

Development of Tissue Kallikrein Inhibitors for the Treatment of Netherton Syndrome

Thèse N° 7676

Présentée le 19 juillet 2019

à la Faculté des sciences de base

Laboratoire de protéines et peptides thérapeutiques

Programme doctoral en chimie et génie chimique

pour l'obtention du grade de Docteur ès Sciences

par

Patrick Andreas GONSCHOREK

Acceptée sur proposition du jury

Dr R. Hovius, président du jury

Prof. C. Heinis, directeur de thèse

Prof. K. Strømgaard, rapporteur

Prof. C. Klein, rapporteur

Prof. H. Lashuel, rapporteur

2019

*“Look deep into nature,
and then you will understand everything better.”*

— Albert Einstein

Acknowledgements

I would like to sincerely thank my thesis supervisor Prof. Christian Heinis for the great opportunity to work on this exciting project. Your vision on translating research into something that could make a change in the future was truly motivating and inspiring. I am grateful for all the precious scientific advices you gave me that will further guide me in my research.

I would like to express my gratitude to my exam jury members Prof. Kristian Strømgaard, Prof. Christian Klein and Prof. Hilal Lashuel for taking time and effort to evaluate my work and to Dr. Ruud Hovius for chairing the committee.

Furthermore, I want to acknowledge all scientist at EPFL who helped to obtain results for my thesis: Dr. David Hacker for expressing proteins, Dr. Jessica Sordet-Dessimoz for her support with the histology, Laroche Thierry for helping me with the microscopy and Dr. Luc Reymond for providing silicon rhodamine. I also greatly thank Dr. Oliver Gouin and Prof. Alain Hovnanian from the institute Imagine in Paris for conducting efficacy studies to show that our molecules really work.

I would like to thank my students Chantal Whitefield, Manon Sciboz, Mathilde Le Jeune, Mischa Schüttel, Denis Vollmar, Rebeca Gómez Ojea and Mathilde Montagnon for their great enthusiasm and valuable contributions to this project.

Huge thanks also go to Béatrice Bliesener-Tong and Anne Lene Odegaard for your kindness and great administrative support.

I wish to thank all LPPT lab members for your help and advice, the warm atmosphere in the lab, and the great time we spent together. Thank you Yuteng, Joao, Carl, Camille, Jonas, Davide, Simon, Kaycie, Christina, Sangram, Xudong, Milan, Manh, Ganesh, Sevan, Gontran, Manuel, Mischa, Edouard, Rakesh, Cristina and Jun. Especially, I want to thank Ale, for his great scientific guidance on this project, Vanessa for all her valuable advice and the Velox adventure, Tamara for handling mice, and Kaycie for taking time to proofread my thesis.

I want to thank my friends Jacob, Jacob, Stefan und Stani for the good time and for making it to Lausanne.

I want to thank Tamara for her unconditional support and sticking with me in this difficult time.

Last but not least, I would like to express my gratitude and love to my Mom who supports me every day in everything I do and my Dad, who always believed in me. Additionally, thanks to my not-so-little-anymore brother Pascal.

None of this would have been possible without you.

Patrick

Abstract

There is currently no approved treatment for Netherton syndrome, which is a severe, genetic skin disease characterized by chronic inflammation and excessive peeling of the skin. Even though the molecular disease mechanism has been assigned to the unopposed activity of the two tissue kallikrein-related peptidases 5 and 7 (KLK5 and KLK7) and has been validated in mouse models, no therapies specifically targeting these two proteases exist or are currently in clinical development. To address this, we have used phage display of cyclic peptides to develop highly potent and selective macrocyclic inhibitors of KLK5 ($K_i = 1$ nM) and KLK7 ($K_i = 7$ nM). To equip these inhibitors with suitable pharmacokinetic properties for systemic *in vivo* application, we applied a half-life extension strategy to the peptides. This improved the plasma stability of the inhibitors from a half-life of less than 5 min to more than 90 hours *in vitro* and prolonged the elimination half-life to 7 hours *in vivo*. The inhibitors diffused surprisingly well into the skin and reached high concentrations in the epidermis. Efficacy studies in mouse models of Netherton syndrome are currently being performed by a collaborator and look promising. Thus, the molecules have the potential to be further development toward the first targeted therapy for Netherton syndrome.

Keywords

Cyclic peptides, protease inhibitors, phage display, half-life extension, pharmacokinetics, therapeutics, Netherton syndrome, skin, tissue kallikreins, KLK5, KLK7, *SPINK5*, LEKT1.

Zusammenfassung

Es gibt derzeit keine zugelassene Behandlungsmöglichkeit für Netherton-Syndrom, eine schwere, genetisch bedingte Hautkrankheit, die durch chronische Entzündung und extremes Abschälen der Haut gekennzeichnet ist. Obwohl der molekulare Mechanismus der unkontrollierten Aktivität der beiden Proteasen Kallikrein 5 und Kallikrein 7 (KLK5 und KLK7) zugeordnet werden konnte und bereits in Mausmodellen validiert wurde, ist derzeit keine Therapie in Entwicklung. Wir haben mit Hilfe von Phagen-Display zyklische Peptide entwickelt, die hocheffektiv und selektiv KLK5 ($K_i = 1 \text{ nM}$) und KLK7 ($K_i = 7 \text{ nM}$) inhibieren. Um die pharmakokinetischen Eigenschaften der Inhibitoren für systemische Anwendung zu optimieren, haben wir eine Strategie zur Verlängerung der Halbwertszeit verwendet. Dies verbesserte die Plasmahalbwertszeit der Inhibitoren von weniger als 5 min auf mehr als 90 Stunden *in vitro* und auf mehr als 7 Stunden *in vivo*. Die Inhibitoren diffundierten überraschend effizient in die Haut und erreichten hohe Konzentrationen in der Epidermis. Wirksamkeitsstudien in Mausmodellen des Netherton-Syndroms werden derzeit von einem Kollaborateur durchgeführt und sehen vielversprechend aus. Die Moleküle haben daher Potenzial zum ersten Medikament für Netherton-Syndrom weiterentwickelt zu werden.

Schlüsselwörter

Zyklische Peptide, Proteaseinhibitoren, Phagen-Display, Verlängerung der Halbwertszeit, Pharmakokinetik, Medikament, Netherton-Syndrom, Haut, Kallikreine, KLK5, KLK7, *SPINK5*, LEKTI.

Contents

Acknowledgements	I
Abstract	III
Zusammenfassung.....	V
Abbreviations.....	XI
Amino acids.....	XIII
1 Introduction	1
1.1 Netherton syndrome	1
1.1.1 Disease	1
1.1.2 Molecular background	2
1.1.3 Tissue kallikreins.....	5
1.1.4 LEKTI	10
1.1.5 KLK5 and KLK7 inhibitors	11
1.1.6 Mouse models	13
1.2 Peptides in drug development	15
1.2.1 Peptides as drug modality	15
1.2.2 Directed evolution in drug development	18
1.2.3 <i>In vitro</i> selection of peptides	21
1.2.4 Phage display of bicyclic peptides	24
1.2.5 Improvement of pharmacological properties	27
2 Aim of this work	33
3 Results & Discussion	35
3.1 Phage selection of bicyclic peptide inhibitors of KLK5 and KLK7	35
3.1.1 Choice of target proteins.....	35
3.1.2 Expression strategies and constructs	36
3.1.3 Expression, purification and characterization	37

3.1.4	Biotinylation of KLKs for immobilization	40
3.1.5	Selection & characterization of bicyclic peptides	41
3.2	Structure activity relationship and affinity improvement	47
3.2.1	Development of monocyclic peptide inhibitors	47
3.2.2	Affinity improvement with unnatural amino acids	49
3.2.3	Selectivity.....	52
3.3	Half-life extension and pharmacokinetics	53
3.3.1	Conjugation to albumin tag	53
3.3.2	Plasma stability of inhibitors	58
3.3.3	Stability improvement of KLK7 inhibitor	58
3.3.4	Replacement of fluorescein in albumin tag.....	60
3.3.5	Final inhibitors	64
3.3.6	Pharmacokinetics in mice	65
3.4	Skin distribution	68
3.4.1	Biodistribution to skin and organs.....	68
3.4.2	Skin distribution of peptide versus antibody.....	70
3.5	Peptides for efficacy studies	76
3.5.1	Inhibition of mouse KLKs	76
3.5.2	Negative control inhibitors	76
3.5.3	Batch synthesis	77
3.5.4	Formulation	79
3.5.5	Efficacy studies	81
4	Conclusion & Outlook.....	83
4.1	Are the inhibitors better than the state-of-the-art?.....	83
4.1.1	Comparison of our inhibitors with LEKTI	83
4.1.2	Comparison of our KLK7 inhibitor to the cyclic depsipeptide	85
4.1.3	Comparison to engineered sunflower trypsin inhibitors.....	86
4.2	Are the inhibitors good enough for therapy?	88
4.2.1	Affinity	88
4.2.2	Specificity and toxicity	89
4.2.3	Stability and pharmacokinetics.....	89
4.2.4	Biodistribution	90
4.2.5	Efficacy.....	90
4.3	Outlook	92

4.3.1	Inhibitors for the systemic treatment of Netherton syndrome	92
4.3.2	Inhibitors for topical administration	94
5	Materials & Methods.....	95
5.1	Materials	95
5.1.1	Peptide synthesis reagents.....	95
5.1.2	Peptide synthesis building blocks.....	96
5.1.3	Biochemical reagents	97
5.1.4	Proteins & substrates	98
5.1.5	Kits & consumables	99
5.2	Protein expression	100
5.2.1	Expression and purification of KLK5 and KLK7	100
5.2.2	Biotinylation of targets.....	101
5.3	Phage display selections	102
5.3.1	Phage display library	102
5.3.2	Phage production and modification.....	102
5.3.3	Biopanning.....	103
5.3.4	Phage sequencing.....	104
5.4	Peptide synthesis	105
5.4.1	Automated peptide synthesis	105
5.4.2	Fluorescein and fatty acid conjugation on solid phase	105
5.4.3	Peptide cleavage and deprotection	105
5.4.4	Palmitic acid and silicon rhodamine coupling in solution	106
5.4.5	Cyclization with thiol-reactive linkers	106
5.4.6	Synthesis of para-nitroanilide substrates.....	106
5.4.7	Peptide purification.....	107
5.5	Labeling of biomolecules	108
5.6	<i>In vitro</i> characterization.....	109
5.6.1	Analysis of peptide mass and purity.....	109
5.6.2	Protease inhibition assays	109
5.6.3	Fluorescence polarization albumin binding assay.....	111
5.6.4	Plasma stability assay	111
5.7	Pharmacokinetic studies in mice	113
5.7.1	Peptide concentrations in blood	113
5.7.2	Fluorescence imaging of organs.....	113

5.7.3	Preparation and fluorescence microscopy of skin sections	114
6	Supporting information	115
6.1	Phylogeny of KLKs	116
6.2	Structure of human serum albumin.....	117
6.3	Synthesis schemes	118
6.4	Raw data	120
6.4.1	Affinity improvement	120
6.4.2	Stability of inhibitors.....	121
6.4.3	Stability improvement of KLK7 inhibitor	122
6.4.4	Stability in mouse plasma	123
6.4.5	Degradation fragments in mice	123
6.4.6	Biodistribution	124
6.5	Overview of binding constants	125
6.6	Peptide purities & masses.....	126
	References	157

Abbreviations

5(6)-FAM	5(6)-Carboxyfluorescein
Ab	antibody
Abs	absorption
Ac	acetyl
AMC	7-amino-4-methylcoumarin
AU	arbitrary unit
BM40	membrane protein 40 secretory signal peptide
Boc	<i>tert</i> -butyloxycarbonyl
BSA	bovine serum albumin
CDR	complementarity-determining regions
CDSN	corneodesmosin
CTRL	control
D1, D2, ...	domain 1, domain 2, ...
Dde	N-(1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl)
DOL	degree of labeling
DSC1	desmocollin 1
DSG1	desmoglein 1
EK	enterokinase
EKcs	enterokinase cleavage site
Em	emission
equiv	equivalent
ESI	electrospray ionization
Ex	excitation
Fab	antigen-binding fragment
Fc	crystallizable fragment
FcRn	neonatal Fc receptor
FDA	Food and Drug Administration of the US
Fmoc	fluorenylmethyloxycarbonyl
FP	fluorescence polarization
F-tag	albumin-binding ligand with N-terminal fluorescein ¹¹⁶
GLP-1	glucagon-like peptide-1
HArg	homoarginine
His₆	hexahistidine-tag
hKLK	human kallikrein-related peptidase
HPLC	high-performance liquid chromatography
HSA	human serum albumin
IC₅₀	half maximal inhibitory concentration
IgG	immunoglobulin gamma
IMAC	nickel-charged immobilized metal affinity chromatography
IP	intraperitoneal
IV	intravenous
K_a	association constant
K_d	dissociation constant
K_i	inhibitory constant
KLK	tissue kallikrein-related peptidase
K_m	Michaelis-Menten constant

LC-MS	liquid chromatography-mass spectrometry
LEKTI	lympho-epithelial Kazal-type-related inhibitor
LG	lamellar granule
mIgG	mouse immunoglobulin gamma
mKLK	mouse tissue kallikrein-related peptidase
MSA	mouse serum albumin
MWCO	molecular weight cut-off
NHS	N-hydroxysuccinimide
Nle	norleucine
NS	Netherton syndrome
P1, P2, ...	non-prime side peptide residue 1, 2, ...
P1', P2', ...	primed side peptide residue 1, 2, ...
PAR2	protease activated receptor 2
Pbf	2,2,4,6,7-Pentamethyl-2,3-dihydro-1-benzofuran-5-sulfonyl
PEG	polyethylene glycol
pK_a	the negative of the base-10 logarithm of the acid dissociation constant (K _a)
pNA	para-nitroanilide
RFU	relative fluorescence unit
ROI	region of interest
RP-HPLC	reversed-phase HPLC
RSA	rat serum albumin
RT	room temperature
S1, S2, ...	non-primed side subsite 1, 2, ...
S1', S2', ...	primed side subsite 1, 2, ...
SAR	structure-activity relationship
SC	subcutaneous / stratum corneum
scFv	single-chain variable fragment
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SiR	6-carboxy-silicon rhodamine
SPINK5	serine protease inhibitor, Kazal-type 5
SPPS	solid-phase peptide synthesis
SSSKSSS-tag	albumin-binding ligand with only Ser residues in the heptapeptide sequence
t_{1/2}	half-life
t_{1/2β}	elimination half-life
tag	albumin-binding ligand from ¹¹⁶
tag-F	albumin-binding ligand with C-terminal fluorescein ¹¹⁶
tBu	<i>tert</i> -Butyl
TCEP	tris-(2-carboxyethyl)-phosphin
Tg-hKLK5	transgenic mouse model expressing hKLK5 in the granular layer of the epidermis
TIC	total ion current
TNF-α	tumor necrosis factor alpha
Trt	triphenylmethyl
U	enzyme unit
UK18	bicyclic peptide inhibitor of uPA ⁹⁷
uPA	urokinase-type plasminogen activator

Amino acids

Name	3-Letter code	1-Letter code	Residue weight (-H ₂ O)	Charge (pH 7.4)	pK _a	Hydrophobicity Index (pH 7)
Alanine	Ala	A	71.08	–	–	41
Arginine	Arg	R	156.19	positive	12.48	-14
Asparagine	Asn	N	114.11	neutral	–	-28
Aspartic acid	Asp	D	115.09	negative	3.65	-55
Cysteine	Cys	C	103.15	neutral	8.18	49
Glutamic acid	Glu	E	129.12	negative	4.25	-31
Glutamine	Gln	Q	128.13	neutral	–	-10
Glycine	Gly	G	57.05	–	–	0
Histidine	His	H	137.14	10% positive 90% neutral	6	8
Isoleucine	Ile	I	113.16	–	–	99
Leucine	Leu	L	113.16	–	–	97
Lysine	Lys	K	128.18	positive	10.53	-23
Methionine	Met	M	131.2	–	–	74
Phenylalanine	Phe	F	147.18	–	–	100
Proline	Pro	P	97.12	–	–	-46
Serine	Ser	S	87.08	neutral	–	-5
Threonine	Thr	T	101.11	neutral	–	13
Tryptophan	Trp	W	186.22	–	–	97
Tyrosine	Tyr	Y	163.18	neutral	10.07	63
Valine	Val	V	99.13	–	–	76

The pK_a value is the negative of the base-10 logarithm of the proton dissociation constant for the side chain of the amino acid.

The hydrophobicity index is a measure of the solubility in water, normalized to Phe (= 100) and relative to Gly (= 0), color coded as following: **very hydrophobic**, **hydrophobic**, neutral, **hydrophilic**.

Data from <https://www.sigmaaldrich.com/life-science/metabolomics/learning-center/amino-acid-reference-chart.html> were used.

1 Introduction

1.1 Netherton syndrome

1.1.1 Disease

Netherton syndrome (NS) is a severe, genetic skin disease that occurs with an autosomal recessive inheritance in 1 : 200,000 newborns and is classified as rare disease¹. The disease is characterized by severe skin inflammation and excessive skin scaling (desquamation) (**Figure 1**). Further complications that arise from the defect skin barrier are a high susceptibility for infections and a high risk for dehydration. Both can be life-threatening for newborns and infants. Even though the condition often partially improves with age and patients have a normal life expectancy, the quality of life is severely impaired by persistent itchiness and severe skin inflammation³.

NS is caused by loss of function mutations in the *SPINK5* (serine protease inhibitor, Kazal-type 5) gene, which encodes the protease inhibitor LEKTI (lympho-epithelial Kazal-type-related inhibitor). It was first described by Comel⁴ in 1949 and Netherton⁵ in 1958, and the association with mutations in the *SPINK5* gene was discovered by Chavanas *et al.*⁶ in 2000. The *SPINK5* gene codes for the multi-domain serine protease inhibitor LEKTI, which is expressed by cells of the stratum granulosum and secreted into the extracellular space of the stratum corneum – the two outermost layers of the skin. LEKTI is the most important regulator of a proteolytic cascade of kallikrein-related peptidases (KLKs), initiated by KLK5 and mediating skin renewal⁷. Insufficient regulation of this proteolytic cascade by LEKTI leads to the uncontrolled activity of two proteolytic pathways causing the two major manifestations of NS: Triggering of protease-activated receptor 2 (PAR2) by KLK5 and KLK14 leads to skin inflammation, while cleavage of cell-cell connecting desmosomes by KLK7 causes desquamation⁸.

More than 70 mutations of the *SPINK5* gene have been reported to date, mostly ascribable to small defects, like insertions, deletions, or substitutions, leading to premature termination codons. Some reported mutations also impair splice sites within the *SPINK5* gene³. The underlying mutation and the level of residual LEKTI expression determines the severity of the disease⁹.

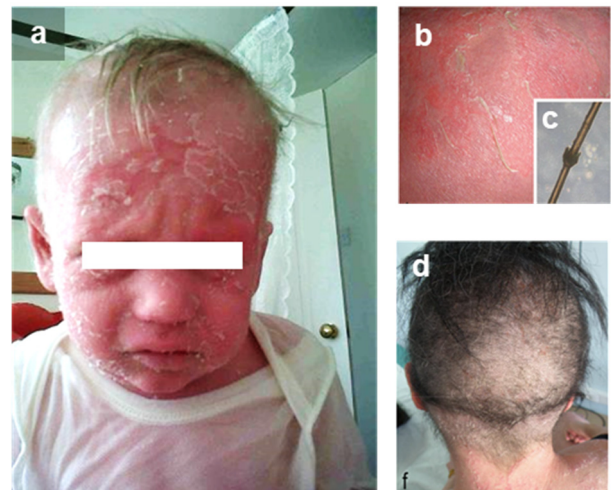


Figure 1 | Clinical manifestation of NS. (a) Severe skin inflammation, scaling (desquamation) and absence of eyebrows in a 21-month-old. **(b)** Inflammation with superficial desquamation. **(c)** Abnormal hair shaft (bamboo hair). **(d)** Short and reduced hair. Picture (a) used by courtesy of ². Pictures (b-d) adapted with permission from ³, reprinting license nr. 4570200976429, © 2013 Springer

Currently, there is no specific treatment available for NS. Treatment approaches that are being tested include intravenous immunoglobulin (IVIg) therapy¹⁰ or anti-TNF- α treatment¹¹. However, neither provides specific treatment, which would ideally restore or replace the LEKTI function, and thus these therapeutic options only provide temporary solutions until a specific therapy has been developed.

For this reason, a specific therapy for the targeted treatment of NS is a great unmet medical need. An approach suggested in literature involves the delivery of recombinantly expressed LEKTI fragments¹². However, no development to equip these fragments with the required pharmacokinetic properties for a systemic delivery to the skin has yet been reported. Nevertheless, the development of synthetic inhibitors of tissue kallikreins in the form of LEKTI fragments or smaller molecules alongside the delivery of the natural inhibitor via gene therapy are the most promising approaches towards a specific therapy for Netherton syndrome¹³. A proof-of-concept phase 1 gene therapy study using transduced skin crafts of 20 cm² has been conducted, but no positive results have yet been reported¹⁴. As the development of a gene therapy for the reconstitution of LEKTI expression in the entire skin will most likely be challenging and take a long time, the current state of the art dictates that a kallikrein-targeted therapy with small molecules, peptides, or antibodies provides the most practicable and realistic option.

1.1.2 Molecular background

The upper part of the human skin can be primarily divided into two main layers: the epidermis and the dermis. The dermis contains nerves, glands, follicles, and blood vessels necessary for supplying the epidermis, whereas the epidermis represents the continuously renewing part of the skin, which has mainly protective function. It serves as water barrier to prevent dehydration and protects the body from chemicals and pathogens from the outside. It contains no nerves or blood vessel as it is

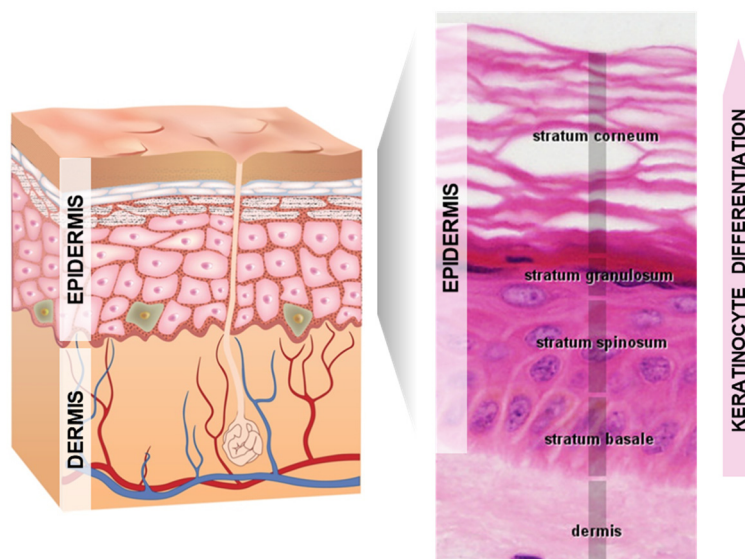


Figure 2 | The skin structure. The primary layers of the skin are the dermis and the epidermis. The epidermis is continuously renewing via keratinocyte differentiation with the sublayers representing the consecutive differentiation states of the keratinocytes. Human skin structure (left) licensed from ¹⁵. Histology picture (right) used by courtesy of L. Slomianka¹⁶.

completely renewed over a period of 48 days in a process called keratinocyte differentiation. The sublayers of the epidermis represent the consecutive differentiation states of keratinocytes during their upwards journey in the epidermis. This journey starts in the lowest layer, the stratum basale, which represents a single layer of undifferentiated epidermal stem cells. During their differentiation process, the cells move upwards to the stratum granulosum, where they flatten and accumulate intracellular keratin. When the keratinocytes finally degrade their nuclei and their cytoplasmic organelles and die off, they have completed their differentiation program and form the stratum corneum. The dead keratinocytes – then called corneocytes – stay filled with keratin to retain water and remain tightly connected through corneodesmosomes. When the corneocytes reach the outermost layer of the stratum corneum, they are finally released via desquamation¹⁷ (**Figure 2**).

As the dead corneocytes of the stratum corneum can no longer react to biological signals, their final release is regulated extracellularly by a cascade of secreted proteases. To regulate the release of the outermost corneocytes in the stratum corneum, a cocktail of different tissue kallikreins is secreted by the keratinocytes of the stratum granulosum into the extracellular space of the stratum corneum. The proteases are secreted as inactive proenzymes (also called zymogens), which become active only after the pro-domain is cleaved by an already active protease. The cascade is initiated by KLK5, which can autoactivate and subsequently cleave the pro-domains off of pro-KLK7, pro-KLK8, and pro-KLK14 to activate them. Active KLK7 cleaves corneocyte-connecting proteins, such as desmoglein 1 (DSG1),

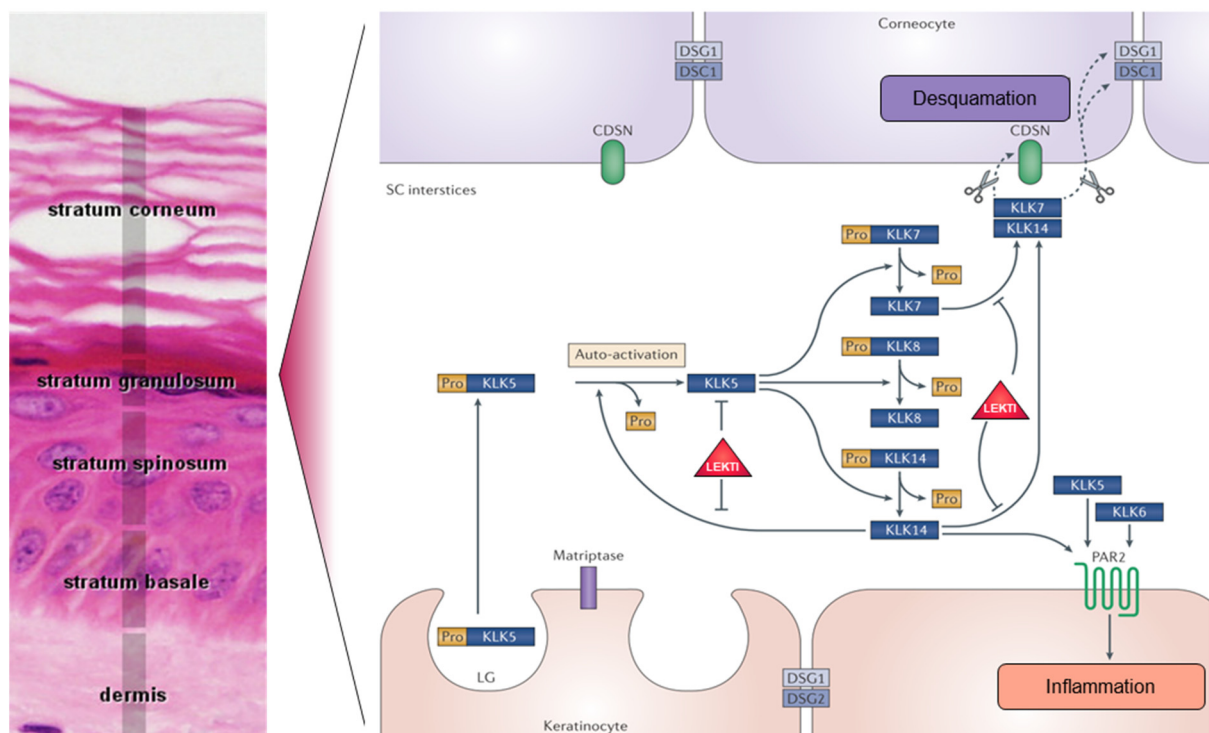


Figure 3 | The epidermal KLK cascade. KLKs are secreted by keratinocytes of the stratum granulosum as inactive proenzymes. KLK5 can autoactivate and subsequently activate KLK7, KLK8, and KLK14. KLK7 mediates desquamation by cleaving desmosomal proteins, like desmoglein 1 (DSG1), desmocollin 1 (DSC1), and corneodesmosin (CDSN). KLK14 can activate KLK5 in a positive feed-back loop and amplify the cascade. Furthermore, together with KLK5, KLK14 is capable of triggering inflammation via proteinase-activated receptor 2 (PAR2). Lamellar granule (LG), stratum corneum (SC), kallikrein (KLK), lympho-epithelial Kazal-type-related inhibitor (LEKTI). Histology picture (left) used by courtesy of L. Slomianka¹⁶. Proteolytic cascade (right) adapted with permission from ⁸, reprinting license nr. 4570210184816, © 2015 Springer Nature.

desmocollin 1 (DSC1), and corneodesmosin (CDSN), leading to desquamation. Via a second pathway, KLK14 cleaves and thereby triggers PAR2, which causes inflammation. Active KLK14 can activate KLK5 again, amplifying the activation of the cascade in a positive feed-back loop⁸ (**Figure 3**).

This entire proteolytic cascade is regulated by the multidomain protease inhibitor LEKTI, which is also produced by the keratinocytes of the stratum granulosum and co-secreted together with the kallikreins. LEKTI inhibits the initiator of the cascade, KLK5, as well as the two effector proteases, KLK7 and KLK14, downstream in the cascade. This tight regulation of the kallikrein cascade allows the detachment of cells in the superficial stratum corneum to ensure renewal of the epidermis, while it prevents destruction of cell-cell connections in the deep stratum corneum to keep the skin barrier intact. The underlying mechanism allowing layer-specific cascade regulation is the pH-dependent LEKTI inhibition of the kallikreins along a pH gradient in the stratum corneum. This pH gradient starts at pH 7.5 at the border of the stratum granulosum (deep stratum corneum), where the kallikreins and LEKTI are secreted, and drops to pH 4.5 at its outermost layers, where cells are released (superficial stratum corneum). LEKTI can inhibit its kallikrein targets in the low nanomolar-range at a neutral pH of 7.5 in the deep stratum corneum but loses its affinity at the acidic pH found towards the superficial layers of the stratum corneum. This pH-dependent inhibition mechanism allows the layer-specific regulation of the epidermal KLK cascade in the absence of any biological signals⁷ (**Figure 4**).

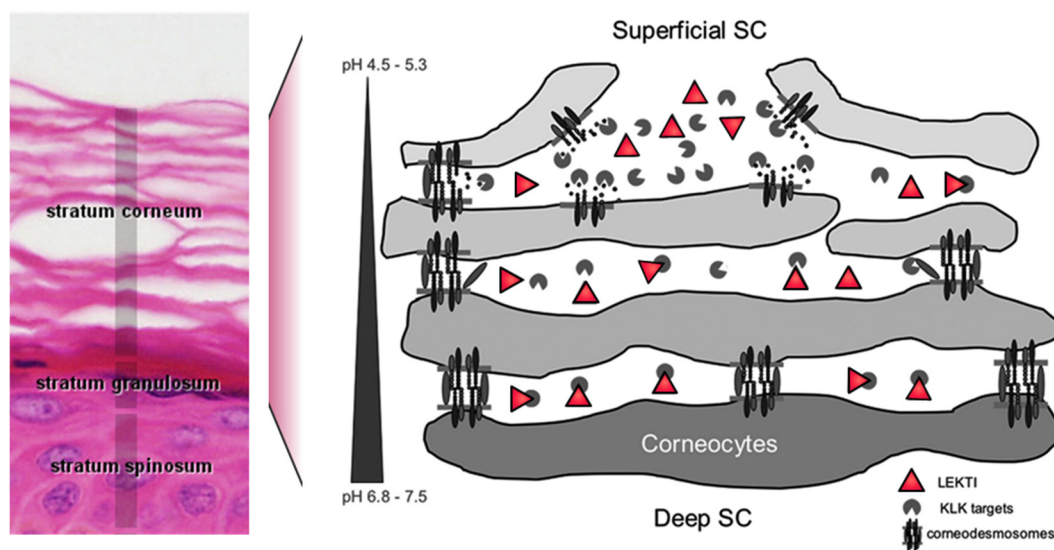


Figure 4 | pH-dependent inhibition of the epidermal KLK cascade by LEKTI. LEKTI inhibits its KLK targets at high pH at the border of the stratum granulosum but loses its affinity at the lower pH found towards the outer layers of the stratum corneum (SC).

Histology picture (left) used by courtesy of L. Slomianka¹⁶. Model of desquamation (right) adapted with permission from ⁷, reprinting license nr. 4570211180188, © 2007 ASCB

1.1.3 Tissue kallikreins

According to their catalytic mechanism, tissue kallikreins belong to the family of serine proteases to which they constitute the homologous gene family of kallikreins – a family next to the thrombins, trypsins, chymotrypsins, elastases, and matriptases¹⁸. The name derives from the Greek word ‘kallikreas’, which means pancreas, referring to the organ where the first kallikrein was discovered¹⁹. Based on their differing location in the body, the kallikrein family can be further divided into plasma kallikrein (KLK1B), which is found in the blood plasma, and the 15 tissue kallikreins (KLK1–KLK15), which can be found exclusively in specific types of tissues²⁰ (**Figure 5a**). In contrast to the plasma kallikrein gene, which is located on chromosome 4, all the tissue kallikreins genes are found on one locus on chromosome 19, which represents the largest protease gene cluster in the human genome²¹ (**Figure 5b**). This makes the tissue kallikrein family an interesting subject for the study of gene duplication events and evolutionary analysis. Phylogenetic reconstruction of all tissue kallikreins homologues revealed that the closest homologue to KLK5 is KLK4 – a product of KLK5 duplication – followed by KLK14. For KLK7, KLK10 and KLK12 are the phylogenetically closest proteases²² (**Figure S1**).

Most of the KLKs have been described to be involved in proteolytic pathways that regulate highly tissue-specific functions: For example KLK4 is involved in the formation of the enamel layer of teeth; KLK1 has several functions in the airway, renal, and cardiovascular system and might be a target for various airway diseases; KLK6 and 8 are involved in synaptic neural plasticity; KLK3 regulates semen liquefaction together with KLK2, 4, and 5 and has been linked to prostate cancer; and KLK5, 7, 8, and 14 are involved in skin desquamation and were linked to Netherton syndrome⁸ (**Figure 5c**).

The activation of KLKs is regulated on several levels through different mechanisms. Through transcription and translation of the gene in the cell an inactive pre-pro-protein of the protease is synthesized. The pre-domain serves as recognition signal for translocation into the endoplasmic reticulum, where it is cleaved off allowing the protease to be secreted as a pro-protein (also called a zymogen). The pro-domain keeps the zymogen from adopting an active conformation and must also be cleaved off by an already active protease⁸. The sequence of the pro-domain determines which proteases can remove it to activate the protease. If the sequence of the pro-domain fits the protease’s own substrate specificity, it also can autoactivate²³. Once mature and active, a protease can be regulated either by endogenous inhibitors, ions, or through further inactivating cleavage events²⁴ (**Figure 6a**).

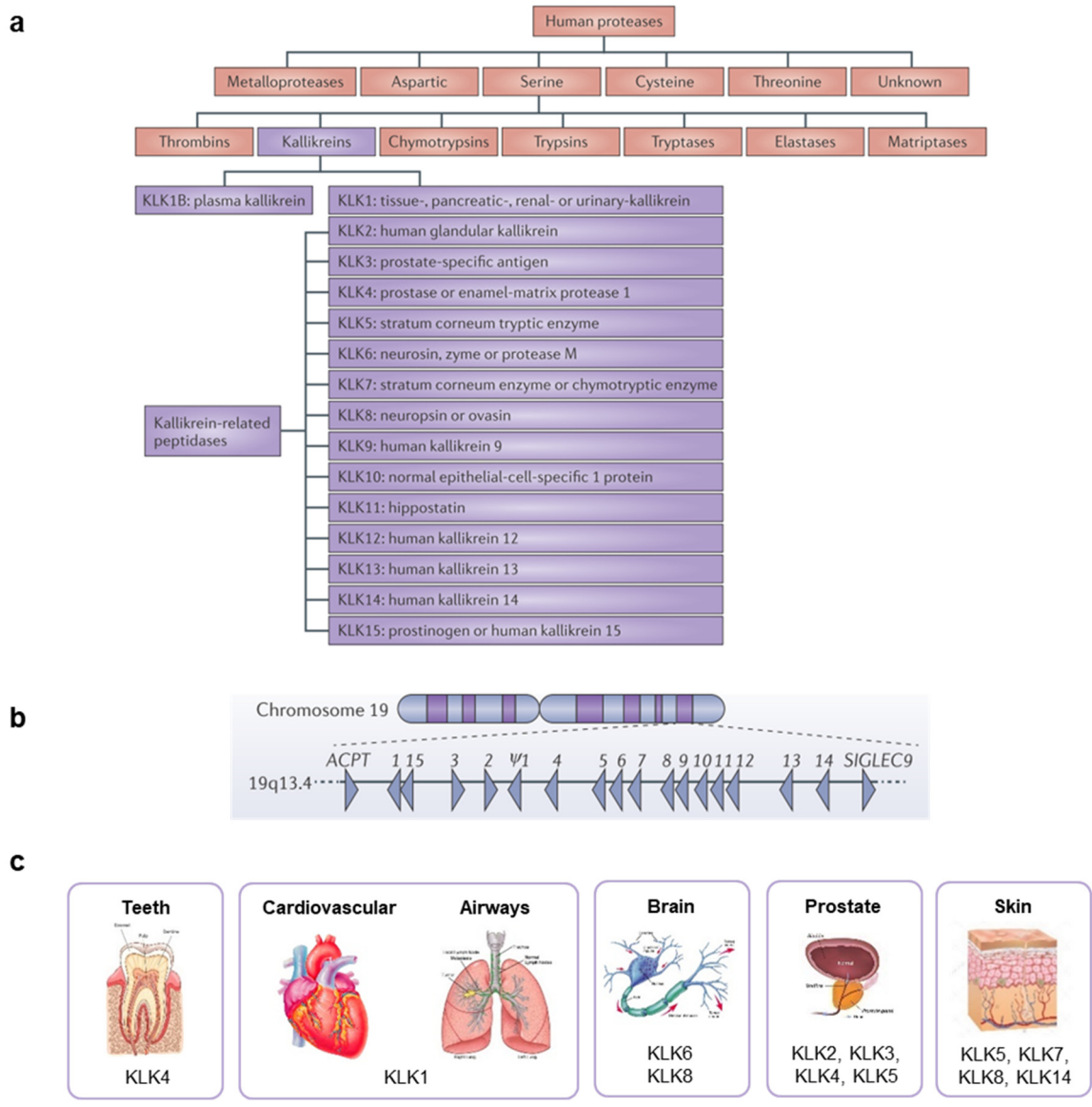


Figure 5 | Tissue kallikreins. (a) Position of KLKs in the human protease family tree. **(b)** Overview of KLK genomic locus on chromosome 19. **(c)** Examples of the functions of several KLKs in the human body. (a) and (b) adapted with permission from ⁸, reprinting license nr. 4570210184816, © 2015 Springer Nature.

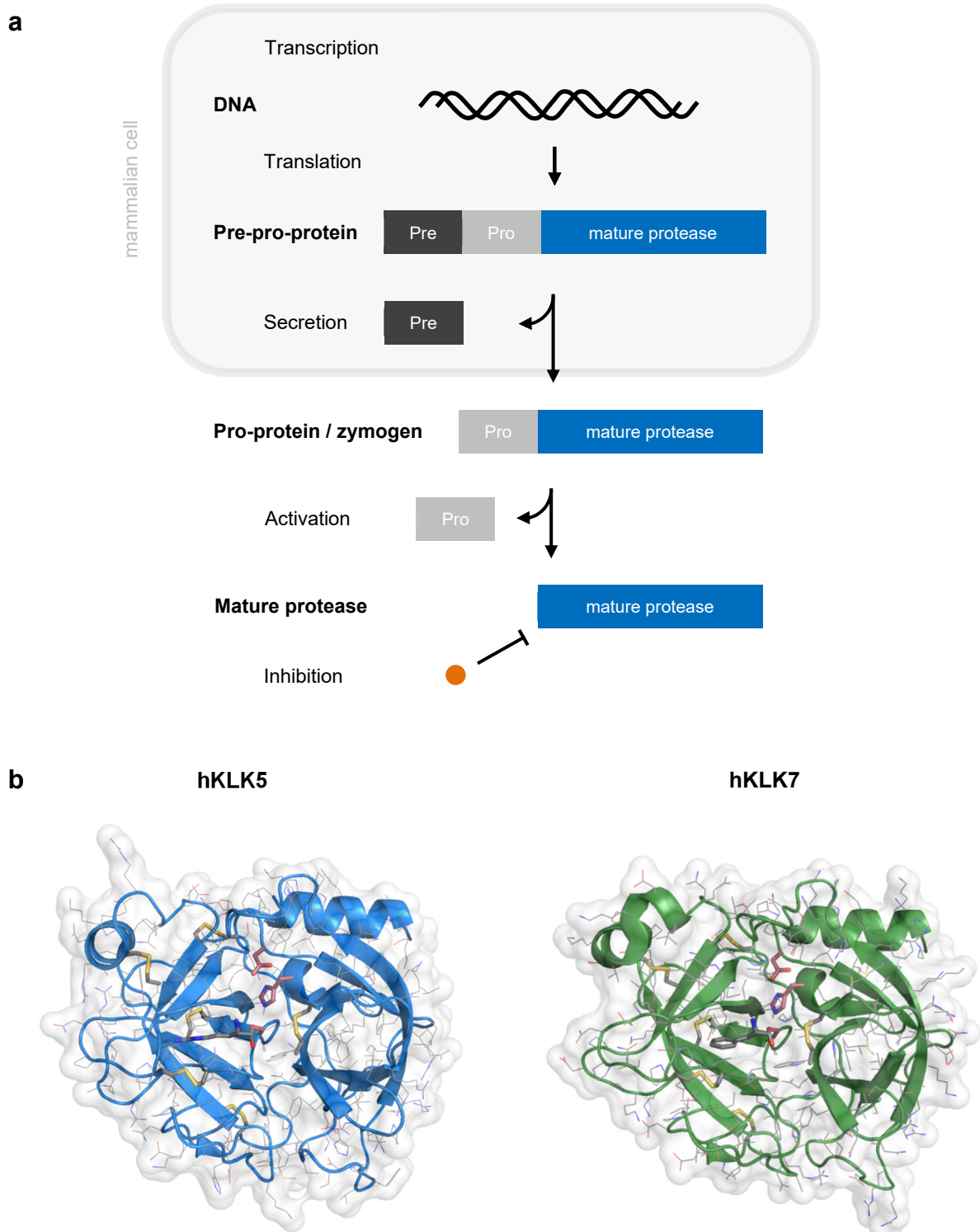


Figure 6 | Expression and structure of KLKs. (a) Tissue kallikreins are synthesized as pre-pro-proteins in the cytoplasm. The pre-domain is cleaved off after translocation into the ER and the pro-protein (also called zymogen) is secreted. Upon cleavage of the pro-domain by an active protease, the mature protease forms when the N-terminus folds in and stabilizes the active site. The activity of the mature protease is regulated by endogenous inhibitors or ions, and it can be inactivated by further cleavage events. **(b)** Structures of hKLK5 (2PSX) and hKLK7 (2QXG) consisting of two six-stranded β barrels. The six disulfide-bonds are shown in yellow. P1 residues of complexed inhibitors (Leupeptin (KLK5) and Ala-Ala-Phe-chloromethylketone (KLK7)) are shown in gray. The side chains of catalytic triads (Ser195, His57, and Asp102) are shown in red.

The catalytic triade consisting of His57, Asp102, and Ser195 is conserved throughout all kallikreins. Furthermore, all kallikreins need to be cleaved to liberate a hydrophobic residue, which inserts into the core of the protease domain to form the oxyanion whole of the active, mature protease²⁵. In the tissue kallikrein family, this cleavage event releases a short pro-domain, which is not part of the mature protease, whereas in plasma kallikrein the N-terminal part of the protein stays connected after cleavage via a disulfide bond, forming a two-subunit protein. Regarding this structural feature, plasma kallikreins are more similar to the coagulation factors, like factor XI (FXI), than to the tissue kallikreins²⁰.

Similar to other serine proteases, the basic structure of KLK5 and KLK7 consists of two six-stranded β barrels, fixed by three trans-domain strands, and two α helices. The catalytic triade is located along the cleft between the two barrels, whereas the substrate recognition sites run vertically to this cleft²⁶ (**Figure 6b**).

Several techniques to characterize the substrate preferences of the 15 kallikreins revealed that most have a trypsin-like specificity, meaning that they cleave exclusively after positivity charged residues, like arginine or lysine. Exceptions are KLK3 and KLK7, which show clear chymotrypsin-like specificity, cleaving large aromatic residues, preferably tyrosine. Some kallikreins like KLK1, KLK10, KLK11, and KLK14 also show dual specificity²⁰.

In general, the protease surface of a pocket, which binds one single side-chain of a substrate, is called subsite. The subsites that bind residues upwards of the scissile bond of the substrate towards its N-terminus are termed non-primed sites and numbered S1-Sn, whereas the subsites in the opposite direction of the scissile bond towards the C-terminus are called primed sites, numbered S1'-Sn'. The corresponding residues of the substrate, which bind the pockets, are labeled P1-Pn and P1'-Pn', respectively²⁷ (**Figure 7a**).

To get an idea for which amino acids might be able to bind the subsites around the catalytic site of the two kallikreins, previous works screening the subsite preferences of hKLK5 and hKLK7 were analyzed. In 2006, Debela *et al.* performed a specificity profiling of seven human tissue kallikreins using a combinatorial library of tetrapeptidic substrates to perform a positional scanning for the non-primed S4-S1 sites²⁸. The results for KLK5 and KLK7 and a consensus logo based on this data are shown in **Figure 7b**. A general observation valid for most proteases is that with an increasing distance to the scissile bond, specificity decreases and a broader range of residues is tolerated. Comparing KLK5 and KLK7, KLK5 can be classified as more specific as only a single residue is tolerated in the S1 site. As a trypsin-like protease, it has a strong preference for arginine at the P1 position, and it prefers serine, threonine, or asparagine (small, polar amino acids) at the P2 position²⁸.

In contrast, KLK7 accepts about four different residues in the S1 site, and the S4 site does not seem selective. As a chymotrypsin-like protease, it cleaves after tyrosine in the P1 position, though also cleaves after a methionine or alanine. At the P2 position, it is less selective than KLK5. Residues that preferentially bind in the S2 site are tyrosine or lysine, so medium-sized, polar, or hydrophobic residues²⁸ (**Figure 7c**).

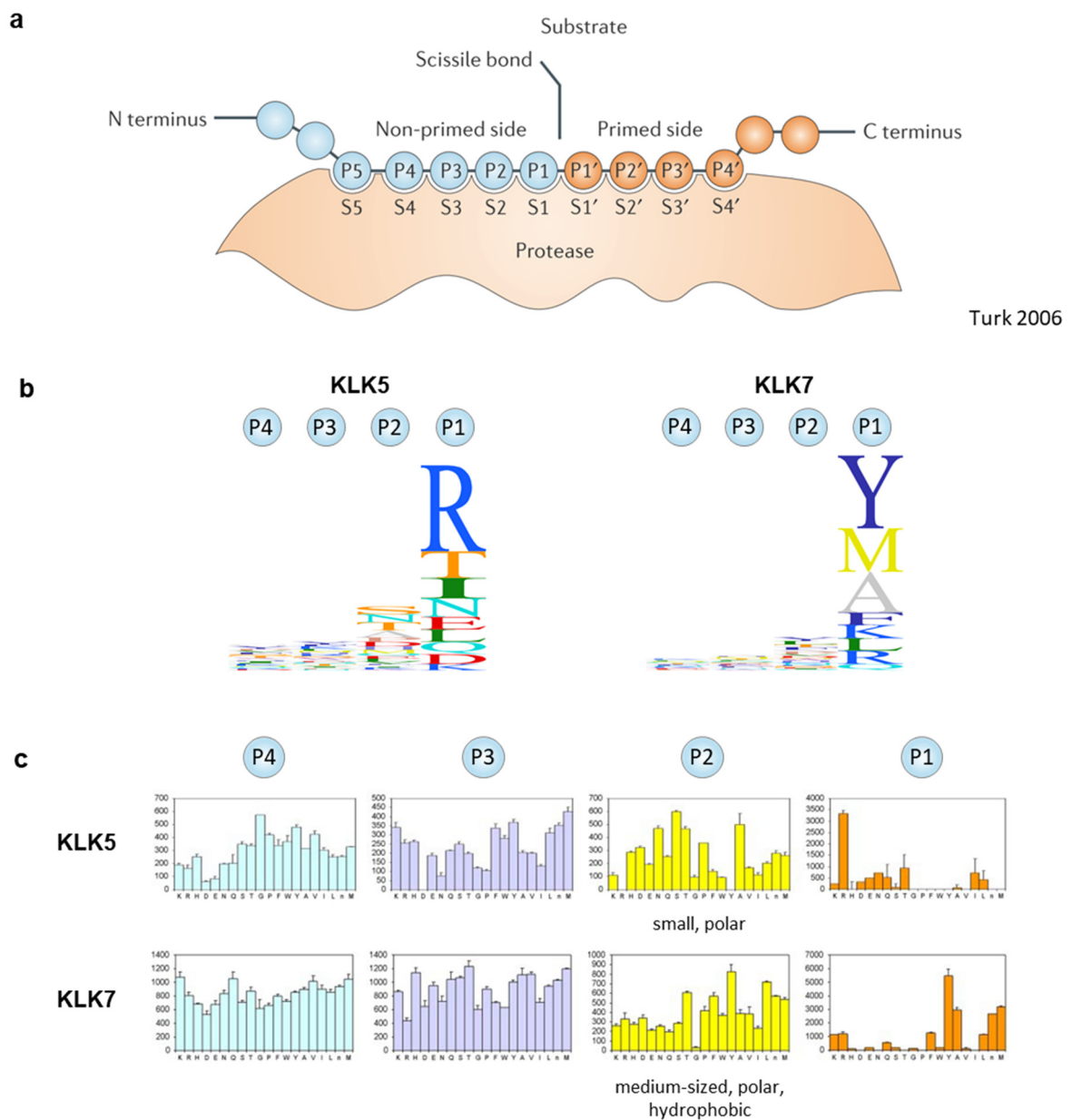


Figure 7 | Specificity profiles for the non-primed side of KLK5 and KLK7. (a) The subsites upwards of the scissile bond on the non-primed side are numbered S1-Sn, whereas the subsites downwards of the scissile bond on the primed side are numbered S1'-Sn'. **(b)** Consensus logo for substrate specificity based on data from ²⁸, which are shown in (c). **(c)** Substrate specificity of non-primed sites of KLK5 and KLK7 represented as cleavage activity of tetrapeptidic substrates for different amino acids at P4-P1 positions.

(a) Reprinted with permission from ²⁷, reprinting license nr. 4570220121179, © 2006 Springer Nature.

(c) Reprinted with permission from ²⁸, reprinting license nr. 4570220887387, © 2006 ASBMB.

1.1.4 LEKTI

The LEKTI protein, which is encoded by the *SPINK5* gene, is intracellularly synthesized as a multidomain protein consisting of 15 similar, Kazal-type inhibitor domains²⁹. Kazal-type inhibitors belong to the standard mechanism inhibitors, which inhibit their protease targets by binding in a substrate-like fashion. These inhibitors are quite rigid and form a very stable enzyme-inhibitor complex due to strong main-chain–main-chain interactions. Even though the families of standard mechanism inhibitors are not homologous, they bind their target proteases with a common main chain conformation. The peptide bond in the inhibitory loop can be hydrolyzed after the P1 residue, though only at an extremely slow rate. The fragments are held together by disulfide bridges so that the hydrolyzed bond can be resynthesized upon complex formation with the enzyme³⁰.

Each of the 15 Kazal-type inhibitory domains of LEKTI has a size of around 7 kDa, and the domains are connected via linkers that contain furin cleavage sites. Some, but not all, of the linkers are cleaved in the endoplasmic reticulum by the protease furin so that LEKTI fragments containing between one to six inhibitory domains are secreted. The final secreted fragments are: D1-D5, the single domain fragments D6 and D7, the dual domain fragment D8-D9, and the D10-D15 fragment³¹ (**Figure 8a**).

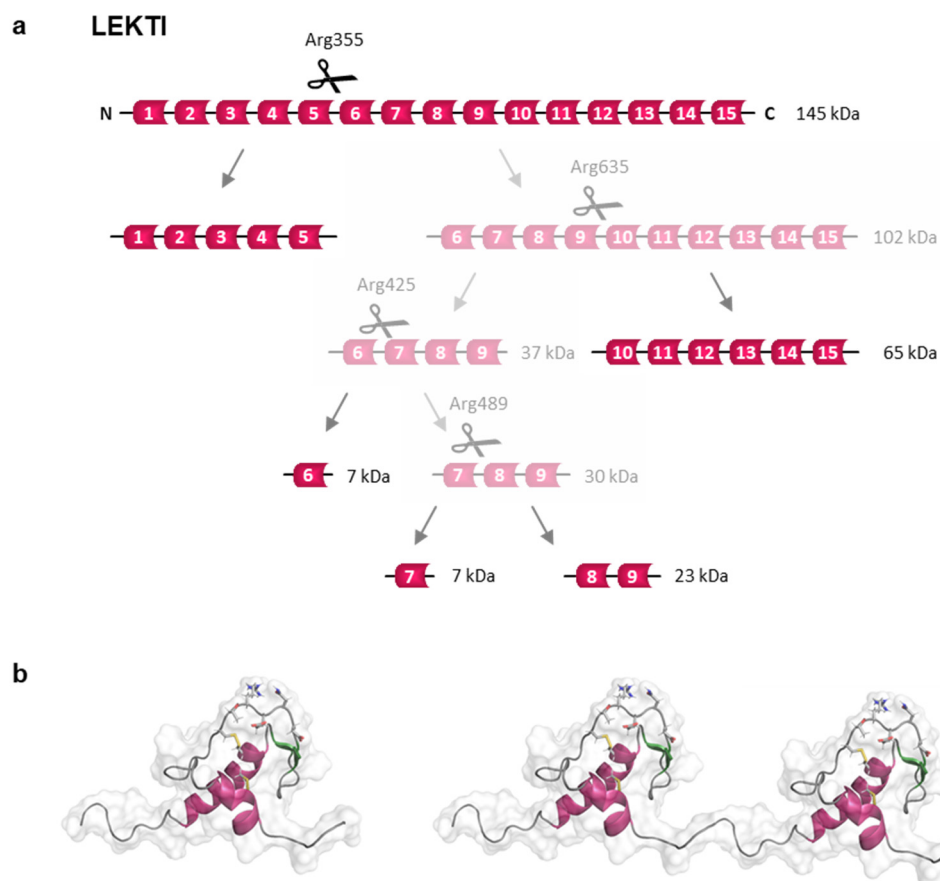


Figure 8 | LEKTI processing and structure of inhibitor domains. (a) LEKTI is synthesized as a multi-domain inhibitor with 15 Kazal-type domains connected by linkers containing a furin cleavage site. The multi-domain protein is processed by furin in the ER, and smaller fragments containing a single or up to six inhibitory domains are secreted. Based on ³¹. **(b)** The structure of a single inhibitory domain (1H0Z) is constrained by two alpha helices, which are connected through a disulfide-bond (red). The inhibitory loop (side chains of P3-P3' residues are shown as sticks) is stabilized via a second disulfide bond on one side and an antiparallel beta sheet on the other side (green).

KLK5, KLK7, and KLK14 have been identified as the major targets of LEKTI fragments. However, all fragments have a different inhibitory profile towards those three epidermal kallikreins and other proteases. Furthermore, their inhibitory profiles are not only determined by the number of domains in the fragment, but also by the primary sequences of the domains⁷.

All the 15 domains contain at least two cysteines at conserved positions, which form two disulfide bridges. Only D2 and D15 contain an additional third disulfide bridge³². The solution structures of D1, D6, and D15 are published^{32, 33}. The basic scaffold for constraining the flexibility of the inhibitory loop of LEKTI D6 consists of two alpha helices, which are interconnected via a disulfide bond. Both ends of the inhibitory loop are stabilized by this scaffold: the left end is linked to the alpha helix via the second disulfide bond in the protein, while the right end is stabilized via an antiparallel beta sheet, which additionally serves as interaction surface with the protease³² (**Figure 8b**).

The most conserved part among all 15 LEKTI domains is the inhibitory loop containing the sequence 'CTR END', with arginine in the P1 position, 'CTR' representing the non-primed side residues, and 'END' representing the primed side residues (**Figure 9a**). With an arginine at the P1 position³², this loop is a good inhibitor of trypsin-like proteases, probably of KLK5 and KLK14 (**Figure 9b**). Also KLK7 has been shown to be inhibited by LEKTI fragments, though to a weaker extent, and the binding mechanism for KLK7 as a chymotrypsin-like protease is not clear, even though some models have been suggested³⁴.

The affinity of the different LEKTI fragments for the three KLKs has been tested by two groups, who have determined different affinities for KLK5, KLK7, and KLK14 in the two publications^{7, 31}. They showed first that D1 to D4 have no relevance for the inhibition of these three kallikreins. The strongest inhibition of trypsin-like activity derived from fragments from the region D6 to D9. The single domain fragments D6 and D7 have the highest affinity for KLK14, with around 1 nM. The best inhibitor for KLK5 is represented by the dual-domain fragment D8-D9, which is in the 10 nM range. The strongest inhibitor of the chymotrypsin-like activity of KLK7 is represented by the last and biggest LEKTI fragment D10-D15, which has an affinity in the 20 nM range^{7, 31}. However, the affinities for KLK7 are inconsistent between the two publications.

1.1.5 KLK5 and KLK7 inhibitors

Some synthetic inhibitor of KLK5 and KLK7 have already been published and it was suggested in literature, the natural inhibitor LEKTI fragments could be engineered for therapeutic application¹². The best reported LEKTI fragments have K_i values of 5.4 nM and 20 nM for KLK5 and KLK7, respectively³¹. Furthermore, several peptide inhibitors based on engineered sunflower trypsin inhibitor (SFTI-1) have been published. Those are an unspecific KLK5/7/14 inhibitor³⁵, but also a KLK5 inhibitor³⁶ (5.1 nM), a KLK7 inhibitor³⁷ (0.14 nM), and KLK14 inhibitor³⁸ (7.0 nM), which were reported to be selective. For KLK7, a cyclic depsipeptide with an IC_{50} value of 0.2 nM is described in a patent as potent inhibitor for topical administration³⁹. Efforts to develop small molecule inhibitors of KLK5 did not result in potent and specific molecules^{40, 41}. Successful approaches using antibodies have not been reported yet and none of the reported molecules was developed or tested for systemic administration.

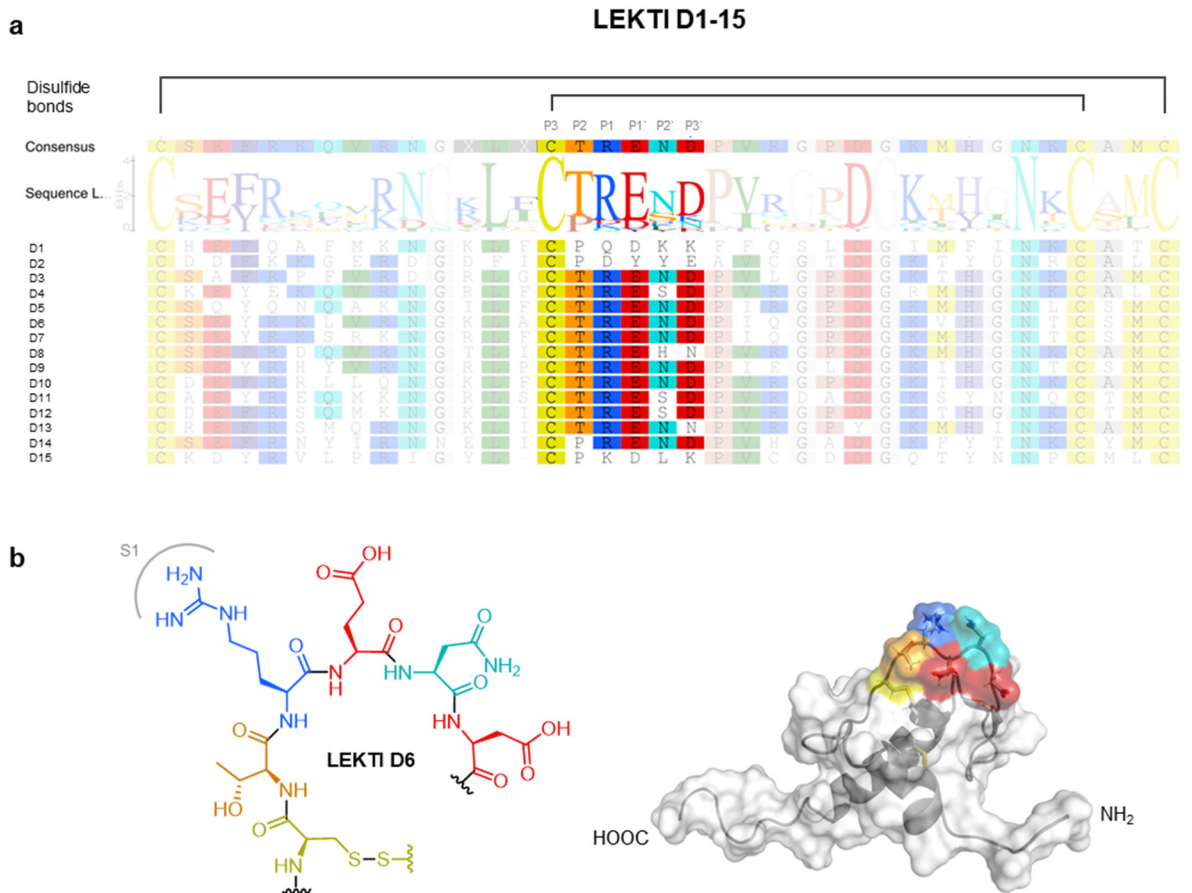


Figure 9 | Sequences and structure of the LEKTI inhibitory loop. (a) Sequence alignment of LEKTI D1-D15 showing a consensus sequence logo, the two disulfide bridges present in all 15 domains, and the P3-P3' residues, serving as the inhibitory loop for trypsin-like protease inhibition. **(b)** Chemical structure of P3-P3' residues of LEKTI D6 and NMR solution structure of LEKTI D6 (1H0Z) with P3-P3' residues shown using the Rasmol color code.

1.1.6 Mouse models

Several mouse models to study the underlying mechanisms of NS have been generated and are described in the following.

The first mouse model of NS was the *Spink5* knock-out model (**Figure 10, *Spink5*^{-/-}**), which reproduces the human genotype in mice. This model was first generated by Descargues *et al.* in 2005 by replacing the translation initiation codon of *Spink5* and was later reproduced by Kasperek *et al.* in 2017 by inducing a mutation that is also found in NS patients^{42, 43}. This mouse model reproduces the key features of human NS with premature desquamation through deep SC detachment and severe skin barrier defects. However, pups die within the first 12 hours after birth due to rapid dehydration, which is also attributed to the fact that mice have a much higher surface-to-volume ratio than humans.

After the main targets of LEKTI were identified, several rescue models were generated to confirm their role *in vivo*. On top of the *Spink5*^{-/-} genetic background, *Klk5*, *Klk7*, or both *Klks* were additionally knocked out to determine whether this would rescue the lethal phenotype. First, a *KLK7* defect alone did not change the survival time, and *Spink5*^{-/-} *KLK7*^{-/-} pups still died within 12 hours after birth⁴³ (**Figure 10, *Spink5*^{-/-} *KLK7*^{-/-}**).

In contrast, *KLK5* ablation prolonged the survival time until post-natal day 5 (**Figure 10, *Spink5*^{-/-} *KLK5*^{-/-}**), though was not sufficient to entirely rescue the lethal phenotype of LEKTI-deficient mice. Even though these mice had less severe barrier defects, some disrupted patches were still present, and an abnormal epidermal differentiation caused by *KLK7* activity was observed. Particularly strong barrier defects could be detected adjacent to hair shafts around the nostrils and could be assigned to a defective separation of the hair shafts from the surrounding tissues⁴³.

Only simultaneous ablation of *KLK5* and *KLK7* completely rescued lethality in these mice, and the resulting mice survived to adulthood (**Figure 10, *Spink5*^{-/-} *KLK5*^{-/-} *KLK7*^{-/-}**). They showed a completely restored skin barrier, though with some growth retardation and alopecia 2–4 weeks after birth, which could potentially be associated with *KLK14* overactivity. This phenotype improved with age, however, and no cutaneous defects were detectable in adulthood. These mice had an entirely functional skin and normal hair growth. Thus, this rescue study suggested that the unregulated activity of both proteases is responsible for the development of NS⁴³.

In addition to the knockout and rescue models, the development of viable mouse models of Netherton syndrome was attempted, which would mimic the phenotype and allow for mechanistic studies over longer time periods or the testing of therapeutic compounds. The first of these models is the *Tg-hKLK5* model, which expresses human *KLK5* in the granular layer of the epidermis⁴⁴ (**Figure 10c, *Tg-hKLK5***). Despite the presence of functional LEKTI, this additional *KLK5* activity overrides the inhibitory capacities of LEKTI and overactivates the epidermal proteolytic cascade, causing strong inflammation, scaling, growth retardation, and hair abnormalities. Interestingly, increased *KLK7*, *KLK14*, and *ELA2* activities were also observed in the skin of these mice, confirming that *KLK5* activates these proteases *in vivo*⁴⁴. Another viable mouse model is based on mosaic inactivation of the *Spink5* gene generated through the microinjection of TALENs into zygotes. These mice showed skin lesions across 20–40% of their body surface, which would be particularly interesting for topical testing of compounds, as it provides control skin on the same animal⁴⁵.

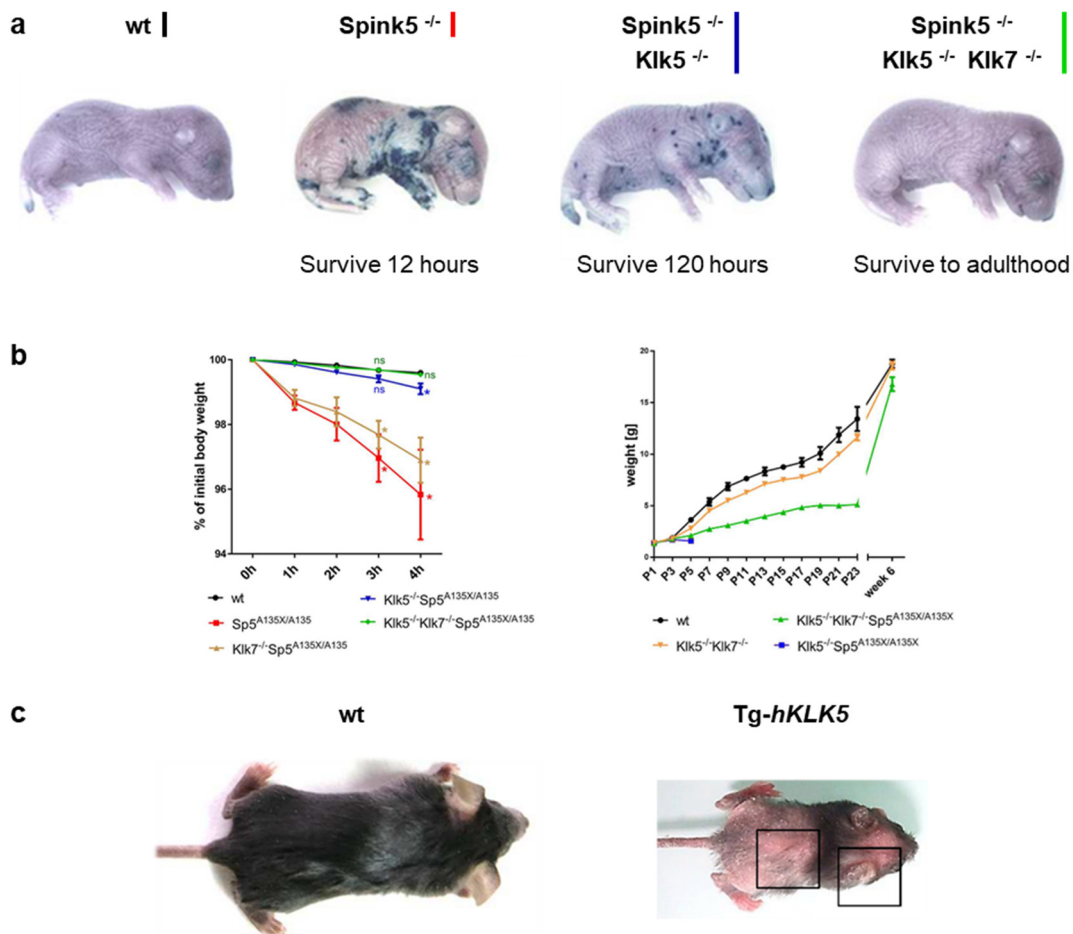


Figure 10 | Mouse models of Netherton syndrome. (a) Skin barrier of newborn pups analyzed via a penetration assay using toluidine blue. **(b)** Left: Relative body weight up to 4 hours after birth ($*p < 0.05$). Right: Body weight development up to 6 weeks after birth. **(c)** Wild-type (wt, 3 weeks old), transgenically expressing hKLK5 mice (Tg-hKLK5, 4 weeks old). (a), (b) and (c, “wt”) adapted from ⁴³ under CC BY 4.0, © 2017 Kasperek *et al.* (c, “Tg-hKLK5”) reprinted with permission from ⁴⁴, reprinting license nr. 4570230469510, © 2014 Rockefeller University Press.

