 1253, 1165, 1023, 980, 871, 780. MS (ESI) *m/z* calculated for $[C_{16}H_{27}NO_5Na]^+$ ($[M+Na]^+$): 336.2, found: 336.3. Purity: 96%.

113, (*tert*-Butoxycarbonyl)-L-leucine-1-^{13}*C*



(Adapted from F. Dutton *et al*.**)**[30]

To a 0 °C cold solution of H-Leu-1-^{13}C-OH (0.98 g, 7.42 mmol, 1 eq) in 1,4-dioxane (5 mL) and water (3 mL) were added NaOH solution (1 M, 15 mL) and Boc_2O (3.66 g, 16.77 mmol, 2.26 eq). The solution was stirred at room temperature overnight and water (30 mL) was added. The mixture was washed with *n*-pentane (4x 80 mL) and acidified to pH = 1 with hydrochloric acid (1 M). The resulting suspension was extracted with EtOAc (4x 50 mL) and the combined organic extract were dried over anhydrous Na_2SO_4. After removing the solvent by distillation under reduced pressure, the desired product was obtained as a colorless oil (1.72 g, 7.42 mmol, quantitative). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 4.89 (d, J = 8.5 Hz, 1H), 4.41–4.05 (m, 1H), 1.82–1.50 (m, 3H), 1.45 (s, 9H), 0.95 (d, J = 6.2 Hz, 6H). ^{13}C NMR (75.5 MHz, $CDCl_3$): δ/ppm = 178.3, 155.9, 80.4, 52.16 (d, J = 58.6 Hz), 41.6, 28.4, 24.9, 23.0, 21.9. $[\alpha]_D^{22}$ = −5 (c

“

4.2 Next Generation of Fluorometric Protease Assays: 7-Nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-Amides) as Class-Spanning Protease Substrates.

4.2.1 Context, Project Summary, and Own Contributions

Since the first approval of a protease inhibitor in 1964, more than 64 protease inhibitors were approved by the FDA for treatment against human-related diseases (1.3.2).[110] One of the most frequently used methods in medicinal chemistry for characterization of the inhibitory activity of drug candidates against proteases represents the fluorometric inhibition assay (1.8).[308] Decisive properties of this method represent a simple and time efficient experimental setup, accompanying with a manageable cost and a steadily growing supply of fluorogenic substrates. The principle underlies the cleavage of an amide bond of substrates, resulting in a change of fluorescence emission, which is a scale for the enzymatic activity. Therefore, the impaired reaction rate of an enzyme in presence of an inhibitor correlates with the change of the fluorescence intensity over time and can be used for evaluation of the inhibitory activity with different methods. However, using the common AMC- and FRET-based fluorogenic substrates, associated assay interferences can appear, leading to false positive or false negative results (1.8.4).[309] Due to the properties of most aromatic drug-like inhibitors, many assay interferences occur in the low wavelength area. For that reason, one promising strategy represents the shift of the fluorophore´s excitation and emission wavelength to the high wavelength range.[310]

Here, we present the 4-amino-7-nitro-benzofurazane (NBD) fluorophore as an alternative approach for fluorescence assays. With excitation and emission maxima in higher wavelength regions (λ_{ex} = 485 nm λ_{ex} = 535 nm) than AMC- (λ_{ex} = 380 nm, λ_{ex} = 460 nm) and FRET-based fluorophores (e.g., Dabcyl/EDANS, λ_{mx} = 340 nm, λ_{ex} = 490 nm), it is possible to mitigate common assay interferences (Figure 29).[290],[292],[311],[312] We designed and synthesized protease substrates with NBD as the fluorophore for overall ten proteases including serine-, cysteine and metalloproteases. Due to the observed SMILES rearrangement of selected NBD-substrates, we addressed not just proteolytic enzymes with carboxypeptidase activity (XXX|NBD) but also endopeptidases (XX|X-NBD). The enzyme-and substrate-specific kinetic parameters (K_M, k_{cat}, v_{max}) as well as inhibition constants of known inhibitors were determined, to confirm the suitability of the new protease substrates in inhibition assays. Further experiments to verify the ability of NBD-based substrates to avoid common assay interferences like the inner filter effect, autofluorescence decay and fluorophore quenching were performed with success.

In summary, this project highlights the NBD-based substrates as potent alternative fluorophore to the commonly used AMC and FRET substrates and may represent a new era of fluorometric assays.

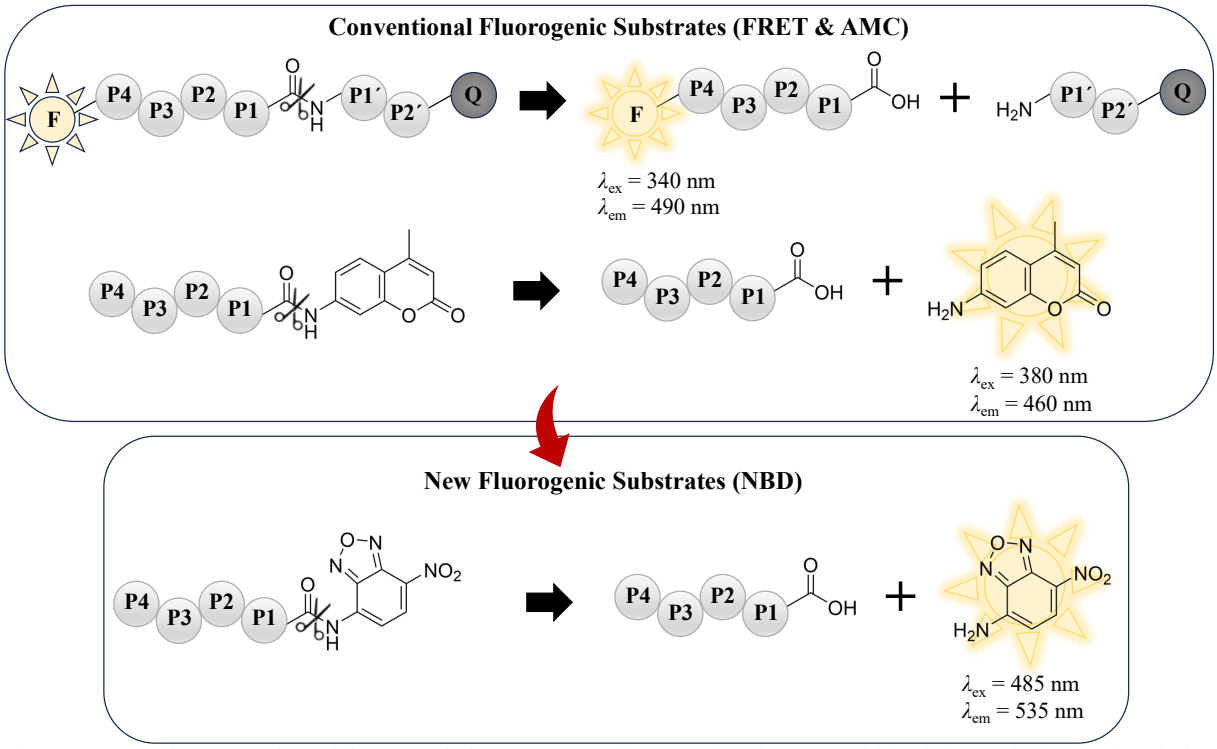


Figure 29: Development of peptidic NBD-derivates as new fluorogenic substrates. Shift of the excitation and emission wavelengths of the NBD-based substrates to longer wavelengths, results in mitigation of assay interferences. Figure created with PowerPoint.

Own contributions: Substrate synthesis (2a–j, 4, 6a–d, 7a–j, 18–20) and writing of the original draft plus editing of the manuscript.

Contribution from other authors: Inhibitor synthesis, protein expression and purification, substrate stability assays, molecular docking, kinetic characterization of substrates, investigations on assay interferences and writing of the original draft plus editing of the manuscript.

This work has been published in Chemistry – A European Journal (impact factor: 5.02).

The appended Supporting Information represents an abridged version. The full version can be accessed online at https://chemistry-europe.onlinelibrary.wiley.com/doi/full/10.1002/chem.202301855.

4.2.2 Publication

The following publication quoted within "" from page 126 to page 147 is the same as the manuscript cited on page 53.

"

Check for updates

Chemistry–A European Journal | Research Article
doi.org/10.1002/chem.202301855

Chemistry Europe
European Chemical Societies Publishing

www.chemeurj.org

Next Generation of Fluorometric Protease Assays: 7-Nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-Amides) as Class-Spanning Protease Substrates

Hannah Maus[+],[a] Patrick Müller[+],[a] Mergim Meta,[a] Sabrina N. Hoba,[a] Stefan J. Hammerschmidt,[a] Robert A. Zimmermann,[a] Collin Zimmer,[a] Natalie Fuchs,[a] Tanja Schirmeister,[a] and Fabian Barthels*[a]

Fluorometric assays are one of the most frequently used methods in medicinal chemistry. Over the last 50 years, the reporter molecules for the detection of protease activity have evolved from first-generation colorimetric *p*-nitroanilides, through FRET substrates, and 7-amino-4-methyl coumarin (AMC)-based substrates. The aim of further substrate development is to increase sensitivity and reduce vulnerability to assay interferences. Herein, we describe a new generation of substrates for protease assays based on 7-nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-amides). In this study, we synthesized and tested substrates for 10 different proteases from the serine-, cysteine-, and metalloprotease classes. Enzyme- and substrate-specific parameters as well as the inhibitory activity of literature-known inhibitors confirmed their suitability for application in fluorometric assays. Hence, we were able to present NBD-based alternatives for common protease substrates. In conclusion, these NBD substrates are not only less susceptible to common assay interference, but they are also able to replace FRET-based substrates with the requirement of a prime site amino acid residue.

Introduction

Fluorescence is one manifestation of the interaction of electromagnetic radiation with matter which is ubiquitously used for the analysis of all kinds of parameters in the life sciences.[1] It is widely used for the characterization of protein structures and protein/ligand interactions.[2] This comprises the analysis of protein structural changes, localization of proteins in cells, organisms, and the determination of binding affinities between proteins and their ligands.[3–6] Proteolytic enzymes represent a current field of research, because of their mechanistic involvement in many diseases from virus infections to cancer progression and immunological disorders.[7–9] Since 1964, more than 64 protease inhibitors were approved by the Food and Drug Administration (FDA) for application in human therapies;[10,11] exemplarily, the recently approved cysteine protease inhibitor Nirmatrelvir is being used to treat the SARS-CoV 2 infection.[12] This implies, that proteases are still important targets, and the scientific community strives for discovering new protease-targeting drugs.[13,14]

Due to their time efficiency and steadily growing supply of commercially available fluorogenic protease substrates, fluorometric assays are one of the most frequently used methods in medicinal chemistry to analyze protease-inhibiting drug candidates.[15,16] The first, in 1973 introduced fluorogenic substrates used for the measurement of protease activities were FRET-based (Förster resonance energy transfer) substrates which harbor a quencher and fluorophore molecular pair (Figure 1).[17] These replaced the previously utilized colorimetric *p*-nitroanilide (pNA) substrates, because of their higher detection sensitivity.[18] Only a few years later, in 1976, the corresponding amides of a fluorescent coumarin derivative (7-amino-4-methyl coumarin, AMC) were implemented as a new fluorogenic protease substrate which decreased the vulnerability to assay interferences compared to the FRET-based and colorimetric pNA substrates.[19]

In general, the principle of fluorometric protease assays is described as follows: The target enzyme cleaves an amide bond of the substrate, which leads to a change in fluorescence properties of the fluorogenic reporter, for example by separation of the quencher from the fluorophore (FRET substrate) or cleavage of the internally quenched fluorogenic residue from the substrate (AMC substrate).[20] The fluorophore is then excited within a cuvette fluorometer or a microplate reader at a specific excitation wavelength and the resulting fluorescence can be detected at its corresponding emission wavelength. The increasing fluorescence intensity over time describes a scale of enzymatic activity. By the addition of an inhibitor, the reaction rate of the enzyme is attenuated which gives information about the inhibitory activity by different evaluation methods. How-

[a] H. Maus,[+] P. Müller,[+] M. Meta, S. N. Hoba, S. J. Hammerschmidt, R. A. Zimmermann, C. Zimmer, N. Fuchs, Prof. Dr. T. Schirmeister, Dr. F. Barthels
Institute of Pharmaceutical and Biomedical Sciences
Johannes Gutenberg-University
Staudingerweg 5, 55128 Mainz (Germany)
E-mail: barthels@uni-mainz.de
Homepage: https://ak-barthels.pharmazie.uni-mainz.de/

[+] These authors contributed equally to this manuscript.

Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202301855

Chemistry–A European Journal | Research Article doi.org/10.1002/chem.202301855 | Chemistry Europe European Chemical Societies Publishing

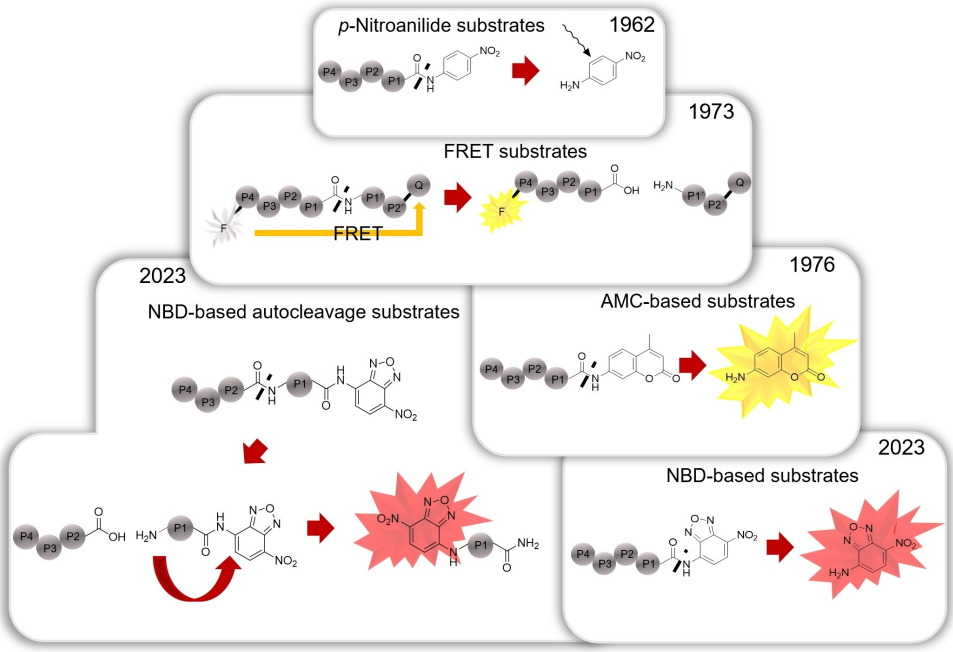


Figure 1. Timeline of protease substrates development with selected example substrates including *p*-nitroanilides (first described in 1962), FRET- (1973), and AMC-based (1976) fluorogenic substrates. In this study, we report the development of newly designed NBD-based protease substrates for biochemical and medicinal chemistry assays.

ever, not less important is the resilience of a fluorometric assay system regarding non-specific interference effects, that originate from various sources of errors like assay components, including the analyzed drug compounds themselves.[21–23] Assay interferences lead to both false positive and false negative results, which makes it necessary to validate potential hits with an additional, orthogonal methodology.[24]

A known strategy for mitigation of intrinsic assay interferences is based on the fact that many of the interfering physicochemical transitions do occur in the near UV-light spectrum, due to the molecular properties of mostly aromatic drug-like inhibitors.[25] One option to avoid inhibitor-induced interferences is to shift the assay wavelengths to the lower-energy range of light, and hence, there are existing biochemical methods using fluorescent dyes (BODIPY, Cy5, etc.) with their excitation and emission maxima in the red wavelength region (> 580 nm), leading to minimized interferences.[26,27] Over the last 30 years, FRET substrates were optimized by an overall shift to longer wavelengths like Dabcyl/EDANS resulting in optimized sensitivity, photochemical, and low-interference properties.[28] In contrast to the most frequently used AMC- ($\lambda_{ex} = 380$ nm, $\lambda_{em} = 460$ nm) and FRET-based (Dabcyl/EDANS; $\lambda_{ex} = 340$ nm, $\lambda_{em} = 490$ nm) fluorogenic substrates, the 4-amino-7-nitrobenzofurazane (NBD) fluorophore displays excitation and emission maxima in the fluorescein wavelength magnitude ($\lambda_{ex} = 485$ nm, $\lambda_{em} = 535$ nm) and might consequently avoid common interferences associated with AMC and EDANS fluorophores.[17,19,29–31]

Applications for NBD-based chemosensors have already been described in the literature, leading to several hundreds of publications and the development of numerous commercial probes which have already been summarized in reviews.[32] For the analysis of enzymatic turnover, however, the use of NBD-based derivates has been only described for the study of histone acetyltransferases, deacetylases, and esterases.[33–37] To expand the scope, herein, we report the synthesis and characterization of peptide-based NBD-amides as substrates for various protease targets. Additionally, due to the observed Smiles rearrangement mechanism of selected NBD derivates, we were able to design substrates suited as corresponding FRET substrate alternatives (Figure 1).[38,39]

To evaluate the substrate affinity and turnover rate of the NBD-based substrates, we determined K_M and k_{cat} values. Both substrate benchmarks were compared with the corresponding literature known, AMC- or FRET-derived parent substrates. To demonstrate that this new assay protocol provides comparable results to the literature-described assays, K_i values were determined for model inhibitors known from the literature. Afterward, the mitigation of common assay interferences of conventional fluorogenic substrates was investigated by comparing a NBD-based substrate with its AMC-based counterpart.

Chemistry–A European Journal | Research Article doi.org/10.1002/chem.202301855 | Chemistry Europe European Chemical Societies Publishing

Results and Discussion

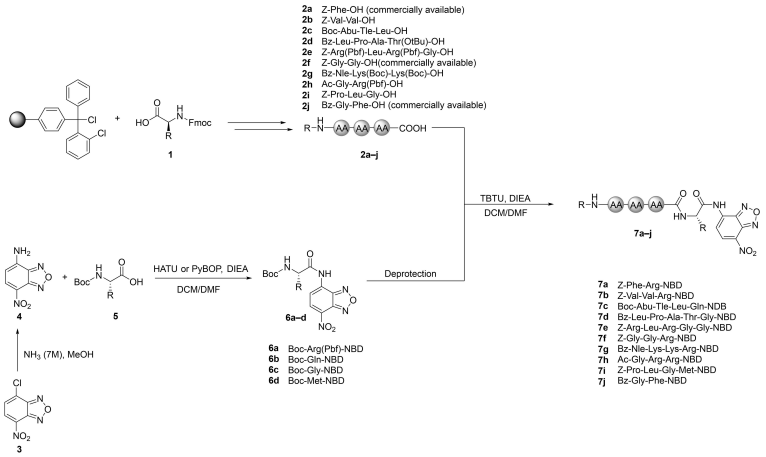
Synthesis

All NBD substrates were prepared in a two-step synthetic procedure. First, the variable peptide sequences **2a–j** spanning from the P2-site to the *N*-capped terminus were synthesized following a standard fluorenylmethoxycarbonyl (Fmoc)-solid phase peptide synthesis (SPPS) protocol using a 2-chlorotrityl resin.[40]

The second step in the synthetic sequence was the preparation of the fluorogenic reporter group through a nucleophilic aromatic substitution of precursor **3** in a methanolic ammonia solution.[41] Subsequently, the NBD amine **4** was coupled with the respective *tert*-butoxycarbonyl (Boc)-protected P1 amino acid **5**, using the *in situ* activating agents *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium-hexafluorophosphate (HATU) or benzotriazole-1-yl-oxytripyrrolidinophosphonium-hexafluorophosphate (PyBOP), and the organic base *N*,*N*-diisopropylethylamine (DIEA). The amino group of the intermediate **6a–d** was deprotected by trifluoroacetic acid, and subsequently, coupled with the remaining peptide sequence **2** to yield the desired fluorogenic substrates **7a–j** (Scheme 1, Table 1).

Kinetic Characterization of NDB-based Substrates

To verify and analyze the suitability of the proposed NBD-based substrates for fluorometric protease assays, the enzyme kinetic parameters K_m, v_{max}, k_{cat}, and the catalytic efficiency k_{cat}/K_M were determined for both the NBD substrates and their corresponding parent substrates. Proteolytic cleavage of the substrate release was detected in 96-well plate format (typically 200 µL). NBD substrates were excited at 485 nm, and the emission was detected at 535 nm.[42,43] The increase in fluorescence intensity over time (reaction rate) is a measure of the activity of the protease, and thus, the fluorescent progress traces follow the Michaelis Menten equation (Eq. (1)) which was used for enzyme kinetic analysis:



Scheme 1. SPPS/solution phase synthesis of the NBD-based substrates suited for cysteine-, serine-, and metalloproteases.

Table 1. Synthesized NBD-based substrates and their corresponding AMC-/FRET-based parent substrates.

Protease	Protease class	NBD-based substrate	Reference/parent substrate
T. brucei rhodesain	Cysteine	Z–Phe–Arg–NBD (**7a**)	Z–Phe–Arg–AMC[44]
Cathepsin S (human)	Cysteine	Z–Val–Val–Arg–NBD (**7b**)	Z–Val–Val–Arg–AMC[45]
SARS-CoV 2 M^{pro}	Cysteine	Boc–Abu–Tle–Leu–Gln–NBD (**7c**)	Boc–Abu–Tle–Leu–Gln–AMC[46]
S. aureus SrtA	Cysteine	Bz–Leu–Pro–Ala–Thr–Gly–NBD (**7d**)	Abz–Leu–Pro–Glu–Thr–Gly–Dap(Dnp)–OH[47]
SARS-CoV 2 PL^{pro}	Cysteine	Z–Arg–Leu–Arg–Gly–Gly–NBD (**7e**)	Z–Arg–Leu–Arg–Gly–Gly–AMC[48]
uPA (human)	Serine	Z–Gly–Gly–Arg–NBD (**7f**)	Z–Gly–Gly–Arg–AMC[49]
DENV NS2B/NS3 ZIKV	Serine	Bz–Nle–Lys–Lys–Arg–NBD (**7g**) Ac–Gly–Arg–Arg–NBD (**7h**)	Bz–Nle–Lys–Lys–Arg–AMC[50] Ac–Gly–Arg–Arg–AMC[51]
MMP9 (human)	Metallo	Z–Pro–Leu–Gly–Met–NBD (**7i**)	Dnp–Pro–Leu–Gly–Met–Trp–Ser–Arg–NH_2[52]
Thermolysin	Metallo	Bz–Gly–Phe–NBD (**7j**)	*N*-[3-(2-furyl)acryloyl]-glycyl–L-leucine amide (FAGLA)[53,54]

Chemistry–A European Journal | Research Article doi.org/10.1002/chem.202301855 | Chemistry Europe European Chemical Societies Publishing

$$v_0 = \frac{v_{max} \cdot [S]}{K_M + [S]} \quad (1)$$

where v_0 is the initial reaction rate, at a given substrate concentration [*S*]. v_{max} is the maximum reaction rate, while K_M indicates [*S*] at which the conversion rate is half-maximal.

The fluorescence progress curves were acquired for at least six different substrate concentrations for the corresponding substrate with at least a technical triplicate. K_M and v_{max} values were then calculated from the respective Michaelis–Menten curves (Table 2 and Figures S55–S63). Exemplarily, fluorescence progress curves for different substrate concentrations and the resulting Michaelis–Menten plot are shown for the cysteine protease rhodesain in Figure 2. The signal-to-noise ratio, as described by the detector-independent reaction rate slope ([au·s^{-1}] vs. [*S*]), was ~3 times higher for NBD than for AMC-based substrates (calculated by the gain factor-independent initial slope of Figure 2C vs. 2D). Consequently, the NBD-based fluorescence assays can be performed at lower enzyme and/or substrate concentrations than with the corresponding parent

Table 2. Enzyme- and substrate-specific kinetic constants for NBD-based and corresponding AMC-based substrates. Michaelis–Menten constant (K_M), maximum turnover rate (v_{max}), turnover number (k_{cat}), catalytic efficiency (k_{cat}/K_M).

protease	NBD substrate				AMC substrate			
	K_M [μM]	v_{max} [nM·min^{-1}]	k_{cat} [min^{-1}]	k_{cat}/K_M [L·μmol^{-1} min^{-1}]	K_M [μM]	v_{max} [nM·min^{-1}]	k_{cat} [min^{-1}]	k_{cat}/K_M [L·μmol^{-1} min^{-1}]
Rhodesain	2.84 ± 0.22	73.8 ± 2.6	59.0 ± 2.1	20.8 ± 1.8	3.08 ± 0.41	679 ± 43	1056 ± 66	343 ± 50
uPA	10.2 ± 0.63	174 ± 6	[a]	[b]	347 ± 35	1847 ± 90	[c]	[d]
CatS	28.0 ± 2.68	70.0 ± 3.1	7.00 ± 0.31	0.25 ± 0.03	34.4 ± 3.06	111 ± 23	0.45 ± 0.09	0.013 ± 0.003
DENV NS2B/NS3	88.1 ± 11.7	762 ± 37	3.05 ± 0.15	0.035 ± 0.005	869 ± 67	41.8 ± 1.9	0.17 ± 0.01	(1.92 ± 0.17) 10^{-4}
ZIKV NS2B/NS3	33.6 ± 1.51	252 ± 5	10.1 ± 0.2	0.30 ± 0.01	1381 ± 371	161 ± 34	1.29 ± 0.27	(9.3 ± 3.3) 10^{-4}
PL^{pro}	40.6 ± 4.11	151 ± 8	1.51 ± 0.08	0.037 ± 0.004	1332 ± 161	7049 ± 427	70.5 ± 4.3	0.053 ± 0.007
M^{pro}	18.9 ± 1.38	10.9 ± 0.4	2.62 ± 0.09	0.14 ± 0.01	56.7 ± 11.6	2.06 ± 0.16	0.50 ± 0.04	0.009 ± 0.002

The concentration of the uPA protein in mol·L^{-1} is unknown but given in units (U) by the vendor: [a] (1.74·10^{-5} ± 1.20·10^{-7}) mol·min^{-1}·U^{-1}; [b] (1.71 ± 0.12) L·min^{-1}·U^{-1}; [c] (1.85·10^{-4} ± 9.00·10^{-6}) mol·min^{-1}·U^{-1}; [d] (0.53 ± 0.06) L·min^{-1}·U^{-1}.

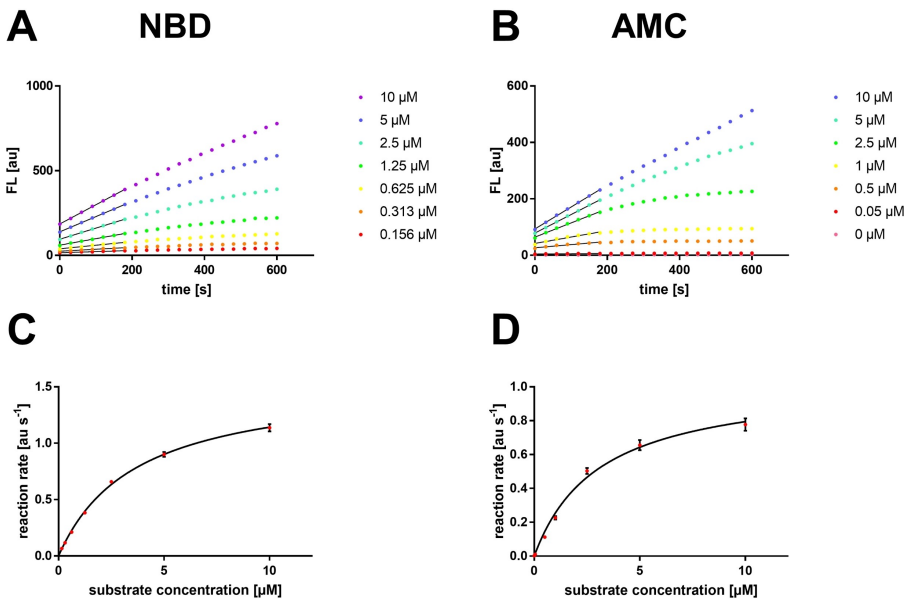


Figure 2. Exemplary fluorometric enzyme assay graphs for rhodesain. (A) Fluorescence progress curves for the conversion of different concentrations of the Z–Phe–Arg–NBD (7a) substrate by rhodesain. (B) Fluorescence progress curves for Z–Phe–Arg–AMC by rhodesain. (C) Michaelis–Menten plot for Z–Phe–Arg–NBD (7a). (D) Michaelis–Menten plot for Z–Phe–Arg–AMC.

Chemistry–A European Journal | Research Article doi.org/10.1002/chem.202301855 | Chemistry Europe European Chemical Societies Publishing

substrate. Especially for highly affine ligands, this might provide an advantage to avoid tight-binding inhibition ([E] > K_i).[55]

In general, the K_M values for the NBD-based substrates were found to be in the same order of magnitude or better than the K_M values determined for the AMC-based parent substrates, indicating similar or increased affinity compared to their AMC counterpart. While the K_M value of the cysteine proteases rhodesain, CatS, and M^{pro} substrates is virtually unchanged by the exchange of the fluorophore from AMC to NBD, the affinity of uPA, DENV/ZIKV NS2B/NS3, and PL^{pro} substrates is increased by up to a factor of ~40. The affinity increase was reasoned by analyzing the binding poses predicted by molecular docking. Comparing the docking poses of the AMC and the NBD substrate, it is noticeable that the positioning of the substrate peptide is mostly not affected by the exchange of the fluorophore residue. However, the positioning of the fluorophore itself might differ depending on the individual topology of the binding pocket. By its chemical nature, the NBD substructure can form more hydrophilic interactions compared to AMC as highlighted by the interaction analysis (Figure 3 and Figure S73). The formation of productive interactions with the charge-polarized catalytic dyad of a protease could be the reason for the improved affinity of NBD-based substrates. Some examples in which the NDB residue mediates increased affinity for protein and lipid surfaces have been previously documented.[32]

Considering the kinetics of catalytic cleavage, for the rhodesain, uPA, and PL^{pro} substrates, k_{cat} is lower for these NBD-based substrates than for the AMC-based substrates (10–100×). For the other enzymes, the k_{cat} values are in the same magnitude or slightly higher for the NBD-based substrates. Thus, no clear trend is apparent regarding the substrate cleavage rate, which seems to be an individual characteristic for each protease probably depending on the positioning of the susceptible amide bond towards the nucleophilic enzyme residue. Except for the rhodesain substrate, the catalytic efficiency k_{cat}/K_M of the NBD-based substrates was found to be as high as or higher than the one of the AMC-based substrates. In the case of the rhodesain NBD substrate (**7a**) the reduction of the k_{cat} value actually represents an advantage over the AMC substrate, since the fluorescence progress curves remain linear for a longer time in the steady-state regiment (Figure 2A) and there is no early saturation of the fluorescence by reaching an equilibrium, as is the case with the AMC substrate (Figure 2B).

Autocleavage of a DENV and ZIKV NS2B/NS3 Protease Substrate

Remarkably, in contrast to all other substrates presented in this study, our first DENV and ZIKV NS2B/NS3 protease substrate **7g** with the sequence Bz–Nle–Lys–Lys–Arg–NBD (Table 1) showed a fluorescence increase even in the absence of any enzyme. In an alkaline buffered solution, this fluorescence increase was found to be higher than at neutral pH values (Figure 4A and B). To analyze the molecular mechanism of this non-enzymatic fluorescence activation, the reaction mixture of the NS2B/NS3 substrate conversion was analyzed by HPLC/MS. By this, two LC peaks of the same mass (405.70 Da) but different fluorescence properties were detected, one corresponding to the substrate educt **7g** and the other to a fluorescent product **8**, suggesting an intramolecular rearrangement mechanism for fluorescence auto-activation (Figure S49). In the literature, an intramolecular Smiles rearrangement for NBD derivatives incorporating primary amine groups has been described during a histone deacetylase assay development.[37,39]

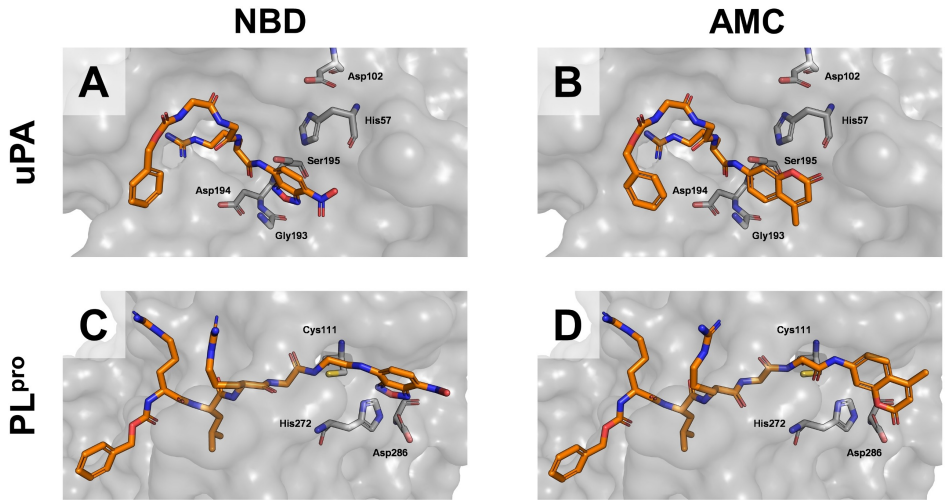


Figure 3. Predicted binding poses for the NBD- and AMC-based substrates for uPA and PL^{pro}. (A) Z–Gly–Gly–Arg–NBD (7f) in the active site of the crystal structure of uPA (pdb: 1LMW). (B) Z–Gly–Gly–Arg–AMC in the crystal structure of uPA (pdb: 1LMW). (C) Z–Arg–Leu–Arg–Gly–Gly–NBD (7e) in the crystal structure of PL^{pro} (pdb: 7RBS). (D) Z–Arg–Leu–Arg–Gly–Gly–AMC in the crystal structure of PL^{pro} (pdb: 7RBS).

Chemistry–A European Journal | Research Article doi.org/10.1002/chem.202301855 | Chemistry Europe European Chemical Societies Publishing

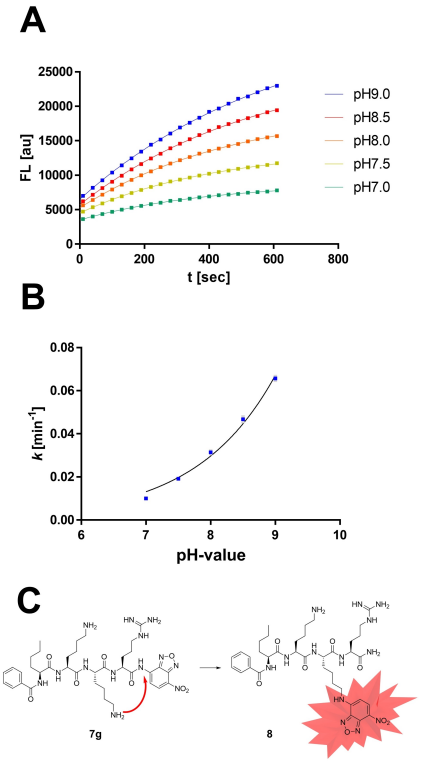


Figure 4. Autocleavage of the Bz–Nle–Lys–Lys–Arg–NBD (7 g) substrate. (A) Enzyme-independent fluorescence increase at different pH values (100 μM Bz–Nle–Lys–Lys–Arg–NBD in DENV NS2B/NS3 buffer). (B) pH-dependence of the non-enzymatic turnover. (C) Putative Smiles rearrangement of lysine-containing substrates.

From these and the literature results, we hypothesized that substrate auto-cleavage is enabled by the presence of free amine groups in the substrate sequence (Figure 4C). Appropriately, we found that substrates without lysine residues are stable in their respective assay buffer. This was confirmed both by mass spectroscopy and by fluorescence spectroscopy (Figures S50–S52). Nonetheless, we aimed to develop a functional substrate for the NS2B/NS3 proteases, and thus, we replaced the lysine residues with arginine residues, resulting in a stable and functional substrate for both proteases (Ac–Gly–Arg–Arg–NBD **7 h**, Table 2). Noteworthy, even high concentrations of buffers containing free amines (e.g., Tris 100 mM) do not lead to instability of the NBD substrates, however, the application scope of NBD-based substrates should be limited to lysine-free protease substrates. However, this does not represent a practical disadvantage in the development of protease substrates since there is no known protease that only tolerates lysine and not arginine.

Harnessing NBD-Amide Autocleavage for FRET Substrate Replacement

We developed the idea of harnessing the characteristic NBD autocleavage behavior for a FRET substrate replacement (Figure 5A). By this strategy, endopeptidases with a requirement for a specific amino acid in their prime site (S1′) might also become accessible to the repertoire of NBD substrates. Such proteases that have specificity for a S1′ amino acid usually must be assayed with a matching FRET substrate because AMC substrates usually do not fulfill the S1′ specificity requirement either. To test the hypothesis if a free *N*-terminus of an H_2N-Xaa-NBD derivative is able to drive the Smiles rearrangement, a minimal substrate (H_2N–Arg(Pbf)-NBD **21 a**) was analyzed fluorometrically and by LC/MS with enzyme-free assay conditions (100 mM Tris, 100 mM NaCl, variable pH) which highlighted that fluorescence auto-activation proceeds as expected (Figure S53), while at pH value between 5 and 8 the chemical turnover rate was moderately efficient ($k = 0.01$ min^{-1}). Outside this pH range the Smiles rearrangement might be less efficient but at very low or high pH values most proteases do not show enzymatic activity either.

We designed several NBD substrates from commonly known FRET-based parent substrates where the *N*-terminus of the prime site (S1′) amino acid can induce a Smiles rearrangement, resulting in fluorescence increase after proteolytic cleavage and subsequent rearrangement (Figure 5A). Exemplary, such substrates were designed for the proteases SrtA (Figure 5B–C), MMP9, and thermolysin. LC/MS spectra of the enzymatic conversion can be found in the Supporting Information (Figure S54). Similar to the above-described non-prime site substrates, the NBD-based substrates for thermolysin, SrtA, and MMP9 showed similar or better affinities (K_M) compared to their parent substrates (Table 3). Both the turnover rates k_{cat} and catalytic efficiency k_{cat}/K_M of these NBD substrates were lower than those of the parent substrate. This might be explained by the two-step reaction mechanism which is limited by the second step of the chemical rearrangement with moderate efficiency. However, the disadvantage of a lower processivity will be outweighed by the higher detection sensitivity and decreased susceptibility to assay interferences (see below).

Applications for Protease Inhibitor Investigation

To confirm the applicability of the NBD-based substrates in drug discovery-relevant fluorometric assays, the inhibition constants K_i of literature-known inhibitors were determined for each protease using NBD-based and the matching parent substrates (Table S2 and Figures S64–S72). Inhibition constants (K_i) were determined from the fluorescence progress curves using the Cheng-Prusoff equation (Eq. (2)).[56]

$$K_i = \frac{IC_{50}}{1 + \left(\frac{[S]}{K_M}\right)} \quad (2)$$

Chemistry–A European Journal | Research Article doi.org/10.1002/chem.202301855 | Chemistry Europe European Chemical Societies Publishing

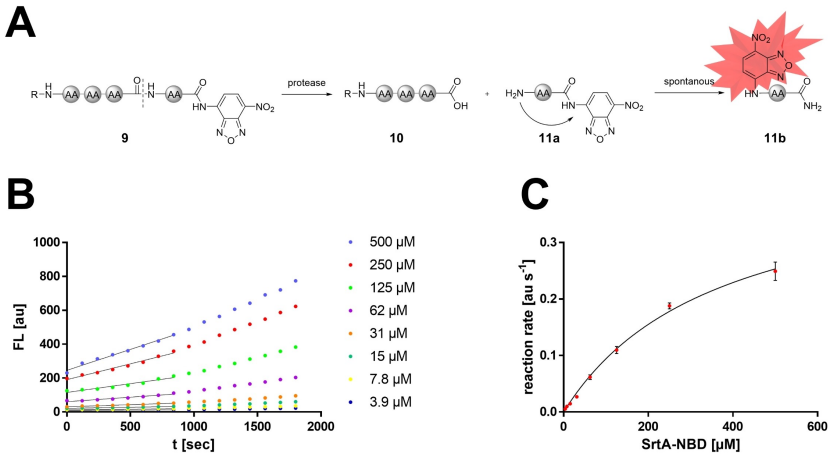


Figure 5. Prime site NBD auto-activating substrates as FRET substrate replacement. (A) Schematic depiction of the substrate activation mechanism. Rearrangement of the substrate after cleavage by the protease yields fluorophore activation. (B) Fluorescence progress curves for the conversion of the Bz–Leu–Pro–Ala–Thr–Gly–NBD (7d) substrate by SrtA. (C) Michaelis–Menten plot for the SrtA substrate.

Table 3. Enzyme- and substrate-specific kinetic constants for NBD-based and corresponding parent substrates. Michaelis–Menten constant (K_M), maximum turnover rate (v_{max}), turnover number (k_{cat}), catalytic efficiency (k_{cat}/K_M).

protease	NBD substrate				Parent substrate			
	K_M [µM]	v_{max} [nM·min^{-1}]	k_{cat} [min^{-1}]	k_{cat}/K_M [L·µmol^{-1} min^{-1}]	K_M [µM]	v_{max} [nM·min^{-1}]	k_{cat} [min^{-1}]	k_{cat}/K_M [L·µmol^{-1} min^{-1}]
Thermolysin	56.4 ± 4.22	52.8 ± 2.3	73.1 ± 3.1	1.30 ± 0.11	703 ± 110	(3.0 ± 0.3) 10^4	(4.2 ± 0.4) 10^4	59.1 ± 10.7
SrtA	365 ± 35[a]	156 ± 7	0.12 ± 0.01	(3.28 ± 0.35) 10^{-4}	92.5 ± 9.62	459 ± 22	0.35 ± 0.02	0.0038 ± 0.0004
MMP9	26.7 ± 7.70	10.3 ± 1.03	4.12 ± 0.40	0.15 ± 0.05	24.7 ± 4.8	241 ± 27	96.5 ± 10.7	3.91 ± 0.88

[a] The apparent K_M value of the SrtA FRET substrate is in accordance with the literature value which is distorted by the inner filter effect (Dnp internal quenching). An HPLC-derived K_M value was found to be in the magnitude of $K_M > 1$ mM.[58]

By this, inhibition values were corrected to the zero-substrate concentration, which 1.) allowed the comparison between different substrate affinities and their concentrations; and 2.) showed that the determined K_M values for NBD substrates (Tables 2 and 3) are valid compared to their reported parent substrates. Here, we could show that K_i values of NBD-based and parent substrates differ usually by less than 10% and in some cases by max. a factor of three. It can therefore be assumed that the NBD-based substrates are a well-suitable alternative for fluorometric assays in drug discovery.

Mitigation of Typical Protease Assay Interferences

The excitation and detection wavelengths of the NBD fluorophore ($\lambda_{ex} = 485$ nm, $\lambda_{em} = 535$ nm) are shifted towards the lower energy range compared to the commonly used AMC substrates ($\lambda_{ex} = 380$ nm, $\lambda_{em} = 460$ nm). In this regard, these assay conditions might be able to suppress the most typical assay interferences, because interfering physicochemical transitions do occur in the near UV-light spectrum.[24,25,27]

Using rhodesain as an example protease with a drug discovery context,[57] we investigated some compounds from our in-house library that are known to lead to false positive results while using the AMC-based rhodesain substrate. By using the newly developed NBD substrate, we aim to reduce these recurring assay interferences (Figure 6 and Figure S48). In this sense, we could show that the apparent inhibition in the AMC-based assay (attenuation of fluorescence increase) is due to common assay interferences, and the NBD-based substrate is not affected. Detailed molecular explanations of the interference effects can be found in the Supporting Information.

Smartphone-Based Fluorometer for Usage in Classroom Applications

The fluorescence of NBD-based dyes is relatively strong in quantum yield and brightness with the maximum sensitivity in the visible region (see above).[59,60] Thus, NBD substrates are potentially suitable for use with DIY or low-cost Vis-fluorometers to perform kinetic protease studies in classroom-like

Chemistry–A European Journal | Research Article doi.org/10.1002/chem.202301855 Chemistry Europe European Chemical Societies Publishing

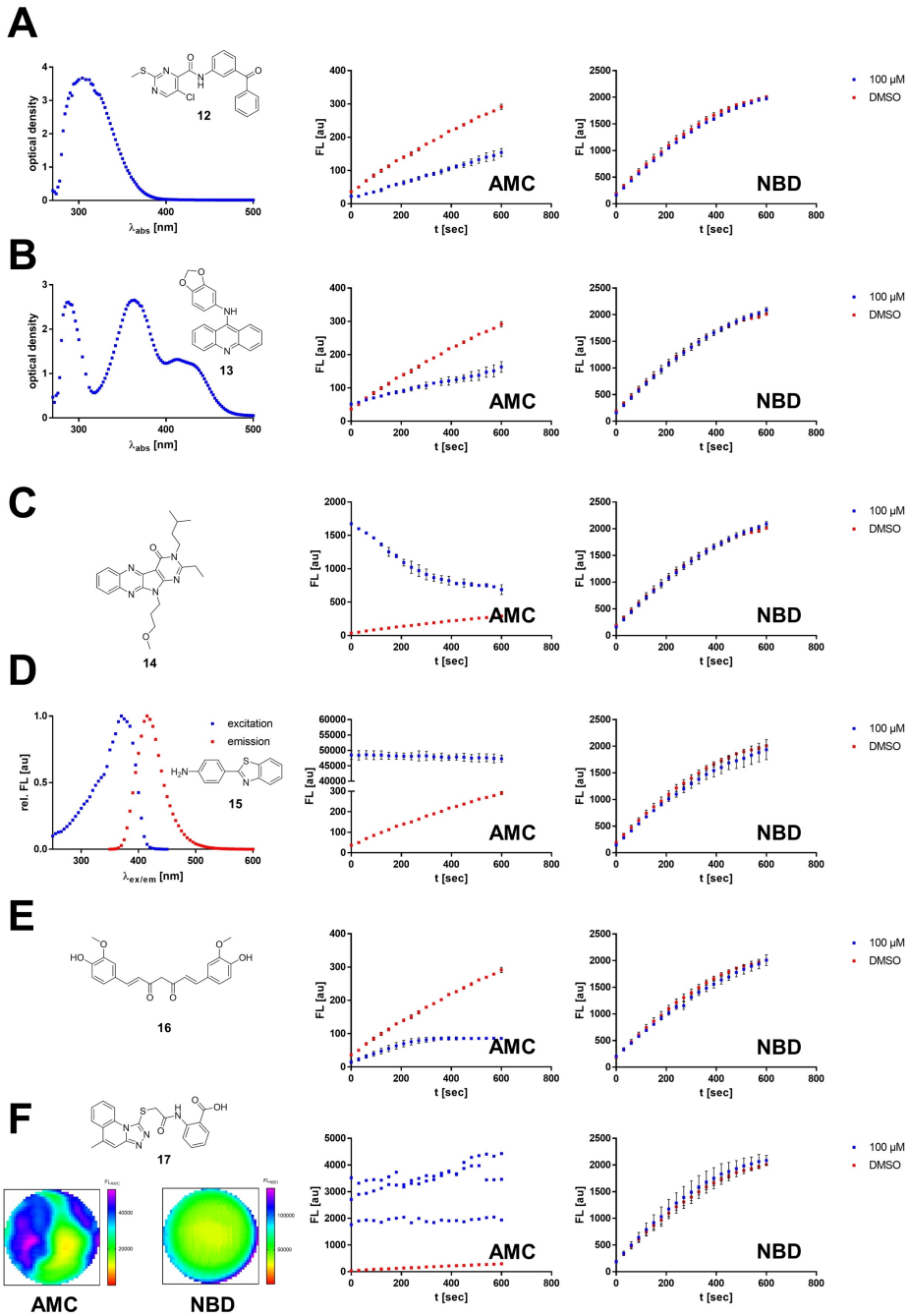


Figure 6. Mitigation of typical AMC assay interferences by an NBD substrate. Each interference example was performed for rhodesain assay conditions with the respective AMC and NBD substrate. (A) Inner filter effect of compound 12 (100 µM): compound absorbance spectrum; AMC-based assay; NBD-based assay. (B) Fluorophore quenching by compound 13 (100 µM): compound absorbance spectrum; AMC- and NBD-based assay. (C) Autofluorescence decay of compound 14 (100 µM) (D) Detector non-linearity by compound 15 (100 µM): compound fluorescence spectrum; AMC- and NBD-based assay. (E) Photosensitization of radical oxygen species by compound 16 (100 µM). (F) Reflecting inhibitor aggregates of compound 17 (100 µM): fluorometric well-homogeneity scan with AMC resp. NBD assay wavelengths; AMC- and NBD-based assay.

Chemistry–A European Journal | Research Article doi.org/10.1002/chem.202301855 | Chemistry Europe European Chemical Societies Publishing

experiments.[61–63] Previous fluorogenic substrates do not allow for such conceptualizations, as monitoring substrate turnover of AMC and FRET substrates requires an excitation light source in the UV range, which is rarely available in the context of educational activities.

In fact, we were able to show that substrate turnover of Z–Phe–Arg–NBD (500 μM) using the cysteine protease rhodesain (1 μM) is possible by using only a conventional smartphone (Xiaomi Redmi 10) as a detection device (Figure 7). For this purpose, the camera lens and the built-in lamp were covered with LeeFilter films (cost < 10 ct): orange film (2×2 cm, LeeFilters 158 Deep Orange) and blue film (2×2 cm, LeeFilters 120 Deep Blue) as excitation resp. emission filters (Figure 7A). Recording a video of the reaction mixture in a 96-well plate with the built-in camera and analyzing the brightness in the well over time enabled the recording of a substrate turnover curve (Figure 7C and Supporting Information Movie 1).

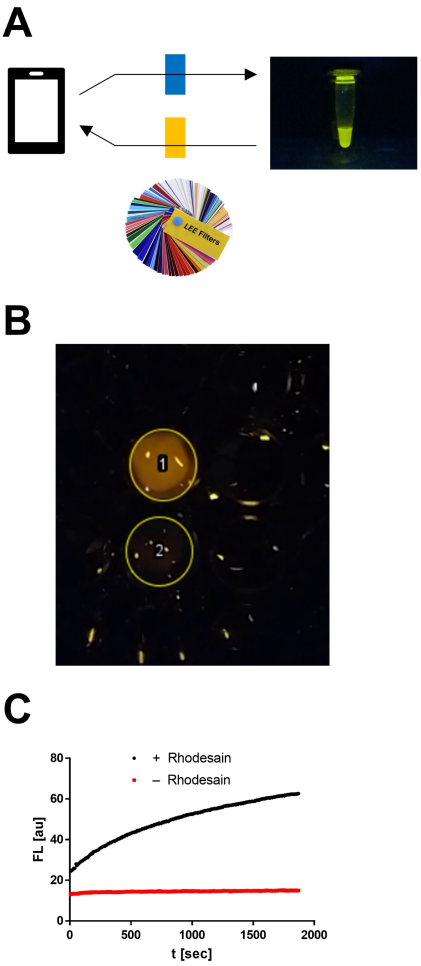


Figure 7. Smartphone-based assay for rhodesain activity. (A) Schematic representation of the experimental setup with a fluorescence-enabled photograph of an NBD-NH_2 solution (1 mM) in rhodesain assay buffer. (B) Screenshot of a reaction mixture containing rhodesain (1 μM) and the Z–Phe–Arg–NBD substrate (500 μM) in a 96-well plate. Circle 1 highlights a well after the addition of rhodesain enzyme, whereas circle 2 shows only NBD substrate in rhodesain assay buffer. (C) Fluorescence progress curves by analysis of the brightness in the wells (ImageJ 1.53).

Conclusions

In this study, protease substrates with NBD as a fluorogenic reporter group were designed, synthesized, and tested for utilization in drug discovery applications. Compared to the previous substrate gold standards, the absorption and emission characteristics of NBD are red-shifted to the visible spectrum. In total, we described the synthesis of substrates for 10 different proteases from the cysteine-, serine-, and metalloprotease classes. By this, we were able to obtain substrates for proteases with carboxypeptidase activity (XXX|NBD) as well as for proteases with endopeptidase activity and specificity requirements in the prime site (XX|**X**–NBD). The suitability of the new substrates was confirmed by determining enzyme- and substrate-specific kinetic parameters. The K_M values determined for the NBD-based substrates were in the same order of magnitude as the parent AMC- resp. FRET substrates. Determination of inhibition constants (K_i and k_{inact}) of known inhibitors confirmed that NBD substrates are well-suitable for various applications. In this regard, typical medicinal chemistry-related assay interferences such as the inner filter effect, fluorophore quenching, and autofluorescence decay could be avoided by using the NBD substrates. Thus, in summary, NBD substrates may not only provide a suitable alternative to previously used fluorophores but may form the basis of a new generation of fluorometric assays.

Experimental Section

Fluorometric assays: Fluorometric assays were performed with a Tecan Spark 10 M plate reader using white flat-bottom 96-well microtiter plates from Greiner Bio-One. Measurements were performed in at least three independent technical replicates. Typically, each well contained 180 μL buffer, 5 μL enzyme stock solution, 10 μL inhibitor in DMSO or pure DMSO for mock treatment, and 5 μL solution of the corresponding protease substrate in DMSO, resulting in a total volume of 200 μL (for MMP9: total volume 100 μL, 1 μL substrate, 98 μL buffer, 1 μL enzyme). The fluorescence was measured for 10 min or for 30 min every 30 s with the following excitation and emission wavelengths.

Emission and Excitation wavelengths. AMC: $\lambda_{ex} = 380$ nm, $\lambda_{em} = 460$ nm; FAGLA: $\lambda_{abs} = 322$ nm; Abz/Dap(dnp): $\lambda_{ex} = 320$ nm, $\lambda_{em} = 430$ nm; NBD: $\lambda_{ex} = 485$ nm, $\lambda_{em} = 535$ nm; Dnp/Trp: $\lambda_{ex} = 280$ nm, $\lambda_{em} = 360$ nm.

