

Charles University
Faculty of Science

Study program: Organic Chemistry



Mgr. Martin Hadzima

Syntéza inhibitorů a substrátů
pro inhibici, vizualizaci a cílení proteáz

Synthesis of inhibitors and substrates
for the inhibition, visualization and targeting of proteases

Doctoral thesis

Supervisor: prof. RNDr. Jan Konvalinka, Ph.D.

Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences

Prague, 2024

PROHLÁŠENÍ

Prohlašuji, že jsem disertační práci zpracoval samostatně, pod vedením školitele prof. RNDr. Jana Konvalinky, Ph.D. a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze dne 01.11.2024

Mgr. Martin Hadzima

IDENTIFIKAČNÍ ZÁZNAM

HADZIMA, Martin. *Syntéza inhibitorů a substrátů pro inhibici, vizualizaci a cílení proteáz. [Synthesis of inhibitors and substrates for the inhibition, visualization and targeting of proteases]*. Praha, 2024, 166. Disertační práce. Univerzita Karlova, Přírodovědecká fakulta; Proteázy lidských patogenů, Ústav organické chemie a biochemie AV ČR, v.v.i., vedoucí práce Konvalinka, Jan.

KLÍČOVÁ SLOVA

proteázy, rakovina, inhibitor, substrát, katepsin, gingipain, syntéza

KEY WORDS

protease, cancer, inhibitor, substrate, cathepsin, gingipain, synthesis

“There are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns — the ones we don't know we don't know.”

Donald H. Rumsfeld

ACKNOWLEDGEMENT

First of all, I would like to thank my supervisor Jan Konvalinka for all the challenging projects and opportunities that emerged on the way. This has truly been a journey that crafted the scientist and the person I am today. I would like to thank Pavel Majer for accommodating me in his group, while providing me with additional exciting projects and opportunities. A big thank you goes out to Aleš Machara, my first scientific mentor, who showed me the true spirit of organic chemistry. I would like to express my gratitude to Barbara Slusher and Matthew Bogyo for welcoming me into their groups and giving me the opportunity to see the world along with world-class science.

My research has always been very collaborative, and I would like to thank all my collaborators, colleagues and friends for their input in the projects I was involved in. Namely, my thanks belong to Radka Houšteká from the group of Michael Mareš for the fruitful collaboration on inhibitors of cathepsin D, Adéla Marcalíková and Tereza Gistrová for their work on gingipains, Franco Faucher and Kristýna Blažková for helping me tremendously with my first author paper, Filip Wichterle for his help with the crystal structures, Martin Dračinský for his help with the interpretation of NMR spectra, Michael Adámek for his help with the statistical analysis, Pavel Šácha, Reza Zamani, Dominik Musil, Jana Pokorná and many others for their work. My thanks also belong to Lukáš Tenora and Kateřina Tenorová for creating a warm environment on “our side of the lab”, and Jitka Bařinková for the technical support.

This section belongs to the people who were more than colleagues to me – friends that helped me not only with my research but also with all the struggles that life brings sometimes. Joshua Smith, Kateřina Radilová, David Kodr, Marta Lubos, Kristýna Krulová, Zuzana Šmahelová, Adéla Šimková, Lucie Svobodová, Jan Malý – thank you for being there for me. And of course, you wouldn't be reading this if Róbert Reiberger would not be here for me when I needed it in the last couple of years. They say you don't pick your family and neighbors – well, I got lucky with the best neighbors that became my second family – Adam Weinzettl, Samuel Hyravý and Markéta Linhartová. NH, EF, BH – You also played an important part in all of this.

Na záver by som chcel poďakovať svojej rodine – Jarke, Lucke, Milanovi a Márii – za ich podporu, morálnu aj finančnú, bez ktorej by som sa svojim cieľom nemohol ani len priblížiť. Ďakujem vám za všetko, čo som len vďaka vám mohol vidieť, zažiť a dosiahnuť.

TABLE OF CONTENTS

ABSTRAKT	1
ABSTRACT	2
LIST OF ABBREVIATIONS.....	3
1 INTRODUCTION	5
1.1 Proteases and Their Role in Pathogenesis	5
1.2 Macrocyclic Inhibitors of Human Cathepsin D	7
1.2.1 Cathepsin D and Cancer Progression	7
1.2.2 Inhibitors of Cathepsin D.....	9
1.2.3 Rational Design of Macrocyclic Inhibitors of Cathepsin D.....	16
1.3 Linker-Modified Gingipain Inhibitors for Probe Development.....	18
1.3.1 <i>Porphyromonas gingivalis</i> Infection – Periodontitis and Beyond	18
1.3.2 Major <i>Porphyromonas gingivalis</i> Proteases – Gingipains and Their Inhibitors	20
1.3.3 Design of Linker-Modified Gingipain Inhibitors for Probe Development.....	25
1.3.4 Passerini Reaction in Preparation of Gingipain Inhibitors.....	29
1.4 Polymer-Tethered Quenched Fluorescent Probes for Imaging of Tumor- Associated Proteases.....	32
1.4.1 Fluorescent Probes in Cancer Imaging.....	32
1.4.2 Design of Polymer-Tethered Quenched Fluorescent Probes for Imaging of Cathepsins	37
2 AIMS OF WORK.....	44
3 RESULTS AND DISCUSSION	45
3.1 Macrocyclic Inhibitors of Human Cathepsin D	45
3.1.1 Preparation of Statine Analogues	46
3.1.2 Preparation of Non-Commercial ω -Amino Aliphatic Acids	50
3.1.3 Preparation of Macrocyclic Inhibitors.....	52
3.1.4 Evaluation of Macrocyclic Inhibitors	56
3.2 Linker-Modified Gingipain Inhibitors for Probe Development.....	61
3.2.1 Preparation of Linker-Modified Lys-Gingipain Inhibitors.....	61
3.2.2 Evaluation of Linker-Modified Lys-Gingipain Inhibitors	68

3.2.3	Preparation of Linker-Modified Arg-Gingipain Inhibitors.....	70
3.2.4	Evaluation of Linker-Modified Arg-Gingipain Inhibitors	77
3.3	Polymer-Tethered Quenched Fluorescent Probes for Imaging of Tumor-Associated Proteases.....	78
3.3.1	Preparation of Low-Molecular Probe Precursors and Corresponding Macromolecular Probes	78
3.3.2	Evaluation of Macromolecular Probes <i>In Vitro</i>	86
3.3.3	Evaluation of Macromolecular Probes <i>In Vivo</i>	88
4	EXPERIMENTAL PART	95
4.1	General information.....	95
4.2	Macrocyclic Inhibitors of Human Cathepsin D	96
4.2.1	Precursors of Macrocyclic Inhibitors.....	96
4.2.2	Macrocyclic Inhibitors	104
4.3	Linker-Modified Gingipain Inhibitors for Probe Development.....	124
4.4	Polymer-Tethered Quenched Fluorescent Probes for Imaging of Tumor-Associated Proteases.....	140
4.4.1	Supplementary Figures and Tables	140
4.4.2	Solid-Phase Peptide Synthesis	142
4.4.3	Preparation of Low-Molecular Probe Precursors.....	143
4.4.4	4T1 Breast Tumor Model	150
5	CONCLUSION.....	152
6	LIST OF AUTHOR'S PUBLICATIONS	154
7	REFERENCES.....	155

ABSTRAKT

Tato disertační práce se zabývá syntézou inhibitorů a substrátů několika klinicky významných proteáz. Práce využívá racionální návrh cílových sloučenin vycházející z dostupných krystalových struktur v kombinaci se známými syntetickými postupy. Práce je rozdělena do tří částí, z nichž každá je zaměřena na jinou proteázu nebo třídu proteáz.

První část popisuje novou třídu biomimetických makrocyclických inhibitorů lidského katepsinu D, inspirovanou krystalovou strukturou s přírodním inhibitorem pepstatinem A. Zvýšená exprese katepsinu D je spojována s rozvojem některých druhů nádorového bujení, proto jsou inhibitory této proteázy potenciálními terapeutickými nástroji. Bylo syntetizováno 27 inhibitorů, které odhalily několik trendů ve vztahu mezi strukturou a aktivitou napříč jednotlivými vazebnými místy. Optimalizací skeletu byly získány inhibitory s nanomolární aktivitou.

Druhá část se zabývá syntézou modifikovaných inhibitorů proteáz z *Porphyromonas gingivalis* (gingipainů). Jde o patogenní bakterie způsobující paradentózu, u kterých je také studován vztah k vývoji Alzheimerovy nemoci. Tyto inhibitory gingipainů byly základem vývoje chemických nástrojů pro detekci a kvantifikaci gingipainů. Struktura těchto sloučenin vychází ze známých inhibitorů KYT-36 a KYT-1. Úspěšně se podařilo syntetizovat čtyři modifikované inhibitory, které byly použity jako sondy v experimentální esaji pro kvantifikaci gingipainů a pro hledání nových inhibitorů.

Třetí část této práce se zaměřuje na modifikaci známého zhášeného fluorescenčního substrátu 6QC, který je štěpen katepsiny a využívá se k zobrazování solidních nádorů. Substrát 6QC byl synteticky modifikován a připojen k makromolekulárnímu nosiči za účelem studia vlastností tohoto typu makromolekulárních zobrazovacích látek. V rámci práce byly syntetizovány makromolekulární sondy s různými spojkami a různou architekturou sondy. Jejich účinnost byla studována *in vitro* a *in vivo* na myším modelu rakoviny prsu. Tato strategie se ukázala jako úspěšná, jelikož většina nově připravených makromolekulárních sond vykázala zlepšení kontrastu a biodistribuce *in vivo*.

ABSTRACT

This dissertation thesis focuses on the synthesis of inhibitors and substrates of various clinically relevant proteases. It employs rational design guided by available crystal structures, along with known synthetic methodologies. The thesis is organized into three sections, each targeting a different protease or class of proteases.

In the first section, a novel class of biomimetic macrocyclic inhibitors of human cathepsin D is reported, inspired by the crystal structure with a natural inhibitor, pepstatin A. Increased expression of cathepsin D is associated with the development of several cancer types, and inhibitors of this protease are thus regarded as potential anti-cancer therapeutics. A total of 27 inhibitors were synthesized, and several trends in structure-activity relationships across individual binding subsites were revealed. Scaffold optimization yielded highly potent inhibitors with low-nanomolar activity.

The second section addresses the synthesis of modified inhibitors of proteases from *Porphyromonas gingivalis* (gingipains). These are pathogenic bacteria causing periodontitis that are also implicated in the development of Alzheimer's disease. The inhibitors served as a basis for the development of novel chemical biology tools for the detection and quantification of gingipains. The structure of these inhibitors is based on the known inhibitors KYT-36 and KYT-1. In total, four modified inhibitors were synthesized and are currently employed as probes in an experimental assay for gingipain quantification and screening of new inhibitors.

The third section focuses on the modification of a known quenched fluorescent substrate 6QC, which is cleaved by cathepsins and used for imaging of solid tumors. 6QC was synthetically modified and attached to a macromolecular carrier to study the properties of such macromolecular imaging agents. Macromolecular probes with various linkers and probe architectures were synthesized and evaluated *in vitro*, in a purified enzyme and cellular assay, and *in vivo* in a mouse breast cancer model. The strategy has proved to be successful since the majority of newly prepared macromolecular probes showed significantly enhanced contrast and biodistribution *in vivo*.

LIST OF ABBREVIATIONS

AAA	Amino acid analysis
AcOH	Acetic acid
AD	Alzheimer's disease
Arg	Arginine
BBB	Blood-brain barrier
Boc	<i>tert</i> -Butyloxycarbonyl
chex	Cyclohexane
CatD	Cathepsin D
CatL	Cathepsin L
Cbz	Carboxybenzyloxy
COPD	Chronic obstructive pulmonary disease
CPP	Cell-penetrating peptide
dALA	δ -Aminolevulinic acid
DCC	<i>N,N'</i> -Dicyclohexylcarbodiimide
DIPA	Diisopropylamine
DIPEA	<i>N,N</i> -Diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMF	<i>N,N</i> -Dimethylformamide
DSC	<i>N,N'</i> -Disuccinimidyl carbonate
DtBAD	Di- <i>tert</i> -butyl azodicarboxylate
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EPR	Enhanced permeability and retention
Eq.	Equivalent
ER	Estrogen receptor
EtOAc	Ethyl acetate
FA	Formic acid
FAP	Fibroblast activation protein
FDA	The U.S. Food and Drug Administration
Fmoc	Fluorenylmethyloxycarbonyl
FR α	Folate receptor α
FRET	Förster resonance energy transfer
GCPII	Glutamate carboxypeptidase II
HATU	Hexafluorophosphate azabenzotriazole tetramethyl uronium
HBTU	Hexafluorophosphate benzotriazole tetramethyl uronium
hCatD	Human cathepsin D
hCatK	Human cathepsin K
HCTU	Hexafluorophosphate chlorobenzotriazole tetramethyl uronium

HPLC	High-performance liquid chromatography
Iva	Isovaleroyl
Kgp	Lysine-specific gingipain
Lys	Lysine
MB	Methylene blue
MCPBA	<i>meta</i> -Chloroperoxybenzoic acid
MeOH	Methanol
MMP	Matrix metalloproteinase
MS	Mass spectrometry
mTORC1	Mechanistic Target of Rapamycin Complex 1
NE	Neutrophil elastase
NHS	<i>N</i> -Hydroxysuccinimide
NIR	Near-infrared
NMR	Nuclear magnetic resonance
PADAM	Passerini reaction – Amine Deprotection – Acyl Migration
PDA	Photodiode array
PEG	Polyethylene glycol
PepA	Pepstatin A
PheSta	Phenylalanine statine analogue
Phg	Phenylglycine
pHPMA	Poly <i>N</i> -(2-Hydroxypropyl)methacrylamide
PpIX	Protoporphyrin IX
PyBrOP	Bromo-tris-pyrrolidino-phosphonium hexafluorophosphate
Rgp	Arginine-specific gingipain
RP	Reversed-phase
SAR	Structure-activity relationship
SBR	Signal-to-background ratio
SPPS	Solid-phase peptide synthesis
Sta	Statine
TBR	Tumor-to-background ratio
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TSTU	<i>N,N,N',N'</i> -Tetramethyl- <i>O</i> -(<i>N</i> -succinimidyl)uronium tetrafluoroborate
UPLC	Ultra-performance liquid chromatography
UPLC-MS	Ultra-performance liquid chromatography – mass spectrometry

1 INTRODUCTION

1.1 Proteases and Their Role in Pathogenesis

Proteases, or proteinases, are enzymes that play a critical role in numerous physiological and pathological processes by catalyzing the hydrolysis of peptide bonds in proteins. These enzymes, classified based on the catalytic residues in their active sites (e.g., serine, cysteine, aspartic, and metalloproteases), are pivotal in cellular regulation, tissue remodeling, immune response, and apoptosis. While proteases maintain essential functions in normal physiology, dysregulated protease activity is associated with numerous diseases, including cancer, inflammatory disorders, neurodegenerative diseases, and infectious conditions. Given their dual role in health and disease, proteases have emerged as attractive targets for therapeutic intervention and as biomarkers in clinical diagnostics.

Numerous examples of the role of human proteases in disease development can be found in the literature. First, several protease classes are known to be involved in tumor progression and metastasis. For example, matrix metalloproteinases (MMPs) are commonly overexpressed in malignant tumors and promote metastasis via degradation of collagen and gelatin in the extracellular matrix.¹⁻³ Elevated MMP levels are correlated with poor prognosis in various cancers, highlighting their potential as biomarkers and therapeutic targets. Another interesting class of proteases are cathepsins, a family of lysosomal cysteine proteases. While human cathepsin D (hCatD) is known to be involved in breast cancer progression,⁴⁻⁶ overactive human cathepsin K (hCatK) has been shown to induce inflammation-related tissue damage.⁷⁻⁹ It contributes to the progression of osteoporosis and osteoarthritis and odanacatib, a selective cathepsin K inhibitor demonstrated efficacy in reducing bone resorption. Even though its development was discontinued due to adverse cardiovascular effects, hCatK remains a relevant clinical target. Another important protease, neutrophil elastase (NE), is also linked to chronic inflammatory diseases, such as chronic obstructive pulmonary disease (COPD) and rheumatoid arthritis.¹⁰⁻¹² In COPD, for example, the overactivity of NE results in tissue damage and airway remodeling, leading to progressive lung function decline. Pharmacological inhibition of NE has been explored as a therapeutic strategy to mitigate the inflammatory damage in COPD, and NE inhibitors are in various stages of clinical development.

In addition to human proteases, the inhibition of pathogenic proteases can often lead to reduction in virulence and, therefore, can be exploited as a treatment strategy. For example, gingipains, cysteine proteases from *Porphyromonas gingivalis* (*P. gingivalis*) are indispensable for the virulence of the bacteria and represent a promising target for the treatment.^{13–15} In viral infections, proteases play a role in viral replication and maturation. The HIV-1 protease, for example, is essential for processing viral polyproteins into functional proteins required for viral replication.^{16–18} Protease inhibitors, such as lopinavir and ritonavir, are core components of antiretroviral therapy, illustrating the effectiveness of targeting viral proteases to control infection. Similarly, protease inhibitors targeting the main protease of SARS-CoV-2 have been explored as potential treatments for COVID-19, with drugs like nirmatrelvir showing promising results in clinical trials.^{19–21}

The diverse roles of proteases in physiological and pathological processes underline their clinical importance. Their involvement in cancer metastasis, inflammation, neurodegeneration, and infection makes them valuable targets for drug discovery and diagnostic development. In this work, three proteases with different clinical relevance will be targeted. First, a new class of macrocyclic hCatD inhibitors will be introduced. Second, linker-modified gingipain inhibitors, serving as functional probes for assay development will be reported. Third, a known fluorescent pan-cathepsin substrate used for tumor imaging was modified and attached to a macromolecular scaffold to create a new class of macromolecular probes. Although the three projects are distinguished by notable differences in the biological background, they are unified by the core concept of protease targeting.

1.2 Macrocyclic Inhibitors of Human Cathepsin D

1.2.1 Cathepsin D and Cancer Progression

Human cathepsin D (hCatD) is a pepsin-family lysosomal aspartic protease, part of the broader cathepsin family, which plays a critical role in the degradation of proteins within the cell. It is ubiquitously expressed in most tissues and is particularly abundant in lysosomes, where it functions optimally due to the acidic environment to break down a wide range of substrates, including both endogenous and exogenous proteins. Besides its major role in protein degradation, hCatD is also involved in several specialized functions, such as the activation of enzymatic precursors, antigen presentation, peptide hormone processing, apoptosis regulation, cytoskeletal protein degradation and the remodeling of the extracellular matrix. Interestingly, it can also be found in the extracellular space⁴ and is one of the cathepsins that retains its activity at neutral pH.²² Lkhider et al. have shown that rat cathepsin D (CatD) was able to cleave prolactin at physiological pH outside the cell.²² While mice knock-out studies have shown that CatD is not required for embryonal development, it is essential for postnatal homeostasis.



Figure 1.1 Crystal structure of hCatD light (orange) and heavy (purple) chains as published by Baldwin et al. (PDB: 1LYA)²³

While proteolytic activity of hCatD is essential for normal physiological processes such as protein turnover and antigen processing, the dysregulation of hCatD has been implicated in various pathological conditions, including cancer, neurodegenerative

disorders, and inflammatory diseases such as acute pancreatitis and osteoarthritis.^{4,22,24} However, the clinical interest in hCatD has surged mostly due to its involvement in cancer progression and metastasis. Overexpression of hCatD has been correlated with the invasive behavior of certain tumor cells, making it an attractive target for therapeutic intervention. For example, hCatD is overexpressed by breast cancer cells and its elevated levels are associated with increased tumor progression.^{4,5} Clinically, hCatD is used as an independent marker of breast cancer progression and its overexpression is associated with the occurrence of clinical metastasis.^{6,25} For example, Alhudiri et al. assessed the expression of hCatD by immunohistochemical staining in a series of early-stage operable breast cancer, as shown in **Figure 1.2a-b**.²⁶ Analysis of breast cancer-specific survival has confirmed that hCatD-positive patients exhibited a significantly lower probability of survival compared to hCatD-negative patients (**Figure 1.2c**). The same trend applied for patients with the estrogen receptor (ER) positive subgroup. Interestingly, patients positive for both ER and hCatD and treated with tamoxifen have shown a higher survival rate than hCatD-negative patients.

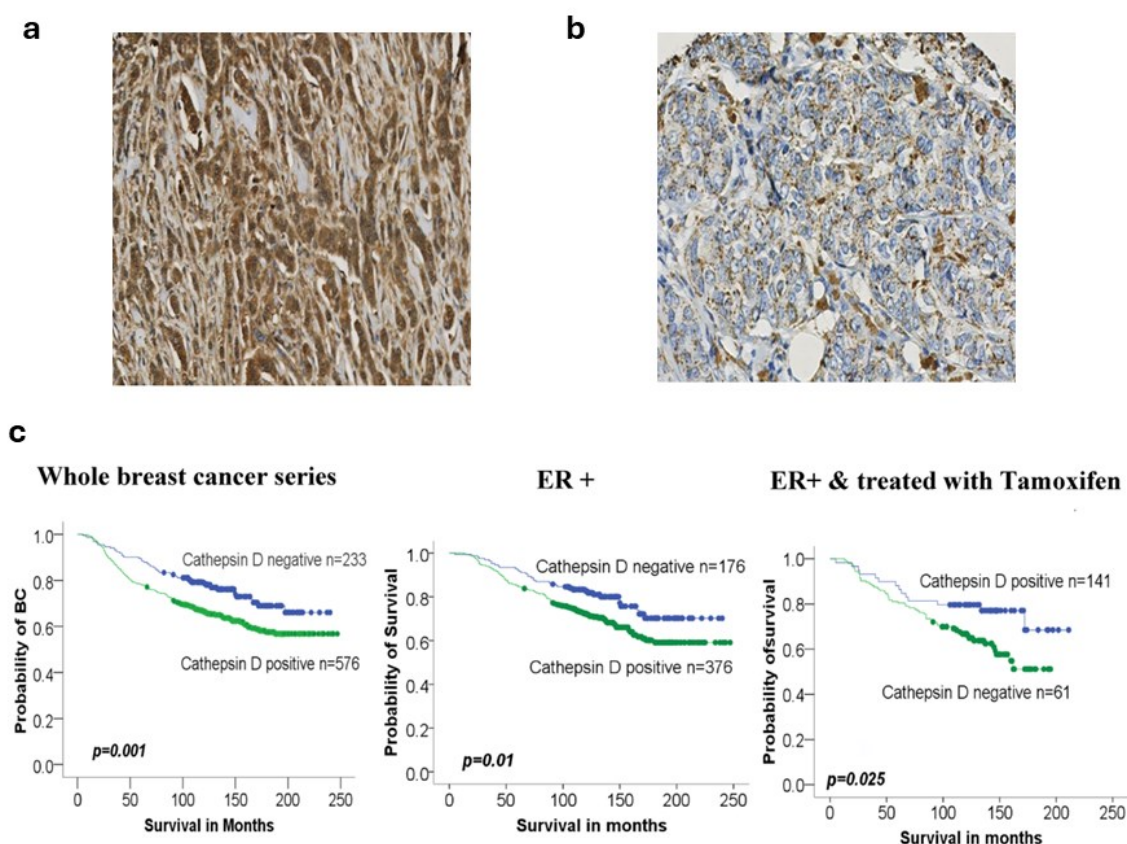


Figure 1.2 Analysis of breast cancer-specific survival in early-stage operable breast cancer by Alhudiri et al.²⁶: (a) hCatD-positive stain tissue sample; (b) hCatD-negative stained tissue sample; (c) Analysis of breast cancer-specific survival, ER+ - estrogen receptor-positive patients.

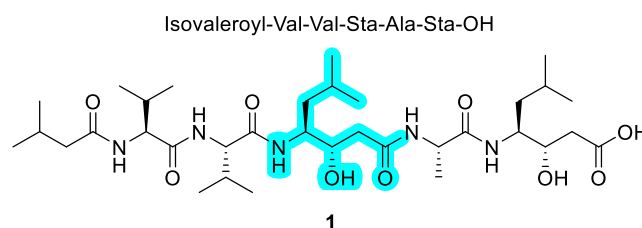
The mechanism of hCatD's contribution to cancer progression has been studied thoroughly. A comprehensive review by Rochefort et al. describes multiple potential mechanisms by which hCatD influences tumor growth and metastasis.²⁷ It is believed that the influence is not only mediated by its proteolytic activity, which can facilitate extracellular matrix degradation, but also by its influence on cell signaling pathways, promoting angiogenesis, and aiding in immune evasion. A recent study by Ketterer et al. used a CatD knockout mouse model to show that CatD deficiency blocked tumor development by manipulating the mechanistic Target of Rapamycin Complex 1 (mTORC1) signaling pathway, which remained permanently disabled in CatD deficient cells.²⁸ In line with other results, CatD deficient mice also shown delayed tumor growth compared to the wild type. Another study proposed that CatD is involved in the regulation of hepsin, a protein responsible for the growth and maintenance of normal tissue morphology among other important functions.²⁹ CatD promotes ubiquitination of hepsin leading to its subsequent proteasomal degradation. According to this study, hepsin expression was negatively correlated with elevated CatD levels. Elevated CatD levels were confirmed to promote metastasis *in vivo*, while this effect was inhibited by the presence of hepsin.

As described above, the literature provides strong and convincing evidence regarding the role of CatD in cancer progression. Despite its critical physiological roles, the dysregulation of CatD presents an opportunity for therapeutic targeting via antibody-based therapy or small-molecule chemotherapy.^{30,31} In the next chapter, known inhibitors of CatD will be introduced.

1.2.2 Inhibitors of Cathepsin D

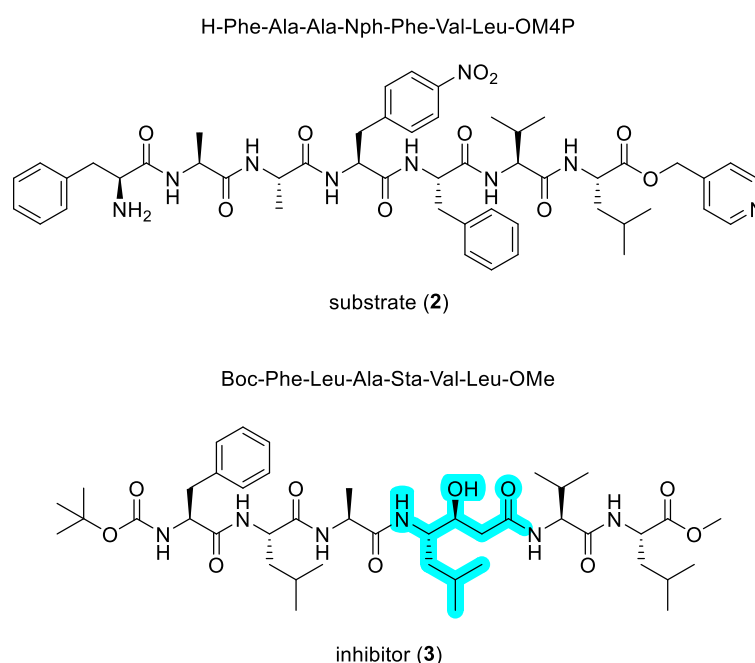
Inhibitors of hCatD have been explored for their potential to modulate its activity in diseases where it is overexpressed or overly active. The design of effective hCatD inhibitors presents several challenges. First, the active site of hCatD is highly conserved among aspartic proteases, which complicates the design of compounds that selectively inhibit hCatD without affecting other proteases in this family, such as cathepsin E or renin. Second, given the enzyme's essential role in protein degradation and cellular homeostasis, complete inhibition of hCatD could lead to unintended side effects, including the disruption of normal lysosomal function. Recent advances in computational drug design and high-throughput screening have enabled the identification of new small-molecule inhibitors with improved properties. Traditional inhibitors of aspartic proteases,

such as pepstatin A (PepA, **1**, **Scheme 1.1**), have been utilized as a starting point for the design of specific hCatD inhibitors.



Scheme 1.1 Structure of pepstatin A (**1**) (the central statine unit is highlighted in light blue)

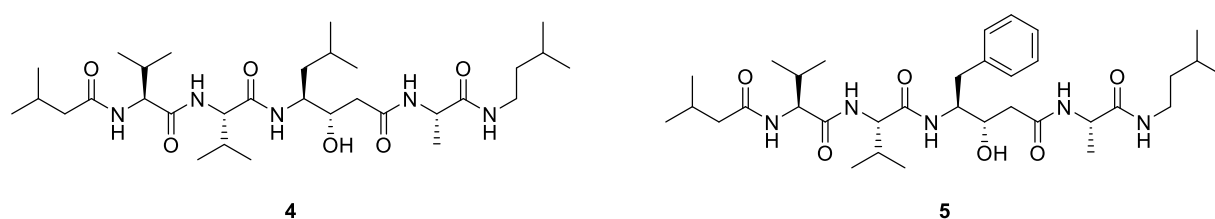
PepA is a naturally occurring hexapeptide isolated from various strains of *Actinomyces*. It is a general non-specific inhibitor of aspartyl proteases and inhibits hCatD with very high affinity ($K_i = 0.5$ nM).³² Inhibition of proteases by PepA is mostly mediated by the statine (Sta; (3*S*,4*S*)-4-amino-3-hydroxy-6-methylheptanoic acid) β -hydroxy- γ -amino acid that acts as a transition state analogue. Early efforts to develop novel CatD inhibitors exploited the structure of PepA by incorporating Sta into various peptides.^{32,33} In 1986, Agarwal et al. reported bovine CatD inhibitors based on analogy with PepA and substrates of the enzyme. In the first series, the scissile dipeptide unit of an efficient CatD substrate was replaced by a Sta or a Sta-Phe unit.



Scheme 1.2 Structure of a hCatD substrate (**2**) and substrate-inspired inhibitor (**3**)³²

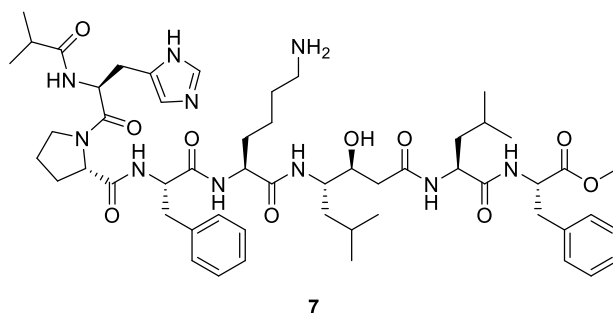
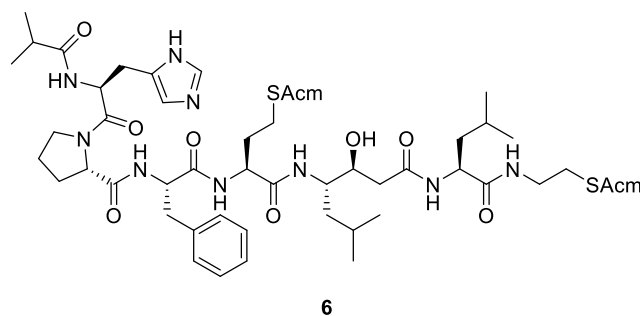
The authors of the study synthesized five inhibitors based on substrate **2**, displayed in **Scheme 1.2**. Inhibitor **3** showed the highest inhibitory potency with a

K_i of 1.1 nM. This compound also exhibited single-digit nanomolar inhibition of pepsin. In the second series, the authors compared PepA to its close analogues. Overall, truncation of the PepA scaffold led to significant reduction in CatD inhibition. One well tolerated substitution was the replacement of the C-terminal Sta unit with isoamyl amide, as shown in inhibitor (**4**, **Scheme 1.3**). This inhibitor exhibited single-digit nanomolar inhibition of CatD ($K_i = 3.7$ nM), comparable to PepA. This further confirmed that the central Sta unit is responsible for CatD inhibition by PepA. Furthermore, replacement of the Sta unit by a phenylalanine analogue (PheSta) led to an almost four-fold increase in inhibitory potency in the case of inhibitor **5**. Another well-tolerated substitution was the replacement of the N-terminal isovaleroyl moiety with a Boc group. The authors have also evaluated PepA analogues with the inverse configuration (*R*) of the Sta hydroxy group. In the case of highly potent CatD inhibitors, the change in configuration led to a significant drop in potency. Surprisingly, in certain cases, inhibitors possessing the unnatural configuration (*S*) at this position exhibited higher inhibitory potency than the naturally occurring isomer (*R*). This could be explained by a change in the binding mode or conformation of the inhibitor, but this hypothesis has not been investigated by the authors.



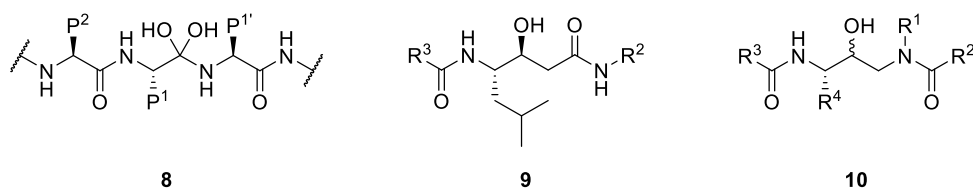
Scheme 1.3 Structure of PepA-derived inhibitors of hCatD

Another study, published in 1990 by Jupp et al., evaluated the specificity of statine-based peptide inhibitors against several human aspartic proteases.³³ These inhibitors were derived from a natural renin substrate by replacing the scissile bond with a Sta unit, similar to the work of Agarwal et al. The authors of this study introduced variations in the neighboring amino acid residues and evaluated changes in inhibitor specificity between renin, pepsin, gastricsin, hCatD, cathepsin E and *Endothia* protease. Furthermore, they also exchanged Sta for PheSta or cyclohexylalanine analogue in certain cases. Overall, since the inhibitors were based on a renin substrate, they displayed the highest inhibitory potency towards renin with the most potent inhibitors reaching subnanomolar affinity. In contrast, most inhibitors exhibited limited ability to inhibit hCatD. Inhibitors (**6**, $K_i = 1$ nM) and (**7**, $K_i = 10$ nM) shown in **Scheme 1.4** were among the most potent hCatD inhibitors in the series.



Scheme 1.4 hCatD inhibitors inspired by natural renin substrate³³

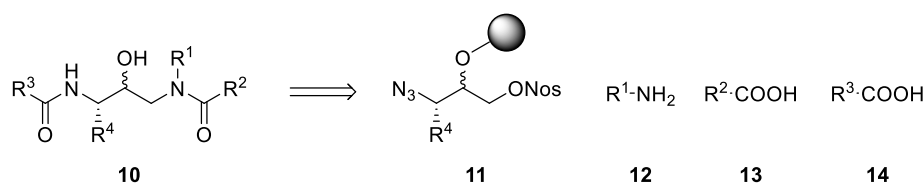
In 1997, Kick et al. published the development of novel non-peptide hCatD inhibitors using a combination of structure-based design and combinatorial chemistry.³⁴ They first proposed to exploit the (hydroxyethyl)amine structural pattern (**10**), suitable for the inhibition of aspartate proteases, inspired by the tetrahedral intermediate of peptide hydrolysis, as shown in **Scheme 1.5**. This design is also analogous to and inspired by the Sta amino acid.



Scheme 1.5 Structure of a peptide hydrolysis tetrahedral intermediate (**8**), Sta-containing peptide inhibitor (**9**) and mechanism-based inhibitor (**10**) as proposed by Kick et al.³⁴

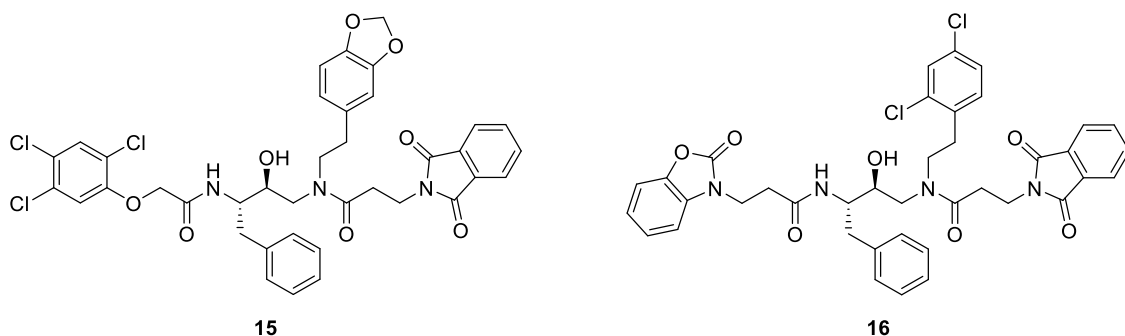
The authors of the paper designed a synthetic route utilizing solid-phase chemistry, enabling the synthesis of a diverse combinatorial library.³⁵ Generally, this approach is based on a multi-step modification of a bead-bound intermediate **11**. First, the alkylation of amine **12** by the bead-bound **11** occurs. The amine intermediate is then connected with acid **13** in a coupling reaction. Next, the reduction of the azide functionality enables attachment of the last fragment **14** in yet another coupling reaction. This approach

allows for variation of four different residues ($R^1 - R^4$) and control over the stereochemistry of the central hydroxy group using a diastereomerically pure precursor of compound **11**.



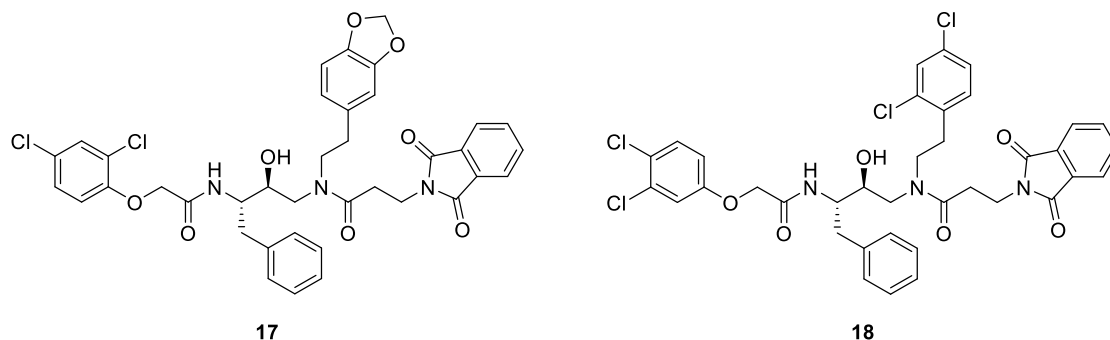
Scheme 1.6 Synthetic approach employed by Kick et al. for synthesis of hCatD inhibitors³⁴

While components **12**, **13** and **14** are generally commercially available, intermediate **11** needs to be synthesized and, therefore, no variations were introduced at this position to expedite the synthesis. According to the authors, a benzyl group was selected as R^4 based on the crystal structure of hCatD with its natural inhibitor PepA. In the preliminary series, the authors synthesized inhibitors with both (*S*) and (*R*) configurations at the carbon possessing the hydroxy group. Interestingly, only compounds with the (*S*) configuration (identical to the Sta stereochemistry) exhibited potent hCatD inhibition. This has also been predicted by computational modeling. Next, the authors identified approx. 700 amines and 1900 carboxylic acids, that were commercially available and compatible with the synthetic sequence. To increase the efficiency of this approach, a computational screening process was employed to create a directed combinatorial library. A structure-based screening process modeled the scaffold in the active site, preserving the binding orientation of PepA in the original crystal structure. This process led to a selection of 50 best-scoring fragments at each position, based on their predicted interaction with CatD. Following this computational selection, individual fragments were clustered, and ten compounds were manually selected for each position to maximize the diversity of the combinatorial library. The authors also designed a control combinatorial library of the same size, without applying the computational screening process. In this case, fragments were clustered without pre-selection and ten compounds were hand-picked for each position to ensure high diversity of functional groups and structural motifs in the synthesized library. Both libraries were then prepared, and the individual compounds were screened in a high-throughput hCatD inhibition assay. The results of this screening confirmed that at all evaluated concentrations (100 nM, 330 nM, 1 μ M, 10 μ M) the directed library yielded more inhibiting compounds than the control library. For example, at 330 nM inhibitor concentration, 23 compounds exhibiting higher than 50% hCatD inhibition were identified using the directed library compared to three in the control library. Compound **15** shown in **Scheme 1.7** inhibited hCatD with a K_i of 73 ± 9 nM, while compound **16** differing in two positions exhibited K_i of 231 ± 31 nM.



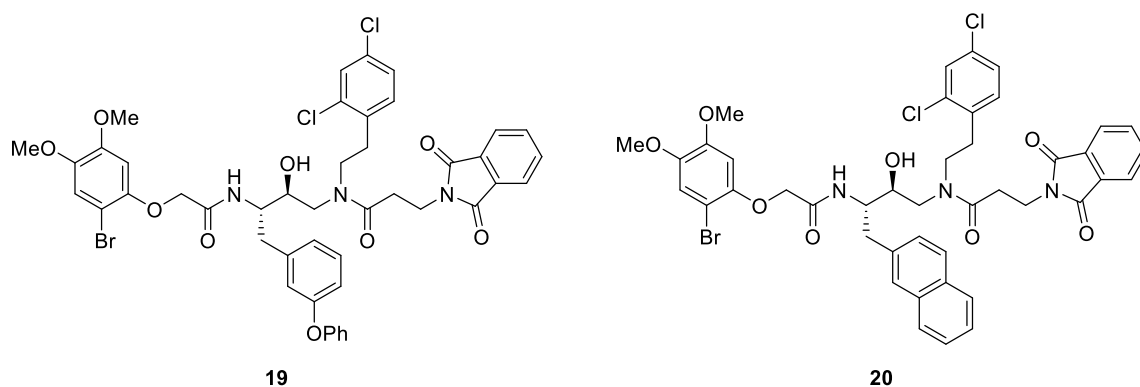
Scheme 1.7 hCatD inhibitors identified in the combinatorial library of Kick et al. – first generation³⁵

Furthermore, the authors designed a second-generation directed library based on the similarity of fragments that yielded the most potent inhibitors in the first-generation library, in order to optimize inhibitory potency. This led to identification of the most potent inhibitors in the series – inhibitor **17** ($K_i = 9$ nM) and **18** ($K_i = 14$ nM).



Scheme 1.8 hCatD inhibitors identified in the combinatorial library of Kick et al. – second generation³⁵

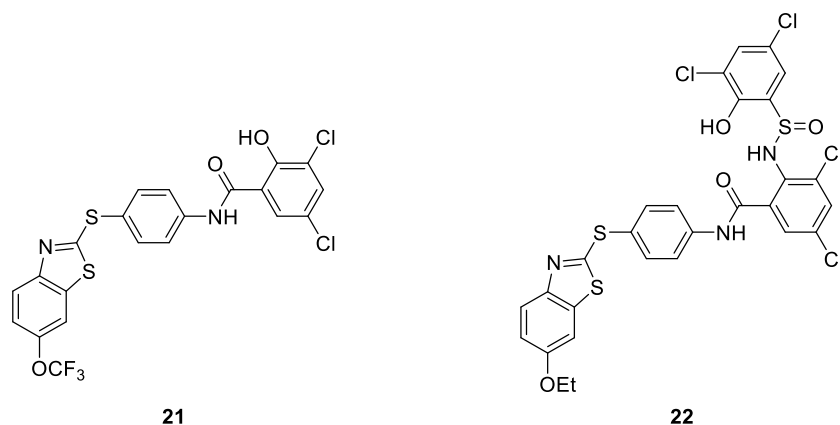
As a follow-up to this study, the authors developed a synthetic approach that enabled the introduction of various substituents at all four positions of scaffold **10**, as shown in **Scheme 1.6**. The authors reported a thorough optimization of the procedure, which includes several interesting steps, such as a Grignard reagent addition to Weinreb's amide and a chelation-controlled stereoselective reduction. The synthetic approach was optimized to yield analytically pure inhibitors in good yields (45 – 64%) over 12 steps performed on solid phase. Using this approach, the authors synthesized a library of 204 compounds to optimize the R^4 substituent of inhibitor scaffold **10** (**Scheme 1.6**). This approach confirmed the hypothesis that the corresponding CatD pocket (P1) can accommodate residues larger than the benzyl group and such substitution can lead to increased inhibitory potency. Inhibitor **19** possessing a diphenyl ether moiety exhibited sub-nanomolar inhibition of hCatD ($K_i = 0.7 \pm 0.3$ nM), while inhibitor **20**, possessing a 2-naphthyl residue, also showed increased inhibitory potency ($K_i = 1.3 \pm 0.2$ nM) compared with compounds shown in **Scheme 1.7**.



Scheme 1.9 hCatD inhibitors identified in the combinatorial library of Lee et al.³⁶

In 2000, the members of the same group also showcased the potential of these compounds to inhibit hCatD in hippocampal slices.³⁷ The study confirmed the relationship between lysosomal activity and the formation of hyperphosphorylated tau fragments, which are connected to Alzheimer's disease progression.

Dumas et al. later identified a novel class of hCatD inhibitors based on a hit (**21**) from a high-throughput library screen, as shown below in **Scheme 1.10**.³⁸ The authors of synthesized approx. 30 analogues of compound **21** to improve its potency. Even though this compound offers an interesting non-peptidomimetic scaffold, the optimized inhibitor **22** only reached an IC_{50} of 250 nM, which is significantly higher than its peptide or peptidomimetic counterparts.



Scheme 1.10 Non-peptidomimetic hCatD inhibitors developed by Dumas et al.³⁸

1.2.3 Rational Design of Macrocyclic Inhibitors of Cathepsin D

In the last chapter, previously known inhibitors of hCatD were introduced. Several classes of peptide, peptidomimetic, or small-molecule inhibitors can be found in the literature. Nevertheless, by closely inspecting the crystal structure (PDB: 1LYB) of hCatD complex with PepA,²³ we hypothesized a novel class of biomimetic macrocyclic inhibitors of hCatD derived from PepA.³⁹ Even though PepA is a highly potent inhibitor of hCatD and is commonly used as a research tool, its therapeutic potential is limited due to unfavorable physicochemical properties, such as low solubility. Furthermore, its large size and high number of hydrogen bond donors and acceptors clearly exceeds Lipinski's rule of five, a rule of thumb commonly used to evaluate the druglikeness of potential drug candidates.⁴⁰

We decided to exploit the U-shaped conformation that PepA adapts upon binding to hCatD, as shown in **Figure 1.3**. We hypothesized that “locking” the inhibitor in the active conformation through macrocyclization would allow us to reduce the size of the inhibitor without a significant loss in inhibitory potency. It has been shown previously that truncation of the PepA linear scaffold inevitably leads to a decrease in biological activity,³² but we proposed that macrocyclization could compensate for this effect.

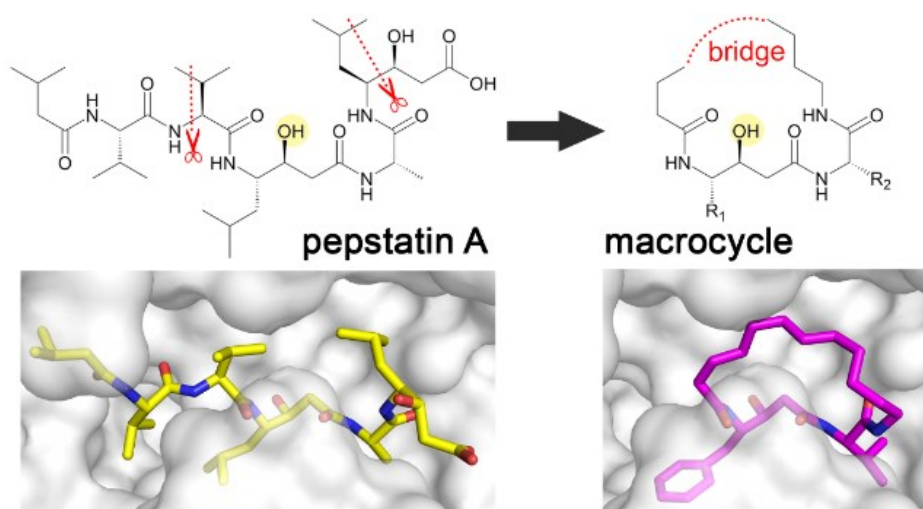


Figure 1.3 Rational biomimetic design of macrocyclic inhibitors of hCatD inspired by the structure of PepA³⁹

On the structural level, we proposed to trim the N-terminal Iva-Val fragment, which occupies the S3 and S4 binding subsites, and the solvent-exposed C-terminal Sta residue, as shown in **Figure 1.4**. Next, we suggested that the P2 and P3' (interacting with the corresponding S2 and S3' subsites) could be connected by an aliphatic linker to yield corresponding macrocycles with high affinity to hCatD. This hypothesis was supported by

the proximity of the two residues in the corresponding crystal structure. The S2 and S3' subsites are quite hydrophobic, so we expected an aliphatic linker to be preferred.⁴¹ This region also contains the S1' subsite, normally accommodating the P1' side chain, which is missing in statine-based inhibitors. The absence of the P1' residue may be the reason that P2 and P3' residues bend toward each other to fill the empty space.

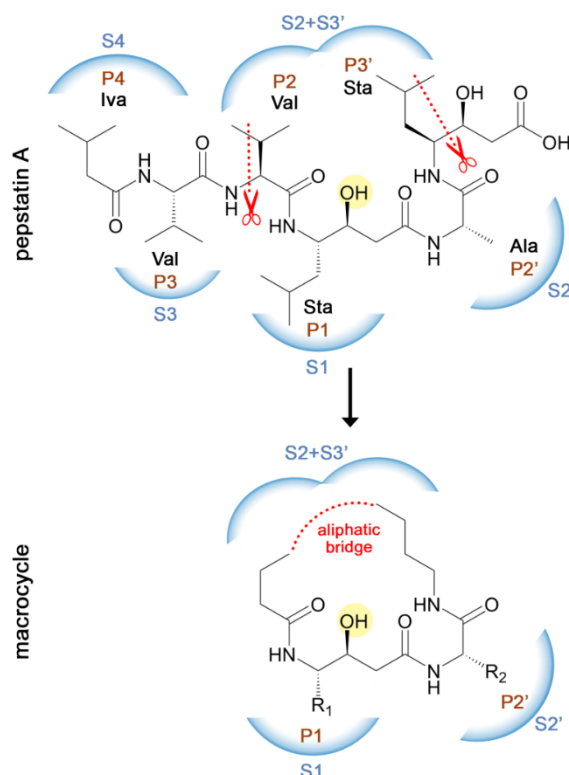


Figure 1.4 Schematic representation of the rational design of the macrocyclic inhibitor scaffold. The scaffold mimics the spatial conformation of PepA in the active site of human hCatD (PDB: 1LYB). Binding subsites (S) are marked and colored in blue; corresponding inhibitor positions (P) are in brown. The central hydroxyl group of the statine residue that interacts with the catalytic aspartates of hCatD is highlighted in yellow. PepA structure: Iva-Val-Val-Sta-Ala-Sta, where Iva is isovaleryl and Sta is statine.

To summarize, we designed a P2-P3' macrocycle that covers a significant portion of the binding site, including the P1 Sta residue, which possesses the transition-state isostere responsible for inhibition. We proposed that locking the inhibitor in the active conformation, along with improving the hydrophobic interactions in the corresponding binding subsites (S2, S1, S2', S3') may compensate for the lost interactions caused by the truncation. The synthesis of the precursors, macrocycles, and evaluation of their inhibitory potency will be reported in this thesis.

1.3 Linker-Modified Gingipain Inhibitors for Probe Development

1.3.1 *Porphyromonas gingivalis* Infection – Periodontitis and Beyond

Periodontitis, a chronic inflammatory disease affecting the supporting structures of the teeth, poses a significant public health challenge worldwide. The pathogenesis of periodontitis is complex, involving an interplay between pathogenic bacteria and the host immune response. Under normal circumstances, over 700 species of commensal bacteria populate the human oral cavity.⁴² Nevertheless, even small changes in composition of the microbiome can lead to a dysbiotic environment affecting the health of an individual.⁴³ In **Figure 1.5**, examples of different stages of periodontal disease caused by such an environment are shown. The prevalence of periodontitis has also been increasing significantly over the past 30 years in most parts of the world.⁴⁴



Figure 1.5 Periodontal disease: Comparison of a healthy tooth with periodontitic tooth (left); gingivitis – early stage of periodontal disease (middle); periodontitis – severe stage of periodontal disease (right). Created with Biorender.com.

Among the diverse microbial community in the periodontal pocket, *P. gingivalis* (**Figure 1.6**) is identified as one of the key etiological agents.^{45–47} This Gram-negative anaerobe exhibits a unique arsenal of virulence factors that enable it to colonize the subgingival environment, evade host defenses, and induce tissue destruction.⁴⁸ Emergence of antibiotic-resistant strains of *P. gingivalis* further highlights the importance of this pathogen.⁴⁹ Recently, multiple studies have demonstrated that *P. gingivalis* is not only a significant pathogen in the development of periodontitis but also has the potential to contribute to the development of other diseases. The role of *P. gingivalis* and other related pathogens in oral and gastrointestinal carcinogenesis has been studied extensively, and a

correlation of their occurrence with progression of multiple types of cancer has been confirmed.⁵⁰ Another important consequence of the systemic inflammation triggered by *P. gingivalis* is the disruption of insulin signaling pathways which leads to reduced insulin sensitivity and insulin resistance.⁵¹ The role of *P. gingivalis* in the development and progression of multiple sclerosis⁵² and atherosclerosis^{53,54} has also been studied extensively.

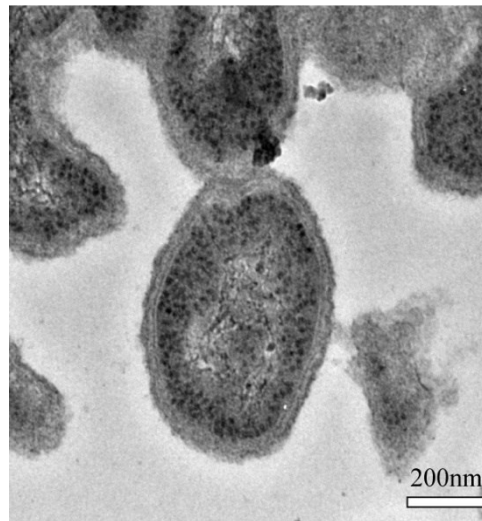


Figure 1.6 Transmission electron microscopy image of *P. gingivalis* as acquired by Fujise et al.⁵⁵

The link between neurodegenerative diseases, such as Parkinson's disease or Alzheimer's disease (AD), and *P. gingivalis* has also been studied thoroughly in recent years.^{56–60} An endotoxin produced by *P. gingivalis* has been found in brain samples of patients with AD in a study published by Poole et al,⁶¹ while Potempa and his group have confirmed the presence of gingipains, major virulence factors of *P. gingivalis*, in such samples.⁶² Bacterial load has also been detected in the cerebrospinal fluid of AD patients. Potempa and his group have performed several interesting *in vitro* and *in vivo* studies to further support the hypothesis that *P. gingivalis* plays a role in development of inflammation in the human central nervous system, correlated with AD progression. For example, oral infection of mice with *P. gingivalis* consistently led to invasion of the bacteria into the brain. Gingipains, proteases secreted by *P. gingivalis* were colocalized with neurons, astrocytes and pathological markers in hippocampus samples of AD patients. Gingipains were also found to cleave tau protein, which has been implicated in the induction of AD pathogenesis. Most importantly, gingipains were proposed as therapeutic targets, as gingipain inhibitors were able to protect neuronal cells from gingipain-induced toxicity both *in vitro* and *in vivo*. Administration of gingipain inhibitors reduced *P. gingivalis* brain infection and prevented loss of interneurons in the

hippocampus. This compelling evidence led to the establishment of the company Cortexyme, Inc. and the submission of an investigational new drug application by the company.⁶³ Unfortunately, **COR388**, the gingipain inhibitor introduced by Cortexyme, Inc. exhibited dose-dependent hepatotoxic properties which led to a full clinical hold by the FDA in January 2022. Interestingly, although **COR388** failed to meet the main treatment goals of the clinical trial, the company reported promising results in a subpopulation of patients with *P. gingivalis* detected by oral swab.⁶⁴

This chapter introduces *P. gingivalis* and its role in the pathogenesis of various diseases. As discussed above, a lot of scientific effort has been invested to investigate the mechanisms of *P. gingivalis*-related pathogenesis, yet many of them remain unclear. Nevertheless, it is indisputable that gingipains play a key role in many of the pathogenic processes and remain a potential diagnostic and therapeutic target. In the next chapter, gingipains and their inhibitors will be introduced in closer detail.

1.3.2 Major *Porphyromonas gingivalis* Proteases – Gingipains and Their Inhibitors

Gingipains, cysteine proteases secreted by *P. gingivalis*, play a pivotal role in bacterial survival, virulence and pathogenicity. *P. gingivalis* is an asaccharolytic bacteria which exploits the cleavage of host proteins to produce amino acids and energy required for metabolic processes. Iron is also acquired through the degradation of heme produced by the host and is crucial for the growth and virulence of the bacteria.⁶⁵ While many factors, such as hemagglutinins, fimbriae and adhesins, mediate invasion of host cells, the lipopolysaccharide produced by *P. gingivalis* and gingipains are particularly interesting, as they exhibit immunomodulatory properties, inducing inflammation and contributing to immune evasion.¹³ According to Potempa et al., 85% of the proteolytic activity of the bacterium is inflicted by the gingipains.⁶⁶ Among various other interesting functions, gingipains have also been shown to influence regulation of kallikrein-like proteinase family members, which are involved in many important processes and implicated in cancer development.⁶⁷

Two unique proteases belonging to the gingipain class are the lysine-specific gingipain (Kgp) and the arginine-specific gingipain (Rgp), and their structures can be seen in **Figure 1.7**. The nomenclature originates from substrate specificity of the proteases.¹⁴ Originally, more than 20 protease genes from various strains of *P. gingivalis* were deposited in gene databases, but many of these genes overlap and they can be simplified

to *kgp*, *rgpA* and *rgpB*. Notably, Kgp is encoded by a single gene *kgp*, while Rgp sequence can be found in two individual genes, *rgpA* and *rgpB*. While *rgpB* encodes for the catalytic part along with a propeptide, *rgpA* also contains associated haemagglutinin domains. The *kgp* gene also contains C-terminal haemagglutinin domains, which are sequence-related to *rgpA* domains. Proteases arising from these two genes undergo complex processing pathways leading to multiple protease isoforms.

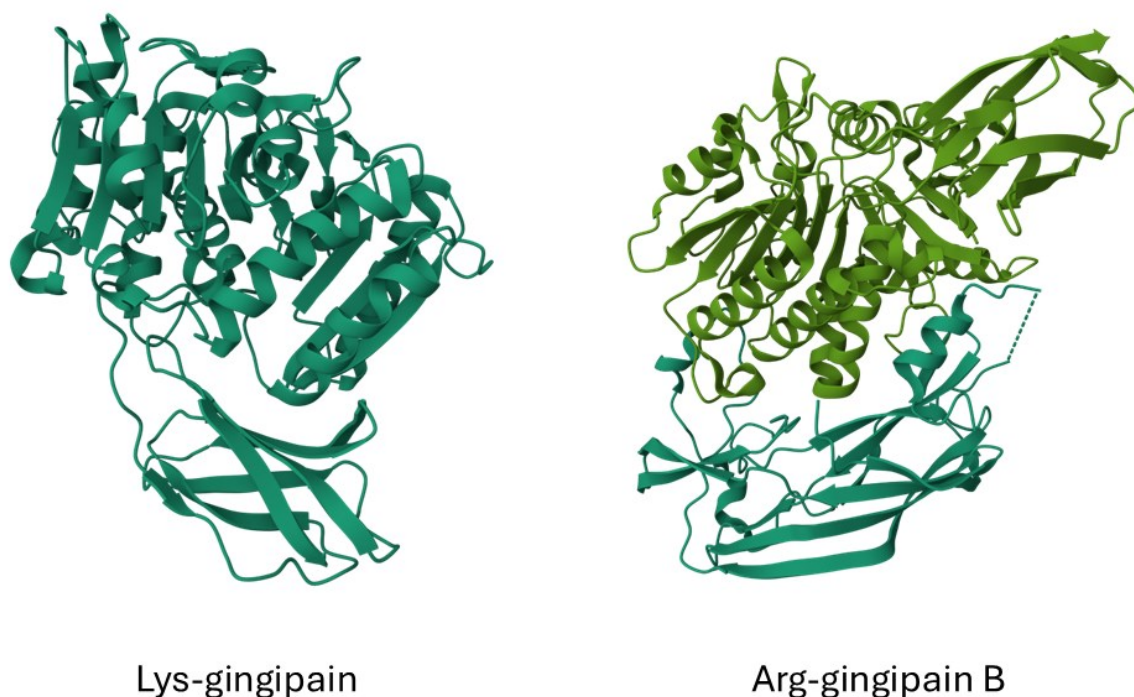
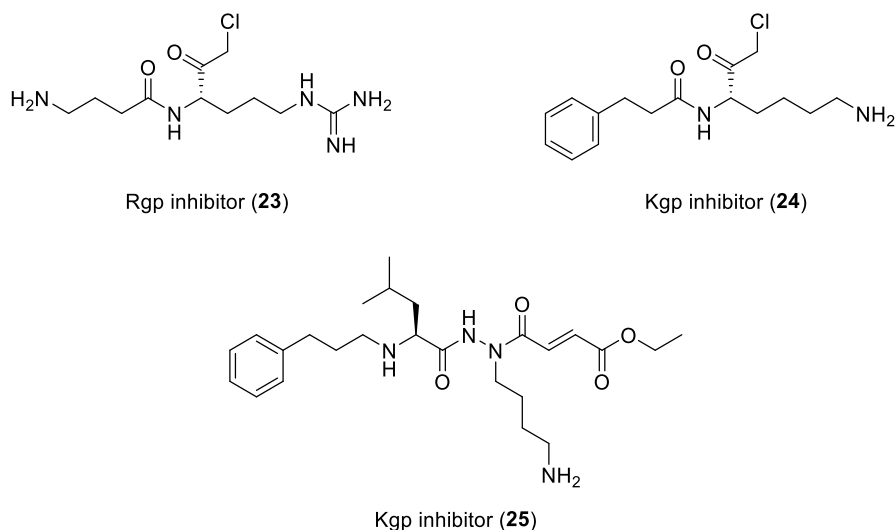


Figure 1.7 Crystal structure of Lys-gingipain (Kgp) catalytic and immunoglobulin superfamily-like domains (PDB: 4RBM, left), and Arg-gingipain B (RgpB) propeptide (teal) and catalytic (olive) domain (PDB: 4IEF, right)⁶⁸

As described in the last two chapters, the role of gingipains in *P. gingivalis* pathogenesis is undeniable. Therefore, platforms for simple and selective detection/quantification of gingipains, as well as high-throughput screening of chemical libraries for the identification of new inhibitors, are highly desired. In our lab, we focus on developing such assays by utilizing high-affinity small-molecule inhibitors combined with a DNA tag, which enables a sensitive and efficient readout.^{69,70} The first step in this process involves identifying a suitable inhibitor and modifying it with a linker to create a chemical probe.

Several classes of gingipain inhibitors can be found in the literature. Potempa et al. have synthesized and tested a wide array of peptidyl chloromethylketones and (acyloxy)-methylketones for the inhibition of gingipains.⁶⁶ As expected, inhibitors

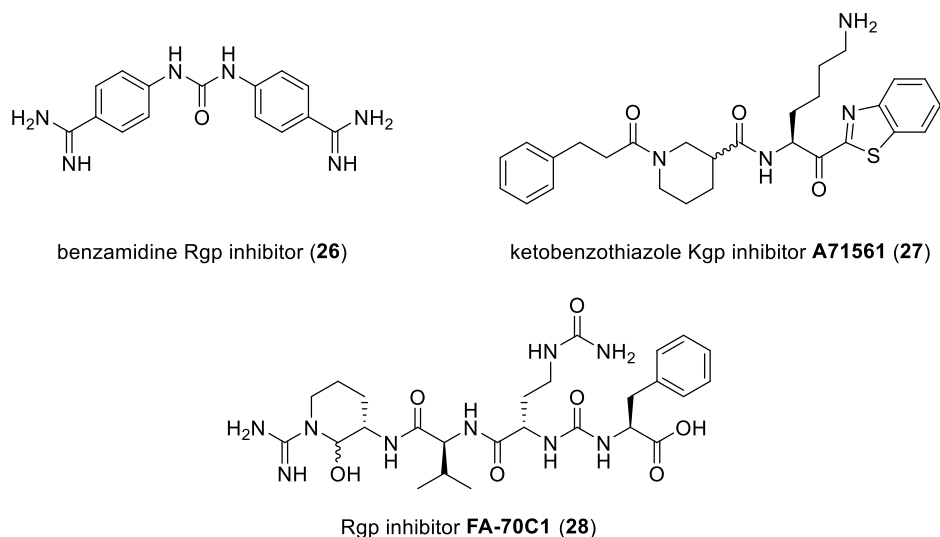
possessing Lys in P1 were preferred for Kgp, while inhibitors with Arg residue in P1 were highly selective for Rgps. Almost 10 years later, the same group exploited a crystal structure of Rgp with covalently bound peptidyl chloromethylketone inhibitor in the active site⁷¹ for the rational design of inhibitors with higher affinity and selectivity.⁷² Structures of selected optimized inhibitors **23** and **24** can be seen in **Scheme 1.11**. Interestingly, while crystal structure of Rgp was available since 1999, Kgp crystal structure was not reported until 2014. Another interesting class of irreversible gingipain inhibitors are aza-peptide Michael acceptors developed by Ekici et al.⁷³ The structure of the lead Kgp inhibitor **25** can also be seen in **Scheme 1.11**. Nevertheless, our previous efforts in assay development were based on competitive reversible inhibitors rather than irreversible covalent inhibitors, making the inhibitors discussed above not optimal for our intended applications.