1.2 Peptides in drug development

1.2.1 Peptides as drug modality

The different modalities currently used in drug development can be roughly classified based on their size into two main classes⁴⁶ (**Figure 11**).

The first class are small molecules, which are chemically synthesized, organic compounds with a molecular weight of less than 0.5 kDa. The majority of drugs used before the 21st century were based on small molecules, they are still considered as the classic drugs. Small molecules almost exclusively hold the property of being efficiently cell permeable making them the modality of choice for intracellular targets, comprising 75% of the pharmaceutical market⁴⁷.

In contrast, larger molecules, also called biologics, revolutionized drug development starting in the 1982 with the FDA approval of the first human insulin that was produced in bacteria⁴⁸. Biologics are polymers of natural building blocks like amino acids, nucleotides, and sugars, and thus can be manufactured in living cells. They can come in the shape of proteins, DNA, or more complex structures, such as viruses and whole cells. Regardless, the most common form of biologics are still large protein-based molecules, such as antibodies⁴⁹. Smaller modalities of biologics are represented by often-engineered intermediate protein scaffolds, while peptides form the link between biologics and small molecules⁵⁰.

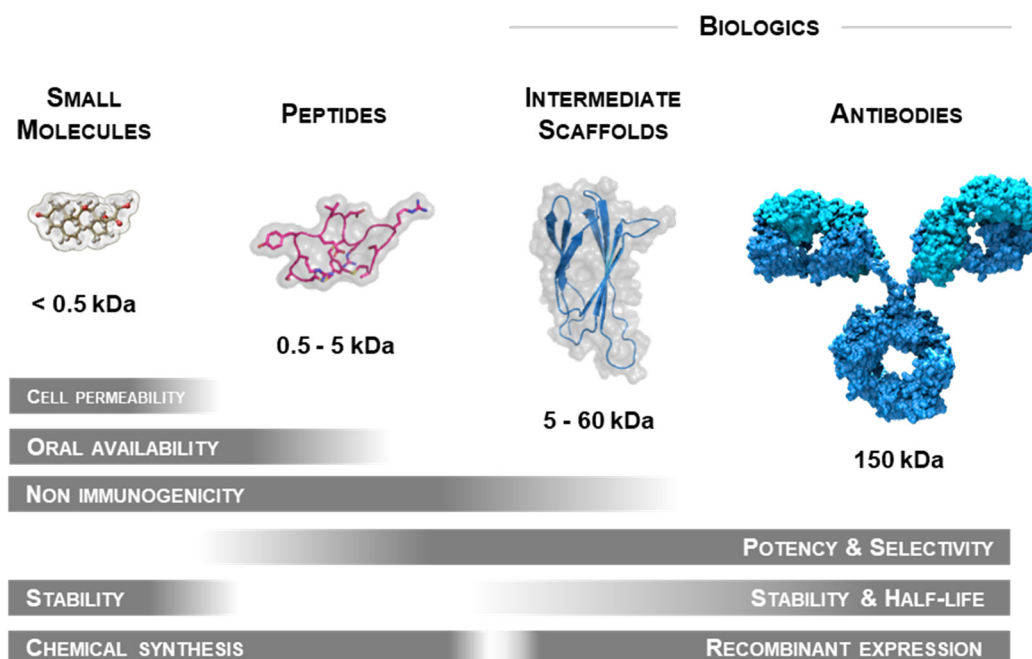


Figure 11 | Size and properties of drug modalities. Small molecules are chemically synthesized and usually orally available, cell permeable and metabolically stable. Peptides can also be chemically synthesized and provide high potency and selectivity, but usually have a short systemic half-life. Intermediate scaffolds and antibodies are produced by recombinant expression and have high potency and selectivity, but only antibodies come with a long systemic half-life. Image of small molecule and antibody used with permission of Amgen Inc.

IgG-type antibodies, due to their high potency and long half-life, are often considered as the magic bullets for human therapies. However, standard antibodies also have some limitations: Their large size impedes their ability to penetrate into certain tissues, such as less-vascularized tissues and organs like are found in the intestine, adipose, and muscle tissue, as well as renal and lung lumina. Also the brain and the central nervous system are difficult to access for antibodies⁵¹. Furthermore, the Fc part of antibodies might trigger physiological effects, which are undesirable for the treatment of some diseases^{52, 53}. Also, the development of multi-specific binders based on standard IgGs is still challenging⁵⁴.

To address some of these shortcomings, engineered antibody-fragments and alternative scaffolds have risen in the last years, which promise a next generation of protein therapeutics with properties superior to standard antibodies for certain applications. These scaffolds are usually based on protein domains of 6 – 30 kDa and can be combined to molecules of up to 60 kDa, such as diabodies⁵⁵ (**Figure 12**). In contrast to antibodies, such protein fragments can be expressed in microorganisms, which lowers their manufacturing costs. Some scaffolds such as adnectins, anticalins, avimers, Fynomers, Kunitz domains, and knottins have exposed flexible loops comparable to antibodies, while some like affibodies or DARPins have their binding regions integrated into their secondary structures⁵⁴. A risk of such molecules is their potential immunogenicity. Thus, for therapeutic applications, only fragments derived from human proteins or those that already have been shown to be non-immunogenic can be used. Another shortcoming for many applications is their shorter half-life compared to antibodies. In addition, despite their smaller size, they cannot reach all parts of the body. Furthermore, they usually are neither orally available nor can reach intracellular targets. As they can be more easily engineered than antibodies, promising applications of such scaffolds include the development of multi-specific binders as well as metabolically more stable and more homogenous antibody drug conjugates^{54, 56}.

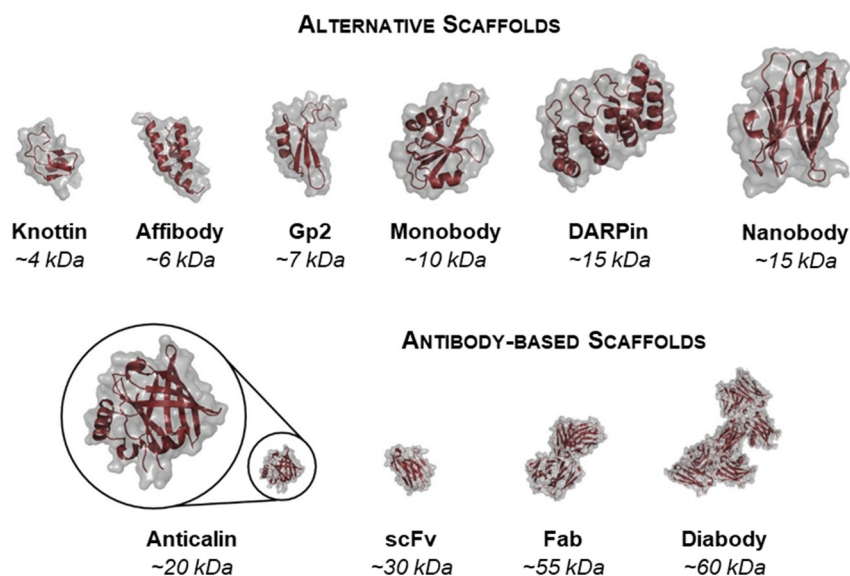


Figure 12 | Intermediate scaffolds. Engineered protein fragments with sizes between peptides and antibodies. These intermediate scaffold can be further classified as antibody-based and alternative scaffolds.

Adapted with permission from ⁵⁵, reprinting license nr. 4570240561447, © 2016 Elsevier.

Peptides form the link between small molecules and protein-based biologics. They are defined as short amino acid chains of up to 50 monomers and can thus be chemically synthesized. With sizes between 0.5 to 5 kDa, they combine some favorable properties of both major drug classes^{50, 57, 58}. Like proteins, peptides have a large surface area that makes them very potent and selective binders, capable of disrupting protein-protein interactions⁵⁰. They also undergo physiological catabolism, are degraded without the accumulation of harmful metabolites⁵⁹, and are considered to be a low risk for triggering immunogenic reactions⁶⁰. Their small size allows deep tissue penetration⁶¹, though their metabolic degradability and small size also comes with a short half-life in the organism⁶².

Because peptides up to 5 kDa can be entirely chemically synthesized or easily modified, their pharmacological properties can be greatly enhanced, which allows to engineer molecules with joint properties of small molecules and biologics. Common modifications are cyclization, peptide backbone modifications, incorporation of unnatural amino acid derivatives, and acylation⁵⁰. The possibility of making peptide drugs orally available or cell permeable is a great opportunity for this modality⁵⁷. The oral availability of a 4-kDa peptide has been achieved through acylation combined with uptake enhancers⁶³, however only a bioavailability of 1.2% could be achieved⁶⁴. Nevertheless, this shows the potential of peptides as orally available biologics. A complete overview of structural modifications that can be used to enhance pharmacological properties of peptides is given in chapter 1.2.5.

Around 100 peptide-based drugs are currently approved for the major pharmaceutical markets⁶², and several have reached blockbuster status, implying annual sales of more than one billion USD. Many of them fall in the category of hormones and hormone analogues, such as the insulin derivatives glargine (Lantus, Sanofi) and lispro (Humalog, Eli Lilly) and the GLP-1 analogues exenatide (Byetta, Astrazeneca) and liraglutide (Victoza, Novo Nordisk), both used for the treatment of diabetes. Teriparatide (Forteo, Eli Lilly) is a parathyroid hormone derivative used for the treatment of osteoporosis. Octreotide (Sandostatin, Novartis) was the best-selling peptide drug in 2018 with sales of 1.6 billion USD⁶⁵. It mimics the function of the natural hormone somatostatin and is used for the treatment of acromegaly, carcinoids, and gastrointestinal and pancreatic tumors. Another class of drugs is represented by the antibiotic daptomycin (Cubicin, Merck)^{66, 67}. The structure of three of the mentioned blockbuster peptide drugs are shown in **Figure 13**. Despite their various sources - natural and engineered - they share common structural features, such as cyclization and acylation. These features are associated with certain pharmacological properties, which will be discussed in the next chapters.

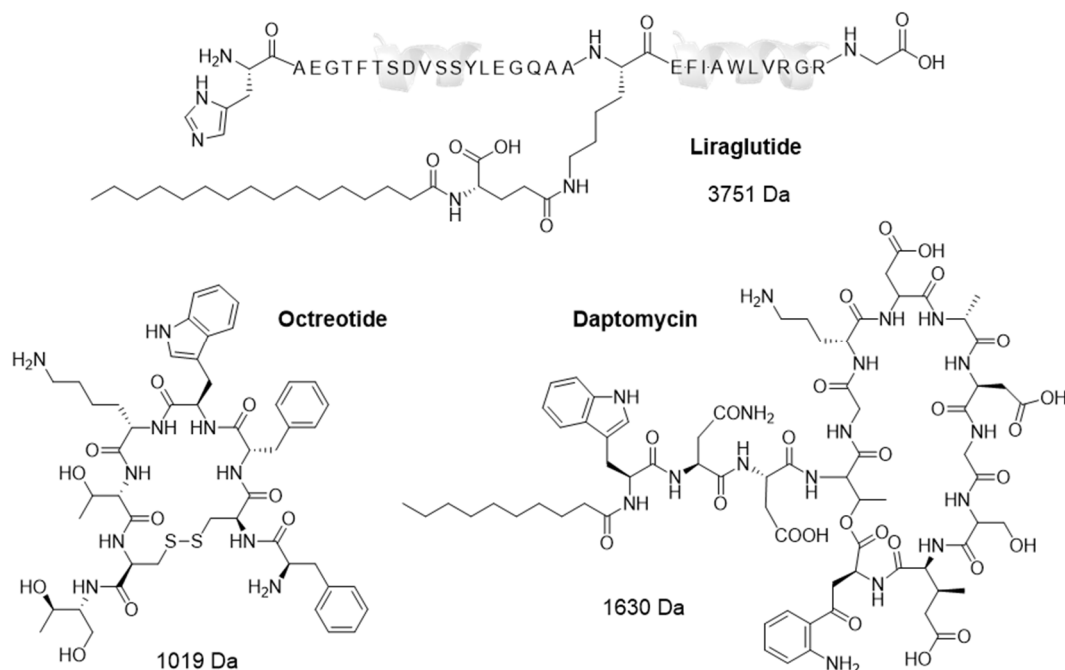


Figure 13 | Three peptide drugs with blockbuster status. Liraglutide (Victoza, Novo Nordisk) is a GLP-1 analogue used for the treatment of diabetes. Octreotide (Sandostatin, Novartis) is a somatostatin mimic used for the treatment acromegaly and certain types of tumors. Daptomycin (Cubicin, Merck) is an antibiotic. Common structural features are cyclization and acylation.

1.2.2 Directed evolution in drug development

The historical way for identifying new bioactive molecule is to isolate, identify, and test molecules from natural sources such as plants, fungi, or animals. Such molecules can then be chemically synthesized and modified to improve their potency, stability, or pharmacokinetic properties. As an example, the first biologics, such as insulin⁶⁸ and antibodies⁶⁹, were initially isolated or derived from animals. All these molecules had in common that they have evolved through natural selection over millions of years to determine their function. However, ligands cannot be easily discovered in nature for all relevant disease targets and natural compounds often do not come with the required pharmacological properties to be used as drugs.

Driven by progression in molecular biology, in the mid-1980s it was discovered that it is possible to mimic the process of natural evolution in the laboratory, leading to a powerful technology, which allows to develop peptide and antibody therapeutics entirely *de novo*⁷⁰. Like natural evolution, this process is driven by generating variation followed by selection according to desired properties, and inheritance to the next generation, making it an accelerated Darwinian process. The important characteristic of this technology is that a link between genotype (DNA code) and phenotype (protein sequence) exists, which allows for the easy identification of the amino acid sequence of the evolved proteins by DNA sequencing⁷¹.

Meanwhile, various directed evolution technologies have been developed since its initial discovery, the first of which being phage display, as described by George P. Smith⁷⁰ in 1985. In 1990, phage display was used for the first time to engineer antibody fragments by McCafferty *et al.*⁷². Since then, it has

been the workhorse for protein engineering, thanks to the high stability of phage particles and the straightforward and robust selection procedure⁵⁴. In 2002, the first drug developed using directed evolution was approved⁷³: Adalimumab (Humira, AbbVie) is a TNF- α neutralizing antibody, which is used for the treatment of arthritis, psoriatic arthritis, plaque psoriasis, Crohn's disease and other inflammatory diseases⁷⁴. With sales of 20 billion USD in 2018, Humira still tops the list of best-selling drug⁷⁵. For the discovery of directed evolution and its application to engineer peptides, antibodies and enzymes, Frances H. Arnold, George P. Smith and Sir Gregory P. Winter were awarded with the Nobel Prize in Chemistry 2018⁷⁶.

In phage display, a peptide or protein library is expressed on the surface of bacteriophages through genetic fusion with the pIII coat protein⁷⁰. Through infection of *E. coli* cells, phages expressing the combinatorial library on their surface are produced. As the DNA, which encodes the protein sequence, is enclosed in the genome of the same phage particle, it is physically linked to the protein. The phage-encoded libraries are then selected in a biopanning process to isolate binders with a high affinity or other desired properties, such as specificity or stability. In this process, immobilized target proteins are used to fish out binders, while non-binders are washed away. Binding phages are eluted and reamplified in bacteria to enrich for these sequences in another selection round. After several selection rounds, DNA sequencing of the phage genome reveals the amino acid sequences of the selected proteins⁷⁷ (**Figure 14**).

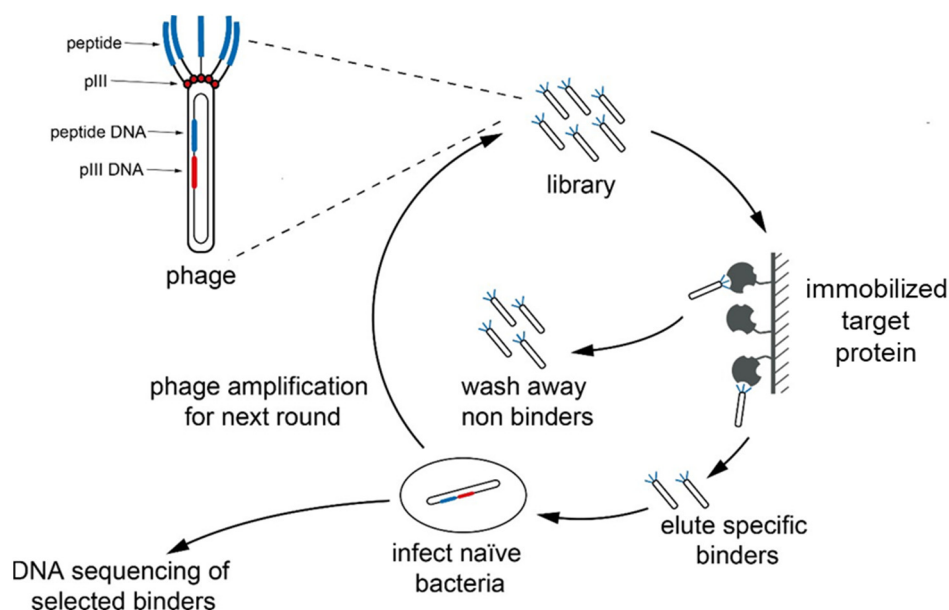


Figure 14 | Phage display technology. The peptide or protein to evolve (blue) is expressed on the surface of a bacteriophage through genetic fusion with the pIII coat protein (red). A combinatorial peptide or protein library is encoded in the phage genome. Through infection of *E. coli* cells, phage expressing the combinatorial library on their surface are produced. The phage library is then selected through exposure to immobilized target, where binders are fished out and non-binders are washed away. Binding phages are eluted and reamplified in bacteria for another selection round. After two to three selection rounds, the amino acids sequences of the binding peptides or proteins are identified by sequencing the phage genome.

With the expansion of the technology, many different options for phage-display libraries are available. To generate these libraries, a plasmid collection has to be transformed into competent *E. coli* cells, with the transformation efficacy usually serving as the library-size-limiting step, with the possible generation of libraries containing a diversity up to 10^8 – 10^{10} variants⁷⁸. The most commonly used phage strains for phage display are the M13, f1, or fd filamentous phage, whose genomes are comprised of single-stranded DNA that encodes five different coat proteins. One viral capsid contains about 2,700 copies of the major coat protein pVIII and five copies of the minor coat protein pIII, which is the most commonly used protein for protein display⁷⁷. Thus five copies of the protein to be selected are displayed on the phage surface (multivalent display), which can cause avidity effects during the selections. This might help in the initial selections rounds to capture weak binders from naïve libraries, but also increases the survival chances of weak binders in later selection rounds. To achieve monovalent display, phage vectors, which additionally express a wild-type version of pIII, can be used to reduce the multivalency of the displayed proteins. The size of proteins that can be fused to the pIII is limited by the total size of the genome that fits into a viral particle. To provide space for the display of bigger proteins, a phagemid that only contains the intergenic region for packaging and the pIII gene can be used for the library. All other proteins required for particle formation can be provided by co-infection with a helper phage⁷⁹ (expressing wt-pIII for monovalent display) or hyper phage (pIII-deleted for multivalent display)⁸⁰. Therefore, a phagemid system allows the user to switch between multivalent and monovalent display between the selection rounds.

To evolve proteins that require post-translational modifications to maintain their activity, yeast display provides an alternative, cell-based technique as its eukaryotic secretory system can perform these modifications. Combinatorial protein libraries can be expressed on the surface of yeast cells, which are selected using flow cytometry. Due to the lower transformation efficiencies of yeast compared to *E. coli*, the library diversity achievable with yeast display (10^{10} clones) is also lower than with phage display. This technique was first described by K. Dane Wittrup⁸¹ in 1997.

Additionally, mRNA and ribosome display are alternative, cell-free display technologies. By avoiding the required transformation of cells to generate libraries, higher diversities of $> 10^{12}$ variants can be obtained. Furthermore, the incorporation of unnatural amino acids is easier in these techniques⁷⁸. In mRNA display, the translated protein is covalently attached to its encoding mRNA molecule via a puromycin adaptor molecule that covalently attaches puromycin to the 3' end of all mRNA strands, which then uses the ribosome to covalently link the polypeptide chain. Reverse transcription and PCR amplification is used to recover selected mRNAs for the next selection round. This PCR amplification step can also be error-prone to introduce additional diversity into the library after each selection round⁷¹ (**Figure 15**). In ribosome display, the link between the protein and its encoding mRNA is mediated by the ribosome itself using an mRNA spacer lacking a stop codon at the end of each mRNA strand. In this way, the polypeptide chain is not released from the tRNA, and the ribosome does not disassemble, keeping translated protein and encoding mRNA linked. All other steps are comparable to mRNA display⁷¹. A challenge with mRNA and ribosome display is the stability of the mRNA, meaning that selections have to be performed under RNase-free conditions and at low temperatures⁵⁴.

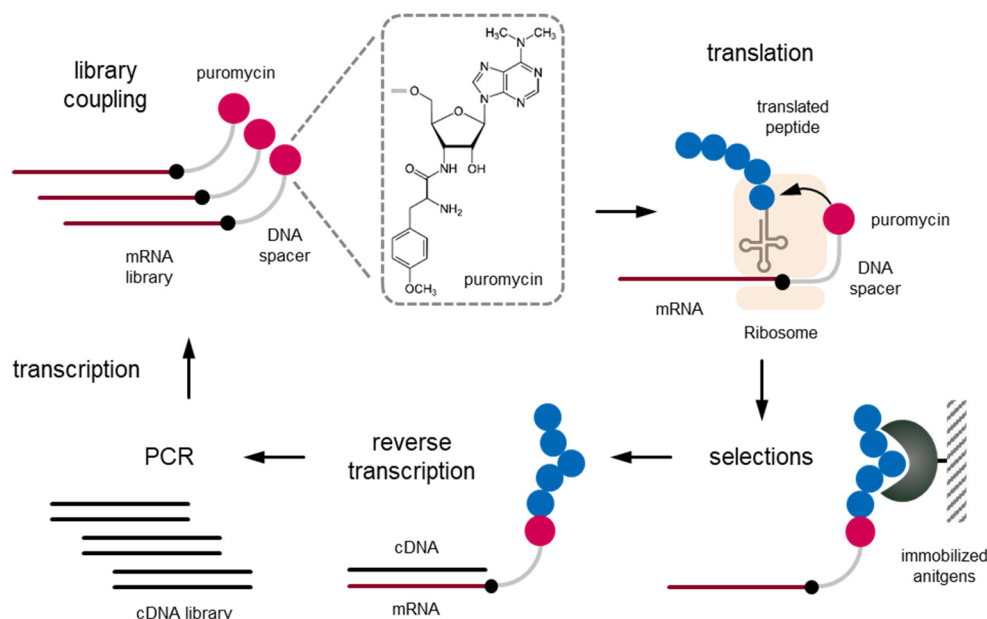


Figure 15 | mRNA display technology. Puromycin is used to covalently link a peptide and its encoding mRNA. This link is achieved by attaching puromycin to the 3' end of all mRNA strands. Puromycin is covalently linked to the polypeptide chain by the ribosome at the end of translation. After selections, reverse transcription and PCR amplification is used to recover selected mRNAs for the next selection round.

1.2.3 *In vitro* selection of peptides

While phage display was successfully applied for the development of therapeutic antibodies, only a few peptides had suitable properties for use in therapeutic applications. One reason for this is that the natural IgG scaffold already comes with many favorable properties like stability and half-life, whereas *in vitro* selection does not develop for these properties alongside target affinity. Therefore, phage display initially identified peptide ligands that bound weakly to their targets and were not resistant to protease pressure, making the technique useful only for antibody epitope mapping.

Linear peptides have been engineered from phage display for use in therapeutics, though. One example of a linear peptide successfully developed towards a drug is romiplostim (Nplate, Amgen). It was approved in 2008 and is used for the treatment of thrombocytopenia, a rare autoimmune disorder. For this drug, phage display originally identified a ligand of the thrombopoietin receptor (TPOR) that was further engineered to a so called “peptibody” by fusing it to an IgG Fc. Two ligands separated by a spacer were C-terminally fused to an IgG1 heavy chain, resulting in a molecule with four peptide ligands (**Figure 16**). Due to the absence of glycosylation, the molecule can be manufactured in *E. coli* cells.

While performing phage selections with linear peptide libraries, it was observed that peptides that were randomly cycled through the oxidation of two cysteines in the sequence were enriched⁸². When characterized, these cyclic peptides often had better affinities and stabilities than their linear counterparts. Thus, it was discovered that structural constraints, which are also characteristic of the complementarity-determining regions (CDRs) of antibodies, could be introduced into peptides to make

them more potent and stable⁸³. Cyclization or the introduction of secondary structures, such as alpha helices, limit the conformational flexibility of peptides, which reduces the entropic penalty upon binding and provides molecules with a higher binding affinity⁸⁴. Furthermore, structural constraints limit the accessibility of the peptide backbone to plasma proteases, which increases the metabolic stability⁸⁴. These insights inspired the second generation of peptide libraries with fixed cysteines that are cyclized with a disulfide bond⁸⁵. Cyclic peptides selected with these libraries bound the same target proteins with 100 to 1,000-fold higher affinities than the best-identified linear peptides⁸⁶.

The first peptide drug selected from disulfide-constrained libraries was peginesatide (Omontys, Affimax/Takeda), an erythropoietin mimic binding the erythropoietin receptor that reached the market in 2012. The phage-display-identified cyclic peptide⁸⁷ was further engineered by dimerization to achieve a stronger *in vivo* effect and was conjugated to two PEG chains of 20 kDa to extend the half-life. This allows for a monthly dosing regimen while keeping the molecule entirely chemically synthesizable (**Figure 16**)⁸⁸. Peginesatide was used for the treatment of anemia due to chronic kidney disease, but had to be withdrawn from the market in 2014 due to lethal adverse reactions. Later studies revealed that the phenol used as a preservative in the formulation, rather than the active molecule itself, likely caused the anaphylactic reactions⁸⁹.

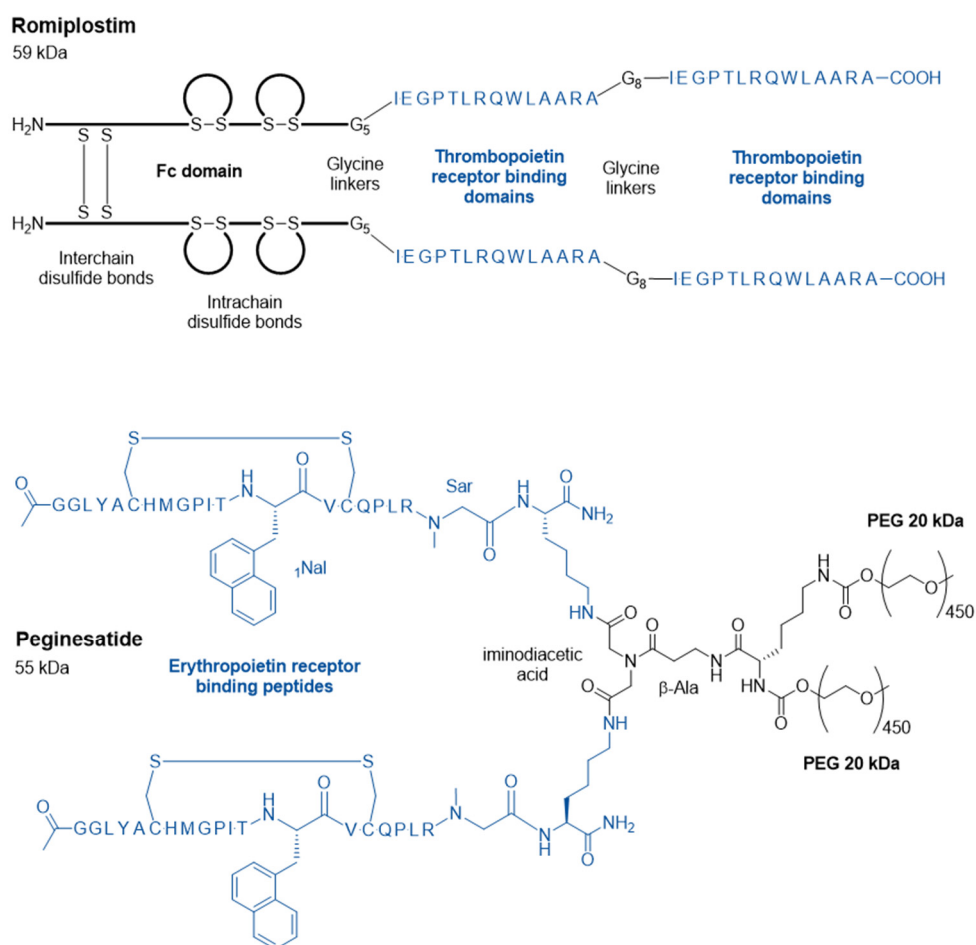


Figure 16 | Examples of first and second generation of phage selected peptide drugs. Romiplostim (approved in 2008) is based on a linear peptides fused to an IgG Fc part and manufactured in *E. coli* cells. Peginesatide (approved in 2012) is based on a PEGylated, disulfide cyclized peptides and is chemically

However, disulfide bridges are known to be susceptible to reduction and disulfide exchange reactions. Thus, a third generation of encoded peptide libraries replaced the disulfide bridges with small, non-reducible linkers. Libraries of non-reducible, cyclic peptides were developed using a variety of different strategies for cyclization in combination with different display technologies⁹⁰ (**Figure 17**). The very first of such libraries were phage-encoded bicyclic peptides, which will be discussed in depth in the next chapter.

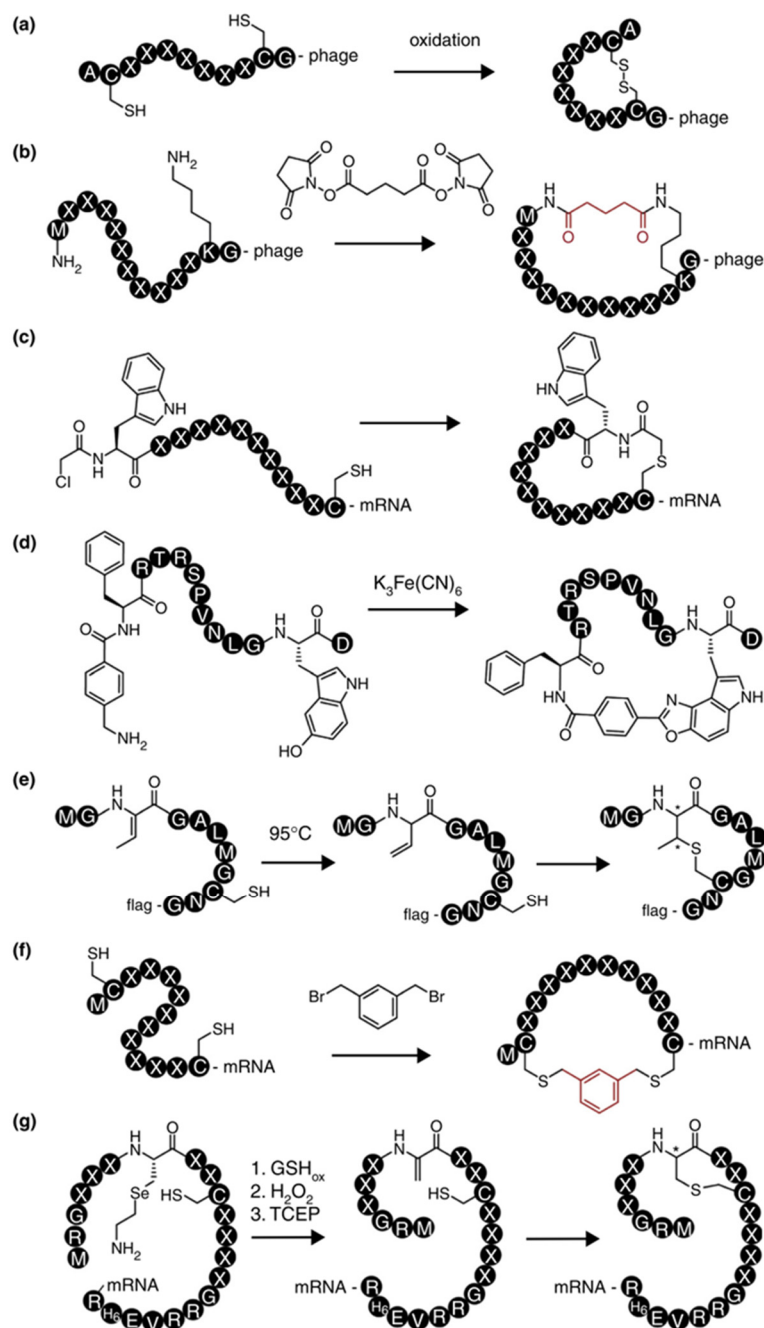


Figure 17 | Cyclization strategies for encoded peptides libraries. (a) Cyclization by oxidation of two cysteines. (b) Linking the N-terminus to the side chain of a Lys with DSG. (c) Spontaneous cyclization of an N-terminally conjugated chloroacetyl group with Cys (d) Cyclization N-terminal benzylamine with unnatural 5-hydroxytryptophane. (e) Cyclization of Cys with dehydrobutyrine, heat-isomerized from unnatural vinylglycine. (f) Cyclization by linking two Cys with α - α' -dibromo-m-xylene. (g) Michael addition of Cys to dehydroalanine generated by oxidative transformation of 4-selenalysine. Reprinted with permission from ⁹⁰, reprinting license nr. 4570240710300, © 2015 Elsevier.

1.2.4 Phage display of bicyclic peptides

Inspired by the idea of further improving the pharmacological properties of peptides by making them more similar to antibodies, in 2009, C. Heinis and G. Winter pioneered a new technology for developing therapeutic peptides⁹¹. This technology relied on the use of bicyclic peptides, which are peptides with three cysteines that, when linked with a trivalent, thiol-reactive small molecule, form peptides with two macrocyclic rings. Such cyclization reactions, previously only used for epitope mapping of non-linear antibody epitopes⁹², were applied by Heinis and Winter to cyclize linear peptide libraries on the surface of bacteriophages. With this strategy, phage-encoded peptide libraries with an enormous diversity could be generated. The first libraries used with this strategy were symmetrical linear libraries of the format XCX_mCX_mCX ($C = \text{Cys}$, $X = \text{any canonical amino acid}$) with various ring sized ($m = 3$ to 6) that were successfully cyclized using the 1,3,5-tris-(bromomethyl)benzene (TBMB) linker⁹¹ (**Figure 18**).

The technology takes advantage of the power of phage display combined with that of chemical peptide modification to produce libraries with improved properties in terms of potency, stability and diversity. To perform this chemical modification, a reduction step with tris-(2-carboxyethyl)-phosphine (TCEP) to reduce the thiols is first performed on the phage after the production and purification of a linear peptide library. This is important, as cysteines were found to be oxidized after expression, which likely already happens in the oxidative environment in the bacteria or in the oxygenated culture medium. After reduction, cyclization in a buffered solution at pH 8 with an optimized linker concentration is performed. Initial iterations of this technology used phage with a disulfide-free pIII protein to avoid undesired reactions of pIII cysteines with TBMB⁹³. However, these phages had a strongly reduced infectivity, so later technology advances optimized the reaction conditions to perform reduction and cyclization on M13 phage with wild-type pIII without significantly reducing phage infectivity. As is normal for phage display, binders are identified via Sanger or next-generation sequencing⁹⁴ after three selection rounds, and peptides are chemically synthesized for characterization. Phages with compromised infection capacity can even be sequenced without infection of *E. coli* cells⁹⁵.

With this technology, inhibitors with a low nanomolar affinity were successfully selected against targets such as plasma kallikrein (PK)⁹⁶, urokinase-type plasminogen activator (uPA)⁹⁷, coagulation factor XII (FXII)⁹⁸. All inhibitors were characterized by a high potency and selectivity for their target; one of the best characterized bicyclic peptide inhibitors being the uPA inhibitor UK18.

The serine protease uPA converts plasminogen into plasmin. It was found to be overexpressed in cancer, degrading the extracellular matrix and promoting metastasis, and thus was considered a potential drug target. Our lab selected a bicyclic peptide inhibitor of uPA called UK18 from a library of the format XCX_6CX_6CX cyclized with TBMB. UK18 has an affinity of 53 nM for human uPA, which was three and one orders of magnitude lower than the affinity of its monocyclic and linear analogues, respectively⁹⁷ (**Figure 19a**). Furthermore, the inhibitor was more than 1000-fold selective over related proteases, including the mouse orthologue muPA (**Figure 19b**). Co-crystallization with uPA demonstrated that both macrocyclic rings engage in target binding and that the inhibitor shares an interaction surface of $> 900 \text{ \AA}^2$ with its target (**Figure 19c**), exemplifying the superior properties of bicyclic peptides over small molecules regarding affinity and selectivity⁹⁷. However, behaving like typical peptides in a pharmacological setting, these molecules still require engineering to obtain pharmacokinetic properties suitable for therapeutic application^{99, 100}.

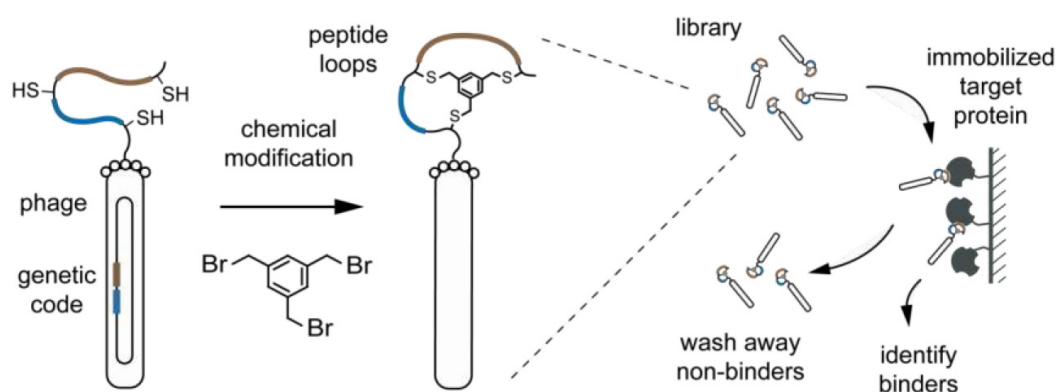


Figure 18 | Phage display selections of bicyclic peptides. Linear peptides with three cysteines are expressed on the surface of phage and cyclized by linking the sidechains of the cysteines with tris(bromomethyl)-benzene (TBMB, left). The combinatorial library of DNA-encoded bicyclic peptides is subsequently subjected to several rounds of affinity selections using immobilized target proteins (right). Bicyclic peptides that bind to the target are enriched throughout the selection rounds, and the peptide sequences can be identified by DNA sequencing. For characterization, peptides are chemically synthesized.

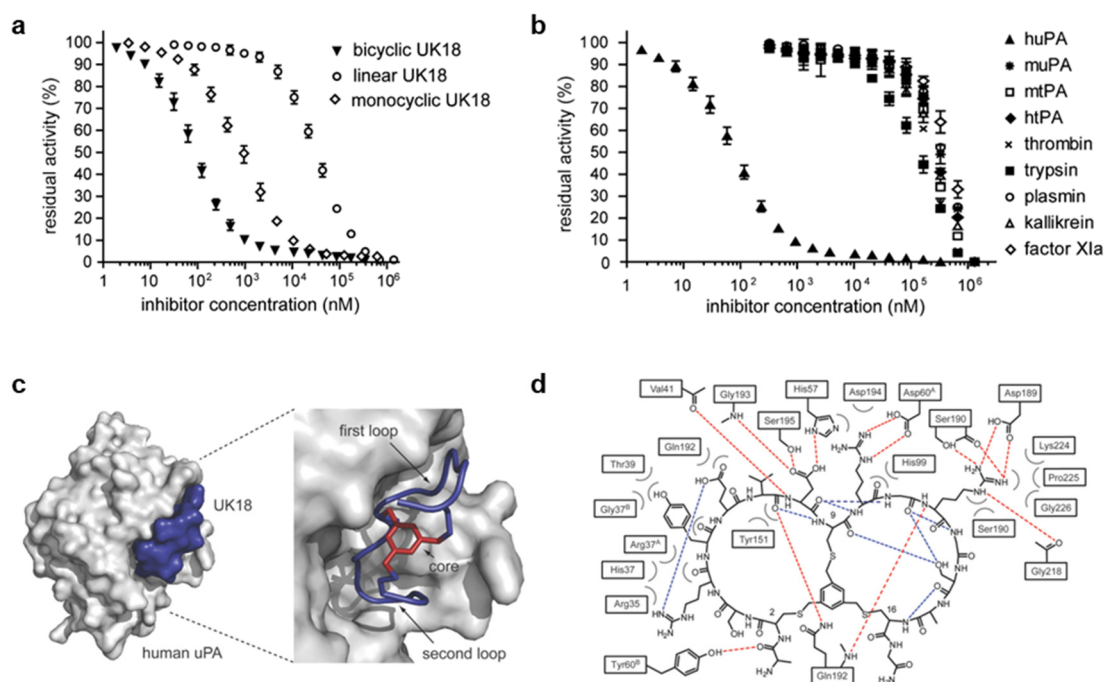


Figure 19 | Bicyclic peptide inhibitor UK18. (a) Residual activity of uPA incubated with bicyclic peptide UK18 or its linear or monocyclic derivatives. (b) Inhibitory activity against uPA and a range of homologous trypsin-like serine proteases were measured. (c) Crystal structure of uPA (gray) in complex with the bicyclic peptide UK18 (peptide ribbon in blue, mesitylene scaffold in red). (d) Schematic representation of molecular interactions between uPA and UK18. Potential intermolecular (red) and intramolecular (blue) hydrogen bonds are shown as dashed lines. Distances shorter than 4.0 Å that are not hydrogen bonds are shown as bent gray lines.

Adapted with permission from ⁹⁷, © 2012 American Chemical Society.

In 2018, a novel bicyclic peptide format that expands upon this previous technology was published by our lab, referred to as double-bridged peptides¹⁰¹. The technology uses an intriguing insight from the bicyclic peptide library screens: In libraries with three cysteines that were selected in an oxidized form as a control, 62% of the selected peptides contained a fourth cysteine in the randomized regions¹⁰². Further analysis revealed that they were selected as cyclized via two disulfide bridges and had a structure similar to knottins, which are natural peptides that are highly constrained, giving them an especially high metabolic stability¹⁰³. However, as previously mentioned, disulfide-bonds are undesirable, though replacing them with non-reducible, chemical linkers could provide peptides with improved properties. Based on this idea, the double-bridged peptide format was developed, which is analogous to bicyclic peptide phage display, though replaces the single tri-cysteine linker with two bi-cysteine linkers¹⁰¹ (**Figure 20**).

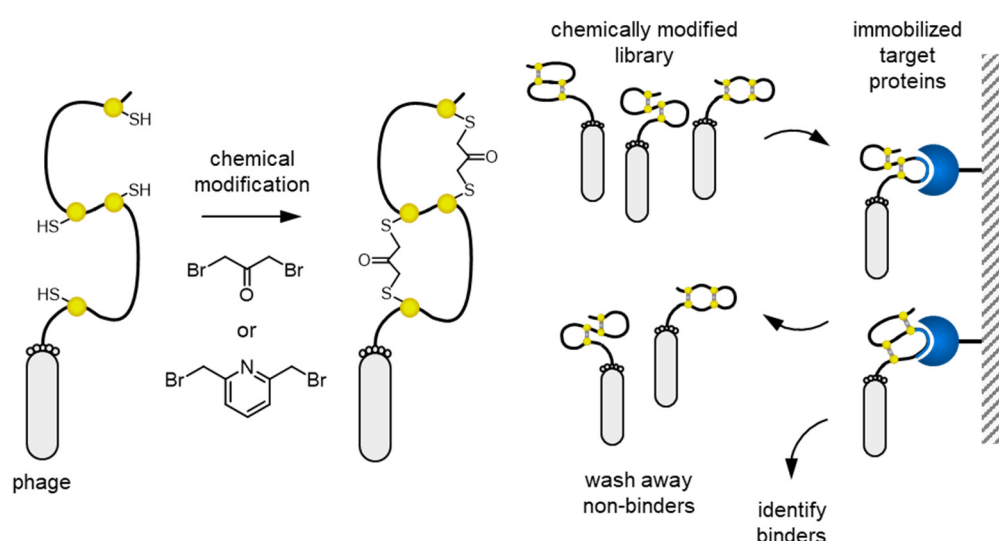


Figure 20 | Phage display selections of double-bridged peptides. Linear peptides containing four cysteines are expressed on the surface of a filamentous M13 phage by N-terminal fusion to the pIII coat protein. After phage expression and reduction, the peptides can be cyclized *in situ* with various thiol-reactive linkers, such as 1,3-dibromoacetone or 2,6-bis(bromomethyl)pyridine. Three regioisomers form during this cyclization reaction, providing libraries of extreme topological diversity. Peptides are selected with immobilized targets and binders are identified by sequencing.

This new, double-bridged peptide format provides several advantages over the bicyclic peptide format previously described: First, during the cyclization reaction of a linear peptides with four cysteines and two linkers, three different regioisomers form depending on which of the two cysteines are connected. This expands the genetically encoded diversity by another factor of 3 in structural diversity. Second, many more bi-valent chemical linkers are commercially available than tri-valent ones, further expanding the theoretical structural diversity. Third, new libraries, which have the four cysteines in variable positions, were generated such that variable ring sizes could be selected from a single library. With this strategy, phage-encoded, cyclic peptide libraries of unprecedented topological diversity could be generated. On the one hand, this is advantageous for challenging targets where no strong binders could be identified with previously used libraries. On the other hand, small, monocyclic ligands can also be identified from such libraries if they fit better into small pockets. Furthermore, such

libraries also allow the identification of strongly constrained peptides, which might be more resistant to protease pressure. Libraries of the format $\text{XCX}_m\text{CX}_n\text{CX}_o\text{CX}$ with $m + n + o = 3-8$ (Kale *et al.*,¹⁰¹) and 12 (Carle *et al.*, unpublished) were already tested with around 20 different linkers, out of which 1,3-dibromoacetone and 2,6-bis(bromomethyl)pyridine turned out to yield the best results.

1.2.5 Improvement of pharmacological properties

The previous chapters provided an overview of how directed evolution technologies can be leveraged to select peptide ligands approaching antibody-like potency and selectivity by introducing structural constraints. Contrary to antibodies, though unmodified peptides usually have poor pharmacokinetic properties such as metabolic stability and systemic half-life that might not be suitable for therapeutic application. The approval of peptide drugs, however, shows that this can be overcome with some engineering. These approved drugs also exemplify the most common strategies for the improvement of pharmacological properties: In liraglutide, acylation is used (**Figure 13**). In romiplostim, Fc-fusion is used (**Figure 16**). PEGylation as well as the replacement of amino acids with unnatural, N-methylated derivatives can be seen in the structure of peginesatide (**Figure 16**). Together with already mentioned structural constraints, which can be found in liraglutide (natural alpha helix) and peginesatide (cyclization), the peptide chemistry toolbox for the enhancement of pharmacological properties is almost complete. An overview of all strategies is shown in **Figure 21**, and they will be discussed alongside the properties they can improve. The oral availability and cell permeability of both peptides and antibodies are still difficult to achieve.

Potency and selectivity are often linked to the scaffold and are limited through its size by the number of potential interactions it can make. The potency and selectivity of a scaffold with a fixed size can be improved through the introduction of structural constraints such as disulfide bonds, chemical bridges, or secondary structural elements. Accordingly, initial shortcomings of peptides in potency and selectivity could be overcome with advanced directed-evolution technologies that embed constraints into the libraries (chapter 1.2.3 and 1.2.4). Thus, for *de novo* engineered peptides, scaffold size and constraints should already be considered before the selections, as they are determined and generally fixed by the peptide library. The later introduction of global constraints into peptides, such as those that might be derived from natural sources, is possible¹⁰⁴, but might be not as efficient as introducing these constraints into an initial library. If no sufficient potency or selectivity over related, critical targets can be achieved, selections with peptide libraries with larger scaffolds providing more randomized residues for interaction might be necessary. Smaller potency improvements can also be achieved after the selections by screening and introducing non-natural amino acid derivatives in an affinity maturation approach¹⁰⁵.

Often problematic with peptides is their low metabolic stability due to a rapid enzymatic proteolysis of the peptide backbone⁶². The introduction of global constraints to limit the accessibility of the peptide backbone to plasma proteases can improve the metabolic stability, and efforts have been made to select early for metabolic stability by exposing the peptide library to protease pressure^{106, 107}. However, selecting for stability often requires a trade-off, as peptides selected under protease pressure can have lower potencies for their desired target. Thus this might only be a reasonable approach if peptides are needed that can resist strong protease pressure, such as for oral

administration. For parenterally administered peptides, stability issues can be addressed in a targeted fashion using structural modifications to the final molecule. Vulnerable bonds can be identified *in vitro* or *in vivo* and can be replaced by amide bond surrogates. Local constraints can also be introduced by replacing natural amino acids around the identified vulnerable peptide bond with non-natural derivatives, N-methyl-, D-, or beta amino acids. This usually improves the metabolic stability, as proteases have evolved to only recognize natural amino acids.

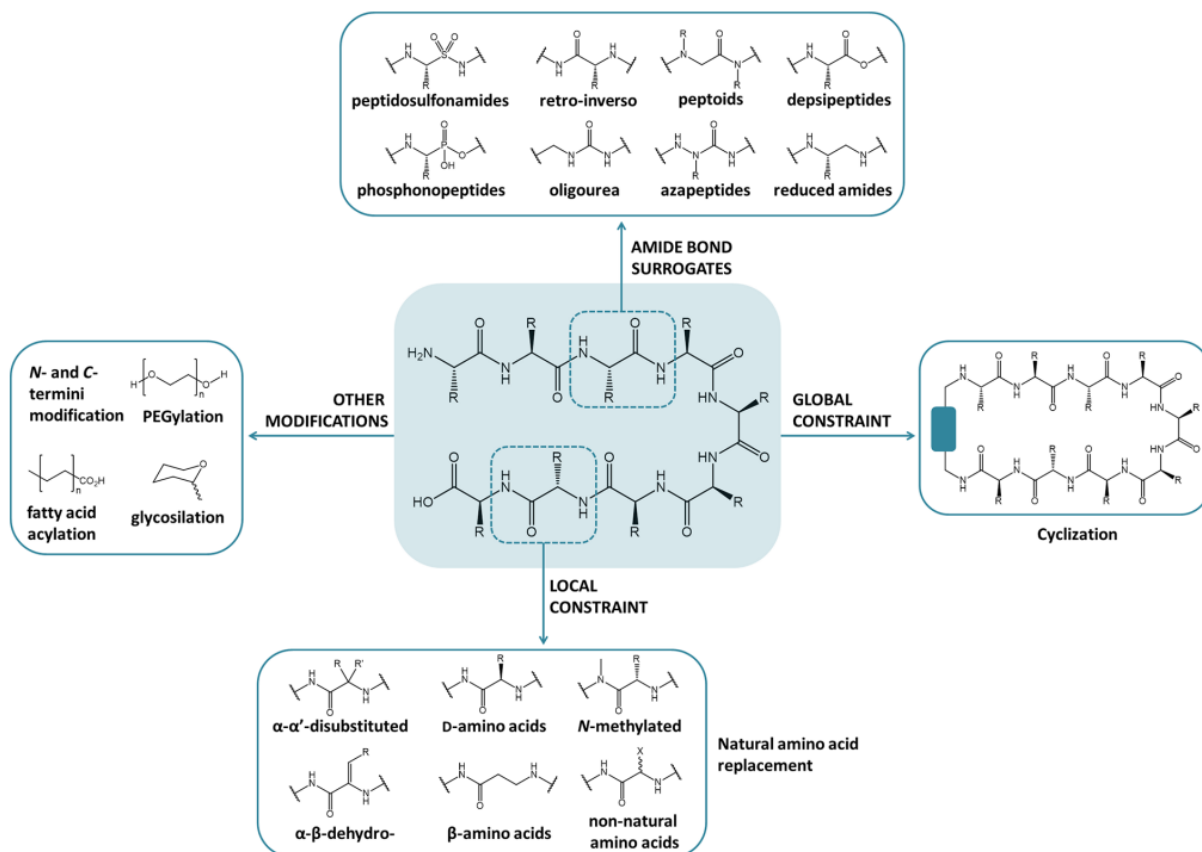


Figure 21 | Peptide modifications to enhance pharmacological properties. The introduction of amide bond surrogates is usually used to improve the metabolic stability. The introduction of global constraints, such as cyclization or secondary structural elements, can improve potency and selectivity as well as the metabolic stability. Local constraints specifically address stability issues and non-natural amino acid derivatives can also improve the potency. Other modifications represent conjugation strategies for half-life extension *in vivo* and are shown in detail in Figure 22. Reprinted with permission from ⁵⁰, © 2014 American Chemical Society.

Another drawback of peptides is their short systemic half-life due to their small size. Molecules with a weight of less than 50 kDa undergo rapid renal clearance, as this is the molecular weight cut-off of the kidneys that filter the blood and transforms it into urine. The filtering occurs at the glomerular epithelium, which has pores with a diameter of around 8 nm and is rich in negatively charged proteoglycans¹⁰⁸. Thus, molecules such as albumin with weights of more than 50 kDa and negative charges are retained in the blood, while molecules such as peptides with weights of only 5 kDa are filtered out to the same extent as water.

Based on this, the first generation of half-life extension strategies aimed to increase the hydrodynamic volume of molecules to reduce renal filtration^{109, 110} (**Figure 22**). This can be achieved by genetic fusion or chemical conjugation to large molecules. Most commonly used is conjugation to the synthetic, hydrophilic polymer polyethylene glycol (PEG)¹¹¹. As adverse reactions to PEG have been reported^{112, 113}, some natural polymers based on polypeptides or saccharides, which are supposed to be safer and less expensive, also exist. One of these is XTENylation, based on a polypeptide with up to 864 amino acids consisting of Ala, Asp, Gly, Pro, Ser, and Thr and that can be genetically fused or chemically conjugated. A similar approach only using Pro, Ala, and Ser residues is called PASylation, which uses polysaccharides such as hydroxyethyl starch (HES), polysialic acid (PAS), hyaluronic acid (HA), and dextran to increase peptide half-lives¹¹¹. Also, naturally occurring post translational modifications such as N- and O-glycosylations can have a half-life extending effect by increasing the molecular weight by 5 kDa per glycan, which has been exploited by introducing additional glycosylation sites into small proteins such as erythropoietin (37 kDa) to reach 50 kDa¹¹⁴. However, the fusion to or conjugation with large polymers usually impairs the pharmacodynamic properties of peptides due to steric hindrance¹¹⁵. Furthermore, ending up with a large molecule, which is difficult to manufacture, wipes out many of the initial advantages of peptides.

In contrast, non-covalent strategies based on conjugation or fusion to a small ligand that binds endogenous plasma proteins and increases the hydrodynamic volume this way allows the final molecule to remain much smaller than 50 kDa. This better preserves tissue penetration capabilities and makes the molecule easier to manufacture. Furthermore, it opens up more options for administration: while large molecules must almost always be administered parenterally, smaller molecules based on binding strategies for half-life extension might be administered orally, nasally, or topically¹¹⁶.

The human blood offers plenty of different plasma proteins as potential carriers for therapeutic molecules, of which albumin (55%) and various globulins (38%) are the most abundant. When comparing the half-lives of different plasma proteins, it was found that serum albumin and IgG have an extraordinarily long plasma half-life of around 20 days. This is enabled through binding to a distinct receptor, the neonatal Fc receptor (FcRn), which protects binding molecules from degradation¹¹⁷. This was already hypothesized for IgG in 1970 by Brambell *et al.*¹¹⁸, while it was discovered for albumin¹¹⁹ only in 2003. The FcRn is widely expressed in the human body and has two major functions: One is the transplacental transport of IgG from the mother to fetus to provide neonatal immunity. The second is to recycle albumin and IgG, reducing the energy consumption for anabolism. FcRn-mediated recycling works using a pH-dependent mechanism: Plasma proteins that are randomly internalized by cells through pinocytosis bind FcRn in acidic conditions in the vesicle and are then transported back to the cell surface and released at a neutral pH. Unbound proteins are degraded within the forming lysosome¹¹⁷. Based on these understandings, the second generation of half-life extension strategies not only aims to increase the volume but additionally take advantage of FcRn-mediated recycling^{109, 110} (**Figure 22**).

This can be achieved again either with covalent or non-covalent binding strategies, both using either albumin^{120, 121} or the Fc part of an IgG. As albumin has a molecular weight of 66.5 kDa and an Fc is 50 kDa, all four options also sufficiently increase the hydrodynamic volume of therapeutic molecule in plasma. If the molecule already has a sufficient hydrodynamic volume, direct FcRn binding using a synthetic or natural FcRn ligand is also an option to further improve the half-life¹²². The direct genetic fusion to Fc or albumin is the most common for large molecules, such as protein fragments or enzymes, that have to be recombinantly expressed. For smaller molecules, such as peptides, which can be chemically synthesized or modified, conjugation to an albumin ligand is a successful strategy¹¹⁰. Despite the diverse approaches that cover all three strategies, the half-life of GLP-1 receptor agonists could be successfully prolonged¹²³ to reduce the administration frequency from twice per day to once per week subcutaneous injections in humans¹²⁴ (Fc fusion: dulaglutide¹²⁵ [Trulicity, Eli Lilly], albumin fusion: albiglutide¹²⁶ [Tanzeum, GSK], albumin-ligand conjugation: semaglutide¹²⁷ [ozempic, Novo Nordisk]).

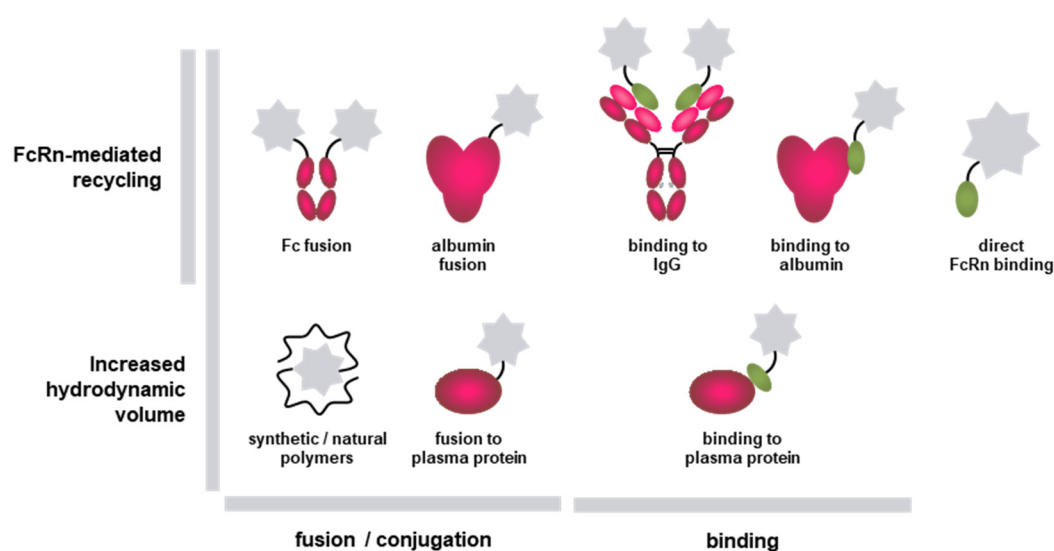


Figure 22 | Strategies for half-life extension. The first generation of half-life extension strategies aim to increase the hydrodynamic volume of molecules to reduce renal filtration. Second generation approaches also take advantage of FcRn mediated recycling, though direct FcRn binding does not necessarily sufficiently increase the hydrodynamic volume, and a molecular weight of 50 kDa still must be achieved.

Oral availability is another highly anticipated property that would make peptide drugs more patient compliant and suitable for everyday use. However, as there are different oral uptake pathways (transcellular, paracellular), there are no simple rules to improve oral availability¹²⁸. Recently, 4 kDa peptides have been made orally available through acylation in combined with an uptake enhancer referred to as SNAC (sodium N-[8-(2-hydroxybenzoyl)amino]caprylate) (oral semaglutide⁶³, FDA approval filed¹²⁹, Novo Nordisk). This nicely demonstrates the flexibility in terms of administration for peptides with non-covalent albumin binding as a half-life extension strategy, as the exact same molecule is approved for subcutaneous administration. Furthermore, this example demonstrates that

if cell permeability is not required, a low molecular weight is not a requirement for oral availability. It also shows the future potential of peptides as orally available biologics.

Also challenging to achieve is cell permeability through passive diffusion^{130, 131}. Approaches to achieve cell permeability are the reduction of size and the replacement of amide bonds⁵⁰. As cell permeability is not required within this work, this will not be discussed further.

In summary, the choice of the right half-life extension strategy is crucial for the development of a peptide drug. Aspects such as size of the final molecule, manufacturing and biophysical properties suitable for the intended application should be considered. The currently available tools allow the development of peptides with two main advantages over biologics: One is the possibility to flexibly and precisely engineer them to obtain multifaceted conjugates. The second is their deep tissue penetration due to their smaller size, allowing to reach targets not accessible to bigger modalities. The next big outstanding goal will be oral availability.

2 Aim of this work

The aim of this work was the development of a molecule that could be further developed towards a drug for the rare, genetic skin disease Netherton syndrome (NS), for which no satisfactory treatment exists. NS is caused by a loss-of-function mutation in the gene coding for the protease inhibitor LEKTI which causes unregulated activity of its targets, the two proteases KLK5 and KLK7. Thus, specific inhibitors of KLK5 and KLK7 could replace the function of genetically impaired LEKTI and provide the first, targeted therapy for NS. To develop such inhibitors, the initial goal was to identify cyclic peptides that bind their target proteases with high affinity by screening double-bridged peptide libraries with phage display. In case good binders could be identified, the plan was to further improve their affinity to the single-digit nanomolar range by incorporation of unnatural amino acids. Given that peptides are rapidly cleared by the kidneys *in vivo*, we planned to apply a half-life extension strategy to the molecules in the next step. Then the metabolic stability of the inhibitor in blood plasma can be assessed *in vitro* and potential stability issues be addressed with structural changes. Then the pharmacokinetic properties of the molecules can be evaluated *in vivo* in wild-type mice. If a satisfying half-life and biodistribution can be achieved, the final goal was to show a therapeutic effect of the inhibitors in a disease model of NS.

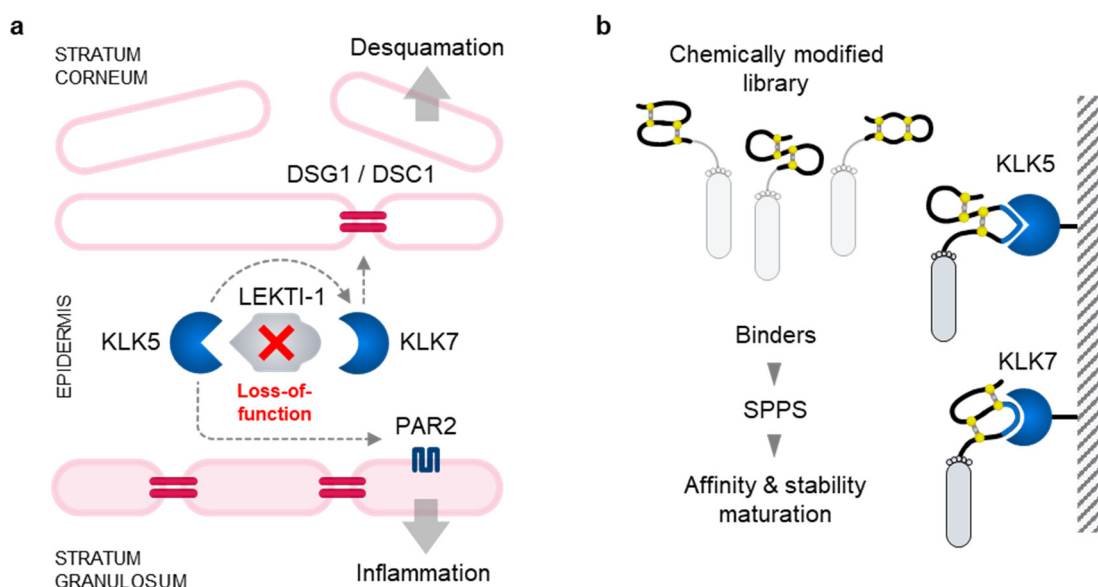


Figure 23 | Development of peptide-based inhibitors for the treatment of NS. (a) NS is caused by missing regulation of the epidermal proteases KLK5 and KLK7 by LEKTI. KLK5 activates KLK7, but also cleaves protease activated receptors (PAR) like PAR2 causing skin inflammation. KLK7 cleavages cell-cell connecting desmosomes like DSG1 or DSC1 and causes skin scaling (desquamation). Thus, KLK5 and KLK7 represent the major therapeutic targets for the treatment of NS. **(b)** To develop novel peptide-based therapeutics, KLK5 and KLK7 should be recombinantly expressed for the screening of cyclic peptide inhibitors in this project. Linear peptide libraries are expressed on the surface of phages and modified with chemical linkers to obtain bicyclic peptides. Affinity selections are performed, and binders are identified by sequencing, chemically synthesized by solid phase peptide synthesis (SPPS), characterized and improved.

3 Results & Discussion

3.1 Phage selection of bicyclic peptide inhibitors of KLK5 and KLK7

3.1.1 Choice of target proteins

The very first question addressed in this project was to determine what form of the KLKs should be expressed and used for phage display selections and further inhibitor development. Either the zymogen or the mature protease of KLK5 and KLK7 could potentially be targeted by peptides. Furthermore, one could either target the human or the mouse orthologue.

The function of a zymogen inhibitor would be to prevent activation of the protease. We decided to develop inhibitors of the active, mature forms of the two KLKs and not the zymogens for the following reasons. The endogenous inhibitor, whose function should be replaced inhibits the active forms of KLK5 and KLK7. This means that the kallikreins are present in the target tissue in their active form as especially KLK5 might already be partially activated intracellularly. Thus, for efficiently inhibiting zymogen activation, a cell-permeable molecule might be required, which is more difficult to develop. From a technical point of view, the development of zymogen activation inhibitors or even inhibitors of both forms would be more challenging as the structure of the zymogen is more flexible and does not have an oxyanion hole. Additionally, inhibitors of the active proteases simplify *in vitro* testing as common substrate-based activity assays can be used to test the binding of peptides.

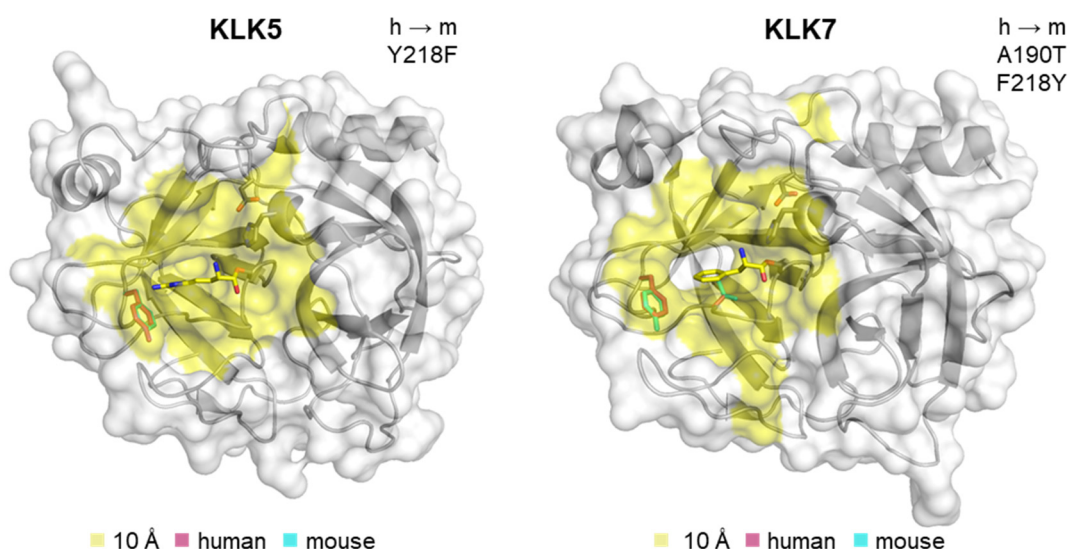


Figure 24 | Structural comparison between human and mouse orthologues of KLK5 and KLK7. Homology models of mKLK5 and mKLK7 aligned to crystal structures of hKLK5 in complex with Leupeptin (2PSX) and hKLK7 in complex with Ala-Ala-Phe-chloromethylketone (2QXG). Side chains of the catalytic triads (Ser195, His57 and Asp102) are shown in gray. P1 residues of the active site ligands are shown as yellow sticks. The side chains of residues within 10 Å from the alpha carbon of the P1 residue (yellow surface) that differ between human and mouse are shown in blue (mouse) and red (human). Identified mutations from human to mouse (h → m) are listed in the top right corner.

Furthermore, it had to be decided whether the human or the mouse orthologues of KLK5 and KLK7 should be used as targets for the inhibitor development. For the treatment of NS patients, inhibitors of the human kallikreins are needed. However, efficacy testing needs to be performed on animals and so far only mouse models of NS are available. Thus, peptides that could inhibit the human and mouse orthologues would be ideal. To assess whether this would be feasible, the crystal structures of human KLK5 and KLK7 were aligned to model structures of the corresponding mouse orthologues (generated with Swiss-Model; <http://swissmodel.expasy.org>). We compared an area of 10 Å around the alpha carbon of the P1 residues of the co-crystallized active site ligands (KLK5: leupeptin; KLK7: Ala-Ala-Phe-chloromethylketone) because peptide inhibitors selected by phage display will most likely bind to this surface. This comparison revealed only one differing residue for KLK5: human KLK5 (hKLK5) has a tyrosine at position 218 whereas mouse KLK5 (mKLK5) has a phenylalanine, a conservative mutation without a predicted structural change. The same amino acid change can be found in KLK7 but inversely: hKLK7 has a phenylalanine at position 218 and mKLK7 a tyrosine, exemplifying the exchangeability that does not impact the structure. hKLK7 and mKLK7 also differ at position 190, which is an alanine in hKLK7 but a serine in mKLK7. This change is located within the S1 subsite and thus might impair the substrate preference of the S1 subsite. Overall, however, we concluded from this comparison that the 10 Å areas around the active sites of KLK5 and KLK7 are well conserved between the human and mouse orthologues such that inhibition of the respective proteases of both species should be feasible using the same peptide. For this reason, the human proteases were chosen as targets for inhibitor development.