Enzyme Buffers and Substrates. **Rhodesain** (enzyme: 1 nM, 50 mM sodium acetate pH 5.5, 5 mM EDTA, 200 mM NaCl, 5 mM DTT, 10 μM Z–Phe–Arg–AMC/10 μM Z–Phe–Arg–NBD); **SARS-CoV 2 M^{pro}** (enzyme: 250 nM, 20 mM Tris pH 7.5, 0.1 mM EDTA, 200 mM NaCl, 1 mM DTT, 60 μM Boc–Abu–Tle–Leu–Gln–AMC/30 μM Boc–Abu–Tle–Leu–Gln–NBD); **DENV2 and ZIKV NS2B/NS3** (en-

Chemistry–A European Journal | Research Article doi.org/10.1002/chem.202301855 | Chemistry Europe European Chemical Societies Publishing

zyme: 250 nM (DENV) & 25 nM (ZIKV), 50 mM Tris pH 7.5, 100 μM Boc–Gly–Arg–Arg–AMC/50 μM Z–Gly–Arg–Arg–NBD); **Urokinase plasminogen activator (uPA)** (enzyme: 10 U, 50 mM Tris pH 7.4, 50 mM NaCl, 0.5 mM EDTA, 240 μM Z–Gly–Gly–Arg–AMC/10 μM Z–Gly–Gly–Arg–NBD); **Sortase A (SrtA)** (enzyme: 1.3 μM, 50 mM Tris pH 7.5, 150 mM NaCl, 5 mM $CaCl_2$, 0.5 mM Gly_4, 25 μM Abz-LPETG-Dap(dnp)-OH/100 μM Bz–Leu–Pro–Ala–Thr–Gly–NBD); **SARS-CoV 2 PL^{pro}** (enzyme: 100 nM, 20 mM Tris pH 7.5, 0.1 mM EDTA, 200 mM NaCl, 1 mM DTT, 50 μM Z–Arg–Leu–Arg–Gly–Gly–AMC/40 μM Z–Arg–Leu–Arg–Gly–Gly–NBD); **Cathepsin S (CatS)** (enzyme: 10 nM, enzyme buffer: 35 mM K_3PO_4 pH 6.5, 35 mM sodium acetate, 2 mM DTT, 2 mM EDTA, assay buffer: 50 mM K_3PO_4 pH 6.5, 2.5 mM DTT, 2.5 mM EDTA, 10 μM Z–Val–Val–Arg–AMC/10 μM Z–Val–Val–Arg–NBD); **Matrixmetalloprotease 9 (MMP9)** (enzyme: 2.5 nM, 150 mM NaCl, 50 mM Tris–HCl pH 7.5, 20 μM $ZnCl_2$, 1 mM $CaCl_2$, Dnp–Pro–Leu–Gly–Met–Trp–Ser–Arg–NH_2/40 μM Z–Pro–Leu–Gly–Met–NBD); **Thermolysin** (enzyme: 0.7 nM, 100 mM Tris, 100 mM NaBr, 2.5 mM $CaCl_2$, 100 μM FAGLA/20 μM Bz–Gly–Phe–NBD).

Chemistry: All reagents and solvents were of analytical grade quality and purchased from Sigma-Aldrich, Carbolution, BLDpharm, or FisherScientific. Chemicals were used without further purification. ^1H and ^{13}C spectra were recorded on a Bruker Fourier 300 or Bruker Avance III 600 using DMSO-d_6 or $CDCl_3$, as solvent. Chemical shifts δ are given in parts per million (ppm) using residual proton peaks of the solvent as an internal standard. UV-chromatograms and mass spectra were obtained by LC-MS consisting of an Agilent 1100 series HPLC system using an Agilent Poroshell 120 EC-C18 150×2.10 mm, 4 μm column or an Agilent Zorbax SB Aq. 150×4.6 mm, 5 μm column. Detection wavelengths were 210, 254, and 396 nm. The tested substrates and inhibitors displayed a purity ≥95% in all cases. The molecular mass was confirmed by an Agilent 1100 series LC/MSD Trap with electron spray ionization (ESI) in positive ionization mode. Melting points (uncorrected) were determined in open capillaries using a Schorpp Gerätetechnik MPM-H3 melting point device. Specific rotations $[\alpha]_D^{20}$ were measured on a P3000 polarimeter from Krüss and are reported in $cm^3\ g^{-1}\ dm^{-1}$. Detailed information on the synthesis and analytical data for characterization of NBD-substrates can be found in the Supporting Information.

Supporting Information

The authors have cited additional references within the Supporting Information.[27,41,47,52–54,56,64–89]

Additional experimental Procedures (Protein Constructs and Expression, Fluorometric Assay, Molecular Docking, Chemistry) and Results and Discussion (Autocleavage of NS2B/NS3 Protease Substrate, Stability of Substrates, Fluorometric Assay Data).

Acknowledgements

T.S. acknowledges financial support from the German Research Foundation (DFG), Collaborative Research Center (CRC) SFB 1066 (projects Q5). F.B. acknowledges financial support from the Johannes Gutenberg University of Mainz. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interests

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: assay interferences · fluorescent probes · medicinal chemistry · nitrobenzofurazane · proteases

[1] S. Udenfriend, in *Fluoresc. Assay Biol. Med.*, Elsevier, **1969**, pp. 1–41.
[2] F. Ciruela, *Curr. Opin. Biotechnol.* **2008**, *19*, 338–343.
[3] T. D. Pollard, *Mol. Biol. Cell* **2010**, *21*, 4061–4067.
[4] M. Leopoldo, E. Lacivita, F. Berardi, R. Perrone, *Drug Discovery Today* **2009**, *14*, 706–712.
[5] H. Maus, G. Hinze, S. J. Hammerschmidt, T. Basché, T. Schirmeister, *Protein Sci.* **2023**, *32*.
[6] C. Götz, G. Hinze, A. Gellert, H. Maus, F. von Hammerstein, S. J. Hammerschmidt, L. M. Lauth, U. A. Hellmich, T. Schirmeister, T. Basché, *J. Phys. Chem. B* **2021**, *125*, 6837–6846.
[7] A. Eatemadi, H. T. Aiyelabegan, B. Negahdari, M. A. Mazlomi, H. Daraee, N. Daraee, R. Eatemadi, E. Sadroddiny, *Biomed. Pharmacother.* **2017**, *86*, 221–231.
[8] P. Müller, H. Maus, S. J. Hammerschmidt, P. M. Knaff, V. Mailänder, T. Schirmeister, C. Kersten, *Curr. Med. Chem.* **2022**, *29*, 635–665.
[9] C. López-Otín, J. S. Bond, *J. Biol. Chem.* **2008**, *283*, 30433–30437.
[10] G. Abbenante, D. Fairlie, *Med. Chem.* **2005**, *1*, 71–104.
[11] P. D. Leeson, A. P. Bento, A. Gaulton, A. Hersey, E. J. Manners, C. J. Radoux, A. R. Leach, *J. Med. Chem.* **2021**, *64*, 7210–7230.
[12] Y. N. Lamb, *Drugs* **2022**, *82*, 585–591.
[13] M. Drag, G. S. Salvesen, *Nat. Rev. Drug Discovery* **2010**, *9*, 690–701.
[14] B. Turk, *Nat. Rev. Drug Discovery* **2006**, *5*, 785–799.
[15] V. Canbay, U. auf dem Keller, *Curr. Opin. Chem. Biol.* **2021**, *60*, 89–96.
[16] I. L. H. Ong, K.-L. Yang, *Analyst* **2017**, *142*, 1867–1881.
[17] A. Carmel, M. Zur, A. Yaron, E. Katchalski, *FEBS Lett.* **1973**, *30*, 11–14.
[18] H. Tuppy, U. Wiesbauer, E. Wintersberger, *Hoppe-Seyler's Z. Physiol. Chem.* **1962**, *329*, 278–288.
[19] M. Zimmerman, E. Yurewicz, G. Patel, *Anal. Biochem.* **1976**, *70*, 258–262.
[20] S. Markossian, A. Grossman, K. Brimacombe, M. Arkin, D. Auld, C. Austin, J. Baell, T. Chung, N. P. Coussens, J. L. Dahlin, *Assay Guidance Manual*, Eli Lilly & Company and The National Center For Advancing Translational Sciences, **2004**.
[21] B. C. Chenna, L. Li, D. M. Mellott, X. Zhai, J. L. Siqueira-Neto, C. Calvet Alvarez, J. A. Bernatchez, E. Desormeaux, E. Alvarez Hernandez, J. Gomez, J. H. McKerrow, J. Cruz-Reyes, T. D. Meek, *J. Med. Chem.* **2020**, *63*, 3298–3316.
[22] T. T. Baird, C. S. Craik, in *Handb. Proteolytic Enzym.*, **2013**, pp. 2594–2600.
[23] A. Jadhav, R. S. Ferreira, C. Klumpp, B. T. Mott, C. P. Austin, J. Inglese, C. J. Thomas, D. J. Maloney, B. K. Shoichet, A. Simeonov, *J. Med. Chem.* **2010**, *53*, 37–51.
[24] J. Comley, *Drug Discov. World Summer 2003* **2003**, *4*, 91–98..
[25] S. K. Grant, J. G. Sklar, R. T. Cummings, *SLAS Discov.* **2002**, *7*, 531–540.
[26] Y. Liu, W. Kati, C.-M. Chen, R. Tripathi, A. Molla, W. Kohlbrenner, *Anal. Biochem.* **1999**, *267*, 331–335.
[27] N. Thorne, D. S. Auld, J. Inglese, *Curr. Opin. Chem. Biol.* **2010**, *14*, 315–324.
[28] F. E. Jernigan, D. S. Lawrence, *Chem. Commun.* **2013**, *49*, 6728.
[29] J.-M. Delaissé, Y. Eeckhout, G. Vaes, *Biochem. Biophys. Res. Commun.* **1984**, *125*, 441–447.
[30] S. G. Deeks, M. Smith, M. Holodniy, J. O. Kahn, *J. Am. Med. Assoc.* **1997**, *277*, 145–153.
[31] S. Jo, H. Kim, S. Kim, D. H. Shin, M. Kim, *Chem. Biol. Drug Des.* **2019**, *94*, 2023–2030.
[32] C. Jiang, H. Huang, X. Kang, L. Yang, Z. Xi, H. Sun, M. D. Pluth, L. Yi, *Chem. Soc. Rev.* **2021**, *50*, 7436–7495.
[33] K. Okada, T. Yamaguchi, K. Dodo, M. Sodeoka, S. Obika, *Bioorg. Med. Chem.* **2019**, *27*, 1444–1448.

Chemistry–A European Journal | Research Article doi.org/10.1002/chem.202301855 | Chemistry Europe European Chemical Societies Publishing

[34] Y. Xie, L. Chen, R. Wang, J. Wang, J. Li, W. Xu, Y. Li, S. Q. Yao, L. Zhang, Q. Hao, H. Sun, *J. Am. Chem. Soc.* **2019**, *141*, 18428–18436.
[35] M. He, Z. Han, J. Qiao, L. Ngo, M. P. Xiong, Y. G. Zheng, *Chem. Commun.* **2018**, *54*, 5594–5597.
[36] I. N. Gober, M. L. Waters, *J. Am. Chem. Soc.* **2016**, *138*, 9452–9459.
[37] Y. Xie, J. Ge, H. Lei, B. Peng, H. Zhang, D. Wang, S. Pan, G. Chen, L. Chen, Y. Wang, Q. Hao, S. Q. Yao, H. Sun, *J. Am. Chem. Soc.* **2016**, *138*, 15596–15604.
[38] J. M. An, S. Kang, E. Huh, Y. Kim, D. Lee, H. Jo, J. F. Joung, V. J. Kim, J. Y. Lee, Y. S. Dho, Y. Jung, J. K. Hur, C. Park, J. Jung, Y. Huh, J.-L. Ku, S. Kim, T. Chowdhury, S. Park, J. S. Kang, M. S. Oh, C.-K. Park, D. Kim, *Chem. Sci.* **2020**, *11*, 5658–5668.
[39] D. Castagnolo, M. Pagano, M. Bernardini, M. Botta, *Tetrahedron Lett.* **2012**, *53*, 5008–5011.
[40] P. M. Knaff, P. Müller, C. Kersten, L. Wettstein, J. Münch, K. Landfester, V. Mailänder, *Biomacromolecules* **2022**, *23*, 2236–2242.
[41] Y. Yang, D. Zhang, M. Xu, J. Wang, J. Chen, L. Wang, *Monatsh. Chem.* **2018**, *149*, 1003–1008.
[42] L. F. Bernal-Perez, L. Prokai, Y. Ryu, *Anal. Biochem.* **2012**, *428*, 13–15.
[43] P. B. Ghosh, M. W. Whitehouse, *Biochem. J.* **1968**, *108*, 155–156.
[44] S. Jung, N. Fuchs, C. Grathwol, U. A. Hellmich, A. Wagner, E. Diehl, T. Willmes, C. Sotriffer, T. Schirmeister, *Eur. J. Med. Chem.* **2022**, *238*, 114460.
[45] D. Dana, A. R. Davalos, S. De, P. Rathod, R. K. Gamage, J. Huestis, N. Afzal, Y. Zavlanov, S. S. Paroly, S. A. Rotenberg, G. Subramaniam, K. J. Mark, E. J. Chang, S. Kumar, *Bioorg. Med. Chem.* **2013**, *21*, 2975–2987.
[46] W. Rut, K. Groborz, L. Zhang, X. Sun, M. Zmudzinski, B. Pawlik, X. Wang, D. Jochmans, J. Neyts, W. Młynarski, R. Hilgenfeld, M. Drag, *Nat. Chem. Biol.* **2021**, *17*, 222–228.
[47] F. Barthels, G. Marincola, T. Marciniak, M. Konhäuser, S. Hammerschmidt, J. Bierlmeier, U. Distler, P. R. Wich, S. Tenzer, D. Schwarzer, W. Ziebuhr, T. Schirmeister, *ChemMedChem* **2020**, *15*, 839–850.
[48] A. Narayanan, M. Narwal, S. A. Majowicz, C. Varricchio, S. A. Toner, C. Ballatore, A. Brancale, K. S. Murakami, J. Jose, *Commun. Biol.* **2022**, *5*, 169.
[49] A. Dal Corso, M. Catalano, A. Schmid, J. Scheuermann, D. Neri, *Angew. Chem.* **2018**, *130*, 17424–17428.
[50] H. Lai, G. Sridhar Prasad, R. Padmanabhan, *Antiviral Res.* **2013**, *97*, 74–80.
[51] H. Maus, F. Barthels, S. J. Hammerschmidt, K. Kopp, B. Millies, A. Gellert, A. Ruggieri, T. Schirmeister, *Bioorg. Med. Chem.* **2021**, *47*, 116392.
[52] S. Netzel-Arnett, S. K. Mallya, H. Nagase, H. Birkedal-Hansen, H. E. Van Wart, *Anal. Biochem.* **1991**, *195*, 86–92.
[53] P. A. Bartlett, C. K. Marlowe, *Biochemistry* **1987**, *26*, 8553–8561.
[54] S. M. Khan, D. W. Darnall, *Anal. Biochem.* **1978**, *86*, 332–336.
[55] D. J. Murphy, *Anal. Biochem.* **2004**, *327*, 61–67.
[56] T. Schirmeister, J. Kesselring, S. Jung, T. H. Schneider, A. Weickert, J. Becker, W. Lee, D. Bamberger, P. R. Wich, U. Distler, S. Tenzer, P. Johé, U. A. Hellmich, B. Engels, *J. Am. Chem. Soc.* **2016**, *138*, 8332–8335.
[57] R. Ettari, S. Previti, L. Tamborini, G. Cullia, S. Grasso, M. Zappalà, *Mini-Rev. Med. Chem.* **2016**, *16*, 1374–1391.
[58] R. G. Kruger, P. Dostal, D. G. McCafferty, *Anal. Biochem.* **2004**, *326*, 42–48.
[59] S. Uchiyama, T. Santa, T. Fukushima, H. Homma, K. Imai, *J. Chem. Soc. Perkin Trans. 2* **1998**, 2165–2174.
[60] B. E. Tebikachew, F. Edhborg, N. Kann, B. Albinsson, K. Moth-Poulsen, *Phys. Chem. Chem. Phys.* **2018**, *20*, 23195–23201.
[61] M. A. Hossain, J. Canning, S. Ast, T. L. Yen, P. Rutledge, A. Jamalipour, in *Adv. Photonics*, OSA, Washington, D. C., **2014**, p. SeTh2 C.1.
[62] L. M. Azad, H. Ehtesabi, A. Rezaei, *Nano-Struct.Nano-Objects* **2021**, *26*, 100722.
[63] A. Jamalipour, M. A. Hossain, **2019**, pp. 15–36.
[64] G. Amendola, R. Ettari, S. Previti, C. Di Chio, A. Messere, S. Di Maro, S. J. Hammerschmidt, C. Zimmer, R. A. Zimmermann, T. Schirmeister, M. Zappalà, S. Cosconati, *J. Chem. Inf. Model.* **2021**, *61*, 2062–2073.
[65] B. Millies, F. Von Hammerstein, A. Gellert, S. Hammerschmidt, F. Barthels, U. Göppel, M. Immerheiser, F. Elgner, N. Jung, M. Basic, C. Kersten, W. Kiefer, J. Bodem, E. Hildt, M. Windbergs, U. A. Hellmich, T. Schirmeister, *J. Med. Chem.* **2019**, *62*, 11359–11382.
[66] L. Schmohl, J. Bierlmeier, N. von Kügelgen, L. Kurz, P. Reis, F. Barthels, P. Mach, M. Schutkowski, C. Freund, D. Schwarzer, *Bioorg. Med. Chem.* **2017**, *25*, 5002–5007.
[67] Y. M. Báez-Santos, S. J. Barraza, M. W. Wilson, M. P. Agius, A. M. Mielech, N. M. Davis, S. C. Baker, S. D. Larsen, A. D. Mesecar, *J. Med. Chem.* **2014**, *57*, 2393–2412.
[68] J. Joossens, P. Van der Veken, G. Surpateanu, A.-M. Lambeir, I. El-Sayed, O. M. Ali, K. Augustyns, A. Haemers, *J. Med. Chem.* **2006**, *49*, 5785–5793.
[69] B. J. Buckley, A. Aboelela, E. Minaei, L. X. Jiang, Z. Xu, U. Ali, K. Fildes, C.-Y. Cheung, S. M. Cook, D. C. Johnson, D. A. Bachovchin, G. M. Cook, M. Apte, M. Huang, M. Ranson, M. J. Kelso, *J. Med. Chem.* **2018**, *61*, 8299–8320.
[70] A. Welker, C. Kersten, C. Müller, R. Madhugiri, C. Zimmer, P. Müller, R. Zimmermann, S. Hammerschmidt, H. Maus, J. Ziebuhr, C. Sotriffer, T. Schirmeister, *ChemMedChem* **2021**, *16*, 340–354.
[71] D. Brömme, *Curr. Protoc. Protein Sci.* **2000**, *21*.
[72] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, *J. Med. Chem.* **2004**, *47*, 1739–1749.
[73] G. N. Lewis, M. Calvin, *Chem. Rev.* **1939**, *25*, 273–328.
[74] L. D. Dias, K. C. Blanco, I. S. Mfouo-Tynga, N. M. Inada, V. S. Bagnato, *J. Photochem. Photobiol. C* **2020**, *45*, 100384.
[75] I. Ahmad, S. Ahmed, Z. Anwar, M. A. Sheraz, M. Sikorski, *Int. J. Photoenergy* **2016**, *2016*, 1–19.
[76] K. T. Kazantzis, K. Koutsonikoli, B. Mavroidi, M. Zachariadis, P. Alexiou, M. Pelecanou, K. Politopoulos, E. Alexandratou, M. Sagnou, *Photochem. Photobiol. Sci.* **2020**, *19*, 193–206.
[77] K. M. Nelson, J. L. Dahlin, J. Bisson, J. Graham, G. F. Pauli, M. A. Walters, *J. Med. Chem.* **2017**, *60*, 1620–1637.
[78] D. Steverding, *Molecules* **2019**, *25*, 143.
[79] S. Ludewig, M. Kossner, M. Schiller, K. Baumann, T. Schirmeister, *Curr. Top. Med. Chem.* **2010**, *10*, 368–382.
[80] Lord Rayleigh, *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1881**, *12*, 81–101.
[81] S. Jung, N. Fuchs, P. Johe, A. Wagner, E. Diehl, T. Yuliani, C. Zimmer, F. Barthels, R. A. Zimmermann, P. Klein, W. Waigel, J. Meyr, T. Opatz, S. Tenzer, U. Distler, H.-J. Räder, C. Kersten, B. Engels, U. A. Hellmich, J. Klein, T. Schirmeister, *J. Med. Chem.* **2021**, *64*, 12322–12358.
[82] R. D. A. Wilkinson, A. Young, R. E. Burden, R. Williams, C. J. Scott, *Mol. Cancer* **2016**, *15*, 29.
[83] L. Fu, S. Shao, Y. Feng, F. Ye, X. Sun, Q. Wang, F. Yu, Q. Wang, B. Huang, P. Niu, X. Li, C. C. L. Wong, J. Qi, W. Tan, G. F. Gao, *mBio* **2021**, *12*.
[84] S. Previti, R. Ettari, E. Calcaterra, C. Di Chio, R. Ravichandran, C. Zimmer, S. Hammerschmidt, A. Wagner, M. Bogacz, S. Cosconati, T. Schirmeister, M. Zappalà, *ACS Med. Chem. Lett.* **2022**, *13*, 1083–1090.
[85] M. A. Abduraman, M. Hariono, R. Yusof, N. A. Rahman, H. A. Wahab, M. L. Tan, *Heliyon* **2018**, *4*, e01023.
[86] K. H. Nam, *J. Inorg. Biochem.* **2021**, *215*, 111319.
[87] M. Johansson, A. J. Brooks, D. A. Jans, S. G. Vasudevan, *J. Gen. Virol.* **2001**, *82*, 735–745.
[88] F. Hammerstein, L. M. Lauth, S. Hammerschmidt, A. Wagner, T. Schirmeister, U. A. Hellmich, *FEBS Lett.* **2019**, *593*, 2204–2213.
[89] C. R. Caffrey, E. Hansell, K. D. Lucas, L. S. Brinen, A. Alvarez Hernandez, J. Cheng, S. L. Gwaltney, W. R. Roush, Y.-D. Stierhof, M. Bogyo, D. Steverding, J. H. McKerrow, *Mol. Biochem. Parasitol.* **2001**, *118*, 61–73.

Manuscript received: June 12, 2023
Accepted manuscript online: June 14, 2023
Version of record online: July 28, 2023

Chemistry–A European Journal

Supporting Information

Next Generation of Fluorometric Protease Assays: 7-Nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-Amides) as Class-Spanning Protease Substrates

Hannah Maus, Patrick Müller, Mergim Meta, Sabrina N. Hoba, Stefan J. Hammerschmidt, Robert A. Zimmermann, Collin Zimmer, Natalie Fuchs, Tanja Schirmeister, and Fabian Barthels*

Table of Contents

Experimental Procedures

Protein Constructs

Urokinase plasminogen activator (uPA), matrix metalloprotease 9 (MMP9), cathepsin S (CatS), and thermolysin were purchased from Sigma Aldrich (SAE0078, 672112, 219343, and P1512).

SARS-CoV 2 M^{pro}. The pMal-c2 plasmid harboring the C-terminal hexahistidine (His_6)-tagged sequence of SARS-CoV 2 M^{pro} (GenBank ID: UCC01618.1) was kindly provided by Prof. John Ziebuhr (Justus Liebig University Gießen, Germany). The sequence contains the nsp4/nsp5 M^{pro} cleavage site between maltose binding protein (MBP) and M^{pro} as well as the nsp5/nsp6 cleavage site between M^{pro} and the His_6-tag, enabling the purification of M^{pro} with native termini.

DENV2/ZIKV NS2B/NS3. The pET-15b expression vector, containing a thrombin-cleavable N-terminal His_6-tagged DENV2 NS2B cofactor domain covalently linked to the NS3 protease domain via a $Gly_4SerGly_4$-linker (GenBank ID: AY037116.1) with two point mutations in the NS3 region (I30A and L31A for better solubility), was kindly provided by the group of Prof. Wibke Diederich (University of Marburg, Germany).[1] The pET-11a vector, containing the N-terminal His_6-tagged sequence of the French Polynesia ZIKV strain (GenBank ID: KJ776791.1) NS2B covalently linked to the NS3 protease domain via a $Gly_4SerGly_4$-linker was designed and ordered from Genscript (New Jersey, USA).[2] To avoid autocatalytic cis-cleavage of the two protease domains, a point mutation (R95A) was introduced into the NS2B sequence.[2] Both His_6 tags can be cleaved by either thrombin (DENV2) or tobacco etch virus (TEV) protease (ZIKV).

***S. aureus* Sortase A (SrtA).** A pET-23b expression construct, containing the codon-optimized sequence for SrtA (GenBank ID: MBE7584187.1) with an uncleavable C-terminal His_6-tag, was kindly provided by the group of Prof. Dirk Schwarzer (University of Tübingen, Germany).[3]

SARS-CoV 2 PL^{pro}. The gene sequence coding for SARS CoV 2 PL^{pro} was obtained from the Protein Databank (PDB-ID: 6W9C). It was synthesized and inserted into a pET-28b vector by Genscript with NcoI and XhoI as restriction sites. Finally, the sequence was verified by Eurofins Genomics. The construct harbors a C-terminal (His_6)-Tag for purification which needs not be removed for protease activity.

***T. brucei* Cathepsin L (rhodesain).** The general construct and the method for heterologous expression in *P. pastoris* have been described in the literature.[4,5] The expression system was *P. pastoris* X-33, stably transfected with a rhodesain gene lacking the C-terminal domain using the pPICZα vector, enabling AOX-controlled expression when methanol is the sole carbon source.

Fluorometric Assay

Fluorometric assays were performed with a Tecan Spark 10M plate reader using white flat-bottom 96-well microtiter plates from Greiner Bio-One. Measurements were performed in at least three independent technical replicates. Typically, each well contained 180 µL buffer, 5 µL enzyme stock solution, 10 µL inhibitor in DMSO or pure DMSO for mock treatment, and 5 µL solution of the corresponding protease substrate in DMSO, resulting in a total volume of 200 µL (for MMP9: total volume 100 µL, 1 µL substrate, 98 µL buffer, 1 µL enzyme). The fluorescence was measured for 10 min or for 30 min every 30 s with the following excitation and emission wavelengths.
Emission and Excitation wavelengths. AMC: λ_{ex} = 380 nm, λ_{em} = 460 nm; FAGLA: λ_{abs} = 322 nm; Abz/Dap(dnp): λ_{ex} = 320 nm, λ_{em} = 430 nm; NBD: λ_{ex} = 485 nm, λ_{em} = 535 nm; Dnp/Trp: λ_{ex} = 280 nm, λ_{em} = 360 nm.
Enzyme Buffers and Substrates. **Rhodesain**[4] (enzyme: 1 nM, 50 mM sodium acetate pH 5.5, 5 mM EDTA, 200 mM NaCl, 5 mM DTT, 10 µM Z-Phe-Arg-AMC/10 µM Z-Phe-Arg-NBD); **SARS-CoV 2 M^{pro}**[6] (enzyme: 250 nM, 20 mM Tris pH 7.5, 0.1 mM EDTA, 200 mM NaCl, 1 mM DTT, 60 µM Boc-Abu-Tle-Leu-Gln-AMC/30 µM Boc-Abu-Tle-Leu-Gln-NBD); **DENV2 and ZIKV NS2B/NS3**[7] (enzyme: 250 nM (DENV) & 25 nM (ZIKV), 50 mM Tris pH 7.5, 100 µM Boc-Gly-Arg-Arg-AMC/50 µM Z-Gly-Arg-Arg-NBD); **Urokinase plasminogen activator (uPA)**[10,11] (enzyme: 10 U, 50 mM Tris pH 7.4, 50 mM NaCl, 0.5 mM EDTA, 240 µM Z-Gly-Gly-Arg-AMC/10 µM Z-Gly-Gly-Arg-NBD); **Sortase A (SrtA)**[3] (enzyme: 1.3 µM, 50 mM Tris pH 7.5, 150 mM NaCl, 5 mM $CaCl_2$, 0.5 mM Gly_4, 25 µM Abz-LPETG-Dap(dnp)–OH/100 µM Bz-Leu-Pro-Ala-Thr-Gly-NBD); **SARS-CoV 2 PL^{pro}**[12] (enzyme: 100 nM, 20 mM Tris pH 7.5, 0.1 mM EDTA, 200 mM NaCl, 1 mM DTT, 50 µM Z-Arg-Leu-Arg-Gly-Gly-AMC/40 µM Z-Arg-Leu-Arg-Gly-Gly-NBD); **Cathepsin S (CatS)**[13] (enzyme: 10 nM, enzyme buffer: 35 mM K_3PO_4 pH 6.5, 35 mM sodium acetate, 2 mM DTT, 2 mM EDTA, assay buffer: 50 mM K_3PO_4 pH 6.5, 2.5 mM DTT, 2.5 mM EDTA, 10 µM Z-Val-Val-Arg-AMC/10 µM Z-Val-Val-Arg-NBD); **Matrixmetalloprotease 9 (MMP9)**[14] (enzyme: 2.5 nM, 150 mM NaCl, 50 mM Tris-HCl pH 7.5, 20 µM $ZnCl_2$, 1 mM $CaCl_2$, Dnp-Pro-Leu-Gly-Met-Trp-Ser-Arg-NH_2/40 µM Z-Pro-Leu-Gly-Met-NBD); **Thermolysin**[15,16] (enzyme: 0.7 nM, 100 mM Tris, 100 mM NaBr, 2.5 mM $CaCl_2$, 100 µM FAGLA/20 µM Bz-Gly-Phe-NBD).

Molecular Docking

A docking protocol (Glide) was conducted within the Schrödinger Maestro 2020.04 worksuite.[17] The protein structures MMP9 (pdb: 4JQG), uPA (pdb: 1LMW), SrtA (pdb: 2KID), M^{pro} (pdb: 6Z2E), PL^{pro} (pdb: 7RBS), NS2B/NS3 (pdb: 3U1I), and thermolysin (pdb: 1TMN) were downloaded from the Protein Databank (PDB). Receptor preparation was performed using the automated binding site, protonation, and energy minimization routine within Maestro “Protein preparation” and “Receptor grid generation.” Ligands were energy minimized using the “LigPrep” routine. The docking protocol was performed under default parameters with peptide docking (SP-peptide) settings. Subsequently, predicted binding modes were analyzed using PyMol 2.0.5.

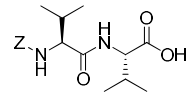
Solid-Phase Peptide Synthesis (SPPS)

A resin loading. Solid-phase peptide synthesis was conducted in a fritted 12 mL polypropylene syringe. 1 g of 2-Chlorotritylchloride resin (2-CTC resin, 1.2 mmol loading capacity) was pre-swelled in 8 mL DCM for 15 min and drained. The first amino acid (3.6 mmol, 3 eq. relative to resin loading capacity) was added in 1.8M *N*-Methylmorpholine (NMM)/DCM (6 mL) and the mixture was swiveled for 12 h. After draining the solution, the resin was washed with DMF (3 × 6 mL for 1 min) and DCM (3 × 6 mL for 1 min).
B resin capping. The remaining free 2-CTC resin linkers were capped with methanol in a solution of 9 mL of DCM/MeOH/DIPEA (9:2:1). The resin was swiveled for 1 h, drained, and washed with DCM (3 × 6 mL).
C Fmoc deprotection. The resin was treated with 20% piperidine in DMF solution (2 × 6 mL for 10 min) and subsequently washed with DMF (3 × 6 mL).
D peptide coupling with HATU. A coupling cocktail was prepared including the specific Fmoc-protected amino acid (3.6 mmol, 3 eq.), HATU (3.6 mmol, 3 eq.), and DIEA (10.8 mmol,9 eq.) in DMF (4.8 mL). The solution was added to the resin and swiveled for 3 h with exception of Fmoc-Arg(Pbf)-OH which was coupled overnight. After the reaction, the resin was drained and washed with DMF (3 × 6 mL for 1 min) and DCM (3 × 6 mL for 1 min).

E resin cleavage with preservation of side chain protecting groups

After the last coupling step, the resin was washed with DMF (3 × 6 mL for 1 min) and DCM (3 × 6 mL for 1 min). Then a cleavage cocktail containing AcOH/Trifluoroethanol/DCM (1:1:4, 6 mL) was added and swiveled for 1.5 h. The resin was drained and washed with DCM (5 mL). The combined eluates were concentrated under reduced pressure to yield the crude linear peptide. The purification of the crude product was performed by flash chromatography on a reversed-phase silica column (POLYPREP 60-50 C_{18}).

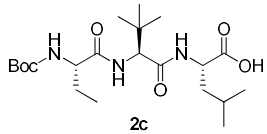
Z-Val-Val-OH 2b



2b

The title compound was prepared according to general procedures A–E on a 1.2 mmol scale. The product was purified by C18 flash chromatography and obtained as a colorless solid after lyophilization. 340 mg, 0.97 mmol, yield: 81%, MS (ESI) m/z $[M + H]^+$ calcd. for $C_{18}H_{26}N_2O_5$ 351.18 (100%), found 351.15 $[M + H]^+$.

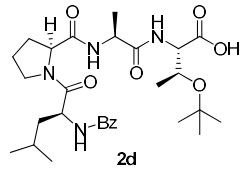
Boc-Abu-Tle-Leu-OH 2c



2c

The title compound was prepared according to general procedures A–E on a 1.2 mmol scale. The product was purified by C18 flash chromatography and obtained as a colorless solid after lyophilization. 334 mg, 0.78 mmol, yield: 65%, MS (ESI) m/z $[M + H]^+$ calcd. for $C_{21}H_{39}N_3O_6$ 429.28 (100%), found 452.24 $[M + Na]^+$.

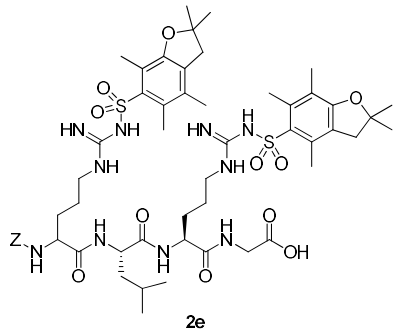
Bz-Leu-Pro-Ala-Thr(*t*Bu)-OH 2d



2d

The title compound was prepared according to general procedures A–E on a 1.2 mmol scale. The product was purified by C18 flash chromatography and obtained as a colorless solid after lyophilization. 269 mg, 0.48 mmol, yield: 40%, MS (ESI) m/z $[M + H]^+$ calcd. for $C_{29}H_{44}N_4O_7$ 560.68 (100%), found 583.25 $[M + Na]^+$.

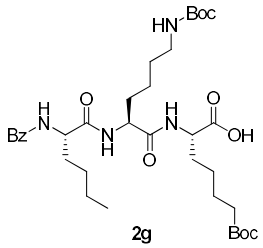
Z-Arg (Pbf)-Leu-Arg(Pbf)-Gly-OH 2e



2e

The title compound was prepared according to general procedures A–E on a 1.2 mmol scale. The product was purified by C18 flash chromatography and obtained as a colorless solid after lyophilization. 274 mg, 0.26 mmol, yield: 22%, MS (ESI) m/z $[M + H]^+$ calcd. for $C_{54}H_{78}N_{10}O_{13}S_2$ 1139.52 (100%), found 570.19 $[(M + 2H)^{2+}/2]$.

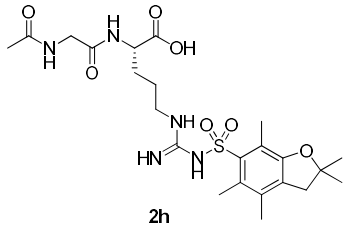
Bz-Nle-Lys(Boc)-Lys(Boc)-OH 2g



2g

The title compound was prepared according to general procedures A–E on a 1.2 mmol scale. The product was purified by C18 flash chromatography and obtained as a colorless solid after lyophilization. 442 mg, 0.64 mmol, yield: 53%, MS (ESI) m/z $[M + H]^+$ calcd. for $C_{35}H_{57}N_5O_9$ 692.42.52 (100%), found 592.30 $[M - Boc + H]^+$.

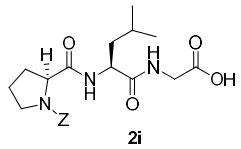
Ac-Gly-Arg(Pbf)-OH 2h



2h

The title compound was prepared according to general procedures A–E on a 1.2 mmol scale. The product was purified by C18 flash chromatography and obtained as a colorless solid after lyophilization. 562 mg, 1.07 mmol, yield: 89%, MS (ESI) m/z $[M + H]^+$ calcd. for $C_{23}H_{35}N_5O_7S$ 526.23 (100%), found 526.12 $[M + H]^+$.

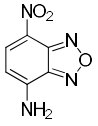
Z-Pro-Leu-Gly-OH 2i



2i

The commercially available Z-Pro-Leu-Gly-OEt (240 mg, 0.54 mmol, 1 eq.) was dissolved in methanol (5 mL). Subsequently, NaOH (1M, 2 mL) was added, and the reaction was stirred for 3 h at rt. The solvent was evaporated under reduced pressure at 40 °C and water (5 mL) was added. The reaction mixture was washed three times with DCM (3 × 15 mL). The aqueous phase was acidified with conc. HCl to pH 1 and extracted with DCM (3 × 15 mL). The organic phase was washed with brine (20 mL), dried over anhydrous Na_2SO_4, filtered, and evaporated under reduced pressure to obtain the desired product as a colorless solid 150 mg, 0.36 mmol, yield: 67%, MS (ESI) m/z $[M + H]^+$ calcd. for $C_{21}H_{30}N_3O_6$ 420.47 (100%), found 420.39 $[M + H]^+$.

4-amino-7-nitrobenzofurazane 4[18]



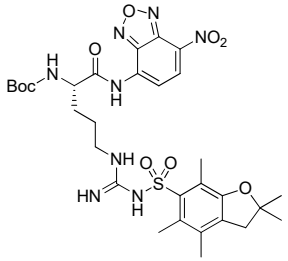
4

4-chloro-7-nitrobenzofurazane **3** (2 g, 10.05 mmol, 1 eq.) was dissolved in methanol (50 mL), and ammonia solution in methanol (7M, 50 mL) was added dropwise at rt. The solution was stirred for 16 h at room temperature and then evaporated under reduced pressure. The crude product was purified by silica column chromatography (equilibrate with Cy:EA = 1:1, elute with EA:MeOH = 10:1) to obtain the titled compound as a brown solid, 1.1 g, 6.11 mmol, yield: 61%, mp: 270–272 °C, ^{1}H NMR: (300 MHz, $CDCl_3$) δ = 8.85 (s, 2H), 8.48 (d, *J* = 8.8 Hz, 1H), 6.38 (d, *J* = 8.8 Hz, 1H) ppm, ^{13}C NMR: (75 MHz, $CDCl_3$) δ = 147.3, 144.3, 144.0, 137.9, 120.4, 102.6 ppm, FT-IR: $\tilde{\nu}$/cm^{-1} = 3356, 1656, 1540, 1425, 1299, 1133, 981, 905, 841, 734. MS (ESI) m/z $[M + H]^+$ calcd. for $C_6H_4N_4O_3$ 181.03 (100%), found 180.89 $[M + H]^+$.

General procedure for the amide bond formation mediated by HATU

The respective carboxylic acid **5**, **2a-j** (1.0 eq.) was dissolved in DCM and cooled to 0 °C with an ice-water bath. DIEA (3 eq.) and HATU (1.2 eq.) were added and stirred for an additional 30 min at 0 °C. Then, the corresponding amine **4**, **6a–d** (after Boc-deprotection) (1 eq.) in DMF was added and stirred for 30 min at 0 °C and 16 h at rt. DCM and water were added and the aqueous phase was extracted three times with DCM. The combined organic phases were dried with Na_2SO_4, filtered, and the solvent was evaporated under reduced pressure.

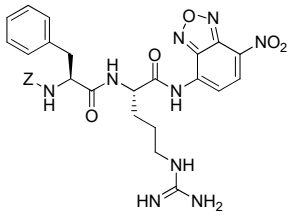
Boc-Arg(Pbf)-NBD 6a



6a

The crude product was purified by silica column chromatography (EA:CH = 3:1), 440 mg, 0.64 mmol, yield: 60%, $[\alpha]_D^{20}$ = −12 (10 mg/mL, $CHCl_3$), ^{1}H NMR: (300 MHz, $CDCl_3$) δ = 8.48 (d, *J* = 8.4 Hz, 1H), 8.34 (d, *J* = 8.4 Hz, 1H), 6.49 (s, 3H), 5.89 (s, 1H), 4.46 (s, 1H), 3.30 (s, 2H), 2.91 (s, 2H), 2.48 (s, 3H), 2.42 (s, 3H), 2.02 (s, 3H), 1.77 (d, *J* = 28.2 Hz, 2H), 1.44 (s, 6H), 1.41 (s, 9H) ppm, ^{13}C NMR: (75 MHz, $CDCl_3$) δ = 172.9, 145.6, 143.5, 134.7, 134.1, 132.9, 131.4, 125.5, 118.3, 114.1, 87.3, 81.8, 55.8, 43.6, 41.3, 29.1, 28.8, 26.0, 19.8, 18.4, 12.9 ppm, FT-IR: $\tilde{\nu}$/cm^{-1} = 3333, 3019, 1703, 1560, 1437, 1321, 1220, 1094, 754, 664, MS (ESI) m/z $[M + H]^+$ calcd. for $C_{30}H_{40}N_8O_9S$ 689.26 (100%), found 689.82 $[M + H]^+$; purity (HPLC, 210 nm): 97.5%.
Subsequently, selective Boc-deprotection was performed in TFA/DCM (16%, 5 mL) for 30 min at rt. The solvent was removed under reduced pressure.

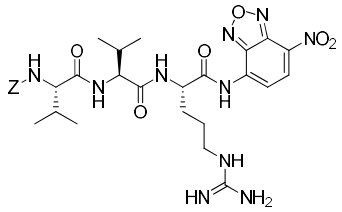
Z-Phe-Arg-NBD 7a



7a

Z-Phe-OH **2a** was commercially obtained. The crude coupling product was used without further purification for the following deprotection of the Pbf group. It was dissolved in TFA/DCM (50%, 2 mL) and stirred for 2 h at rt. The solution was evaporated under reduced pressure and purified by HPLC (ACN/H_2O/TFA (0.1%), gradient), 28.6 mg, 0.046 mmol, yield: 56%, ^{1}H NMR (600 MHz, DMSO-d_6) δ = 8.78 (d, *J* = 8.5 Hz, 1H), 8.57 (d, 1H), 8.38 (d, *J* = 8.5 Hz, 1H), 7.63 – 7.56 (m, 2H), 7.36 – 7.28 (m, 5H), 7.26 – 7.21 (m, 4H), 7.19 – 7.13 (m, 2H), 4.93 (s, 1H), 4.85 – 4.80 (m, 1H), 4.37 – 4.32 (m, 1H), 3.19 – 3.13 (m, 1H), 3.05 – 2.99 (m, 1H), 2.77 – 2.70 (m, 1H), 1.87 – 1.81 (m, 1H), 1.72 – 1.52 (m, 3H) ppm, ^{13}C NMR: (600 MHz, DMSO-d_6) δ = 173.1, 172.3, 158.3, 158.1, 156.6, 155.9, 145.5, 143.4, 137.9, 137.0, 135.8, 134.2, 130.2, 129.2, 128.3, 128.1, 127.7, 127.5, 126.3, 114.1, 65.2, 55.8, 53.2, 40.4, 40.0, 37.4, 28.6, 25.2 ppm, FT-IR: $\tilde{\nu}$/cm^{-1} = 2355, 1664, 1524, 1442, 1330, 1203, 1133, 998, 799, 689. MS (ESI) m/z $[M + H]^+$ calcd. for $C_{29}H_{32}N_9O_7$ 618.23 (100%), found 618.2 $[M + H]^+$; purity (HPLC, 210 nm): 98.8%.

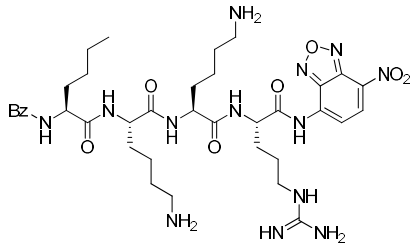
Z-Val-Val-Arg-NBD 7b



7b

The crude coupling product was used without further purification for the following deprotection of the Pbf-protecting group. It was dissolved in TFA/DCM (50%, 2 mL) and stirred for 2 h at rt. The solution was evaporated under reduced pressure and purified by HPLC (ACN/H_2O/TFA (0.1%), gradient), 15.64 mg, 0.023 mmol, yield: 34%, ^{1}H NMR (600 MHz, DMSO-d_6) δ = 8.78 – 8.73 (m, 1H), 8.47 – 8.39 (m, 1H), 8.37 – 8.34 (m, 1H), 8.02 – 7.76 (m, 1H), 7.60 – 7.55 (m, 1H), 7.42 – 7.28 (m, 7H), 5.05 – 4.97 (m, 2H), 4.86 – 4.74 (m, 1H), 4.24 (dd, *J* = 8.6, 7.1 Hz, 1H), 4.03 – 3.90 (m, 1H), 3.17 – 3.07 (m, 2H), 1.99 – 1.90 (m, 1H), 1.84 – 1.46 (m, 5H), 0.89 – 0.81 (m, 12H) ppm, ^{13}C NMR: (600 MHz, DMSO-d_6) δ = 173.1, 171.4, 171.2, 158.3, 156.6, 156.1, 145.5, 143.4, 137.0, 135.9, 134.2, 130.1, 128.3, 127.6, 113.9, 65.4, 60.2, 57.4, 57.2, 53.2, 52.9, 40.3, 30.7, 30.3, 28.8, 28.3, 25.2, 19.2, 18.2, 18.0, 17.8 ppm, FT-IR: $\tilde{\nu}$/cm^{-1} = 3291, 2968, 2336, 1653, 1538, 1451, 1347, 1304, 1178, 992. MS (ESI) m/z [M + H]$^+$ calcd. for $C_{30}H_{41}N_{10}O_8$ 669.30 (100%), found 669.07 [M + H]$^+$; purity (HPLC, 210 nm): 96.9%.

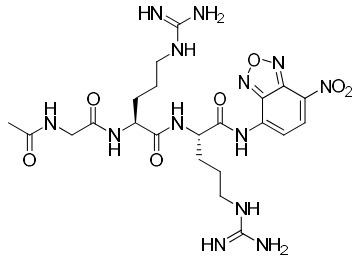
Bz-Nle-Lys-Lys-Arg-NBD 7g



7g

The crude coupling product was used without further purification for the following deprotection of the Boc and Pbf-protecting groups. It was dissolved in TFA/DCM (50%, 5 mL) and stirred for 2 h at rt. The solution was evaporated under reduced pressure and purified by HPLC (ACN/H_2O, gradient), 12.5 mg, 0.016 mmol, yield: 5%, ^{1}H NMR (600 MHz, DMSO-d_6) δ = 8.76 (d, *J* = 8.4 Hz, 1H), 8.50 (d, *J* = 7.4 Hz, 1H), 8.42 (d, *J* = 7.4 Hz, 1H), 8.35 (d, *J* = 8.4 Hz, 1H), 8.11 (d, *J* = 7.8 Hz, 1H), 8.02 (d, *J* = 7.8 Hz, 1H), 7.88 – 7.85 (m, 2H), 7.79 – 7.65 (m, 8H), 7.53 (s, 1H), 7.48 – 7.43 (m, 2H), 4.83 – 4.79 (m, 1H), 4.47 – 4.43 (m, 1H), 4.36 – 4.31 (m, 1H), 4.26 – 4.21 (m, 1H), 3.11 (dp, *J* = 20.2, 6.9 Hz, 2H), 2.73 (tq, *J* = 12.7, 6.1 Hz, 5H), 1.85 – 1.79 (m, 1H), 1.72 – 1.61 (m, 6H), 1.60 – 1.41 (m, 10H), 1.36 – 1.21 (m, 6H), 0.88 (d, *J* = 6.3 Hz, 3H), 0.85 (d, *J* = 6.4 Hz, 3H) ppm, ^{13}C NMR: (600 MHz, DMSO-d_6) δ = 173.0, 172.5, 171.8, 166.7, 158.3, 158.1, 156.7, 145.5, 143.4, 135.7, 134.1, 134.0, 131.4, 130.2, 128.2, 127.5, 118.2, 116.2, 114.1, 52.9, 52.6, 52.1, 51.9, 40.3, 38.6, 31.8, 31.1, 28.7, 26.6, 26.5, 25.1, 24.4, 23.1, 22.3, 22.2, 21.4 ppm, FT-IR: $\tilde{\nu}$/cm^{-1} = 2948, 2367, 1636, 1535, 1439, 1327, 1200, 1122, 838, 717. MS (ESI) m/z [M + H]$^+$ calcd. for $C_{37}H_{56}N_{13}O_8$ 810.43 (100%), found 405.66 [(M + 2H)$^{2+}$/2]; purity (HPLC, 210 nm): 99.3%.

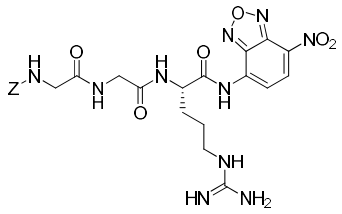
Ac-Gly-Arg-Arg-NBD 7h



7h

The general coupling protocol was used with modifications. PyBOP (1.2 eq.) HOBt (1.2 eq.), and DIEA (3 eq.) were used as coupling reagents. The crude coupling product was used without further purification for the following deprotection of the Pbf-protecting groups. It was dissolved in TFA/DCM (50%, 4 mL) and stirred for 3 h at rt. The solution was evaporated under reduced pressure and purified by HPLC (ACN/H_2O, gradient), 10.1 mg, 0.017 mmol, yield: 8%, ^{1}H NMR (600 MHz, DMSO-d_6) δ = 8.77 (d, *J* = 8.4 Hz, 1H), 8.45 (d, *J* = 6.9 Hz, 1H), 8.34 (d, *J* = 8.4 Hz, 1H), 8.16 (t, *J* = 5.8 Hz, 1H), 8.06 (d, *J* = 8.0 Hz, 1H), 7.63 (t, *J* = 5.8 Hz, 1H), 7.56 (t, *J* = 5.8 Hz, 1H), 4.79 – 4.74 (m, 1H), 4.37 (m, 1H), 3.70 (d, *J* = 5.8 Hz, 2H), 3.17 – 3.06 (m, 4H), 1.85 (s, 4H), 1.71 – 1.44 (m, 8H) ppm, ^{13}C NMR: (600 MHz, DMSO-d_6) δ = 173.1, 171.8, 169.9, 169.0, 158.4, 158.2, 145.5, 143.4, 135.8, 134.2, 130.2, 118.2, 116.2, 113.9, 53.2, 51.7, 42.0, 29.3, 28.5, 25.2, 24.9, 22.4 ppm, FT-IR: *v*/cm^{-1} = 3173, 2367, 1667, 1535, 1456, 1313, 1265, 1203, 1116, 950. MS (ESI) m/z $[M + H]^+$ calcd. for $C_{22}H_{34}N_{13}O_7$ 592.26 (100%), found 296.56 $[(M + 2H)^{2+}/2]$; purity (HPLC, 396 nm): 97.1%.

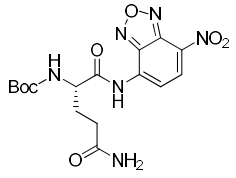
Z-Gly-Gly-Arg-NBD 7f



7f

Z-Gly-Gly-OH **2f** was commercially obtained. The crude coupling product was used without further purification for the following deprotection of the Pbf-protecting group. It was dissolved in TFA/DCM (50%, 6 mL) and stirred for 2 h at rt. The solution was evaporated under reduced pressure and purified by HPLC (ACN/H_2O, gradient), 15.6 mg, 0.016 mmol, yield: 39%, ^{1}H NMR (600 MHz, DMSO-d_6) δ = 8.77 (d, *J* = 8.4 Hz, 1H), 8.36 (dd, *J* = 8.4 Hz, 2H), 8.15 (m, 1H), 7.59 (m, 1H), 7.54 (m, 1H), 7.36 – 7.28 (m, 6H), 5.02 (s, 2H), 4.87 – 4.82 (m, 1H), 3.86 – 3.73 (m, 2H), 3.64 (d, *J* = 6.1 Hz, 2H), 3.13 (p, *J* = 6.5 Hz, 2H), 1.84 – 1.77 (m, 1H), 1.70 – 1.49 (m, 3H) ppm, ^{13}C NMR: (600 MHz, DMSO-d_6) δ = 173.2, 169.5, 169.3, 156.7, 156.6, 145.5, 143.4, 136.9, 135.8, 134.2, 130.2, 128.4, 127.8, 127.7, 114.1, 65.5, 53.0, 43.5, 41.6, 40.3, 28.8, 25.1 ppm, FT-IR: *v*/cm^{-1} = 3185, 2358, 1661, 1515, 1434, 1333, 1203, 1125, 998, 717, MS (ESI) m/z $[M + H]^+$ calcd. for $C_{24}H_{29}N_{10}O_8$ 585.21 (100%), found 585.23.66 $[M + H]^+$; purity (HPLC, 396 nm): 100%.

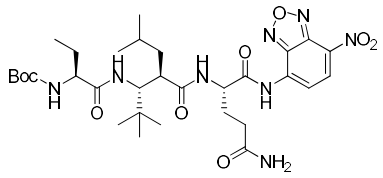
Boc-Gln-NBD 6b



6b

The crude coupling product was purified by silica column chromatography (EA:CH = 6:1), 62 mg, 0.152 mmol, yield: 57%, ^{1}H NMR: (300 MHz, $CDCl_3$) δ = 8.76 (d, *J* = 8.5 Hz, 1H), 8.37 (d, *J* = 8.5 Hz, 1H), 7.36 – 7.25 (m, 2H), 6.85 – 6.78 (m, 1H), 4.40 (s, 1H), 2.28 – 1.70 (m, 4H), 1.39 (s, 9H) ppm, ^{13}C NMR: (75 MHz, $CDCl_3$) δ = 173.6, 155.7, 145.5, 143.3, 135.8, 134.4, 130.0, 113.8, 78.4, 54.9, 31.6, 28.1, 26.6 ppm, FT-IR: *v*/cm^{-1} = 3350, 2918, 1647, 1518, 1442, 1324, 1164, 1088, 998, 863, MS (ESI) m/z $[M + H]^+$ calcd. for $C_{36}H_{21}N_6O_7$ 408.14 (100%), found 431.19 $[M + Na]^+$. Subsequently, Boc-deprotection was performed in TFA/DCM(16%, 2 mL) for 30 min at rt. The solvent was removed under reduced pressure.

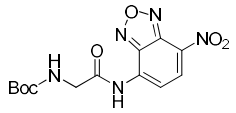
Boc-Abu-Tle-Leu-Gln-NBD 7c



7c

The crude coupling product was purified by HPLC (ACN/H_2O/TFA (0.1%), gradient), 10.2 mg, 0.014 mmol, yield: 10%, ^{1}H NMR (600 MHz, DMSO-d_6) δ =. 8.76 (d, *J* = 8.4 Hz, 1H), 8.37 – 8.33 (m, 2H), 8.07 (d, *J* = 8.4 Hz, 1H), 7.46 (d, *J* = 9.5 Hz, 1H), 7.32 (d, *J* = 2.5 Hz, 1H), 7.07 (d, *J* = 8.2 Hz, 1H), 6.86 (d, *J* = 2.4 Hz, 1H), 4.69 – 4.65 (m, 1H), 4.41 – 4.36 (m, 1H), 4.27 (d, *J* = 9.5 Hz, 1H), 3.88 – 3.83 (m, 1H), 2.29 – 2.14 (m, 2H), 2.02 – 1.96 (m, 1H), 1.92 – 1.85 (m, 1H), 1.62 – 1.41 (m, 5H), 1.37 (s, 8H), 1.33 (s, 1H) ppm, ^{13}C NMR: (600 MHz, DMSO-d_6) δ = 173.4, 172.9, 172.3, 171.7, 169.7, 155.3, 145.5, 143.4, 135.9, 134.4, 130.0, 113.8, 78.0, 59.1, 56.0, 53.3, 50.5, 40.7, 34.7, 31.4, 28.2, 28.0, 27.0, 26.6, 24.9, 24.1, 23.0, 21.4 ppm, FT-IR: *v*/cm^{-1} = 2965, 2353, 1673, 1529, 1439, 1338, 1206, 1127, 995, 796, MS (ESI) m/z [M + H]$^+$ calcd. for $C_{32}H_{50}N_9O_{10}$ 720.36 (100%), found 620.23 [M - Boc + H]$^+$; purity (HPLC, 210 nm): 99.9%.