Scheme 1.11 Selected irreversible covalent inhibitors of Kgp and Rgp^{72,73}

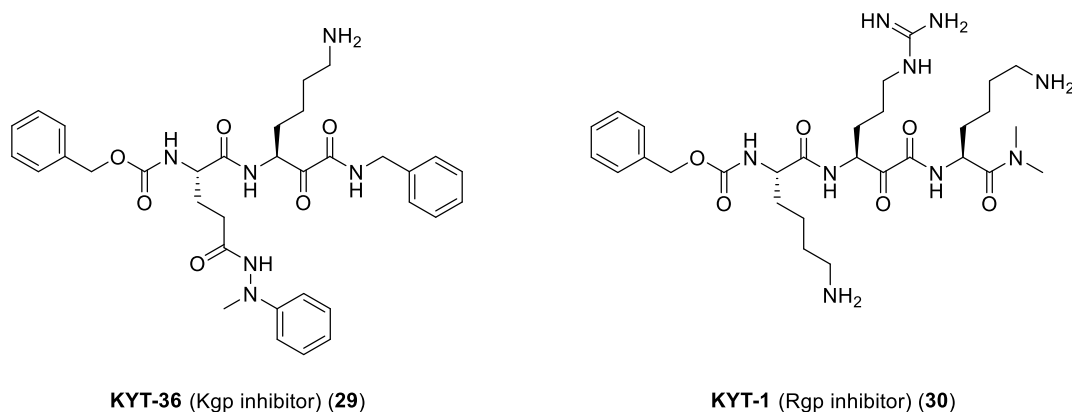
Several examples of reversible gingipain inhibitors can be seen in **Scheme 1.12**. First, Krauser et al. reported a series of benzamidine inhibitors, such as compound **26**, for inhibition of RgpB. Unfortunately, even though this structural motif is interesting, inhibitors in the series exhibited micromolar-range inhibition. Another interesting inhibitor found in the literature is **A71561 (27)**, a ketobenzothiazole inhibitor with a similar mechanism of action as the ketoamide inhibitors.⁷⁴ This compound exhibits very high sub-nanomolar affinity towards Kgp ($K_i = 0.9$ nM) and high selectivity over comparable human enzymes. Unfortunately, no Rgp-inhibiting analogue has been reported. **FA-70C1 (28)** is a natural gingipain inhibitor, an antipain analog isolated from *Streptomyces*, with nanomolar affinity ($K_i = 4.5$ nM) towards Rgp.⁷⁵ This compound

contains an interesting arginine aldehyde motif commonly observed in natural protease inhibitors, such as antipain or leupeptin. However, it also inhibits human cathepsins with double-digit nanomolar affinity.



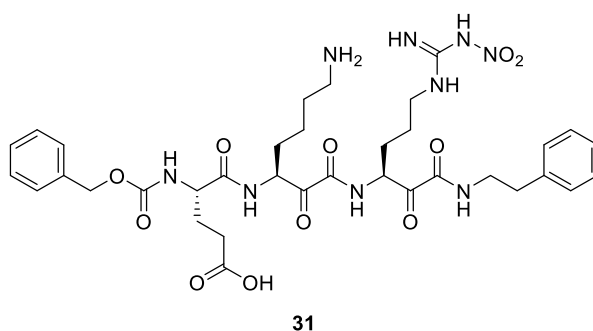
Scheme 1.12 Structures of various reversible gingipain inhibitors

Inhibitors **A71561** (**27**) and **FA-70C1** (**28**) would meet most requirements for linker modification and probe development, but little structural information is available about their binding to the respective targets. Therefore, we turned our attention to inhibitors developed in the group of Kenji Yamamoto.¹⁵ First, high-affinity inhibitors for both Kgp (**KYT-36**, **29**, $K_i(\text{Kgp}) = 75 \text{ pM}$) and Rgp (**KYT-1**, **30**, $K_i(\text{Rgp}) = 130 \text{ pM}$) were reported in their work. These inhibitors share high structural similarity, as shown in **Scheme 1.13**, so the design of the linker-modified analogues can be performed analogously. Furthermore, a crystal structure of Kgp with **KYT-36** in the active site has been reported in the literature.⁷⁶ Structural information provided by the crystal structure can be exploited in the design and increases the chance for inhibitor modification without loss of inhibitory potency.



Scheme 1.13 Structures of gingipain inhibitors developed in the group of K. Yamamoto¹⁵

While **KYT-1** and **KYT-36** are highly selective for Rgp and Kgp, respectively, the group of Kenji Yamamoto has also developed a high-affinity dual inhibitor for both gingipains, shown in **Scheme 1.14**. Interestingly, this dual inhibitor, **KYT-41** (**31**), was rationally designed as a hybrid of **KYT-1** and **KYT-36**, possessing two ketoamide moieties, each targeted toward one of the gingipains.⁷⁷



Scheme 1.14 Structure of **KYT-41**, dual Kgp and Rgp inhibitor

Nevertheless, our goal was the development of probes that would be able to distinguish between Kgp and Rgp. Therefore, we picked **KYT-36** and **KYT-1**, shown in **Scheme 1.13**, as the inhibitors of choice for the development. In the next chapter, the design of linker-modified analogues based on these inhibitors will be discussed.

1.3.3 *Design of Linker-Modified Gingipain Inhibitors for Probe Development*

Chemical probes are valuable tools widely used in chemical biology, biochemistry and many other related fields. The increasing number of new potential drug targets identified via genome sequencing and the widespread availability of new screening technologies have shown that the development of small-molecule tools – chemical probes – is essential to clarify the functions of targeted proteins in both healthy and diseased tissue and to confirm therapeutic potential of the protein targets.⁷⁸ Chemical probes are reagents based on small-molecule ligands or inhibitors that selectively target a protein of interest and enable the investigation of mechanistic and phenotypic questions in various biological systems. Although probes relying solely on inhibition/binding to the target protein are valuable, probes carrying a reporter as part of the molecule or a clickable handle offer investigation of more complex questions, such as target localization, interactions with neighboring proteins and, finally, screening for new inhibitors in chemical libraries.^{79,80} Typically, such probe development begins with a high-affinity ligand and the final probe is composed of the ligand modified with a linker which possesses an appropriate affinity tag (such as biotin), fluorophore or a reactive handle (such as azide or alkyne; suitable for a biocompatible click reaction), as depicted in **Figure 1.8a**. Identifying a suitable position for the linker attachment without disrupting the protein-ligand interaction is a crucial step in the process. In the case of a hypothetical interaction between a protein and a ligand, according to **Figure 1.8b**, only certain modifications of the ligand would lead to a functional chemical probe. As shown in **Figure 1.8c**, modification at position A or B would disrupt the protein-ligand interaction and inevitably reduce the binding affinity of the modified ligand. On the other hand, linker attachment at position C would preserve the affinity and lead to a functional chemical probe.

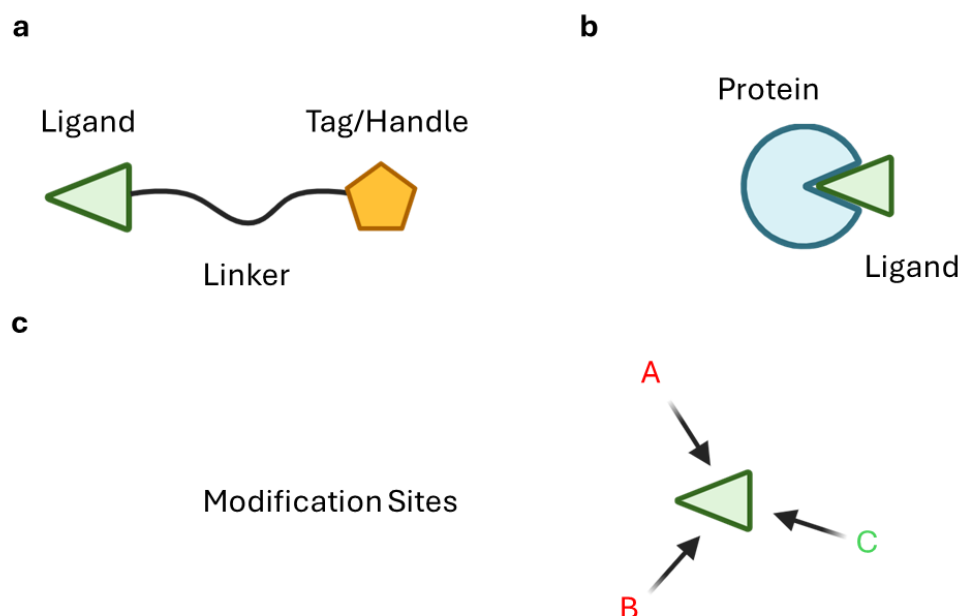
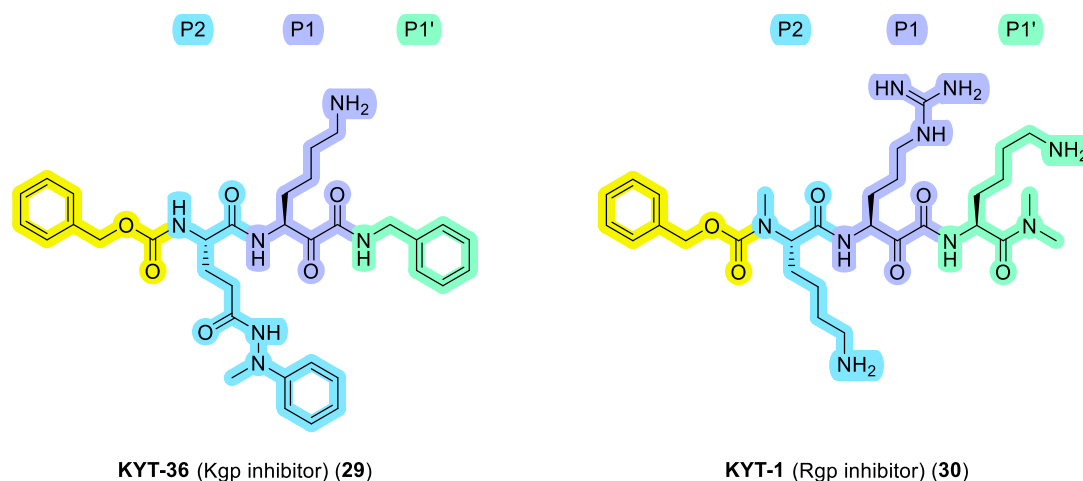


Figure 1.8 Design of a chemical probe: (a) General scheme of a chemical probe composed of a ligand (green arrow), a linker (black line) and a functional tag/handle (orange pentagon); (b) Hypothetical interaction between a ligand (green arrow) and a target protein (blue pacman); (c) Modification sites of a hypothetical ligand, modifications disrupting ligand affinity are shown in red letters, modifications preserving ligand affinity are shown in green letters. Created with Biorender.com

Based on these considerations, we proceeded with the design of linker-modified gingipain inhibitors. Generally, for probe development, any available structural information about binding of the ligand to the target protein or a potential structure-activity relationship (SAR) is valuable. Taking this into account, we decided to develop linker-modified gingipain inhibitors based on inhibitors **KYT-1** and **KYT-36**. The structures of **KYT-1** and **KYT-36** are synthetically available and highly modular. These inhibitors offer many sites where linker attachment is possible, allowing us to design and synthesize multiple derivatives if necessary. Furthermore, a crystal structure of Kgp with **KYT-36** in the active site is available in the literature and can provide detailed information about the binding mode of the inhibitor to its protein target. Both inhibitors can be considered as peptidomimetic substrate analogues possessing a ketoamide warhead, and their respective residues can be labelled as P2, P1 and P1' according to conventional protease substrate nomenclature, as depicted in **Scheme 1.15**.



Scheme 1.15 Structures of **KYT-36** and **KYT-1** gingipain inhibitors. Inhibitors are divided into respective amino acid-like residues (P2, P1, P1') according to the protease substrate nomenclature.

Looking at the structure of **KYT-36**, direct modification of the inhibitor would only be feasible at the P1 Lys-like residue via amide coupling. As discussed earlier, the names of Kgp (Lys-gingipain) and Rgp (Arg-gingipain) are based on the substrate specificity of these proteases and the amino acids preferred in the P1 position. Therefore, we can assume that the Lys-like residue in **KYT-36** plays a key role in its binding to Kgp, as well as the Arg-like residue in **KYT-1** binding to Rgp, and modification at these positions would disrupt the binding affinity severely. The three benzene rings present in the structure of **KYT-36** are not available for direct modification of the final inhibitor, but they come from simple fragments assembled during the synthesis of **KYT-36**. Thus, introduction of a handle for linker attachment would be possible at any of these positions. As mentioned above, the crystal structure of Kgp with **KYT-36** in the active site (PDB: 6I9A) provided us with extensive structural information about the binding mode of **KYT-36** (**Figure 1.9**).⁷⁶ The crystal structure confirms reversible covalent binding of **KYT-36** via the ketoamide warhead to the catalytic cysteine residue (Cys-477). The ketone group responsible for covalent inhibition is pyramidalized and is only 2.02 Å away from the catalytic cysteine residue. As expected, the P1 Lys-like residue reaches deep into the binding pocket responsible for the substrate specificity of Kgp and is subject to many important interactions. On the other hand, all benzene rings are located on the surface of the protein and partially solvent-exposed, as seen in **Figure 1.9c-d**. The Cbz and benzylamine fragments are tracing the surface of Kgp, and the authors of the crystal structure propose multiple hydrophobic interactions between these fragments and hydrophobic amino acid residues on the surface of the protein. In contrast, the hydrazide fragment is directed away from the surface of the protein towards the bulk solvent and only interacts with one hydrophobic residue on the protein surface (**Figure 1.9d**). Also,

the fragment can be found in two different conformations in the crystal structure. This added flexibility, further emphasized by the incomplete electron density map of the fragment – compared to a well-defined map for the rest of the inhibitor, as shown in **Figure 1.9b** – suggests no strong interactions are taking place at this fragment. Therefore, we expected that the fragment, even after linker attachment, could exploit its flexibility to achieve stable enzyme-bound conformation, thereby providing high binding affinity towards Kgp.

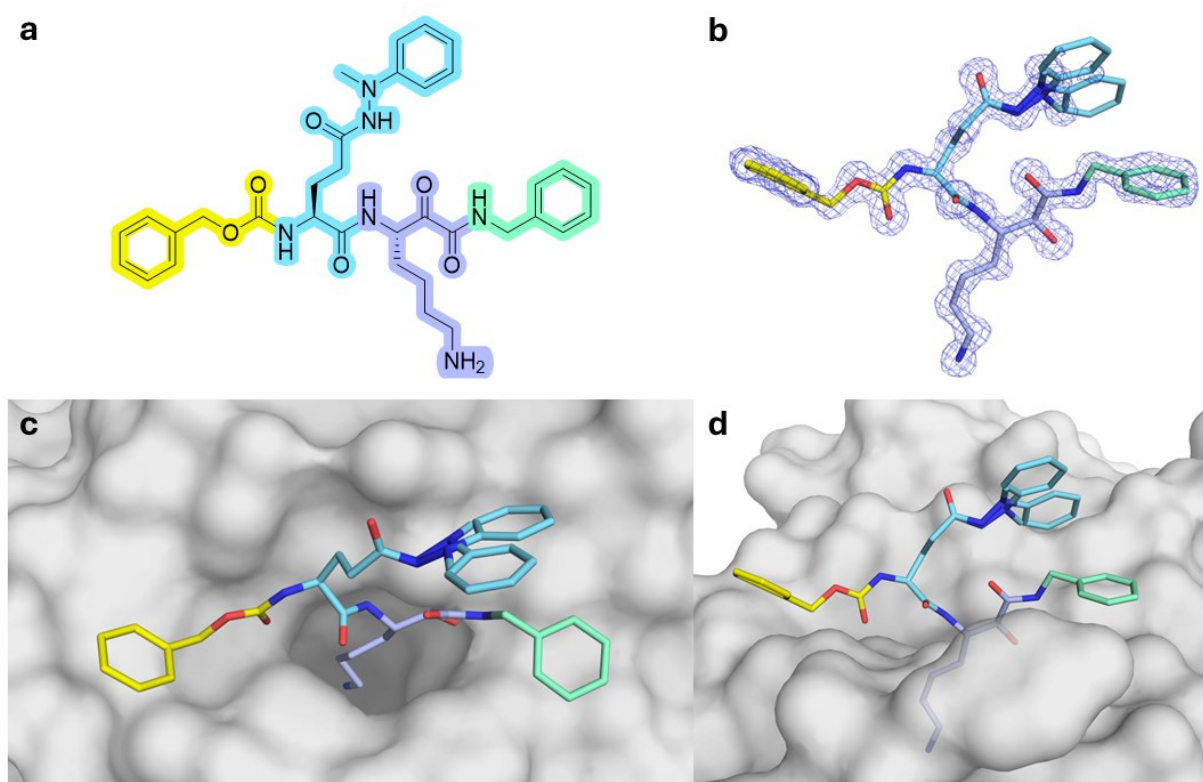


Figure 1.9 Crystal structure of Kgp with inhibitor **KYT-36** (PDB:6I9A): (a) Structure of **KYT-36** (**29**); (b) Electron density map of **KYT-36**, contoured to 1.5 σ ; (c) **KYT-36** in the active site (top view); (d) **KYT-36** in the active site (side view).⁷⁶

In the case of **KYT-1**, the situation is notably different. No structural information is available about the binding of **KYT-1** to Rgp making it difficult to fully rationalize the design of linker-modified analogues. Nevertheless, **KYT-1** contains two Lys residues along with one Arg residue which can be directly modified via amide coupling. Based on the analogy with **KYT-36**, we can assume that the Arg residue is important for the binding of **KYT-1** to Rgp and modification at this position is not desired. The reactivity of the guanidine moiety in Arg residue is significantly different from the amine groups present in Lys residues, and this difference allows for selective modification of the Lys residues

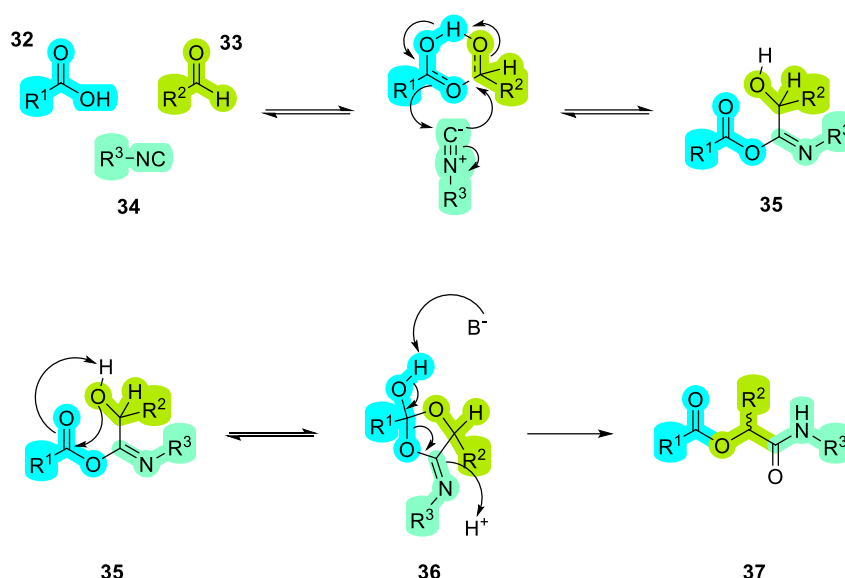
without affecting the Arg residue in most cases. Therefore, we proposed that using optimized conditions, we should be able to modify **KYT-1** at the Lys residues (P2, P1') and generate a mixture of two linker-modified **KYT-1** analogues. We also hypothesized that at least partial separation of the two isomers could be achieved. This would allow us to identify individual isomers and test them in an enzymatic assay for Rgp inhibition. Several outcomes are possible: both isomers could be active, only one could be active, or both could be inactive. We believed that there was a high probability that at least one of the isomers would be active, but even in case of both isomers being inactive, this experiment would cover two out of three available modification sites. In this case, modification at the Cbz moiety would be considered as the next prospective modification site. Based on these considerations, modification of **KYT-1** Lys residues has been selected as the most straightforward and efficient approach to identify suitable positions for linker attachment and the following probe development.

1.3.4 Passerini Reaction in Preparation of Gingipain Inhibitors

To proceed with the preparation of linker-modified gingipain inhibitors, the design of a suitable synthetic approach was required. Interestingly, even though **KYT-1** and **KYT-36** are widely used, and their biological activity has been studied extensively, very little can be found about the synthesis of these compounds. A procedure for their preparation can only be found in a single patent.⁸¹ This procedure is based on oxidation of α -amino acids to their corresponding β -amino- α -keto acids, serving as precursors of the α -ketoamide warhead, followed by step-by-step peptide coupling to yield the final inhibitors. Nevertheless, we decided to develop our own synthetic strategy based on a Passerini reaction, which seemed more elegant, convenient and modular.

The Passerini reaction is a well-known multicomponent reaction between an isonitrile, an aldehyde (or ketone) and a carboxylic acid. This reaction is commonly used in combinatorial chemistry as it allows for facile preparation of diverse advanced intermediates in a single step. Multiple mechanisms have been proposed for the course of the Passerini reaction, and the widely accepted concerted mechanism – prevalent in apolar solvents with high concentration of the reactants – is depicted below in **Scheme 1.16**.^{82–84} This mechanism suggests the addition of the carbonyl component **33** (acts as an electrophile) and the carboxylic acid **32** (as a nucleophile) to the isonitrile species **34**. This

leads to formation of an imidate intermediate **35**, which then undergoes Mumm's rearrangement to form the α -acyloxyamide **37**.

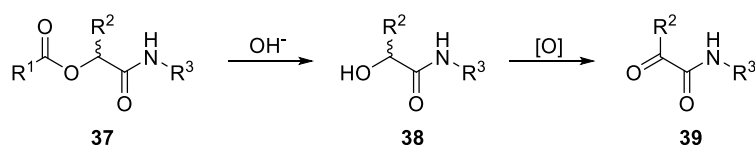


Scheme 1.16 Mechanism of the Passerini Reaction

Generally, mechanisms of multicomponent reactions are hard to elucidate, and small adjustments are still being introduced. Maeda and co-workers proposed that Passerini reaction is a four-component reaction.⁸⁵ They suggested that an extra molecule of the carboxylic acid **32** mediates the proton transfer in conversion of imidate **35** to intermediate **36** and in the decomposition of this intermediate to the final product **37**. This subtle change in the reaction mechanism makes the Passerini reaction an organocatalytic reaction. Another iteration has been introduced in 2015 by Ramozzi and Morokuma,⁸⁶ that proposed the existence of an explicit nitrilium ion before formation of imidate **35**. The existence of the nitrilium ion has been previously denied because of its low stability. Even though this study did not provide experimental evidence of the nitrilium ion's existence, it did provide a rationale for its existence based on Gibbs free energy profile calculations.

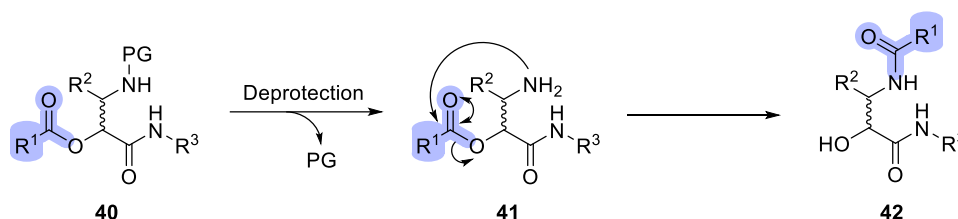
Various post-Passerini modifications are known, as the α -acyloxyamide **37** motif offers several reactivity pathways.^{82,83} Cyclization reactions to form various heterocycles, such as oxazoles, furans and many others, are common and generally require the presence of additional functional groups in one of the reaction components. From the medicinal chemist's point of view, the α -acyloxy group is not a preferred motif due to its metabolic instability. The easiest way to address this issue is deacylation via hydrolysis to release an α -hydroxyketone. Another potential challenge of the Passerini reaction in the context of inhibitor design is the lack of stereoselectivity. The attack of the isonitrile (**34**) on the

aldehyde (**33**), as depicted in **Scheme 1.16**, defines the stereochemistry of the product and generally takes place from both sides. Nevertheless, α -acyloxyamide **37** is a convenient template for preparation of α -ketoamide inhibitors that have proven to be a valuable tool for inhibition of relevant protease targets, such as fibroblast activation protein (FAP),⁸⁷ rhomboid proteases⁸⁸ or coronavirus main protease (Mpro).^{89,90} After hydrolysis of the ester bond, the α -hydroxyamide **38** can be easily oxidized to ketone **39**, as shown in **Scheme 1.17**.



Scheme 1.17 Conversion of the α -acyloxyamide **37** to α -ketoamide **39**

When protected α -amino aldehyde is used, the deprotection of the amino functionality of **40** triggers spontaneous acyl migration converting the O-acyl **41** to a more thermodynamically stable N-acyl **42**. This reaction sequence is referred to as Passerini Reaction – Amine Deprotection – AcyL Migration (PADAM). General mechanism of this post-Passerini modification is depicted below in **Scheme 1.18**.



Scheme 1.18 Mechanism of the post-Passerini Amine Deprotection-Acyl Migration

In the following step, α -hydroxyamide **42** can be oxidized to yield an α -ketoamide inhibitor as shown in **Scheme 1.17**. Therefore, we decided to employ this strategy for preparation of both gingipain inhibitors, **KYT-1** and **KYT-36**.

1.4 Polymer-Tethered Quenched Fluorescent Probes for Imaging of Tumor-Associated Proteases

1.4.1 Fluorescent Probes in Cancer Imaging

Surgical resection remains a major treatment strategy for the treatment of solid tumors. Typically, the surgery is performed under bright-field imaging and the surgeon is striving to delineate borders of the tumor precisely using the limited information provided by the simple white light image. After the surgery, the resected tissue is sent for histological analysis to visualize real tumor margins. Two potential outcomes, as illustrated in **Figure 1.10**, may occur. In instances where a tumor is successfully resected, a negative surgical margin is observed. This indicates that the margin is free from tumor cells and that the entire tumor has been extracted. When tumor margins were not accurately delineated, a positive surgical margin is observed, indicating the presence of cancerous cells at the edges of the resected tissue. In most cases, this means that a portion of the cancerous cells have been left behind, which can result in tumor regrowth and the necessity for repeated surgical procedures. It is standard practice for surgeons to aim to maintain larger margins to reduce the risk of relapse. However, this can prove challenging in the context of tissues that are of functional importance (such as vital organs) or when aesthetic considerations are a factor.

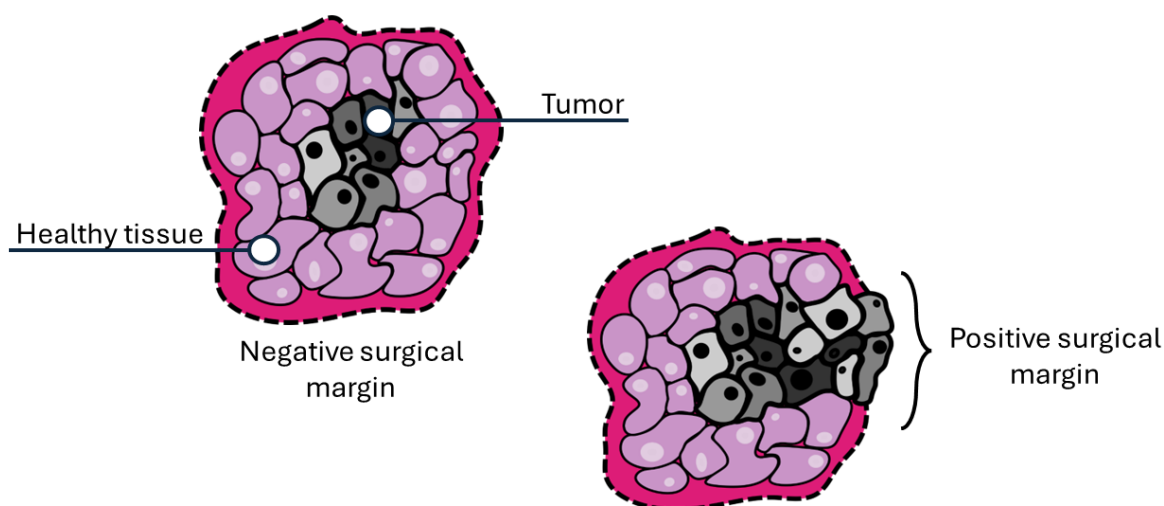
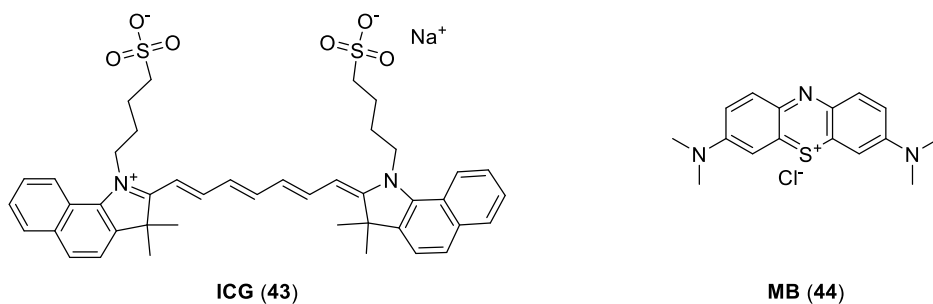


Figure 1.10 Histological analysis of a resected sample – negative and positive surgical margin

Achieving accurate tumor resection remains challenging due to difficulties in detecting tumor boundaries in real-time with high precision. Fluorescent probes have

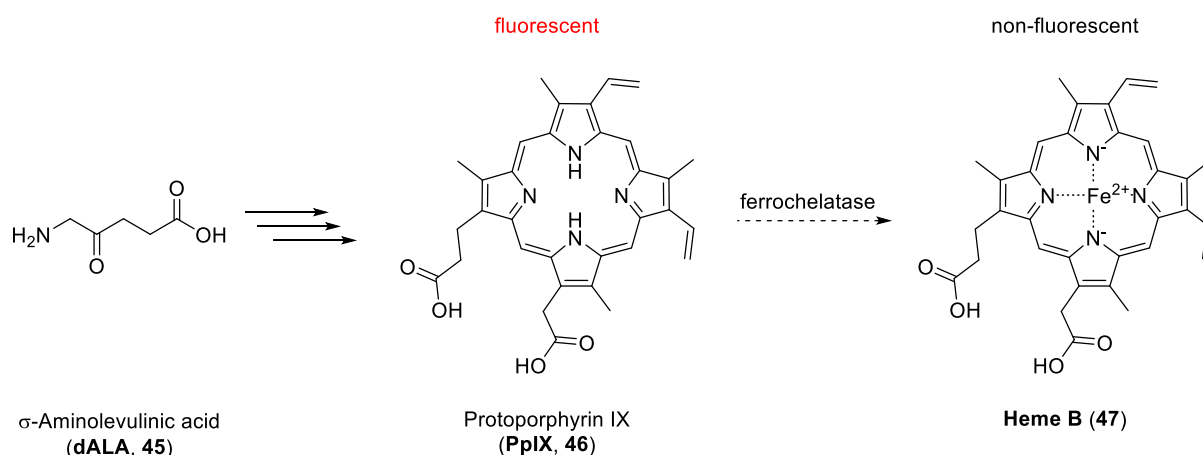
emerged as powerful tools for cancer imaging, providing strong signal and contrast, generally improving resection outcomes.^{91–95} In one of many available studies, Hoogstins et al. describe the use of an affinity-based fluorescent agent, **OTL38**, which enabled the detection of an additional 30% of malignant lesions in patients with epithelial ovarian cancer compared to conventional techniques.⁹⁶ The integration of fluorescence imaging modules into common robotic-assisted surgical devices, such as the da Vinci Surgical System⁹⁷, and the development of numerous surgical camera systems^{98,99}, also highlights the potential of fluorescence-guided surgery. These probes can not only be used for the *in vivo* surgical guidance, but they can also be exploited for the diagnostics and the pre-operative planning, as well as the *ex vivo* analysis of the extracted samples.

Early fluorescent agents approved by the U.S. Food and Drug Administration (FDA) more than 50 years ago, such as indocyanine green (**ICG**, **43**)^{95,100} and methylene blue (**MB**, **44**)^{95,101} (**Scheme 1.19**), visualize solid tumors due to distinct tumor vasculature and blood flow, resulting in effects such as the enhanced permeability and retention (EPR) effect commonly observed in malignant tissues.^{102,103} These agents provide enhanced contrast, however, their performance is limited by lack of specific targeting mechanisms. Both **ICG** and **MB** are near-infrared (NIR) fluorophores with emission in the NIR-I region (700 to 1000 nm). This part of the spectrum is commonly used in imaging of biological samples for multiple reasons.⁹⁵ First, mammalian tissues are generally opaque to visible light (400 to 700 nm), so photon penetration depth in this area is low. Second, absorption of photons in this area is coupled with autofluorescence of biological tissues which creates additional interference. Lastly, light scattering processes turn signal photons into stray light adding to the background and, therefore, influence resolution significantly. It has been shown that the degree of scattering is inversely proportional to the wavelength for most biological samples, which also explains prioritization of longer wavelength fluorophores.



Scheme 1.19 Structures of early NIR fluorophores approved by FDA

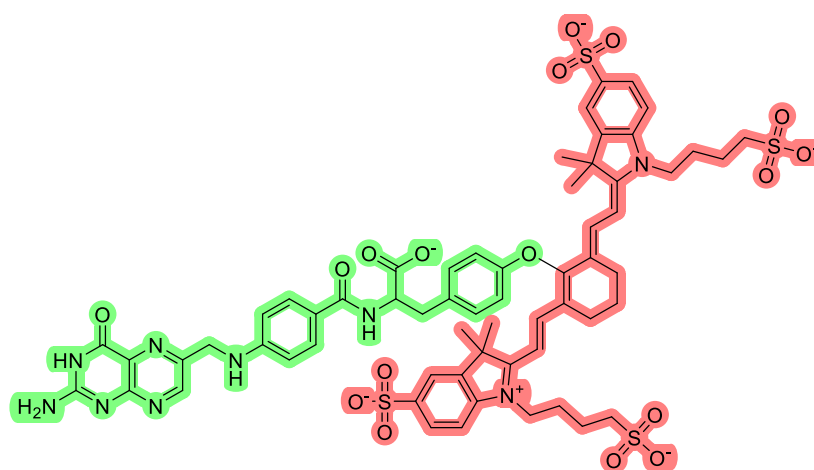
An interesting compound in the field of fluorescence-guided surgery is δ -aminolevulinic acid (**dALA**, **45**, **Scheme 1.20**). Notably, this acid does not possess any aromatic or conjugated π systems and cannot be inherently fluorescent. Nevertheless, it has been approved by the FDA for use in fluorescence-guided surgery and can generate fluorescence in tumors through a unique mechanism. As shown in **Scheme 1.20**, **dALA** is a natural metabolite involved in the hemoglobin metabolic pathway. Exogenous **dALA** acts as a so-called “pro-agent”, a precursor of the fluorescent protoporphyrin IX (**PpIX**, **46**) which is detected after metabolic processing.¹⁰⁴ In healthy tissue, **PpIX** is further converted into non-fluorescent **heme B** (**47**) by ferrochelatase but malignant cells often exhibit reduced ferrochelatase activity which leads to accumulation of **PpIX** and increased tissue fluorescence. Even though ferrochelatase activity is considered to be the major factor responsible for the accumulation of **PpIX** in malignant tissues, other factors, such as tumor cells proliferative activity, selective uptake of **dALA** by an ATP-binding cassette transporter and its high blood-brain barrier (BBB) permeability also play a role in its effectiveness.¹⁰⁴



Scheme 1.20 Hemoglobin metabolic pathway involving **dALA**

The FDA approved **dALA** as an optical imaging agent for patients with gliomas in 2017. The unprecedented BBB permeability of **dALA** makes it an ideal agent for the visualization of gliomas as options for brain delivery of more complex compounds, such as fluorophores, are very limited. Phase III clinical trial reported convincing results with complete tumor resection in 65% of patients after **dALA** tumor imaging compared with 36% of patients when bright-field imaging was applied.¹⁰⁵ Furthermore, 41% of patients in the **dALA** group had 6-month progression free survival compared with 21% in the control group.

Despite the long clinical use of untargeted fluorescent probes, such as **ICG** or **MB**, targeted probes are only recently reaching late-stage clinical trials. To our knowledge, **OTL38** (**48**, also known as Pafolacianine, sold under the brand name Cytalux[®]) is the first targeted fluorescent contrast agent approved by the FDA (approved in November 2021).^{96,106} **OTL38** (**Scheme 1.21**) is an affinity-based contrast agent with high affinity towards folate receptor α (FR α), which is commonly overexpressed in various solid tumors. It is a folic acid derivative equipped with a NIR fluorophore **S0456**, an analog of a common dye **IRDye800CW** with additional sulfonate groups for improved water solubility. **OTL38** is mainly used for intraoperative imaging of FR α positive ovarian cancer, where sensitivity of 85% has been demonstrated.⁹³ It has also been successful in the imaging of endometrial tumors, pulmonary adenocarcinomas and pituitary adenomas.



OTL38 (48)

Scheme 1.21 Structure of **OTL38** (Pafolacianine). The folic acid derivative is highlighted in green, and the fluorophore (**S0456**) is highlighted in red.

Another interesting class of contrast agents are quenched fluorescent probes, also known as activity-based probes. These probes typically consist of a fluorophore, a quencher and a cleavable linker, as depicted in **Figure 1.11**. The process of quenching is enabled by the Förster resonance energy transfer (FRET), which leads to non-radiative energy transfer between a donor (an excited fluorophore) and an acceptor (a quencher). When quencher and fluorophore are in proximity, FRET occurs after fluorophore excitation and no (or very limited) fluorescence can be detected. FRET is highly distance dependent and, therefore, the quencher-fluorophore distance plays a key role in quenching efficiency. Following the enzymatic cleavage of the cleavable linker in the target tissue, typically by proteases, the covalent bond between the quencher and the fluorophore is cleaved and the

quencher-fluorophore distance increases (**Figure 1.11**). This results in a reduction in FRET efficiency and an increase in fluorescence upon excitation.

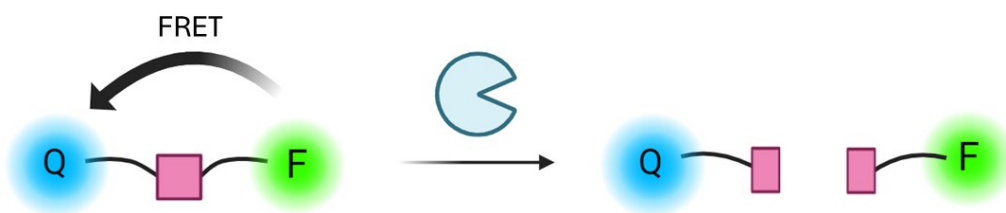
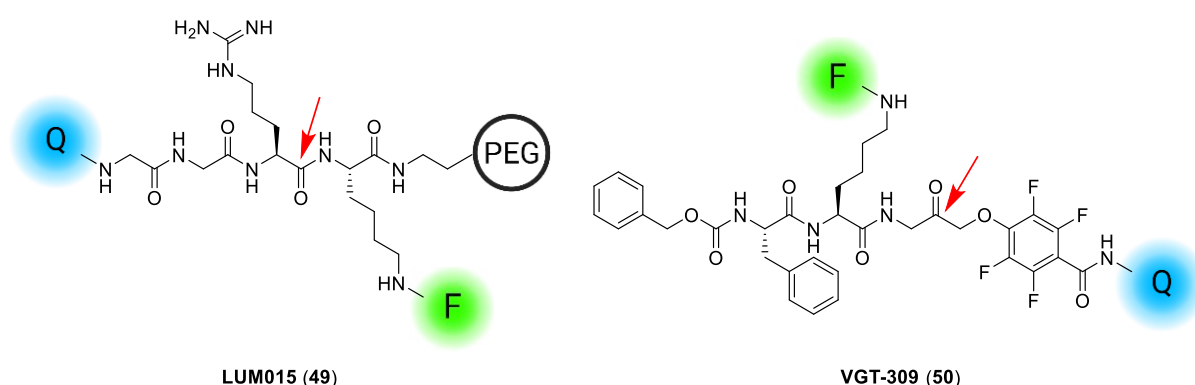


Figure 1.11 General scheme of a quenched fluorescent probe and its activation by a protease. Q (blue sphere) represents a quencher, F (green sphere) represents a fluorophore, the pink box represents a protease-cleavable linker, and the light blue pacman represents a protease. Created with Biorender.com

Affinity-based probes have already reached the market, with **OTL38** being the first successful example. Several examples of activity-based probes reaching late-stage clinical trials, such as **LUM015 (49)** or **VGT-309 (50)** (**Scheme 1.22**), can be found.^{107,108} Both **LUM015**,^{107,109} a PEGylated peptidic substrate, and **VGT-309**^{108,110}, an irreversible covalent probe, are targeting cathepsin protease family. **VGT-309** acts via a mechanism analogous to inhibition of proteases by α -chloroketones. Upon protease attack on the ketone moiety, the tetrafluorophenoxy group acts as a leaving group and is replaced by the protease catalytic residue. After the reaction, the quencher moiety diffuses away while the fluorophore-containing part stays covalently attached to the protease.



Scheme 1.22 Structures of **LUM015 (49)** and **VGT-309 (50)**. Q (blue sphere) represents a quencher, F (green sphere) represents a fluorophore, and the red arrow indicates the protease attack site.

Another activity-based probe under clinical investigation is **AVB-620**, targeting matrix metalloproteinases (MMP).¹¹¹ **AVB-620** possesses multiple interesting design traits, based on a combination of pioneering work of the Weissleder lab,¹¹² and the activatable cell-penetrating peptides (CPP) developed in the Tsien lab.¹¹³ First, **AVB-620** is a approx. dodecamer peptide possessing two arms – poly-D-arginine octamer CPP arm

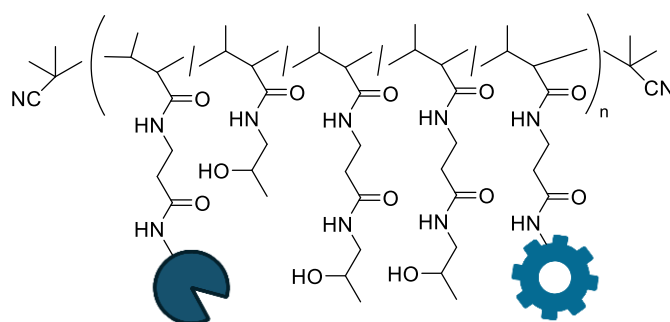
and a poly-D-glutamate pentamer masking arm – connected via sequence cleavable by MMPs commonly overexpressed in tumor tissues. This peptide contains two different dyes at the ends of the peptide chain. In the native state, the peptide adopts a U-shaped conformation due to electrostatic interactions between both arms which hold the two dyes in proximity to enable efficient FRET. Second, the probe does not possess a typical quencher-fluorophore pair, but it contains two individual fluorophores with spectral overlap. This allows for a so-called ratiometric readout where the fluorescence of both fluorophores, which is affected by efficiency of the FRET, is measured simultaneously. After successful cleavage of the MMP substrate sequence, the fluorophores are separated and significant changes in both channels can be observed. This approach is very efficient in reducing false positives caused by probe accumulation in unwanted areas and, therefore, increases contrast.^{111,114} Furthermore, the CPP is unmasked after substrate cleavage and leads to internalization of one of the fluorophores for enhanced signal retention. Finally, **AVB-620** also contains a PEG tail for improved water solubility and hydrophilicity.

As discussed in this chapter, fluorescent probes have proven their potential in the clinical setting but further optimization of their structure, contrast and biodistribution is desired. Two of the discussed probes, **LUM015** and **AVB-620**, exploit PEGylation as a tool to increase molecular weight and hydrophilicity in order to improve solubility, biodistribution and to exploit the EPR effect present in malignant tissues. In our work, we hypothesized that combination of a biocompatible, inert and hydrophilic macromolecular scaffold with an activity-based fluorescent substrate could lead to improved properties of the small-molecule counterparts.

1.4.2 *Design of Polymer-Tethered Quenched Fluorescent Probes for Imaging of Cathepsins*

In 2016, an antibody mimetics platform called iBodies has been developed in the group of Jan Konvalinka at the Institute of Organic Chemistry and Biochemistry in collaboration with the Institute of Macromolecular Chemistry.¹¹⁵ This platform is based on a *N*-Hydroxypropylmethacrylamide copolymer (pHPMA), which is water-soluble, biocompatible, and has already been used in the development of drug delivery systems and imaging agents.^{116,117} Typically, polymeric precursors are synthesized with amine-reactive thiazolidine-2-thione groups in the side chains. In the next step, the side chains

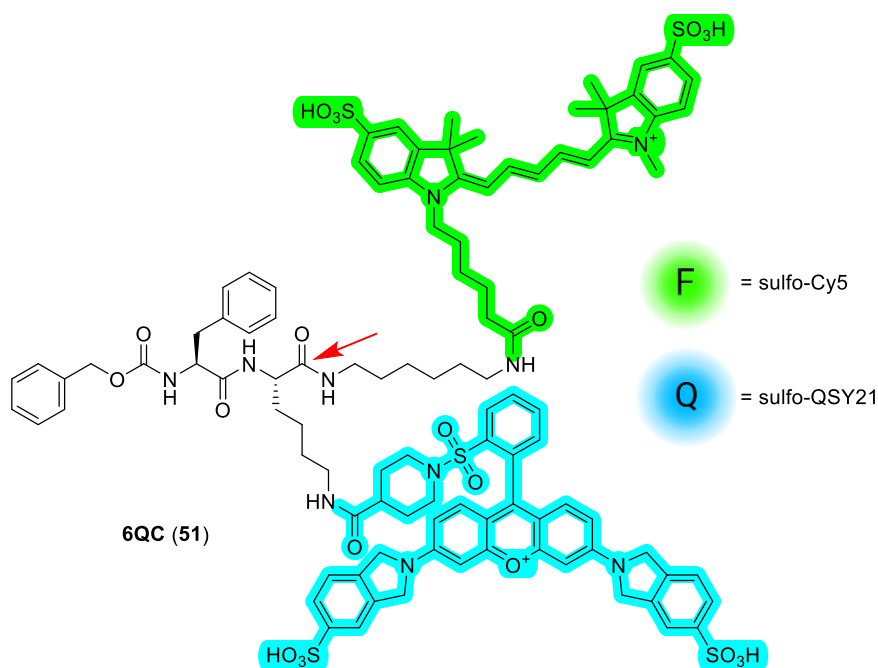
are modified by appropriate functional moieties to form a functional conjugate with the general structure shown in **Scheme 1.23**. Multiple copies of each moiety are usually present on a single iBody molecule, and their abundance and ratio can be controlled during the synthesis. This can be used for fine-tuning of properties of the final construct. In the original design, iBodies are modified with three different moieties – a high-affinity ligand or inhibitor, a fluorescent dye and an affinity anchor. The ligand (or inhibitor) is responsible for targeting the macromolecular conjugate, the fluorophore enables visualization, and the affinity tag allows for immobilization of the conjugate and its protein target. Such functional conjugates can be used in biochemical experiments for enzyme inhibition, staining, quantification and many other applications.^{118–122} iBodies for molecular recognition of His-tagged proteins have also been developed.¹²³



Scheme 1.23 General structure of a pHPMA functional conjugate. The functional moieties are represented by the gear and pacman icons. Created with Biorender.com

Our group has been successful in developing functional high-affinity conjugates targeting glutamate carboxypeptidase II (GCP II), HIV-1 protease, aspartic proteases, carbonic anhydrases and fibroblast activation protein (FAP). In our last work, we have also demonstrated that iBodies could be applied for inhibition of protein-protein interactions and potentially replace antibodies in cancer immunotherapy.¹²⁴ Interestingly, in certain cases, iBodies exhibit a strong avidity effect caused by the presence of multiple copies of the targeting ligand on a single iBody molecule. Due to this effect, even low-affinity ligands tethered to iBodies can lead to high-affinity conjugates. In the case of GCP II, the low-molecular inhibitor exhibited K_i of 2.0 nM compared to 4.3 pM of the GCP II-targeted iBody. This drop in K_i reaches far beyond the multiplicity of ligands on the macromolecular conjugate (average of 19 inhibitors per polymer molecule). Even though iBodies were originally designed as antibody mimetics, in recent years we also focused on expanding their use for applications beyond this scope.

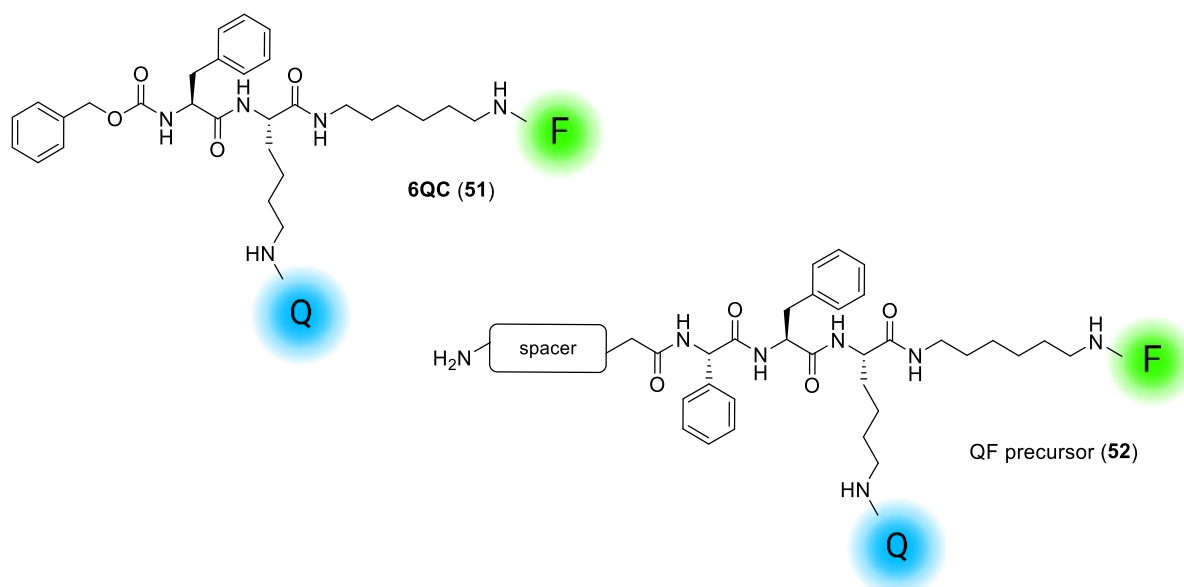
In this study, we proposed that the pHMPA backbone could provide a promising template for enhancing the contrast and biodistribution of low-molecular-weight activity-based probes. We decided to tether a quenched fluorescent substrate to the pHMPA backbone to examine the potential of such probes in reference to their small-molecular counterparts. These probes could offer enhanced imaging properties due to the EPR effect, improved solubility and hydrophilicity, and tunable signal intensity via control of substrate content on the pHMPA scaffold. Furthermore, additional affinity-based targeting elements could be added to drive the conjugates into the target tissue. This study serves as a proof of concept, and in the event of success, the concept could be readily applied to various activity-based probes. The probe selected as the starting point for this study is a previously published quenched pan-cathepsin substrate **6QC** (**51**, **Scheme 1.24**). It was originally developed in the group of Prof. Bogoy at Stanford University, with whom we collaborated on this project.¹²⁵ **6QC** contains sulfo-Cy5 as the fluorophore and sulfo-QSY21 as the corresponding quencher. As discussed earlier in this chapter, probes targeting cathepsins are valuable tools for cancer imaging^{109,110,126–128} and **6QC** has been thoroughly studied and optimized in this context.^{125,129,130}



Scheme 1.24 Structure of **6QC**. The fluorophore (sulfo-Cy5) is highlighted in green, the quencher (sulfo-QSY21) is highlighted in light blue, and the red arrow indicates the cathepsin cleavage site.

The very first step of the process was the design of **6QC** analogues that would possess a suitable handle for attachment to the polymeric carrier and retain high efficiency and selectivity for cleavage by cathepsins. It can be expected that the pHMPA scaffold

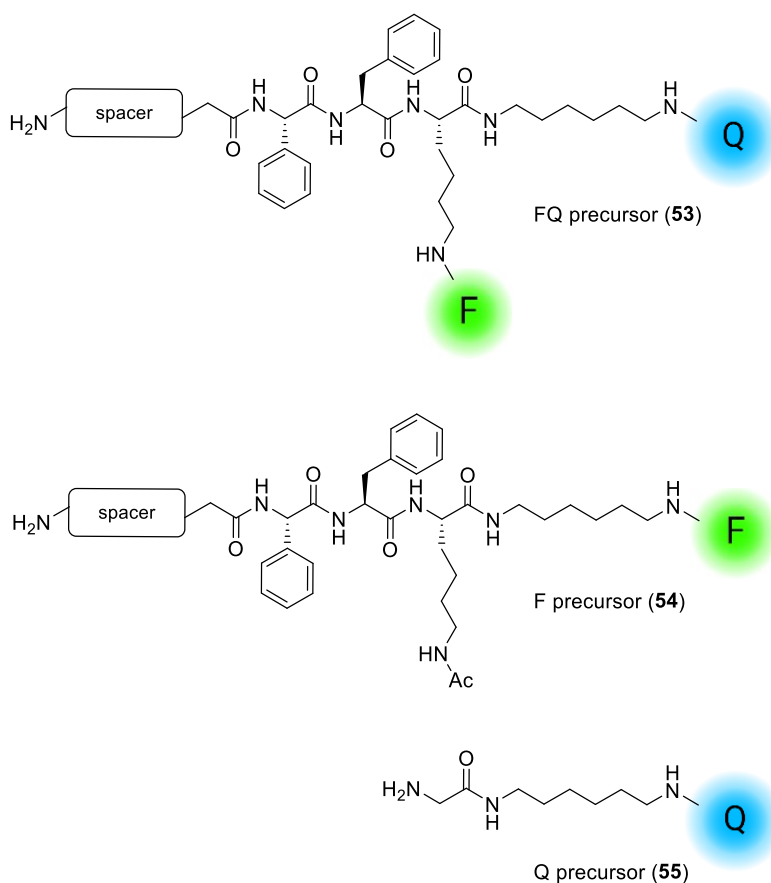
might cause steric hindrance and, therefore, affect cleavage efficiency. For this reason, we planned to introduce different linkers between the macromolecular scaffold and the protease substrate to eliminate this factor and preserve efficient cleavage. The N-terminus of **6QC** appeared to be an ideal option for spacer attachment. From previous data obtained in the group of Prof. Bogyo, we knew that the Cbz group is important for **6QC** recognition by cathepsins.¹²⁸ Based on this knowledge, we decided to replace the Cbz group with phenylglycine which also performed well in the original screening, but the α -amino group would allow for chain extension and spacer attachment. Combining all this information led us to design of the first **6QC** analog **52** shown in **Scheme 1.25**, which features the original quencher-fluorophore arrangement of **6QC**. This probe architecture will be referred to as the QF architecture and, therefore, the low-molecular precursor is designated as the QF precursor (**52**).



Scheme 1.25 Structure of **6QC** and QF architecture precursor (**52**)

Then, we proposed an inverse FQ architecture where the quencher would be cleaved off the polymer, instead of the fluorophore in the case of a QF architecture. Originally, **6QC** was designed to release the amine-containing fluorophore fragment upon cleavage. This design should increase cellular retention due to the latent cationic lysosomotropic effect, as described by Ofori et al.¹²⁵ This plays an important role in the case of small molecules, which can easily diffuse away, but with macromolecular probes the situation is notably different. Therefore, we planned to synthesize both QF and FQ probes to investigate if differences in contrast or signal retention could be observed. The structure of the FQ precursor **53** is shown in **Scheme 1.26**. In order to fully exploit the

modularity enabled by the pHPMA carrier, we were intrigued about an idea to separate the quencher and the fluorophore attaching them to the macromolecular scaffold separately. We proposed a Qx:1F architecture, where the quencher would be attached directly to the polymer backbone, via a short non-cleavable linker, and the fluorophore would occupy the same position as in the case of **6QC**. The lysine amino residue was acetylated to mask this group and avoid cross-reactivity. The structures of both F (**54**) and Q (**55**) precursors can be seen in **Scheme 1.26**. In the case of Q precursor **55**, glycine was added to the end of the spacer as a quantification tag for amino acid analysis (AAA). Typically, we use AAA for the quantification of individual moieties on the final iBodies conjugates. The 6QC analogues can be easily quantified via multiple amino acids present in the substrate sequence.



Scheme 1.26 Structure of FQ (**53**), F (**54**) and Q (**55**) precursors

In case of success, the Qx:1F architecture could simplify the development of new quenched probes significantly, as the attachment of multiple dyes to a single peptide fragment generally tends to reduce solubility and complicate the synthesis or purification processes. Furthermore, we believed that removal of the quencher from the protease cleavage site vicinity could reduce steric hindrance caused by presence of the bulky

quencher moiety and lead to a more efficient cleavage by the target protease(s). Nevertheless, as discussed in **Chapter 1.4.1**, we were aware and cautious about the fact that the quenching-enabling FRET is highly distance dependent and, therefore, this architecture will inevitably affect quenching efficiency of such probes. To overcome this problem, we proposed increasing the quencher-fluorophore ratio to compensate for the loss caused by distance increase. As discussed earlier in this chapter, this can be easily achieved during the synthesis of iBodies.

The design of probe architectures discussed in this chapter is visually summarized in **Figure 1.12**. This cartoon representation shows pHPMA carrier as a red ball with low-molecular precursors attached to the polymeric core. In this representation, each polymer contains two fluorophore moieties, for simplification, but this will inevitably vary on the molecular level. The Q1:1F and Q2:1F architecture probes possess two and four quencher moieties, accordingly. Cleavable bonds are highlighted in red, and the reactions show the fluorescent product upon processing by a corresponding protease (represented by the light blue pacman icon).

Additionally, we designed a non-quenched probe nQ which served as a control. As discussed in the beginning of this chapter, we believe that the EPR effect will play a significant role in improving the contrast of low-molecular probes after attachment to the pHPMA carrier. Nevertheless, we need to confirm that the contrast is provided by both the EPR effect along with contribution from the cathepsin cleavage. To address this issue, we designed the nQ probe, which is only decorated with the fluorophore (sulfo-Cy5) attached via a short non-cleavable linker (**56**), according to **Figure 1.13**. Even though it is hard to fully deconvolute individual contributions of the two effects *in vivo*, we believe the nQ probe should provide us with information about the extent of the EPR effect and its contribution to contrast of our macromolecular probes.

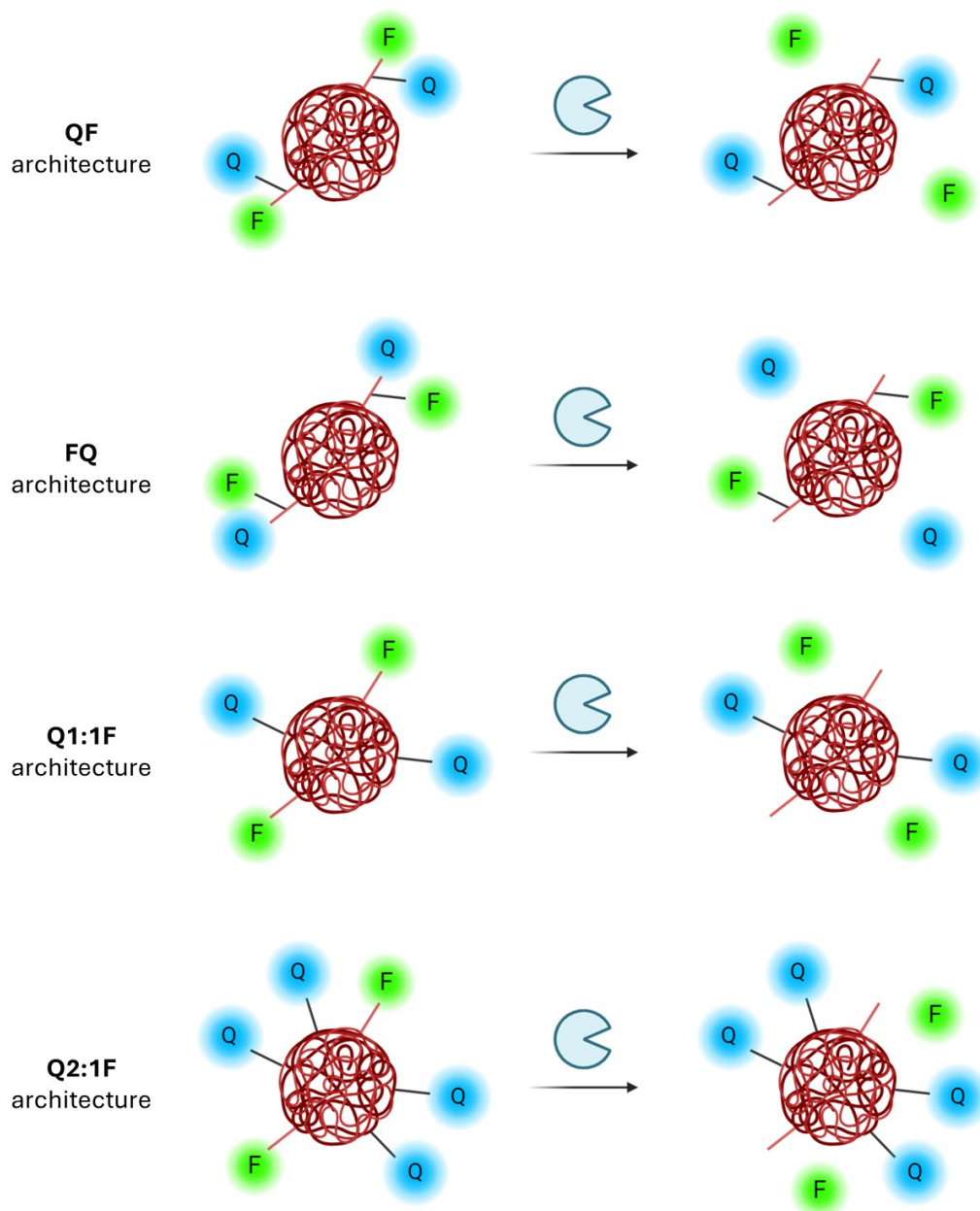


Figure 1.12 Schematic representation of the proposed probe architectures. The pHPMA carrier is represented by the red sphere, the protease responsible for probe processing is represented by the light blue pacman, and cleavable bonds are highlighted in red color. Created with Biorender.com

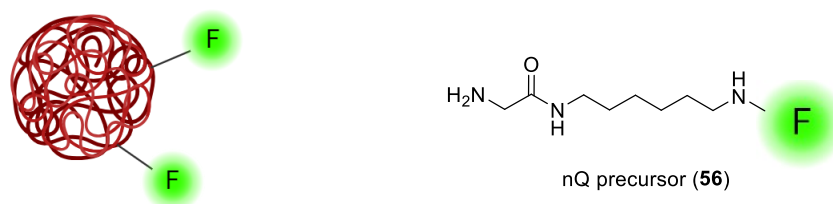


Figure 1.13 Schematic representation of the nQ probe (left) and structure of nQ probe precursor (56) (right). Created with Biorender.com.

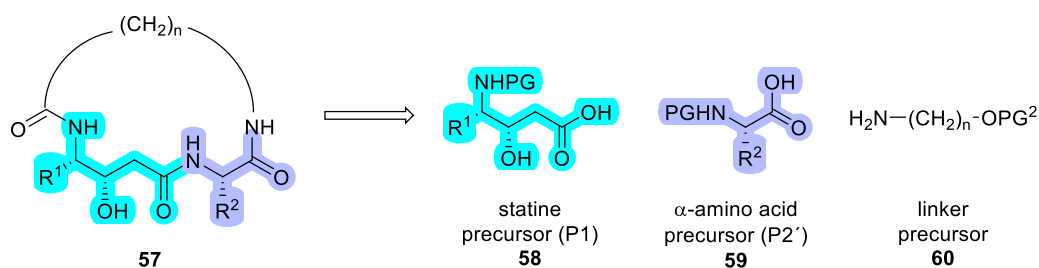
2 AIMS OF WORK

- Design and synthesize macrocyclic inhibitors of hCatD inspired by PepA
- Design and synthesize linker-modified gingipain inhibitors based on KYT-1 and KYT-36
- Synthesize low-molecular precursors of HPMA-based probes for imaging of tumor-associated proteases
- Evaluate the HPMA-based probes *in vitro* and *in vivo*

3 RESULTS AND DISCUSSION

3.1 Macrocyclic Inhibitors of Human Cathepsin D

The synthesis of the proposed macrocyclic inhibitors of hCatD requires three precursors shown in **Scheme 3.1** – a statine precursor (P1), an α -amino acid precursor (P2'), and a linker connecting the first two precursors. The precursors must possess suitable protecting groups to enable a successful peptide coupling strategy leading to the final macrocyclic inhibitor **57**. In general, we aimed to work with Boc-protected statine **58** and α -amino acid **59** precursors due to their availability and facile deprotection using TFA. For the linker (**60**), we selected an orthogonal methyl ester protecting group which can be easily removed by mild basic hydrolysis.



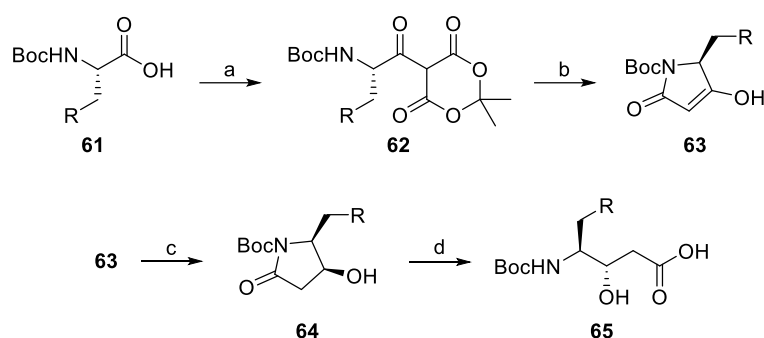
Scheme 3.1 General structure of a proposed macrocyclic inhibitor **57** with individual precursors

The commercial availability of individual precursors varied significantly. Boc-protected derivatives of a wide range of α -amino acids could be purchased. In the case of linker precursors, most derivatives were commercially available, but some were either not commercially available or expensive, and had to be synthesized. Nevertheless, most of our synthetic efforts focused on preparing various statine precursors. The commercial availability of statine analogues is generally low and only a few derivatives could be purchased when we began working on this project. Besides the natural statine derived from leucine, analogues with phenylalanine, cyclohexylalanine and tryptophan side chain residues were commercially available. Nevertheless, we aimed to extend the scope of our SAR beyond these compounds, and thus we synthesized multiple novel statine analogues that had not been previously described.

In the following chapters, the strategy for the synthesis of individual precursors and the macrocyclic inhibitors will be discussed. The synthetic section will be followed by the *in vitro* evaluation of the prepared inhibitors.

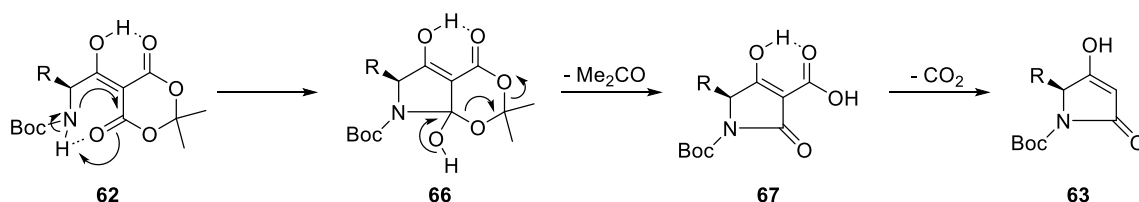
3.1.1 Preparation of Statine Analogues

As discussed in the previous chapter, we aimed to synthesize a wide range of statine analogues and incorporate them into our macrocyclic inhibitors to extend inhibitor interactions within the P1 binding subsite of hCatD. We were inspired by the work of Jouin et al.,¹³¹ who described an elegant, stereospecific synthetic route for the synthesis of *N*-protected statine analogues, as shown in **Scheme 3.2**.