3.1.2 Expression strategies and constructs

Expression of the active proteases was expected to be more challenging. It can result in low yields if the protease is toxic to the cells because it cleaves intracellular targets and degrades the host cells. Furthermore, some proteases auto-degrade when highly concentrated in the cell or supernatant. Several strategies addressing these issues and allowing for the efficient expression of tissue kallikreins are reported in literature. KLK5 and KLK7 were previously expressed in insect cells using a baculovirus expression system while replacing the endogenous pro-domain by ubiquitin followed by an enterokinase (EK) cleavage site, to keep the zymogen conformation before purification^{23, 26, 132}. Expression as ubiquitin fusion protein is also a common approach in mammalian cells to improve yields. However, in this case ubiquitin is site-specifically cleaved off by ubiquitin-specific proteases before secretion¹³³. Furthermore, native pro-KLK5 was expressed in yeast (*P. pastoris*) and was found to be active after purification, suggesting that it either autoactivated or was activated by a yeast protease¹³⁴. Also KLK7 was expressed already in active conformation through secreted, periplasmatic expression in *E. coli*³⁴. Given that our lab has no experience with the yeast and insect expression systems, we decided to express the two kallikreins as secreted proteins in mammalian cells using transient transfection.

To obtain good expression yields, the proteases would ideally be kept in an inactive conformation by a pro-domain during expression, which then can be removed during the purification procedure. Whether a protease is capable of autoactivation depends on whether the sequence of the last amino acids of the pro-domain match with the substrate specificity of the protease itself²³. For KLK5, the last two amino acids of the pro-domain (Ser-Arg) perfectly match with its substrate specificity for the S2

and S1 subsites (Ser-Arg), meaning it is likely to be able to autoactivate. In contrast, the last two amino acids of the pro-domain of KLK7 (Asp-Lys) do not match with its substrate specificity for the S2 and S1 (Tyr-Tyr) subsites, indicating that KLK7 cannot autoactivate.

Based on the differing autoactivation capabilities, different expression constructs were initially designed for KLK5 and KLK7. For KLK5, which can autoactivate, the sequences of the whole human pro-protein (UniProtKB Q9Y337, Asn30-Ser293) was used. The DNA sequence additionally encoding a C-terminal poly-histidine tag was cloned into a mammalian expression vector after a BM40 secretion signal (**Figure 25a, pPG001**). For KLK7, which must be activated by a different protease, only the sequence of the mature protease (UniProtKB P49862, Ile30-Arg253) was used while the endogenous pro-domain was entirely replaced by an EK cleavage site (Val-Asp-Asp-Asp-Asp-Lys) to allow manual activation after expression. The DNA sequence additionally encoding a C-terminal poly-histidine tag was also cloned into a mammalian expression vector after a BM40 signal and bovine ubiquitin (bUB), which is also cleaved off during secretion and was expected to improve expression yields (**Figure 25a, pPG002**). However, as described in the next paragraph, the manual activation of ProKLK7 with EK turned out to be expensive and time consuming and the additional purification step to remove EK drastically reduced the yield. Thus, another construct that allowed a one-step purification of autoactivating KLK7, comparable to KLK5, was cloned. To do this, the last two amino acids of the pro-domain of KLK5 were adjusted to Leu-Tyr to match the substrate specificity of KLK7. The pro-domain of KLK5 (4 kDa) was used, as it is larger than the pro-domain of KLK7 (0.8 kDa) and thus the removal of the pro-domain is better detectable by SDS-PAGE. The sequence of mature hKLK7 was then cloned downstream of this engineered pro-domain (**Figure 25a, pPG007**).

3.1.3 Expression, purification and characterization

All constructs could be successfully expressed in CHO suspension cell cultures at 100 mL scale using 300 µg plasmid DNA for transient transfection. After the first Ni-NTA purification, buffer exchange, and concentration 5.4 mg KLK5 could be obtained from pPG001 and 1.5 mg VDDDDK-KLK7 could be obtained from pPG002. After overnight activation of VDDDDK-KLK7 with EK and a second Ni-NTA purification, 0.2 mg active KLK7 could be recovered. KLK7 expressed in 100 mL scale from pPG007 and purified with a one-set procedure on a Ni-NTA column yielded 0.1 mg pure KLK7. To confirm the identity of the expressed proteins by size, they were analyzed by SDS-PAGE (**Figure 25b**). To confirm their activity, specific, literature-reported substrates were used (**Figure 25c**).

For KLK5, up to four bands of different sizes were detected, representing proteins with a varying number of glycosylations. KLK5 has four potential N-glycosylation sites and an unglycosylated molecular weight of 26 kDa. Taking 2.5 kDa for each N-glycosylation into account, up to 5 bands between 26 and 36 kDa were expected. The pro-domain of KLK5 has a molecular weight of 5 kDa and thus inactive KLK5 would have a molecular weight of 31 kDa (plus additional 2.5 kDa for each glycosylation) and would not be distinguishable from mature protein with two or more glycosylations. However, all bands < 31 kDa must be the active form. Thus, all visible bands corresponded to the expected sizes and no additional bands indicated the presence of impurities. A band at 25 kDa indicated that at least a fraction of the protein was mature and active. Specific activity on the substrate Boc-VPR-AMC confirmed the partial autoactivation of the KLK5 zymogen during expression or

purification, making additional activation steps obsolete. Furthermore, a K_m value of 216 μM could be measured on Boc-VPR-AMC, matching the literature value¹³⁴. Based on this characterization, self-expressed hKLK5 was considered suitable for phage display selections and activity assays, and was further used for this work.

KLK7 expressed from plasmid pPG002 has an unglycosylated molecular weight of 26 kDa and one potential N-glycosylation site, which should result in bands with sizes of 26 and 28.5 kDa. Both expected bands were clearly visible. As expected, the purified KLK7 pro-protein with an EK cleavage-site of 1 kDa as synthetic pro-domain did not show any activity. This confirms that the EK cleave-site was not removed by any endogenous mammalian proteases or by autoactivation, keeping KLK7 in an inactive conformation. After overnight activation with EK and a second, subsequent Ni-NTA purification, a decrease in size of around 1 kDa was observed, and activity on the substrate Suc-LLVY-AMC became measurable. However, due to the poor solubility of this substrate in the assay buffer, the K_m value could not be determined. Nevertheless, KLK7 expressed from plasmid pPG002 and characterized for its purity and activity with Suc-LLVY-AMC was used for phage display selections. However, Suc-LLVY-AMC was not used further but exchanged to the absorbance-based, self-synthesized substrate KHLY-pNA, which was identified as good KLK7 substrate by Veer *et al.*³⁷. The synthesis of the substrate is described in section 5.4.6 and shown in **Figure S3**. Analysis of the identity and purity of the synthesized substrate is shown in **Figure S11**.

For all KLK7 inhibition assays, the more easily purified, autoactivating KLK7 expressed from pPG007 was used in combination with KHLY-pNA as a substrate. On the SDS gel, this protein showed the expected band of 25 kDa (unglycosylated) corresponding to the mature protease, but also two bands at 30 and 32.5 kDa corresponding to inactive proprotein (unglycosylated and monoglycosylated). Specific activity on KHLY-pNA, confirmed the partial auto-activation during purification. Additionally, a K_m value of 107 μM could be measured with KHLY-pNA that matched the expected literature value³⁷. The concentration of mature KLK7 was not determined as it is not relevant for the activity assays, as the determination of K_i values is not enzyme concentration dependent.

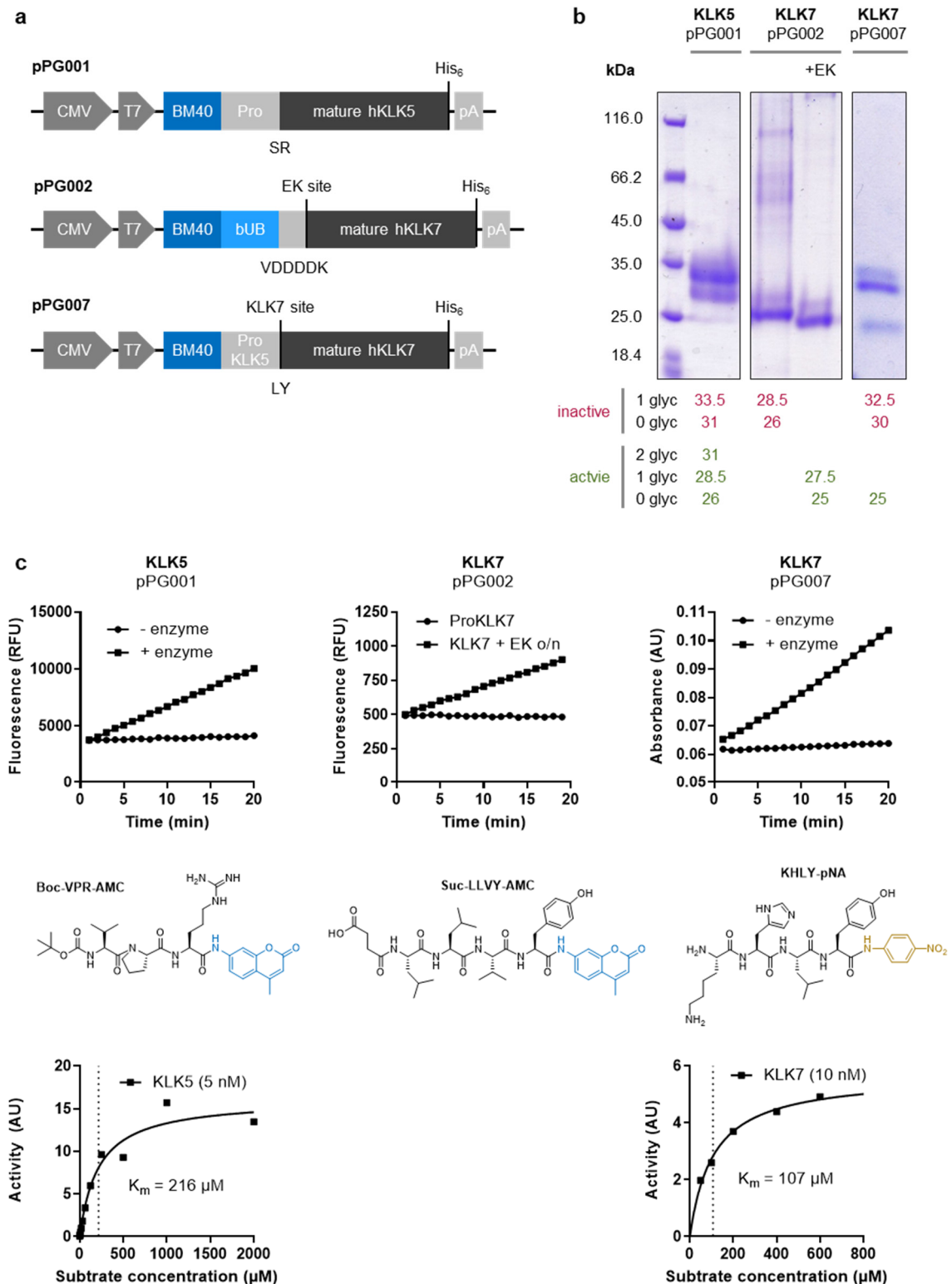


Figure 25 | Expression and characterization of hKLK5 and hKLK7. (a) Constructs used for secreted expression in mammalian cells. hKLK5 was expressed as a self-activating proprotein (pPG001). hKLK7 was expressed as a pro-protein that must be activated with enterokinase (EK; pPG002) and with an engineered pro-domain as autoactivating pro-protein (pPG007). CMV/T7: promoters from cytomegalovirus/T7 RNA polymerase; BM40: membrane protein 40 secretory signal peptide; bUB: bovine ubiquitin; pA: polyadenylation signaling sequence. (b) Reducing SDS gel of expressed proteins. KLK7 from pPG002 is shown before and after activation with EK followed by a second purification. Expected molecular weights are indicated. (c) The activity of expressed kallikreins on specific substrates, and the determination of K_m values.

3.1.4 Biotinylation of KLKs for immobilization

To immobilize the target proteins on magnetic beads as necessary for phage display selection, KLK5 and KLK7 were biotinylated at random primary amines using a 20-fold molar excess of Sulfo-NHS-LC-Biotin. After biotinylation and desalting, a capture assay was performed to confirm that the biotinylation was successful. This was done by testing if the biotinylated proteins bind to streptavidin magnetic beads, but not the negative control protein, which was not biotinylated.

After incubation of biotinylated or unbiotinylated protein with streptavidin magnetic beads, the beads and supernatant were separated. Beads were washed and supernatant and beads were loaded onto a reducing SDS gel. Only the biotinylated protein was detected in the bead fraction, while the unbiotinylated control protein appeared in the supernatant, indicating a successful biotinylation (Figure 26).

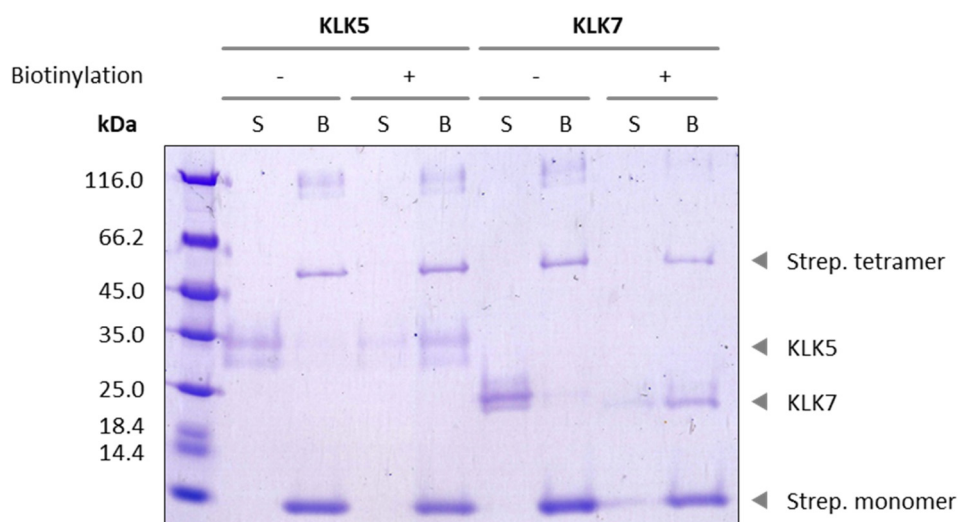


Figure 26 | Capture assay with biotinylated versus unbiotinylated KLK5 and KLK7. 1 µg protein was added to 25 µl streptavidin (Strep.) beads. 'S' shows the supernatant while 'B' shows the bead fraction. KLK5 runs at 35 kDa and KLK7 at 25 kDa. The biotinylated proteins are mostly detected in the bead fraction, while the unbiotinylated proteins remain in the supernatant.

3.1.5 Selection & characterization of bicyclic peptides

We used phage display to identify cyclic-peptide inhibitors of KLK5 and KLK7. A phage encoded library of linear peptides of 18 amino acids of the format $XCX_mCX_nCX_oCX$ ($m + n + o = 12$) developed by Carle *et al.* in our lab (unpublished) was used for the selections (**Figure 27a**). All peptides in this library were designed to contain four cysteines for cyclization with two thiol-reactive linkers based on the novel, double-bridged peptide format recently published by our lab¹⁰¹. This library format offers the possibility to identify double-bridged peptides, but also monocyclic peptides, in case that only one ring of the peptide contributes to target binding.

The first and last cysteines (C_1 and C_4) in this library are fixed at position 2 and 17 while the second and third cysteines (C_2 and C_3) can appear at any position between 3 and 16. This gives a theoretical combinatorial diversity of 1.5×10^{20} . However, the realized diversity after cloning is limited by the transformation efficiency and the vector size and was determined to be 3×10^{10} for the library that was used. This library was encoded in a phagemid as linear peptide N-terminally fused to the PIII protein of the M13 phage. The library was expressed with the help of a hyperphage, leading to multivalent display. On the surface of the phage, the expressed, linear peptides were cyclized with either of the two thiol-reactive linkers 2,6-bis(bromomethyl)pyridine or 1,3-dibromoacetone (**Figure 27b**). As two linker molecules can react with one linear peptide, three different regioisomers form during the cyclization reaction, depending on which of the two cysteines is connected (**Figure 27c**). This expands the genetically encoded diversity by another factor of 3 based on the structural diversity.

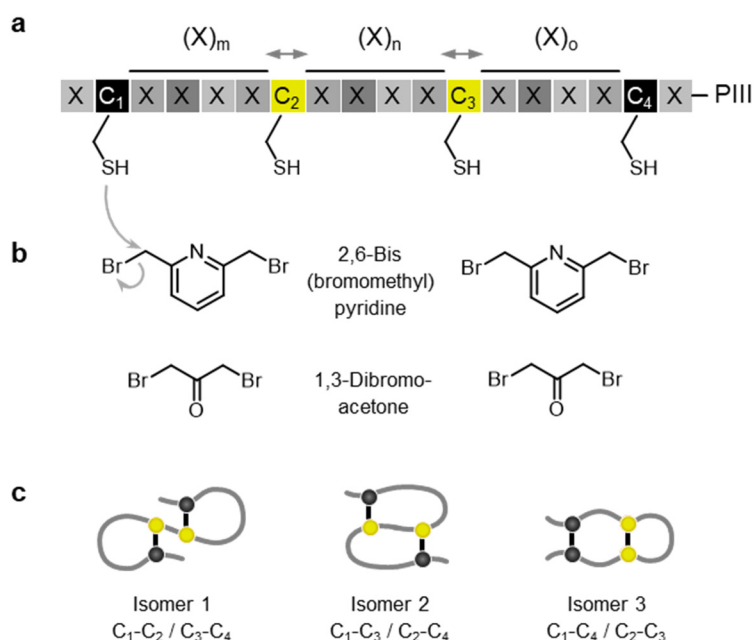


Figure 27 | Phage display of double-bridged peptides. (a) Library format used for the selections. The first and the last cysteine (C_1 and C_4) were fixed at position 2 and 18 while C_2 and C_3 were designed to appear variably between position 3 and 16. (b) Structures of thiol-reactive linkers used for cyclization of the linear peptide library. Two moles of linker react with one mole of peptide. (c) Isomers that potentially form during the cyclization reaction.

Three rounds of phage-display selections against immobilized hKLK5 and hKLK7 (10 µg, 5 µg, 2.5 µg – first, second, third round) were performed with this library. Samples to determine the phage titers were taken at different steps of the phage purification, peptide modification, and phage selection procedure during each of the three selection rounds (**Figure 28**). These titers give valuable information about the success of the phage selection procedure. The ‘L1’ and ‘L2’ titers indicate the number of *E. coli* after library inoculation and hyperphage infection and thus give information of how much of the library diversity is covered in the culture. The ‘S0’ titer shows the amount of phages in the supernatant of the culture medium after overnight phage production and indicates how often each clone of the library is represented. The following phage titers ‘S1’ to ‘S3’ follow up the phage titers through the purification and modification to ensure that the amount of infective phages is not severely reduced by any of the steps: ‘S1’: phage titer after PEG precipitation from the supernatant, ‘S2’: phage titer after phage reduction with TCEP and precipitation; ‘S3’: phage titer after peptide cyclization and precipitation. The ‘S4’ titers finally indicate the number of phages that were captured and eluted from the beads after target panning. To ensure the majority of phages were captured by the target protein and not by the beads or the plastic tubes, negative control selections using beads without bound target were performed during the second round.

The titer ‘L2’ of the first round indicates that 2.5×10^{10} phage producing cells were inoculated and successfully infected with hyperphage, meaning that 2.5×10^{10} peptides of the library were represented in the selections (**Figure 28a**). Negative control selections with beads not bound to any target were performed during the second selection round. An enrichment between 385-fold (KLK5 with 2,6-bis[bromomethyl]pyridine) and 5000-fold (KLK7 with 1,3-dibromoacetone) could be observed, indicating successful, target specific selections (**Figure 28b**). Positive capture titers were target- and linker-specific, which confirms target specificity and indicates successful selections. As a target specific enrichment could already be confirmed during the second selection round, no negative control selections were performed during the third round. Furthermore, the titers confirmed that the loss of infective phages during the preparation procedure did not exceed one order of magnitude in any of the three selection rounds.

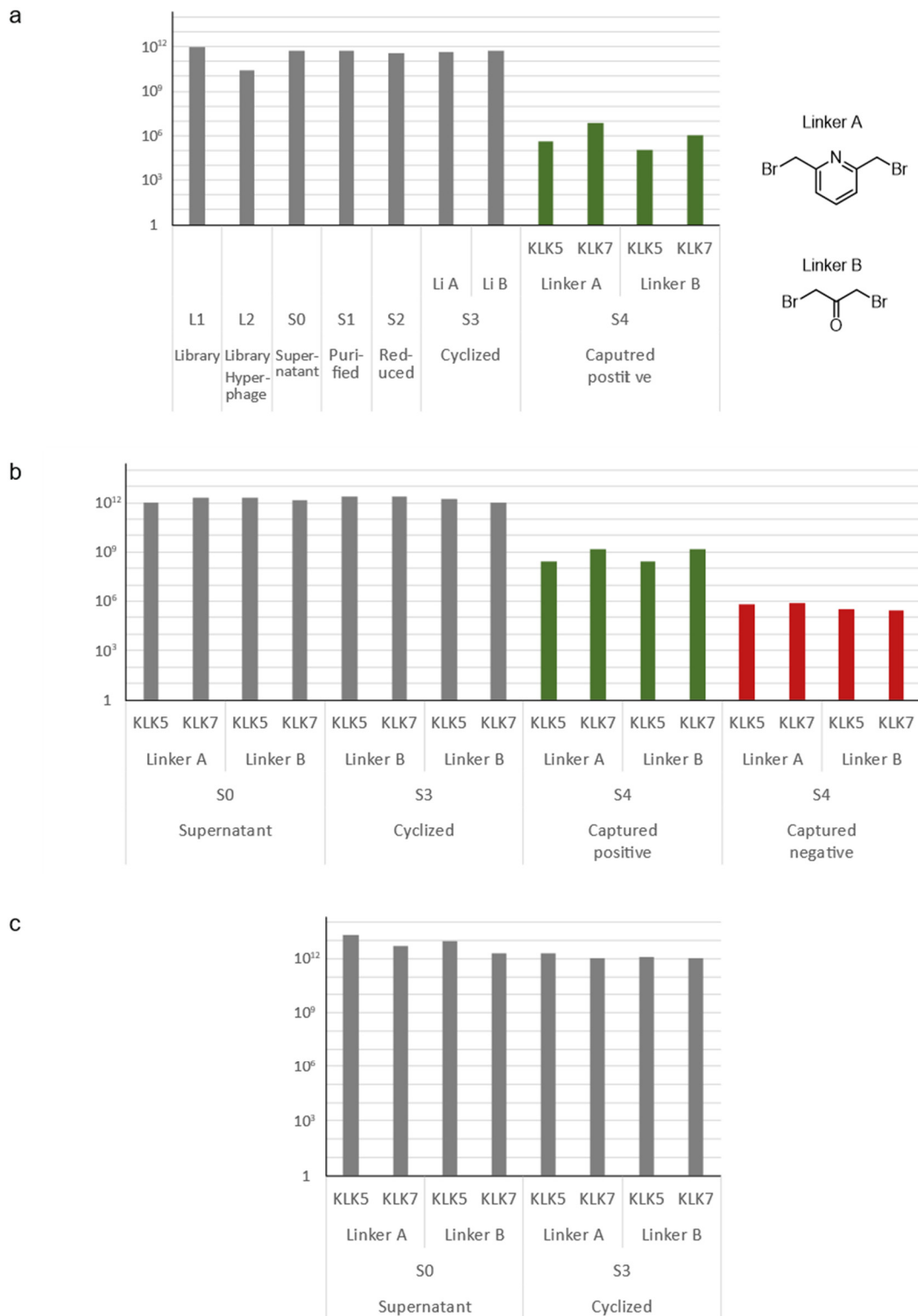


Figure 28 | Titers of phage display selections. (a) First selection round. **(b)** Second selection round. Negative selections were performed with target-free beads to confirm phage enrichment through target binding. **(c)** Third selection round. Linker A: 2,6-bis(bromomethyl)pyridine; linker B: 1,3-dibromoacetone; L1: *E. coli* titer after library inoculation and growth until $OD_{600}=0.5$; L2: *E. coli* titer after hyperphage infection and addition of second antibiotic; S0: phage titer in supernatant of 2xYT medium after overnight phage production; S1: phage titer after PEG precipitation from the supernatant, S2: phage titer after phage reduction with TCEP and precipitation; S3: phage titer after peptide cyclization and precipitation; S4: Titers of phages captured and eluted from the beads after target panning.

As the titers indicated the successful selection of target specific binders, the sequences of 96 randomly selected clones for each target and linker were identified by Sanger sequencing (**Figure 29**). For KLK5 a single consensus group per linker could be identified, while for KLK7 two or three consensus groups per linker were found. For KLK5, threonine followed by arginine was common in all sequences, while for KLK7 leucine followed by tyrosine appeared in most of the sequences. Both motifs match with the reported substrate specificities for the S2 and S1 subsites of KLK5 and KLK7, respectively, which implies active site binding with a standard mechanism of inhibition. Furthermore, it can be observed that a consensus in peptides selected against KLK5 can only be found N-terminally between the first and second cysteine.

To assess the activity of the found peptides, three randomly selected bicyclic peptides per target and linker were synthesized on 25 μ mol scale using 9-fluorenylmethyloxycarbonyl (Fmoc) solid-phase peptide synthesis (SPPS). The synthesis scheme is shown in **Figure S3** and the detailed protocol can be found in section 5.4. The peptides were not synthesized as pure regioisomers, but if possible, the three isomers were separated during the second RP-HPLC purification. The yields of the pure compounds were between 1 – 5 mg per isomer, corresponding to yields between 12 - 60%. However, not for all peptides all three isomers could be separated, meaning they either did not form during the cyclization reaction due to steric reasons or they did not elute separately from the column. For this reason, the purity of the peptides could not determine (**Figure S12**). Thus, one must consider that the measured activity of the peptide might be not precise, for examples if the isomers could not be separated at all, the measure activity values could deviate by a factor three. For this reason, the measured activity values were only taken as indicative and differences of less than one order of magnitude were not considered. The inhibitory activity of the peptides was tested in substrate based-activity assays. The inhibitory constant (K_i) values of the best isomer or fraction that could be separated is shown next to the corresponding peptide sequences (**Figure 29**). With K_i values as low as 2 nM, inhibitors in the low nanomolar range could be successfully identified for KLK5 (peptide 001, 002, 004). With K_i values as low as 9 nM, inhibitors in the low-nanomolar range were also found for KLK7 (peptide 011). The high potencies of these bicyclic inhibitors confirm the successful phage display selections, and the identified inhibitors provided a good starting point for the development of KLK5- and KLK7-targeted therapeutics.

As the consensus sequence covers less than half of the 18 amino acids of the peptides, great parts of each peptide probably do not contribute to target binding. We have previously shown that the peptides of one consensus group are selected as any of three possible isomers and that this selected isomer shows the strongest inhibitory activity¹⁰¹. However, as the cysteine connectivity is not genetically encoded, it has to be unraveled by synthesizing all three isomers using orthogonal cysteine protecting groups and testing their activity. For all obtained consensus groups, it is apparent that the consensus sequence was located between two cysteines, of which one is a fixed cysteine and the other a variable cysteine. This suggests that the two, consensus surrounding cysteines are connected, and peptides were selected as isomer one (C_1 - C_2 and C_3 - C_4 connected). Isomer one represents the least constrained scaffold, with two, separate cycles. As a consensus sequence can only be found in one cycle, it seems that only one cycle provides the major binding interaction while the rest of the peptide does not contribute considerably to target binding. Another indicator for this was, that the second variable cysteine in each sequence appeared in random positions. This suggests, that monocyclic peptides might suffice for target binding and provide the same inhibitory activities as the synthesized bicyclic peptides.

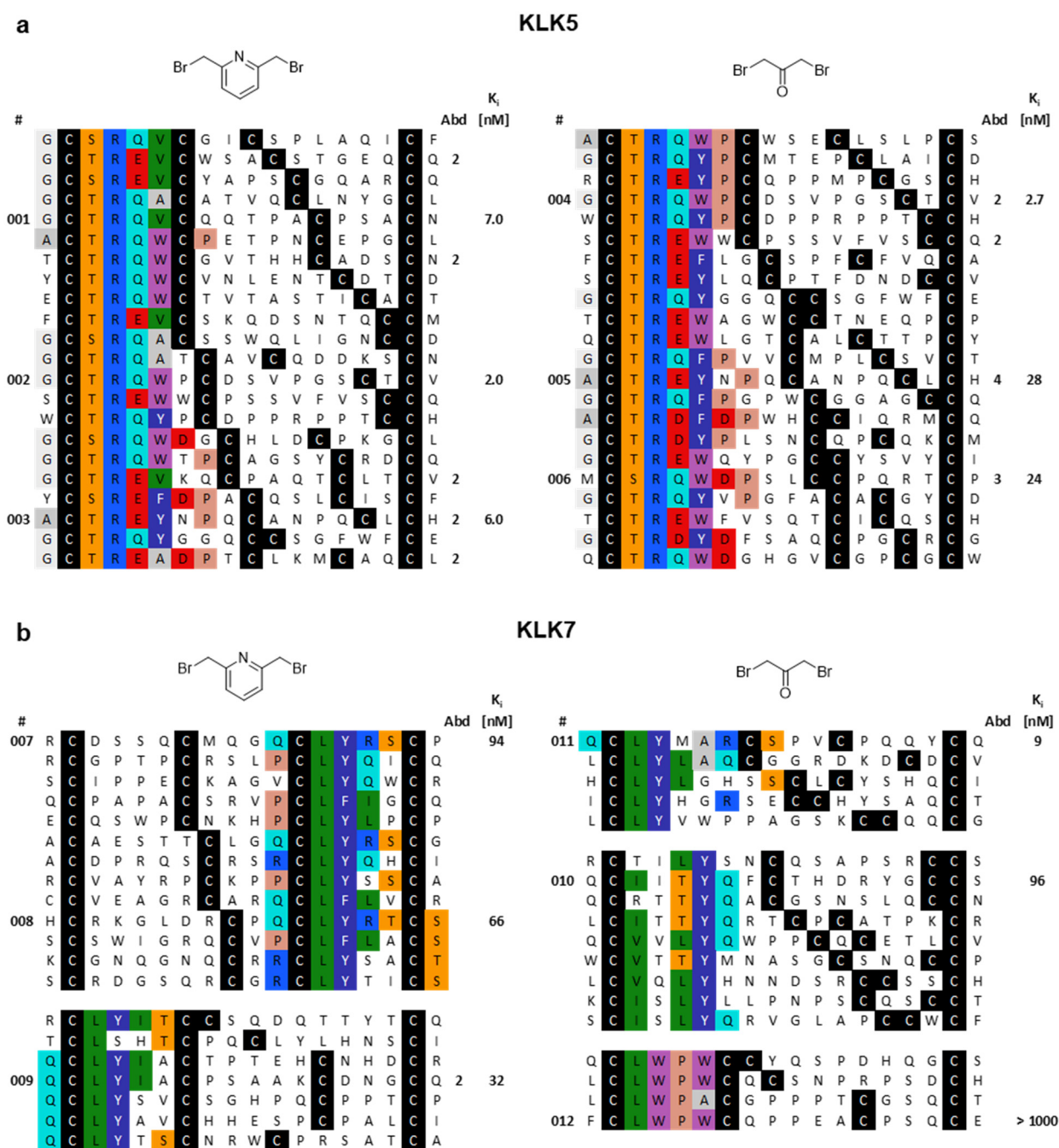


Figure 29 | Sequences of phage-selected double-bridged peptides. Amino acid sequences found after three rounds of selections with an 18 amino acid linear peptide library cyclized with either 2,6-bis(bromomethyl)pyridine or 1,3-dibromoacetone. **(a)** Sequences found after selections against hKLK5. **(b)** Sequences found after selections against hKLK7. Sequences that show a consensus are shown in groups and similar amino acids are highlighted using the Rasmol color code. Inhibitory constant (K_i) values of selected, synthesized peptides are shown as the mean of three measurements next to the corresponding peptide sequences. Identity and purity analysis of the synthesized peptides is shown in Figure S12.

Based on the hypothesis that all binders were selected as isomer one (C_1 - C_2 and C_3 - C_4 connected), we analyzed the ring sizes of the consensus containing monocycles. Even though various ring sizes appeared in the sequences, the minimal ring size that could be found was linker specific. It was apparent from the consensus for KLK5 as well as for KLK7 that the smallest cycles selected with 2,6-bis(bromomethyl)pyridine contained six amino acids in the ring, while when selected with 1,3-dibromoacetone they contained at least seven amino acids. It has been described previously that macrocycles with similar binding motifs often have comparable ring sizes, which suggests a correlation between a certain binding motif and its cycle topology¹³⁵. This would explain that with 2,6-bis(bromomethyl)pyridine as the linker no cycles with less than six amino acids were selected, while with 1,3-dibromoacetone, which is two carbons shorter, the smallest ring size was seven amino acids.

Based on this assumption, the consensus sequences of the smallest monocyclic peptides, which most likely are sufficient to provide the major binding interactions, were derived from the initially discovered sequences. To understand which amino acids are tolerated in the different positions in this smallest found macrocycles depending on target and linker, we visualized them as sequence maps (Figure 30). We were particularly interested in the smallest possible cycles, as in contrast to the bigger cycles, most of the positions in these cycles showed conserved amino acids. Additionally, their smaller size makes them more attractive for drug development than their bigger, initially selected, double-bridged progenitors, as their synthesis is much easier. Thus, they can also be easier conjugated and combined with other moieties, which was important for the further development.

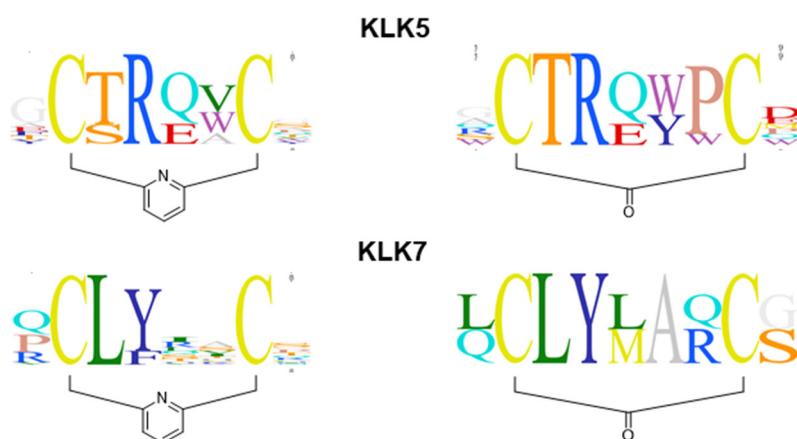


Figure 30 | Sequence maps of consensus monocycles. Based on the assumption that the consensus surrounding cysteines were connected, sequence maps of the smallest identified monocycles were generated.

3.2 Structure activity relationship and affinity improvement

3.2.1 Development of monocyclic peptide inhibitors

To validate the hypothesis that only one macrocycle of the peptide mediates target binding, several monocyclic inhibitors of KLK5 and KLK7 were synthesized (**Figure 31a**). If different amino acids appeared at one position in the consensus, based on the consensus maps (**Figure 30**) peptides with these variants were synthesized to assess which amino acids provided the best activity. Additionally, we tried to analyze the impact of the two extra-cyclic positions, by synthesizing peptides lacking these amino acids (peptide 108, 109 and 114). For the KLK5 inhibitors, where the cycle size was not as clearly conserved as for the KLK7 inhibitors, several peptides with increasing cycle size were synthesized to assess the impact of cycle size on affinity (peptide 031, 051, 034 and 054).

Peptides could be successfully synthesized on 25 μ mol scale with high purities of more than 95% for all peptides (**Figure S13**) except of peptide 081, where an N-terminal glutamine formed pyroglutamate for around 50% of the product. The synthesis scheme is shown in **Figure S3** and the detailed protocol can be found in section 5.4.

With K_i values of 2.2 nM for KLK5 (peptide 035) and 16 nM for KLK7 (peptide 096), monocyclic peptides inhibited their targets with a similar efficacy as their bicyclic progenitors, while having a much smaller molecular weight of only 1 kDa versus 3 kDa for the bicyclic peptides.

For both targets, peptides based on 2,6-bis(bromomethyl)pyridine showed inhibitory activities in the same range as peptides based on 1,3-dibromoacetone. As cyclic peptides based on aromatic linkers are patent protected¹³⁶, we decided to only continue the development of 1,3-dibromoacetone based peptides for intellectual property reasons. Nevertheless, the development of inhibitors based on two different linkers was valuable until this stage as it provided entirely independent controls, and as the discovered peptides are even similar, the information provided seems to be transferable between the inhibitors based on the two different linkers – in particular for position one to five.

The structure-activity relationship (SAR) analysis provided information that allowed us to narrow in on the best inhibitor based on the consensus sequences. In the fifth position of 1,3-dibromoacetone-based KLK5 inhibitors, glutamine had better K_i values than glutamic acid. Additionally, glutamine was favored due to its neutral charge. In the sixth position, tyrosine outperformed tryptophan. Inhibitors with a C-terminal extra-cyclic amino acid or a bigger ring size did not significantly improve the inhibitory activity (1.15-fold and 1.47-fold, respectively) and thus peptide 035, including a glutamine, a tyrosine, and the smaller ring size to keep the molecular weight low, was chosen as the lead for the further development of KLK5 inhibitors (**Figure 31b**).

For the KLK7, several amino acids were compared, especially in the first, fifth, and sixth position. In the first, N-terminal, extra-cyclic position, glutamine and arginine had better K_i values than proline. Nevertheless, proline was chosen for the lead peptide, as arginine could lead to stability issues, while a free, N-terminal glutamine can cyclize with the free $-NH_2$ to form pyroglutamate and thus would cause purity problems during synthesis. In position five, leucine displayed the best activity after methionine, which was avoided due to oxidation reasons. In position six, which was not very well conserved in the sequences, serine, threonine and alanine were compared. With both linkers, alanine provided a great improvement in activity compared to the polar amino acids. Peptide 096, which has

as an extra-cyclic, C-terminal serine showed a 2-fold better activity than peptide 114 without serine, and thus peptide 096 was chosen as lead for the further development of KLK7 inhibitors (**Figure 31b**).

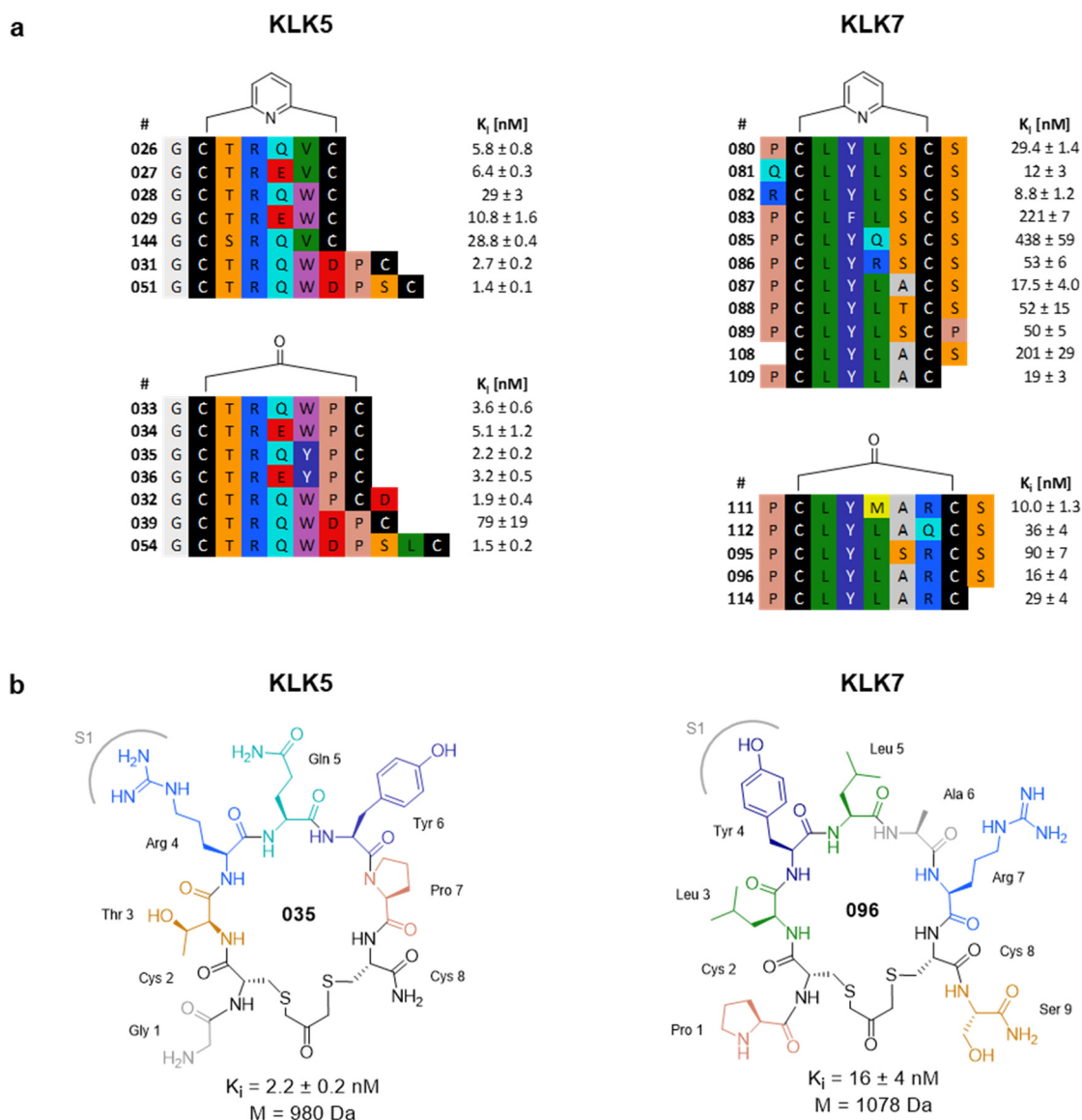


Figure 31 | Development of monocyclic inhibitors. (a) Peptides were synthesized based on amino acids found in the consensus sequences. In contrast to the first generation, they contain only two cysteines cyclized with one linker. The K_i values are shown as mean of three measurements \pm SD. **(b)** The structures of peptides 035 and peptide 096, which were chosen as leads for further development. All structures and purities of the compounds can be found in Figure S13.

3.2.2 Affinity improvement with unnatural amino acids

To exploit the so far untouched chemical space of non-natural amino acids for affinity improvement, peptides with non-natural side-chain derivatives in one position were synthesized and tested on their impact on inhibitory activity. Unnatural derivatives were identified using SciFinder (<https://scifinder.cas.org>), searching for commercially available compounds with up to 90% similarity to the N-Fmoc-protected, natural derivative. Based the output of this search, derivatives were manually chosen based on availability and price. The structures of the chosen and tested unnatural amino acids are shown in **Figure 32c**. Additionally to the unnatural derivatives, for the KLK5 inhibitor Ser and Lys were tested in position 3 and 4, respectively. Peptide 035 served as the reference for the KLK5 inhibitor and the side chain of one residue was changed at a time. For the KLK7 inhibitor, we used peptide 114 and not lead peptide 096 as reference because the C-terminal serine of peptide 096 had to be removed later for stability reasons (see 3.3.3).

As for the natural monocyclic inhibitors, the peptides were synthesized on 25 μ mol scale. The synthesis scheme shown in **Figure S3** and the detailed protocol can be found in section 5.4. High purities of more than 95% could be obtained for almost all peptides with unnatural side-chain derivatives (**Figure S14**).

For KLK5, no derivatives providing any further improvement in inhibitory activity could be identified (**Figure 32a**). As it was already partially apparent from the consensus, changes in positions three and four, which provide the non-primed side residues, led to big losses in the inhibitory activity, while positions five and six, which provide primed side residues, are more tolerant to structural changes. Position seven is particularly sensitive to mutations to D-amino acids and therefore probably has an important impact on the topology of the macrocycle.

Also for the KLK7 inhibitor, positions three and four were highly sensitive to structural changes, while positions five and six, were more tolerant. In position five, one derivative that improves the inhibitory activity was found: Norleucine exploits the chemical space between the previously used leucine and methionine, which also appeared in the consensus at position five but was avoided due to oxidation concerns and improves the K_i value to 7 ± 0.7 nM (**Figure 32b**).

This part of the SAR study confirmed that our developed KLK5 and KLK7 inhibitors both exploit a standard mechanism of inhibition, binding the active site like a substrate. The results confirmed the assumption made base on the phage display results, that Arg 4 and Tyr 4 provide the P1 residues in the KLK5 and KLK7 inhibitor, respectively, as they were the residues which were most conserved in the phage display consensus, and also the residues most sensitive to structural changes in the affinity maturation screen. Furthermore, this means, that also the other respective residues of the two inhibitors most likely bind to homologues subsites in the two proteases. The identified residues match well with previously in literature described subsite preferences of KLK5 and KLK7 and thus allow a clear assignment of most of the side chains to the subsites of their targets.

For KLK5, a strong preference for P1-Arg over Lys is described in literature. For P2, in literature a preference for small or polar residues such as Ser, Ala, Thr and Asn is described²⁸. We found P2 as the second most conserved residue in the phage display sections with a strong preference of Thr over Ser, with Ser leading to a 5-fold loss in affinity. Also in literature described is a preference for Gly in P4, which matches with the found Gly in the N-terminal, extra-cyclic position, later Gly 1 in the KLK5 inhibitor. Veer *et al.* describe GFSR-pNA as the best KLK5 substrate from their screen with tetrapeptide

substrate libraries and suggest an involvement of the free N-terminus of the substrate in KLK5 binding³⁶.

KLK7 prefers Tyr followed by Ala and has a low preference for Phe, Arg, and Lys according to literature²⁸. We could confirm a clear preference of Tyr over Phe, which appeared in all binders and thus allows the clear assignment of Tyr 4 in the KLK7 inhibitor as P1 residue, however no peptides with Ala or Met as potential P1 residue were found. Phe as P1 residue led to a 7.5-fold loss in affinity compared to P1-Tyr. For the S2 pocket a preference for medium-sized polar and hydrophobic residues, such as Tyr, Leu, Met, and Thr is described and thus seems rather unspecific according to literature. In our selections, most binders had Leu as P2 residue. In the context of our further developed inhibitor also Neopentylglycine was tolerated, indicating some tolerance in this subsite. In another consensus group, which was not further followed, Thr appeared equally frequent as Leu, however Met or Tyr were not found. Veer *et al.* described KHLY-pNA as the best KLK7 substrate from their screen³⁷, matching very well with the preferences for P2-Leu and P1-Tyr observed by us.

Our data confirm and further sharpen the understanding of the subsite preferences of KLK5 and KLK7. The found scaffold, which was four-times independently selected with two different linkers and two different targets, proves the suitability of the technology published by Kale *et al.* to identify target-tailored scaffolds with phage display¹⁰¹.

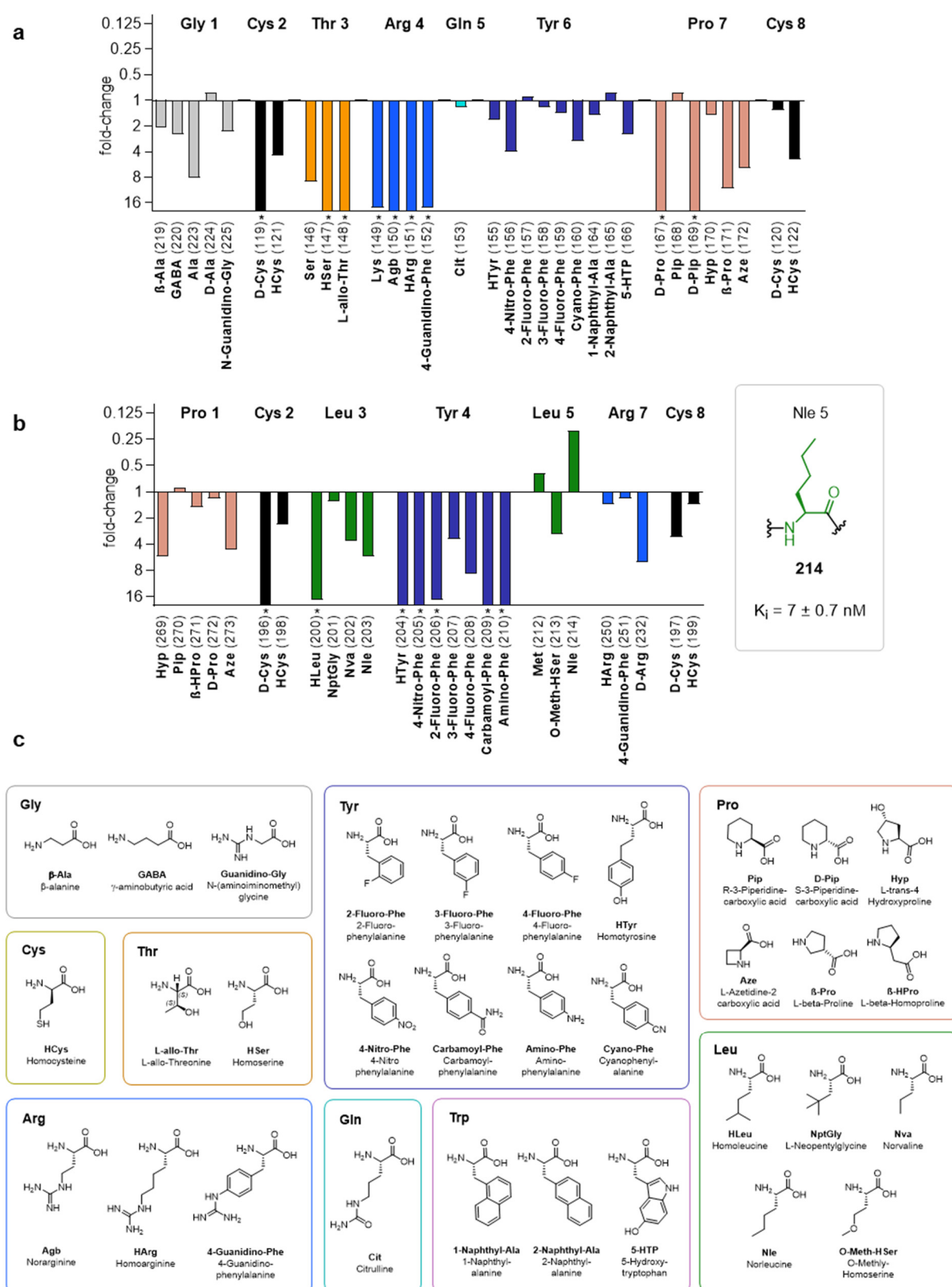


Figure 32 | Affinity improvement with unnatural amino acids. Macrocycles with single unnatural amino acid substitutions were synthesized and tested. The K_i values relative to the reference peptide from single measurements are shown. **(a)** For KLK5, 035 served as the reference. **(b)** For KLK7, 114 (not lead peptide 096) served as the reference because the C-terminal serine of peptide 096 had to be removed later for stability reasons. **(c)** Structures of tested unnatural derivatives. * fold-change > 16. Complete structures and purities of the compounds can be found in Figure S14. Raw data in Figure S5.

3.2.3 Selectivity

Even though phage-selected peptides are in general specific for their targets and usually provide selective inhibitors, we wanted to confirm this for our molecule. We tested the inhibitory activity of our lead KLK5 inhibitor in substrate-based activity assays against a group of various trypsin-like serine proteases that were available in the lab.

None of the seven tested proteases inhibited KLK5 at nanomolar concentrations, demonstrating a high selectivity of the KLK5 inhibitor (**Figure 33**). FXIIa showed a weak off-target inhibition with a K_i value of 40 μ M, which is around 4300-fold weaker than the K_i value for KLK5. Furthermore, trypsin is inhibited at almost the same concentration, which shows that this is not a specific inhibition, but already in the concentration range where many trypsin-like serine proteases might be unspecifically inhibited by arginine containing peptides.

This shows that our lead KLK5 inhibitor is selective over distantly related proteases. The specificity profiling within the kallikrein family has not been performed yet. To get an idea of the selectivity over their closest phylogenetic homologues, the KLK5 inhibitor should be tested against hKLK4 and hKLK14 while the KLK7 inhibitor should be tested against hKLK10 and hKLK12 (**Figure S1**). However, as the aforementioned, closely related kallikreins have no essential functions in the body or are, as KLK14, even involved in the epidermal proteolytic cascade, this profiling was not given high priority.

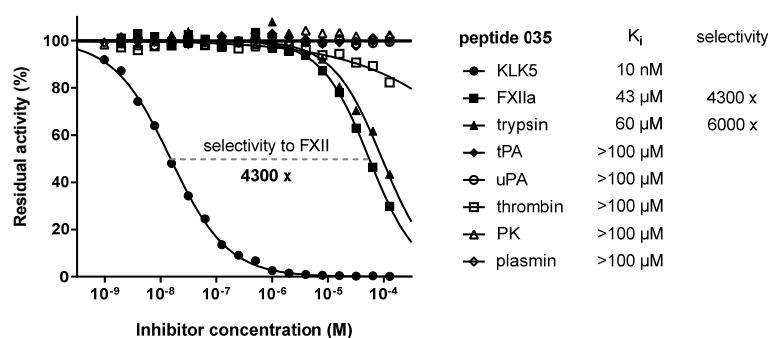


Figure 33 | Selectivity of KLK5 inhibitor. The selectivity of the lead KLK5 inhibitor 035 to other human trypsin-like serine proteases was tested in activity assays. For each protease, their specific substrates were used.

3.3 Half-life extension and pharmacokinetics

3.3.1 Conjugation to albumin tag

Even though peptides furnish potent binders, they typically suffer *in vivo* from short half-lives due to fast renal clearance. To still profit from their enormous potential for therapeutic applications, several strategies to improve their half-life have been applied to peptide-based molecules (section 1.2.5).

We chose non-covalent albumin binding as half-life extension strategy for our tissue kallikrein inhibitors, the rationale being that the inhibitors remain small in size and thus promise a good diffusion into tissue. To equip our inhibitors with an albumin binding moiety, we decided to use an albumin-binding ligand recently developed in our lab (called “albumin tag”), which is based on a fatty acid peptide chimera, has a molecular weight of 2 kDa, and binds human serum albumin (HSA) with low nanomolar affinity of 50 nM^{116, 137} (Figure 34a).

Conjugation of inhibitors to an albumin ligand might affect the binding constants of the inhibitors. We observed in the past, for example with a FXII inhibitor, that conjugation or binding to high molecular weight proteins like albumin can dramatically reduce their activity¹³⁷. One reason might be that an

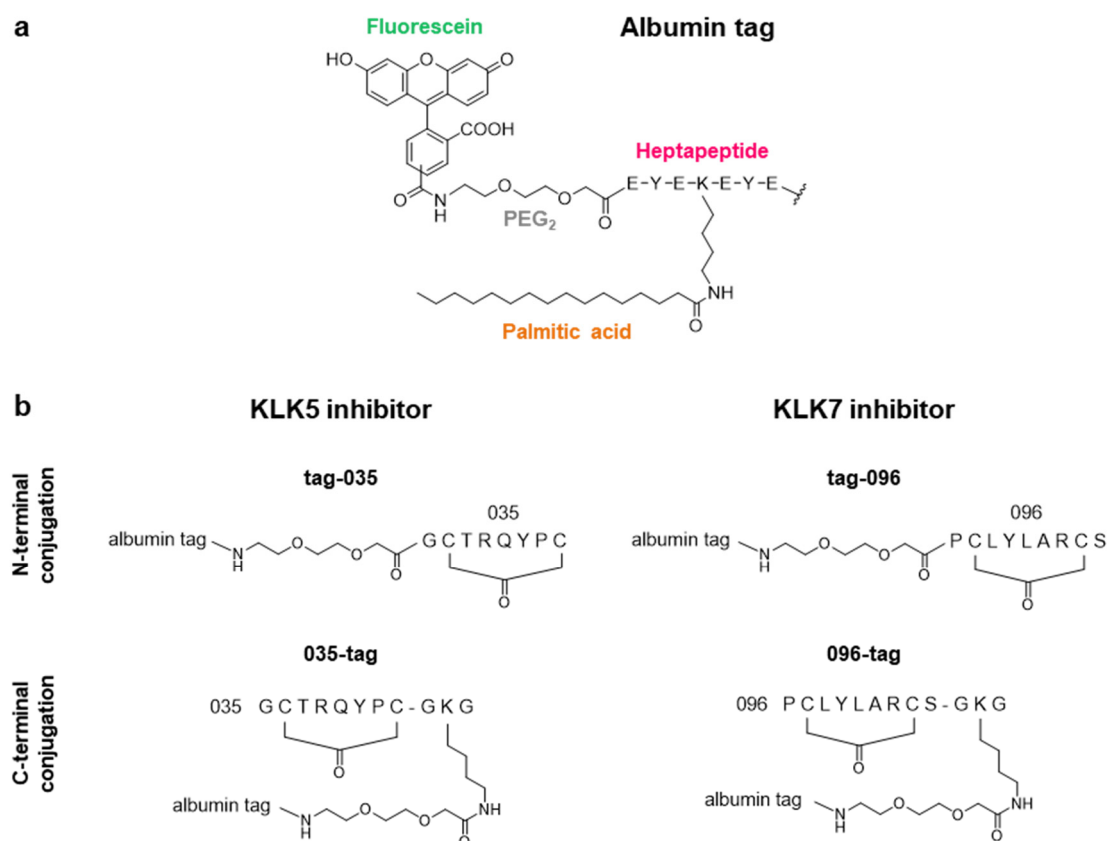


Figure 34 | Conjugation strategies for albumin tag. (a) Structure of the used albumin binding ligand called “albumin tag”, which is based on an acylated heptapeptide. (b) Two different conjugation strategies were compared: For N-terminal conjugation the albumin tag was attached through a PEG₂ linker. C-terminal conjugation was realized through sidechain coupling to a lysine in a C-terminally attached Gly-Lys-Gly linker. Complete structures and purities of the compounds can be found in Figure S15.

increase in molecular weight reduces the diffusion rate and thus affects the k_{on} value of an inhibitor. Furthermore, the increase in size might have steric impacts or intermolecular interactions with albumin can also impair the binding constants. While an increase in molecular weight is difficult to circumvent, steric impacts can be minimized by choosing the best site for conjugation.

We decided to test two different conjugation strategies and to compare which results in molecules with better properties. The KLK5 inhibitor 035 and the KLK7 inhibitor 096 were conjugated N-terminally and C-terminally through the sidechain of a lysine to the albumin tag (**Figure 34b**).

To synthesize the conjugates, I used the previously in our lab by Zorzi *et al.* established synthesis protocol based on Fmoc SPPS¹¹⁶ (**Figure 35**).

N-terminal conjugation was realized by first synthesizing the linear inhibitor and continuing the synthesis on its free N-terminus. First, a PEG₂ unit was introduced as spacer to the tag. Then the heptapeptide sequence of the tag was attached. Dde (N-(1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl)) was used as protection group for the side chain of the Lys within the heptapeptide as it can be selectively removed using 1% hydrazine. However, Fmoc is not stable under these conditions, thus the N-terminus must be protected with protection groups stable in 1% hydrazine such as Boc (*tert*-butoxycarbonyl). For the synthesis of the N-terminal conjugates this was not necessary as 5(6)-carboxyfluorescein (5(6)-FAM) was coupled to the N-terminus. Next, the Dde on the side chain of the Lys was selectively removed and palmitic acid coupled on this primary amine. Then the molecule was cleaved off from the solid phase and cyclization of the two free Cys with 1,3-dibromoacetone was performed on the cured product in solution. Then the peptides were purified by RP-HPLC.

For synthesizing C-terminal conjugates, the inhibitors were synthesized with a C-terminal Gly-Lys-Gly linker to attach the tag on the side chain of the Lys. Thus, Lys with Dde-protected side chain was used. As last amino acid at the N-terminus, N-Boc protected amino acids were introduced. Then the side chain of the Lys within the Gly-Lys-Gly linker was selectively deprotected with hydrazine and the synthesis of the tag continued until 5(6)-FAM. Again, Lys(Dde) was integrated in the middle of the heptapeptide. In the last step, the side chain of this Lys was selectively deprotected and coupled with palmitic acid. After cleavage, cyclization with 1,3-dibromoacetone was performed in solution and the peptides were purified by RP-HPLC.

The inhibitors were synthesized on 50 μ mol scale and around 15 mg pure peptide could be obtained, giving a satisfactory yield of around 10% regarding the length of the synthesis. Purities of more than 95% could be reached (**Figure S15**). For more details on the synthesis protocol see section 5.4.

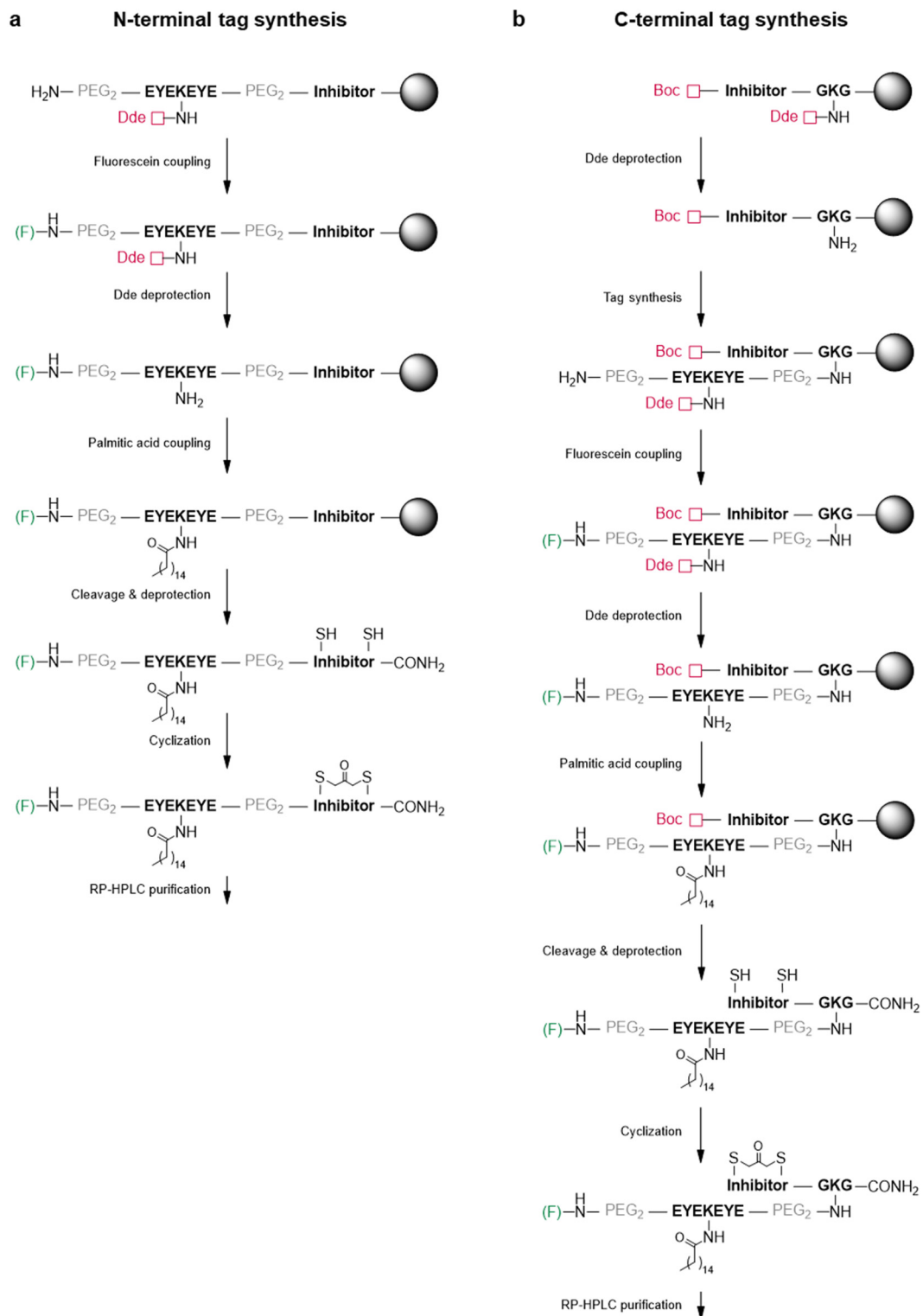


Figure 35 | Synthesis of albumin tag inhibitors. Fmoc SPPS was used. **(a)** Synthesis of N-terminal conjugates. After linear synthesis until the N-terminal 5(6)-carboxyfluorescein (F), the side chain of the Lys within the heptapeptide sequence was selectively deprotected and coupled with palmitic acid. After cleavage, cyclization was performed in solution. **(b)** Synthesis of C-terminal conjugates. The inhibitor was synthesized with a N-Boc protected N-terminal amino acid. The side chain of the Lys within the GKG linker was selectively deprotected and the synthesis of the tag continued until 5(6)-carboxyfluorescein (F). In the last step the side chain of the Lys within the heptapeptide sequence was selectively deprotected and coupled with palmitic acid. After cleavage, cyclization was performed in solution and peptides were purified by RP-HPLC.

After successful synthesis of the molecules, the K_i values for KLK5 or KLK7 in the presence of 25 μM HSA were determined (**Figure 36a**). At this albumin concentration more than 99% of the inhibitor is bound to albumin (at the highest inhibitor concentration), so that the K_i value realistically represent the inhibitory activity of the molecule in the presence of albumin. Even though the actual concentration of albumin in human blood is at around 600 μM and thus much higher, a higher albumin concentration in the assay was not possible, as it interfered with the protease activity.

For the N-terminal conjugates a K_i value of 23 ± 5 nM was measured for KLK5 and 106 ± 12 nM for KLK7. In comparison to the cyclic peptide alone, this means losses in the affinity of both inhibitors with a particularly strong, 10-fold loss observable for the KLK5 inhibitor and a 6.6-fold loss for the KLK7 inhibitor. In contrast, C-terminal conjugation preserved the affinity of both inhibitors at 1.2 ± 0.01 nM for KLK5 and 32 ± 6 nM for KLK7. Regarding the SAR, this confirms the contribution of the free $-\text{NH}_2$ of the N-terminus to target binding, particularly for the KLK5 inhibitor. This assumption was first made based on the consensus sequences and could already be confirmed by the affinity optimization study. What seems rather surprising is that instead of the expected losses in inhibitory activity, a 1.8-fold improvement in the K_i value could be noted for the KLK5 inhibitor. A viable explanation for this improvement is that the peptide backbone of the C-terminally attached Gly-Lys-Gly linker or the tag contributes to KLK5 binding.

Previous work showed, that conjugation also can lead to losses in the affinity for albumin. Thus, the dissociation constants (K_d) of the molecules for HSA were tested with a fluorescence polarization (FP) assay using the fluorescein moiety in the albumin tag. The fluorescence anisotropy of the molecules in a dilution series of human serum was measured and K_d values calculated as described in 5.6.3 (Equation 5). The K_d value for HSA of all conjugates could be mostly preserved, independent of the conjugation strategy (**Figure 36b**). With acceptable 2.1 and 2.9-fold losses, the C-terminally conjugated KLK5 and KLK7 inhibitors, reached final K_d values of 119 ± 9 nM and 164 ± 12 nM for HSA, respectively. The losses in affinity seem to be more dependent on the conjugated inhibitor than on the terminus that was used for conjugation. This might be due to intramolecular interactions of amino acids in the inhibitor with the tag. Hydrophobic amino acids like in the KLK7 inhibitor might interact with the fatty acid moiety, which could slightly impair albumin binding.

Based on these results for both inhibitors the C-terminus was identified as suitable for conjugation without losses in affinity and thus the C-terminally conjugated inhibitors were chosen for further development.