Boc-Gly-NBD 6c

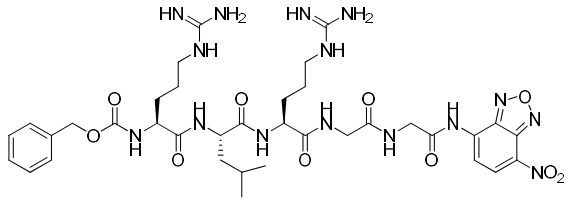


6c

The crude coupling product was purified by silica column chromatography (EA:CH = 1:3), 140 mg, 0.60 mmol, yield: 64%, ^{1}H NMR: (300 MHz, $CDCl_3$) δ = 8.75 (d, *J* = 8.5 Hz, 1H), 8.37 (d, *J* = 8.5 Hz, 1H), 7.21 (t, *J* = 6.1 Hz, 1H), 3.99 (d, *J* = 6.1 Hz, 2H), 1.40 (s, 9H) ppm, ^{13}C NMR: (75 MHz, $CDCl_3$) δ = 170.6, 156.0, 145.4, 143.3, 135.9, 134.3, 129.8, 113.5, 78.2, 44.2, 28.1 ppm, FT-IR: *v*/cm^{-1} = 2985, 1673, 1518, 1442, 1327, 1153, 1085, 995, 855, 796, MS (ESI) m/z [M + H]$^+$ calcd. for $C_{30}H_{40}N_8O_9S$ 338.10 (100%), found 237.95 [M – Boc + H]$^+$.

Subsequently, Boc-deprotection was performed in TFA/DCM(16%, 2 mL) for 30 min at rt. The solvent was removed under reduced pressure.

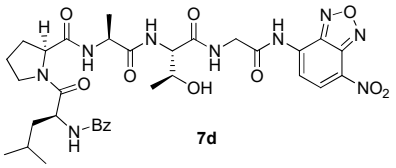
Z-Arg-Leu-Arg-Gly-Gly-NBD 7e



7e

The crude coupling product was used without further purification for the following deprotection of the Pbf-protecting groups. It was dissolved in TFA/DCM (50%, 7 mL) and stirred for 2 h at rt. The solution was evaporated under reduced pressure and purified by HPLC (ACN/H_2O/TFA (0.1%), gradient). Afterward, the product was dissolved in 0.5 mL TFA and precipitated in diethyl ether as the Di-TFA salt, 9.23 mg, 0.011 mmol, yield: 5%, ^{1}H NMR (600 MHz, DMSO-d_6) δ = 8.76 (d, *J* = 8.5 Hz, 1H), 8.38 – 8.35 (m, 1H), 8.34 (d, *J* = 8.5Hz, 1H), 8.24 – 8.22 (m, 1H), 8.10 (d, *J* = 7.5 Hz, 1H), 7.93 (d, *J* = 8.0 Hz, 1H), 7.57 – 7.52 (m, 2H), 7.51 (d, *J* = 8.0 Hz, 1H), 7.36 – 7.30 (m, 6H), 5.02 (s, 2H), 4.31 (td, *J* = 8.0, 5.6 Hz, 1H), 4.27 – 4.22 (m, 1H), 4.21 – 4.14 (m, 2H), 4.01 – 3.97 (m, 1H), 3.81 (d, *J* = 5.7 Hz, 2H), 3.09 – 3.04 (m, 4H), 1.72 – 1.44 (m, 12H), 0.88 – 0.81 (m, 6H) ppm, ^{13}C NMR: (600 MHz, DMSO-d_6) δ = 172.2, 171.6, 170.2, 169.3, 158.4, 158.2, 156.6, 156.0, 145.5, 143.4, 136.9, 135.9, 134.3, 130.0, 128.4, 127.9, 127.7, 113.6, 65.5, 54.3, 52.3, 50.9, 43.0, 41.7, 40.7, 40.4, 29.0, 25.1, 24.9, 24.1, 23.2, 21.4 ppm, FT-IR: *v*/cm^{-1} = 3190, 1653, 1524, 1445, 1324, 1214, 1119, 998, 804, 717, MS (ESI) m/z [M + H]$^+$ calcd. for $C_{35}H_{50}N_{15}O_{10}$ 854.39 (100%), found 427.66 [(M + 2H)$^{2+}$/2]; purity (HPLC, 210 nm): 95.2%.

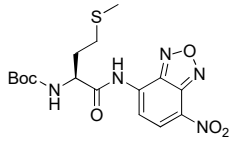
Bz-Leu-Pro-Ala-Thr-Gly-NBD 7d



7d

The crude coupling product was used without further purification for the following deprotection of the O^{t}Bu-protecting group. It was dissolved in TFA/DCM (50%, 5 mL) and stirred for 30 min at rt. The solution was evaporated under reduced pressure and purified by HPLC (ACN/H_2O, gradient), 5.1 mg, 0.01 mmol, yield: 7%, ^{1}H NMR (600 MHz, DMSO-d_6) δ = 8.76 (d, *J* = 8.5 Hz, 1H), 8.51 (d, *J* = 7.7 Hz, 1H), 8.41 – 8.31 (m, 2H), 8.20 – 8.10 (m, 1H), 7.95 – 7.84 (m, 3H), 7.58 – 7.40 (m, 4H), 4.75 – 4.68 (m, 1H), 4.40 – 4.11 (m, 8H), 3.75 – 3.70 (m, 1H), 3.58 – 3.54 (m, 1H), 2.07 – 1.64 (m, 5H), 1.50 – 1.42 (m, 1H), 1.27 – 1.19 (m, 4H), 1.12 – 1.06 (m, 3H), 0.94 – 0.88 (m, 3H), 0.87 – 0.83 (m, 3H) ppm, ^{13}C NMR: (600 MHz, DMSO-d_6) δ = 172.5, 171.61, 170.7, 170.0, 166.3, 145.4, 143.3, 135.9, 134.2, 133.8, 131.3, 130.0, 128.2, 127.5, 113.7, 66.9, 59.1, 58.2, 49.6, 48.4, 46.7, 43.2, 39.5, 28.9, 24.5, 24.3, 23.1, 21.2, 19.7, 19.1, 17.7, 17.5 ppm, FT-IR: *v*/cm^{-1} = 3288, 1633, 1540, 1434, 1330, 1209, 1127, 992, 807, 720, MS (ESI) m/z [M + H]$^+$ calcd. for $C_{33}H_{42}N_9O_{10}$ 724.3 (100%), found 724.1.66 [M + H]$^+$; purity (HPLC, 210 nm): 98.7%.

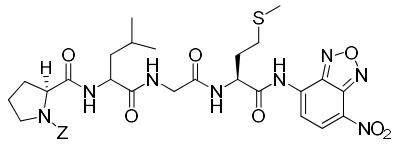
Boc-Met-NBD 6d



6d

The crude coupling product was purified by silica column chromatography (EA:CH = 1:3), 199 mg, 0.48 mmol, yield: 52%, ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 8.74 (d, *J* = 8.5 Hz, 1H), 8.35 (d, *J* = 8.5 Hz, 1H), 7.41 (d, *J* = 7.5 Hz, 1H), 4.54 – 4.42 (m, 1H), 2.65 – 2.51 (m, 2H), 2.06 (s, 3H), 2.03 – 1.82 (m, 2H), 1.39 (s, 9H) ppm, ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 173.5, 155.8, 145.5, 143.3, 135.8, 134.3, 130.0, 113.9, 78.5, 54.7, 31.1, 29.8, 28.1, 14.6 ppm, FT-IR: *v*/cm^{-1} =3331, 1701, 1574, 1518, 1316, 1164, 998, 863, 810, 731, MS (ESI) m/z [M + H]$^+$ calcd. for $C_{16}H_{22}N_5O_6S$ 412.43 (100%), found 433.99 [M + Na]$^+$.
Subsequently, Boc-deprotection was performed in HCl/Dioxan (4M, 4 mL) for 4.5 h at rt. The solvent was removed under reduced pressure.

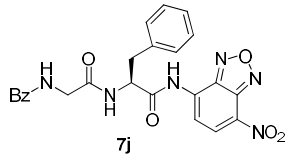
Z-Pro-Leu-Gly-Met-NBD 7i



7i

The crude coupling product was purified by RP-HPLC (ACN/H_2O/TFA(0.1%), gradient), 25 mg, 0.035 mmol, yield: 11%, ^{1}H NMR (600 MHz, DMSO-d_6) δ = 8.78 – 8.71 (m, 1H), 8.42 – 8.25 (m, 2H), 8.14 – 8.06 (m, 2H), 7.37 – 7.24 (m, 5H), 5.07 – 5.03 (m, 1H), 4.91 (d, *J* = 13.1 Hz, 1H), 4.86 – 4.79 (m, 1H), 4.30 – 4.19 (m, 2H), 3.82 – 3.69 (m, 2H), 3.49 – 3.38 (m, 2H), 2.61 – 2.52 (m, 2H), 2.17 – 2.01 (m, 5H), 1.96 – 1.57 (m, 5H), 1.53 – 1.46 (m, 1H), 1.41 – 1.37 (m, 1H), 0.80 (ddd, *J* = 61.8, 28.9, 6.6 Hz, 6H) ppm, ^{13}C NMR: (600 MHz, DMSO-d_6) δ = 172.3, 172.1, 169.2, 154.2, 153.8, 145.6, 143.4, 136.9, 135.8, 128.4, 128.2, 127.7, 127.5, 126.9, 114.1, 65.7, 59.8, 59.0, 53.1, 51.0, 47.1, 46.6, 41.6, 40.8,40.4, 31.5, 31.2, 29.6, 24.0, 22.9, 21.4, 14.6 ppm, FT-IR: *v*/cm^{-1} = 3288, 2948, 2344, 1656, 1504, 1437, 1327, 1200, 990, 734, MS (ESI) m/z [M + H]$^+$ calcd. for $C_{32}H_{41}N_8O_9S$ 713.77 (100%), found 713.06 [M + H]$^+$; purity (HPLC, 210 nm): 99.9%.

Bz-Gly-Phe-NBD 7j

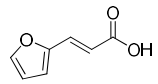


7j

Bz-Gly-Phe-OH **2j** was commercially obtained. The crude coupling product was purified by RP-HPLC (ACN/H_2O/TFA(0.1%), gradient), 8 mg, 0.016 mmol, yield: 14%, ^{1}H NMR (600 MHz, DMSO-d_6) δ = 8.77 (d, *J* = 8.4 Hz, 1H), 8.69 (t, *J* = 5.9 Hz, 1H), 8.48 (d, *J* = 7.6 Hz, 1H), 8.38 (d, *J* = 8.4 Hz, 1H), 7.85 – 7.81 (m, 2H), 7.54 – 7.51 (m, 1H), 7.48 – 7.44 (m, 2H), 7.38 – 7.34 (m, 2H), 7.32 – 7.22 (m, 3H), 7.21 – 7.17 (m, 1H), 5.15 – 5.08 (m, 1H), 3.93 (dd, *J* = 16.5, 5.9 Hz, 1H), 3.83 (dd, *J* = 16.5, 5.9 Hz, 1H), 3.13 (dd, *J* = 13.8, 4.4 Hz, 1H), 2.92 (dd, *J* = 13.8, 10.1 Hz, 1H) ppm, ^{13}C NMR: (600 MHz, DMSO-d_6) δ = 173.0, 169.5, 166.3, 145.5, 143.4, 137.1, 135.8, 134.3, 133.9, 131.4, 130.2, 129.3, 128.3, 128.1, 127.3, 126.5, 114.1, 54.8, 42.1, 37.0 ppm, FT-IR: *v*/cm^{-1} = 3286, 2370, 1720, 1667, 1532, 1442, 1335, 1200, 992, 706, MS (ESI) m/z [M + H]$^+$ calcd. for $C_{24}H_{21}N_6O_6$ 489.14 (100%), found 489.04 [M + H]$^+$; purity (HPLC, 210 nm): 99.9%.

Synthesis of (*E*)-2-(3-(furan-2-yl)acrylamido)acetic acid (FAGLA) 20

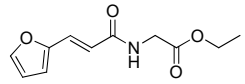
(*E*)-3-(furan-2-yl)acrylic acid **18**



18

According to Kudelko *et al*., furfural (3.65 g, 38.00 mmol, 1 eq.), malonic acid (9.36 g, 90 mmol, 2.4 eq.), and piperidine (0.3 mL) were dissolved in pyridine (18 mL). The reaction mixture was stirred for 90 min at 50 °C and refluxed for further 20 min. Subsequently, the solution was poured into cold water (20 mL) and acidified with conc. HCl to pH 4. The precipitate was filtered and washed with water to obtain a beige solid, 4.05 g, 29.34 mmol, yield: 77%. mp = 136–138 °C (Lit.: 138–141 °C), ^{1}H NMR (300 MHz, DMSO-d_6) δ = 7.85 – 7.77 (m, 1H), 7.39 (d, *J* = 15.8 Hz, 1H), 6.91 (d, *J* = 3.5 Hz, 1H), 6.61 (dd, *J* = 3.4, 1.8 Hz, 1H), 6.16 (d, *J* = 15.8 Hz, 1H) ppm, ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 167.4, 150.3, 145.7, 130.8, 116.0, 115.4, 112.7 ppm, FT-IR: *v*/cm^{-1} = 1681, 1628, 1484, 1307, 1259, 1223, 1026, 852, 762, 666.

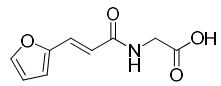
(*E*)-ethyl 2-(3-(furan-2-yl)acrylamido)acetate **19**



19

(*E*)-3-(furan-2-yl)acrylic acid **18** (1.50 g, 10.86 mmol, 1 eq.) was amide coupled with Glycine ethyl ester hydrochloride (1.51 g, 10.86 mmol, 1 eq.) in an analogous way to the "General procedure for the amide bond formation mediated by HATU". The crude product was purified by silica gel column chromatography (EA:CH = 1:2) to obtain the desired product as a colorless oil, 1.71 g, 7.66 mmol, yield: 71%. ^{1}H NMR (300 MHz, DMSO-d_6) δ = 7.45 – 7.37 (m, 2H), 6.53 (d, *J* = 3.3 Hz, 1H), 6.46 – 6.32 (m, 3H), 4.26 – 4.13 (m, 4H), 1.28 (t, *J* = 7.1 Hz, 3H) ppm, ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 170.1, 166.1, 151.2, 144.2, 128.7, 117.6, 114.2, 112.2, 61.6, 41.7, 14.2 ppm, FT-IR: *v*/cm^{-1} = 3356, 2982, 1743, 1664, 1521, 1372, 1206, 1009, 976, 751. MS (ESI) m/z [M + H]$^+$ calcd. for $C_{11}H_{13}NO_4$ 224.08 (100%), found 223.76 [M + H]$^+$.

(*E*)-2-(3-(furan-2-yl)acrylamido)acetic acid (FAGLA) **20**



20

(*E*)-ethyl 2-(3-(furan-2-yl)acrylamido)acetate **19** (0.04 g, 0.18 mmol, 1 eq.) was dissolved in dioxane (2 mL) and 0.35 mL of a NaOH-solution (1M) was added. The mixture was stirred for 4 h at rt and the solution was evaporated under reduced pressure. The crude product was purified by silica gel column chromatography (EA:CH = 1:1 +TFA(0.1%)) to obtain a colorless oil, 0.03 g, 0.14 mmol, yield: 80%. ^{1}H NMR (300 MHz, DMSO-d_6) δ = 8.48 (t, *J* = 5.9 Hz, 1H), 7.77 (d, *J* = 1.7 Hz, 1H), 7.26 (d, *J* = 15.6 Hz, 1H), 6.78 (d, *J* = 3.3 Hz, 1H), 6.61 – 6.44 (m, 2H), 3.88 (d, *J* = 5.9 Hz, 2H) ppm, ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 171.2, 165.1, 150.8, 144.8, 126.5, 118.8, 113.9, 112.3 ppm, FT-IR: *v*/cm^{-1} =3269, 2906, 1746, 1647, 1639, 1538, 1198, 1040, 762, 640. MS (ESI) m/z [M + H]$^+$ calcd. for $C_9H_{10}NO_4$ 196.05 (100%), found 195.9 [M + H]$^+$; purity (HPLC, 210 nm): 99.9%.

"

5 Chapter B: Evaluation of Covalent Inhibitors for Trypsin-Like Serine Proteases

5.1 Peptidomimetic inhibitors of TMPRSS2 block SARS-CoV-2 infection in cell culture

5.1.1 Context, Project Summary, and Own Contributions

Since the outbreak of SARS-CoV-2 in Wuhan (China), the virus rapidly spread out to a pandemic dimension and caused over 770 million infections with 7 million deaths worldwide so far (1.5).[233] Besides the viral proteases M^{pro} and PL^{pro}, the transmembrane serine protease subtype 2 (TMPRSS2) presents an additional target for treatment of the infection, due to the important role in the viral cell entry of SARS-CoV-2 (1.5.2).[211],[245],[249] Most of the published inhibitors yet, display an aromatic small molecule scaffold with a basic guanidine/amidine moiety, that protrudes into the S1 pocket (1.5.4).[212],[250],[252],[255] Noteworthy is nafamostat that is actually used as treatment for chronic pancreatitis in Japan and shows remarkable inhibition in subnanomolar range against TMPRSS2 (1.5.4, Figure 19).[252] In contrast, only few substrate analog inhibitors have been described yet.[255]

Within this study we used computational modeling and docking of combinatorial peptide libraries to discover potential suitable recognition sequences towards the binding site of TMPRSS2 and combined them with a α-ketobenzothiazole moiety as the warhead, which resulted in covalent reversible inhibitors. After fluorometric inhibition studies, we identified four compounds as potential hits with high affinity to TMPRSS2 and moderate selectivity to other trypsin-like serine proteases. Further cell-based measurements confirmed the inhibitory activity towards TMPRSS2 in cells of the four lead compounds and additionally displayed the ability of blocking the spike protein-driven cell entry and viral infection by different SARS-CoV-2 variants. Stability tests of two inhibitors in different body fluids for 10 days and subsequent observed inhibitory activity suggest remarkable metabolic resistance.

Summarized, the high potency of the selected inhibitors against TMPRSS2 and anti-SARS-CoV-2 activity, combined with the high stability in body fluids present promising lead structures for further development and may synergize with other protease inhibitors like nirmatrelvir (FDA approved M^{pro} inhibitor) for treatment against the SARS-CoV-2 infection.[313]

Own contributions: Inhibitor synthesis (S1−S5) and writing of the original draft plus editing of the manuscript.

Contributions from other authors: Inhibitor synthesis (1−8), inhibition-, cytotoxicity- and stability assay, molecular docking, pseudovirus inhibition assay, SARS-CoV-2 inhibition assay and writing of the original draft plus editing of the manuscript.

This work has been published in Communications Biology (impact factor: 5.90).

5.1.2 Publication

The following publication quoted within "" from page 150 to page 161 is the same as the manuscript cited on page 53.

"

communications biology

ARTICLE

https://doi.org/10.1038/s42003-022-03613-4 OPEN

Peptidomimetic inhibitors of TMPRSS2 block SARS-CoV-2 infection in cell culture

Lukas Wettstein[1,9], Philip Maximilian Knaff[2,3,9], Christian Kersten[4,9], Patrick Müller[4], Tatjana Weil[1], Carina Conzelmann[1], Janis A Müller[1,5], Maximilian Brückner[2,3], Markus Hoffmann[6,7], Stefan Pöhlmann[6], Tanja Schirmeister[4], Katharina Landfester[3], Jan Münch[1,8,10]✉ & Volker Mailänder[2,3,10]✉

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 inhibit TMPRSS2 while demonstrating good off-target selectivity against selected coagulation proteases. Lead candidates are stable in blood serum and plasma for at least ten days. Finally, we show that selected peptidomimetics inhibit SARS-CoV-2 Spike-driven pseudovirus entry and authentic SARS-CoV-2 infection with comparable efficacy as camostat mesylate. The peptidomimetic TMPRSS2 inhibitors also prevent entry of recent SARS-CoV-2 variants of concern Delta and Omicron BA.1. 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.

[1] Institute of Molecular Virology, Ulm University Medical Center, 89081 Ulm, Germany. [2] Dermatology Clinic of the University Medicine of the Johannes Gutenberg University Mainz, 55131 Mainz, Germany. [3] Max Planck Institute for Polymer Research, 55128 Mainz, Germany. [4] Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, 55128 Mainz, Germany. [5] Institute of Virology, Philipps University Marburg, Marburg, Germany. [6] Infection Biology Unit, German Primate Center, 37077 Göttingen, Germany. [7] Faculty of Biology and Psychology, Georg-August-University Göttingen, 37073 Göttingen, Germany. [8] Core Facility Functional Peptidomics, Ulm University Medical Center, 89081 Ulm, Germany. [9] These authors contributed equally: Lukas Wettstein, Philip Maximilian Knaff, Christian Kersten. [10] These authors jointly supervised this work: Jan Münch, Volker Mailänder. ✉email: jan.muench@uni-ulm.de; mailaend@mpip-mainz.mpg.de

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the causative agent of coronavirus disease 2019 (COVID-19). By the end of March 2022, the WHO reported more than 470 million confirmed SARS-CoV-2 infections worldwide, resulting in more than 6 million deaths since its first occurrence in late 2019[1]. 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 such as cardiovascular disease, diabetes, chronic respiratory disorders, and cancer are more likely to develop serious illness, characterized by respiratory failure, shock, and multiorgan dysfunction[2–5]. SARS-CoV-2 is primarily transmitted through aerosols and droplets of saliva. The 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[6–9]. Measures to prevent or mitigate SARS-CoV-2 spread include lockdown strategies, social distancing, quarantining, use of face masks, and hygiene concepts. The 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[10–14]. Currently, the antivirals nirmatrelvir (in combination with ritonavir for CYP-inhibition, Paxlovid®) and molnupiravir (Lagevrio®) are available for the treatment of COVID-19[15,16]. Nevertheless, it remains imperative to develop further potent therapeutic interventions for COVID-19 therapy.

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[17–19]. 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[19]. 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[19–25]. TMPRSS2 priming is essential for triggering fusion of these virions with target cells 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[26–28].

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[29]. The family of TTSP comprises a total of 17 members in four subfamilies (matriptase, corin, hepsin/TMPRSS, and HAT/DESC subfamily) with a common 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[30]. TMPRSS2 belongs to the hepsin/TMPRSS subfamily with a total of seven serine proteases including TMPRSS2–5, MSPL, hepsin, and enteropeptidase. The development of TMPRSS2 inhibitors is hampered by the fact that no crystal structure was available until early 2021. Up to date, only few substrate analog inhibitors of TMPRSS2 have been described[31,32]. 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 the treatment of chronic pancreatitis in Japan, as well as the endogenous protease inhibitor alpha-1 antitrypsin (α_1-AT), and the mucolytic cough suppressant bromhexine[19,33–35]. 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 development of peptidomimetic inhibitors for TMPRSS2 for the treatment of 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 Alpha and Beta. The tested peptidomimetics are stable in human plasma and serum for at least 10 days, suggesting that these TMPRSS2 inhibitors are promising leads for further development as antiviral drugs in COVID-19 therapy and other viral diseases.

Results

Structure-based design of TMPRSS2 inhibitors. For the identification of peptide-based TMPRSS2 inhibitors as a potential treatment of SARS-CoV-2 infection, molecular docking studies were performed. As no crystal structure of TMPRSS2 was available in the protein data bank (PDB) at the beginning of the study, we employed a matriptase crystal structure to build a surrogate model[36,37]. Matripase shares 41% sequence identity with TMPRSS2 and previously described substrate analog TMPRSS2 inhibitors showed no selectivity over matriptase[31]. Computational redocking of the crystallographic ligand and the matriptase surrogate yielded a ligand orientation comparable to that in the crystal structure (RMSD of 1.8 Å), thus validating the surrogate model (Supplementary Fig. 3a). In addition, a homology model of TMPRSS2 was built using hepsin (43% sequence identity) as a template. Both models were tested for selectivity by docking of known binders and non-binders. Subsequent receiver operating characteristic (ROC) analysis indicated reliable discrimination between known binders and decoys for both, the homology and the surrogate model (Supplementary Fig. 3b, c). Previous studies characterized 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 (Supplementary Table 1)[31]. 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 (Fig. 1a and Supplementary Table 3) and the TMPRSS2 homology model (Fig. 1b). Our dockings show that the reference binder ace-D-Arg-Pro-Arg-aldehyde binds to the reactive center of the surrogate and the homology model. To optimize the binding affinity of the reference binder, we successively altered the residues at P1-P3 position using proteinogenic and non-proteinogenic amino acids and docked the resulting structures to both, the matriptase

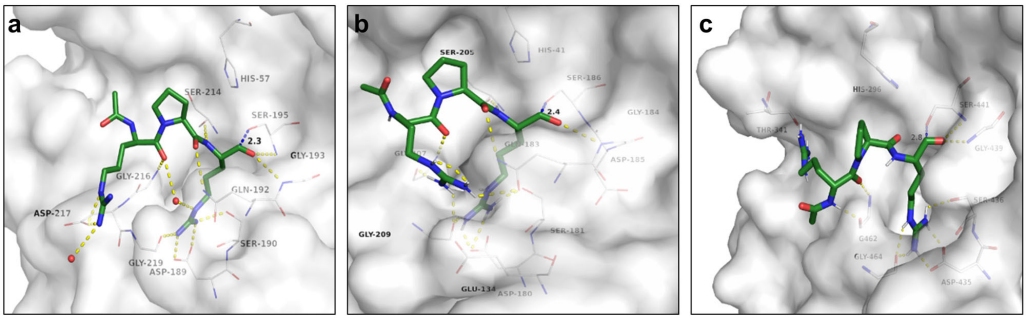


Fig. 1 Predicted binding of reference binder. a Docking of ace-D-Arg-Pro-Arg-aldehyde reference binder to matriptase surrogate model (white carbon atoms and surface). For a clear view, only residues forming polar interactions (yellow dashed lines), and the catalytic residues Ser-195 and His-57 are depicted. **b** Docking of ace-D-Arg-Pro-Arg-aldehyde reference binder to hepsin-based TMPRSS2 homology model (white carbon atoms and surface). For a clear view, only residues forming polar interactions (yellow dashed lines) and the catalytic residues Ser-186 and His-41 are depicted. **c** Docking of ace-D-Arg-Pro-Arg-aldehyde reference binder to TMPRSS2 crystal structure (PDB-ID: 7MEQ, white carbon atoms and surface). For a clear view, only residues forming polar interactions (yellow dashed lines), and the catalytic residues Ser-441 and His-296 are depicted. For all panels, carbon atoms of docked ligands are shown in green, oxygen in red, and nitrogen in blue. The distance between the nucleophilic serine oxygen and the electrophilic carbon atom of the serine trap in angstrom is illustrated by a dashed blue line.

Table 1 Assembled peptidomimetic inhibitor library selected for synthesis.

Compound	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

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.

surrogate and TMPRSS2 homology model. Compounds were then ranked based on their binding score (Supplementary Tables 2–4), the plausibility of their binding mode to the S1-S3 sub-pockets of the TMPRSS2 active site, and the proximity of the aldehyde serine trap to the catalytic Ser-186/195 (TMPRSS2 homology model/matriptase enumeration), as well as commercial availability of their building blocks. Overall, the D-configuration for P3 residue was favored to improve metabolic stability. The most promising compounds were chosen for solid-phase synthesis, whereby the aldehyde serine trap used for in silico modeling was exchanged by a well-established ketobenzothiazole, yielding a library of peptidomimetic inhibitors (Table 1 and Supplementary Fig. 4).

We retrospectively validated our molecular docking results by employing a TMPRSS2 crystal structure (PDB-ID: 7MEQ)[38] that became available during the course of this study. Superposition with the TMPRSS2 structure confirmed the accuracy of the homology- and the matriptase surrogate model, with overall C_α-RMSD values of 0.6 Å for both models (Supplementary Fig. 5). The TMPRSS2 structure allowed successful redocking of its crystallographic ligand and enabled discrimination of binders and decoys (Supplementary Fig. 3d, e). Accordingly, the reference binder ace-D-Arg-Pro-Arg-aldehyde accommodated the substrate-binding pocket of the TMPRSS2 crystal structure, with the electrophilic serine trap being in close proximity to the catalytic serine 441 (Fig. 1c). Retrospective docking studies of the selected peptide sequences (Table 1) against this crystal structure confirmed the predicted binding modes found in the homology model and matriptase surrogate and their classification as potential TMPRSS2 binders (Supplementary Tables 3 and 4).

Designed peptidomimetic inhibitors block TMPRSS2 and matriptase activity. We next investigated the impact of the 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 (Fig. 2a, b and Table 2). 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. The activity of the inhibitors against TMPRSS2 correlated with their activity against matriptase (Fig. 3). Yet, compound 1 showed the highest selectivity (~51-fold) for matriptase while FOY-251 showed the highest selectivity (~18-fold) for TMPRSS2.

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 (Fig. 2c, d and Table 2). All compounds excluding compound 1 displayed a >100-fold selectivity against thrombin (Table 2). Compound 1 showed no selectivity over factor Xa while compounds 2, 3, 4, 5, 6, and 7 revealed 1.6–38.9-fold selectivity compared to TMPRSS2. 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 (Supplementary Fig. 6 and Supplementary Table 5). Considering the inhibitory constants and selectivity over potential off-target coagulation proteases, the

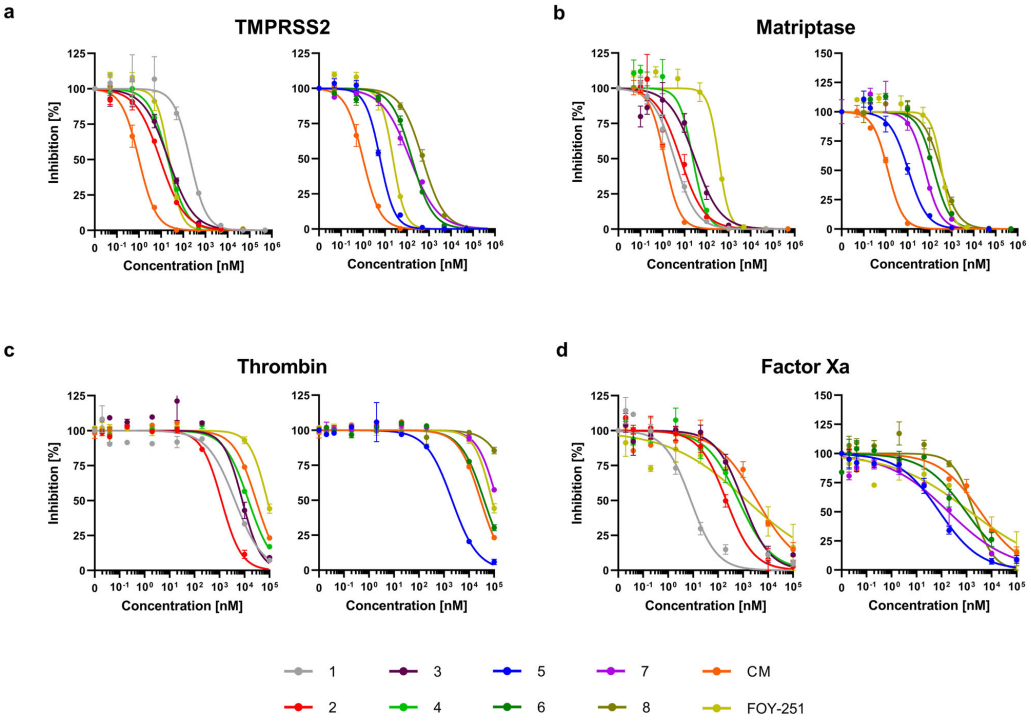


Fig. 2 Peptidomimetic inhibitors block activity of purified proteases. Isolated TMPRSS2 (**a**), matriptase (**b**), thrombin (**c**), and factor Xa (**d**) were mixed with peptidomimetic inhibitors, camostat mesylate (CM), and FOY-251. After 30 min, the fluorogenic reference substrate Boc-Gln-Ala-Arg-AMC was added to TMPRSS2 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 $n = 1$ experiment performed in triplicates.

Table 2 Inhibitory activity (K_i) of synthesized TMPRSS2 inhibitors 1–8 against TMPRSS2, matriptase, thrombin, and factor Xa.

	K_i [nM]				Selectivity indices		
Compound	TMPRSS2	Matriptase	Thrombin	Factor Xa	Matriptase	Thrombin	Factor Xa
1	86.7	1.7	2077	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	4088	271.3	1.5	449	29.8
4	8.5	11.5	7217	331	1.4	849	38.9
5	2.5	5.2	1046	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	1785.5	1.5	>12,500	4,464
FOY-251	9.7	173.4	>50,000	697	17.9	>5155	71.9

Selectivity indices represent the quotient of K_i values of matriptase, thrombin, and factor Xa by the K_i value of TMPRSS2.

compounds 2, 4, 5, and 7 were further analyzed for inhibition of cellular TMPRSS2 activity.

Having demonstrated that the designed peptidomimetics inhibit cell-free TMPRSS2 activity, we next analyzed inhibition of cell-associated protease activity. For this, we used SARS-CoV-2 permissive Caco-2 cells, which show high levels of TMPRSS2 mRNA and express TMPRSS2 on the cell surface (Supplementary Fig. 7)[39]. 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 (Fig. 4a and Supplementary Table 6). To ensure that the reduction in cellular protease activity is due to inhibition of TMPRSS2, HEK293T cells transiently expressing TMPRSS2 were treated with the respective inhibitors and fluorogenic protease substrate. The signals were corrected for the protease activity of mock-transfected HEK293T cells and revealed a dose-dependent reduction of cellular TMPRSS2

activity (Fig. 4b and Supplementary Table 6), with compounds 2 ($IC_{50} = 3.5$ nM) and 5 ($IC_{50} = 2.2$ nM) as the most active synthesized inhibitors. 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 TMPRSS2 activity.

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 wild-type 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 (Fig. 5a). Compound 5 was most efficient with an IC_{50} value of 467.2 nM and was even more potent than CM (IC_{50} ~ 747.5 nM). Compounds 1–4 and 7 suppressed spike-driven entry with IC_{50} values between 1200 and 2068 nM while compounds 6 and 8 were the least antivirally active with IC_{50} values between 5604–12,085 nM, respectively (Supplementary Table 7).

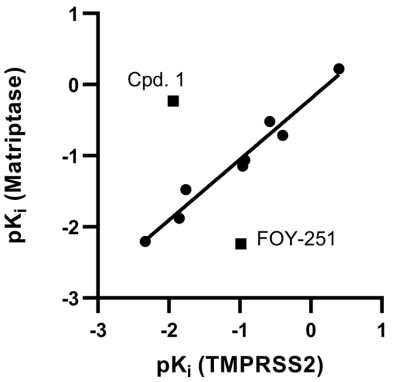


Fig. 3 Correlation of the pK_i values from compounds 1-8, CM and FOY-251 against TMPRSS2 and matriptase. A linear regression fit ($R^2 = 0.96$) was plotted through the data points excluding cpd. 1 and FOY-251.

As the wild-type virus has largely been replaced by SARS-CoV-2 variants of concern (VOC) with increased transmissibility, virulence, or immune escape, we also determined the activity of selected compounds against the Alpha, Beta, Delta, and Omicron BA.1 spike protein. TMPRSS2 inhibitors 2, 4, 5, and 7 suppressed cell entry mediated by all spike variants in a dose-dependent manner with IC_{50} values ranging between 260.7 and 2367 nM (Fig. 5b–d and Supplementary Table 7). Collectively these data show that the designed TMPRSS2 inhibitors suppress SARS-CoV-2 spike-driven viral transduction.

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 SARS-CoV-2 strain Wuhan-Hu-1. Infection rates were determined 2 days later by quantifying intracellular viral protein expression by ELISA[40]. All compounds including CM suppressed SARS-CoV-2 infection in a concentration-dependent manner (Fig. 6a). The IC_{50} values were, however, generally higher as compared to the pseudotype experiment (Supplementary Table 8). 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. The DMSO solvent control neither affected the transduction of cells with SARS-CoV-2 pseudoparticles nor the infection with wild-type virus (Supplementary Fig. 8) and we did not observe cytotoxic effects from the compounds tested that exceeded the toxicity of the solvent control (Fig. 6e and Supplementary Fig. 9).

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 (Fig. 6b), and the VOCs Alpha (Fig. 6c) and Beta (Fig. 6d)[41]. The four selected compounds as well as CM inhibited all three tested SARS-CoV-2 isolates. Compounds 2 and 5 suppressed the infection of the SARS-CoV-2 Wuhan-Hu-1 D614G strain with IC_{50} values of 17.1 and 11.9 µM, respectively, and were even more active than CM (26.7 µM) (Supplementary Table 8). SARS-CoV-2 variants Alpha and Beta 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 (Supplementary Table 8) but were still capable of blocking infection entirely. Thus, the designed peptidomimetic TMPRSS2 inhibitors prevent SARS-CoV-2 infection with comparable antiviral activity as

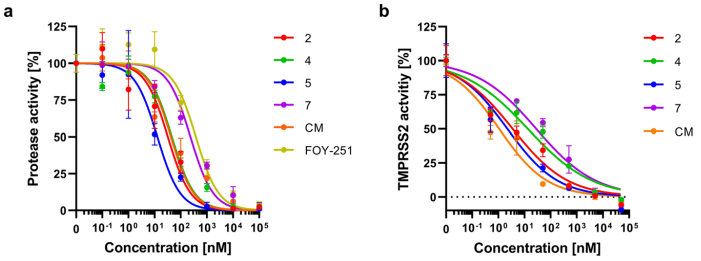


Fig. 4 Peptidomimetic inhibitors block cellular protease activity. a Peptidomimetic compounds 2, 4, 5, and 7 as well as camostat mesylate (CM) and FOY-251 were added to Caco-2 cells. After 30 min, the fluorogenic reference substrate Boc-Gln-Ala-Arg-AMC was added, and the reaction rate of substrate degradation was assessed by recording the fluorescence intensity within 2 h. **b** Peptidomimetic compounds 2, 4, 5, and 7 as well as camostat mesylate (CM) were added to HEK293T cells transiently expressing TMPRSS2, followed by addition of fluorogenic reference substrate Boc-Gln-Ala-Arg-AMC. Graph shows normalized fluorescence intensities after 2 h, corrected for the signal of mock-transfected HEK293T cells. Shown are the means ± SD of $n = 1$ experiment performed in triplicates (**a**) or duplicates (**b**). Calculated IC_{50} values for each compound are presented in Supplementary Table 6.

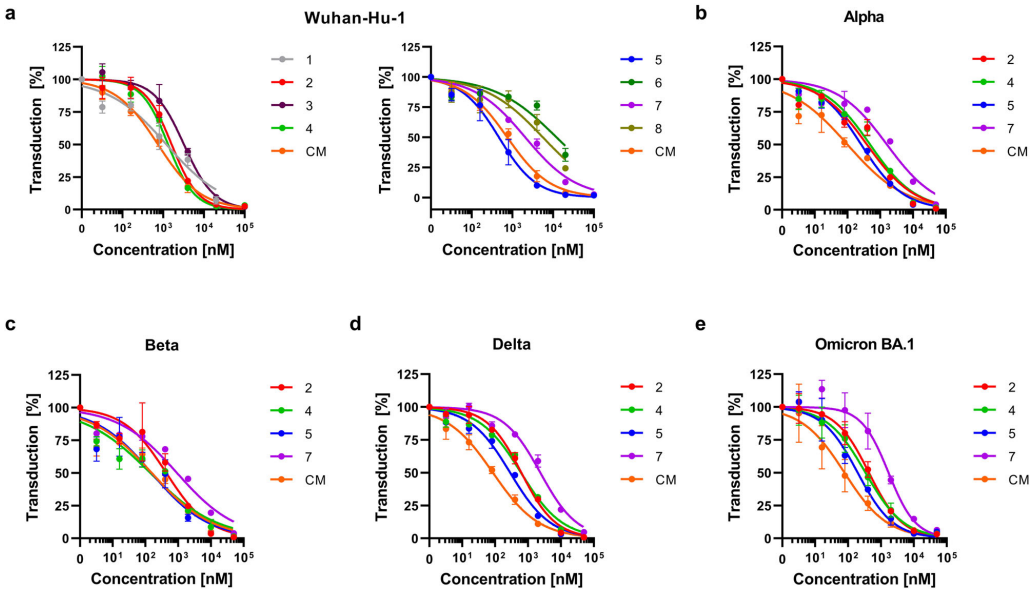


Fig. 5 Peptidomimetic inhibitors reduce SARS-CoV-2 spike-driven entry. Peptidomimetic inhibitors and the small molecule camostat mesylate (CM) were added to Caco-2 cells. After 1 h, cells were transduced with lentiviral SARS-CoV-2 pseudoparticles carrying the spike protein of SARS-CoV-2 wild-type (**a**), Alpha (**b**), Beta (**c**), Delta (**d**), or Omicron BA.1 (**e**) variant of concern. Transduction rates were assessed 2 days post transduction by measuring luciferase activity in cell lysates. Shown are the means ± SEM of $n = 2$ independent experiments, each performed in triplicates. Calculated IC_{50} values for each compound are presented in Supplementary Table 7.

CM, which is currently evaluated in clinical trials as COVID-19 therapeutic.

In vitro stability of selected leads in body fluids and epimerization studies. The stability of peptidomimetic inhibitors is the 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 selectivity as well as potency in suppressing genuine SARS-CoV-2 infection, and compound **7** was selected since the structure comprises non-proteinogenic amino acids with potential resistance to proteolysis.

In the first experiment, compound 7 was spiked into 25% human serum and chromatographic analysis of peptide content confirmed the presence of residual inhibitor for up to 10 days (Supplementary Fig. 10). Due to the strong electrophilicity of the ketobenzothiazole serine trap[42], the P1 Arg C_α-atom of compound 7 atom readily epimerized within 30 min (Supplementary Fig. 10). Of note, both epimers displayed similar inhibitory activity, as shown for compound 2 (Supplementary Fig. 11). The assay buffer used for in vitro activity studies did not alter the epimerization ratio significantly (Supplementary Fig. 12). We then determined the residual TMPRSS2 inhibitory activity of compounds 2 and 7 and camostat mesylate after incubation in human serum (Fig. 7a–c). Compounds 2 and 7 displayed similar activity than in assay buffer and both inhibitors retained their activity against TMPRSS2 for all timepoints tested, with an increase in K_i values of 1.4–1.8-fold during the course of 10 days (Fig. 7d). In contrast, the activity of the control inhibitor camostat mesylate reduced by around 60-fold upon addition to serum. Further reduction in activity was observed up to day 1, while the inhibition vanished entirely within 10 days in human serum. Similar observations were made for incubation of the compounds in human plasma and serum-free cell culture medium (Supplementary Figs. 13 and 14). Therefore, the stability of the peptidomimetic inhibitors supports further studies in in vivo models.

Discussion

We here describe 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 serine trap moiety as a reactive functional group to yield a panel of inhibitors[43–45]. 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 were used for comparison since they

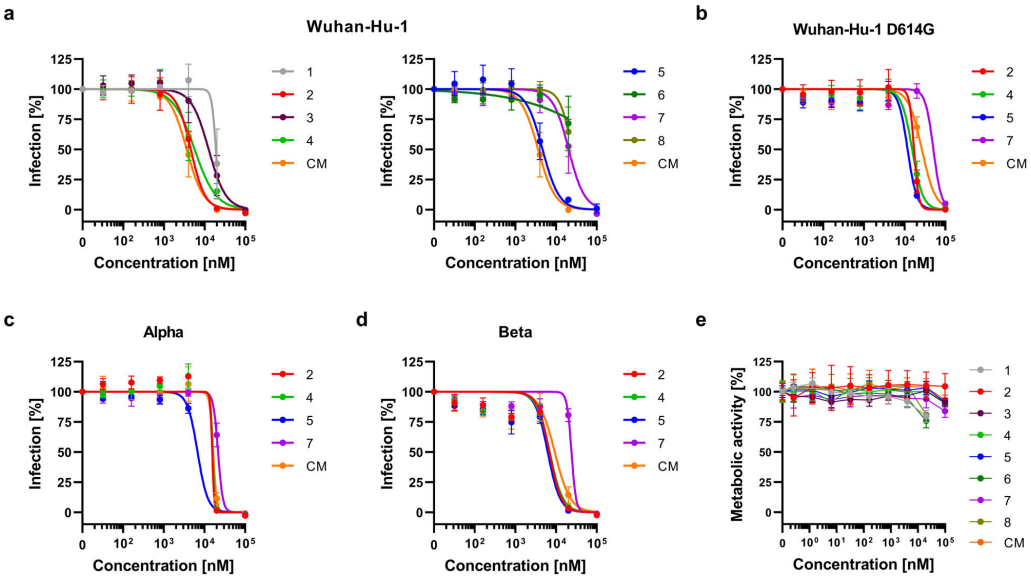


Fig. 6 Peptidomimetic inhibitors reduce SARS-CoV-2 infection. Peptidomimetic inhibitors and the small molecule camostat mesylate (CM) were added to Caco-2 cells. After 1 h, cells were infected with SARS-CoV-2 Wuhan-Hu-1 (**a**), SARS-CoV-2 bearing the spike D614G mutation (**b**), or the variants of concern Alpha (**c**) and Beta (**d**). Infection rates were determined 2 days post infection by in cell ELISA for the viral N protein. (**e**) Cytotoxicity of peptidomimetic inhibitors. 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. Shown are the means ± SEM of $n = 3$ independent experiments (**a**–**d**) or mean ± SD of $n = 1$ experiment (**e**), each performed in triplicates. Calculated IC_{50} values for each compound are listed in Supplementary Table 8.

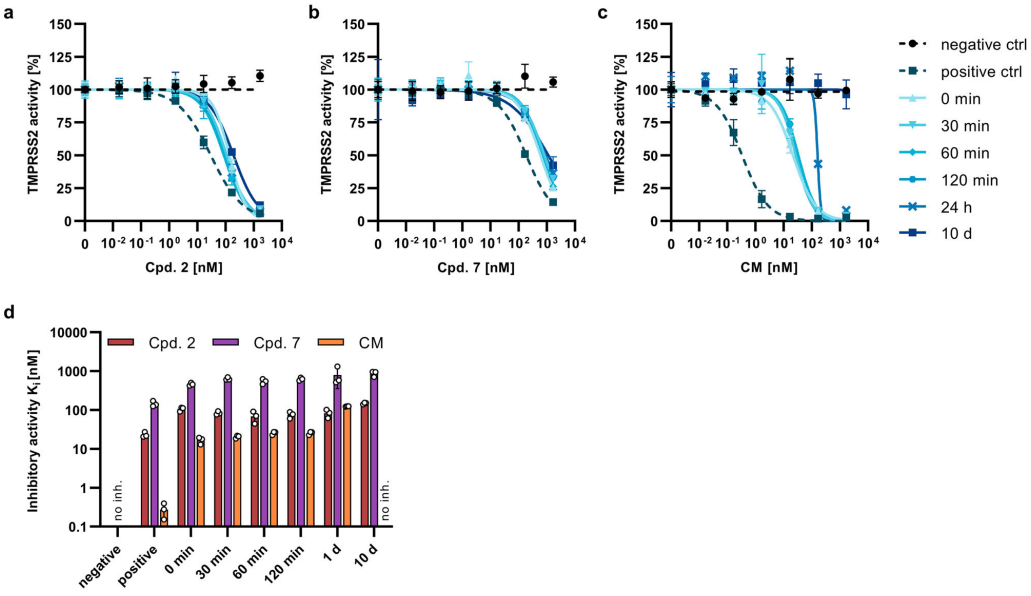


Fig. 7 Serum stability of inhibitors. Compound 2 (**a**), compound 7 (**b**), or camostat mesylate (**c**) were incubated in human serum for indicated timepoints. Samples were mixed with recombinant TMPRSS2, followed by the addition of the fluorogenic reference substrate BOC-Gln-Ala-Arg-AMC. Graph shows normalized fluorescence intensities after incubation for 2 h. Negative ctrl: no inhibitor, positive ctrl: inhibitor in assay buffer. **d** Inhibitory constants K_i as determined from (**a**–**c**). Shown are the means ± SD of $n = 1$ experiment performed in triplicates. No inh. no inhibition.

have been shown previously to efficiently inhibit TMPRSS2 proteolytic activity and CM is currently evaluated in clinical trials for COVID-19[19,46]. 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 of TMPRSS2, in contrast to the less hydrophobic and smaller ketothiazole. The addition of further amino acids or mimetics on the ketobenzothiazole moiety could improve the interaction with the S'-sites and might be of interest in further studies[43].

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, and TMPRSS2 expressing HEK293T cells. Finally, the inhibitors blocked SARS-CoV-2 spike-driven viral entry into and infection of Caco-2 cells by authentic SARS-CoV-2 wild-type 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 the tested SARS-CoV-2 VOCs are still dependent on TMPRSS2 as essential cofactor for cell entry and demonstrate that VOCs that escape from pre-existing immunity are equally sensitive to entry inhibitors, as previously shown for soluble ACE2 or fusion-inhibiting peptide EK1 and EK1C4[11,47].

Two compounds (2 and 7) were incubated in body fluids for up to 10 days and retained inhibitory activity in the sub-nanomolar range which is remarkable considering the literature known stability issues of peptide therapeutics[48,49]. Accordingly, CM was found inactive when incubated for the same period of time. 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 warrants further preclinical development of selected compounds. Furthermore, a combination with drugs targeting viral replication, such as the protease inhibitor nirmatrelvir (in combination with ritonavir, Paxlovid®), could yield synergistic effects. 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 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.

Methods

Molecular modeling. When modeling was performed, no crystal structure of TMPRSS2 was freely available in the protein data bank (PDB)[37]. Thus, a TMPRSS2 homology model was built using the Swiss-Model web server[50]. The template structure was selected based on the serine-protease hepsin in complex with *N*-acetyl-6-ammonio-L-norleucyl-L-glutaminyl-*N*-[(1 S)-4-w-1-(chloroacetyl)butyl]-L-leucinamide (PDB-ID: 1Z8G)[51] 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*-[(1 S,2 S)-1-(1,3-benzothiazol-2-yl)-5-carbamimidamido-1-hydroxypentan-2-yl]-L-valinamide (PDB-ID: 6N4T)[36] as a surrogate model were used. The focused serine-protease inhibitor library was derived from the ZINC15 database[52]. Sequences of tripeptides for docking studies on TMPRSS2 homology/surrogate were generated using CycloPs and included proteinogenic and non-proteinogenic amino acids (aa)[53]. The generated SMILES were modified to carry a N-terminal acetyl cap (ace) and a C-terminal aldehyde serine trap. Prior to docking, all molecules were protonated and energetically minimized using MOE2019[54]. Hereby, the MMFF94x[55] forcefield was used for small molecules and AMBER14:EHT[56] 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)[57]. 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[58] 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 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)[31,32,59]. Results were analyzed by FlexX score and visual pose inspection to select molecules for purchase and synthesis. During the course of this study the crystal structure of TMPRSS2 in a covalent complex with nafamostat became available (PDB-ID: 7MEQ)[38]. Hence, retrospective docking studies were performed with this structure as well. After untethering the covalent bond, the binding site was defined to include residues 8 Å around the reference ligand and additionally residues Leu-419, Lys-340, Thr-341 from the S2-S4 sites and water molecules forming at least three interactions with the protease (water molecules 703,735 and 792 within the S1 pocket) were included. The LeadIT docking parameters were as described for the hepsin-based homology model and the surrogate matriptase. The docking setup was validated by redocking of 4-guanidinobenzoic acid (from nafamostat) and binder-decoy discrimination. Figures were made with PyMOL[60].

Chlorination of tritylhydroxide resin (Trt-OH). The chlorination of tritylhydroxide resin was performed based on a modified method described elsewhere[61]. 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 followed by 10 mL acetyl chloride. 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.

Addition of first amino acid to Trt-Cl. In all, 15 mL of dry DCM was added to 0.8 g of Trt-Cl[61] 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 were 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[62]. 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 × 10 mL), then filtrated. The coupling was done by addition of 7.5 mL of 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 min. Afterward, 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 diethyl ether.

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)[63]. 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 (Supplementary Fig. 1). 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, $[2\,M + H]^+$ 1059.6, found $[M + H]^+$ 530.2, $[2\,M + H]^+$ 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-oxo-pentan-2-yl)carbamate)[64]. 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 min, Boc-Arg(Mtr) Weinreb amide (1.94 g, 3.21 mmol, 1 equiv.) was dissolved in THF (15 mL) and added slowly over 50 min. The mixture was stirred at −78.8 °C for 3 h. The reaction was quenched with sat. aqueous NH_4Cl (30 mL) and the aqueous layer was extracted with EtOAc (40 mL). The organic phase was collected, dried with Na_2SO_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 (Supplementary Fig. 1). 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 $C_{28}H_{37}N_5O_6S_2$ $[M + H]^+$ 604.2, $[2\,M + H]^+$ 1207.4, found $[M + H]^+$ 603.9, $[2\,M + H]^+$ 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)[43]. 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 (Supplementary Fig. 1). 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)[65]. 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_4Cl (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_2SO_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 (Supplementary Fig. 1). 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 $C_{24}H_{35}N_5O_6S_2$ $[M + H]^+$ 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 (Supplementary Fig. 1). 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 min 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 colorless oil (Supplementary Fig. 1). 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 (Supplementary Fig. 1).

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_2O) and further agitated for 8 h at room temperature. After concentrating, the crude inhibitor was precipitated in 50 mL cold diethyl ether and afterward purified using RP-HPLC.

Purification, lyophilization, and analysis of peptide- and serine trap 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 for deprotection was dissolved in the water to a concentration of 0.1%. Column used: Zorbax Eclipse XDB C-18 9.4 × 250 mm 5 µm, company: Agilent Technologies. Detector: UV Vis detector model S-3702, company: 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. Model used: expression-L compact mass spectrometer, company 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 TMPRSS2 was purchased from Creative BioMart (New York, USA) or LSBio (Seattle, USA), factor Xa was obtained from Bio-Techne GmbH (Wiesbaden, Germany). Recombinant human thrombin 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 min 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-Homo-pro-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 min for 2 h or as endpoint after 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[66]. The K_M value was determined to be 77 µM for TMPRSS2 (Supplementary Fig. 2).