Scheme 3.2 Synthesis of *N*-protected statine analogues (**65**) as reported by Jouin et al.¹³¹: (a) coupling reagent, DMAP, Meldrum's acid, DCM, rt; (b) various solvents, reflux; (c) reducing agent; (d) NaOH, acetone, rt.

In the first step, Boc-protected α -amino acid **61** is coupled with Meldrum's acid, which serves as a convenient building block for elongating the carbon chain. Next, intermediate **62** is refluxed in ethyl acetate or methanol, leading to cyclization to a tetramic acid **63**. The authors of the paper proposed a mechanism for this transformation, as shown in **Scheme 3.3**. The reaction begins with a nucleophilic attack by the Boc-protected amine on one of the ester functionalities. The tetrahedral intermediate **66** then collapses to release amide **67** and acetone. In the final step, β -keto acid **67** spontaneously decarboxylates to yield the final product of the reaction, tetramic acid **63**.

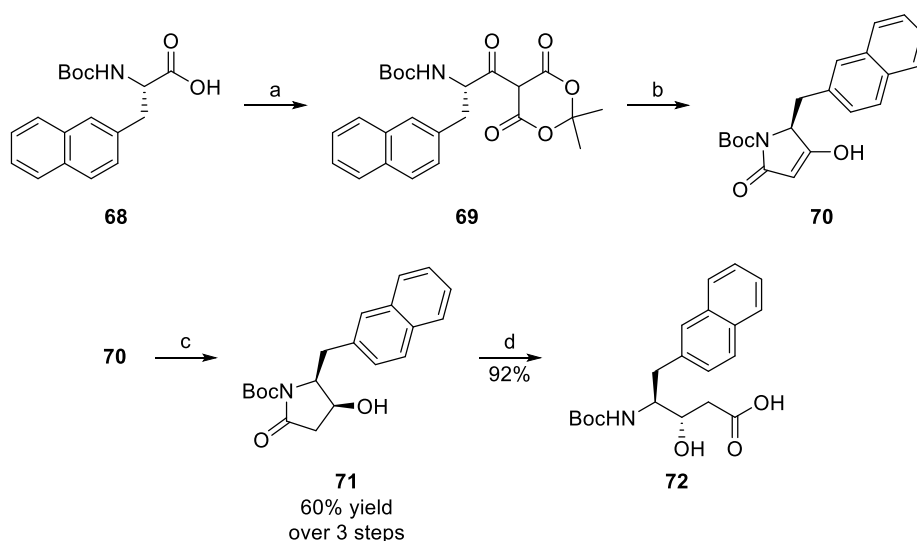


Scheme 3.3 Cyclization of intermediate **62** to a tetramic acid **63**

Tetramic acid **63** can be stereospecifically reduced to yield cyclic statine precursor **64** in the desired *threo* conformation. This can be achieved using either catalytic hydrogenation or hydride-based reducing reagents. In the final step, cyclic statine **64** is

regioselectively hydrolyzed to yield statine **65**. The authors of the paper also reported preparation of Cbz-protected statine analogue using this synthetic route.

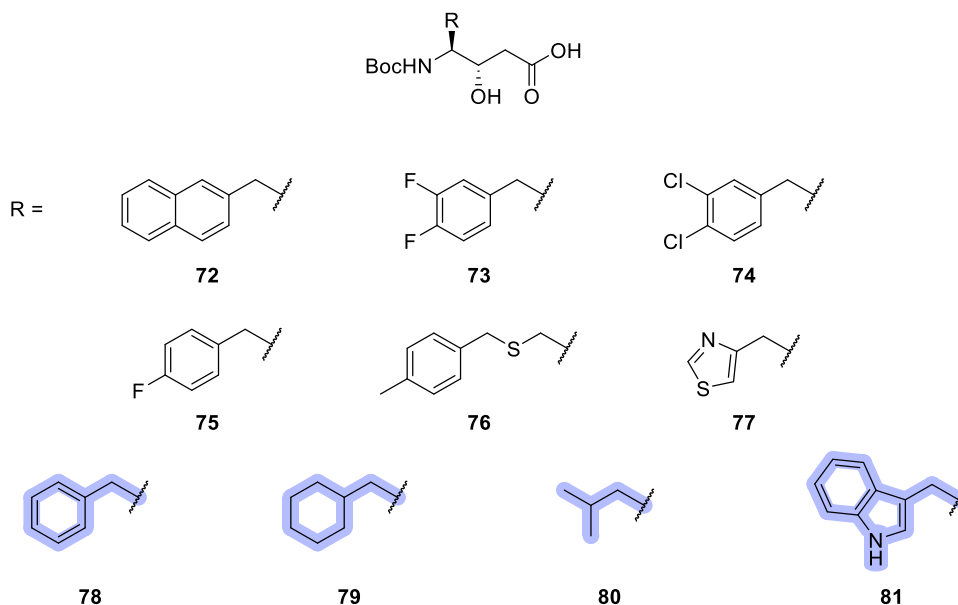
In this work, we provide a full description of the synthetic route leading to the 3-(2-naphthyl)-L-alanine statine analogue **72**. As shown in **Scheme 3.4**, *N*-Boc-protected 3-(2-naphthyl)-L-alanine acid **68** was coupled with Meldrum's acid using DCC to yield intermediate **69**. From the reagents reported by Jouin et al.,¹³¹ I preferred DCC over isobutyl chloroformate due to easier handling and lower toxicity. Intermediates **69** and **70** were only purified by work-up and filtration due to their high acidity, caused by the presence of conjugated carbonyl groups. In the next step, intermediate **69** was refluxed in EtOAc to yield tetramic acid **70**, which was then stereospecifically reduced to cyclic statine **71** using NaBH₄ in acidic conditions. In this case, I selected NaBH₄, due to the high pressure necessary for catalytic hydrogenation of tetramic acids, as reported in the paper.¹³¹ After reduction of the keto group, the reaction mixture could be easily separated by a normal-phase chromatography to yield cyclic statine **71** in 60% yield over three steps. The hydrolytic opening of cyclic statine **71** yielded the 2-naphthylalanine statine analogue **72**.



Scheme 3.4 Preparation of 3-(2-naphthyl)-L-alanine statine analogue **72**: (a) DCC, DMAP, Meldrum's acid, DCM, rt; (b) EtOAc, reflux; (c) NaBH₄, DCM:AcOH / 10:1, 0 °C to rt; (d) NaOH, acetone, rt.

By employing this general procedure, a series of statine analogues, as illustrated in **Scheme 3.5**, was successfully prepared to supplement the commercially available statines (highlighted in steel). All prepared statine analogues contain aromatic side chain residues, due to the high lipophilicity of the S1 subsite. Moreover, the crystal structure of hCatD with PepA indicated that the original leucine residue of PepA only partially

populates the S1 subsite. Our set of statine analogues was also inspired by previous SAR studies published in the literature.^{33–36} To summarize, in addition to the commercially available statines with phenylalanine (**78**), cyclohexylalanine (**79**), leucine (**80**) and tryptophan (**81**) residues, I synthesized the 2-naphthylalanine analogue (**72**), three phenylalanine analogues with various halogen substituents (**73**, **74**, **75**), a *p*-tolyl-protected cysteine analogue (**76**) and a 4-thiazolealanine analogue (**77**).

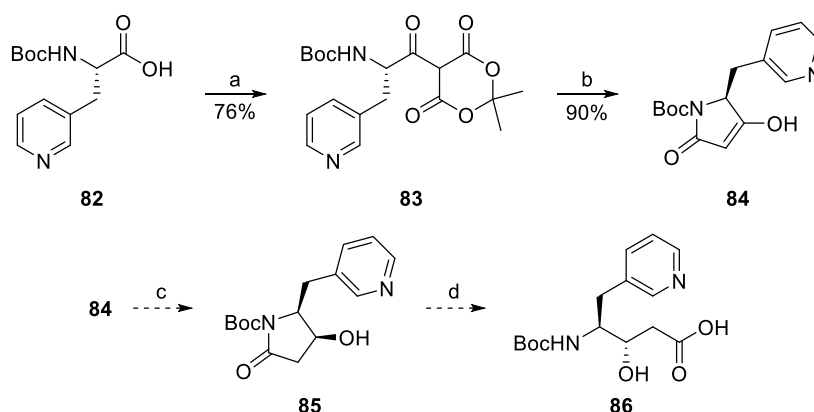


Scheme 3.5 Structures of synthesized (no highlight) and commercially available (highlighted in steel) statine analogues

Even though the general procedure used for the synthesis of statine analogues yielded all the desired compounds in sufficient quantity and purity, I decided to further optimize this methodology by introducing certain modifications. For example, I successfully replaced DCC with EDC in the first step of the synthetic sequence. This substitution replaces the dicyclohexylurea byproduct with an EDC-derived urea, which is easily washed away during the acidic work-up after the reaction, reducing the number of purification steps. Additionally, in certain cases, the reduction of tetramic acid to the cyclic statine did not proceed to full conversion. To facilitate separation of the tetramic acid derivatives from the neutral cyclic statines, I implemented an acid-base work-up procedure which allowed their efficient separation.

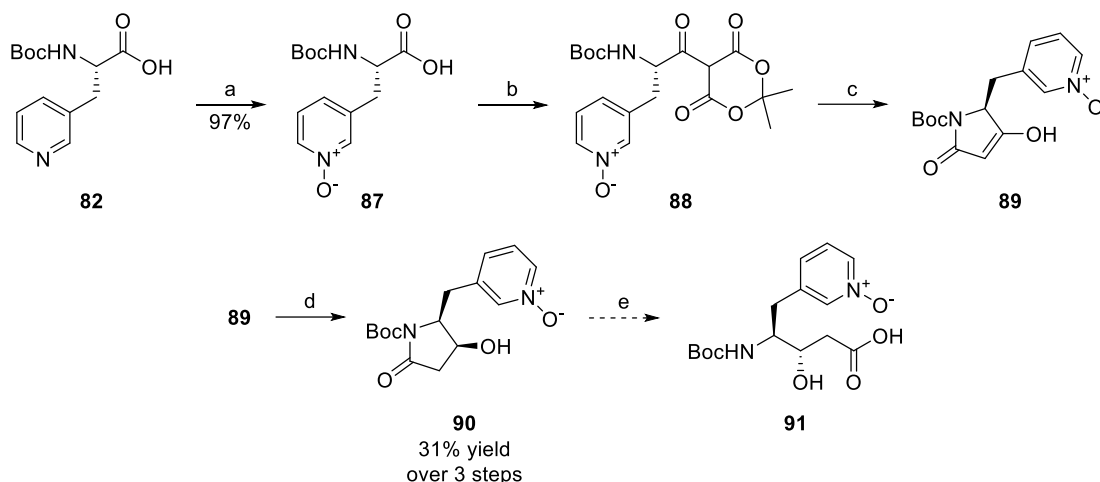
Furthermore, we were intrigued by the idea of developing statine analogues with basic heteroaromatic residues, such as pyridine, which, to the best of our knowledge have not been reported in the literature. Therefore, I tried to apply our optimized procedure for statine synthesis to Boc-protected 3-(3-pyridyl)-L-alanine, according to **Scheme 3.6**. Due

to the basicity of the pyridine moiety, an acidic work-up could not be applied to separate intermediate **83** from the coupling reagents and byproducts. Nevertheless, reversed-phase chromatography allowed us to isolate intermediate **83** in high purity. Interestingly, this compound was highly insoluble after the chromatography. After refluxing intermediate **83** in EtOAc until dissolution, tetramic acid **84** was isolated in very good yield. Unfortunately, despite several attempts and condition optimizations, the reduction of tetramic acid **84** did not proceed efficiently and cyclic statine **85** was only isolated in trace quantities.



Scheme 3.6 Preparation of the 3-(3-pyridyl)-L-alanine statine analogue **86**: (a) EDC, DMAP, Meldrum's acid, DCM, rt; (b) EtOAc, reflux; (c) NaBH₄, DCM:AcOH / 10:1, 0 °C to rt; (d) NaOH, acetone, rt.

We hypothesized that the reactivity issues observed with tetramic acid **84** may be associated with the presence of the basic pyridine moiety. Therefore, I decided to apply a synthetic strategy involving oxidation of the pyridine nitrogen to generate an *N*-oxide, which would serve as a masking functionality. First, Boc-protected 3-(3-pyridyl)-L-alanine **82** was oxidized to the corresponding *N*-oxide **87** in excellent yield. Next, *N*-oxide **87** was coupled with Meldrum's acid to yield intermediate **88**. Purification by acidic work-up was possible in this case due to the masked pyridine nitrogen and no chromatographic purification was necessary. Refluxing intermediate **88** in EtOAc yielded tetramic acid **89** in high purity. In the next step, it was reduced to cyclic statine **90** using previously optimized conditions. This approach yielded the cyclic statine in 30% overall yield over four reaction steps. Despite the reported yield being low, we believe that this procedure can be optimized to afford a higher yield and, most importantly, it may provide access to a new class of statine analogues with basic heteroaromatic substituents.

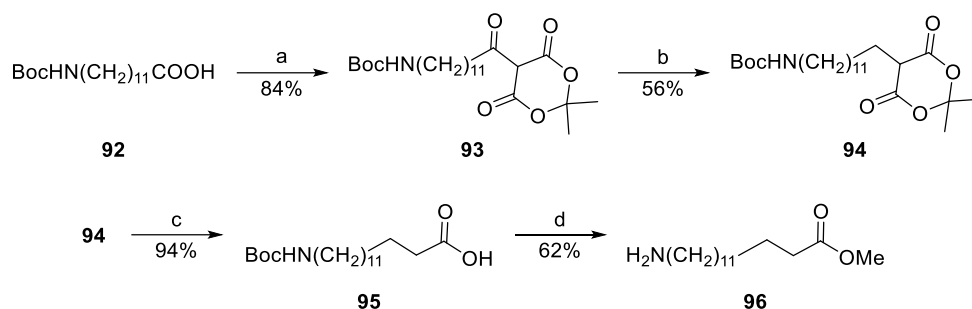


Scheme 3.7 Preparation of the 3-(3-pyridyl)-L-alanine statine analogue via *N*-oxide: (a) MCPBA, DCM, rt; (b) EDC, DMAP, Meldrum's acid, DCM, rt; (c) EtOAc, reflux; (d) NaBH₄, DCM:AcOH / 10:1, 0 °C to rt; (e) NaOH, acetone, rt.

3.1.2 Preparation of Non-Commercial ω -Amino Aliphatic Acids

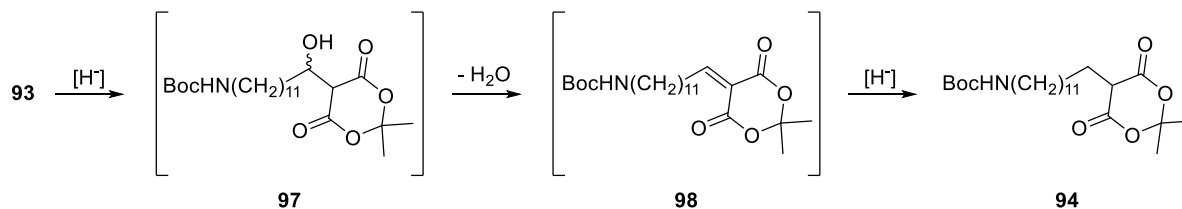
Another important class of precursors to our macrocyclic inhibitors were ω -amino aliphatic acids used as linkers. We aimed to synthesize macrocycles with different linker lengths to identify the optimal macrocyclic ring size for the inhibition of hCat. As a result, various ω -amino aliphatic acids were required. While most of the necessary linkers were commercially available, some, such as methyl 14-aminotetradecanoate (**96**), had to be synthesized.

In this case, we employed a carbon chain extension approach starting from Boc-12-aminododecanoate (**92**). This approach follows a similar synthetic sequence as the statine analogue synthesis described in **Chapter 3.1.1** and was inspired by a procedure published by Smrcina et al.¹³² According to **Scheme 3.8**, aliphatic acid **92** was coupled with Meldrum's acid to yield intermediate **93** in very good yield. Next, in the presence of NaBH₄, intermediate **93** undergoes complete reduction of the exocyclic ketone functionality leading to compound **94**.



Scheme 3.8 Synthesis of methyl 14-aminotetradecanoate **96**: (a) DCC, DMAP, Meldrum's acid, DCM, rt; (b) NaBH₄, DCM:AcOH / 10:1, 0 °C to rt; (c) DMF:H₂O / 10:1, 110 °C; (d) (i) SOCl₂, MeOH, 0 °C; (ii) TFA, rt.

Conversion of **93** to **94** proceeds via the mechanism described in **Scheme 3.9**, as reported by the authors of the paper.¹³² The reduction of **93** yields the corresponding β -hydroxydiester **97**, which undergoes spontaneous dehydration to form an α,β -unsaturated diester **98**. This intermediate is finally reduced to **94** via Michael addition of hydride ion to the conjugated double bond. Heating intermediate **94** in the presence of water leads to ester hydrolysis and subsequent spontaneous decarboxylation, yielding Boc-14-aminotetradecanoate (**95**). Acid **95** was then converted to the corresponding methyl ester and the Boc protecting group was removed in two subsequent steps to yield the desired methyl 14-aminotetradecanoate **96**.

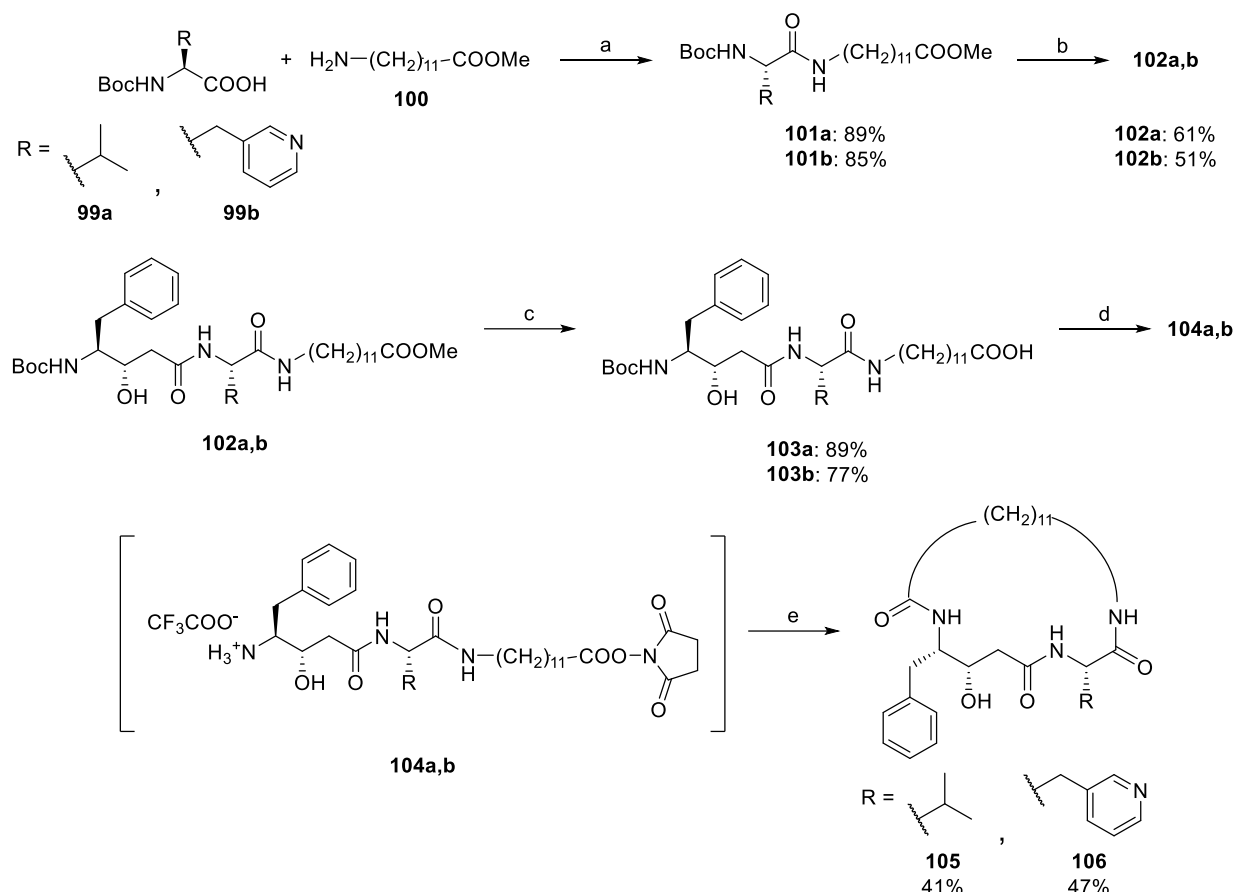


Scheme 3.9 Mechanism of reduction of **93** to **94**

3.1.3 Preparation of Macrocyclic Inhibitors

After gathering all the necessary precursors, we proceeded with the synthesis of the final macrocyclic inhibitors. Our approach to synthesizing macrocycles with an aliphatic linker was based on a series of consecutive peptide coupling reactions and deprotections, as shown in **Scheme 3.10**. Here, I provide a complete description of the synthetic route leading to two macrocycles (**105**, **106**) with a C₁₁ aliphatic linker and a phenylalanine statine analogue (PheSta, **78**), which differ in the side chain residue of the α -amino acid fragment.

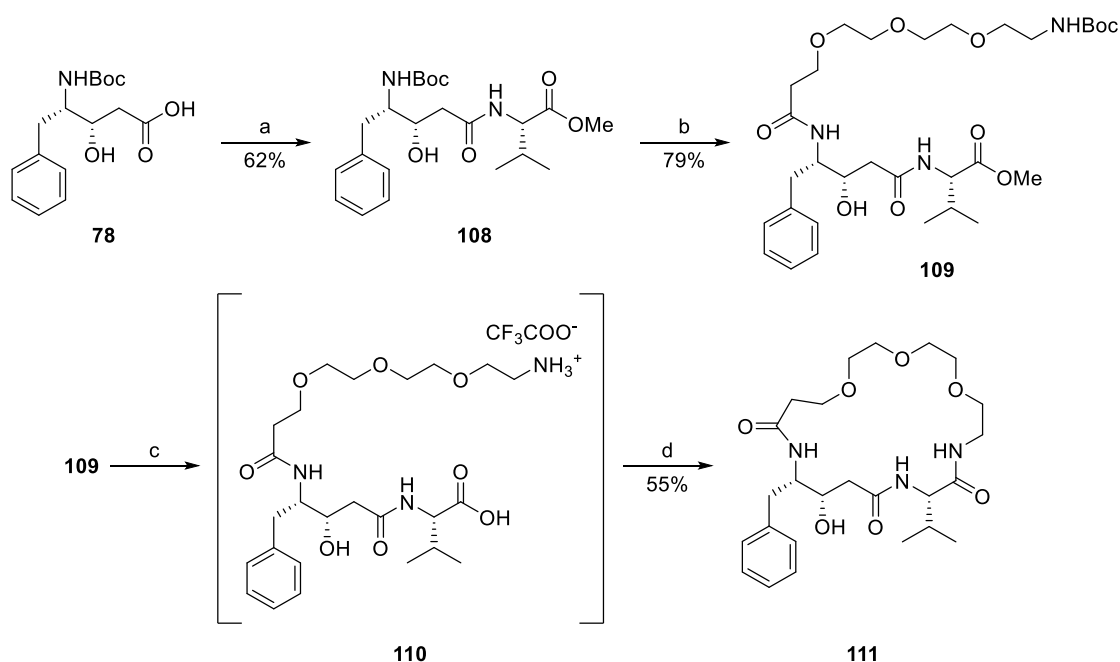
In the first step, Boc-protected α -amino acid (**99a,b**) was coupled with methyl 11-aminoundecanoate (**100**) using HBTU as the coupling reagent. Generally, the reaction proceeded with very good yields. Next, intermediate **101a,b** was deprotected and coupled with the corresponding Boc-protected statine analogue, PheSta (**78**) in this case, yielding intermediate **102a,b** in average yields. LiOH was then used to convert the methyl ester to the corresponding free acid **103a,b** under mild conditions. Next, acid **103a,b** was converted to NHS ester using TSTU, and excess reagent was removed by aqueous work-up. The crude NHS ester was then dissolved in TFA to remove the Boc-protecting group leading to macrocyclic inhibitor precursor **104a,b**. This precursor was used in the final cyclization step without further purification. To suppress intermolecular coupling reactions and minimize solvents usage, we performed the cyclization reaction by slowly adding precursor **104a,b** to a solution of DIPEA in DMF. The amino group of the intermediate is fully protonated after the Boc deprotection reaction and, therefore, unreactive towards the NHS ester until a base is added. This allowed us to dissolve intermediate **104a,b** in a small volume of DMF without triggering the reaction. Upon dropwise addition to a basic solution, intermediate **104a,b** spontaneously cyclizes to macrocycle **105**, **106**. The final macrocycles were typically purified by reversed-phase (RP) HPLC and isolated in average yields. Using this procedure, I successfully synthesized over 20 macrocyclic inhibitors (**Table 3.1**, **3.2**, **3.3**) with different α -amino acids, statine analogues and linker lengths.



Scheme 3.10 General procedure for synthesis of macrocycles with aliphatic chain: (a) HBTU, DIPEA, DMF, rt; (b) (i) DCM, TFA, rt; (ii) Boc-PheSta-OH (**78**), HBTU, DIPEA, DMF, rt; (c) LiOH, THF:H₂O / 1:1, rt; (d) (i) TSTU, DIPEA, DMF, rt; (ii) TFA, rt; (e) DIPEA, DMF, rt.

Next, we aimed to synthesize a macrocycle possessing a PEG linker to increase the overall hydrophilicity and solubility of the inhibitor. In this case, we selected a different cyclization site due to the availability of a Boc-protected PEG derivative in our laboratory. Also, we were interested in verifying whether the cyclization would occur with efficiency comparable to that observed with the original cyclization site. First, Boc-PheSta-OH (**78**) was coupled with valine methyl ester using HBTU to yield intermediate **108**, as shown in **Scheme 3.11**. Next, the Boc protecting group was removed and the intermediate amine was coupled with a PEG-3 derivative, yielding protected linear precursor **109** in very good yield. After hydrolysis of the methyl ester and removal of the Boc protecting group, the fully deprotected linear precursor **110** was obtained and used in the final cyclization step without further purification. In this instance, I selected DPPA as the cyclization reagent of choice due to its ability to suppress racemization and its low reactivity towards amines. Analogous to the preparation of macrocycles **105** and **106**, a solution of linear precursor **110** was slowly added to the basic solution of DPPA to achieve pseudo-dilution. This procedure yielded the final macrocycle **111** in very good yield.

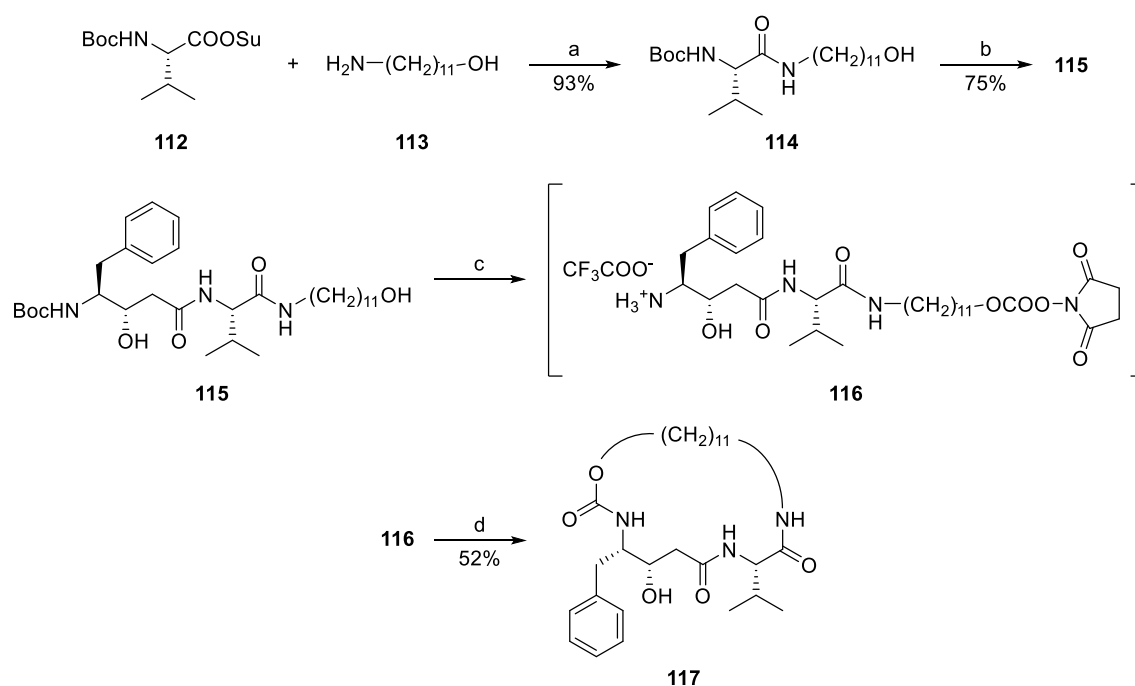
Overall, the yield of this approach was comparable to that of the original procedure described in **Scheme 3.10**. Additionally, we did not observe significant racemization during hydrolysis of the methyl ester **109** or the activation of the acid **110**. However, this could change in the case of different amino acids, as valine is generally highly resistant to racemization due to the steric hinderance caused by its isopropyl side chain. Therefore, caution would be required when applying this methodology to other, more racemization-prone, amino acids.



Scheme 3.11 Synthesis of a macrocycle with PEG chain: (a) (i) HBTU, DIPEA, DMF, rt; (ii) H-Val-OMe, DMF, rt; (b) (i) TFA, rt; (ii) BocHN-PEG-3-CH₂CH₂COOH, HBTU, DIPEA, DMF, rt; (c) (i) LiOH, THF:H₂O / 1:1, rt; (ii) TFA, rt; (d) DPPA, DIPEA, DMF, rt.

As an intermediate between the polarity of macrocycles with aliphatic linkers and the macrocycle with a PEG linker, we designed a macrocycle with a carbamate moiety. We hypothesized that the additional oxygen could slightly increase polarity and water solubility without disrupting the lipophilic interactions of the linker. The synthetic strategy leading to the macrocycle with a carbamate moiety **117** is shown in **Scheme 3.12** and is based on the approach, used for the synthesis of macrocycles with aliphatic chains (**Scheme 3.10**). First, Boc-Val-OSu (**112**) was coupled with aliphatic aminoalcohol **113**, yielding intermediate **114** in excellent yield. In the next step, the Boc protecting group was removed and the corresponding amine was coupled with Boc-PheSta-OH (**78**) to yield the protected linear precursor **115**. In this case, TSTU was used instead of HBTU to avoid any potential cross reactivity with the primary alcohol present in the structure. Next, the

primary alcohol of **115** was activated using DSC and an aqueous work-up was used to remove excess reagent. Interestingly, the highest selectivity for activation was observed when no base was used in the reaction. After Boc deprotection, the reaction yielded crude linear precursor **116**, which was used in the cyclization step without further purification. In the final step, macrocyclic precursor **116** was cyclized to carbamate-containing macrocycle **117** under the pseudo-dilution conditions previously described in this chapter. This approach yielded macrocycle **117** in good overall yield, comparable to yields reported for macrocycles without the carbamate moiety.



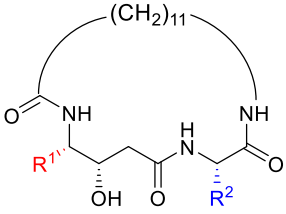
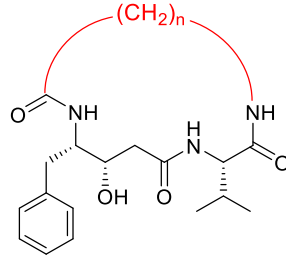
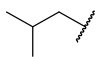
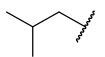
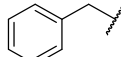
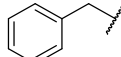
Scheme 3.12 Synthesis of a macrocycle with carbamate moiety: (a) DIPEA, DMF, rt; (b) (i) TFA, DCM, rt; (ii) Boc-PheSta-OH (**78**), TSTU, DIPEA, DMF, rt; (c) (i) DSC, DMF, rt; (ii) TFA, DCM, rt; (d) DIPEA, DMF, rt.

This chapter presents the general procedures developed for the synthesis of peptidic macrocycles containing various amino acids, statine analogues and linkers. In total, approx. 25 macrocycles (**Table 3.1, 3.2, 3.3**) were synthesized and characterized. These macrocycles and their biological properties will be presented in the following chapter.

3.1.4 Evaluation of Macrocyclic Inhibitors

To confirm our hypothesis that both the P1 and P2' residues of PepA only partially occupy the corresponding binding pockets, we synthesized an initial set of macrocycles (**105**, **118-124**) shown in **Table 3.1**. While macrocycle **118** contains the original isopropylmethyl (R^1) and methyl (R^2) substituents of PepA, the methyl group was replaced with an isopropyl group in compound **119** and the isopropylmethyl group was replaced with a benzyl group in compound **120**. The inhibitory potency of both macrocycles **119** and **120** was significantly higher than that of compound **118**. The combination of both residues in compound **105** led to a further decrease in IC_{50} to 5.8 ± 0.7 nM. Next, we synthesized five macrocycles with fixed substituents R^1 and R^2 , and varying linker lengths ranging from 10 to 14 methylene units. The IC_{50} values of these macrocycles varied within one order of magnitude. Compounds **105** and **122** both exhibited optimal inhibition with single-digit nanomolar potency. We decided to proceed with the smaller ring size ($n = 11$, **105**) for further inhibitor optimization.

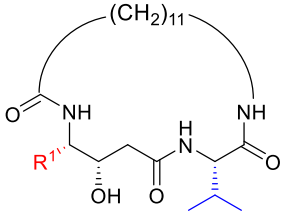
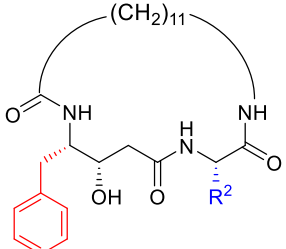
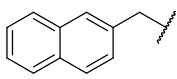
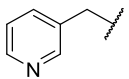
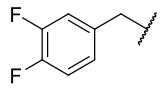
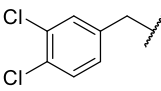
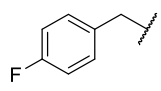
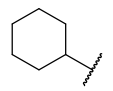
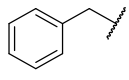
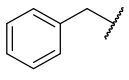
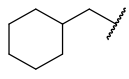
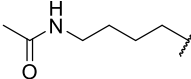
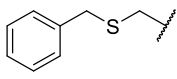
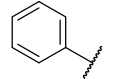
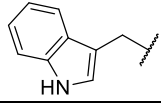
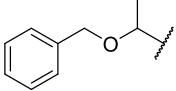
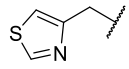
Table 3.1 Initial design of PepA-mimicking macrocyclic inhibitors of hCatD. Compound **118** contains the original PepA substituents (R^1, R^2). Modification of the two substituents (left) and optimization of macrocycle length (right) improved the inhibitory potency of the macrocyclic scaffold

							
compound	R^1	R^2	IC_{50} (nM)	compound	n	ring size	IC_{50} (nM)
118			310 ± 61	121	10	20	40 ± 11
119			22 ± 4.5	105	11	21	5.8 ± 0.7
120			68 ± 19.3	122	12	22	3.5 ± 0.5
105			5.8 ± 0.7	123	13	23	11 ± 1.4
				124	14	24	32 ± 4.4

Next, we conducted an extensive screen of R¹ and R² substituents to map the hCatD subsite preferences. As shown in **Table 3.2**, we fixed a valine in the P2' position (R² = isopropyl) to evaluate various statine analogues in the P1 position. For R² variations, we fixed PheSta in the P1 position (R¹ = benzyl). As discussed earlier, the binding pockets accommodating these two substituents are highly lipophilic, therefore, we designed the series to contain primarily aromatic and non-polar aliphatic substituents. We began by incorporating various statine analogues to optimize the R¹ substituent. Interestingly, a strong preference for aromatic residues in this position was observed. Macrocycles (**105**, **126–128**) with a benzyl moiety in the R¹ position all exhibited single-digit nanomolar inhibition. No significant differences in IC₅₀ values were observed for different halogen substituents on the aromatic ring in the case of macrocycles **126–128**. Notably, macrocycles with increased polarity at the R¹ position, such as **133** possessing cysteine residue or (**132**, **134**) possessing heterocyclic residues, showed a significant drop in potency. While macrocycle **129** possessing the cyclohexylmethyl residue in R¹ exhibited comparable potency to the benzyl group-containing macrocycle **105**, compound **125** equipped with the 2-naphthylmethyl in R¹ showed sub-nanomolar IC₅₀ of 0.6 ± 0.2 nM.

In the case of R², lipophilic residues were also preferred, similar to R¹. However, no strong preference for aromatic residues was observed, both aromatic (benzyl, **137**; 3-pyridylmethyl, **106**) and aliphatic (isopropyl, **105**; 2-butyl, **135**; cyclohexyl, **136**) substituents were well tolerated and the corresponding macrocycles exhibited IC₅₀ values in the low nanomolar range. Interestingly, macrocycle **139** with a phenyl residue showed over three-fold decrease in potency compared to macrocycle **136** possessing a cyclohexyl residue. The introduction of polar substituents in R², such as the 1-hydroxyethyl in macrocycle **141** caused a significant loss in potency. Benzylation of this position in macrocycle **142** did not restore the inhibitory activity, likely due to the excessive steric volume of the benzyl group in the relatively small P2' pocket. Rather surprisingly, the introduction of the basic pyridine moiety in macrocycle **106** did not lead to a decrease in inhibitory potency. In fact, an over 13-fold increase in inhibitory potency was observed compared with the benzyl analog **137**. This is the only example from **Table 3.2** where the introduction of a more polar moiety increased the inhibitory potency. Macrocycle **106** reached the lowest IC₅₀ in this series.

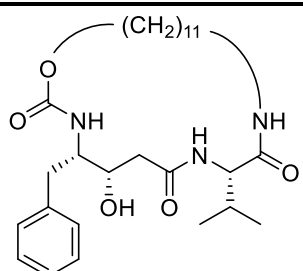
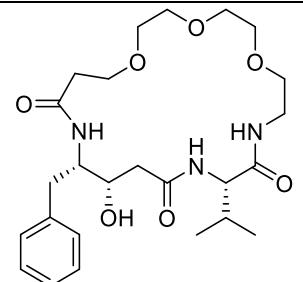
Table 3.2 Modification of R¹ and R² substituents of macrocyclic inhibitors of hCatD. Compounds with variation at the R¹ substituent (left) and the R² substituent (right) are ordered according to their potency. The series also includes compounds from **Table 3.1** for comparison.

					
compound	R ¹	IC ₅₀ (nM)	compound	R ²	IC ₅₀ (nM)
125		0.6 ± 0.2	106		1.1 ± 0.2
126		3.1 ± 0.6	135		3.8 ± 0.7
127		4.2 ± 0.7	105		5.8 ± 0.7
128		4.7 ± 0.7	136		8.2 ± 0.7
105		5.8 ± 0.7	137		15 ± 3.4
129		9.0 ± 0.7	138		16 ± 1.1
130		20 ± 4.2	139		27 ± 2.8
131		22 ± 4.5	120		68 ± 19
132		190 ± 20	141		500 ± 55
133		340 ± 38	142		560 ± 130
134		600 ± 150			

As discussed above, both P1 and P2' showed a strong preference for lipophilic residues. Given the high lipophilicity of the most potent inhibitors, we sought to introduce backbone modifications to increase their polarity and, potentially, water solubility. Macrocycle **117**, a carbamate analogue of **122** with an identical ring size (22 atoms), exhibited a two-fold reduction in IC₅₀ compared to the original "full carbon" analogue **122**. We also synthesized macrocycle **111** possessing a PEG linker to investigate whether the aliphatic linker could be replaced by a more hydrophilic ethylene glycol-based linker without a loss in potency. This alteration resulted in a significant change in the IC₅₀, with a two-order-of-magnitude drop in potency compared with inhibitor **122**. This finding highlights the significance of hydrophobic interactions in the P2/P3' region occupied by the linker.

This chapter evaluates the inhibition of hCatD by our macrocyclic inhibitors. Several interesting trends were identified in the SAR, elucidating hCatD preferences within the P1 and P2' binding pockets, as well as the P2/P3' region that accommodates the linker. Further structural biology and computational studies have been conducted to establish the binding characteristics of our macrocyclic inhibitors to hCatD. Complexes of hCatD with inhibitors **105**, **125** and **117** were successfully crystallized and the corresponding crystal structures were resolved. The crystallographic complexes were then used to evaluate the interaction energies of the macrocycles with hCatD by performing quantum chemical calculations. These studies can be found in the original publication and are not discussed in further detail in this work as they were performed by our collaborators from various groups at IOCB and are beyond the scope of this doctoral thesis.³⁹

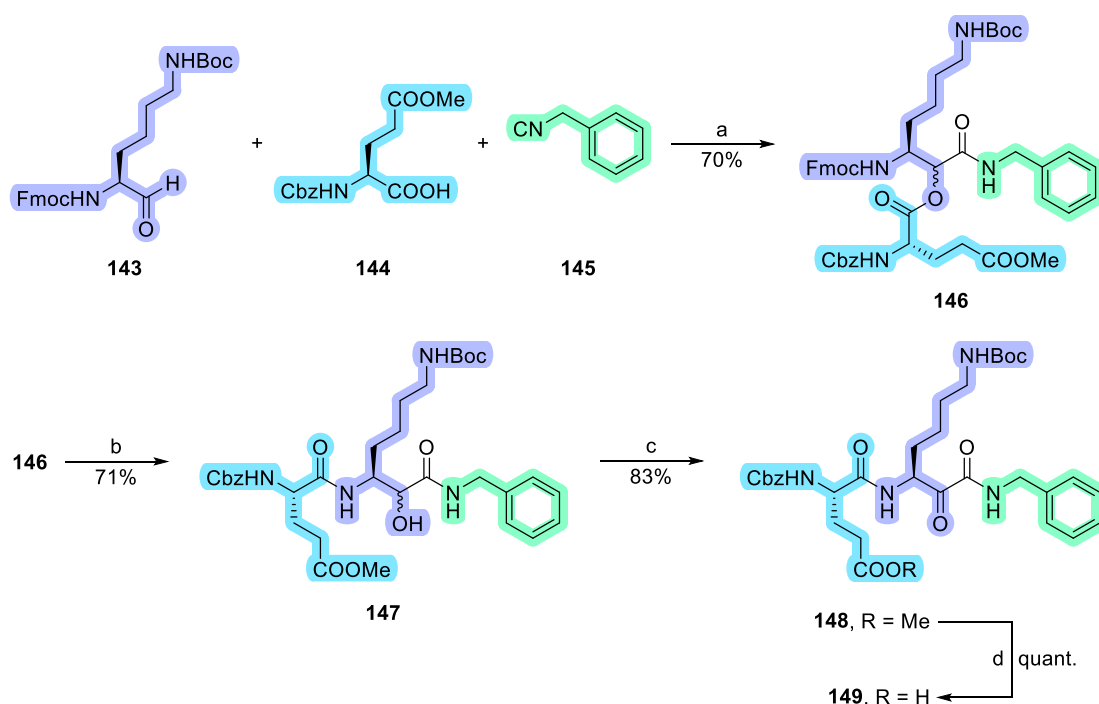
Table 3.3 Selected backbone modifications of macrocyclic inhibitors of human CatD

compound	structure	IC ₅₀ (nM)
117		2.5 ± 0.7
111		850 ± 310

3.2 Linker-Modified Gingipain Inhibitors for Probe Development

3.2.1 Preparation of Linker-Modified Lys-Gingipain Inhibitors

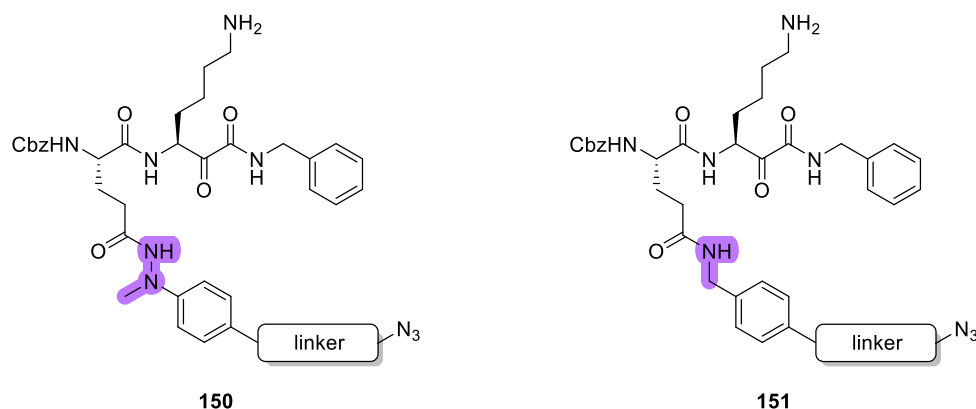
The design of the linker-modified Kgp inhibitor **KYT-36** was based on structural analysis of the enzyme crystal structure with the inhibitor, as described in **Chapter 1.3.3**. The inhibitor core was prepared by a PADAM reaction sequence which was introduced in detail in **Chapter 1.3.4**. The preparation of the peptidomimetic core of inhibitor **KYT-36** is described in **Scheme 3.13**.



Scheme 3.13 Preparation of **KYT-36** inhibitor core: (a) DCM, rt; (b) piperidine, DCM, rt; (c) Dess-Martin periodinane, DCM, rt; (d) NaOH, MeOH:H₂O, rt.

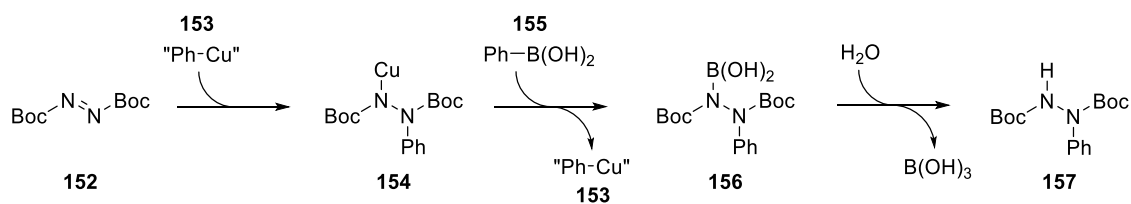
This work was carried out by a bachelor student under my supervision and experimental details can be found in her bachelor's thesis.¹³³ In the first step, Fmoc-Lys(Boc)-H (**143**), Cbz-Glu(OMe)-OH (**144**) and benzyl isonitrile (**145**) reacted in a Passerini reaction to yield intermediate **146** in a very good yield. In the next step, the Fmoc protecting group was removed using standard conditions and the acyl migration occurred spontaneously to yield α -hydroxyamide **147**. After oxidation to an α -ketoamide **148** using Dess-Martin Periodinane, the glutamate methyl ester was hydrolyzed to yield free acid **149**. The free acid was utilized for the attachment of the linker-containing fragment of the inhibitor. The original **KYT-36** inhibitor contains an *N*-methyl-*N*-

phenylhydrazine substituent at this position. We expected that this group might pose difficulties during the synthesis, but from the available SAR data we could not fully predict its importance for the inhibitory activity. Therefore, we designed two inhibitor analogues **150** and **151**, depicted in **Scheme 3.14**, a version containing the original *N*-methyl-*N*-phenylhydrazine moiety (**150**) along with a simplified version containing benzylamine moiety (**151**) instead of the hydrazine fragment.



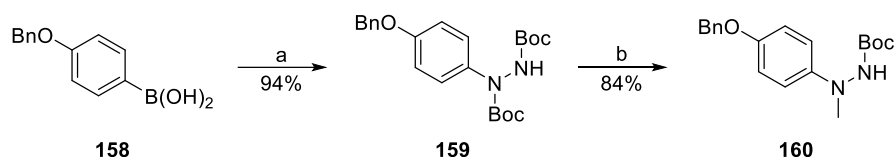
Scheme 3.14 Design of inhibitor analogues – the original design (**150**) and the simplified design (**151**)

The simplified version of the inhibitor **151** was synthesized by Tereza Gistrová under my supervision, while I focused on the synthesis of the version possessing the original design **150**. As expected, not many methods were found in the literature for preparation of the densely functionalized *N*-methyl-*N*-phenylhydrazine moiety and its analogues. Nevertheless, various phenylhydrazines can be prepared by nucleophilic addition of arylboronic acids to azodicarboxylates.^{134,135} According to the authors, this method is robust and tolerates a wide array of substrates. The proposed mechanism of the catalyzed nucleophilic addition to azodicarboxylates is described in **Scheme 3.15**. It starts with a transmetalation of a phenyl boronate **155** to yield a phenyl cuprate **153** that adds to azodicarboxylate **152** to yield intermediate **154**.¹³⁵ In the next step, another molecule of phenyl boronate **155** transmetalates to produce boronate intermediate **156** that finally hydrolyzes to a Boc-protected phenylhydrazine **157**.



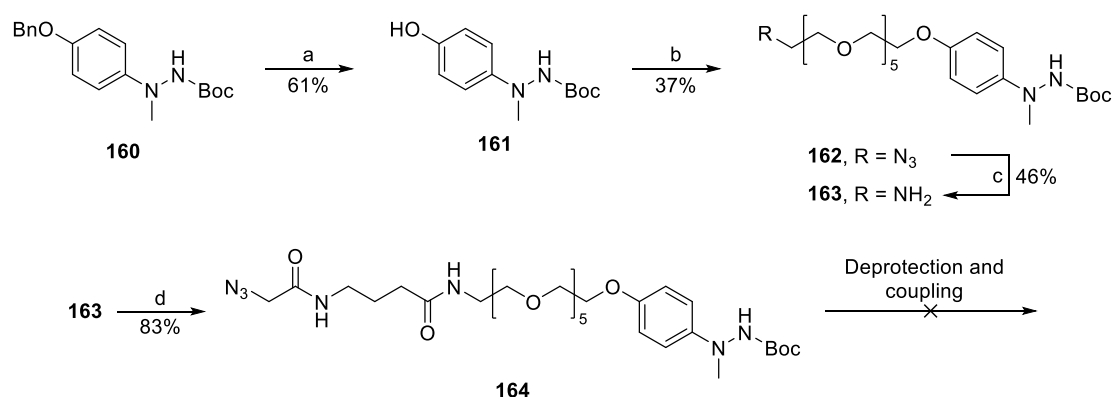
Scheme 3.15 Mechanism of the copper-catalyzed boronate addition to an azodicarboxylate¹³⁵

This method was employed and further modified by de Graaff et al. to yield substituted *N*₆-Boc-*N*_α-methyl-phenylhydrazine **160**, as described in **Scheme 3.16**.¹³⁶ The reduction of *N*_α-Boc-protected anilines to *N*_α-methylanilines by LiAlH₄ seems to be a well-known method commonly used in the literature,^{137–139} and combination of the two aforementioned methodologies by de Graaf et al. leads to an elegant synthetic route for the preparation of substituted *N*₆-Boc-*N*_α-methyl-phenylhydrazines.



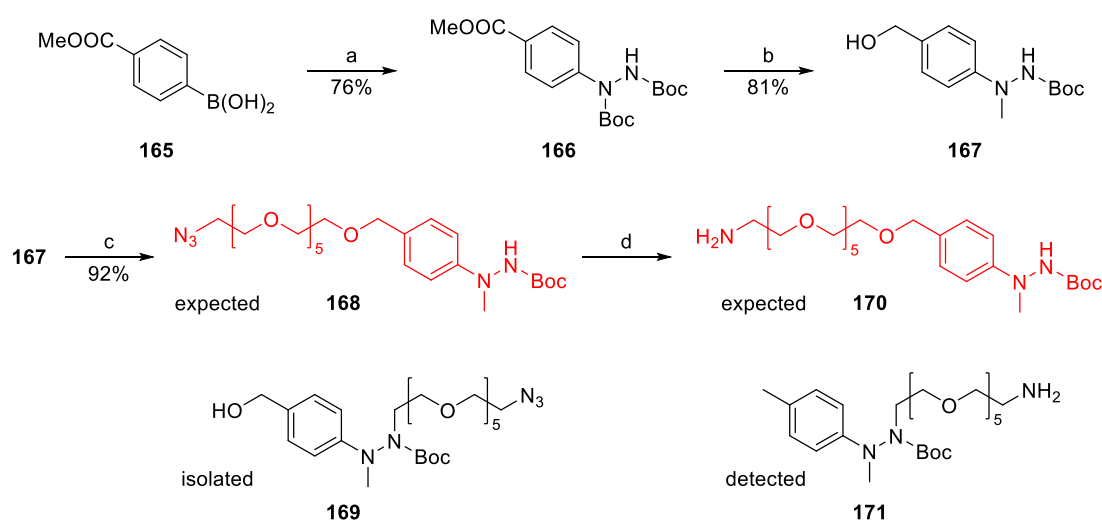
Scheme 3.16 Synthesis of *N*₆-Boc-*N*_α-methyl-phenylhydrazine motif¹³⁶: (a) DtBAD, Cu(OAc)₂, MeOH, rt; (b) LiAlH₄, THF, 0 °C to rt.

This method seemed ideal for preparation of the **KYT-36** inhibitor fragment possessing *N*-methyl-*N*-phenylhydrazine moiety along with an additional functional group for attachment of the linker. In the primary design, I decided to follow the exact synthetic pathway published by de Graaf et al., using the *para* hydroxy group as a site for linker attachment. Therefore, I prepared compound **160** according to **Scheme 3.16** and proceeded with a catalytic debenzoylation to obtain compound **161**. In the next step, phenol **161** was alkylated by a linker tosylate to obtain ether **162**, as described in **Scheme 3.17**. The last step of this sequence consisted of catalytic hydrogenation to convert azide **162** to amine **163**, which was then coupled with a terminal linker part N₃-Ac-GABA-OH to yield intermediate **164**. This fragment serves as a universal tag that we use for convenient quantification of the ligand content on the polymer using amino acid analysis (AAA).



Scheme 3.17 Preparation of linker-modified *N*-methyl-*N*-phenylhydrazine fragment **164** – phenol variant: (a) H₂, 10% Pd/C, AcOH, rt; (b) N₃-PEG-5-OTs, K₂CO₃, DMF, 70 °C; (c) H₂, 10% Pd/C, THF, rt; (d) N₃Ac-GABA-OH, HBTU, DIPEA, DMF, rt.

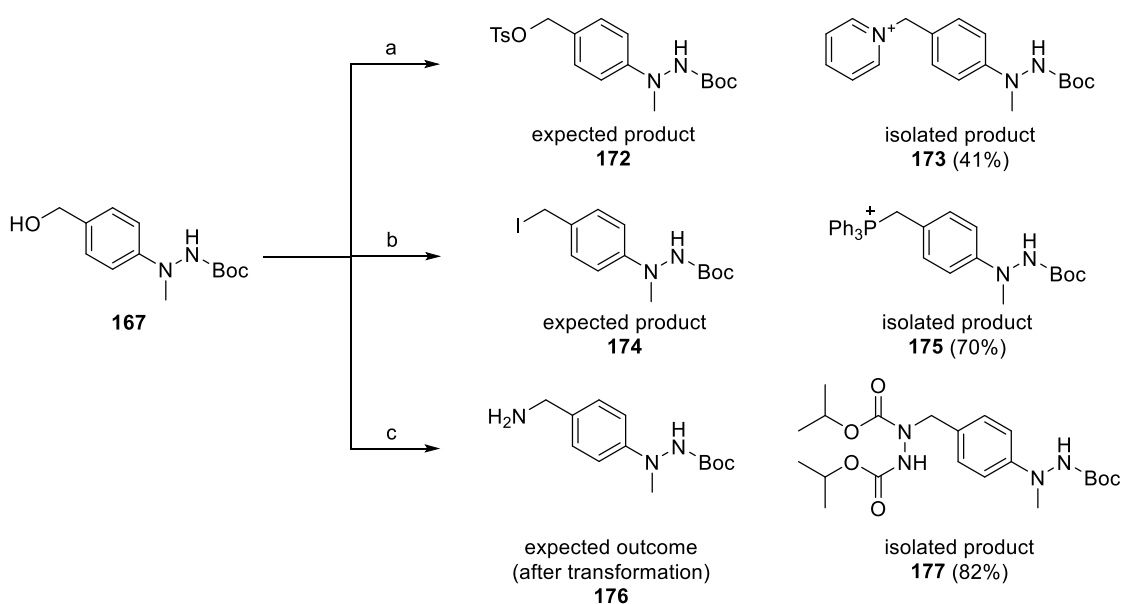
Generally, this reaction sequence was problematic, and yields were dropping significantly after the deprotection of the phenol group. In the final step of the sequence, deprotection of the Boc protecting group was accompanied by the formation of various byproducts, and despite several attempts and condition optimizations, coupling of this fragment with a carboxylic acid was not successful or very inefficient. Therefore, change of the synthetic strategy was necessary and I decided to exchange the phenol moiety for a benzylic alcohol to reduce the electron density of the aromatic core, as this seemed like a potential cause of instability and, therefore, low yields in the sequence shown in **Scheme 3.17**.



Scheme 3.18 Preparation of linker-modified *N*-methyl-*N*-phenylhydrazine fragment **170** – benzyl alcohol variant: (a) DtBAD, Cu(OAc)₂, MeOH, rt; (b) LiAlH₄, THF, rt; (c) N₃-PEG-5-OTs, KO^tBu, DMF, rt; (d) H₂, 10% Pd/C, THF, rt.

As a starting material for the preparation of the benzyl alcohol variant, commercially available boronate **165** was chosen (**Scheme 3.18**), possessing carboxylic acid methyl ester in the *para* position. This methyl ester seemed like an ideal precursor moiety for generating the benzyl alcohol in the LiAlH₄ reduction with no further protection/deprotection steps necessary. The addition of the boronate **165** to an azodicarboxylate proceeded with a very good yield and the intermediate **166** was reduced to the key benzyl alcohol intermediate **167** without any significant problems or side reactions. In the next step, alkylation of the benzylic alcohol with a PEG tosylate was expected to generate ether **168**. After analysis by MS and NMR, the reaction initially seemed to yield the desired ether. Proceeding with the hydrogenation step, reaction yielded multiple byproducts, one of them with a mass loss of 16 units compared to the desired product **170**. After thorough analysis of the reaction mixture, this byproduct was

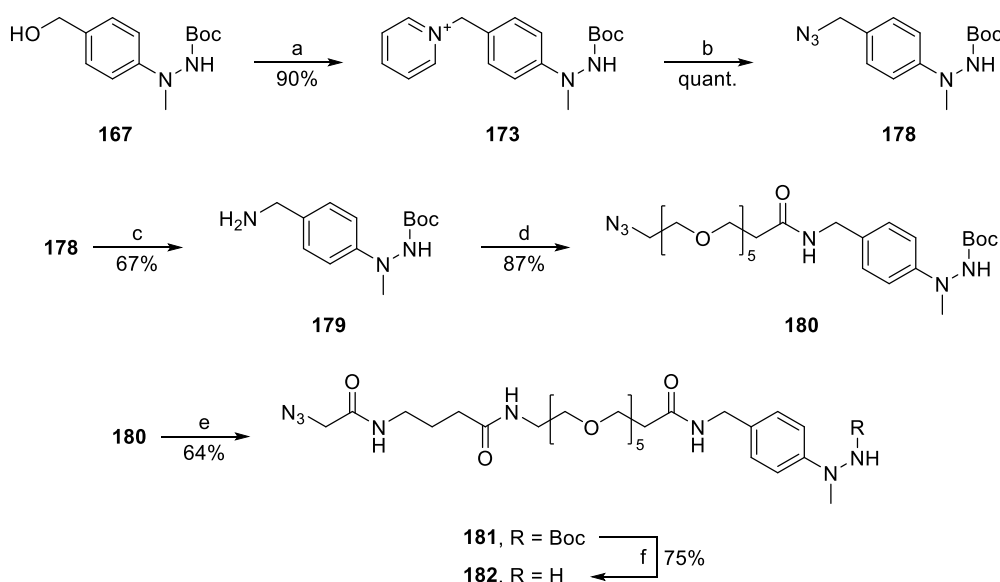
identified as compound **171**, leading us to the conclusion that the alkylation occurred primarily at the carbamate nitrogen (**169**) rather than the benzylic alcohol. This finding was rather surprising as one would generally expect higher reactivity from a primary benzylic alcohol rather than a Boc-protected amine/hydrazine. Analysis of the pK_a values confirmed that the acidity of the two functional groups is comparable and could explain the cross-reactivity. Based on these findings, changing the strategy appeared more reasonable than a time-demanding optimization that would not necessarily lead to the desired product.



Scheme 3.19 Modification of benzyl alcohol **167**: (a) TsCl, pyridine, rt; (b) I₂, PPh₃, DCM, rt; (c) DIAD, PPh₃, phthalimide, THF.

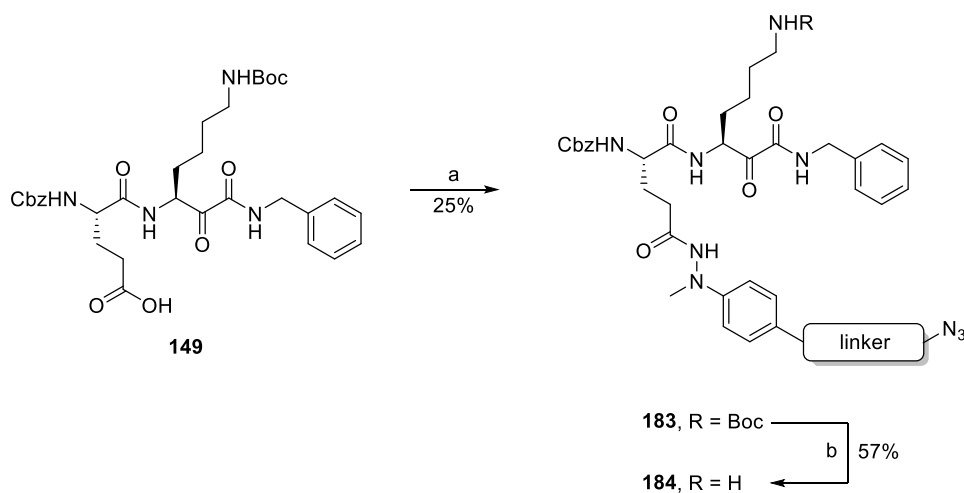
In this case, I decided to take advantage of the intermediate **167** and convert the benzylic alcohol to a benzylamine moiety. A primary amine should undoubtedly offer higher reactivity compared to other functional groups present in the molecule and allow for efficient linker attachment. Nevertheless, compound **167** continued to showcase its unconventional reactivity as shown in **Scheme 3.19**. Three different methods were employed for modification of benzyl alcohol **167**, such as conversion to a tosylate, a variant of the Appel reaction, and the Mitsunobu reaction. In all three cases, the major isolated compound was not the expected “textbook” product of the reaction. After conversion to tosylate **172**, an additional substitution reaction with pyridine – which acted as a base in the reaction – took place and yielded pyridinium **173**. In the case of the Appel reaction, different conditions were tried, but the most abundant product observed was compound **175**, indicating that activation of the benzylic alcohol was occurring, but PPh₃ acted as a

nucleophile and attacked the activated intermediate. Based on this consideration, I decided to try the reaction in the presence of excess sodium azide to outcompete the reaction with PPh_3 . This modification was partially successful, with a benzylic azide being formed predominantly. However, the reaction never reached high conversion and was ultimately abandoned. The Mitsunobu reaction also did not yield the expected product; instead, a well-known byproduct **177** caused by the attack of the hydrazine intermediate to the activated alcohol was formed. This is a common process that occurs in the case of slow protonation, due to lower acidity of the acid component used in the Mitsunobu reaction. Even though phthalimide is often used as acid in this reaction setup, it did not provide the desired outcome in this case. After careful analysis of all these results, I decided to test whether compound **173** would be reactive enough to be attacked by an azide anion in a nucleophilic substitution. Indeed, the pyridine turned out to be a sufficient leaving group for this reaction and heating compound **173** in DMF in the presence of NaN_3 led to formation of the desired azide **178** with a quantitative yield. After confirming this hypothesis, reaction conditions for the formation of the pyridinium intermediate **173** were optimized to yield the intermediate in an excellent yield, as shown in **Scheme 3.20**. Next, azide **178** was reduced to amine **179** via catalytic hydrogenation and the amine was coupled with an NHS ester of a PEG linker to yield derivative **180**. In the last step of this sequence, azide **180** was reduced and the terminal part of the linker, that serves as a tag for AAA, was attached to yield compound **181**.



Scheme 3.20 Preparation of linker-modified *N*-methyl-*N*-phenylhydrazine fragment **182** – benzylamine variant: (a) TsCl , pyridine, rt; (b) NaN_3 , DMF, 80 °C; (c) H_2 , 10% Pd/C, THF, rt; (d) $\text{N}_3\text{-PEG-5-CH}_2\text{CH}_2\text{COOSu}$, DIPEA, DMF; (e) (i) H_2 , 10% Pd/C, THF, rt; (ii) $\text{N}_3\text{-Ac-GABA-OH}$, HBTU, DIPEA, DMF, rt; (f) TFA, DCM, rt.

Next, the Boc protecting group was removed and the linker-modified *N*-methyl-*N*-phenylhydrazine fragment **182** was attached to the inhibitor core **149** via amide coupling, as depicted in **Scheme 3.21**. Yields in this step were generally poor, even in the case of the simplified version **151** prepared by Tereza Gistrová. One of the issues of this coupling reaction is the cyclization of the γ -carboxylate of the glutamate with the α -amino group. Although this amine is protected, and its reactivity is significantly reduced, intramolecular cyclization to form a 5-membered ring proceeds rapidly and competes with the desired coupling reaction. This process is described in the literature and generally requires strong activation reagent (such as PyBrOP) and strong bases to proceed.¹⁴⁰ However, this is probably strongly sequence dependent and might change significantly with higher molecular weight and, therefore, lower reactivity of the amine coupling partner. In our case, we were able to detect two different compounds with corresponding mass, which suggests that even cyclization with the α -amino group of the adjacent lysine proceeded well to form a 6-membered ring. Authors of the above-mentioned paper used an excess of the amino component to outcompete the intramolecular cyclization, but this was not a viable option in our case. Therefore, we performed a small conditions screen to identify optimal reagents for this reaction. Surprisingly, the use of TSTU as a mild activation reagent did not suppress this reaction significantly. Next, 2,4,6-collidine was used as a base and this change affected the course of the reaction strongly. The intramolecular cyclization was suppressed, while coupling with the amino partner proceeded efficiently. HBTU and PyBOP were tested, as PyBOP does not require activation of the acid without presence of the amino component, but both reagents showed similar results. Ultimately, HBTU was selected as the reagent of choice for this reaction. After the coupling reaction, the Boc protecting group was removed in the final step to yield KYT-36 analogue **184**. This inhibitor was later tested in a Kgp inhibition assay to confirm its inhibitory activity and was used as a starting point for the development of an assay for Kgp quantification and novel inhibitor screening.