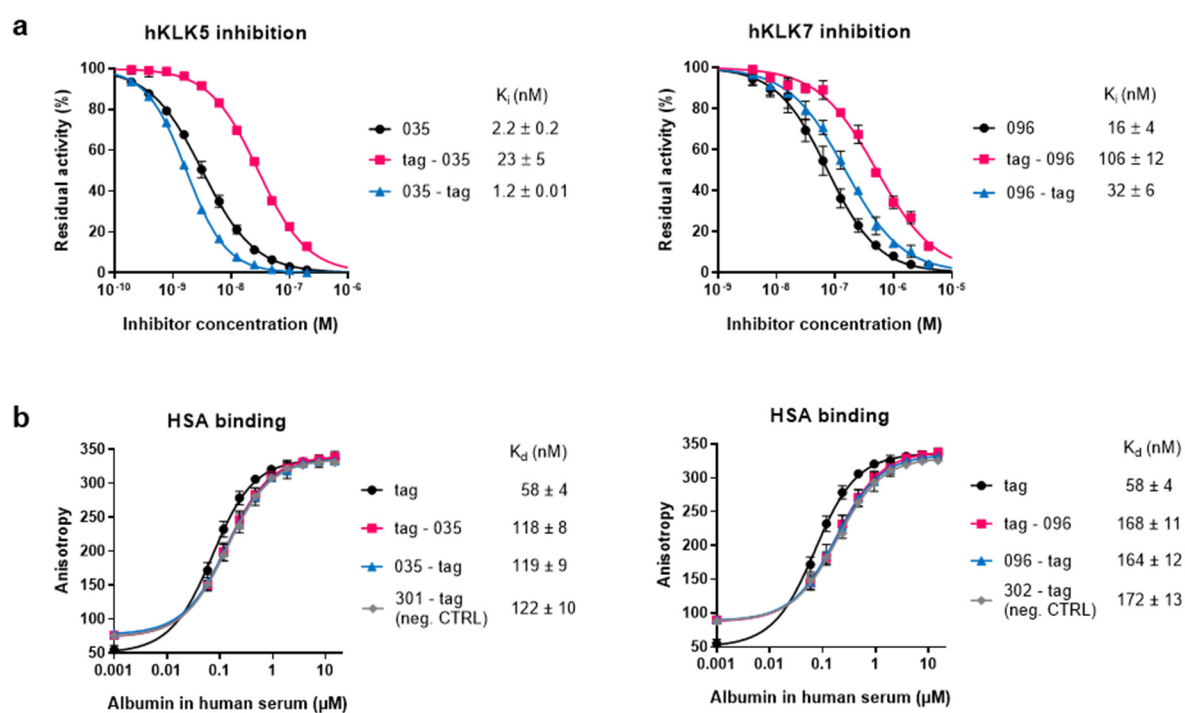


Figure 36 | Affinities of albumin tag conjugates. (a) K_i values of N- and C-terminal tag conjugates in comparison to unconjugated inhibitors measured by inhibitor assays in the presence of HSA. **(b)** K_d values of conjugates for HSA assessed by fluorescence polarization. A dilution series of human serum was used. Additionally to the peptides described above, the peptide 301-tag and 302-tag were tested, which are negative controls, further described in section 3.5.2. All data represent the mean of three measurements \pm SD.

3.3.2 Plasma stability of inhibitors

The metabolic stability of the inhibitor-tag conjugates in human blood plasma was tested *in vitro* and compared to the original inhibitors without tag (**Figure 37**). The peptides were added to human blood plasma at a final concentration of 80 μ M and were incubated at 37°C. The relative concentration of intact peptide was assessed over time by LC-MS. Both inhibitors without the albumin tag showed a very short plasma half-lives in the range of minutes. In contrast, the half-life of both inhibitors with albumin tag was greatly prolonged. The KLK5 inhibitor 035-tag had a half-life of 90 h, meaning that the attachment of the tag extended it more than 1,000-fold (**Figure 37a**). The KLK7 inhibitor 096-tag showed a half-life of 30 min. In this case the conjugation to the tag extended the half-life only 6-fold, but still significantly (**Figure 37b**). While the half-life of the KLK5 inhibitor was already satisfying and suited for *in vivo* use, we had to further engineer the KLK7 inhibitor to achieve comparable stability. As both molecules are similar except for the exact amino acid sequences inhibitors, the result exemplifies the major impact of the amino acid sequence on the proteolytic stability of molecules without secondary structure.

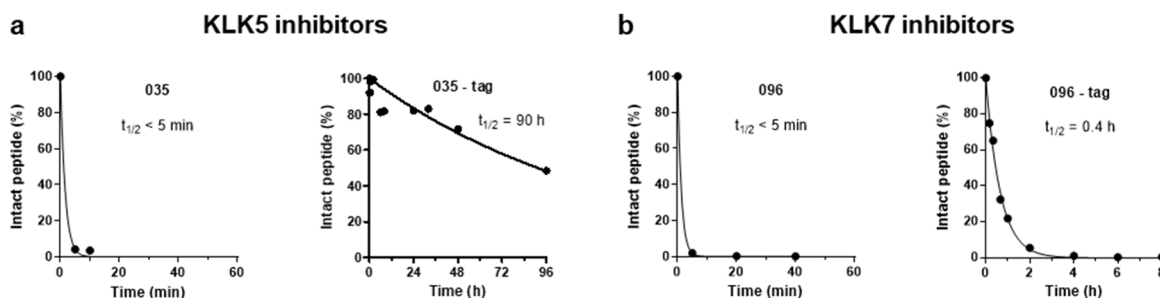


Figure 37 | Metabolic stability of inhibitor-tag conjugates in human plasma. Inhibitors were incubated in human plasma, and the relative concentration of the intact peptide over time was assessed by LC-MS. **(a)** The KLK5 inhibitor without (035) and with the albumin tag (035-tag). **(b)** The KLK7 inhibitor without (096) and with the albumin tag (096-tag). The raw data used for this figure can be found in Figure S6.

3.3.3 Stability improvement of KLK7 inhibitor

To further improve the proteolytic stability of the KLK7 inhibitor 096-tag, its degradation pathway was analyzed by identifying degradation products in the LC-MS data-set of the stability experiment. Subsequently, vulnerable peptide bonds were identified and amino acids in this region modified (**Figure 38**). Peptide synthesis was performed as previously described and high purities could be obtained (**Figure S16**).

A clearly identifiable proteolytic fragment of the KLK7 inhibitor 096-tag was the cyclic peptide alone, with hydrolysis occurring after the C-terminal serine that connected it to the GKG linker and albumin tag (**Figure 38a**). To improve this link, serine 9 was removed from the molecule as it only contributes a factor of 2 to the target binding when comparing peptide 096 and peptide 114 in the monocycle SAR screening (**Figure 31a**). A new KLK7 inhibitor without serine 9 (214-tag) was synthesized and tested for its stability in human plasma. It had a 10-fold improved half-life of 4 hours. As affinity and stability improvement studies were conducted in parallel, it is noteworthy that in this peptide also leucine 5

was changed to norleucine, which was found to improve the affinity but was not expected to change the stability, but a contribution to the stability improvement cannot be entirely excluded.

As plasma half-life of 4 hours was still not satisfactory for therapeutic application we wanted to further improve the plasma stability. Analysis of degradation fragments of peptide 214-tag revealed that cleavage of the amide bonds on both sides of arginine 7 causes proteolytic instability (**Figure 38b**). In the affinity optimization study with unnatural amino acids, we already had identified derivatives for arginine 7, that do not reduce the inhibitory activity: 4-guanidino-phenylalanine and homoarginine. As homoarginine provides the economically favorable alternative from a manufacturing point of view, this derivative was chosen. Furthermore, it was compared to glutamine and its derivative citrulline, which also appeared in the consensus but had not yet been tested. When testing inhibitors with alternative amino acids in position 7, all three alternatives to arginine provided a significantly better stability with half-lives of more than 96 hours. However, from these three alternatives, only homoarginine (peptide 278-tag) did not negatively impair the inhibitory activity (**Figure 38c**), so peptide 278-tag with homoarginine instead of arginine in position 7 was chosen as the final KLK7 inhibitor.

After this stability optimization, both inhibitors had a metabolic half-life of more than 90 hours in human plasma and were considered suited for *in vivo* studies regarding their inhibitory activity and plasma stability.

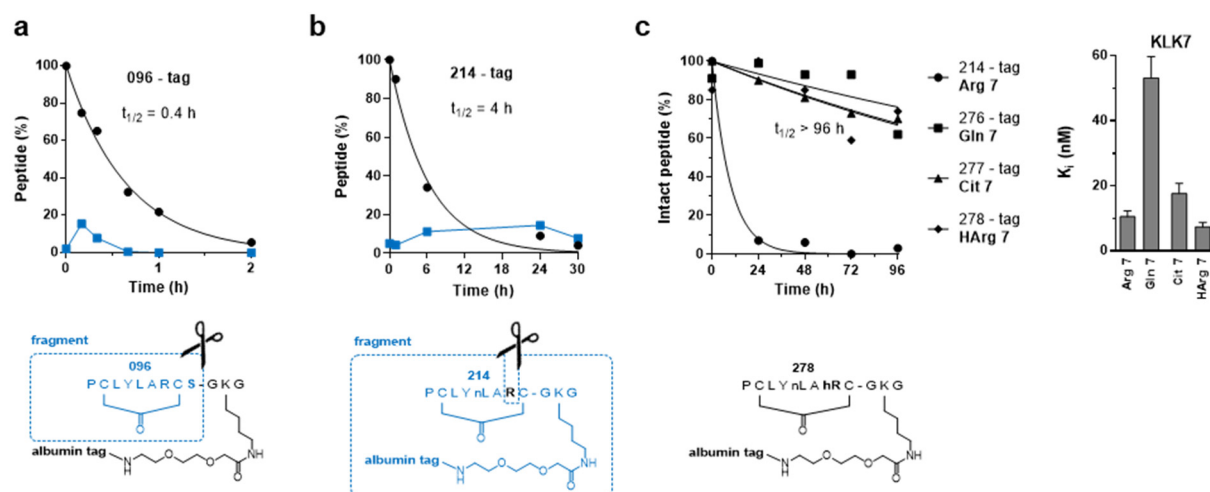


Figure 38 | Stability improvement of the KLK7 inhibitor. (a) Degradation fragments were identified by mass, and Ser 9 was identified as the cause of hydrolytic instability. (b) The removal of Ser 9 improved the half-life to 4 h. Noteworthy, at this point Nle 5 (identified in the affinity optimization study) was incorporated in the inhibitor, which should not impair the plasma stability. Further stability issues were caused by Arg 7. (c) Replacement of Arg 7 finally led to stable inhibitors. Homoarginine was chosen as best derivative based on affinity. All raw data used for this figure are shown in Figure S7.

3.3.4 Replacement of fluorescein in albumin tag

Because the future application of these molecules will be therapeutics, they should have suitable pharmaceutical properties. Regarding the structures of both inhibitors, however, the fluorescein moiety in the albumin tag is not ideal for this application. Even though fluorescein is FDA approved, it has been reported to cause skin discolorations in humans after only a single injection for angiography¹³⁸. Therefore, it is probably not ideal to have a fluorescein moiety in molecules which are intended for chronic therapy in regards of patient compliance, as the chronic application of the molecule might cause discolorations.

It has previously been shown that the fluorescein moiety – even though initially installed for detection reasons only – strongly contributes to HSA binding. Removal or displacement of fluorescein from the albumin tag is associated with an around 30-fold loss in affinity for HSA¹¹⁶. Even slight changes in its position in the albumin tag lead to losses in affinity for HSA (Zorzi *et al.*, data not shown). We searched for a solution to eliminate fluorescein from the molecules without losing affinity for HSA, while the tag should still be easily synthesizable using SPPS and commercially available building blocks. To minimize the number of molecules that had to be synthesized and tested, we addressed this issue with an initial SAR experiment and subsequent rational design.

To get an idea of which moieties of fluorescein contribute to albumin binding, several known albumin-binding small molecules were tested in a competitive FP assay to identify those that compete with the tag for albumin binding. It is known, that albumin has seven binding sites for fatty acids with 10 to 18 carbons and two small molecule binding sites, called ‘drug site one’ and ‘drug site two’¹³⁹ (**Figure S2**). ‘Drug site one’ was characterized to accommodate bulky heterocycles with negative charge, while for ‘drug site two’ aromatic carboxylic acids fit best. We chose a set of molecules that bind albumin in ‘drug site one’ or ‘drug site two’ with reported affinities¹⁴⁰ (**Figure 39a**). The results showed, that the drug site two binders, diflunisal and diclofenac outcompeted the albumin tag at lower concentrations than all drug site one binders (**Figure 39b**). As diflunisal is one of the rather weaker albumin binders tested, but the best competitor, it could be excluded that this is due to a higher affinity for albumin. However, as diflunisal and diclofenac have multiple binding sites in HSA, it is not clear from this assay whether they compete with the fluorescein or the acylated heptapeptide part of the tag. Thus, the binding site of fluorescein in HSA in context of the albumin tag cannot be clearly identified. However, as the only common moieties in both molecules are two aromatic rings and a carboxylic acid, which are moieties that can also be found in fluorescein, we hypothesized, that these are the three moieties in fluorescein that form the main interaction with albumin in context of the tag (**Figure 39c**).

To design molecules that could provide similar target interactions as fluorescein, we attempted to replace it with a free carboxylic acid and two aromatic rings. The strategy was to reintroduce these three groups one by one, first optimizing their position and then introducing the next group. Thus, three generations of molecules were synthesized, starting with the carboxylic acid, followed by a first, and then a second benzoic ring.

To be able to measure the affinity of these tag variants for HSA via FP assays, fluorescein was still required in the molecules. Thus, the synthesis of the tag had to be slightly changed and fluorescein was attached to the side chain of an additional lysine at the C-terminus of the tag (**Figure 40**). First, Fmoc-Lys(Dde) was coupled on the solid support. Then the heptapeptide sequence with standard Lys(Boc) in heptapeptide sequence was attached. After that, the N-terminal variants were coupled to

the amine of the PEG₂ unit. Then the side chain of the C-terminal Lys was selectively deprotected and 5(6)-FAM was coupled on solid phase. After cleavage, palmitic acid was coupled to the side chain of the Lys within the heptapeptide sequence using palmitic acid NHS-ester.

The peptides could be successfully synthesized on 25 μmol scale and satisfactory purities could be obtained (**Figure S18**). The two control peptides F-tag and tag-F were previously synthesized by Dr. Alessandro Zorzi and their identity and purity were confirmed (**Figure S17**).

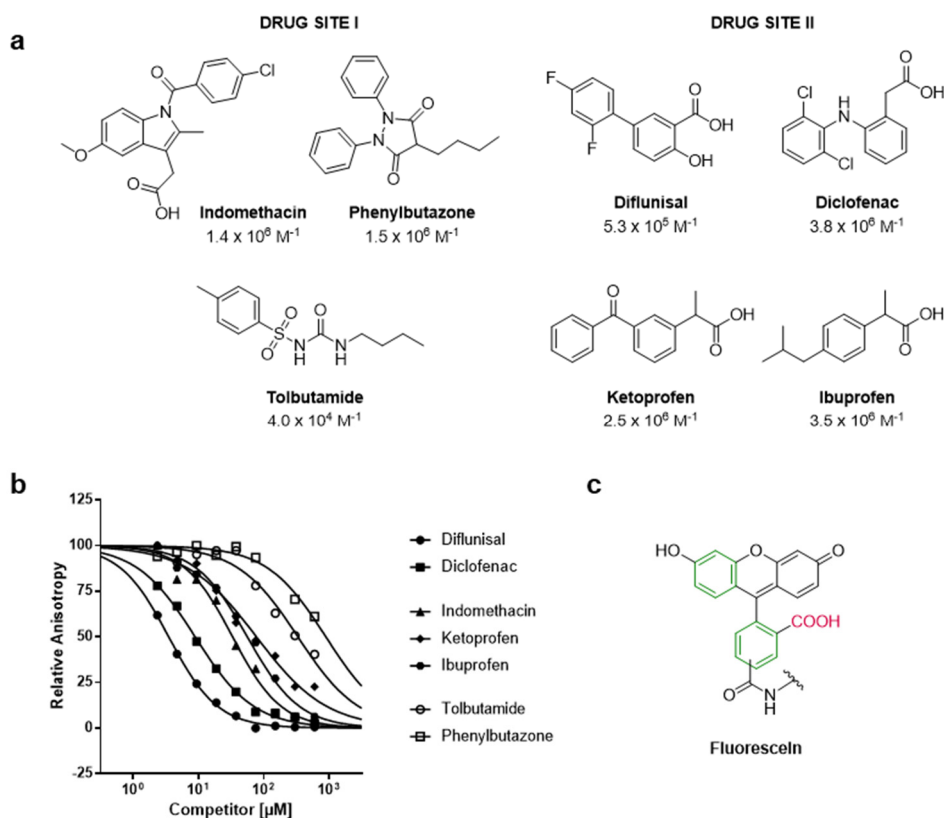


Figure 39 | Competition of small molecules with albumin tag for HSA binding. (a) Structure, binding sites and association constant (K_a) values of known albumin-binding small molecules. Values from ¹⁴⁰. **(b)** Competitive fluorescence polarization assay with albumin tag. **(c)** Moieties of fluorescein that potentially provide interactions with HSA in the context of the albumin tag.

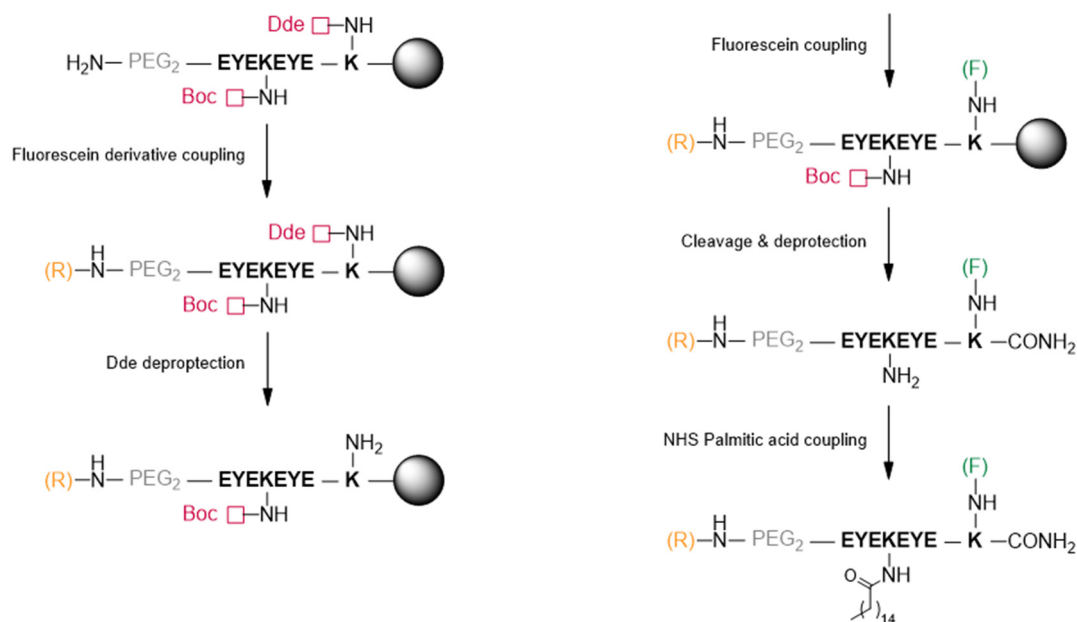


Figure 40 | Synthesis of albumin tag variants. Fmoc SPPS was used. First, N-terminal variants (R) were coupled to the -NH_2 of PEG_2 . Second, the side chain of the C-terminal Lys was selectively deprotected and 5(6)-carboxyfluorescein (F) was coupled on solid phase. After cleavage, palmitic acid was coupled to the side chain of the Lys within the heptapeptide sequence using palmitic acid NHS-ester.

In order to reintroduce only the carboxylic acid, butanedioic acid and pentanedioic acid were appended to the N-terminus of the tag separated by a PEG_2 unit as previously the 5(6)-FAM (peptide 262, 263). Pentanedioic acid performed slightly better than butanedioic acid and already recovered a factor of two in the HSA affinity compared to tag-F. To introduce the first aromatic ring while keeping the carboxylic acid in a comparable position, two strategies each required two coupling steps were compared: either benzenedicarboxylic acid or benzoic acid was coupled to the amine of a gamma-coupled glutamic acid (peptide 264 and 266). The molecule assembled from the benzoic acid (peptide 266) performed better, and the first aromatic ring again recovered another factor of two in affinity. Based on peptide 266, four variants introducing a second aromatic ring were compared (peptide 282, 283, 284, 285). Out of these, biphenyl-3-carboxylic (peptide 282) and 2-naphthalenecarboxylic acid (peptide 284) performed the best, and both improved the affinity by another factor of four. With K_d values of 132 ± 18 nM and 147 ± 8 nM, respectively, the two tag variants 282 and 284 afford an affinity for HSA comparable to the fluorescein-tag (F-tag). Peptide 289 confirms the binding of the fluorescein competitor diflunisal in context of the albumin tag.

Both dye-free tag alternatives are SPPS compatible, require only cheap, commercially available building blocks and one extra coupling step. Thus, they provide dye-free albumin tag variants suitable for therapeutic application. Additionally, the whole set of tag variants allows for a tunable affinity for HSA in the range of 100 – 1000 nM.

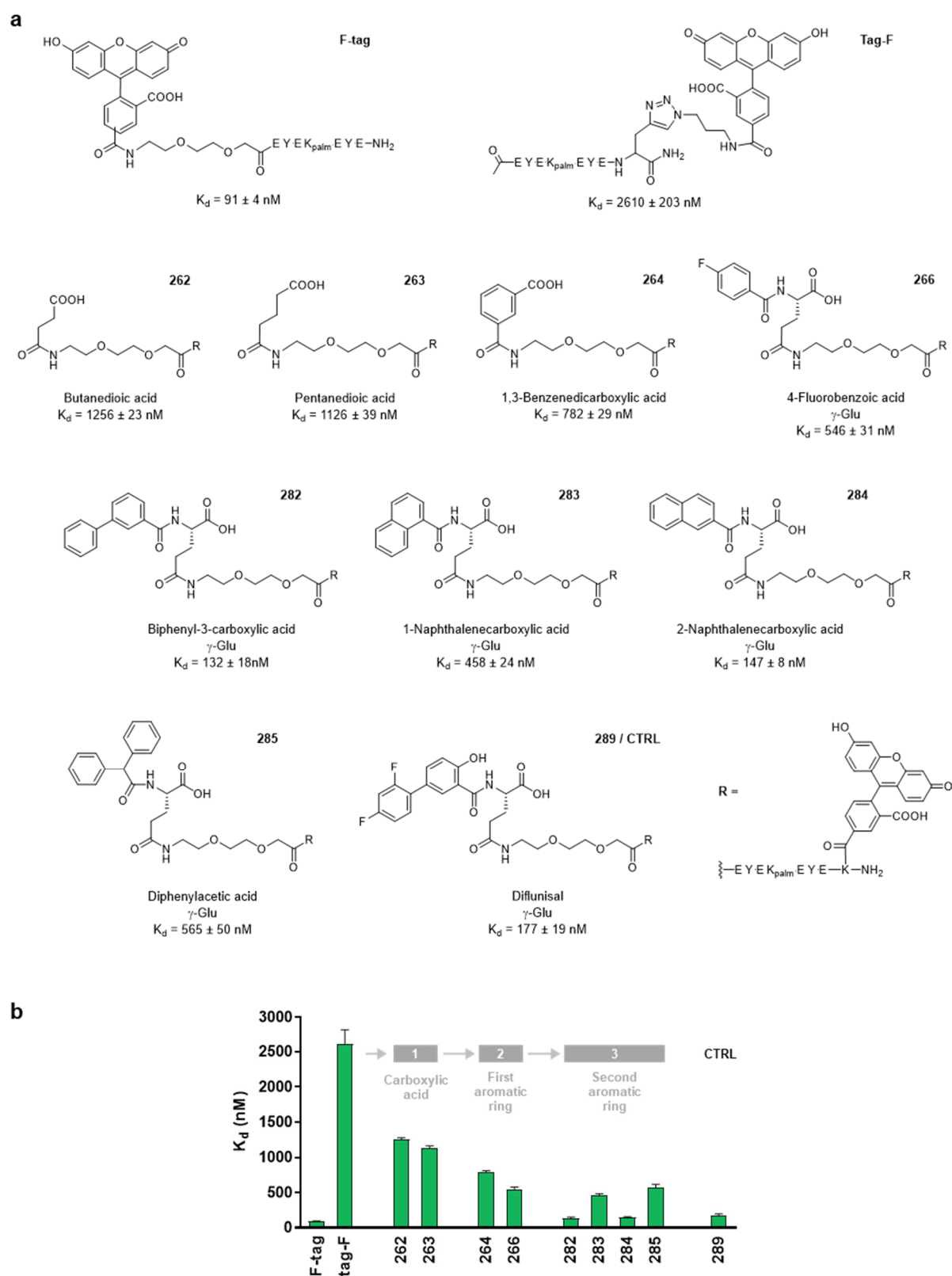


Figure 41 | Fluorescein replacement in albumin tag. (a) Structures and K_d values of tag variants for HSA as measured by FP. R stands for the acylated heptapeptide part of the tag with fluorescein linked to the side chain of an additional lysine at the C-terminus. Complete structures and purities of the compounds can be found in Figure S18. **(b)** Comparison to K_d values of original tag (F-tag), tag with C-terminal fluorescein (tag-F), and three generations of tags with N-terminal derivatives. Data represent the mean of three measurements \pm SD.

3.3.5 Final inhibitors

The structures of the final KLK5 inhibitor 035-tag and the final KLK7 inhibitor 278-tag after *in vitro* optimization, which were used in the *in vivo* studies, are shown in **Figure 42**. The conjugates contain the cyclic peptides found in the phage display selections. The KLK7 inhibitor was further improved with two unnatural amino acid derivatives: norleucine was incorporated in position 5 for affinity reasons and homoarginine in position 7 for stability reasons. Both cyclic peptide inhibitors are C-terminally prolonged by a Gly-Lys-Gly tripeptide serving as linker for the installation of a previously developed albumin tag through the side chain of the lysine. The albumin tag consists of a PEG₂ unit, followed by the heptapeptide Glu-Tyr-Glu-Lys-Glu-Tyr-Glu with the side chain of lysine acylated with palmitic acid, and a N-terminal fluorescein moiety linked through another PEG₂ unit. Derivates of the N-terminal fluorescein moiety were developed but have not been incorporated yet due to the need for sensitive detection in the various *in vivo* studies. The inhibitors have respective K_i values of 1.2 ± 0.01 nM for KLK5 and 7.4 ± 1.3 nM for KLK7. They bind HSA with K_d values of 119 ± 8 nM (KLK5 inhibitor) and 164 ± 12 nM (KLK7 inhibitor). The albumin binding moiety equips the inhibitors with *in vitro* plasma half-lives of more than 90 hours in human plasma. As during the pharmacokinetic and efficacy studies of this work sensitive detection of the molecules was still essential, the fluorescein-containing albumin tag was further used in this work.

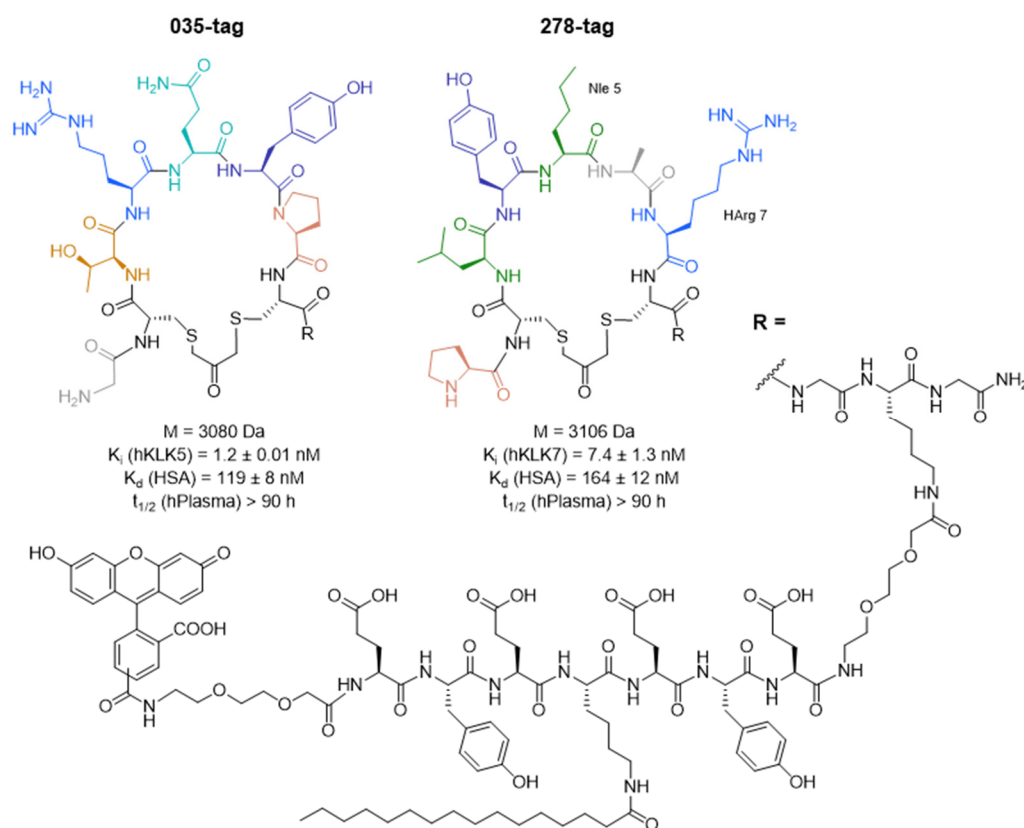


Figure 42 | Structures of the final KLK5 and KLK7 inhibitors. The molecules are composed of specific, cyclic-peptide-based kallikrein inhibitors and an acylated-heptapeptide-based albumin binding ligand for half-life extension. They have single-digit nanomolar affinity for hKLK5 and hKLK7, respectively, and nanomolar affinities for HSA affording high plasma stability with a half-life of more than 90 h in human plasma.

3.3.6 Pharmacokinetics in mice

Before the efficacy of the molecules could be tested in a disease model, their pharmacokinetic profiles had to be analyzed *in vivo* in wild-type mice. Mice were chosen for the first pharmacokinetic studies as they represent the model organism that will be used for the efficacy studies. The results of the pharmacokinetic studies determine the dosing and administration route for the efficacy studies.

The albumin binding property of the peptides is expected to prevent fast renal clearance of the inhibitors *in vivo*. However, the albumin tag was developed to have a high affinity for human albumin and thus might have a weaker affinity for mouse serum albumin (MSA). Before performing mouse pharmacokinetic experiments, we determined the affinity of the albumin tag for MSA. The affinity for rat serum albumin (RSA) was previously tested, and a K_d value of 220 nM, around 5-times weaker than for HSA, was measured¹¹⁶. The K_d value of the F-tag for MSA was 4.2 μ M and thus two orders of magnitude higher than for HSA, indicating a much weaker affinity of the albumin tag for MSA than for HSA (**Figure 43**).

In order to understand which parts of the tag not forming interactions with MSA, we tested additional control peptides, such as a tag composed of Ser residues (SSSKSSS-tag) and the tag with fluorescein at the C-terminus (tag-F). Additionally, the KLK5 inhibitor O35-tag was tested. With K_d values between 7 and 10 μ M, the control peptides as well as the inhibitor did not perform much worse than the F-tag. The conclusion was that only the fatty acid moiety of the albumin tag efficiently binds MSA, while it was shown previously, that the heptapeptide as well as the fluorescein moiety of the tag contribute to HSA binding¹¹⁶.

For this reason, the albumin tag is not ideal for mouse studies and the inhibitors were expected to strongly underperform in mice compared to other species in terms of pharmacokinetics. These results made a pharmacokinetic pre-study before efficacy testing very important, to understand whether efficacy testing in mice was feasible at all, and if yes, how frequently injections would need to be performed to obtain an effect *in vivo*.

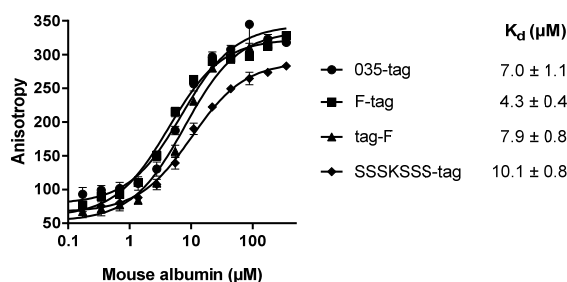


Figure 43 | Binding to MSA assayed by FP. SSSKSSS-tag is a tag variant with all amino acids (Tyr and Glu) replaced by Ser. The structure and purity of this molecules can be found in Figure S19. A dilution series auf mouse albumin in PBS was used.

The pharmacokinetics of the KLK5 inhibitor 035-tag and KLK7 inhibitor 278-tag were investigated in wild-type C57BL/6J mice, comparing different routes of administration. Mice ($n=3$) were either injected intravenously (IV), intraperitoneal (IP) and subcutaneously (SC) with 100 μ L KLK5 inhibitor 035-tag at a concentration of 400 μ M and IV ($n=3$) with the same dose of KLK7 inhibitor 278-tag. For both inhibitors, this corresponds to a dose of 6.2 mg/kg for a mouse of 20 g. For the IV injection this dose was expected to result in an inhibitor blood concentration of 20 μ M, assuming 2 mL of blood, and a plasma concentration of 40 μ M, assuming 1 mL of plasma. After injection, the concentration of the inhibitor in plasma was analyzed at different time points over a period of 24 hours. Blood samples were taken from the tail vein into EDTA coated tubes, plasma proteins were precipitated, and the supernatant was analyzed by HPLC equipped with a fluorescence detector. After IV injection, the KLK5 inhibitor 035-tag and the KLK7 inhibitor 278-tag inhibitor showed an elimination half-life of 3.9 and 5.9 h, respectively (**Figure 44a**). The residual plasma concentrations 24 hours post injection were around 0.2 μ M for the KLK5 inhibitor and around 1 μ M for the KLK7 inhibitor, which seems to have a slightly longer half-life *in vivo*. For the KLK5 inhibitor 035-tag, SC and IP injection routes showed lower peak concentrations and higher residual concentrations 24 hours post injection with around 0.6 μ M.

Two degradation fragments could be observed in the HPLC tracks (**Figure 44b**): While the intact peptide elutes at 18 min, one fragment elutes at 21 min and the other one at 22 min. Both fragments were also overserved when testing the stability of peptide 035-tag in mouse plasma *in vitro* (**Figure S8**). Both fragments must still contain the fluorescein moiety as they were detected with the fluorescence detector. Furthermore, they must still contain the fatty acid as they elute later from the column than the intact inhibitor and thus are more hydrophobic. Further analysis revealed, that the main fragment, which elutes at 21 min is probably a fragment of the tag, cleaved after the second Tyr within the heptapeptide sequence. By replacing this Tyr with Glu (280-tag) the degradation fragment disappeared. However, the stability of the peptide in mouse plasma could not be improved through this replacement (**Figure S9**). That the peptide part of the tag is metabolically less stable in mouse plasma than in human plasma is most likely also due to the weaker affinity of the tag for MSA than for HSA.

To test, whether the inhibitor would have a longer half-life, if it had a better affinity for MSA, a subcutaneous co-injection with one equivalent HSA was performed. The results showed that co-injection with HSA prolongs the absorption phase and the elimination half-life to 7 hours. A plasma concentration of 0.1 μ M could still be measured 48 hours post injection. RSA, used as a control, only influenced the absorption phase, but the elimination half-life was not prolonged (**Figure 44c**).

These results demonstrate the strong relationship between HSA affinity and half-life. It suggests that the half-life of the inhibitors in other species might be better than would be expected by allometric scaling and will strongly depend on the affinity of the tag for the albumin of the respective species. Furthermore, we demonstrated that the absorption phase can be prolonged through administration with adjuvants that bind the inhibitors and retard its distribution from the injection site.

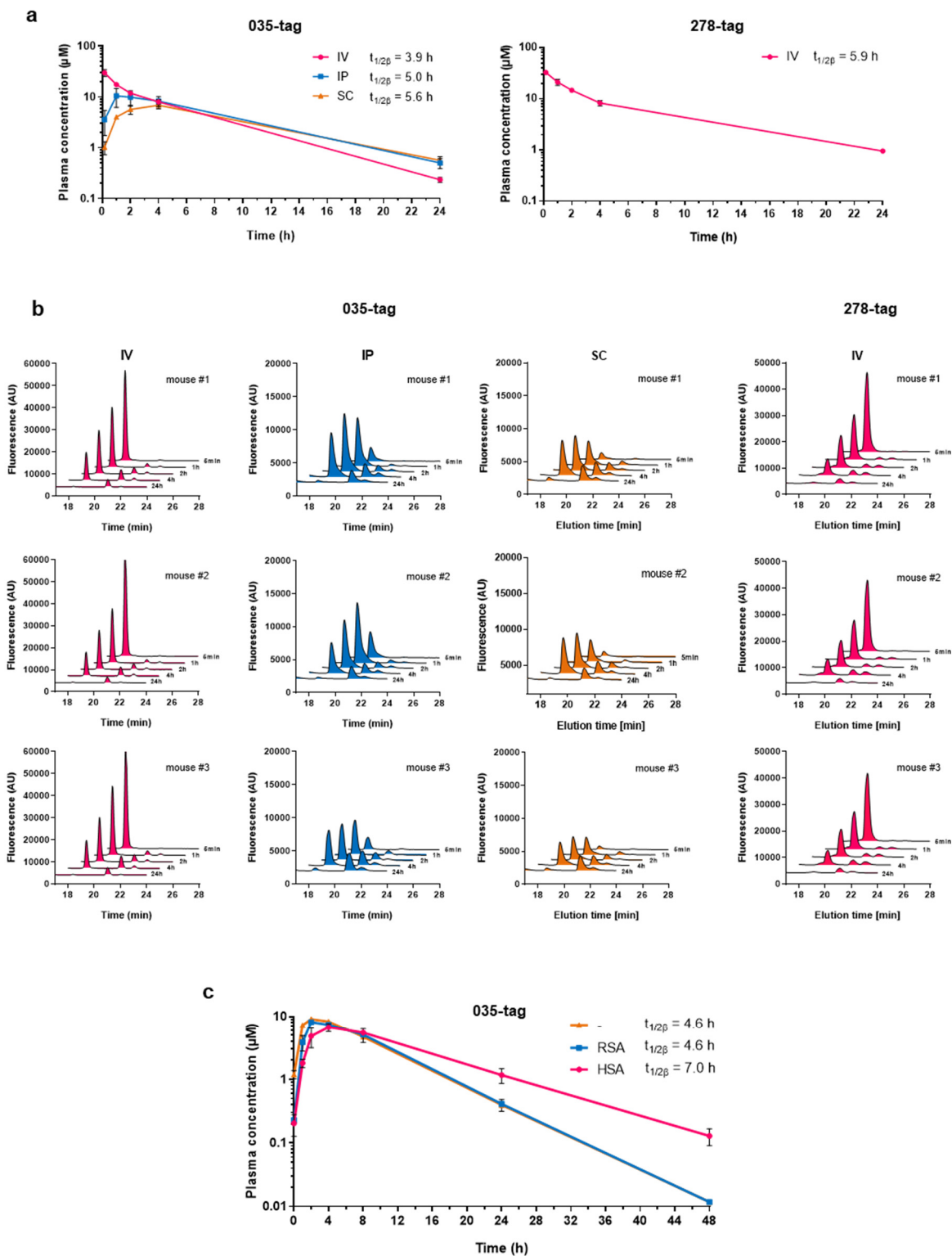


Figure 44 | Pharmacokinetics in C57BL/6J mice. (a) Mice ($n = 3$) were injected intravenously (IV), intraperitoneally (IP) or subcutaneously (SC) with 6.2 mg/kg of KLK5 inhibitor or IV ($n = 3$) with the same dose of KLK7 inhibitor. The expected plasma concentration after IV injection was 40 μM . Blood samples were taken at different time points, plasma proteins were precipitated, and the supernatant was analyzed by analytical HPLC equipped with a fluorescence detector. Graph depicts the concentrations of intact peptide in blood plasma at different time points after injection. (b) Raw HPLC data of all mice. (c) Mice were injected SC with 6.2 mg/kg of KLK5 inhibitor ($n = 2$) or inhibitor coinjected with 1 equivalent (26.4 mg/mL) RSA ($n = 2$) or HSA ($n = 3$). Blood samples were taken at different time points, plasma proteins were precipitated, and the supernatant was analyzed by analytical HPLC equipped with a fluorescence detector. Graph depicts the concentrations of intact peptide in blood plasma at different time points after injection.

3.4 Skin distribution

3.4.1 Biodistribution to skin and organs

In addition to the plasma concentration, we wanted to assess the biodistribution of the inhibitors. The fluorescein moiety in the albumin tag theoretically allows to monitor the biodistribution with a sensitive CCD camera. However the emission signal of fluorescein (Em: 565 nm) is confined to superficial structures with < 3 mm of tissue depth due to absorption and light scattering¹⁴¹. To get an idea of the concentration of the inhibitor in skin in comparison to other organs, we imaged skin samples and all visceral organs.

We intravenously injected mice (n = 3 per timepoint) with the KLK5 inhibitor 035-tag, F-tag (without inhibitor), 035-fluorescein (non-HSA binding control), and PBS (background control) with a dose of 2 $\mu\text{mol/kg}$ (corresponding to 6.2 mg/kg for the KLK5 inhibitor). Mice were imaged at various time points up to 48 h post injection (**Figure 45a**). The signal of 035-fluorescein reached to almost background level only 2 hours after injection, indicating the rapid clearance of non-albumin-bound molecules. In contrast, the KLK5 inhibitor and albumin tag showed an increase in signal up to 8 hours post injection, and a significant signal over background was still detectable 48 hours post injection. First, the increase in the signal intensity up to 8 hours after IV injection confirms that superficial tissue and not the vascular system is being imaged, as IV injected molecules reach their peak concentration in the vascular system seconds after injection. In the first 4-8 hours post IV injection, a strong drop in concentration was observed in the blood. This imaging experiment confirmed that the molecules are not eliminated but instead distribute to the tissue in this time. An increase in signal originating from extravasation that was only observed for the two albumin binding molecules and not for the peptide with fluorescein only shows that extravasation to the tissue is strongly driven by albumin binding. This underlines the importance of the albumin tag for both the half-life as well as for the distribution of the molecules to the tissue.

To confirm the distribution of the inhibitor to the skin, we sacrificed mice at different time points (1 h, 4 h, 12 h) after IV injection and imaged the isolated skin (**Figure 45b**). The strongest signal in the skin was detected in the 4 hour sample, matching with the observed peak concentration in the tissue between 4 and 8 hours post injection.

To compare the distribution of the molecule between skin and visceral organs, mice were euthanized 48 hours after injection of 12.4 mg/kg of inhibitor, and the organs were collected and imaged. The strongest fluorescence signal was observed in the skin, while only a low signal intensity was measured in the liver, spleen, kidneys, heart and lung. The gastro-intestinal tract showed moderate fluorescence compared to the skin and other organs (**Figure 45c**). These results indicate that the molecule is secreted through the digestive tract rather than the kidneys, which can also be attributed to albumin binding.

Taken together, these data strongly indicate that the inhibitor not only reaches, but primarily distributes, to the skin. This favorable biodistribution can be attributed to the non-covalent albumin-binding strategy. Based on the pharmacokinetic and biodistribution studies, the optimal dose, administration route, and frequency of injection were determined to be 3x/week subcutaneous injections of 6.2 mg/kg of inhibitor for the first efficacy study.

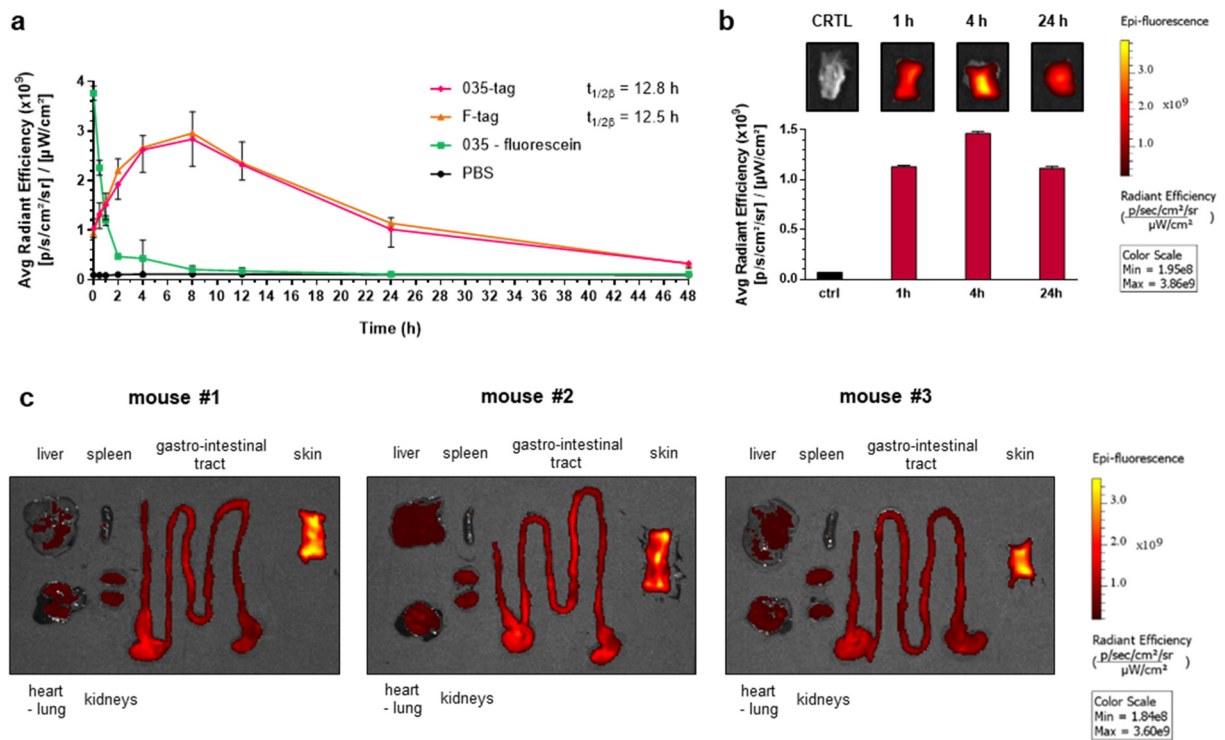


Figure 45 | Biodistribution assessed by fluorescence imaging. (a) Mice ($n = 3$ per time point) were IV injected with KLK5 inhibitor, albumin tag (without inhibitor), 035-fluorescein (non-HSA binding control) ($2 \mu\text{mol/kg}$), and vehicle (background control, $n = 1$ per time point), sacrificed and imaging at different time points after injection. The raw data, which also show the region of interest (ROI) used for quantification, are shown in Figure S10. Structure and purity of the peptide 035-fluorescein can be found in Figure S19. (b) Fluorescence imaging of skin patches from mice sacrificed at different time points after IV injection of 6.2 mg/kg of KLK5 inhibitor. (c) Fluorescence imaging of skin and visceral organs of three mice 48 h after IP injection of 12.4 mg/kg of KLK5 inhibitor.

3.4.2 Skin distribution of peptide versus antibody

The *in vivo* fluorescence imaging experiment confirmed that the inhibitors reach the skin, though this is not sufficient evidence that they also reach KLK5 and KLK7, which are located in the outer layers of the epidermis. The skin tissue, which can be separated from the mouse and has been imaged in the previous experiment, comprises not only the epidermis but also the vascularized dermis, and subcutaneous muscular and adipose tissue, which is easier to reach through systemic administration. Even reaching the epidermis is not sufficient, as the two kallikreins are located in the extracellular space of the outermost layer of the epidermis, the stratum corneum, which consists of tightly connected corneocytes embedded in a lipid matrix. The barrier function of the stratum corneum that protects the body from external pathogens and chemicals and internally against dehydration makes it challenging to reach KLK5 and KLK7 from the vascular system.

To get an idea of how deep our inhibitors can penetrate the skin after systemic administration, we performed fluorescence microscopy of skin sections. A second aim of this experiment was to compare whether the small, peptide-based drug modality has an advantage over bigger drug modalities, such as antibodies, in penetrating the epidermal tissue. As KLK5- or KLK7-specific antibodies could not be afforded in the required amounts, we used a monoclonal antibody that does not bind KLK5 or KLK7 specifically. We used mIgG anti-Flag, which binds the flag epitope (DYKDDDDK), an artificial epitope that is not present in mice. Given that KLK5 and KLK7 are present at concentrations of 2-4 and 7 -14 ng/mg stratum corneum dry weight, respectively¹⁴² and we aimed to inject 0.128 mg inhibitor or 6 mg antibody, we assumed that the amount of injected molecule should significantly exceed the amount of available target in the organism. Therefore, the biodistribution should not be primarily driven by target binding, but rather by the biophysical properties of the drug modality. To exclude potential impact of target binding on the biodistribution of the peptide, we additionally studied the distribution of the albumin tag without the inhibitor. Additionally, we tested the biodistribution of MSA. This experiment was to see whether albumin can reach the epidermis, as the biodistribution of our molecules is influenced by albumin binding. Taken together, mIgG anti-Flag-SiR, MSA-SiR, O35-SiR-tag, and SiR-tag were compared on the skin distribution by fluorescence microscopy.

Previously we used fluorescein as a dye for *in vivo* imaging because fluorescein-containing molecules were already available, and we only aimed to image superficial tissue layers. However, fluorescein is not well suited for microscopy due to the strong autofluorescence of tissue at its emission maximum of around 520 nm. Much better signal-to-background ratios can be obtained with red-shifted dyes, like the by Spirochrome developed silicon rhodamine (SiR) with an emission maximum at 674 nm^{143, 144} (**Figure 46b**). For this reason, we used SiR conjugates to assess the skin distribution of our molecules by microscopy. As only milligrams of the dye were available, the synthesis of the inhibitor and tag had to be slightly changed: The N-terminus of the inhibitor was acetylated and Boc-PEG₂-OH was coupled as last building block on the tag. After Lys(Dde) deprotection and palmitic acid coupling, peptides were cleaved from the resin and RP-HPLC purified. Pure inhibitor or tag was coupled at the free N-terminus of the albumin tag using the SiR-NHS ester in DMSO with DIPEA, and it was purified again by RP-HPLC (**Figure 47**). The chemical structure of the final inhibitor-SiR-tag conjugate is shown in **Figure 48a**. Antibody and albumin were randomly labeled with an excess of 5 equivalents SiR-NHS ester in sodium bicarbonate buffer, aiming for a degree of labeling (DOL) of one (**Figure 46a**). However, for MSA, even with 20 equivalents, only a DOL of 0.4 could be obtained. The conjugates were characterized by their

UV spectra (**Figure 48b**) and by SDS-PAGE, confirming their correct size and labeling (**Figure 48c and d**). The peptide-conjugates were additionally checked by LC-MS and analytical HPLC (**Figure S20**).

We aimed at injecting equimolar doses of 2 $\mu\text{mol/kg}$ for all molecules. Additionally, dye only was used as background control. As the mice had a weight of around 20 g, 200 μL at a concentration of 200 μM were injected. These doses correspond to 300 mg/kg of antibody or MSA, 6.4 mg/kg of 035-SiR-tag, 3.6 mg/kg of SiR-tag, and 1.0 mg/kg of SiR. The concentrations of the injection solutions (**Figure 46c**) were calculated based on labeled protein as determined by absorbance at 652 nm (**Figure 48b**), so that the lower DOL of MSA was considered.

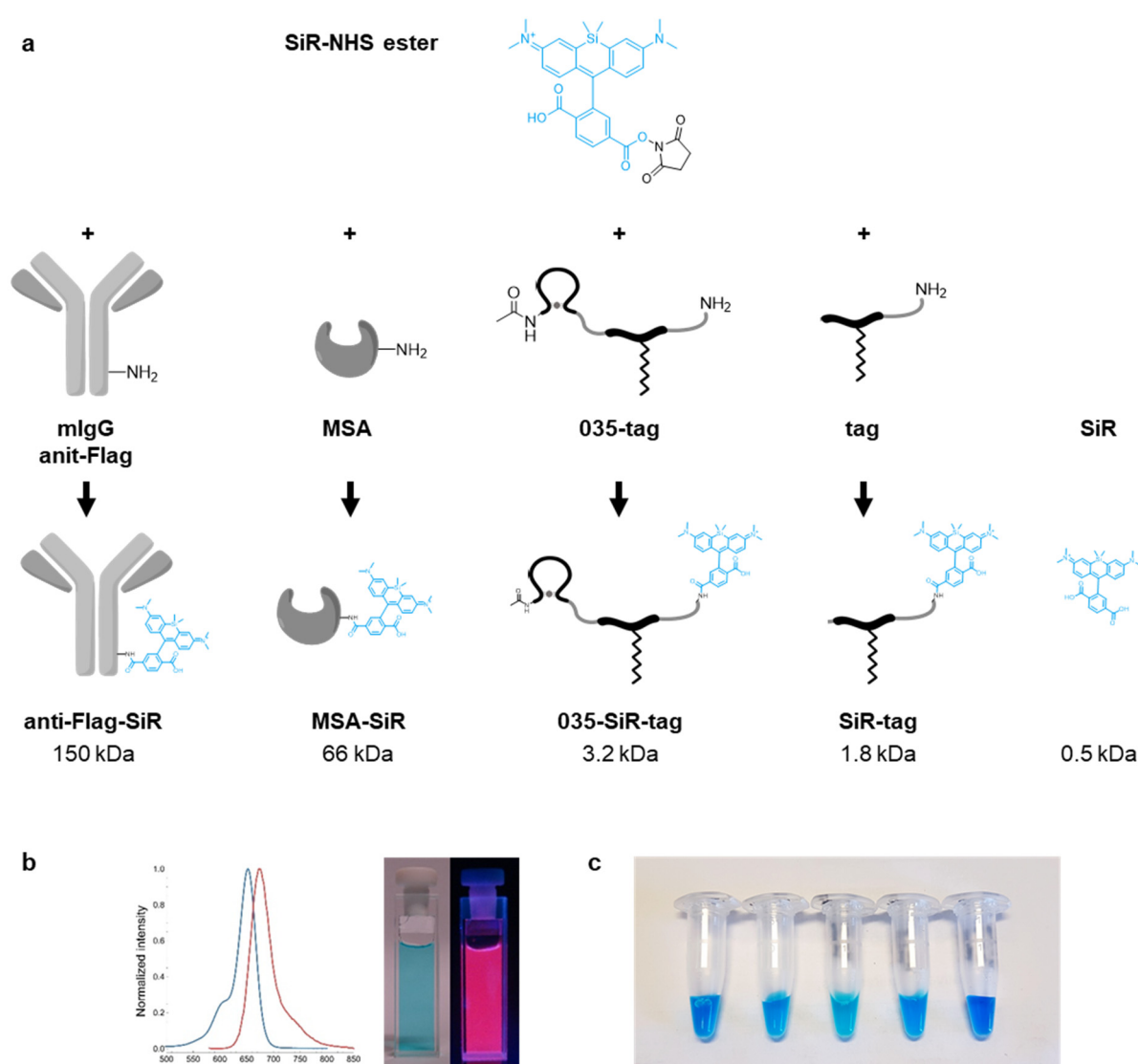


Figure 46 | Silicon rhodamine (SiR) labeling of antibody, albumin, and peptides. (a) All molecules were labeled on free amines using an SiR-NHS ester. Antibody and albumin were randomly labeled with an excess of SiR NHS ester in sodium bicarbonate buffer, aiming for a DOL of one. After labeling, they were purified by size-exclusion chromatography. The inhibitor and tag were labeled on the free N-terminal amine of the tag. The N-terminus of the inhibitor was acetylated for this experiment. Labeling was performed in DMSO with DIPEA followed by RP-HPLC purification. The chemical structure of 035-SiR-tag is shown in Figure 48a. Purities of the two peptide conjugates 035-SiR-tag and SiR-tag are shown in Figure S20. **(b)** Ex/Em spectra of SiR. Abs/Em: 652/674 nm. **(c)** Injection solutions of all five molecules with a volume of 200 μL at a concentration of 200 μM . (b) Used with permission of Spirochrome SA.

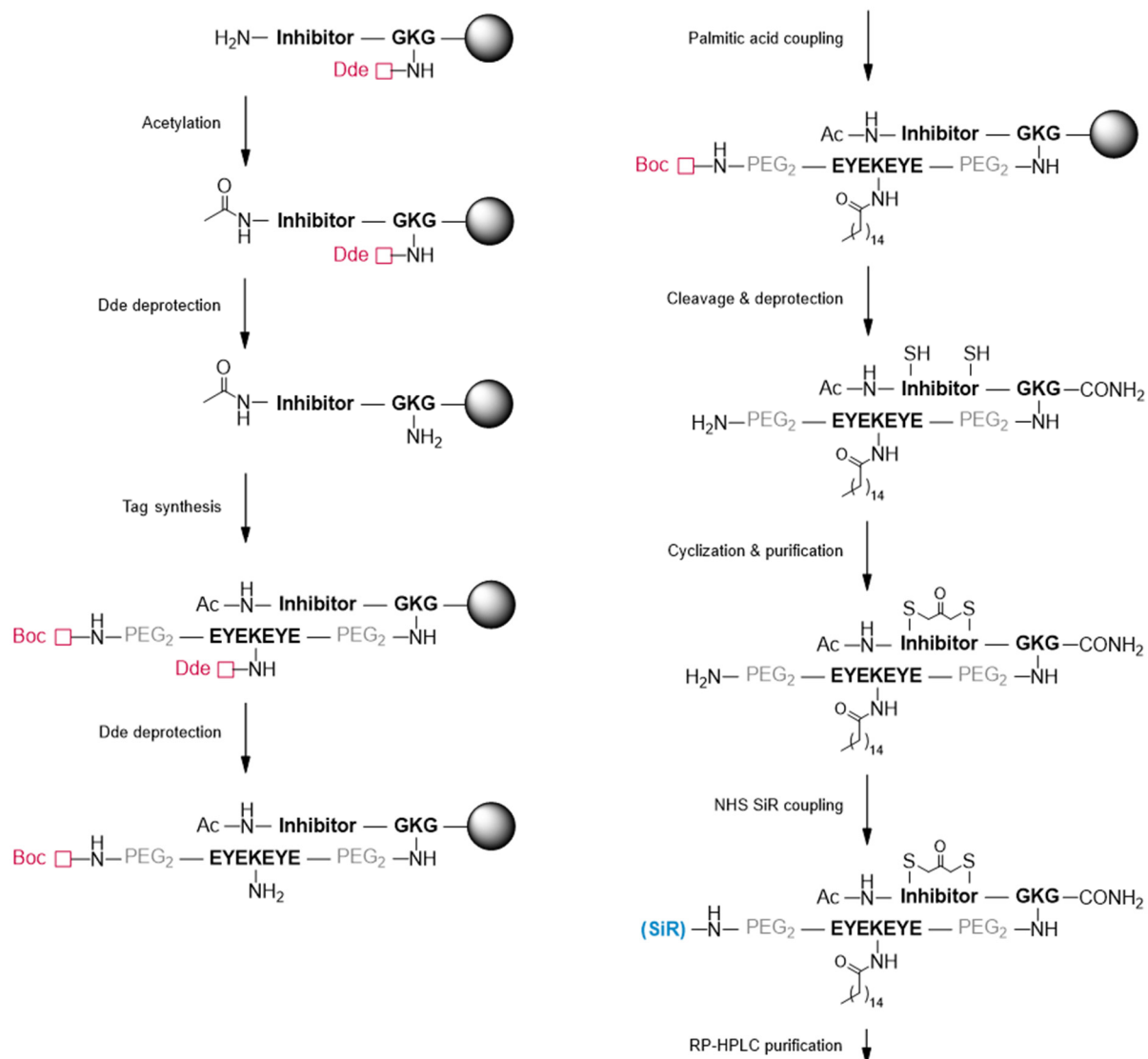


Figure 47 | Synthesis of SiR conjugated peptides. Fmoc SPPS was used. To obtain a molecule with a single free amine after cleavage, the inhibitor was N-terminally acetylated. After selective side chain deprotection of the Lys within the GKG-linker, the synthesis of the tag was continued until a N-Boc protected PEG2. After, the side chain of the Lys within the heptapeptide sequence was selectively deprotected and coupled with palmitic acid. After cleavage, cyclization was performed in solution. Carboxy silicon rhodamine (SiR) labeling was performed on the RP-HPLC purified peptide in solution using its SiR-NHS ester.

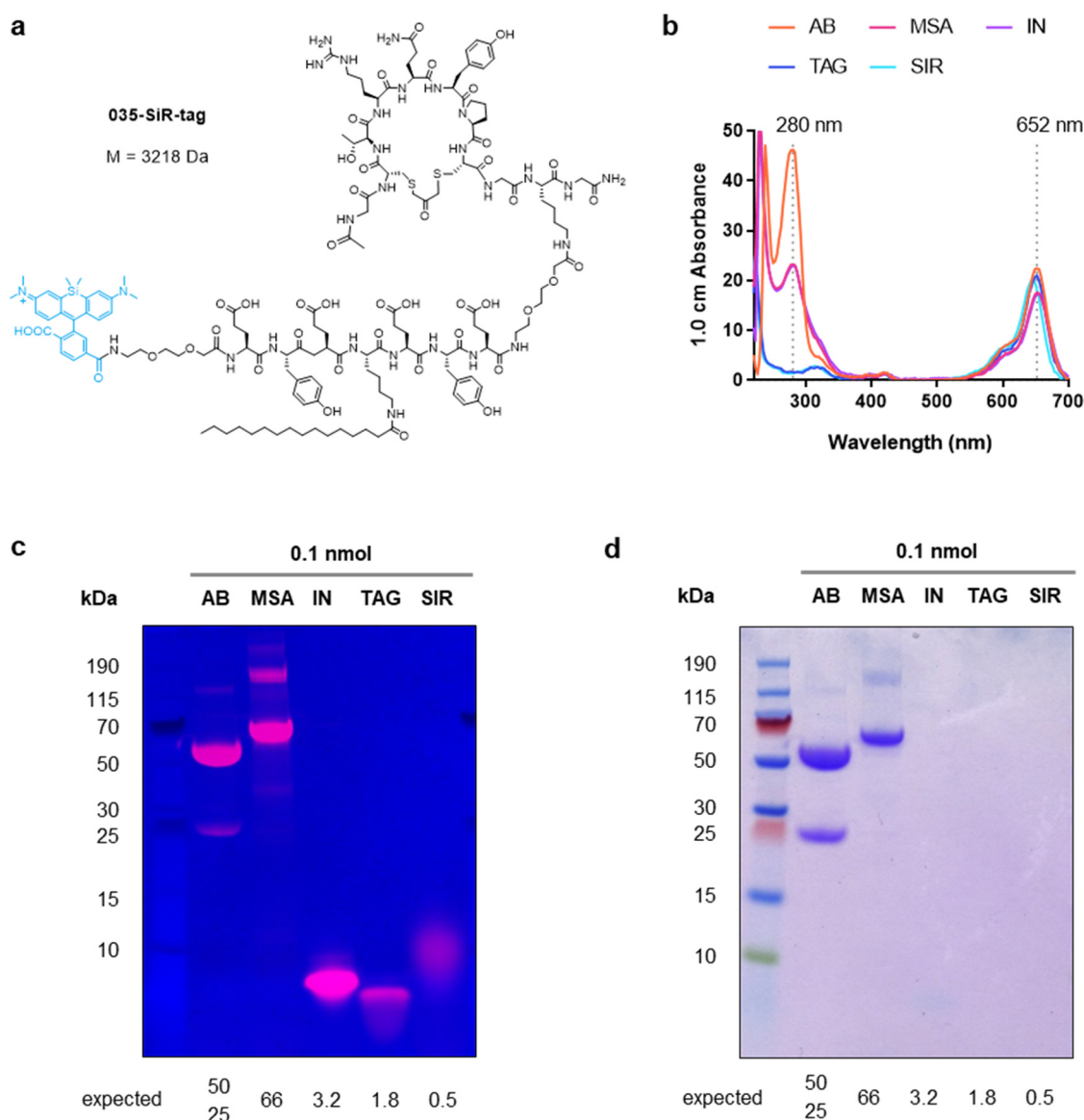


Figure 48 | Characterization of SiR conjugates. (a) Chemical structure of 035-SiR-tag conjugate. (b) UV spectra of SiR-conjugate injection solutions. (c) SDS page gel of 0.1 nmol of SiR-conjugates imaged by excitation at 628 ± 10 nm and capturing at 678 ± 28 nm. (d) Coomassie blue staining of the same gel.

AB: mlgG anti-Flag-SiR; MSA: MSA-SiR; IN: 035-SiR-tag; TAG: SiR-tag; SIR: 6-Carboxy-silicon rhodamine.

Mice ($n = 1$) were injected IP with the prepared solution. We aimed at analyzing the skin distribution at the time when we expected a high concentration of molecule in the skin. Thus mice injected with antibody or MSA were sacrificed 24 hours after injection, while mice injected with inhibitor, tag, or SiR were sacrificed 8 hours after injection. 8 hours correspond to the previously overserved peak concentrations of the inhibitors in the tissue. The injection with dye only served as a background control, as the dye was not expected to distribute to the tissue and should get eliminated within 2 hours after injection, as it was previously observed for fluorescein. Ventral skin tissue was isolated, cryopreserved, and analyzed by fluorescence microscopy (methods in section 5.7.3).

Regarding the section overviews of all five samples, the SiR signal from the injected molecules was only visible for the antibody, the inhibitor, and the tag, but not for MSA (**Figure 49a**). As the MSA-SiR conjugate was characterized for size and labeling, either a problem with the injection or an unexpected weak distribution to the tissue could explain the unexpectedly weak signal. As only one mouse was injected per molecule, issues with the injection cannot be excluded, and more replicates of the experiment would give further information. As the antibody, inhibitor, tag, and SiR were visible, we further analyzed their distribution in the skin by having a closer look at the dermis and epidermis. The inhibitor and the tag both showed a strong signal intensity particularly localized to the epidermis. In contrast, the distribution of the antibody seemed to be limited to the dermis and the signal intensity decreased towards the outer layers (**Figure 49b**). To confirm the stronger epidermal localization of the inhibitor as opposed to the antibody, we generated intensity profile plots along arrows through the epidermis and dermis (**Figure 49c**). In these plots, the beginning of the epidermis is identifiable by the strong DAPI signal originating from the compact, mono-layered stratum basale. While the signal of the antibody was very weak in the epidermis and did not overlay with the DAPI signal, the signal originating from the inhibitor was strong and completely overlaid with the DAPI signal originating from the epidermis / stratum basale.

These initial skin distribution data suggest a clear advantage of our molecular format as opposed to a standard IgG in terms of epidermal targeting. Two apparent differences in the molecules might contribute to their very different skin distribution: First, they have very different biophysical properties. While the 3-kDa inhibitor has only 2% of the size of an IgG, it is also much more hydrophobic due to its fatty acid moiety. The small size might be advantageous for diffusing into the extracellular space of the epidermis, as all keratinocytes are tightly connected through desmosomes, which might limit the accessibility of large molecules. Further the hydrophobicity might be an advantage as the keratinocytes are embedded into a matrix of various lipids, which might limit the accessibility of hydrophilic molecules. The second difference in the molecules that might influence their skin distribution is their mode of biodistribution: While the inhibitor is distributed with albumin, the distribution of IgG is mostly driven by its Fc region. Both, albumin and IgG can bind the FcRn receptor, but they bind to different sites with different affinities, which, for example leads to different levels of transplacental transport¹²⁰ and might also lead to different levels of transport to the epidermis. It is noteworthy that this experiment should be repeated with more replicates to confirm these results.

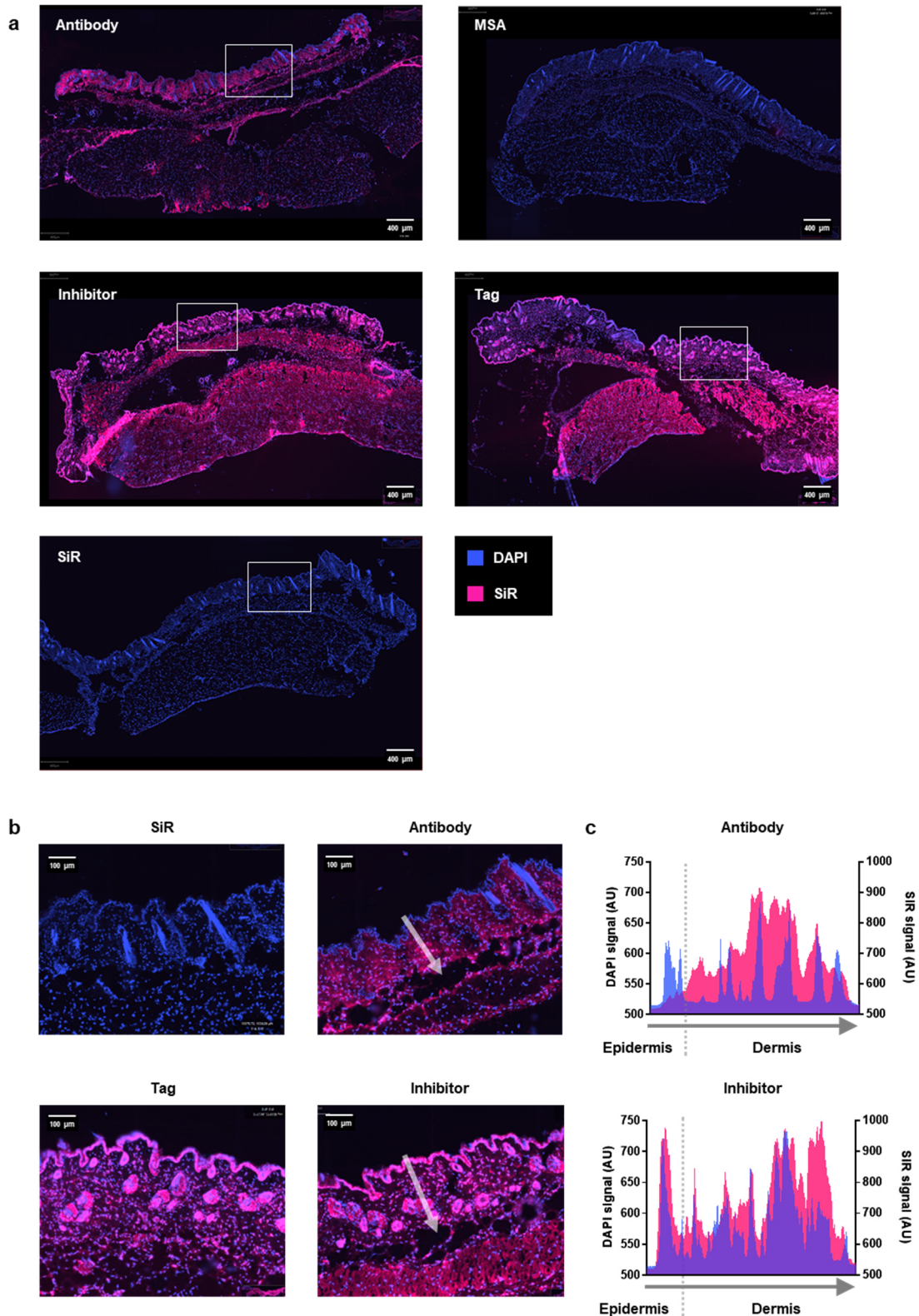


Figure 49 | Skin distribution of inhibitors in comparison with antibodies. Mice ($n = 1$) were IP injected with a dose of 2 $\mu\text{mol/kg}$ of SiR-labeled molecules corresponding to 300 mg/kg of antibody (mIgG anti-Flag) and MSA, 6.4 mg/kg of KLK5 inhibitor, 3.6 mg/kg of tag, and 1.0 mg/kg of SiR. Mice injected with antibody or MSA were sacrificed 8 h post injection. Mice injected with KLK5 inhibitor, tag, or SiR were sacrificed 8 h post injection. Skin was isolated and cryofixed. The DAPI signal is shown in blue and the SiR signal in red. **(a)** Section overview. **(b)** View of epidermis. **(c)** Intensity profile plots of DAPI and SiR signal along the arrows shown in (b). Antibody: mIgG anti-Flag-SiR; MSA: MSA-SiR; Inhibitor: 035-SiR-tag; Tag: SiR-tag; SiR: 6-Carboxy-silicon rhodamine.