Analysis of cellular TMPRSS2 expression. The human colorectal adenocarcinoma cell line Caco-2 from the 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 the TMPRSS2 antibody (ThermoFisher Scientific, PA5-14264) at final concentrations of 10, 20, 40, and 100 µg/mL for 30 min 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 min at 4 °C. Following a final centrifugation (200 × *g* for 3 min), the cells were resuspended in 1 mL DPBS and analyzed by flow cytometry. The measurements were performed on an Attune™ NxT cytometer (ThermoFisher) 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. In total, 10,000 Caco-2 cells (ATCC) in 100 µL EMEM 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 4 days until full confluency of the cells. The cells were washed two times with PBS and EMEM medium without FBS was added. For the determination of IC_{50} values, 1 µL of inhibitor was incubated for 30 min 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. In all, 20,000 HEK293T cells (ATCC) were seeded in 100 µL DMEM supplemented with 10% FBS, 100 U/mL penicillin, 100 mg/mL streptomycin, and 2 mM glutamine. The next day, cells were transfected with 100 ng TMPRSS2 expression plasmid (Addgene 53887, kindly provided by Roger Reeves, Johns Hopkins University, Baltimore, USA) using polyethyleneimine (PEI). Briefly, DNA was mixed with PEI at a DNA:PEI ratio of 1:3 in serum-free medium, incubated for 20 min at room temperature, and added to the cells. After 14 h, transfection mix was removed from the cells and 80 µL fresh medium without FBS and 10 µL inhibitors were added (final concentration 50 µM). After 15 min, 20 µL of fluorogenic reference substrate Boc-Gln-Ala-Arg-AMC were added (final concentration of 100 µM). Fluorescence intensity was measured after 2 h at 37 °C as described above.

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 Alpha (B.1.1.7), Beta (B.1.351), Delta (B.1.617.2) or Omicron BA.1 (B.1.1.529) 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 min at room temperature, and added to cells dropwise. At 8 h post transfection, the 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 min at 1500 rpm.

SARS-CoV-2 strains and propagation. Viral isolates BetaCoV/Netherlands/01/NL/2020 (SARS-CoV-2 D614G variant, #010V-03903), BetaCoV/France/IDF0372/2020 (SARS-CoV-2 WT, #014V-03890) and variant of concern (VOC) Alpha hCoV-19/Netherlands/NH-RIVM-20432/2020 (B.1.1.7, #014V-04031) were obtained from the European Virus Archive global. The VOC Beta 2102-cov-IM-r1-164 (B.1.351) was isolated, sequenced, and kindly provided by Michael Schindler (Tübingen Medical Center, Germany). All strains were 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 strong cytopathic effect (CPE) was visible. Supernatants were centrifuged for 5 min at 1000 × *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. Overall, 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 1 mM 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 min at room temperature. In all, 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. Overall, 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 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 min at room temperature and permeabilized by incubation with 100 µL 0.1% Triton X-100 for 5 min. After washing once with PBS, cells were stained with 1:5000 diluted anti-spike protein antibody 1A9 (Biozol GTX-GTX632604) or anti-nucleocapsid antibody (Sinobiological 40143-MM05) in antibody buffer (10% FBS and 0.3% Tween 20 in PBS) for 1 h at 37 °C. After two washes with 0.3% Tween 20 in PBS, the secondary HRP-conjugated antibody (ThermoFisher #A16066) (1:15,000) was incubated for 1 h at 37 °C. Cells were washed three times with 0.3% Tween 20 in PBS, TMB peroxidase substrate (Medac #52-00-04) was added for 5 min and the reaction stopped using 0.5 M H_2SO_4. The optical density (OD) was recorded at 450–620 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. In all, 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 min at room temperature. In total, 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 ref. [67]. In short, 10 µL of 10 mM inhibitor solutions were added to 1 mL of 25% (v/v) human serum or plasma in RPMI 1640 and incubated at 37 °C. At indicated timepoints, 100 µL samples were taken and mixed with 200 µL of ethanol to precipitate proteins. The cloudy solution was cooled at 4 °C for 15 min and centrifuged at 14,800 rpm for 2 min. The supernatant was aspirated and analyzed using analytical HPLC. The residual inhibitory constants K_i after incubation was assessed by mixing serial dilutions of the inhibitor with recombinant TMPRSS2 as described above.

Statistics and reproducibility. Sample sizes and the number of replicates are indicated in the respective figure legend. The number of experiments is described as $n = X$, triplicates/duplicates describe individual replicates within each experiment.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Crystal structures were obtained from Protein Data Bank with accession codes PDB-ID: 1Z8G, 6N4T, 7MEQ, and from SWISS-MODEL repository (https://swissmodel.expasy.org/repository/uniprot/P05981?template=1z8g). Source data are provided with this paper as Supplementary Data 1. Analytical data are presented in Supplementary Fig. 1.

Received: 26 October 2021; Accepted: 21 June 2022;
Published online: 08 July 2022

References

1. WHO. COVID-19 weekly epidemiological update. *World Health Organization* 1–23 https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---29-march-2022 (2022).
2. Ji, H.-L., Zhao, R., Matalon, S. & Matthay, M. A. Elevated plasmin(ogen) as a common risk factor for COVID-19 susceptibility. *Physiol. Rev.* **100**, 1065–1075 (2020).
3. Li, G. et al. Mortality risk of COVID-19 in elderly males with comorbidities: a multi-country study. *Aging* **13**, 27–60 (2021).
4. 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.* **20**, 363–374 (2020).
5. Gupta, A. et al. Extrapulmonary manifestations of COVID-19. *Nat. Med.* **26**, 1017–1032 (2020).
6. Hou, Y. J. et al. SARS-CoV-2 reverse genetics reveals a variable infection gradient in the respiratory tract. *Cell* **182**, 429–446.e14 (2020).
7. Xiao, F. et al. Evidence for gastrointestinal infection of SARS-CoV-2. *Gastroenterology* **158**, 1831–1833.e3 (2020).
8. Lin, L. et al. Gastrointestinal symptoms of 95 cases with SARS-CoV-2 infection. *Gut* **69**, 997–1001 (2020).
9. Lindner, D. et al. Association of cardiac infection with SARS-CoV-2 in confirmed COVID-19 autopsy cases. *JAMA Cardiol.* **5**, 1281 (2020).

10. Sah, P. et al. Accelerated vaccine rollout is imperative to mitigate highly transmissible COVID-19 variants. *EClinicalMedicine* **35**, 100865 (2021).
11. Hoffmann, M. et al. SARS-CoV-2 variants B.1.351 and P.1 escape from neutralizing antibodies. *Cell* **184**, 2384–2393.e12 (2021).
12. Wang, Z. et al. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature* **592**, 616–622 (2021).
13. Frampton, D. et al. 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.* **21**, 1246–1256 (2021).
14. Davies, N. G. et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* **372**, eabg3055 (2021).
15. EMA. COVID-19: EMA recommends conditional marketing authorisation for Paxlovid. https://www.ema.europa.eu/en/news/covid-19-ema-recommends-conditional-marketing-authorisation-paxlovid (2022).
16. EMA. EMA issues advice on use of Lagevrio (molnupiravir) for the treatment of COVID-19. https://www.ema.europa.eu/en/news/ema-issues-advice-use-lagevrio-molnupiravir-treatment-covid-19 (2021).
17. Ou, X. et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat. Commun.* **11**, 1620 (2020).
18. 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.* **5**, 562–569 (2020).
19. Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280.e8 (2020).
20. Limburg, H. et al. 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.* **93**, 649–668 (2019).
21. Shirogane, Y. et al. Efficient multiplication of human metapneumovirus in vero cells expressing the transmembrane serine protease TMPRSS2. *J. Virol.* **82**, 8942–8946 (2008).
22. Matsuyama, S. et al. Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2. *J. Virol.* **84**, 12658–12664 (2010).
23. Shirato, K., Kawase, M. & Matsuyama, S. Middle East respiratory syndrome coronavirus infection mediated by the transmembrane serine protease TMPRSS2. *J. Virol.* **87**, 12552–12561 (2013).
24. Bertram, S. et al. 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.* **87**, 6150–6160 (2013).
25. Wettstein, L., Kirchhoff, F. & Münch, J. The transmembrane protease TMPRSS2 as a therapeutic target for COVID-19 treatment. *Int. J. Mol. Sci.* **23**, 1351 (2022).
26. Hatesuer, B. et al. Tmprss2 is essential for influenza H1N1 virus pathogenesis in mice. *PLoS Pathog.* **9**, e1003774 (2013).
27. Iwata-Yoshikawa, N. et al. TMPRSS2 contributes to virus spread and immunopathology in the airways of murine models after coronavirus infection. *J. Virol.* **93**, e01815-18 (2019).
28. Kim, T. S., Heinlein, C., Hackman, R. C. & Nelson, P. S. Phenotypic analysis of mice lacking the Tmprss2-encoded protease. *Mol. Cell. Biol.* **26**, 965–975 (2006).
29. Szabo, R. & Bugge, T. H. Membrane-anchored serine proteases in vertebrate cell and developmental biology. *Annu. Rev. Cell Dev. Biol.* **27**, 213–235 (2011).
30. Ramsay, A. J., Hooper, J. D., Folgueras, A. R., Velasco, G. & Lopez-Otin, C. Matriptase-2 (TMPRSS6): a proteolytic regulator of iron homeostasis. *Haematologica* **94**, 840–849 (2009).
31. Meyer, D. et al. Identification of the first synthetic inhibitors of the type II transmembrane serine protease TMPRSS2 suitable for inhibition of influenza virus activation. *Biochem. J.* **452**, 331–343 (2013).
32. Sielaff, F. et al. Development of substrate analogue inhibitors for the human airway trypsin-like protease HAT. *Bioorg. Med. Chem. Lett.* **21**, 4860–4864 (2011).
33. Wettstein, L. et al. Alpha-1 antitrypsin inhibits TMPRSS2 protease activity and SARS-CoV-2 infection. *Nat. Commun.* **12**, 1726 (2021).
34. 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.* **157**, 104837 (2020).
35. 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.* **15**, 801–812 (2020).
36. Béliveau, F. et al. Discovery and development of TMPRSS6 inhibitors modulating hepcidin levels in human hepatocytes. *Cell Chem. Biol.* **26**, 1559–1572.e9 (2019).
37. Berman, H. M. The protein data bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
38. Fraser, B. J. et al. Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. *Nat. Chem. Biol.* (2022) https://doi.org/10.1038/s41589-022-01059-7.
39. Bertram, S. et al. TMPRSS2 and TMPRSS4 facilitate trypsin-independent spread of influenza virus in Caco-2 cells. *J. Virol.* **84**, 10016–10025 (2010).
40. Conzelmann, C. et al. An enzyme-based immunodetection assay to quantify SARS-CoV-2 infection. *Antivir. Res.* **181**, 104882 (2020).
41. Yurkovetskiy, L. et al. Structural and functional analysis of the D614G SARS-CoV-2 spike protein variant. *Cell* **183**, 739–751.e8 (2020).
42. Costanzo, M. J. et al. In-depth study of tripeptide-based α-ketoheterocycles as inhibitors of thrombin. Effective utilization of the S 1 ' subsite and its implications to structure-based drug design. *J. Med. Chem.* **48**, 1984–2008 (2005).
43. Han, Z. et al. α-Ketobenzothiazole serine protease inhibitors of aberrant HGF/c-MET and MSP/RON kinase pathway signaling in cancer. *ChemMedChem* **11**, 585–599 (2016).
44. Damalanka, V. C., Wildman, S. A. & Janetka, J. W. Piperidine carbamate peptidomimetic inhibitors of the serine proteases HGFA, matriptase and hepsin. *Medchemcomm* **10**, 1646–1655 (2019).
45. Colombo, É. et al. Design and synthesis of potent, selective inhibitors of matriptase. *ACS Med. Chem. Lett.* **3**, 530–534 (2012).
46. Hoffmann, M. et al. Camostat mesylate inhibits SARS-CoV-2 activation by TMPRSS2-related proteases and its metabolite GBPA exerts antiviral activity. *EBioMedicine* **65**, 103255 (2021).
47. Zhou, D. et al. Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine-induced sera. *Cell* **184**, 2348–2361.e6 (2021).
48. Nguyen, L. T. et al. Serum stabilities of short tryptophan- and arginine-rich antimicrobial peptide analogs. *PLoS ONE* **5**, e12684 (2010).
49. Li, H. et al. Novel retro-inverso peptide inhibitor reverses angiotensin receptor autoantibody-induced hypertension in the rabbit. *Hypertension* **65**, 793–799 (2015).
50. Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
51. Herter, S. et al. 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.* **390**, 125–136 (2005).
52. Sterling, T. & Irwin, J. J. ZINC 15—ligand discovery for everyone. *J. Chem. Inf. Model.* **55**, 2324–2337 (2015).
53. Duffy, F. J. et al. CycloPs: generating virtual libraries of cyclized and constrained peptides including nonnatural amino acids. *J. Chem. Inf. Model.* **51**, 829–836 (2011).
54. Chemical Computing Group ULC, 1010 Sherbrooke St. WEst, Suite #910, Montreal, QC, Canada, H. 2R7. (Molecular Operating Environment (MOE), 2019).
55. Halgren, T. A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **17**, 490–519 (1996).
56. Maier, J. A. et al. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
57. Rarey, M., Kramer, B., Lengauer, T. & Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **261**, 470–489 (1996).
58. Bietz, S., Urbaczek, S., Schulz, B. & Rarey, M. Protoss: a holistic approach to predict tautomers and protonation states in protein-ligand complexes. *J. Cheminform.* **6**, 12 (2014).
59. 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.* **55**, 6582–6594 (2012).
60. The PyMOL Molecular Graphics System, Version 2.2.0. (Schrödinger, LLC, 2015).
61. CEM Corporation. Trityl-Hydroxide (TRT-OH) SpheriTide Resin Loading Procedure. http://cem.com/media/contenttype/media/literature/525_TechNote_SpheriTide_TRT-OHLoading.pdf. PCT/EP2012/057264, (2017).
62. Coin, I., Beyermann, M. & Bienert, M. Solid-phase peptide synthesis: from standard procedures to the synthesis of difficult sequences. *Nat. Protoc.* **2**, 3247–3256 (2007).
63. Lin, J. et al. Design, synthesis, and biological evaluation of peptidomimetic inhibitors of factor XIa as novel anticoagulants. *J. Med. Chem.* **49**, 7781–7791 (2006).
64. Damalanka, V. C. et al. Discovery of selective matriptase and hepsin serine protease inhibitors: useful chemical tools for cancer cell biology. *J. Med. Chem.* **62**, 480–490 (2019).
65. Adang, A. E. P. et al. Unique overlap in the prerequisites for thrombin inhibition and oral bioavailability resulting in potent oral antithrombotics. *J. Med. Chem.* **45**, 4419–4432 (2002).
66. Yung-Chi, C. & Prusoff, W. H. Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem. Pharmacol.* **22**, 3099–3108 (1973).
67. Jenssen, H. & Aspmo, S. I. Serum stability of peptides. *Methods Mol. Biol.* **494**, 177–186 (2008).

Acknowledgements
This work was supported by grants from the MWK Baden-Württemberg (to J.M.), the EU's Horizon 2020 research and innovation program (Fight-nCoV, 101003555 to J.M.) and by the DFG (CRC1279 to J.M.).

Author contributions
P.K. and P.M. synthesized the compounds. P.K. and L.W. performed in vitro characterization, stability analyses, and pseudovirus assays. C.K. performed the molecular docking, hit prioritization, and final compound selection for synthesis in agreement with P.K and M.B. T.W., C.C., and J. Müller performed infection assays. P.K., L.W., J.M., and V.M. wrote the manuscript with contributions from all authors. C.K., M.H., S.P., T.S., and K.L. advised and edited the manuscript.

Funding
Open Access funding enabled and organized by Projekt DEAL.

Competing interests
The authors declare the following competing interests: P.K., L.W., J.M., and V.M. filed a patent that claims to use peptidomimetics described herein to inhibit protease activity and cure viral infections. The remaining authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-022-03613-4.

Correspondence and requests for materials should be addressed to Jan Münch or Volker Mailänder.

Peer review information *Communications Biology* thanks Shibo Jiang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Anam Akhtar and Christina Karlsson Rosenthal. Peer reviewer reports are available.

"

5.2 Structure-Based Design of High-Affinity and Selective Peptidomimetic Hepsin Inhibitors

5.2.1 Context, Project Summary, and Own Contributions

Hepsin presents another trypsin-like transmembrane serine protease and has been shown to be overexpressed in prostate cancer.[261],[307] Due to the activation of oncogenic proteins like pro-hepatocyte growth factor (pro-HGF) by hepsin, downstream signal pathways are triggered that promote tumor progression (1.6.1).[314] For targeting hepsin-related tumors, the group of Mailänder and Landfester *et al*. investigated the development of peptide-based nanocapsules for selective release of cargo molecules.[315],[316] Novel substrates with increased affinity and stability properties were designed by virtual screening and synthesis.[315]

As a follow-up project, we designed new potential hepsin inhibitors. We combined the best binders of the aforementioned designed substrates with the electrophilic *α*-ketobenzothiazole moiety as the warhead.[315] After docking simulations of the inhibitors, we synthesized the best potential compounds and tested their affinity and selectivity via fluorometric assays towards hepsin and other related trypsin-like serine proteases. Additionally, we exchanged the warhead moiety with different electrophilic traps and truncated the peptide recognition sequence to elucidate their influence regarding their pharmacodynamic properties. Further stability tests in body fluids of the most promising inhibitors, showed high remaining potency after one day of incubation. We demonstrated that this combinatorial methodology, beginning with a virtual screening of peptide libraries and combination with a suitable warhead, lead to potent and selective inhibitors for hepsin.

Own contributions: Inhibitor synthesis (S1−S5) and writing of the original draft plus editing of the manuscript.

Contributions from other authors: Inhibitor synthesis (1−10), inhibition-, cytotoxicity- and stability-assay, molecular docking and writing of the original draft plus editing of the manuscript.

This work was published in Biomacromolecules (impact factor: 6.09).

5.2.2 Publication

The following publication quoted within “” from page 163 to page 169 is the same as the manuscript cited on page 53.

“

BioMACROMOLECULES

pubs.acs.org/Biomac Article

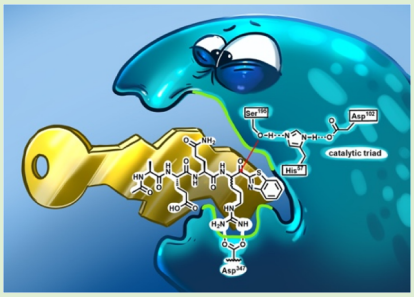
Structure-Based Design of High-Affinity and Selective Peptidomimetic Hepsin Inhibitors

Philip Maximilian Knaff, Patrick Müller, Christian Kersten, Lukas Wettstein, Jan Münch, Katharina Landfester, and Volker Mailänder*

Cite This: *Biomacromolecules* 2022, 23, 2236−2242 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

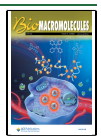
ABSTRACT: In many solid tumors, increased upregulation of transmembrane serine proteases (TTSPs) leads to an overactivation of growth factors, which promotes tumor progression. Here, we have used a combinatorial methodology to develop high-affinity tetrapeptidic inhibitors. A previous virtual screening of 8000 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. A 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 their stability in biological fluids as well as the influence of the warhead and truncated inhibitors on the inhibitory potency. Overall, this methodology yielded compounds as selective inhibitors for potential cancer drug development, where hepsin is overexpressed.



INTRODUCTION

Cancer emerges from a complex variety of gene mutations. Most frequently, the genes encoding for the protein class of kinases and proteases are highly upregulated.[1,2] 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 (TTSPs) as an alternative drug target has come into research focus.[3] Hepsin is a TTSP which shows, in healthy prostate tissues, low expression levels, while high expression levels have been shown to play critical roles in cancer such as androgen-independent prostate cancer.[4,5] 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. The 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 tissues, the activity of TTSP is regulated by endogenous inhibitors, HAI-1 and HAI-2 proteins. In vivo hepsin transgenic mice showed primary prostate tumors and progression to systemic cancer by promoting metastasis. Contrarily, hepsin knockout mice were 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.[6,7] 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.[7] The in silico approach included the design of a P- and a P′-site peptide library, consisting of 8000 combinations each, which target the respective S- and S′-sites of the binding pocket of hepsin (Figure S1). 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 recognition sequences identified previously (Figure S2). For this purpose, the top binders of the S-site library were combined with an electrophilic ketobenzothiazole (kbt) moiety to yield covalent reversible inhibitors.[8] 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 a

Received: August 5, 2021
Revised: April 14, 2022
Published: May 20, 2022



https://doi.org/10.1021/acs.biomac.1c01011
Biomacromolecules 2022, 23, 2236−2242

natural manner simply by the sequential choice of amino acids. Still, previous inhibitors showed activity among similar proteases like matriptase.[9−11] 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.[12] Toxicity effects of hepsin and matriptase short-term inhibition still need 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 elucidate 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 (SAR).

MATERIALS AND METHODS

Molecular Docking. Molecular docking studies were performed as described previously with slight modifications.[7] 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).[13−16] 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.[17] 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 kbt warhead.[18] Energetically minimized structures for docking were generated using the AMBER14/EHT force field within MOE2018.[19,20] Differently from the previous approach, the docking was performed without any pharmacophoric constraints.[7] 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).[21] Figures were made with PyMOL.[22]

Preparation of resin, addition of first amino acid, peptide synthesis, N-Cap modification, and cleavage from the resin as well as synthesis of the inhibitory head groups and coupling of the inhibitory heads with the peptides are described in the Supporting Information.

Purification, Lyophilization, and Analysis of Peptide and Warhead Precursors and Inhibitors. All precursor compounds were purified with a semi-preparative reversed-phase high-performance liquid chromatography (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%. Column used: Zorbax Eclipse XDB C-18 9.4 × 250 mm 5 μm, company: Agilent Technologies. Detector: UV−vis detector model S-3702, company: 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. Model used: expression-L compact mass spectrometer, company 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 min 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 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 the activity of the compounds (Table S1). However, due to consistency with the published literature, 30 min of 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 the 20 μL total volume and 0.6 nM (thrombin), 0.35 nM (factor Xa) in the 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. 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.[23]

Cytotoxicity Assay. 10,000 Caco-2 cells were seeded in 100 μL 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. The next day, the medium was replaced by 80 μL of fresh medium and cells were treated with 20 μL of serial dilutions of peptidomimetic hepsin inhibitors. Cell viability was assessed after 48 h with a commercially available kit (Promega). Briefly, the medium was removed and cells were lysed with 100 μL of CellTiter-Glo reagent for 10 min 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 the 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 Jenssen et al.[24] 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 the indicated intervals, 100 μL of samples was taken and mixed with 200 μL of ethanol for the precipitation of proteins. The cloudy solution was cooled at 4 °C for 15 min and centrifuged at 14,800 rpm for 2 min. The supernatant was aspirated and analyzed using analytical HPLC. The residual inhibitory constants K_i after 1 day of incubation in biological fluids was measured as described above.

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.[7] In this regard, 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

https://doi.org/10.1021/acs.biomac.1c01011
Biomacromolecules 2022, 23, 2236−2242

(Table S2) 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.[14,25] The common structure is as follows: N-terminal acetyl cap (ace), proteinogenic amino acids at P4−P2, arginine at the P1 position, and a kbt warhead (Table 1). The designed inhibitor library was virtually prepared and docked in the active binding cleft of hepsin. All compounds revealed plausible binding modes with a likeliness for the reaction and high calculated binding affinities (Table S3). In the next step, we synthesized the designed inhibitor library using a two-step approach, previously reported in the literature (Scheme 1).[9−11] 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 S4).

Table 1. Assembled Peptidomimetic Inhibitor Library for Synthesis

compound	N-cap	P4	P3	P2	P1	warhead
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

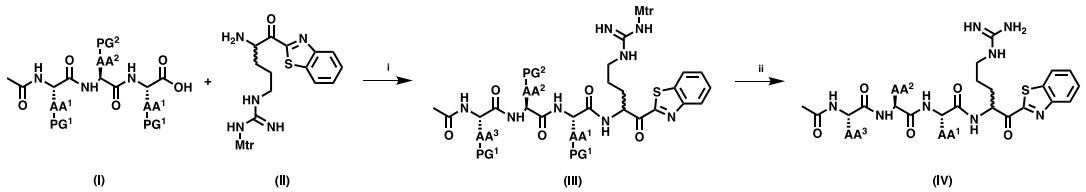
[a]Recognition sequence of previously identified high-affinity substrate (RS). [b]Recognition sequence of natural substrate (NS), PDB ID: 1Z8G. For RS, rank 2 docking pose was evaluated due to an inverted binding mode of the rank 1 pose. kbt = ketobenzothiazole, kt = ketothiazole, ol = alcohol, carb = carboxyl, pvs = phenylvinylsulfone.

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 the addition of a reporter substrate. Overall, multiple members of the P-site library showed a high potency in the low nM concentration range against hepsin and the closely related protease, matriptase, in the fluorometric inhibition assay (Table 2 and Figure S5). 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 and K_i = 0.5, 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 compounds **2** and **3** are more selective for matriptase (3.5- and 5.2-fold, respectively). 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 the 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. Compounds **RS**, **NS**, **1**, and **2** inhibit factor Xa in the low nM range with selectivity ratios varying between 12- and 73.6-fold. The highest factor Xa selectivity reveal compound **3** with 1194-fold and compound **4** with 248.4-fold. We did not observe cytotoxic effects from the compounds tested that exceed the toxicity of the solvent control (Figure S6).

To gain a better understanding of the SAR 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 overall sequence similarities 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), a 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 the selective inhibitors.[26] Based on the molecular docking studies, the arginine in the 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 catalytic Ser353, which facilitates covalent-reversible binding

Scheme 1. Synthesis of Tripeptidic Inhibitors[a]



[a]Reagents and conditions: (i) PyBOP/DIEA, RT, 2 h; (ii) TFA/TIPS/H_2O (93/3/4), 8 h.

https://doi.org/10.1021/acs.biomac.1c01011
Biomacromolecules 2022, 23, 2236−2242

Biomacromolecules pubs.acs.org/Biomac Article

Table 2. Inhibition of Hepsin and Off-Target Proteases Matriptase, Thrombin, and Factor Xa in a Fluorogenic Inhibition Assay[a]

	K_i [nM]				selectivity indices		
compound	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	1668	0.2	n.d	1194
4	5.1	79.1	>50,000	19643	15.5	n.d	248.4
5	21.7	55.8	>50,000	1167	2.6	n.d	20.9
6	467.2	374	>50,000	2647	0.8	n.d	7.1
7	20.2	157.6	>50,000	3081	7.8	n.d	19.6
3.1	19.8	3.9	>50,000	34529	0.2	n.d	8847
3.2	14,364	4971	>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	1147
9	72.4	5	>50,000	12626	0.1	n.d	2544
10	4623	144.6	>50,000	>50,000	0.03	n.d	n.d

[a]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 = phenylvinylsulfone.

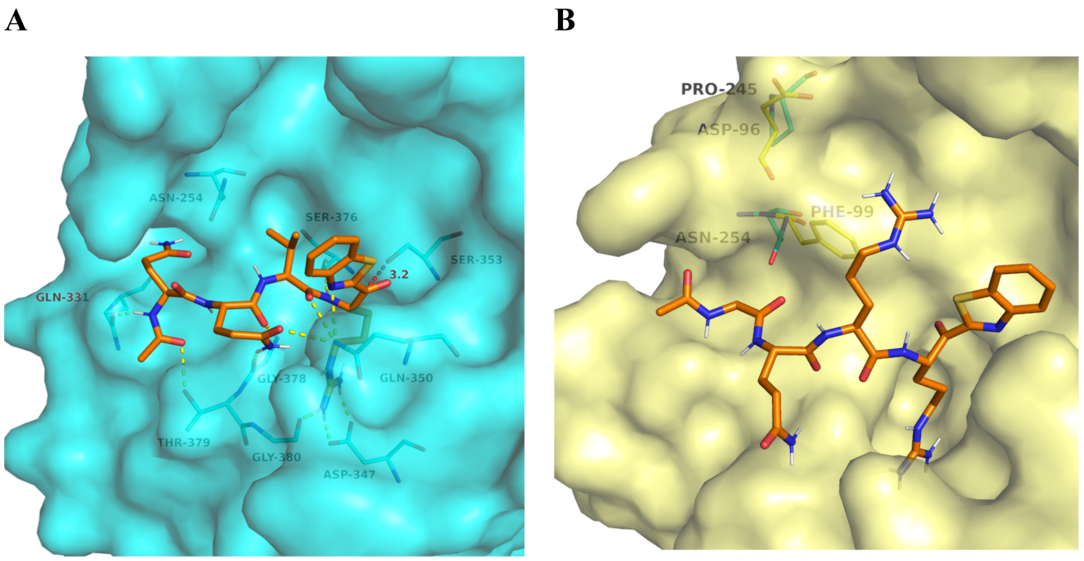


Figure 1. (A) Predicted binding mode of hepsin-selective compound **4** in complex with hepsin (PDB ID: 1Z8G). (B) Predicted binding mode of matriptase-selective compound **3** in complex with matriptase (PDB ID: 4O9V) with superposed residues of hepsin. Ligand carbon atoms are depicted in orange, hepsin carbon atoms and surface in cyan, and matriptase carbon atoms and surface in yellow. Polar interactions are shown as yellow dashed lines and the distance between electrophilic carbon of kbt and nucleophilic oxygen of Ser as red dashed line. For a clear view, only residues that are described in the text are depicted and labeled.

(Figure 1A). 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 1B). 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.[9] 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 contributions. Similar binding modes were reported for multi-basic high affinity matriptase inhibitors before.[21,27]

https://doi.org/10.1021/acs.biomac.1c01011
Biomacromolecules 2022, 23, 2236−2242

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 the removal of the S2 and S3 residues led to a decrease in inhibitory potency (Figure S7). 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, the truncation of the kbt (**3**) moiety to a ketothiazole (**3.1**) led to a ∼3-fold decrease in activity against hepsin and matriptase, while the 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 (Figure S8 and Table 2).

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.[28] Two compounds were selected due to their high activity against hepsin and good selectivity over off-target serine proteases for in vitro stability studies. Compound **3** which reveals the highest selectivity against factor Xa (1194-fold) and compound **4** with the highest matriptase selectivity (15.5-fold). Both compounds were spiked in the 25% blood serum and the inhibitory activity was tested at different time points from 0 min to 10 days. As shown in Figure S9, even after 1 day of incubation, both compounds still have the potential to fully prevent the degradation of the fluorogenic reference substrate at a concentration of 1 μM. After 10−20 days, the peptidomimetic inhibitors start to lose their potency. Analysis of the residual activity of compound **3** and **4** after 1 day of incubation in body fluids revealed only a slight loss of initial potency (Figure 2).

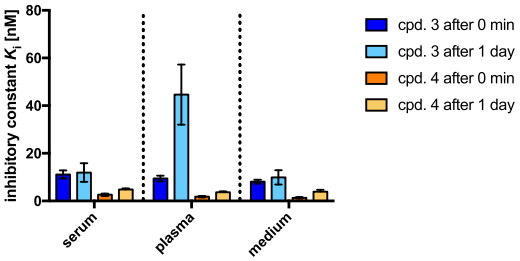


Figure 2. Residual inhibitory activity for compound **3** and compound **4** against hepsin after 0 min and 1 day incubation in 25% serum, 25% plasma, and pure RPMI cell medium.

The most prominent loss can be seen for compound **3**, which had an initial inhibitory potency of K_i = 9.4 nM versus K_i = 44.6 nM after incubation with plasma for 1 day, which results in a 5-fold loss in activity during the course of incubation.

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 tissues. 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′-sites of the binding cavity of hepsin. Assembling of those binders led to efficient substrates for hepsin. This study uses slight 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 the 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 revealed compounds **RS** and **NS** with K_i < 1 nM; however, 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 to be upregulated in cancer and to activate the endogenous ligand HGF. However, the 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** to stability tests in body fluids which showed high remaining potency even after 1 day of incubation. The kbt moiety of the serine trap proved to be the most active, which is consistent with previous findings in the literature.[10,11,29] 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 and potency. Our results might warrant deeper analysis of pharmacokinetics, broader protease selectivity or in vivo stability studies. Of note, structurally 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.

CONCLUSIONS

We are demonstrating here how a combinatorial methodology can be used to develop high affinity tetrapeptidic inhibitors. From a previous virtual screening of 8000 peptide combinations against the crystal structure of the TTSP hepsin, we had identified a series of recognition sequences, customized for the non-prime substrate binding (P) sites of this serine protease.

https://doi.org/10.1021/acs.biomac.1c01011
Biomacromolecules 2022, 23, 2236−2242

Here, we then combined these successfully with an electrophilic warhead. We demonstrated that highly potent inhibitors with good selectivity can be generated by this approach. We also demonstrated that stability in biological fluids is highly advantageous with this approach. Overall, this methodology resulted in selective inhibitors for hepsin-overexpressing tumors..

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.biomac.1c01011.

Schematic illustration of the Schechter and Berger nomenclature, scheme for the development of tailored peptide inhibitors, analytical data (structure, HPLC, MS, and NMR data) of precursors and serine traps, analytical data (structure, HPLC, and MS) of peptidomimetic inhibitors, peptidomimetic inhibitors block the activity of purified hepsin, matriptase, thrombin, and factor Xa, cytotoxicity of hepsin inhibitors, influence of truncation on the inhibitory activity of compound **3** against hepsin, matriptase, thrombin, and factor Xa, influence of serine trap on the biological activity of compound **3** against hepsin, matriptase, thrombin, and factor Xa, inhibition of reference substrate degradation by hepsin with compounds **3** and **4** at c = 1000 nM after spiking in 25% blood serum at various time points, influence of preincubation on the inhibitory activity of compounds **3** and **4** against hepsin, docking result for P-site (P1−P4)-designed molecules with the structure of acetyl-X-X-X-R-methylamide, and docking studies of assembled peptidomimetic inhibitor library (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Volker Mailänder − *Max Planck Institute for Polymer Research, Mainz 55128, Germany; Dermatology Clinic of the University Medicine of the Johannes Gutenberg University Mainz, Mainz 55131, Germany;* orcid.org/0000-0001-6583-8136; Email: mailaend@mpip-mainz.mpg.de

Authors

Philip Maximilian Knaff − *Max Planck Institute for Polymer Research, Mainz 55128, Germany; Dermatology Clinic of the University Medicine of the Johannes Gutenberg University Mainz, Mainz 55131, Germany*

Patrick Müller − *Institute for Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Mainz 55128, Germany*

Christian Kersten − *Institute for Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Mainz 55128, Germany;* orcid.org/0000-0001-9976-7639

Lukas Wettstein − *Institute of Molecular Virology, Ulm University Medical Center, Ulm 89081, Germany*

Jan Münch − *Institute of Molecular Virology, Ulm University Medical Center, Ulm 89081, Germany;* orcid.org/0000-0001-7316-7141

Katharina Landfester − *Max Planck Institute for Polymer Research, Mainz 55128, Germany;* orcid.org/0000-0001-9591-4638

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.biomac.1c01011

Funding

This work was supported by the Deutsche Forschungsgemeinschaft (DFG project HEPSIN-NANOLYSIS, MA3271/10-1) and the CRC1066.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Cicenas, J.; Zalyte, E.; Bairoch, A.; Gaudet, P. Kinases and Cancer. *Cancers (Basel)* **2018**, *10*, 63.

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

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

(4) 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*, 16948−16953.

(5) 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*, 187−191.

(6) 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*, 4829−4839.

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

(8) 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*, 1984−2008.

(9) 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*, 480−490.

(10) Han, Z.; Harris, P. K. W.; 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*, 585−599.

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

(12) 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*, 1017−1025.

(13) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470−489.

(14) 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*, 125−136.

(15) 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*, 235−242.
(16) LeadIT. *BioSolveIT*, version 2.3.2; GmbH: Germany, 2017.
(17) Bietz, S.; Urbaczek, S.; Schulz, B.; Rarey, M. Protoss: a holistic approach to predict tautomers and protonation states in protein-ligand complexes. *J. Cheminf.* **2014**, *6*, 12.
(18) 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*, 829−836.
(19) 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*, 3696−3713.
(20) Chemical Computing Group. *Molecular Operating Environment (MOE)*; Chemical Computing Group, Inc.: Montreal, QC, Canada, 2019.
(21) 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*, 3187−3203.
(22) Schrödinger, Inc.. *The PyMOL Molecular Graphics System*, version 2.2.0; Schrödinger LLC, 2017.
(23) 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*, 3099−3108.
(24) Jenssen, H.; Aspmo, S. I. Serum stability of peptides. *Methods Mol. Biol.* **2008**, *494*, 177−186.
(25) Béliveau, F.; Désilets, A.; Leduc, R. Probing the substrate specificities of matriptase, matriptase-2, hepsin and DESC1 with internally quenched fluorescent peptides. *FEBS J.* **2009**, *276*, 2213−2226.
(26) 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*, D503−D509.
(27) 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*, 610−625.
(28) Powell, M. F.; Stewart, T.; Otvos, L., Jr.; Urge, L.; Gaeta, F. C. A.; 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*, 1268−1273.
(29) Han, Z.; Harris, P. K. W.; 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*, 1219−1224.

https://doi.org/10.1021/acs.biomac.1c01011
Biomacromolecules 2022, 23, 2236−2242

5.3 Ligand-Based Design of Selective Peptidomimetic uPA and TMPRSS2 Inhibitors with Arg Bioisosteres

5.3.1 Context, Project Summary, and Own Contributions

The urokinase-type plasminogen activator (uPA) belongs also to the trypsin-like serine protease family but is not anchored to the cell membrane like the before mentioned hepsin and TMPRSS2 proteases. However, the uPA can bind to its specific uPA-receptor, which represent together the uPA-system (uPAS) that influences the ECM degradation, cell motility, proliferation and survival (1.4.1).[169],[170] Dysregulation of the uPA and overall uPAS, can lead to tumor progression, metastasis and angiogenesis of different cancer-types (1.4.2).[83],[169],[175] Therefore, the uPA presents a promising target for the design of inhibitors as anti-metastatic drugs. The current state of the uPA inhibitor development is represented mostly by (non-)covalent reversible or irreversible small molecule inhibitors, first and foremost by upamostat as the most promising drug candidate in clinical trials (1.4.4).[203]

Here, we describe the ligand-based development of peptidomimetic covalent reversible inhibitors with a α-ketobenzothiazole warhead. Starting with a previously discovered uPA inhibitor in project 4.1 and inspired by the potent serine protease inhibitor nafamostat, we substituted the P1 arginine residue with different bioisosteres and additionally modified the benzothiazole structure of the warhead moiety.[195] Thereby, the cyclohexyl- and phenylguanidine derivates represented the most promising inhibitors, regarding their pharmacodynamic properties.

In 2021, a high structure similarity was discovered between the uPA and TMPRSS2.[317] Therefore, we translated the abovementioned cyclohexyl- and phenylguanidine moiety of the uPA inhibitors to the previously published TMPRSS2 compounds in project 5.1, to enhance the selectivity of the peptidomimetic covalent reversible inhibitors.

Indeed, the replacement of the P1 arginine side chain by biosiosteres successfully led to nanomolar and even subnanomolar potent inhibitors towards uPA and TMPRSS2. Additionally, the general selectivity issue of trypsin-like serine proteases described in 1.7 could be overcome, displaying a promising selectivity profile against other structurally similar proteases (Figure 30).

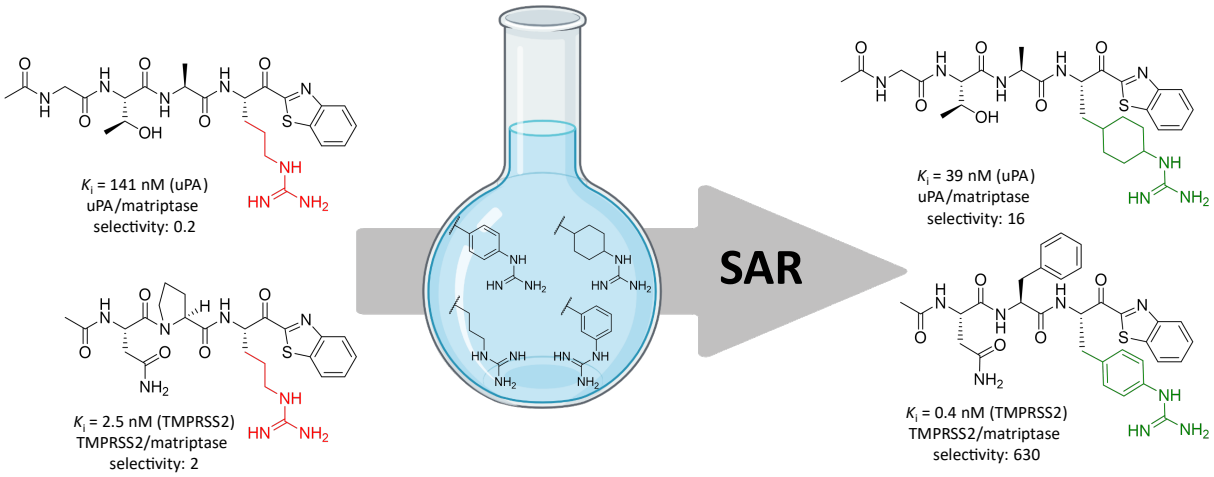


Figure 30: Ligand-based design of uPA and TMPRSS2 inhibitors, by combining previously published inhibitors (4.1 & 5.1) with arginine bioisosteres. Structure-activity relationship studies (SAR) resulting in optimized inhibitors with improved pharmacodynamic properties towards both target proteases. Figure created with Biorender.com.

Own contributions: Inhibitor synthesis (3,5, 9, 11, 12a, 13, 14a, 15, 21–45), inhibition assay and writing of the original draft plus editing of the manuscript.

Contributions of others: Inhibitor synthesis (6b–e, 10b–f, 12b–f, 14b–f, 17–20), parallel artificial membrane permeation assay (PAMPA), absorption spectra, calculation of physicochemical properties, calculation of protein similarity, stability studies and writing of the original draft plus editing of the manuscript.

Manuscript submitted to International Journal of Molecular Sciences (impact factor: 5.60).

5.3.2 Publication

The following publication quoted within “” from page 172 to page 211 is the same as the manuscript cited on page 53.

“

International Journal of
Molecular Sciences

MDPI

Article

Ligand-Based Design of Selective Peptidomimetic uPA and TMPRSS2 Inhibitors with Arg Bioisosteres

Patrick Müller, Collin Zimmer, Ariane Frey, Gideon Holzmann, Annabelle Carolin Weldert and Tanja Schirmeister *

Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Staudinger Weg 5, D-55128 Mainz, Germany; muelpat@uni-mainz.de (P.M.); cozimmer@uni-mainz.de (C.Z.); arfrey@uni-mainz.de (A.F.); gholzman@students.uni-mainz.de (G.H.); anwelder@uni-mainz.de (A.C.W.)
* Correspondence: schirmei@uni-mainz.de; Tel.: +49-6131-39-25742

Abstract: Trypsin-like serine proteases are involved in many important physiological processes like blood coagulation and remodeling of the extracellular matrix. On the other hand, they are also associated with pathological conditions. The urokinase-pwlasminogen activator (uPA), which is involved in tissue remodeling, can increase the metastatic behavior of various cancer types when overexpressed and dysregulated. Another member of this protease class that received attention during the SARS-CoV 2 pandemic is TMPRSS2. It is a transmembrane serine protease, which enables cell entry of the coronavirus by processing its spike protein. A variety of different inhibitors have been published against both proteases. However, the selectivity over other trypsin-like serine proteases remains a major challenge. In the current study, we replaced the arginine moiety at the P1 site of peptidomimetic inhibitors with different bioisosteres. Enzyme inhibition studies revealed that the phenylguanidine moiety in the P1 site led to strong affinity for TMPRSS2, whereas the cyclohexylguanidine derivate potently inhibited uPA. Both inhibitors exhibited high selectivity over other structurally similar and physiologically important proteases.

Keywords: trypsin-like serine proteases; covalent reversible inhibitors; enzyme inhibition study; protease inhibitors; peptidomimetic sequence; arginine bioisosteres

Citation: Müller, P.; Zimmer, C.; Frey, A.; Holzmann, G.; Weldert, A.C.; Schirmeister, T. Ligand-Based Design of Selective Peptidomimetic uPA and TMPRSS2 Inhibitors with Arg Bioisosteres. *Int. J. Mol. Sci.* **2024**, *25*, 1375. https://doi.org/10.3390/ijms25031375

Academic Editor: Asim Debnath

Received: 3 January 2024
Revised: 20 January 2024
Accepted: 21 January 2024
Published: 23 January 2024

1. Introduction

With over 600 different proteins, proteases represent an important class of enzymes [1]. Approximately one-third of all known proteolytic enzymes are serine proteases [2]. According to the MEROPS database of peptidases, these enzymes are classified into clans by their catalytic mechanism and into families on the basis of a common ancestry [3]. The largest family of serine proteases are the trypsin-like proteases (TLPs). The catalytic triad of TLPs harbors a nucleophilic serine residue in combination with aspartate and histidine, which increase the nucleophilicity of the serine. The trypsin-like substrate specificity is characterized by the positively charged side chain of arginine or lysine in the P1 position [3,4]. Numerous important physiological processes rely on trypsin-like serine proteases. This includes hemostasis, the immune response system and extracellular matrix remodeling [5–8]. Dysregulation of these enzymes can lead to severe pathological incidents, which range from cardiovascular disorders to cancer progression or neurodegenerative and inflammation processes [8–10]. Moreover, proteases often are virulence factors in infectious diseases. As an example, tropical and subtropical countries are heavily affected by dengue virus infections, where the viral NS2B-NS3 trypsin-like serine protease is essential for the replication process of the virus [11]. Undoubtedly, this class of enzymes includes promising targets in various diseases, and the scientific community still strives to discover more drug candidates [7].

The urokinase-type plasminogen activator (uPA) is one member of the trypsin-like serine proteases. The enzyme is involved in the fibrinolytic system [12]. The binding of

uPA to its specific glycolipid-anchored uPA receptor (uPAR) on cell surfaces enables the conversion of plasminogen to the serine protease plasmin [13]. This mediates extracellular proteolysis and the activation of several further proteases, like activating growth factors and metalloproteases, which catalyze the degradation and remodeling of extracellular matrix components [14,15]. Unfortunately, pathophysiological mechanisms like tumor angiogenesis, tumor progression and metastasis profit from these events, and therefore, inhibition of this protease could be beneficial for the mitigation, or even prevention, of tumor proliferation (Figure 1 left side) [14,15]. Blocking of the catalytic activity was achieved by specific antibodies, overexpression of the endogenous inhibitors PAI-1 and small-molecule inhibitors [16–18]. One of the most promising peptidomimetic inhibitors, mesupron® (upamostat, WX-671, RHB-107, Wilex AG, Heidelberg, Germany), led to reduced metastasis and extended lifespan in clinical trials on pancreatic and breast cancer patients [19]. Hence, uPA can be considered as a promising drug target to block tumor dissemination.

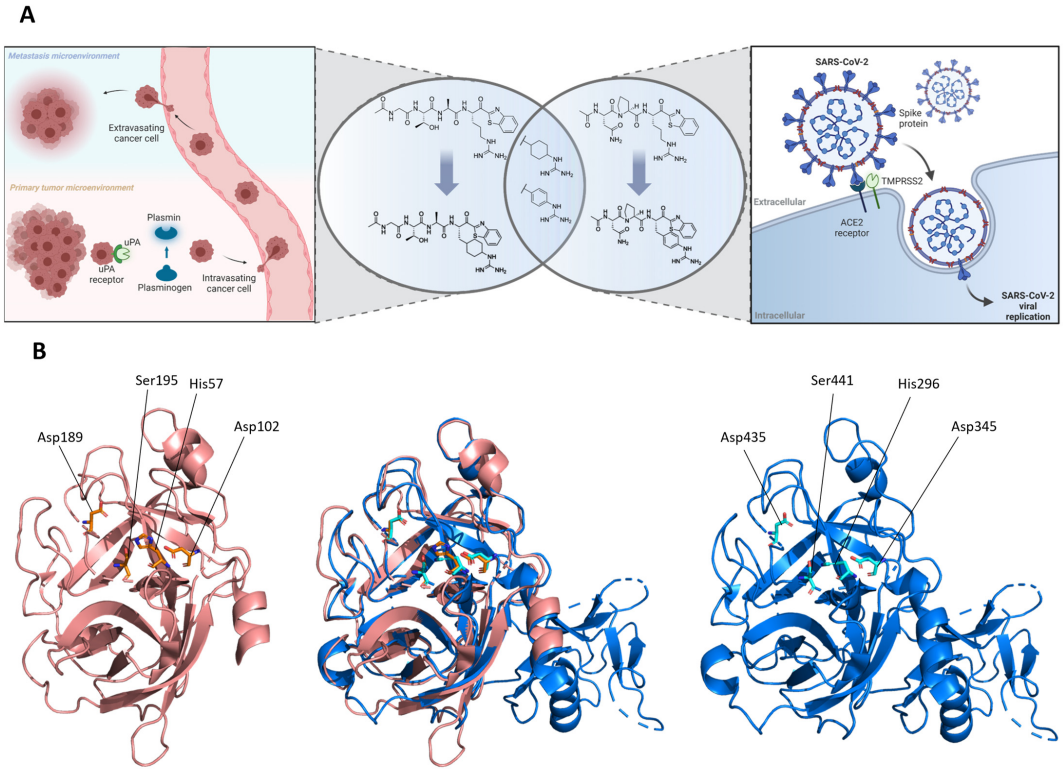


Figure 1. (**A**) Ligand-based design of covalent-reversible uPA (**left**) and TMPRSS2 (**right**) inhibitors with arginine bioisosteres. (**A**) Created with BioRender.com. (**B**) Visualization of the uPA structure (red, PDB: 7VM4) containing the catalytic triad (Asp102, His57, Ser195) and Asp189 in the S1 pocket, TMPRSS2 (blue, PDB: 7MEQ) containing the catalytic triad (Asp345, His296, Ser441) and Asp435 in the S1 pocket and the superimposition of both proteases (red/blue). (**B**) Created with PyMOL (Version 2.4.0, Schrödinger, LLC, New York, NY, USA).

Another proteolytic enzyme that belongs to the trypsin-like serine proteases is the human transmembrane protease serine subtype 2 (TMPRSS2). It has been shown to play an important role for viral host cell entry, and received increased attention during the SARS-CoV-2 pandemic due to its ability to enable cell entry and spread of the coronaviruses SARS-CoV-2, SARS-CoV and MERS-CoV [20–24]. The entry of these viruses is mediated by

the spike protein, which is located at the viral cell surface. TMPRSS2 processes the spike protein after binding of the virus to the angiotensin-converting enzyme 2 receptor (ACE2), initiating the entry into lung cells (Figure 1 right side) [23,25]. Additionally, viral cell entry can occur via the endosomal pathway, whereby the spike protein is processed by cathepsin L [26]. Studies have demonstrated that inhibition of TMPRSS2 blocks the viral host cell entry and replication of SARS-CoV-2 in lung epithelial Calu-3 cells [27,28]. Previous work, in cooperation with Mailänder et al. showed that peptidomimetic inhibitors efficiently reduce TMPRSS2 activity, block SARS-CoV-2 spike-driven entry and prevent SARS-CoV-2 infection in CaCo-2 cells [29]. This highlights the opportunity for an alternative therapeutic strategy, besides targeting of the viral host proteases papain-like protease (PL^{pro}) and the 3C-like- or "main protease" (3CL- or M^{pro}) [30–32].

In the past decades, several uPA inhibitors have been disclosed, most of them with non-covalent reversible or covalent-irreversible inhibition mode [16,33]. On the contrary, only few covalent-reversible inhibitors are found in the literature [34]. Such inhibitors could combine the benefits from both concepts: the high-affinity properties and extended residence time by covalent modification of the catalytic serine residue and the reduced risk for unwanted side effects and toxicity by a reversible binding mechanism [35–37]. Furthermore, in order to minimize the risk for side effects, it is of great importance to inhibit the target protease selectively. This, however, is a major challenge due to the high structural similarity within the trypsin-like serine protease family.

In 2021, the group around Huang et al. created a homology structure model of the TMPRSS2 serine protease domain, and revealed a high similarity between the homology model and the structure of the uPA [38]. This led to the idea to transfer the design of the synthesized uPA inhibitors to the previously published TMPRSS2 inhibitors, to receive an improved set of inhibitors in terms of off-target selectivity (Figure 1) [29].

Herein, we describe the ligand-based development of peptidomimetic inhibitors, which started with Ac-Gly-L-Thr-L-Ala-L-Arg-ketobenzothiazole (kbt) as a covalent-reversible uPA inhibitor discovered in previous work [39]. We substituted the P1-arginine moiety with a variety of bioisosteres, inspired by the serine protease inhibitor camostat, and furthermore modified the benzothiazole structure [40,41]. The cyclohexyl-and phenylguanidine moiety presented the most promising results during the enzyme inhibition studies. Therefore, we translated this structure motif to the suitable peptide sequence Ac-L-Asn-L-Pro-L-Arg-kbt from our previous work towards TMPRSS2 [29]. Within this study, we successfully enhanced the affinity and selectivity for both main-target proteases by systematic variation of different structural elements.

2. Results

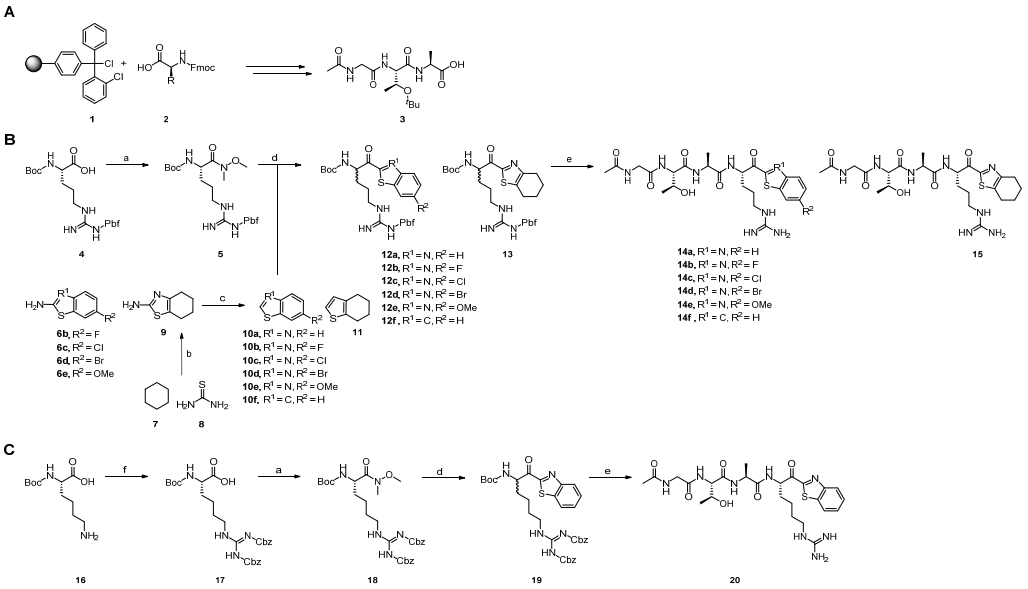
2.1. Chemistry

All tested peptidomimetic inhibitors were synthesized in multistep reactions. First, the peptide sequences (P2–P4) of the inhibitors were prepared via a standard fmoc solid phase peptide synthesis (SPPS) protocol, which is described in detail in the Supporting Information. The P1 derivatives with the ketobenzothiazole moiety as warhead were prepared as described in Schemes 1 and 2.