Scheme 3.21 Final inhibitor **184** assembly: (a) (i) HBTU, 2,4,6-collidine, DMF; (ii) compound **182**, rt; (b) TFA, DCM, rt.

3.2.2 Evaluation of Linker-Modified Lys-Gingipain Inhibitors

Two variants of linker-modified Kgp inhibitor **KYT-36** were prepared, as discussed in **Chapter 3.2.1** (structures can be seen in **Scheme 3.14**). Both the simplified version **151** prepared by T. Gistrová and inhibitor **184** possessing the original *N*-methyl-*N*-phenylhydrazine moiety were tested in a chromogenic substrate assay, along with the parental inhibitor **KYT-36**. This assay was optimized and performed by Adéla Marcalíková from Jan Konvalinka's group at IOCB. As shown in **Figure 3.1**, **KYT-36** exhibited K_i of 8.6 nM, while both linker-modified inhibitors showed slightly reduced inhibitory potency. In the case of the simplified version (**151**), a three-fold drop in inhibitory potency was observed ($K_i = 21.5$ nM), while the original version (**184**) showed a five-fold drop in potency ($K_i = 39.4$ nM). Interestingly, the simplified version (**151**) reached a similar K_i as the original version (**184**) and, therefore, we confirmed that the *N*-methyl-*N*-phenylhydrazine moiety could be replaced with a benzylamine moiety without significant loss in inhibitory potency.

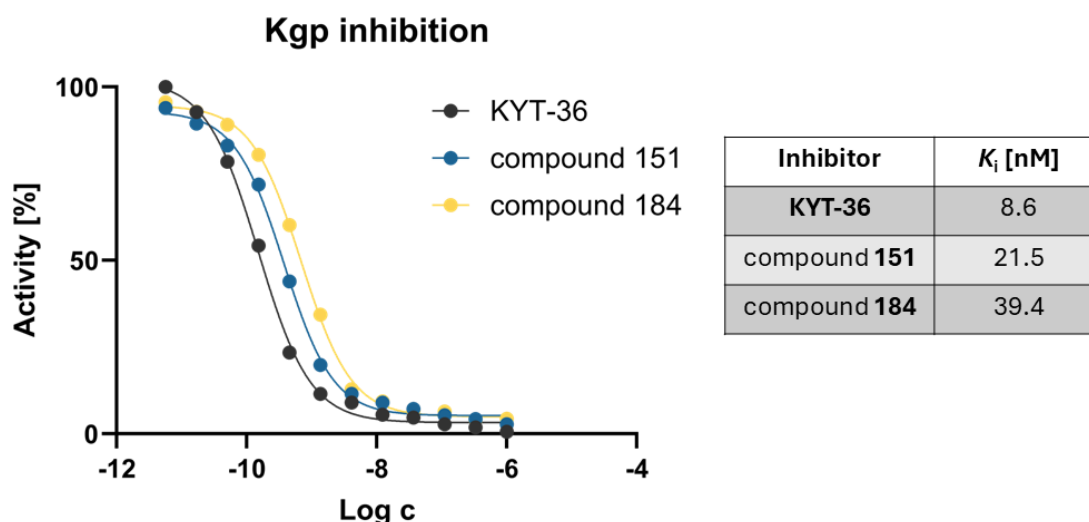
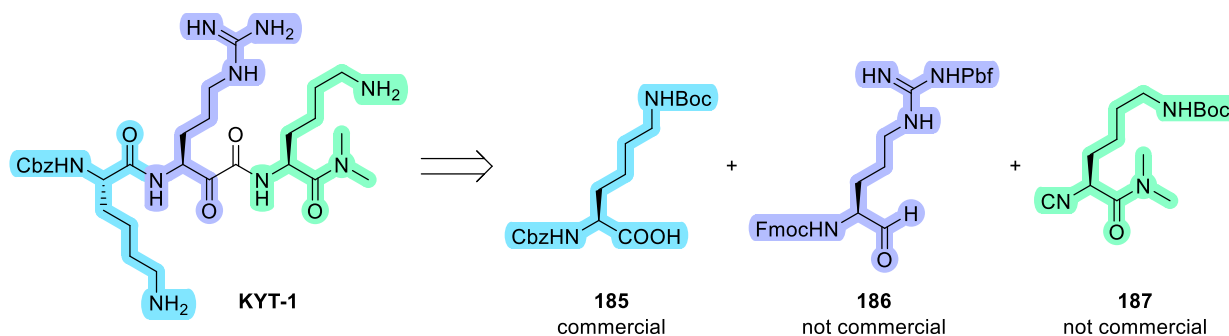


Figure 3.1 Kgp inhibitors comparison: XY plot with connecting line showing enzyme activity dependence on inhibitor concentration (left) and corresponding K_i values (right). 330 ng/ul of Kgp was incubated in 200 μ l reaction buffer (200 mM Tris-Cl, pH 7.6, 150 mM NaCl, 5 mM CaCl₂, 20 mM cysteine and 0.05% Tween-20) with inhibitor (0.005 – 1000 nM dilution series) in 37 °C for 5 min. The reaction was started by the addition of Tos-Gly-Pro-Lys-4-nitroanilide chromogenic substrate (100 μ M final concentration, Merck). Change of absorbance in time (405 nm wavelength) was recorded over 30 minutes using Tecan Spark. 50% inhibitory concentration was determined using GraphPad Prism and the inhibition constant (K_i) values were calculated using the Cheng-Prusoff equation. Representative experiment is shown.

As discussed in **Chapter 1.3.3**, a linker-modified inhibitor suitable for probe development should retain high inhibitory potency. Small changes in K_i , as observed for **151** and **184**, can be expected after significant structural changes to an optimized inhibitor. Nevertheless, linker-modified inhibitors with double-digit nanomolar potency should be suitable for the development of a functional probe. Currently, A. Marcalíková is developing and optimizing a versatile assay for the detection and quantification of Kgp, based on the DIANA technology.^{69,70} So far, promising preliminary results were observed, but further optimization is ongoing, and the results have yet to be reported.

3.2.3 Preparation of Linker-Modified Arg-Gingipain Inhibitors

The Rgp inhibitor **KYT-1** that we selected for the modification with a linker comes from the same inhibitor series as **KYT-36**. They share many structural features; therefore, the synthetic strategy was rather similar. Again, the most important step involves the Passerini reaction and the PADAM sequence. Nevertheless, there were many significant differences in the synthesis of these two compounds. The first major difference was that, in the case of **KYT-36** synthesis, all components for the Passerini reaction were commercially available, and most of the synthetic efforts were made in the preparation of the linker-containing fragment and attaching it to the inhibitor core. In the case of **KYT-1**, the situation was quite different with only one of the components for the Passerini reaction being commercially available, as depicted in **Scheme 3.22**.

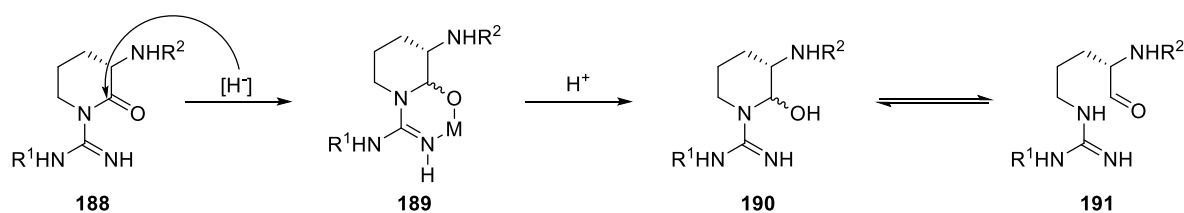


Scheme 3.22 Retrosynthetic analysis of the KYT-1 inhibitor

The acid component (**185**) necessary for the Passerini reaction is a simple lysine derivative commonly used in peptide synthesis. The aldehyde component (**186**) is an aldehyde derived from arginine. These compounds are easily accessible from arginine but are generally not commercially available. The biggest challenge was the preparation of the isonitrile component (**187**) derived from lysine. Preparation of this component required several steps and careful preparation and handling of the isonitrile due to its relatively low stability. Regarding the protecting groups strategy, similar to **KYT-36**, **KYT-1** contains the hydrogenation-labile Cbz group as part of the inhibitor. Therefore, I decided to use acid-labile protecting groups for both lysine residues and the guanidine moiety, which would be deprotected at the end of the inhibitor synthesis. Another orthogonal protecting group was necessary for the aldehyde α -amino group, which would allow selective deprotection of this residue after the Passerini reaction to trigger the acyl migration. For this purpose, I selected the Fmoc protecting group, which fulfilled all the

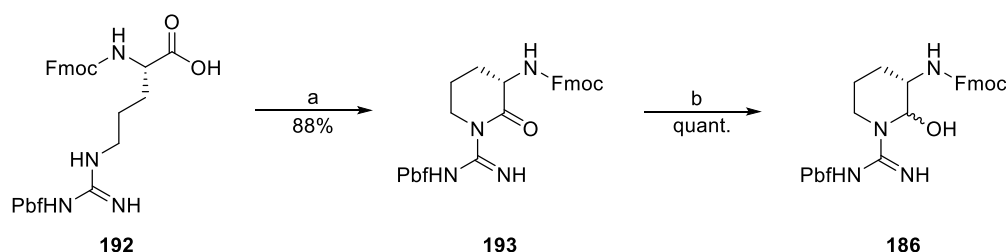
orthogonality requirements and allowed for mild deprotection conditions. Generally, this protecting group strategy was analogous to the one used in synthesis of **KYT-36**.

Arginine-derived aldehydes can be prepared from arginine via a lactam intermediate with the guanidine moiety.^{141–143} These lactams exhibit behavior of a Weinreb's amide and are therefore only reduced to their respective aldehydes, even when excess of LiAlH_4 or another reducing agent is present. We can assume that this is caused by similar coordination effects as in the case of the Weinreb's amide, as proposed in **Scheme 3.23**. The metal ion in intermediate **189** is coordinated by the newly formed hydroxy group and the guanidine nitrogen, preventing the addition of a second hydride equivalent. Only after hydrolysis of the reaction mixture, is the hydroxy group protonated, and the metal ion is released to form hemiaminal **190**, which exists as two distinct diastereoisomers. This hemiaminal is in equilibrium with the aldehyde **191** and these compounds can generally be detected in all three forms. This can be confirmed by the presence of multiple peaks during HPLC analysis and purification. It can also be seen in the NMR spectra and often complicates their interpretation.



Scheme 3.23 Proposed mechanism of arginine lactam **188** reduction to aldehyde **191**

As shown in **Scheme 3.24**, in the first step of the synthesis, arginine derivative **192** was cyclized to lactam **193** using HATU as the activation reagent. This intermediate was purified and isolated in very good yield. Then, it was reduced to arginine aldehyde/hemiaminal **186** by LiAlH_4 .

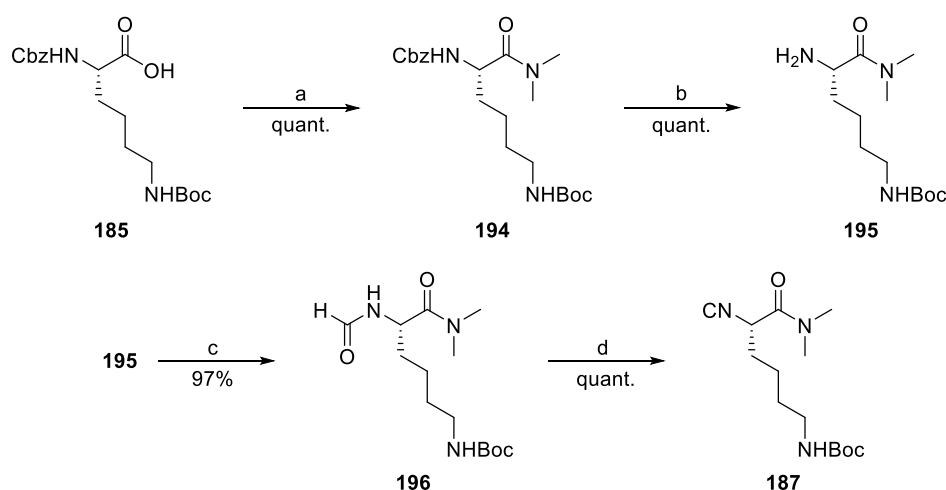


Scheme 3.24 Preparation of arginine aldehyde **186**: (a) HATU, DIPEA, DMF, rt; (b) LiAlH_4 , THF, $-20\text{ }^\circ\text{C}$.

Due to the existence of multiple forms of this compound, which complicate separation from other species, it was used in the next step without chromatographic

purification. Extraction with Rochelle's salt was performed to isolate compound **186** as the major product of the reaction, in purity sufficient for the Passerini reaction.

Preparation of the isonitrile component **187** was done in four steps from a commercially available lysine derivative **185**, as shown in **Scheme 3.25**. In the first step, dimethylamide was installed in a simple peptide coupling reaction using HATU as the coupling reagent to yield intermediate **194**. In the next step, α -amino group was deprotected via catalytic hydrogenation, and amine **195** was converted into formamide **196** in another coupling reaction using a mixture of formic acid and acetic anhydride. Formamides are commonly used as isonitrile precursors, as they can be easily dehydrated to yield their respective isonitriles. This reaction was performed according to procedures available in the literature.^{144,145}



Scheme 3.25 Preparation of lysine-derived isonitrile **187**: (a) Me_2NH , HATU, DIPEA, DMF, rt; (b) H_2 , 10% Pd/C, EtOH, rt; (c) HCOOH , Ac_2O , DCM, rt; (d) POCl_3 , DIPA, DCM, $-20\text{ }^\circ\text{C}$.

Numerous experiments were necessary to optimize the formation and isolation of isonitrile **187**. Consistent with expectations, relatively low stability of isonitrile **187** was confirmed and, therefore, storage and purification of this compound was avoided. It was possible to obtain ^1H NMR spectra of a pure sample of compound **187**, but decomposition was observed during storage in both solution and solid state. As shown in **Figure 3.2**, the compound was obtained in high purity after the reaction. This sample was then evaporated and stored at $5\text{ }^\circ\text{C}$ overnight in solid state. Upon remeasuring the NMR spectra, partial decomposition was observed with new signals appearing in the spectra. POCl_3 was selected as the dehydration reagent of choice due to the fast and clean formation of the product, along with byproducts that could easily be removed by extraction without the need for chromatography. The reaction was also highly sensitive to the quality of the

reagents, particularly POCl_3 . After the reaction conditions were optimized to achieve full conversion, the reaction was performed on a larger scale and isonitrile **187** was used in the next step without further purification.

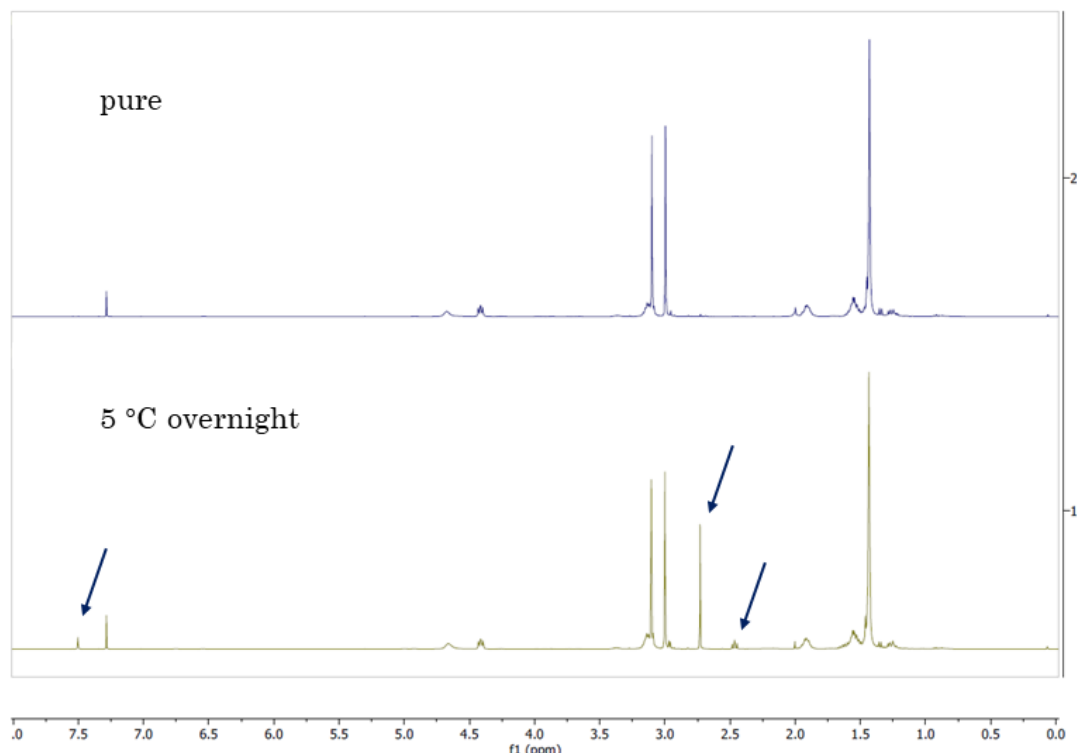
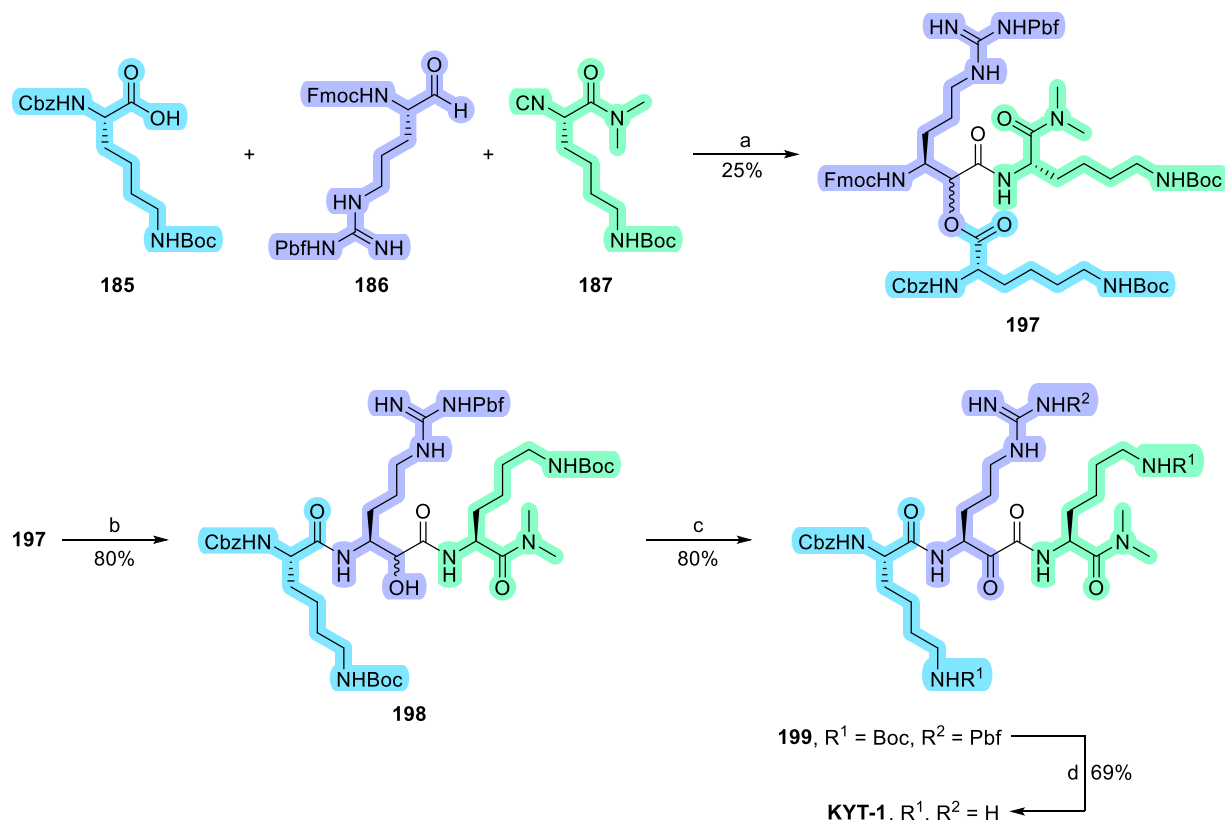


Figure 3.2 NMR spectra of a pure isonitrile **187** (top) and a sample after overnight storage at 5 °C (bottom). New peaks are highlighted using blue arrows.

The next step of the **KYT-1** preparation was the Passerini reaction, according to **Scheme 3.26**. Interestingly, in this case the reaction proceeded with a poor yield (25%) compared to the Passerini reaction used for the preparation of the **KYT-36** (70%) inhibitor core. Both experiments were performed multiple times, and the yields were reproducible. One could argue that low yields in the case of **KYT-1** are caused by using non-commercial and potentially impure materials. In my opinion, the situation is much more complex, with the root cause lying in the discrepancy between high/low reactivity of individual components. The isonitrile **187** exhibits stability issues, as discussed above, caused by the high reactivity of the intermediate. This suggests that the compound might be susceptible to alternative reaction pathways leading to unwanted byproducts. On the other hand, the behavior of aldehyde **186** is specific due to presence of the guanidine moiety, as shown in **Scheme 3.23**. Cyclization of aldehyde **186** to the respective hemiaminal form stabilizes it, thereby reducing the reactivity of the aldehyde. Combination of the two factors might have a significant impact on the yields of the Passerini reaction. Nevertheless, by increasing

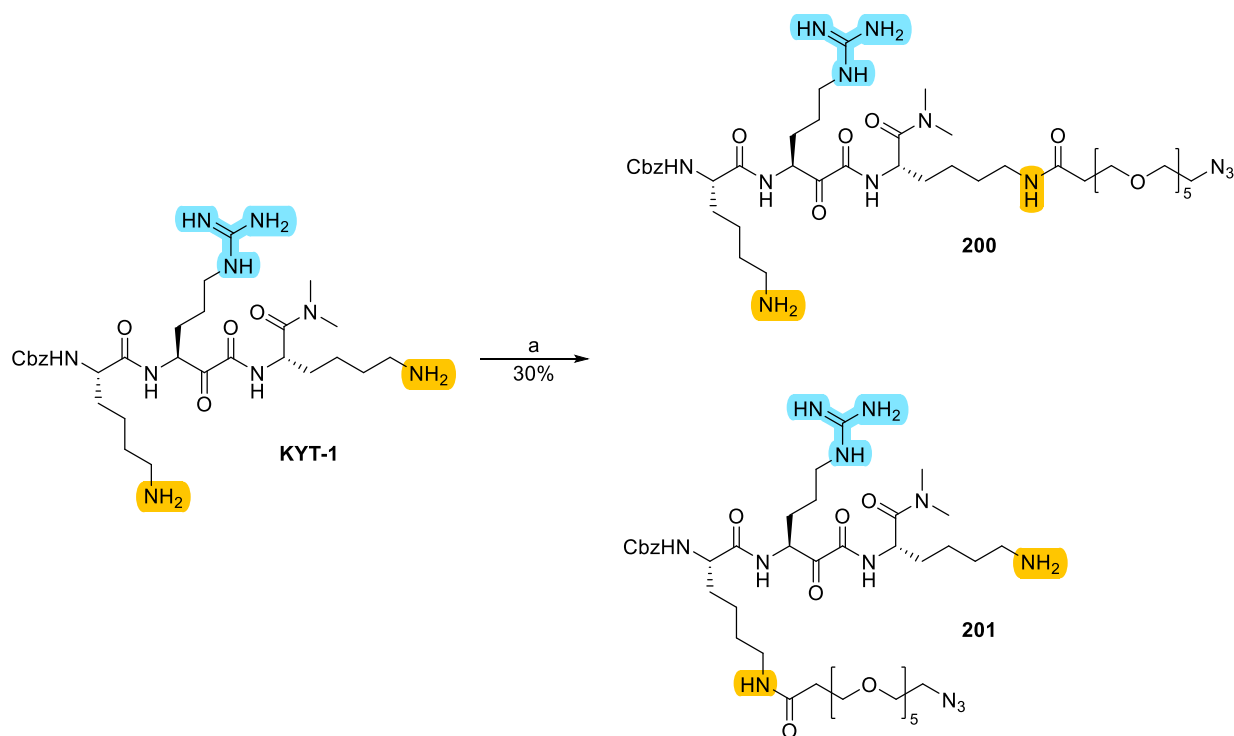
the reaction scale it was possible to obtain sufficient quantities of intermediate **197**. Subsequent reaction steps proceeded with very good yields, similar to the preparation of **KYT-36**. Finishing the PADAM sequence after deprotection of the Fmoc group yielded intermediate **198**. In the next step, oxidation to α -ketoamide **199** was performed using Dess-Martin periodinane and protecting groups were removed to yield the final Rgp inhibitor **KYT-1**.



Scheme 3.26 Preparation of **KYT-1** inhibitor: (a) DCM, rt; (b) piperidine, DCM, rt; (c) Dess-Martin periodinane, DCM, rt; (d) TFA, DCM, rt.

Next, I proceeded with linker modification of the **KYT-1** inhibitor. **KYT-1** possesses three nucleophilic functionalities that could potentially react with an activated carboxylic acid derivative. These moieties are highlighted in **Scheme 3.27**. Two lysine residues (orange) can be modified selectively in the presence of the arginine residue (light blue) due to significant differences in basicity ($pK_a(\text{Lys}) = 10.7$, $pK_a(\text{Arg}) = 12.1$) and, therefore, nucleophilicity. Generally, using a less reactive carboxylic acid derivative and a weaker base for an amide coupling reaction leads to highly selective modification of lysine residues. We can assume that the arginine residue plays a key role in the inhibitory potency towards Rgp, so avoiding modification at this position is desired. Based on these considerations, an NHS ester linker derivative was chosen as a coupling partner for its

mild reactivity, along with 2,4,6-collidine as a weak base ($pK_a = 7.4$). This strategy yielded two isomers of **KYT-1** modified at P1' (**200**) and P2 (**201**) lysine residues (based on the protease substrate nomenclature, P1' refers to the C-terminal lysine and P2 to the N-terminal lysine), as shown in **Scheme 3.27**.



Scheme 3.27 Active ester modification of **KYT-1** inhibitor: (a) $\text{N}_3\text{-PEG-5-CH}_2\text{CH}_2\text{COOSu}$, 2,4,6-collidine, DMF, rt. The amine groups of **KYT-1** are highlighted in orange and the guanidine group is highlighted in light blue.

The composition of the reaction mixture, in the case of “random” modification at multiple reactive sites, will inevitably depend on multiple factors. Besides the two desired products **200** and **201**, the presence of the starting material (**KYT-1**) and a byproduct with modification at both reactive sites can be expected. The relative reactivity of the individual residues, which might also depend on steric factors, will influence the product ratio along with the ratio between individual reactants and the reaction setup. To maximize the formation of the two desired products, the experimental procedure was designed carefully. First, 0.7 equivalents of the active ester were used to achieve a balance between mono- and bis-modification. While the starting material (**KYT-1**) can be recovered and used in another experiment, the bis-substituted product cannot be further exploited. In order to limit the formation of the bis-substituted byproduct, a linear syringe pump was used to achieve slow and controlled addition of the activated ester. After the active ester was consumed, the reaction mixture was analyzed using UPLC-MS. A representative

chromatogram can be seen below in **Figure 3.3**. Consistent with expectations, two distinguishable isomers of the mono-modified product and small amount of the bis-modified byproduct were detected using UPLC-MS. Relative ratio of these compounds was approximately 0.35:0.50:0.15 (**KYT-1**:mono-modification, both isomers:bis-modification), based on integration of the peaks in the photodiode array (PDA) chromatogram, which also corresponded with the isolated quantities. Thanks to difference in retention between the two mono-modified isomers, they were separated quantitatively using HPLC, and both isomers were isolated in sufficient quantities and analytically pure.

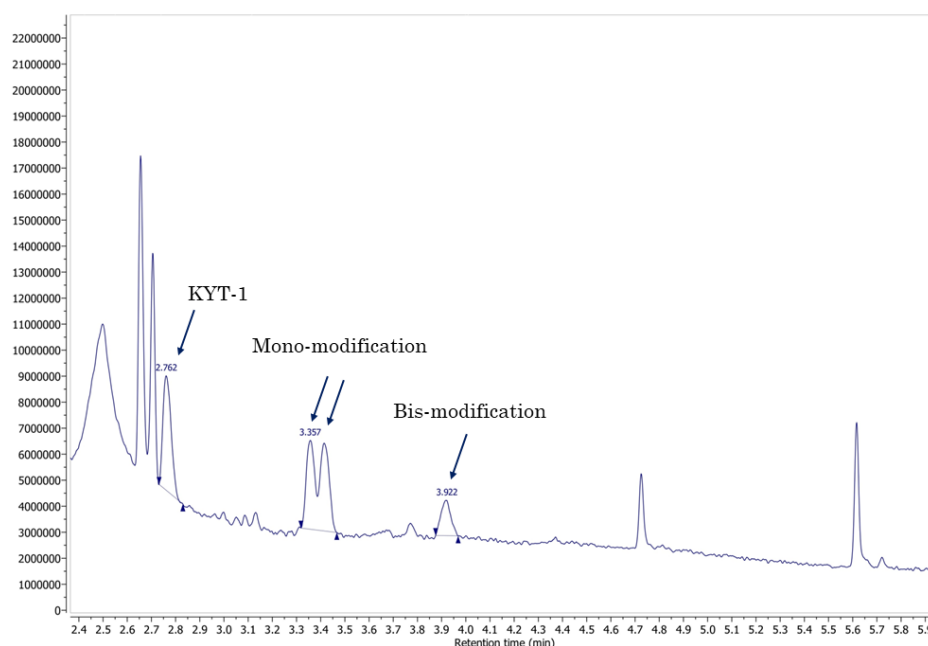


Figure 3.3 UPLC chromatogram of a crude reaction mixture after active ester modification of **KYT-1**. Relevant peaks are highlighted using blue arrows, other peaks are reagents or instrument impurities. Chromatogram shows signal from PDA detector, peaks were assigned based on MS signals.

After successful separation of the two product isomers, I aimed to assign the structure and, therefore, position of the linker for both isomers. NMR was selected as the method of choice for structure assignment and help with the analysis was provided by Dr. M. Dračinský. This analysis was rather complex and is not discussed in this thesis in further detail. Briefly, using two-dimensional NMR spectra signals corresponding to C-terminal lysine and surrounding CH(α) and CH₂ groups were assigned. In the case of one isomer, correlation was confirmed between CH₂(ϵ) protons of the C-terminal lysine and the CO carboxylic amide group belonging to the linker. This served as direct proof of linker position for the faster-eluting isomer (**200**) and the slower-eluting isomer was, therefore, assigned as compound (**201**). Both isomers were then tested for their inhibition activity

towards Rgp and used for the development of a DIANA-based assay for target detection, quantification and screening of libraries in search of new potential Rgp inhibitors.

3.2.4 Evaluation of Linker-Modified Arg-Gingipain Inhibitors

Analogous to Kgp, two linker-modified derivatives of Rgp inhibitor **KYT-1** were synthesized and tested for inhibition. In this case, inhibitors **200** and **201** differ in linker attachment position, as shown in **Scheme 3.27**. These compounds were tested in a chromogenic substrate assay (analogous to Kgp) performed by A. Marcalíková. In our hands, **KYT-1** exhibited K_i of 3.3 nM towards RgpB, as shown in **Figure 3.4**. Both linker-modified **KYT-1** analogues showed an approximately fifteen-fold increase in K_i . Although the increase in K_i values was more significant than in the case of Kgp, the K_i values stayed in the high-affinity nanomolar range. This confirmed our hypothesis that the lysine residues of **KYT-1** were suitable for the linker attachment. Interestingly, the two isomers exhibited almost identical inhibitory potency. Similar to Kgp, a platform for Rgp detection, quantification, inhibitor evaluation and chemical library screening based on these compounds is currently being developed.

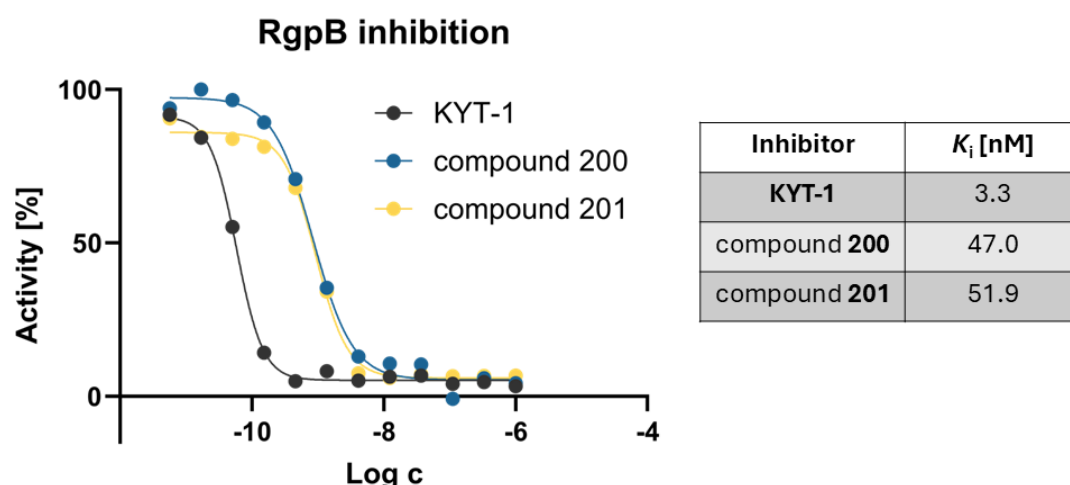
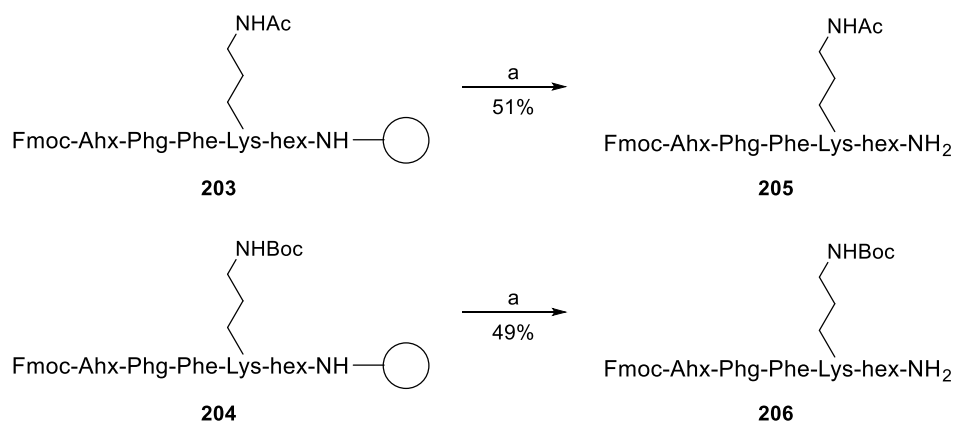


Figure 3.4 Rgp inhibitors comparison: XY plot with connecting line showing enzyme activity dependence on inhibitor concentration (left) and corresponding K_i values (right). 75 ng/ul of RgpB was incubated in 200 μ l reaction buffer (200 mM Tris-Cl, pH 7.6, 150 mM NaCl, 5 mM CaCl₂, 20 mM cysteine and 0.05% Tween-20) with inhibitor (0.005–1000 nM dilution series) in 37 °C for 5 min. The reaction was started by addition of Na-benzoyl-L-arginine-4-nitroanilide chromogenic substrate (100 μ M final concentration, Merck). Change of absorbance in time (405 nm wavelength) was recorded over 30 minutes using Tecan Spark. 50% inhibitory concentration was determined using GraphPad Prism and the inhibition constant (K_i) values were calculated using the Cheng-Prusoff equation. Representative experiment is shown.

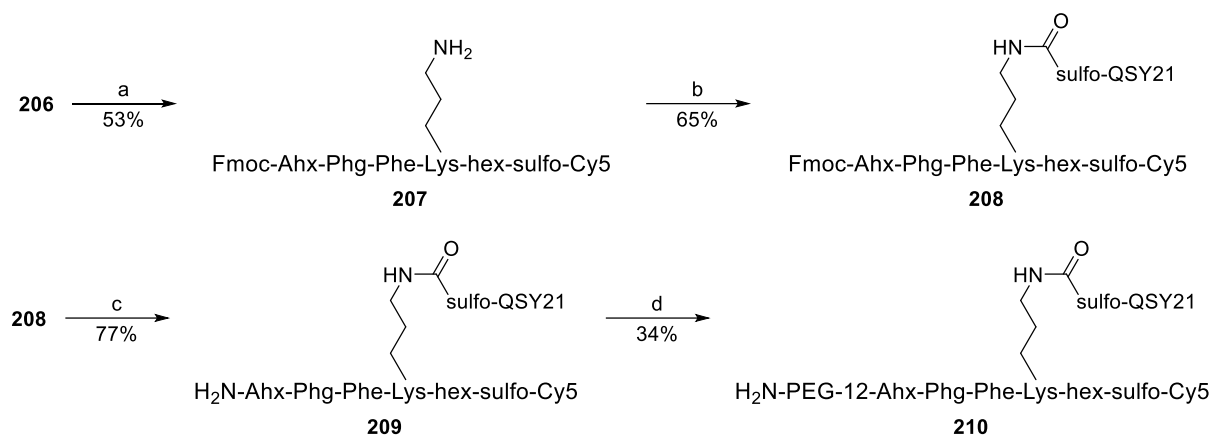
Next, specific conditions were introduced for the peptide coupling of the phenylglycine (Phg) derivative. Phg is prone to racemization during the coupling reaction and Fmoc deprotection due to the proximity of the aromatic core. In the case of deprotonation in the α -position, which leads to racemization, the anion is stabilized by conjugation over the aromatic core. This increases acidity of the α -position and accelerates racemization in basic conditions. Originally, the standard conditions mentioned above (HBTU/2,4,6-collidine) were used and led to full racemization at this position. Liang et al. have studied this process on several model peptides and optimized reaction conditions to minimize it.¹⁴⁶ Multiple coupling reagents were able to suppress racemization, and we decided to use HATU/2,4,6-collidine (3 eq./6 eq., respectively). Additionally, we used 2-methylpiperidine, a more sterically hindered base, instead of piperidine for Fmoc deprotection of Phg. In the last step, 6-(Fmoc-amino)hexanoic acid was attached to the N-terminus as a first part of the linker to yield bead-bound intermediates **203** and **204**. These intermediates were then cleaved off the resin using a mixture of DCM:HFIP to yield peptides **205** and **206**, as shown in **Scheme 3.29**. In both cases, the peptides were obtained in very good purity and were used without further purification. The yields were lower than typically expected for synthesis of short peptides on SPPS. This was probably caused by the low solubility of mono-Fmoc 1,6-diaminohexane (**202**) which affected the substitution of the resin.



Scheme 3.29 Cleavage of 6QC-like peptide fragments **205** and **206** from the resin: a. DCM:HFIP / 3:1, rt.

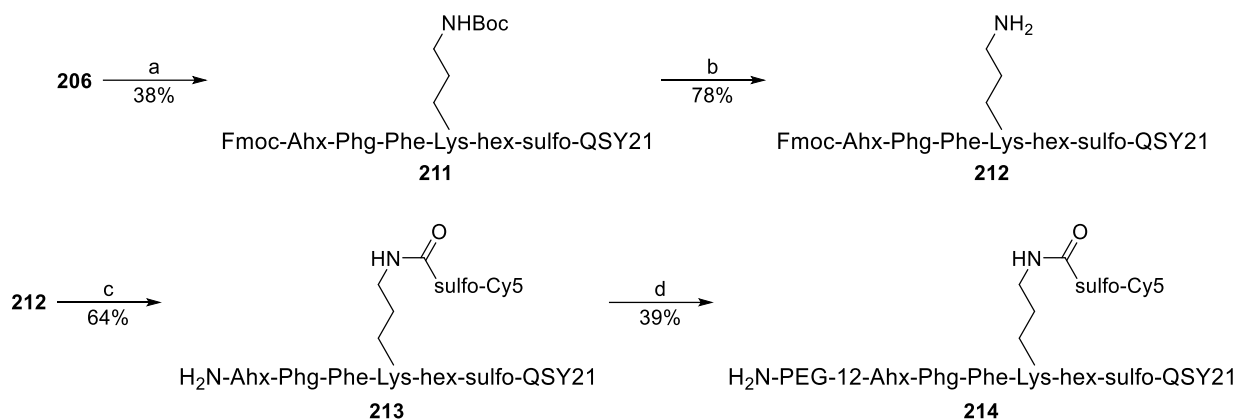
Next, peptide fragment **206** was used for the preparation of a QF probe precursor, as introduced in **Chapter 1.4.2**. First, according to **Scheme 3.30**, sulfo-Cy5 NHS ester was used to attach the fluorophore to the free amine at the C-terminus and the Boc protecting group in the Lys side chain was subsequently removed using TFA to yield intermediate **207** in average yield. Generally, yields reported in this chapter are rather

low due to complexity of the fluorophores and individual probe precursors. Typically, the compounds were purified and isolated after attachment of a new functional moiety and deprotection of the intermediate to streamline the synthetic process and reduce amounts of expensive fluorophores used. Nevertheless, this was not always possible depending on the crude purity of the intermediate. In the next step, sulfo-QSY21 NHS ester was used to attach the quencher to the Lys side chain to yield intermediate **208**. In this case, excess of sulfo-QSY21 overlapped with amine **209** during purification and, therefore, purification was performed in both steps. After the Fmoc-protected precursor **208** was isolated, the Fmoc group was removed using 1% DBU in DMF to yield amine **209**. While Boc deprotection proceeded without any severe side reactions, multiple factors needed to be taken into account in the case of Fmoc deprotection. First, fluorophores consist of a complex system of conjugated double bonds which can be reactive towards nucleophilic bases commonly used for the removal of the Fmoc protecting group. Second, as discussed earlier in this chapter, Phg is sensitive towards base-catalyzed racemization and, therefore, Fmoc deprotection reactions should be followed carefully and quenched as soon as possible after achieving sufficient conversion. Originally, I attempted the Fmoc deprotection using the standard 20% piperidine in DMF solution. This led to a change in the color of the reaction mixture from the original blue to green and indicated structural changes in the fluorophore moieties. Byproducts of green color were observed during purification and yields of the transformation were low (less than 50%). Therefore, I performed a literature search for alternative Fmoc deprotection reagents and found many reports of 1–2% DBU in DMF used for the reaction.^{146,147} Interestingly, a review by Li et al. mentions that DBU also efficiently reduces racemization of sensitive amino acids.¹⁴⁷ This was rather surprising due to high basicity of DBU, which is probably counterbalanced by the low concentration used and high steric demands of the reagent. Based on this information, DBU seemed like an ideal reagent for our use case and the very good yield in the preparation of amine **209** confirmed this hypothesis. In the last step, the PEG linker (Boc-PEG-12-CH₂CH₂COOH) was attached using HATU and 2,4,6-collidine and the Boc protecting group was removed using standard deprotection conditions to yield the final QF probe precursor **210**.



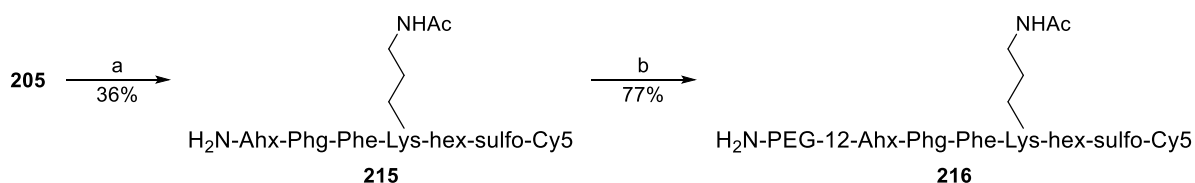
Scheme 3.30 Preparation of QF probe precursor **210**: (a) (i) sulfo-Cy5-COOSu, DIPEA, DMF, rt; (ii) TFA, DCM, rt; (b) sulfo-QSY21-COOSu, DIPEA, DMF, rt; (c) 1% DBU, DMF, rt; (d) (i) Boc-PEG-12-CH₂CH₂COOH, HATU, 2,4,6-collidine, DMF, rt; (ii) TFA, DCM, rt.

The preparation of inverse FQ probe precursor **214** was analogous to the preparation of QF precursor **210** and proceeded according to **Scheme 3.31**. Starting from the same peptide fragment **206**, sulfo-QSY21 NHS ester was used to attach the quencher to the C-terminus to yield compound **211**. Reaction of amines with sulfo-QSY-21 NHS ester is generally less efficient than reaction with sulfo-Cy5 NHS ester, which is also reflected in the lower yield of the reaction. Due to the higher complexity of the reaction mixture, intermediate **211** was purified first, then the Boc protecting group was removed using standard procedure and amine **212** was purified again. Sulfo-Cy5 was then attached to the Lys side chain and the Fmoc group was removed in a one-pot procedure by adding DBU to the reaction mixture after achieving sufficient conversion. After successful removal of the Fmoc protecting group, the reaction could be easily quenched by addition of AcOH to eliminate any additional racemization, and directly submitted for reversed-phase purification to yield amine **213**. Good yield of this transformation confirms higher reactivity of sulfo-Cy5, compared with sulfo-QSY21, and efficient removal of the Fmoc protecting group with very little byproduct formation. In the last step, the PEG linker was attached and deprotected, analogously to preparation of the QF precursor, to yield the final FQ precursor **214**. Interestingly, yields in this step were lower in both cases. This is likely due to the higher molecular size and complexity of both coupling partners.



Scheme 3.31 Preparation of FQ probe precursor **214**: (a) sulfo-QSY21-COOH, PyBOP, DIPEA, DMF, rt; (b) TFA, DCM, rt; (c) (i) sulfo-Cy5-COOSu, DIPEA, DMF, rt; (ii) 1% DBU, DMF, rt; (d) (i) Boc-PEG-12-CH₂CH₂COOH, HATU, 2,4,6-collidine, DMF, rt; (ii) TFA, DCM, rt.

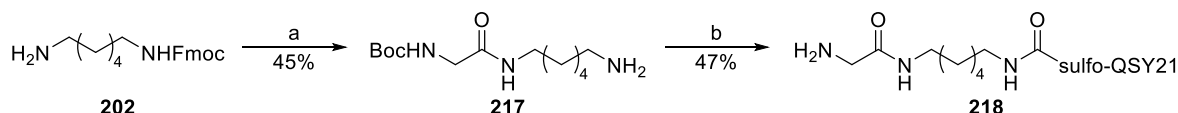
Next, I proceeded with the preparation of precursor **216** required for the Qx:1F probe architecture. This precursor is acetylated at the Lys residue, where the second fluorophore would normally be attached. One of the arguments that led us to the Qx:1F architecture was the simplification of probe synthesis, as discussed in **Chapter 1.4.2**. The synthetic sequence for the preparation of precursor **216**, as shown in **Scheme 3.32**, is notably shorter than the synthesis of precursors **210** (QF) or **214** (FQ). In the first step, peptide **205** was modified with sulfo-Cy5 NHS ester, and the Fmoc group was removed using the one-pot procedure described above to yield intermediate **215**. Next, the PEG linker was attached and deprotected using conditions used for both QF and FQ precursors to yield Qx:1F precursor **216**. Surprisingly, in this case, the last step of the sequence proceeded with a very good yield.



Scheme 3.32 Preparation of F precursor **216**: (a) (i) sulfo-Cy5-COOSu, DIPEA, DMF, rt; (ii) 1% DBU, DMF, rt; (b) (i) Boc-PEG-12-CH₂CH₂COOH, HATU, 2,4,6-collidine, DMF, rt; (ii) TFA, DCM, rt.

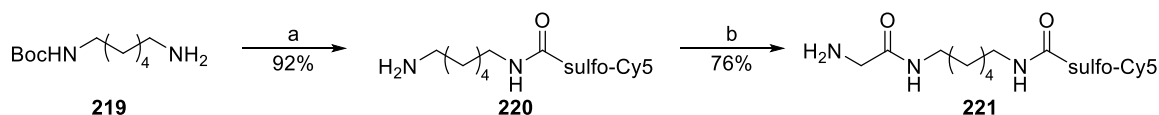
The Qx:1F architecture requires a quencher precursor, which enables individual attachment of the quencher to the polymeric backbone. This obviously adds more synthetic work, but this precursor is universal and can be applied for other substrate sequences without the need for modification. The preparation of Q precursor **218** is described below in **Scheme 3.33**. Starting from mono-Fmoc 1,6-diaminohexane (**202**), Boc-protected glycine was attached to the free amine using HCTU and 2,4,6-collidine. The protected

intermediate was partially purified by aqueous work-up, the Fmoc protecting group was removed using standard 20% solution of piperidine in DMF, and amine **217** was purified using reversed-phase flash chromatography. Next, amine **217** reacted with sulfo-QSY21 NHS ester and, finally, the Boc protecting group was removed using standard conditions to yield Q precursor **218** in good yield.



Scheme 3.33 Preparation of Q precursor **218**: (a) (i) Boc-Gly-OH, HCTU, 2,4,6-collidine, DMF, rt; (ii) 20% piperidine, DMF, rt; (b) (i) sulfo-QSY21-COOSu, DIPEA, DMF, rt; (ii) TFA, DCM, rt.

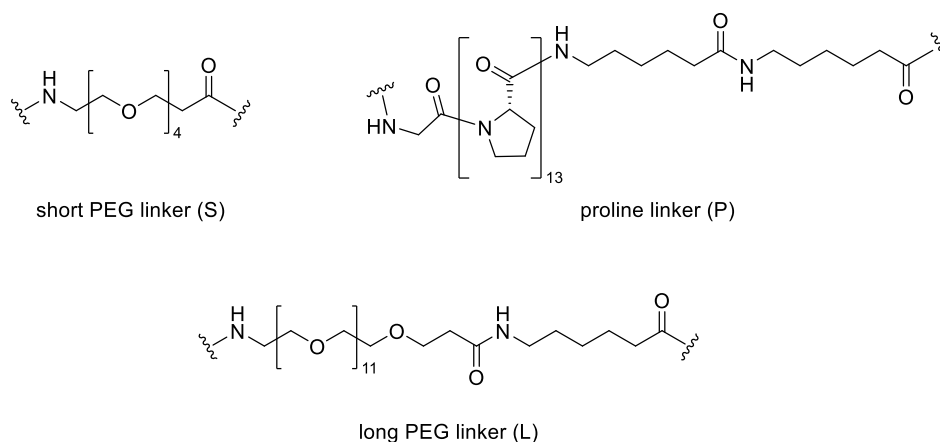
Analogously to the Q precursor **218**, I also prepared the nQ (non-quenched) probe precursor **221**, according to **Scheme 3.34**. Due to the low atom economy and yield in preparation of **218**, the nQ precursor **221** was prepared in a slightly modified manner. First, mono-Boc 1,6-diaminohexane **219** was coupled with sulfo-Cy5 NHS ester and the Boc protecting group was removed under standard conditions to yield compound **220** in excellent yield. Next, Boc-protected glycine was attached using the corresponding NHS ester and the Boc protecting group was removed to yield the final nQ precursor **221**.



Scheme 3.34 Preparation of nQ probe precursor **221**: (a) (i) sulfo-Cy5-COOSu, DIPEA, DMF, rt; (ii) TFA, DCM, rt; (b) (i) Boc-Gly-OSu, DIPEA, DMF, rt; (ii) TFA, DCM, rt.

This chapter provides a detailed description of the synthesis of probe precursors leading to macromolecular probes with various architectures. The QF, FQ and F precursors (**210**, **214**, **216**) described above all possess a PEG-12 linker. As discussed in **Chapter 1.4.2**, we aimed to test different linkers to ensure efficient protease cleavage of the substrates on the polymeric carrier. Besides PEG-12, we also designed a shorter PEG-4 alternative, and a more rigid linker composed of 13 prolines. We hypothesized that the additional rigidity of the proline linker could be beneficial for effective display of the substrate sequences and allowing for less hindered protease access. The linkers will be referred to as S (short, PEG-4), L (long, PEG-12) and P (13 prolines) and their structure can be seen below in **Scheme 3.35**. Probe precursors possessing S and P linkers were prepared by Prof. M. Bogoy's group at Stanford University and, therefore, their

preparation is not discussed in detail in this thesis. Generally, they were prepared by methods analogous to those described above and detailed procedures can be found in the corresponding paper.¹⁴⁸



Scheme 3.35 Structures of employed linkers

Macromolecular probes based on all prepared precursors were synthesized by our collaborators at the Institute of Macromolecular Chemistry using previously optimized procedures.^{115,118–122} The pHPMA precursors are decorated with thiazolidine-2-thione reactive groups which react readily with amine-terminated probe precursors to yield final macromolecular probes. As mentioned earlier, multiple different components can be tethered to the macromolecular carrier and their abundance can be controlled in this step to yield probes with different fluorophore loadings and quencher-fluorophore ratios in the case of the Qx:1F architecture. A summary of the prepared probes and their main characteristics can be found below in **Table 3.4**.

Table 3.4 Compositions and characteristics of evaluated polymeric probes: M_w' – calculated as M_w of the copolymer precursor increased by the sum of apparent molecular weights of all conjugated ligands; F – weight percentage of the fluorophore moiety on the polymeric probe, as determined by amino acid analysis after polymer hydrolysis; No. of F – average number of fluorophore moieties per unit of polymer; Q – weight percentage of the quencher moiety on the polymeric probe, as determined by amino acid analysis after polymer hydrolysis; No. of Q – average number of quencher moieties per unit of polymer; Q:F Ratio – ratio between No. of Q and No. of F.

Probe	Linker	Architecture	M_w' [kg.mol ⁻¹]	F [wt%]	No. of F	Q [wt%]	No. of Q	Q:F Ratio
S-QF	PEG-3	QF	75	6.3	1.9	-	-	-
S-Q1:1F	PEG-3	Q 1:1 F	78	5.1	2.8	4.5	3.5	1.3
S-Q3:1F	PEG-3	Q 3:1 F	82	5.0	2.9	9.4	7.8	2.7
P-QF	polyPro	QF	73	9.9	2.1	-	-	-
L-QF	PEG-12	QF	72	8.2	2.2	-	-	-
L-FQ	PEG-12	FQ	67	7.7	1.8	-	-	-
L-Q1:1F	PEG-12	Q 1:1 F	71	5.2	1.9	2.1	1.5	0.8
L-Q2:1F	PEG-12	Q 2:1 F	75	6.1	2.4	6.2	4.7	1.9
nQ	-	F	67	2.7	2.3	-	-	-

As summarized in the table above, we have prepared eight probes possessing three different linkers (S, P, L) along with the control non-quenched probe **nQ**. Probes with the QF architecture were prepared in the case of all three linkers. Probes from the short PEG series (S) and the long PEG series (L) were also prepared with the Qx:1F architecture and different quencher-fluorophore ratios. In the case of the long PEG linker, we also prepared a probe possessing the inverse FQ architecture **L-FQ**. Generally, our objective was to maintain the molecular weight as well as the fluorophore loading (No. of F) uniform across all macromolecular probes in order to minimize impact of these variables on probe performance. In this study, we focused mostly on the impact of the linker and the architecture on probe performance and, therefore, we tried to minimize other contributing factors.

3.3.2 Evaluation of Macromolecular Probes *In Vitro*

Before proceeding with the *in vivo* evaluation of our pHPMA probes, we planned to verify their cleavage by a recombinant purified protease *in vitro*. We believe that the pHPMA carrier will offer enhanced biodistribution but cathepsin cleavage remains crucial for the generation of a specific fluorescent signal. I evaluated cleavage of the macromolecular probes by cathepsin L (CatL) using a simple plate reader assay. The probes were incubated in 96-well plates with CatL, and fluorescence values were recorded in a kinetic measurement over a 90 min period. In the case of negative samples, probes were mixed with an assay buffer without CatL to confirm the increase in fluorescence is cathepsin-mediated. Also, **6QC** was included in the experiment as a standard and a positive control. I performed the experiment at two different probe concentrations [10, 2.5 μ M] to minimize any concentration-related effects such as aggregation or conformational changes of the macromolecules. Specific increase of fluorescence was observed for all analyzed probes. To quantify cleavage efficiency, I applied linear regression to the initial part of the kinetic curve to obtain initial cleavage rate (k_i) as the slope of the regression curve. All k_i values were expressed as relative to **6QC** to yield k_{rel} . Interestingly, several noteworthy trends emerged from the data shown in **Figure 3.5a**. First, consistent with our expectations, cleavage efficiency was highly linker dependent. Probes from the short PEG linker series **S-QF**, **S-Q1:1F** and **S-Q3:1F** all exhibited similar k_{rel} values around half of the **6QC** value. Notably, the relative cleavage rate was higher for all macromolecular probes in the lower probe concentration. Next, we expected that probes possessing longer linkers, such as **P-QF** and **L-QF**, might achieve higher cleavage efficiency due to higher accessibility of the substrate for protease processing. This has been confirmed, both **P-QF** ($k_{rel} = 0.90 \pm 0.03$, 2.5 μ M) and **L-QF** ($k_{rel} = 1.24 \pm 0.07$, 2.5 μ M) reached higher k_{rel} values than any of the short PEG linker probes. This trend could be seen for both probe concentrations. Interestingly, the rigidity of the proline linker did not lead to any additional benefits in cleavage efficiency. On the contrary, **L-QF** was the only probe from the QF architecture series that reached a higher cleavage rate than **6QC**. While the flipped **L-FQ** probe showed comparable cleavage efficiency to **L-QF**, both **L-Q1:1F** and **L-Q2:1F** exhibited a significant increase in k_{rel} with the latter reaching a value of 2.20 ± 0.07 at 2.5 μ M concentration. This could suggest that the absent quencher in the vicinity of the protease cleavage site increases cleavage efficiency due to the reduction of steric hindrance caused by the bulky quencher moiety. Nevertheless, this trend was not

observed in the case of the short PEG series and, therefore, it cannot be fully explained based on our data.

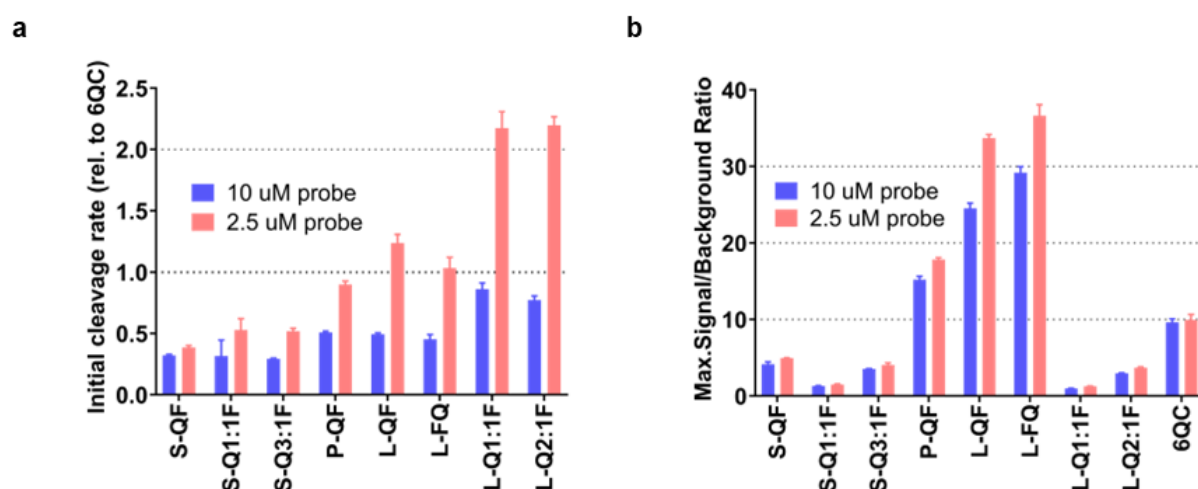


Figure 3.5 Effect of linker and architecture on *in vitro* probe activation by cathepsin L (CatL) (a) Bar graph showing initial cleavage rate of selected probes by purified CatL. Rates were measured at 2 probe concentrations [10, 2.5 μ M] and normalized to 6QC as a standard; (b) Bar graph showing maximal signal-to-background ratio (SBR) calculated from the cleavage assay as a ratio between end point signal (90 min) in positive (CatL) and negative (no CatL) samples. SBR was measured at 2 concentrations [10, 2.5 μ M]. All experiments were performed in 96-well plates in triplicates. Data were collected over 90 min using a Biotek Cytation 3 plate reader. Error bars represent standard deviation. All values can be found in **Table 4.1**.

To maximize information extracted from the *in vitro* assay, I also decided to analyze the ratio between the fluorescence intensity generated by a probe and the background fluorescence of a probe. The maximal signal-to-background (SBR) was defined as a ratio between the end point signal (90 min) in positive (CatL) and negative (no CatL) samples. SBR is a function of cleavage efficiency, the number of fluorophores present on the polymer and, most importantly, quenching efficiency. I proposed that SBR could provide us with important information about how changing the architecture influences quenching efficiency. Comparing SBR of **S-QF**, **P-QF** and **L-QF** in **Figure 3.5b** shows how cleavage efficiency translates into SBR and confirms the trends observed in **Figure 3.5a**. Swapping the quencher-fluorophore arrangement in the case of **L-FQ** did not influence the cleavage nor the quenching efficiency significantly, resulting in comparable SBR to **L-QF**. We expected reduction of quenching efficiency in the case of probes possessing Qx:1F architecture but we could not predict the magnitude of the effect and its impact on *in vivo* performance. In the short PEG series, we could observe a drop in SBR of **S-Q1:1F** (SBR = 1.5 ± 0.1 , 2.5 μ M) compared to **S-QF** (SBR = 5.0 ± 0.1 , 2.5 μ M), which was partially compensated for by additional quencher moieties of **S-Q3:1F** (SBR = $4.1 \pm$

0.3, 2.5 μM). Nevertheless, this effect was rather subtle and is more pronounced in the case of the long PEG series probes. While **L-QF** (SBR = 33.7 ± 0.5 , 2.5 μM) and **L-FQ** (SBR = 36.6 ± 1.5 , 2.5 μM) reached the highest SBR values, with an over three-fold increase compared to **6QC** (SBR = 9.9 ± 0.7 , 2.5 μM), **L-Q1:1F** (SBR = 1.3 ± 0.1 , 2.5 μM) and **L-Q2:1F** (SBR = 3.7 ± 0.1 , 2.5 μM) exhibited dramatically reduced SBRs.

To summarize this chapter, many valuable trends were observed in the data from the *in vitro* assay. The cleavage efficiency was highly linker-dependent but also exhibited architecture dependence in the case of long PEG series. Analysis of SBR values confirmed the cleavage efficiency trends observed in initial cleavage rate values and revealed trends in changes in quenching efficiency. The expected drop in quenching efficiency and, therefore, SBR in the case of Qx:1F probes due to increased quencher-fluorophore distance was observed. In the case of the long PEG series, this effect caused a significant drop in SBR. Nevertheless, encouraged by the interesting SAR observed *in vitro*, we sought to evaluate our probes *in vivo* in an animal model.

3.3.3 Evaluation of Macromolecular Probes In Vivo

After confirming the successful cleavage of our macromolecular probes *in vitro*, we proceeded with the evaluation of their *in vivo* performance. We picked a mouse model of an orthotopic 4T1 triple negative breast cancer established in the group of our collaborator, Prof. M. Bogoyo at Stanford University. The study was conducted by the members of Prof. Bogoyo's group, as I do not possess appropriate training for animal work. Nevertheless, I assisted with the experiments, prepared injection solutions and captured images using the Pearl Trilogy small animal imaging system. After the experiments were completed, I analyzed all the images manually, plotted the data and performed statistical analysis.

Proceeding to the *in vivo* experiments, we reduced the number of evaluated probes to minimize the number of sacrificed animals and simplify the experimental setup. Nevertheless, we tried to pick a set of probes that would provide us with as much information and potential SAR as possible within the scope of the study. We picked all three representatives of the QF architecture series (**S-QF**, **P-QF**, **L-QF**) to verify if the *in vitro* cleavage efficiency translates into enhanced *in vivo* performance. Also, we included **L-QF** and **L-Q2:1F** to study the impact of the architecture on probe performance. **L-Q2:1F** has been picked as the representative of the Qx:1F architecture due to its

efficient cleavage *in vitro* and higher SBR compared to **L-Q1:1F**. Finally, the **nQ** probe served as a control to verify the contrast is not generated solely by the EPR effect, and DMSO was included as a negative control.

The *in vivo* study proceeded according to the following procedure (detailed procedure can be found in **Chapter 4.4.4**). First, mice were subcutaneously injected with 4T1 cancer cells, and the tumors were left to grow for one week. Next, probes were administered intravenously retro-orbitally, tumor areas were shaved, and images were recorded at multiple time points to examine the evolution of probe performance over time. In the next step, images were analyzed manually using Image Studio software (version 5.2.5, LI-COR, Inc.). Regions of interest (tumor, background) were selected using a freehand drawing tool and white light image only. The mean fluorescence in selected areas was calculated by the software and the tumor mean fluorescence was divided by the mean fluorescence of the background to obtain the tumor-to-background ratio (TBR). We used TBR as the main metric of probe performance *in vivo*. Generally, tumor areas are easily identifiable under white light. We defined the background area as the medial chest between and slightly superior to the tumors. This option seemed superior to the abdominal region where various organs could influence the background fluorescence intensity. Representative images with the tumor (white dotted line) and background (blue dotted line) highlighted can be seen below in **Figure 3.6**, and detailed images can be found in **Chapter 4.4.1**.

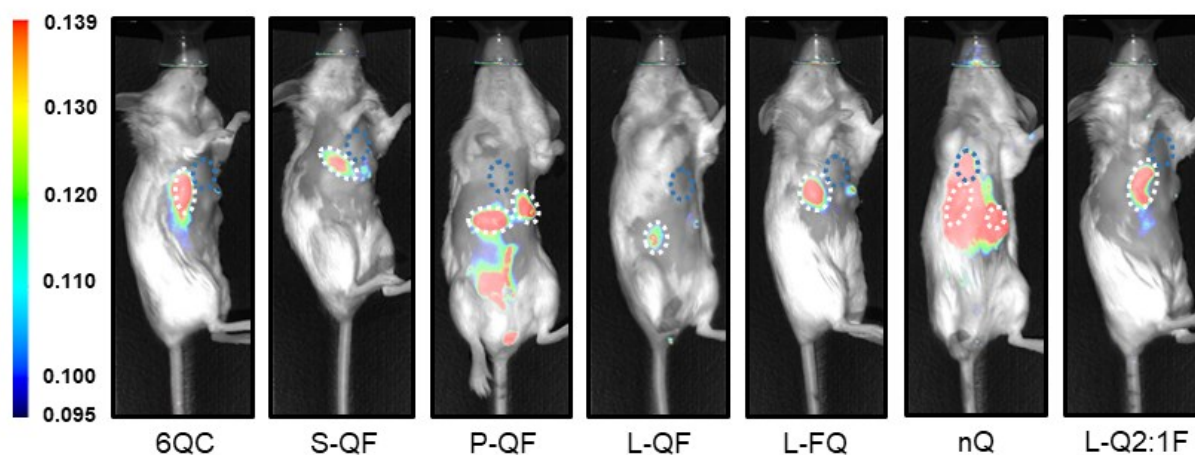


Figure 3.6 Representative fluorescence images showing a mouse with two breast tumors after injection of indicated probe [6.25 nmol] 24 hours prior to imaging. Tumors are highlighted using a white dotted line, background area is highlighted using blue dotted lines. The color bar shows fluorescence intensity in arbitrary units.