3.5 Peptides for efficacy studies

3.5.1 Inhibition of mouse KLKs

To be able to test the efficacy of the inhibitors in NS mouse models other than Tg-*hKLK5*, which only express mKLKs, it was important that mKLK5 and mKLK7 were inhibited. For this reason, the IC_{50} values of the inhibitors for mKLK5 and mKLK7 had to be determined.

The IC_{50} value of peptide 035-tag for mKLK5 is 2 nM and the IC_{50} value of peptide 096-tag for mKLK7 is 20 nM (**Figure 50**). For both inhibitors, the IC_{50} values for their respective mouse KLKs lie in the same range as for the human KLKs. This confirms the structural comparison-based assumption that both orthologues are comparably inhibitable with the same peptide. Thus, the inhibitors are expected to show activity also in mouse models expressing mKLK5 and mKLK7.

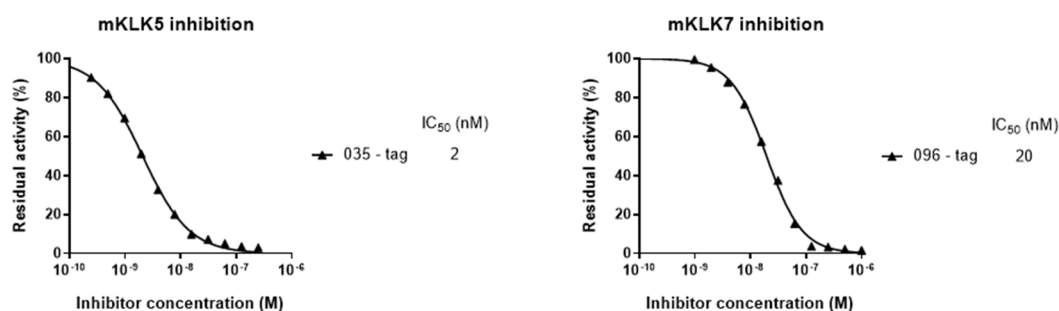


Figure 50 | Inhibition of mouse KLKs. Activity of KLK5 inhibitor 035-tag and 096-tag against mKLK5 and mKLK7. The same assay conditions as for the hKLK activity assays were used. Note that IC_{50} values are shown, as the K_m values have not been determined.

3.5.2 Negative control inhibitors

Two inactive control peptides were synthesized for the *in vivo* efficacy studies. The inhibitors were rendered inactive by replacing the P2 amino acids with alanine and swapping the P1 and P2 positions. Even though swapping the positions would be sufficient to render the inhibitors inactive, the additional amino acid change gives active and inactive peptides of different masses that allows the clear identification of the molecules by mass to be able to detect contaminations or possible confusions also later in samples.

The inactive peptides (301 and 302) were compared to their corresponding active forms (035 and 278) on target inhibition. For both control peptides IC_{50} values $> 100 \mu\text{M}$ could be confirmed for their respective targets (**Figure 51**). After confirming their inactivity, the peptides were synthesized with C-terminal albumin tag for the *in vivo* studies (301-tag and 302-tag).

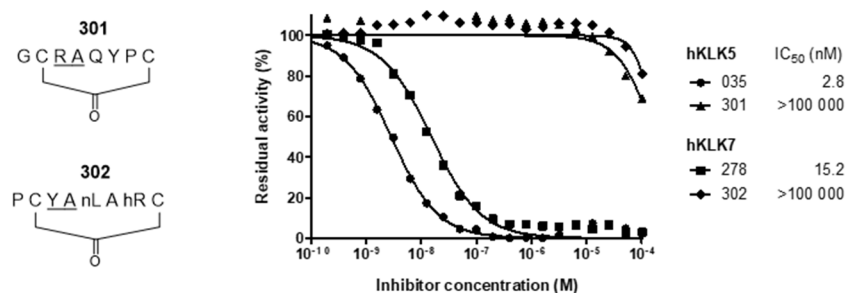


Figure 51 | Characterization of inactive control peptides. 301 is the negative control for KLK5, 302 for KLK7. Inhibitors were rendered inactive by replacing the P2 amino acids with alanine and swapping the P1 and P2 positions (underlined).

3.5.3 Batch synthesis

To obtain sufficient material for *in vivo* efficacy studies, a batch synthesis was performed without changing the synthesis protocol (**Figure 35b**). The whole synthesis involved 23 coupling steps, of which 22 were performed on solid phase and one in solution. 20 out of the 22 solid phase couplings are based on standard Fmoc-chemistry, while two are based on Dde-deprotection with hydrazine to branch the peptide chain.

The synthesis was performed in the scale of $6 \times 50 \mu\text{mol} = 300 \mu\text{mol}$. This corresponded to 900 mg ($M \approx 3000 \text{ g/mol}$) and around 90 mg pure peptide could be obtained, giving a satisfactory yield of around 10% (**Figure 52a**).

The identity of the compound was validated by ESI-MS and the purity was analyzed with RP-HPLC. All peptides used for *in vivo* studies had a purity of $> 95\%$. (**Figure 52b**)

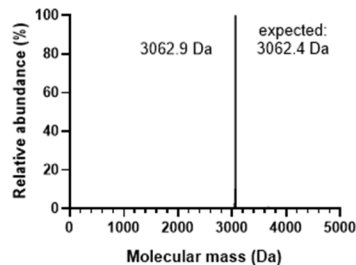
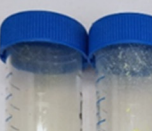


Figure 52 | Peptide inhibitors used for *in vivo* studies. (a) 20 mg of each of the four peptides synthesized for efficacy studies is shown. In total around 90 mg were synthesized. (b) Structure, purity and mass of the four peptides synthesized for efficacy studies.

3.5.4 Formulation

For the *in vivo* studies, a suitable formulation had to be found. First, the solubility of both molecules in ddH₂O was tested. 2 mg of peptide were dissolved in an increasing volume of ddH₂O, while visually observing whether the peptide dissolved or precipitated. While peptide 035-tag was entirely soluble already in 20 µl ddH₂O, which gives a good water solubility of >100 mg/mL, peptide 278-tag was only poorly water soluble and did not entirely dissolve in 2 mL, giving a water solubility of <1 mg/mL (**Figure 53**). As expected, the amino acid sequence in the inhibitor strongly impairs the water solubility, with more than half of the amino acid composition of the KLK7 inhibitor coming from hydrophobic amino acids.

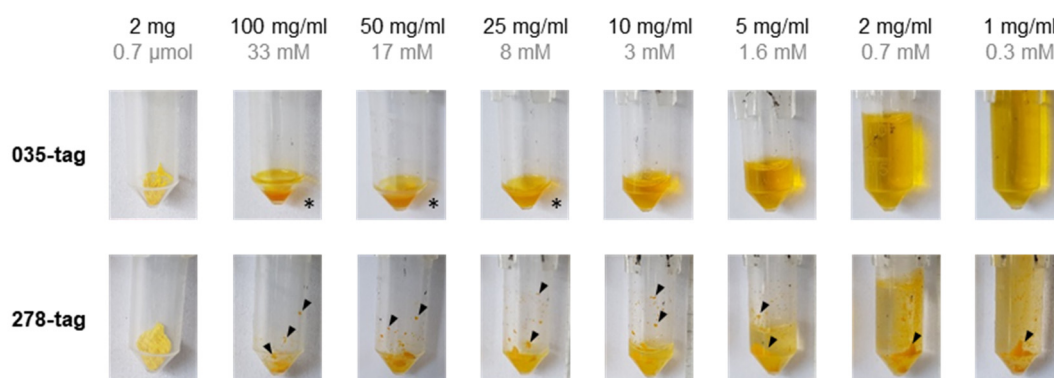


Figure 53 | H₂O solubility of inhibitors. 2 mg peptide inhibitor ($M \approx 3000$ Da) were dissolved in increasing volumes of ddH₂O from 20 µl to 2 mL. After thorough vortexing, the solution was centrifuged for 5 min. Visible precipitation is indicated with black arrow heads. * No precipitation was visible.

For parenteral administration in animals, an aqueous solution with pH 7.4 and salts for tonicity is required. Various buffers were tested. A buffer published for the peptide drug semaglutide, which is based on disodium phosphate with propylene glycol as a tonicity agent and a pH of 7.4 was taken as starting point¹⁴⁵. Then it was tried to be optimized further the buffer for both peptide by addition of SDS, PEG400, DMSO or HSA.

For the experiment, 10 mM peptide stock solutions in ddH₂O with 15% DMSO were diluted into different buffers to a final concentration of 0.4 mM, which corresponds to the desired concentration of the injection solution. The dilution of the stock solution with buffer reduces the final DMSO concentration to 0.6%. The solution was centrifuged to identify and remove precipitated peptide, and the concentration of peptide in the supernatant was spectrophotometrically determined to quantify how much of the expected concentration of 0.4 mM could be reached (**Figure 54**).

First, the reference buffer 1, was compared to PBS (buffer 2). Buffer 1 did provide a slightly better solubility for peptide 035-tag than PBS and the solubility to the required 0.4 mM could already be reached. Peptide 278-tag was comparably soluble in buffer 1 and PBS, but the aimed concentration of 0.4 mM could not be reached. Then it was tried to improve the solubility of peptide 278-tag by addition of SDS and PEG400 to PBS (buffer 3 and 4), but both rather reduced the solubility. Finally, the aimed 0.4 mM concentration of peptide 278-tag could be achieved by either increasing the DMSO concentration to final 1.6% (buffer 5 and 6) or adding 0.1 mM HSA (buffer 8).

For peptide 035-tag, buffer 1 with 0.6% final DMSO concentration was used for all studies. For peptide 278-tag, buffer 6 with a final DMSO concentration of 1.6% was used for the pharmacokinetic experiments, for which only single injections were performed and thus no DMSO toxicity was expected. However, for the efficacy studies wherein repetitive injections were planned to be performed, we wanted to avoid possible DMSO toxicity and used buffer 1 with additional 0.4 mM HSA and a final DMSO concentration of 0.6%. The preparation of the injection solution for *in vivo* experiments is shown in **Figure 55**. The final DMSO concentration could be slightly further reduced to 1.4% (without HSA) or 0.4 % (with 0.4 mM HSA) by preparing 10 mM stock solution in ddH₂O with only 10% DMSO.

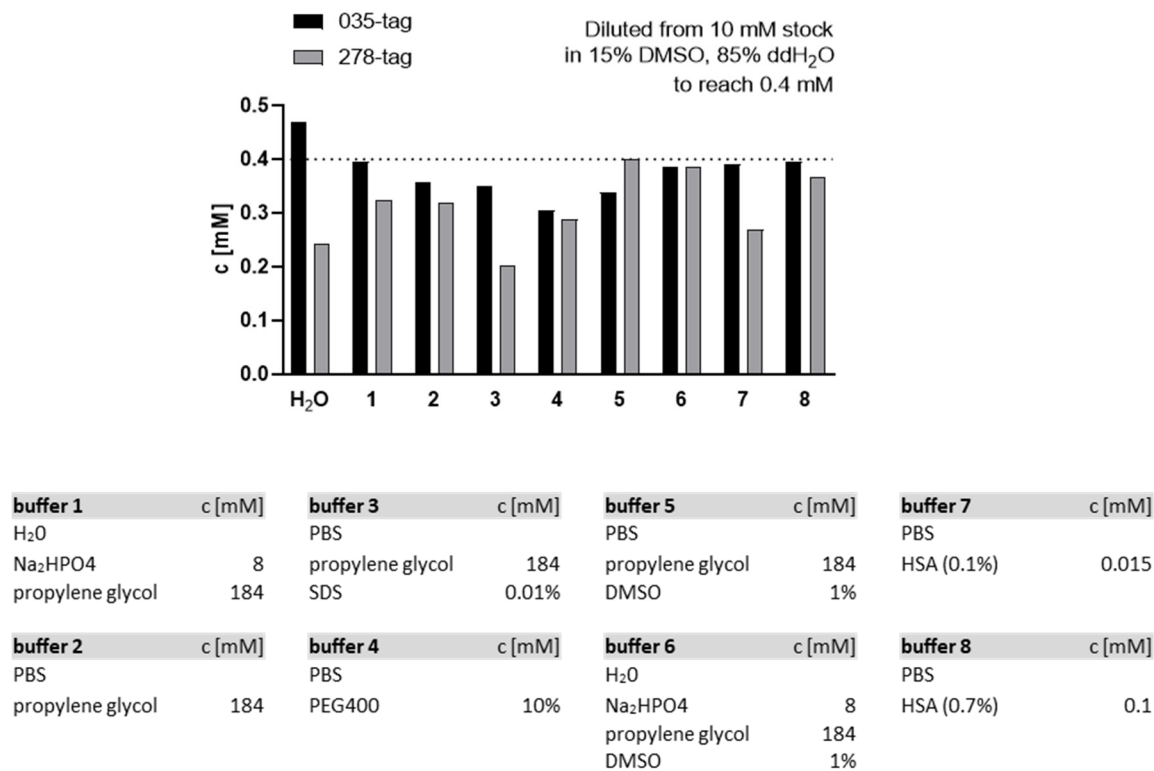


Figure 54 | Solubility of inhibitors different buffers. 10 mM stock solutions in ddH₂O with 15% DMSO were diluted to 0.4 mM in different buffers, all at a pH of 7.4. The solution was centrifuged, and the peptide concentration in the supernatant was spectrophotometrically determined.

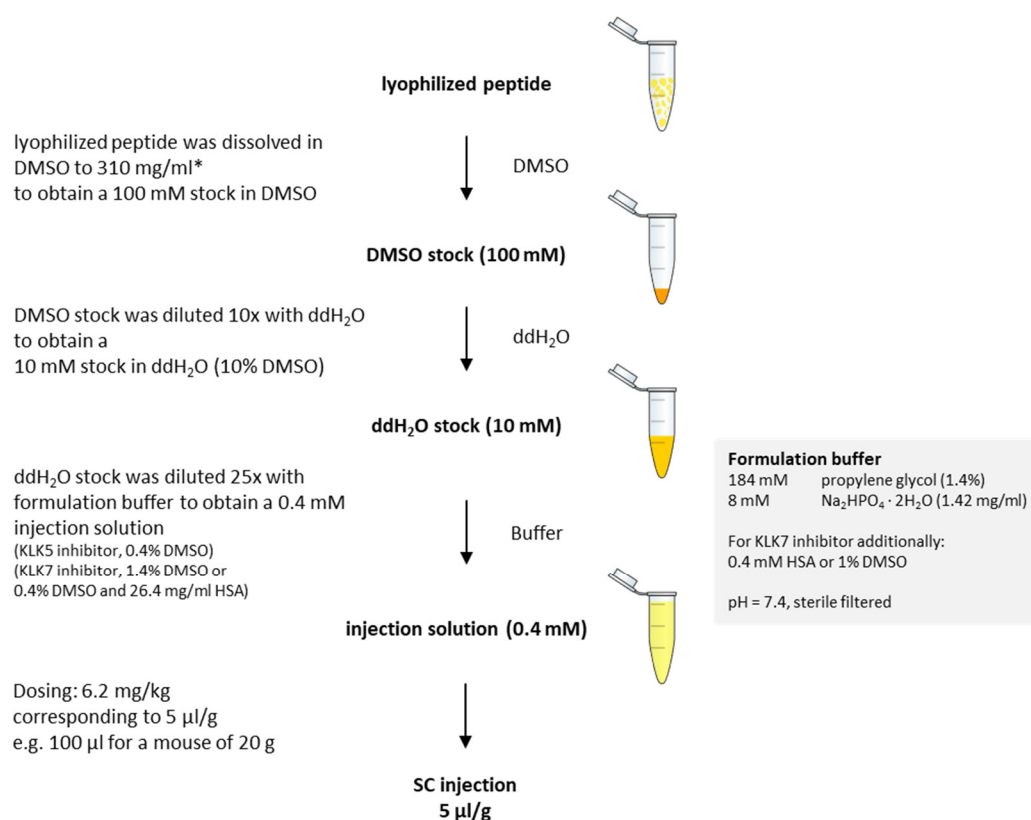


Figure 55 | Preparation of injection solutions for *in vivo* studies. * due to residual TFA and H₂O in the powder, in practice 400 mg/ml were necessary to obtain a 100 mM solution

3.5.5 Efficacy studies

Efficacy studies in various mouse model are currently ongoing in collaboration with Prof. Alain Hovnanian at the IMAGINE Institute for Genetic Diseases in Paris.

To conduct these studies, 0.4 mM injection solution of 30 mg KLK5 inhibitor 035-tag and 30 mg corresponding negative control 301-tag as well as 15 mg KLK7 inhibitor 278-tag and 15 mg corresponding negative control 302-tag were prepared and sent to Prof. Hovnanian.

The studies are still ongoing but the already obtained results look promising (data not shown).

4 Conclusion & Outlook

4.1 Are the inhibitors better than the state-of-the-art?

The aim of the project was the development of specific inhibitors of the two tissue kallikreins KLK5 and KLK7 for therapeutic application in the rare genetic skin disease Netherton syndrome. This chapter will provide a comparison of our molecules to already reported molecules.

A potential therapeutic approach for NS suggested in literature is the delivery of recombinantly expressed LEKTI fragments. Thus, a comparison of our molecules to a LEKTI fragment will be provided. Furthermore, several peptide-based inhibitors designed from engineered sunflower trypsin inhibitor (SFTI-1) have been published and offer nanomolar inhibitory activity against KLK5 and KLK7. For KLK7, a cyclic depsipeptide is described and patented as potent inhibitor. Selective, potent small molecules or therapeutic antibodies have not yet been described.

4.1.1 Comparison of our inhibitors with LEKTI

Due to the 15 slightly different domains that are processed into fragments containing various numbers of domains, the characterization of LEKTI is complex. Furthermore, only a few studies trying to characterize its inhibitory activity against different kallikreins are available^{7, 31}. The solution structure of LEKTI domain 6 (D6), though, has been studied by NMR and the P1-P1' site has been experimentally identified as Arg383 - Glu384³². This allows the assignment of the residue of the inhibitory loops of LEKTI to the subsites of KLK5 for D3 to D14, as they all have an Arg in the same position as Arg383.

Our identified KLK5 inhibitor 036 shares all residues from P3 to P1' with the inhibitory loops of LEKTI D3 to D13, while in D14 only the P2 residue differs. This convergent evolutionary selection mutually confirms both KLK5 as main target of LEKTI D3 to D14 and our inhibitor as a suitable, miniaturized form of LEKTI to replace its function. The LEKTI fragment with the best activity against KLK5 is the D8-D11 fragment with a K_i value of 5.4 nM⁷ (**Table S1**). Our KLK5 inhibitor 035-tag provides with 1.2 nM a 5-fold better K_i than the best LEKTI fragment. Furthermore, it is compared to a single LEKTI domain seven times smaller. In contrast to LEKTI fragments, our inhibitors are entirely chemically synthesizable. A comparison of our most similar KLK5 inhibitor 036 to LEKTI D6, which has been described as a 83 nM KLK5 inhibitor, is shown in **Figure 56**.

The LEKTI fragment with the best activity against KLK7 is the D10-D15 fragment with a K_i value of 20 nM³¹ (**Table S1**). Our KLK7 inhibitor 278-tag provides with 7.4 nM a 2.5-fold better inhibitory activity than the best LEKTI fragment.

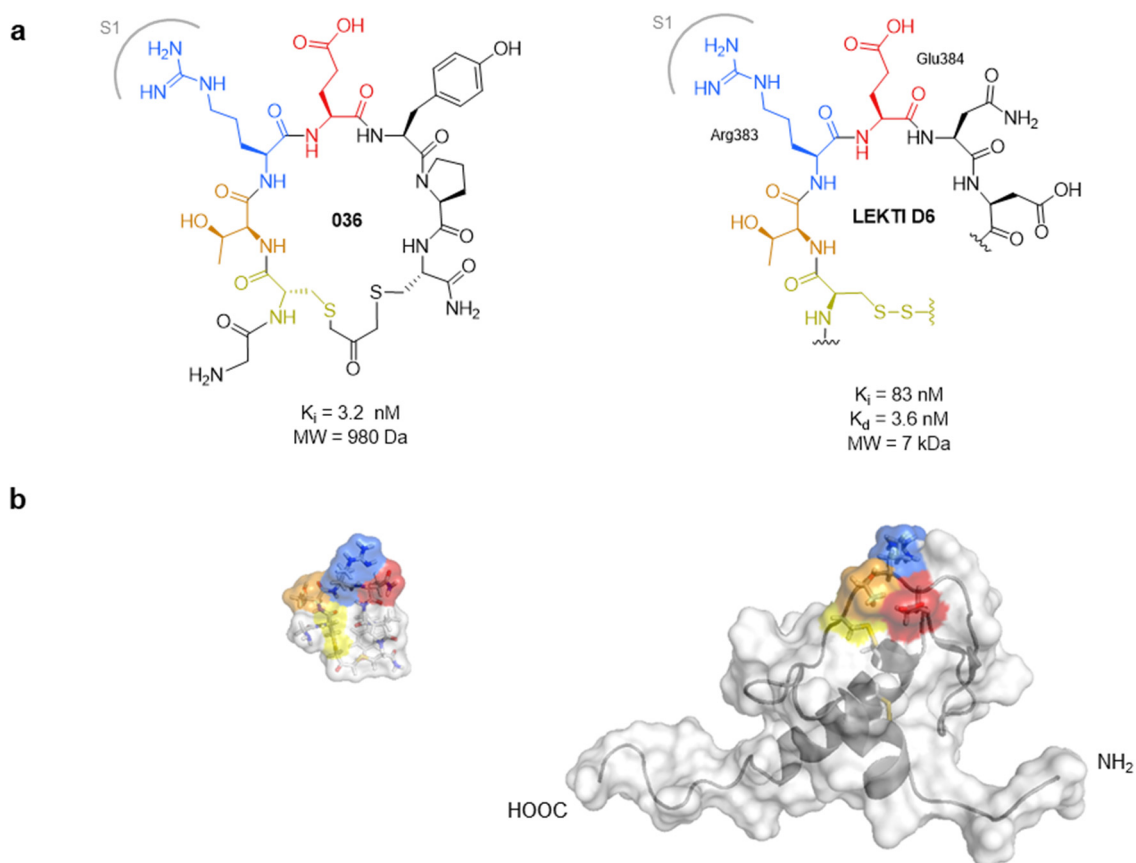


Figure S6 | Comparison of KLK5 inhibitor with LEKTI D6. (a) Chemical structures. Comparison of peptide inhibitor 036 with LEKTI residues 381–386 with Arg383 as the P1 residue. **(b)** 3D structures. Energy minimized model structure of peptide 036 in comparison to the NMR solution structure of LEKTI D6 (1H0Z) with the P3-P1' residues Cys301, Thr 382, Arg383 and Glu384 highlighted in Rasmol color code.

4.1.2 Comparison of our KLK7 inhibitor to the cyclic depsipeptide

The best known non-SFTI inhibitor of KLK7 is a cyclic depsipeptide described in a patent³⁹. It is a natural product isolated from the *Chondromyces* species, wherein further chemical modification of this depsipeptide lead to an inhibitor with an IC_{50} value of 0.2 nM for KLK7. When comparing the structure of the depsipeptide to our KLK7 inhibitor, it is apparent that at least the P2-P1' residues are binding the same subsites of KLK7 (**Figure 57**).

The P1-hydroxy-modification found in the modified, 5-fold better depsipeptide variant could also be introduced in our KLK7 inhibitor. However, this was shown to make the inhibitor less specific and thus might not be ideal for the systemic application in Netherton syndrome, as it might increase the probability for side effects. It could, however, be interesting for the application of the inhibitor for topical administration for other indications.

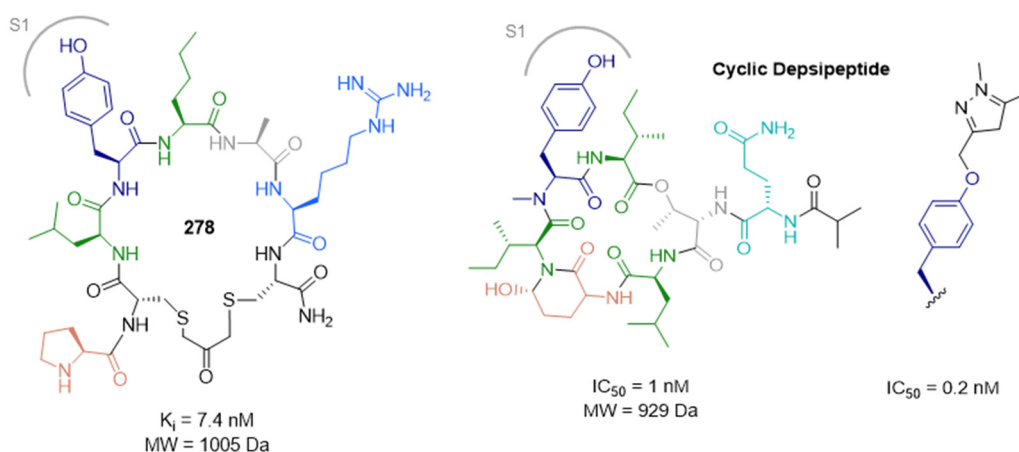


Figure 57 | Comparison of KLK7 inhibitor to cyclic depsipeptide. Chemical structure of KLK7 inhibitor 278 in comparison to a natural (left) and hydroxy-modified (right) cyclic depsipeptide described in US patents 2009/0156472 A1³⁹.

4.1.3 Comparison to engineered sunflower trypsin inhibitors

In addition to the discussed LEKTI fragments for KLK5 and the depsipeptide for KLK7, engineered SFTI-based inhibitors have been published for both kallikreins. The published SFTI-based KLK5 inhibitor³⁶ has a K_i value of 5.1 nM, while the SFTI-based KLK7 inhibitor³⁷ has a K_i value of 0.14 nM.

All SFTI-based molecules are composed of 14 amino acids, making them 1.5-fold heavier than our scaffolds. Their selectivity relies on three to four specific amino acids crafted in certain positions on the SFTI-1 scaffold. The scaffold is N- to C-cyclized, making it more difficult to synthesize and fuse to other molecules than the scaffold identified in this work. Also comparable to our scaffold, the P3 position is occupied with a bridged cysteine, while P4, P2, and P1 render the inhibitor selective. While the introduction of TCTR (P4-P1) resulted in a very unspecific molecule³⁵ inhibiting KLK5, KLK7, KLK14, trypsin and most likely several other proteases, YCQR resulted in the already mentioned a 5.2 nM KLK5 inhibitor and KCLF resulted in the 0.14 nM KLK7 inhibitor (**Figure 58**).

While the SFTI-KLK5 has a 5-fold lower activity than our KLK5 inhibitor (035-tag, 1.2 nM), SFTI-KLK7 is reported to have a 50-fold better activity than our KLK7 inhibitor (278-tag, 7.4 nM) and thus deserves a closer look. This seems rather surprising, as in comparison to our inhibitor, it does not seem to have additional residues that might interact with KLK7, and it contains P1-Phe, which provides less affinity than P1-Tyr, according to our and published data²⁸. An apparent difference to our inhibitor is the Lys in the P4 position, wherein our inhibitor contains Pro. Even though we also found Arg (which is comparable to Lys) in the consensus sequences and it also showed a better affinity than Pro, extra-cyclic Arg or Lys were avoided for stability reasons. However an unnatural Lys derivative like Ornithine or Homolysine might provide a better affinity for KLK7 without causing stability issues and could still be tested if our KLK7 inhibitor, in case it would still need to be further improved for other applications.

In general, all SFTI-based inhibitors rely only on three specifically interacting amino acid side chains and thus might also inhibit other proteases with nanomolar affinity. Thus might have a risk of side effects *in vivo*, particularly when administrated systemically. Furthermore, no *in vivo* studies providing pharmacokinetic data are published. However, due to their low molecular weight they would only have a short plasma half-life due to rapid renal clearance. Thus, without half-life extension strategy they would probably not be suitable for systemic administration.

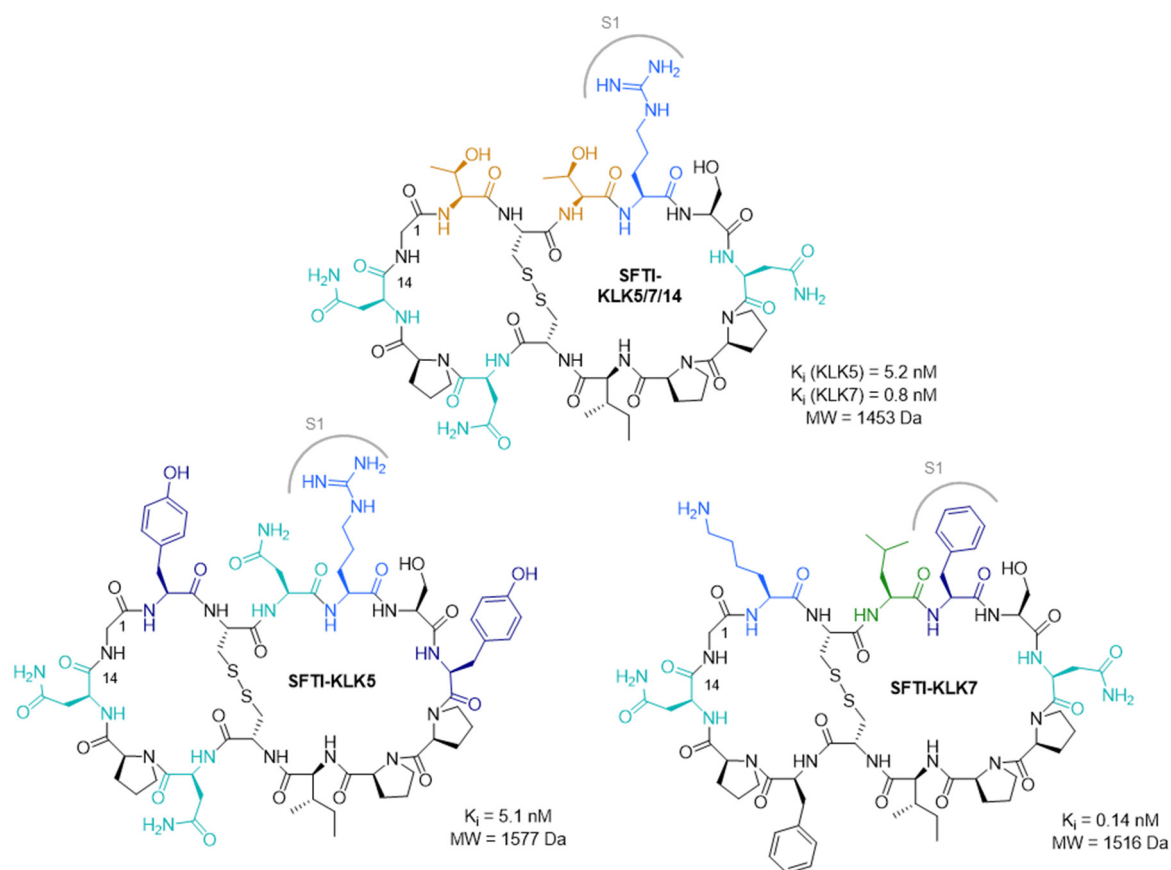


Figure 58 | Sunflower trypsin inhibitor (SFTI)-based KLK inhibitors. The by the authors replaced, target-specific residues, which differ from SFTI-1, are highlighted in Rasmol color code.

4.2 Are the inhibitors good enough for therapy?

The peptide-based molecules developed in this work are the first inhibitors developed for a systemic therapy of Netherton syndrome. Our work provides the first evidence of the feasibility of treating Netherton syndrome using a systemic approach.

Therefore, we have performed pharmacokinetic studies to gain better insights into the efficacy of non-covalent albumin binding as a half-life extension strategy and to understand the differences in biodistribution of different drug modalities. This study provides new insights regarding the suitability of different drug modalities for targeting the epidermis in skin diseases.

This chapter will discuss whether the molecules provide all the required properties for the successful development towards a drug in terms of affinity, specificity, stability & pharmacokinetics, biodistribution and efficacy.

4.2.1 Affinity

The affinity of the inhibitor for its target determines its effective concentration *in vitro* and eventually the dose required to obtain a therapeutic effect *in vivo*. To understand whether the affinity of our inhibitors is sufficient, one needs to estimate which molar concentration would be required in the target tissue to obtain a therapeutic effect. Based on this required concentration, one can calculate the necessary inhibitor dose to evaluate whether it would be realistic.

First one needs to estimate the required molar concentration in the target tissue: In the human epidermis, KLK5 is present at concentrations of 2–4 μg per g of stratum corneum dry weight¹⁴². In the intact epidermis, LEKTI fragments and KLK5 are present in a nearly 1:1 molar ratio³¹, and our inhibitor provides comparable inhibitor activity to LEKTI fragments (**Figure 56, Table S1**). Using a molecular weight of 30 kDa for KLK5, this gives a concentration of 60–130 pmol per g of stratum corneum dry weight for both KLK5 and LEKTI. In the intact tissue, this would correspond to concentrations of around 20–40 nM, calculating with a wet-to-dry weight ratio of 3.5.

Second, to calculate the dose for humans, the volume of epidermis in the human body was estimated: Given that each square millimeter of epidermis contains 29.7 μg of dry material distributed equally between the stratum corneum and the nucleated compartment¹⁴⁶ and that the human body has a surface of around $2 \times 10^6 \text{ mm}^2$, one can estimate the total dry weight of epidermis in the human body to around 60 g, of which 30 g are stratum corneum, which corresponds to 100 mL of intact tissue (wet-to-dry weight ratio of 3.5).

Based on volume and concentration, the dose can be calculated: To achieve at least 90% inhibition, one would need a 10-fold molar excess of inhibitor over target. This gives a total of 20–40 nmol of inhibitor (20 nM*100 mL) corresponding to 60–120 μg that would need to reach the stratum corneum. As the stratum corneum makes up 0.15% of the human body mass (100 g of 70 kg) one can assume that 0.15% of the injected molecule could reach the stratum corneum when administered parenterally. Based on this calculation, a dose of 40–80 mg would need to be administered to a human.

If an equal concentration in the whole body is assumed, the dose could also be easily calculated as: $m = M * c * V$ with $3000 \frac{g}{mol} * 200 \frac{nmol}{L} * 70 L = 42 mg$. For calculation with other biodistributions the results that the stratum corneum makes up 0.15% of the human body mass might be valuable though.

As volumes up to 1 mL are realistic for subcutaneous injections in humans, the concentration of the injection solution would need to be 40–80 mg/mL, which is 30–60-fold higher than used in the pharmacokinetic studies in this work. Because the KLK5 inhibitor has a good water solubility of > 100 mg/mL (**Figure 53**) such concentrations should certainly be feasible, though the formulation would need to be improved.

However, the first realistic estimations to address this question can only be made based on the results of a phase I clinical trial.

4.2.2 Specificity and toxicity

In the specificity testing, no specific inhibition of other, tested proteases was observed. Potential side effects of the inhibitors could be caused by the unspecific inhibition of plasma proteases, which could cause bleeding side effects. The concentration at which proteases of the coagulation pathway were inhibited was around 40 μM (**Figure 33**). To reach 40 μM in an average volume of 3 L of human plasma, an IV injection of 0.36 g peptide inhibitor would be required. As SC injections gave roughly four-times lower peak concentrations, SC injections of 1.44 g would be needed to reach a blood concentration in the range of 40 μM . Thus, the estimated required human doses of 0.04–0.08 g should be entirely unproblematic in terms of plasma protease inhibition.

Furthermore, during the pharmacokinetics experiments in mice, blood concentrations of 30 μM were already reached, and no apparent toxicity was observed *in vivo* for the KLK5 or the KLK7 inhibitor.

4.2.3 Stability and pharmacokinetics

The association of the inhibitors to serum albumin protects the cyclic peptides from exposure to plasma proteases, affording a half-life of more than 90 h in human plasma *in vitro*. Given that peptides in plasma are generally degraded within minutes, as we also observed for the inhibitors without the tag, this is a remarkable stability. *In vivo* albumin binding also prevents the fast renal clearance of the molecule. In contrast to other plasma proteins, HSA has a particular long half-life of around three weeks, and it has been shown for therapeutic molecules that an increase in affinity for albumin correlates with a longer half-life.

The affinity of our KLK5 inhibitor for MSA (7 μM) is around 60-fold weaker than for HSA. *In vitro*, this was associated with a shorter half-life in mouse plasma of around 10–20 h. Despite the rather weak affinity for MSA, the KLK5 inhibitor had a half-life of 6 h in mice (6.2 mg/kg, SC). In the tissue, the molecule could be detected up to 48 hours after injection (6.2 mg/kg, IV).

The peptide drug semaglutide has an affinity for MSA of $1.8 \mu\text{M}$ ¹⁴⁷ (**Table S1**), and the half-life in mice was 7.5 h (1 mg/kg, SC)¹⁴⁸. A duration of action of 48 h was demonstrated¹²⁷ (40 $\mu\text{g/kg}$, SC, Lau 2015). In humans, the half-life of 165 h was reported¹⁴⁹, while the affinity for HSA (1070 nM) is not much stronger¹⁴⁷ (**Table S1**). Semaglutide is administrated once-weekly subcutaneously. Considering that the pharmacokinetic properties of semaglutide and our peptide inhibitor in mice are comparable and that the affinity of our inhibitors for HSA (120 nM) is 9-times better than the one of semaglutide, weekly SC injections of our molecules in humans should be feasible.

4.2.4 Biodistribution

In terms of the biodistribution, we demonstrated an advantage of our peptide-based molecule over an IgG modality for systemic epidermal targeting (**Figure 49**). Efficient biodistribution of the molecules is particularly important for the treatment of NS because the stratum corneum is difficult to reach and normally not targeted by drugs through the systemic route.

When performing efficacy testing in mice, it must also be considered that mouse skin differs from human skin. Particularly, the stratum corneum is much thicker in humans than in mice (**Figure 59**). Thus, efficient and deep penetration of a therapeutic molecule into the stratum corneum is even more important and difficult in human skin than it is in mice. For this reason, some molecules that show efficacy for NS in mouse models might not work well in humans if they do not diffuse efficiently into the stratum corneum, as might be the case for antibodies.

These results also have an impact beyond the application for Netherton syndrome, as they demonstrate the effects of the drug modality on the delivery efficiency to a certain target tissue. It highlights the importance of a careful consideration of size and other biophysical properties as well as the employed mechanism of biodistribution for the development of biologics. Furthermore, these results open new perspectives for the application of small, peptide-based biologics in drug development for targets in tissues that might be difficult to access with antibodies.

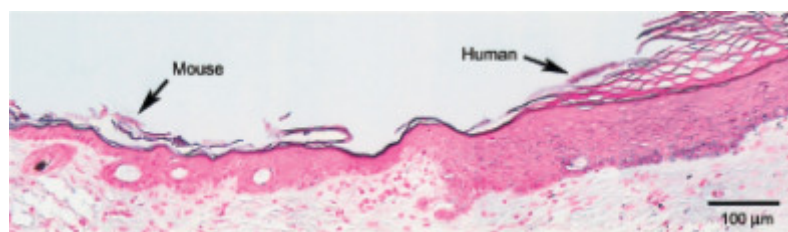


Figure 59 | Comparison of mouse and human skin. Human skin following transplantation to a mouse. The transplanted human skin exhibits a thick epidermis and a lack of skin appendages compared to the thin interfollicular epidermis of the mouse.

Reprinted with permission from ¹⁴⁶, reprinting license nr. 4570241437949, © 2003 Elsevier.

4.2.5 Efficacy

The efficacy of the KLK5 inhibitor is currently being evaluated in a mouse model of Netherton syndrome, and the results look promising. This shows the therapeutic potential of our developed KLK5 inhibitor. A question remaining is whether or not an additional therapeutic effect could be achieved by the combined use with the KLK7 inhibitor. To answer this question, a different animal model than Tg-*hKLK5* that represents the *Spink5*^{-/-} genetic background of Netherton syndrome should be used. As

our peptides can inhibit mKLK5 and mKLK7, they should also be efficacious in mouse models and not require recombinant expressing the human homologues.

As complete *Spink5* knock-out is lethal in mice, either a knock-down or a postnatal induced knock-out would be needed. Treatment of pregnant *Spink5*^{+/-} mice to see whether the survival time of *Spink5*^{-/-} pups could be increased would be an option that would not require developing a new animal model. Otherwise, bigger preclinical test species with a lower surface-volume ratio like rats, where a *SPINK5* knock-out might not be lethal, would also be an option.

4.3 Outlook

4.3.1 Inhibitors for the systemic treatment of Netherton syndrome

Two questions regarding the final chemical structure of the molecules remain to be addressed to use the inhibitors for the systemic treatment of Netherton syndrome

1. *Should the final molecule include a KLK7 inhibitor?*
2. *Should fluorescein be used in the albumin tag?*

An overview of structurally open questions highlighted in dashed boxes is shown **Figure 60a**.

Whether the final molecule should include the KLK7 inhibitor will depend on whether an additional therapeutic effect can be demonstrated by combined administration. Approaches to address this question have already been discussed in the last chapter, though no data that is not that of the collaborators are available to answer this question yet. However, literature indicates that KLK7 activity is not solely dependent on activation by KLK5, and for this reason, both proteases should become targets for NS therapy.

If a KLK7 inhibitor will be required for the therapy, it is important to consider that due to the high costs of toxicology studies and manufacturing development, only a single molecule can be brought into clinical trial, and thus both inhibitors would need to be combined in a single molecule. This would be feasible, but would take some development time. As both inhibitors showed better affinities when coupled C-terminally, the obvious way to combine both in one molecule would be to introduce another branching unit within the GKG-linker region (**Figure 60a**). To avoid isomerization during the cyclization, the straightforward solution for the synthesis of such a molecule would be to synthesize the cyclic inhibitors with C-terminal carboxylic acids and to retain the side-chain protection groups during cleavage. In a second SPPS, the cyclic inhibitors could be assembled, and the albumin tag attached (**Figure S4**). Such a bispecific peptide inhibitor (example shown in **Figure 60b**) would have comparable functional domains to a bispecific antibody with an almost 40-times lower molecular weight. It would have better biophysical properties for reaching the epidermis and would be entirely chemically synthesizable.

The second question left to answer is whether dye-free tag variants should be used in the final molecule. The best dye-free tag variant we developed is based on gamma Glu and biphenyl-3-carboxylic acid, and it has a 1.4-fold weaker affinity for HSA than the F-tag. However, this tag variant has not been entirely characterized yet. For example, it has neither been synthesized without C-terminal fluorescein nor together with the inhibitor. The synthesis of this tag variant without C-terminal fluorescein would allow for the confirmation of its affinity for HSA in a competitive FP assay with the F-tag. Subsequently, it should also be synthesized as a fusion with the inhibitors so that the inhibitor-tag conjugates can be characterized.

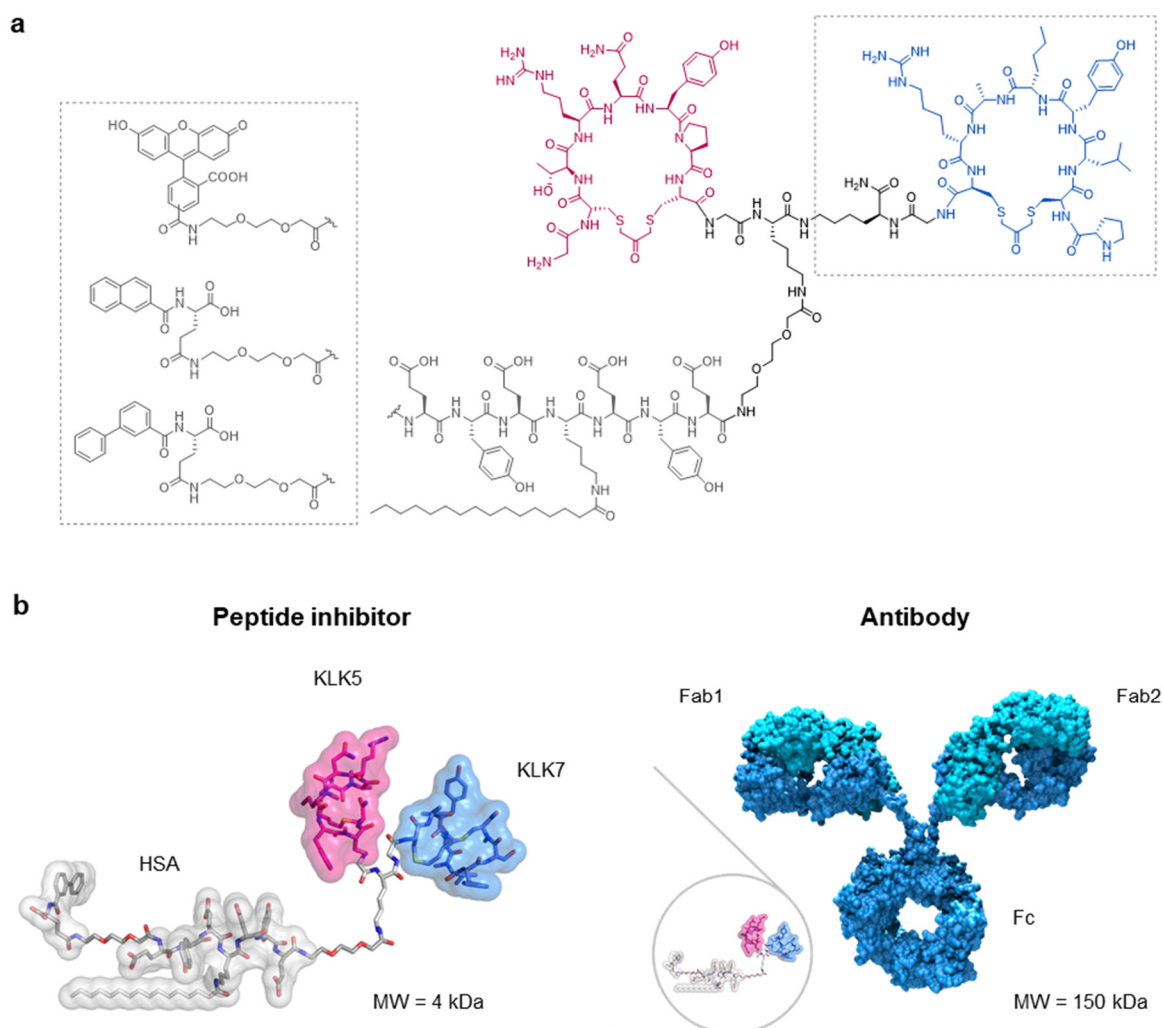


Figure 60 | Antibody-like, bispecific inhibitor for systemic administration. (a) Chemical structure of the inhibitor with open questions to address shown in dashed boxes. **(b)** Example structure of bispecific inhibitor in comparison to bispecific antibody.

Image of antibody used with permission of Amgen Inc.

4.3.2 Inhibitors for topical administration

In this work, we have shown that the KLK5 and the KLK7 inhibitors provide suitable pharmacokinetic properties for a systemic treatment of Netherton syndrome. For some indications, a topical administration might be preferable and thus different properties might be required.

As size is important for topical administration, the cyclic inhibitors without the albumin tag would probably be better suited. Further optimization for a topical administration route would comprise, for example, a reduction in size and a removal of polarity. Furthermore, route-specific stability issues might need to be addressed. To guide the medicinal chemistry efforts for optimizations, the crystal structures of the two inhibitors in complex with their targets would be valuable. As crystal structures of the kallikreins are already available, the crystallization conditions are known and a co-crystallization with the inhibitors should be feasible. Alternatively, as the P1 residues of both inhibitors are clear, molecular docking might be possible and provide further insights.

Additionally, the inhibitors with the albumin tag could also be considered for topical administration due to their very good epidermal distribution. This would provide the additional advantage that transdermally delivered molecules could bind to albumin after epidermal passage and thus would be protected from degradation and renal clearance.

The development of a topical formulation would also be interesting for both peptide inhibitors for application in less severe diseases than Netherton syndrome. Diseases that have been linked with KLK5 or KLK7 overactivity and that could be treated with a topical formulation would be, for example, atopic dermatitis, rosacea, or psoriasis¹⁵⁰.

5 Materials & Methods

5.1 Materials

5.1.1 Peptide synthesis reagents

Abbreviation	Name	Product code	Supplier
-	Fmoc-protected standard amino acids	-	GL Biochem (Shanghai, China)
ACN	Acetonitrile, HPLC gradient grade	A/0627/17	Thermo Fisher Scientific (Waltham, MA, USA)
Ac ₂ O	Acetic anhydride, 99+%, pure	14949	Acros Organics, Thermo Fisher Scientific (Waltham, MA, USA)
DCM	Dichloromethane	32222	Sigma-Aldrich (Steinheim, Germany)
DIPEA	N,N-Diisopropylethylamine, ≥99 %, for synthesis	4105	Carl Roth (Karlsruhe, Germany)
DMF	N,N-Dimethylformamide ≥99,5 %, for synthesis	6251	Carl Roth (Karlsruhe, Germany)
DMF technical	Dimethylformamide	246-VL54K	Thommen-Furler AG (Büren, Switzerland)
DMSO	Dimethyl sulfoxide	D5879	Sigma-Aldrich (Steinheim, Germany)
EDT	1,2-Ethanedithiol	E0032	TCI (Zwijndrecht, Belgium)
Et ₂ O	Diethyl ether (PG)	527-VL53TS	Thommen-Furler AG (Büren, Switzerland)
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate	-	GL Biochem (Shanghai, China)
HOBt	1-Hydroxybenzotriazole	-	GL Biochem (Shanghai, China)
Lutidine	2,6-Lutidine	L3900	Sigma-Aldrich (Steinheim, Germany)
NH ₂ -NH ₂	Hydrazine hydrate (hydrazine 64%)	19671	Acros Organics, Thermo Fisher Scientific (Waltham, MA, USA)
NMM	4-Methylmorpholine	67870	Sigma-Aldrich (Steinheim, Germany)
NMP	1-Methyl-2-pyrrolidinone	M79603	Sigma-Aldrich (Steinheim, Germany)
Phenol	Phenol, 99%, extra pure	18078	Acros Organics, Thermo Fisher Scientific (Waltham, MA, USA)
Piperidine	Piperidine, 99%, extra pure	147180010	Acros Organics, Thermo Fisher

			Scientific (Waltham, MA, USA)
Resin	Fmoc-rink amide AM resin	-	GL Biochem (Shanghai, China)
TCEP	Tris(2-carboxyethyl)phosphine Hydrochloride	T1656	TCI (Zwijndrecht, Belgium)
TFA	Trifluoroacetic acid	T6508	Sigma-Aldrich (Steinheim, Germany)
TFA for HPLC	Trifluoroacetic acid, HPLC grade	T/3258/PB05	Thermo Fisher Scientific (Waltham, MA, USA)
Thioanisole	Thioanisole	T0191	TCI (Zwijndrecht, Belgium)

5.1.2 Peptide synthesis building blocks

Abbreviation	Name	Product code	Supplier
5(6)-FAM	5(6)-Carboxyfluorescein	40410	Acros Organics, Thermo Fisher Scientific (Waltham, MA, USA)
Bis(bromo-methyl)pyridine	2,6-Bis(bromomethyl)pyridine, 98%	405426	Sigma-Aldrich (Steinheim, Germany)
Boc-Pro	Boc-L-Pro-OH	BAA1130	Iris Biotech GmbH (Marktredwitz, Germany)
Boc-Gly	Boc-Gly-OH	BAA1062	Iris Biotech GmbH (Marktredwitz, Germany)
Dibromo-acetone	1,3-Dibromoacetone, 95%	OR16995	Apollo Scientific (Manchester, UK)
Dichloro-acetone	1,3-Dichloroacetone, 99%	173630	Acros Organics, Thermo Fisher Scientific (Waltham, MA, USA)
γ-Glu	Fmoc-Glu-OtBu	045463	Fluorochem (Hadfield, UK)
hArg	Fmoc-L-HomoArg(Pbf)-OH	M03211	Fluorochem (Hadfield, UK)
Lys(Dde)	Fmoc-L-Lys(Dde)-OH	FAA1390	Iris Biotech GmbH (Marktredwitz, Germany)
Nle	Fmoc-L-Nle-OH	FAA1413	Iris Biotech GmbH (Marktredwitz, Germany)
Palmitic acid	Palmitic acid	P0002	TCI (Zwijndrecht, Belgium)
Palmitic acid NHS ester	Palmitic acid N-hydroxysuccinimide ester	P1162	Sigma-Aldrich (Steinheim, Germany)
PEG₂	Fmoc-O2Oc-OH	FAA1435	Iris Biotech GmbH (Marktredwitz, Germany)
SiR-NHS ester	6-Carboxy-silicon rhodamine N-hydroxysuccinimide ester		Spirochrome (Lausanne, Switzerland)
-	4-(tert-butoxy)-4-oxobutanoic acid	AB285664	abcr (Karlsruhe, Germany)

-	5-(tert-Butoxy)-5-oxopentanoic acid	210043	Fluorochem (Hadfield, UK)
-	3-(tert-Butoxycarbonyl)benzoic acid	AB245675	abcr (Karlsruhe, Germany)
-	4-Fluorobenzoic acid	128384	Sigma-Aldrich (Steinheim, Germany)
-	Biphenyl-3-carboxylic acid	715204	Sigma-Aldrich (Steinheim, Germany)
-	1-Naphthoic acid	N1909	Sigma-Aldrich (Steinheim, Germany)
-	2-Naphthoic acid	180246	Sigma-Aldrich (Steinheim, Germany)
-	Diphenylacetic acid	D204307	Sigma-Aldrich (Steinheim, Germany)

Other, non-canonical amino acids used for screening were purchased from abcr, Ark Pharm, Bachem, Fluorochem, Iris Biotech, Sigma-Aldrich and TCI.

5.1.3 Biochemical reagents

Abbreviation	Name	Product code	Supplier
2xYT	YT-Medium powder (2X)	A0981	Applchem (Darmstadt, Germany)
Agar-Agar	Agar-Agar	6494.3	Carl Roth (Karlsruhe, Germany)
Agarose	Ultrapure Agarose	16500500	Thermo Fisher Scientific (Waltham, MA, USA)
Amp	Ampicillin Sodium Salt	BP1760-25	Thermo Fisher Scientific (Waltham, MA, USA)
BSA	Albumin Fraction V (pH 7.0)	A1391	Applchem (Darmstadt, Germany)
CaCl₂	Calcium chloride, dehydrated	CN93	Carl Roth (Karlsruhe, Germany)
Coomassie blue	Coomassie® Brilliant Blue R-250	A1092	Applchem (Darmstadt, Germany)
Cryomatrix	Shandon™ Cryomatrix™ embedding resin	6769006	Thermo Fisher Scientific (Waltham, MA, USA)
DAPI	2-(4-Amidinophenyl)-6-indolecarbamidine dihydrochloride	D9542	Sigma-Aldrich (Steinheim, Germany)
DTT	Dithiothreitol	A1101	Applchem (Darmstadt, Germany)
EDTA	Ethylenediaminetetraacetic acid for molecular biology	A5097	Applchem (Darmstadt, Germany)
EPPS	EPPS	AB116401	abcr (Karlsruhe, Germany)
Glucose	D(+)-Glucose anhydrous	A1422	Applchem (Darmstadt, Germany)
Glycerol	Glycerol 87 % pure	A0970	Applchem (Darmstadt, Germany)
Glycine	Glycine for molecular biology	A1067	Applchem (Darmstadt, Germany)
GND	Guanidinium hydrochloride	G0162	TCI (Zwijndrecht, Belgium)
Imidazole	Imidazole for buffer solutions	A1073	Applchem

Kan	Kanamycinsulfat	A1493	(Darmstadt, Germany) Applichem (Darmstadt, Germany)
MgCl₂	Magnesium chloride, anhydrous	KK36	Carl Roth (Karlsruhe, Germany)
Mounting medium	Fluoromount-G®	0100-01	SouthernBiotech (Birmingham, AL, USA)
NaCl	Natriumchlorid	3957.1	Carl Roth (Karlsruhe, Germany)
Na₂HPO₄	di-Sodium hydrogen phosphate 2-hydrate	A3905	Applichem (Darmstadt, Germany)
NH₄HCO₃	Ammonium bicarbonate	11213	Sigma-Aldrich (Steinheim, Germany)
NHS-Biotin	EZ-Link™ Sulfo-NHS-LC-Biotin	21335	Thermo Fisher Scientific (Waltham, MA, USA)
PBS	PBS, pH 7.4, - CaCl ₂ , - MgCl ₂	10010023	Gibco, Thermo Fisher Scientific (Waltham, MA, USA)
PEG 6000	Polyethylene Glycol 6000	A1387	Applichem (Darmstadt, Germany)
SDS	Sodium Dodecyl Sulfate, Lauryl	28365	Thermo Fisher Scientific (Waltham, MA, USA)
SDS gels	RunBlue™ TEO-Tricine SDS Gels	NXG01212	Expedeon, Sygnis (Heidelberg, Germany)
SDS sample buffer	RunBlue LDS Sample Buffer (TEO-Tricine)	NXB31010	Expedeon, Sygnis (Heidelberg, Germany)
TE buffer	TE buffer, pH 7.4, RNase free	J60234	Alfa Aesar, Thermo Fisher Scientific (Waltham, MA, USA)
Tris	Tris for buffer solutions	A1379	Applichem (Darmstadt, Germany)
Tris-HCl	Tris Hydrochloride for buffer solutions	A1087	Applichem (Darmstadt, Germany)
Triton X-100	Triton® X-100 molecular biology grade	A4975	Applichem (Darmstadt, Germany)
Tween 20	Tween® 20 for molecular biology	A4974	Applichem (Darmstadt, Germany)

5.1.4 Proteins & substrates

Abbreviation	Name	Product code	Supplier
EK	Enterokinase from porcine intestine	E0885	Sigma-Aldrich (Steinheim, Germany)
HSA	Albumin from human serum	A3782	Sigma-Aldrich (Steinheim, Germany)
Human serum	Human serum	H4522	Sigma-Aldrich (Steinheim, Germany)
Human plasma	Single Donor Human Plasma Anticoagulant: Na Citrate	IPLA-N-S	Innovative Research, (Novi, MI, USA)
mKLK5	Recombinant Mouse Kallikrein 5	7236-SE-010	R&D Systems, Bio-Techne (Minneapolis, MN, USA)

mKLK7	Recombinant Mouse Kallikrein 7	7688-SE-010	R&D Systems, Bio-Techne (Minneapolis, MN, USA)
MSA	Albumin from mouse serum	MSA-100MG	Molecular Innovations (Novi, MI, USA)
NeutrAvidin	NeutrAvidin Protein	31000	Thermo Fisher Scientific (Waltham, MA, USA)
-	Dynabeads™ M-280 Tosylactivated	14204	Thermo Fisher Scientific (Waltham, MA, USA)
-	Dynabeads™ M-280 Streptavidin	11205D	Thermo Fisher Scientific (Waltham, MA, USA)
Substrate VPR-AMC	Boc-Val-Pro-Arg-AMC	I-1120	Bachem (Bubendorf, Switzerland)
Substrate LLVY-AMC	Suc-Leu-Leu-Val-Tyr-AMC	I-1395	Bachem (Bubendorf, Switzerland)

5.1.5 Kits & consumables

Abbreviation	Name	Product code	Supplier
Centrifugal concentrators	Macrosep® Advance, 10K MWCO	MAP010C38	Pall Corporation (New York, USA)
Desalting columns	PD 10 Desalting Columns	17085101	GE Healthcare (Chicago, USA)
Desalting columns	Zeba Spin Desalting Columns, 7K MWCO, 5 mL	89892	Thermo Fisher Scientific (Waltham, MA, USA)
Gel extraction	QIAquick Gel Extraction Kit	28115	Quiagen (Venlo, Netherlands)
IMAC column	HisTrap™ excel, 1 ml	17371205	GE Healthcare (Chicago, USA)
Maxiprep	NucleoBond® Xtra Maxi	740414	Macherey-Nagel (Düren, Germany)
Miniprep	NucleoSpin® Plasmid	740499	Macherey-Nagel (Düren, Germany)
Microvettes	Microvette® 100 µL, EDTA	20.1278	Sarstedt (Nümbrecht, Germany)

5.2 Protein expression

5.2.1 Expression and purification of KLK5 and KLK7

For KLK5, which can autoactivate, the protein sequences of human ProKLK5 (UniProtKB Q9Y337, Asn30-Ser293) was codon optimized for mammalian expression using a codon optimization tool (Integrated DNA Technologies). The DNA sequence additionally encoding a C-terminal poly-histidine tag was cloned into pEXPR-IBA42 using *NheI* and *HindIII* restriction sites for expression in mammalian cells as a secreted protein. The following coding sequence was expressed:

MRAWIFFLLCLAGRALAASNNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDDSSRIINGSDCDMHTQPWQAA
LLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLSPVYESGQQMFQGVKSIPHPGYSHPGHSNDLMLI
KLNRRIPTKDVPRPINVSSHCPASAGTKCLVSGWGTTKSPQVHFQVQLCLNISVLSQKRCEDAYPRQIDDTMFCA
GDKAGRDSQCQDGGPVCNGLSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANSSGSGSHHHHHH
 (BM40 secretory signal peptide, mature protease).

The protein was expressed by transient transfection of Chinese hamster ovary (CHO) cells cultured in solution (EPFL Protein Expression Core Facility, Dr. David Hacker) as follows: 1.5 mg plasmid DNA was transfected into 500 mL CHO cell suspension culture using polyethylenimine (PEI). Cells were incubated for seven days at 37°C, 5% CO₂ under shaking conditions and then were removed by centrifugation. hKLK5 was purified from the supernatant (pH adjust to 8.0 with NaOH) by nickel-charged immobilized metal affinity chromatography (IMAC) using a HisTrap™ Excel column and in below shown binding and elution buffers. The purification was done according to the manufactures suggested protocol using a one-step elution procedure.

For KLK7, which needs to be activated by a different protease, the protein sequences of mature hKLK7 (UniProtKB P49862, Ile30-Arg253) was codon optimized for mammalian expression and the endogenous pro-domain was replaced by an enterokinase cleavage site (EKcs, Val-Asp-Asp-Asp-Lys) or by the pro-domain of hKLK5, with the last two amino acids change to Leu-Tyr. The DNA sequence additionally encoding a C-terminal poly-histidine tag was cloned into pEXPR-IBA42 using *NheI* and *HindIII* restriction sites for expression in mammalian cells as a secreted protein. The following coding sequences were expressed:

EKcs-hKLK7:

MRAWIFFLLCLAGRALAASMQIFVKTLTGKTITLEVEPSDTIENVKGKIQEKEGIPPDQQRILIFAGKQLEDGRTLSDYN
IQKESTLHLVLRRLGGADPVDDDDKIIDGAPCARGSHPWQVALLSGNQLHCGGVLVNERWVLTAAHCKMNEYT
VHLGSDTLGDRRAQRIKASKSFRHPGYSTQTHVNDLMLVKLNSQARLSSMVKKVRLPSRCEPPGTTCTVSGWG
TTTSPDVTFPSDLMCVDVKLISPQDCTKVYKDLENSMLCAGIPDSKKNACNGDSGGPLVCRGTLQGLVSWGTF
PCGQPNDPGVYTQVCKFTKWINDTMKKHRGSGSHHHHHH (BM40 secretory signal peptide, bovine ubiquitin, mature protease).

Pro5-hKLK7:

MRAWIFFLLCLAGRALAASNNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDDSSLYIIDGAPCARGSHPWQVAL
LSGNQLHCGGVLVNERWVLTAAHCKMNEYTVHLGSDTLGDRRAQRIKASKSFRHPGYSTQTHVNDLMLVKLNS
SQARLSSMVKKVRLPSRCEPPGTTCTVSGWGTTTSPDVTFPSDLMCVDVKLISPQDCTKVYKDLENSMLCAGIP
DSKKNACNGDSGGPLVCRGTLQGLVSWGTFPCGQPNDPGVYTQVCKFTKWINDTMKKHRGSGSHHHHHH
 (BM40 secretory signal peptide, mature protease)

Both proteins were expressed by transient transfection of CHO cells cultured in solution as described above and were purified from the supernatant by IMAC. EKcs-hKLK7 was activated by treatment with enterokinase (0.04 U / μ g KLK7) at 37 °C overnight in activation buffer and purified again by Ni-NTA affinity chromatography.

The buffer of the from the column eluted proteases was exchanged to storage buffer and the concentration adjusted to around 1 mg/mL using centrifugal concentrators (MacroSep Advance 10K MWCO, Pall Corporation). Proteins were frozen in liquid nitrogen and stored at -80 °C. 5 μ g protein was analyzed by SDS-PAGE under reducing conditions.

Binding Buffer	15 mM imidazole 200 mM NaCl 10 mM EPPS pH 8.0, filtered
Elution Buffer	500 mM imidazole 200 mM NaCl 10 mM EPPS pH 8.0, filtered
KLK storage buffer	200 mM NaCl 10 mM EPPS pH 8.0, filtered
Activation buffer	20 mM Tris-HCl 50 mM NaCl 2 mM CaCl ₂ pH 8.0, filtered

5.2.2 Biotinylation of targets

Recombinantly expressed hKLK5 and hKLK7 were biotinylated by incubating the proteins at a concentration of 10 μ M with 20 equiv of Sulfo-NHS-LC-Biotin in an amine-free buffer (KLK storage buffer, pH 8.0). Sulfo-NHS-LC-Biotin was dissolved at 10 mg/mL in DMSO and added to the reaction mix with a final concentration of 200 μ M. Then the protein was added with final concentration of 10 μ M and incubated for 1h at RT. Excess biotinylation reagent was removed by size exclusion chromatography using a disposable PD-10 desalting column (GE Healthcare). 10 μ g aliquots of biotinylated protein were prepared, frozen in liquid nitrogen and stored at -80 °C. To check for successful biotinylation, a capture assay was performed. 25 μ L streptavidin magnetic beads were washed twice with 250 μ L W1 buffer and resuspended in 25 μ L W1 buffer. 1 μ g of biotinylated or unbiotinylated control protein was diluted to 0.2 mg/mL with KLK storage buffer, added to the beads and incubated for 30 min on a rotating wheel at RT. The supernatant was separated using a magnetic rack. Beads were washed 2x with 250 μ L W1 buffer and finally resuspended in 25 μ L W1 buffer. Samples were complemented with DDT containing SDS loading buffer and cooked for 10 min at 95 °C.

5.3 Phage display selections

5.3.1 Phage display library

A library of the format $XCX_mCX_nCX_oCX$ ($m + n + o = 12$) previously developed in our lab by Carle *et al.* (unpublished) was used for the selections. For the library the phagemid pSEX81 (Progen) was used, which is a vector for the expression of scFv libraries as pIII fusion proteins on the surface of M13 filamentous phage. The peptide library was cloned N-terminal of the pIII gene connected through a Ser-Gyl-Ser linker, replacing the originally in the plasmid encoded scFv gene. The library was generated with 91 separate whole vector PCRs using 91 different degenerated forward primers with the sequence

5'-AGCGCCATGGCCNNKTGT(NNK)_mTGT(NNK)_nTGT(NNK)_oTGTNNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC-3'

and the reverse primer with the sequence

5'-CATGCCATGGCCGGCTGAGCTGCC-3'.

The PCR reactions were run on a gel, extracted, digested with *Nco*I, self-ligated, and transformed into electrocompetent TG1 *E. coli* cells.