2.1.1. Synthesis of the (Homo)arginine-Based Inhibitors

Boc-protected N_ω-2,2,4,6,7-pentamethyl-dihydrobenzofuran-5-sulfonyl (pbf)-L-arginine **4**, which was used as the starting material for the arginine-based inhibitors **14a**–**f** and **15**, was modified to the Weinreb amide **5**. The ketobenzothiazole derivatives **12a**–**f** and **13** were obtained by alkylation of **5** with the respective heterocycles **10a**–**f** and **11**. The benzothiazole **10a** and the benzothiophene **10f** were commercially available, whereas the 6-fluoro-, 6-chloro-, 6-bromo-, 6-methoxybenzothiazoles **10b**–**e** and 4,5,6,7-tetrahydrothiazole **11** had to be synthesized by desamination of the commercially available 2-amino precursors **6b**–**e** and **9**. The 2-amino-4,5,6,7-tetrahydrothiazole **9** was prepared from cyclohexane **7** and thiourea **8** with iodine. The preparation of the homoarginine inhibitor **20** started with

the guanylation of boc-protected L-lysine **16** with *N*,*N*′-bis-(carbobenzoxy)-1-*H*-pyrazole-1-carboxamidine, yielding compound **17**. Afterwards, **17** was converted to the ketobenzothiazole **19**, in analogy to the arginine derivatives. After boc-deprotection of the amino group, the P1 precursor derivatives were coupled with the Ac-Gly-L-Thr(O^{t}Bu)-L-Ala-OH peptide **3** using 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo [4,5-b]pyridinium 3-oxide hexafluorophosphate (HATU) as the coupling reagent. Final deprotection of the (homo)arginine and threonine side chain under acidic conditions and purification via RP-HPLC yielded the inhibitors **14a–f**, **15** and **20**.



Scheme 1. Synthesis of the (homo)arginine-based inhibitors **14a–f**, **15**, **20**. (**A**) Solid phase peptide synthesis of peptide sequence **3**. (**B**) Preparation of the final arginine compounds **14a–f**, **15**. (**C**) Synthesis of the homoarginine inhibitor **20**. Reaction conditions: (a) *N*,*O*-dimethylhydroxylamine·HCl, TBTU, DIEA, DCM, rt, 12 h, 70–90%; (b) cyclohexane, thiourea, iodine, 130 °C, 12 h, 53%; (c) isopentyl nitrite, THF, reflux, 2–5 h, 46–76%; (d) benzothiazole-derivative, *n*-BuLi, THF, −78 °C, 2 h, 50–88%; (e) 1. TFA, DCM, rt, 0.5 h, 2. HATU, DIEA, DMF, DCM, rt, 12 h, 3. TFA, DCM, rt, 2 h, 4–10%; (f) *N*,*N*′ bis-(carbobenzoxy)-1-*H*-pyrazole-1-carboxamidine, Et_3N, DMF, rt, 72 h, quant.

2.1.2. Synthesis of the Phenyl/Cyclohexylguanidine-Based Inhibitors

The preparation of the *p*-phenyl- and *p*-cyclohexylguanidine-based inhibitors both started with boc-protected *p*-nitro-L-phenylalanine **24**. The reduction of the nitro group was carried out with 5% Pd/C in methanol to yield **25**, whereas the hydrogenation of the benzene ring and the nitro group using the Adam's catalyst under acidic conditions yielded the cyclohexane derivative **34**. The amine group of both compounds was guanylated with *N*,*N*′-bis-(carbobenzoxy)-1-*H*-pyrazole-1-carboxamidine. The bis-cbz-protected intermediates (**26**, **35**) were converted, in a similar way to the arginine-based inhibitors (Section 2.1.1), to the Weinreb amides **27** and **36** and later to the ketobenzothiazole derivatives **28** and **37.** After removal of the boc-protecting groups, the respective peptide sequences **3**, **21–23**, which were synthesized via a standard fmoc solid-phase synthesis (SPPS), were coupled with the *p*-phenyl- and cyclohexylguanidine precursor derivatives. After final deprotection of the side chains in TFA/DCM and purification via RP-HPLC, the inhibitors **29–33** and **38–39** were obtained. Starting with *m*-nitro-L-phenylalanine **40**, the inhibitor **45** was prepared in analogy to the 5-step synthetic process of the *p*-phenylguanidine derivatives **29–33**.

The synthesized final compounds **14a–f**, **15**, **20**, **29–33**, **38–39**, **45** showed two peaks with identical m/z ratio and similar retention times in initial chromatographic analyses. This is due to the partial epimerization of the α-carbon in the P1 amino acid portion during the reaction of the Weinreb amide with lithium-benzothiazole solution. Since the faster eluting epimer was always isolated via RP-HPLC in very large excess, while the other diastereomer was obtained only in traces, we supposed the first one to be the L-epimer, and used it for all inhibition studies [42].

Scheme 2. Synthesis of the phenyl/cyclohexylguanidine-based inhibitors **29–33**, **38–41**, **45**. (**A**) Solid phase peptide synthesis of peptide sequences **3**, **21–23**. (**B**) Preparation of the final *p*-phenyl/cyclohexylguanidine compounds **29–33**, **38–41**. (**C**) Synthesis of the *m*-phenylguanidine inhibitor **45**. Reaction conditions: (a) 1. Pd/C 5%, H_2 (3 bar), MeOH, 3 h, (b) PtO_2, H_2 (3 bar), AcOH, MeOH, rt, 24 h, (c) *N*,*N*′-bis-(carbobenzoxy)-1-*H*-pyrazole-1-carboxamidine, Et_3N, DMF, rt, 6 h, 90%–quant.; (d) *N*,*O*-dimethylhydroxylamine·HCl, TBTU, DIEA, DCM, rt, 12 h, 51–53%; (e) benzothiazole, *n*-BuLi, THF, −78 °C, 2 h, 64–84%, (f) 1. TFA, DCM, rt, 0.5 h, 2. HATU, DIEA, DMF, DCM, rt, 12 h, 3. TFA, DCM, rt, 2 h, 10–40%.

2.2. Enzyme Inhibition Studies

The inhibitory activity of the synthesized compounds towards the respective main- and off-target proteases was measured via fluorometric and colorimetric assays. Thus, fluorogenic AMC- or colorimetric *p*NA-based substrates with a peptide sequence suitable for the tested protease were utilized (see Supplementary Figure S2). At first, the compounds were screened against five proteases (uPA, TMPRSS2, matriptase, tPA, thrombin, factor Xa) at 20 µM, and a cut-off value of 80% inhibition at this concentration was set, for the differentiation between nonactive (n.a.) and active inhibitors. Due to the reversible inhibition mechanism of the ketobenzothiazole derivates, the IC_{50} values were determined with Graphpad Prism 9, and afterwards converted to the corresponding K_i values for an adequate comparison between the inhibitory activities of the compounds toward all tested proteases. The K_i values were calculated using the Cheng–Prusoff equation [43].

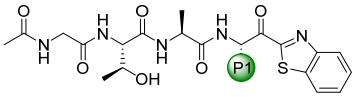
2.2.1. Inhibition Studies with uPA Inhibitors

At first, we investigated the selectivity profile of the starting compound **14a,** which exhibited good inhibition of uPA with a K_i value of 141 nM. **14a** was originally synthesized for the analysis of reactivity and selectivity studies of peptidomimetic covalent inhibitors [40]. The peptide sequence Ac-L-Gly-L-Thr-L-Ala-L-Arg was used, because of its literature-known selectivity for uPA vs. tPA [44]. Due to the similar and important physiological roles of uPA and tPA, a good selectivity is necessary to avoid severe side effects concerning ECM degradation and cell proliferation [5,45,46]. Furthermore, the trypsin-like serine proteases thrombin and factor Xa were chosen because of their important roles in blood coagulation, as well as matriptase as a representative of a transmembrane protease, which is involved in the remodeling of plasma membranes and other lipid matrix formations [47–49]. Due to their structural similarity (calculated sequence similarity is given in Tables 1–3) to uPA and their physiological roles, they resemble important off-targets. As expected, **14a** did not show inhibition of tPA, and only moderate selectivities for uPA towards thrombin (**14a** K_i = 4390 nM) and factor Xa (**14a** K_i = 3360 nM) with inhibition constants in the low micromolar range. In contrast, a lower K_i value was obtained for mapriptase (**14a** K_i = 32 nM). Exchanging the arginine side chain with a *p*-phenyl- or cyclohexylguanidine moiety enhanced the inhibitory properties. Both derivates resulted in more affine inhibitors (**29** K_i = 29 nM, **38** K_i = 39 nM), with a significant improvement in their selectivity profiles. The inhibitors **29**, **38** did not inhibit tPA, thrombin and factor Xa, and the selectivity indices for matriptase (**29** K_i = 132 nM, **38** K_i = 626 nM) were improved. The inhibitors **20** and **45**, which contain the homoarginine and *m*-substituted phenylguanidine moiety, did not show inhibition of all tested proteases at 20 µM, which highlights the importance of the alkyl chain length and the *p*-position of the guanidine element for proper binding into the S1 pocket. Additionally, all compounds were tested against the TMPRSS2 because of the aforementioned structure similarity to uPA [38]. The results indicated a strong affinity to the TMPRSS2 protease with K_i values in the nanomolar range of the arginine, phenyl- and cyclohexyl derivates (**14a** K_i = 5 nM, **29** K_i = 10 nM, **38** K_i = 73 nM). Based on these results, a SAR study with the phenyl- and cyclohexylguanidine moiety as arginine bioisosteres for new TMPRSS2 inhibitors was performed, which is described in Section 2.2.2 [29].

Besides the arginine replacement in the P1 position, we also evaluated the influence of modifications of the benzothiazole moiety (cpds. **14b–f**, **15**). The introduction of the electronegative halogen atoms fluorine, chlorine and bromine in position 6 led to an approximately two-fold increase in the affinity for the chloro- and bromo-derivates (**14c** K_i = 82 nM, **14d** K_i = 60 nM), and a three-fold loss of affinity for the fluoro-derivate (**14b** K_i = 388 nM). Other modifications, like the electron-donating methoxy group in position 6 (**14e** K_i = 178 nM), the exchange of the benzene ring system with a cyclohexyl ring (**15** K_i = 435 nM) or the replacement of the benzothiazole with a benzothiophene ring, led to a decrease in or complete loss of the affinity towards the uPA. Selectivity studies were performed with the chloro- and bromo-derivates **14c–d**, because they were

the only inhibitors with slightly better affinity than the nonsubstituted ketobenzothiazole inhibitor **14a**. They revealed similar affinity to TMPRSS2 (**14c** K_i = 9 nM, **14d** K_i = 6 nM) and matriptase (**14c** K_i = 59 nM, **14d** K_i = 38 nM), and reduced selectivity vs. thrombin (**14c** K_i = 456 nM, **14d** K_i = 450 nM) and factor Xa (**14c** K_i = 2447 nM, **14d** K_i = 2847 nM) in comparison to **14a**.

Table 1. Inhibition data (K_i values and selectivity indices [SI]) of the synthesized uPA inhibitors **14a**, **20**, **29**, **38**, **45** towards uPA, TMPRSS2, matriptase, tPA, thrombin and factor Xa.



14a, 20, 29, 38, 45

Compound		K_i [nM] [SI]					
P1		uPA	TMPRSS2 55.6%	Matriptase 68.9%	tPA 66.7	Thrombin 60.0%	Factor Xa 62.2%
	**14a**	141 ± 28	5 ± 1 [0.04]	32 ± 14 [0.2]	n.a.	4390 ± 1480 [31]	3360 ± 320 [24]
	**29**	29 ± 2	10 ±4 [0.3]	133 ± 3 [4]	n.a.	n.a.	n.a.
	**38**	39 ± 5	73 ± 16 [1.8]	626 ± 74 [16]	n.a.	n.a.	n.a.
	**20**	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
	**45**	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

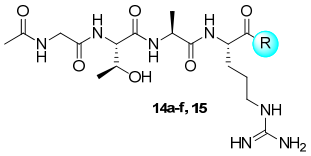
Selectivity indices [SI] represent the quotient of the K_i values for TMPRSS2, matriptase, tPA, thrombin and factor Xa to the K_i value for uPA; n.a. not active, i.e., <80% inhibition at 20 μM; calculated sequence similarity between the binding sites of uPA and the other proteases are given in %.

2.2.2. Inhibition Studies with TMPRSS2 Inhibitors

The selection of the peptide sequence Ac-L-Asn-L-Pro-L-Arg was based on the results of previous work. The published inhibitor Ac-L-Asn-L-Pro-L-Arg-kbt showed K_i values in the single-digit nanomolar range (K_i = 2.5 nM) and a good selectivity vs. thrombin (K_i = 1046 nM) [29]. Unfortunately, only slight selectivity could be observed over factor Xa (K_i = 41.1 nM), and almost no difference in inhibition potency between TMPRSS2 and matriptase (K_i = 5.2 nM). Therefore, we tried to improve the selectivity profile by substituting the arginine side chain in the P1 position with the previously used phenyl- and cyclohexylguanidine moiety. Both derivates **30**, **39** showed an increase in selectivity for TMPRSS2 towards matriptase, with the phenylguanidine-based compound being more affine for TMPRSS2 (**30** K_i = 5 nM) than the cyclohexyl derivate (**39** K_i = 44 nM), but also showing a better inhibition of the matriptase for the phenylguanidine derivate (**30** K_i = 60 nM, **39** K_i = 1198 nM). **30**–**33** and **39** did not inhibit tPA, thrombin and factor Xa. In addition, **30** and **39** showed a moderate selectivity for TMPRSS2 over uPA (**30** K_i = 479 nM, **39** K_i = 936 nM). Based on these results, and due to the overall good affinity and selectivity parameters, we decided to maintain the phenylguanidine moiety in the P1 position and implement P2 modifications with phenyl- and cyclohexylalanine (Phe, Cha) instead of proline. The latter is based on results obtained with hepsin inhibitors from the group of

Kwon et al. [40]. The inhibitor **31** with the P2 phenylalanine residue showed inhibition of TMPRSS2 in the subnanomolar range (**31** K_i = 0.4 nM) and a significant increase in selectivity over matriptase (**31** K_i = 252 nM) and uPA (**31** K_i = 3574 nM). The inhibitor **32** with the cyclohexylalanine residue in P2 position also showed very good selectivity over matriptase (**32** K_i = 3333 nM) and uPA (**32** K_i = 2688 nM), but less affinity to TMPRSS2 (**32** K_i = 34 nM). In an attempt to improve the drug-like properties of the designed inhibitors, we synthesized the shortened compound **33**. This led to a slightly less active TMPRSS2 inhibitor, but still in the low nanomolar range (**33** K_i = 5 nM). The selectivity profile for TMPRSS2 inhibition over matriptase (**33** K_i = 1443 nM) and uPA (**33** K_i = 5264 nM) is still very promising.

Table 2. Inhibition data (K_i values and selectivity indices [SI]) of the synthesized uPA inhibitors **14a–f**, **15** towards uPA, TMPRSS2, matriptase, tPA, thrombin and factor Xa.



Compound		K_i [nM] [SI]					
R		uPA	TMPRSS2 55.6%	Matriptase 68.9%	tPA 66.7	Thrombin 60.0%	Factor Xa 62.2%
	**14a**	141 ± 28	5 ± 1 [0.04]	32 ± 14 [0.2]	n.a.	4390 ± 1480 [31]	3360 ± 320 [23]
	**14b**	388 ± 95	-	-	-	-	-
	**14c**	82 ± 6	9 ± 1 [0.1]	59 ± 9 [0.7]	n.a.	456 ± 72 [5]	2447 ± 99 [30]
	**14d**	60 ± 15	6 ± 1 [0.1]	38 ± 2 [0.6]	n.a.	450 ± 69 [8]	2847 ± 844 [47]
	**14e**	178 ± 19	-	-	-	-	-
	**14f**	n.a.	-	-	-	-	-
	**15**	435 ± 33	-	-	-	-	-

Selectivity indices [SI] represent the quotient of the K_i values for TMPRSS2, matriptase, tPA, thrombin and factor Xa to the K_i value for uPA; n.a., not active, i.e., <80% inhibition at 20 µM; "-", not tested; calculated sequence similarity between the binding sites of uPA and the other proteases are given in %.

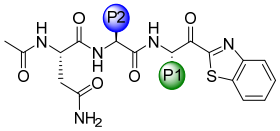
2.3. Parallel Artificial Membrane Permeation Assay (PAMPA)

Following the envisioned applications of the presented inhibitors as drug leads for the treatment of cancer or viral infections, cell permeability is an important factor in the characterization process. Since both main targets (uPA and TMPRSS2) are membrane-located, extracellular structures, inhibitors seemingly do not require cell permeation to address their target. However, in the organismic contexts of oral bioavailability (facilitated application) and biodistribution (reaching target tissue), adequate permeation is an important quality. To assess this characteristic, PAMPA was used as a suitable model for passive permeation.

Generally, the inhibitor scaffold combines some favorable features: The Arg-like P1 amino acid (in combination with other hydrophilic amino acids like Asp and Thr)

ensures high aqueous solubility, even in the presence of the rather hydrophobic benzoheteroarenes. The latter motif conveys reliable detectability by spectroscopy-based methods (λ_{max} = 305–350 nm, depending on substitution pattern). The inhibitors also were found to be sufficiently stable in the utilized aqueous system (50 mM TRIS, pH = 7.4) over the course of the assay (7 h at room temperature) and under elevated temperature conditions (17 h at 37 °C). Computed physicochemical properties, absorption spectra and stability studies are depicted in Supplementary Figures S4 and S5 and Table S1.

Table 3. Inhibition data (K_i values and selectivity indices [SI]) of the synthesized TMPRSS2 inhibitors **30**–**33**, **39** towards TMPRSS2, uPA, matriptase, tPA, thrombin and factor Xa.



30-33, 39

Compound			K_i [nM] [SI]					
P2	**P1**		**TMPRSS2**	**uPA 53.2%**	**Matriptase 66.7%**	**tPA 62.2%**	**Thrombin 57.8%**	**Factor Xa 66.7%**
Pro	Arg [31]		2.5	-	5.2 [2]	-	1046 [418]	41.1 [16]
Pro		**30**	5 ± 1	479 ± 17 [96]	60 ± 8 [12]	n.a.	n.a.	n.a.
Pro		**39**	44 ± 7	936 ± 100 [21]	1198 ± 181 [27]	n.a.	n.a.	n.a.
Phe		**31**	0.4 ± 0.1	3574 ± 421 [>5000]	252 ± 67 [630]	n.a.	n.a.	n.a.
Cha		**32**	34 ± 4	2688 ± 217 [79]	3333 ± 1141 [98]	n.a.	n.a.	n.a.
		**33**	5 ± 2	5264 ± 922 [1052]	1443 ± 90 [289]	n.a.	n.a.	n.a.

Selectivity indices [SI] represent the quotient of the K_i values for uPA, matriptase, tPA, thrombin and factor Xa to the K_i value for TMPRSS2; n.a., not active, i.e., <80% inhibition at 20 µM; "-", not tested; calculated sequence similarity between the binding sites of TMPRSS2 and the other proteases are given in %.

However, all presented compounds were found to have very low permeabilities ($P_e < 1 \times 10^{-6}$ cm/s) without any indication of improvement between the structural modifications (as exemplified for **38** in Figure 2). This result is not surprising. The pK_a (of the protonated guanidine function) of all compounds is calculated to be ≥10 (Marvin JS 23.11.0), meaning that in assay (or physiologic) conditions, all compounds are expected to be fully (≥99.75%) protonated, and therefore remarkably hydrophilic. Most of the presented compounds have negative $logD_{7.4}$ values, with **33** being the exception ($logD_{7.4}$ = 1.2; compare Supplementary Table S1). This level of lipophilicity, however, was still not enough to exert measurable permeability. For approved drugs with similar structural characteristics (e.g., camostat, melagatran, xylometazoline, metformin), only very limited permeabilities are described as well [50–53]. All this indicates the pronounced hindering effect of the guanidine group for passive permeation.

The discussed properties of the presented compounds can be paralleled to BCS class III compounds, namely their high aqueous solubility and low permeability. For these types of drugs, one major option to improve permeability is to remove charge from the molecule. In amidine-containing drugs, where charge is almost pH-independent due to their immense basicity, this was addressed by conversion to the amidoxime (ximelagatran or mesupron®) or carbamate prodrugs (dabigatran) with lower basicities [19,50,54]. For the guanidine moiety, the conversion to *N*-hydroxyguanidine is possible [55]. In a technological approach

to improved absorption, possible options for oral application are the formulation with permeation enhancing agents, or lipophilic counter ions [56]. For intravenous applications, nanoparticular formulations can be applied (e.g., for doxorubicin or for protease inhibitors) [57,58]. Of course, combinations of both chemical and technological approaches should be employed for optimization.

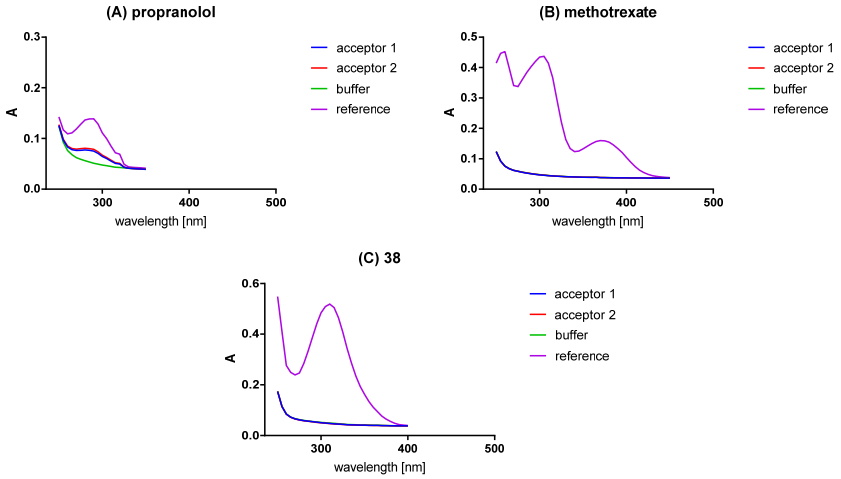


Figure 2. (**A**) Analysis of PAMPA permeable control propranolol HCl. (**B**) Analysis of PAMPA impermeable control methotrexate. (**C**) Analysis of PAMPA for **38**, representative of all synthesized final compounds of this study. All absorption spectra were baseline-corrected with the measured buffer spectrum at $\lambda \geq 450$ nm prior to AUC calculations. The reference spectra represent manually prepared samples, equivalent to a maximally permeated experimental sample. The acceptor spectra result from membrane permeation experiments and their AUC are proportional to the concentration of permeated compound.

3. Discussion

Trypsin-like serine proteases present attractive drug targets for treatment against many diseases, which can be of malignant cellular or viral origin [9,11]. Over the past decades many potent inhibitors were designed with remarkable affinity for the target protease. But most of them lack selectivity because of the highly structural similarity between the proteases. Within our study, we describe a systematic ligand-based approach to enhance affinities and selectivities. Starting from the previous published covalent reversible ketobenzothiazole inhibitor Ac-Gly-L-Thr-L-Ala-L-Arg-kbt **14a**, we modified the P1 arginine side chain with different bioisosteres [39]. The results indicate that the cyclohexylguanidine moiety fits best for uPA inhibition. The inhibitor **38** showed remarkable inhibition with a K_i value of 39 nM and a very good selectivity profile towards the other trypsin-like serine proteases. The modification of the benzothiazole moiety did not improve either the inhibitory properties nor the selectivity profiles, rendering the original kbt warhead the most promising.

The transfer of the P1-arginine replacement with the promising phenyl- and cyclohexylguanidine moieties to the previously published TMPRSS2 inhibitor Ac-L-Asn-L-Pro-L-Arg-kbt was a success, leading to a subnanomolar TMPRSS2 inhibitor **31** (K_i = 0.4 nM), with significantly increased selectivity over other trypsin-like serine proteases [29]. Furthermore, the shortened peptide sequence of the TMPRSS2 inhibitor **31** led to the more drug-like candidate **33,** with still very good inhibitory and selectivity properties. In terms of permeability, the inhibitor scaffold (and especially the shortened compound **33**) leaves the opportunity for improvement in a focused structure-permeability relationship study.

4. Materials and Methods

The materials as well as the methods used for this study are described in the Supporting Information. The authors have cited additional references within the Supporting Information [29,39,40,43,59–72]. Supplementary Figures of the protein similarity calculation (Figure S1), fluorometric inhibition assays (Figures S2 and S3), absorption spectra (Figure S4), stability studies (Figure S5), NMR-spectra and HPLC-chromatograms (Figures S6a–S21c) and Table S1 of the computation of physicochemical parameters can be accessed in the supporting information.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms25031375/s1.

Author Contributions: P.M.: design, synthesis and enzyme inhibition studies, writing—original draft; C.Z.: parallel artificial membrane permeation assay (PAMPA), absorption spectra, calculation of physicochemical properties, calculation of protein similarity, stability studies, writing—original draft; A.F.: synthesis and enzyme inhibition studies; G.H.: synthesis and enzyme inhibition studies, A.C.W.: protein preparation and purification; T.S.: validation, review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article and Supplementary Materials.

Acknowledgments: We thank Michael Klein for creating the graphical abstract and figures with BioRender.com. The authors wish to thank Torsten Steinmetzer for sharing the pQE-vector containing human matriptase and Simon Huber for sharing his expertise in matriptase 1 preparation.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

Ac	acetyl
Ac-Gly	acetylglycine
AcOH	acetic acid
Ala	alanine
AMC	7-amino-4-methyl coumarine
Asn	asparagine
Boc:	*t*-butyloxy carbonyl
Cbz	benzyloxycarbonyl
Cha	cyclohexylalanine
DCM	dichloromethane
DIEA	*N*,*N*-diisopropylethylamine
DMSO	dimethylsulfoxide
Et_3N	triethylamine
Fmoc	9-Fluorenylmethoxycarbonyl
h	hours
HATU	2-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate
MeOH	methanol
MERS-CoV	Middle East Respiratory Syndrome
n-BuLi	*n*-butyllithium
O^{t}Bu	*O*-*tert*-butyl
Pbf	N_ω-2,2,4,6,7-pentamethyl-dihydrobenzofuran-5-sulfonyl
Phe	phenylalanine
*p*NA	*para*-nitroanilide
Pro	proline
quant	quantitative
RP-HPLC	reversed phase high pressure liquid chromatography

rt	room temperature
SARS-CoV-2	severe acute respiratory syndrome coronavirus type 2
SPPS	solid phase peptide synthesis
TA	thioanisole
TBTU	2-(1*H*-Benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate
TFA	trifluoroacetic acid
THF	tetrahydrofurane
TLP	trypsin-like serine protease
TMPRSS2	transmembrane protease serine subtype 2
tPA	tissue-type plasminogen activator
trt	triphenylmethyl
uPA	urokinase-type plasminogen activator
uPAR	urokinase-type plasminogen activator receptor

References

1. López-Otín, C.; Overall, C.M. Protease degradomics: A new challenge for proteomics. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 509–519. [CrossRef] [PubMed]
2. Di Cera, E. Serine proteases. *IUBMB Life* **2009**, *61*, 510–515. [CrossRef] [PubMed]
3. Rawlings, N.D.; Morton, F.R.; Kok, C.Y.; Kong, J.; Barrett, A.J. MEROPS: The peptidase database. *Nucleic Acids Res.* **2007**, *36*, 320–325. [CrossRef] [PubMed]
4. Ma, W.; Tang, C.; Lai, L. Specificity of Trypsin and Chymotrypsin: Loop-Motion-Controlled Dynamic Correlation as a Determinant. *Biophys. J.* **2005**, *89*, 1183–1193. [CrossRef] [PubMed]
5. Lu, P.; Takai, K.; Weaver, V.M.; Werb, Z. Extracellular Matrix Degradation and Remodeling in Development and Disease. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, a005058. [CrossRef]
6. Oncul, S.; Afshar-Kharghan, V. The interaction between the complement system and hemostatic factors. *Curr. Opin. Hematol.* **2020**, *27*, 341–352. [CrossRef]
7. Ferguson, T.E.G.G.; Reihill, J.A.; Martin, S.L.; Walker, B. Novel Inhibitors and Activity-Based Probes Targeting Trypsin-Like Serine Proteases. *Front. Chem.* **2022**, *10*, 782608. [CrossRef]
8. Yaron, J.R.; Zhang, L.; Guo, Q.; Haydel, S.E.; Lucas, A.R. Fibrinolytic Serine Proteases, Therapeutic Serpins and Inflammation: Fire Dancers and Firestorms. *Front. Cardiovasc. Med.* **2021**, *8*, 648947. [CrossRef]
9. Eatemadi, A.; Aiyelabegan, H.T.; Negahdari, B.; Mazlomi, M.A.; Daraee, H.; Daraee, N.; Eatemadi, R.; Sadroddiny, E. Role of protease and protease inhibitors in cancer pathogenesis and treatment. *Biomed. Pharmacother.* **2017**, *86*, 221–231. [CrossRef]
10. Wu, Q.; Kuo, H.-C.; Deng, G.G. Serine proteases and cardiac function. *Biochim. Biophys. Acta-Proteins Proteom.* **2005**, *1751*, 82–94. [CrossRef]
11. Nitsche, C.; Holloway, S.; Schirmeister, T.; Klein, C.D. Biochemistry and Medicinal Chemistry of the Dengue Virus Protease. *Chem. Rev.* **2014**, *114*, 11348–11381. [CrossRef] [PubMed]
12. Dreymann, N.; Wuensche, J.; Sabrowski, W.; Moeller, A.; Czepluch, D.; Van, D.V.; Fuessel, S.; Menger, M.M. Inhibition of Human Urokinase-Type Plasminogen Activator (uPA) Enzyme Activity and Receptor Binding by DNA Aptamers as Potential Therapeutics through Binding to the Different Forms of uPA. *Int. J. Mol. Sci.* **2022**, *23*, 4890–4912. [CrossRef] [PubMed]
13. Mahmood, N.; Mihalcioiu, C.; Rabbani, S.A. Multifaceted Role of the Urokinase-Type Plasminogen Activator (uPA) and Its Receptor (uPAR): Diagnostic, Prognostic, and Therapeutic Applications. *Front. Oncol.* **2018**, *8*, 24. [CrossRef] [PubMed]
14. Mekkawy, A.H.; Pourgholami, M.H.; Morris, D.L. Involvement of Urokinase-Type Plasminogen Activator System in Cancer: An Overview. *Med. Res. Rev.* **2014**, *34*, 918–956. [CrossRef] [PubMed]
15. Masucci, M.T.; Minopoli, M.; Di Carluccio, G.; Motti, M.L.; Carriero, M.V. Therapeutic Strategies Targeting Urokinase and Its Receptor in Cancer. *Cancers* **2022**, *14*, 498. [CrossRef] [PubMed]
16. Buckley, B.J.; Aboelela, A.; Minaei, E.; Jiang, L.X.; Xu, Z.; Ali, U.; Fildes, K.; Cheung, C.-Y.; Cook, S.M.; Johnson, D.C.; et al. 6-Substituted Hexamethylene Amiloride (HMA) Derivatives as Potent and Selective Inhibitors of the Human Urokinase Plasminogen Activator for Use in Cancer. *J. Med. Chem.* **2018**, *61*, 8299–8320. [CrossRef] [PubMed]
17. Ma, D.; Gerard, R.D.; Li, X.-Y.; Alizadeh, H.; Niederkorn, J.Y. Inhibition of Metastasis of Intraocular Melanomas by Adenovirus-Mediated Gene Transfer of Plasminogen Activator Inhibitor Type 1 (PAI-1) in an Athymic Mouse Model. *Blood* **1997**, *90*, 2738–2746. [CrossRef] [PubMed]
18. Ossowski, L.; Russo-Payne, H.; Wilson, L.E. Inhibition of Urokinase-type Plasminogen Activator by Antibodies: The Effect on Dissemination of a Human Tumor in the Nude Mouse. *Cancer Res.* **1991**, *51*, 274–281.
19. Schmitt, M.; Harbeck, N.; Brünner, N.; Jänicke, F.; Meisner, C.; Mühlenweg, B.; Jansen, H.; Dorn, J.; Nitz, U.; Kantelhardt, E.J.; et al. Cancer therapy trials employing level-of-evidence-1 disease forecast cancer biomarkers uPA and its inhibitor PAI-1. *Expert Rev. Mol. Diagn.* **2011**, *11*, 617–634. [CrossRef]
20. Leow, M.K.-S. Correlating Cell Line Studies With Tissue Distribution of DPP4/TMPRSS2 and Human Biological Samples May Better Define the Viral Tropism of MERS-CoV. *J. Infect. Dis.* **2013**, *208*, 1350–1351. [CrossRef]

21. Bertram, S.; Heurich, A.; Lavender, H.; Gierer, S.; Danisch, S.; Perin, P.; Lucas, J.M.; Nelson, P.S.; Pöhlmann, 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*, e35876. [CrossRef] [PubMed]
22. Simmons, G.; Zmora, P.; Gierer, S.; Heurich, A.; Pöhlmann, S. Proteolytic activation of the SARS-coronavirus spike protein: Cutting enzymes at the cutting edge of antiviral research. *Antivir. Res.* **2013**, *100*, 605–614. [CrossRef] [PubMed]
23. Mahoney, M.; Damalanka, V.C.; Tartell, M.A.; Chung, D.H.; Lourenço, A.L.; Pwee, D.; Mayer Bridwell, A.E.; Hoffmann, M.; Voss, J.; Karmakar, P.; et al. A novel class of TMPRSS2 inhibitors potently block SARS-CoV-2 and MERS-CoV viral entry and protect human epithelial lung cells. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2108728118. [CrossRef] [PubMed]
24. Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T.S.; Herrler, G.; Wu, N.H.; Nitsche, A.; et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181*, 271–280.e8. [CrossRef]
25. Jackson, C.B.; Farzan, M.; Chen, B.; Choe, H. Mechanisms of SARS-CoV-2 entry into cells. *Nat. Rev. Mol. Cell Biol.* **2022**, *23*, 3–20. [CrossRef]
26. Zhao, M.M.; Yang, W.L.; Yang, F.Y.; Zhang, L.; Huang, W.J.; Hou, W.; Fan, C.F.; Jin, R.H.; Feng, Y.M.; Wang, Y.C.; et al. Cathepsin L plays a key role in SARS-CoV-2 infection in humans and humanized mice and is a promising target for new drug development. *Signal Transduct. Target. Ther.* **2021**, *6*, 134. [CrossRef]
27. Bestle, D.; Heindl, M.R.; Limburg, H.; Pilgram, O.; Moulton, H.; Stein, D.A.; Hardes, K.; Eickmann, M.; Dolnik, O.; Rohde, C.; et al. TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Sci. Alliance* **2020**, *3*, e1–e14. [CrossRef]
28. Li, F.; Han, M.; Dai, P.; Xu, W.; He, J.; Tao, X.; Wu, Y.; Tong, X.; Xia, X.; Guo, W.; et al. Distinct mechanisms for TMPRSS2 expression explain organ-specific inhibition of SARS-CoV-2 infection by enzalutamide. *Nat. Commun.* **2021**, *12*, 866. [CrossRef]
29. Wettstein, L.; Knaff, P.M.; Kersten, C.; Müller, P.; Weil, T.; Conzelmann, C.; Müller, J.A.; Brückner, M.; Hoffmann, M.; Pöhlmann, S.; et al. Peptidomimetic inhibitors of TMPRSS2 block SARS-CoV-2 infection in cell culture. *Commun. Biol.* **2022**, *5*, 681. [CrossRef]
30. Sanders, B.C.; Pokhrel, S.; Labbe, A.D.; Mathews, I.I.; Cooper, C.J.; Davidson, R.B.; Phillips, G.; Weiss, K.L.; Zhang, Q.; O'Neill, H.; et al. Potent and selective covalent inhibition of the papain-like protease from SARS-CoV-2. *Nat. Commun.* **2023**, *14*, 1733. [CrossRef]
31. Welker, A.; Kersten, C.; Müller, C.; Madhugiri, R.; Zimmer, C.; Müller, P.; Zimmermann, R.; Hammerschmidt, S.; Maus, H.; Ziebuhr, J.; et al. Structure-Activity Relationships of Benzamides and Isoindolines Designed as SARS-CoV Protease Inhibitors Effective against SARS-CoV-2. *ChemMedChem* **2021**, *16*, 340–354. [CrossRef]
32. Kincaid, J.R.; Caravez, J.C.; Iyer, K.S.; Kavthe, R.D.; Fleck, N.; Aue, D.H.; Lipshutz, B.H. A sustainable synthesis of the SARS-CoV-2 Mpro inhibitor nirmatrelvir, the active ingredient in Paxlovid. *Commun. Chem.* **2022**, *5*, 156. [CrossRef]
33. Joossens, J.; Ali, O.M.; El-Sayed, I.; Surpateanu, G.; Van der Veken, P.; Lambeir, A.-M.; Setyono-Han, B.; Foekens, J.A.; Schneider, A.; Schmalix, W.; et al. Small, Potent, and Selective Diaryl Phosphonate Inhibitors for Urokinase-Type Plasminogen Activator with In Vivo Antimetastatic Properties. *J. Med. Chem.* **2007**, *50*, 6638–6646. [CrossRef]
34. Zeslawska, E.; Jacob, U.; Schweinitz, A.; Coombs, G.; Bode, W.; Madison, E. Crystals of Urokinase Type Plasminogen Activator Complexes Reveal the Binding Mode of Peptidomimetic Inhibitors. *J. Mol. Biol.* **2003**, *328*, 109–118. [CrossRef]
35. Lonsdale, R.; Ward, R.A. Structure-based design of targeted covalent inhibitors. *Chem. Soc. Rev.* **2018**, *47*, 3816–3830. [CrossRef]
36. Baillie, T.A. Targeted Covalent Inhibitors for Drug Design. *Angew. Chem. Int. Ed.* **2016**, *55*, 13408–13421. [CrossRef]
37. De Cesco, S.; Kurian, J.; Dufresne, C.; Mittermaier, A.K.; Moitessier, N. Covalent inhibitors design and discovery. *Eur. J. Med. Chem.* **2017**, *138*, 96–114. [CrossRef]
38. Sun, G.; Sui, Y.; Zhou, Y.; Ya, J.; Yuan, C.; Jiang, L.; Huang, M. Structural Basis of Covalent Inhibitory Mechanism of TMPRSS2-Related Serine Proteases by Camostat. *J. Virol.* **2021**, *95*, 10–1128. [CrossRef]
39. Müller, P.; Meta, M.; Meidner, J.L.; Schwickert, M.; Meyr, J.; Schwickert, K.; Kersten, C.; Zimmer, C.; Hammerschmidt, S.J.; Frey, A.; et al. Investigation of the Compatibility between Warheads and Peptidomimetic Sequences of Protease Inhibitors—A Comprehensive Reactivity and Selectivity Study. *Int. J. Mol. Sci.* **2023**, *24*, 7226. [CrossRef]
40. Kwon, H.; Ha, H.; Jeon, H.; Jang, J.; Son, S.-H.; Lee, K.; Park, S.-K.; Byun, Y. Structure-activity relationship studies of dipeptide-based hepsin inhibitors with Arg bioisosteres. *Bioorg. Chem.* **2021**, *107*, 104521. [CrossRef]
41. Hoffmann, M.; Hofmann-Winkler, H.; Smith, J.C.; Krüger, N.; Arora, P.; Sørensen, L.K.; Søgaard, O.S.; Hasselstrøm, J.B.; Winkler, M.; Hempel, T.; et al. Camostat mesylate inhibits SARS-CoV-2 activation by TMPRSS2-related proteases and its metabolite GBPA exerts antiviral activity. *EBioMedicine* **2021**, *65*, 103255. [CrossRef] [PubMed]
42. 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*, 585–599. [CrossRef] [PubMed]
43. Ludewig, S.; Kossner, M.; Schiller, M.; Baumann, K.; Schirmeister, T. Enzyme Kinetics and Hit Validation in Fluorimetric Protease Assays. *Curr. Top. Med. Chem.* **2010**, *10*, 368–382. [CrossRef] [PubMed]
44. Li, C.Y.; de Veer, S.J.; Law, R.H.P.; Whisstock, J.C.; Craik, D.J.; Swedberg, J.E. Characterising the Subsite Specificity of Urokinase-Type Plasminogen Activator and Tissue-Type Plasminogen Activator using a Sequence-Defined Peptide Aldehyde Library. *ChemBioChem* **2019**, *20*, 46–50. [CrossRef] [PubMed]

45. Green, K.A.; Lund, L.R. ECM degrading proteases and tissue remodelling in the mammary gland. *BioEssays* **2005**, *27*, 894–903. [CrossRef] [PubMed]
46. Ortiz-Zapater, E.; Peiró, S.; Roda, O.; Corominas, J.M.; Aguilar, S.; Ampurdanés, C.; Real, F.X.; Navarro, P. Tissue Plasminogen Activator Induces Pancreatic Cancer Cell Proliferation by a Non-Catalytic Mechanism That Requires Extracellular Signal-Regulated Kinase 1/2 Activation through Epidermal Growth Factor Receptor and Annexin A2. *Am. J. Pathol.* **2007**, *170*, 1573–1584. [CrossRef] [PubMed]
47. Tanaka, K.A.; Key, N.S.; Levy, J.H. Blood Coagulation: Hemostasis and Thrombin Regulation. *Anesth. Analg.* **2009**, *108*, 1433–1446. [CrossRef]
48. Borensztajn, K.; Spek, C.A. Blood coagulation factor Xa as an emerging drug target. *Expert Opin. Ther. Targets* **2011**, *15*, 341–349. [CrossRef]
49. List, K.; Bugge, T.H.; Szabo, R. Matriptase: Potent Proteolysis on the Cell Surface. *Mol. Med.* **2006**, *12*, 1–7. [CrossRef]
50. Gustafsson, D.; Nyström, J.E.; Carlsson, S.; Bredberg, U.; Eriksson, U.; Gyzander, E.; Elg, M.; Antonsson, T.; Hoffmann, K.J.; Ungell, A.L.; et al. The Direct Thrombin Inhibitor Melagatran and Its Oral Prodrug H 376/95: Intestinal Absorption Properties, Biochemical and Pharmacodynamic Effects. *Thromb. Res.* **2001**, *101*, 171–181. [CrossRef]
51. Sibinovska, N.; Žakelj, S.; Trontelj, J.; Kristan, K. Applicability of RPMI 2650 and Calu-3 Cell Models for Evaluation of Nasal Formulations. *Pharmaceutics* **2022**, *14*, 369. [CrossRef]
52. Elezović, A.; Marić, A.; Biščević, A.; Hadžiabdić, J.; Škrbo, S.; Špirtović-Halilović, S.; Rahić, O.; Vranić, E.; Elezović, A. In vitro pH dependent passive transport of ketoprofen and metformin. *ADMET DMPK* **2020**, *9*, 57–68. [CrossRef]
53. Butnarasu, C.; Caron, G.; Pacheco, D.P.; Petrini, P.; Visentin, S. Cystic Fibrosis Mucus Model to Design More Efficient Drug Therapies. *Mol. Pharm.* **2022**, *19*, 520–531. [CrossRef]
54. Hauel, N.H.; Nar, H.; Priepke, H.; Ries, U.; Stassen, J.-M.; Wienen, W. Structure-Based Design of Novel Potent Nonpeptide Thrombin Inhibitors. *J. Med. Chem.* **2002**, *45*, 1757–1766. [CrossRef]
55. Schade, D.; Kotthaus, J.; Riebling, L.; Kotthaus, J.; Müller-Fielitz, H.; Raasch, W.; Hoffmann, A.; Schmidtke, M.; Clement, B. Zanamivir Amidoxime- and N-Hydroxyguanidine-Based Prodrug Approaches to Tackle Poor Oral Bioavailability. *J. Pharm. Sci.* **2015**, *104*, 3208–3219. [CrossRef]
56. Dave, V.S.; Gupta, D.; Yu, M.; Nguyen, P.; Varghese Gupta, S. Current and evolving approaches for improving the oral permeability of BCS Class III or analogous molecules. *Drug Dev. Ind. Pharm.* **2017**, *43*, 177–189. [CrossRef]
57. Fuchs, N.; Meta, M.; Schuppan, D.; Nuhn, L.; Schirmeister, T. Novel Opportunities for Cathepsin S Inhibitors in Cancer Immunotherapy by Nanocarrier-Mediated Delivery. *Cells* **2020**, *9*, 2021. [CrossRef]
58. Lee, J.; Choi, M.-K.; Song, I.-S. Recent Advances in Doxorubicin Formulation to Enhance Pharmacokinetics and Tumor Targeting. *Pharmaceuticals* **2023**, *16*, 802. [CrossRef]
59. Kansy, M.; Senner, F.; Gubernator, K. Physicochemical High Throughput Screening: Parallel Artificial Membrane Permeation Assay in the Description of Passive Absorption Processes. *J. Med. Chem.* **1998**, *41*, 1007–1010. [CrossRef]
60. St-Georges, C.; Désilets, A.; Béliveau, F.; Ghinet, M.; Dion, S.P.; Colombo, É.; Boudreault, P.-L.; Najmanovich, R.J.; Leduc, R.; Marsault, É. Modulating the selectivity of matriptase-2 inhibitors with unnatural amino acids. *Eur. J. Med. Chem.* **2017**, *129*, 110–123. [CrossRef]
61. Costanzo, M.J.; Yabut, S.C.; Almond, H.R.; Andrade-Gordon, P.; Corcoran, T.W.; de Garavilla, L.; Kauffman, J.A.; Abraham, W.M.; Recacha, R.; Chattopadhyay, D.; et al. Potent, Small-Molecule Inhibitors of Human Mast Cell Tryptase. Antiasthmatic Action of a Dipeptide-Based Transition-State Analogue Containing a Benzothiazole Ketone. *J. Med. Chem.* **2003**, *46*, 3865–3876. [CrossRef]
62. Furlan, A.; Colombo, F.; Kover, A.; Issaly, N.; Tintori, C.; Angeli, L.; Leroux, V.; Letard, S.; Amat, M.; Asses, Y.; et al. Identification of new aminoacid amides containing the imidazo[2,1-b]benzothiazol-2-ylphenyl moiety as inhibitors of tumorigenesis by oncogenic Met signaling. *Eur. J. Med. Chem.* **2012**, *47*, 239–254. [CrossRef]
63. Capaldo, L.; Quadri, L.L.; Merli, D.; Ravelli, D. Photoelectrochemical cross-dehydrogenative coupling of benzothiazoles with strong aliphatic C–H bonds. *Chem. Commun.* **2021**, *57*, 4424–4427. [CrossRef]
64. Kerns, E.H.; Di, L.; Petusky, S.; Farris, M.; Ley, R.; Jupp, P. Combined Application of Parallel Artificial Membrane Permeability Assay and Caco-2 Permeability Assays in Drug Discovery. *J. Pharm. Sci.* **2004**, *93*, 1440–1453. [CrossRef]
65. Avdeef, A. *Absorption and Drug Development*; Wiley: Hoboken, NJ, USA, 2003. [CrossRef]
66. Sugano, K.; Hamada, H.; Machida, M.; Ushio, H. High Throughput Prediction of Oral Absorption: Improvement of the Composition of the Lipid Solution Used in Parallel Artificial Membrane Permeation Assay. *SLAS Discov.* **2001**, *6*, 189–196. [CrossRef]
67. Hammerschmidt, S.J.; Maus, H.; Weldert, A.C.; Gütschow, M.; Kersten, C. Improving binding entropy by higher ligand symmetry?—A case study with human matriptase. *RSC Med. Chem.* **2023**, *14*, 969–982. [CrossRef]
68. Ehnebom, J.; Pusa, S.; Björquist, P.; Deinum, J. Comparison of chromogenic substrates for tissue plasminogen activator and the effects on the stability of plasminogen activator inhibitor type-1. *Fibrinolysis Proteolysis* **1997**, *11*, 287–293. [CrossRef]
69. Tapp, H.J.; Grundmann, C.; Kusch, M.; Konig, H. Calibrating Thrombin Generation in Different Samples: Less Effort with a Less Efficient Substrate. *Open Atheroscler. Thromb. J.* **2009**, *2*, 6–11. [CrossRef]
70. Edwards, S.T.; Betz, A.; James, H.L.; Thompson, E.; Yonkovich, S.J.; Sinha, U. Differences between human and rabbit coagulation factor X—Implications for in vivo models of thrombosis. *Thromb. Res.* **2002**, *106*, 71–79. [CrossRef]

71. Wilkinson, D.J.; Habgood, A.; Lamb, H.K.; Thompson, P.; Hawkins, A.R.; Désilets, A.; Leduc, R.; Steinmetzer, T.; Hammami, M.; Lee, M.S.; et al. Matriptase Induction of Metalloproteinase-Dependent Aggrecanolysis In Vitro and In Vivo: Promotion of Osteoarthritic Cartilage Damage by Multiple Mechanisms. *Arthritis Rheumatol.* **2017**, *69*, 1601–1611. [CrossRef]
72. Steinmetzer, T.; Schweinitz, A.; Stürzebecher, A.; Dönnecke, D.; Uhland, K.; Schuster, O.; Steinmetzer, P.; Müller, F.; Friedrich, R.; Than, M.E.; et al. Secondary Amides of Sulfonylated 3-Amidinophenylalanine. New Potent and Selective Inhibitors of Matriptase. *J. Med. Chem.* **2006**, *49*, 4116–4126. [CrossRef]

Supporting information

Ligand-Based Design of Selective Peptidomimetic uPA and TMPRSS2 Inhibitors with Arg Bioisosteres

Patrick Müller, Collin Zimmer, Ariane Frey, Gideon Holzmann, Annabelle Carolin Weldert and Tanja Schirmeister*

Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University Mainz, Staudinger Weg 5, D-55128 Mainz, Germany; muelpat@uni-mainz.de (P.M.); cozimmer@uni-mainz.de (C.Z.); arfrey@uni-mainz.de (A.F.); gholzman@students.uni-mainz.de (G.H.); anwelder@uni-mainz.de (A.C.W.)

* Correspondence: schirmei@uni-mainz.de; Tel.: +49-6131-39-25742

Content

Fluorometric inhibition assays

The inhibitory activities of the compounds against the proteases were determined by enzyme inhibition assays with fluorogenic or colorimetry-based substrates. The fluorescence, caused by cleaved AMC from the fluorogenic substrates (uPA, matriptase, TMPRSS2, thrombin), was measured in white flat-bottom 96-well plates from Greiner Bio-One using a Tecan Infinite F200 Pro plate reader. The absorption, caused by cleaved *p*NA from colorimetric substrates (factor Xa and tPA), was measured in transparent flat-bottom 96-well plates from Greiner Bio-One using a Tecan Spark 10M plate reader. All measurements with the main-target proteases (uPA and TMRPSS2) were performed as triplicates and with the off-target proteases (matriptase, factor Xa, tPA, thrombin) as duplicates. The substrates and the compounds were prepared as stock solutions in DMSO. Each well contained a total volume of 200 µL, consisting of 185 µL buffer, 5 µL inhibitor in DMSO or pure DMSO as negative control, 5 µL substrate in DMSO and 5 µL enzyme solution in buffer. Dilution series were prepared for the determination of the inhibition constants. The fluorescence signal and absorbance were measured every 30 s for 10 min at 25 or 37 °C with the corresponding excitation/emission (λ_{ex} = 380 nm / λ_{em} = 460 nm) and absorbance (λ_{abs} = 405 nm) wavelengths, depending on the targeted protease. IC_{50} values for the reversible inhibitors were calculated with Graphpad Prism 9 by fitting the remaining enzymatic activity to the four parameter IC_{50} equation with Y [%] as the residual enzyme activity , Y_{max} as the maximum value of the dose response curve at inhibitor concentrations [I] = 0 µM, Y_{min} as the minimum value at high inhibitor concentrations and *s* as the hill coefficient.[1] The fluorescence progress curves and sigmoidal dose-response curve for inhibition of uPa by compound **38** is shown exemplary in Figure **S2**.

$$y = \frac{y_{\max} - y_{\min}}{1 + \left(\frac{[\mathrm{I}]}{\mathrm{IC}_{50}}\right)^{s}} + y_{\min} \qquad 1$$

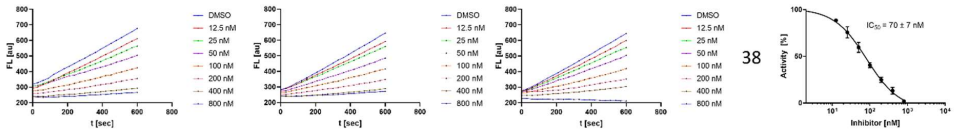


Figure S2: (left) Fluorescence progress curves for inhibition of uPa by cpd. **38**. (right) Plot showing the respective IC_{50} value from sigmoidal fit.

Due to the dependence of the IC_{50} value on the substrate affinity and concentration, the K_i values were calculated with the Cheng-Prusoff equation, using the final substrate concentration [*S*] and the Michaelis-Menten constant K_M, (2) for appropriate comparison of the inhibitory activities to the other enzymes and inhibitors.[1]

$$K_{\mathrm{i}} = \frac{\mathrm{IC}_{50}}{1 + \frac{[S]}{K_{\mathrm{M}}}} \qquad 2$$

K_M values were determined by fitting to the Michaelis-Menten equation using Graphpad Prism 9 with *v* [Δ*F*/min] as the substrate hydrolysis rate, v_{max} as the maximum slope of the dose-response curve, and the substrate [*S*] concentration.[1] The different substrates were serially diluted in DMSO. The measurement was done in analogy to the determination of the inhibitory activity, but with 5 µL pure DMSO instead of inhibitor solution. The fluorescence progress curves and Michaelis-Menten curve of the specific substrate of uPA is shown exemplarily in Figure **S3**.

$$v = \frac{v_{\max} \cdot [\mathrm{S}]}{K_{\mathrm{M}} + [\mathrm{S}]} \qquad 3$$

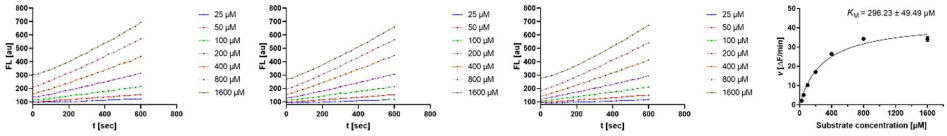


Figure S3: (left) Fluorescence progress curves of cleavage of Z-Gly-Gly-Arg-AMC by uPA. (right) Michaelis-Menten curve showing the respective K_M value.