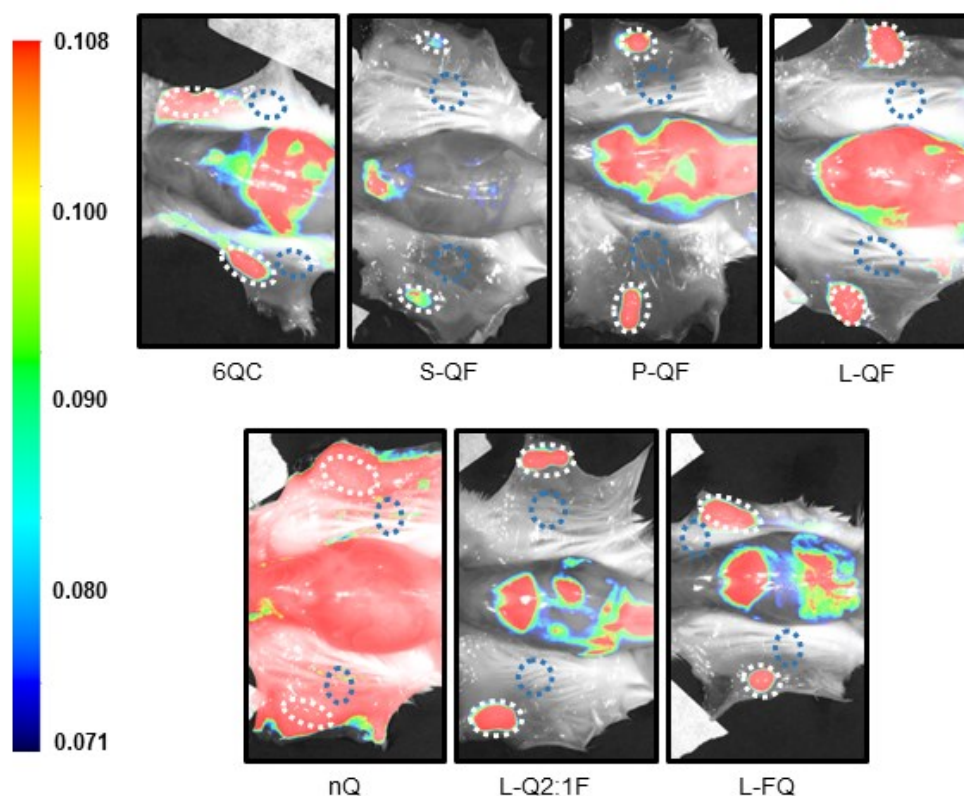


Figure 3.7 Representative fluorescence images showing a splayed mouse with two tumors after injection of indicated probe [6.25 nmol] 24 hours prior to imaging. Tumors are highlighted using white dotted lines, background area is highlighted using blue dotted lines. Color bar shows fluorescence intensity in arbitrary units.

Images were acquired at three different time points (2 h, 8 h, 24 h). After the final time point, mice were sacrificed, and we splayed open their skin to reveal the undersurface of the tumors. Again, images were acquired and analyzed using the same procedure as for the living mice. In this case, surrounding skin was selected as the background area, as shown in **Figure 3.7** and detailed images can also be found in **Chapter 4.4.1**. Finally, to assess biodistribution of the probes, tumors and various organs were harvested, and the mean fluorescence intensity was analyzed for all harvested organs. Images in **Figure 3.6** and **Figure 3.7** offer visual insight into tumor labeling by our macromolecular probes, but due to different fluorescence intensity of individual probes, it is hard to display images for all probes which would accurately represent their performance. Therefore, it is more relevant to analyze the TBR values of individual probes which represent true contrast of the probe rather than the images showing absolute fluorescence intensity.

After analyzing all images, TBR values were plotted using box plots and statistical significance was calculated via the Brown-Forsythe and Welch ANOVA test

with Dunnett's multiple comparison test. The results can be seen in **Figure 3.8** and **Figure 3.9**. To summarize the data in **Figure 3.8**, all macromolecular probes, except **S-QF** and **nQ**, reached significantly higher TBR than **6QC** at all analyzed time points. Probe **S-QF** did not achieve a significant difference from **6QC** at any of the time points, while **nQ** only reached a significant difference from **6QC** at the 24 h time point. This confirms that probe **nQ**, driven solely by the EPR effect, is able to achieve higher contrast than **6QC**, but requires up to 24 hours to attain this contrast. At the 24 h time point, **6QC** (TBR = 2.3 ± 0.5) exhibited the lowest TBR of all studied probes, while **L-Q2:1F** (TBR = 3.9 ± 0.9) reached the highest TBR in our dataset. Multiple interesting trends can be drawn from **Figure 3.8**. First, comparing the three macromolecular probes possessing the QF architecture (**S-QF**, **P-QF**, **L-QF**) suggests that the differences in cleavage efficiency did affect contrast of the probes *in vivo* as the TBR values followed cleavage efficiency trends observed *in vitro*, even though the TBR differences are subtle. Probe **S-QF** showed TBR of 2.7 ± 0.6 at the 24 h time point, compared to 2.9 ± 0.4 for **P-QF** and 3.1 ± 0.8 for **L-QF**. Looking at other time points, TBRs for all probes possessing QF architecture were generally stable and fluctuated around similar values following the trend discussed above. Next, probe **L-FQ** (TBR = 3.2 ± 0.3 , 24h), releasing the quencher moiety, showed similar TBR values as the **L-QF** probe (TBR = 3.1 ± 0.8 , 24 h). Interestingly, it seems that the TBR values of **L-FQ** were declining over time from 3.7 ± 1.0 at the 2 h time point to 3.2 ± 0.3 at the 24 h time point. Nevertheless, it is not possible to attribute this trend to the FQ architecture based on a single probe. On the other hand, probes **nQ** and **L-Q2:1F** share an increasing trend in TBR values over time. This can be explained by the gradual exclusion of probes circulating in the bloodstream, increasing background fluorescence in the case of poorly or non-quenched probes and, therefore, affecting the TBR values. TBR values for **nQ** increased from 2.1 ± 0.5 at the 2 h time point to 3.0 ± 0.8 at the final 24 h time point. As discussed in **Chapter 3.3.2**, probe **L-Q2:1F** exhibits low SBR due to low quenching efficiency caused by increased quencher-fluorophore distance. Due to its high background, this probe also requires up to 24 hours to reach peak TBR values. Probe **L-Q2:1F** reached TBR values higher than **nQ** at all analyzed time points, ranging from 2.9 ± 0.6 (2 h) to the highest TBR in live mice from all probes tested 3.9 ± 0.9 (24 h). This confirms that quenching, even if only partial, acts synergistically with the EPR effect provided by the macromolecular carrier.

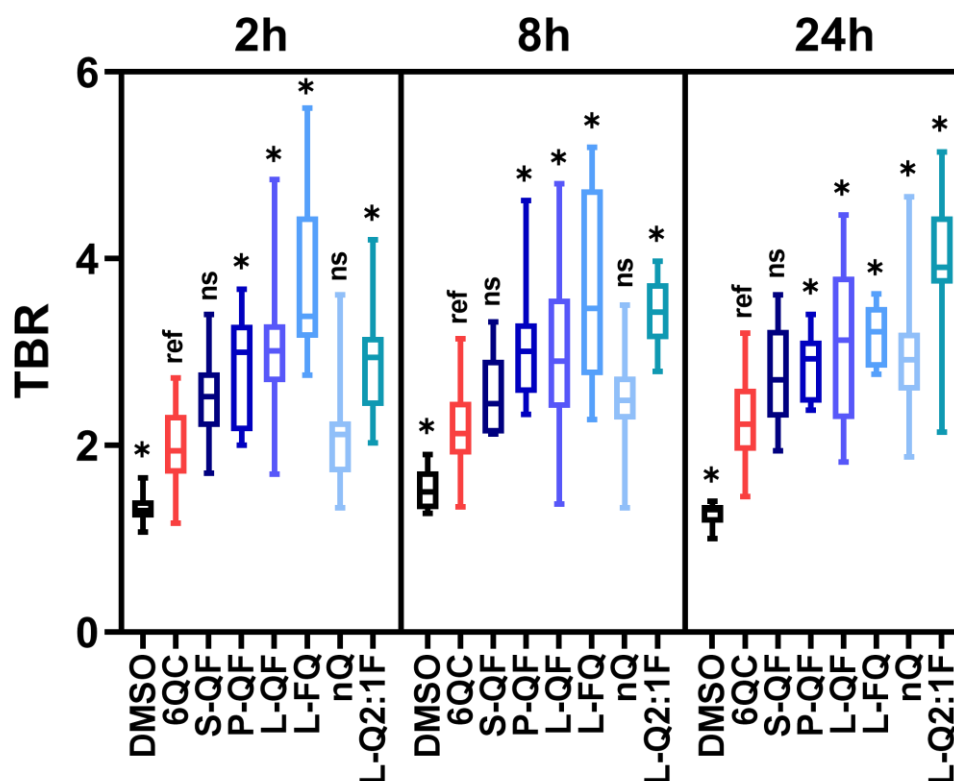


Figure 3.8 *In vivo* Probe Evaluation: Quantification of Tumor-to-Background Ratio (TBR) for individual probes in live mice: Box plots showing tumor-to-background ratio (TBR) quantification of selected probes in live mice after 2, 8 and 24 hours. *N*: 3 to 10 mice, 6 to 20 tumors per condition. Images were acquired using the Pearl Trilogy small animal imaging system at the indicated time points. Statistics were calculated via the Brown-Forsythe and Welch ANOVA test with Dunnett's multiple comparison test, asterisk represents statistical significance ($p < 0.05$), ns – not significant. Boxes extend from the 25th to the 75th percentile with whiskers extending to minimal and maximal values. The line in the middle of the box is plotted at the median value for each sample set. All values can be found in **Table 4.2**.

The discrepancy between low SBR of **L-Q2:1F** *in vitro* and high TBR *in vivo* might seem confusing but can be rationalized by the simplicity of the *in vitro* model which fails to replicate *in vivo* effects, such as blood circulation or the EPR effect. Due to the EPR effect, part of the signal that would be considered background *in vitro* converts into true tumor signal *in vivo* after passive probe accumulation in the tumor, dramatically increasing the contrast.

Next, we analyzed TBR values for all probes in splayed mice images shown in **Figure 3.7** and the results can be seen in **Figure 3.9** below. All macromolecular probes exhibited significantly higher TBR than **6QC** ($TBR = 2.3 \pm 0.4$), ranging from 3.4 ± 1.0 for **L-FQ** to 5.1 ± 1.9 for **S-QF** and 5.1 ± 1.3 for **L-Q2:1F**. The control probe **nQ** ($TBR = 3.7 \pm 1.4$) reached lower TBR than most of the quenched probes. Interestingly, probe **S-QF**, which did not exhibit significantly higher TBR than **6QC** in live mice, showed

highest TBR values in splayed mice along with probe **L-Q2:1F**. Another interesting observation was that all macromolecular probes showed higher TBR in splayed mice than in live mice. This is particularly interesting and underlines the potential of our macromolecular probes, as the splayed mouse setup mimics a surgical procedure more than the live mice imaging. Additionally, **6QC** did not demonstrate this effect (TBR = 2.3 ± 0.5 , live mice, 24 h; TBR = 2.3 ± 0.5 , splayed mice).

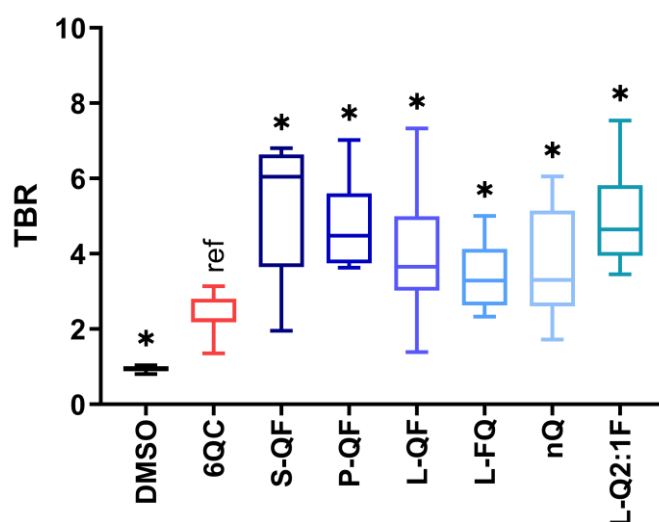


Figure 3.9 *In vivo* Probe Evaluation: Quantification of Tumor-to-Background Ratio (TBR) for individual probes in splayed mice: Box plot showing tumor-to-background ratio (TBR) quantification of selected probes in splayed mice. *N*: 3 to 10 mice, 6 to 20 tumors per condition. Images were acquired using the Pearl Trilogy small animal imaging system. Statistics were calculated via the Brown-Forsythe and Welch ANOVA test with Dunnett's multiple comparison test, asterisk represents statistical significance ($p < 0.05$). Boxes extend from the 25th to the 75th percentile with whiskers extending to minimal and maximal values. The line in the middle of the box is plotted at the median value for each data set. All values can be found in **Table 4.2**.

Finally, to assess biodistribution of our probes, we analyzed the mean fluorescence intensity in harvested organs and plotted the signal intensity relative to the tumor in **Figure 3.10**. Probe **6QC** does not accumulate significantly in fat pads or spleen and no further reduction was observed for any of the macromolecular probes. In the case of the heart, the accumulation of **6QC** is also limited, however, a significant reduction was observed for the majority of macromolecular probes, with the exception of **P-QF** and **nQ**. Accumulation of **6QC** was slightly higher in lungs (0.85 ± 0.19), but significant reduction (below 0.5) was observed for all probes. The highest accumulation of **6QC** could be observed in kidneys and liver. Notably, all macromolecular probes exhibited reduced accumulation in kidneys. For most probes, relative fluorescence intensity dropped below 1, compared to 2.6 ± 0.5 for **6QC**. Interestingly, in the case of **P-QF**, relative intensity was only reduced to 2.0 ± 0.3 . We can speculate that this is caused by the proline linker, as **P-**

QF is the only macromolecular probe with increased kidney accumulation. On the other hand, most macromolecular probes did not attain reduced fluorescence intensity in the liver. We hypothesized that this could be caused by high proteolytic activity and release of the fluorophore in the liver. Nevertheless, probe **L-FQ**, releasing the quencher instead of the fluorophore, showed similar relative fluorescence intensity in the liver as the rest of the quenched probes. The only two probes with significantly reduced liver accumulation compared to **6QC** (2.1 ± 0.8) were **S-QF** (1.1 ± 0.3) and **nQ** (0.5 ± 0.1). Based on this result, one could hypothesize that liver accumulation increases with increasing lipophilicity of individual constructs. The most hydrophilic construct **nQ**, possessing only the fluorophore moiety, exhibited the lowest liver accumulation, while probe **S-QF**, containing the substrate fragment with a short PEG linker, showed slightly elevated liver accumulation. This trend could be extended to other probes containing longer and more lipophilic linkers. This is an interesting observation, but further investigation would have to be carried out to fully understand the essence of this trend.

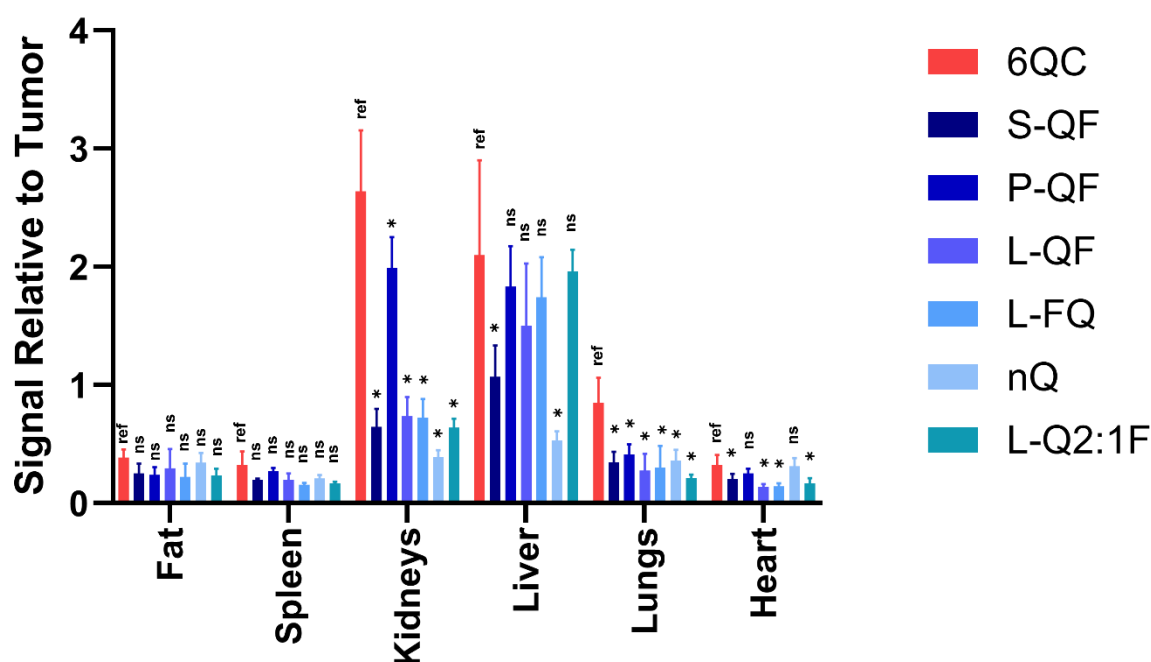


Figure 3.10 *Ex vivo* Probe Evaluation: Bar graph showing fluorescence distribution in selected organs and tissues. *N*: 3 to 10 mice, 6 to 20 tumors per condition. Images were acquired using the Pearl Trilogy small animal imaging system. Statistics were calculated via the Brown-Forsythe and Welch ANOVA test with Dunnett's multiple comparison test, asterisk represents statistical significance ($p < 0.05$), ns – not significant. Error bars represent standard deviation.

4 EXPERIMENTAL PART

4.1 General information

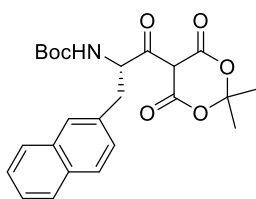
Commercially available HPLC grade acetonitrile, catalysts and reagent grade materials were used for the synthesis of described compounds. All amino acids have L stereochemistry, unless stated otherwise. TLC was performed on Silica gel 60 F254-coated aluminum sheets (Merck). Flash chromatography was performed on CombiFlash NextGen 100 (Teledyne Isco) with a C18 RediSep Rf Gold columns. All chemicals were purchased from Sigma-Aldrich, TCI, Combi-Blocks or Iris Biotech GmbH and were used without further purification. The ^1H and ^{13}C NMR spectra were measured using instruments Bruker Avance IIITM HD 400 MHz, Bruker Avance IIITM HD 500 MHz and Bruker Avance IIITM 600 MHz. For standardization of ^1H NMR spectra, the internal signal of TMS (δ 0.0, CDCl_3) or residual signals of CDCl_3 (δ 7.26), $\text{DMSO-}d_6$ (δ 2.50, 3.33) or CD_3OD (δ 3.31, 4.87) were used. In the case of ^{13}C NMR spectra (APT experiments) the residual signal of CDCl_3 (δ 77.16), $\text{DMSO-}d_6$ (δ 39.52) or CD_3OD (δ 49.00) were used. The NMR spectra were recorded at the room temperature, unless noted otherwise. The chemical shifts are given in δ scale; the coupling constants J are given in Hz. The ESI mass spectra were recorded using ZQ micromass mass spectrometer (Waters) equipped with an ESCi multimode ion source and controlled by MassLynx software. The low-resolution ESI mass spectra were recorded using a quadrupole orthogonal acceleration time-of-flight tandem mass spectrometer (Q-Tof micro, Waters) and high resolution ESI mass spectra using a hybrid FT mass spectrometer combining a linear ion trap MS and the Orbitrap mass analyzer (LTQ Orbitrap XL, Thermo Fisher Scientific). The conditions were optimized for suitable ionization in the ESI Orbitrap source (sheath gas flow rate 35 a.u., aux gas flow rate 10 a.u. of nitrogen, source voltage 4.3 kV, capillary voltage 40 V, capillary temperature 275 °C, tube lens voltage 155 V). The samples were dissolved in methanol and applied by direct injection. Final inhibitors were purified using preparative scale HPLC Jasco PU-975 (flow rate 10 mL/min) equipped with UV detector UV-975 and with column Waters YMC-PACK ODS-AM C18 Prep Column, 5 μm , 20 $\times$ 250 mm. The purity of compounds was tested on analytical Jasco PU-1580 HPLC (flow rate 1 mL/min, invariable gradient 2 – 100 % ACN in 30 minutes) with column Watrex C18 Analytical Column, 5 μm , 250 $\times$ 5 mm. The final inhibitors were all at least of 95% purity. Optical rotations were measured using an Autopol IV instrument (Rudolph Research Analytical).

4.2 Macrocyclic Inhibitors of Human Cathepsin D

4.2.1 Precursors of Macrocyclic Inhibitors

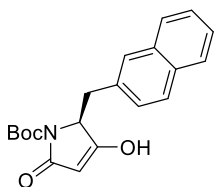
Representative procedure for preparation of statine analogues:

***tert*-Butyl (S)-(1-(2,2-dimethyl-4,6-dioxo-1,3-dioxan-5-yl)-3-(naphthalen-2-yl)-1-oxopropan-2-yl)carbamate (69)**



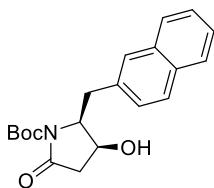
To a solution of Boc-3-(2-Naphthyl)-L-alanine (0.32 g, 1.00 mmol, 1.0 eq.), Meldrum's acid (0.15 g, 1.05 mmol, 1.05 eq.) and DMAP (0.24 g, 2.00 mmol, 2.0 eq.) in anhydrous DCM (5 mL) was added DCC (0.22 g, 1.05 mmol, 1.05 eq.). The reaction mixture was stirred overnight, filtered and the solid was washed with DCM (2 × 10 mL). The filtrate was then washed with 10% KHSO₄ (2 × 20 mL), water (2 × 20 mL) and brine (20 mL). The solvent was evaporated to yield crude compound **69**, which was used in the next step without further purification.

***tert*-Butyl (S)-3-hydroxy-2-(naphthalen-2-ylmethyl)-5-oxo-2,5-dihydro-1H-pyrrole-1-carboxylate (70)**



The crude product **69** (0.44 g, 1.00 mmol, 1.0 eq.) was dissolved in EtOAc (10 mL), and the solution was heated to reflux for 30 minutes. After cooling, the product was extracted into aqueous phase using 5% NaHCO₃ (3 × 10 mL). The aqueous phase was then acidified using KHSO₄ to pH 2. The product was then extracted to EtOAc (3 × 15 mL). The extract was evaporated to dryness to yield crude compound **70**.

***tert*-Butyl (2S,3S)-3-hydroxy-2-(naphthalen-2-ylmethyl)-5-oxopyrrolidine-1-carboxylate (71)**

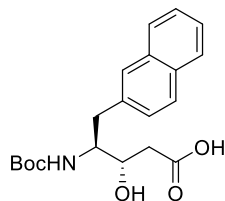


A solution of the crude compound **70** (0.34 g, 1.00 mmol, 1.0 eq.) in a mixture of DCM (10 mL) and AcOH (1 mL) was cooled down to 0 °C and NaBH₄ (0.08 g, 2.00 mmol, 2.0 eq.) was added portionwise. The mixture was stirred overnight at room temperature. It was then poured into cold water and the organic layer was washed with water, dried with MgSO₄ and evaporated to dryness. Chromatography of the residue on silica gel (chex:EtOAc / 2:1) afforded compound **71** (0.20 g, 60% over 3 steps).

¹H NMR (400 MHz, DMSO-*d*₆): 1.46 (s, 9H), 2.10 (bs, 1H), 2.40 – 2.50 (m, 1H), 2.55 – 2.63 (m, 1H), 3.29 (d, *J* = 6.0 Hz, 2H), 4.43 – 4.57 (m, 2H), 7.40 – 7.50 (m, 3H), 7.73 – 7.83 (m, 4H).

¹³C NMR (100 MHz, DMSO-*d*₆): 27.9, 34.2, 40.1, 62.6, 65.6, 83.4, 125.7, 126.2, 127.5, 127.6, 127.9, 128.2 (2C), 132.2, 133.5, 135.3, 149.8, 171.3.

(3*S*, 4*S*)-4-*tert*-Butoxycarbonylamino-3-hydroxy-5-(2-naphthyl)pentanoic acid (72**)**



A solution of compound **71** (0.20 g, 0.60 mmol, 1.0 eq.) in acetone (5 mL) was treated dropwise with NaOH (1M, 1.20 mL, 1.20 mmol, 2.0 eq.). After stirring for 30 minutes at room temperature, the solution was diluted with water (10 mL), carefully acidified to pH 3 with 1M HCl and the product extracted with EtOAc (3 × 10 mL) to afford compound **72** (0.20 g, 92%).

¹H NMR (400 MHz, DMSO-*d*₆): 1.26 (s, 9H), 2.23 (dd, *J* = 15.5, 9.2 Hz, 1H), 2.41 (dd, *J* = 15.5, 3.6 Hz, 1H), 2.76 (dd, *J* = 13.6, 9.4 Hz, 1H), 3.00 (dd, *J* = 13.6, 5.1 Hz, 1H), 3.74 – 3.84 (m, 1H), 3.90 – 3.98 (m, 1H), 6.56 (d, *J* = 9.3 Hz, 1H), 7.38 – 7.51 (m, 3H), 7.71 (s, 1H), 7.79 – 7.88 (m, 3H).

¹³C NMR (100 MHz, DMSO-*d*₆): 28.1, 36.1, 38.6, 55.7, 68.4, 77.6, 125.2, 125.8, 127.1, 127.2, 127.3, 127.4, 128.6, 131.9, 133.5, 135.3, 155.4, 173.1.

EI MS: 359 ([M]⁺).

HR EI MS: calcd for C₂₀H₂₅O₅N 359.1732; found 359.1725.

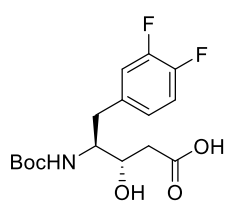
MP: 180 - 181 °C

[α]_D²⁰(c = 2.0, EtOH): -26.0°

EA: calcd for C, 66.83; H, 7.01; N, 3.90; found C, 66.68; H, 7.08; N, 3.89.

Characterization of statine analogues:

(3*S*, 4*S*)-4-*tert*-Butoxycarbonylamino-3-hydroxy-5-(3,4-difluorophenyl)pentanoic acid (73**)**



¹H NMR (400 MHz, DMSO-*d*₆): 1.29 (s, 9H), 2.18 (dd, *J* = 15.6, 9.1 Hz, 1H), 2.38 (dd, *J* = 15.6, 3.7 Hz, 1H), 2.57 (dd, *J* = 13.5, 10.2 Hz, 1H), 2.80 (dd, *J* = 13.5, 4.5 Hz, 1H), 3.54 – 3.73 (m, 1H), 3.81 - 3.97 (m, 1H), 4.90 (bs, 1H), 6.50 (d, *J* = 9.4 Hz, 1H), 6.99 – 7.10 (m, 1H), 7.17 – 7.40 (m, 2H), 11.87 (bs, 1H).

¹³C NMR (100 MHz, DMSO-*d*₆): 28.0, 35.2, 38.5, 55.7, 68.6, 77.6, 116.9 (d, *J*_{CCCF} = 16.3), 117.4 (dd, *J*_{CCF} = 51.3 Hz, *J*_{CCCF} = 16.6 Hz), 125.9 (dd, *J*_{CCCF} = 6.2 Hz, *J*_{CCCCF} = 3.1 Hz),

137.4 (dd, $J_{\text{CCCF}} = 5.8$ Hz, $J_{\text{CCCCF}} = 3.5$ Hz), 148.4 (dd, $J_{\text{CF}} = 243.6$ Hz, $J_{\text{CCF}} = 52.5$ Hz), 148.6 (dd, $J_{\text{CF}} = 243.8$ Hz, $J_{\text{CCF}} = 52.5$ Hz), 155.5, 173.0.

CI MS: 368 ($[\text{M} + \text{Na}]^+$).

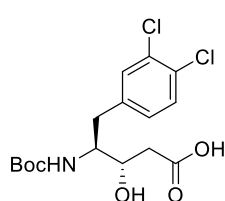
HR EI MS: calcd for $\text{C}_{16}\text{H}_{21}\text{O}_5\text{NF}_2$ 345.1315; found 345.1387.

MP: 155 - 157 °C

$[\alpha]_{\text{D}}^{20}$ (c = 1.33, EtOH): -26.4°

EA: calcd for C, 55.65; H, 6.13; N, 4.06; F, 11.00; found C, 56.09; H, 6.25; N, 4.00; F, 10.79.

(3*S*, 4*S*)-4-*tert*-Butoxycarbonylamino-3-hydroxy-5-(3,4-dichlorophenyl)-pentanoic acid (74)



^1H NMR (400 MHz, $\text{DMSO}-d_6$): 1.29 (s, 9H), 2.19 (dd, $J = 15.5, 9.1$ Hz, 1H), 2.39 (dd, $J = 15.5, 3.6$ Hz, 1H), 2.56 (dd, $J = 13.4, 10.1$ Hz, 1H), 2.82 (dd, $J = 13.4, 4.2$ Hz, 1H), 3.56 – 3.72 (m, 1H), 3.83 – 3.96 (m, 1H), 5.00 (bs, 1H), 6.54 (d, $J = 9.4$ Hz, 1H), 7.21 (dd, $J = 8.2, 1.9$ Hz, 1H), 7.45 – 7.55 (m, 2H), 11.20 (bs, 1H).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$): 28.0, 35.1, 38.4, 55.6, 68.7, 77.7, 128.5, 129.7, 130.1, 130.5, 131.1, 140.9, 155.5, 173.0.

CI MS: 400 ($[\text{M} + \text{Na}]^+$).

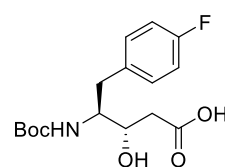
HR EI MS: calcd for $\text{C}_{16}\text{H}_{21}\text{O}_5\text{NCl}_2$ 377.0795; found 377.0796.

MP: 170 - 172 °C

$[\alpha]_{\text{D}}^{20}$ (c = 1.2, EtOH): -25.0°

EA: calcd for C, 50.81; H, 5.60; N, 3.70; Cl, 18.75; found C, 50.55; H, 5.60; N, 3.58; Cl, 18.43.

(3*S*, 4*S*)-4-*tert*-Butoxycarbonylamino-3-hydroxy-5-(4-fluorophenyl)pentanoic acid (75)



^1H NMR (400 MHz, $\text{DMSO}-d_6$): 1.29 (s, 9H), 2.18 (dd, $J = 15.5, 9.2$ Hz, 1H), 2.37 (dd, $J = 15.5, 3.6$ Hz, 1H), 2.55 (dd, $J = 13.6, 9.7$ Hz, 1H), 2.80 (dd, $J = 13.6, 4.8$ Hz, 1H), 3.54 – 3.71 (m, 1H), 3.88 (dddd, $J = 8.8, 8.7, 3.2, 3.1$ Hz, 1H), 6.48 (d, $J = 9.4$ Hz, 1H), 7.01 – 7.15 (m, 2H), 7.18 – 7.28 (m, 2H).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$): 28.1, 35.1, 38.5, 55.8, 68.4, 77.6, 114.7 (d, $J_{\text{CCCF}} = 20.9$ Hz), 130.9 (d, $J_{\text{CCCF}} = 7.8$ Hz), 135.7 (d, $J_{\text{CCCCF}} = 2.8$ Hz), 155.5, 160.8 (d, $J_{\text{CF}} = 241.4$ Hz), 173.1.

CI MS: 334 ($[\text{M} + \text{Li}]^+$).

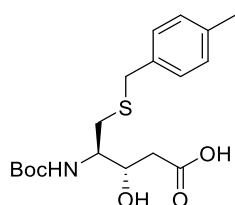
HR EI MS: calcd for $\text{C}_{16}\text{H}_{22}\text{O}_5\text{NF}$ 327.1482; found 327.1468.

MP: 162 - 163 °C

$[\alpha]_D^{20}$ (c = 2.5, EtOH): -31.2°

EA: calcd for C, 58.71; H, 6.77; N, 4.28; F, 5.80; found C, 58.40; H, 6.71; N, 4.22; F, 5.67.

(3*S*,4*R*)-4-((*tert*-Butoxycarbonyl)amino)-3-hydroxy-5-((4-methylbenzyl)thio)pentanoic acid (76)



$^1\text{H NMR}$ (400 MHz, DMSO- d_6): 1.41 (s, 9H), 2.18 (dd, J = 15.6, 8.9 Hz, 1H), 2.27 (s, 3H), 2.31 – 2.41 (m, 2H), 2.57 (dd, J = 13.4, 6.0 Hz, 1H), 3.58 – 3.74 (m, 3H), 3.99 – 4.07 (m, 1H), 6.42 (d, J = 9.3 Hz, 1H), 7.10 (d, J = 7.7 Hz, 2H), 7.20 (d, J = 8.0 Hz, 1H).

$^{13}\text{C NMR}$ (100 MHz, DMSO- d_6): 20.7, 28.2, 31.9, 34.6, 38.6, 53.3, 67.7,

77.8, 128.8, 128.9, 135.4, 135.8, 155.6, 172.8.

ESI MS: 392 ($[\text{M} + \text{Na}]^+$).

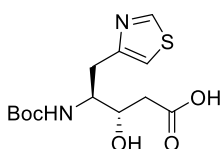
HR ESI MS: calcd for $\text{C}_{18}\text{H}_{27}\text{O}_5\text{NNaS}$ 392.15021; found 392.14989.

MP: 95-96 °C

$[\alpha]_D^{20}$ (c = 8.8, MeOH): -23.1°

EA: calcd for C, 58.51; H, 7.37; N, 3.79; S, 8.68; found C, 58.58; H, 7.44; N, 3.78; S, 8.55.

(3*S*,4*S*)-4-((*tert*-Butoxycarbonyl)amino)-3-hydroxy-5-(thiazol-4-yl)pentanoic acid (77)



$^1\text{H NMR}$ (400 MHz, DMSO- d_6): 1.32 (s, 9H), 2.16 (dd, J = 15.5, 9.3 Hz, 1H), 2.30 (dd, J = 15.5, 3.2 Hz, 1H), 2.80 (dd, J = 14.5, 9.1 Hz, 1H), 3.00 (dd, J = 14.5, 4.8 Hz, 1H), 3.82 – 3.90 (m, 2H), 6.42 (d, J = 9.2 Hz, 1H), 7.28 (d, J = 1.9 Hz, 1H), 9.00 (d, J = 1.9 Hz, 1H).

$^{13}\text{C NMR}$ (100 MHz, DMSO- d_6): 28.2, 32.1, 39.0, 53.8, 68.6, 77.6, 114.5, 153.1, 154.9, 155.3, 174.5.

EI MS: 316 ($[\text{M}]^+$).

HR EI MS: calcd for $\text{C}_{13}\text{H}_{20}\text{O}_5\text{N}_2\text{S}$ 316.1093; found 316.1081.

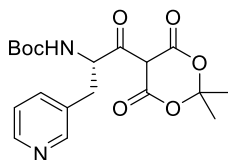
MP: 145 - 148 °C

$[\alpha]_D^{20}$ (c = 1.9, EtOH): -17.7°

EA: calcd for C, 49.35; H, 6.37; N, 8.85; S, 10.13; found C, 48.96; H, 6.32; N, 8.48; S, 9.78

Preparation of 3-pyridylalanine statine analogue:

tert-Butyl (S)-(1-(2,2-dimethyl-4,6-dioxo-1,3-dioxan-5-yl)-3-(naphthalen-2-yl)-1-oxopropan-2-yl)carbamate (**83**)



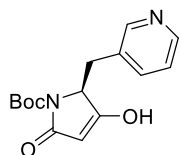
To a solution of Boc-L-3-Pal-OH (100 mg, 0.38 mmol, 1.0 eq.), Meldrum's acid (57 mg, 0.39 mmol, 1.05 eq.) and DMAP (69 mg, 0.56 mmol, 1.5 eq.) in anhydrous DCM (5 mL) was added EDC (83 mg, 0.43 mmol, 1.15 eq.). The reaction mixture was stirred at room temperature for 3 hours. The solvent was then evaporated and the product was purified by RP flash chromatography (C18, gradient 5 – 60% MeCN in H₂O+0.1% TFA) to afford **83** (112 mg, 76%) as a white lyophilisate.

¹H NMR (400 MHz, D₂O) δ 1.09 – 1.36 (m, 8H), 1.64 (s, 6H), 2.57 – 2.73 (m, 1H), 3.05 – 3.21 (m, 1H), 5.49 – 5.75 (m, 1H), 7.34 – 7.52 (m, 1H), 7.78 – 7.91 (m, 1H), 8.21 – 8.61 (m, 2H). The ¹³C NMR spectrum was not recorded due to the limited solubility of compound **83**.

ESI MS: 415 ([M + Na]⁺).

HR ESI MS: calcd for C₁₉H₂₄O₇N₂Na 415.14757, found 415.14771.

tert-Butyl (S)-3-hydroxy-5-oxo-2-(pyridin-3-ylmethyl)-2,5-dihydro-1H-pyrrole-1-carboxylate (**84**)

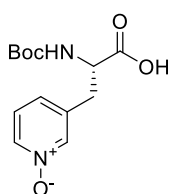


Intermediate **83** (0.30 g, 0.77 mmol, 1.0 eq.) was suspended in EtOAc (40 mL), and the solution was heated to reflux until dissolution. After cooling, the solvent was evaporated and the product was purified by RP flash chromatography (C18, gradient 5 – 50% MeCN in H₂O+0.1% TFA) to afford **84** (0.20 g, 90%) as a white lyophilisate.

¹H NMR (400 MHz, MeOD) δ 1.61 (s, 9H), 3.43 (dd, *J* = 14.2, 3.0 Hz, 1H), 3.64 (dd, *J* = 14.2, 5.6 Hz, 1H), 7.87 – 7.92 (m, 1H), 8.17 (dt, *J* = 8.1, 1.7 Hz, 1H), 8.51 (s, 1H), 8.71 (d, *J* = 5.3 Hz, 1H).

The ¹³C NMR spectrum was not recorded due to the limited stability of the compound in solution.

(S)-3-(2-((*tert*-Butoxycarbonyl)amino)-2-carboxyethyl)pyridine 1-oxide (**87**)



To a solution of Boc-L-3-Pal-OH (0.20 g, 0.75 mmol, 1.0 eq.) in DCM (15 mL) was added 3-chloroperoxybenzoic acid (0.51 g, 3.0 mmol, 4.0 eq.). The reaction mixture was refluxed overnight. After cooling, the solvent was evaporated and the product was purified by RP flash chromatography

(C18, gradient 5 – 30% MeCN in H₂O+0.1% TFA) to afford **87** (0.21 g, 97%) as a clear glassy solid.

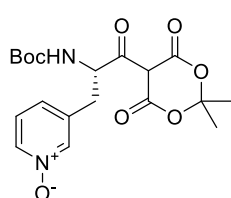
¹H NMR (400 MHz, CD₃OD) δ 1.38 (s, 9H), 2.98 (dd, *J* = 14.1, 9.7 Hz, 1H), 3.26 – 3.34 (m, 2H), 4.40 – 4.45 (m, 1H), 7.56 – 7.61 (m, 1H), 7.71 (d, *J* = 8.0 Hz, 1H), 8.33 – 8.37 (m, 1H), 8.38 – 8.41 (m, 1H).

¹³C NMR (100 MHz, CD₃OD) δ 28.6, 35.6, 55.1, 80.8, 127.9, 134.4, 138.7, 140.1, 141.0, 157.7, 174.1.

ESI MS: 281 ([M – H][–]).

HR ESI MS: calcd for C₁₃H₁₇O₅N₂ 281.11429, found 281.11418.

(S)-3-(2-((*tert*-Butoxycarbonyl)amino)-3-(2,2-dimethyl-4,6-dioxo-1,3-dioxan-5-yl)-3-oxopropyl)pyridine 1-oxide (88)

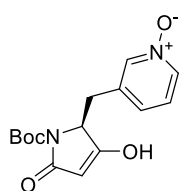


To a solution of intermediate **87** (0.20 g, 0.71 mmol, 1.0 eq.), Meldrum's acid (0.11 g, 0.74 mmol, 1.05 eq.) and DMAP (0.24 g, 1.06 mmol, 1.5 eq.) in anhydrous DCM (10 mL) was added EDC (0.16 g, 0.82 mmol, 1.15 eq.). The reaction mixture was stirred at room temperature for 3 hours. Then, the reaction mixture was diluted with DCM (10 mL). The organic phase was washed with 10% KHSO₄ (2 × 20 mL) and brine (20 mL). The solvent was evaporated to yield crude compound **88**, which was used in the next step without further purification.

ESI MS: 431 ([M + Na]⁺).

HR ESI MS: calcd for C₁₉H₂₄O₈N₂Na 431.14249, found 431.14264.

(S)-3-((1-(*tert*-Butoxycarbonyl)-3-hydroxy-5-oxo-2,5-dihydro-1H-pyrrol-2-yl)methyl)pyridine 1-oxide (89)

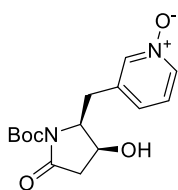


Intermediate **88** (1.0 g, 2.45 mmol, 1.0 eq.) was dissolved in EtOAc (30 mL), and the solution was refluxed for 30 minutes. After cooling, the solvent was evaporated to yield crude intermediate **89**, which was used in the next step without further purification.

ESI MS: 329 ([M + Na]⁺).

HR ESI MS: calcd for C₁₅H₁₈O₅N₂Na 329.11079, found 329.11060.

3-(((2*S*,3*S*)-1-(*tert*-Butoxycarbonyl)-3-hydroxy-5-oxopyrrolidin-2-yl)methyl)pyridine 1-oxide (90)



A solution of the crude compound **89** (0.75 g, 2.45 mmol, 1.0 eq.) in a mixture of DCM (30 mL) and AcOH (3 mL) was cooled down to 0 °C and NaBH₄ (0.37 g, 9.80 mmol, 4.0 eq.) was added portionwise. The reaction mixture was then stirred at room temperature for 1.5 hours. Next, the reaction was quenched by cautious addition of water and volatiles were evaporated. The product was purified by RP flash chromatography (C18, gradient 5 – 30% MeCN in H₂O+0.1% TFA) to afford **90** (0.23 g, 31% over 3 steps) as a colorless amorphous solid.

¹H NMR (400 MHz, CD₃OD) δ 1.47 (s, 9H), 2.45 – 2.55 (m, 1H), 2.61 – 2.71 (m, 1H), 3.09 (dd, *J* = 13.9, 4.5 Hz, 1H), 3.27 (dd, *J* = 13.9, 6.9 Hz, 1H), 4.41 – 4.53 (m, 2H), 7.55 (dd, *J* = 8.0, 6.3 Hz, 1H), 7.65 (d, *J* = 8.1 Hz, 1H), 8.26 – 8.31 (m, 1H), 8.39 – 8.43 (m, 1H).

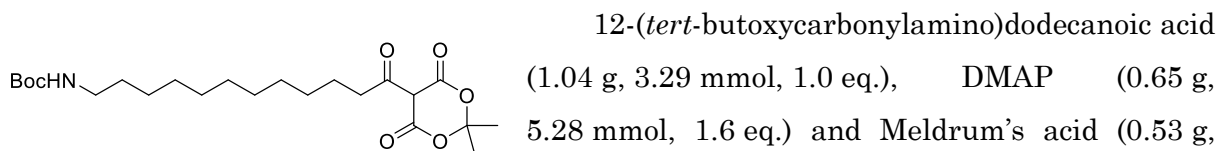
¹³C NMR (100 MHz, CD₃OD) δ 28.2, 32.3, 40.9, 63.5, 65.7, 84.7, 127.7, 133.8, 138.4, 140.8, 141.5, 151.3, 173.7.

ESI MS: 331 ([*M* + Na]⁺).

HR ESI MS: calcd for C₁₅H₂₀O₅N₂Na 331.12644, found 331.12631.

Preparation of methyl 14-aminotetradecanoate:

***tert*-Butyl 12-(2,2-dimethyl-4,6-dioxo-1,3-dioxan-5-yl)-12-oxododecylcarbamate (93)**



12-(*tert*-butoxycarbonylamino)dodecanoic acid (1.04 g, 3.29 mmol, 1.0 eq.), DMAP (0.65 g, 5.28 mmol, 1.6 eq.) and Meldrum's acid (0.53 g, 3.62 mmol, 1.1 eq.) were partially dissolved in anhydrous DCM (30 mL) under inert atmosphere. Mixture was cooled to 0 °C and solution of DCC (0.75 g, 3.62 mmol, 1.1 eq.) in DCM (20 mL) was added slowly to the suspension. The reaction was stirred at 0 °C under inert atmosphere overnight. DCU was filtered off and washed with DCM (2 × 10 mL). Solvent was evaporated and the residue was chromatographed on silica gel (CHCl₃:acetone / gradient 10:1 to 5:1) to afford desired derivative **93** (1.22 g, 84%) as a yellowish oil.

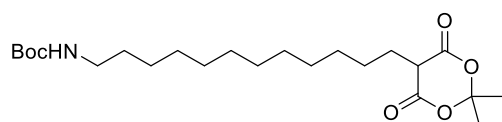
¹H NMR (400 MHz, CDCl₃) 1.19 – 1.36 (m, 14H), 1.42 (s, 9H), 1.50 – 1.58 (m, 2H), 1.61 (s, 3H), 1.65 – 1.70 (m, 2H), 1.72 (s, 3H), 1.84 – 1.92 (m, 2H), 3.00 – 3.14 (m, 2H), 3.38 (s, 1H), 4.54 (s, 1H).

¹³C NMR (100 MHz, CDCl₃) 25.0, 25.7, 26.3, 26.9, 28.6, 29.4, 29.6, 29.8, 30.2, 31.1, 33.8, 35.9, 40.8, 42.3, 49.4, 79.2, 104.9, 157.2, 170.7, 198.5.

ESI MS: 464 ([M + Na]⁺).

HR ESI MS: calcd for C₂₃H₃₉O₇NNa 464.26187, found 464.26224.

***tert*-Butyl 12-(2,2-dimethyl-4,6-dioxo-1,3-dioxan-5-yl)dodecylcarbamate (**94**)**



Intermediate **93** (1.22 g, 2.72 mmol, 1.0 eq.) was dissolved in DCM (30 mL) and cooled down to -5 °C. Acetic acid (0.96 mL, 16.58 mmol, 6.0 eq.)

was added dropwise and NaBH₄ (0.31 g, 8.15 mmol, 3.0 eq.) was added in three portions. Mixture was stirred at -5 °C overnight and then washed with brine (2 × 20mL). Organic phase was dried using anhydrous MgSO₄, filtered and evaporated. Column chromatography (CHCl₃:acetone / 25:1) afforded desired derivative **94** (0.65 g, 56%) as a white solid.

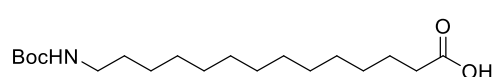
¹H NMR (400 MHz, CDCl₃) 1.22 – 1.37 (m, 16H), 1.38 – 1.50 (m, 4H), 1.43 (s, 9H), 1.75 (s, 3H), 1.78 (s, 3H), 2.05 – 2.14 (m, 2H), 3.09 (t, *J* = 7.0 Hz, 2H), 3.49 (dd, *J* = 6.8, 3.3 Hz, 1H), 4.48 (s, 1H).

¹³C NMR (100 MHz, CDCl₃) 26.5, 26.7, 26.8, 27.0, 28.4, 29.2, 29.3, 29.5 (5C), 30.1, 40.7, 46.1, 79.0, 104.8, 156.0, 165.7.

ESI MS: 450 ([M + Na]⁺).

HR ESI MS: calcd for C₂₃H₄₁O₆NNa 450.28261, found 450.28273.

14-(*tert*-Butoxycarbonylamino)tetradecanoic acid (95**)**



Intermediate **94** (0.55 g, 1.28 mmol, 1.0 eq.) was heated to 110 °C for 3 hours in DMF:H₂O / 10:1

(5 mL). After cooling, solvents were evaporated to obtain compound **95** (0.42 g, 94%) as a white solid.

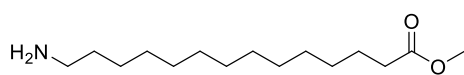
¹H NMR (400 MHz, CD₃OD) 1.30 (s, 18H), 1.43 (s, 9H), 1.45 – 1.50 (m, 2H), 1.55 – 1.65 (m, 2H), 2.27 (t, *J* = 7.4 Hz, 2H), 2.62 (s, 1H), 3.01 (t, *J* = 7.0 Hz, 2H).

¹³C NMR (100 MHz, CD₃OD) 26.2, 27.9, 28.8, 30.3, 30.4, 30.6, 30.7 (6C), 31.0, 35.1, 41.4, 79.7, 158.6, 177.8.

ESI MS: 342 ([M - H]⁻).

HR ESI MS: calcd for C₁₉H₃₆O₄N 342.26498, found 342.26445.

Methyl 14-aminotetradecanoate (**96**)



A solution of compound **95** (0.36 g, 1.04 mmol, 1.0 eq.) in MeOH (5 mL) was added slowly to a cooled (0 °C) solution of SOCl₂ (0.31 mL, 4.17 mmol, 4.0 eq.) in MeOH (10 mL). The reaction mixture was stirred at 0 °C for 1 hour, then the volatiles were evaporated. The residue was dissolved in TFA (2 mL) and stirred at room temperature for 15 minutes. TFA was then removed by a flow of nitrogen. The residue was partially dissolved in saturated solution of NaHCO₃ (10 mL) and extracted with EtOAc (3 × 15 mL). Organic phases were combined, dried using anhydrous MgSO₄. The solvent was then evaporated to obtain compound **96** (0.17 g, 62%) as a white solid.

¹H NMR (400 MHz, CD₃OD): 1.25 – 1.39 (m, 18H), 1.50 – 1.64 (m, 4H), 2.31 (t, *J* = 7.4 Hz, 2H), 3.75 (t, *J* = 7.4 Hz, 2H), 3.65 (s, 3H).

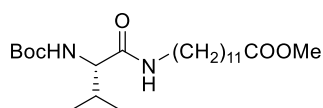
¹³C NMR (100 MHz, CD₃OD): 26.0, 27.8, 30.2, 30.4, 30.5, 30.6 (2C), 30.7 (3C), 31.4, 34.8, 41.7, 51.9, 176.0.

ESI MS: 258 ([M + H]⁺).

HR ESI MS: calcd for C₁₅H₃₁O₂NH 258.42276, found 258.42278.

4.2.2 Macrocyclic Inhibitors

Methyl (S)-12-((*tert*-butoxycarbonyl)amino)-3-methylbutanamido) dodecanoate (**101a**)



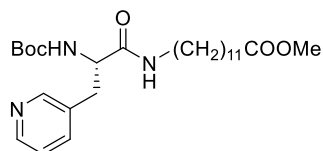
Boc-L-Val-OH (0.84 g, 3.88 mmol, 1.0 eq.) and HBTU (1.48 g, 3.88 mmol, 1.0 eq.) were dissolved in anhydrous DMF (15 mL) and DIPEA (2.0 mL, 11.32 mmol, 3.0 eq.) was added. The reaction mixture was stirred at room temperature for 25 minutes and then solution of methyl 12-aminododecanoate (0.86 g, 3.23 mmol, 0.8 eq.) in anhydrous DMF (5 mL) was added. The reaction mixture was stirred at room temperature for 6 hours under inert atmosphere. The reaction mixture was diluted with CHCl₃ (30 mL) and washed with saturated aq. NaHCO₃ (2 × 25 mL), 10% KHSO₄ (2 × 25 mL), brine (35 mL), dried over MgSO₄ and the organic solvent was evaporated. Column chromatography on silica gel (CHCl₃:MeOH / 30:1) afforded the desired product **101a** (1.23 g, 89%) as a white solid.

¹H NMR (400 MHz, CDCl₃): 0.90 – 0.95 (m, 6H), 1.20 – 1.34 (m, 14H), 1.42 (s, 9H), 1.45 – 1.51 (m, 2H), 1.56 – 1.64 (m, 2H), 2.03 – 2.14 (m, 1H), 2.29 (t, *J* = 7.7 Hz, 2H), 3.14 – 3.31 (m, 2H), 3.66 (s, 3H), 3.86 (dd, *J* = 8.8, 6.6 Hz, 1H), 5.20 (s, 1H), 6.31 (s, 1H).

¹³C NMR (100 MHz, CDCl₃): 18.1, 19.4, 25.0, 27.0, 28.4, 29.2, 29.3 (3C), 29.5, 29.6, 29.7, 31.0, 34.2, 39.5, 51.5, 60.2, 156.0, 171.7, 174.4.

ESI MS: 451 ([M + Na]⁺).

Methyl (S)-12-(2-((*tert*-butoxycarbonyl)amino)-3-(pyridin-3-yl)propanamido) dodecanoate (101b)



Boc-L-(3)Pal-OH (0.18 g, 0.68 mmol, 1.0 eq.) and HBTU (0.26 g, 0.68 mmol, 1.0 eq.) were dissolved in anhydrous DMF (3 mL) and DIPEA (0.34 mL, 1.98 mmol, 3.0 eq.) was added. The reaction mixture was stirred at room temperature for 25 minutes and then solution of methyl 12-aminododecanoate (0.15 g, 0.56 mmol, 0.8 eq.) in anhydrous DMF (2 mL) was added. The reaction mixture was stirred at room temperature for 6 h under inert atmosphere. The reaction mixture was diluted with CHCl₃ (15 mL) and washed with saturated NaHCO₃ (2 × 15 mL), 10% KHSO₄ (2 × 15 mL), brine (15 mL), dried over MgSO₄ and the organic solvent was evaporated. Column chromatography (CHCl₃:MeOH / 25:1) afforded the desired product **101b** (0.23 g, 85%) as an amorphous solid.

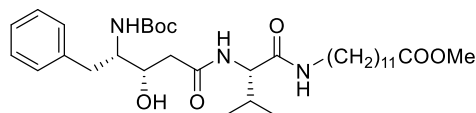
¹H NMR (400 MHz, CDCl₃): 1.12 – 1.29 (m, 14H), 1.34 (s, 9H), 1.35 – 1.37 (m, 2H), 1.58 (p, *J* = 7.5 Hz, 2H), 2.27 (t, *J* = 7.5 Hz, 2H), 2.91 – 2.98 (m, 1H), 3.05 – 3.21 (m, 3H), 3.63 (s, 3H), 4.29 – 4.36 (m, 1H), 5.38 (d, *J* = 8.6 Hz, 1H), 6.46 (s, 1H), 7.20 (dd *J* = 7.9, 4.8 Hz, 1H), 7.54 – 7.57 (m, 1H), 8.41 – 8.45 (m, 2H).

¹³C NMR (101 MHz, CDCl₃): 25.0, 26.9, 28.3, 29.2, 29.3 (2C), 29.5 (2C), 29.6, 29.8, 34.2, 36.1, 39.7, 51.5, 55.5, 80.3, 123.6, 133.1, 137.5, 147.8, 150.2, 155.5, 170.7, 174.4

ESI MS: 478 ([M + H]⁺).

HR ESI MS: calcd for C₂₆H₄₄O₅N₃ 478.32755; found 478.32760.

Methyl (6*S*,7*S*,11*S*)-6-benzyl-7-hydroxy-2,2-dimethyl-4,9,12-trioxo-11-isopropyl-3-oxa-5,10,13-triazapentacosan-25-oate (102a)



Compound **101a** (0.18 g, 0.42 mmol, 1.0 eq.) was dissolved in DCM (2 mL) and TFA (0.2 mL) was added. The reaction mixture was stirred at room temperature for 1 hour. The reaction mixture was diluted with DCM (10 mL), washed with saturated aq. NaHCO₃ (2 × 10 mL), dried over MgSO₄ and the organic solvent was evaporated to obtain crude amine. Boc-PheSta-OH (0.16 g, 0.51 mmol) and HBTU (0.2 g, 0.51 mmol, 1.0 eq.) were dissolved in anhydrous DMF (4 mL), and Et₃N (0.19 mL, 1.36 mmol, 2.5 eq.) was added. The reaction mixture was stirred at room temperature for

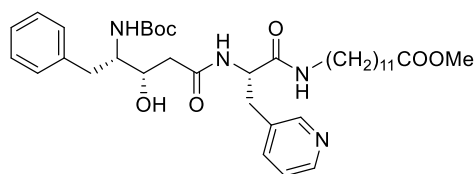
30 minutes and then solution of the amine (0.14 g, 0.42 mmol, 0.8 eq.) in anhydrous DMF (2 mL) was added. The reaction mixture was stirred at room temperature for 14 hours under inert atmosphere. The organic solvent was evaporated in vacuo. The residue was chromatographed on silica gel (CHCl₃:MeOH / 10:1) to obtain the desired product **102a** (0.16 g, 61%) as a white solid.

¹H NMR (400 MHz, CDCl₃): 0.87 – 0.94 (m, 6H), 1.21 – 1.33 (m, 14H), 1.37 (s, 9H), 1.42 – 1.49 (m, 2H), 1.56 – 1.64 (m, 2H), 2.01 – 2.14 (m, 1H), 2.25 – 2.33 (m, 3H), 2.50 (dd, *J* = 15.0, 9.2 Hz, 1H), 2.80 (s, 1H), 2.89 (d, *J* = 7.5 Hz, 2H), 3.09 – 3.27 (m, 2H), 3.65 (s, 3H), 3.68 – 3.74 (m, 1H), 3.94 – 4.01 (m, 1H), 4.17 (dd, *J* = 8.4, 7.0 Hz, 1H), 5.05 (d, *J* = 9.7 Hz, 1H), 6.46 (s, 1H), 6.77 (d, *J* = 8.4 Hz, 1H), 7.14 – 7.29 (m, 5H).

¹³C NMR (101 MHz, CDCl₃): 18.3, 19.4, 25.1, 27.0, 29.3 (2C), 29.4, 29.5 (2C), 29.6 (2C), 31.0, 34.2, 38.7, 39.8, 40.5, 51.6, 55.6, 59.0, 68.7, 79.7, 126.5, 128.5, 129.4, 138.3, 156.4, 171.0, 172.6, 174.5.

ESI MS: 642 ([M + Na]⁺).

Methyl (6*S*,7*S*,11*S*)-6-benzyl-7-hydroxy-2,2-dimethyl-4,9,12-trioxo-11-(pyridin-3-ylmethyl)-3-oxa-5,10,13-triazapentacosan-25-oate (102b)



Compound **101b** (0.18 g, 0.38 mmol, 1.0 eq.) was dissolved in DCM (2 mL) and TFA (0.2 mL) was added. The reaction mixture was stirred at room temperature for 1 hour. The reaction mixture was

diluted with DCM (10 mL), washed with saturated aq. NaHCO₃ (2 × 10 mL), dried over MgSO₄ and the organic solvent was evaporated in vacuo to obtain crude amine. Boc-PheSta-OH (0.14 g, 0.46 mmol, 1.2 eq.) and HBTU (0.17 g, 0.46 mmol, 1.2 eq.) were dissolved in anhydrous DMF (3 mL) and triethylamine (0.17 mL, 1.23 mmol, 3.2 eq.) was added. The reaction mixture was stirred at room temperature for 30 minutes and then solution of the amine (0.15 g, 0.38 mmol, 0.8 eq.) in anhydrous DMF (2 mL) was added. The reaction mixture was stirred at room temperature for 14 hours under inert atmosphere. The organic solvent was evaporated and the residue was chromatographed on silica gel (CHCl₃:MeOH / 15:1) to obtain the desired product **102b** (0.13 g, 51%) as an amorphous solid.

¹H NMR (400 MHz, CDCl₃): 1.11 – 1.36 (m, 25H), 1.55 – 1.62 (m, 2H), 2.23 – 2.28 (m, 1H), 2.27 (t, *J* = 7.5 Hz, 2H), 2.42 (dd, *J* = 14.9, 9.0 Hz, 1H), 2.82 – 2.88 (m, 2H), 2.96 – 3.15 (m, 4H), 3.64 (s, 3H), 3.67 – 3.74 (m, 1H), 3.92 – 3.97 (m, 1H), 4.77 (dd, *J* = 7.5, 7.2 Hz, 1H),

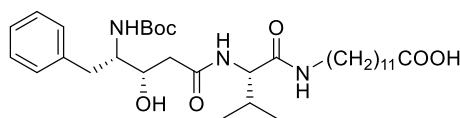
4.84 – 4.89 (m, 1H), 5.11 (d, $J = 9.7$ Hz, 1H), 5.79 – 5.81 (m, 1H), 7.03 – 7.08 (m, 1H), 7.11 – 7.24 (m, 6H), 7.51 – 7.55 (m, 1H), 8.36 – 8.40 (m, 2H).

^{13}C NMR (100 MHz, CDCl_3): 25.0, 26.9, 28.4, 29.2, 29.3 (2C), 29.5 (2C), 29.6 (2C), 34.2, 36.0, 38.7, 39.7, 40.9, 51.5, 54.0, 55.6, 68.5, 79.6, 123.6, 126.4, 128.5, 129.4, 132.8, 137.6, 138.3, 147.8, 150.2, 156.3, 170.2, 172.1, 174.4.

ESI MS: 691 ($[\text{M} + \text{Na}]^+$).

HR ESI MS: calcd for $\text{C}_{37}\text{H}_{56}\text{O}_7\text{N}_4\text{Na}$ 691.40412; found 691.40431.

(6*S*,7*S*,11*S*)-6-Benzyl-7-hydroxy-11-isopropyl-2,2-dimethyl-4,9,12-trioxo-3-oxa-5,10,13-triazapentacosan-25-oic acid (103a)



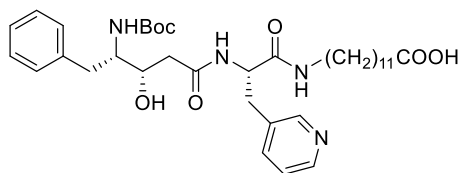
To a solution of methylester **102a** (0.15 g, 0.24 mmol, 1.0 eq.) in THF (2 mL), solution of 1M LiOH (2 mL, 2.0 mmol, 8.3 eq.) was added dropwise. The reaction mixture was stirred at room temperature for 2 hours, then acidified with 1M hydrochloric acid (5 mL) and CHCl_3 (10 mL) was added. Layers were separated and the product was extracted with CHCl_3 (3×10 mL). Organic layers were combined, dried over MgSO_4 , solvent was removed in vacuo. Residue was chromatographed on silica gel (CHCl_3 :MeOH / 20:1) to yield the desired acid **103a** (0.13 g, 89%) as a white amorphous solid.

^1H NMR (400 MHz, CDCl_3): 0.87 – 0.94 (m, 6H), 1.21 – 1.35 (m, 16H), 1.38 (s, 9H), 1.41 – 1.50 (m, 2H), 1.57 – 1.66 (m, 2H), 2.04 (s, 1H), 2.29 – 2.36 (m, 3H), 2.47 – 2.57 (m, 1H), 2.89 (d, $J = 7.4$ Hz, 2H), 3.10 – 3.30 (m, 2H), 3.48 (s, 1H), 3.67 – 3.76 (m, 1H), 3.94 – 4.02 (m, 1H), 4.18 – 4.24 (m, 1H), 5.01 – 5.12 (m, 1H), 6.73 (m, 1H), 7.16 – 7.29 (m, 5H).

^{13}C NMR (101 MHz, CDCl_3): 18.5, 19.4, 24.8, 26.8, 28.5, 28.9, 29.0 (2C), 29.1, 29.2, 29.3, 29.8, 31.0, 34.1, 38.7, 39.8, 40.2, 55.7, 59.2, 68.6, 79.7, 126.5, 128.3, 129.5, 138.3, 156.4, 171.5, 173.1, 175.6, 178.0.

ESI MS: 628 ($[\text{M} + \text{Na}]^+$).

(6*S*,7*S*,11*S*)-6-Benzyl-7-hydroxy-2,2-dimethyl-4,9,12-trioxo-11-(pyridin-3-ylmethyl)-3-oxa-5,10,13-triazapentacosan-25-oic acid (103b)



To a solution of methylester **102b** (0.13 g, 0.19 mmol, 1.0 eq.) in THF (2 mL), solution of 1M LiOH (2 mL, 2.0 mmol, 10.5 eq.) was added dropwise. The reaction mixture was stirred at room temperature for 2 hours,

then acidified with 1M hydrochloric acid (5 mL) and CHCl_3 (10 mL) was added. Layers

were separated and the product was extracted with CHCl_3 (3×10 mL). Organic layers were combined, dried over MgSO_4 and the solvent was removed. The residue was chromatographed on silica gel (CHCl_3 :MeOH / gradient 15:1 to 5:1) to yield the desired acid **103b** (0.10 g, 77%) as an amorphous solid.

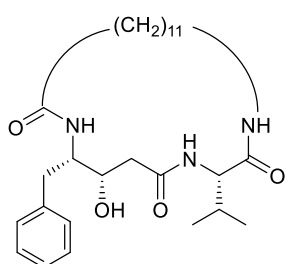
^1H NMR (400 MHz, CDCl_3): 1.12 – 1.35 (m, 25H), 1.54 – 1.62 (m, 2H), 2.21 – 2.27 (m, 1H), 2.28 (t, $J = 7.5$ Hz, 2H), 2.44 (dd, $J = 14.9, 9.0$ Hz, 1H), 2.83 – 2.88 (m, 2H), 2.96 – 3.14 (m, 4H), 3.68 – 3.75 (m, 1H), 3.92 – 3.98 (m, 1H), 4.78 (dd, $J = 7.5, 7.3$ Hz, 1H), 4.84 – 4.88 (m, 1H), 5.11 (d, $J = 9.7$ Hz, 1H), 5.77 – 5.81 (m, 1H), 7.02 – 7.09 (m, 1H), 7.12 – 7.24 (m, 6H), 7.51 – 7.54 (m, 1H), 8.36 – 8.41 (m, 2H).