5.3.2 Phage production and modification

The phage library was inoculated with an $OD_{600} = 0.1$ in 1 L 2xYT completed with 100 µg/mL ampicillin and 100 mM glucose and grown at 200 rpm, 37 °C. At $OD_{600} = 0.5$ the culture was infected with Hyperphage (Progen) with a MOI of 10 and incubated for 15 min at 37 °C without shaking. After infection, the culture was incubated another 45 min, 250 rpm, 37 °C to allow expression of the additional resistance gene. To remove the glucose, the bacteria were pelleted by centrifugation for 10 min, 2000 g, 4 °C and resuspended in 1 L of 2xYT with 100 µg/mL ampicillin and 50 µg/mL kanamycin. Phages were produced overnight at 200 rpm, 30 °C. In the second and third round the volume of the expression culture was reduced to 100 mL.

Reduction was performed for 30 min at RT with 1 mM TCEP. A 20 mM TCEP stock solution in H₂O was prepared, aliquoted and stored at 20 °C. 500 µL TCEP stock solution was added to 10 mL phage suspension. The tubes were inverted several times and incubated for 30 min at RT. Phages are precipitated by addition of ¼ volume of ice-cold buffer N and subsequent incubation on ice for 15 min. The phages are pelleted by centrifugation for 15 min, 4500 rpm, 4 °C. The supernatant was removed completely, and phage pellets were resuspended in 9 mL buffer R by pipetting.

Cyclization was performed with final linker concentrations of 40 µM 2,6-bis(bromomethyl)pyridine or 1,3-dibromoacetone for 1 h at RT. 40 mM linker stock solutions were prepared in ACN, then diluted to 400 µM with 90% H₂O / 10 % ACN and 1 mL was added to 9 mL phage suspension. The tubes are inverted several times and incubated in a water bath for 1 h at 30 °C. Phages were precipitated by addition of ¼ volume of ice-cold buffer N and incubation for 15 min on ice. The phages are pelleted by centrifugation for 15 min, 4500 rpm, 4 °C. The supernatant was removed completely, and phage

pellets were resuspended in 5 mL buffer W2 by pipetting. Modified phage suspensions were stored at 4 °C o/n.

2xYT medium	31.6 g/L 2xYT powder in H ₂ O autoclaved
Agar plates	31.6 g/L 2xYT powder in H ₂ O + 15 g/L agar-agar autoclaved
Glucose	1 M glucose autoclaved
Ampicillin (1000x)	100 mg/mL ampicillin in H ₂ O
Kanamycin (1000x)	50 mg/mL kanamycin in H ₂ O

5.3.3 Biopanning

Per target and linker either 50 µL streptavidin magnetic beads (first and third round) or 20 µL neutravidin magnetic beads (second round) were used. Neutravidin beads were generated by coupling neutravidin to tosyl-activated magnetic beads according to the manufacturer's recommended protocol. Beads were washed 2x with W1 buffer and resuspended in 50 µL buffer W1. 10 µg, 5 µg or 2.5 µg (first, second and third round, respectively) recombinantly expressed hKLK5 or hKLK7 was added to the beads and incubated for 10 min on a rotating wheel. Beads were washed 3x with 1 mL buffer W1. For blocking, the beads were incubated in 500 µL buffer W2 for 30 min at 5 rpm on a rotating wheel.

The beads were added to 5 mL modified phage suspension and incubated for 30 min at 5 rpm on a rotating wheel. The bead-phage suspension was added stepwise to a 2 mL tube in a magnetic rack and the supernatant was discarded after 2 min in the rack. The beads were subsequently washed 3x with 1 mL buffer W4 and 6x with 1 mL buffer 1. The tube was changed after every third wash to avoid carrying along phages attached to the surface of the tube.

For elution, beads were resuspended in 100 µL buffer G and incubated for 5 min. Afterwards the beads were placed in a magnetic rack the supernatant was transferred into a new tube containing 100 µL buffer F.

Eluted phages were added to a TG1 *E. coli* culture with OD₆₀₀ = 0.5 and incubated for 30 min at 37 °C without shaking. The infected cells were centrifuge for 10 min, 3000 g, 4 °C and resuspended in 1 mL 2xYT medium by pipetting. The resuspended cells were plated on 15 cm 2xYT plates with 100 µg/mL ampicillin.

The next day, cells were harvested with 4 mL 2xYT medium and glycerol was added to a final concentration of 10 %. Aliquots of 500 µL were frozen in liquid nitrogen, stored at -80 °C and used for subsequent selection rounds or sequencing.

Buffer R	20 mM NH ₄ HCO ₃ 5 mM EDTA pH 8.0, degassed
Buffer N	20 % PEG-6000 2.5 M NaCl store at 4°C
Buffer W1	10 mM Tris-HCl 150 mM NaCl 10 mM MgCl ₂ 1 mM CaCl ₂ pH 7.4 (HCl adj.)
Buffer W2	= Buffer W1 + 1 % BSA + 0.1 % Tween
Buffer W4	= Buffer W1 + 0.1 % Tween
Buffer G	50 mM glycine pH 2.2
Buffer F	1 M Tris-HCl pH 8.0

5.3.4 Phage sequencing

15 µL of glycerol stocked *E. coli* were suspended in 1.5 mL 2xYT and the OD₆₀₀ was measured. The cells were diluted to 500 cells/mL in serial dilutions and 1 mL was plated on a 15 cm 2xYT plates with 100 µg/mL ampicillin. After overnight incubation at 37 °C, 96 clones are picked manually and inoculated in 100 µL 2xYT with 100 µg/mL ampicillin in a 96-well plate. The cells were incubated for 2-3 h at 37 °C under shaking conditions and 2 µL of the culture was used as template for the subsequent PCR reaction.

For the PCR reaction, the two primers pSEX81_seq0 (5'-GTGTGGAATTGTGAGCGGATAAC-3') and B-insert (5'-CACCACCAGAGCCGCCGCCAGCATTGACAGGAGGTTGAG-3') were used.

The PCR plate was spot-checked on an agarose gel and the amplified inserts were subsequently sent for Sanger Sequencing (Macrogen, Amsterdam, Netherlands). For the sequencing reaction, the primer pSEX81_seq0 was used. For sequence analysis the alignment tool MEGA5 was used¹⁵¹. Consensus sequences were further analyzed manually.

PCR mix	3975 µL H ₂ O
96 reactions	500 µL 5x buffer 100 µL dNTPs (10 mM) 25 µL Taq polymerase 100 µL primer 1 (10µM) 100 µL primer 2 (10µM)
1 reaction	48 µL +2 µL template

5.4 Peptide synthesis

5.4.1 Automated peptide synthesis

Solid phase peptide synthesis (SPPS) was carried out on 25 μ M or 50 μ M scale using an Intavis MultiPep RSi peptide synthesizer on rink amide AM resin (loading 0.3 mmol/g) with standard Fmoc (9-fluorenylmethyloxycarbonyl) chemistry and DMF as solvent. N $^{\alpha}$ -deprotection was performed twice with 800 μ L of 20% (v/v) piperidine for 5 min, 500 rpm. The deprotection step was followed by seven washings of the resin.

Fmoc-protected amino acids were dissolved in DMF (4.2 equiv, 0.5 M dissolved, 0.2 M final concentration) and coupling was performed twice for 30 min at 500 rpm using a HATU/HOBt mixture (4 equiv each, 0.5 M dissolved, 0.2 M final concentration) as coupling reagents, NMM as base (8 equiv, 4 M dissolved, 0.4 M final concentration) and a catalytic amount of NMP. Coupling was performed twice for 30 min at 500 rpm. To prevent the formation of truncated peptides, a capping step was carried out using a solution of 5 % (v/v) acetic anhydride and 6% (v/v) 2,6-lutidine in DMF (5 min, 800 μ L). The capping step was followed by seven washings of the resin.

Branches in the peptide were introduced with orthogonally protected Fmoc-Lys(Dde)-OH. Dde is stable under Fmoc-deprotection conditions and was removed with 2% (v/v) hydrazine hydrate in DMF (4 mL, 3 min, 700 rpm, 3 cycles) followed by five washings of the resin. As 2% hydrazine hydrate would also remove Fmoc-groups, the prior synthesized branch was finished with a N $^{\alpha}$ -Boc-protected amino acid as final, N-terminal amino acid.

5.4.2 Fluorescein and fatty acid conjugation on solid phase

If not indicated differently, coupling of 5(6)-carboxyfluorescein (5(6)-FAM) or palmitic acid to the N-terminus or the side chain of a lysine was performed on solid phase with DMF as solvent. 3.5 equiv (0.05 M dissolved, 0.05 M final concentration) were coupled twice for 90 min at 700 rpm using a HATU/HOBt mixture (3 equiv each, 0.5 M dissolved, 0.04 M final concentration) as coupling reagents, NMM as base (6 equiv, 4 M dissolved, 0.05 M final concentration) and a catalytic amount of NMP.

5.4.3 Peptide cleavage and deprotection

Before cleavage, the resin was washed twice with DCM. Peptides were cleaved from the resin under reducing conditions by addition of 5 mL cleavage mixture for 2 h under shaking conditions. The cleavage mixture was separated from the resin by vacuum filtration. For fatty acid-coupled peptides a second cleavage round was performed with another 5 mL cleavage mixture for 2 h under shaking conditions while the first 5 mL were still incubated at RT separated from the resin. The mixture was again separated by vacuum filtration and pooled with the previously filtered 5 mL. Peptides were precipitated with 40 mL ice-cold diethyl ether and incubated for 30 min at -20 $^{\circ}$ C. Precipitated peptides were pelleted by centrifugation (10 min, 4500 rpm, 4 $^{\circ}$ C). The supernatant was removed, and the pellet was washed twice with 30 mL and 20 mL diethyl ether.

Cleavage mixture	90 % (v/v) TFA 2.5 % (v/v) 1,2-ethanedithiol 2.5 % (w/v) phenol 2.5 % (v/v) thioanisole 2.5 % (v/v) H ₂ O
-------------------------	--

5.4.4 Palmitic acid and silicon rhodamine coupling in solution

For solution coupling the preactivated N-hydroxysuccinimide (NHS) ester of palmitic acid or 6-carboxy silicon rhodamine (SiR) were used. SiR-NHS ester was kindly provided by Dr. Luc Reymond (EPFL, Switzerland).

For palmitic acid coupling 10 μ mol crude peptide was dissolved in 0.5 mL DMF with 5% (v/v) DIPEA. 4 equiv palmitic acid NHS ester dissolved in 0.5 mL DMF was added. The final concentrations in the reaction mixture were 10 mM peptide, 40 mM palmitic acid NHS ester, 2.5% (v/v) DIPEA. The reaction mixture was incubated overnight at 30°C under shaking conditions and purified by RP-HPLC the next day.

For SiR coupling 1 μ mol pure peptide was dissolved in 0.5 mL DMSO with 5% (v/v) DIPEA and 1.2 equiv SiR-NHS ester dissolved in 0.5 mL DMSO was added. The final concentrations in the reaction mixture were 1 mM peptide, 1.2 mM SiR-NHS ester, 2.5% (v/v) DIPEA. The reaction mixture was incubated for 1 h at RT and subsequently purified by RP-HPLC.

5.4.5 Cyclization with thiol-reactive linkers

50 μ mol peptide was dissolved in 9 mL 40% ACN and 60% buffer (buffer R synthesis) and 1.5 equiv 2,6-bis(bromomethyl)pyridine or 1,3-dichloroacetone dissolved in 1 mL ACN was added. The final concentrations in the reaction mixture were 5 mM peptide, 7.5 mM 2,6-bis(bromomethyl)pyridine or 1,3-dichloroacetone, 50% aqueous buffer, and 50% ACN. For not conjugated to a fatty acid, 75% aqueous buffer and 25% ACN was used. The pH was checked and readjusted to 8.0 with NH₄HCO₃, the reaction was incubated at 30°C for 1 h and stopped with 200 μ L acetic acid. The solution was lyophilized, and peptides purified by RP-HPLC.

Buffer R synthesis	60 mM NH ₄ HCO ₃ 5 mM EDTA pH 8.0, degassed
---------------------------	---

5.4.6 Synthesis of para-nitroanilide substrates

The synthesis of the KLK7 substrate KHLy-pNA was performed according to Abbenenate *et al.*¹⁵². The synthesis scheme is shown in **Figure S3**. In difference to the published protocol, 2-chlorotriyl chloride resin was used. The resin was loaded overnight with 1,4-phenylenediamine (8 equiv) and DIPEA (8 equiv) in DMF. After washing with DMF, unreacted sites were capped with a mixture of DMF / MeOH (9 : 1) for 30 min. Then the resin was washed with DCM, dried and standard SPPS was performed to couple the tetrapeptide Lys-His-Leu-Tyr. The peptide was cleaved

with DCM / TFA (99 : 1) for 2 h, the resin was removed by filtration and DCM was evaporated under reduced pressure. The peptide was redissolve in ACN / H₂O (1 : 1) and oxidized with 6 equiv Oxone[®], stirring at RT for 48 h. After, the peptide was extracted into ethyl acetate, which was removed through evaporation under reduced pressure. Cleavage was performed with the standard cleavage mixture for 1 h at RT. Peptides were precipitated with 40 mL ice-cold diethyl ether and incubated for 30 min at -20 °C. Precipitated peptides were pelleted by centrifugation (10 min, 4500 rpm, 4°C). The supernatant was removed, and the pellet was washed once with 30 mL diethyl ether, dried and purified by RP-HPLC.

5.4.7 Peptide purification

Purification was performed by semi-preparative RP-HPLC (PrepLC 4000-Waters system) on a C18 column (Vydac C18 TP1022, 250 x 22 mm, 10 µm, 300 Å) with a linear solvent gradient of solvent B (ACN, 0.1% TFA) in solvent A (H₂O, 0.1% TFA) from 0 – 40 % in 40 min (mono- and bicyclic inhibitors) or 30 – 60 % in 30 min (inhibitor-tag conjugates) or 50 – 90 % in 40 min (albumin tags) at a flow rate of 20 mL/min. Fractions were collected automatically by absorbance at 220 nm. Product containing fractions were identified by ESI-MS (Shimadzu LCMS-2020), pooled and lyophilized.

5.5 Labeling of biomolecules

Anti-FLAG mouse IgG (mIgG) was provided by the EPFL Protein Expression Core Facility (Dr. David Hacker). Labeling of proteins was performed through acylation at primary amines using the 6-carboxy-SiR-NHS ester, kindly provided by Dr. Luc Reymond (EPFL, Switzerland). An amine-free buffer at pH 8.2 - 8.4. (labeling buffer) was used. The proteins were diluted to 1 mg/mL in labeling buffer. The acylating reagent was dissolved to 10 mM in anhydrous, amine-free DMSO and a 5-fold molar excess to the protein was added, aiming for a degree of labeling (DOL) of one. The pH of the reaction mixture was checked to be 8.2 – 8.4 and the mixture was incubated for 3 h at 4 °C, gently shaking. Excess acylating reagent was removed by passing the antibody conjugate through a Zeba Spin Desalting Column (7K MWCO) using PBS for equilibration and elution. Antibody was concentrated to 30 mg/mL and MSA to 13 mg/mL using centrifugal concentrators (10K MWCO). 0.1 nmol conjugate was analyzed by SDS-PAGE under reducing conditions.

Labeling buffer	0.1 M NaHCO ₃ 1% (v/v) DMSO pH 8.2 - 8.4
------------------------	---

The protein concentration and the DOL were spectrophotometrically measured and calculated using **Equation 1** and **Equation 2**.

Equation 1 | Mass concentration of antibody-SiR conjugates. A: Absorbance; $\epsilon_{280\text{ nm}}$: molar extinction coefficient of protein ($\text{M}^{-1}\text{cm}^{-1}$); CF: correction factor of dye; d: pathlength (cm); M: molecular weight (g/mol); CF = 0.147 for SiR; $\epsilon_{280\text{ nm}} = 210,000$ and M = 150,000 for IgG.

$$\gamma \text{ (mg/mL)} = \frac{A_{280\text{ nm}} - (CF * A_{652\text{ nm}})}{\epsilon_{280\text{ nm}} * d} * M$$

Equation 2 | Degree of labeling (DOL) of antibody-SiR conjugates. A: Absorbance; $\epsilon_{652\text{ nm}}$: molar extinction coefficient of dye ($\text{M}^{-1}\text{cm}^{-1}$); d: pathlength (cm); γ : mass concentration of conjugate (mg/mL); M: molecular weight (g/mol); $\epsilon_{652\text{ nm}} = 100,000$ for SiR; M = 150,000 for IgG.

$$DOL = \frac{A_{652\text{ nm}}}{\epsilon_{652\text{ nm}} * d * \frac{\gamma}{M}}$$

5.6 In vitro characterization

5.6.1 Analysis of peptide mass and purity

The mass of synthesized peptides was verified by ESI-MS in positive ion mode on a single quadrupole LC-MS (Shimadzu LCMS-2020). The purity was determined by RP-HPLC (Agilent 1260 HPLC system) using a C18 column (Agilent Zorbax 300SB-C18, 4.6 mm x 250 mm, 5 μ m) with a linear gradient of solvent B (94.9% ACN, 5% H₂O, 0.1% TFA) in solvent A (94.5% H₂O, 5% ACN, 0.1% TFA) from 0 – 50% in 15 min (mono- and bicyclic inhibitors) or 0 – 100% in 30 min (inhibitor-tag conjugates and tags) at a flow rate of 1 mL/min. Peptide stock solutions were prepared by weight and concentrations were validated using a microvolume spectrophotometer (NanoDrop™ 8000, Witec AG, Sursee, Switzerland) for peptides either containing aromatic residues (Tyr: $\epsilon_{280\text{ nm}} = 1490\text{ M}^{-1}\text{cm}^{-1}$, Trp: $\epsilon_{280\text{ nm}} = 5500\text{ M}^{-1}\text{cm}^{-1}$), fluorescein ($\epsilon_{495\text{ nm}} = 68,000\text{ M}^{-1}\text{cm}^{-1}$) or SiR ($\epsilon_{652\text{ nm}} = 100,000$). For fluorescein containing peptides the measurement was performed at a concentration of around 0.5 mM at pH 7.4 (diluted in TE buffer). Concentrations were calculated using **Equation 3**. 1 mM stock solutions in H₂O were prepared for unconjugated peptides and 10 mM stock solutions in H₂O / DMSO (90 : 10) were prepared for peptides with albumin tag.

Equation 3 | Molar concentration according to Lambert–Beer. A: Absorbance; ϵ : molar extinction coefficient ($\text{M}^{-1}\text{cm}^{-1}$); d: pathlength (cm);

$$c\text{ (M)} = \frac{A}{\epsilon * d}$$

5.6.2 Protease inhibition assays

The K_i values of the inhibitors were determined in residual enzymatic activity assays with varying peptide concentrations.

Table 1 | KLK activity assay conditions.

	KLK5 assay	KLK7 assay
Peptide	200 nM to 0.2 nM	4000 nM to 4 nM
Enzyme	1 nM	10 nM
Substrate	Boc-VPR-AMC: 100 μ M $K_m = 216\text{ }\mu$ M	KHLY-pNA: 400 μ M $K_m = 107\text{ }\mu$ M

For the KLK5 inhibition assay, peptide stock solutions were diluted to 0.6 μ M in KLK assay buffer and ten 2-fold serial dilutions were prepared. The fluorogenic substrate Boc-VPR-AMC (100 mM stock in DMSO) was diluted to 300 μ M and hKLK5 to 3 nM in KLK assay buffer. 50 μ L peptide, 50 μ L enzyme and 50 μ L substrate were pipetted into a transparent, 96 well flat-bottom plate resulting in final concentrations of 200 nM to 0.2 nM peptide, 1 nM enzyme and 100 μ M substrate in the assay. The residual enzymatic activity was measured as rate of AMC formation recorded with a Tecan plate reader (Tecan Infinite M200 Pro) by emission at 467 nm and excitation at 368 nm at 25 °C over 30 min.

For the KLK7 inhibition assay, peptide stock solutions were diluted to 12 μM in KLK assay buffer and ten 2-fold serial dilutions were prepared. The chromogenic substrate KHLY-pNA (100 mM stock in DMSO) was diluted to 1200 μM and hKLK7 to 30 nM in KLK assay buffer. 50 μL peptide, 50 μL enzyme and 50 μL substrate were pipetted into a transparent, 96 well flat-bottom plate resulting in final concentrations of 4 μM to 4 nM peptide, 10 nM enzyme and 400 μM substrate in the assay. The residual enzymatic activity was measured as rate of pNA formation recorded with a Tecan plate reader (Tecan Infinite M200 Pro) by absorbance at 405 nm at 25 $^{\circ}\text{C}$ over 30 min.

KLK assay buffer	200 mM NaCl 10 mM TRIS 0.01 % (v/v) Triton-X100 pH 8.0, filtered 0.1 % (m/v, 1 mg/ml) BSA, freshly added
-------------------------	--

The K_i values of inhibitors with albumin tag were measured with a final concentration of 25 μM HSA in the assay buffer instead of 0.1% (m/v) BSA.

For the specificity assay with different proteases the conditions shown in **Tab. X.** and peptide concentration from 128 μM to 1 nM were used. Enzymes were bought from Molecular Innovations or Innovative Research and substrates from Bachem.

Table 2 | Specificity assay substrates and concentrations.

Enzyme	[E] (nM)	Substrate	[S] (nM)	K_m (μM)
β FXIIa	2	Z-Gly-Gly-Arg-AMC	50	180
Trypsin	0.1	Z-Gly-Gly-Arg-AMC	50	90
tPA	7.5	Z-Gly-Gly-Arg-AMC	50	359
uPA	1.5	Z-Gly-Gly-Arg-AMC	50	122
Thrombin	2	Z-Gly-Gly-Arg-AMC	50	168
PK	0.25	Z-Phe-Arg-AMC	50	165
Plasmin	2.5	H-D-Val-Leu-Lys-AMC	50	323

All data were fitted using GraphPad Prism software and **Equation 4** was used to calculate K_i values. The K_m value of Boc-VPR-AMC on hKLK5 was measured to be 216 μM . The K_m value of KHLY-pNA on hKLK7 was 107 μM . K_m values of other proteases (**Table 2**) used for the specificity assay were taken from literature.

Equation 4 | Inhibitory constant values according to Cheng-Prusoff. IC_{50} : half maximal inhibitory concentration (M); [S]: substrate concentration (M); K_m : Michaelis-Menten constant (M). $K_m = 216 \mu\text{M}$ for Boc-VPR-AMC on KLK5. $K_m = 107 \mu\text{M}$ for KHLY-pNA on KLK7.

$$K_i (M) = \frac{IC_{50}}{1 + \frac{[S]}{K_m}}$$

5.6.3 Fluorescence polarization albumin binding assay

2-fold serial dilutions of human serum (assumed HSA concentration = 600 μ M) or MSA (dissolved 600 μ M in PBS) were made in FP assay buffer. Serial diluted albumin (45 μ L) and fluorescent peptide (15 μ L, 100 nM in FP assay buffer) were pipetted into a black, 96 well half-area microtiter plate. The final protein concentration in the plate ranged from around 30 μ M to 30 nM for HSA in human serum and from around 350 μ M to 170 nM for MSA, whereas the final fluorescent peptide concentration was constantly 25 nM. Fluorescence anisotropy was measured using a plate reader (Tecan Infinite M200 Pro) with an excitation filter at 485 nm and an emission filter at 535 nm. Using **Equation 5**, the K_d was determined by nonlinear regression analysis of anisotropy versus protein concentration.

For the competitive FP assay, HSA and F-tag were diluted to 300 nM and 75 nM, respectively, in FP assay buffer. Small molecules were serial-diluted (2-fold) in PBS, FP assay buffer starting from a concentration of 1.8 mM. 20 μ L of each dilution was added into black 96 well half-area microtiter plate. The concentrations in the assay plate were 100 nM for human albumin and 25 nM for fluorescent tag in all wells, and ranging from 600 μ M to 2.3 μ M for small molecules.

FP assay buffer

PBS (10 mM Na_2HPO_4 , 1.8 mM KH_2PO_4 , pH 7.4, 137 mM NaCl, 2.7 mM KCl)
0.01% (v/v) Tween 20

Data were fitted and values calculated using GraphPad prism software.

Equation 5 | Experimental fluorescence anisotropy. A_f and A_b : fluorescence anisotropy for free and the fully bound fluorescent ligand. $[L]_T$ and $[P]_T$: total fluorescent ligand and protein concentration; K_d : dissociation constant.

$$A = A_f + (A_b - A_f) * \left[\frac{([L]_T + K_d + [P]_T) - \sqrt{([L]_T + K_d + [P]_T)^2 - 4[L]_T * [P]_T}}{2[L]_T} \right]$$

5.6.4 Plasma stability assay

To assess the plasma stability of peptides *in vitro*, peptides were added to a final concentration of 80 μ M in human plasma (citrate), which was incubated at 37 $^{\circ}$ C. 15 μ L sample were taken at various time points and 5 μ L guanidinium hydrochloride solution (7M, pH 2.0, HCl adjusted, filtered) was added to each sample taken to denature plasma proteins. Samples were vortexed, incubated for 10 min at RT and stored at -20 $^{\circ}$ C until the end of the experiments, when all samples were analyzed together. To each 20 μ L sample 200 μ L ice-cold EtOH / TFA (99.9 : 0.1) was added and the samples were incubated for 30 min at 4 $^{\circ}$ C to precipitate plasma proteins. Precipitated proteins were pelleted by centrifugation (9000 g, 4 $^{\circ}$ C, 20 min) and the supernatants were transferred into fresh tubes. EtOH was evaporated at 50 $^{\circ}$ C using a vacuum concentrator. The dry pellets were redissolved in 40 μ L H_2O / formic acid (99.9 : 0.1) and 5 μ L were analyzed by LC-MS (Shimadzu LCMS-2020) on a C18

column (Phenomenex Kinetex® column, 50 × 2.1 mm, 2.6 μm, 100 Å) using a linear gradient of solvent B (94.95% ACN, 5% H₂O, 0.05% HCOOH) in solvent A (94.95% H₂O, 5% ACN, 0.05% HCOOH) from 0 – 30% (mono- and bicyclic inhibitors) or 10 – 65% (inhibitor-tag conjugates) in 4.5 min at a flow rate of 1 mL/min. ESI-MS detection was done in positive mode.

Quantification was done based on the total ion current (TIC) chromatograms: The absolute intensities of the intact peptide peak were normalized to its intensity at timepoint zero. Thus, the amount of intact peptide at each timepoint is given as percentage of initial peptide. Fragments were identified manually by mass and intensities were also normalized to the intensity of the intact peptide peak at timepoint zero. Data were fitted and half-lives calculated with a one-phase decay model using GraphPad Prism software.

5.7 Pharmacokinetic studies in mice

All experiments in mice were conducted in accordance with the terms of the Swiss animal protection law and were approved by the animal experimentation committee of the cantonal veterinary service (Canton of Berne, Switzerland). For all experiments male C57BL/6J mice (Charles River) with a weight of around 20 g were used.

5.7.1 Peptide concentrations in blood

Mice were IV, IP or SC injected with 6.2 mg/kg KLK5 inhibitor or KLK7 inhibitor dissolved in 100 μ L formulation buffer (0.4 % DMSO for KLK5 inhibitor, 1.4 % DMSO for KLK7 inhibitor) at a concentration of 400 μ M. The preparation of the injection solutions is described in Figure 55. 20 μ L blood samples were taken at different time points into EDTA coated tubes (Microvette® 100 μ L, EDTA). The samples were centrifuged (5 min, 4 °C, 2000 g) and plasma was separated and stored at -80 °C.

5 μ L plasma were taken for analysis and 2 μ L albumin binding peptide SA21 (N-terminally 5(6)-FAM labeled, 0.1 mM in PBS) was added as standard to each plasma sample. 5 μ L guanidinium hydrochloride solution (7M, pH 2.0, HCl adjusted, filtered) was added and the samples were vortexed and incubated for 10 min at RT. Plasma proteins were precipitated with EtOH / TFA and removed by centrifugation as described above and EtOH was evaporated. Dried pellets were redissolved in 113 μ L H₂O / TFA (99.9 : 0.1) and 95 μ L were analyzed by analytical RP-HPLC (Agilent 1260 HPLC system) equipped with a fluorescence detector (Shimadzu RF-10AXL detector, excitation at 445 nm, emission at 535 nm) using a C18 column (Agilent Zorbax 300SB-C18, 4.6 mm x 250 mm, 5 μ m) with a linear gradient from 10 – 65% of solvent B (94.9% ACN, 5% H₂O, 0.1% TFA) in solvent A (94.5% H₂O, 5% ACN, 0.1% TFA) in 15 min at a flow rate of 1 mL/min.

The peptide concentrations in the plasma samples were determined by integrating the area under the intact peptide peak (ChemStation software, Agilent Technologies) and recalculating the concentration with a linear standard curve derived from a dilution series of known peptide concentrations in mouse plasma analyzed with the same protocol. Pharmacokinetic parameters were calculated based on a two-compartment model.

Formulation buffer	184 mM propylene glycol 8 mM Na ₂ HPO ₄ · 2 H ₂ O pH 7.4, sterile filtered
---------------------------	---

5.7.2 Fluorescence imaging of organs

Prior to injection, the ventral fur of mice was shaved off with a surgery prep shaver under anesthesia. Mice were IV injected with KLK5 inhibitor (6.2 mg/kg) and equimolar doses (2 μ mol/kg) of peptide-fluorescein conjugate (035-fluorescein, 3.2 mg/kg) and albumin tag (3.3 mg/kg) as described above. One mouse injected with 100 μ L PBS was used as background control. Mice were imaged using a CCD camera (IVIS Spectrum, PerkinElmer, USA; excitation filter: 495 nm, emission filter: 519 nm) at the indicated time points. To image skin samples, one mouse was euthanized by cervical dislocation 1,

4 h and 24 h after injection and 1 cm² patches of ventral skin was taken and imaged immediately as before with the same imaging system.

To assess the distribution to visceral organs mice were IP injected with 12.4 mg/kg KLK5 inhibitor and euthanized by cervical dislocation 48 h after injection. Organs were harvested and imaged immediately using the imaging setup described above.

All data were processed using the Living Image software (PerkinElmer, USA) taking the shaved ventral region or the whole skin patch as region of interest (ROI) and plotting the average radiant efficiency.

5.7.3 Preparation and fluorescence microscopy of skin sections

Mice were IP injected with equimolar doses (2 µmol/kg) of mIgG-anti-Flag-SiR (DOL 1.1, 300 mg/kg), MSA-SiR (DOL 0.4, 300 mg/kg), 035-tag-SiR (DOL 1.0, 6.4 mg/kg), Tag-SiR (DOL 1.0, 3.6 mg/kg) and 6-carboxy-SiR (1.0 mg/kg). 200 µL injection solutions with a concentration of 200 µM were prepared in in PBS (mIgG-anti-Flag-SiR and MSA-SiR) or formulation buffer (035-tag-SiR, Tag-SiR and 6-carboxy-SiR). Mice injected with mIgG-anti-Flag-SiR or MSA-SiR were sacrificed 24 h after injection, mice injected with 035-tag-SiR, Tag-SiR or 6-carboxy-SiR were sacrificed 8 h after injection by cervical dislocation. 1 cm² patches of ventral skin were taken and fixed by embedding in cryomatrix within a cryomold, and freezing using isopentane cooled down with dry ice. Samples were stored at -20 °C and microscopy sides were prepared by the EPFL Histology core facility, (Dr. Jessica Sordet-Dessimoz). The skin samples were cut into 10 µm sections using a cryostat (Leica CM3050S). Sections were directly mounted on Superfrost+ slides (Menzel Gläser) and slides were stored at -20 °C until stained. DAPI staining was performed as following: slides were air dried for 30 minutes at RT, stained for 10 minutes in DAPI solution (diluted 1 : 10,000 in PBS from a 1 mg/ml stock solution in water), washed 3-times for 5 min PBS and mounted using Fluoromount-G mounting medium.

Fluorescence microscopy was conducted at the EPFL Bioimaging and Optics Platform (Dr. Arne Seitz) on a slide scanner (Olympus VS120-L100, 20x UPLSAPO objective, DAPI and Cy5 channel settings) using consistent illumination times for all samples. Images were analyzed using QuPath software¹⁵³ setting identical contrast and brightness scales for all samples. Intensity profiles were generated using ImageJ software and plotted with GraphPad Prism.

6 Supporting information

6.1 Phylogeny of KLKs

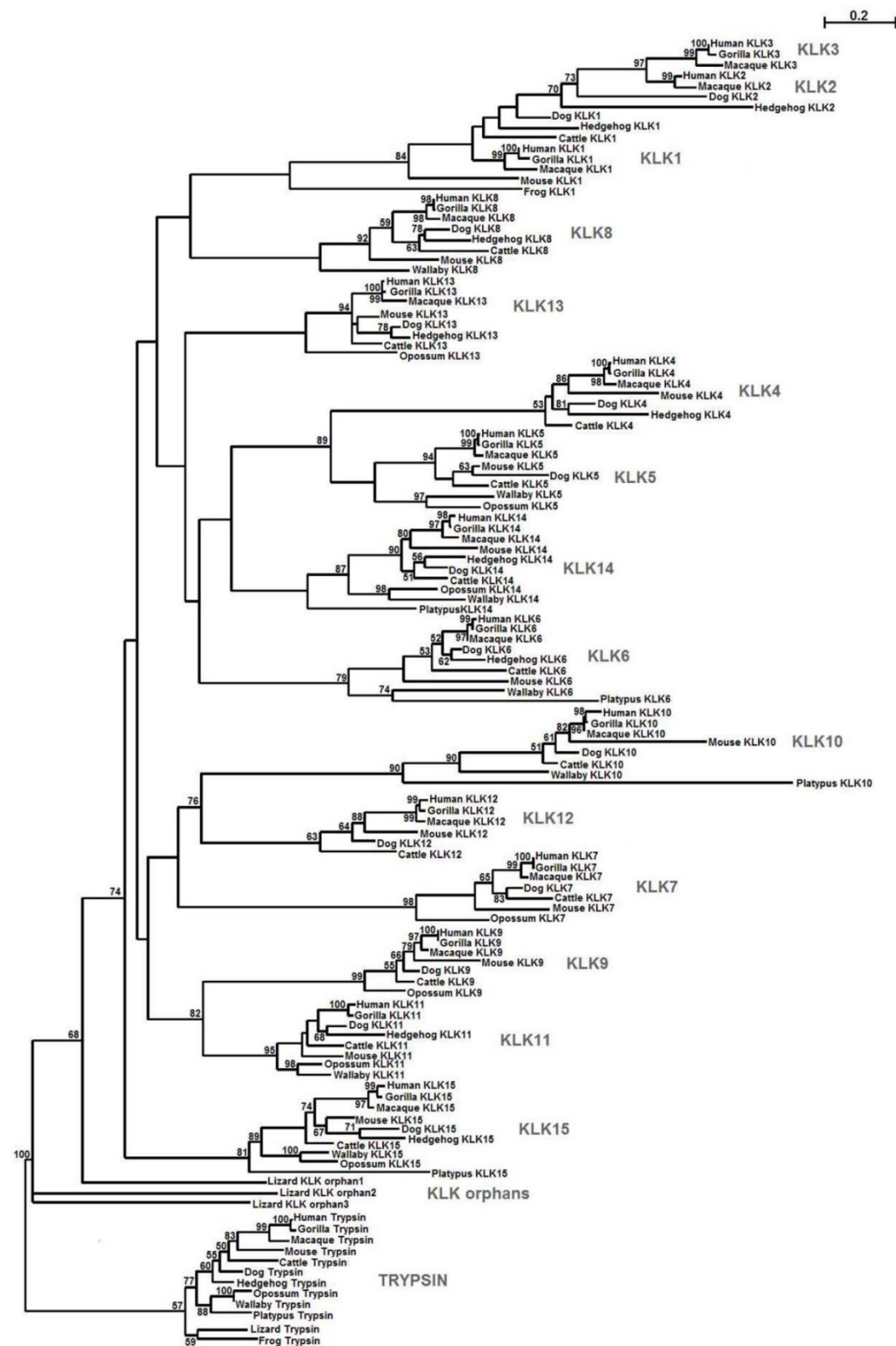


Figure S1 I Representative maximum likelihood phylogenetic tree of KLK homologues. Bootstrap values (.50%) are indicated at the nodes. The branch lengths depict evolutionary distance. The trypsin proteins are used as outgroup. The scale bar at the upper right denotes evolutionary distance of 0.2 amino acids per position. Used from ²² under CC BY 4.0, © 2010 Pavlopoulou *et al.*

6.2 Structure of human serum albumin

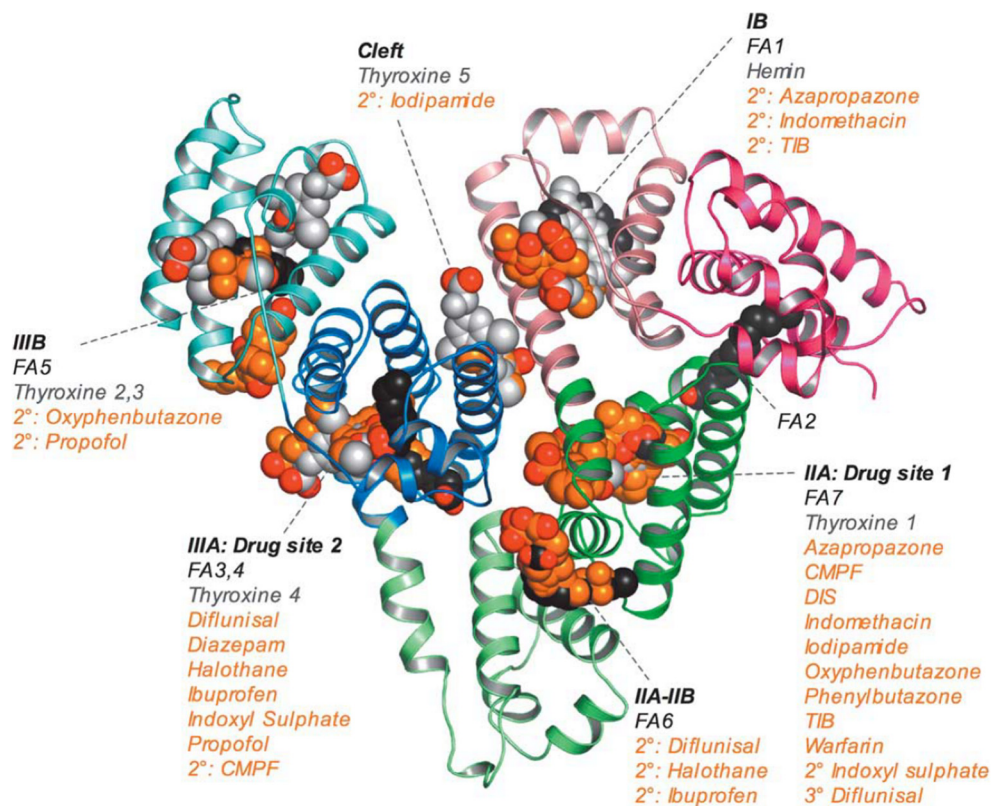
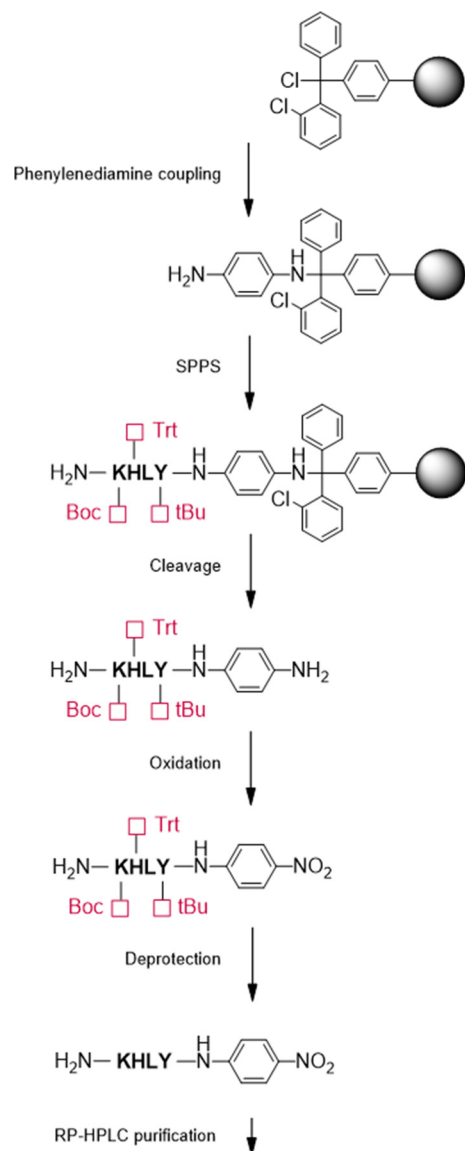


Figure S2 | Structure of human serum albumin with binding sites. The three domains of albumin are shown in red (IA), pink (IB), green (IIA), light green (IIB), blue (IIIA) and turquoise (IIIB). Ligands are shown in space-filling representation: fatty acids (black), endogenous ligands hemin and thyroxine (gray) and drugs (orange). All oxygen atoms are shown in red.

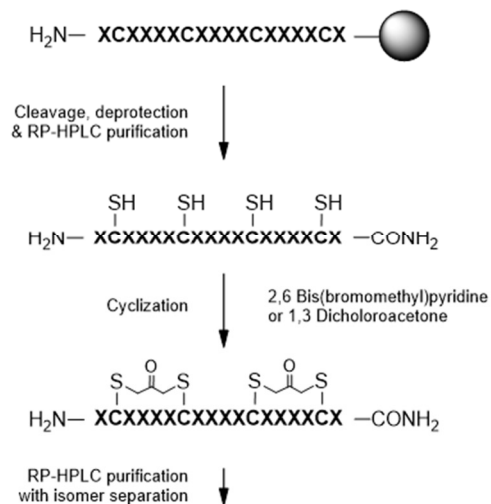
Reprinted with permission from ¹³⁹, reprinting license nr. 4570250087933, © 2005 Elsevier.

6.3 Synthesis schemes

a Para-nitroanilide substrate synthesis



b Synthesis of bicyclic peptides



c Synthesis of monocyclic peptides

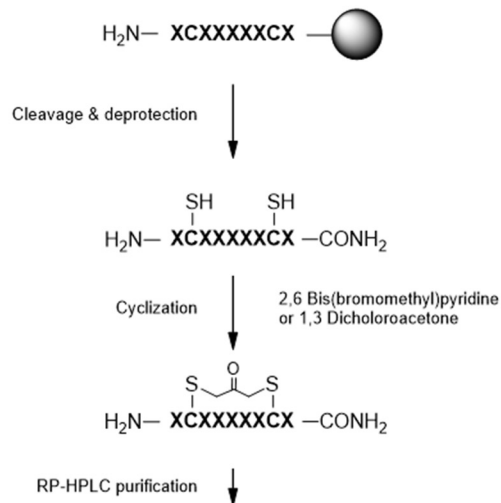


Figure S3 | Synthesis schemes of substrates and cyclic peptides. Fmoc SPPS was used. **(a)** Synthesis of KHLY-pNA according to ¹⁵². **(b)** Synthesis of bicyclic peptides. **(c)** Synthesis of monocyclic peptides.

d

Potential synthesis of bispecific inhibitor

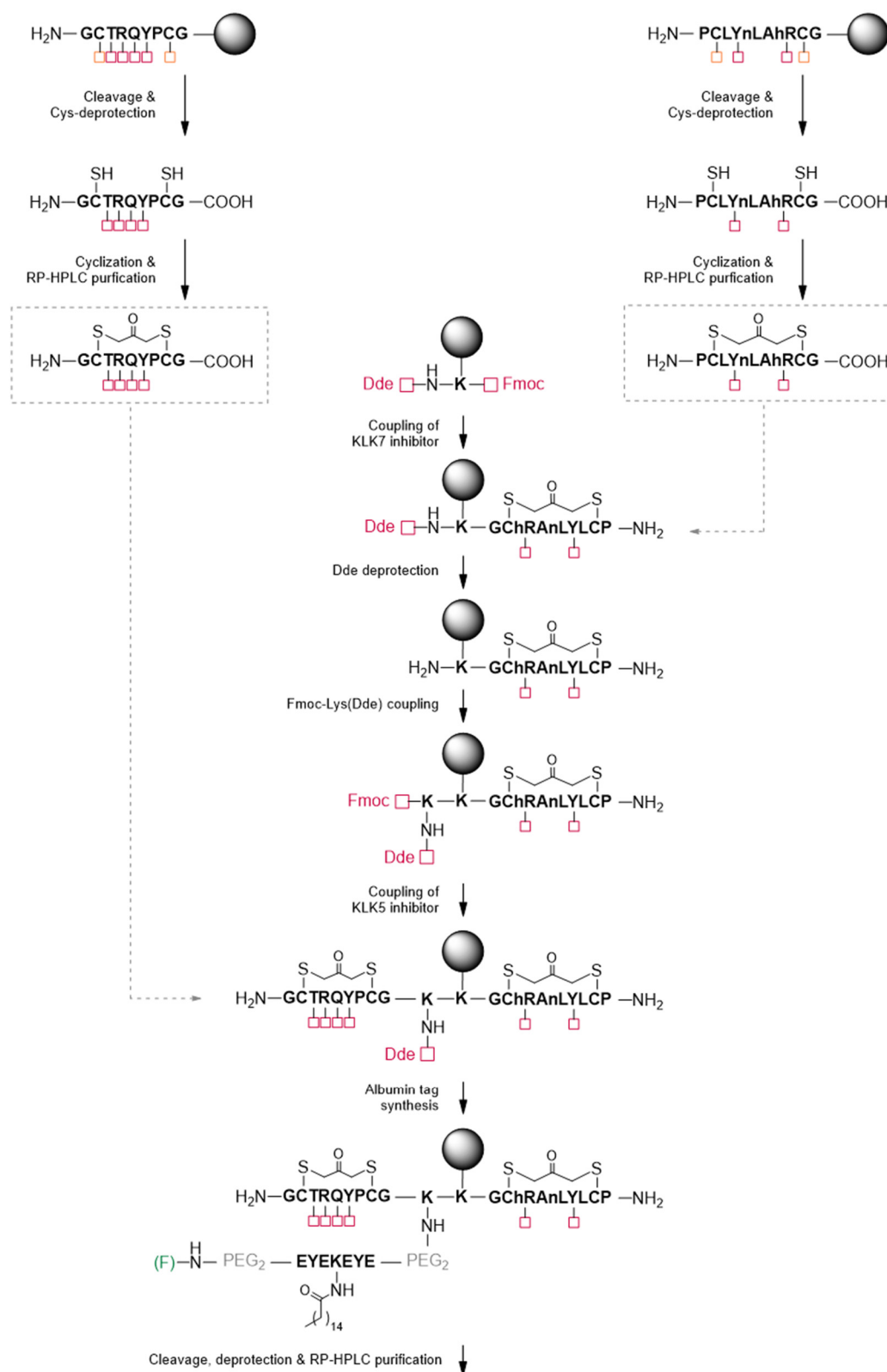


Figure S4 | Synthesis schemes of potential bispecific inhibitor. Based on Fmoc SPPS.

6.4 Raw data

6.4.1 Affinity improvement

hKLK5				hKLK7			
	Substitution	Ki (nM)	fold-change		Substitution	Ki (nM)	fold-change
Gly 1	219 β-Ala	3.6	1.6	Pro 1	269 Hyp	161.9	5.5
	220 GABA	4.4	2.0		270 Pip	25.7	0.9
	223 Ala	14.0	6.3		271 β-Pro	44.3	1.5
	224 D-Ala	1.4	0.6		272 D-Pro	34.8	1.2
	225 N-Guanidino-Gly	4.1	1.9		273 Aze	134.6	4.6
Cys 2	119 D-Cys	254.3	115.1	Cys2	196 D-Cys	844.2	28.8
	121 HCys	7.8	3.5		198 HCys	69.9	2.4
Thr 3	146 Ser	15.8	7.2	Leu 3	200 HLeu	506.3	17.3
	147 HSer	41.4	18.7		201 NptGly	39.5	1.3
	148 L-allo-Thr	200.1	90.5		202 Nva	109.3	3.7
Arg 4	149 Lys	32.0	14.5		203 Nle	162.5	5.5
	150 Agb	691.2	312.7	Tyr 4	204 HTyr	844.2	28.8
	151 Homo-arginine	288.9	130.7		205 4-Nitro-Phe	1547.8	52.8
	152 4-Guanidino-Phe	31.7	14.4		206 2-Fluoro-Phe	512.0	17.5
5	153 Cit	2.1	0.9		207 3-Fluoro-Phe	101.7	3.5
					208 4-Fluoro-Phe	258.3	8.8
Tyr 6	155 HTyr	3.0	1.4		209 Carbamoyl-Phe	844.2	28.8
	156 4-Nitro-Phe	7.0	3.2		210 Amino-Phe	1040.9	35.5
	157 2-Fluoro-Phe	1.6	0.7	Leu 5	212 Met	17.8	0.6
	158 3-Fluoro-Phe	2.2	1.0		213 O-meth-HSer	89.7	3.1
	159 4-Fluoro-Phe	2.5	1.1		214 Nle	7.0	0.2
	160 Cyano-Phe	5.3	2.4	Arg 7	250 HArg	42.2	1.4
	164 1-Naphthyl-Ala	2.6	1.2		251 4-Guanidino-Phe	34.4	1.2
	165 2-Naphthyl-Ala	1.4	0.6		232 D-Arg	188.9	6.4
Pro 7	166 5-HTP	4.4	2.0	Cys 8	197 D-Cys	97.1	3.3
	167 D-Pro	138.3	62.6		199 HCys	40.3	1.4
	168 Pip	1.3	0.6		114 Reference	29 ± 4	
	169 D-Pip	58.9	26.6				
	170 Hyp	2.6	1.2				
	171 β-Pro	18.9	8.5				
	172 Aze	11.1	5.0				
Cys 8	120 D-Cys	2.2	1.0				
	122 HCys	8.5	3.9				
	035 Reference	2.2 ± 0.2					

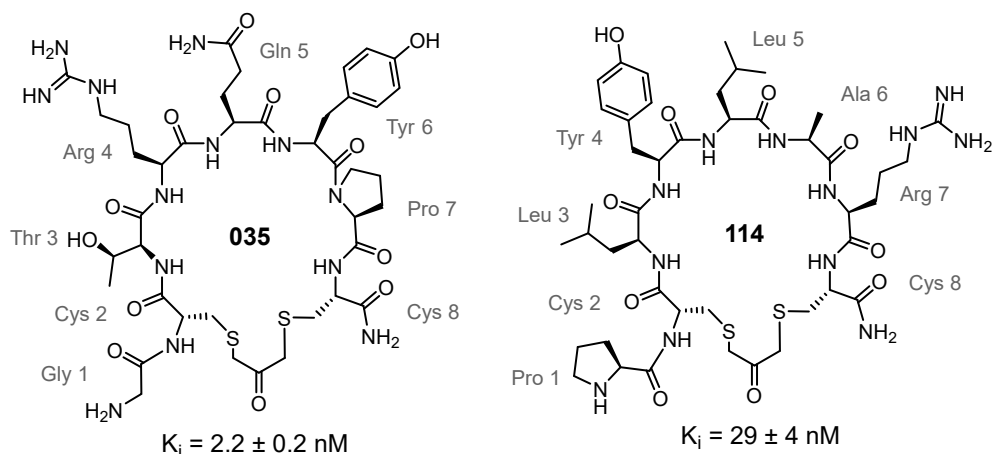


Figure S5 | Raw data – Affinity improvement with unnatural amino acids. Macrocycles with single unnatural amino acid substitutions were tested. For KLK5 peptide 035 and for KLK7 peptide 114 served as the reference. The processed data can be found in Figure 32.

6.4.2 Stability of inhibitors

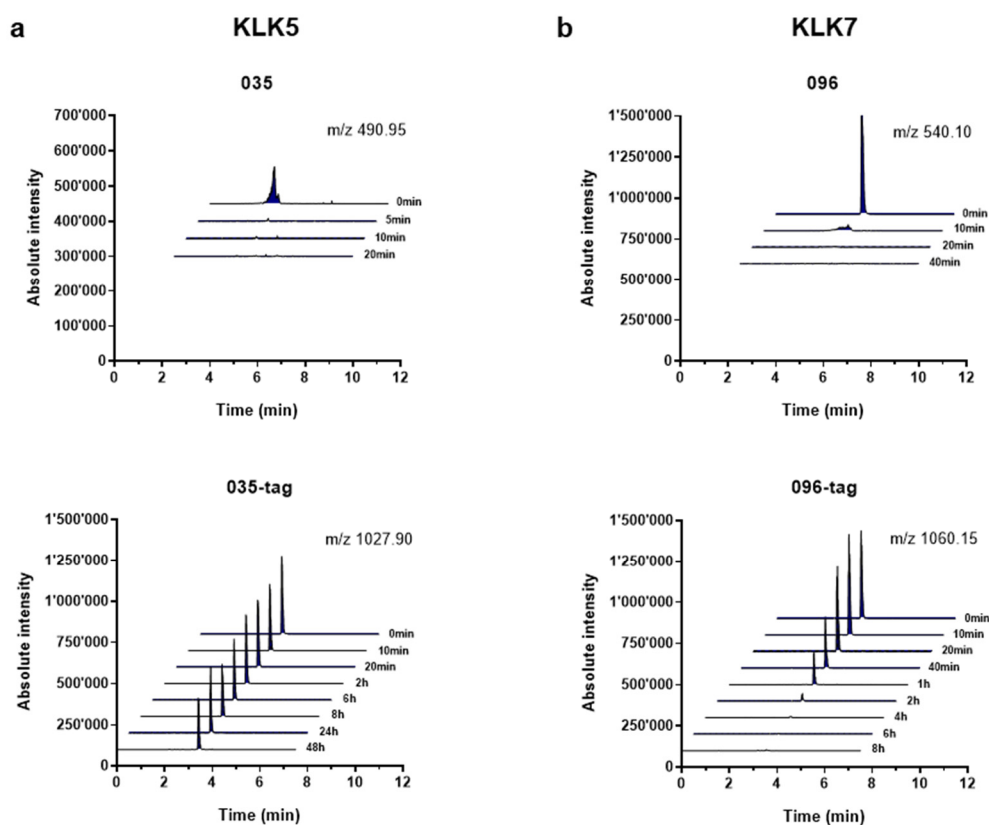


Figure S6 | Raw data - Stability inhibitors in human plasma. Shown are the TIC traces filtered for the indicated m/z values. For peptide 035 and 096 the used values correspond to the m+2 ions. For peptide 035-tag and 096-tag the used values correspond to the m+3 ions. The processed data can be found in Figure 37.

6.4.3 Stability improvement of KLK7 inhibitor

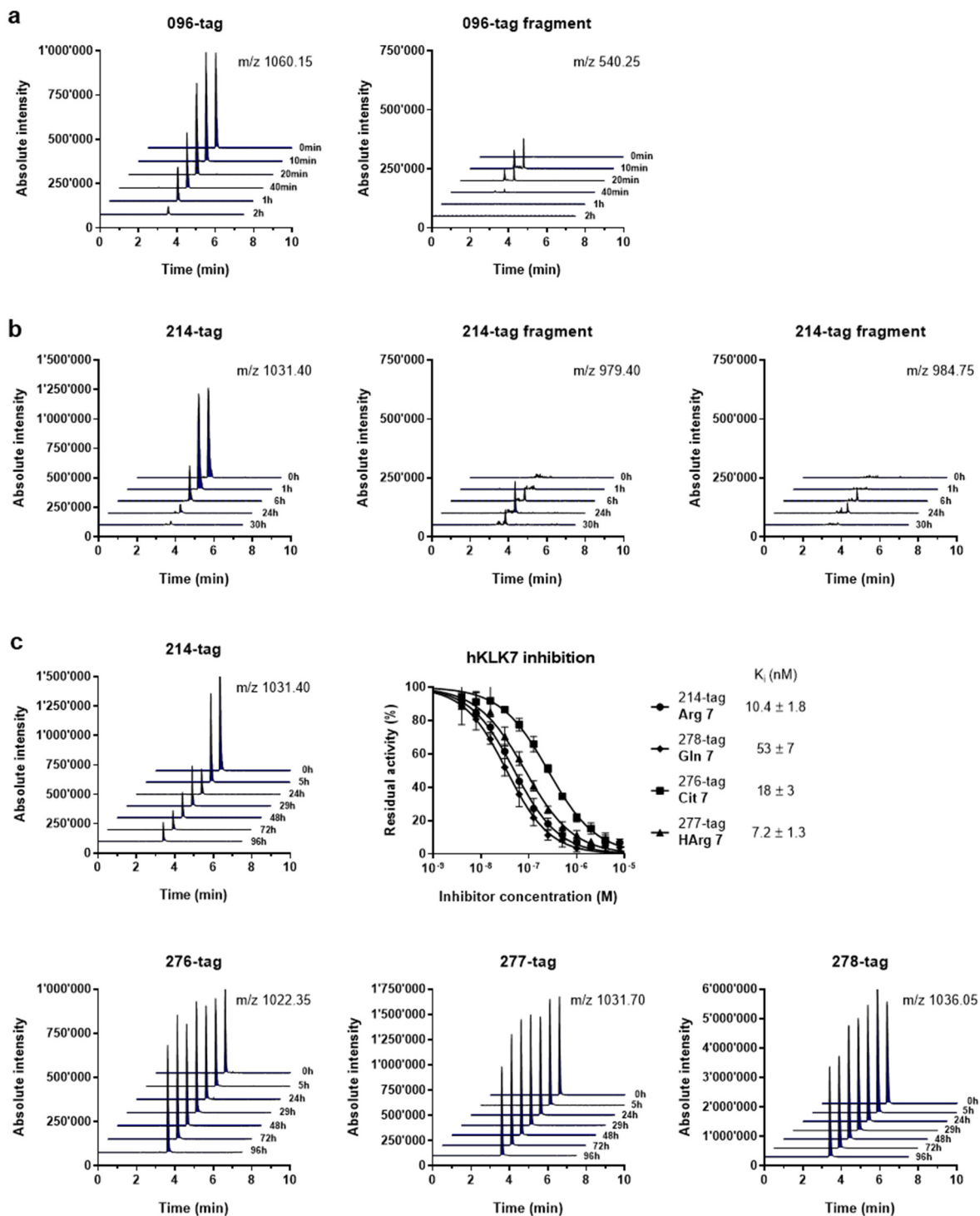


Figure S7 | Raw data - Stability of KLK7 inhibitors in human plasma. Shown are the TIC traces filtered for the indicated m/z values. For the 096-tag fragment the used values correspond to the m+2 ion. For all other peptides and fragments the used values correspond to the m+3 ions. The processed data are shown in Figure 38.

6.4.4 Stability in mouse plasma

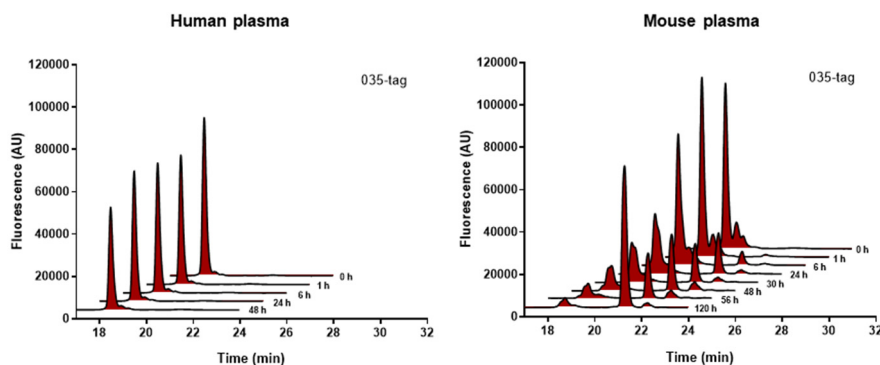


Figure S8 | Stability in human versus mouse plasma. Peptide 035-tag was incubated at a concentration of 80 μM in human or mouse plasma. The plasma was incubated at 37°C, samples were taken at various time points, plasma proteins were precipitated, and the supernatant was analyzed by analytical HPLC equipped with a fluorescence detector. While the half-life in human plasma was more than 90 h, a half-life of only 10-20 h was reached in mouse plasma.

6.4.5 Degradation fragments in mice

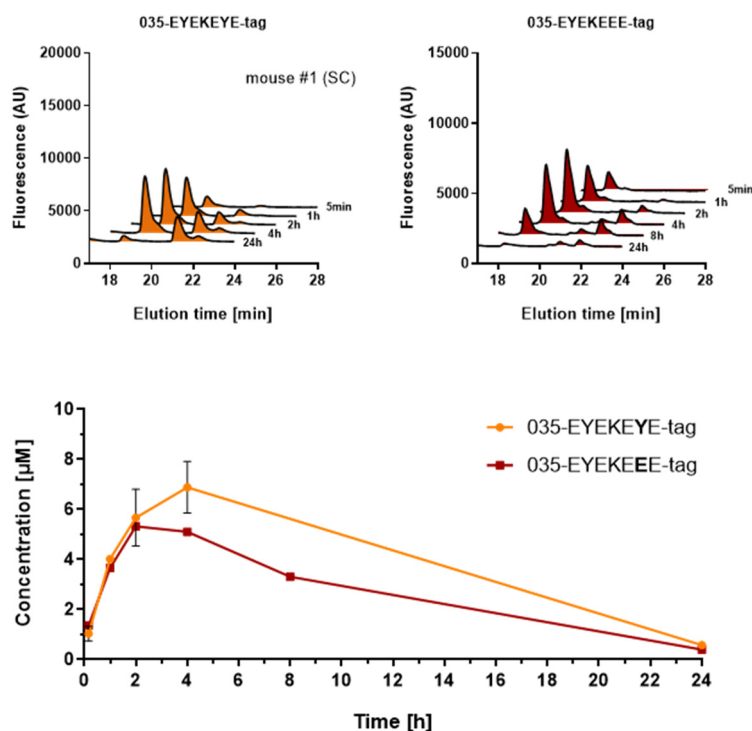


Figure S9 | Identification of degradation fragment through amino acid exchange. One mouse was injected subcutaneously (SC) with 6.2 mg/kg peptide 035-tag with the tag variant (EYEKEEE). Blood samples were taken at different time points, plasma proteins were precipitated, and the supernatant was analyzed by analytical HPLC equipped with a fluorescence detector. Data for 035-EYEKEYE-tag are taken from Figure 44 (a) Raw HPLC data of one mouse. (b) Concentrations of intact peptide in blood plasma at different time points after injection.

6.4.6 Biodistribution

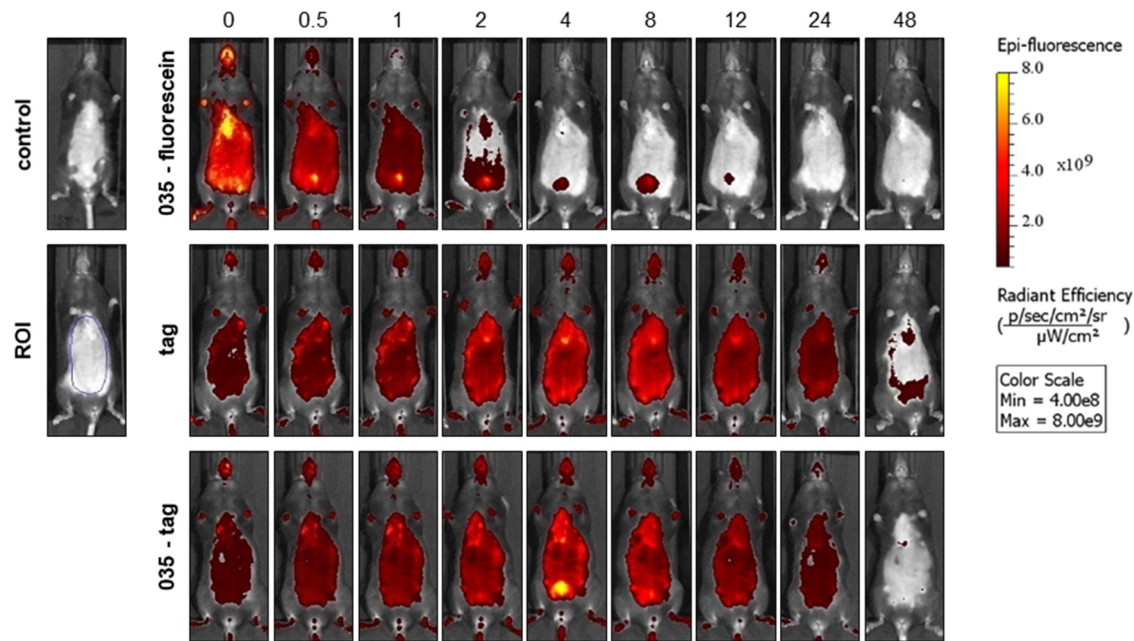


Figure S10 | Raw data - Biodistribution assessed by fluorescence imaging. Representative images of mice IV injected with KLK5 inhibitor, albumin tag, 035-fluorescein (2 μmol/kg), or vehicle sacrificed and imaged at different time points after injection. The quantified data are shown in Figure 45.

6.5 Overview of binding constants

Table S1 | Overview of binding constants. Data sources are color coded:

This work (black), *Deraison et al.* ⁷ (blue), *Fortugno et al.* ³¹ (green), *Zorzi et al.* ¹¹⁶ (red), and European Medicines Agency ¹⁴⁷ (gray). If not indicated differently, K_i values are given for KLKs and K_d value for albumins.

	hKLK5	mKLK5	hKLK7	mKLK7	HSA	MSA	RSA	BSA
LEKTI D5	33 nM		77 nM					
LEKTI D6	83 nM		296 nM					
LEKTI D6-D9	7.7 nM							
LEKTI D7	17 nM							
LEKTI D7-D9	8.4 nM							
LEKTI D8-D9	5.4 nM							
LEKTI D8-D11	3.7 nM		35 nM					
LEKTI D9-15	119 nM		-					
LEKTI D10-15			20 nM					
035	2.2 nM	IC ₅₀ = 2 nM	-					
096	-		16 nM	IC ₅₀ = 20 nM				
114	-		29 nM					
214	-		7 nM					
035-tag	1.2 nM		-		119 nM	7000 nM		
096-tag	-		32 nM		164 nM			
278-tag	-		7.4 nM		-			
Semaglutide					1070 nM	1780 nM	2800 nM	
F-tag					91 nM	4300 nM	220 nM	200 nM
Tag-F					2610 nM	7900 nM	500 nM	
F-SSSKSSS-tag					1080 nM	10,000 nM		
282					132 nM			

6.6 Peptide purities & masses

All peptides assigned with numbers were developed and synthesized in this work. The peptides F-tag, Tag-F, SSSKSSS-tag were developed and synthesized by Dr. Alessandro Zorzi¹¹⁶. The KLK7 substrate KHLY-pNA synthesized and used in this work was published by Veer et al³⁷. For simplification 5(6)-FAM is shown as 5-FAM in the figures of chapter 6.6.

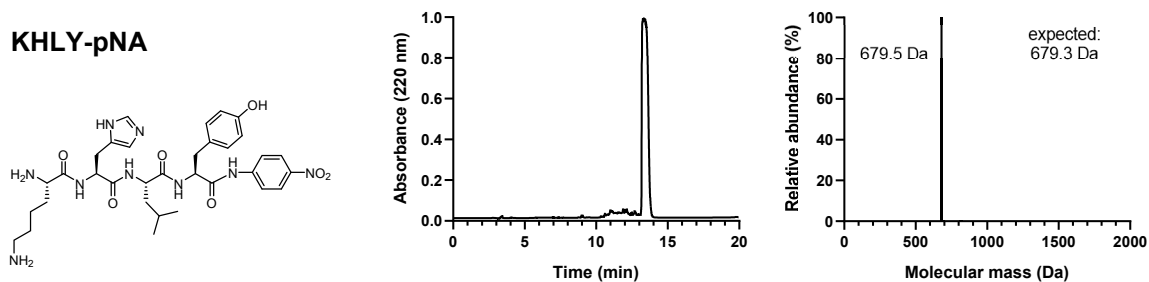


Figure S11 | Structure, purity and mass of KLK7 substrate KHLY-pNA.