Buffers and substrates

The following buffers and substrates were used for the respective assays: **uPA** (50 mM Tris HCl pH = 7.9, 150 mM NaCl, 10 mM $CaCl_2$, 0.005% TX-100, 240 µM Z-Gly-Gly-Arg-AMC (K_M: 296.23 ± 49.49 µM), 37 °C, 9 U/mL uPA)[2]; **TMPRSS2** (25 mM Tris HCl pH = 8.0, 150 mM NaCl, 5 mM $CaCl_2$, 0.001% TX-100, 100 µM Boc-Gln-Ala-Arg-AMC (K_M: 68.63 ± 6.64 µM), 25 °C, 3 nM TMPRSS2)[3]; **matriptase** (50 mM Tris HCl pH = 8.0, 150 mM NaCl, 5 mM $CaCl_2$, 0.0001% TX-100, 100 µM Boc-Leu-Arg-Arg-AMC (K_M: 36.13 ± 5.78 µM, 25 °C, 6 nM matriptase)[4]; **tPA** (50 mM Tris HCl pH = 8.3, 250 µM *N*-methylsulfonyl-*D*-Phe-Gly-Arg-*p*NA (K_M: 148.89 ± 29.72 µM), 37 °C, 0.4 ng/µL tPA)[5]; **thrombin** (50 mM Tris HCl pH = 8.0, 100 mM NaCl, 5 mM $CaCl_2$, 0.01% Tween-20, 200 µM Z-Gly-Gly-Arg-AMC (K_M: 49.70 ± 7.15 µM), 25 °C, 10 nM thrombin)[6]; **factor Xa** (50 mM Tris HCl pH = 7.5, 150 mM NaCl, 10 mM $CaCl_2$, 0.005% Brij-35, 100 µM Z-*D*-Arg-Gly-Arg-*p*NA (K_M: 249.12 ± 21.59 µM), 37 °C, 0.1 ng/µL factor Xa)[7].

Human uPa and tPA was purchased from Sigma Aldrich, TMPRSS2 from Cusabio, factor Xa and thrombin from Bio-techne. Matriptase was expressed under the conditions described below.

Protein expression and purification of matriptase 1

Recombinant human matriptase 1 was expressed as described previously with slight adaptations.[8,9] In short, the pQE-30 vector containing the zymogen of the catalytic domain of matriptase (uniport: Q9Y5Y6 aa 596–855) with an *N*-terminal hexa-histidine tag, was transformed into competent *Escherichia coli* (*E. coli*) BL21 Gold (DE3) cells (Agilent Technologies, Santa Clara, CA, USA). The transformed bacteria were grown at 37 °C and 160 rpm in LB medium (10 L) containing 100 µg/mL ampicillin. After reaching an optical density (OD_{600}) of ~0.8 overexpression was induced by adding 1 mM isopropyl-β-D-thiogalactopyranosid (IPTG) and the cells were incubated over night at 20 °C. Cells were harvested by centrifugation (9000 rpm at 4 °C for 15 min), resuspended in 250 mL cold lysis buffer (50 mM TRIS·HCl pH 8.0, 10% (v/v) glycerol, 300 mM NaCl, 0.1% (v/v) $Triton_{X\text{-}100}$, RNase, DNase, lysozyme, 1 mM dithiothreitol (DTT)) and lysed by sonication (Sonoplus HD 2200; Bandelin, Berlin, Germany). Inclusion bodies were isolated by centrifugation (20000 rpm at 4 °C for 1 h) and resuspended in 500 mL solubilization buffer (50 mM TRIS·HCl pH 8.0, 5% (v/v) glycerol, 6 M urea, 20 mM imidazole). After stirring the suspension overnight at 4 °C the mixture was centrifuged (20000 rpm at 4 °C for 1 h) again to remove insoluble impurities. Matriptase 1 was purified by immobilized metal affinity chromatography (IMAC) using a HisTrap HP 5 ml column (Cytiva Europe GmbH, Freiburg im Breisgau, Germany). The protein was washed with 5 column volumes (CV) wash buffer (50 mM TRIS·HCl pH 8.0, 6 M urea, 20 mM imidazole), and elution was achieved with a linear gradient of elution buffer (50 mM TRIS·HCl pH 8.0, 6 M urea, 200 mM imidazole). Matriptase 1 containing fractions were pooled and refolded by two-step dialysis against 2 L of dialysis buffer 1 (50 mM TRIS·HCl pH 9.0, 1 mM *β*-ME, 3 M urea) and 2 L of dialysis buffer 2 (50 mM TRIS·HCl pH 9.0, 1 mM β-ME) at 4 °C for > 8 h each. For further purification, anion exchange chromatography (AEX) on a HiTrap 5 mL column (Cytiva Europe GmbH) was utilized. Dialysis buffer 2 was used as AEX wash buffer (5 CV) and Matriptase was eluted with a linear gradient of dialysis buffer 2 and AEX elution buffer (50 mM TRIS·HCl pH 9.0, 1 mM β-ME, 1 M NaCl). Fractions containing matriptase, were pooled, flash-frozen in liquid nitrogen and stored at −80°C.

Calculations of effective permeability P_e were performed using the following equation with V_D and V_A as volumes of donor and acceptor solutions (0.15 cm^3 and 0.4 cm^3, respectively), AUC_{acc} and AUC_{ref} as the area of the measured and baseline-corrected spectrum of acceptor and reference solutions after incubation, A as the porosity-corrected filter area (0.3019 cm^2*0.7 = 0.2113 cm^2) and t as the incubation time given in seconds.

$$P_e = -\frac{V_D * V_A * \ln\left(1 - \frac{AUC_{acc}}{AUC_{ref}}\right)}{(V_D + V_A) * A * t} \qquad 4$$

Lit. values for propranolol (Lit-P_e(propranolol) = 8.3–13.9 *10^{-6} cm/s [11,12], measured P_e = 9.0 ± 0.2*10^{-6} cm/s) and methotrexate (Lit-P_e(methotrexate) = 0.0 *10^{-6} cm/s [13], measured P_e = 0*10^{-6} cm/s) were replicated ensuring reliability of data for novel compounds.

Synthesis

General Methods and Materials:

All reagents and solvents were purchased commercially and used as provided by the supplier without further purification. Solvents for synthesis, extraction, and chromatography were of analytical grade. Moisture-sensitive reactions were carried out under argon atmosphere as indicated, and anhydrous solvents were used as provided by the commercial supplier. Reaction progress was monitored by thin-layer chromatography using Alugram Xtra F254 silica plates from Macherey-Nagel and/or LC-MS (Agilent 1100 series HPLC system and an Agilent Poroshell 120 EC-C18, 150 x 2.10 mm, 4 μm column coupled to an Agilent 1100 series LC/MSD Trap with electron spray ionization (ESI)). The identities and purities of compounds were determined by the same LC-MS system with a gradient of acetonitrile and water (+0.1% formic acid). Signals were detected at 210/254 nm with quantitation by AUC and masses were determined in positive ionization mode (ESI). HPLC purification was performed with the Agilent 1290 II Infinity Preparative LC System using an InfinityLab Pursuit XRs C18, 30 x 250mm, 5 µm, preparative LC column and a gradient method (10% ACN (0.1% formic acid) to 100% ACN (0.1% formic acid)). Flash chromatography was performed with the Biotage Isolera™ One system using prepacked columns from Biotage. Silica gel (0.040 – 0.063 mm) from Macherey-Nagel was used for column chromatography. Optical rotations $[\alpha]_D^{20}$ were measured on an P3000 polarimeter from Krüss at 20 °C and are reported in ml·dm^{-1}·g^{-1} with the concentration c being g/100 ml. Fourier-transformed ATR-corrected IR spectra were measured on an Avatar 330 single crystal spectrometer from ThermoNicolet. Melting points (uncorrected) were measured with an MPM-H3 using semi-open capillaries. NMR spectra were recorded as stated individually on Bruker Fourier 300 MHz, Bruker Avance DSX 400 MHz and Bruker Avance III 600 MHz. Chemical shifts are indicated in parts per million (ppm), with the solvent resonance ($CDCl_3$ or DMSO-d_6 from Deutero GmbH) as internal standard. The purity of all compounds tested in biological assays was ≥95% as determined by LC-MS.

Solid-Phase Peptide Synthesis (SPPS)[2]

A Resin loading. Solid-phase peptide synthesis was conducted in a fritted 12 mL polypropylene syringe. 1 g of 2-chlorotritylchloride resin (2-CTC resin, 1.2 mmol loading capacity) was pre-swelled in 8 mL DCM for 15 min and drained. The first amino acid (3.6 mmol, 3 eq. relative to resin loading capacity) was added in 1.8 M *N*-methylmorpholine (NMM)/DCM (6 mL) and the mixture was swiveled for 12 h. After draining the solution, the resin was washed with DMF (3 × 6 mL for 1 min) and DCM (3 × 6 mL for 1 min).

B Resin capping. The remaining free 2-CTC resin linkers were capped with MeOH in a solution of 9 mL of DCM/MeOH/DIPEA (9:2:1). The resin was swiveled for 1 h, drained, and washed with DCM (3 × 6 mL).

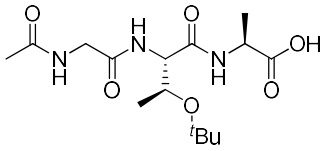
C Fmoc deprotection. The resin was treated with 20% piperidine in DMF solution (2 × 6 mL for 10 min) and subsequently washed with DMF (3 × 6 mL).

D Peptide coupling with HATU. A coupling cocktail was prepared including the specific Fmoc-protected amino acid (3.6 mmol, 3 eq.), HATU (3.6 mmol, 3 eq.), and DIEA (10.8 mmol, 9 eq.) in DMF (4.8 mL). The solution was added to the resin and swiveled for 3 h with exception of Fmoc-Arg(Pbf)-OH which was coupled overnight. After the reaction, the resin was drained and washed with DMF (3 × 6 mL for 1 min) and DCM (3 × 6 mL for 1 min).

E Acetylation of the *N*-terminus. The free amino terminus was acetylated using 8 mL of Ac_2O, DIEA and DCM (3:2:3). The reaction was shaken overnight and subsequently drained and washed with DMF (3 × 6 mL for 1 min) and DCM (3 × 6 mL for 1 min).

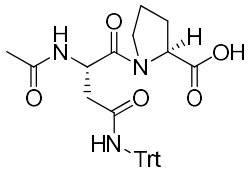
F Resin cleavage with preservation of side chain protecting groups. After the last coupling or acetylation step, the resin was washed with DMF (3 × 6 mL for 1 min) and DCM (3 × 6 mL for 1 min). Then a cleavage cocktail containing AcOH/trifluoroethanol/DCM (1:1:4, 6 mL) was added and swiveled for 1.5 h. The resin was drained and washed with DCM (5 mL). The combined eluates were concentrated under reduced pressure to yield the crude linear peptide. The purification of the crude product was performed by flash chromatography on a reversed-phase silica column (POLYPREP 60-50 C_{18}).

3, (*S*)-2-((2*S*,3*S*)-2-(2-Acetamidoacetamido)-3-(*tert*-butoxy)butanamido)propanoic acid



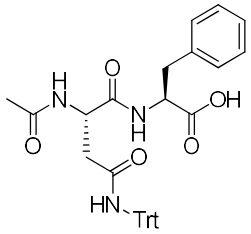
The title compound was prepared according to general procedures A, B, C, D, and F on 1.2 mmol scale. The product was purified by reversed phase flash chromatography (H_2O/MeCN, gradient) and obtained as a colorless solid after lyophilization (314 mg, 0.90 mmol, 75%) ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 8.19 (t, *J* = 5.8 Hz, 1H), 7.95 (d, *J* = 7.0 Hz, 1H), 7.53 (d, *J* = 8.5 Hz, 1H), 4.26 – 4.16 (m, 2H), 3.92 (qd, *J* = 6.2, 3.9 Hz, 2H), 3.74 (dd, *J* = 7.2, 5.8 Hz, 2H), 1.86 (s, 3H), 1.28 (d, *J* = 7.2 Hz, 3H), 1.12 (s, 9H), 1.01 (d, *J* = 6.2 Hz, 3H), ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 173.8, 169.9, 169.3, 168.9, 73.8, 67.2, 57.3, 47.7, 42.3, 28.1, 22.4, 19.3, 17.5, FT-IR: ν/cm^{-1} = 668, 701, 1081, 1121, 1159, 1192, 1214, 1371, 1518, 1636, mp: 87 – 95 °C. $[\alpha]_D^{20}$ = +12 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{15}H_{28}N_3O_6]^+$ ($[M+H]^+$): 346.2, found 368.1 ($[M+Na]^+$).

21, (*S*)-1-((*S*)-2-Acetamido-4-oxo-4-(tritylamino)butanoyl)pyrrolidine-2-carboxylic acid



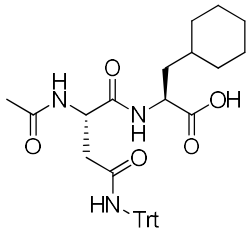
The title compound was prepared according to general procedures A, B, C, D, E and F on 1.2 mmol scale. The product was purified by reversed phase flash chromatography (H_2O/MeCN, gradient) and obtained as a colorless solid after lyophilization (326 mg, 0.63 mmol, 52%) ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 7.71 – 7.41 (m, 2H), 7.22 – 7.16 (m, 15H), 5.02 – 4.92 (m, 1H), 4.33 – 4.20 (m, 1H), 3.70 – 3.42 (m, 3H), 2.77 – 2.58 (m, 2H), 1.99 – 1.75 (m, 7H), ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 170.5, 160.5, 154.9, 144.3, 128.9, 127.9, 127.0, 123.8, 70.9, 59.9, 58.1, 48.6, 47.6, 39.1, 31.1, 28.7, 24.9, 23.0, FT-IR: ν/cm^{-1} = 689, 903, 1198, 1437, 1521, 1628, 1726, 2240, 3047, 3303, mp: 184 – 188 °C, $[\alpha]_D^{20}$ = −16 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{30}H_{32}N_3O_5]^+$ ($[M+H]^+$): 513.5, found 536.1 ($[M+Na]^+$).

22, (*S*)-2-((*S*)-2-Acetamido-4-oxo-4-(tritylamino)butanamido)-3-phenylpropanoic acid



The title compound was prepared according to general procedures A, B, C, D, E and F on 1.2 mmol scale. The product was purified by reversed phase flash chromatography (H_2O/MeCN, gradient) and obtained as a colorless solid after lyophilization (514 mg, 0.91 mmol, 76%) ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 8.51 – 8.37 (m, 1H), 8.22 – 8.08 (m, 1H), 7.89 – 7.79 (m, 1H), 4.69 – 4.57 (m, 1H), 4.50 – 4.39 (m, 1H), 3.10 – 3.01 (m, 1H), 2.96 – 2.84 (m, 1H), 2.65 – 2.54 (m, 1H), 2.03 (s, 1H), 1.85 (s, 3H), ^{13}C NMR (75 MHz, DMSO-d_6): δ/ppm = 172.6, 171.2, 169.2, 168.7, 144.8, 137.3, 129.2, 129.1, 128.5, 128.2, 127.4, 126.5, 126.3, 69.3, 53.4, 49.8, 36.5, 22.5, FT-IR: ν/cm^{-1} = 697, 748, 903, 1040, 1181, 1439, 1484, 1639, 3041, 3274, mp: 226 – 228 °C, $[\alpha]_D^{20}$ = −14 (*c* 1.00, DMSO), MS (ESI) *m/z* calculated for $[C_{34}H_{34}N_3O_5]^+$ ($[M+H]^+$): 563.6, found 586.1 ($[M+Na]^+$).

23, (*S*)-2-((*S*)-2-Acetamido-4-oxo-4-(tritylamino)butanamido)-3-cyclohexylpropanoic acid

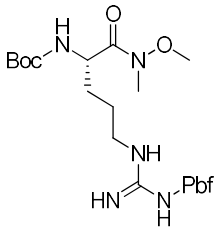


The title compound was prepared according to general procedures A, B, C, D, E and F on 1.2 mmol scale. The product was purified by reversed phase flash chromatography (H_2O/MeCN, gradient) and obtained as a colorless solid after lyophilization (467 mg, 0.82 mmol, 68%) ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 7.61 – 7.44 (m, 2H), 7.28

– 7.14 (m, 15H), 4.90 – 4.67 (m, 2H), 4.40 – 4.21 (m, 1H), 2.93 – 2.54 (m, 2H), 1.87 (s, 3H), 1.71 – 1.04 (m, 12H), 0.85 (q, *J* = 11.3 Hz, 2H), ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 175.4, 171.7, 171.5, 171.3, 171.1, 144.3, 144.1, 128.8, 128.0, 127.1, 70.9, 50.9, 50.1, 38.6, 37.7, 34.1, 33.4, 32.4, 26.4, 26.1, 26.0, 23.0, FT-IR: ν/cm^{-1} = 699, 736, 914, 1204, 1435, 1510, 1656, 1722, 2365, 2926, mp: 180 – 182 °C, $[\alpha]_D^{20}$ = −15 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{34}H_{40}N_3O_5]^+$ ($[M+H]^+$): 569.6, found 592.1 ($[M+Na]^+$).

Synthesis of P1-precursor molecules

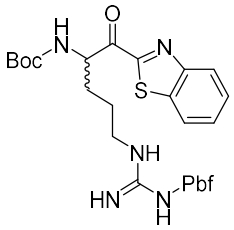
5, *Tert*-Butyl (*S*)-(1-(methoxy(methyl)amino)-1-oxo-5-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)pentan-2-yl)carbamate



(Adapted from St-Georges *et al*.)[14]

Boc-Arg(Pbf)-OH **4** (4.00 g, 7.59 mmol, 1 eq) was dissolved in DCM (30 mL) under argon atmosphere. At 0 °C DIEA (10.58 mL, 60.76 mmol, 8 eq) and TBTU (2.92 g, 9.11 mmol, 1.2 eq) were added. The reaction mixture was stirred for 30 min at 0 °C, *N*,*O*-dimethyl hydroxylamine · HCl (4.44 g, 45.54 mmol, 6 eq) was added and stirred for 12 h. The solution was diluted with DCM (15 mL) and washed with saturated $NaHCO_3$ solution (3x 20 mL) and saturated NaCl solution (3x 20 mL). The organic solution was dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The crude product was purified by column chromatography (EA) to yield the desired product as a colorless solid (3.89 g, 6.83 mmol, 90%). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 6.38 (s, 1H), 5.49 (d, *J* = 8.8 Hz, 1H), 4.61 (s, 1H), 3.73 (s, 3H), 3.18 (s, 3H), 2.95 (s, 3H), 2.56 (s, 3H), 2.50 (s, 3H), 2.08 (s, 3H), 1.74 – 1.52 (m, 4H), 1.45 (s, 6H), 1.41 (s, 9H), ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 156.4, 156.0, 138.8, 132.7, 124.8, 117.7, 86.6, 80.3, 61.7, 43.3, 41.0, 30.9, 28.7, 28.4, 24.9, 19.4, 18.0, 12.5, mp: 76 – 84 °C. $[\alpha]_D^{20}$ = −23 (*c* 1.00, $CHCl_3$). FT-IR: ν/cm^{-1} = 668, 756, 1106, 1165, 1456, 1556, 1652, 2362, 2974, 2981. MS (ESI) *m/z* calculated for $[C_{26}H_{44}N_5O_7S]^+$ ($[M+H]^+$): 570.3, found 570.3.

12a, *Tert*-Butyl (*S*/*R*)-(1-(benzo[*d*]thiazol-2-yl)-1-oxo-5-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)pentan-2-yl)carbamate

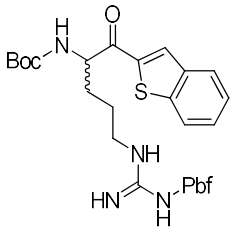


(Adapted from Costanzo *et al*.)[15]

Benzothiazole **10a** (4.83 mL, 44.67 mmol, 10.2 eq) was dissolved in dry THF (80 mL) under argon atmosphere and cooled to −78 °C. *n*-BuLi (2.5 M in hexanes, 17.52 mL, 43.80 mmol, 10 eq) was added dropwise and the reaction stirred for 1.5 h at the same temperature. Afterwards, **5** (2.50 g, 4.38 mmol, 1 eq) was dissolved in dry THF (30 mL), added slowly to the reaction mixture and stirred for two hours at −78 °C. The reaction mixture was

quenched with saturated aqueous NH_4Cl solution (20 mL). The organic phase was separated and washed three times each with saturated $NaHCO_3$ solution (25 mL) and saturated NaCl solution (25 mL). The organic phase was dried over anhydrous $NaSO_4$ and concentrated under reduced pressure. The crude product was purified by column chromatography (CH/EA, 1:1 to 0:1) to yield the product as a colorless solid (1.11 g, 3.06 mmol, 70%). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 8.23 – 8.15 (m, 1H), 8.00 – 7.91 (m, 1H), 7.60 – 7.49 (m, 2H), 6.31 (s, 1H), 6.21 (s, 2H), 5.65 (d, *J* = 8.6 Hz, 1H), 5.61 – 5.49 (m, 1H), 3.57 – 3.18 (m, 2H), 2.92 (s, 3H), 2.54 (s, 3H), 2.49 (s, 3H), 2.06 (s, 3H), 1.72 (s, 4H), 1.45 (s, 6H), 1.41 (s, 9H), ^{13}C NMR (75 MHz, $CDCl_3$): δ/ppm = 158.8, 156.2, 153.5, 138.5, 137.3, 133.0, 132.5, 128.3, 127.4, 126.0, 124.7, 122.5, 117.5, 86.4, 80.6, 43.3, 40.8, 28.7, 28.4, 25.5, 19.3, 18.0, 12.5, FT-IR: ν/cm^{-1} = 668, 756, 1102, 1164, 1216, 1369, 1483, 1552, 1622, 1694, mp: 105 – 115 °C. $[\alpha]_D^{20}$ = +6 (*c* 1.00, $CHCl_3$, MS (ESI) *m/z* calculated for $[C_{31}H_{42}N_5O_6S_2]^+$ ($[M+H]^+$): 644.3, found 644.3.

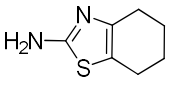
12f, *Tert*-Butyl (*S*/*R*)-(1-(benzo[*b*]thiophene-2-yl)-1-oxo-5-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)pentan-2-yl)carbamate



(Adapted from Costanzo *et al.*)[15]

1-Benzothiophene **10f** (0.71 g, 5.30 mmol, 10 eq) was dissolved in dry THF (20 mL) under argon atmosphere and cooled to −78 °C. *n*-BuLi (2.5 M in hexanes, 1.70 mL, 4.24 mmol, 8 eq) was added dropwise and the reaction stirred for 2 h letting the temperature rise to rt. Afterwards, **5** (0.30 g, 0.53 mmol, 1 eq) was dissolved in dry THF (10 mL), added slowly to the again cooled reaction mixture at −78 °C and stirred for 2.5 h. The reaction mixture was quenched with saturated NH_4Cl solution (5 mL). The organic phase was separated and washed three times each with saturated $NaHCO_3$ solution (10 mL) and saturated NaCl solution (10 mL). The organic phase was dried over anhydrous $NaSO_4$ and concentrated under reduced pressure. The crude product was purified by column chromatography (CH/EA, 1:1 to 1:3) to yield the product as a yellow solid (0.27 g, 0.42 mmol, 80%). ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = 8.41 (s, 1H), 8.31 (s, 1H), 8.08–7.99 (m, 2H), 7.58–7.44 (m, 3H), 4.92 (s, 1H), 3.12–3.02 (m, 2H), 2.90 (s, 2H), 2.45 (s, 3H), 2.39 (s, 3H), 1.96 (s, 3H), 1.79–1.48 (m, 5H), 1.42–1.32 (m, 15H), ^{13}C NMR (75 MHz, DMSO-d_6): δ/ppm = 194.3, 157.4, 156.1, 155.5, 141.4, 141.2, 139.1, 137.2, 134.1, 131.4, 130.5, 127.8, 126.4, 125.3, 124.3, 123.1, 116.2, 86.2, 79.2, 78.4, 55.8, 40.4, 28.3, 28.1, 18.9, 17.5, 12.2, FT-IR: ν/cm^{-1} = 667, 763, 816, 1024, 1097, 1148, 1252, 1561, 2966, 3334, mp: 105 – 126 °C. $[\alpha]_D^{20}$ = +7 (*c* 1.00, DMSO), MS (ESI) *m/z* calculated for $[C_{32}H_{42}N_4O_6S_2]^+$ ($[M+H]^+$): 643.3, found 643.2.

9, 2-Amino-4,5,6,7-tetrahydrobenzothiazole



(Adapted from Furlan *et al.*)[16]

Cyclohexane **7** (3 g, 30.6 mmol, 1 eq), thiourea **8** (4.65 g, 61.2 mmol, 2 eq) and iodine (7.74 g, 30.6 mmol, 1 eq) were stirred at 130 °C under argon atmosphere for 12 h. To the cooled mixture was added water (50 mL) and stirred for 30 min. The aqueous phase was washed three times with ether (50 mL) and neutralized with $NaHCO_3$ (pH = 7). The precipitate was filtered and solved in hot saturated Na_2CO_3 solution. The cooled mixture was extracted with DCM (30 mL) dried over Na_2SO_4 and evaporated under reduced pressure to yield the product as

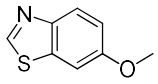
a yellowish solid (2.5 g, 16.2 mmol, 53%). ^{1}H NMR (300 MHz, DMSO-d_6): δ/ppm = δ 6.56 (s, 2H), 2.51 – 2.41 (m, 2H), 2.39 – 2.30 (m, 2H), 1.76 – 1.63 (m, 4H), ^{13}C NMR (75 MHz, DMSO-d_6): δ/ppm = 165.3, 144.7, 114.4, 26.2, 23.2, 22.7, 22.5, FT-IR: ν/cm^{-1} = 692, 892, 1018, 1062, 1112, 1236, 1309, 1365, 1524, 1634, mp: 86 – 90 °C, MS (ESI) *m/z* calculated for $[C_7H_{13}N_2S]^+$ ($[M+H]^+$): 155.0., found 155.9.

General procedure for desamination of the 2-aminobenzoheterocycle derivatives

(Adapted from Capaldo *et al*.)[17]

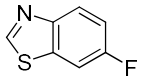
The respective 2-aminobenzoheterocycle **6b**-**e**, **9** (1 eq) was dissolved in dry THF (25 mL) under argon atmosphere. Isopentyl nitrite (2.2 eq) was added dropwise under light protection and the reaction stirred for 2 – 5 h under reflux. Afterwards, ice-water was added (40 mL) and the aqueous solution was extracted three times with EA (30 mL). The organic phase was dried over anhydrous $NaSO_4$ and concentrated under reduced pressure.

10e, 6-Methoxybenzo[*d*]thiazole



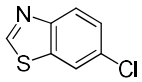
The crude product was purified by column chromatography (CH/EA, 20:1) to yield the product as a yellow solid (1.98 g, 11.96 mmol, 72%). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 8.84 (s, 1H), 8.01 (dd, J = 9.0, 0.5 Hz, 1H), 7.38 (d, J = 2.5 Hz, 1H), 7.12 (dd, J = 9.0, 2.5 Hz, 1H), 3.87 (s, 3H), ^{13}C NMR (75 MHz, $CDCl_3$): δ/ppm = 158.2, 151.6, 147.5, 135.1, 124.0, 116.1, 104.1, 55.9, FT-IR: ν/cm^{-1} = 653, 822, 900, 1018, 1049, 1193, 1246, 1476, 1608, 2952, mp: 72 – 75 °C, MS (ESI) *m/z* calculated for $[C_8H_7NOS]^+$ ($[M+H]^+$): 166.0, found 165.9.

10b, 6-Fluorobenzo[*d*]thiazole



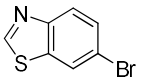
The crude product was purified by column chromatography (CH/EA, 20:1) to yield the product as a yellow solid (1.99 g, 12.96 mmol, 66%). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 8.96 (s, 1H), 8.08 (dd, J = 9.0, 4.8 Hz, 1H), 7.62 (dd, J = 8.1, 2.5 Hz, 1H), 7.26 (td, J = 8.9, 2.5 Hz, 1H), ^{13}C NMR (75 MHz, $CDCl_3$): δ/ppm = 162.5, 159.3, 149.8, 134.9, 124.6, 115.2, 108.1, FT-IR: ν/cm^{-1} = 779, 807, 838, 971, 1218, 1260, 1451, 1482, 1569, 3081, mp: 57 – 61 °C, MS (ESI) *m/z* calculated for $[C_7H_4FNS]^+$ ($[M+H]^+$): 154.0, found 153.9.

10c, 6-Chlorobenzo[*d*]thiazole



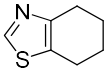
The crude product was purified by column chromatography (CH/EA, 20:1) to yield the product as a yellow solid (1.26 g, 7.43 mmol, 46%). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 9.00 (s, 1H), 8.05 (dd, J = 8.7, 0.5 Hz, 1H), 7.93 (dd, J = 2.1, 0.5 Hz, 1H), 7.48 (dd, J = 8.7, 2.1 Hz, 1H), ^{13}C NMR (75 MHz, $CDCl_3$): δ/ppm = 154.5, 151.6, 135.0, 131.9, 127.3, 124.4, 121.6, FT-IR: ν/cm^{-1} = 757, 802, 855, 889, 1046, 1108, 1392, 1426, 1474, 3042, mp: 42 – 47 °C, MS (ESI) *m/z* calculated for $[C_7H_4ClNS]^+$ ($[M+H]^+$): 170.0, found 169.9.

10d, 6-Bromobenzo[*d*]thiazole



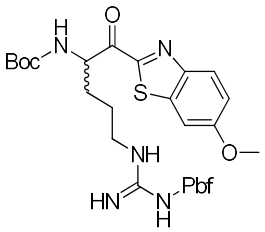
The crude product was purified by column chromatography (CH/EA, 30:1) to yield the product as a yellow solid (2.14 g, 9.99 mmol, 76%). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 9.01 (s, 1H), 8.10 (dd, J = 1.9, 0.5 Hz, 1H), 8.00 (dd, J = 8.7, 0.5 Hz, 1H), 7.63 (dd, J = 8.7, 1.9 Hz, 1H), ^{13}C NMR (75 MHz, $CDCl_3$): δ/ppm = 154.6, 151.8, 135.5, 130.0, 124.7, 124.6, 119.7, FT-IR: ν/cm^{-1} = 740, 799, 858, 886, 1077, 1294, 1386, 1459, 1547, 3039, mp: 50 – 55 °C, MS (ESI) *m/z* calculated for $[C_7H_4BrNS]^+$ ($[M+H]^+$): 213.9, found 213.8.

11, 4,5,6,7-Tetrahydrobenzothiazole



The crude product was purified by column chromatography (CH/EA, 7:1) to yield the product as a yellow oil (1.08 g, 9.99 mmol, 76%). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 8.60 (s, 1H), 2.87 – 2.74 (m, 4H), 1.90 – 1.80 (m, 4H), ^{13}C NMR (75 MHz, $CDCl_3$): δ/ppm = 150.8, 149.6, 128.8, 26.7, 23.5, 22.9, FT-IR: ν/cm^{-1} = 712, 807, 859, 891, 1138, 1209, 1379, 1413, 1458, 1591, MS (ESI) *m/z* calculated for $[C_7H_9NS]^+$ ($[M+H]^+$): 140.05, found 139.9.

12e, *Tert*-butyl (*S*/*R*)-(1-(6-methoxybenzo[d]thiazol-2-yl)-1-oxo-5-(3-((2,2,4,5,7-pentamethyl-2,3-dihydrobenzofuran-6-yl)sulfonyl)guanidino)pentan-2-yl)carbamate

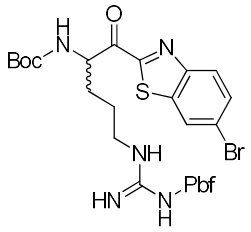


(Adapted from Costanzo *et al*.)[15]

6-methoxybenzo[*d*]thiazole **10e** (1.94 g, 11.72 mmol, 10.2 eq) was dissolved in dry THF (40 mL) under argon atmosphere and cooled to −78 °C. *n*-BuLi (2.5 M in hexanes, 4.60 mL, 11.50 mmol, 10 eq) was added dropwise and the reaction stirred for 2 h at the same temperature. Afterwards, **5** (0.66 g, 1.15 mmol, 1 eq) was dissolved in dry THF (20 mL), added slowly to the reaction mixture at −78 °C and stirred for 2 h. The reaction was quenched with saturated NH_4Cl solution (15 mL). The organic phase was separated and washed three times each with saturated $NaHCO_3$ solution (20 mL) and saturated NaCl solution (20 mL). The organic phase was dried over anhydrous $NaSO_4$ and concentrated under reduced pressure. The crude product was purified by column chromatography ($CHCl_3$/MeOH, 100:1) to yield the product as a brown oil (0.68 g, 1.01 mmol, 88%). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 8.04 (d, J = 9.1 Hz, 1H), 7.34 (d, J = 2.5 Hz, 1H), 7.15 (dd, J = 9.1, 2.5 Hz, 1H), 6.52 (s, 2H), 5.68 (d, J = 8.6 Hz, 1H), 5.52 (s, 1H), 3.90 (s, 3H), 3.53 (s, 1H), 3.35–3.23 (m, 1H), 2.93 (s, 2H), 2.54 (s, 3H), 2.49 (s, 3H), 2.06 (s, 3H), 1.83–1.58 (m, 5H), 1.45 (s, 6H), 1.41 (s, 9H), ^{13}C NMR (75 MHz, $CDCl_3$): δ/ppm = 193.0, 160.3, 159.3, 156.5, 156.0, 152.4, 148.2, 139.6, 139.0, 133.0, 131.9, 126.8, 124.9, 118.2, 117.8, 103.7, 86.7, 80.6, 56.0, 55.1, 43.3, 40.9, 31.6, 28.7, 28.5, 25.4, 19.4, 18.0, 12.6, FT-IR: ν/cm^{-1} = 723, 903, 1091, 1164, 1260, 1485, 1555, 1682, 2974, 3334, $[\alpha]_D^{20}$ = +9 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{32}H_{44}N_5O_7S_2]^+$ ($[M+H]^+$): 674.3, found 674.2.

156.3, 156.1, 152.1, 138.9, 138.4, 134.7, 132.8, 131.8,128.5, 126.8, 124.9, 122.0, 117.8, 86.7, 80.7, 55.5, 43.3, 40.8, 28.7, 28.5, 25.5, 19.4, 18.0, 12.6, FT-IR: ν/cm^{-1} = 726, 911, 1097, 1162, 1252, 1482, 1552, 1696, 2974, 3337, , $[\alpha]_D^{20}$ = +5 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{31}H_{41}ClN_5O_6S_2]^+$ ($[M+H]^+$): 678.2, found 678.2.

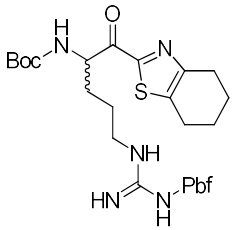
12d, *Tert*-butyl (*S*/*R*)-(1-(6-bromobenzo[d]thiazol-2-yl)-1-oxo-5-(3-((2,2,4,5,7-pentamethyl-2,3-dihydrobenzofuran-6-yl)sulfonyl)guanidino)pentan-2-yl)carbamate



(Adapted from Costanzo *et al*.)[15]

6-Bromobenzo[*d*]thiazole **10d** (2.11 g, 9.9 mmol, 10.1 eq) was dissolved in dry THF (30 mL) under argon atmosphere and cooled to −78 °C. *n*-BuLi (2.5 m in hexanes, 3.9 mL, 9.70 mmol, 10 eq) was added dropwise and the reaction stirred for 1.5 h under the same temperature. Afterwards, **5** (0.55 g, 0.97 mmol, 1 eq) was dissolved in dry THF (10 mL), added slowly to the reaction mixture at −78 °C and stirred for 2 hours. The reaction was quenched with saturated NH_4Cl solution (5 mL). The organic phase was separated and washed three times each with saturated $NaHCO_3$ solution (10 mL) and saturated NaCl solution (10 mL). The organic phase was dried over anhydrous $NaSO_4$ and concentrated under reduced pressure. The crude product was purified by column chromatography ($CHCl_3$/MeOH, 70:1) to yield the product as a yellow oil (0.51 g, 0.51 mmol, 73%). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 8.23–7.88 (m, 3H), 7.58–7.49 (m, 1H), 6.44 (s, 2H), 5.67 (t, *J* = 8.8 Hz, 1H), 5.50 (s, 1H), 3.42 (s, 1H), 3.32–3.20 (m, 1H), 2.92 (s, 2H), 2.52 (s, 3H), 2.46 (d, *J* = 2.1 Hz, 3H), 2.05 (s, 3H), 1.82–1.58 (m, 4H), 1.45 (s, 6H), 1.40 (s, 9H), ^{13}C NMR (75 MHz, $CDCl_3$): δ/ppm = 217.9, 193.4, 159.1, 156.2, 153.5, 152.3, 138.8, 137.3, 132.6, 131.1, 128.3, 127.4, 127.0, 126.0, 125.0, 124.8, 122.5, 117.7, 86.6, 80.6, 55.6, 43.3, 40.8, 28.7, 28.4, 25.5, 19.4, 18.1, 12.6, FT-IR: ν/cm^{-1} = 723, 807, 886, 1097, 1246, 1398, 1459, 1558, 1690, 3042, $[\alpha]_D^{20}$ = +9 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{31}H_{41}BrN_5O_6S_2]^+$ ($[M+H]^+$): 722.2, found 722.1.

13, *Tert*-butyl (*S*/*R*)-(1-oxo-5-(3-((2,2,4,5,7-pentamethyl-2,3-dihydrobenzofuran-6-yl)sulfonyl)guanidino)-1-(4,5,6,7-tetrahydrobenzo[d]thiazol-2-yl)pentan-2-yl)carbamate

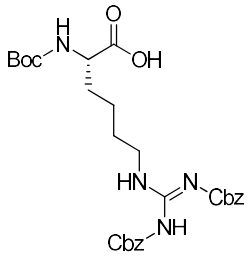


(Adapted from Costanzo *et al*.)[15]

4,5,6,7-Tetrahydrobenzothiazole **11** (0.75 g, 5.40 mmol, 10.1 eq) was dissolved in dry THF (18 mL) under argon atmosphere and cooled to −78 °C. TMEDA (0.79 mL, 5.28 mmol, 10 eq) and *n*-BuLi (2.5 m in hexanes, 2.1 mL, 5.28 mmol, 10 eq) were added dropwise and the reaction stirred for 1.5 h under the same temperature. Afterwards, **5** (0.30 g, 0.52 mmol, 1 eq) was dissolved in dry THF (8 mL), added slowly to the reaction mixture at −78 °C and stirred for 2 hours. The reaction was quenched with saturated NH_4Cl solution (5 mL). The organic phase was separated and washed three times each with saturated $NaHCO_3$ solution (10 mL) and saturated NaCl

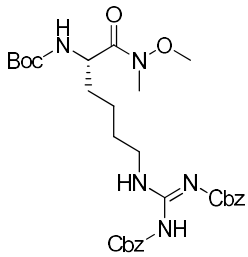
solution (10 mL). The organic phase was dried over anhydrous $NaSO_4$ and concentrated under reduced pressure. The crude product was purified by column chromatography ($CHCl_3$/MeOH, 40:1) to yield the product as a colorless oil (0.17 g, 0.26 mmol, 50%). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 7.28 (d, J = 7.4 Hz, 1H), 5.05 (s, 1H), 3.02 (d, J = 5.9 Hz, 2H), 2.95 (s, 2H), 2.87 (s, 2H), 2.78 (d, J = 6.1 Hz, 2H), 2.46 (s, 3H), 2.41 (s, 3H), 1.99 (s, 3H), 1.84 – 1.47 (m, 8H), 1.40 (s, 6H), 1.35 (s, 9H), ^{13}C NMR (75 MHz, $CDCl_3$): δ/ppm = 192.4, 160.5, 157.4, 156.0, 155.4, 153.2, 138.6, 137.2, 134.1, 131.4, 124.2, 116.2, 86.2, 78.2, 55.2, 42.4, 28.1, 26.4, 23.6, 22.4, 22.1, 18.9, 17.5, 12.2, FT-IR: ν/cm^{-1} = 623, 812, 1118, 1164, 1249, 1366, 1425, 1540, 1677, 1719, $[\alpha]_D^{20}$ = +5 (c 1.00, $CHCl_3$), MS (ESI) m/z calculated for $[C_{31}H_{46}N_5O_6S_2]^+$ ([M+H]$^+$): 648.2, found 648.2.

17, (*S*)-6-(2,3-Bis((benzyloxy)carbonyl)guanidino)-2-((*tert*-butoxycarbonyl)amino)hexanoic acid



Boc-Lys-OH **16** (2.00 g, 8.12 mmol, 1 eq) and *N*,*N*´-bis-(carbobenzoxy)-1*H*-pyrazole-1-carboxamidine (3.38 , 8.93 mmol, 1.1 eq) were dissolved in DMF (58 mL) and Et_3N (1.69 mL, 12.18 mmol, 1.5 eq) was added under argon atmosphere. The reaction stirred for 72 h and got quenched by addition of NH_4Cl (20 mL). The mixture was extracted with EA (50 mL), dried over $MgSO_4$ and the organic solvents were removed under reduced pressure. The crude product was purified by column chromatography (CH/EA, 2:1 + 0.1% TFA) to yield the product as a colorless oil (4.94 g, 8.12 mmol, quant.). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 7.40 – 7.28 (m, 10H), 5.18 (d, J = 16.0 Hz, 4H), 4.27 (s, 1H), 3.47 (s, 2H), 1.92 – 1.56 (m, 6H), 1.43 (s, 9H), ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 175.6, 155.2, 153.7, 135.7, 134.2, 129.1, 128.9, 128.7, 128.6, 128.4, 128.3, 69.0, 68.3, 53.1, 41.8, 31.8, 28.4, 22.4, FT-IR: ν/cm^{-1} = 698, 752, 1052, 1140, 1160, 1204, 1248, 132, 1640, 1729, $[\alpha]_D^{20}$ = +16 (c 1.00, $CHCl_3$), MS (ESI) m/z calculated for $[C_{28}H_{37}N_4O_8]^+$ ([M+H]$^+$): 557.2, found 557.2.

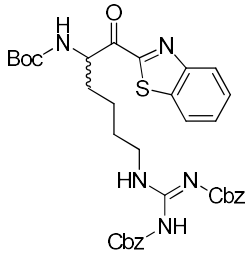
18, (*S*)-(1-(Methoxy(methyl)amino)-1-oxo-6-(2,3-bis((benzyloxy)carbonyl)guanidino)-2-((tert-butoxycarbonyl)amino)carbamate



17 (4.94 g, 8.12 mmol, 1 eq), TBTU (3.64 g, 11.36 mmol, 1.4 eq) and *N*,*O*-dimethyl hydroxylamine · HCl (1.58 g, 16.24 mmol, 2 eq) were dissolved in THF (30 mL) under argon atmosphere and cooled to 0 °C. DIEA (8.48 mL, 48.72 mmol, 6 eq) was added and the reaction stirred overnight. The mixture was extracted with EA (40 mL) and washed with saturated NaCl solution, dried over $MgSO_4$ and the organic solvents were removed under reduced pressure. The crude product was purified by column chromatography (CH/EA, 1:1) to yield the product as a colorless oil (3.38 g, 5.68 mmol, 70%). ^{1}H NMR (300 MHz, $CDCl_3$): δ/ppm = 7.42 – 7.24 (m, 10H), 5.19 – 5.10 (m, 4H), 4.65 (s, 1H), 3.74 (s, 3H), 3.49 – 3.32 (m, 2H), 3.18 (s, 3H), 1.77 – 1.48 (m, 6H), 1.43 (s, 9H), ^{13}C NMR (75 MHz,

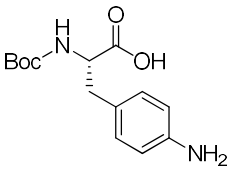
CDCl3): δ/ppm = 163.8, 156.0, 155.6, 153.9, 136.9, 134.7, 128.9, 128.8, 128.6, 128.5, 128.2, 128.0, 79.1, 68.2, 67.2, 61.7, 50.2, 41.0, 32.7, 28.7, 28.4, 22.7, FT-IR: ν/cm^{-1} = 1049, 1137, 1167, 1206, 1249, 1323, 1365, 1383, 1425, 1638, $[\alpha]_D^{20}$ = +14 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{31}H_{42}N_5O_8]^+$ ($[M+H]^+$): 600.3, found 600.3.

19, (*S*/*R*)-1-Oxo-6-(2,3-bis((benzyloxy)carbonyl)guanidino)-2-((*tert*-butoxycarbonyl) amino)-2-benzo[d]thiazol



Benzothiazole **10a** (6.20 mL, 57.26 mmol, 10.2 eq) and TEMED (8.47 mL, 56.14 mmol, 10 eq) was dissolved in dry THF (90 mL) under argon atmosphere and cooled to −78 °C. *n*-BuLi (2.5 M in hexanes, 22,45 mL, 56.14 mmol, 10 eq) was added dropwise and the reaction stirred for 1.5 h at the same temperature. Afterwards, **18** (3.36 g, 5.61 mmol, 1 eq) was dissolved in dry THF (10 mL) and was added slowly to the reaction mixture and stirred for two hours at −78 °C. The reaction mixture was quenched with saturated NH_4Cl solution (15 mL). The organic phase was separated and washed three times each with saturated $NaHCO_3$ solution (20 mL) and saturated NaCl solution (20 mL). The organic phase was dried over anhydrous $NaSO_4$ and concentrated under reduced pressure. The crude product was purified by column chromatography (CH/EA, 7:1 to 3:1) to yield the product as an orange oil (2.52 g, 3.75 mmol, 67%). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 8.36 – 8.27 (m, 1H), 8.24 – 8.15 (m, 1H), 8.00 – 7.93 (m, 1H), 7.60 – 7.49 (m, 2H), 7.41 – 7.24 (m, 10H), 5.44 – 5.33 (m, 1H), 5.13 (d, *J* = 14.6 Hz, 4H), 1.80 – 1.51 (m, 6H), 1.44 (s, 9H), ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 163.6, 156.0, 153.9, 153.6, 137.3, 136.8, 134.7, 128.9, 128.8, 128.6, 128.5, 128.2, 128.1, 128.0, 127.2, 126.0, 122.4, 80.0, 68.2, 67.2, 56.3, 40.9, 32.8, 28.6, 28.4, 22.9, FT-IR: ν/cm^{-1} = 754, 1050, 1139, 1163, 1215, 1256, 1319, 1382, 1638, 1695, $[\alpha]_D^{20}$ = +3 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{35}H_{40}N_5O_7S]^+$ ($[M+H]^+$): 674.3, found 674.2.

25, (*S*)-2-((*Tert*-butoxycarbonyl)amino)-3-(4-aminophenyl)propanoic acid

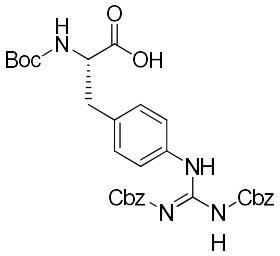


(Adapted from Kwon *et al.*)[18]

24 (1.50 g, 4.84 mmol, 1 eq) was dissolved in methanol (40 mL) and Pd/C (5%, 0.075 g) was added and the reaction stirred for 3 h under H_2-atmosphere (3 bar). The mixture was then filtered over Celite® and the solvent was evaporated under reduced pressure. The crude product was used without further purification for the next step.

MS (ESI) *m/z* calculated for $[C_{14}H_{21}N_2O_4]^+$ ($[M+H]^+$): 281.1, found 303.1 ($[M+Na]^+$).

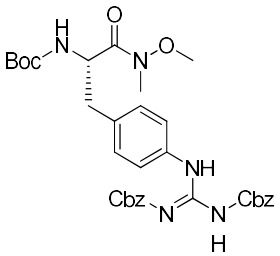
26, (*S*)-3-(4-(2,3-Bis((benzyloxy)carbonyl)guanidino)phenyl)-2-((*tert*-butoxycarbonyl) amino)propanoic acid



(Adapted from Kwon *et al.*)[18]

25 (1.35 g, 4.84 mmol, 1 eq), *N*,*N*´-bis-(carbobenzoxy)-1*H*-pyrazole-1-carboxamidine (2.19 g, 5.80 mmol, 1.2 eq) were dissolved in DMF (30 mL) and Et_3N (1.00 mL, 7.25 mmol, 1.5 eq) was added under argon atmosphere. The reaction stirred for 6 h and got quenched by addition of NH_4Cl (10 mL). The mixture was extracted with EA (30 mL), dried over $MgSO_4$ and the organic solvents were removed under reduced pressure. The crude product was purified by column chromatography (CH/EA, 4:1 to 2:1 + 0.1% TFA) to yield the product as a colorless solid (2.86 g, 4.84 mmol, quant.). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 7.46 – 7.27 (m, 10H), 7.14 (d, *J* = 7.9 Hz, 4H), 5.28 – 5.10 (m, 4H), 4.53 (s, 1H), 3.13 – 2.98 (m, 2H), 1.41 (s, 9H), ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 175.0, 161.9, 156.1, 154.5, 149.3, 135.2, 135.1, 130.7, 129.4, 129.3, 129.1, 129.0, 128.7, 123.9, 69.0, 56.0, 38.8, 28.8, FT-IR: ν/cm^{-1} = 752, 911, 1003, 1172, 1378, 1456, 1496, 1532, 1726, 1792, mp: 117 – 122 °C, $[\alpha]_D^{20}$ = +42 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{31}H_{35}N_4O_8]^+$ ([M+H]$^+$): 591.2, found 591.0.

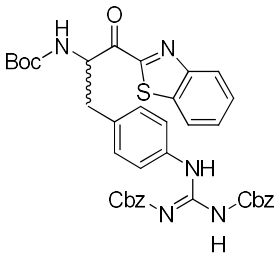
27, (*S*) -(1-(Methoxy(methyl)amino)-1-oxo -3-(4-(2,3-bis((benzyloxy)carbonyl) guanidino)phenyl)-2-((*tert*-butoxycarbonyl)amino)carbamate



(Adapted from Kwon *et al.*)[18]

26 (3.00 g, 5.08 mmol, 1 eq), TBTU (2.28 g, 7.12 mmol, 1.4 eq) and *N*,*O*-dimethyl hydroxylamine · HCl (0.99 g, 10.16 mmol, 2 eq) were dissolved in THF (15 mL) and DIEA (1.16 mL, 6.66 mmol, 6 eq) was added under argon atmosphere. The reaction stirred for 12 h and the mixture was extracted with EA (45 mL), washed with saturated NaCl solution (20 mL), dried over $MgSO_4$ and the was removed under reduced pressure. The crude product was purified by column chromatography (CH/EA, 2:1) to yield the product as a colorless oil (1.64 g, 2.85 mmol, 51%.). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 11.91 (s, 1H), 10.25 (s, 1H), 7.54 – 7.45 (m, 2H), 7.44 – 7.28 (m, 10H), 7.19 – 7.10 (m, 2H), 5.24 (s, 2H), 5.14 (s, 2H), 4.91 (s, 1H), 3.67 (s, 3H), 3.15 (s, 3H), 3.02 (dd, *J* = 13.7, 6.0 Hz, 1H), 2.85 (dd, *J* = 13.8, 7.0 Hz, 1H), 1.39 (s, 9H), ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 176.8, 172.9, 169.5, 160.3, 155.3, 154.1, 153.7, 130.1, 128.9, 128.6, 128.2, 128.1, 122.6, 68.7, 67.59, 61.7, 51.6, 38.3, 32.2, 28.5, FT-IR: ν/cm^{-1} = 754, 1055, 1220, 1265, 1294, 1367, 1383, 1422, 1608, 1639, $[\alpha]_D^{20}$ = +21 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{33}H_{40}N_5O_8]^+$ ([M+H]$^+$): 634.3, found 634.0.

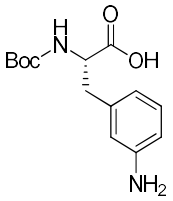
28, (*S*/*R*) -1-Oxo -3-(4-(2,3-bis((benzyloxy)carbonyl)guanidino)phenyl)-2-((*tert*-butoxycarbonyl)amino)2-benzo[*d*]thiazol



(Adapted from Kwon *et al.*)[18]

Benzothiazole **10a** (2.86 mL g, 26.42 mmol, 10.2 eq) and TEMED (3.91 mL, 25.90 mmol, 10 eq) were dissolved in dry THF (35 mL) under argon atmosphere and was cooled to −78 °C. *n*-BuLi (2.5 M in hexanes, 10.36 mL, .25.90 mmol, 10 eq) was added dropwise and the reaction stirred for 1.5 h at the same temperature. Afterwards, **27** (1.64 g, 2.59 mmol, 1 eq) was dissolved in dry THF (15 mL) and was added slowly to the reaction mixture and stirred for two hours at −78 °C. The reaction mixture was quenched with saturated NH_4Cl solution (25 mL). The organic phase was separated and washed three times each with saturated $NaHCO_3$ solution (25 mL) and saturated NaCl solution (25 mL). The organic phase was dried over anhydrous $NaSO_4$ and concentrated under reduced pressure. The crude product was purified by column chromatography (CH/EA, 9:1 to 2:1) to yield the product as a brown solid (1.18 g, 1.66 mmol, 64%). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 8.24 (dd, *J* = 7.4, 1.7 Hz, 1H), 8.02 – 7.96 (m, 1H), 7.58 (pd, *J* = 7.2, 1.5 Hz, 2H), 7.46 (d, *J* = 8.4 Hz, 2H), 7.41 – 7.28 (m, 10H), 7.09 (d, *J* = 8.0 Hz, 2H), 5.86 (d, *J* = 7.2 Hz, 1H), 5.24 (s, 2H), 5.13 (s, 2H), 3.43 (dd, *J* = 14.1, 5.1 Hz, 1H), 3.25 – 3.15 (m, 1H), 1.41 (s, 9H), ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 183.7, 164.0, 155.2, 154.6, 154.0, 153.7, 153.5, 137.4, 127.3, 129.1, 128.9, 128.6, 128.2, 126.0, 122.6, 68.8, 67.8, 57.7, 38.5, 28.4, FT-IR: ν/cm^{-1} = 733, 755, 1055, 1218, 1294, 1367, 1422, 1607, 1632, 1694, mp: 175 – 178 °C, $[\alpha]_D^{20}$ = +6 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{38}H_{38}N_5O_7S]^+$ ([M+H]$^+$): 708.2, found 708.3.

41, (*S*)-2-((*Tert*-butoxycarbonyl)amino)-3-(3-aminophenyl)propanoic acid

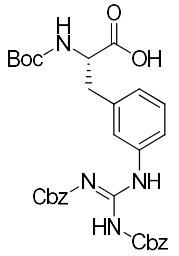


(Adapted from Kwon *et al.*)[18]

40 (1.00 g, 3.22 mmol, 1 eq) was dissolved in methanol (26 mL) and Pd/C (5%, 0.05 g) was added and the reaction stirred for 3 h under H_2-atmosphere (3 bar). The mixture was then filtered over Celite® and the solvent was evaporated under reduced pressure. The crude product was used without further purification for the next step.

MS (ESI) *m/z* calculated for $[C_{14}H_{21}N_2O_4]^+$ ([M+H]$^+$): 281.1, found 181.0 ([M+H–Boc]$^+$).