^{13}C NMR (100 MHz, CDCl_3): 25.0, 26.9, 28.4, 29.2, 29.3 (2C), 29.5 (2C), 29.6 (2C), 34.2, 36.1, 38.7, 39.7, 41.0, 54.0, 55.7, 68.5, 79.6, 123.6, 126.4, 128.5, 129.3, 132.8, 137.6, 138.3, 147.8, 150.2, 156.4, 170.3, 172.1, 172.2.

ESI MS: 677 ($[\text{M} + \text{Na}]^+$).

HR ESI MS: calcd for $\text{C}_{36}\text{H}_{54}\text{O}_7\text{N}_4\text{Na}$ 677.38847; found 677.38849.

(3*S*,7*S*,8*S*)-8-Benzyl-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (105)



To a solution of acid **103a** (92 mg; 0.14 mmol, 1.0 eq.) and TSTU (42 mg, 0.14 mmol, 1.0 eq.) in anhydrous DMF (2 mL), DIPEA (26 μL , 0.14 mmol, 1.0 eq.) was added. The reaction mixture was stirred at room temperature for 2 hours under inert atmosphere. The reaction mixture was diluted with CHCl_3 (15 mL), washed with brine (2×10 mL), dried over MgSO_4 and the organic solvent

was evaporated in vacuo. The residue was dissolved in TFA (1 mL), and the reaction mixture was stirred at room temperature for 15 minutes. TFA was removed by a flow of nitrogen to obtain the crude amine. The crude amine was dissolved in anhydrous DMF (5 mL) and dropwise added to a solution of DIPEA (5 mL) in anhydrous DMF (5 mL) overnight. The reaction mixture was stirred overnight, evaporated and the crude product was purified by preparative HPLC to afford compound **105** (29 mg, 41%).

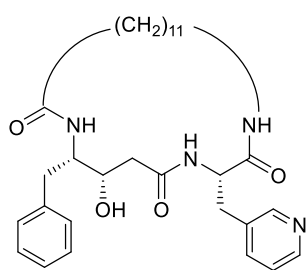
^1H NMR (400 MHz, $\text{DMSO}-d_6$): 0.84 (d, $J = 6.8$ Hz, 6H), 0.97 – 1.07 (m, 2H), 1.09 – 1.25 (m, 12H), 1.26 – 1.45 (m, 4H), 1.92 – 2.17 (m, 3H), 2.18 – 2.34 (m, 2H), 2.64 – 2.75 (m, 2H), 2.80 (dd, $J = 13.8, 4.1$ Hz, 1H), 3.29 (dt, $J = 13.8, 6.9$ Hz, 1H), 3.51 (s, 1H), 3.84 – 3.91 (m, 1H), 3.92 – 4.02 (m, 2H), 7.11 – 7.25 (m, 5H), 7.60 (d, $J = 8.3$ Hz, 1H), 7.66 (d, $J = 9.2$ Hz, 1H), 7.96 (dd, $J = 7.0, 4.4$ Hz, 1H).

¹³C NMR (100 MHz, DMSO-*d*₆): 18.0, 19.3, 24.9, 25.1, 26.3, 26.8, 27.2, 27.3, 27.4, 27.6, 28.0, 29.0, 29.5, 34.8, 36.6, 38.2, 53.8, 58.8, 69.4, 125.8, 128.0, 129.0, 139.6, 170.5, 170.8, 172.5.

ESI MS: 510 ([M + Na]⁺).

HR ESI MS: calcd for C₂₈H₄₅O₄N₃Na 510.33023; found 510.33010.

(3*S*,7*S*,8*S*)-8-benzyl-7-hydroxy-3-(pyridin-3-ylmethyl)-1,4,9-triazacyclohenicosane-2,5,10-trione (106)



To a solution of acid **103b** (90 mg, 0.14 mmol, 1.0 eq.) and TSTU (42 mg, 0.14 mmol, 1.0 eq.) in anhydrous DMF (2 mL) and DIPEA (24 μ L, 0.14 mmol, 1.0 eq.) was added. The reaction mixture was stirred at room temperature for 2 hours under inert atmosphere. The reaction mixture was diluted with CHCl₃ (15 mL), washed with brine (2 $\times$ 10 mL), dried over MgSO₄ and the organic solvent

was evaporated. The residue was dissolved in TFA (1 mL), and the reaction mixture was stirred at room temperature for 15 minutes. TFA was removed by a flow of nitrogen to obtain the crude amine. The crude amine was then dissolved in anhydrous DMF (5 mL) and dropwise added to a solution DIPEA (5 mL) in anhydrous DMF (5 mL) overnight. The reaction mixture was evaporated and the crude product was purified by preparative HPLC to afford compound **106** (35 mg, 47%).

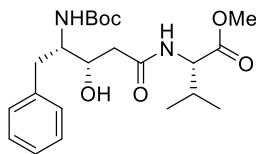
¹H NMR (400 MHz, CD₃OD): 0.98 – 1.07 (m, 2H), 1.12 – 1.32 (m, 14H), 1.35 – 1.53 (m, 4H), 2.10 – 2.22 (m, 2H), 2.26 – 2.33 (m, 2H), 2.75 – 2.89 (m, 3H), 3.06 – 3.14 (m, 1H), 3.34 – 3.41 (m, 2H), 3.94 – 3.99 (m, 1H), 4.05 (td, *J* = 7.1, 1.8 Hz, 1H), 4.64 (dd, *J* = 9.3, 5.4 Hz, 1H), 7.11 – 7.16 (m, 1H), 7.17 – 7.24 (m, 4H), 7.65 (d, *J* = 9.3 Hz, 1H), 7.92 (dd, *J* = 8.1, 5.6 Hz, 1H), 8.33 – 8.37 (m, 1H), 8.42 – 8.47 (m, 1H), 8.69 (d, *J* = 5.6 Hz, 1H), 8.72 (s, 1H).

¹³C NMR (100 MHz, CD₃OD): 26.6, 27.9, 28.5, 28.6, 28.8, 29.0, 29.1, 29.2, 35.4, 36.7, 39.0, 40.7, 40.8, 41.7, 55.6, 55.8, 71.5, 127.6, 128.1, 129.6, 130.4, 140.0, 140.5, 142.0, 144.2, 147.9, 172.5, 173.8, 176.2.

ESI MS: 559 ([M + Na]⁺).

HR ESI MS: calcd for C₃₁H₄₄O₄N₄Na 559.32548; found 559.32554.

(S)-Methyl 2-((3S,4S)-4-(*tert*-butoxycarbonylamino)-3-hydroxy-5-phenylpentanamido)-3-methylbutanoate (108)



Boc-PheSta-OH (0.19 g, 0.61 mmol, 1.0 eq.) and HBTU (0.23 g, 0.61 mmol, 1.0 eq.) were dissolved in anhydrous DMF (2.5 mL) and DIPEA (0.36 mL, 2.05 mmol, 3.4 eq.) was added. The reaction mixture was stirred at room temperature for 30 minutes and the solution of H-Val-OMe (90 mg, 0.50 mmol, 0.8 eq.) in anhydrous DMF (2 mL) was added. The reaction mixture was stirred at room temperature for 12 hours under inert atmosphere. The reaction mixture was then diluted with CHCl₃ (30 mL) and washed with saturated aq. NaHCO₃ (2 × 20 mL), 10% KHSO₄ (2 × 20 mL), brine (30 mL), dried over MgSO₄ and the organic solvent was evaporated in vacuo. The residue was chromatographed on silica gel (CHCl₃:MeOH / 20:1) to obtain the desired product **108** (0.16 g, 62%).

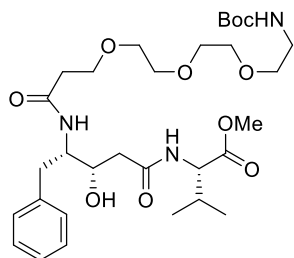
¹H NMR (400 MHz, CDCl₃): 0.90 (d, *J* = 6.9 Hz, 3H), 0.93 (d, *J* = 6.9 Hz, 3H), 1.40 (s, 9H), 2.08 – 2.20 (m, 1H), 2.29 (dd, *J* = 15.2, 3.1 Hz, 1H), 2.52 (dd, *J* = 15.2, 9.6 Hz, 1H), 2.85 (s, 1H), 2.91 (d, *J* = 7.6 Hz, 2H), 3.72 (s, 3H), 3.73 – 3.81 (m, 1H), 3.96 (d, *J* = 8.1 Hz, 1H), 4.49 (dd, *J* = 8.6, 4.9 Hz, 1H), 4.99 (d, *J* = 9.4 Hz, 1H), 6.53 (d, *J* = 8.4 Hz, 1H), 7.16 – 7.31 (m, 5H).

¹³C NMR (101 MHz, CDCl₃): 17.9, 19.1, 28.5, 31.2, 38.7, 40.0, 52.4, 55.6, 57.3, 68.2, 79.7, 126.5, 128.6, 129.5, 138.3, 156.4, 172.3, 172.9.

ESI MS: 445 ([M + Na]⁺).

HR ESI MS: calcd for C₂₂H₃₄O₆N₂Na 445.23091, found 445.23092.

(19S,20S,24S)-Methyl 19-benzyl-20-hydroxy-24-isopropyl-2,2-dimethyl-4,17,22-trioxo-3,8,11,14-tetraoxa-5,18,23-triazapentacosan-25-oate (109)



Compound **108** (0.15 g, 0.33 mmol, 1.0 eq.) was dissolved in TFA (1 mL) and the reaction mixture was stirred at room temperature for 15 minutes. TFA was removed by a flow of nitrogen to obtain crude amine. BocHN-PEG-3-CH₂CH₂COOH (0.10 g, 0.33 mmol, 1.0 eq.) and HBTU (0.12 g, 0.33 mmol, 1.0 eq.) were dissolved in anhydrous DMF (2.5 mL), and DIPEA (0.17 mL, 0.99 mmol, 3.0 eq.) was added. The reaction mixture was stirred at room temperature for 30 minutes. Then solution of the amine (0.10 g, 0.33 mmol, 1.0 eq.) in anhydrous DMF (2 mL) was added. The reaction mixture was stirred at room temperature for 16 hours under inert

atmosphere. Column chromatography on silica gel (CHCl₃:MeOH / 20:1) afforded the desired product **109** (0.15 g, 79%).

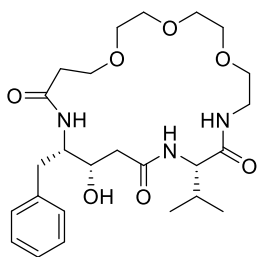
¹H NMR (400 MHz, CDCl₃): 0.92 (d, *J* = 6.9, 3H), 0.94 (d, *J* = 6.9 Hz, 3H), 1.43 (s, 9H), 2.10 – 2.21 (m, 1H), 2.38 (dd, *J* = 14.9, 4.7 Hz, 1H), 2.47 – 2.51 (m, 2H), 2.53 (dd, *J* = 15.0, 8.9 Hz, 1H), 2.93 – 2.98 (m, 4H), 3.29 (t, *J* = 5.1 Hz, 2H), 3.53 (t, *J* = 5.2 Hz, 2H), 3.58 – 3.65 (m, 8H), 3.70 (s, 3H), 3.99 – 4.04 (m, 1H), 4.15 (dd, *J* = 16.1, 8.4 Hz, 1H), 4.44 (dd, *J* = 8.4, 5.2 Hz, 1H), 7.23 – 7.28 (m, 5H).

¹³C NMR (100 MHz, CDCl₃): 18.0, 19.1, 28.6, 30.8, 36.8, 37.8, 40.4, 52.3, 54.8, 57.7, 67.3, 69.0, 70.3, 70.4 (4C), 70.5, 79.6, 126.6, 128.6, 129.4, 138.2, 156.3, 172.3, 172.9, 173.3.

ESI MS: 648 ([M + Na]⁺).

HR ESI MS: calcd for C₃₁H₅₁O₁₀N₃Na 648.34667, found 648.34678.

(12S,16S,17S)-17-Benzyl-16-hydroxy-12-isopropyl-1,4,7-trioxa-10,13,18-triazacyclo henicosane-11,14,19-trione (111)



To a solution of corresponding methylester **109** (65 mg, 0.10 mmol, 1.0 eq.) in THF (2 mL), a solution of LiOH (7 mg, 0.15 mmol, 1.5 eq.) in H₂O (2 mL) was added dropwise. The reaction mixture was stirred at room temperature for 2 hours and then the organic solvent was evaporated. The residue was dissolved in TFA (2 mL) and the reaction mixture was stirred at

room temperature for 15 minutes. TFA was removed by a flow of nitrogen to obtain crude amino acid salt. A solution of DIPEA (0.37 mL, 2.00 mmol, 20 eq.), DPPA (30 mg, 0.11 mmol, 1.1 eq.) in anhydrous DMF (10 mL) was slowly added to a solution of the crude amino acid salt in DMF (5 mL). The reaction mixture was stirred overnight, evaporated and the crude product was purified by preparative HPLC to afford the desired compound **111** (27 mg, 55%).

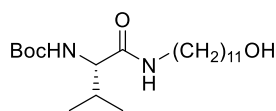
¹H NMR (400 MHz, DMSO-*d*₆): 0.82 (d, *J* = 6.9 Hz, 6H), 2.04 – 2.14 (m, 1H), 2.14 – 2.46 (m, 5H), 2.63 – 2.74 (m, 1H), 2.84 (dd, *J* = 13.9, 4.9 Hz, 1H), 2.91 – 3.02 (m, 1H), 3.36 – 3.48 (m, 10H), 3.56 – 3.63 (m, 3H), 3.93 – 4.01 (m, 2H), 4.05 (dd, *J* = 8.6, 5.7 Hz, 1H), 7.12 – 7.18 (m, 1H), 7.18 – 7.28 (m, 4H), 7.62 (d, *J* = 8.7 Hz, 1H), 7.76 (d, *J* = 9.2 Hz, 1H), 7.91 (t, *J* = 5.3 Hz, 1H).

¹³C NMR (100 MHz, DMSO-*d*₆): 17.6, 19.3, 29.4, 36.2, 36.4, 39.9, 53.3, 58.0, 66.9, 68.4, 68.9, 69.6, 69.7, 70.1, 70.4, 125.8, 128.0, 128.9, 139.5, 170.6, 170.7, 171.1.

ESI MS: 516 ([M + Na]⁺).

HR ESI MS: calcd for C₂₅H₃₉O₇N₃Na 516.26802, found 516.26782.

***tert*-Butyl (S)-((11-hydroxyundecyl)amino)-3-methyl-1-oxobutan-2-yl)carbamate (114)**



Boc-L-Val-OSu (1.00 g, 3.18 mmol, 1.0 eq.) and 11-aminoundecan-1-ol (0.65 g, 3.50 mmol, 1.1 eq.) were dissolved in anhydrous DMF (10 mL) and DIPEA (1.19 mL, 6.40 mmol, 2.0 eq.) was added. The reaction mixture was stirred at room temperature overnight. The reaction mixture was diluted with EtOAc (50 mL) and washed with saturated aq. NaHCO₃ (3 × 50 mL), 10% KHSO₄ (3 × 50 mL), brine (35 mL), dried over MgSO₄ and the organic solvent was evaporated in vacuo to yield the desired product **114** (1.13 g, 93%) as a white solid.

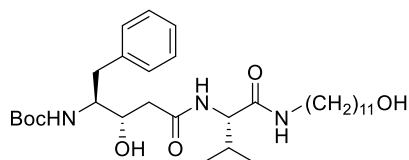
¹H NMR (400 MHz, CDCl₃): 0.90 (d, *J* = 6.8 Hz, 3H), 0.93 (d, *J* = 6.8 Hz, 3H), 1.20 – 1.34 (m, 12H), 1.42 (s, 9H), 1.44 – 1.58 (m, 4H), 1.78 – 1.94 (m, 1H), 2.00 – 2.15 (m, 1H), 3.13 – 3.30 (m, 2H), 3.61 (t, *J* = 6.6 Hz, 2H), 3.82 (dd, *J* = 9.0, 6.6 Hz, 1H), 5.16 (d, *J* = 9.1 Hz, 1H), 6.21 (t, *J* = 5.8 Hz, 1H).

¹³C NMR (100 MHz, CDCl₃): 18.1, 19.4, 25.8, 26.9, 28.4, 29.3, 29.5 (3C), 29.6, 30.9, 32.9, 39.6, 60.3, 63.0, 79.9, 156.1, 171.7.

ESI MS: 409 ([M + Na]⁺).

HR ESI MS: calcd for C₂₁H₄₂O₄N₂Na 409.30368, found 409.30348.

***tert*-Butyl ((2S,3S)-3-hydroxy-5-(((S)-1-((11-hydroxyundecyl)amino)-3-methyl-1-oxobutan-2-yl)amino)-5-oxo-1-phenylpentan-2-yl)carbamate (115)**



Compound **114** (1.0 g, 2.59 mmol, 1.0 eq.) was dissolved in DCM (5 mL) and TFA (5 mL) was added. The reaction mixture was stirred at room temperature for 1 hour. The organic solvent was evaporated to obtain the crude amine.

Boc-PheSta-OH (0.72 g, 2.33 mmol) and TSTU (0.70 g, 2.33 mmol, 1.0 eq.) were dissolved in anhydrous DMF (5 mL) and DIPEA (1.44 mL, 7.77 mmol, 3.0 eq.) was added. The reaction mixture was stirred at room temperature for 1 hour. Then solution of the amine in anhydrous DMF (2 mL) was added and the reaction mixture was stirred at room temperature for 2 hours. The reaction mixture was then diluted with EtOAc (50 mL) and washed with saturated aq. NaHCO₃ (3 × 50 mL), 10% KHSO₄ (3 × 50 mL), brine (35 mL), dried over MgSO₄ and the solvent was evaporated. Column chromatography on silica gel (DCM:MeOH / 20:1) afforded the desired product **115** (1.0 g, 75%).

¹H NMR (400 MHz, CD₃OD): 0.94 (d, *J* = 6.5 Hz, 3H), 0.96 (d, *J* = 6.5 Hz, 3H), 1.21 – 1.33 (m, 14H), 1.35 (s, 9H), 1.41 – 1.57 (m, 5H), 2.12 (dq, *J* = 13.5, 6.8 Hz, 1H), 2.45 (dd, *J* = 7.0, 3.1 Hz, 1H), 2.77 (dd, *J* = 13.7, 9.5 Hz, 1H), 2.87 (dd, *J* = 13.7, 5.4 Hz, 1H), 3.15 (ddp,

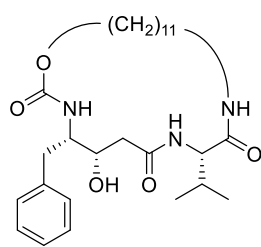
$J = 20.3, 13.4, 7.1$ Hz, 2H), 3.53 (t, $J = 6.7$ Hz, 2H), 3.74 (ddd, $J = 9.6, 5.4, 2.1$ Hz, 1H), 4.06 (td, $J = 7.0, 2.0$ Hz, 1H), 4.11 (d, $J = 6.0$ Hz, 1H), 7.11 – 7.28 (m, 5H).

^{13}C NMR (100 MHz, CD_3OD): 18.2, 19.7, 27.0, 28.0, 28.8, 30.4 (2C), 30.6, 30.7 (2C), 31.4, 33.7, 39.2, 40.4, 41.5, 57.0, 60.7, 63.0, 70.8, 80.1, 127.2, 129.2, 130.4, 140.2, 158.4, 173.6, 174.0.

ESI MS: 600 ($[\text{M} + \text{Na}]^+$).

HR ESI MS: calcd for $\text{C}_{32}\text{H}_{55}\text{O}_6\text{N}_3\text{Na}$ 600.39831, found 600.39754.

(4*S*,5*S*,9*S*)-4-Benzyl-5-hydroxy-9-isopropyl-1-oxa-3,8,11-triazacyclodocosane-2,7,10-trione (117)



Compound **115** (50 mg, 0.09 mmol, 1.0 eq.) and DSC (33 mg, 0.13 mmol, 1.5 eq.) were dissolved in anhydrous DMF (2 mL). The reaction mixture was stirred at room temperature overnight. The reaction mixture was then diluted with EtOAc (25 mL) and washed with saturated aq. NaHCO_3 (3×25 mL), 10% KHSO_4 (3×25 mL), brine (20 mL), dried over MgSO_4 and the organic solvent was evaporated in vacuo. The intermediate was then dissolved in DCM (2 mL) and TFA (2 mL) was added. The reaction mixture was stirred at room temperature for 20 minutes. The organic solvent was evaporated to obtain the crude amine. It was then dissolved in anhydrous DMF (2 mL) and dropwise added to a solution of DIPEA (1 mL) in anhydrous DMF (5 mL) overnight. The volatiles were evaporated, and the crude product was purified by preparative HPLC to afford compound **117** (23 mg, 52%).

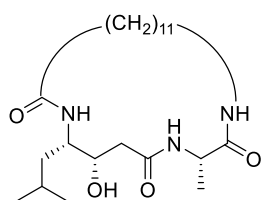
^1H NMR (400 MHz, $\text{DMSO}-d_6$): 0.83 (d, $J = 6.5$ Hz, 3H), 0.85 (d, $J = 6.5$ Hz, 3H), 1.09 – 1.53 (m, 16H), 1.85 – 1.99 (m, 1H), 2.19 – 2.30 (m, 2H), 2.57 (dd, $J = 13.8, 10.6$ Hz, 1H), 2.72 – 2.90 (m, 3H), 3.23 – 3.26 (m, 2H), 3.59 – 3.76 (m, 2H), 3.94 – 4.01 (m, 2H), 4.10 (dt, $J = 11.4, 5.9$ Hz, 1H), 6.92 (d, $J = 9.1$ Hz, 1H), 7.11 – 7.29 (m, 5H), 7.70 (d, $J = 8.5$ Hz, 1H), 7.75 (dd, $J = 7.0, 4.5$ Hz, 1H).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$): 18.3, 19.3, 25.0, 25.5, 27.3, 27.6, 27.9, 28.0, 28.2, 28.4 (2C), 29.6, 35.1, 38.1, 40.0, 56.5, 58.6, 63.3, 68.9, 125.8, 128.0, 129.0, 139.8, 156.2, 170.6, 171.0.

ESI MS: 526 ($[\text{M} + \text{Na}]^+$).

HR ESI MS: calcd for $\text{C}_{28}\text{H}_{45}\text{O}_5\text{N}_3\text{Na}$ 526.32514; found 526.32488.

(3*S*,7*S*,8*S*)-7-Hydroxy-8-isobutyl-3-methyl-1,4,9-triazacyclohenicosane-2,5,10-trione (118)



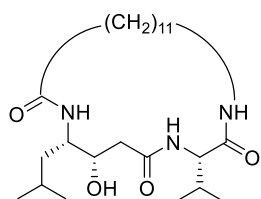
¹H NMR (400 MHz, DMSO-*d*₆): 0.80 (d, *J* = 6.4 Hz, 3H), 0.86 (d, *J* = 6.5 Hz, 3H), 1.10 – 1.64 (m, 21H), 2.06 – 2.26 (m, 4H), 2.76 – 2.86 (m, 1H), 3.18 – 3.28 (m, 1H), 3.72 – 3.82 (m, 1H), 3.88 (q, *J* = 6.7 Hz, 1H), 4.05 (p, *J* = 7.2 Hz, 1H), 4.79 (d, *J* = 5.3 Hz, 1H), 7.44 (d, *J* = 9.3 Hz, 1H), 7.93 (d, *J* = 7.0 Hz, 1H), 7.97 (t, *J* = 5.8 Hz, 1H).

¹³C NMR (101 MHz, DMSO-*d*₆): 17.6, 21.7, 23.3, 24.4, 24.9, 26.1, 26.6, 26.7, 27.1, 27.3, 27.8, 34.9, 38.3, 39.8, 48.9, 49.1, 69.5, 170.3, 172.2, 172.5.

ESI MS: 448 ([M + Na]⁺).

HR ESI MS: calcd for C₂₃H₄₃O₄N₃Na 448.31458; found 448.31401.

(3*S*,7*S*,8*S*)-7-Hydroxy-8-isobutyl-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (119)



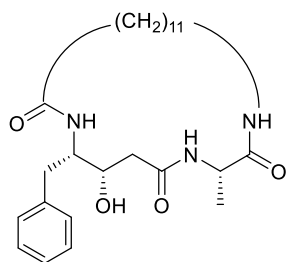
¹H NMR (400 MHz, DMSO-*d*₆): 0.80 (d, *J* = 6.5 Hz, 3H), 0.84 (d, *J* = 6.8 Hz, 6H), 0.86 (d, *J* = 6.5 Hz, 3H), 1.14 – 1.57 (m, 21H), 1.95 – 2.03 (m, 1H), 2.16 – 2.21 (m, 2H), 2.13 (dd, *J* = 14.5, 6.1 Hz, 1H), 2.21 (dd, *J* = 14.5, 8.0 Hz, 1H), 2.77 – 2.85 (m, 1H), 3.19 – 3.28 (m, 1H), 3.77 (dddd, *J* = 10.7, 10.5, 4.1, 1.6 Hz, 1H), 3.93 – 3.98 (m, 1H), 3.86 – 3.90 (m, 1H), 4.79 (bs, 1H), 7.44 (d, *J* = 9.4 Hz, 1H), 7.93 (d, *J* = 6.9 Hz, 1H), 7.97 (dd, *J* = 6.4, 5.1 Hz, 1H).

¹³C NMR (100 MHz, DMSO-*d*₆): 18.0, 19.2, 21.6, 23.3, 24.4, 24.8, 24.9, 26.1, 26.6, 26.7, 27.0, 27.2, 27.8, 29.5, 34.8, 38.3, 39.7, 48.8, 49.1, 69.5, 170.3, 172.1, 172.5.

ESI MS: 476 ([M + Na]⁺).

HR ESI MS: calcd for C₂₅H₄₇O₄N₃Na 476.34588; found 476.34557.

(3*S*,7*S*,8*S*)-8-Benzyl-7-hydroxy-3-methyl-1,4,9-triazacyclohenicosane-2,5,10-trione (120)

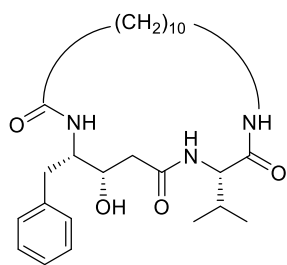


¹H NMR (600 MHz, DMSO-*d*₆): 0.92 – 1.02 (m, 2H), 1.09 – 1.45 (m, 18H), 2.06 – 2.12 (m, 3H), 2.16 – 2.26 (m, 2H), 2.68 – 2.81 (m, 3H), 3.21 (dq, *J* = 14.0, 7.1 Hz, 1H), 3.86 – 3.92 (m, 1H), 3.99 – 4.03 (m, 1H), 4.07 (p, *J* = 7.2 Hz, 1H), 5.02 (bs, 1H), 7.11 – 7.25 (m, 5H), 7.64 (d, *J* = 9.3 Hz, 1H), 7.91 – 7.99 (m, 2H).

¹³C NMR (150 MHz, DMSO-*d*₆): 17.6, 24.8 (2C), 26.0, 26.5, 26.6, 27.0, 27.1, 27.2, 27.6, 30.7, 34.8, 36.9, 38.3, 39.9, 40.4, 48.9, 53.2, 69.4, 125.8, 127.9, 128.9, 139.5, 170.1, 172.2, 172.4.
ESI MS: 482 ([M + Na]⁺).

HR ESI MS: calcd for C₂₆H₄₁O₄N₃Na 482.29893; found 482.29871.

(3*S*,7*S*,8*S*)-8-Benzyl-7-hydroxy-3-isopropyl-1,4,9-triazacycloicosane-2,5,10-trione (121)



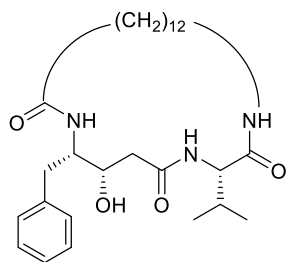
¹H NMR (400 MHz, DMSO-*d*₆): 0.82 (d, *J* = 5.2 Hz, 3H), 0.84 (d, *J* = 6.0 Hz, 3H), 0.92 – 1.04 (m, 1H), 1.04 – 1.28 (m, 11H), 1.29 – 1.45 (m, 4H), 1.89 – 2.04 (m, 2H), 2.12 (dt, *J* = 13.2, 6.3 Hz, 1H), 2.18 – 2.30 (m, 2H), 2.58 – 2.69 (m, 1H), 2.76 – 2.89 (m, 2H), 3.22 – 3.29 (m, 1H), 3.91 – 4.04 (m, 3H), 4.99 (d, *J* = 5.2 Hz, 1H), 7.12 – 7.26 (m, 5H), 7.64 (d, *J* = 8.7 Hz, 1H), 7.68 (d, *J* = 8.9 Hz, 1H), 7.86 (dd, *J* = 7.2, 4.4 Hz, 1H).

¹³C NMR (100 MHz, DMSO-*d*₆): 18.2, 19.3, 24.6, 26.9, 27.0, 27.3, 27.5, 27.8, 28.0, 29.4, 35.0, 35.6, 37.6, 40.0, 54.0, 58.5, 69.9, 125.8, 128.0, 128.9, 139.7, 170.6, 170.9, 172.2.

ESI MS: 496 ([M + Na]⁺).

HR ESI MS: calcd for C₂₇H₄₃O₄N₃Na 496.31458; found 496.31396.

(3*S*,7*S*,8*S*)-8-Benzyl-7-hydroxy-3-isopropyl-1,4,9-triazacyclodocosane-2,5,10-trione (122)



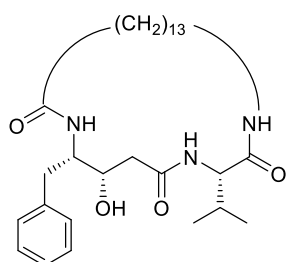
¹H NMR (400 MHz, DMSO-*d*₆): 0.77 – 0.88 (m, 6H), 1.03 – 1.52 (m, 20H), 1.92 – 2.03 (m, 2H), 2.07 – 2.16 (m, 1H), 2.18 – 2.30 (m, 2H), 2.60 – 2.70 (m, 1H), 2.74 – 2.88 (m, 2H), 3.23 – 3.33 (m, 1H), 3.91 – 3.99 (m, 2H), 4.04 (dd, *J* = 8.7, 6.3 Hz, 1H), 7.11 – 7.29 (m, 5H), 7.58 (d, *J* = 8.7 Hz, 1H), 7.64 (d, *J* = 9.0 Hz, 1H), 7.91 (dd, *J* = 7.1, 4.5 Hz, 1H).

¹³C NMR (100 MHz, DMSO-*d*₆): 18.0, 19.3, 24.9, 25.0, 27.0 (2C), 27.2, 27.3, 27.4, 27.6, 28.3, 29.8, 35.2, 36.1, 38.1, 40.0, 53.8, 58.1, 68.9, 125.8, 128.0, 128.9, 139.6, 170.5, 170.9, 172.3.

ESI MS: 524 ([M + Na]⁺).

HR ESI MS: calcd for C₂₉H₄₇O₄N₃Na 524.34588; found 524.34564.

(3*S*,7*S*,8*S*)-8-benzyl-7-hydroxy-3-isopropyl-1,4,9-triazacyclotricosane-2,5,10-trione (123)



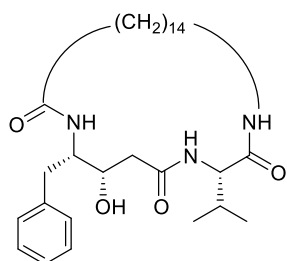
¹H NMR (400 MHz, CD₃OD): 0.94 (d, *J* = 6.9 Hz, 3H), 0.95 (d, *J* = 6.9 Hz, 3H), 1.11 – 1.34 (m, 18H), 1.40 – 1.52 (m, 4H), 2.08 – 2.23 (m, 3H), 2.37 – 2.49 (m, 2H), 2.78 – 2.94 (m, 3H), 3.41 (dt, *J* = 14.0, 7.0 Hz, 1H), 4.03 (ddd, *J* = 10.0, 4.9, 1.7 Hz, 1H), 4.11 – 4.17 (m, 2H), 7.13 – 7.25 (m, 5H).

¹³C NMR (100 MHz, CD₃OD): 16.8, 18.3, 25.2, 25.7, 26.8, 27.0, 27.2, 27.4, 27.5, 27.8, 27.9, 28.1, 28.6, 29.9, 35.6, 37.5, 39.0, 40.2, 54.1, 59.4, 69.9, 125.9, 127.9, 128.9, 138.9, 172.4 (2C), 174.6.

ESI MS: 538 ([M + Na]⁺).

HR ESI MS: calcd for C₃₀H₄₉O₄N₃Na 538.36153; found 538.36136.

(3*S*,7*S*,8*S*)-8-benzyl-7-hydroxy-3-isopropyl-1,4,9-triazacyclotetracosane-2,5,10-trione (124)



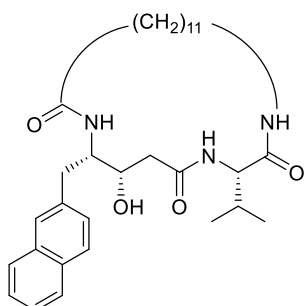
¹H NMR (400 MHz, CD₃OD): 0.93 (d, *J* = 7.1 Hz, 3H), 0.95 (d, *J* = 7.1 Hz, 3H), 1.11 – 1.35 (m, 20H), 1.39 – 1.49 (m, 4H), 2.05 – 2.24 (m, 3H), 2.38 – 2.52 (m, 2H), 2.78 – 3.00 (m, 3H), 3.34 – 3.36 (m, 1H), 4.04 – 4.09 (m, 1H), 4.12 – 4.21 (m, 2H), 7.12 – 7.25 (m, 5H).

¹³C NMR (100 MHz, CD₃OD): 18.0, 19.7, 26.9, 27.4, 27.9, 28.3, 28.5, 28.6, 28.7, 29.0, 29.2, 29.6, 29.7, 30.3, 31.3, 37.1, 38.9, 40.4, 41.5, 55.4, 60.5, 71.2, 127.3, 129.3, 130.3, 140.3, 173.7, 173.8, 176.0.

ESI MS: 552 ([M + Na]⁺).

HR ESI MS: calcd for C₃₁H₅₁O₄N₃Na 552.37718; found 552.37729.

(3*S*,7*S*,8*S*)-7-Hydroxy-3-isopropyl-8-(naphthalen-2-ylmethyl)-1,4,9-triazacyclohenicosane-2,5,10-trione (125)



¹H NMR (400 MHz, DMSO-*d*₆): 0.82 (d, *J* = 6.7 Hz, 3H), 0.83 (d, *J* = 6.7 Hz, 3H), 0.94 – 1.41 (m, 18H), 1.93 – 2.03 (m, 1H), 2.03 – 2.16 (m, 2H), 2.25 (dd, *J* = 14.0, 6.1 Hz, 1H), 2.33 (dd, *J* = 14.0, 7.8 Hz, 1H), 2.61 – 2.68 (m, 1H), 2.88 (dd, *J* = 13.6, 10.3 Hz, 1H), 2.95 (dd, *J* = 12.6, 4.2 Hz, 1H), 3.24 – 3.32 (m, 1H), 3.93 (dd, *J* = 8.0, 6.2 Hz, 1H), 3.93 – 3.98 (m, 1H), 4.03 – 4.08 (m,

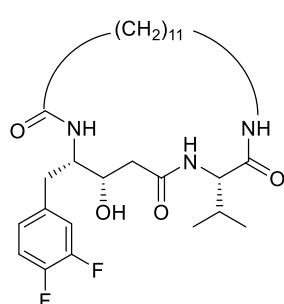
1H), 5.10 (bs, 1H), 7.38 (dd, $J = 8.4, 1.5$ Hz, 1H), 7.39 – 7.47 (m, 2H), 7.59 (d, $J = 8.0$ Hz, 1H), 7.70 (bs, 1H), 7.71 – 7.87 (m, 4H), 7.98 (dd, $J = 7.0, 4.4$ Hz, 1H).

^{13}C NMR (100 MHz, DMSO- d_6): 17.9, 19.2, 24.6, 24.8, 25.7, 26.4, 26.7, 27.0, 27.2, 27.7, 29.4, 34.8, 37.2, 38.2, 40.2, 53.7, 58.8, 69.8, 125.0, 125.7, 127.1, 127.2, 127.3, 127.4, 127.8, 131.6, 133.0, 137.2, 170.4, 170.7, 172.5.

ESI MS: 560 ($[\text{M} + \text{Na}]^+$).

HR ESI MS: calcd for $\text{C}_{32}\text{H}_{47}\text{O}_4\text{N}_3\text{Na}$ 560.34588; found 560.34551.

(3*S*,7*S*,8*S*)-8-(3,4-Difluorobenzyl)-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (126)



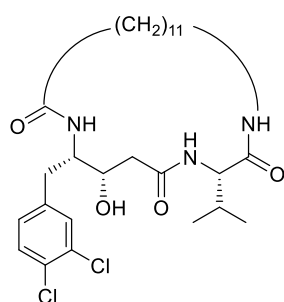
^1H NMR (400 MHz, DMSO- d_6): 0.83 (d, $J = 6.8$ Hz, 6H), 0.88 – 1.48 (m, 18H), 1.95 – 2.05 (m, 1H), 2.09 – 2.19 (m, 2H), 2.21 (dd, $J = 14.0, 5.8$ Hz, 1H), 2.32 (dd, $J = 14.0, 8.3$ Hz, 1H), 2.62 – 2.70 (m, 1H), 2.70 – 2.79 (m, 2H), 3.24 – 3.32 (m, 1H), 3.75 – 3.83 (m, 1H), 3.93 (dd, $J = 7.9, 5.9$ Hz, 1H), 3.98 – 4.03 (m, 1H), 5.11 (d, $J = 5.2$ Hz, 1H), 7.01 – 7.06 (m, 1H), 7.21 – 7.30 (m, 2H), 7.60 (d, $J = 7.9$ Hz, 1H), 7.70 (d, $J = 9.4$ Hz, 1H), 7.98 (dd, $J = 7.0, 4.4$ Hz, 1H).

^{13}C NMR (100 MHz, DMSO- d_6): 17.9, 19.2, 24.6, 24.8, 25.6, 26.3, 26.5, 27.0, 27.2, 27.7, 29.3, 34.7, 36.3, 38.2, 40.3, 53.5, 58.8, 69.9, 116.8 (d, $J_{\text{CCF}} = 17$ Hz), 117.3 (dd, $J_{\text{CCF}} = 135.6$ Hz, $J_{\text{CCCF}} = 16.4$ Hz), 125.7 (dd, $J_{\text{CCCF}} = 6$ Hz, $J_{\text{CCCCF}} = 3.2$ Hz), 137.4 (dd, $J_{\text{CCCF}} = 5.8$ Hz, $J_{\text{CCCCF}} = 3.2$ Hz), 147.9 (dd, $J_{\text{CF}} = 243$ Hz, $J_{\text{CCF}} = 12.6$ Hz), 148.9 (dd, $J_{\text{CF}} = 244.3$ Hz, $J_{\text{CCF}} = 12.6$ Hz), 170.3, 170.7, 172.6.

ESI MS: 546 ($[\text{M} + \text{Na}]^+$).

HR ESI MS: calcd for $\text{C}_{28}\text{H}_{43}\text{O}_4\text{N}_3\text{F}_2\text{Na}$ 546.31138; found 546.31104.

(3*S*,7*S*,8*S*)-8-(3,4-Dichlorobenzyl)-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (127)



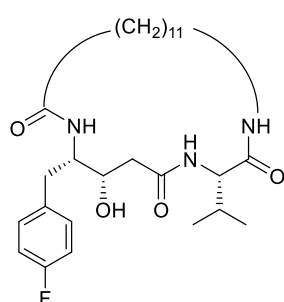
^1H NMR (400 MHz, DMSO- d_6): 0.83 (d, $J = 6.8$ Hz, 6H), 0.84 – 1.49 (m, 18H), 1.95 – 2.06 (m, 1H), 2.08 – 2.20 (m, 2H), 2.20 (dd, $J = 14.0, 5.5$ Hz, 1H), 2.33 (dd, $J = 14.0, 8.6$ Hz, 1H), 2.61 – 2.68 (m, 1H), 2.69 – 2.79 (m, 2H), 3.25 – 3.34 (m, 1H), 3.79 (dddd, $J = 9.7, 9.7, 4.6, 1.2$ Hz, 1H), 3.92 (dd, $J = 7.8, 5.8$ Hz, 1H), 3.99 – 4.04 (m, 1H), 5.12 (bs, 1H), 7.20 (dd, $J = 8.3, 1.9$ Hz, 1H), 7.46 (d, $J = 8.3$ Hz, 1H), 7.47 (d, $J = 1.9$ Hz, 1H), 7.61 (d, $J = 7.9$ Hz, 1H), 7.72 (d, $J = 9.4$ Hz, 1H), 7.98 (dd, $J = 7.1, 4.5$ Hz, 1H).

¹³C NMR (100 MHz, DMSO-*d*₆): 17.8, 19.2, 24.5, 24.7, 25.6, 26.2, 26.4, 26.9, 27.0, 27.1, 27.6, 29.3, 34.7, 36.4, 38.2, 40.3, 53.4, 58.9, 70.0, 128.6, 129.5, 130.0, 130.5, 131.2, 140.9, 170.3, 170.6, 172.6.

ESI MS: 578 ([M + Na]⁺).

HR ESI MS: calcd for C₂₈H₄₃O₄N₃Cl₂Na 578.25228; found 578.25196.

(3*S*,7*S*,8*S*)-8-(4-Fluorobenzyl)-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (128)



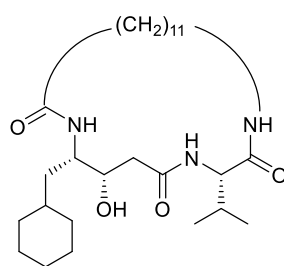
¹H NMR (400 MHz, DMSO-*d*₆): 0.84 (d, *J* = 6.8 Hz, 6H), 0.94 – 1.47 (m, 18H), 1.95 – 2.03 (m, 1H), 2.03 – 2.10 (m, 1H), 2.10 – 2.17 (m, 1H), 2.22 (dd, *J* = 14.0, 6.4 Hz, 1H), 2.29 (dd, *J* = 14.0, 7.5 Hz, 1H), 2.65 – 2.72 (m, 2H), 2.77 (dd, *J* = 13.7, 4.0 Hz, 1H), 3.25 – 3.34 (m, 1H), 3.82 (dddd, *J* = 11.0, 10.9, 3.8, 1.7 Hz, 1H), 3.94 (dd, *J* = 8.0, 6.2 Hz, 1H), 3.96 – 4.01 (m, 1H), 5.05 (bs, 1H), 7.00 – 7.06 (m, 2H), 7.20 – 7.25 (m, 2H), 7.60 (d, *J* = 8.1 Hz, 1H), 7.67 (d, *J* = 9.2 Hz, 1H), 7.97 (dd, *J* = 7.0, 4.4 Hz, 1H).

¹³C NMR (100 MHz, DMSO-*d*₆): 17.9, 19.2, 24.8, 25.0, 26.0, 26.6, 26.9, 27.1, 27.2, 27.4, 27.8, 29.4, 34.7, 36.0, 38.1, 40.1, 53.8, 58.7, 69.5, 114.6 (d, *J*_{CCF} = 20.9 Hz), 130.7 (d, *J*_{CCCF} = 8 Hz), 135.7 (d, *J*_{CCCCF} = 3 Hz), 160.7 (d, *J*_{CF} = 240.9 Hz), 170.4, 170.7, 172.5.

ESI MS: 528 ([M + Na]⁺).

HR ESI MS: calcd for C₂₈H₄₄O₄N₃FNa 528.32081; found 528.32039.

(3*S*,7*S*,8*S*)-8-(Cyclohexylmethyl)-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (129)



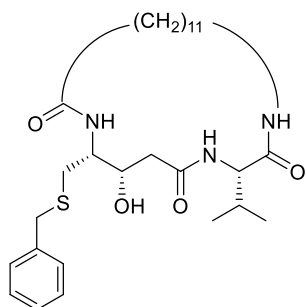
¹H NMR (400 MHz, DMSO-*d*₆): 0.84 (d, *J* = 6.8 Hz, 3H), 0.85 (d, *J* = 6.8 Hz, 3H), 1.06 – 1.73 (m, 31H), 1.97 – 2.05 (m, 1H), 2.15 (dd, *J* = 14.2, 6.3 Hz, 1H), 2.19 (dd, *J* = 6.3, 6.1 Hz, 2H), 2.28 (dd, *J* = 14.2, 7.8 Hz, 1H), 2.71 – 2.79 (m, 1H), 3.28 – 3.37 (m, 1H), 3.74 – 3.81 (m, 1H), 3.93 (dd, *J* = 8.0, 5.9 Hz, 1H), 3.82 – 3.87 (m, 1H), 4.80 (d, *J* = 5.5 Hz, 1H), 7.45 (d, *J* = 9.2 Hz, 1H), 8.02 (dd, *J* = 7.1, 4.4 Hz, 1H).

¹³C NMR (100 MHz, DMSO-*d*₆): 17.9, 19.2, 24.9, 25.0, 25.8, 26.0, 26.1, 26.6, 27.0, 27.1, 27.2, 27.4, 27.9, 29.4, 31.9, 33.6, 33.9, 34.8, 38.2, 48.6, 58.7, 69.8, 170.6, 170.7, 172.5.

ESI MS: 516 ([M + Na]⁺).

HR ESI MS: calcd for C₂₈H₅₂O₄N₃ 494.39523; found 494.39501.

(3*S*,7*S*,8*R*)-8-((Benzylthio)methyl)-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (130)



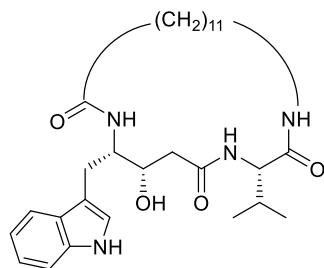
¹H NMR (400 MHz, DMSO-*d*₆): 0.85 (d, *J* = 6.7 Hz, 3H), 0.86 (d, *J* = 6.7 Hz, 3H), 1.15 – 1.62 (m, 18H), 1.94 – 2.05 (m, 1H), 2.13 – 2.24 (m, 4H), 2.48 – 2.55 (m, 1H), 2.59 (dd, *J* = 13.4, 5.9 Hz, 1H), 2.78 – 2.85 (m, 1H), 3.27 – 3.37 (m, 1H), 3.69 (s, 2H), 3.93 – 3.97 (m, 1H), 3.98 (dd, *J* = 8.1, 6.6 Hz, 1H), 4.06 – 4.11 (m, 1H), 4.99 (bs, 1H), 7.21 – 7.26 (m, 1H), 7.28 – 7.31 (m, 4H), 7.63 (d, *J* = 9.2 Hz, 1H), 7.66 (d, *J* = 8.2 Hz, 1H), 7.92 (dd, *J* = 7.0, 4.5 Hz, 1H).

¹³C NMR (100 MHz, DMSO-*d*₆): 18.0, 19.2, 24.9, 25.2, 26.6, 27.2, 27.4, 27.5, 27.9, 28.1, 29.5, 31.9, 34.7, 34.8, 38.1, 39.8, 51.0, 58.6, 68.0, 126.7, 128.3, 128.8, 138.5, 170.4, 170.8, 172.7.

ESI MS: 556 ([M + Na]⁺).

HR ESI MS: calcd for C₂₉H₄₇O₄N₃NaS 556.31795; found 556.31774.

(3*S*,7*S*,8*S*)-8-((1*H*-Indol-3-yl)methyl)-7-hydroxy-3-isopropyl-1,4,9-triazacyclohenicosane-2,5,10-trione (132)



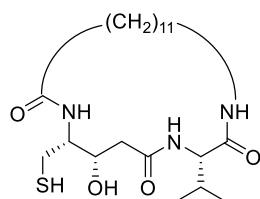
¹H NMR (400 MHz, DMSO-*d*₆): 0.83 (d, *J* = 6.8 Hz, 6H), 1.08 – 1.55 (m, 18H), 1.92 – 2.01 (m, 1H), 2.06 (ddd, *J* = 13.8, 7.2, 4.7 Hz, 1H), 2.16 (ddd, *J* = 13.8, 8.7, 4.8 Hz, 1H), 2.23 (dd, *J* = 14.2, 6.2 Hz, 1H), 2.28 (dd, *J* = 14.2, 7.4 Hz, 1H), 2.65 – 2.72 (m, 1H), 2.75 (dd, *J* = 14.8, 8.9 Hz, 1H), 2.96 (dd, *J* = 14.8, 5.4 Hz, 1H), 3.22 – 3.30 (m, 1H), 3.95 (dd, *J* = 8.4, 6.7 Hz, 1H), 4.00 – 4.05 (m, 1H), 5.01 (bs, 1H), 6.94 (ddd, *J* = 9.6, 8.7, 0.9 Hz, 1H), 7.03 (ddd, *J* = 9.0, 8.0, 1.0 Hz, 1H), 7.08 (d, *J* = 2.2 Hz, 1H), 7.25 – 7.35 (m, 1H), 7.54 (d, *J* = 8.0 Hz, 1H), 7.59 (d, *J* = 8.6 Hz, 1H), 7.62 (d, *J* = 9.0 Hz, 1H), 7.88 (dd, *J* = 7.0, 4.6 Hz, 1H).

¹³C NMR (100 MHz, DMSO-*d*₆): 18.1, 19.2, 24.9, 25.2, 26.3, 26.6, 27.1, 27.4, 27.6, 27.9, 28.0, 29.5, 34.9, 38.0, 39.7, 52.9, 58.6, 68.7, 111.2, 111.7, 118.0, 118.3, 119.8, 120.7, 122.0, 122.9, 127.5, 128.9, 136.2, 170.6, 170.8, 172.4.

ESI MS: 549 ([M + Na]⁺).

HR ESI MS: calcd for C₃₀H₄₆O₄N₄Na 549.34113; found 549.34082.

(3*S*,7*S*,8*R*)-7-Hydroxy-3-isopropyl-8-(mercaptomethyl)-1,4,9-triazacyclohenicosane-2,5,10-trione (133)



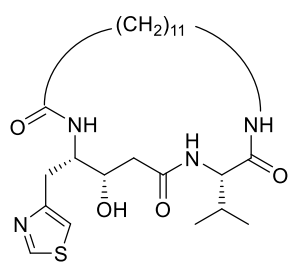
¹H NMR (400 MHz DMSO-*d*₆): 0.84 (d, *J* = 7.0 Hz, 3H), 0.85 (d, *J* = 7.0 Hz, 3H), 1.15 – 1.62 (m, 18H), 1.92 – 2.03 (m, 1H), 2.10 – 2.25 (m, 5H), 2.48 – 2.55 (m, 1H), 2.67 (ddd, *J* = 13.7, 9.1, 5.5 Hz, 1H), 2.78 – 2.86 (m, 1H), 3.27 – 3.34 (m, 1H), 3.72 (dddd, *J* = 8.7, 8.6, 5.7, 2.1 Hz, 1H), 3.98 (dd, *J* = 7.6, 7.4 Hz, 1H), 4.07 – 4.13 (m, 1H), 4.99 (bs, 1H), 7.60 (d, *J* = 9.1 Hz, 1H), 7.62 (d, *J* = 8.7 Hz, 1H), 7.88 (dd, *J* = 6.6, 4.7 Hz, 1H).

¹³C NMR (100 MHz, DMSO-*d*₆): 18.2, 19.2, 24.7, 25.0, 25.3, 26.8, 27.3, 27.5, 27.6, 27.7, 28.0, 28.2, 29.6, 34.9, 38.1, 39.6, 55.5, 58.5, 67.9, 170.3, 170.8, 172.8.

ESI MS: 442 ([M - H]⁻).

HR ESI MS: calcd for C₂₂H₄₀O₄N₃S 442.27450; found 442.27416.

(3*S*,7*S*,8*S*)-7-Hydroxy-3-isopropyl-8-(thiazol-4-ylmethyl)-1,4,9-triazacyclohenicosane-2,5,10-trione (134)



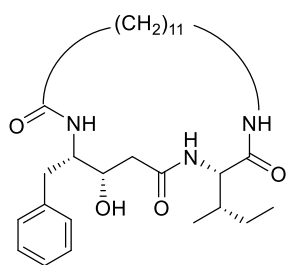
¹H NMR (400 MHz, DMSO-*d*₆): 0.84 (d, *J* = 6.8 Hz, 3H), 0.85 (d, *J* = 6.8 Hz, 3H), 1.03 – 1.48 (m, 18H), 1.96 – 2.04 (m, 1H), 2.04 – 2.10 (m, 1H), 2.10 – 2.17 (m, 1H), 2.23 (dd, *J* = 14.2, 6.8 Hz, 1H), 2.29 (dd, *J* = 14.2, 7.0 Hz, 1H), 2.73 – 2.81 (m, 1H), 2.92 (dd, *J* = 14.8, 10.0 Hz, 1H), 2.99 (dd, *J* = 14.8, 4.4 Hz, 1H), 3.27 – 3.38 (m, 1H), 4.09 (dddd, *J* = 10.0, 9.8, 4.4, 2.0 Hz, 1H), 3.96 (dd, *J* = 8.2, 6.2 Hz, 1H), 3.99 (ddd, *J* = 6.9, 6.8, 2.0 Hz, 1H), 5.05 (bs, 1H), 7.25 (d, *J* = 2.0 Hz, 1H), 7.62 (d, *J* = 8.2 Hz, 1H), 7.65 (d, *J* = 9.5 Hz, 1H), 7.96 (dd, *J* = 7.3, 4.4 Hz, 1H), 8.97 (d, *J* = 2.0 Hz, 1H).

¹³C NMR (100 MHz, DMSO-*d*₆): 18.0, 19.2, 24.7, 25.1, 26.2, 26.8, 27.2, 27.3, 27.5, 28.0, 29.5, 32.5, 34.8, 38.1, 39.8, 51.7, 58.7, 69.3, 114.3, 153.1, 154.9, 170.4, 170.7, 172.5.

ESI MS: 517 ([M + Na]⁺).

HR ESI MS: calcd for C₂₅H₄₂O₄N₄NaS 517.28190; found 517.28161.

(3*S*,7*S*,8*S*)-8-Benzyl-3-((*R*)-*s*-butyl)-7-hydroxy-1,4,9-triazacyclohenicosane-2,5,10-trione (135)



¹H NMR (400 MHz, DMSO-*d*₆): 0.76 – 0.84 (m, 6H), 0.99 – 1.47 (m, 20H), 1.66 – 1.76 (m, 1H), 1.99 – 2.17 (m, 2H), 2.18 – 2.28 (m, 2H), 2.62 – 2.73 (m, 2H), 2.81 (dd, *J* = 13.8, 4.3 Hz, 1H), 3.25 – 3.31 (m, 1H), 3.83 – 3.93 (m, 1H), 3.93 – 4.02 (m, 2H), 5.02 (d, *J* = 5.3 Hz, 1H), 7.10 – 7.26 (m, 5H), 7.60 (d, *J* = 8.2 Hz, 1H), 7.65 (d, *J* = 9.1 Hz, 1H), 7.95 (dd, *J* = 7.3, 4.3 Hz, 1H).

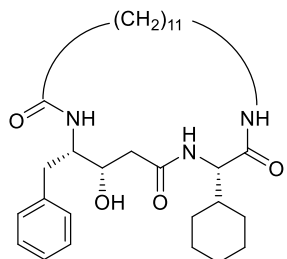
¹³C NMR (100 MHz, DMSO-*d*₆): 11.4, 15.5, 24.4, 24.8, 25.1, 26.3, 26.8, 27.2, 27.3, 27.4, 27.7, 28.0, 34.8, 35.9, 36.5, 38.1, 40.0, 53.8, 57.7, 69.4, 125.8, 128.0, 128.9, 139.6, 170.4, 170.8, 172.4.

ESI MS: 524 ([M + Na]⁺).

HR ESI MS: calcd for C₃₁H₄₅O₄N₄Na 524.34588; found 524.34528.

(3*S*,7*S*,8*S*)-8-Benzyl-3-cyclohexyl-7-hydroxy-1,4,9-triazacyclohenicosane-2,5,10-trione (136)

NMR spectra for this compound were recorded at 80 °C, due to limited solubility of the compound.



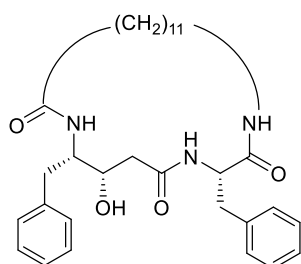
¹H NMR (600 MHz, Acetic acid-*d*₄): 1.00 – 1.87 (m, 31H), 2.13 – 2.19 (m, 1H), 2.22 – 2.33 (m, 1H), 2.50 – 2.63 (m, 1H), 2.85 – 3.03 (m, 3H), 3.44 – 3.63 (m, 1H), 4.12 – 4.40 (m, 2H), 7.16 – 7.36 (m, 5H).

¹³C NMR (150 MHz, Acetic acid-*d*₄): 11.4, 15.5, 24.4, 26.5, 26.8, 26.9 (2C), 27.1, 28.6, 28.8, 29.1, 29.2, 29.4, 29.5, 29.7, 30.6, 36.8, 38.6, 40.4, 42.1, 56.1, 60.4, 71.4, 127.6, 129.6, 130.3, 139.6, 171.5, 173.8, 173.9.

ESI MS: 550 ([M + Na]⁺).

HR ESI MS: calcd for C₃₁H₄₉O₄N₃Na 550.36153; found 550.36104.

(3*S*,7*S*,8*S*)-3,8-Dibenzyl-7-hydroxy-1,4,9-triazacyclohenicosane-2,5,10-trione (137)



¹H NMR (400 MHz, DMSO-*d*₆): 0.91 – 1.05 (m, 2H), 1.01 – 1.50 (m, 16H), 2.04 – 2.14 (m, 3H), 2.19 (dd, *J* = 14.4, 7.7 Hz, 1H), 2.64 – 2.80 (m, 4H), 3.05 (dd, *J* = 14.0, 4.6 Hz, 1H), 3.18 – 3.28 (m, 1H), 3.83 – 3.97 (m, 2H), 4.28 – 4.36 (m, 1H), 4.97 (d, *J* = 5.1 Hz,

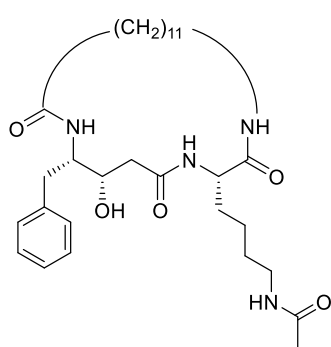
1H), 7.08 – 7.30 (m, 10H), 7.65 (d, $J = 9.2$ Hz, 1H), 7.96 (d, $J = 8.2$ Hz, 1H), 7.98 – 8.06 (m, 1H).

^{13}C NMR (100 MHz, $\text{DMSO-}d_6$): 24.8, 24.9, 26.1, 26.6, 26.7, 27.0, 27.1, 27.3, 27.7, 34.9, 36.9, 37.0, 38.4, 40.0, 53.3, 54.6, 69.2, 125.8, 126.2, 128.1, 128.9, 129.0, 138.4, 139.5, 170.3, 170.8, 172.5.

ESI MS: 558 ($[\text{M} + \text{Na}]^+$).

HR ESI MS: calcd for $\text{C}_{32}\text{H}_{45}\text{O}_4\text{N}_3\text{Na}$ 558.33023; found 558.33006.

***N*-(4-((3*S*,7*S*,8*S*)-8-Benzyl-7-hydroxy-2,5,10-trioxo-1,4,9-triazacyclohenicosan-3-yl)butyl)acetamide (138)**



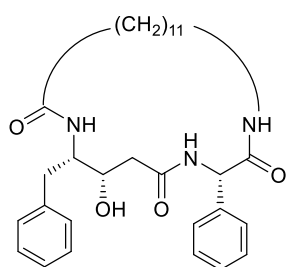
^1H NMR (500 MHz, $\text{DMSO-}d_6$): 0.92 – 1.02 (m, 2H), 1.04 – 1.50 (m, 20H), 1.58 – 1.69 (m, 1H), 1.77 (s, 3H), 2.05 – 2.14 (m, 2H), 2.17 – 2.26 (m, 2H), 2.36 (p, $J = 1.8$ Hz, 1H), 2.63 (p, $J = 1.8$ Hz, 1H), 2.66 – 2.80 (m, 3H), 2.97 (q, $J = 6.6$ Hz, 2H), 3.83 – 3.90 (m, 1H), 3.95 – 4.02 (m, 2H), 5.07 (d, $J = 5.1$ Hz, 1H), 7.11 – 7.24 (m, 5H), 7.68 (d, $J = 9.3$ Hz, 1H), 7.80 (t, $J = 5.6$ Hz, 1H), 7.85 (d, $J = 7.5$ Hz, 1H), 7.99 (t, $J = 5.9$ Hz, 1H).

^{13}C NMR (125 MHz, $\text{DMSO-}d_6$): 22.7, 23.1, 24.9, 26.0, 26.6, 26.7, 27.0, 27.1, 27.4, 27.7, 28.9, 31.1, 34.8, 36.9, 38.3, 38.4, 40.0, 53.4, 53.5, 69.5, 125.8, 128.0, 129.0, 139.6, 170.0, 170.5, 171.6, 172.4.

ESI MS: 581 ($[\text{M} + \text{Na}]^+$).

HR ESI MS: calcd for $\text{C}_{31}\text{H}_{50}\text{O}_5\text{N}_4\text{Na}$ 581.36734; found 581.36685.

(3*S*,7*S*,8*S*)-8-Benzyl-7-hydroxy-3-phenyl-1,4,9-triazacyclohenicosane-2,5,10-trione (139)



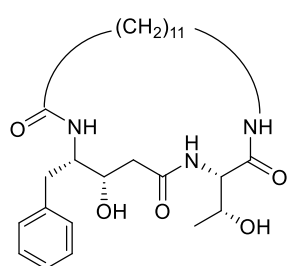
^1H NMR (500 MHz, $\text{DMSO-}d_6$): 1.03 – 1.46 (m, 16H), 1.99 – 2.08 (m, 1H), 2.10 – 2.18 (m, 1H), 2.29 (dd, $J = 6.9, 2.8$ Hz, 2H), 2.36 (p, $J = 1.8$ Hz, 1H), 2.63 (p, $J = 1.9$ Hz, 1H), 2.68 (dd, $J = 13.7, 10.1$ Hz, 1H), 2.71 – 2.78 (m, 2H), 2.81 (dd, $J = 13.8, 4.4$ Hz, 1H), 3.88 – 3.94 (m, 1H), 3.95 – 4.00 (m, 1H), 5.18 (bs, 1H), 5.22 (d, $J = 7.2$ Hz, 1H), 7.12 – 7.41 (m, 10H), 7.72 (d, $J = 9.1$ Hz, 1H), 8.37 (d, $J = 7.2$ Hz, 1H), 8.41 (dd, $J = 7.9, 4.6$ Hz, 1H).

^{13}C NMR (125 MHz, $\text{DMSO-}d_6$): 24.3, 24.8, 25.1, 26.9, 27.3, 27.4, 27.6, 27.9, 28.0, 34.7, 36.5, 38.3, 40.0, 53.8, 57.8, 69.3, 125.8, 127.6, 128.0, 128.3, 129.0, 138.3, 139.7, 169.9, 170.3, 172.4.

ESI MS: 522 ($[M + H]^+$).

HR ESI MS: calcd for $C_{31}H_{44}O_4N_3$ 522.33263; found 522.33248.

(3*S*,7*S*,8*S*)-8-Benzyl-7-hydroxy-3-((*R*)-1-hydroxyethyl)-1,4,9-triazacyclohenicosane-2,5,10-trione (141)



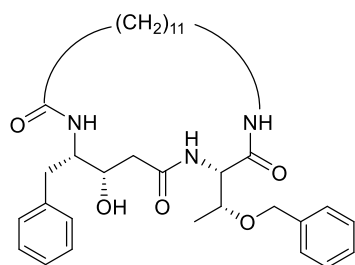
^1H NMR (500 MHz, $DMSO-d_6$): 1.00 (d, $J = 6.4$ Hz, 1H), 1.04 – 1.47 (m, 16H), 2.10 (t, $J = 6.5$ Hz, 2H), 2.23 (dd, $J = 14.1, 6.0$ Hz, 1H), 2.30 (dd, $J = 14.2, 8.1$ Hz, 1H), 2.34 (p, $J = 1.9$ Hz, 1H), 2.62 (p, $J = 1.9$ Hz, 1H), 2.67 – 2.85 (m, 3H), 3.07 – 3.17 (m, 1H), 3.82 – 3.88 (m, 1H), 3.94 – 4.01 (m, 2H), 4.02 – 4.04 (m, 1H), 4.76 (d, $J = 5.6$ Hz, 1H), 5.06 (d, $J = 5.2$ Hz, 1H), 7.08 – 7.23 (m, 5H), 7.48 (d, $J = 8.5$ Hz, 1H), 7.69 (d, $J = 9.2$ Hz, 1H), 7.88 (t, $J = 5.8$ Hz, 1H).

^{13}C NMR (125 MHz, $DMSO-d_6$): 20.6, 24.8, 25.0, 25.9, 26.4, 26.5, 26.9, 27.2 (2C), 27.4, 34.9, 37.2, 38.6, 39.9, 53.4, 59.0, 65.8, 69.7, 125.9, 128.0, 129.0, 139.5, 170.3, 170.8, 172.6.

ESI MS: 512 ($[M + Na]^+$).

HR ESI MS: calcd for $C_{27}H_{43}O_5N_3Na$ 512.30949; found 512.30878.

(3*S*,7*S*,8*S*)-8-Benzyl-3-((*R*)-1-(benzyloxy)ethyl)-7-hydroxy-1,4,9-triazacyclohenicosane-2,5,10-trione (142)



^1H NMR (400 MHz, $DMSO-d_6$): 0.88 – 1.00 (m, 3H), 1.06 – 1.48 (m, 19H), 1.40 – 1.54 (m, 1H), 2.13 (t, $J = 6.4$ Hz, 2H), 2.24 (dd, $J = 14.1, 8.7$ Hz, 1H), 2.38 (dd, $J = 14.1, 8.7$ Hz, 1H), 2.64 – 2.80 (m, 2H), 3.19 – 3.32 (m, 1H), 3.82 – 3.90 (m, 1H), 3.99 – 4.05 (m, 1H), 4.09 (qd, $J = 6.3, 2.6$ Hz, 1H), 4.16 (dd, $J = 8.7, 2.6$ Hz, 1H), 4.39 (d, $J = 11.8$ Hz, 1H), 4.50 (d, $J = 11.9$ Hz, 1H), 7.09 – 7.35 (m, 10H), 7.56 (d, $J = 8.7$ Hz, 1H), 7.73 (d, $J = 9.3$ Hz, 1H), 8.13 (t, $J = 5.8$ Hz, 1H).