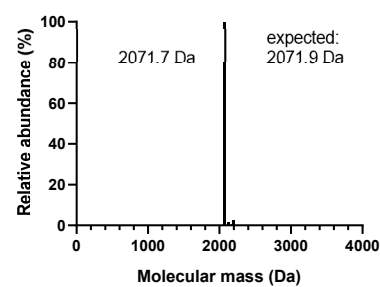
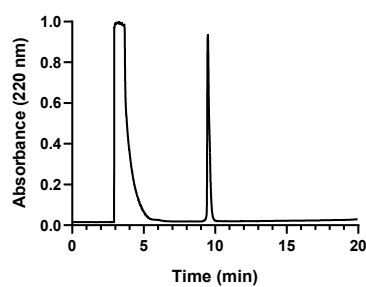
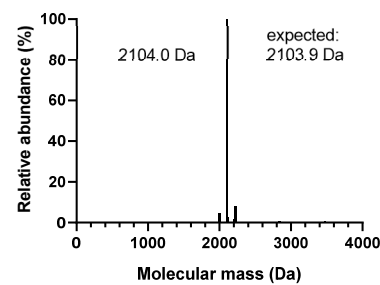
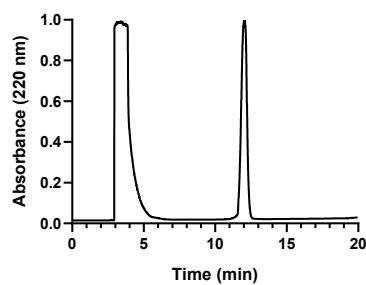
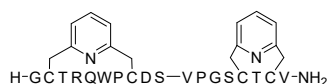
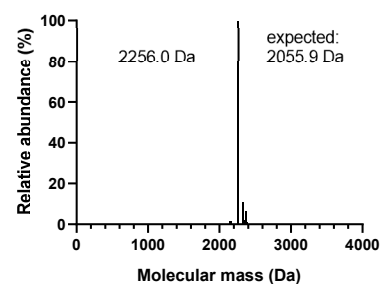
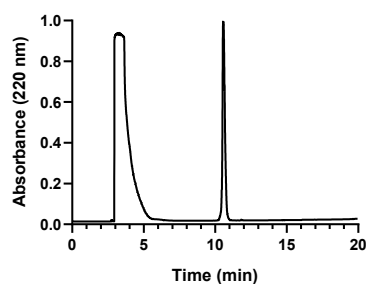
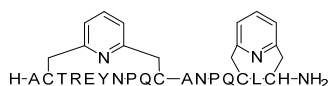
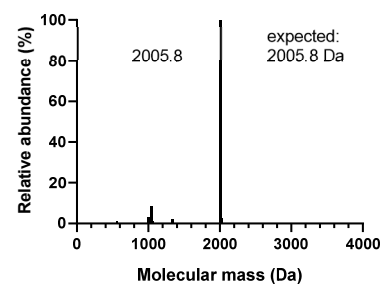
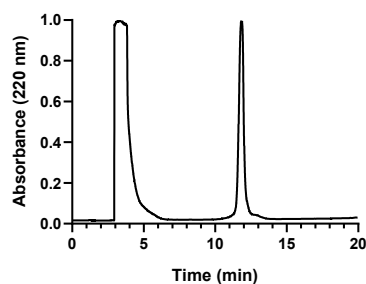
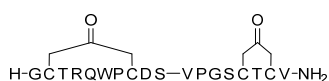
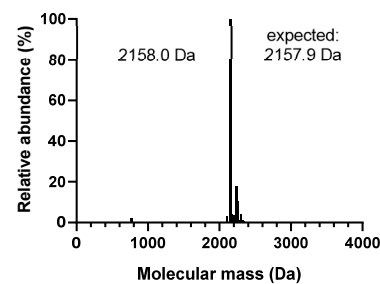
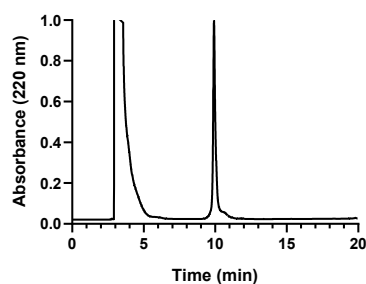
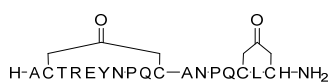
001**002****003****004****005**

Figure S12 | Structure, purity and mass of bicyclic peptides.

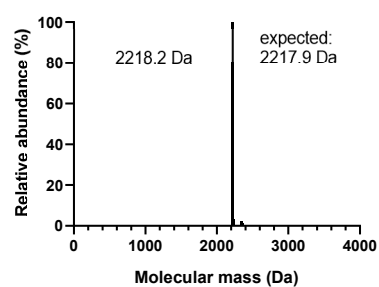
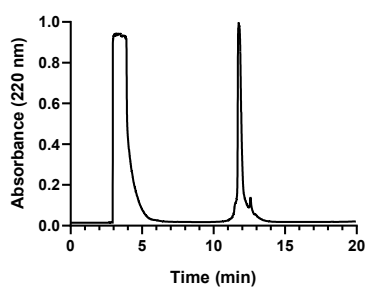
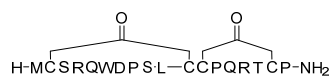
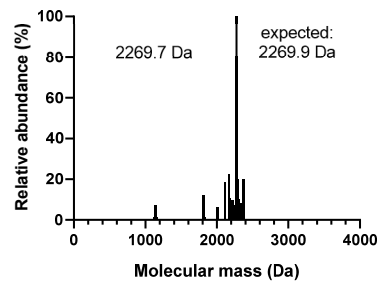
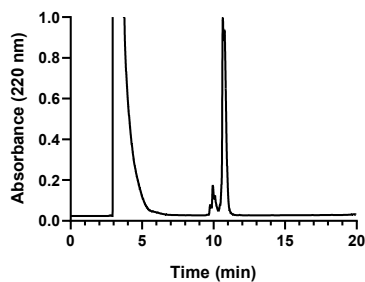
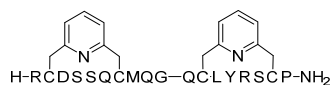
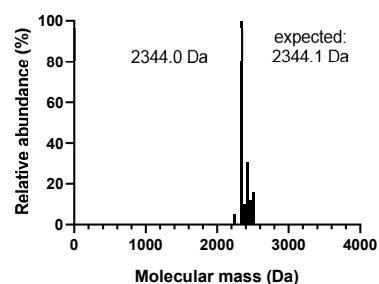
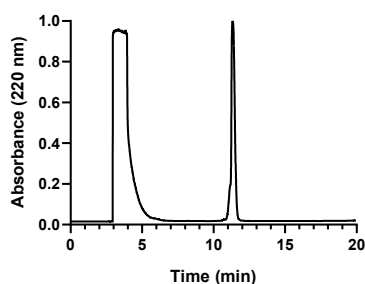
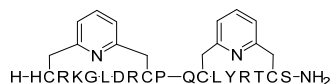
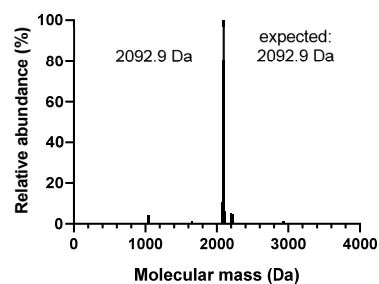
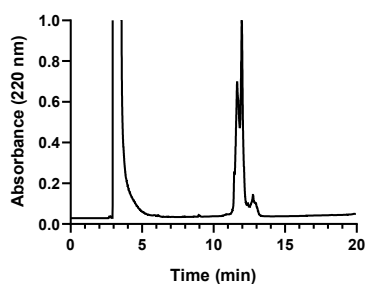
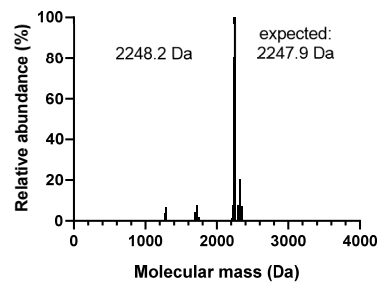
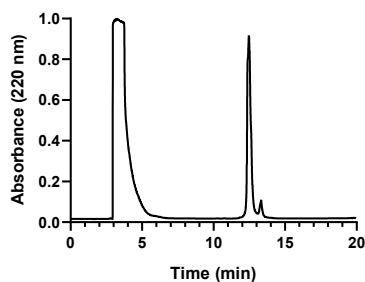
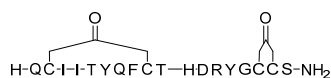
006**007****008****009****010**

Figure S12 cont. | Structure, purity and mass of bicyclic peptides.

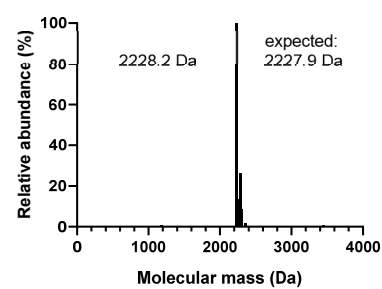
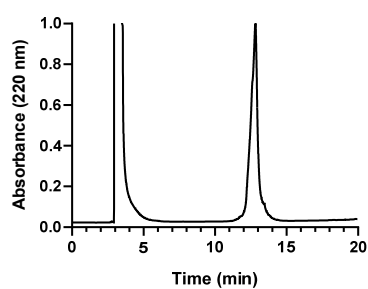
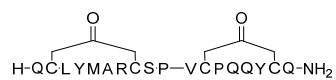
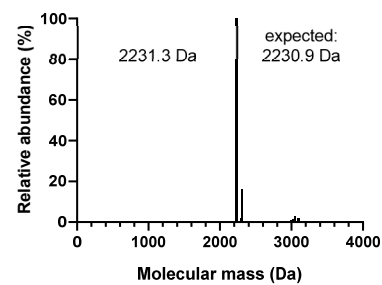
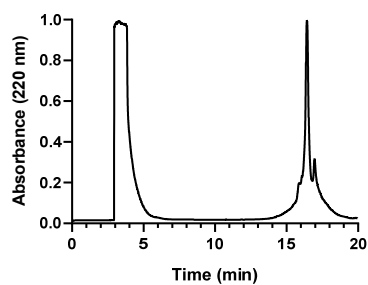
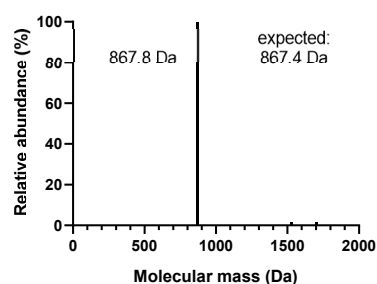
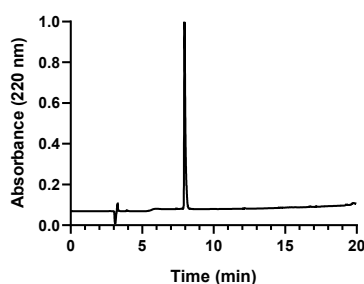
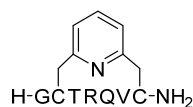
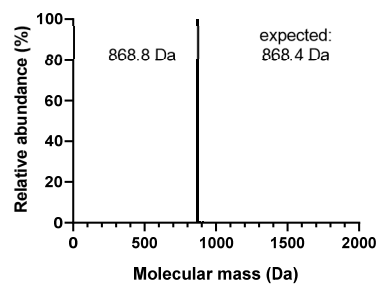
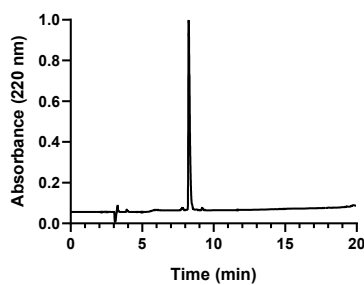
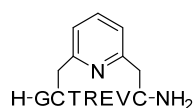
011**012**

Figure S12 cont. I Structure, purity and mass of bicyclic peptides.

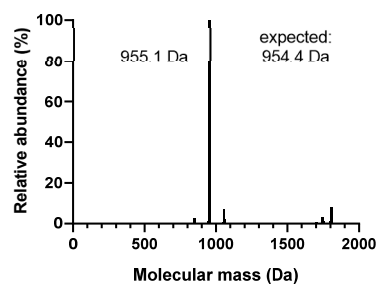
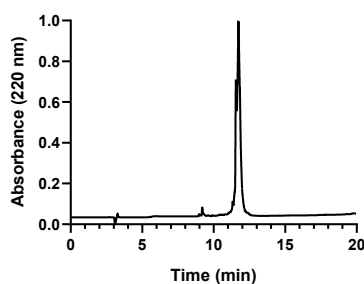
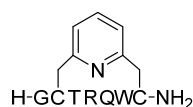
026



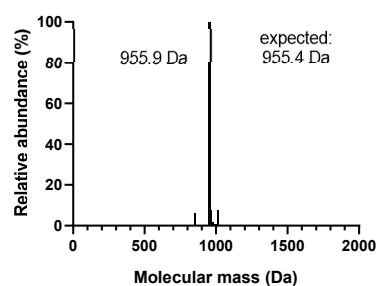
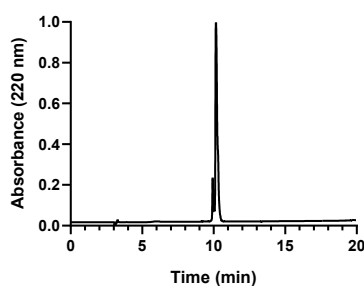
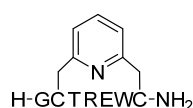
027



028



029



144

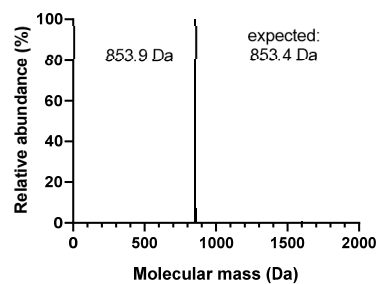
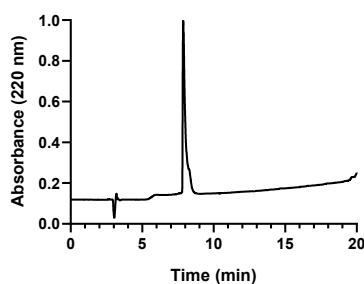
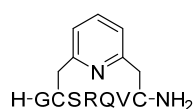
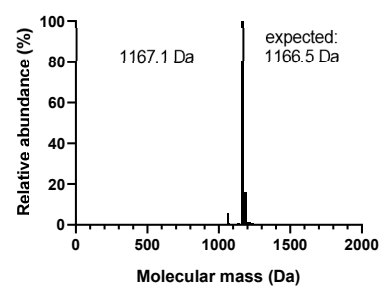
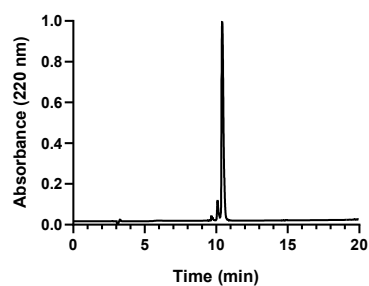
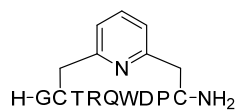
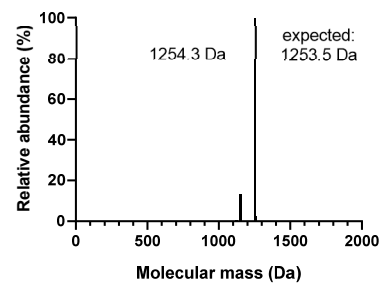
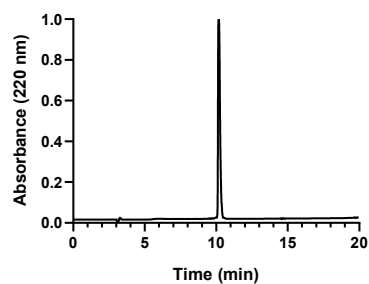
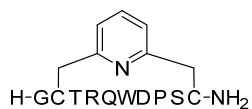


Figure S13 | Structure, purity and mass of monocyclic peptides.

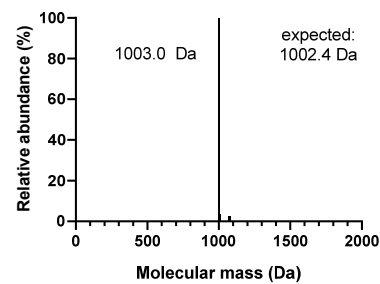
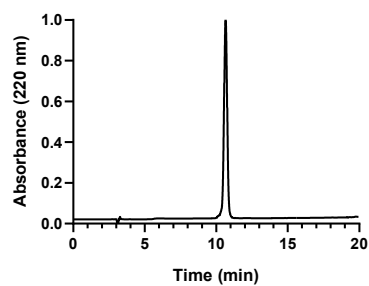
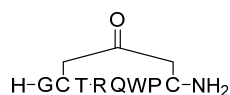
031



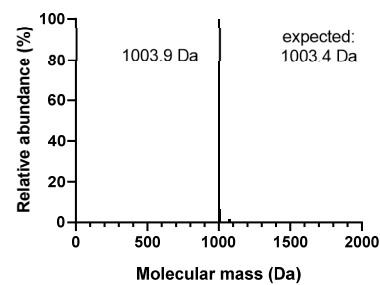
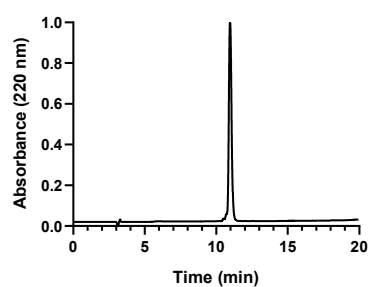
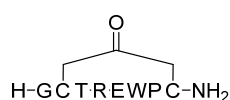
051



033



034



035

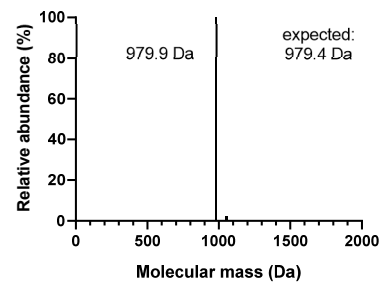
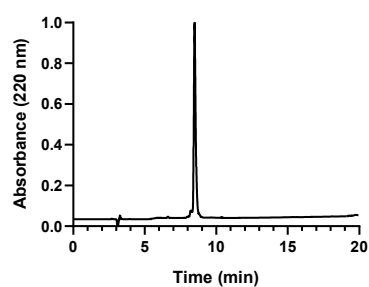
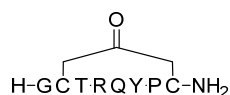
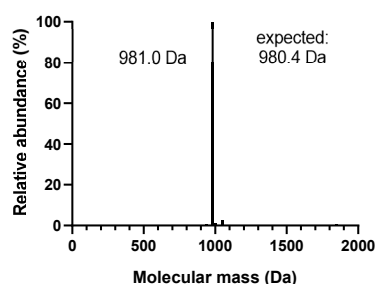
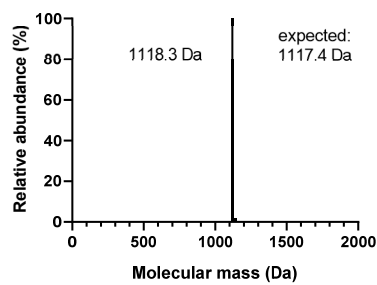
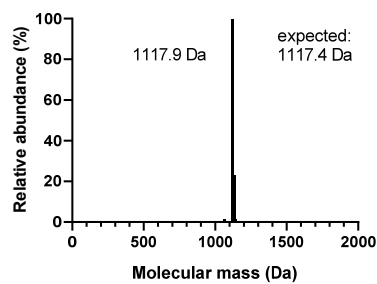
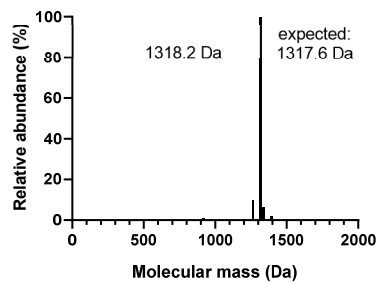
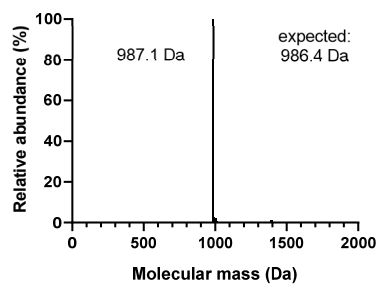


Figure S13 cont. I Structure, purity and mass of monocyclic peptides.

$$\text{H-GCTREYPC-NH}_2$$

$$\text{H-GCTRQWP} \begin{array}{c} \text{O} \\ \parallel \end{array} \text{CD-NH}_2$$

$$\text{H-GCTRQWDPC-NH}_2$$

$$\text{H-GCTRQWDPS-L-C-NH}_2$$
NCCSCC1C=CC=CC=N1CC

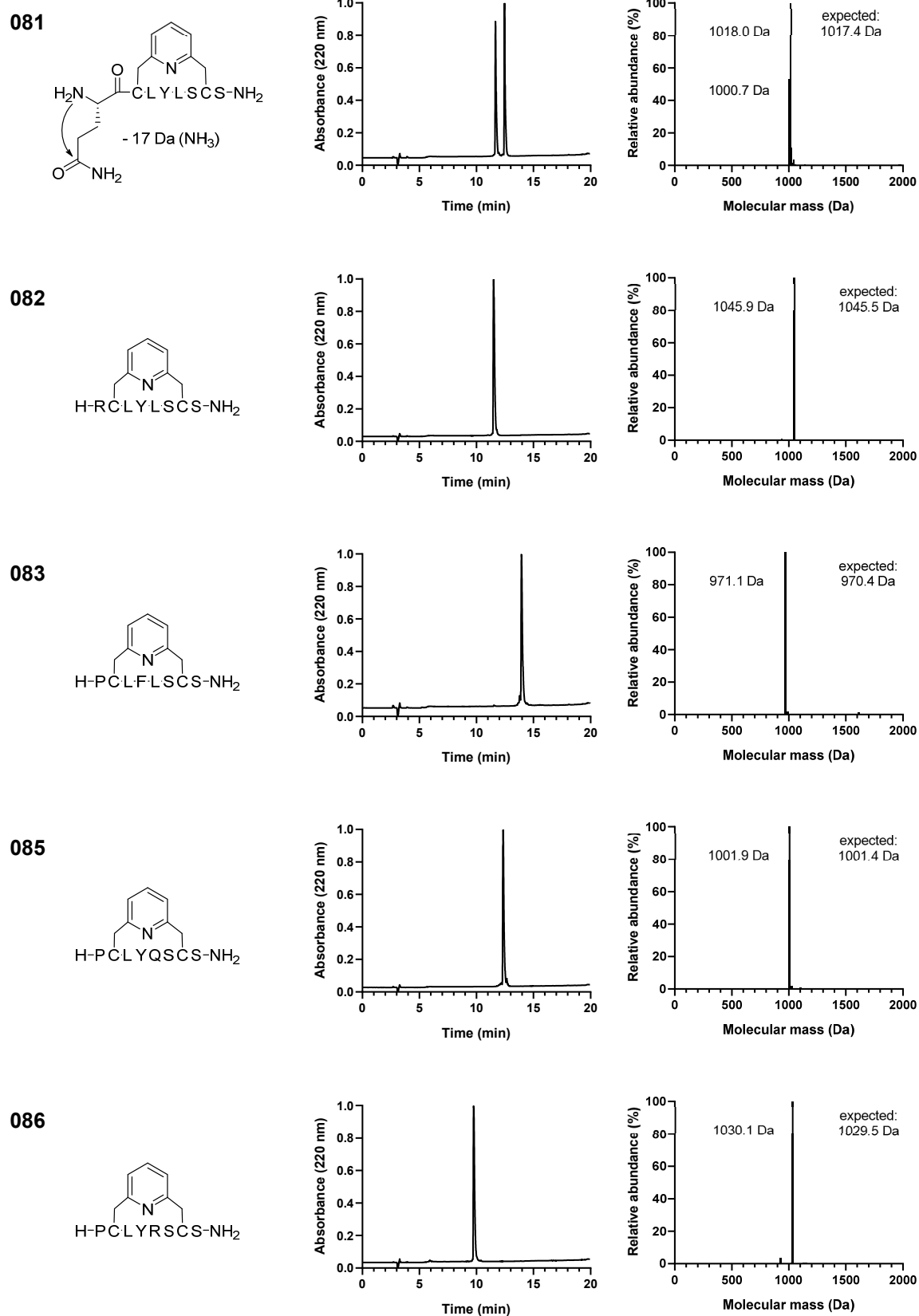
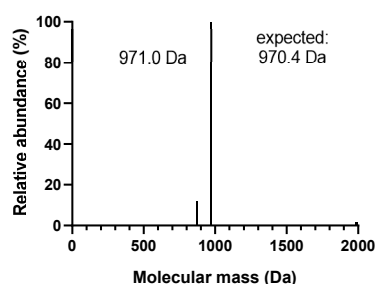
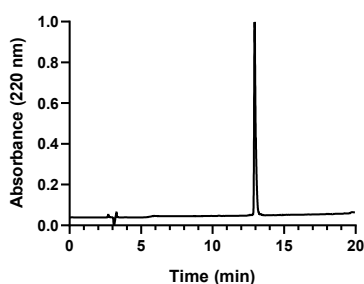
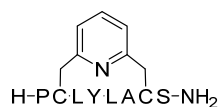
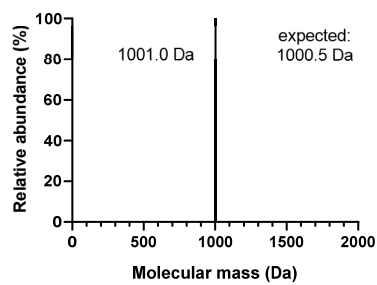
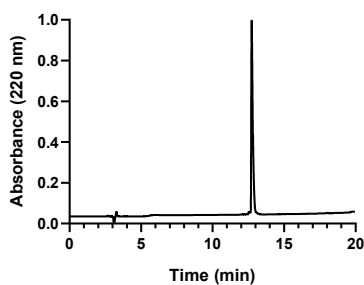
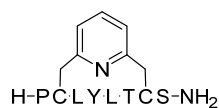


Figure S13 cont. | Structure, purity and mass of monocyclic peptides.

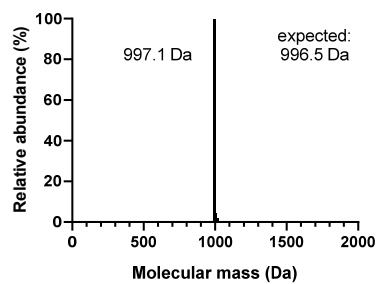
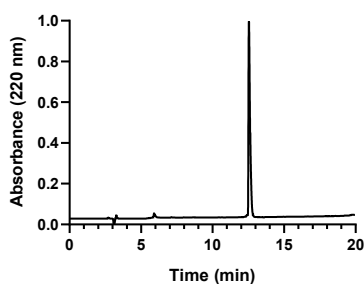
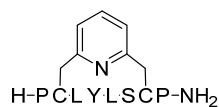
087



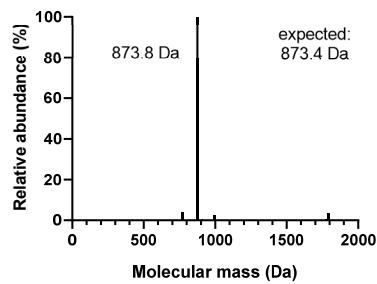
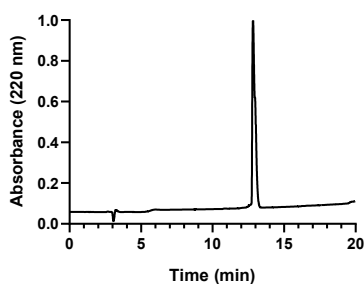
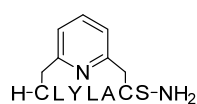
088



089



108



109

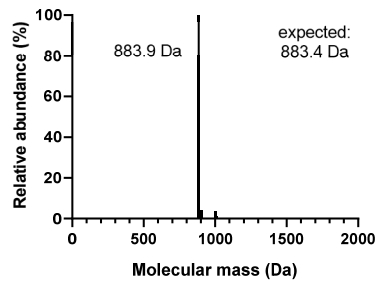
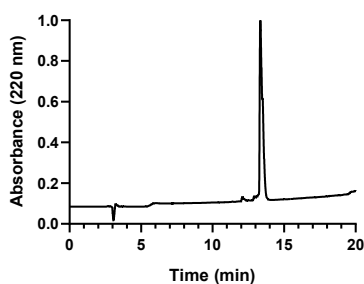
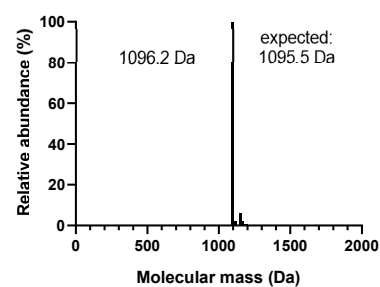
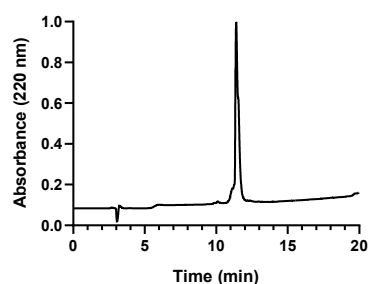
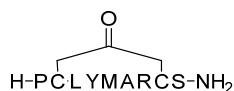
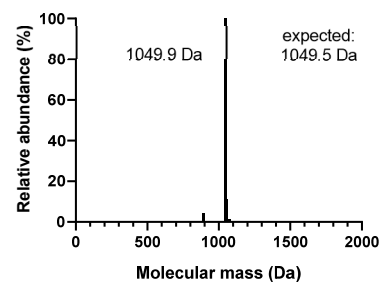
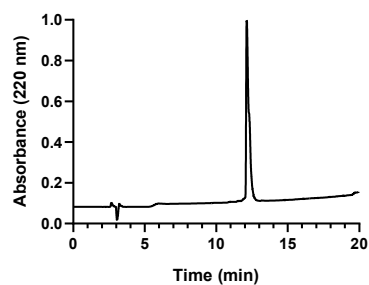
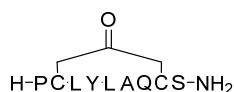


Figure S13 cont. | Structure, purity and mass of monocyclic peptides.

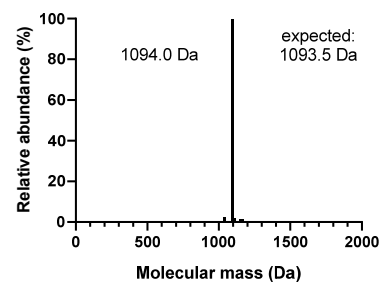
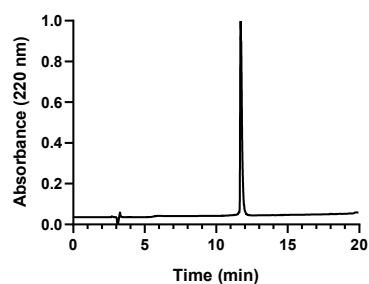
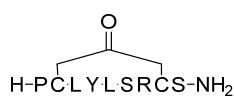
111



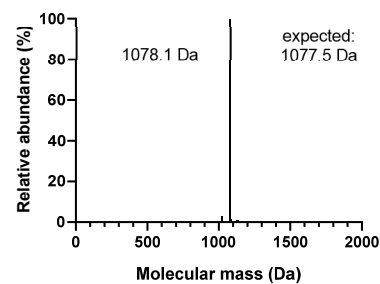
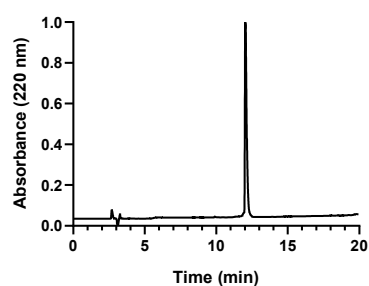
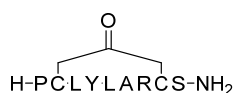
112



095



096



114

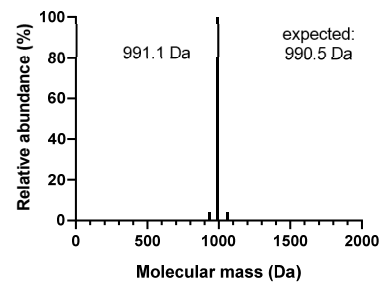
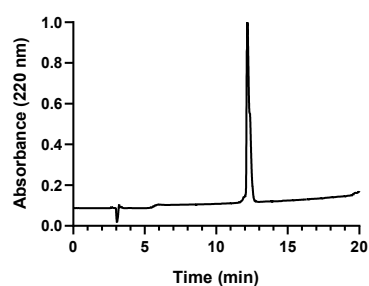
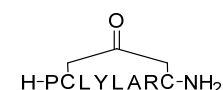
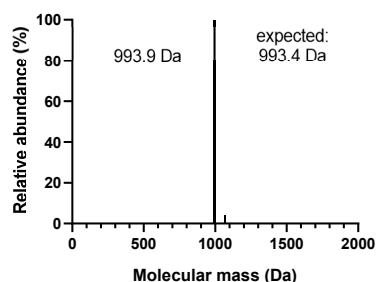
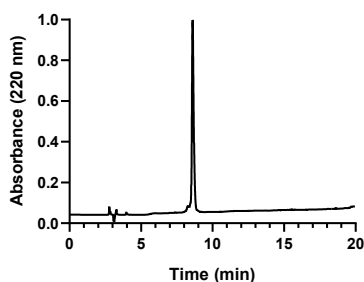
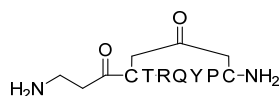
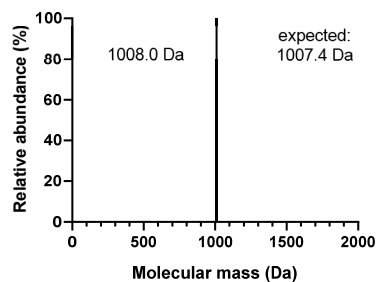
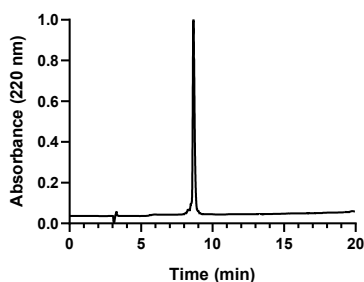
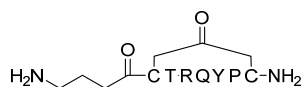
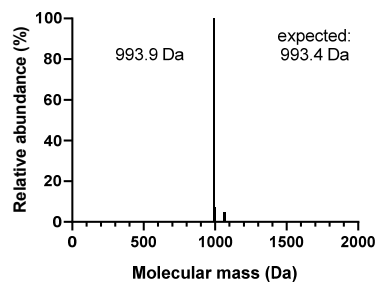
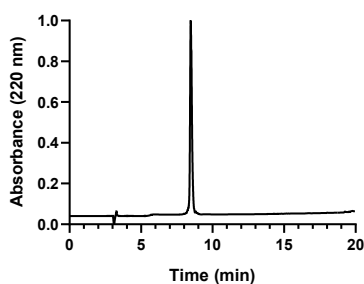
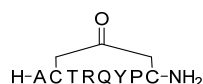
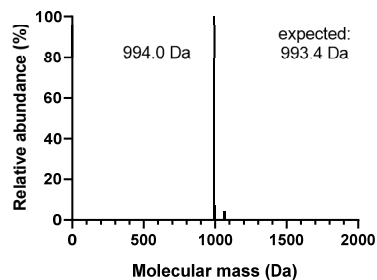
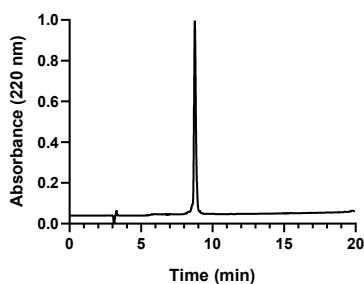
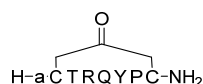
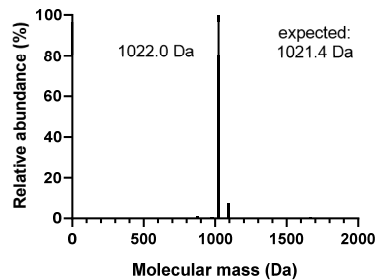
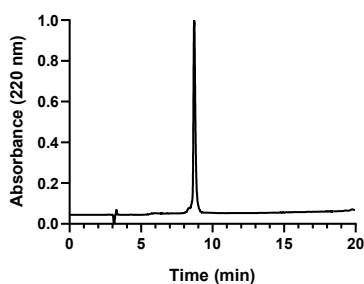
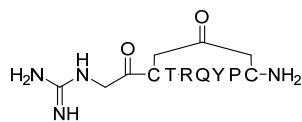


Figure S13 cont. | Structure, purity and mass of monocyclic peptides.

219**220****223****224****225****Figure S14 | Structure, purity and mass of monocyclic peptides with unnatural derivatives.**

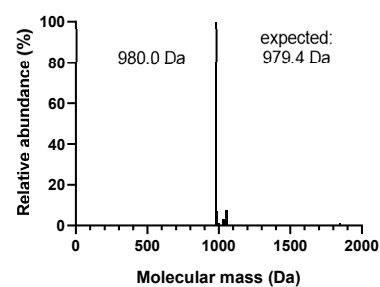
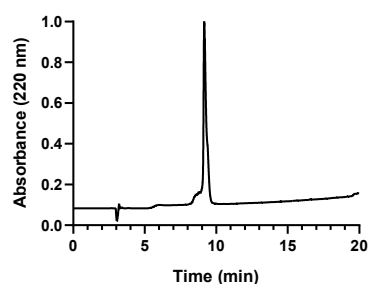
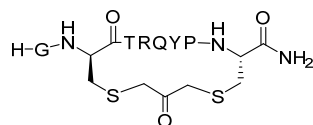
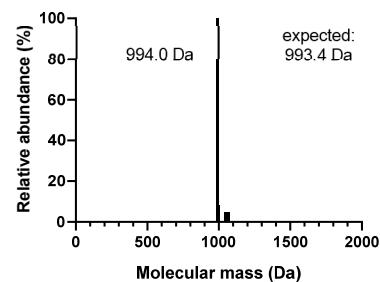
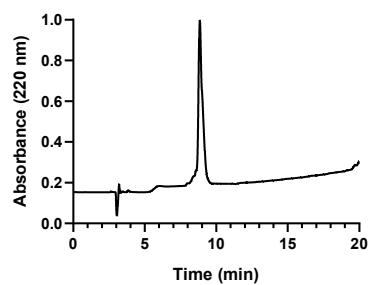
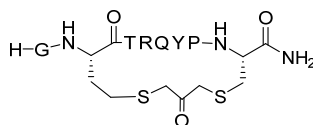
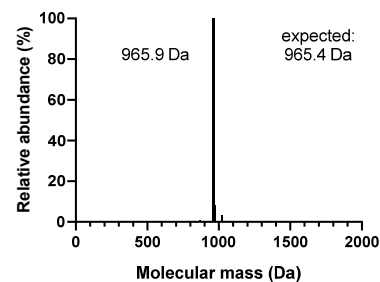
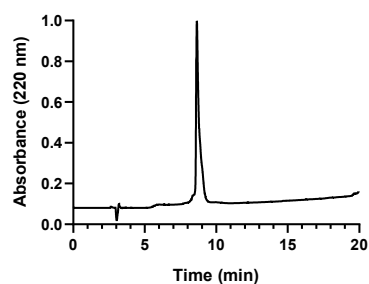
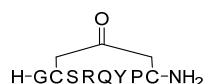
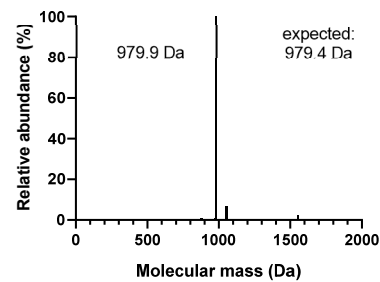
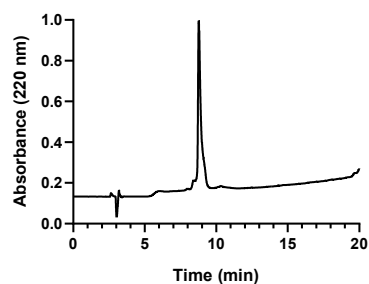
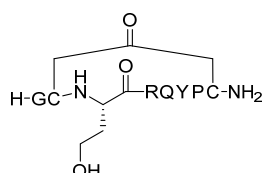
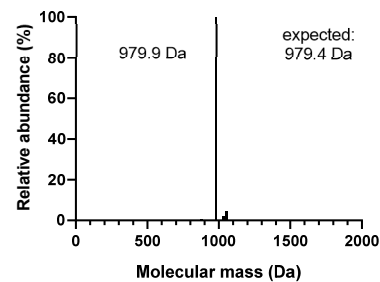
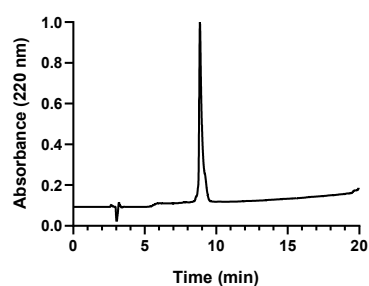
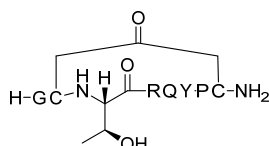
119**121****146****147****148**

Figure S14 cont. | Structure, purity and mass of monocyclic peptides with unnatural derivatives.

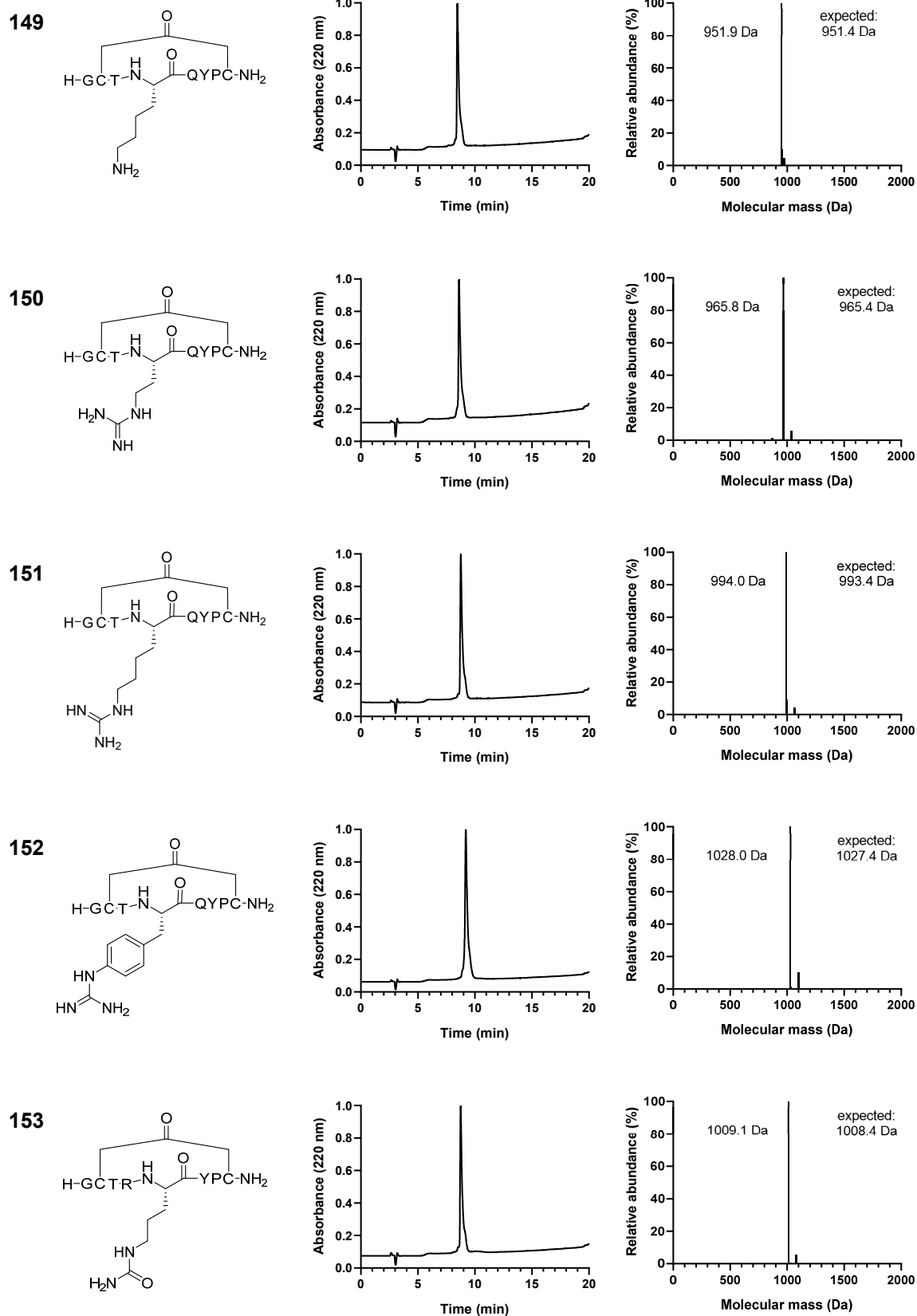


Figure S14 cont. | Structure, purity and mass of monocyclic peptides with unnatural derivatives.

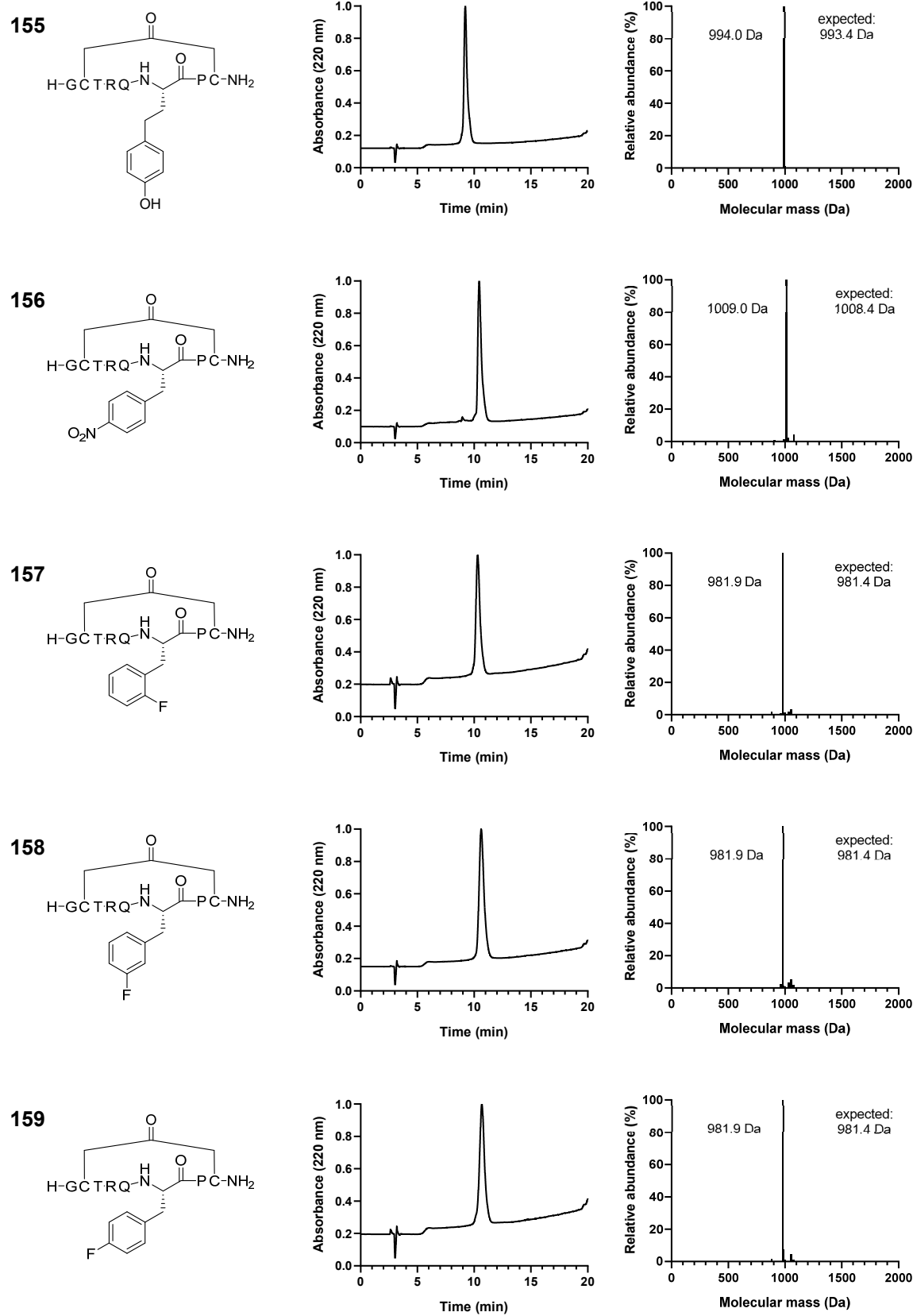


Figure S14 cont. I Structure, purity and mass of monocyclic peptides with unnatural derivatives.

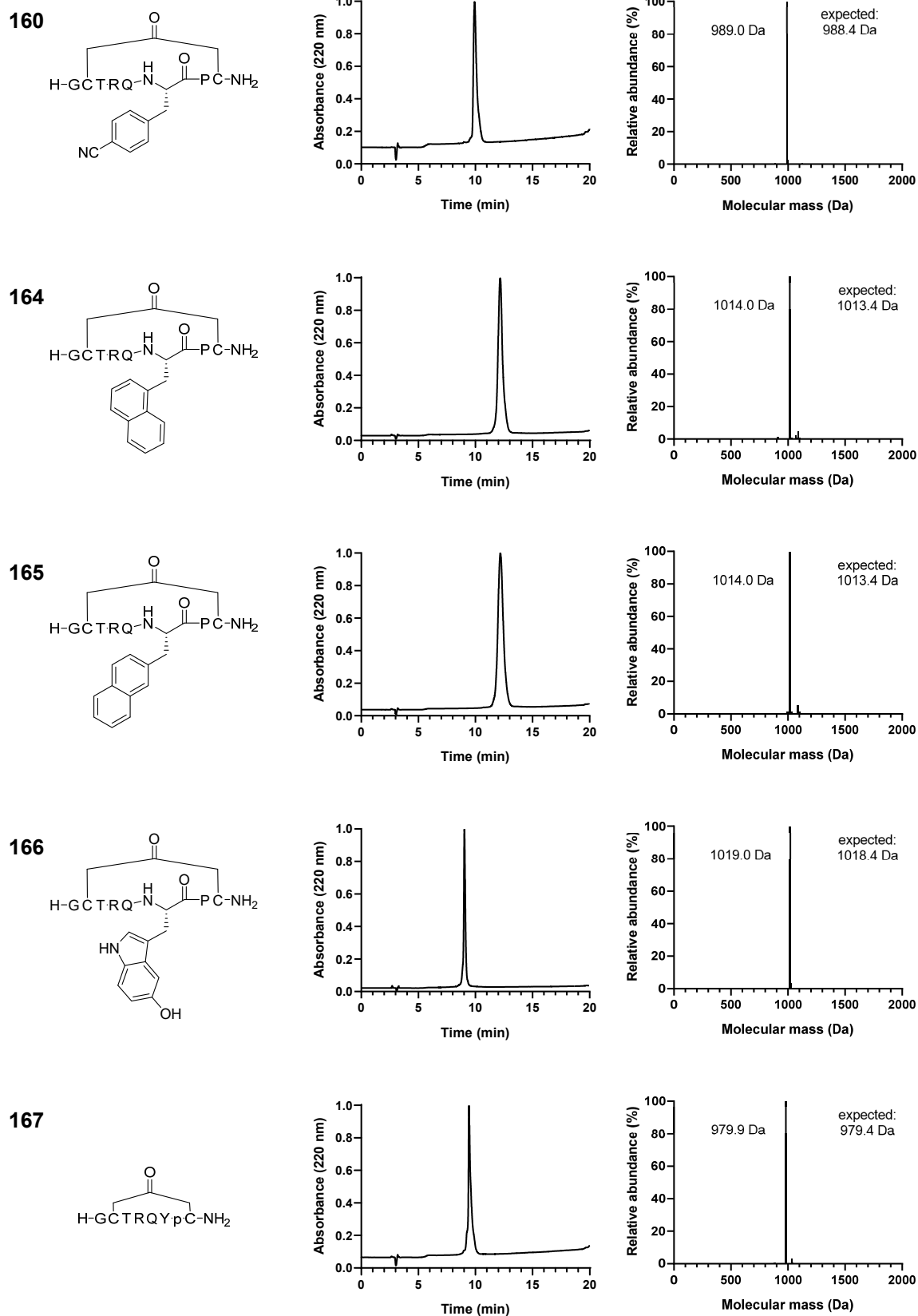
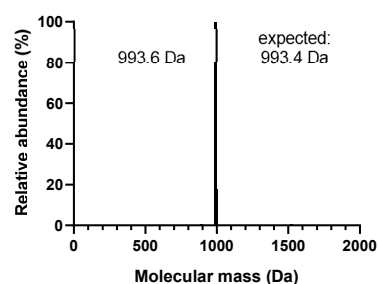
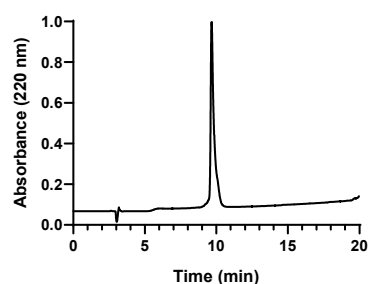
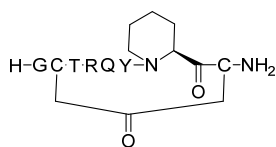
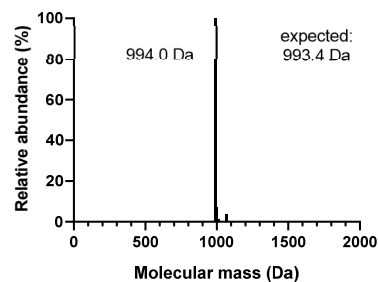
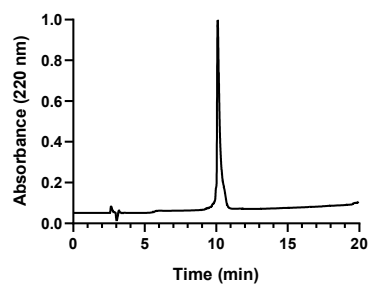
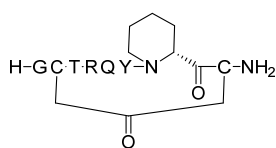


Figure S14 cont. | Structure, purity and mass of monocyclic peptides with unnatural derivatives.

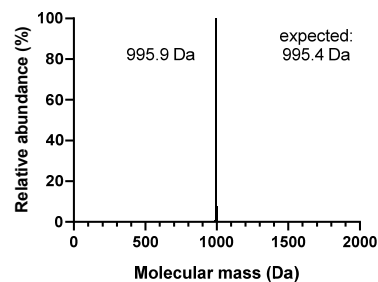
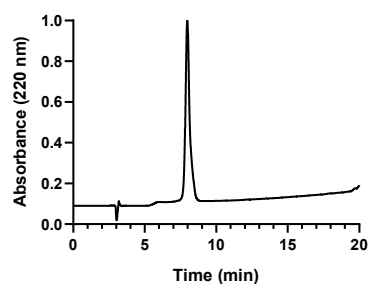
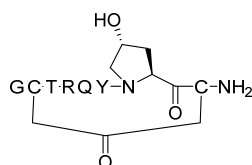
168



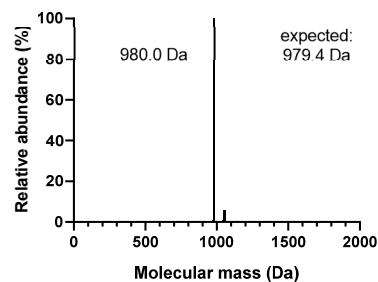
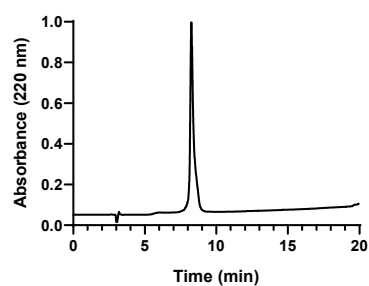
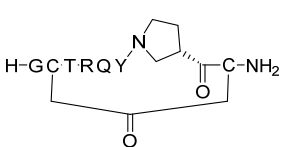
169



170



171



172

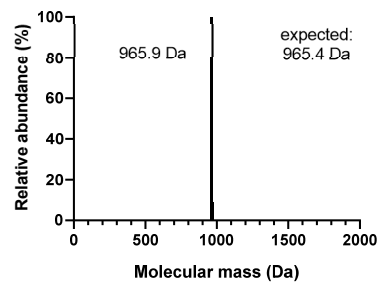
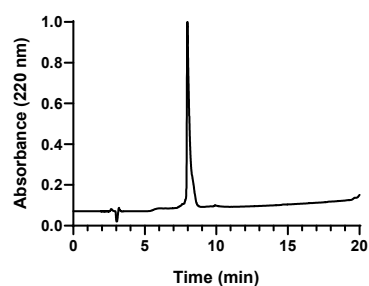
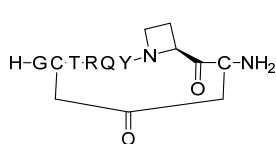


Figure S14 cont. | Structure, purity and mass of monocyclic peptides with unnatural derivatives.

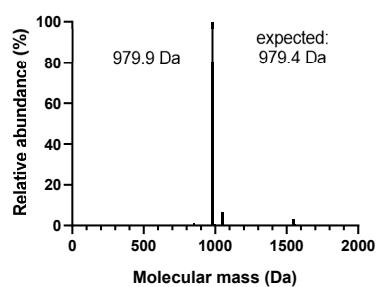
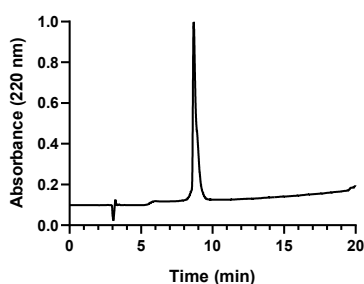
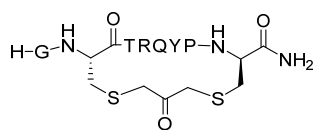
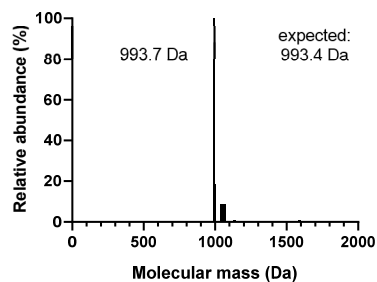
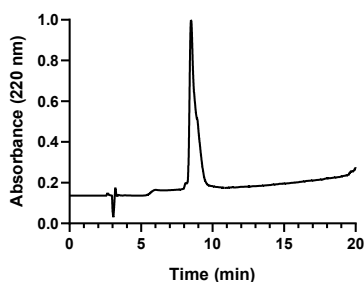
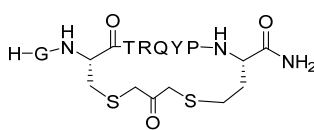
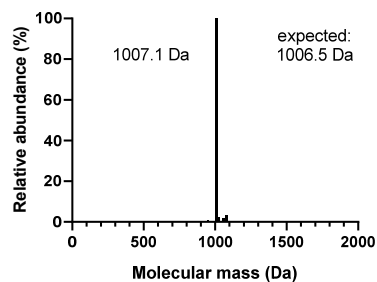
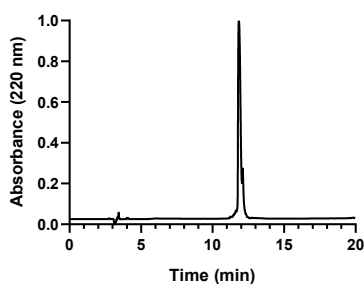
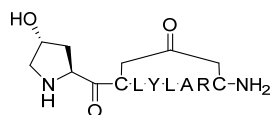
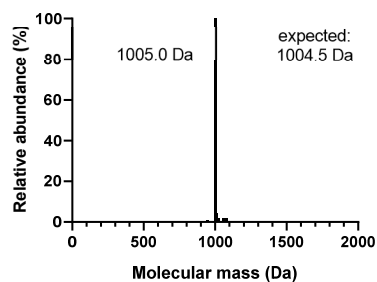
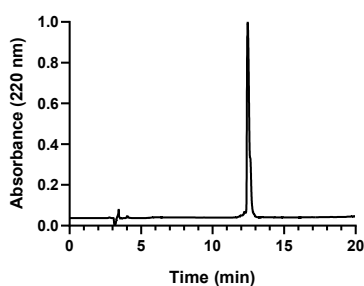
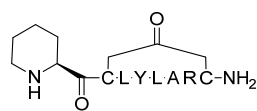
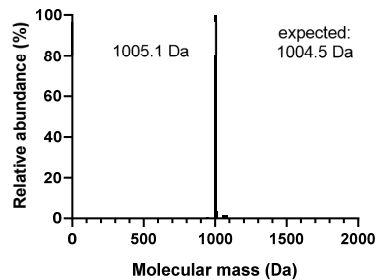
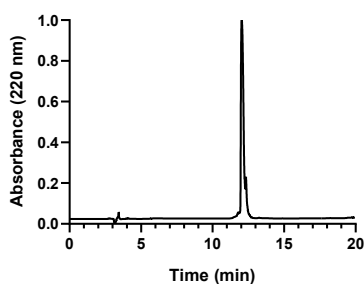
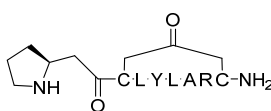
120**122****269****270****271**

Figure S14 cont. | Structure, purity and mass of monocyclic peptides with unnatural derivatives.

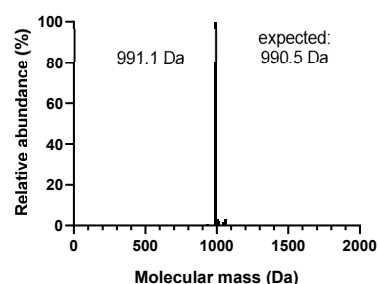
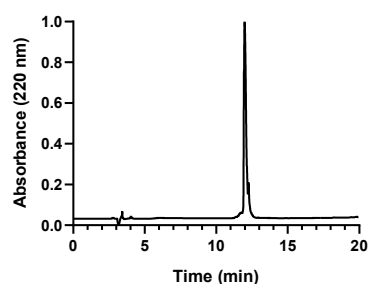
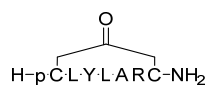
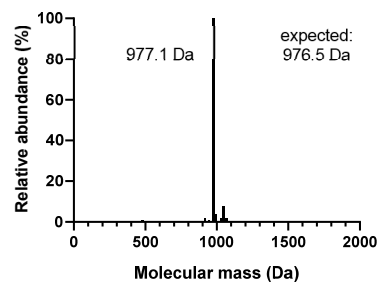
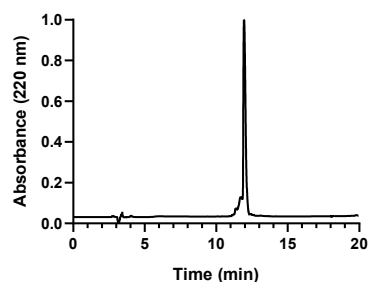
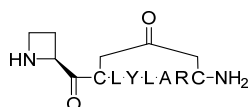
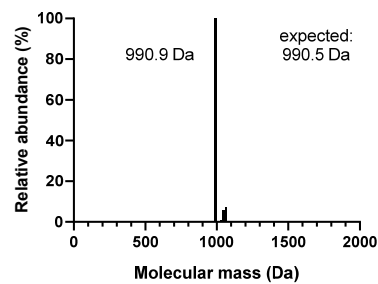
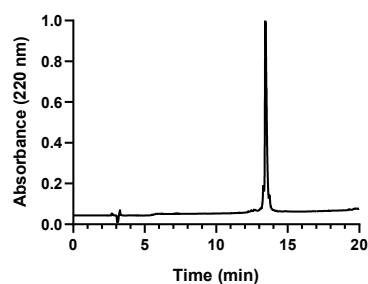
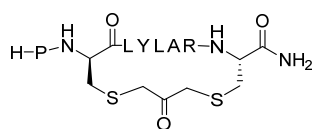
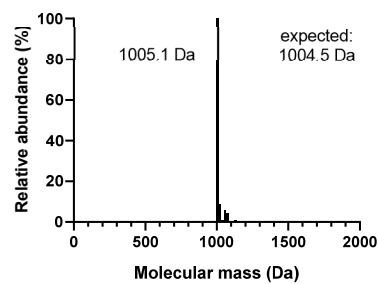
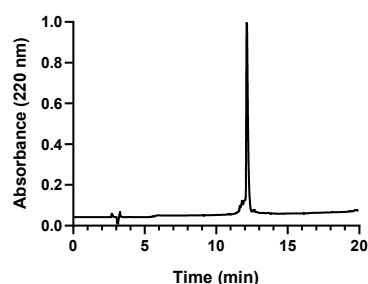
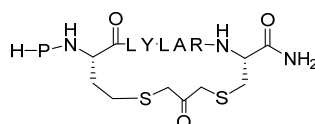
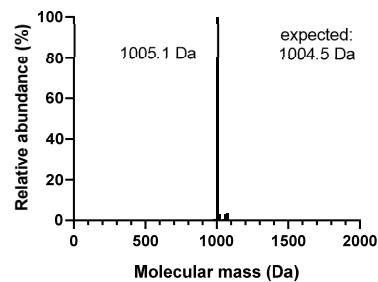
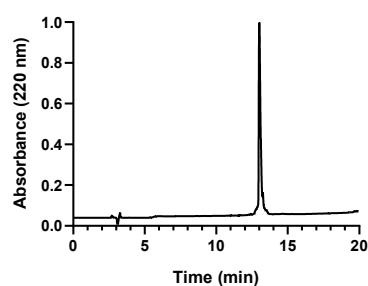
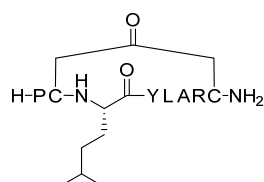
272**273****196****198****200**

Figure S14 cont. | Structure, purity and mass of monocyclic peptides with unnatural derivatives.

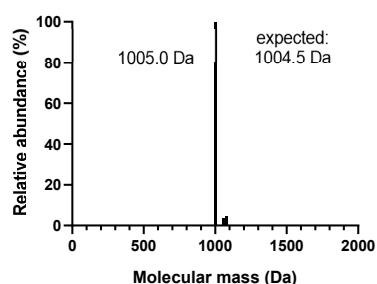
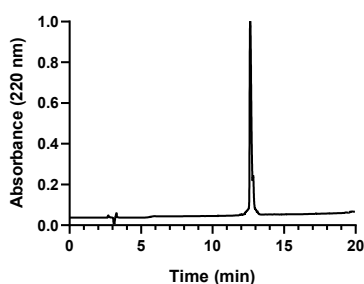
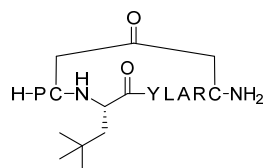
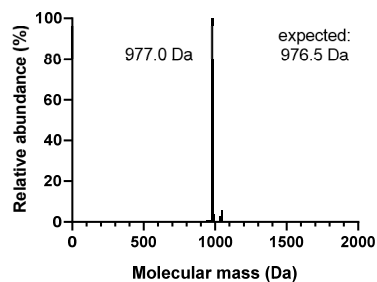
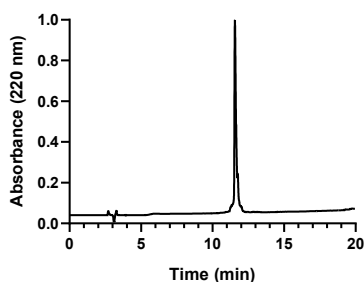
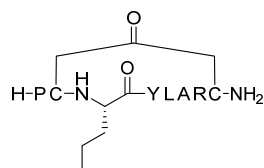
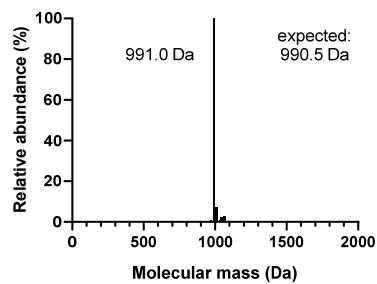
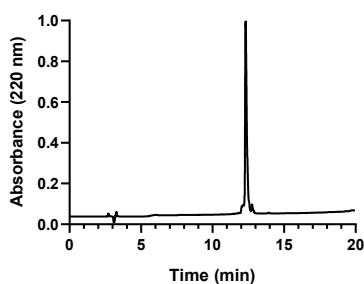
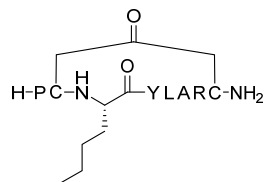
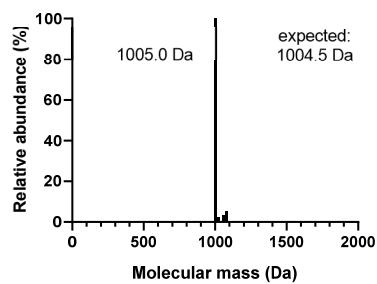
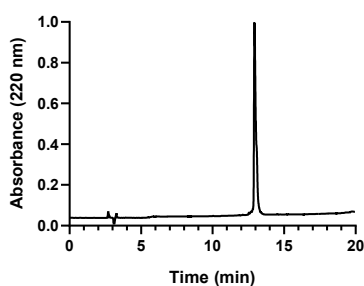
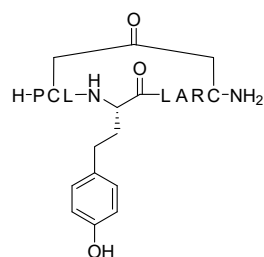
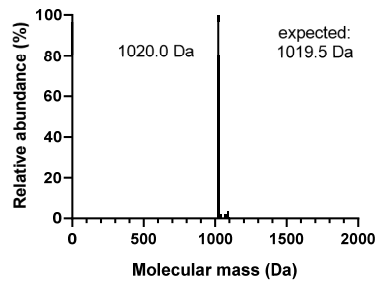
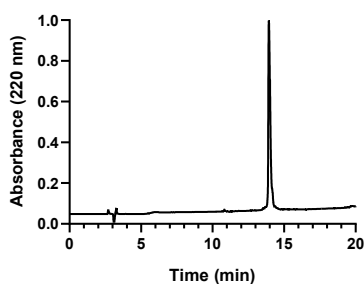
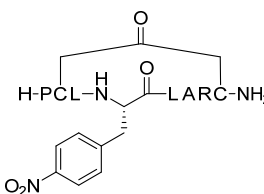
201**202****203****204****205**

Figure S14 cont. | Structure, purity and mass of monocyclic peptides with unnatural derivatives.