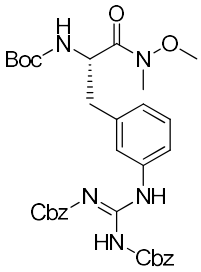
42, (*S*)-3-(3-(2,3-Bis((benzyloxy)carbonyl)guanidino)phenyl)-2-((tert-butoxycarbonyl) amino)propanoic acid



(Adapted from Kwon *et al.*)[18]

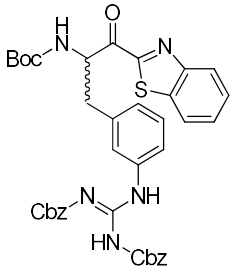
41 (0.90 g, 3.22 mmol, 1 eq), *N*,*N*´-bis-(carbobenzoxy)-1*H*-pyrazole-1-carboxamidine (1.46 g, 3.86 mmol, 1.2 eq) were dissolved in DMF (30 mL) and Et_3N (0.67 mL, 4.83 mmol, 1.5 eq) was added under argon atmosphere. The reaction stirred for 6 h and got quenched by addition of NH_4Cl (10 mL). The mixture was extracted with EA (30 mL), dried over $MgSO_4$ and the organic solvents were removed under reduced pressure. The crude product was purified by column chromatography (CH/EA, 3:1 to 1:1 + 0.1% TFA) to yield the product as a colorless oil (1.70 g, 2.89 mmol, 90%). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 7.49 – 7.08 (m, 14H), 5.35 – 5.26 (m, 2H), 5.22 – 5.10 (m, 2H), 4.59 – 4.37 (m, 1H), 3.10 (s, 2H), 1.50 – 1.29 (m, 9H), ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 173.9, 173.0, 160.6, 160.1, 153.8, 146.8, 130.3, 129.5, 129.0, 128.7, 128.6, 117.3, 113.5, 80.4, 70.3, 69.9, 54.7, 38.6, 28.4, FT-IR: ν/cm^{-1} = 731, 900, 1051, 1164, 1209, 1318, 1504, 1608, 1701, 1748, $[\alpha]_D^{20}$ = +20 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{31}H_{35}N_4O_8]^+$ ($[M+H]^+$): 591.2, found 591.1.

43, (*S*) -(1-(Methoxy(methyl)amino)-1-oxo -3-(3-(2,3-bis((benzyloxy)carbonyl) guanidino)phenyl)-2-((*tert*-butoxycarbonyl)amino)carbamate



42 (2.26 g, 3.83 mmol, 1 eq), TBTU (1.47 g, 4.60 mmol, 1.2 eq) and *N*,*O*-dimethyl hydroxylamine · HCl (2.24 g, 23.00 mmol, 6 eq) were dissolved in THF/DMF (4:1, 20 mL) and DIEA (2.67 mL, 15.32 mmol, 4 eq) was added under argon atmosphere. The reaction was stirred for 12 h and the mixture was extracted with EA (40 mL), washed with saturated NaCl solution (20 mL), dried over $MgSO_4$ and the was removed under reduced pressure. The crude product was purified by column chromatography (CH/EA, 2:1) to yield the product as a colorless oil (1.28 g, 2.02 mmol, 53%.). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 7.55 – 7.20 (m, 14H), 6.97 (d, *J* = 7.6 Hz, 1H), 5.27 – 5.11 (m, 4H), 4.91 (s, 1H), 3.63 (s, 3H), 3.13 (s, 3H), 3.05 (dd, *J* = 13.6, 5.7 Hz, 1H), 2.85 (dd, *J* = 13.7, 7.6 Hz, 1H), 1.39 (s, 9H), ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 172.2, 163.8, 155.2, 153.6, 137.8, 136.4, 134.6, 129.0, 128.8, 128.6, 128.5, 128.2, 128.0, 126.5, 123.4, 121.2, 79.7, 68.6, 67.5, 61.6, 51.6, 38.8, 32.1, 28.4, FT-IR: ν/cm^{-1} = 695, 731, 905, 1049 1116, 1161, 1231, 1422, 1600, 1729, $[\alpha]_D^{20}$ = +26 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{33}H_{40}N_5O_8]^+$ ($[M+H]^+$): 634.3, found 634.2.

44, (*S*/*R*) -1-Oxo -3-(3-(2,3-bis((benzyloxy)carbonyl)guanidino)phenyl)-2-((*tert*-butoxycarbonyl)amino)2-benzo[*d*]thiazol

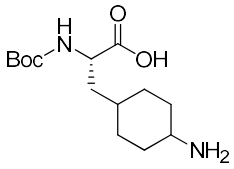


(Adapted from Kwon *et al.*)[18]

Benzothiazole **10a** (2.22 mL g, 20.60 mmol, 10.2 eq) and TEMED (3.04 mL, 20.20 mmol, 10 eq) were dissolved in dry THF (30 mL) under argonatmosphere and was cooled to −78 °C. *n*-BuLi (2.5 M in hexanes, 8.08 mL, .20.20 mmol, 10 eq) was added dropwise and the reaction stirred for 1.5 h at the same temperature. Afterwards, **43** (1.28 g, 2.02 mmol, 1 eq) was dissolved in dry THF (10 mL) and was added slowly to the reaction mixture and stirred for two hours at −78 °C. The reaction mixture was quenched with saturated NH_4Cl solution (20 mL). The organic phase was separated and washed three times each with saturated $NaHCO_3$ solution (20 mL) and saturated NaCl solution (20 mL). The organic phase was dried over anhydrous $NaSO_4$ and concentrated under reduced pressure. The crude product was purified by column chromatography (CH/EA, 7:1 to 3:1) to yield the product as a brown solid (1.20 g, 1.69 mmol, 84%). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 8.23 (d, *J* = 7.7 Hz, 1H), 8.02 – 7.93 (m, 1H), 7.63 – 7.51 (m, 3H), 7.44 – 7.17 (m, 12H), 6.95 (d, *J* = 7.6 Hz, 1H), 5.92 – 5.79 (m, 1H), 5.38 – 5.06 (m, 5H), 3.50 – 3.05 (m, 2H), 1.40 (s, 9H) ppm, ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 192.9, 163.8, 155.2, 153.9, 153.5, 137.3, 137.2, 136.1, 129.4, 129.0, 128.8, 128.7, 128.5, 128.1, 127.3, 126.9, 126.0, 123.6, 122.5, 121.6, 86.3, 80.1, 68.8, 67.8, 57.5, 52.8, 38.6, 28.3 ppm, FT-IR: ν/cm^{-1} = 726, 1046, 1167, 1240, 1366, 1422, 1484, 1639, 1687, 2974, mp: 164 – 166 °C, $[\alpha]_D^{20}$ = +5 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{38}H_{38}N_5O_7S]^+$ ($[M+H]^+$): 708.2, found 708.1.

w

34, (*S*)-2-((*Tert*-butoxycarbonyl)amino)-3-(4-aminocyclohexyl)propanoic acid

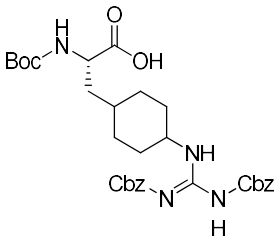


(Adapted from Kwon *et al.*)[18]

24 (1.50 g, 4.84 mmol, 1 eq) was dissolved in AcOH/MeOH (1:1, 30 mL) and PtO_2 (0.06 g, 0.48 mmol, 0.1 eq) was added. The reaction stirred for 24 h under H_2 atmosphere (3 bar) at rt. The mixture was filtered over Celite® and the solution was removed under reduced pressure. The crude product was used without further purification for the next step.

MS (ESI) *m/z* calculated for $[C_{14}H_{26}N_2O_4]^+$ ($[M+H]^+$): 286.2, found 326.0 ($[M+K]^+$).

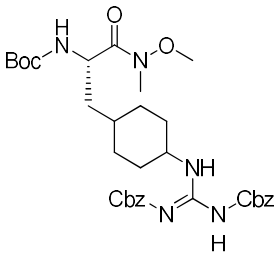
35, (*S*)-3-(4-(2,3-Bis((benzyloxy)carbonyl)guanidino)cyclohexyl)-2-((tert-butoxycarbonyl)amino)propanoic acid



(Adapted from Kwon *et al.*)[18]

34 (1.19 g, 4.16 mmol, 1 eq), *N*,*N´* bis-(carbobenzoxy)-1-*H*-pyrazol-1-carboxamidin (2.01 g, 5.32 mmol, 1.1 eq) were dissolved in DMF (30 mL) and Et_3N (0.86 mL, 6.25 mmol, 1.5 eq) was added under argon atmosphere. The reaction stirred for 6 h and got quenched by addition of NH_4Cl (10 mL). The mixture was extracted with EA (30 mL), dried over $MgSO_4$ and the organic solvents were removed under reduced pressure. The crude product was purified by column chromatography (DCM/MeOH, 100:1 to 11:1) to yield the product as a colorless solid (2.48 g, 4.16 mmol, quant.). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 7.41 – 7.33 (m, 10H), 5.21 (d, *J* = 8.3 Hz, 2H), 5.02 (s, 2H), 4.10 – 3.95 (m, 1H), 3.88 (d, *J* = 7.5 Hz, 1H), 1.90 – 1.41 (m, 6H), 1.34 (d, *J* = 17.7 Hz, 9H), 1.03 (ddd, *J* = 50.6, 24.1, 12.0 Hz, 5H) ppm, ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 163.1, 155.5, 155.1, 154.3, 152.8, 135.3, 128.5, 128.4, 128.2, 122.2, 77, 67.8, 66.4, 56.8, 51.9, 38.4, 31.6, 30.2, 30.1, 29.4, 28.3, 28.2 ppm, FT-IR: ν/cm^{-1} = 696, 733, 1053, 1124, 1162, 1206, 1295, 1367, 1420, 1635, mp: 98 – 112 °C, $[\alpha]_D^{20}$ = +18 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{31}H_{41}N_4O_8]^+$ ([M+H]$^+$): 597.3, found 597.3.

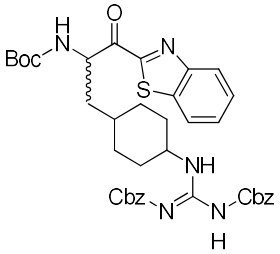
36, (*S*) -(1-(Methoxy(methyl)amino)-1-oxo -3-(4-(2,3 bis((benzyloxy)carbonyl) guanidino)cyclohexyl)-2-((tert-butoxycarbonyl)amino)carbamate



(Adapted from Kwon *et al.*)[18]

35 (0.66 g, 1.11 mmol, 1 eq), TBTU (0.49 g, 1.55 mmol, 1.4 eq) and *N*,*O*-dimethyl hydroxylamine · HCl (0.21 g, 2.22 mmol, 2 eq) were dissolved in THF (7 mL) and DIEA (1.16 mL, 6.66 mmol, 6 eq) was added under argon atmosphere. The reaction stirred for 12 h and and the mixture was extracted with EA (20 mL), washed with saturated NaCl solution (10 mL), dried over $MgSO_4$ and the was removed under reduced pressure. The crude product was purified by column chromatography (CH/EA, 2:1) to yield the product as a colorless solid (1.64 g, 2.85 mmol, 51%.). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 7.49 (d, *J* = 8.5 Hz, 1H), 7.43 – 7.28 (m, 10H), 7.13 (d, *J* = 8.5 Hz, 1H), 5.24 – 5.05 (m, 4H), 4.71 (d, *J* = 9.2 Hz, 1H), 3.77 (d, *J* = 5.8 Hz, 2H), 3.67 (s, 1H), 3.24 – 3.13 (m, 3H), 3.02 (dd, *J* = 13.7, 6.0 Hz, 1H), 2.87 (d, *J* = 7.1 Hz, 1H), 2.07 – 1.51 (m, 7H), 1.41 (d, *J* = 14.9 Hz, 9H), 1.28 – 1.03 (m, 4H) ppm, ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 156.4, 155.1, 154.0, 149.6, 136.9, 128.8, 128.5, 128.2, 79.8, 67.4, 61.7, 52.7, 32.7, 32.3, 32.3, 29.5, 28.5 ppm, FT-IR: ν/cm^{-1} = 1053,1170, 1227, 1295, 1337, 1366, 1383, 1428, 1639, 1711, mp: 70 – 77 °C, $[\alpha]_D^{20}$ = –20 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{33}H_{46}N_5O_8]^+$ ([M+H]$^+$): 640.3, found 640.3.

37, (*S*/*R*) -1-Oxo -3-(4-(2,3-bis((benzyloxy)carbonyl)guanidino)cyclohexyl)-2-((tert-butoxycarbonyl)amino)2-benzo[*d*]thiazol



(Adapted from Kwon *et al.*)[18]

Benzothiazole **10a** (0.88 mL g, 8,02 mmol, 10.2 eq) and TEMED (1.19 mL, 7.86 mmol, 10 eq) were dissolved in dry THF (13 mL) under argonatmosphere and was cooled to −78 °C. *n*-BuLi (2.5 M in hexanes, 3.15 mL, .7.86 mmol, 10 eq) was added dropwise and the reaction stirred for 40 min at the same temperature. Afterwards, **36** (0.50 g, 0.78 mmol, 1 eq) was dissolved in dry THF (5 mL) and was added slowly to the reaction mixture and stirred for two hours at −78 °C. The reaction mixture was quenched with saturated NH_4Cl solution (20 mL). The organic phase was separated and washed three times each with saturated $NaHCO_3$ solution (20 mL) and saturated NaCl solution (20 mL). The organic phase was dried over anhydrous $NaSO_4$ and concentrated under reduced pressure. The crude product was purified by column chromatography (CH/EA, 9:1 to 4:1) to yield the product as a yellow solid (0.42 g, 0.59 mmol, 76%). ^{1}H NMR (300 MHz, CDCl3): δ/ppm = 8.26 – 8.18 (m, 1H), 8.02 – 7.95 (m, 1H), 7.57 (dddd, *J* = 8.8, 6.7, 4.0, 2.1 Hz, 2H), 7.42 – 7.27 (m, 10H), 5.26 (d, *J* = 17.0 Hz, 2H), 5.20 – 5.11 (m, 3H), 3.43 (dd, *J* = 14.0, 5.1 Hz, 1H), 3.21 (dd, *J* = 14.1, 6.9 Hz, 1H), 2.30 – 1.48 (m, 8H), 1.43 (d, *J* = 9.0 Hz, 9H), 1.32 – 1.12 (m, 3H), ^{13}C NMR (75 MHz, CDCl3): δ/ppm = 193.0, 164.0, 160.3, 156.3, 154.9, 149.7, 136.7, 136.6, 128.9, 128.6, 128.1, 127.3, 126.0, 122.6, 80.1, 68.6, 67.6, 50.4, 33.7, 30.5, 28.4, 28.3, FT-IR: ν/cm^{-1} = 733, 755, 1055, 1218, 1294, 1367, 1422, 1607, 1632, 1694, mp: 82 – 93 °C, $[\alpha]_D^{20}$ = +6 (*c* 1.00, $CHCl_3$), MS (ESI) *m/z* calculated for $[C_{38}H_{43}N_5O_7S]^+$ ($[M+H]^+$): 714.2, found 714.0.

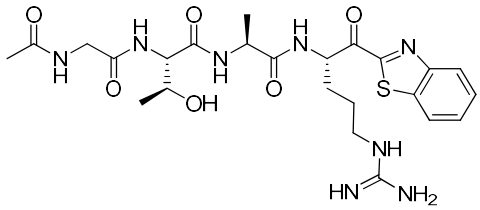
General procedure for the amide bond formation between the H_2N-P1-warhead moiety and the peptide residue

The respective Boc-NH-P1-warhead moieties **12a**–**f**, **13**, **19, 27, 36, 43** were deprotected by using a deprotection mixture of 4 M HCl in dioxane (4 ml) or 16% TFA in DCM (4 mL) and stirred for 30 min at rt. The solvent was removed under reduced pressure to yield the deprotected amines as hydrochloride or trifluoroacetate salts.

The carboxylic acid of different peptide backbones **3**, **21**–**23** (1.2 eq) was dissolved in DCM and cooled to 0 °C with an ice-water bath. DIEA (6 eq) and HATU (1.2 eq) were added and stirring was continued for 30 min at 0 °C. Then, the respective deprotected amine (1 eq) as a solution in DMF were added, stirred for 30 min at 0 °C and 16 h at room temperature. DCM and water were added, and the aqueous phase was extracted three times with DCM. The combined organic phases were dried over Na_2SO_4, and the solvent was removed under reduced pressure. The crude product was purified with reversed-phase flash chromatography performed with the Biotage Isolera™ and used without further analysis for the next step.

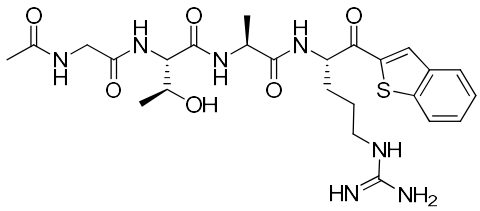
Synthesis of uPA targeting compounds

14a, (2*S*,3*S*)-2-(2-Acetamidoacetamido)-*N*-((*S*)-1-(((*S*)-1-(benzo[*d*]thiazol-2-yl)-5-guanidino-1-oxopentan-2-yl)amino)-1-oxopropan-2-yl)-3-hydroxybutanamide



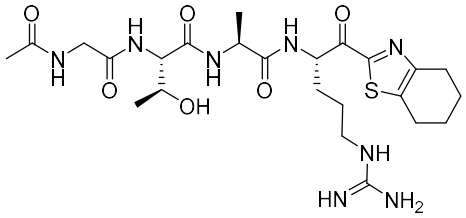
After the coupling of the respective peptide backbone **3** with the P1-warhead moiety **12a**, the deprotection of the Pbf- and O^{t}Bu-protecting group was performed in a TFA/DCM solution (50%, 2 mL) and stirred for 2 h at room temperature. The solution was evaporated under reduced pressure and purified by preparative HPLC (12.8 mg, 0.017 mmol, 10%). ^{1}H NMR (400 MHz, DMSO-d_6): δ/ppm = 8.57 – 8.39 (m, 1H), 8.31 – 8.24 (m, 2H), 8.23 – 8.16 (m, 1H), 8.08 – 7.96 (m, 1H), 7.83 – 7.72 (m, 1H), 7.71 – 7.63 (m, 2H), 7.60 – 7.54 (m, 1H), 5.54 – 5.41 (m, 1H), 5.00 (d, *J* = 35.3 Hz, 1H), 4.36 (ddd, *J* = 10.0, 7.2, 2.7 Hz, 1H), 4.17 (tdd, *J* = 12.1, 8.0, 4.2 Hz, 1H), 3.97 (ddt, *J* = 16.9, 11.2, 5.2 Hz, 1H), 3.83 – 3.69 (m, 2H), 3.14 (p, *J* = 6.6 Hz, 2H), 1.97 (td, *J* = 13.7, 11.4, 6.7 Hz, 2H), 1.90 – 1.82 (m, 2H), 1.79 – 1.69 (m, 0H), 1.65 – 1.54 (m, 2H), 1.27 – 1.17 (m, 3H), 1.06 – 0.98 (m, 3H). ^{13}C NMR (400 MHz, DMSO-d_6): δ/ppm = 193.1, 172.6, 169.9, 169.7, 169.5, 164.5, 156.8, 153.0, 136.5, 128.3, 127.6, 125.3, 123.3, 66.7, 58.2, 54.3, 47.9, 42.2, 42.2, 27.8, 25.2, 22.4, 19.7, 18, mp: 68 – 75 °C, $[\alpha]_D^{20}$ = +34 (*c* 1.00, DMSO), FT-IR: ν/cm^{-1} = 660, 668, 756, 1136, 1215, 1673, 1684, 2922, 2964, 3327, MS (ESI) *m/z* calculated for $[C_{24}H_{35}N_8O_6S]^+$ ([M+H]$^+$): 563.2, found 563.2. Purity: 97%.

14f, (2*S*,3*S*)-2-(2-Acetamidoacetamido)-*N*-((*S*)-1-(((*S*)-1-(benzo[b]thiophen-2-yl)-5-guanidino-1-oxopentan-2-yl)amino)-1-oxopropan-2-yl)-3-hydroxybutanamide



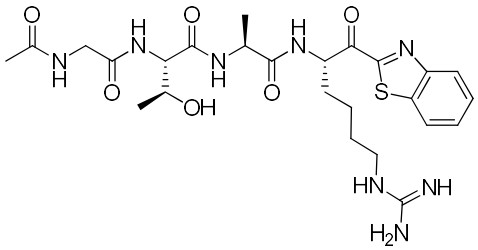
After the coupling of the respective peptide backbone **3** with the P1-warhead moiety **12f**, the deprotection of the Pbf- and O^{t}Bu-protecting group was performed in a TFA/DCM solution (50%, 2 mL) and stirred for 2 h at room temperature. The solution was evaporated under reduced pressure and purified by preparative HPLC (16.8 mg, 0.03 mmol, 8%). ^{1}H NMR (600 MHz, DMSO-d_6): δ/ppm = 8.53 (d, *J* = 7.0 Hz, 1H), 8.46 (d, *J* = 4.7 Hz, 1H), 8.43 – 8.35 (m, 1H), 8.33 – 8.25 (m, 1H), 8.19 (dt, *J* = 8.6, 6.1 Hz, 1H), 8.11 – 7.97 (m, 3H), 7.81 (dd, *J* = 17.4, 8.0 Hz, 1H), 7.58 – 7.48 (m, 2H), 5.27 (dddd, *J* = 32.2, 9.1, 7.5, 4.8 Hz, 1H), 5.00 (dd, *J* = 29.7, 5.0 Hz, 1H), 4.35 – 4.26 (m, 1H), 4.22 – 4.11 (m, 1H), 4.04 – 3.91 (m, 1H), 3.81 – 3.69 (m, 2H), 3.17 – 3.08 (m, 2H), 1.86 (dd, *J* = 7.3, 6.1 Hz, 4H), 1.71 – 1.50 (m, 3H), 1.24 – 1.10 (m, 4H), 1.02 (d, *J* = 6.3 Hz, 2H) ppm, ^{13}C NMR (151 MHz, DMSO-d_6): δ/ppm = 193.3, 172.4, 169.9, 169.5, 156.6, 141.6, 141.1, 141.0, 139.1, 131.2, 130.9, 128.0, 126.5, 125.4, 123.1, 66.7, 58.7, 54.1, 48.4, 42.2, 40.3, 28.3, 25.0, 22.5, 19.7, 18.2 ppm, FT-IR: ν/cm^{-1} = 716, 757, 929, 1028, 1129, 1202, 1368, 1512, 1644, 3291, mp: 74 – 80 °C, $[\alpha]_D^{20}$ = +23 (*c* 1.00, DMSO), MS (ESI) *m/z* calculated for $[C_{25}H_{36}N_7O_6S]^+$ ([M+H]$^+$): 562.2, found 562.3. Purity: 99%.

15, (2*S*,3*S*)-2-(2-Acetamidoacetamido)-*N*-((*S*)-1-(((*S*)-5-guanidino-1-oxo-1-(4,5,6,7-tetrahydrobenzo[d]thiazol-2-yl)pentan-2-yl)amino)-1-oxopropan-2-yl)-3-hydroxybutanamide



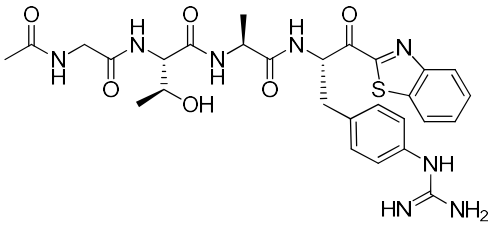
After the coupling of the respective peptide backbone **3** with the P1-warhead moiety **13**, the deprotection of the Pbf- and O^{t}Bu-protecting group was performed in a TFA/DCM solution (50%, 4 mL) and stirred for 2 h at room temperature. The solution was evaporated under reduced pressure and purified by preparative HPLC (16.9 mg, 0.03 mmol, 8%). ^{1}H NMR (600 MHz, DMSO-d_6): δ/ppm = 8.67 – 8.55 (m, 1H), 8.48 – 8.42 (m, 1H), 8.31 – 8.21 (m, 1H), 8.09 (d, *J* = 7.6 Hz, 1H), 7.89 (dd, *J* = 55.4, 8.2 Hz, 1H), 7.62 (s, 3H), 5.33 – 5.22 (m, 2H), 4.37 – 4.28 (m, 1H), 4.25 – 4.21 (m, 1H), 4.05 – 3.94 (m, 1H), 3.82 – 3.68 (m, 2H), 3.05 (p, *J* = 8.2, 7.4 Hz, 2H), 2.91 – 2.78 (m, 4H), 1.88 – 1.77 (m, 8H), 1.69 – 1.49 (m, 3H), 1.26 – 1.20 (m, 3H), 1.03 (d, *J* = 6.4 Hz, 3H) ppm, ^{13}C NMR (151 MHz, DMSO-d_6): δ/ppm = 191.2, 172.4, 170.0, 169.8, 169.5, 169.3, 167.4, 160.4, 157.2, 153.3, 138.8, 66.5, 58.0, 53.9, 48.3, 47.8, 42.1, 28.1, 26.4, 25.0, 23.6, 22.4, 19.5, 18.0 ppm, FT-IR: ν/cm^{-1} =773, 886, 939, 1026, 1231, 1369, 1428, 1532, 1642, 2937, mp: 110 – 114 °C, $[\alpha]_D^{20}$ = +27 (*c* 1.00, DMSO), MS (ESI) *m/z* calculated for $[C_{24}H_{39}N_8O_6S]^+$ ($[M+H]^+$): 567.6, found 567.1. Purity: 95%.

20, (2*S*,3*S*)-2-(2-Acetamidoacetamido)-*N*-((*S*)-1-(((*S*)-1-(benzo[d]thiazol-2-yl)-6-guanidino-1-oxohexan-2-yl)amino)-1-oxopropan-2-yl)-3-hydroxybutanamide



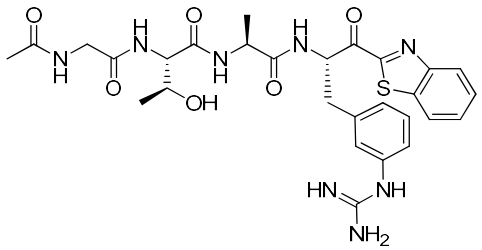
After the coupling of the respective peptide backbone **3** with the P1-warhead moiety **19**, the deprotection of the Cbz- and O^{t}Bu-protecting groups was performed in a TFA/TA solution (5:1, 10 mL) and stirred for 12 h at room temperature. The solution was evaporated under reduced pressure and purified by preparative HPLC (17.3 mg, 0.03 mmol, 8%). ^{1}H NMR (600 MHz, DMSO-d_6): δ/ppm = 8.54 – 8.33 (m, 1H), 8.25 (m, 2H), 8.18 (m, 1H), 8.06 – 7.93 (m, 1H), 7.87 – 7.54 (m, 4H), 5.56 – 5.39 (m, 1H), 4.44 – 4.31 (m, 1H), 4.25 – 4.13 (m, 1H), 4.05 – 3.92 (m, 1H), 3.80 – 3.69 (m, 2H), 3.18 – 2.98 (m, 2H), 1.95 (s, 1H), 1.89 – 1.83 (m, 3H), 1.81 – 1.34 (m, 6H), 1.28 – 1.19 (m, 3H), 1.06 – 0.98 (m, 3H), ^{13}C NMR (151 MHz, DMSO-d_6): δ/ppm = 193.3, 172.6, 169.9, 169.8, 169.7, 169.6, 169.5, 169.4, 164.5, 156.8, 153.0, 136.4, 128.3, 127.6, 126.6, 125.3, 123.3, 66.7, 54.7, 44.6, 42.2, 30.0, 28.0, 22.7, 22.4, 19.6, 19.5, 18.2, 18.0, FT-IR: ν/cm^{-1} = 668, 722, 762, 800, 1130, 1203, 1482, 1549, 1634, 3284, mp: 105 – 110 °C, $[\alpha]_D^{20}$ = +31 (*c* 1.00, DMSO), MS (ESI) *m/z* calculated for $[C_{25}H_{37}N_8O_6S]^+$ ($[M+H]^+$): 577.6, found 577.3. Purity: 98%.

29, (2*S*,3*S*)-2-(2-Acetamidoacetamido)-*N*-((*S*)-1-(((*S*)-1-(benzo[d]thiazol-2-yl)-3-(4-guanidinophenyl)-1-oxopropan-2-yl)amino)-1-oxopropan-2-yl)-3-hydroxybutanamide



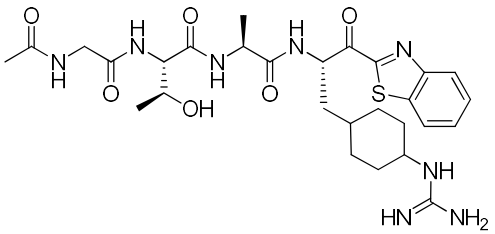
After the coupling of the respective peptide backbone **3** with the P1-warhead moiety **28**, the deprotection of the Cbz- and O^{t}Bu-protecting groups was performed in a TFA/TA solution (5:1, 10 mL) and stirred for 12 h at room temperature. The solution was evaporated under reduced pressure and purified by preparative HPLC (42.7 mg, 0.07 mmol, 25%). ^{1}H NMR (600 MHz, DMSO-d_6): δ/ppm = 8.65 – 8.51 (m, 2H), 8.31 – 8.25 (m, 2H), 8.16 (m, 1H), 7.97 – 7.90 (m, 2H), 7.70 – 7.66 (m, 2H), 7.52 (s, 4H), 7.42 – 7.37 (m, 2H), 7.20 – 7.16 (m, 2H), 5.74 – 5.67 (m, 1H), 4.98 (s, 1H), 4.38 – 4.12 (m, 3H), 4.04 – 3.90 (m, 1H), 3.75 (dd, 2H), 2.99 – 2.88 (m, 1H), 1.85 (m, 3H), 1.13 (m, 3H), 1.02 – 0.98 (m, 3H), ^{13}C NMR (151 MHz, DMSO-d_6): δ/ppm = 192.4, 172.6, 172.3, 170.0, 169.8, 169.6, 169.4, 164.2, 159.2, 158.8, 155.8, 152.9, 136.4, 135.6, 133.8, 130.4, 128.3, 127.6, 125.3, 124.4, 123.2, 58.2, 47.8, 42.1, 35.7, 22.4, 19.5, 18.2, 18.0, FT-IR: ν/cm^{-1} = 668, 721, 762, 801, 1132, 1183, 1202, 1515, 1552, 1645, mp: 125 – 133 °C, $[\alpha]_D^{20}$ = −23 (*c* 1.00, DMSO), MS (ESI) *m/z* calculated for $[C_{28}H_{35}N_8O_6S]^+$ ($[M+H]^+$): 611.2, found 611.2. Purity: 97%.

45, (2*S*,3*S*)-2-(2-Acetamidoacetamido)-*N*-((*S*)-1-(((*S*)-1-(benzo[d]thiazol-2-yl)-3-(3-guanidinophenyl)-1-oxopropan-2-yl)amino)-1-oxopropan-2-yl)-3-hydroxybutanamide



After the coupling of the respective peptide backbone **3** with the P1-warhead moiety **44**, the deprotection of the Cbz- and O^{t}Bu-protecting groups was performed in a TFA/TA solution (5:1, 10 mL) and stirred for 12 h at room temperature. The solution was evaporated under reduced pressure and purified by preparative HPLC (36.6 mg, 0.06 mmol, 40%). ^{1}H NMR (600 MHz, DMSO-d_6): δ/ppm = 8.52 (dd, *J* = 13.2, 7.4 Hz, 1H), 8.44 (d, *J* = 1.2 Hz, 1H), 8.33 – 8.28 (m, 2H), 8.21 (dd, *J* = 26.0, 5.8 Hz, 1H), 8.01 – 7.89 (m, 4H), 7.71 – 7.66 (m, 2H), 7.35 – 7.30 (m, 1H), 7.23 – 7.14 (m, 2H), 7.08 – 7.02 (m, 1H), 5.77 – 5.67 (m, 1H), 5.28 (s, 0H), 5.03 (s, 1H), 4.33 – 4.11 (m, 3H), 3.99 – 3.91 (m, 1H), 3.78 – 3.68 (m, 2H), 2.96 – 2.91 (m, 1H), 1.87 – 1.83 (m, 3H), 1.17 – 1.08 (m, 3H), 1.00 (t, *J* = 6.2 Hz, 3H), ^{13}C NMR (151 MHz, DMSO-d_6): δ/ppm = 192.4, 172.5, 172.4, 170.0, 169.7, 169.2, 167.2, 164.2, 156.1, 152.9, 138.9, 136.4, 129.6, 128.3, 127.6, 125.4, 124.6, 123.3, 122.2, 66.6, 57.7, 56.5, 48.0, 42.1, 36.2, 22.4, 19.7, 18.1, FT-IR: ν/cm^{-1} = 751, 886, 1023, 1130, 1251, 1347, 1473, 1512, 1580, 2783, mp: 103 – 106 °C, $[\alpha]_D^{20}$ = +18 (*c* 1.00, DMSO), MS (ESI) *m/z* calculated for $[C_{28}H_{35}N_8O_6S]^+$ ($[M+H]^+$): 611.2, found 611.1. Purity: 96%.

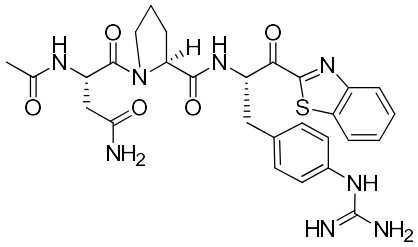
38, (2*S*,3*S*)-2-(2-Acetamidoacetamido)-*N*-((*S*)-1-(((*S*)-1-(benzo[d]thiazol-2-yl)-3-(4-guanidinocyclohexyl)-1-oxopropan-2-yl)amino)-1-oxopropan-2-yl)-3-hydroxybutanamide



After the coupling of the respective peptide backbone **3** with the P1-warhead moiety **37**, the deprotection of the Cbz- and O^{t}Bu-protecting groups was performed in a TFA/TA solution (5:1, 10 mL) and stirred for 12 h at room temperature. The solution was evaporated under reduced pressure and purified by preparative HPLC (24.6 mg, 0.04 mmol, 20%). ^{1}H NMR (600 MHz, DMSO-d_6): δ/ppm = 8.42 – 8.39 (m, 1H), 8.30 – 8.21 (m, 2H), 8.07 (d, J = 7.6 Hz, 1H), 8.02 – 7.89 (m, 1H), 7.79 (d, J = 8.2 Hz, 1H), 7.71 – 7.65 (m, 2H), 5.63 – 5.51 (m, 1H), 5.19 – 5.05 (m, 1H), 4.39 – 4.16 (m, 2H), 4.05 – 3.93 (m, 1H), 3.84 – 3.66 (m, 2H), 3.29 – 3.22 (m, 1H), 1.96 – 1.38 (m, 9H), 1.36 – 1.06 (m, 5H), 1.06 – 1.00 (m, 3H), ^{13}C NMR (151 MHz, DMSO-d_6): δ/ppm = 193.7, 172.5, 170.0, 169.4, 167.2, 164.5, 157.9, 155.9, 152.9, 136.4, 135.9, 128.3, 127.7, 125.2, 123.3, 66.6, 58.0, 52.7, 52.5, 49.7, 48.3, 47.8, 42.3, 32.9, 29.9, 22.4, 19.6, 18.1, FT-IR: ν/cm^{-1} = 657, 668, 699, 730, 764, 1139, 1203, 1551, 1655, 2343, mp: 148 – 155 °C, $[\alpha]_D^{20}$ = −25 (c 1.00, DMSO), MS (ESI) m/z calculated for $[C_{28}H_{41}N_8O_6S]^+$ ([M+H]$^+$): 617.3, found 617.3. Purity: 97%.

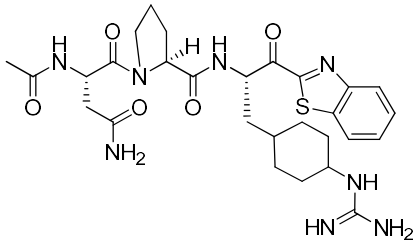
Synthesis of TMPRSS2 targeting compounds

30, (*S*)-1-((*S*)-2-Acetamido-4-amino-4-oxobutanoyl)-*N*-((*S*)-1-(benzo[d]thiazol-2-yl)-3-(4-guanidinophenyl)-1-oxopropan-2-yl)pyrrolidine-2-carboxamide



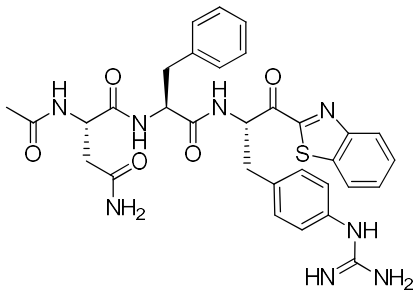
After the coupling of the respective peptide backbone **21** with the P1-warhead moiety **28**, the deprotection of the Cbz- and O^{t}Bu-protecting groups was performed in a TFA/TA solution (5:1, 10 mL) and stirred for 12 h at room temperature. The solution was evaporated under reduced pressure and purified by preparative HPLC (50.0 mg, 0.08 mmol, 19%). ^{1}H NMR (600 MHz, DMSO-d_6): δ/ppm = δ 8.45 – 8.41 (m, 1H), 8.39 – 8.37 (m, 1H), 8.32 – 8.26 (m, 2H), 7.85 (s, 2H), 7.72 – 7.64 (m, 2H), 7.63 – 7.59 (m, 1H), 7.46 – 7.41 (m, 1H), 7.34 (d, J = 8.3 Hz, 1H), 7.20 – 7.10 (m, 2H), 7.08 – 7.05 (m, 1H), 5.71 – 5.54 (m, 1H), 4.81 – 4.75 (m, 1H), 4.29 (ddd, J = 7.9, 4.8, 2.4 Hz, 1H), 3.78 – 3.64 (m, 2H), 3.33 – 3.28 (m, 2H), 3.06 – 2.96 (m, 1H), 2.71 – 2.59 (m, 1H), 2.43 – 2.34 (m, 1H), 1.96 – 1.61 (m, 7H), ^{13}C NMR (151 MHz, DMSO-d_6): δ/ppm = 192.6, 172.3, 172.1, 171.7, 171.5, 170.6, 170.3, 168.9, 167.7, 164.4, 164.1, 156.1, 153.0, 136.4, 130.3, 128.3, 127.6, 125.3, 124.0, 123.7, 59.7, 59.2, 56.8, 47.0, 37.6, 34.9, 29.5, 23.6, 22.3, FT-IR: ν/cm^{-1} = 759, 877, 1021, 1127, 1203, 1313, 1375, 1420, 1619, 2364, mp: 200 – 203 °C, $[\alpha]_D^{20}$ = −59 (c 1.00, DMSO), MS (ESI) m/z calculated for $[C_{28}H_{33}N_8O_5S]^+$ ([M+H]$^+$): 593.2, found 593.1. Purity: 97%.

39, (*S*)-1-((*S*)-2-Acetamido-4-amino-4-oxobutanoyl)-*N*-((*S*)-1-(benzo[d]thiazol-2-yl)-3-(4-guanidinocyclohexyl)-1-oxopropan-2-yl)pyrrolidine-2-carboxamide



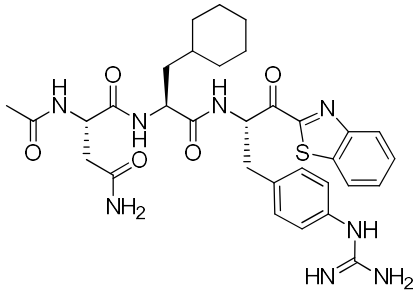
After the coupling of the respective peptide backbone **21** with the P1-warhead moiety **37**, the deprotection of the Cbz- and O^{t}Bu-protecting groups was performed in a TFA/TA solution (5:1, 10 mL) and stirred for 12 h at room temperature. The solution was evaporated under reduced pressure and purified by preparative HPLC (27.8 mg, 0.05 mmol, 10%). ^{1}H NMR (600 MHz, DMSO-d_6): δ/ppm = δ 8.40 – 8.31 (m, 1H), 8.29 – 8.20 (m, 2H), 8.17 – 8.13 (m, 1H), 7.70 – 7.65 (m, 2H), 7.59 – 7.47 (m, 1H), 7.43 – 7.36 (m, 1H), 5.67 – 5.58 (m, 1H), 5.52 – 5.38 (m, 1H), 4.80 – 4.74 (m, 1H), 4.35 – 4.28 (m, 1H), 3.86 – 3.55 (m, 3H), 3.32 – 3.25 (m, 1H), 2.74 – 2.60 (m, 1H), 2.43 – 2.33 (m, 1H), 2.05 – 0.93 (m, 20H), ^{13}C NMR (151 MHz, DMSO-d_6): δ/ppm = 193.9, 172.8, 172.2, 170.6, 169.3, 165.1, 164.7, 158.6, 156.2, 153.3, 136.7, 128.6, 128.1, 125.5, 123.8, 118.7, 116.7, 60.4, 59.6, 53.3, 50.2, 47.7, 37.7, 33.4, 29.6, 24.1, 23.5, 22.7, FT-IR: ν/cm^{-1} = 720, 835, 990, 1023, 1119, 1200, 1431, 1538, 1619, 2347, mp: 113 – 116 °C, $[\alpha]_D^{20}$ = −60 (*c* 1.00, DMSO), MS (ESI) *m/z* calculated for $[C_{28}H_{38}N_8O_5S]^+$ ([M+H]$^+$): 599.2, found 599.1. Purity: 98%.

31, (*S*)-2-Acetamido-*N*-((*S*)-1-(((*S*)-1-(benzo[d]thiazol-2-yl)-3-(4-guanidinophenyl)-1-oxopropan-2-yl)amino)-1-oxo-3-phenylpropan-2-yl)succinamide



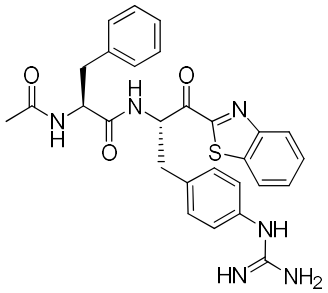
After the coupling of the respective peptide backbone **22** with the P1-warhead moiety **28**, the deprotection of the Cbz- and O^{t}Bu-protecting groups was performed in a TFA/TA solution (5:1, 10 mL) and stirred for 12 h at room temperature. The solution was evaporated under reduced pressure and purified by preparative HPLC (20.5 mg, 0.03 mmol, 11%). ^{1}H NMR (600 MHz, DMSO-d_6): δ/ppm = 8.80 – 8.69 (m, 1H), 8.34 – 8.29 (m, 1H), 8.13 – 7.89 (m, 2H), 7.75 – 7.68 (m, 2H), 7.49 – 7.36 (m, 6H), 7.33 – 7.25 (m, 1H), 7.21 – 7.14 (m, 5H), 7.03 (ddt, *J* = 5.9, 4.4, 1.7 Hz, 1H), 6.97 – 6.92 (m, 1H), 5.79 – 5.71 (m, 1H), 4.55 – 4.44 (m, 2H), 3.10 – 2.61 (m, 4H), 2.35 – 2.13 (m, 2H), 1.84 – 1.73 (m, 3H), ^{13}C NMR (151 MHz, DMSO-d_6): δ/ppm = 192.5, 171.6, 171.3, 169.4, 164.3, 164.2, 155.7, 153.0, 137.5, 136.4, 136.0, 133.9, 130.4, 129.2, 128.4, 128.0, 127.7, 126.3, 125.4, 124.4, 123.8, 118.2, 116.3, 56.4, 53.7, 53.4, 49.9, 49.5, 37.3, 35.5, 22.5, FT-IR: ν/cm^{-1} = 700, 728, 779, 877, 1122, 1195, 1377, 1552, 1656, 2358, mp: 136 – 138 °C, $[\alpha]_D^{20}$ = −57 (*c* 1.00, DMSO), MS (ESI) *m/z* calculated for $[C_{32}H_{35}N_8O_5S]^+$ ([M+H]$^+$): 643.2, found 643.1. Purity: 96%.

32, (*S*)-2-Acetamido-*N*-((*S*)-1-(((*S*)-1-(benzo[d]thiazol-2-yl)-3-(4-guanidinophenyl)-1-oxopropan-2-yl)amino)-1-oxo-3-cyclohexylpropan-2-yl)succinamide



After the coupling of the respective peptide backbone **23** with the P1-warhead moiety **28**, the deprotection of the Cbz- and O^{t}Bu-protecting groups was performed in a TFA/TA solution (5:1, 10 mL) and stirred for 12 h at room temperature. The solution was evaporated under reduced pressure and purified by preparative HPLC (35.2 mg, 0.05 mmol, 19%). ^{1}H NMR (600 MHz, DMSO-d_6): δ/ppm = 8.63 – 8.52 (m, 1H), 8.41 (s, 1H), 8.32 – 8.26 (m, 1H), 8.19 – 8.13 (m, 1H), 7.98 – 7.85 (m, 4H), 7.74 – 7.62 (m, 2H), 7.39 – 7.33 (m, 2H), 7.16 – 7.10 (m, 2H), 6.99 – 6.92 (m, 1H), 5.68 – 5.59 (m, 1H), 4.56 – 4.48 (m, 1H), 4.31 – 4.23 (m, 1H), 3.32 – 3.26 (m, 1H), 3.08 (s, 1H), 3.05 – 2.94 (m, 1H), 2.48 – 2.32 (m, 2H), 1.87 – 1.75 (m, 3H), 1.60 – 1.00 (m, 11H), 0.82 – 0.68 (m, 2H), ^{13}C NMR (151 MHz, DMSO-d_6): δ/ppm = 192.4, 172.4, 171.6, 171.1, 169.6, 169.3, 167.6, 164.2, 161.5, 156.1, 155.1, 153.6, 152.9, 143.7, 136.4, 135.1, 130.3, 129.7, 128.3, 127.6, 125.4, 124.2, 123.7, 123.3, 56.6, 49.8, 37.0, 35.2, 33.1, 31.5, 26.0, 22.6, FT-IR: ν/cm^{-1} =664, 754, 1015, 1203, 1377, 1512, 1538, 1636, 2355, 2920, mp: 153 – 155 °C, $[\alpha]_D^{20}$ = −60 (*c* 1.00, DMSO), MS (ESI) *m/z* calculated for $[C_{32}H_{41}N_8O_5S]^+$ ([M+H]$^+$): 649.2, found]$^+$): 649.2, Purity: 96%.

33, (*S*)-2-Acetamido-*N*-((*S*)-1-(benzo[d]thiazol-2-yl)-3-(4-guanidinophenyl)-1-oxopropan-2-yl)-3-phenylpropanamide



After the coupling of Ac-Phe-OH with the P1-warhead moiety **28**, the deprotection of the Cbz- and O^{t}Bu-protecting groups was performed in a TFA/TA solution (5:1, 10 mL) and stirred for 12 h at room temperature. The solution was evaporated under reduced pressure and purified by preparative HPLC (16.3 mg, 0.03 mmol, 28%). ^{1}H NMR (400 MHz, DMSO-d_6): δ/ppm = 8.76 (dd, *J* = 28.0, 7.4 Hz, 1H), 8.41 (s, 1H), 8.32 – 8.27 (m, 2H), 8.08 (dd, *J* = 8.5, 4.4 Hz, 1H), 7.91 (s, 2H), 7.72 – 7.64 (m, 2H), 7.38 – 7.31 (m, 2H), 7.23 – 7.11 (m, 8H), 5.77 – 5.68 (m, 0H), 4.60 – 4.50 (m, 1H), 3.32 – 3.29 (m, 1H), 3.00 – 2.52 (m, 3H), 1.73 – 1.66 (m, 3H), ^{13}C NMR (101 MHz, DMSO-d_6): δ/ppm = 193.0, 172.0, 169.6, 167.9, 164.7, 156.6, 153.4, 138.3, 136.9, 135.4, 129.5, 128.4, 128.1, 126.7, 125.8, 124.3, 123.7, 56.7, 54.2, 37.8, 36.1, 22.8, FT-IR: ν/cm^{-1} = 711, 773, 793, 832, 1125, 1195, 1397, 1510, 1636, 2355, mp: 140 – 143 °C, $[\alpha]_D^{20}$ = +57 (*c* 1.00, DMSO), MS (ESI) *m/z* calculated for $[C_{28}H_{29}N_6O_3S]^+$ ([M+H]$^+$): 529.1, found 529.0. Purity: 95%.

“

6 Summary and Outlook

In the framework of ongoing trends of covalent protease inhibitor development in medicinal chemistry, this doctoral thesis presents a meaningful addition, which can be divided into chapter A “Investigation and Characterization of Warheads and Substrates for Proteases and their Inhibitors”, and chapter B “Evaluation of Covalent Inhibitors for Trypsin-Like Serine Proteases”.

The majority of the FDA approved covalent serine protease inhibitors as well as new potential drugs in current drug design display a peptidomimetic scaffold combined with an electrophilic trap at the *C*-terminus. Within the extensive systematic study in project 4.1 important properties of both structure elements were characterized. Commonly used warhead-types were evaluated regarding their reactivity towards proteases with different catalytic-active amino acid residues and their raw chemoselectivity in test systems. The results reveal preferences of the different warheads to specific nucleophiles in the enzymes and in solution, but also shows disadvantages in case of very high-reactive electrophilic moieties regarding selectivity and possible safety issues. Additionally, the effectiveness of well suited peptidomimetic recognition sequences was highlighted, due to very promising selectivity profiles of some designed inhibitors. However, structurally very similar proteases still showed cross-reactivity and highlight the challenges of protease inhibitor development. Moreover, for some proteases only the combination of the right warhead targeting the specific nucleophilic residue of the enzyme with a suitable peptidomimetic recognition sequence showed affinity in first place. This study presents important information about peptidomimetic covalent inhibitors with an assortment of specific warheads, which helps the scientific community for further drug design towards different targets. Furthermore, it represents an example of well-executed inhibitor development, which might be a guideline in general, for first characterization-steps in drug research.

The inhibition studies of all inhibitors towards their target protease were performed with a fluorometric assay. The generally used AMC-or FRET-based substrates can cause assay interferences, which could lead to false information about the inhibitory activity. Therefore, we designed and synthesized NBD-based fluorogenic substrates for overall ten different proteases. Evaluation of their kinetic properties and their suitability for inhibition studies were proven successfully. Due to the shift of the excitation and emission wavelengths of NBD to a longer wavelength area, common assay interferences could be mitigated. Further optimization of this new type of fluorogenic substrates should focus on the current state of the synthetic procedure. Improving the synthesis by different strategies that lead to higher yields would increase the accessibility of the NBD-based substrates as well as enhance their scope of application.

Chapter B represents the rational design of specific peptidomimetic covalent inhibitors, mostly connected with a α-ketobenzothiazole moiety as a warhead, against the trypsin-like serine proteases hepsin, TMPRSS2 and uPA. The first two projects, regarding TMPRSS2- (5.1) and hepsin- (5.2)

inhibitors in cooperation with the group around Mailänder *et al*. resulted in the development of potent peptidomimetic, covalent reversible inhibitors with good selectivity profiles towards other structure-related serine proteases. Cell-based assays of the TMPRSS2 inhibitors also proved the ability to impair the spike protein-driven cell entry of the coronavirus and block the SARS-CoV-2 infection.

The ligand-based design of uPA and TMPRSS2 inhibitors, started with the discovered hits of the uPA-inhibitor in project 4.1 and TMPRSS2 inhibitor in project 5.1. The substitution of the P1 arginine side chain with a cyclohexyl/phenylguanindine moiety yielded inhibitors with nanomolar/subnanomolar affinities to their respective target protease and display improved selectivity profiles.

In summary, the developed inhibitors within these three projects (5.1–5.3) represent promising protease inhibitors with high potencies and advantageous selectivity properties. However, pharmacokinetic properties like oral bioavailability (facilitated drug application) and biodistribution (reaching target tissue) are heavily impaired by the positively charged guanidine structure elements. Therefore, next steps should focus on further drug optimization to improve the abovementioned pharmacokinetic issues for example by converting the positively charged groups into prodrugs or using nanoparticle formulations for a target-oriented drug transport (1.7).

7 Bibliography

1. Turk, B. Targeting proteases: successes, failures and future prospects. *Nat. Rev. Drug Discov.* **2006** 5, 785–799.

2. Barrett, Alan J. , Rawlings Neil D., Woessner, J. F. *Handbook of Proteolytic Enzymes*. (Elsevier, **2004**).

3. Macfarlane, R. G. An Enzyme Cascade in the Blood Clotting Mechanism, and its Function as a Biochemical Amplifier. *Nature* **1964** 202, 498–499.

4. Davie, E. W. & Ratnoff, O. D. Waterfall Sequence for Intrinsic Blood Clotting. *Science (80-.).* **1964** 145, 1310–1312.

5. Davie, E. W. & Neurath, H. Identification of a Peptide released during Autocatalytic Activation of Trypsinogen. *J. Biol. Chem.* **1955** 212, 515–529.

6. López-Otín, C. & Overall, C. M. Protease degradomics: A new challenge for proteomics. *Nat. Rev. Mol. Cell Biol.* **2002** 3, 509–519.

7. Patel, T., Gores, G. J. & Kaufmann, S. H. The role of proteases during apoptosis. *FASEB J.* **1996** 10, 587–597.

8. Scher, W. The role of extracellular proteases in cell proliferation and differentiation. *Lab. Invest.* **1987** 57, 607–33.

9. Yaron, J. R., Zhang, L., Guo, Q., Haydel, S. E. & Lucas, A. R. Fibrinolytic Serine Proteases, Therapeutic Serpins and Inflammation: Fire Dancers and Firestorms. *Front. Cardiovasc. Med.* **2021** 8, 648947.

10. Oncul, S. & Afshar-Kharghan, V. The interaction between the complement system and hemostatic factors. *Curr. Opin. Hematol.* **2020** 27, 341–352.

11. Lu, P., Takai, K., Weaver, V. M. & Werb, Z. Extracellular Matrix Degradation and Remodeling in Development and Disease. *Cold Spring Harb. Perspect. Biol.* **2011** 3, 1–25.

12. Khan, A. R., Khazanovich-Bernstein, N., Bergmann, E. M. & James, M. N. G. Structural aspects of activation pathways of aspartic protease zymogens and viral 3C protease precursors. *Proc. Natl. Acad. Sci.* **1999** 96, 10968–10975.

13. Bird, P. I., Trapani, J. A. & Villadangos, J. A. Endolysosomal proteases and their inhibitors in immunity. *Nat. Rev. Immunol.* **2009** 9, 871–882.

14. Vidoni, C., Follo, C., Savino, M., Melone, M. A. B. & Isidoro, C. The Role of Cathepsin D in the Pathogenesis of Human Neurodegenerative Disorders. *Med. Res. Rev.* **2016** 36, 845–870.

15. Wu, Q., Kuo, H.-C. & Deng, G. G. Serine proteases and cardiac function. *Biochim. Biophys. Acta - Proteins Proteomics* **2005** 1751, 82–94.

16. Eatemadi, A. *et al.* Role of protease and protease inhibitors in cancer pathogenesis and treatment. *Biomed. Pharmacother.* **2017** 86, 221–231.

17. Puente, X. S., Sánchez, L. M., Overall, C. M. & López-Otín, C. Human and mouse proteases: a comparative genomic approach. *Nat. Rev. Genet.* **2003** 4, 544–558.

18. Tavano, O. L. Protein hydrolysis using proteases: An important tool for food biotechnology. *J. Mol. Catal. B Enzym.* **2013** 90, 1–11.