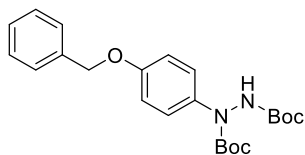
^{13}C NMR (100 MHz, $DMSO-d_6$): 16.8, 24.6, 24.7, 25.5, 26.0, 26.4, 26.8, 27.0, 27.1, 27.3, 34.9, 37.5, 40.0, 40.6, 53.3, 58.0, 70.1, 70.7, 74.4, 125.8, 127.3, 127.9 (2C), 129.0, 138.6, 139.5, 170.0, 170.9, 172.8.

ESI MS: 602 ($[M + Na]^+$).

HR ESI MS: calcd for $C_{34}H_{49}O_5N_3Na$ 602.35644; found 602.35600.

4.3 Linker-Modified Gingipain Inhibitors for Probe Development

Di-*tert*-butyl 1-(4-(benzyloxy)phenyl)hydrazine-1,2-dicarboxylate (**159**)



Di-*tert*-butyl azodicarboxylate (1.28 g, 5.55 mmol, 1.0 eq.), 4-benzyloxyphenylboronic acid (1.90 g, 8.33 mmol, 1.5 eq.) and copper(II) acetate (0.05 g, 0.28 mmol, 0.05 eq.) were dissolved in MeOH (90 mL). The reaction mixture was stirred at room temperature overnight. Solvent was then evaporated and the product was purified by chromatography on silica gel (hex:EtOAc / 7:1) to afford **159** (2.20 g, 94%) as an off-white amorphous solid.

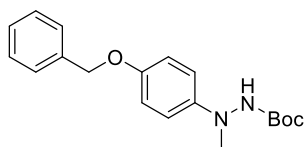
¹H NMR (400 MHz, CDCl₃) δ 1.40 – 1.57 (m, 18H), 5.05 (s, 2H), 6.77 (bs, 1H), 6.89 – 6.95 (m, 2H), 7.27 – 7.45 (m, 7H).

¹³C NMR (100 MHz, CDCl₃) δ 28.3, 28.4, 70.3, 82.1, 114.8, 116.1, 127.6, 128.1, 128.7, 135.7, 137.0, 154.2, 155.6, 157.0.

ESI MS: 437 ([M + Na]⁺).

HR ESI MS: calcd for C₂₃H₃₀O₅N₂Na 437.20469, found 437.20481.

tert-Butyl 2-(4-(benzyloxy)phenyl)-2-methylhydrazine-1-carboxylate (**160**)



Intermediate **159** (2.16 g, 5.22 mmol, 1.0 eq.) was dissolved in anhydrous THF (20 mL) under inert atmosphere and the reaction mixture was cooled down to 0 °C. Solution of LiAlH₄ in THF (1M, 15.7 mL, 15.70 mmol, 3 eq.) was added cautiously and the reaction mixture was stirred at room temperature for 4 hours. Then, the reaction mixture was cooled down to 0 °C and the reaction was quenched by slow addition of saturated aqueous solution of potassium sodium tartrate (20 mL). Next, the product was extracted with DCM (3 × 20 mL), the organic phases were combined, washed with brine (20 mL) and dried using anhydrous MgSO₄. The solvent was then evaporated and the product was purified by chromatography on silica gel (hex:EtOAc / 5:1) to afford **160** (1.43 g, 84%) as an off-white amorphous solid.

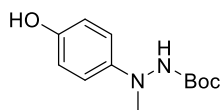
¹H NMR (400 MHz, CDCl₃) δ 1.26 – 1.59 (m, 9H), 3.12 (s, 3H), 5.02 (s, 2H), 6.43 (bs, 1H), 6.75 – 6.83 (m, 2H), 6.87 – 6.94 (m, 2H), 7.29 – 7.35 (m, 1H), 7.35 – 7.45 (m, 4H).

¹³C NMR (100 MHz, CDCl₃) δ 28.4, 41.7, 70.7, 80.9, 114.5, 115.8, 127.6, 127.9, 128.6, 137.5, 144.6, 152.8, 155.0.

ESI MS: 351 ([M + Na]⁺).

HR ESI MS: calcd for C₁₉H₂₄O₃N₂Na 351.16791, found 351.16822.

***tert*-Butyl 2-(4-hydroxyphenyl)-2-methylhydrazine-1-carboxylate (**161**)**



Intermediate **160** (120 mg, 0.37 mmol, 1.0 eq.) was dissolved in glacial AcOH (7 mL) and 10% Pd/C (25 mg) was added. The reaction mixture was stirred under H₂ atmosphere at room temperature for 5 hours. Next, Pd/C was removed by filtration and the solvent was evaporated. The residue was redissolved in DCM (10 mL), washed with saturated aq. NaHCO₃ (15 mL) and the product was extracted with DCM (4 × 10 mL). The organic phases were combined, washed with brine (20 mL) and dried using anhydrous MgSO₄. The solvent was then evaporated and the product was purified by chromatography on silica gel (hex:EtOAc / 2:1) to afford desired derivative **161** (54 mg, 61%) as a clear glassy solid.

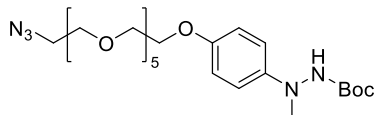
¹H NMR (401 MHz, CDCl₃) δ 1.44 – 1.52 (m, 9H), 3.07 (s, 3H), 6.45 (bs, 1H), 6.62 – 6.74 (m, 4H).

¹³C NMR (101 MHz, CDCl₃) δ 28.5, 41.9, 81.2, 114.5, 116.3, 143.1, 148.6, 155.5.

ESI MS: 261 ([M + Na]⁺).

HR ESI MS: calcd for C₁₂H₁₈O₃N₂Na 261.12096, found 261.12099.

***tert*-Butyl 2-(4-((17-azido-3,6,9,12,15-pentaoxaheptadecyl)oxy)phenyl)-2-methylhydrazine-1-carboxylate (**162**)**



Phenol **161** (54 mg, 0.23 mmol, 1.0 eq.), N₃-PEG-5-OTs (157 mg, 0.34 mmol, 1.5 eq.) and K₂CO₃ (47 mg, 0.34, 1.5 eq.) were dissolved in anhydrous DMF (2 mL) under inert atmosphere in a sealed tube. The reaction mixture was heated to 70 °C for 48 hours. The product was purified by RP HPLC (C18, gradient 5 – 90% MeCN in H₂O+0.1% TFA) to afford **162** (51 mg, 43%) as a clear oil.

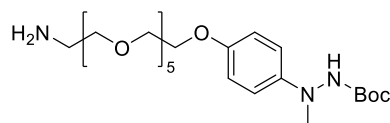
¹H NMR (400 MHz, CD₃OD) δ 1.49 (s, 9H), 3.05 (s, 3H), 3.33 – 3.39 (m, 2H), 3.53 – 3.71 (m, 18H), 3.77 – 3.84 (m, 2H), 4.02 – 4.09 (m, 2H), 6.77 – 6.82 (m, 2H), 6.82 – 6.88 (m, 2H).

¹³C NMR (100 MHz, CD₃OD) δ 27.3, 40.6, 50.4, 67.8, 68.4, 69.6, 69.7, 70.1 (2C), 70.2 (4C), 70.3, 79.9, 114.4, 115.0, 144.9, 152.7, 156.7.

ESI MS: 550 ([M + Na]⁺).

HR ESI MS: calcd for C₂₄H₄₁O₈N₅Na 550.28473, found 550.28469.

***tert*-Butyl 2-(4-((17-amino-3,6,9,12,15-pentaoxaheptadecyl)oxy)phenyl)-2-methylhydrazine-1-carboxylate (**163**)**



Intermediate **162** (30 mg, 57 μ mol, 1.0 eq.) was dissolved in THF (5 mL) and 10% Pd/C (5 mg) was added. The reaction mixture was stirred under H₂ atmosphere at room

temperature for 3 hours. Next, Pd/C was removed by filtration and the solvent was evaporated. The product was purified by RP HPLC (C18, gradient 20 – 50% MeCN in H₂O+0.1% TFA) to afford **163** (13 mg, 37%) as a clear oil.

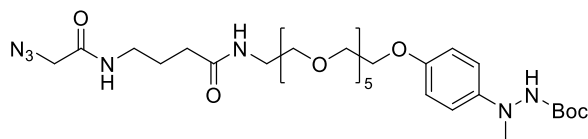
¹H NMR (400 MHz, CD₃OD) δ 1.49 (s, 9H), 3.06 (s, 3H), 3.08 – 3.13 (m, 2H), 3.57 – 3.77 (m, 18H), 3.80 – 3.86 (m, 2H), 4.05 – 4.12 (m, 2H), 6.78 – 6.91 (m, 4H).

¹³C NMR (100 MHz, CD₃OD) δ 28.7, 40.5, 41.9, 67.8, 69.1, 70.7, 71.0, 71.1 (2C), 71.2, 71.3, 71.4 (2C), 115.8, 116.4, 146.4, 153.7, 156.3.

ESI MS: 502 ([M + H]⁺).

HR ESI MS: calcd for C₂₄H₄₄O₈N₃ 502.31229, found 502.31212.

***tert*-Butyl 2-(4-((25-azido-19,24-dioxo-3,6,9,12,15-pentaoxa-18,23-diazapentacosyl)oxy)phenyl)-2-methylhydrazine-1-carboxylate (**164**)**



N₃Ac-GABA-OH (3.7 mg, 20 μ mol, 1.1 eq.) and HBTU (7.5 mg, 20 μ mol, 1.1 eq.) were dissolved in anhydrous DMF (1 mL) under

inert atmosphere. Next, the reaction mixture was cooled down to 0 °C, a drop of DIPEA was added and the reaction mixture was stirred at 0 °C for 10 minutes. Then, a solution of amine **163** (9 mg, 18 μ mol, 1.0 eq.) in anhydrous DMF (1 mL) was added. The reaction mixture was stirred for 2 hours at room temperature. Next, the solvent was evaporated, the residue was redissolved in DCM (10 mL), washed with saturated aq. NaHCO₃ (15 mL) and the product was extracted with DCM (4 $\times$ 10 mL). The organic phases were combined, washed with brine (20 mL) and dried using anhydrous MgSO₄. The solvent was then evaporated to afford compound **164** (10 mg, 83%) as a clear oil.

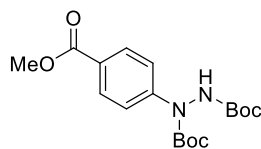
¹H NMR (400 MHz, CD₃OD) δ 1.49 (s, 9H), 1.76 – 1.84 (m, 2H), 2.22 (t, *J* = 7.5 Hz, 2H), 3.06 (s, 3H), 3.24 (t, *J* = 7.0 Hz, 2H), 3.35 (t, *J* = 5.5 Hz, 2H), 3.52 (t, *J* = 5.5 Hz, 2H), 3.57 – 3.72 (m, 16H), 3.77 – 3.84 (m, 2H), 3.87 (s, 2H), 4.02 – 4.09 (m, 2H), 6.77 – 6.90 (m, 4H).

¹³C NMR (100 MHz, CD₃OD) δ 26.6, 28.7, 34.3, 39.9, 40.4, 42.0, 48.4, 48.6, 48.8, 49.0, 49.2, 49.4, 49.6, 53.0, 69.2, 70.5, 71.0, 71.2, 71.5 (3C), 71.6 (2C), 71.7, 81.3, 112.1, 115.8, 116.4, 146.2, 154.1, 156.5, 170.2, 175.5.

ESI MS: 692 ($[M + Na]^+$).

HR ESI MS: calcd for $C_{30}H_{51}O_{10}N_7Na$ 692.35896, found 692.35874.

Di-*tert*-butyl 1-(4-(methoxycarbonyl)phenyl)hydrazine-1,2-dicarboxylate (166)



Di-*tert*-butyl azodicarboxylate (0.43 g, 1.85 mmol, 1.0 eq.), 4-benzyloxyphenylboronic acid (0.50 g, 2.78 mmol, 1.5 eq.) and copper(II) acetate (17 mg, 0.09 mmol, 0.05 eq.) were dissolved in MeOH (30 mL). The reaction mixture was stirred at room

temperature overnight. Solvent was then evaporated and the product was purified by chromatography on silica gel (chex:EtOAc / 6:1) to afford **166** (0.51 g, 76%) as an off-white amorphous solid.

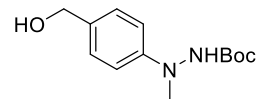
^1H NMR (400 MHz, $CDCl_3$) δ 1.42 – 1.54 (m, 18H), 3.89 (s, 3H), 6.78 (bs, 1H), 7.52 (d, J = 8.5 Hz, 2H), 7.94 – 8.02 (m, 2H).

^{13}C NMR (100 MHz, $CDCl_3$) δ 28.2, 28.3, 52.2, 83.2, 121.8, 126.4, 130.2, 146.2, 153.0, 155.4, 166.8.

ESI MS: 389 ($[M + Na]^+$).

HR ESI MS: calcd for $C_{18}H_{26}O_6N_2Na$ 389.16831, found 389.16803.

***tert*-Butyl 2-(4-(hydroxymethyl)phenyl)-2-methylhydrazine-1-carboxylate (167)**



Intermediate **166** (0.72 g, 1.97 mmol, 1.0 eq.) was dissolved in anhydrous THF (10 mL) under inert atmosphere and the reaction mixture was cooled down to 0 °C. Solution of $LiAlH_4$ in THF (1M, 19.7 mL, 19.70 mmol, 10 eq.) was added cautiously and the reaction mixture was stirred at room temperature for 4 hours. Then, the reaction mixture was cooled down to 0 °C and the reaction was quenched by slow addition of saturated aqueous solution of potassium sodium tartrate (20 mL). Next, the product was extracted with THF (3×20 mL), the organic phases were combined and dried using anhydrous $MgSO_4$. The solvent was then evaporated and the product was purified by chromatography on silica gel (chex:EtOAc / 2:1) to afford **167** (0.40 g, 81%) as an off-white amorphous solid.

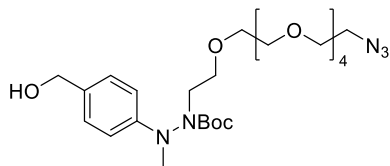
^1H NMR (400 MHz, $CDCl_3$) δ 1.30 – 1.56 (m, 9H), 3.18 (s, 3H), 4.57 (s, 2H), 6.53 (bs, 1H), 6.79 (d, J = 8.1 Hz, 2H), 7.22 – 7.26 (m, 2H).

^{13}C NMR (100 MHz, $CDCl_3$) δ 28.4, 41.1, 65.3, 81.2, 112.8, 128.6, 131.8, 149.7.

ESI MS: 275 ($[M + Na]^+$).

HR ESI MS: calcd for $C_{13}H_{20}O_3N_2Na$ 275.13661, found 275.13658.

***tert*-Butyl 19-azido-2-((4-(hydroxymethyl)phenyl)(methyl)amino)-5,8,11,14,17-pentaoxa-2-azanonadecanoate (**169**)**

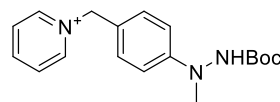


Benzyl alcohol **167** (49 mg, 0.19 mmol, 1.0 eq.) and N₃-PEG-5-OTs (90 mg, 0.19 mmol, 1.0 eq.) were dissolved in anhydrous DMF (2 mL) under inert atmosphere. The reaction mixture was cooled down to 0 °C and *t*BuOK (44 mg, 0.38 mmol, 2.0 eq.) was added slowly. The reaction mixture was stirred at room temperature for 2 hours. Next, saturated aq. NaHCO₃ (20 mL) was added, and the product was extracted with DCM (4 × 10 mL). The organic phases were combined, washed with brine (20 mL) and dried using anhydrous MgSO₄. The product was purified by RP flash chromatography (C18, gradient 30 – 80% MeCN in H₂O+0.1% TFA) to afford **169** (95 mg, 92%) as a clear oil.

¹H NMR (400 MHz, CD₃OD) δ 1.16 – 1.57 (m, 9H), 3.16 (s, 3H), 3.36 (t, *J* = 4.9 Hz, 2H), 3.49 – 3.58 (m, 1H), 3.58 – 3.70 (m, 20H), 3.78 – 3.88 (m, 1H), 4.50 (s, 2H), 6.70 (d, *J* = 8.6 Hz, 2H), 7.23 (d, *J* = 8.6 Hz, 2H).

¹³C NMR (100 MHz, CD₃OD) δ 28.5, 39.7, 51.8, 65.0, 69.0, 71.1, 71.2, 71.5, 71.6 (3C), 81.9, 112.9, 129.4, 132.7, 150.5, 157.4.

1-(4-(2-(*tert*-Butoxycarbonyl)-1-methylhydrazineyl)benzyl)pyridin-1-ium (173**)**



Benzyl alcohol **167** (0.16 g, 0.63 mmol, 1 eq.) was dissolved in anhydrous pyridine (5 mL) and TsCl (0.30 g, 1.58 mmol, 2.5 eq.) was added. The reaction mixture was stirred at room temperature overnight. The solvent was the evaporated and the product was purified by RP flash chromatography (C18, gradient 15 – 50% MeCN in H₂O+0.1% TFA) to afford **173** (0.25 g, 90%) as a yellow oil.

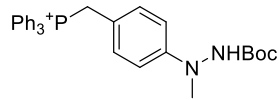
¹H NMR (401 MHz, CDCl₃) δ 1.20 – 1.53 (m, 9H), 3.11 (s, 3H), 5.68 (s, 2H), 6.67 (d, *J* = 8.3 Hz, 2H), 7.28 (d, *J* = 8.3 Hz, 2H), 7.48 (s, 1H), 7.79 – 7.86 (m, 2H), 8.27 (t, *J* = 7.7 Hz, 1H), 9.13 – 9.22 (m, 2H).

¹³C NMR (101 MHz, CDCl₃) δ 28.4, 40.5, 64.5, 81.2, 112.9, 122.3, 128.2, 130.8, 144.7, 144.9, 151.2, 155.3.

ESI MS: 314 ([M]⁺).

HR ESI MS: calcd for C₁₈H₂₄O₂N₃ 314.18630, found 314.18634.

(4-(2-(*tert*-Butoxycarbonyl)-1-methylhydrazineyl)benzyl)triphenylphosphonium (175)



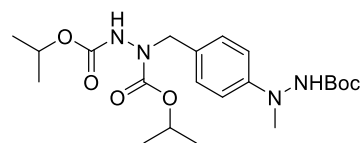
Iodine (12 mg, 96 μ mol, 2.4 eq.), PPh₃ (22 mg, 80 μ mol, 2.0 eq.) and imidazole (6 mg, 80 μ mol, 2.0 eq.) were dissolved in anhydrous DCM (1 mL). Solution of benzyl alcohol **167** (10 mg, 40 μ mol, 1.0 eq.) in anhydrous DCM (1 mL) was added dropwise and the reaction mixture was stirred in dark at room temperature overnight. The product was purified by chromatography on silica gel (DCM:MeOH / gradient 30:1 to 20:1) to afford **175** (14 mg, 70%) as an off-white amorphous solid.

¹H NMR (400 Hz, CDCl₃) δ 1.19 – 1.53 (m, 9H), 3.10 (s, 3H), 5.03 (d, J = 13.2 Hz, 2H), 6.51 – 6.57 (m, 2H), 6.65 (bs, 1H), 6.86 – 6.92 (m, 3H), 7.59 – 7.71 (m, 12H), 7.72 – 7.80 (m, 3H).

ESI MS: 497 ([M]⁺).

HR ESI MS: calcd for C₃₁H₃₄O₂N₂P 497.23524, found 497.23464.

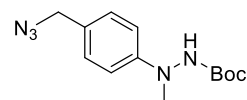
Diisopropyl 1-(4-(2-(*tert*-butoxycarbonyl)-1-methylhydrazineyl)benzyl)hydrazine-1,2-dicarboxylate (177)



PPh₃ (13 mg, 48 μ mol, 1.2 eq.) and DIAD (9 μ L, 48 μ mol, 1.2 eq.) were dissolved in anhydrous THF (2 mL). The reaction mixture was cooled down to 0 °C. Next, benzyl alcohol **167** (10 mg, 40 μ mol, 1.0 eq.) and Phthalimide (7 mg, 48 μ mol, 1.2 eq) were added and the reaction mixture was stirred at 0 °C for 5 hours. The product was purified by chromatography on silica gel (hex:EtOAc / 2:1) to afford **177** (18 mg, 82%) as a yellow oil.

¹H NMR (400 MHz, MeOD) δ 1.25 (d, J = 6.2 Hz, 12H), 1.49 (s, 9H), 3.12 (s, 3H), 4.41 – 4.64 (m, 2H), 4.88 – 4.95 (m, 2H), 6.75 – 6.82 (m, 2H), 7.12 – 7.19 (m, 2H).

***tert*-Butyl 2-(4-(azidomethyl)phenyl)-2-methylhydrazine-1-carboxylate (178)**



Intermediate **173** (0.28 g, 0.65 mmol, 1.0 eq.) and NaN₃ (0.13 g, 1.94 mmol, 3.0 eq.) were suspended in anhydrous DMF (5 mL) under inert atmosphere. The reaction mixture was stirred at 80 °C for 3 hours. Next, the solvent was evaporated, the residue was suspended in H₂O (15 mL) and the product was extracted with EtOAc (3 $\times$ 10 mL). The organic phases were combined and dried using anhydrous MgSO₄. The solvent was then evaporated to afford compound **178** (0.18 g, quant.) as a yellow oil.

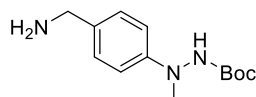
¹H NMR (400 MHz, CDCl₃) δ 1.21 – 1.53 (m, 9H), 3.19 (s, 3H), 4.23 (s, 2H), 6.52 (bs, 1H), 6.77 – 6.85 (m, 2H), 7.14 – 7.24 (m, 2H).

¹³C NMR (100 MHz, CDCl₃) δ 28.4, 41.0, 54.6, 81.3, 112.8, 125.9, 129.6, 150.0, 154.8.

ESI MS: 300 ([M + Na]⁺).

HR ESI MS: calcd for C₁₃H₁₉O₂N₅Na 300.14310, found 300.14301.

***tert*-Butyl 2-(4-(aminomethyl)phenyl)-2-methylhydrazine-1-carboxylate (179)**



Azide **178** (50 mg, 0.18 mmol, 1.0 eq.) was dissolved in THF (5 mL) and 10% Pd/C (5 mg) was added. The reaction mixture was stirred under H₂ atmosphere at room temperature for 1 hour. Next, Pd/C was removed by filtration and the solvent was evaporated. The product was purified by RP HPLC (C18, gradient 5 – 50% MeCN in H₂O+0.1% TFA) to afford **179** (45 mg, 67%) as a clear oil.

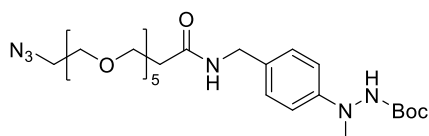
¹H NMR (400 MHz, CD₃OD) δ 1.24 – 1.56 (m, 9H), 3.15 (s, 3H), 4.00 (s, 2H), 6.86 (d, *J* = 8.8 Hz, 1H), 7.30 (d, *J* = 8.8 Hz, 1H).

¹³C NMR (100 MHz, CD₃OD) δ 28.6, 40.9, 44.0, 81.7, 114.0, 124.2, 131.0, 152.3, 157.9.

ESI MS: 274 ([M + Na]⁺).

HR ESI MS: calcd for C₁₃H₂₁O₂N₃Na 274.15260, found 274.15281.

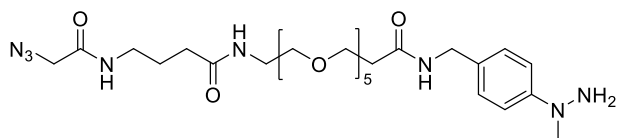
***tert*-Butyl 2-(4-(20-azido-3-oxo-6,9,12,15,18-pentaoxa-2-azaicosyl)phenyl)-2-methylhydrazine-1-carboxylate (180)**



Amine **179** (45 mg, 0.12 mmol, 1.0 eq.) and N₃-PEG-5-CH₂CH₂COOSu (53 mg, 0.12 mmol, 1.0 eq.) were dissolved in anhydrous DMF (2 mL). Then, DIPEA (63 μL, 0.36 mmol, 2.0 eq.) was added and the reaction mixture was stirred at room temperature for 2 hours. Next, the solvent was evaporated, the residue was redissolved in DCM (10 mL), washed with 10% aq. KHSO₄ (2 × 10 mL), and the combined aqueous layers were re-extracted with DCM (2 × 10 mL). Then, the combined organic layers were washed with saturated aq. NaHCO₃ (2 × 15 mL) and the combined aqueous layers were re-extracted with DCM (2 × 10 mL). The organic phases were combined and dried using anhydrous MgSO₄ to yield compound **180** (60 mg, 87%) as a pale-yellow oil.

¹H NMR (400 MHz, CD₃OD) δ 1.26 – 1.56 (m, 9H), 2.47 (t, *J* = 6.1 Hz, 2H), 3.11 (s, 3H), 3.36 (t, *J* = 4.9 Hz, 2H), 3.56 – 3.69 (m, 18H), 3.74 (t, *J* = 6.1 Hz, 2H), 4.29 (s, 2H), 6.79 (d, *J* = 8.2 Hz, 2H), 7.14 – 7.19 (m, 2H).

1-(4-(2-Azidoacetamido)butanamido)-N-(4-(1-methylhydrazineyl)benzyl)-3,6,9,12,15-pentaoxaoctadecan-18-amide (182)



Intermediate **181** (60 mg, 84 μ mol, 1.0 eq.) was dissolved in DCM (1 mL) and TFA (1 mL) was added. The reaction mixture was stirred for 30 minutes. The volatiles were evaporated and the product was purified by RP HPLC (C18, gradient 2 – 30% MeCN in H₂O+0.1% TFA) to afford **182** (45 mg, 75%) as a yellow oil.

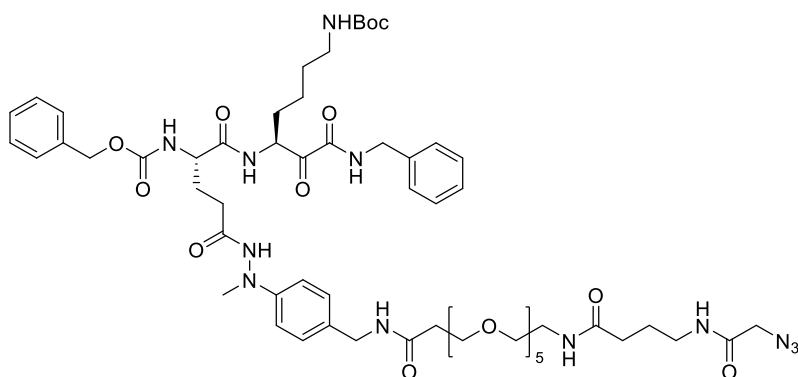
¹H NMR (400 MHz, CD₃OD) δ 1.80 (p, J = 7.2 Hz, 2H), 2.22 (t, J = 7.5 Hz, 2H), 2.50 (t, J = 6.0 Hz, 2H), 3.17 (s, 3H), 3.23 (t, J = 7.0 Hz, 2H), 3.35 (t, J = 5.5 Hz, 2H), 3.49 – 3.61 (m, 2H), 3.57 – 3.66 (m, 16H), 3.75 (t, J = 6.0 Hz, 2H), 3.88 (s, 2H), 4.37 (s, 2H), 7.13 – 7.18 (m, 2H), 7.32 – 7.37 (m, 2H).

¹³C NMR (100 MHz, CD₃OD) δ 25.2, 32.9, 36.3, 38.5, 39.0, 41.9, 42.8, 51.6, 51.6, 63.8, 66.9, 69.2, 69.8, 69.9, 70.0, 70.1 (3C), 117.5, 128.3, 135.2, 147.3, 168.8, 172.6, 174.1.

ESI MS: 633 ([M + Na]⁺).

HR ESI MS: calcd for C₂₇H₄₆O₈N₈Na 633.33308, found 633.33325

Boc-Protected Linker Modified KYT-36 (183)



Acid **149** (17 mg, 27 μ mol, 1.0 eq.) and HBTU (10 mg, 27 μ mol, 1.0 eq.) were dissolved in anhydrous DMF (0.5 mL) under inert atmosphere. Next, the reaction mixture was cooled down to 0 °C, 2,4,6-collidine (11 μ L, 81 μ mol, 3.0 eq.) was added and the reaction mixture was stirred at 0 °C for 10 min. Then, a solution of amine **182** (17 mg, 23 μ mol, 0.85 eq.) in anhydrous DMF (0.5 mL) was added. The reaction mixture was stirred at room temperature overnight. Next, the solvent was evaporated and the product was purified by RP HPLC (C18, gradient 30 – 70% MeCN in H₂O+0.1% TFA) to afford **183** (7 mg, 75%) as a yellow oil.

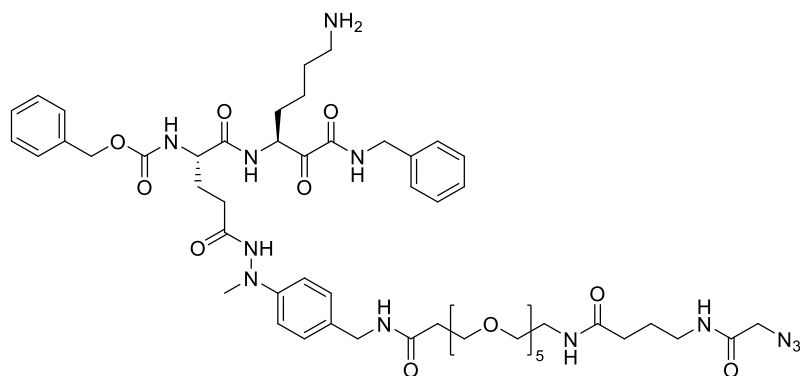
¹H NMR (400 MHz, CD₃OD) δ 1.24 – 1.67 (m, 13H), 1.79 (p, *J* = 7.2 Hz, 2H), 1.87 – 2.17 (m, 2H), 2.21 (t, *J* = 7.6 Hz, 2H), 2.33 – 2.43 (m, 2H), 2.46 (t, *J* = 6.1 Hz, 2H), 2.90 – 3.05 (m, 2H), 3.09 (s, 2H), 3.23 (t, *J* = 7.0 Hz, 2H), 3.32 – 3.37 (m, 2H), 3.52 (t, *J* = 5.5 Hz, 2H), 3.56 – 3.66 (m, 18H), 3.69 – 3.78 (m, 2H), 3.86 (s, 2H), 4.09 – 4.50 (m, 6H), 5.03 – 5.18 (m, 2H), 6.76 – 6.85 (m, 2H), 7.12 – 7.21 (m, 2H), 7.21 – 7.40 (m, 10H).

The **¹³C NMR** spectrum was not recorded due to the high complexity of the compound along with low available quantities of compound **183**.

ESI MS: 1242 ([M + Na]⁺).

HR ESI MS: calcd for C₅₉H₈₆O₁₆N₁₂Na 1241.61770, found 1241.61853.

Linker Modified KYT-36 (**184**)



Intermediate **183** (7 mg, 5.7 μmol, 1.0 eq.) was dissolved in DCM (1 mL) and TFA (1 mL) was added. The reaction mixture was stirred for 30 minutes. The volatiles were evaporated and the product was purified by RP HPLC (C18, gradient 20 – 60% MeCN in H₂O+0.1% TFA) to afford **184** (4 mg, 57%) as a pale-yellow lyophilisate.

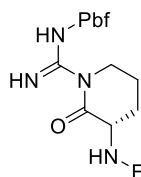
¹H NMR (400 MHz, CD₃OD) δ 1.23 – 1.37 (m, 2H), 1.42 – 1.73 (m, 2H), 1.79 (p, *J* = 7.3 Hz, 2H), 1.88 – 2.18 (m, 2H), 2.21 (t, *J* = 7.5 Hz, 2H), 2.35 – 2.50 (m, 4H), 2.75 – 2.96 (m, 2H), 3.10 (s, 3H), 3.22 (t, *J* = 6.9 Hz, 2H), 3.34 (t, *J* = 5.6 Hz, 2H), 3.52 (t, *J* = 5.4 Hz, 2H), 3.56 – 3.63 (m, 18H), 3.73 (t, *J* = 6.0 Hz, 2H), 3.87 (s, 2H), 4.17 – 4.50 (m, 6H), 5.04 – 5.16 (m, 2H), 6.77 – 6.86 (m, 2H), 7.12 – 7.20 (m, 2H), 7.22 – 7.40 (m, 10H).

The **¹³C NMR** spectrum was not recorded due to the high complexity of the compound along with low available quantities of compound **184**.

ESI MS: 1120 ([M + H]⁺).

HR ESI MS: calcd for C₅₄H₇₉O₁₄N₁₂ 1119.58332, found 1119.58376.

(9H-Fluoren-9-yl)methyl (S)-(2-oxo-1-(N-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)carbamimidoyl)piperidin-3-yl)carbamate (193)



Fmoc-Arg(Pbf)-OH (1.00 g, 1.54 mmol, 1.0 eq.) and HATU (0.70 g, 1.84 mmol, 1.2 eq.) were dissolved in anhydrous DMF (25 mL) under inert atmosphere. Next, DIPEA (0.66 mL, 3.85 mmol, 2.5 eq.) was added and the reaction mixture was stirred overnight at room temperature.

The solvent was evaporated, the residue was redissolved in EtOAc (50 mL), and the organic phase was washed with 10% aq. KHSO₄ (2 × 50 mL), H₂O (50 mL), saturated aq. NaHCO₃ (2 × 50 mL) and brine (30 mL). The organic layer was then dried using anhydrous MgSO₄ and the solvent was evaporated. The product was purified by chromatography on silica gel (hex:EtOAc / 3:2) to afford **193** (0.85 g, 88%) as an off-white amorphous solid.

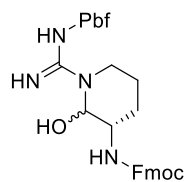
¹H NMR (400 MHz, CDCl₃) δ 1.47 (s, 6H), 1.53 – 1.64 (m, 1H), 1.78 – 1.97 (m, 1H), 2.11 (s, 3H), 2.43 – 2.51 (m, 1H), 2.53 (s, 3H), 2.58 (s, 3H), 2.97 (s, 2H), 3.31 – 3.45 (m, 1H), 4.23 (t, J = 6.7 Hz, 1H), 4.32 – 4.72 (m, 3H), 5.54 (d, J = 6.8 Hz, 1H), 7.28 – 7.36 (m, 2H), 7.40 (t, J = 7.5 Hz, 2H), 7.53 – 7.63 (m, 2H), 7.76 (d, J = 7.6 Hz, 2H), 7.91 (bs, 1H), 9.40 (bs, 1H).

¹³C NMR (100 MHz, CDCl₃) δ 12.6, 18.1, 19.4, 19.7, 25.3, 28.7, 43.3, 47.2, 52.8, 86.8, 117.9, 120.1, 125.0, 127.2, 127.9, 131.7, 133.0, 139.0, 141.4, 143.8, 143.9, 153.9, 156.0, 159.4, 175.8.

ESI MS: 631 ([M + H]⁺).

HR ESI MS: calcd for C₃₄H₃₉O₆N₄S 631.25848, found 631.25883.

(9H-Fluoren-9-yl)methyl ((3S)-2-hydroxy-1-(N-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)carbamimidoyl)piperidin-3-yl)carbamate (186)



Cyclic arginine derivative **193** (0.15 g, 0.24 mmol, 1.0 eq.) was dissolved in anhydrous THF (10 mL) under inert atmosphere and the reaction mixture was cooled down to -20 °C. Solution of LiAlH₄ in THF (1M, 0.24 mL, 0.24 mmol, 1.0 eq.) was added cautiously and the reaction mixture was

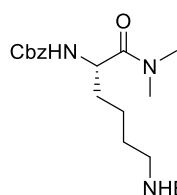
stirred at -20 °C for 45 minutes. Then, the reaction was quenched by slow addition of saturated aqueous solution of potassium sodium tartrate (20 mL) and the product was extracted with DCM (3 × 20 mL), the organic phases were combined, washed with brine (30 mL) and dried using anhydrous MgSO₄. The solvent was then evaporated to afford **186** (0.11 g, 73%) as an off-white amorphous solid. The compound was stored at -20 °C to minimize any potential decomposition or racemization.

The ^1H and ^{13}C NMR spectra for this compound were recorded but not interpreted due to the high complexity of the spectra caused by the presence of multiple forms of the compound.

ESI MS: 655 ($[\text{M} + \text{Na}]^+$).

HR ESI MS: calcd for $\text{C}_{34}\text{H}_{40}\text{O}_6\text{N}_4\text{NaS}$ 655.25608, found 655.25552.

Benzyl *tert*-butyl (6-(dimethylamino)-6-oxohexane-1,5-diyl)(*S*)-dicarbamate (194**)**



Cbz-Lys(Boc)-OH (1.00 g, 2.63 mmol, 1.0 eq.) and HATU (1.00 g, 2.63 mmol, 1.0 eq.) were dissolved in anhydrous DMF (25 mL) under inert atmosphere. Next, DIPEA (1.40 mL, 7.89 mmol, 3.0 eq.) was added and the reaction mixture was stirred for 15 minutes. Then, a solution of Me_2NH in EtOH (33%, 0.47 mL, 2.63 mmol, 1.0 eq.) was added and the reaction mixture was stirred for 2 hours at room temperature. The solvent was evaporated, the residue was redissolved in EtOAc (50 mL), and the organic phase was washed with 10% aq. KHSO_4 (2×50 mL), saturated aq. NaHCO_3 (2×50 mL) and brine (30 mL). The organic layer was then dried using anhydrous MgSO_4 and the solvent was evaporated to afford **194** (1.07 g, quant.) as a yellow oil.

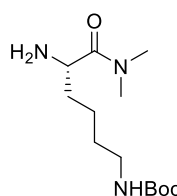
^1H NMR (400 MHz, CD_3OD) δ 1.28 – 1.53 (m, 13H), 1.53 – 1.74 (m, 2H), 2.94 (s, 3H), 3.03 (q, $J = 6.1$ Hz, 2H), 3.14 (s, 3H), 4.50 – 4.62 (m, 1H), 5.02 – 5.13 (m, 2H), 6.55 (t, $J = 6.0$ Hz, 1H), 7.24 – 7.40 (m, 5H).

^{13}C NMR (100 MHz, CD_3OD) δ 22.5, 27.6, 29.2, 31.1, 34.9, 36.1, 39.5, 51.0, 66.3, 78.5, 127.4, 127.6, 128.1, 136.8, 157.1, 157.2, 173.1.

ESI MS: 430 ($[\text{M} + \text{Na}]^+$).

HR ESI MS: calcd for $\text{C}_{21}\text{H}_{33}\text{O}_5\text{N}_3\text{Na}$ 430.23124, found 430.23107.

***tert*-Butyl (*S*)-(5-amino-6-(dimethylamino)-6-oxohexyl)carbamate (**195**)**



Intermediate **194** (1.07 g, 2.62 mmol, 1.0 eq.) was dissolved in THF (25 mL) and 10% Pd/C (0.10 g) was added. The reaction mixture was stirred under H_2 atmosphere at room temperature for 2 hours. Next, Pd/C was removed by filtration and the solvent was evaporated to yield amine **195** (0.71 g, quant.) as a tan foam.

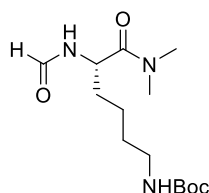
^1H NMR (401 MHz, CDCl_3) δ 1.32 – 1.53 (m, 11H), 1.55 – 1.69 (m, 4H), 2.95 (s, 3H), 3.01 (s, 3H), 3.05 – 3.15 (m, 2H), 3.65 (dd, $J = 7.3, 4.5$ Hz, 1H), 4.61 (bs, 1H).

^{13}C NMR (100 MHz, CDCl_3) δ 23.0, 28.4, 30.0, 34.9, 35.8, 36.8, 40.3, 51.0, 79.1, 156.0, 175.5.

ESI MS: 274 ($[M + H]^+$).

HR ESI MS: calcd for $C_{13}H_{28}O_3N_3$ 274.21252, found 274.21226.

***tert*-Butyl (S)-(6-(dimethylamino)-5-formamido-6-oxohexyl)carbamate (**196**)**



Acetic anhydride (0.86 mL) and formic acid (0.70 mL) were mixed and stirred for 1 hour at room temperature. Next, amine **195** (0.64 g, 2.34 mmol, 1 eq.) was dissolved in the previously prepared formylating solution (0.40 mL, 4.68 mmol of $HCOOH$, 2.0 eq.) and stirred for 1.5 hours. The reaction was quenched with saturated aq. $NaHCO_3$ (20 mL), and the product was extracted with DCM (3×20 mL). The combined organic layers were washed with saturated aq. $NaHCO_3$ (20 mL) and brine (20 mL). Then, it was dried using anhydrous $MgSO_4$ and the solvent was evaporated to afford **196** (0.68 g, 97%) as a yellow oil.

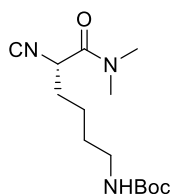
^1H NMR (400 MHz, $CDCl_3$) δ 1.30 – 1.64 (m, 14H), 1.69 – 1.81 (m, 1H), 2.95 (s, 3H), 3.03 – 3.13 (m, 5H), 4.60 (bs, 1H), 4.99 (td, $J = 8.2, 4.7$ Hz, 1H), 6.77 (d, $J = 8.3$ Hz, 1H), 8.15 – 8.18 (m, 1H).

^{13}C NMR (100 MHz, $CDCl_3$) δ 22.2, 28.4, 29.7, 32.5, 35.8, 37.1, 47.3, 79.1, 156.0, 160.6, 171.2.

ESI MS: 302 ($[M + H]^+$).

HR ESI MS: calcd for $C_{14}H_{28}O_4N_3$ 302.20743, found 302.20729.

***tert*-Butyl (S)-(6-(dimethylamino)-5-isocyano-6-oxohexyl)carbamate (**187**)**



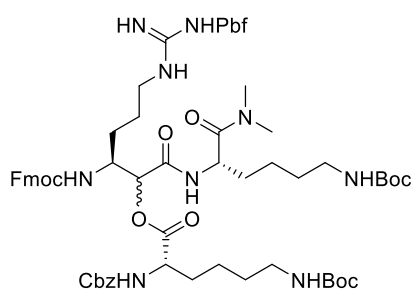
Formamide **196** (0.24 g, 0.78 mmol, 1.0 eq.) and DIPA (0.74 mL, 5.26 mmol, 6.75 eq.) were dissolved in anhydrous DCM (5 mL) under inert atmosphere.

The reaction mixture was cooled down to -20 °C and $POCl_3$ (0.21 mL, 2.34 mmol, 3.0 eq.) was added. The reaction mixture was stirred for

1.5 hours at -20 °C until disappearance of the starting material according to TLC. Then, the reaction was quenched with saturated aq. $NaHCO_3$ (20 mL) and the product was extracted with DCM (3×20 mL). The combined organic layers were then dried using anhydrous $MgSO_4$. The solvent was evaporated to yield isonitrile **187** (0.22 g, quant.) as a bright-yellow oil.

^1H NMR (400 MHz, $CDCl_3$) δ 1.41 (s, 9H), 1.47 – 1.61 (m, 4H), 1.82 – 1.94 (m, 2H), 2.97 (s, 3H), 3.03 – 3.15 (m, 5H), 4.39 (dd, $J = 8.0, 5.8$ Hz, 1H), 4.65 (bs, 1H).

The **^{13}C NMR** spectrum was not recorded due to the low stability of the compound.

Fmoc-Arg(Pbf)-CH(OOC(Lys-*N*^α-Cbz-*N*^ε-Boc)CO-Lys(Boc)-NMe₂) (197)

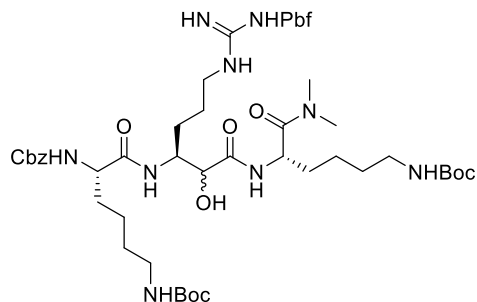
Cbz-Lys(Boc)-OH (0.26 g, 0.68 mmol, 1.0 eq.), aldehyde **186** (0.43 g, 0.68 mmol, 1.0 eq.) and isonitrile **187** (0.19 g, 0.68 mmol, 1.0 eq.) were dissolved in anhydrous DCM (5 mL) and the reaction mixture was stirred at room temperature overnight. The solvent was evaporated and the product was purified by chromatography on silica gel (hex:EtOAc / gradient 1:2 to 0:1) to afford **197** (0.22 g, 25%) as a white foam.

¹H NMR (400 MHz, CD₃OD) δ 1.31 – 1.75 (m, 38H), 1.76 – 1.97 (m, 2H), 2.04 (s, 3H), 2.07 (s, 3H), 2.52 (s, 3H), 2.59 (s, 3H), 2.91 – 2.95 (m, 5H), 2.96 – 3.34 (m, 9H), 4.15 – 4.41 (m, 4H), 4.74 – 4.99 (m, 2H), 5.01 – 5.18 (m, 2H), 5.65 – 6.05 (m, 1H), 6.13 – 6.34 (m, 2H), 7.19 – 7.41 (m, 9H), 7.52 – 7.62 (m, 2H), 7.73 (t, *J* = 6.8 Hz, 2H).

The ¹³C NMR spectrum was recorded but not interpreted due to the high complexity of the spectra caused by the presence of two diastereoisomers.

ESI MS: 1319 ([M + Na]⁺).

HR ESI MS: calcd for C₆₇H₉₃O₁₅N₉NaS 1318.64040, found 1318.64069.

Cbz-Lys(Boc)-Arg(Pbf)-CH(OH)CO-Lys(Boc)-NMe₂ (198)

Intermediate **197** (0.22 g, 0.17 mmol, 1 eq.) was dissolved in DCM (5 mL) and piperidine (1 mL) was added. The reaction mixture was stirred at room temperature for 2 hours. The volatiles were evaporated and the product was purified by chromatography on silica gel (DCM:MeOH / 15:1) to

afford **198** (0.15 g, 80%) as a white foam.

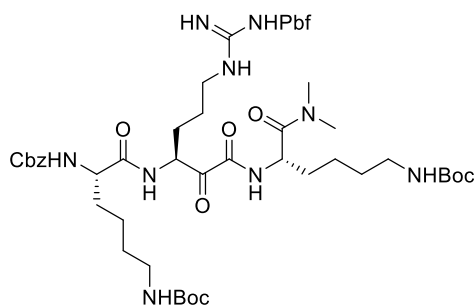
¹H NMR (400 MHz, CD₃OD) δ 1.22 – 1.82 (m, 40H), 2.07 (s, 3H), 2.51 (s, 3H), 2.57 (s, 3H), 2.93 (s, 3H), 2.95 – 3.06 (m, 6H), 3.07 – 3.20 (m, 5H), 3.97 – 4.13 (m, 2H), 4.15 – 4.32 (m, 1H), 4.99 – 5.10 (m, 2H), 6.48 – 6.59 (m, 1H), 7.20 – 7.37 (m, 5H).

The ¹³C NMR spectrum was recorded but not interpreted due to the high complexity of the spectra caused by the presence of two diastereoisomers.

ESI MS: 1097 ([M + Na]⁺).

HR ESI MS: calcd for C₅₂H₈₃O₁₃N₉NaS 1096.57233, found 1096.57137.

Cbz-Lys(Boc)-Arg(Pbf)-COCO-Lys(Boc)-NMe₂ (**199**)



Intermediate **198** (150 mg, 0.14 mmol, 1.0 eq.) was dissolved in anhydrous DCM (8 mL). Dess-Martin Periodinane (89 mg, 0.21 mmol, 1.5 eq.) was added and the reaction mixture was stirred at room temperature for 1.5 hours. Next, 10% aq. thiosulfate (20 mL) was added, and the product was extracted with DCM (20 mL). Then, the organic phase was dried using anhydrous MgSO₄ and the solvent was evaporated. The product was purified by chromatography on silica gel (DCM:MeOH / 20:1) to afford **199** (120 mg, 80%) as a white foam.

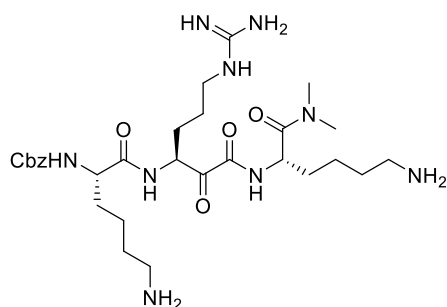
¹H NMR (400 MHz, CD₃OD) δ 1.23 – 1.89 (m, 40H), 2.06 (s, 3H), 2.51 (s, 3H), 2.57 (s, 3H), 2.93 (s, 3H), 2.96 (s, 2H), 2.99 – 3.08 (m, 4H), 3.08 – 3.18 (m, 5H), 3.98 – 4.13 (m, 1H), 4.13 – 4.30 (m, 1H), 4.68 – 4.80 (m, 1H), 4.96 – 5.13 (m, 2H), 7.17 – 7.38 (m, 5H).

¹³C NMR (100 MHz, CD₃OD) δ 12.6, 18.4, 19.6, 23.8, 24.0, 24.2, 26.2, 28.7, 28.8, 30.5, 30.6, 32.3, 32.7, 36.2, 37.7, 41.0, 41.8, 43.9, 51.0, 55.2, 56.6, 67.7, 79.8, 87.6, 118.3, 125.9, 128.8, 128.9, 129.4, 133.5, 134.5, 138.1, 139.4 (2C), 158.0, 158.4 (2C), 159.7, 161.4, 171.8, 173.4, 196.3.

ESI MS: 1095 ([M + Na]⁺).

HR ESI MS: calcd for C₅₂H₈₁O₁₃N₉NaS 1094.55668, found 1094.55565.

Cbz-Lys(H)-Arg(H)-COCO-Lys(H)-NMe₂ (**KYT-1**, **30**)



Intermediate **199** (60 mg, 56 μmol, 1.0 eq.) was dissolved in DCM (2 mL). TFA (2 mL) and TIS (30 μL, 146 μmol, 2.60 eq.) were added and the reaction mixture was stirred at room temperature for 1 hour. The volatiles were evaporated, and the product was purified by RP HPLC (C18, gradient 5 – 35% MeCN in H₂O+0.1% TFA) to afford **KYT-1** (37 mg, 69%) as a white lyophilisate.

¹H NMR (400 MHz, CD₃OD) δ 1.32 – 1.97 (m, 16H), 2.82 – 3.03 (m, 7H), 3.03 – 3.25 (m, 5H), 3.99 – 4.29 (m, 2H), 4.71 – 4.82 (m, 1H), 5.03 – 5.19 (m, 2H), 7.26 – 7.40 (m, 5H).

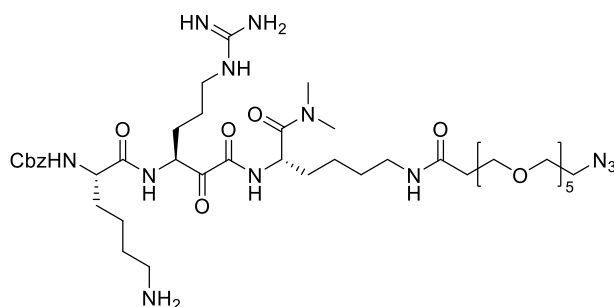
The spectrum corresponds to cyclic hemiaminal which exists as two distinct diastereoisomers.

The ¹³C NMR spectrum was recorded but not interpreted due to the high complexity of the spectra caused by the presence of the two hemiaminal diastereoisomers.

ESI MS: 620 ([M + H]⁺).

HR ESI MS: calcd for C₂₉H₅₀O₆N₉ 620.38786, found 620.38742.

Cbz-Lys(-OC-PEG-5-N₃)-Arg(H)-COCO-Lys(H)-NMe₂ (200)



KYT-1 (19 mg, 20 μ mol, 1.0 eq.) was dissolved in anhydrous DMF (1 mL) and 2,4,6-collidine (13 μ L, 100 μ L, 5.0 eq.) was added. The reaction mixture was stirred at room temperature and solution of N₃-PEG-5-CH₂CH₂COOSu (6 mg, 14 μ mol, 0.7 eq.) in anhydrous DMF (0.5 mL) was added over 2 hours. The reaction mixture was stirred overnight. The solvent was then evaporated and the reaction mixture was separated by RP HPLC (C18, gradient 5 – 45% MeCN in H₂O+0.1% TFA) to afford **200** (3.5 mg, 15%) as a white lyophilisate.

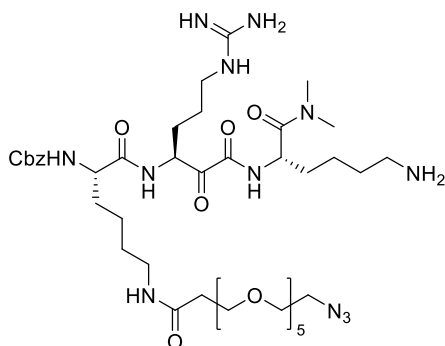
¹H NMR (400 MHz, CD₃OD) δ 1.21 – 1.87 (m, 16H), 2.44 (t, J = 6.0 Hz, 2H), 2.87 – 3.00 (m, 5H), 3.08 – 3.26 (m, 7H), 3.36 – 3.41 (m, 2H), 3.57 – 3.69 (m, 18H), 3.72 (t, J = 6.0 Hz, 2H), 4.02 – 4.29 (m, 2H), 4.70 – 4.82 (m, 1H), 5.03 – 5.14 (m, 2H), 7.16 – 7.45 (m, 5H).

The **¹³C NMR** spectrum was recorded but not interpreted due to the high complexity of the spectra caused by the presence of the two hemiaminal diastereoisomers.

ESI MS: 960 ([M + Na]⁺).

HR ESI MS: calcd for C₄₂H₇₂O₁₂N₁₂Na 959.52849, found 959.52789.

Cbz-Lys(H)-Arg(H)-COCO-Lys(-OC-PEG-5-N₃)-NMe₂ (201)



Compound **201** was prepared according to the procedure described for compound **200**. The isolated yield was 3.5 mg (15%). The **¹H NMR** spectrum was identical to compound **200**. The **¹³C NMR** spectrum was recorded but not interpreted due to the high complexity of the spectra caused by the presence of the two hemiaminal diastereoisomers.

ESI MS: 960 ([M + Na]⁺).

HR ESI MS: calcd for C₄₂H₇₂O₁₂N₁₂Na 959.52849, found 959.52768.

4.4 Polymer-Tethered Quenched Fluorescent Probes for Imaging of Tumor-Associated Proteases

4.4.1 Supplementary Figures and Tables

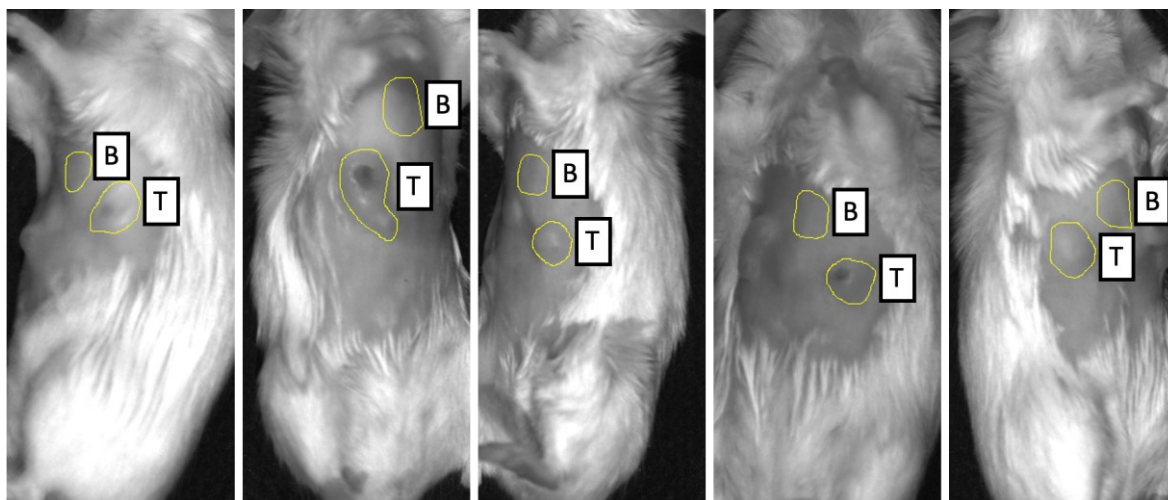


Figure 4.1 Representative images of selection of regions of interest in live mice: T – Tumor area, B – Background area.

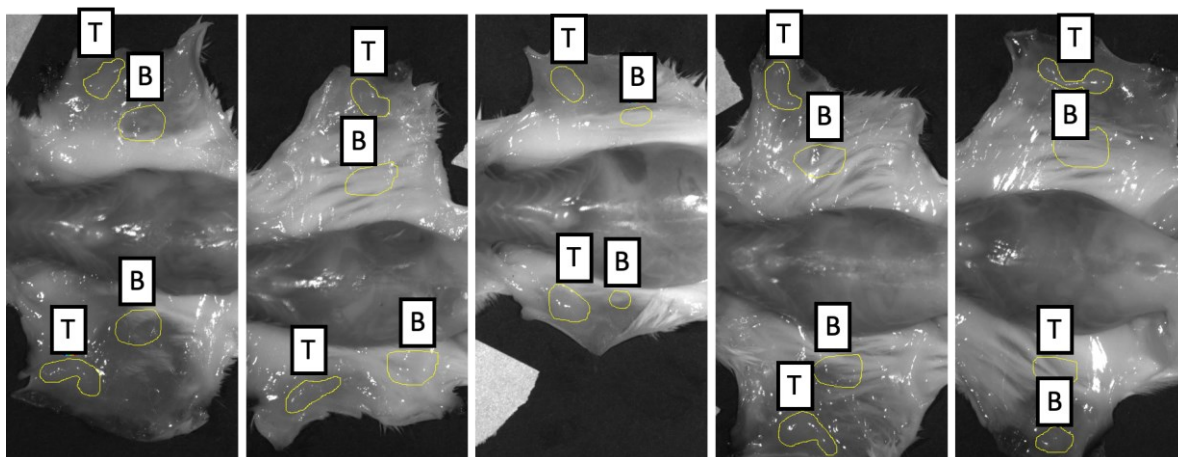


Figure 4.2 Representative images of selection of regions of interest in splayed mice: T – Tumor area, B – Background area.

Table 4.1 Evaluated probes and their *in vitro* characteristics (Initial cleavage rate rel. to 6QC and Signal-to-background ratio): k_{rel} – initial cleavage rate, relative to 6QC. Initial cleavage rate was evaluated as slope of the regression curve of the initial linear phase of the curve. k_{rel} was obtained as ratio between initial cleavage rate of the specified probe and 6QC; SBR - Signal-to-background ratio was evaluated as ratio between absorbance of the positive sample at 90 min and mean absorbance of negative samples at 90 min. All experiments were performed in triplicates.

Probe	Linker	Architecture	k_{rel} [10 μM]	k_{rel} [2.5 μM]	SBR [10 μM]	SBR [2.5 μM]
S-QF	PEG-4	QF	0.32 ± 0.01	0.39 ± 0.01	4.0 ± 0.3	5.0 ± 0.1
S-Q1:1F	PEG-4	Q1:1F	0.32 ± 0.13	0.53 ± 0.09	1.3 ± 0.1	1.5 ± 0.1
S-Q3:1F	PEG-4	Q3:1F	0.29 ± 0.01	0.52 ± 0.03	3.5 ± 0.1	4.1 ± 0.3
P-QF	polyPro	QF	0.51 ± 0.01	0.90 ± 0.03	15.2 ± 0.4	17.8 ± 0.2
L-QF	PEG-12	QF	0.49 ± 0.01	1.24 ± 0.07	24.5 ± 0.7	33.7 ± 0.5
L-FQ	PEG-12	FQ	0.45 ± 0.03	1.04 ± 0.09	29.2 ± 0.8	36.6 ± 1.5
L-Q1:1F	PEG-12	Q1:1F	0.86 ± 0.05	2.18 ± 0.14	1.0 ± 0.1	1.3 ± 0.1
L-Q2:1F	PEG-12	Q2:1F	0.77 ± 0.03	2.20 ± 0.07	3.0 ± 0.1	3.7 ± 0.1
6QC	-	-	1	1	9.6 ± 0.5	9.9 ± 0.7

Table 4.2 Evaluated probes and their *in vivo* characteristics: TBR – tumor-to-background ratio. All values are show as mean $\pm$ standard deviation. Data are calculated from $N = 4$ -10 mice (8-20 tumors) over multiple experiments.

Probe	Linker	Architecture	TBR (2h)	TBR (8h)	TBR (24h)	TBR (splay)
S-QF	PEG-4	QF	2.5 ± 0.5	2.5 ± 0.5	2.7 ± 0.6	5.1 ± 1.9
P-QF	polyPro	QF	2.8 ± 0.6	3.1 ± 0.7	2.9 ± 0.4	4.7 ± 1.2
L-QF	PEG-12	QF	3.0 ± 0.7	3.0 ± 0.9	3.1 ± 0.8	3.9 ± 1.5
L-FQ	PEG-12	FQ	3.7 ± 1.0	3.6 ± 1.1	3.2 ± 0.3	3.4 ± 1.0
L-Q2:1F	PEG-12	Q2:1F	2.9 ± 0.6	3.4 ± 0.4	3.9 ± 0.9	5.1 ± 1.3
nQ	-	-	2.1 ± 0.5	2.5 ± 0.5	3.0 ± 0.8	3.7 ± 1.4
6QC	-	-	2.0 ± 0.4	2.1 ± 0.5	2.3 ± 0.5	2.3 ± 0.4

4.4.2 Solid-Phase Peptide Synthesis

2-Chlorotrityl chloride resin (256 mg, 1.5 mmol/g) was swelled in DCM (15 mL) for 15 minutes. Next, the resin was washed with DCM (2×15 mL) solution of FmocHN-hexane-NH₂ (400 mg, 1.2 mmol) and DIPEA (410 μ L, 2.3 mmol) in DMF (7 mL) was added. This reaction mixture was rocked overnight. Then, coupling solution was discarded and the resin was washed with DMF (3×15 mL), DCM (3×15 mL) and DMF (3×15 mL). Unbound beads were capped with MeOH (15 mL) for 50 minutes. MeOH was then discarded and the resin was washed again with DMF (3×15 mL), DCM (3×15 mL) and DMF (3×15 mL). Fmoc protecting group was removed with 20% piperidine in DMF (2×10 mL, 10 min) and the resin was washed with DMF (3×15 mL), DCM (3×15 mL) and DMF (3×15 mL). Next, coupling solution of Fmoc-Lys(Boc)-OH (562 mg, 1.2 mmol), HBTU (450 mg, 1.2 mmol) and 2,4,6-collidine (310 μ L, 2.4 mmol) in DMF (10 mL) was added and the mixture was rocked for 1 hour. The solution was then discarded, and the resin was washed with DMF (3×15 mL), DCM (3×15 mL) and DMF (3×15 mL).

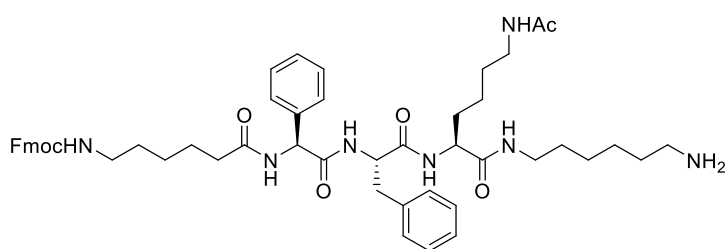
The deprotection/coupling procedure was then repeated to extend the peptide chain to yield the desired compounds. Acids and amino acids employed were: Fmoc-Phe-OH, Fmoc-Phg-OH, Fmoc-Ahx-OH. Modifications were introduced for coupling of Fmoc-Phg-OH to limit racemization. In this case, HBTU was replaced with HATU and Fmoc-Phg-peptide was deprotected using 20% 2-methylpiperidine in DMF (2×10 mL, 10 min).

The resin was washed with DCM (5×15 mL) and the peptide was cleaved from the resin using DCM:HFIP / 3:1 (10 mL). After 2 hours, the liquid was collected, and the resin was washed with DCM:HFIP / 3:1 (2×5 mL). Volatiles were then evaporated, and the solid residue was resuspended in 0.1% TFA in H₂O:MeCN / 1:1 (30 mL) and freeze-dried to yield crude peptide. The peptides were used without further purification if not stated otherwise.

4.4.3 Preparation of Low-Molecular Probe Precursors

The compounds reported in this part were not characterized by the NMR spectroscopy due to the high complexity of the compounds caused by the presence of the fluorophores in combination with the peptides. The identity of intermediates was confirmed by HPLC-MS/UPLC-MS and the final compounds were characterized by ESI MS, HR ESI MS, and the purity was confirmed by analytical HPLC. The HPLC-MS (Method A) setup used was Agilent 1100 Series HPLC equipped with a Luna 4251-E0 C18 column (3 μ m, 4.6 $\times$ 150 mm) coupled to a PE SCIEX API 3000 mass spectrometer (flow 1 mL/min, gradient 5% – 95% MeCN in H₂O+0.1% FA over 20 min). The UPLC-MS (Method B) setup used was Waters UPLC-MS Acquity equipped with Acquity UPLC BEH C18 column (1.7 μ m, 2.1 $\times$ 100 mm + 2.1 $\times$ 5 mm column guard) coupled to a QDa mass detector (flow 0.5 mL/min, gradient 5% – 95% MeCN in H₂O+0.1% FA over 7 min).

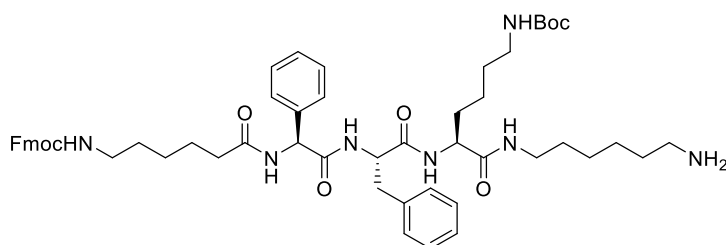
Fmoc-Ahx-Phg-Phe-Lys(Ac)-hex-NH₂ (205)



This compound was prepared according to the general procedure for SPPS. Fmoc-Ahx-Phg-Phe-Lys(Ac)-hex-NH₂ (**205**) was isolated as a white powder (175 mg, 51%).