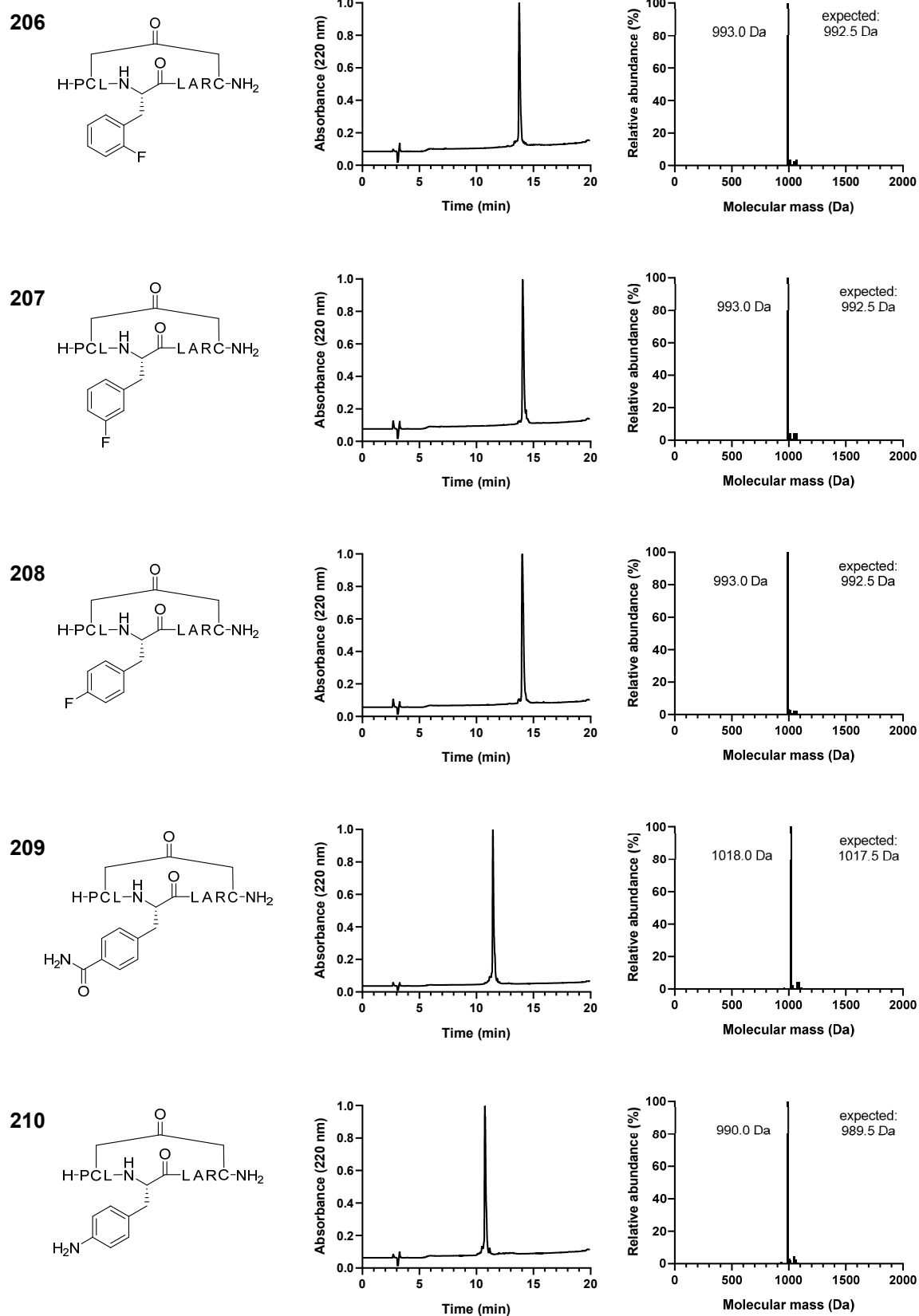


Figure S14 cont. | Structure, purity and mass of monocyclic peptides with unnatural derivatives.

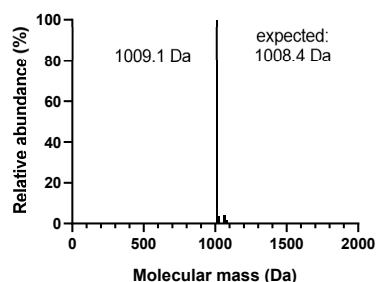
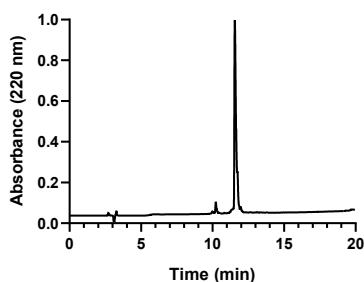
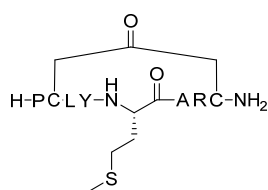
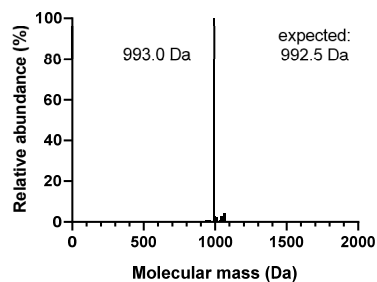
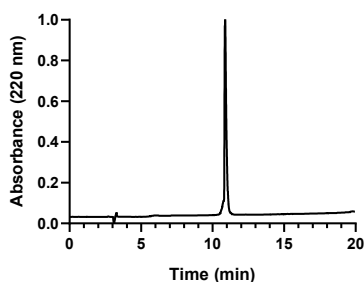
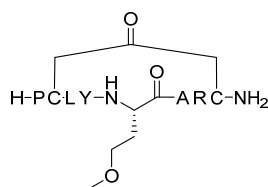
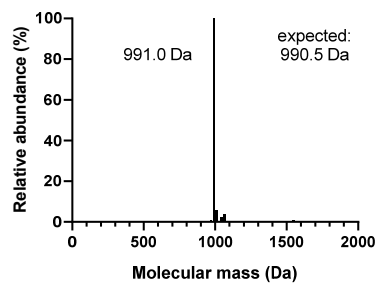
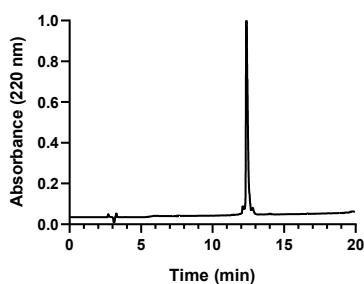
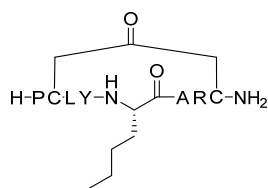
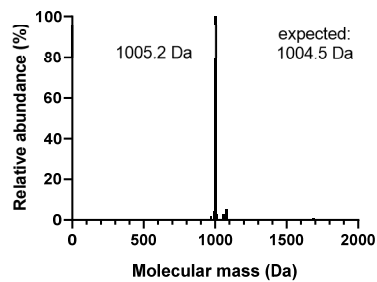
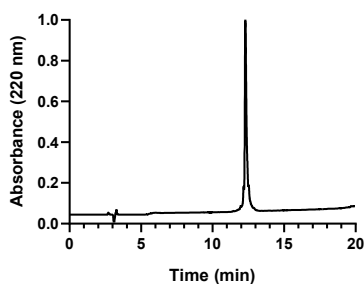
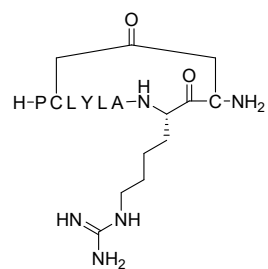
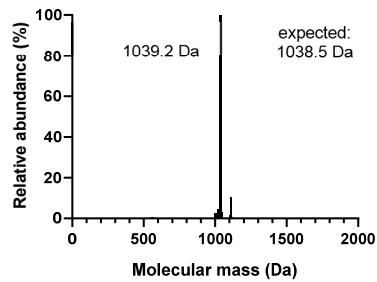
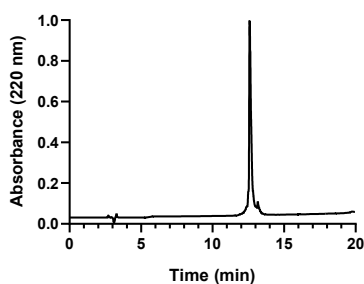
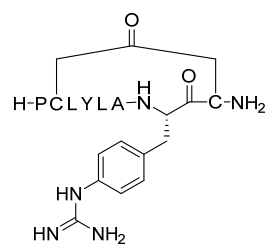
212**213****214****250****251**

Figure S14 cont. | Structure, purity and mass of monocyclic peptides with unnatural derivatives.

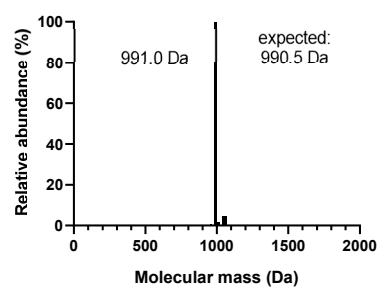
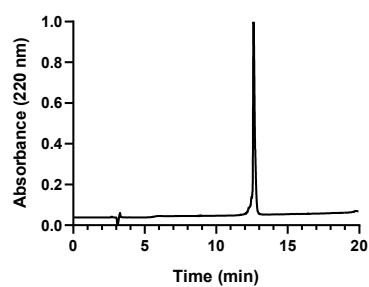
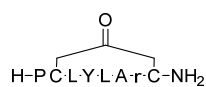
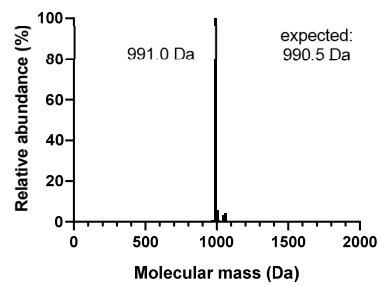
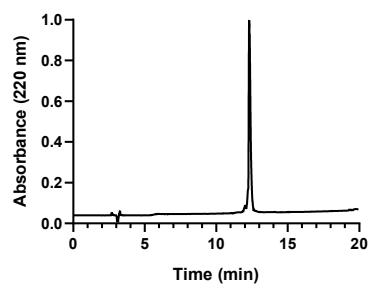
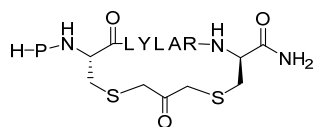
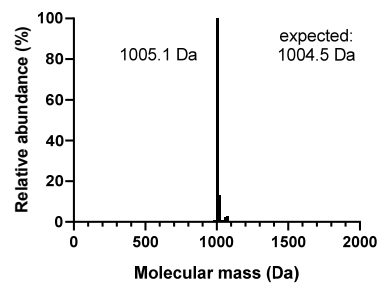
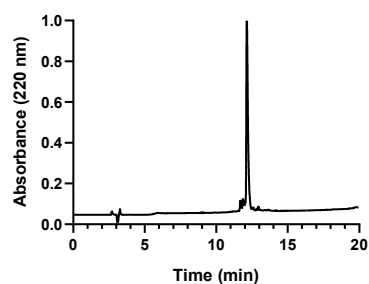
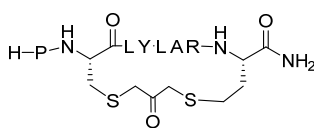
232**197****199**

Figure S14 cont. I Structure, purity and mass of monocyclic peptides with unnatural derivatives.

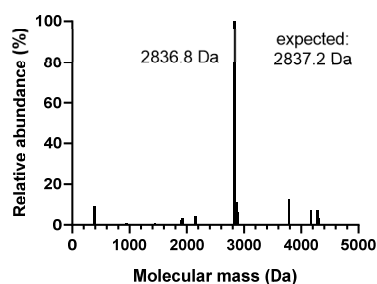
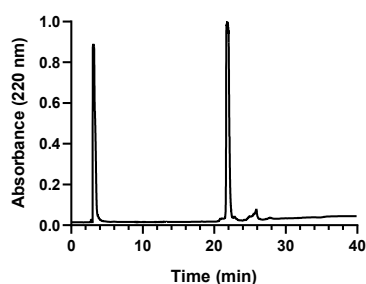
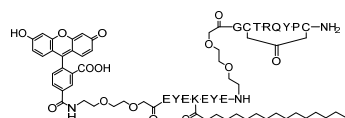
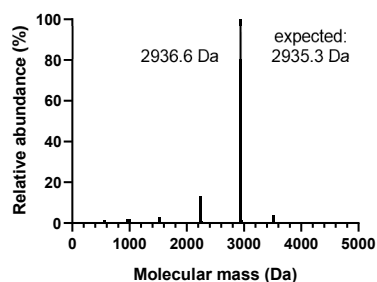
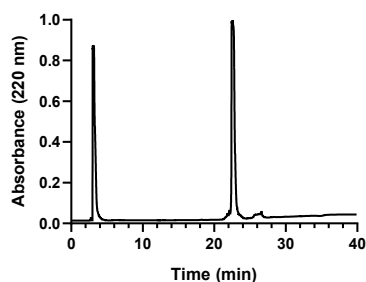
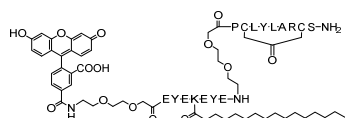
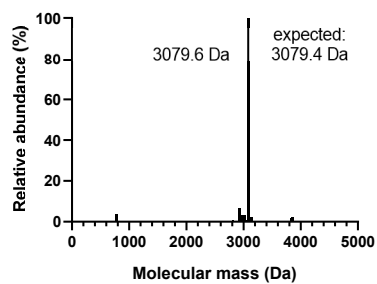
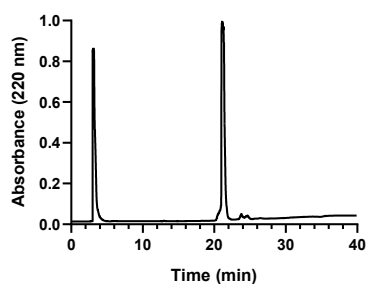
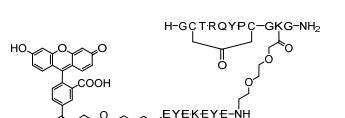
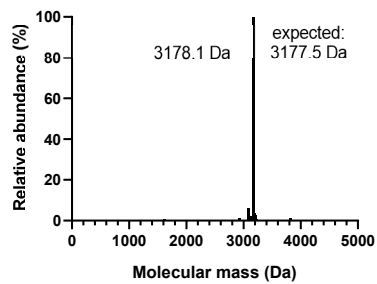
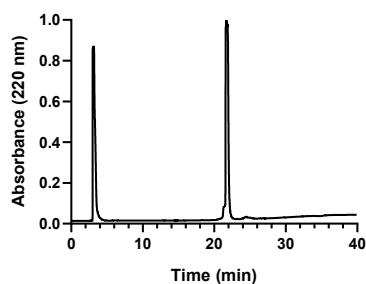
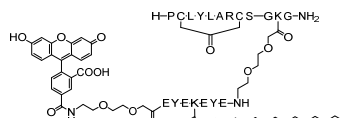
Tag-035**Tag-096****035-tag****096-tag**

Figure S15 | Structure, purity and mass of N- and C-terminal albumin tag conjugates.

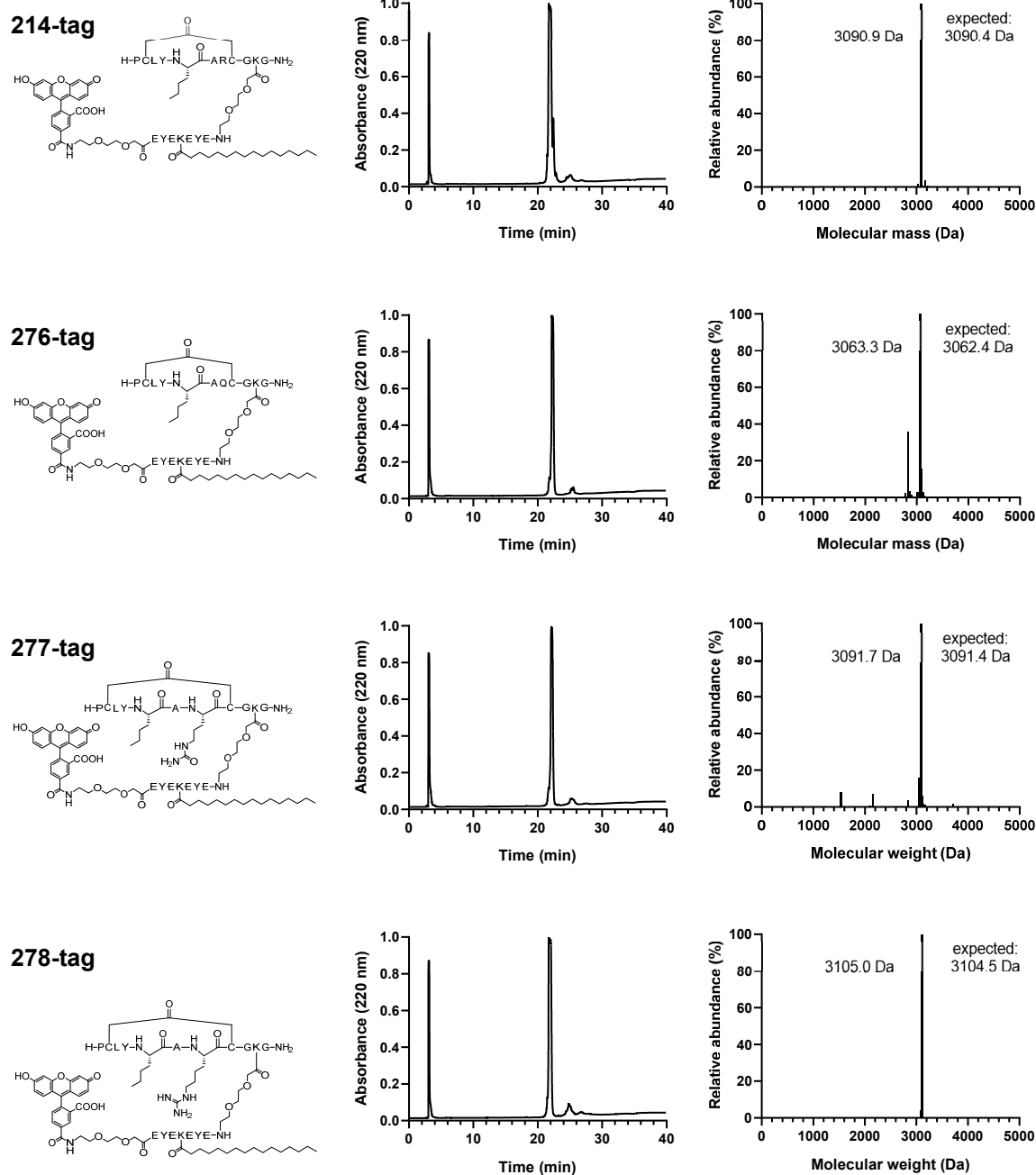


Figure S16 | Structure, purity and mass of KLK7 inhibitor tag conjugates with unnatural derivatives.

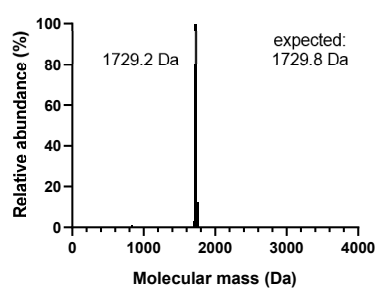
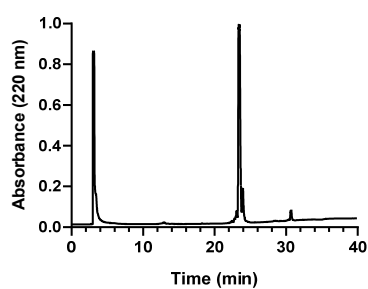
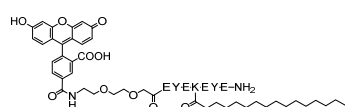
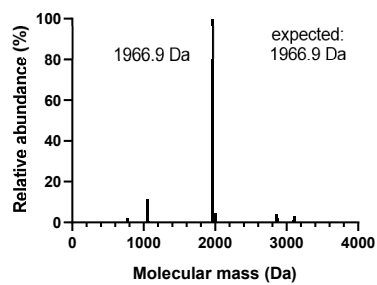
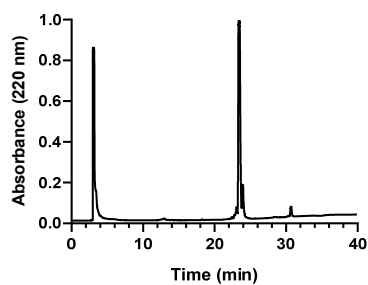
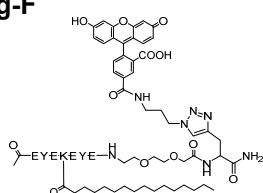
F-tag**Tag-F**

Figure S17 | Structure, purity and mass of albumin tags with N- and C-terminal fluorescein.

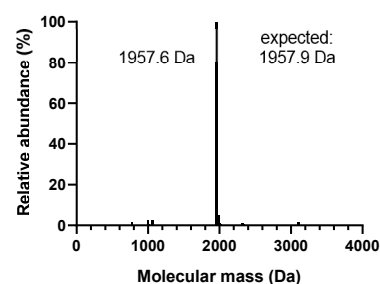
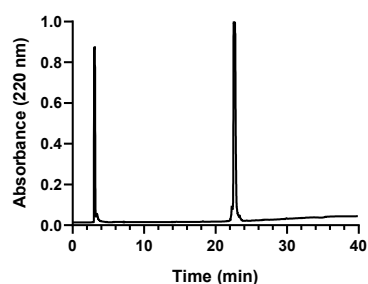
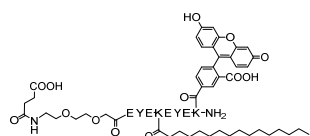
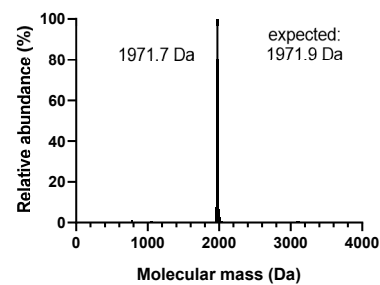
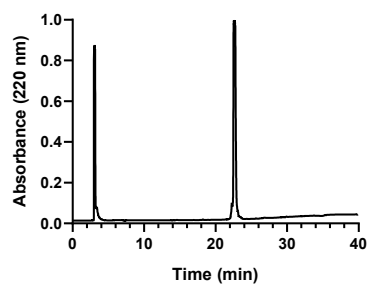
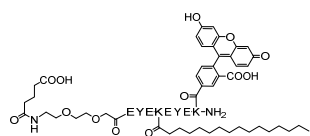
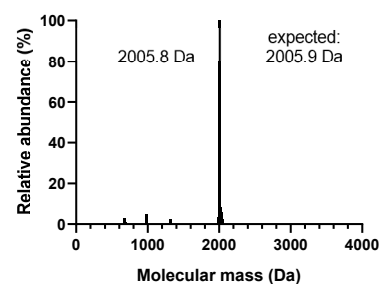
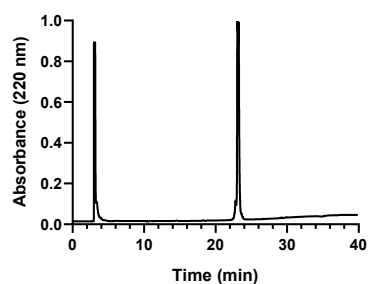
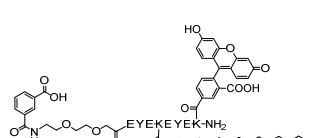
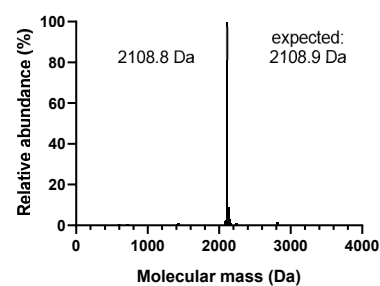
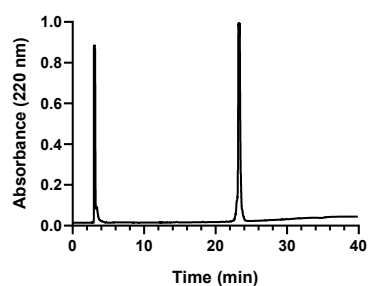
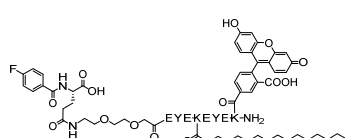
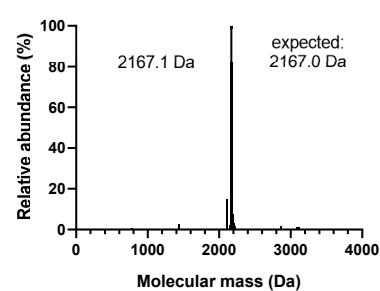
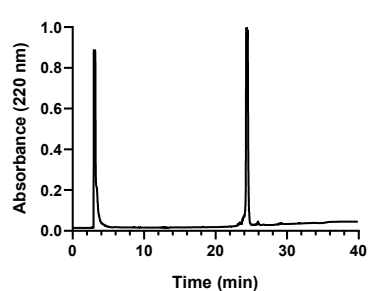
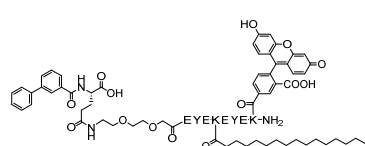
262**263****264****266****282**

Figure S18 | Structure, purity and mass of albumin tags with N-terminal fluorescein replacements.

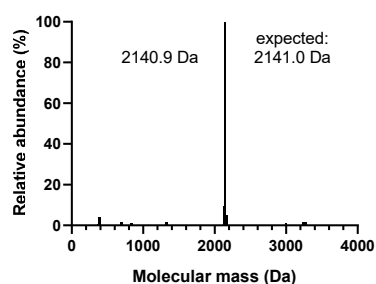
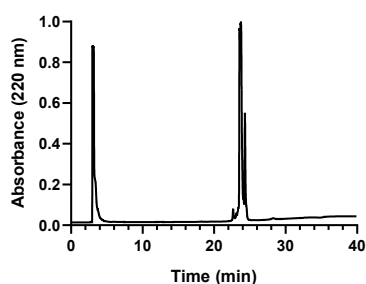
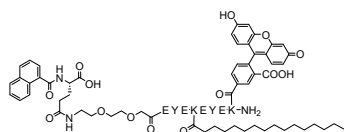
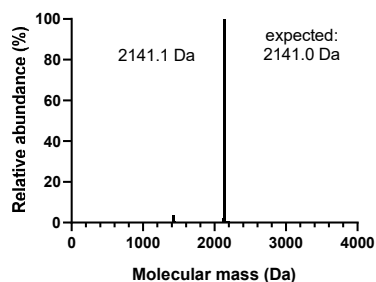
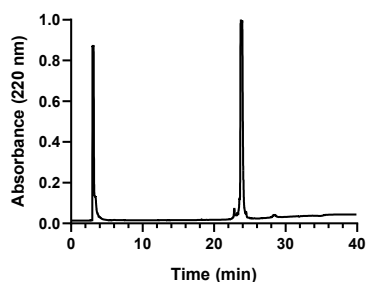
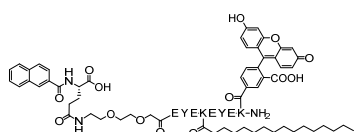
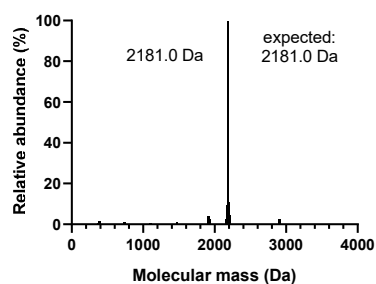
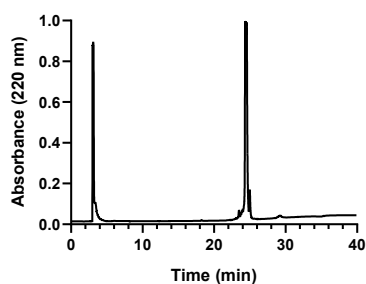
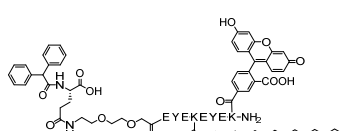
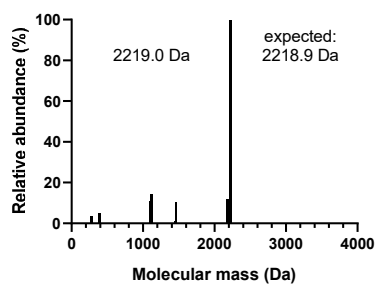
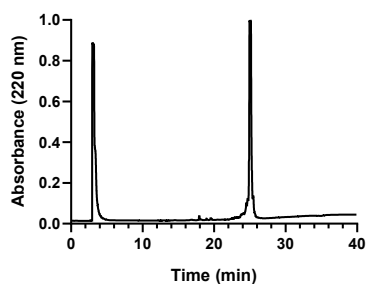
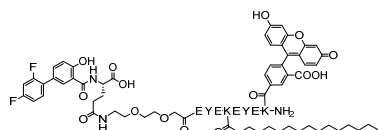
283**284****285****289**

Figure S18 cont. | Structure, purity and mass of albumin tags with N-terminal fluorescein replacements.

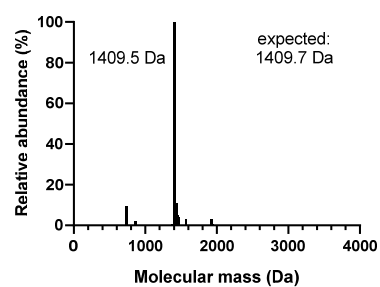
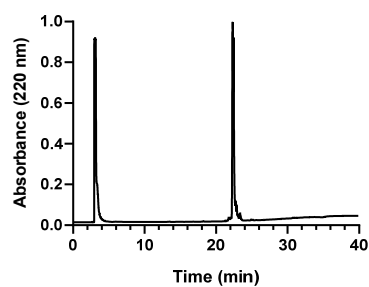
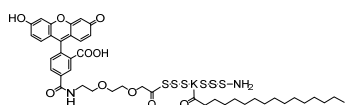
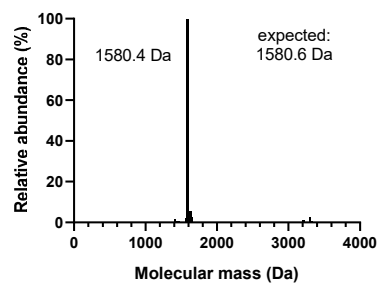
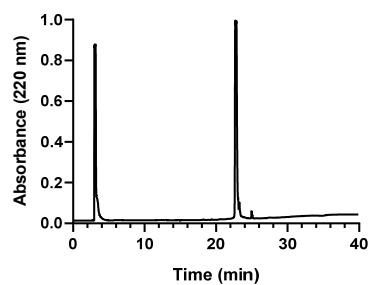
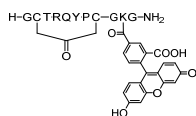
SSSKSSS-Tag**035-fluorescein**

Figure S19 | Structure, purity and mass of the two additional control peptides used in the pharmacokinetics experiments.

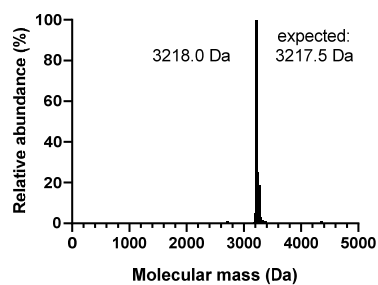
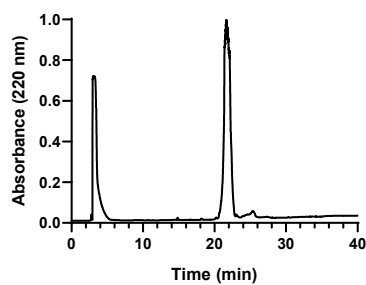
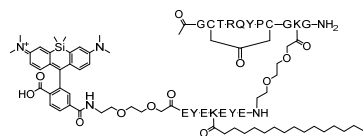
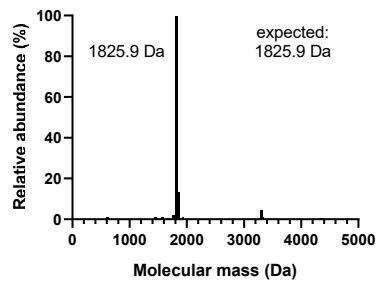
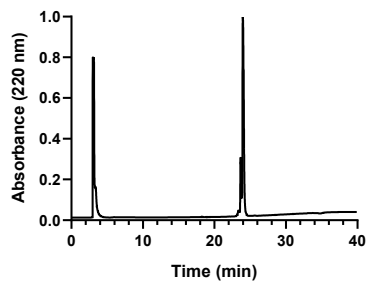
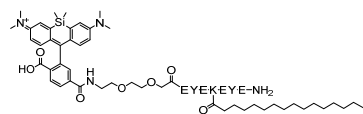
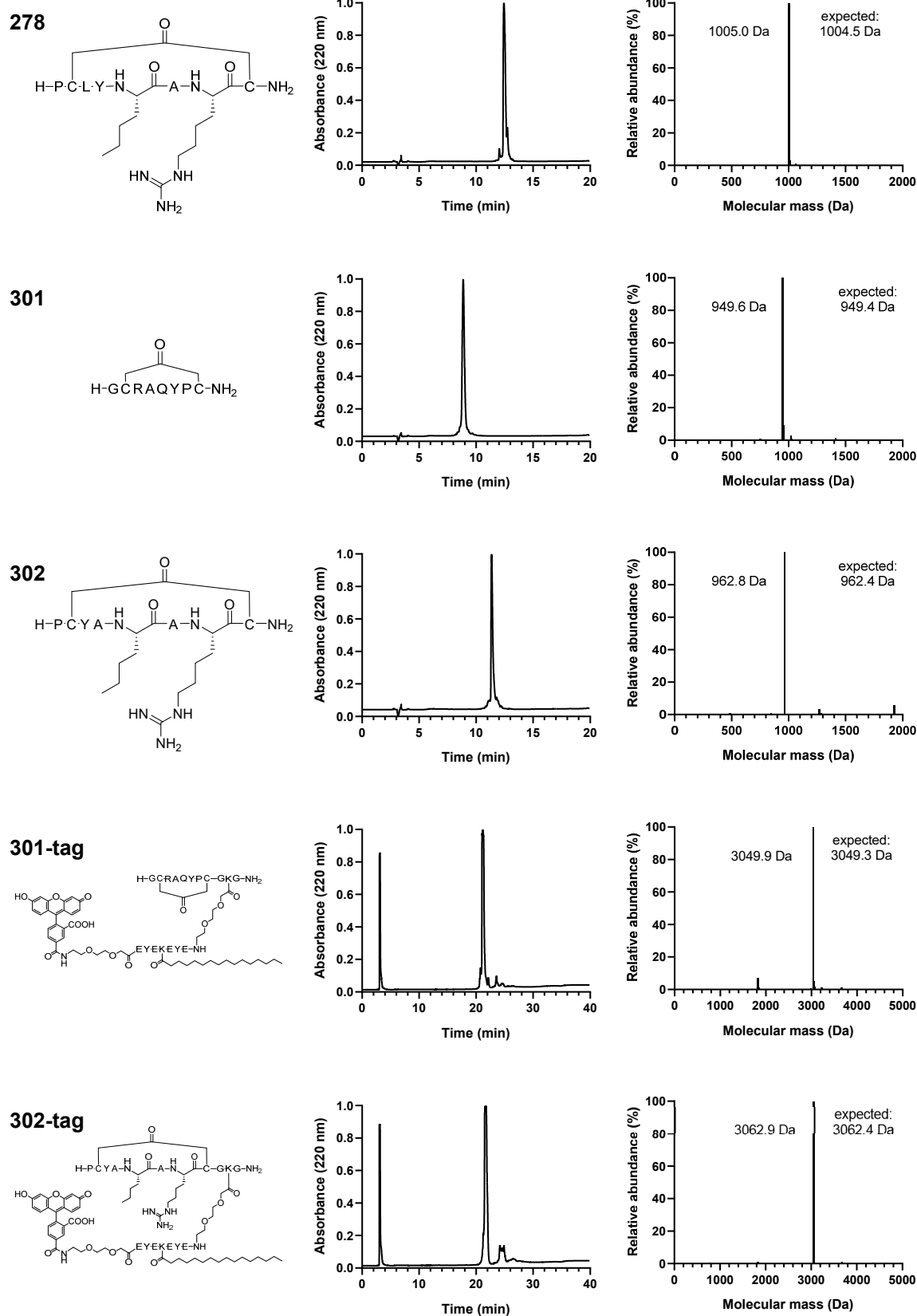
035-SiR-tag**SiR-tag**

Figure S20 | Structure, purity and mass of SiR conjugates used for skin distribution experiments. The synthesis of molecules is described in Figure 46 and Figure 47.



References

1. Orphanet. Netherton syndrome (2008). Available from: https://www.orpha.net/consor/cgi-bin/OC_Exp.php?Expert=634.
2. The Netherton Syndrome Group. About Netherton Syndrome. Available from: <http://www.nethertonsyndrome.com/about-nethertons.php>.
3. **Hovnanian, A.** Netherton syndrome: skin inflammation and allergy by loss of protease inhibition. *Cell Tissue Res* 351, 289-300 (2013).
4. **Comel, M.** Ichthyosis linearis circumflexa. *Dermatologica* 98, 133-136 (1949).
5. **Netherton, E.W.** A unique case of trichorrhexis nodosa; bamboo hairs. *AMA Arch Derm* 78, 483-487 (1958).
6. **Chavanas, S. et al.** Mutations in SPINK5, encoding a serine protease inhibitor, cause Netherton syndrome. *Nat Genet* 25, 141-142 (2000).
7. **Deraison, C. et al.** LEKTI fragments specifically inhibit KLK5, KLK7, and KLK14 and control desquamation through a pH-dependent interaction. *Mol Biol Cell* 18, 3607-3619 (2007).
8. **Prassas, I., Eissa, A., Poda, G. & Diamandis, E.P.** Unleashing the therapeutic potential of human kallikrein-related serine proteases. *Nat Rev Drug Discov* 14, 183-202 (2015).
9. **Hachem, J.-P. et al.** Serine Protease Activity and Residual LEKTI Expression Determine Phenotype in Netherton Syndrome. *Journal of Investigative Dermatology* 126, 1609-1621 (2006).
10. **Renner, E.D. et al.** Comel-Netherton syndrome defined as primary immunodeficiency. *The Journal of allergy and clinical immunology* 124, 536-543 (2009).
11. **Fontao, L. et al.** Infliximab infusions for Netherton syndrome: sustained clinical improvement correlates with a reduction of thymic stromal lymphopoietin levels in the skin. *J Invest Dermatol* 131, 1947-1950 (2011).
12. **Sotiropoulou, G. & Pampalakis, G.** Targeting the kallikrein-related peptidases for drug development. *Trends Pharmacol Sci* 33, 623-634 (2012).
13. **Xiang, N.N. & Di, W.-L.** Therapeutic interventions for Netherton syndrome. *Expert Review of Dermatology* 7, 319-321 (2012).
14. Gene Therapy for Netherton Syndrome (2012). Available from: <https://ClinicalTrials.gov/show/NCT01545323>.
15. Human skin structure, vector image ID: 136222271, © shutterstock.
16. L. Slomianka. Histology of human epidermis. Available from: <http://www.lab.anhb.uwa.edu.au/mb140/corepages/integumentary/integum.htm>.
17. **Watt, F.M.** Terminal differentiation of epidermal keratinocytes. *Current Opinion in Cell Biology* 1, 1107-1115 (1989).

18. **Rawlings, N.D., Morton, F.R., Kok, C.Y., Kong, J. & Barrett, A.J.** MEROPS: the peptidase database. *Nucleic Acids Res* 36, D320-325 (2008).
19. **Yousef, G.M. & Diamandis, E.P.** The new human tissue kallikrein gene family: structure, function, and association to disease. *Endocrine reviews* 22, 184-204 (2001).
20. **Pathak, M., Wong, S.S., Dreveny, I. & Emsley, J.** Structure of plasma and tissue kallikreins. *Thromb Haemost* 110, 423-433 (2013).
21. **Yousef, G.M., Chang, A., Scorilas, A. & Diamandis, E.P.** Genomic organization of the human kallikrein gene family on chromosome 19q13.3-q13.4. *Biochemical and biophysical research communications* 276, 125-133 (2000).
22. **Pavlopoulou, A., Pampalakis, G., Michalopoulos, I. & Sotiropoulou, G.** Evolutionary history of tissue kallikreins. *PLoS One* 5, e13781 (2010).
23. **Yoon, H. et al.** Activation profiles and regulatory cascades of the human kallikrein-related peptidases. *J Biol Chem* 282, 31852-31864 (2007).
24. **Goettig, P., Magdolen, V. & Brandstetter, H.** Natural and synthetic inhibitors of kallikrein-related peptidases (KLKs). *Biochimie* 92, 1546-1567 (2010).
25. **Polgar, L.** The catalytic triad of serine peptidases. *Cellular and molecular life sciences : CMLS* 62, 2161-2172 (2005).
26. **Debela, M. et al.** Structural basis of the zinc inhibition of human tissue kallikrein 5. *J Mol Biol* 373, 1017-1031 (2007).
27. **Turk, B.** Targeting proteases: successes, failures and future prospects. *Nat Rev Drug Discov* 5, 785-799 (2006).
28. **Debela, M. et al.** Specificity profiling of seven human tissue kallikreins reveals individual subsite preferences. *J Biol Chem* 281, 25678-25688 (2006).
29. **Magert, H.J. et al.** LEKTI, a novel 15-domain type of human serine proteinase inhibitor. *J Biol Chem* 274, 21499-21502 (1999).
30. **Laskowski, M. & Qasim, M.A.** What can the structures of enzyme-inhibitor complexes tell us about the structures of enzyme substrate complexes? *Biochim Biophys Acta* 1477, 324-337 (2000).
31. **Fortugno, P. et al.** Proteolytic activation cascade of the Netherton syndrome-defective protein, LEKTI, in the epidermis: implications for skin homeostasis. *J Invest Dermatol* 131, 2223-2232 (2011).
32. **Lauber, T., Schulz, A., Schweimer, K., Adermann, K. & Marx, U.C.** Homologous Proteins with Different Folds: The Three-dimensional Structures of Domains 1 and 6 of the Multiple Kazal-type Inhibitor LEKTI. *Journal of Molecular Biology* 328, 205-219 (2003).
33. **Tidow, H. et al.** The solution structure of a chimeric LEKTI domain reveals a chameleon sequence. *Biochemistry* 43, 11238-11247 (2004).
34. **Fernandez, I.S. et al.** Crystal structure of human epidermal kallikrein 7 (hK7) synthesized directly in its native state in *E. coli*: insights into the atomic basis of its inhibition by LEKTI domain 6 (LD6). *J Mol Biol* 377, 1488-1497 (2008).
35. **de Veer, S.J. et al.** Engineered protease inhibitors based on sunflower trypsin inhibitor-1 (SFTI-1) provide insights into the role of sequence and conformation in Laskowski mechanism inhibition. *Biochem J* 469, 243-253 (2015).

36. **de Veer, S.J., Swedberg, J.E., Brattsand, M., Clements, J.A. & Harris, J.M.** Exploring the active site binding specificity of kallikrein-related peptidase 5 (KLK5) guides the design of new peptide substrates and inhibitors. *Biol Chem* 397, 1237-1249 (2016).
37. **de Veer, S.J. et al.** Selective Substrates and Inhibitors for Kallikrein-Related Peptidase 7 (KLK7) Shed Light on KLK Proteolytic Activity in the Stratum Corneum. *J Invest Dermatol* 137, 430-439 (2017).
38. **de Veer, S.J., Wang, C.K., Harris, J.M., Craik, D.J. & Swedberg, J.E.** Improving the Selectivity of Engineered Protease Inhibitors: Optimizing the P2 Prime Residue Using a Versatile Cyclic Peptide Library. *J Med Chem* 58, 8257-8268 (2015).
39. **Krastel, P., Liechty, B.-M., Meingassner, J.G., Shmitt, E. & Schreiner, E.P.** Novartis Institutes for Biomedical Research Inc. Cyclic Depsipeptides, US 2009/0156472 A1. (2009).
40. Imagine Institute and GlaxoSmithKline (GSK) Agree on a Collaborative Partnership for Drug Discovery for rare diseases. [Press Release]. (2012). updated 12 Nov. Available from: <https://www.inserm-transfert.com/imagine-institute-and-glaxosmithkline-gsk-agree-on-a-collaborative-partnership-for-drug-discovery-for-rare-diseases/>.
41. **White, G.V. et al.** Kallikrein 5 inhibitors identified through structure based drug design in search for a treatment for Netherton Syndrome. *Bioorg Med Chem Lett* 29, 821-825 (2019).
42. **Descargues, P. et al.** Spink5-deficient mice mimic Netherton syndrome through degradation of desmoglein 1 by epidermal protease hyperactivity. *Nat Genet* 37, 56-65 (2005).
43. **Kasperek, P. et al.** KLK5 and KLK7 Ablation Fully Rescues Lethality of Netherton Syndrome-Like Phenotype. *PLoS Genet* 13, e1006566 (2017).
44. **Furio, L. et al.** Transgenic kallikrein 5 mice reproduce major cutaneous and systemic hallmarks of Netherton syndrome. *J Exp Med* 211, 499-513 (2014).
45. **Kasperek, P. et al.** A viable mouse model for Netherton syndrome based on mosaic inactivation of the Spink5 gene. *Biol Chem* 397, 1287-1292 (2016).
46. Bayer. Small and large molecules. Available from: <http://pharma.bayer.com/en/innovation-partnering/technologies-and-trends/small-and-large-molecules/>.
47. Kent D, R.S., Di Biase S,. Disruption and maturity: The next phase of biologics: QuintilesIMS; (2017). Available from: https://www.iqvia.com/-/media/iqvia/pdfs/nemea/uk/disruption_and_maturity_the_next_phase_of_biologics.pdf.
48. Junod, S.W. Celebrating a Milestone: FDA's Approval of First Genetically-Engineered Product (2007). Available from: <https://www.fda.gov/media/110447/download>.
49. **Mullard, A.** 2018 FDA drug approvals. *Nat Rev Drug Discov* 18, 85-89 (2019).
50. **Gongora-Benitez, M., Tulla-Puche, J. & Albericio, F.** Multifaceted roles of disulfide bonds. Peptides as therapeutics. *Chem Rev* 114, 901-926 (2014).
51. **Shah, D.K. & Betts, A.M.** Antibody biodistribution coefficients. *mAbs* 5, 297-305 (2013).
52. **Irani, V. et al.** Molecular properties of human IgG subclasses and their implications for designing therapeutic monoclonal antibodies against infectious diseases. *Mol Immunol* 67, 171-182 (2015).
53. **Vafa, O. et al.** An engineered Fc variant of an IgG eliminates all immune effector functions via structural perturbations. *Methods* 65, 114-126 (2014).
54. **Simeon, R. & Chen, Z.** In vitro-engineered non-antibody protein therapeutics. *Protein Cell* 9, 3-14 (2018).

55. **Kintzing, J.R., Filsinger Interrante, M.V. & Cochran, J.R.** Emerging Strategies for Developing Next-Generation Protein Therapeutics for Cancer Treatment. *Trends Pharmacol Sci* 37, 993-1008 (2016).
56. **Beck, A., Goetsch, L., Dumontet, C. & Corvaia, N.** Strategies and challenges for the next generation of antibody-drug conjugates. *Nat Rev Drug Discov* 16, 315-337 (2017).
57. **Craik, D.J., Fairlie, D.P., Liras, S. & Price, D.** The future of peptide-based drugs. *Chemical biology & drug design* 81, 136-147 (2013).
58. **Uhlig, T. et al.** The emergence of peptides in the pharmaceutical business: From exploration to exploitation. *EuPA Open Proteomics* 4, 58-69 (2014).
59. **Tsomaia, N.** Peptide therapeutics: targeting the undruggable space. *Eur J Med Chem* 94, 459-470 (2015).
60. **Diao, L. & Meibohm, B.** Pharmacokinetics and pharmacokinetic-pharmacodynamic correlations of therapeutic peptides. *Clinical pharmacokinetics* 52, 855-868 (2013).
61. **Zahnd, C. et al.** Efficient tumor targeting with high-affinity designed ankyrin repeat proteins: effects of affinity and molecular size. *Cancer research* 70, 1595-1605 (2010).
62. **Lau, J.L. & Dunn, M.K.** Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorganic & Medicinal Chemistry* 26, 2700-2707 (2018).
63. **Davies, M. et al.** Effect of Oral Semaglutide Compared With Placebo and Subcutaneous Semaglutide on Glycemic Control in Patients With Type 2 Diabetes: A Randomized Clinical Trial. *JAMA* 318, 1460-1470 (2017).
64. **Buckley, S.T. et al.** Transcellular stomach absorption of a derivatized glucagon-like peptide-1 receptor agonist. *Science translational medicine* 10 (2018).
65. Novartis. Full Year 2018 Product Sales. Available from: <https://www.novartis.com/investors/financial-data/product-sales>.
66. Pharmacompass. Product Sales Data From Annual Reports of Major Pharmaceutical Companies (2017). Available from: <https://www.pharmacompass.com/data-compilation/product-sales-data-from-annual-reports-of-major-pharmaceutical-companies-2017>.
67. **Zorzi, A., Deyle, K. & Heinis, C.** Cyclic peptide therapeutics: past, present and future. *Curr Opin Chem Biol* 38, 24-29 (2017).
68. **Banting, F.G., Best, C.H., Collip, J.B., Campbell, W.R. & Fletcher, A.A.** Pancreatic Extracts in the Treatment of Diabetes Mellitus. *Canadian Medical Association journal* 12, 141-146 (1922).
69. **Köhler, G. & Milstein, C.** Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 256, 495-497 (1975).
70. **Smith, G.P.** Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science (New York, N.Y.)* 228, 1315-1317 (1985).
71. **Davis, A.M., Plowright, A.T. & Valeur, E.** Directing evolution: the next revolution in drug discovery? *Nat Rev Drug Discov* 16, 681-698 (2017).
72. **McCafferty, J., Griffiths, A.D., Winter, G. & Chiswell, D.J.** Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* 348, 552-554 (1990).
73. **Jespers, L.S., Roberts, A., Mahler, S.M., Winter, G. & Hoogenboom, H.R.** Guiding the selection of human antibodies from phage display repertoires to a single epitope of an antigen. *Bio/Technology* 12, 899-903 (1994).

-
74. Drugbank. Adalimumab. Available from: <https://www.drugbank.ca/drugs/DB00051>.
 75. **Urquhart, L.** Top drugs and companies by sales in 2018. *Nat Rev Drug Discov* (2019).
 76. The Nobel Prize in Chemistry 2018. Available from: <https://www.nobelprize.org/prizes/chemistry/2018/summary/>.
 77. **Smith, G.P. & Petrenko, V.A.** Phage Display. *Chem Rev* 97, 391-410 (1997).
 78. **Obexer, R., Walport, L.J. & Suga, H.** Exploring sequence space: harnessing chemical and biological diversity towards new peptide leads. *Curr Opin Chem Biol* 38, 52-61 (2017).
 79. **Chasteen, L., Ayriss, J., Pavlik, P. & Bradbury, A.R.** Eliminating helper phage from phage display. *Nucleic Acids Res* 34, e145 (2006).
 80. **Rondot, S., Koch, J., Breitling, F. & Dubel, S.** A helper phage to improve single-chain antibody presentation in phage display. *Nat Biotechnol* 19, 75-78 (2001).
 81. **Boder, E.T. & Wittrup, K.D.** Yeast surface display for screening combinatorial polypeptide libraries. *Nat Biotechnol* 15, 553-557 (1997).
 82. **Devlin, J.J., Panganiban, L.C. & Devlin, P.E.** Random peptide libraries: a source of specific protein binding molecules. *Science (New York, N.Y.)* 249, 404 (1990).
 83. **Katz, B.A.** Structural and mechanistic determinants of affinity and specificity of ligands discovered or engineered by phage display. *Annual Review of Biophysics and Biomolecular Structure* 26, 27-45 (1997).
 84. **Li, P. & Roller, P.P.** Cyclization strategies in peptide derived drug design. *Current topics in medicinal chemistry* 2, 325-341 (2002).
 85. **Ladner, R.C.** Constrained peptides as binding entities. *Trends Biotechnol* 13, 426-430 (1995).
 86. **Giebel, L.B. et al.** Screening of cyclic peptide phage libraries identifies ligands that bind streptavidin with high affinities. *Biochemistry* 34, 15430-15435 (1995).
 87. **Wrighton, N.C. et al.** Small Peptides as Potent Mimetics of the Protein Hormone Erythropoietin. *Science (New York, N.Y.)* 273, 458 (1996).
 88. **Sinclair, A.M.** Erythropoiesis stimulating agents: approaches to modulate activity. *Biologics : targets & therapy* 7, 161-174 (2013).
 89. **Hermanson, T., Bennett, C.L. & Macdougall, I.C.** Peginesatide for the treatment of anemia due to chronic kidney disease - an unfulfilled promise. *Expert opinion on drug safety* 15, 1421-1426 (2016).
 90. **Heinis, C. & Winter, G.** Encoded libraries of chemically modified peptides. *Curr Opin Chem Biol* 26, 89-98 (2015).
 91. **Heinis, C., Rutherford, T., Freund, S. & Winter, G.** Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat Chem Biol* 5, 502-507 (2009).
 92. **Timmerman, P., Beld, J., Puijk, W.C. & Meloen, R.H.** Rapid and quantitative cyclization of multiple peptide loops onto synthetic scaffolds for structural mimicry of protein surfaces. *Chembiochem* 6, 821-824 (2005).
 93. **Rentero Rebollo, I. & Heinis, C.** Phage selection of bicyclic peptides. *Methods* 60, 46-54 (2013).
 94. **Rentero Rebollo, I., Sabisz, M., Baeriswyl, V. & Heinis, C.** Identification of target-binding peptide motifs by high-throughput sequencing of phage-selected peptides. *Nucleic Acids Res* 42, e169 (2014).

95. **Villequey, C., Kong, X.D. & Heinis, C.** Bypassing bacterial infection in phage display by sequencing DNA released from phage particles. *Protein Eng Des Sel* 30, 761-768 (2017).
96. **Baeriswyl, V. et al.** Bicyclic peptides with optimized ring size inhibit human plasma kallikrein and its orthologues while sparing paralogous proteases. *ChemMedChem* 7, 1173-1176 (2012).
97. **Angelini, A. et al.** Bicyclic peptide inhibitor reveals large contact interface with a protease target. *ACS Chem Biol* 7, 817-821 (2012).
98. **Baeriswyl, V. et al.** Development of a selective peptide macrocycle inhibitor of coagulation factor XII toward the generation of a safe antithrombotic therapy. *J Med Chem* 56, 3742-3746 (2013).
99. **Angelini, A. et al.** Chemical macrocyclization of peptides fused to antibody Fc fragments. *Bioconjug Chem* 23, 1856-1863 (2012).
100. **Angelini, A., Morales-Sanfrutos, J., Diderich, P., Chen, S. & Heinis, C.** Bicyclization and tethering to albumin yields long-acting peptide antagonists. *J Med Chem* 55, 10187-10197 (2012).
101. **Kale, S.S. et al.** Cyclization of peptides with two chemical bridges affords large scaffold diversities. *Nat Chem* 10, 715-723 (2018).
102. **Chen, S. et al.** Bicyclic Peptide Ligands Pulled out of Cysteine-Rich Peptide Libraries. *Journal of the American Chemical Society* 135, 6562-6569 (2013).
103. **Colgrave, M.L. & Craik, D.J.** Thermal, Chemical, and Enzymatic Stability of the Cyclotide Kalata B1: The Importance of the Cyclic Cystine Knot. *Biochemistry* 43, 5965-5975 (2004).
104. **White, C.J. & Yudin, A.K.** Contemporary strategies for peptide macrocyclization. *Nat Chem* 3, 509-524 (2011).
105. **Middendorp, S.J. et al.** Peptide Macrocycle Inhibitor of Coagulation Factor XII with Subnanomolar Affinity and High Target Selectivity. *J Med Chem* 60, 1151-1158 (2017).
106. **Howell, S.M. et al.** Serum stable natural peptides designed by mRNA display. *Scientific reports* 4, 6008 (2014).
107. **Baeriswyl, V. & Heinis, C.** Phage selection of cyclic peptide antagonists with increased stability toward intestinal proteases. *Protein Eng Des Sel* 26, 81-89 (2013).
108. **Tryggvason, K. & Wartiovaara, J.** How does the kidney filter plasma? *Physiology (Bethesda, Md.)* 20, 96-101 (2005).
109. **Kontermann, R.E.** Strategies for extended serum half-life of protein therapeutics. *Curr Opin Biotechnol* 22, 868-876 (2011).
110. **Kontermann, R.E.** Half-life extended biotherapeutics. *Expert opinion on biological therapy* 16, 903-915 (2016).
111. **van Witteloostuijn, S.B., Pedersen, S.L. & Jensen, K.J.** Half-Life Extension of Biopharmaceuticals using Chemical Methods: Alternatives to PEGylation. *ChemMedChem* 11, 2474-2495 (2016).
112. **Gaberc-Porekar, V., Zore, I., Podobnik, B. & Menart, V.** Obstacles and pitfalls in the PEGylation of therapeutic proteins. *Current opinion in drug discovery & development* 11, 242-250 (2008).
113. **Knop, K., Hoogenboom, R., Fischer, D. & Schubert, U.S.** Poly(ethylene glycol) in drug delivery: pros and cons as well as potential alternatives. *Angew Chem Int Ed Engl* 49, 6288-6308 (2010).

114. **Egrie, J.C. & Browne, J.K.** Development and characterization of darbepoetin alfa. *Oncology (Williston Park, N.Y.)* 16, 13-22 (2002).
115. **Fishburn, C.S.** The pharmacology of PEGylation: balancing PD with PK to generate novel therapeutics. *J Pharm Sci* 97, 4167-4183 (2008).
116. **Zorzi, A.** Development of an albumin-binding ligand for prolonging the plasma half-life of peptide therapeutics. 206 (2017).
117. **Roopenian, D.C. & Akilesh, S.** FcRn: the neonatal Fc receptor comes of age. *Nat Rev Immunol* 7, 715-725 (2007).
118. **Brambell, F.W.R.** The transmission of passive immunity from mother to young. *North Holland Research Monographs Frontiers of Biology* 18, xvi + 385 pp. (1970).
119. **Chaudhury, C. et al.** The major histocompatibility complex-related Fc receptor for IgG (FcRn) binds albumin and prolongs its lifespan. *J Exp Med* 197, 315-322 (2003).
120. **Sand, K.M. et al.** Unraveling the Interaction between FcRn and Albumin: Opportunities for Design of Albumin-Based Therapeutics. *Front Immunol* 5, 682 (2014).
121. **Larsen, M.T., Kuhlmann, M., Hvam, M.L. & Howard, K.A.** Albumin-based drug delivery: harnessing nature to cure disease. *Mol Cell Ther* 4, 3 (2016).
122. **Seijsing, J. et al.** An engineered affibody molecule with pH-dependent binding to FcRn mediates extended circulatory half-life of a fusion protein. *Proc Natl Acad Sci U S A* 111, 17110-17115 (2014).
123. **Meier, J.J.** GLP-1 receptor agonists for individualized treatment of type 2 diabetes mellitus. *Nat Rev Endocrinol* 8, 728-742 (2012).
124. **Bailey, T.S.** Clinical Efficacy of Once-weekly Glucagonlike Peptide-1 Receptor Agonists in Patients with Type 2 Diabetes. *The Journal of family practice* 67, S14-s24 (2018).
125. **Scheen, A.J.** Dulaglutide for the treatment of type 2 diabetes. *Expert opinion on biological therapy* 17, 485-496 (2017).
126. **Rendell, M.S.** Albiglutide for the management of type 2 diabetes. *Expert review of endocrinology & metabolism* 13, 1-8 (2018).
127. **Lau, J. et al.** Discovery of the Once-Weekly Glucagon-Like Peptide-1 (GLP-1) Analogue Semaglutide. *J Med Chem* 58, 7370-7380 (2015).
128. **Rader, A.F.B. et al.** Orally Active Peptides: Is There a Magic Bullet? *Angew Chem Int Ed Engl* 57, 14414-14438 (2018).
129. Cision. Novo Nordisk files for US FDA approval of oral semaglutide for blood sugar control and cardiovascular risk reduction in adults with type 2 diabetes (2019). Available from: <https://www.prnewswire.com/news-releases/novo-nordisk-files-for-us-fda-approval-of-oral-semaglutide-for-blood-sugar-control-and-cardiovascular-risk-reduction-in-adults-with-type-2-diabetes-300815668.html>.
130. **Wang, C.K. & Craik, D.J.** Cyclic peptide oral bioavailability: Lessons from the past. *Biopolymers* 106, 901-909 (2016).
131. **Di, L.** Strategic approaches to optimizing peptide ADME properties. *The AAPS journal* 17, 134-143 (2015).
132. **Brattsand, M., Stefansson, K., Lundh, C., Haasum, Y. & Egelrud, T.** A proteolytic cascade of kallikreins in the stratum corneum. *J Invest Dermatol* 124, 198-203 (2005).

133. **Matentzoglou, K. & Scheffner, M.** Ubiquitin-fusion protein system: a powerful tool for ectopic protein expression in mammalian cells. *Biotechniques* 46, 21-22, 24, 26 passim (2009).
134. **Michael, I.P. et al.** Biochemical and enzymatic characterization of human kallikrein 5 (hK5), a novel serine protease potentially involved in cancer progression. *J Biol Chem* 280, 14628-14635 (2005).
135. **Josephson, K., Ricardo, A. & Szostak, J.W.** mRNA display: from basic principles to macrocycle drug discovery. *Drug Discov Today* 19, 388-399 (2014).
136. **Timmerman, P., Puijk, W.C. & Meloen, R.H.** Immunogenic compounds and protein mimics, US 2010/0322945 A1 (2007).
137. **Zorzi, A., Middendorp, S.J., Wilbs, J., Deyle, K. & Heinis, C.** Acylated heptapeptide binds albumin with high affinity and application as tag furnishes long-acting peptides. *Nat Commun* 8, 16092 (2017).
138. **Banerjee, P.J. & Davies, N.P.** An Innocuous Adverse Effect of Routine Fundus Fluorescein AngiographyAn Innocuous Adverse Effect of Routine Fundus Fluorescein AngiographyAn Innocuous Adverse Effect of Routine Fundus Fluorescein Angiography. *JAMA Ophthalmology* 134, e153624-e153624 (2016).
139. **Ghuman, J. et al.** Structural basis of the drug-binding specificity of human serum albumin. *J Mol Biol* 353, 38-52 (2005).
140. **Yamasaki, K., Chuang, V.T., Maruyama, T. & Otagiri, M.** Albumin-drug interaction and its clinical implication. *Biochim Biophys Acta* 1830, 5435-5443 (2013).
141. **van Willigen, D.M. et al.** Multispectral fluorescence guided surgery; a feasibility study in a phantom using a clinical-grade laparoscopic camera system. *American journal of nuclear medicine and molecular imaging* 7, 138-147 (2017).
142. **Komatsu, N. et al.** Quantification of human tissue kallikreins in the stratum corneum: dependence on age and gender. *J Invest Dermatol* 125, 1182-1189 (2005).
143. **Lukinavicius, G. et al.** A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins. *Nat Chem* 5, 132-139 (2013).
144. Spirochrome. SiR-NHS ester. Available from: <https://spirochrome.com/product/sir-nhs-ester/>.
145. **Jensen, L. et al.** Absorption, metabolism and excretion of the GLP-1 analogue semaglutide in humans and nonclinical species. *Eur J Pharm Sci* 104, 31-41 (2017).
146. **Hoath, S.B. & Leahy, D.G.** The organization of human epidermis: functional epidermal units and phi proportionality. *J Invest Dermatol* 121, 1440-1446 (2003).
147. EMA. Assessment report, Ozempic, INN-semaglutide (14 December 2017). Available from: https://www.ema.europa.eu/en/documents/assessment-report/ozempic-epar-public-assessment-report_en.pdf.
148. FDA. Non-clinical review 209637Orig1s000 (2017). Available from: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2017/209637Orig1s000PharmR.pdf.
149. **Kapitza, C. et al.** Semaglutide, a once-weekly human GLP-1 analog, does not reduce the bioavailability of the combined oral contraceptive, ethinylestradiol/levonorgestrel. *Journal of clinical pharmacology* 55, 497-504 (2015).
150. **Furio, L. et al.** KLK5 Inactivation Reverses Cutaneous Hallmarks of Netherton Syndrome. *PLoS Genet* 11, e1005389 (2015).

151. **Tamura, K. et al.** MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution* 28, 2731-2739 (2011).
152. **Abbenante, G., Leung, D., Bond, T. & Fairlie, D.P.** An efficient Fmoc strategy for the rapid synthesis of peptide para-nitroanilides. *Letters in Peptide Science* 7, 347-351 (2000).
153. **Bankhead, P. et al.** QuPath: Open source software for digital pathology image analysis. *Scientific reports* 7, 16878 (2017).

PATRICK GONSCHOREK

Ph.D. in Biochemistry

Route du Lac 14
CH-1026 Denges
+49 151 51925163
patrick.gonschorek90@gmail.com



PROFESSIONAL EXPERIENCE

02/2016 – 07/2019

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Doctoral assistant in the Laboratory of Therapeutic Proteins and Peptides of Prof. Christian Heinis
Doctoral program in Chemistry and Chemical Engineering
PhD project: Development of cyclic peptide-based tissue kallikrein inhibitors for the treatment of Netherton syndrome.

04/2015 – 01/2016

BIOSS CENTRE FOR BIOLOGICAL SIGNALLING STUDIES FREIBURG

Research scientist in the group of Prof. Wilfried Weber
Projects: Development of novel optogenetic tools for research applications.

ACADEMIC EDUCATION

10/2012 – 01/2015

UNIVERSITY OF FREIBURG

Master studies in Translational and Synthetic Biology

Master's thesis in the lab of Prof. Wilfried Weber

Title: "Modularity in Synthetic Biology - Engineering Tools for Biomedical Application"

Projects: TALE-induced apoptosis as novel approach for the treatment of virus-infected cells;
Development of a novel enzyme-free cloning method.

08.01.2015

Degree: Master of Science (1.4 / "Very Good")

10/2009 – 09/2012

UNIVERSITY OF TÜBINGEN

Bachelor studies in Biology

Bachelor's Thesis in the lab of Prof. Ulrike Zentgraf, ZMBP Tübingen

Project: Regulation of the WRKY53 promoter through the transcription factors WRKY38 and WRKY62 in *Arabidopsis thaliana*.

13.09.2012

Degree: Bachelor of Science (1.6 / "Good")

ADDITIONAL PROJECTS

- 02/2018 – 06/2018 **INNOSUISSE START-UP TRAINING**
Module: Business Concept
Award: Best project
- 01/2013 – 10/2013 **iGEM 2013 COMPETITION**
Participation at the "International Genetically Engineered Machine Competition" (iGEM)
Project: Target specific gene regulation with the CRISPR/Cas9 system in mammalian cells
- Awards won at the European Jamboree in Lyon, FR:
- Best New BioBrick Device, Engineered (Overgraduates)
 - 1st runner up (Overgraduates)
 - Gold Medal and advancement to World Championship
- Awards won at the World Championship in Cambridge, MA, USA:
- Best Foundational Advance (Overgraduates)

SCIENTIFIC PUBLICATIONS

Dual-controlled optogenetic system for the rapid down-regulation of protein levels in mammalian cells.

Baaske J*, Gonschorek P*, Engesser R, Dominguez-Monedero A, Raute K, Fischbach P, Müller K, Cachat E, Schamel WWA, Minguet S, Davies JA, Timmer J, Weber W, Zurbriggen MD.

* equal contribution

Scientific Reports (2018) DOI: [10.1038/s41598-018-32929-7](https://doi.org/10.1038/s41598-018-32929-7)

AQUA Cloning: a versatile and simple enzyme-free cloning approach.

Beyer HM, Gonschorek P, Samodelov SL, Meier M, Weber W, Zurbriggen MD.

PLoS One (2015) DOI: [10.1371/journal.pone.0137652](https://doi.org/10.1371/journal.pone.0137652)

Modularized CRISPR/dCas9 Effector Toolkit for Target-Specific Gene Regulation.

Agne M, Blank I, Emhardt AJ, Gäbelein CG, Gawlas F, Gillich N, Gonschorek P, Juretschke TJ, Krämer SD, Louis N, Müller A, Rudolf A, Schäfer LM, Scheidmann MC, Schmunk LJ, Schwenk PM, Stammnitz MR, Warmer PM, Weber W, Fischer A, Kaufmann B, Wagner HJ, Radziwill G.

ACS Synthetic Biology (2014) DOI: [10.1021/sb500035y](https://doi.org/10.1021/sb500035y)

REFERENCES

Prof. Christian Heinis

École polytechnique fédérale de Lausanne
christian.heinis@epfl.ch

Prof. Wilfried Weber

University of Freiburg
wilfried.weber@biologie.uni-freiburg.de