19. Venter, J. C. *et al.* The Sequence of the Human Genome. *Science (80-.).* **2001** 291, 1304–1351.

20. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **2001** 409, 860–921.

21. Webb, E. C. Enzyme Nomenclature: A Personal Retrospective. *FASEB J.* **1993** 7, 1192–1194.

22. Rawlings, N. D. Twenty-five years of nomenclature and classification of proteolytic enzymes. *Biochim. Biophys. Acta - Proteins Proteomics* **2020** 1868, 140345.

23. Rawlings, N. D., Barrett, A. J. & Bateman, A. Asparagine Peptide Lyases. *J. Biol. Chem.* **2011** 286, 38321–38328.

24. Hedstrom, L. Serine Protease Mechanism and Specificity. *Chem. Rev.* **2002** 102, 4501–4524.

25. Schauperl, M. *et al.* Characterizing Protease Specificity: How Many Substrates Do We Need? *PLoS One* **2015** 10, 1–17.

26. Frey, P. A. Strong hydrogen bonding in chymotrypsin and other serine proteases. *J. Phys. Org. Chem.* **2004** 17, 511–520.

27. Waldner, B. J. *et al.* Electrostatic recognition in substrate binding to serine proteases. *J. Mol. Recognit.* **2018** 31, 1–12.

28. Tyndall, J. D. A., Nall, T. & Fairlie, D. P. Proteases Universally Recognize Beta Strands In Their Active Sites. *Chem. Rev.* **2005** 105, 973–1000.

29. Sharma, A. K. & Bansal, P. *Serine Proteases*. (De Gruyter, **2023**).

30. Di Cera, E. Serine proteases. *IUBMB Life* **2009** 61, 510–515.

31. Matthews, B. W., Sigler, P. B., Henderson, R. & Blow, D. M. Three-dimensional Structure of Tosyl-α-chymotrypsin. *Nature* **1967** 214, 652–656.

32. Dodson, G. Catalytic triads and their relatives. *Trends Biochem. Sci.* **1998** 23, 347–352.

33. Wright, C. S., Alden, R. A. & Kraut, J. Structure of Subtilisin BPN′ at 2.5 Å Resolution. *Nature* **1969** 221, 235–242.

34. 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, 503–509.

35. Page, M. J. & Di Cera, E. Serine peptidases: Classification, structure and function. *Cell. Mol. Life Sci.* **2008** 65, 1220–1236.

36. Abdul Razzaq, Khadija Quddoos, and A. A. EndopeptidasesProteases Enzymes in Proteins. *Mol Biol* **2021** 10, 274.

37. Page, M. J. & Di Cera, E. Evolution of Peptidase Diversity. *J. Biol. Chem.* **2008** 283, 30010–30014.

38. Ralf, T., Bernd, L. & Martine, S. StructuralBasisofProline-Specific ExopeptidaseActivityasObserved inHumanDipeptidylPeptidase-IV. *Structure* **2003** 11, 947–959.

39. Lee, M. & Rhee, M. K. Sitagliptin for Type 2 diabetes: a 2015 update. *Expert Rev. Cardiovasc. Ther.* **2015** 13, 597–610.

40. Ripka, A. S. & Rich, D. H. Peptidomimetic design. *Curr. Opin. Chem. Biol.* **1998** 2, 441–452.

41. Gante, J. Peptidomimetics—Tailored Enzyme Inhibitors. *Angew. Chemie Int. Ed. English* **1994** 33, 1699–1720.

42. Liu, R., Li, X. & Lam, K. S. Combinatorial chemistry in drug discovery. *Curr. Opin. Chem. Biol.* **2017** 38, 117–126.

43. Nicholls, A. *et al.* Molecular Shape and Medicinal Chemistry: A Perspective. *J. Med. Chem.* **2010** 53, 3862–3886.

44. Sharma, K., Sharma, K. K., Sharma, A. & Jain, R. Peptide-based drug discovery: Current status and recent advances. *Drug Discov. Today* **2023** 28, 103464.

45. Lu, J. *et al.* D- and Unnatural Amino Acid Substituted Antimicrobial Peptides With Improved Proteolytic Resistance and Their Proteolytic Degradation Characteristics. *Front. Microbiol.* **2020** 11, 1–17.

46. Cruz, L., Soares, P. & Correia, M. Ubiquitin-Specific Proteases: Players in Cancer Cellular Processes. *Pharmaceuticals* **2021** 14, 848.

47. Vandermarliere, E., Mueller, M. & Martens, L. Getting intimate with trypsin, the leading protease in proteomics. *Mass Spectrom. Rev.* **2013** 32, 453–465.

48. Barry, M. & Bleackley, R. C. Cytotoxic T lymphocytes: all roads lead to death. *Nat. Rev. Immunol.* **2002** 2, 401–409.

49. Diamandis, E. P., Yousef, G. M., Luo, L.-Y., Magklara, A. & Obiezu, C. V. The New Human Kallikrein Gene Family: Implications in Carcinogenesis. *Trends Endocrinol. Metab.* **2000** 11, 54–60.

50. Lin, C.-Y., Anders, J., Johnson, M., Sang, Q. A. & Dickson, R. B. Molecular Cloning of cDNA for Matriptase, a Matrix-degrading Serine Protease with Trypsin-like Activity. *J. Biol. Chem.* **1999** 274, 18231–18236.

51. Pereira, P. J. B. *et al.* Human β-tryptase is a ring-like tetramer with active sites facing a central pore. *Nature* **1998** 392, 306–311.

52. Schreuder, H. & Matter, H. Serine Proteinases from the Blood Coagulation Cascade. in *Structural Biology in Drug Discovery* (Wiley, **2020**). 395–422.

53. Sotiropoulou, G., Pampalakis, G. & Diamandis, E. P. Functional Roles of Human Kallikrein-related Peptidases. *J. Biol. Chem.* **2009** 284, 32989–32994.

54. De Oliveira, E. B. & Salgado, M. C. O. Pancreatic Elastases. in *Handbook of Proteolytic Enzymes* (Elsevier, **2013**). 2639–2645.

55. Bender, M. L., Killheffer, J. V. & Cohen, S. Chymotrypsin. *CRC Crit. Rev. Biochem.* **1973** 1, 149–199.

56. Hirota, M., Ohmuraya, M. & Baba, H. The role of trypsin, trypsin inhibitor, and trypsin receptor in the onset and aggravation of pancreatitis. *J. Gastroenterol.* **2006** 41, 832–836.

57. Steinhoff, M. *et al.* Agonists of proteinase-activated receptor 2 induce inflammation by a neurogenic mechanism. *Nat. Med.* **2000** 6, 151–158.

58. Whitcomb, D. C. & Lowe, M. E. Human Pancreatic Digestive Enzymes. *Dig. Dis. Sci.* **2007** 52, 1–17.

59. Page, M. J., Macgillivray, R. T. A. & Di Cera, E. Determinants of specificity in coagulation proteases. *J. Thromb. Haemost.* **2005** 3, 2401–2408.

60. Alan J. Barrett, J. Fred Woessner, N. D. R. *Handbook of Proteolytic Enzymes, Volume 1*. (**2004**).

61. Fehlhammer, H., Bode, W. & Huber, R. Crystal structure of bovine trypsinogen at 1·8 Å resolution. *J. Mol. Biol.* **1977** 111, 415–438.

62. Fehlhammer, H. & Bode, W. The refined crystal structure of bovine β-trypsin at 1·8 Å resolution. *J. Mol. Biol.* **1975** 98, 683–692.

63. Renatus, M. Lysine 156promotes the anomalous proenzyme activity of tPA: X-ray crystal structure of single-chain human tPA. *EMBO J.* **1997** 16, 4797–4805.

64. Blow, D. M. Structure and Substrate Specificity in the Serine Enzymes. *Isr. J. Chem.* **1974** 12, 483–494.

65. Evnin, L. B., Vásquez, J. R. & Craik, C. S. Substrate specificity of trypsin investigated by using a genetic selection. *Proc. Natl. Acad. Sci.* **1990** 87, 6659–6663.

66. Shotton, D. M. & Watson, H. C. Three-dimensional Structure of Tosyl-elastase. *Nature* **1970** 225, 811–816.

67. Gurina, T. & Mohiuddin, S. Biochemistry, Protein Catabolism. *Eur. PMC* **2020**.

68. Freiburghaus, A. U., Roduner, J. & Hadorn, H. B. Activation of Human Pancreatic Proteolytic Enzymes: The Role of Enteropeptidase and Trypsin. *JPGN Reports* **2021** 2, 1–6.

69. Ji, B. & Logsdon, C. D. Digesting New Information About the Role of Trypsin in Pancreatitis. *Gastroenterology* **2011** 141, 1972–1975.

70. Kumari, A. Citric Acid Cycle. in *Sweet Biochemistry* (Elsevier, **2018**). 7–11.

71. Li, X., Zheng, S. & Wu, G. Amino Acid Metabolism in the Kidneys: Nutritional and Physiological Significance. in (**2020**). 71–95.

72. Morrissey, J., Macik, B., Neuenschwander, P. & Comp, P. Quantitation of activated factor VII levels in plasma using a tissue factor mutant selectively deficient in promoting factor VII activation. *Blood* **1993** 81, 734–744.

73. Esmon, C. T. Targeting factor Xa and thrombin: impact on coagulation and beyond. *Thromb. Haemost.* **2014** 111, 625–633.

74. Mosseson, M. W. Fibrinogen and fibrin structure and functions. *J. Thromb. Haemost.* **2005** 3, 1894–1904.

75. Wang, Y., Luo, W. & Reiser, G. Trypsin and trypsin-like proteases in the brain: Proteolysis and cellular functions. *Cell. Mol. Life Sci.* **2008** 65, 237–252.

76. Heutinck, K. M., ten Berge, I. J. M., Hack, C. E., Hamann, J. & Rowshani, A. T. Serine proteases of the human immune system in health and disease. *Mol. Immunol.* **2010** 47, 1943–1955.

77. Heng, H. H. Q. *et al.* The evolutionary mechanism of cancer. *J. Cell. Biochem.* **2010** 109, 1072–

1084.

78. Klein, C. A. Cancer progression and the invisible phase of metastatic colonization. *Nat. Rev. Cancer* **2020** 20, 681–694.

79. Martin, C. E. & List, K. Cell surface–anchored serine proteases in cancer progression and metastasis. *Cancer Metastasis Rev.* **2019** 38, 357–387.

80. Poddar, N. K., Maurya, S. K. & Saxena, V. Role of Serine Proteases and Inhibitors in Cancer. in *Proteases in Physiology and Pathology* (Springer Singapore, **2017**). 257–287.

81. Tagirasa, R. & Yoo, E. Role of Serine Proteases at the Tumor-Stroma Interface. *Front. Immunol.* **2022** 13, 1–12.

82. Dass, K., Ahmad, A., Azmi, A. S., Sarkar, S. H. & Sarkar, F. H. Evolving role of uPA/uPAR system in human cancers. *Cancer Treat. Rev.* **2008** 34, 122–136.

83. Duffy, M. The Urokinase Plasminogen Activator System: Role in Malignancy. *Curr. Pharm. Des.* **2004** 10, 39–49.

84. Gao, L., Chao, L. & Chao, J. A novel signaling pathway of tissue kallikrein in promoting keratinocyte migration: Activation of proteinase-activated receptor 1 and epidermal growth factor receptor. *Exp. Cell Res.* **2010** 316, 376–389.

85. Harapan, H., Michie, A., Sasmono, R. T. & Imrie, A. Dengue: A Minireview. *Viruses* **2020** 12, 829.

86. Barreto-Vieira, D. F., Couto-Lima, D., Jácome, F. C., Caldas, G. C. & Barth, O. M. Dengue, Yellow Fever, Zika and Chikungunya epidemic arboviruses in Brazil: ultrastructural aspects. *Mem. Inst. Oswaldo Cruz* **2020** 115, 1–7.

87. Yun, S.-I. & Lee, Y.-M. Zika virus: An emerging flavivirus. *J. Microbiol.* **2017** 55, 204–219.

88. Nitsche, C., Holloway, S., Schirmeister, T. & Klein, C. D. Biochemistry and Medicinal Chemistry of the Dengue Virus Protease. *Chem. Rev.* **2014** 114, 11348–11381.

89. Boldescu, V., Behnam, M. A. M., Vasilakis, N. & Klein, C. D. Broad-spectrum agents for flaviviral infections: dengue, Zika and beyond. *Nat. Rev. Drug Discov.* **2017** 16, 565–586.

90. Rao, M. B., Tanksale, A. M., Ghatge, M. S. & Deshpande, V. V. Molecular and Biotechnological Aspects of Microbial Proteases. *Microbiol. Mol. Biol. Rev.* **1998** 62, 597–635.

91. Sasahara, A. A. Urokinase: Past, present, and future. *Tech. Vasc. Interv. Radiol.* **1998** 1, 170–178.

92. Nikitin, D. *et al.* Development and Testing of Thrombolytics in Stroke. *J. Stroke* **2021** 23, 12–36.

93. Medved, L. & Nieuwenhuizen, W. Molecular mechanisms of initiation of fibrinolysis by fibrin. *Thromb. Haemost.* **2003** 89, 409–419.

94. Di Cera, E. Thrombin. *Mol. Aspects Med.* **2008** 29, 203–254.

95. Craik, C. S., Page, M. J. & Madison, E. L. Proteases as therapeutics. *Biochem. J.* **2011** 435, 1–16.

96. Jin, Z. *et al.* Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020** 582, 289–293.

97. Irving, J. A. Phylogeny of the Serpin Superfamily: Implications of Patterns of Amino Acid Conservation for Structure and Function. *Genome Res.* **2000** 10, 1845–1864.

98. Quinsey, N. S., Greedy, A. L., Bottomley, S. P., Whisstock, J. C. & Pike, R. N. Antithrombin: in control of coagulation. *Int. J. Biochem. Cell Biol.* **2004** 36, 386–389.

99. De Serres, F. & Blanco, I. Role of alpha-1 antitrypsin in human health and disease. *J. Intern. Med.* **2014** 276, 311–335.

100. Xu, P. & Huang, M. Small Peptides as Modulators of Serine Proteases. *Curr. Med. Chem.* **2020** 27, 3686–3705.

101. Rahaghi, F. F. Alpha-1 antitrypsin deficiency research and emerging treatment strategies: what's down the road? *Ther. Adv. Chronic Dis.* **2021** 12, 77–90.

102. Chapman, K. R. *et al.* Intravenous augmentation treatment and lung density in severe α1 antitrypsin deficiency (RAPID): a randomised, double-blind, placebo-controlled trial. *Lancet* **2015** 386, 360–368.

103. Ranasinghe, S. & McManus, D. P. Structure and function of invertebrate Kunitz serine protease inhibitors. *Dev. Comp. Immunol.* **2013** 39, 219–227.

104. Waxler, B. & Rabito, S. Aprotinin: A Serine Protease Inhibitor with Therapeutic Actions: Its Interaction with ACE Inhibitors. *Curr. Pharm. Des.* **2003** 9, 777–787.

105. Scott, C. J. & Taggart, C. C. Biologic protease inhibitors as novel therapeutic agents. *Biochimie* **2010** 92, 1681–1688.

106. Muheem, A. *et al.* A review on the strategies for oral delivery of proteins and peptides and their clinical perspectives. *Saudi Pharm. J.* **2016** 24, 413–428.

107. Åstedt, B. Clinical Pharmacology of Tranexamic Acid. *Scand. J. Gastroenterol.* **1987** 22, 22–25.

108. Ladehoff, A. & Otte, E. Inhibitory Effect of Epsilon Aminocaproic Acid on Fibrinolytic Activity and Bleeding in Transvesical Prostatectomy. *Scand. J. Clin. Lab. Invest.* **1963** 15, 239–247.

109. Yasuzo, N. F. E. M. M. M. I. S. Method of Treating Diseases Associated with Plasmin Activity. *US Pat.* **1960** 1–8.

110. Leeson, P. D. *et al.* Target-Based Evaluation of "Drug-Like" Properties and Ligand Efficiencies. *J. Med. Chem.* **2021** 64, 7210–7230.

111. Brockway, W. J. & Castellino, F. J. The Mechanism of the Inhibition of Plasmin Activity by ε-Aminocaproic Acid. *J. Biol. Chem.* **1971** 246, 4641–4647.

112. Tyagi, M., Begnini, F., Poongavanam, V., Doak, B. C. & Kihlberg, J. Drug Syntheses Beyond the Rule of 5. *Chem. – A Eur. J.* **2020** 26, 49–88.

113. Doak, B. C. & Kihlberg, J. Drug discovery beyond the rule of 5 - Opportunities and challenges. *Expert Opin. Drug Discov.* **2017** 12, 115–119.

114. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **1997** 23, 3–25.

115. Prudovsky, I. *et al.* Tranexamic acid: Beyond antifibrinolysis. *Transfusion* **2022** 62, 301–312.

116. Walenga, J. M. An Overview of the Direct Thrombin Inhibitor Argatroban. *Pathophysiol. Haemost. Thromb.* **2002** 32, 9–14.

117. Fredenburgh, J. C., Gross, P. L. & Weitz, J. I. Emerging anticoagulant strategies. *Blood* **2017** 129, 147–154.

118. Lekura, J. & Kalus, J. S. Overview of betrixaban and its role in clinical practice. *Am. J. Heal. Pharm.* **2018** 75, 1095–1102.

119. Rawal, A. *et al.* Current status of oral anticoagulant reversal strategies: a review. *Ann. Transl. Med.* **2019** 7, 411–411.

120. DeLoughery, E. P. & Shatzel, J. J. A comparative analysis of the safety profile of direct oral anticoagulants using the FDA adverse event reporting system. *Eur. J. Haematol.* **2019** 103, 43–46.

121. Awad, A. J. *et al.* Dabigatran, intracranial hemorrhage, and the neurosurgeon. *Neurosurg. Focus* **2013** 34, 1–6.

122. Fisman, E. Z. & Tenenbaum, A. Antidiabetic treatment with gliptins: focus on cardiovascular effects and outcomes. *Cardiovasc. Diabetol.* **2015** 14, 129.

123. McKeage, K. Trelagliptin: First Global Approval. *Drugs* **2015** 75, 1161–1164.

124. Chung, J. Y. L. *et al.* Evolution of a Manufacturing Route to Omarigliptin, A Long-Acting DPP-4 Inhibitor for the Treatment of Type 2 Diabetes. *Org. Process Res. Dev.* **2015** 19, 1760–1768.

125. Tan, X. & Hu, J. Evogliptin: a new dipeptidyl peptidase inhibitor for the treatment of type 2 diabetes. *Expert Opin. Pharmacother.* **2016** 17, 1285–1293.

126. Neumiller, J. J. & Setter, S. M. Review of Linagliptin for the Treatment of Type 2 Diabetes Mellitus. *Clin. Ther.* **2012** 34, 993–1005.

127. Thareja, S. *et al.* Saxagliptin: A New Drug for the Treatment of Type 2 Diabetes. *Mini-Reviews Med. Chem.* **2010** 10, 759–765.

128. Kirk, J. & Oldham, E. Sitagliptin and Vildagliptin: Efficacy of Therapy in Type 2 Diabetes. *Recent Pat. Endocr. Metab. Immune Drug Discov.* **2008** 2, 162–171.

129. Lamb, Y. N. Glecaprevir/Pibrentasvir: First Global Approval. *Drugs* **2017** 77, 1797–1804.

130. Ganta, N. M. *et al.* A review on HCV inhibitors: Significance of non-structural polyproteins. *Eur. J. Med. Chem.* **2019** 164, 576–601.

131. Farkas, H. & Balla, Z. A review of berotralstat for the treatment of hereditary angioedema. *Expert Rev. Clin. Immunol.* **2023** 19, 145–153.

132. Wang, M. *et al.* Deep learning approaches for de novo drug design: An overview. *Curr. Opin. Struct. Biol.* **2022** 72, 135–144.

133. Sharma, K. *et al.* Synthetic amino acids-derived peptides target Cryptococcus neoformans by inducing cell membrane disruption. *Bioorg. Chem.* **2023** 130, 106252.

134. Chingle, R., Proulx, C. & Lubell, W. D. Azapeptide Synthesis Methods for Expanding Side-Chain Diversity for Biomedical Applications. *Acc. Chem. Res.* **2017** 50, 1541–1556.

135. Vinogradov, A. A., Yin, Y. & Suga, H. Macrocyclic Peptides as Drug Candidates: Recent Progress and Remaining Challenges. *J. Am. Chem. Soc.* **2019** 141, 4167–4181.

136. Blech, S., Ebner, T., Ludwig-Schwellinger, E., Stangier, J. & Roth, W. The Metabolism and Disposition of the Oral Direct Thrombin Inhibitor, Dabigatran, in Humans. *Drug Metab. Dispos.* **2008** 36, 386–399.

137. Lonsdale, R. & Ward, R. A. Structure-based design of targeted covalent inhibitors. *Chem. Soc.*

Rev. **2018** 47, 3816–3830.

138. Baillie, T. A. Targeted Covalent Inhibitors for Drug Design. *Angew. Chemie Int. Ed.* **2016** 55, 13408–13421.

139. Singh, J., Petter, R. C., Baillie, T. A. & Whitty, A. The resurgence of covalent drugs. *Nat. Rev. Drug Discov.* **2011** 10, 307–317.

140. Chen, D., Frezza, M., Schmitt, S., Kanwar, J. & P. Dou, Q. Bortezomib as the First Proteasome Inhibitor Anticancer Drug: Current Status and Future Perspectives. *Curr. Cancer Drug Targets* **2011** 11, 239–253.

141. Uetrecht, J. Immune-Mediated Adverse Drug Reactions. *Chem. Res. Toxicol.* **2009** 22, 24–34.

142. Ramadan, M. Review of Serine Protease Inhibitors: Development and Applications. *Bull. Pharm. Sci. Assiut Univ.* **2023** 46, 835–854.

143. Kim, H., Hwang, Y. S., Kim, M. & Park, S. B. Recent advances in the development of covalent inhibitors. *RSC Med. Chem.* **2021** 12, 1037–1045.

144. Du, D. *et al.* Multidrug efflux pumps: structure, function and regulation. *Nat. Rev. Microbiol.* **2018** 16, 523–539.

145. Fletcher, J. I., Haber, M., Henderson, M. J. & Norris, M. D. ABC transporters in cancer: more than just drug efflux pumps. *Nat. Rev. Cancer* **2010** 10, 147–156.

146. Popovic-Djordjevic, J. B., Jevtic, I. I. & Stanojkovic, T. P. Antidiabetics: Structural Diversity of Molecules with a Common Aim. *Curr. Med. Chem.* **2018** 25, 2140–2165.

147. De Leuw, P. & Stephan, C. Protease inhibitors for the treatment of hepatitis C virus infection. *GMS Infect. Dis.* **2017** 5, 1–14.

148. Maryanoff, B. E. & Costanzo, M. J. Inhibitors of proteases and amide hydrolases that employ an α-ketoheterocycle as a key enabling functionality. *Bioorg. Med. Chem.* **2008** 16, 1562–1595.

149. Robello, M. *et al.* The Alpha Keto Amide Moiety as a Privileged Motif in Medicinal Chemistry: Current Insights and Emerging Opportunities. *J. Med. Chem.* **2021** 64, 3508–3545.

150. Fleming, F. F., Yao, L., Ravikumar, P. C., Funk, L. & Shook, B. C. Nitrile-Containing Pharmaceuticals: Efficacious Roles of the Nitrile Pharmacophore. *J. Med. Chem.* **2010** 53, 7902–7917.

151. Thompson, R. C. Peptide aldehydes: Potent inhibitors of serine and cysteine proteases. in *Methods in Enzymology* (**1977**). vol. 693 220–225.

152. Smoum, R., Rubinstein, A., Dembitsky, V. M. & Srebnik, M. Boron Containing Compounds as Protease Inhibitors. *Chem. Rev.* **2012** 112, 4156–4220.

153. Pace, V., Castoldi, L. & Pregnolato, M. Halomethylketones: Synthetic Methodologies and Pharmaceutical Applications as Serine and Cysteine Protease Inhibitors. *Mini-Reviews Med. Chem.* **2013** 13, 988–996.

154. Maślanka, M. & Mucha, A. Recent Developments in Peptidyl Diaryl Phoshonates as Inhibitors and Activity-Based Probes for Serine Proteases. *Pharmaceuticals* **2019** 12, 86.

155. Powers, J. C., Asgian, J. L., Ekici, Ö. D. & James, K. E. Irreversible Inhibitors of Serine, Cysteine, and Threonine Proteases. *Chem. Rev.* **2002** 102, 4639–4750.

156. Ji, S. & Verhelst, S. H. L. Furin-targeting activity-based probes with phosphonate and phosphinate esters as warheads. *Org. Biomol. Chem.* **2023** 21, 6498–6502.

157. Béné, M. C. *et al.* CD87 (urokinase-type plasminogen activator receptor), function and pathology in hematological disorders: a review. *Leukemia* **2004** 18, 394–400.

158. Blasi, F. & Carmeliet, P. uPAR: a versatile signalling orchestrator. *Nat. Rev. Mol. Cell Biol.* **2002** 3, 932–943.

159. Choong, P. F. M. & Nadesapillai, A. P. W. Urokinase Plasminogen Activator System. *Clin. Orthop. Relat. Res.* **2003** 415, 46–58.

160. Ulisse, S., Baldini, E., Sorrenti, S. & D'Armiento, M. The Urokinase Plasminogen Activator System: A Target for Anti-Cancer Therapy. *Curr. Cancer Drug Targets* **2009** 9, 32–71.

161. Andreasen, P. A., Kjøller, L., Christensen, L. & Duffy, M. J. The urokinase-type plasminogen activator system in cancer metastasis: A review. *Int. J. Cancer* **1997** 72, 1–22.

162. Behrens, M. A. *et al.* Activation of the Zymogen to Urokinase-Type Plasminogen Activator Is Associated with Increased Interdomain Flexibility. *J. Mol. Biol.* **2011** 411, 417–429.

163. M, S. *et al.* Biological and clinical relevance of the urokinase-type plasminogen activator (uPA) in breast cancer. *Biochim. Biophys. Acta* **1991** 50, 731–741.

164. Mahmood, N., Mihalcioiu, C. & Rabbani, S. A. Multifaceted Role of the Urokinase-Type Plasminogen Activator (uPA) and Its Receptor (uPAR): Diagnostic, Prognostic, and Therapeutic Applications. *Front. Oncol.* **2018** 8, 1–21.

165. Mukhina, S. *et al.* The Chemotactic Action of Urokinase on Smooth Muscle Cells Is Dependent on Its Kringle Domain. *J. Biol. Chem.* **2000** 275, 16450–16458.

166. Ploug, M. & Ellis, V. Structure—function relationships in the receptor for urokinase-type

plasminogen activator Comparison to other members of the Ly-6 family and snake venom α-neurotoxins. *FEBS Lett.* **1994** 349, 163–168.

167. Ploug, M. *et al.* Cellular receptor for urokinase plasminogen activator. Carboxyl-terminal processing and membrane anchoring by glycosyl-phosphatidylinositol. *J. Biol. Chem.* **1991** 266, 1926–1933.

168. Smith, H. W. & Marshall, C. J. Regulation of cell signalling by uPAR. *Nat. Rev. Mol. Cell Biol.* **2010** 11, 23–36.

169. Kumar, A. A., Buckley, B. J. & Ranson, M. The Urokinase Plasminogen Activation System in Pancreatic Cancer: Prospective Diagnostic and Therapeutic Targets. *Biomolecules* **2022** 12, 152.

170. Olson, L. J., Yammani, R. D., Dahms, N. M. & Kim, J.-J. P. Structure of uPAR, plasminogen, and sugar-binding sites of the 300 kDa mannose 6-phosphate receptor. *EMBO J.* **2004** 23, 2019–2028.

171. Degryse, B. The Urokinase Receptor and Integrins Constitute a Cell Migration Signalosome. in *The Cancer Degradome* (Springer New York). 451–474.

172. LeBeau, A. M. *et al.* Targeting uPAR with Antagonistic Recombinant Human Antibodies in Aggressive Breast Cancer. *Cancer Res.* **2013** 73, 2070–2081.

173. De Bock, C. E. & Wang, Y. Clinical significance of urokinase-type plasminogen activator receptor (uPAR) expression in cancer. *Med. Res. Rev.* **2004** 24, 13–39.

174. Agrawal, A., Romero-Perez, D., Jacobsen, J. A., Villarreal, F. J. & Cohen, S. M. Zinc-Binding Groups Modulate Selective Inhibition of MMPs. *ChemMedChem* **2008** 3, 812–820.

175. Ismail, A. A., Shaker, B. T. & Bajou, K. The Plasminogen–Activator Plasmin System in Physiological and Pathophysiological Angiogenesis. *Int. J. Mol. Sci.* **2021** 23, 337.

176. Plesner, T., Behrendt, N. & Ploug, M. Structure, Function and Expression on Blood and Bone Marrow Cells of the Urokinase-Type Plasminogen Activator Receptor, uPAR. *Stem Cells* **1997** 15, 398–408.

177. Rijken, D. C. & Sakharov, D. V. Basic Principles in Thrombolysis: Regulatory Role of Plasminogen. *Thromb. Res.* **2001** 103, 41–49.

178. Krieger, M. & Krieger, J. Structures and Functions of Multiligand Lipoprotein Receptors: Macrophage Scavenger Receptors and LDL Receptor-Related Protein (LRP). *Annu. Rev. Biochem.* **1994** 63, 601–637.

179. Chi Ki Ngo, J. *et al.* Structural Basis for Therapeutic Intervention of uPA/uPAR System. *Curr.*

Drug Targets **2011** 12, 1729–1743.

180. Zeslawska, E. *et al.* Crystals of Urokinase Type Plasminogen Activator Complexes Reveal the Binding Mode of Peptidomimetic Inhibitors. *J. Mol. Biol.* **2003** 328, 109–118.

181. Bevan, P. & Mala, C. The Role of uPA and uPA Inhibitors in Breast Cancer. *Breast Care* **2008** 3, 1–2.

182. Chiang, S. P. H., Cabrera, R. M. & Segall, J. E. Tumor cell intravasation. *Am. J. Physiol. Physiol.* **2016** 311, 1–14.

183. Ossowski, L. Plasminogen activator dependent pathways inthe dissemination of human tumor cells in the chick embryo. *Cell* **1988** 52, 321–328.

184. Rifkin, D. B. *et al.* Growth factor control of extracellular proteolysis. *Cell Differ. Dev.* **1990** 32, 313–318.

185. Mackman, R. L. *et al.* Exploiting Subsite S1 of Trypsin-Like Serine Proteases for Selectivity: Potent and Selective Inhibitors of Urokinase-Type Plasminogen Activator. **2001** 44, 3856–3871.

186. Rockway, T., Nienaber, V. & Giranda, V. Inhibitors of the Protease Domain of Urokinase-Type Plasminogen Activator. *Curr. Pharm. Des.* **2002** 8, 2541–2558.

187. Katz, B. A. *et al.* Dissecting and Designing Inhibitor Selectivity Determinants at the S1 Site Using an Artificial Ala190 Protease (Ala190 uPA). *J. Mol. Biol.* **2004** 344, 527–547.

188. Kruithof, E., Baker, M. & Bunn, C. Biological and clinical aspects of plasminogen activator inhibitor type 2. *Blood* **1995** 86, 4007–4024.

189. Dellas, C. & Loskutoff, D. J. Historical analysis of PAI-1 from its discovery to its potential role in cell motility and disease. *Thromb. Haemost.* **2005** 93, 631–640.

190. Lijnen, H. R. Pleiotropic functions of plasminogen activator inhibitor-1. *J. Thromb. Haemost.* **2005** 3, 35–45.

191. Stefansson, S., McMahon, G., Petitclerc, E. & Lawrence, D. Plasminogen Activator Inhibitor-1 in Tumor Growth, Angiogenesis and Vascular Remodeling. *Curr. Pharm. Des.* **2003** 9, 1545–1564.

192. Ann, G. & J., D. P. Structure-function Relationships in Serpins: Current Concepts and Controversies. *Thromb. Haemost.* **1998** 80, 531–541.

193. Wendt, M. D. *et al.* Interaction with the S1β-pocket of urokinase: 8-heterocycle substituted and 6,8-disubstituted 2-naphthamidine urokinase inhibitors. *Bioorg. Med. Chem. Lett.* **2004** 14, 3063–3068.

194. Nienaber, V. L. *et al.* Structure-directed discovery of potent non-peptidic inhibitors of human urokinase that access a novel binding subsite. *Structure* **2000** 8, 553–563.

195. Zhou, Y. *et al.* Structural study of the uPA-nafamostat complex reveals a covalent inhibitory mechanism of nafamostat. *Biophys. J.* **2022** 121, 3940–3949.

196. Künzel, S., Schweinitz, A., Reißmann, S., Stürzebecher, J. & Steinmetzer, T. 4-Amidinobenzylamine-Based inhibitors of urokinase. *Bioorg. Med. Chem. Lett.* **2002** 12, 645–648.

197. Joossens, J., Van der Veken, P., Lambeir, A.-M., Augustyns, K. & Haemers, A. Development of Irreversible Diphenyl Phosphonate Inhibitors for Urokinase Plasminogen Activator. *J. Med. Chem.* **2004** 47, 2411–2413.

198. McDonald, A. G. & Tipton, K. F. Enzymes: Irreversible Inhibition. in *Encyclopedia of Life Sciences* (Wiley, **2020**). 1–17.

199. Garattini, E. & Terao, M. The role of aldehyde oxidase in drug metabolism. *Expert Opin. Drug Metab. Toxicol.* **2012** 8, 487–503.

200. Sanoh, S., Tayama, Y., Sugihara, K., Kitamura, S. & Ohta, S. Significance of aldehyde oxidase during drug development: Effects on drug metabolism, pharmacokinetics, toxicity, and efficacy. *Drug Metab. Pharmacokinet.* **2015** 30, 52–63.

201. Banke, I. J. *et al.* Increase of Anti-Metastatic Efficacy by Selectivity- But Not Affinity-Optimization of Synthetic Serine Protease Inhibitors. *Biol. Chem.* **2003** 384, 1515–1525.

202. Schade, D. *et al.* Zanamivir Amidoxime- and N-Hydroxyguanidine-Based Prodrug Approaches to Tackle Poor Oral Bioavailability. *J. Pharm. Sci.* **2015** 104, 3208–3219.

203. Setyono-Han, B. *et al.* Suppression of rat breast cancer metastasis and reduction of primary tumour growth by the small synthetic urokinase inhibitor WX-UK1. *Thromb. Haemost.* **2005** 93, 779–786.

204. Rudzińska, M. *et al.* Current Status and Perspectives of Protease Inhibitors and Their Combination with Nanosized Drug Delivery Systems for Targeted Cancer Therapy. *Drug Des. Devel. Ther.* **2021** Volume 15, 9–20.

205. Zeslawska, E. *et al.* Crystals of the urokinase type plasminogen activator variant βc-uPA in complex with small molecule inhibitors open the way towards structure-based drug design 1 1Edited by A. Fersht. *J. Mol. Biol.* **2000** 301, 465–475.

206. Davie, E. & Kulman, J. An Overview of the Structure and Function of Thrombin. *Semin. Thromb.*

Hemost. **2006** 32, 3–15.

207. Bugge, T. H., Antalis, T. M. & Wu, Q. Type II Transmembrane Serine Proteases. *J. Biol. Chem.* **2009** 284, 23177–23181.

208. Barré, O. *et al.* Cleavage Specificity Analysis of Six Type II Transmembrane Serine Proteases (TTSPs) Using PICS with Proteome-Derived Peptide Libraries. *PLoS One* **2014** 9, 1–13.

209. Hohenester, E., Takako, S. & Rupert, T. Crystal structure of a scavenger receptor cysteine-rich domain sheds light on an ancient superfamily. *Nat. Struct. Mol. Biol.* **1999** 6, 228–232.

210. 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, 309–320.

211. Müller, P. *et al.* Interfering with Host Proteases in SARS-CoV-2 Entry as a Promising Therapeutic Strategy. *Curr. Med. Chem.* **2022** 29, 635–665.

212. Wettstein, L., Kirchhoff, F. & Münch, J. The Transmembrane Protease TMPRSS2 as a Therapeutic Target for COVID-19 Treatment. *Int. J. Mol. Sci.* **2022** 23, 1351.

213. 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, 8682–8689.

214. Afar, D., Vivanco, I., Hubert, R. S., Kuo, J. & Chen, E. Catalytic Cleavage of the Androgen-regulated TMPRSS2 Protease Results in Its Secretion by Prostate and Prostate Cancer Epithelia. *Cancer Res.* **2001** 61, 1686–1692.

215. Lin, B., Ferguson, C., White J., T. & Wang, S. Prostate-localized and androgen-regulated expression of the membrane-bound serine protease TMPRSS2. *Cancer Res.* **1999** 59, 4180–4184.

216. Bertram, S. *et al.* Influenza and SARS-Coronavirus Activating Proteases TMPRSS2 and HAT Are Expressed at Multiple Sites in Human Respiratory and Gastrointestinal Tracts. *PLoS One* **2012** 7, 1–8.

217. Donaldson, S. H. *et al.* Regulation of the Epithelial Sodium Channel by Serine Proteases in Human Airways. *J. Biol. Chem.* **2002** 277, 8338–8345.

218. Vaarala, H. & Porvari, S. Expression of transmembrane serine protease TMPRSS2 in mouse and human tissues. *J. Pathol.* **2001** 193, 134–140.

219. Wilson, S., GREER, B. & HOOPER, J. The membrane-anchored serine protease, TMPRSS2,

activates PAR-2 in prostate cancer cells. *Biochem. J.* **2005** 388, 967–972.

220. Mackie, E. J. *et al.* Protease-Activated Receptors: A Means of Converting Extracellular Proteolysis into Intracellular Signals. *IUBMB Life* **2002** 53, 277–281.

221. Pawar, N. R., Buzza, M. S. & Antalis, T. M. Membrane-Anchored Serine Proteases and Protease-Activated Receptor-2–Mediated Signaling: Co-Conspirators in Cancer Progression. *Cancer Res.* **2019** 79, 301–310.

222. Ko, C.-J. *et al.* Androgen-Induced TMPRSS2 Activates Matriptase and Promotes Extracellular Matrix Degradation, Prostate Cancer Cell Invasion, Tumor Growth, and Metastasis. *Cancer Res.* **2015** 75, 2949–2960.

223. R. D'Andrea, M. *et al.* Characterization of Protease-activated Receptor-2 Immunoreactivity in Normal Human Tissues. *J. Histochem. Cytochem.* **1998** 46, 157–164.

224. Cocks, T. M. *et al.* A protective role for protease-activated receptors in the airways. *Nature* **1999** 398, 156–160.

225. Garty, H. & Palmer, L. G. Epithelial sodium channels: function, structure, and regulation. *Physiol. Rev.* **1997** 77, 359–396.

226. 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, 965–975.

227. V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat. Rev. Microbiol.* **2021** 19, 155–170.

228. Böttcher-Friebertshäuser, E., Stein, D. A., Klenk, H.-D. & Garten, W. Inhibition of Influenza Virus Infection in Human Airway Cell Cultures by an Antisense Peptide-Conjugated Morpholino Oligomer Targeting the Hemagglutinin-Activating Protease TMPRSS2. *J. Virol.* **2011** 85, 1554–1562.

229. Laporte, M. & Naesens, L. Airway proteases: an emerging drug target for influenza and other respiratory virus infections. *Curr. Opin. Virol.* **2017** 24, 16–24.

230. Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* **2020** 382, 727–733.

231. Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **2020** 579, 265–269.

232. Lu, R. *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **2020** 395, 565–574.

233. WHO Coronavirus Disease (COVID-19) Dashboard. (**2024**).

234. Majid, M. Evaluation of Coronavirus Families & Covid-19 Proteins: Molecular Modeling Study. *Biointerface Res. Appl. Chem.* **2020** 10, 6039–6057.

235. Yao, H. *et al.* Molecular Architecture of the SARS-CoV-2 Virus. *Cell* **2020** 183, 730–738.

236. Astuti, I. & Ysrafil. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): An overview of viral structure and host response. *Diabetes Metab. Syndr. Clin. Res. Rev.* **2020** 14, 407–412.

237. Chen, S., Jonas, F., Shen, C. & 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, 59–74.

238. Shin, D. *et al.* Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature* **2020** 587, 657–662.

239. Tortorici, M. A. *et al.* Structural basis for human coronavirus attachment to sialic acid receptors. *Nat. Struct. Mol. Biol.* **2019** 26, 481–489.

240. 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. *Mol. Cell* **2020** 78, 779–784.

241. 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, 562–569.

242. Simmons, G. *et al.* Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. *Proc. Natl. Acad. Sci.* **2005** 102, 11876–11881.

243. Shulla, A. *et al.* A Transmembrane Serine Protease Is Linked to the Severe Acute Respiratory Syndrome Coronavirus Receptor and Activates Virus Entry. *J. Virol.* **2011** 85, 873–882.

244. Glowacka, I. *et al.* 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, 4122–4134.

245. Jackson, C. B., Farzan, M., Chen, B. & Choe, H. Mechanisms of SARS-CoV-2 entry into cells. *Nat. Rev. Mol. Cell Biol.* **2022** 23, 3–20.

246. Jackson, C. B., Farzan, M., Chen, B. & Choe, H. Mechanisms of SARS-CoV-2 entry into cells. *Nat. Rev. Mol. Cell Biol.* **2022** 23, 3–20.

247. Park, J.-E. *et al.* Proteolytic processing of Middle East respiratory syndrome coronavirus spikes

expands virus tropism. *Proc. Natl. Acad. Sci.* **2016** 113, 12262–12267.

248. Iwata-Yoshikawa, N. *et al.* TMPRSS2 Contributes to Virus Spread and Immunopathology in the Airways of Murine Models after Coronavirus Infection. *J. Virol.* **2019** 93, 1–15.

249. Fraser, B. J. *et al.* Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. *Nat. Chem. Biol.* **2022** 18, 963–971.

250. Hu, X. *et al.* Discovery of TMPRSS2 Inhibitors from Virtual Screening as a Potential Treatment of COVID-19. *ACS Pharmacol. Transl. Sci.* **2021** 4, 1124–1135.

251. Yamamoto, M. *et al.* The Anticoagulant Nafamostat Potently Inhibits SARS-CoV-2 S Protein-Mediated Fusion in a Cell Fusion Assay System and Viral Infection In Vitro in a Cell-Type-Dependent Manner. *Viruses* **2020** 12, 629.

252. Shrimp, J. H. *et al.* An Enzymatic TMPRSS2 Assay for Assessment of Clinical Candidates and Discovery of Inhibitors as Potential Treatment of COVID-19. *ACS Pharmacol. Transl. Sci.* **2020** 3, 997–1007.

253. Li, K., Meyerholz, D. K., Bartlett, J. A. & McCray, P. B. The TMPRSS2 Inhibitor Nafamostat Reduces SARS-CoV-2 Pulmonary Infection in Mouse Models of COVID-19. *MBio* **2021** 12, 1–11.

254. Sun, Y. J. *et al.* Structure-based phylogeny identifies avoralstat as a TMPRSS2 inhibitor that prevents SARS-CoV-2 infection in mice. *J. Clin. Invest.* **2021** 131, 1–13.

255. Meyer, D. *et al.* 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, 331–343.

256. Seccia, T. M. *et al.* RAndomized Clinical Trial Of NAfamostat Mesylate, A Potent Transmembrane Protease Serine 2 (TMPRSS2) Inhibitor, in Patients with COVID-19 Pneumonia. *J. Clin. Med.* **2023** 12, 6618.

257. Kincaid, J. R. A. *et al.* A sustainable synthesis of the SARS-CoV-2 Mpro inhibitor nirmatrelvir, the active ingredient in Paxlovid. *Commun. Chem.* **2022** 5, 156.

258. Antalis, T. M., Bugge, T. H. & Wu, Q. Membrane-Anchored Serine Proteases in Health and Disease. *PMBTS* **2011** 49, 1–50.

259. Li, S. *et al.* Hepsin enhances liver metabolism and inhibits adipocyte browning in mice. *Proc. Natl. Acad. Sci.* **2020** 117, 12359–12367.

260. 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, 1067–1074.

261. Li, S., Wang, L., Sun, S. & Wu, Q. Hepsin: a multifunctional transmembrane serine protease in pathobiology. *FEBS J.* **2021** 288, 5252–5264.

262. Herter, S. *et al.* 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, 125–136.

263. Hsu, Y.-C. *et al.* Serine protease hepsin regulates hepatocyte size and hemodynamic retention of tumor cells by hepatocyte growth factor signaling in mice. *Hepatology* **2012** 56, 1913–1923.

264. Kirchhofer, D. *et al.* Hepsin activates pro-hepatocyte growth factor and is inhibited by hepatocyte growth factor activator inhibitor-1B (HAI-1B) and HAI-2. *FEBS Lett.* **2005** 579, 1945–1950.

265. Wu, Q. & Parry, G. Hepsin and prostate cancer. *FBS* **2007** 12, 5052–5059.

266. Wilkinson, D. J. *et al.* The serine proteinase hepsin is an activator of pro-matrix metalloproteinases: molecular mechanisms and implications for extracellular matrix turnover. *Sci. Rep.* **2017** 7, 16693.

267. Mukai, S. *et al.* Dysregulation of Type II Transmembrane Serine Proteases and Ligand-Dependent Activation of MET in Urological Cancers. *Int. J. Mol. Sci.* **2020** 21, 2663.

268. Comoglio, P. M., Giordano, S. & Trusolino, L. Drug development of MET inhibitors: targeting oncogene addiction and expedience. *Nat. Rev. Drug Discov.* **2008** 7, 504–516.

269. Kwon, H., Han, J., Lee, K.-Y., Son, S.-H. & Byun, Y. Recent Advances of Hepsin-Targeted Inhibitors. *Curr. Med. Chem.* **2017** 24,.

270. Goswami, R. *et al.* Structure-guided discovery of 2-aryl/pyridin-2-yl-1H-indole derivatives as potent and selective hepsin inhibitors. *Bioorg. Med. Chem. Lett.* **2015** 25, 5309–5314.

271. Franco, F. M. *et al.* Structure-based discovery of small molecule hepsin and HGFA protease inhibitors: Evaluation of potency and selectivity derived from distinct binding pockets. *Bioorg. Med. Chem.* **2015** 23, 2328–2343.

272. Han, Z. *et al.* α-Ketobenzothiazole Serine Protease Inhibitors of Aberrant HGF/c-MET and MSP/RON Kinase Pathway Signaling in Cancer. *ChemMedChem* **2016** 11, 585–599.

273. Gladysz, R. *et al.* Discovery and SAR of Novel and Selective Inhibitors of Urokinase Plasminogen Activator (uPA) with an Imidazo[1,2- a]pyridine Scaffold. *J. Med. Chem.* **2015** 58, 9238–9257.

274. List, K., Bugge, T. H. & Szabo, R. Matriptase: Potent Proteolysis on the Cell Surface. *Mol. Med.* **2006** 12, 1–7.

275. Borensztajn, K. & Spek, C. A. Blood coagulation factor Xa as an emerging drug target. *Expert Opin. Ther. Targets* **2011** 15, 341–349.

276. Tanaka, K. A., Key, N. S. & Levy, J. H. Blood Coagulation: Hemostasis and Thrombin Regulation. *Anesth. Analg.* **2009** 108, 1433–1446.

277. Drag, M. & Salvesen, G. S. Emerging principles in protease-based drug discovery. *Nat. Rev. Drug Discov.* **2010** 9, 690–701.

278. Liang, G. & Phillip Bowen, J. Development of Trypsin-Like Serine Protease Inhibitors as Therapeutic Agents: Opportunities, Challenges, and their Unique Structure-Based Rationales. *Curr. Top. Med. Chem.* **2016** 16, 1506–1529.

279. Nar, H. The role of structural information in the discovery of direct thrombin and factor Xa inhibitors. *Trends Pharmacol. Sci.* **2012** 33, 279–288.

280. Duchêne, D. *et al.* Analysis of Subpocket Selectivity and Identification of Potent Selective Inhibitors for Matriptase and Matriptase-2. *J. Med. Chem.* **2014** 57, 10198–10204.

281. Surabhi, S. & Singh, B. Computer Aided Drug Design: An Overview. *J. Drug Deliv. Ther.* **2018** 8, 504–509.

282. Eriksson, B. I., Quinlan, D. J. & Weitz, J. I. Comparative Pharmacodynamics and Pharmacokinetics of Oral Direct Thrombin and Factor Xa Inhibitors in Development. *Clin. Pharmacokinet.* **2009** 48, 1–22.

283. Hauel, N. H. *et al.* Structure-Based Design of Novel Potent Nonpeptide Thrombin Inhibitors. *J. Med. Chem.* **2002** 45, 1757–1766.

284. Kumar, A. A., Vine, K. L. & Ranson, M. Recent Advances in Targeting the Urokinase Plasminogen Activator with Nanotherapeutics. *Mol. Pharm.* **2023** 20, 2766–2780.

285. Lee, J., Choi, M.-K. & Song, I.-S. Recent Advances in Doxorubicin Formulation to Enhance Pharmacokinetics and Tumor Targeting. *Pharmaceuticals* **2023** 16, 802.

286. Adeniyi, A. A., Muthusamy, R. & Soliman, M. E. New drug design with covalent modifiers. *Expert Opin. Drug Discov.* **2016** 11, 79–90.

287. Williams, M. A. Protein–Ligand Interactions: Fundamentals. in (**2013**). 3–34.

288. Ludewig, S., Kossner, M., Schiller, M., Baumann, K. & Schirmeister, T. Enzyme Kinetics and Hit Validation in Fluorimetric Protease Assays. *Curr. Top. Med. Chem.* **2010** 10, 368–382.

289. Zimmerman, M., Ashe, B., Yurewicz, E. C. & Patel, G. Sensitive assays for trypsin, elastase, and chymotrypsin using new fluorogenic substrates. *Anal. Biochem.* **1977** 78, 47–51.

290. Carmel, A., Zur, M., Yaron, A. & Katchalski, E. Use of substrates with fluorescent donor and acceptor chromophores for the kinetic assay of hydrolases. *FEBS Lett.* **1973** 30, 11–14.

291. Tuppy, H., Wiesbauer, U. & Wintersberger, E. Aminosäure- p -nitroanilide als Substrate für Aminopeptidasen und andere proteolytische Fermente. *Hoppe-Seyler´s Zeitschrift für Physiol. Chemie* **1962** 329, 278–288.

292. Zimmerman, M., Yurewicz, E. & Patel, G. A new fluorogenic substrate for chymotrypsin. *Anal. Biochem.* **1976** 70, 258–262.

293. Hu, H. *et al.* FRET-based and other fluorescent proteinase probes. *Biotechnol. J.* **2014** 9, 266–281.

294. Auld, D. S. Fluorometric Assays. in *Proteolytic Enzymes* (Springer Berlin Heidelberg, **1999**). 30–48.

295. Walters, W. P. & Namchuk, M. Designing screens: how to make your hits a hit. *Nat. Rev. Drug Discov.* **2003** 2, 259–266.

296. Rishton, G. M. Molecular diversity in the context of leadlikeness: compound properties that enable effective biochemical screening. *Curr. Opin. Chem. Biol.* **2008** 12, 340–351.

297. Pacholarz, K. J., Garlish, R. A., Taylor, R. J. & Barran, P. E. Mass spectrometry based tools to investigate protein–ligand interactions for drug discovery. *Chem. Soc. Rev.* **2012** 41, 4335.

298. Yung-Chi, C. & Prusoff, W. H. Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem. Pharmacol.* **1973** 22, 3099–3108.

299. Dixon, M. The graphical determination of Km and Ki. *Biochem. J.* **1972** 129, 197–202.

300. Klein, P. *et al.* Naphthoquinones as Covalent Reversible Inhibitors of Cysteine Proteases—Studies on Inhibition Mechanism and Kinetics. *Molecules* **2020** 25, 2064.

301. Copeland, R. A., Basavapathruni, A., Moyer, M. & Scott, M. P. Impact of enzyme concentration and residence time on apparent activity recovery in jump dilution analysis. *Anal. Biochem.* **2011** 416, 206–210.

302. Mons, E., Roet, S., Kim, R. Q. & Mulder, M. P. C. A Comprehensive Guide for Assessing Covalent Inhibition in Enzymatic Assays Illustrated with Kinetic Simulations. *Curr. Protoc.* **2022** 2, 1–85.

303. Comley, J. Assay Interference a limiting factor in HTS? *Drug Discovery Wolrd Summer* **2003** 91–98.

304. Thorne, N., Auld, D. S. & Inglese, J. Apparent activity in high-throughput screening: origins of compound-dependent assay interference. *Curr. Opin. Chem. Biol.* **2010** 14, 315–324.

305. Jadhav, A. *et al.* Quantitative Analyses of Aggregation, Autofluorescence, and Reactivity Artifacts in a Screen for Inhibitors of a Thiol Protease. *J. Med. Chem.* **2010** 53, 37–51.

306. Maus, H. *et al.* Next Generation of Fluorometric Protease Assays: 7-Nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-Amides) as Class-Spanning Protease Substrates. *Chem. – A Eur. J.* **2023** 29,.

307. STEPHAN, C. *et al.* Hepsin is Highly Over Expressed in and a New Candidate for a Prognostic Indicator in Prostate Cancer. *J. Urol.* **2004** 171, 187–191.

308. Ong, I. L. H. & Yang, K.-L. Recent developments in protease activity assays and sensors. *Analyst* **2017** 142, 1867–1881.

309. John, C. Assay interference. *Drug Discov. Wolrd Summer* **2003** 91–98.

310. Grant, S. K., Sklar, J. G. & Cummings, R. T. Development of Novel Assays for Proteolytic Enzymes Using Rhodamine-Based Fluorogenic Substrates. *SLAS Discov.* **2002** 7, 531–540.

311. Jiang, C. *et al.* NBD-based synthetic probes for sensing small molecules and proteins: design, sensing mechanisms and biological applications. *Chem. Soc. Rev.* **2021** 50, 7436–7495.

312. Jo, S., Kim, H., Kim, S., Shin, D. H. & Kim, M. Characteristics of flavonoids as potent MERS-CoV 3C-like protease inhibitors. *Chem. Biol. Drug Des.* **2019** 94, 2023–2030.

313. Navitha Reddy, G., Jogvanshi, A., Naikwadi, S. & Sonti, R. Nirmatrelvir and ritonavir combination: an antiviral therapy for COVID-19. *Expert Rev. Anti. Infect. Ther.* **2023** 21, 943–955.

314. Klezovitch, O. *et al.* Hepsin promotes prostate cancer progression and metastasis. *Cancer Cell* **2004** 6, 185–195.

315. Knaff, P. M. *et al.* From In Silico to Experimental Validation: Tailoring Peptide Substrates for a Serine Protease. *Biomacromolecules* **2020** 21, 1636–1643.

316. Fuchs, A. V. *et al.* Enzyme cleavable nanoparticles from peptide based triblock copolymers. *Nanoscale* **2013** 5, 4829.

317. Sun, G. *et al.* Structural Basis of Covalent Inhibitory Mechanism of TMPRSS2-Related Serine Proteases by Camostat. *J. Virol.* **2021** 95, 1–9.

Curriculum Vitae