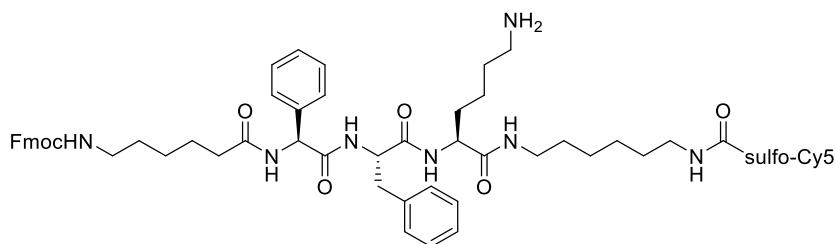
HPLC-MS (Method A): t_r = 8.67 min; observed mass: 902.4 ([M + H]⁺).

Fmoc-Ahx-Phg-Phe-Lys(Boc)-hex-NH₂ (206)



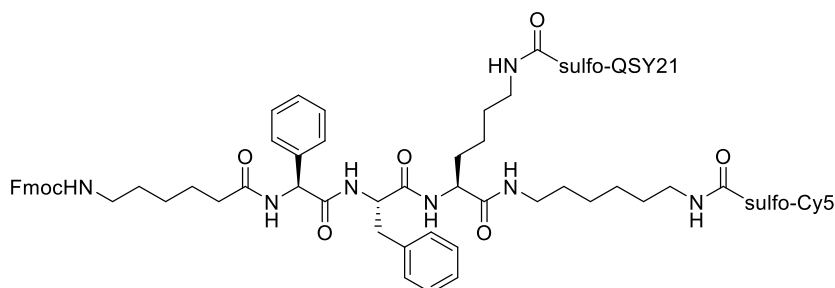
This compound was prepared according to the general procedure for SPPS and cleavage of peptides from the resin. Fmoc-Ahx-Phg-Phe-Lys(Ac)-hex-NH₂ (**206**) was isolated as a white powder (175 mg, 49%).

HPLC-MS (Method A): t_r = 9.56 min; observed mass: 960.5 ([M + H]⁺).

Fmoc-Ahx-Phe-Lys(H)-hex-sulfo-Cy5 (207)

Peptide **206** (10 mg, 9.3 μmol) was dissolved in DMF (1 mL) and DIPEA (9 μL , 47 μmol) was added to the mixture. Sulfo-Cyanine5 NHS ester (8.7 mg, 12 μmol) was then added and the reaction mixture was stirred at room temperature. The reaction was followed by HPLC-MS until full consumption of starting materials. Next, volatiles were evaporated, and the residue was redissolved in DCM (0.5 mL) and TFA (0.5 mL). The reaction mixture was stirred for 30 minutes. Then, volatiles were evaporated again. The residue was purified by RP flash chromatography (50 g C18, gradient 20 – 60% MeCN in H_2O +0.1% TFA) to obtain compound **207** as a blue powder (8.0 mg, 53%).

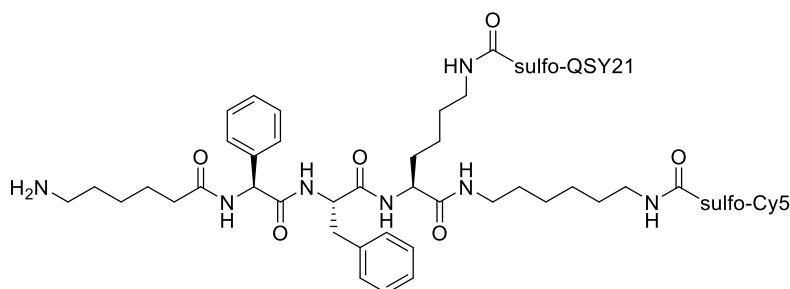
HPLC-MS (Method A): t_r = 10.10 min; observed mass: 1485.2 ($[\text{M} + \text{H}]^+$).

Fmoc-Ahx-Phe-Lys(sulfo-QSY21)-hex-sulfo-Cy5 (208)

Compound **207** (8.0 mg, 5.0 μmol) was dissolved in DMF (1 mL) and DIPEA (9 μL , 50 μmol) was added to the mixture. Sulfo-QSY21 NHS ester (6.6 mg, 7.0 μmol) was then added and the reaction mixture was stirred at room temperature. The reaction was followed by HPLC-MS until full consumption of starting materials. Next, the product was purified by RP flash chromatography (50 g C18, gradient 20 – 50% MeCN in H_2O +0.1% TFA) to obtain compound **208** as a blue powder (7.5 mg, 65%).

HPLC-MS (Method A): t_r = 10.34 min; observed mass: 1155.4 ($[\text{M} + 2\text{H}]^{2+}$).

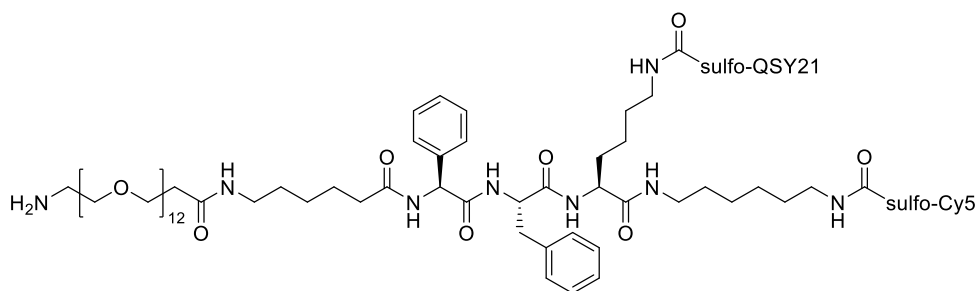
H-Ahx-Phg-Phe-Lys(sulfo-QSY21)-hex-sulfo-Cy5 (209)



Compound **208** (7.5 mg, 3.2 μmol) was dissolved in 1% DBU in DMF (1 mL) and stirred at room temperature. Reaction was quenched after 15 min by addition of AcOH (20 μL). Next, the product was purified by RP flash chromatography (50 g C18, gradient 20 – 50% MeCN in H_2O +0.1% TFA) to obtain compound **209** as a blue powder (5.5 mg, 77%).

HPLC-MS (Method A): t_r = 8.98 min; observed mass: 1044.4 ($[\text{M} + 2\text{H}]^{2+}$).

H-PEG-12-OCH₂CH₂-Ahx-Phg-Phe-Lys(QSY21)-hex-sulfo-Cy5 (210)

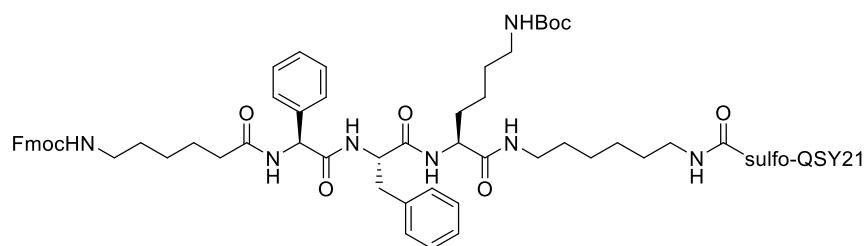


BocHN-PEG-12-OCH₂CH₂COOH (3.6 mg, 5.0 μmol) and HATU (1.8 mg, 4.8 μmol) were dissolved in DMF (1 mL). 2,4,6-Collidine (3.3 μL , 25 μmol) was added and the reaction mixture was stirred for 5 min at room temperature. Next, amine **209** (5.5 mg, 2.5 μmol) was added and the reaction was stirred for 3 hours. Then, volatiles were evaporated, and the residue was redissolved in DCM (0.5 mL) and TFA (0.5 mL). The reaction mixture was stirred for 15 min. Then, volatiles were evaporated again. The residue was purified by RP flash chromatography (50 g C18, gradient 20 – 60% MeCN in H_2O +0.1% TFA) to obtain compound **210** as a blue powder (2.4 mg, 34%).

HPLC-MS (Method A): t_r = 8.87 min; observed mass: 1344.7 ($[\text{M} + 2\text{H}]^{2+}$).

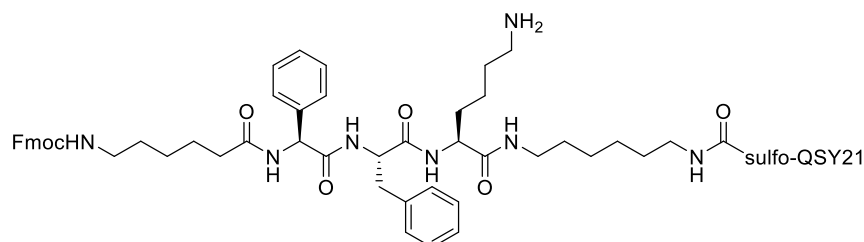
ESI MS: 1341.1 ($[\text{M} - 2\text{H}]^{2-}$).

HR ESI MS: calcd for $\text{C}_{135}\text{H}_{175}\text{O}_{34}\text{N}_{13}\text{S}_5$ 1341.04894, found 1341.04940.

Fmoc-Ahx-Phg-Phe-Lys(Boc)-hex-sulfo-QSY21 (211)

Peptide **206** (13 mg, 12 μ mol), sulfo-QSY21-COOH (10 mg, 12 μ mol) and PyBOP (8.0 mg, 15 μ mol) were dissolved in DMF (2 mL). Then, DIPEA (17 μ L, 95 μ mol) was added to the mixture and the reaction mixture was stirred at room temperature. The reaction was followed by HPLC-MS until full consumption of starting materials. Next, volatiles were evaporated and the residue was purified by RP flash chromatography (50 g C18, gradient 20 – 60% MeCN in H₂O+0.1% TFA) to obtain compound **211** as a blue powder (8.0 mg, 38%).

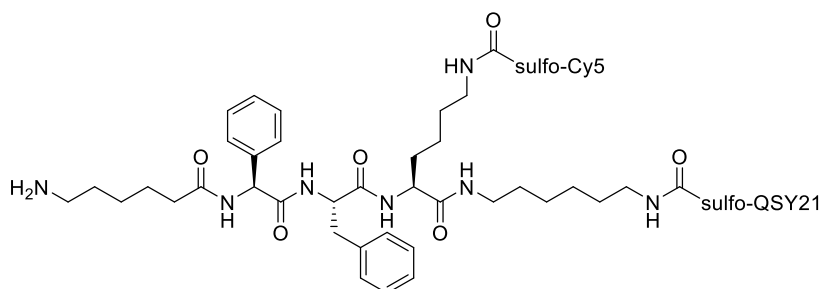
HPLC-MS (Method B): t_r = 4.84 min; observed mass: 842.2 ($[M + 2H - Boc]^{2+}$).

Fmoc-Ahx-Phg-Phe-Lys(H)-hex-sulfo-QSY21 (212)

Compound **211** (23 mg, 13 μ mol) was dissolved in DCM (0.5 mL) and TFA (0.5 mL). The reaction mixture was stirred for 30 minutes at room temperature. Then, volatiles were evaporated, and the residue was purified by RP flash chromatography (50 g C18, gradient 20 – 50% MeCN in H₂O+0.1% TFA) to obtain compound **212** as a blue powder (18 mg, 78%).

HPLC-MS (Method B): t_r = 4.12 min; observed mass: 842.8 ($[M + 2H]^{2+}$).

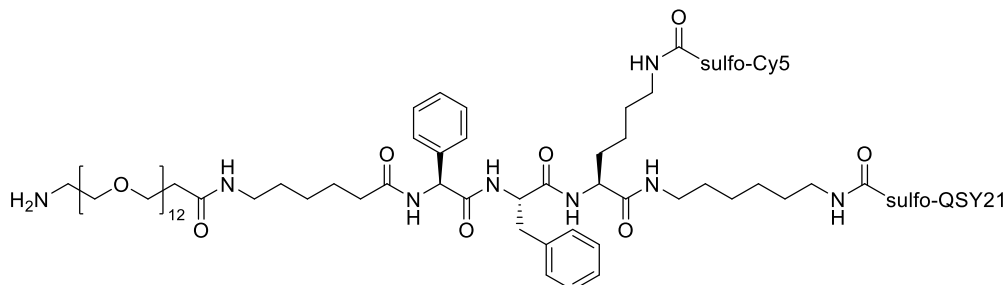
H-Ahx-Phg-Phe-Lys(Cy5)-hex-sulfo-QSY21 (**213**)



Compound **212** (18 mg, 10 μ mol) was dissolved in DMF (2 mL) and DIPEA (26 μ L, 150 μ mol) was added to the mixture. Sulfo-Cyanine5 NHS ester (8.9 mg, 12 μ mol) was then added and the reaction mixture was stirred at room temperature. The reaction was followed by HPLC-MS until full consumption of starting materials. Next, 2% solution of DBU in DMF (2 mL) was added and the reaction mixture was stirred for 15 minutes. Then, 150 μ L of AcOH were added to quench the reaction. Then, volatiles were evaporated, and the residue was purified by RP flash chromatography (50 g C18, gradient 10 – 40% MeCN in H₂O+0.1% TFA) to obtain compound **213** as a blue powder (14 mg, 64%).

HPLC-MS (Method B): t_r = 3.45 min; observed mass: 696.2 ($[M + 3H]^{3+}$).

H-PEG-12-OCH₂CH₂-Ahx-Phg-Phe-Lys(Cy5)-hex-sulfo-QSY21 (**214**)



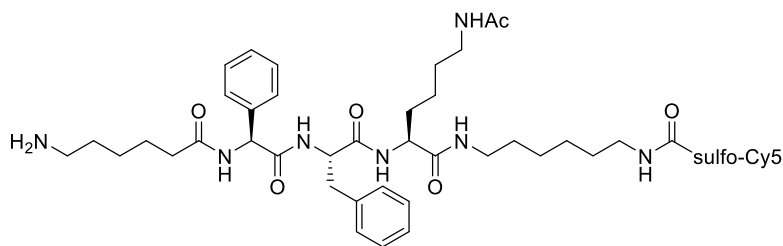
BocHN-PEG-12-OCH₂CH₂COOH (9.1 mg, 13 μ mol) and HATU (4.4 mg, 11 μ mol) were dissolved in DMF (2 mL). 2,4,6-Collidine (13 μ L, 95 μ mol) was added and the reaction mixture was stirred for 5 min at room temperature. Next, amine **213** (14 mg, 6.4 μ mol) was added and the reaction was stirred for 3 hours. Then, volatiles were evaporated, and the residue was redissolved in DCM (0.5 mL) and TFA (0.5 mL). The reaction mixture was stirred for 15 minutes. Then, volatiles were evaporated again, and the residue was purified by RP flash chromatography (50 g C18, gradient 15 – 50% MeCN in H₂O+0.1% TFA) to obtain compound **214** as a blue powder (7.0 mg, 39%).

HPLC-MS (Method B): t_r = 3.52 min; observed mass: 896.6 ($[M + 3H]^{3+}$).

ESI MS: 1341.6 ($[M - 2H]^{2-}$).

HR ESI MS: calcd for C₁₃₅H₁₇₅O₃₄N₁₃S₅ 1341.04894, found 1341.04955.

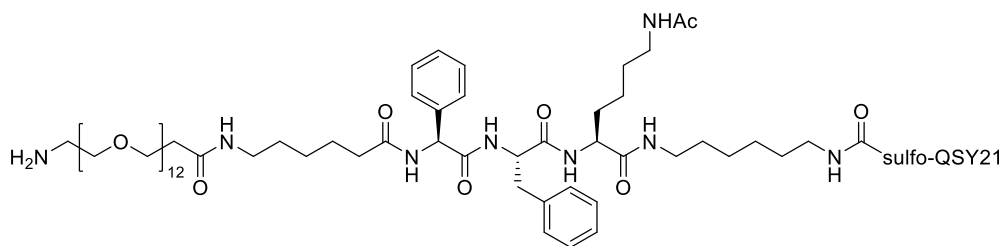
H-Ahx-Phg-Phe-Lys(Ac)-hex-sulfo-Cy5 (215)



Peptide **205** (10 mg, 9.8 μmol) was dissolved in DMF (1 mL) and DIPEA (9 μL , 49 μmol) was added to the mixture. Sulfo-Cyanine5 NHS ester (9.2 mg, 12 μmol) was then added and the reaction mixture was stirred at room temperature. The reaction was followed by HPLC-MS until full consumption of starting materials. Next, DBU (10 μL) was added to achieve Fmoc deprotection and the reaction mixture was stirred for 15 minutes at room temperature. It was then quenched by addition of AcOH (20 μL) and purified by RP flash chromatography (50 g C18, gradient 20 – 40% MeCN in H_2O +0.1% TFA) to obtain compound **215** as a blue powder (5.0 mg, 36%).

HPLC-MS (Method A): t_r = 8.89 min; observed mass: 1305.1 ($[\text{M} + \text{H}]^+$).

H-PEG-12-OCH₂CH₂-Ahx-Phg-Phe-Lys(Ac)-hex-sulfo-Cy5 (216)



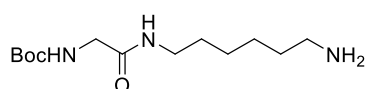
BocHN-PEG-12-OCH₂CH₂COOH (5.0 mg, 7.0 μmol) and HATU (2.4 mg, 6.3 μmol) were dissolved in DMF (1 mL). 2,4,6-Collidine (5.0 μL , 35 μmol) was added and the reaction mixture was stirred for 5 minutes at room temperature. Next, amine **215** (5.0 mg, 3.5 μmol) was added and the reaction was stirred for 1 hour. Then, volatiles were evaporated, and the residue was redissolved in DCM (0.5 mL) and TFA (0.5 mL). The reaction mixture was stirred for 15 minutes. Then, volatiles were evaporated again, and the residue was purified by RP flash chromatography (50 g C18, gradient 20 – 40% MeCN in H_2O +0.1% TFA) to obtain compound **216** as blue powder (5.5 mg, 77%).

HPLC-MS (Method A): t_r = 8.69 min; observed mass: 953.5 ($[\text{M} + 2\text{H}]^{2+}$).

ESI MS: 951.0 ($[\text{M} - 2\text{H}]^{2-}$).

HR ESI MS: calcd for $\text{C}_{96}\text{H}_{144}\text{O}_{25}\text{N}_{10}\text{S}_2$ 950.48782, found 950.48743.

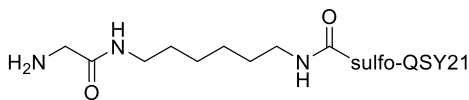
Boc-Gly-hex-NH₂ (217)



Boc-Gly-OH (187 mg, 1.1 mmol) and HCTU (440 mg, 1.1 mmol) were dissolved in DMF (5 mL). 2,4,6-Collidine (0.85 mL, 6.4 mmol) was added, and the reaction mixture was stirred for 5 minutes at room temperature. Next, *N*-Fmoc-1,6-diaminohexane (400 mg, 1.1 mmol) was added and the reaction was stirred overnight. Then, the solvent was evaporated, DCM was added (20 mL), and the organic phase was washed with 10% KHSO₄ (2 × 20 mL), H₂O (2 × 20 mL), saturated aq. NaHCO₃ (2 × 20 mL) and brine (10 mL). Then, volatiles were evaporated, and the residue was redissolved in 20% piperidine in DMF (5 mL). The reaction was followed by HPLC-MS until full consumption of the Fmoc-protected intermediate. Volatiles were evaporated and the residue was purified by RP flash chromatography (50 g C18, gradient 10 – 40% MeCN in H₂O+0.1% TFA) to obtain compound **217** as a white powder (187 mg, 45%).

HPLC-MS (Method A): t_r = 7.23 min; observed mass: 274.3 ([M + H]⁺).

H-Gly-hex-sulfo-QSY21 (218)



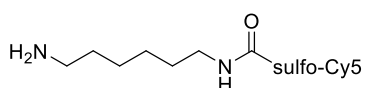
Compound **217** (6.0 mg, 22 μmol) was dissolved in DMF (1 mL) and DIPEA (57 μL, 329 μmol) was added to the mixture. QSY21 NHS ester (10 mg, 11 μmol) was then added and the reaction was stirred at room temperature. The reaction was followed by HPLC-MS until full consumption of starting materials. Then, volatiles were evaporated, and the residue was redissolved in DCM (0.5 mL) and TFA (0.5 mL). The reaction mixture was stirred for 15 minutes. Then, volatiles were evaporated again, and the product was purified by RP flash chromatography (50 g C18, gradient 20 – 60% MeCN in H₂O+0.1% TFA) to obtain compound **218** as a blue powder (5.2 mg, 47%).

HPLC-MS (Method A): t_r = 7.74 min; observed mass: 997.7 ([M + H]⁺).

ESI MS: 995.3 ([M – H]⁻).

HR ESI MS: calcd for C₄₉H₅₁O₁₁N₆S₃ 995.27834, found 995.27899.

H-hex-sulfo-Cy5 (220)



N-Boc-1,6-diaminohexane (7.0 mg, 32 μmol) was dissolved in DMF (1 mL) and DIPEA (24 μL, 135 μmol) was added to the mixture. Sulfo-Cyanine5 NHS ester (20 mg, 27 μmol) was then added and the reaction was stirred at room temperature. The reaction was followed by HPLC-MS until full consumption of starting materials. Then, volatiles were evaporated, and the residue was

redissolved in DCM (0.5 mL) and TFA (0.5 mL). The reaction mixture was stirred for 15 minutes. Then, volatiles were evaporated again, and the product was purified by RP flash chromatography (50 g C18, gradient 10 – 40% MeCN in H₂O+0.1% TFA) to obtain compound **220** as blue powder (21 mg, 92%).

HPLC-MS (Method B): (Boc-protected intermediate) $t_r = 3.68$ min; observed mass: 841.7 ([M + H]⁺).

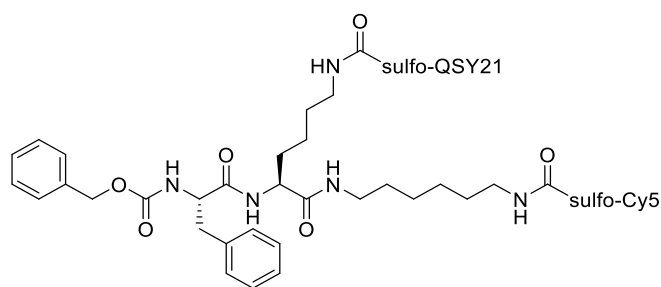
H-Gly-hex-sulfo-Cy5 (**221**)

Compound **220** (16 mg, 19 μmol) was dissolved in DMF (1 mL) and DIPEA (16 μL, 93 μmol) was added to the mixture. Boc-Gly-OSu (6.1 mg, 22 μmol) was then added and the reaction was stirred at room temperature. The reaction was followed by HPLC-MS until full consumption of starting materials. Then, volatiles were evaporated, and the residue was redissolved in DCM (0.5 mL) and TFA (0.5 mL). The reaction mixture was stirred for 15 minutes. Next, volatiles were evaporated again, and the product was purified by RP flash chromatography (50 g C18, gradient 10 – 40% MeCN in H₂O+0.1% TFA) to obtain compound **221** as blue powder (13 mg, 76%).

ESI MS: 796.3 ([M – H]⁻).

HR ESI MS: Calculated for C₄₀H₅₄O₈N₅S₂ 796.34193. Found 796.34137.

6QC (**222**)



6QC was synthesized according to a published procedure.¹²⁵

4.4.4 4T1 Breast Tumor Model

Prior to injecting the cells for the tumor model, 4T1 cells were cultured to confluency in treated 10 cm dishes. 4T1 cells were detached using pre-warmed Trypsin (Gibco, Cat. No.: 25300062) via incubation for 3 min at 37 °C. Cells were detached with additional pipetting after the addition of cell culture media. Cells were then centrifuged for 3 min at 250 G and

the media removed. Cells were resuspended in PBS, centrifuged, and PBS removed a total of three times to wash the cells. The 4T1 cells were then resuspended in PBS to a dilution of 1×10^6 cells per mL. While under isoflurane anesthesia, mice were subcutaneously injected in the third and eighth mammary fat pads of BALB/c female mice (aged 6–8 weeks; Jackson Laboratory) with 100 μ L of the diluted 4T1 cells (1×10^5 cells per fat pad). The seeded tumors were allowed to grow for 7 – 10 days. Body fur was removed using Nair lotion in the mammary fat pads and surrounding area while under anesthesia prior to probe injection. Once the tumors were developed, mice were injected intravenously with 6.25 nmols of probe in 100 μ L using a 28-gauge 1 mL insulin syringe into the tail vein. **6-QC** was dissolved in a 100 μ L formulation of 1 $\times$ PBS with 10% DMSO and 30% PEG 400. All macromolecular probes were dissolved in a formulation of 1 $\times$ PBS with 10% DMSO. Post injection, the mice were non-invasively imaged using the LI-COR Pearl Trilogy imaging system at multiple time points. Mice were positioned to align tumor to the center of the camera field of view. The 700 nm channel and accompanying standard excitation/emission settings were used to image Cyanine5 fluorescence along with the standard white light channel filters at a resolution of 85 μ m. After the live imaging, mice were euthanized using cervical dislocation under isoflurane anesthesia. Mice were then splayed and imaged. Finally, through dissection organs were collected (liver, kidney, spleen, lung, heart) as well as the primary tumors and non-injected fat pad controls. These tissues were then imaged *ex-vivo*. Fluorescence intensity was measured using the built-in LI-COR Image Studio software. The regions of interest (tumor, background) were selected using the white light image only with the researcher not being blinded to the condition. Mean fluorescence values in these areas were calculated by the software and divided to obtain TBR values. Both regions were selected using a hand drawn region of interest with the tumor easily identifiable by eye. The background was defined as a medial chest between and slightly superior to the tumors in live mice and as neighboring healthy tissue in splayed mice. For quantification of *ex-vivo* tissues, the average fluorescence intensity was measured in the tissue of interest using the largest ellipsoid region of interest possible for each tissue. Fluorescence intensity color plots were generated using the built in LI-COR Image Studio software. All images shown were linked to display the same brightness and contrast settings for a given condition (example: 24-hour images for all probes are linked). Plots were generated and statistics evaluated using GraphPad Prism 10.2.0 (392).

5 CONCLUSION

Three separate scientific problems were addressed within the scope of this dissertation thesis. Overall, the three projects were centered on the design and synthesis of various protease inhibitors and substrates with particular biological properties.

First, a novel class of macrocyclic peptidomimetic inhibitors of hCatD inspired by PepA was introduced (**Chapter 3.1**). A previously published synthetic approach was employed for the preparation of six novel statine analogues (**72-77**). An approach for the synthesis of statine analogues with basic heteroaromatic residues was also investigated. Exploiting the newly synthesized statine analogues and commercially available amino acid derivatives, 27 macrocyclic inhibitors of hCatD were synthesized, characterized and evaluated for hCatD inhibition. Several inhibitors exhibited single-digit inhibitory potency, with **125** ($IC_{50} = 0.6 \pm 0.2$ nM) and **106** ($IC_{50} = 1.1 \pm 0.2$ nM) reaching the highest potency.

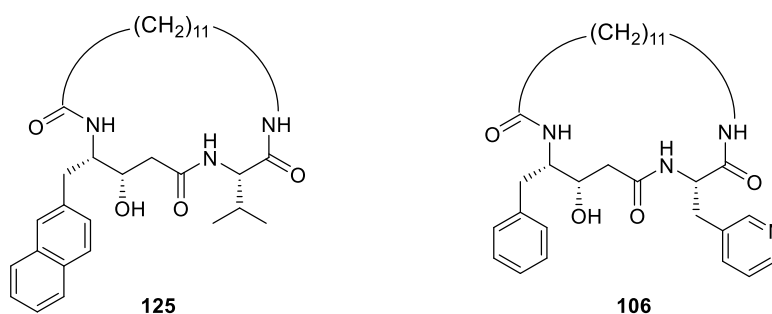


Figure 5.1 Structures of the most potent macrocyclic inhibitors reported in this work

Second, two known gingipain inhibitors, **KYT-36** (Kgp) and **KYT-1** (Rgp), were modified with a linker to yield chemical probes for the development of an assay for the detection and quantification of gingipains in various samples (**Chapter 3.2**). The design was based on a structural analysis of the crystal structure of Kgp with **KYT-36** in the active site. A synthetic methodology, involving Passerini reaction, was designed for the synthesis of modified **KYT-36** and **KYT-1**. The methodology was successfully applied to yield linker-modified **KYT-36** (compound **184**) and two regioisomers of modified **KYT-1** (compound **200** and **201**). The synthesized compounds were evaluated *in vitro*, and high-affinity inhibition of the respective proteases was confirmed for all three compounds.

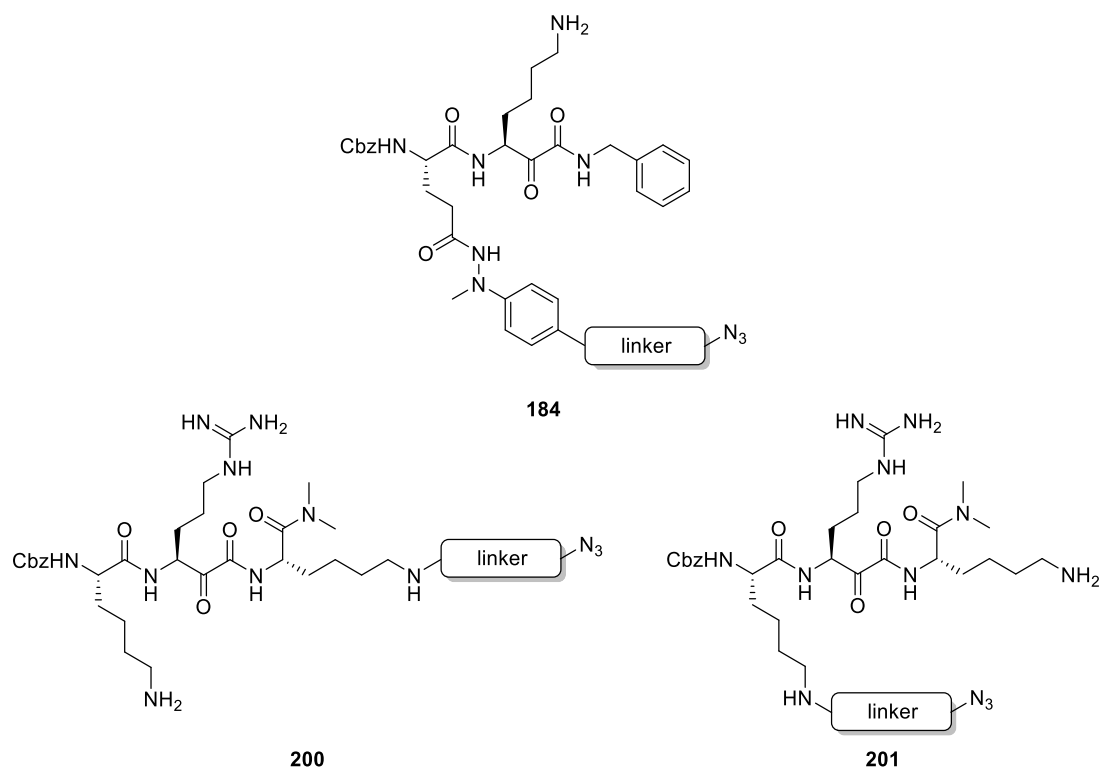


Figure 5.2 Structures of synthesized linker-modified gingipain inhibitors **184**, **200** and **201**

Third, a previously reported fluorescent pan-cathepsin substrate **6QC**, used for tumor imaging, was successfully modified with an appropriate linker to allow tethering to a macromolecular pHPMA scaffold (**Chapter 3.3**). Multiple probe architectures were designed, and the required low-molecular probe precursors were successfully synthesized (**210**, **214**, **216**, **218**, **221**). The resulting macromolecular probes were evaluated *in vitro* in a purified enzyme assay, revealing several interesting SAR trends in cleavage efficiency. Lead probes were tested in a cellular assay and, most importantly, *in vivo* in a mouse breast cancer model. Overall, the majority of synthesized pHPMA probes exhibited enhanced imaging properties compared to **6QC**, as well as improved biodistribution. This proves that pHPMA is a promising scaffold for the development of novel fluorescent probes for protease imaging *in vitro* and *in vivo*.

To summarize, the three projects presented in this thesis were finalized from the synthetic point of view and the biological properties of synthesized compounds were evaluated. In the case of the inhibitors of hCatD and probes for cancer imaging, the biological evaluation was finished and published successfully. Optimization of the assay for quantification of gingipains is underway and yet to be published.

6 LIST OF AUTHOR'S PUBLICATIONS

Articles

Hadzima, M., Faucher, F., Blažková, K., Yim, J., Guerra, M., Chen, S., Woods, E., Park, K. W., Šácha, P., Šubr, V., Kostka, L., Etrych, T., Majer, P., Konvalinka, J., Bogyo, M. Polymer-tethered quenched fluorescent probes for enhanced imaging of tumor-associated proteases. *ACS Sens.* **9**, 3720-3729 (2024).

Zamani, M. R., **Hadzima, M.**, Blažková, K., Šubr, V., Ormsby, T., Celis-Gutierrez, J., Malissen, B., Etrych, T., Kostka, L., Šácha, P., Konvalinka, J. Polymer-based antibody mimetics (iBodies) target human PD-L1 and function as a potent immune checkpoint blocker. *J. Biol. Chem.* **6**, 107325 (2024).

Kugler, M., **Hadzima, M.**, Dzijak, R., Rampmaier, R., Voburka, Z., Majer, P., Řezáčová, P., Vrábel, M. Identification of advanced carbonic anhydrase inhibitors via in situ click chemistry, phage-displayed and synthetic peptide libraries. Comparison of the methods and structural biology study. *RSC Med. Chem.* **14**, 144-153 (2023).

Houštecká, R., **Hadzima, M.**, Fanfrlík, J., Brynda, J., Pallová, L., Hánová, I., Mertlíková-Kaiserová, H., Lepšík, M., Horn, M., Smrčina, M., Majer, P., Mareš, M. Biomimetic Macrocyclic Inhibitors of Human Cathepsin D: Structure–Activity Relationship and Binding Mode Analysis. *J. Med. Chem.* **63**, 1576-1596, (2020).

Kutil, Z., Škultétyová, L., Rauh, D., Meleshin, M., Šnajdr, I., Nováková, Z., Mikešová, J., Pavlíček, J., **Hadzima, M.**, Baranová, P., Havlínová, B., Majer, P., Schutkowski, M., Bařinka, C. The unraveling of substrate specificity of histone deacetylase 6 domains using acetylome peptide microarrays and peptide libraries. *FASEB J.* **33**, 4035–4045, (2019).

Patents

Bell, B., Rais, R., Rangaramanujam, K., Sharma, A., Slusher, B., Tsukamoto, T., Gao, R.-D., Majer, P., Novotná, K., Tenora, L., **Hadzima, M.**, Kaiser, M., Tichý, T., Zhu, X. Dendrimer-glutamine antagonist conjugates and methods of use thereof. WO2022221490A1, (2022).

Majer, P., Šnajdr, I., **Hadzima, M.**, Tenora, L., Lawson, J. P., Wild, R. C., Shao, Y., Chu, J., Weng J. Method of preparing a DON prodrug from *L*-glutamic acid. WO2020167831A1, (2020).

7 REFERENCES

1. Quintero-Fabián, S. *et al.* Role of Matrix Metalloproteinases in Angiogenesis and Cancer. *Front. Oncol.* **9**, 480558 (2019).
2. Egeblad, M. & Werb, Z. New functions for the matrix metalloproteinases in cancer progression. *Nat. Rev. Cancer* **2**, 161–174 (2002).
3. Mustafa, S., Koran, S. & Alomair, L. Insights Into the Role of Matrix Metalloproteinases in Cancer and its Various Therapeutic Aspects: A Review. *Front. Mol. Biosci.* **9**, 896099 (2022).
4. Benes, P., Vetvicka, V. & Fusek, M. Cathepsin D—Many functions of one aspartic protease. *Crit. Rev. Oncol. Hematol.* **68**, 12–28 (2008).
5. Liaudet-Coopman, E. *et al.* Cathepsin D: newly discovered functions of a long-standing aspartic protease in cancer and apoptosis. *Cancer Lett.* **237**, 167–179 (2006).
6. Foekens, J. A. *et al.* Cathepsin-D in primary breast cancer: prognostic evaluation involving 2810 patients. *Br. J. Cancer* **79**, 300–307 (1998).
7. Hao, L. *et al.* Deficiency of cathepsin K prevents inflammation and bone erosion in rheumatoid arthritis and periodontitis and reveals its shared osteoimmune role. *FEBS Lett.* **589**, 1331–1339 (2015).
8. Cathepsin K is a potential therapeutic target for autoimmune arthritis. *Nat. Clin. Pract. Rheumatol.* **4**, 231–231 (2008).
9. Suzuki, N., Takimoto, K. & Kawashima, N. Cathepsin K Inhibitor Regulates Inflammation and Bone Destruction in Experimentally Induced Rat Periapical Lesions. *J. Endod.* **41**, 1474–1479 (2015).
10. Jugniot, N., Voisin, P., Bentaher, A. & Mellet, P. Neutrophil Elastase Activity Imaging: Recent Approaches in the Design and Applications of Activity-Based Probes and Substrate-Based Probes. *Contrast Media Mol. Imaging* 2019, 7417192 (2019).
11. Krotova, K., Khodayari, N., Oshins, R., Aslanidi, G. & Brantly, M. L. Neutrophil elastase promotes macrophage cell adhesion and cytokine production through the integrin-Src kinases pathway. *Sci. Rep.* **10**, 1–10 (2020).
12. Döring, G. The Role of Neutrophil Elastase in Chronic Inflammation. *Am. J. Respir. Crit. Care Med.* **150**, S114–S117 (2012).

13. O'Brien-Simpson, N. M., Veith, P. D., Dashper, S. G. & Reynolds, E. C. Antigens of bacteria associated with periodontitis. *Periodontol.* **2000** **35**, 101–134 (2004).
14. Curtis, M. A. *et al.* Molecular genetics and nomenclature of proteases of *Porphyromonas gingivalis*. *J. Periodontal. Res.* **34**, 464–472 (1999).
15. Kadowaki, T. *et al.* Suppression of Pathogenicity of *Porphyromonas gingivalis* by Newly Developed Gingipain Inhibitors. *Mol. Pharmacol.* **66**, 1599–1606 (2004).
16. Lv, Z., Chu, Y. & Wang, Y. HIV protease inhibitors: a review of molecular selectivity and toxicity. *HIV/AIDS (Auckl.)* **7**, 95–104 (2015).
17. Brik, A. & Wong, C. H. HIV-1 protease: mechanism and drug discovery. *Org. Biomol. Chem.* **1**, 5–14 (2003).
18. Weber, I. T., Wang, Y. F. & Harrison, R. W. HIV Protease: Historical Perspective and Current Research. *Viruses* **13**, 839 (2021).
19. Pang, X., Xu, W., Liu, Y., Li, H. & Chen, L. The research progress of SARS-CoV-2 main protease inhibitors from 2020 to 2022. *Eur. J. Med. Chem.* **257**, 115491 (2023).
20. Jin, Z. *et al.* Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **582**, 289–293 (2020).
21. Chan, H. T. H. *et al.* Discovery of SARS-CoV-2 Mpro peptide inhibitors from modelling substrate and ligand binding. *Chem. Sci.* **12**, 13686–13703 (2021).
22. Lkhider, M., Castino, R., Bouguyon, E., Isidoro, C. & Ollivier-Bousquet, M. Cathepsin D released by lactating rat mammary epithelial cells is involved in prolactin cleavage under physiological conditions. *J. Cell. Sci.* **117**, 5155–5164 (2004).
23. Baldwin, E. T. *et al.* Crystal structures of native and inhibited forms of human cathepsin D: implications for lysosomal targeting and drug design. *Proc. Natl. Acad. Sci.* **90**, 6796–6800 (1993).
24. Aghdassi, A. A. *et al.* Cathepsin D regulates cathepsin B activation and disease severity predominantly in inflammatory cells during experimental pancreatitis. *J. Biol. Chem.* **293**, 1018–1029 (2018).
25. Ferrandina, G. *et al.* Relationship between cathepsin D content and disease-free survival in node-negative breast cancer patients: a meta-analysis. *Br. J. Cancer* **76**, 661–666 (1997).
26. Alhudiri, I. *et al.* Expression of Cathepsin D in early-stage breast cancer and its prognostic and predictive value. *Breast Cancer Res. Treat.* **206**, 143–153 (2024).
27. Rochefort, H. *et al.* Cathepsin D in breast cancer: mechanisms and clinical applications, a 1999 overview. *Clin. Chim. Acta* **291**, 157–170 (2000).

28. Ketterer, S. *et al.* Cathepsin D deficiency in mammary epithelium transiently stalls breast cancer by interference with mTORC1 signaling. *Nat. Commun.* **11**, 1–18 (2020).
29. Zhang, C., Zhang, M. & Song, S. Cathepsin D enhances breast cancer invasion and metastasis through promoting hepsin ubiquitin-proteasome degradation. *Cancer Lett.* **438**, 105–115 (2018).
30. Ashraf, Y. *et al.* Immunotherapy of triple-negative breast cancer with cathepsin D-targeting antibodies. *J. Immunother. Cancer* **7**, 29 (2019).
31. Masson, O. *et al.* Pathophysiological functions of cathepsin D: Targeting its catalytic activity versus its protein binding activity? *Biochimie* **92**, 1635–1643 (2010).
32. Agarwal, N. S. & Rich, D. H. Inhibition of cathepsin D by substrate analogs containing statine and by analogs of pepstatin. *J. Med. Chem.* **29**, 2519–2524 (1986).
33. Jupp, R. A. *et al.* The selectivity of statine-based inhibitors against various human aspartic proteinases. *Biochem. J.* **265**, 871–878 (1990).
34. Kick, E. K. *et al.* Structure-based design and combinatorial chemistry yield low nanomolar inhibitors of cathepsin D. *Chem. Biol.* **4**, 297–307 (1997).
35. Kick, E. K. & Ellman, J. A. Expedient Method for the Solid-Phase Synthesis of Aspartic Acid Protease Inhibitors Directed toward the Generation of Libraries. *J. Med. Chem.* **38**, 1427–1430 (1995).
36. Lee, C. E., Kick, E. K. & Ellman, J. A. General solid-phase synthesis approach to prepare mechanism-based aspartyl protease inhibitor libraries. Identification of potent cathepsin D inhibitors. *J. Am. Chem. Soc.* **120**, 9735–9747 (1998).
37. Bi, X. *et al.* Novel cathepsin D inhibitors block the formation of hyperphosphorylated tau fragments in hippocampus. *J. Neurochem.* **74**, 1469–1477 (2000).
38. Dumas, J. *et al.* Synthesis and structure activity relationships of novel small molecule cathepsin D inhibitors. *Bioorg. Med. Chem. Lett.* **9**, 2531–2536 (1999).
39. Houšteká, R. *et al.* Biomimetic Macrocyclic Inhibitors of Human Cathepsin D: Structure-Activity Relationship and Binding Mode Analysis. *J. Med. Chem.* **63**, 1576–1596 (2020).
40. 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.* **46**, 3–26 (2001).
41. Majer, P., Collins, J. R., Gulnik, S. V. & Erickson, J. W. Structure-based subsite specificity mapping of human cathepsin D using statine-based inhibitors. *Protein Sci.* **6**, 1458–1466 (1997).

42. Paster, B. J., Olsen, I., Aas, J. A. & Dewhirst, F. E. The breadth of bacterial diversity in the human periodontal pocket and other oral sites. *Periodontol. 2000* **42**, 80–87 (2006).
43. Mohanty, R. *et al.* Red complex: Polymicrobial conglomerate in oral flora: A review. *J. Family Med. Prim. Care* **8**, 3480 (2019).
44. Nocini, R., Lippi, G. & Mattiuzzi, C. Periodontal disease: the portrait of an epidemic. *J. Public Health Emerg.* **4**, (2020).
45. Howard, K. C., Gonzalez, O. A. & Garneau-Tsodikova, S. *Porphyromonas gingivalis* : where do we stand in our battle against this oral pathogen? *RSC Med. Chem.* **12**, 666–704 (2021).
46. How, K. Y., Song, K. P. & Chan, K. G. *Porphyromonas gingivalis*: An Overview of Periodontopathic Pathogen below the Gum Line. *Front. Microbiol.* **7**, (2016).
47. Fiorillo, L. *et al.* *Porphyromonas gingivalis*, Periodontal and Systemic Implications: A Systematic Review. *Dent. J. (Basel)* **7**, 114 (2019).
48. Socransky, S. S., Haffajee, A. D., Cugini, M. A., Smith, C. & Kent, R. L. Microbial complexes in subgingival plaque. *J. Clin. Periodontol.* **25**, 134–144 (1998).
49. Rams, T. E., Sautter, J. D. & van Winkelhoff, A. J. Emergence of Antibiotic-Resistant *Porphyromonas gingivalis* in United States Periodontitis Patients. *Antibiotics* **12**, 1584 (2023).
50. Wang, B. *et al.* The Roles and Interactions of *Porphyromonas gingivalis* and *Fusobacterium nucleatum* in Oral and Gastrointestinal Carcinogenesis: A Narrative Review. *Pathogens* **13**, 93 (2024).
51. Jia, S., Li, X. & Du, Q. Host insulin resistance caused by *Porphyromonas gingivalis*-review of recent progresses. *Front. Cell. Infect. Microbiol.* **13**, (2023).
52. Shapira, L., Ayalon, S. & Brenner, T. Effects of *Porphyromonas gingivalis* on the Central Nervous System: Activation of Glial Cells and Exacerbation of Experimental Autoimmune Encephalomyelitis. *J. Periodontol.* **73**, 511–516 (2002).
53. Beck, J. D., Pankow, J., Tyroler, H. A. & Offenbacher, S. Dental infections and atherosclerosis. *Am. Heart J.* **138**, S528–S533 (1999).
54. Afzoon, S., Amiri, M. A., Mohebbi, M., Hamedani, S. & Farshidfar, N. A systematic review of the impact of *Porphyromonas gingivalis* on foam cell formation: Implications for the role of periodontitis in atherosclerosis. *BMC Oral Health* **23**, 1–13 (2023).
55. Fujise, K., Kikuchi, Y., Kokubu, E., Okamoto-Shibayama, K. & Ishihara, K. Effect of extracytoplasmic function sigma factors on autoaggregation, hemagglutination,

- and cell surface properties of *Porphyromonas gingivalis*. *PLoS One* **12**, e0185027 (2017).
56. Singhrao, S. K., Harding, A., Poole, S., Kesavalu, L. & Crean, S. J. *Porphyromonas gingivalis* Periodontal Infection and Its Putative Links with Alzheimer's Disease. *Mediators Inflamm.* **2015**, 137357 (2015).
 57. Li, D., Ren, T., Li, H., Liao, G. & Zhang, X. *Porphyromonas gingivalis*: A key role in Parkinson's disease with cognitive impairment? *Front. Neurol.* **13**, 945523 (2022).
 58. Olsen, I., Kell, D. B. & Pretorius, E. Is *Porphyromonas gingivalis* involved in Parkinson's disease? *Eur. J. Clin. Microbiol. Infect. Dis.* **39**, 2013–2018 (2020).
 59. Ermini, F. *et al.* Ultrastructural localization of *Porphyromonas gingivalis* gingipains in the substantia nigra of Parkinson's disease brains. *NPJ Parkinson's Dis.* **10**, 1–16 (2024).
 60. Li, D. *et al.* Oral Microbiota and *Porphyromonas gingivalis* Kgp Genotypes Altered in Parkinson's Disease with Mild Cognitive Impairment. *Mol. Neurobiol.* **61**, 8631–8639 (2024).
 61. Poole, S., Singhrao, S. K., Kesavalu, L., Curtis, M. A. & Crean, S. J. Determining the Presence of Periodontopathic Virulence Factors in Short-Term Postmortem Alzheimer's Disease Brain Tissue. *J. Alzheimer's Dis.* **36**, 665–677 (2013).
 62. Dominy, S. S. *et al.* *Porphyromonas gingivalis* in Alzheimer's disease brains: Evidence for disease causation and treatment with small-molecule inhibitors. *Sci. Adv.* **5**, (2019).
 63. GAIN Trial: Phase 2/3 Study of COR388 in Subjects With Alzheimer's Disease, NCT03823404, (2023). <https://clinicaltrials.gov/study/NCT03823404> (accessed 2024-10-31)
 64. Sabbagh, M. N. & Decourt, B. COR388 (atuzaginstat): an investigational gingipain inhibitor for the treatment of Alzheimer disease. *Expert Opin. Investig. Drugs* **31**, 987–993 (2022).
 65. McKee, A. S. *et al.* Effect of hemin on the physiology and virulence of *Bacteroides gingivalis* W50. *Infect. Immun.* **52**, 349–355 (1986).
 66. Potempa, J., Pike, R. & Travis, J. Titration and Mapping of the Active Site of Cysteine Proteinases from *Porphyromonas gingivalis* (Gingipains) Using Peptidyl Chloromethanes. *Biol. Chem.* **378**, 223–230, (1997).
 67. Plaza, K. *et al.* Gingipains of *Porphyromonas gingivalis* Affect the Stability and Function of Serine Protease Inhibitor of Kazal-type 6 (SPINK6), a Tissue Inhibitor of Human Kallikreins. *J. Biol. Chem.* **291**, 18753–18764 (2016).

68. de Diego, I. *et al.* Structure and Mechanism of Cysteine Peptidase Gingipain K (Kgp), a Major Virulence Factor of *Porphyromonas gingivalis* in Periodontitis. *J. Biol. Chem.* **289**, 32291–32302 (2014).
69. Kožíšek, M. *et al.* DNA-linked inhibitor antibody assay (DIANA) as a new method for screening influenza neuraminidase inhibitors. *Biochem. J.* **475**, 3847–3860 (2018).
70. Navrátil, V. *et al.* DNA-linked Inhibitor Antibody Assay (DIANA) for sensitive and selective enzyme detection and inhibitor screening. *Nucleic Acids Res.* **45**, e10 (2017).
71. Eichinger, A. Crystal structure of gingipain R: an Arg-specific bacterial cysteine proteinase with a caspase-like fold. *EMBO J.* **18**, 5453–5462 (1999).
72. Białas, A. *et al.* Exploring the *Sn* Binding Pockets in Gingipains by Newly Developed Inhibitors: Structure-Based Design, Chemistry, and Activity. *J. Med. Chem.* **49**, 1744–1753 (2006).
73. Ekici, Ö. D. *et al.* Aza-Peptide Michael Acceptors: A New Class of Inhibitors Specific for Caspases and Other Clan CD Cysteine Proteases. *J. Med. Chem.* **47**, 1889–1892 (2004).
74. Curtis, M. A. *et al.* Attenuation of the Virulence of *Porphyromonas gingivalis* by Using a Specific Synthetic Kgp Protease Inhibitor. *Infect. Immun.* **70**, 6968–6975 (2002).
75. Kadowaki, T. *et al.* Isolation and characterization of a novel and potent inhibitor of Arg-gingipain from *Streptomyces* sp. strain FA-70. *Biol. Chem.* **384**, 911–920 (2003).
76. Guevara, T. *et al.* Structural determinants of inhibition of *Porphyromonas gingivalis* gingipain K by KYT-36, a potent, selective, and bioavailable peptidase inhibitor. *Sci. Rep.* **9**, 1–8 (2019).
77. Kataoka, S. *et al.* A novel, potent dual inhibitor of Arg-gingipains and Lys-gingipain as a promising agent for periodontal disease therapy. *The FASEB J.* **28**, 3564–3578 (2014).
78. Arrowsmith, C. H. *et al.* The promise and peril of chemical probes. *Nat. Chem. Biol.* **11**, 536–541 (2015).
79. Li, R., Medchem, /, Licciardello, M. P. & Workman, P. The era of high-quality chemical probes. *RSC Med. Chem.* **13**, 1446–1459 (2022).
80. Jo, J. *et al.* Development of Oxadiazolone Activity-Based Probes Targeting FphE for Specific Detection of *Staphylococcus aureus* Infections. *J. Am. Chem. Soc.* **146**, 6880–6892 (2024).

81. Kanaumi, E., Konno, J., Miyazaki, M., Maeda, H. JP2010270061A, (2010)
82. Banfi, L., Basso, A., Lambruschini, C., Moni, L. & Riva, R. The 100 facets of the Passerini reaction. *Chem. Sci.* **12**, 15445–15472 (2021).
83. Reza Kazemizadeh, A. & Ramazani, A. Synthetic Applications of Passerini Reaction. *Curr. Org. Chem.* **16**, 418–450 (2012).
84. Dömling, A. & Ugi, I. Multikomponentenreaktionen mit Isocyaniden. *Angew. Chem.* **112**, 3300–3344 (2000).
85. Maeda, S., Komagawa, S., Uchiyama, M. & Morokuma, K. Finding Reaction Pathways for Multicomponent Reactions: The Passerini Reaction is a Four-Component Reaction. *Angew. Chem. Int. Ed.* **50**, 644–649 (2011).
86. Ramozzi, R. & Morokuma, K. Revisiting the Passerini Reaction Mechanism: Existence of the Nitrilium, Organocatalysis of Its Formation, and Solvent Effect. *J. Org. Chem.* **80**, 5652–5657 (2015).
87. Šimková, A. *et al.* Structure-activity relationship and biochemical evaluation of novel fibroblast activation protein and prolyl endopeptidase inhibitors with α -ketoamide warheads. *Eur. J. Med. Chem.* **224**, 113717 (2021).
88. Tichá, A. *et al.* General and Modular Strategy for Designing Potent, Selective, and Pharmacologically Compliant Inhibitors of Rhomboid Proteases. *Cell. Chem. Biol.* **24**, 1523-1536.e4 (2017).
89. Zhang, L. *et al.* α -Ketoamides as Broad-Spectrum Inhibitors of Coronavirus and Enterovirus Replication: Structure-Based Design, Synthesis, and Activity Assessment. *J. Med. Chem.* **63**, 4562–4578 (2020).
90. Huang, Q. *et al.* Discovery of α -Ketoamide inhibitors of SARS-CoV-2 main protease derived from quaternized P1 groups. *Bioorg. Chem.* **143**, 107001 (2024).
91. Barth, C. W. & Gibbs, S. Fluorescence image-guided surgery: a perspective on contrast agent development. in *Molecular-Guided Surgery: Molecules, Devices, and Applications VI* (eds. Gibbs, S. L., Pogue, B. W. & Gioux, S.) 18 (SPIE, 2020)
92. Garland, M., Yim, J. J. & Bogoy, M. A Bright Future for Precision Medicine: Advances in Fluorescent Chemical Probe Design and Their Clinical Application. *Cell Chem. Biol.* **23**, 122–136 (2016).
93. Seah, D., Cheng, Z. & Vendrell, M. Fluorescent Probes for Imaging in Humans: Where Are We Now? *ACS Nano* **17**, 19478–19490 (2023)
94. Olson, M. T., Ly, Q. P. & Mohs, A. M. Fluorescence Guidance in Surgical Oncology: Challenges, Opportunities, and Translation. *Mol. Imaging Biol.* **21**, 200–218 (2018).

95. Hong, G., Antaris, A. L. & Dai, H. Near-infrared fluorophores for biomedical imaging. *Nat. Biomed. Eng.* **1**, 1–22 (2017).
96. Hoogstins, C. E. S. *et al.* A Novel Tumor-Specific Agent for Intraoperative Near-Infrared Fluorescence Imaging: A Translational Study in Healthy Volunteers and Patients with Ovarian Cancer. *Clin. Cancer Res.* **22**, 2929–2938 (2016).
97. Buchs, N. C. *et al.* Intra-operative fluorescent cholangiography using indocyanin green during robotic single site cholecystectomy. *Int. J. Med. Robot.* **8**, 436–440 (2012).
98. Tummers, Q. R. J. G. *et al.* The Value of Intraoperative Near-Infrared Fluorescence Imaging Based on Enhanced Permeability and Retention of Indocyanine Green: Feasibility and False-Positives in Ovarian Cancer. *PLoS One* **10**, e0129766 (2015).
99. Yoshida, M. *et al.* Indocyanine green injection for detecting sentinel nodes using color fluorescence camera in the laparoscopy-assisted gastrectomy. *J. Gastroenterol. Hepatol.* **27**, 29–33 (2012).
100. Peltrini, R. *et al.* Intraoperative use of indocyanine green fluorescence imaging in rectal cancer surgery: The state of the art. *World J. Gastroenterol.* **27**, 6374 (2021).
101. Seitzkazina, A., Yang, J.-K. & Kim, S. Clinical effectiveness and prospects of methylene blue: A systematic review. *Precis. Future Med.* **6**, 193–208 (2022).
102. Maeda, H. Macromolecular therapeutics in cancer treatment: The EPR effect and beyond. *J. Control. Release* **164**, 138–144 (2012).
103. Fang, J., Nakamura, H. & Maeda, H. The EPR effect: Unique features of tumor blood vessels for drug delivery, factors involved, and limitations and augmentation of the effect. *Adv. Drug Deliv. Rev.* **63**, 136–151 (2011).
104. Hadjipanayis, C. G., Widhalm, G. & Stummer, W. What is the surgical benefit of utilizing 5-aminolevulinic acid for fluorescence-guided surgery of malignant gliomas? *Neurosurgery* **77**, 663–673 (2015).
105. Stummer, W. *et al.* Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial. *Lancet Oncol.* **7**, 392–401 (2006).
106. Tanyi, J. L. *et al.* A Phase III Study of Pafolacianine Injection (OTL38) for Intraoperative Imaging of Folate Receptor-Positive Ovarian Cancer (Study 006). *J. Clin. Oncol.* **41**, 276–284 (2023).
107. Investigation of Novel Surgical Imaging for Tumor Excision (INSITE), NCT03686215, (2023). <https://clinicaltrials.gov/study/NCT03686215> (accessed 2024-03-29).

108. Phase 2 Study of VGT-309 in Lung Cancer, NCT05400226, (2023). <https://classic.clinicaltrials.gov/ct2/show/NCT05400226> (accessed 2024-03-29).
109. Whitley, M. J. *et al.* A mouse-human phase 1 co-clinical trial of a protease-activated fluorescent probe for imaging cancer. *Sci. Transl. Med.* **8**, 320ra4 (2016).
110. Kennedy, G. T. *et al.* A Cathepsin-Targeted Quenched Activity-Based Probe Facilitates Enhanced Detection of Human Tumors during Resection. *Clin. Cancer Res.* **28**, 3729–3741 (2022).
111. Miampamba, M. *et al.* Sensitive *in vivo* Visualization of Breast Cancer Using Ratiometric Protease-activatable Fluorescent Imaging Agent, AVB-620. *Theranostics* **7**, 3369–3386 (2017).
112. Weissleder, R., Tung, C.-H., Mahmood, U. & Bogdanov, A. In vivo imaging of tumors with protease-activated near-infrared fluorescent probes. *Nat. Biotechnol.* **17**, 375–378 (1999).
113. Jiang, T. *et al.* Tumor imaging by means of proteolytic activation of cell-penetrating peptides. *Proc. Natl. Ac. Sci.* **101**, 17867–17872 (2004).
114. Faucher, F. F. *et al.* Protease Activated Probes for Real-Time Ratiometric Imaging of Solid Tumors. *ACS Cent. Sci.* **9**, 1059–1069 (2023).
115. Šácha, P. *et al.* iBodies: Modular Synthetic Antibody Mimetics Based on Hydrophilic Polymers Decorated with Functional Moieties. *Angew. Chem. Int. Ed.* **55**, 2356–2360 (2016).
116. Haag, R. & Kratz, F. Polymer Therapeutics: Concepts and Applications. *Angew. Chem. Int. Ed.* **45**, 1198–1215 (2006).
117. Kopeček, J. & Kopečková, P. HPMA copolymers: Origins, early developments, present, and future. *Adv. Drug Deliv. Rev.* **62**, 122–149 (2010).
118. Pospíšilová, K. *et al.* Inhibitor–Polymer Conjugates as a Versatile Tool for Detection and Visualization of Cancer-Associated Carbonic Anhydrase Isoforms. *ACS Omega* **4**, 6746–6756 (2019).
119. Blažková, K. *et al.* The development of a high-affinity conformation-sensitive antibody mimetic using a biocompatible copolymer carrier (iBody). *J. Biol. Chem.* **297**, 101342 (2021).
120. Šimon, P. *et al.* Identification of Protein Targets of Bioactive Small Molecules Using Randomly Photomodified Probes. *ACS Chem. Biol.* **13**, 3333–3342 (2018).
121. Dvořáková, P. *et al.* Inhibitor-Decorated Polymer Conjugates Targeting Fibroblast Activation Protein. *J. Med. Chem.* **60**, 8385–8393 (2017).

122. Šubr, V. *et al.* The role of the biotin linker in polymer antibody mimetics, iBodies, in biochemical assays. *Polym. Chem.* **12**, 6009–6021 (2021).
123. Beranová, J. *et al.* Tris-(Nitrilotriacetic Acid)-Decorated Polymer Conjugates as Tools for Immobilization and Visualization of His-Tagged Proteins. *Catalysts* **9**, 1011 (2019).
124. Zamani, M. R. *et al.* Polymer-based antibody mimetics (iBodies) target human PD-L1 and function as a potent immune checkpoint blocker. *J. Biol. Chem.* **300**, 107325 (2024).
125. Ofori, L. O. *et al.* Design of Protease Activated Optical Contrast Agents That Exploit a Latent Lysosomotropic Effect for Use in Fluorescence-Guided Surgery. *ACS Chem. Biol.* **10**, 1977–1988 (2015).
126. Verdoes, M. *et al.* Improved Quenched Fluorescent Probe for Imaging of Cysteine Cathepsin Activity. *J. Am. Chem. Soc.* **135**, 14726–14730 (2013).
127. Poreba, M. *et al.* Fluorescent probes towards selective cathepsin B detection and visualization in cancer cells and patient samples. *Chem. Sci.* **10**, 8461–8477 (2019).
128. Tholen, M. *et al.* Design of optical imaging probes by screening of diverse substrate libraries directly in disease tissue extracts. *Angew. Chem. Int. Ed. Engl.* **59**, 19143 (2020).
129. Yim, J. J., Tholen, M., Klaassen, A., Sorger, J. & Boggyo, M. Optimization of a Protease Activated Probe for Optical Surgical Navigation. *Mol. Pharm.* **15**, 750–758 (2018).
130. Yim, J. J. *et al.* A protease-activated, near-infrared fluorescent probe for early endoscopic detection of premalignant gastrointestinal lesions. *Proc. Natl. Ac. of Sci.* **118**, (2021).
131. Jouin, P., Castro, B. & Nisato, D. Stereospecific synthesis of N-protected statine and its analogues via chiral tetramic acid. *J. Chem. Soc. Perkin. 1* 1177–1182 (1987) doi:10.1039/P19870001177.
132. Smrčina, M., Majer, P., Majerová, E., Guerassina, T. A. & Eissenstat, M. A. Facile stereoselective synthesis of γ -substituted γ -amino acids from the corresponding α -amino acids. *Tetrahedron* **53**, 12867–12874 (1997).
133. Gistrová, T. Preparation, synthesis and activity of new inhibitors of cysteine proteases from *Porphyromonas gingivalis*, bachelor's thesis (2021)
134. Kisseljova, K., Tšubrik, O., Sillard, R., Mäeorg, S. & Mäeorg, U. Addition of Arylboronic Acids to Symmetrical and Unsymmetrical Azo Compounds. *Org. Lett.* **8**, 43–45 (2006).

135. Uemura, T. & Chatani, N. Copper Salt Catalyzed Addition of Arylboronic Acids to Azodicarboxylates. *J. Org. Chem.* **70**, 8631–8634 (2005).
136. De Graaff, C. *et al.* Asymmetric Synthesis of Tetracyclic Pyrroloindolines and Constrained Tryptamines by a Switchable Cascade Reaction. *Angew. Chem. Int. Ed.* **54**, 14133–14136 (2015).
137. Garcia, P., Lau, Y. Y., Perry, M. R. & Schafer, L. L. Phosphoramidate Tantalum Complexes for Room-Temperature C–H Functionalization: Hydroaminoalkylation Catalysis. *Angew. Chem. Int. Ed.* **52**, 9144–9148 (2013).
138. Laufer, R. *et al.* Discovery of inhibitors of the mitotic kinase TTK based on N-(3-(3-sulfamoylphenyl)-1H-indazol-5-yl)-acetamides and carboxamides. *Bioorg. Med. Chem.* **22**, 4968–4997 (2014).
139. Blackburn, C. *et al.* Identification and characterization of amino-piperidinequinolones and quinazolinones as MCHR1 antagonists. *Bioorg. Med. Chem. Lett.* **16**, 2621–2627 (2006).
140. Tofteng, A. P., Sørensen, K. K., Conde-Frieboes, K. W., Hoeg-Jensen, T. & Jensen, K. J. Fmoc Solid-Phase Synthesis of C-Terminal Peptide Thioesters by Formation of a Backbone Pyroglutamyl Imide Moiety. *Angew. Chem. Int. Ed.* **48**, 7411–7414 (2009).
141. Harmat, N. J. S. *et al.* 1,2-Disubstituted cyclohexane derived tripeptide aldehydes as novel selective thrombin inhibitors. *Bioorg. Med. Chem. Lett.* **8**, 1249–1254 (1998).
142. Shuman, R. T. *et al.* Structure-Activity Study of Tripeptide Thrombin Inhibitors Using α -Alkyl Amino Acids and Other Conformationally Constrained Amino Acid Substitutions. *J. Med. Chem.* **38**, 4446–4453 (1995).
143. Shuman, R. T. *et al.* Highly selective tripeptide thrombin inhibitors. *J. Med. Chem.* **36**, 314–319 (1993).
144. Buyck, T., Wang, Q. & Zhu, J. Triple Role of Phenylselenonyl Group Enabled a One-Pot Synthesis of 1,3-Oxazinan-2-ones From α -Isocyanoacetates, Phenyl Vinyl Selenones, and Water. *J. Am. Chem. Soc.* **136**, 11524–11528 (2014).
145. Odriozola, A., Oiarbide, M. & Palomo, C. Enantioselective Synthesis of Quaternary Δ^4 - and Δ^5 -Dehydroprolines Based on a Two-Step Formal [3+2] Cycloaddition of α -Aryl and α -Alkyl Isocyano(thio)acetates with Vinyl Ketones. *Chem. Eur. J.* **23**, 12758–12762 (2017).

146. Liang, C., Behnam, M. A. M., Sundermann, T. R. & Klein, C. D. Phenylglycine racemization in Fmoc-based solid-phase peptide synthesis: Stereochemical stability is achieved by choice of reaction conditions. *Tetrahedron Lett.* **58**, 2325–2329 (2017).
147. Li, W., O'Brien-Simpson, N. M., Hossain, M. A. & Wade, J. D. The 9-Fluorenylmethoxycarbonyl (Fmoc) Group in Chemical Peptide Synthesis – Its Past, Present, and Future. *Aust. J. Chem.* **73**, 271–276 (2019).
148. Hadzima, M. *et al.* Polymer-Tethered Quenched Fluorescent Probes for Enhanced Imaging of Tumor-Associated Proteases. *ACS Sens.* **9**, 3720–3729 (2024).