Four Models for Peptide Engineering

Dissertation zur Erlangung des akademischen Grades des
Doktors der Naturwissenschaften (Dr. rer. nat.)

eingereicht im Fachbereich Biologie, Chemie, Pharmazie
der Freien Universität Berlin
vorgelegt von

Kristin Folmert
aus Leipzig

Berlin, Dezember 2017

Why, sometimes I believe in as many as six impossible things before breakfast."

Lewis Carroll - ***Alice in Wonderland*** **(1865)**

Erklärung

Die vorliegende Arbeit wurde auf Anregung und unter Anleitung von Prof. Dr. Beate Koksch in der Zeit von Mai 2012 bis Dezember 2017 am Institut für Chemie und Biochemie des Fachbereiches Biologie, Chemie, Pharmazie der Freien Universität Berlin angefertigt. Hiermit versichere ich, die vorliegende Arbeit mit dem Titel „Four Models for Peptide Engineering" ohne Benutzung anderer als der zugelassenen Hilfsmittel selbständig angefertigt zu haben. Alle angeführten Zitate sind als solche kenntlich gemacht. Die Dissertation wurde in englischer Sprache verfasst. Die vorliegende Arbeit wurde in keinem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

_____

Berlin, Dezember 2017   Kristin Folmert

1. Gutachterin: Prof. Dr. Beate Koksch (Freie Universität Berlin)
2. Gutachter: Prof. Dr. Kevin Pagel (Freie Universität Berlin)

Disputation am: 16.10.2018

Publications

- **K. Folmert**, M. Broncel, H. v. Berlepsch, C. H. Ullrich, M. Siegert and B. Koksch, Inhibition of Peptide Aggregation by Means of Enzymatic Phosphorylation, *Beilstein J. Org. Chem.* **2016**, 12, 2462–2470.

- W. Hoffmann, **K. Folmert**, J. Moschner, X. Huang, H. v. Berlepsch, B. Koksch, M. T. Bowers, G. v. Helden and K. Pagel; NFGAIL Amyloid Oligomers: The Onset of Beta-Sheet Formation and the Mechanism for Fibril Formation, *J. Am. Chem. Soc.*, **2017**, accepted.

- S. Huhmann, A.-K. Stegemann, **K. Folmert**, D. Klemczak, J. Moschner, M. Kube and B. Koksch, Hexafluoroleucine and trifluoroisoleucine help to increase protease stability of peptides, *Beilstein J. Org. Chem., 2017,* accepted.

Oral Presentations

- "Inhibition of Peptide Aggregation by Means of Enzymatic Phosphorylation"; German Peptide Symposium, in München/Deutschland am 18.03.**2013**

- „Rosi, Fritz und Torsten – eine Liebesgeschichte aus der Peptidchemie, oder warum *klein* nicht immer schlechter ist"; ChromForum 2015, in Leipzig/Deutschland am 09.07.**2015**

- "Polymer Caging for Aggregating Peptides – Intracellular Delivery, Light-induced Release and Conformational Switching"; Symposium: "Protein Folding, Dynamics and Stability", in Halle/Deutschland am 21.10.**2016**

Poster Presentations

- "Inhibition of Peptide Aggregation by Means of Enzymatic Phosphorylation"; German Peptide Symposium, in München/Deutschland am 20.03.**2013**

- "Inhibition of Peptide Aggregation by Means of Enzymatic Phosphorylation"; Berliner Chemie Symposium, in Berlin/Deutschland am 09.04.**2013**

- "Polymer Caging for Aggregating Peptides – Intracellular Delivery, Light-induced Release and Conformational Switching"; Berliner Chemie Symposium, in Berlin/Deutschland am 13.04.**2016**

- "Polymer Caging for Aggregating Peptides – Intracellular Delivery, Light-induced Release and Conformational Switching"; Tag der Chemie, in Berlin/Deutschland am 18.06.**2016**

- "Polymer Caging for Aggregating Peptides – Intracellular Delivery, Light-induced Release and Conformational Switching"; European Peptide Symposium, in Leipzig/Deutschland am 04.09.**2016**

Danksagung

Ich danke Frau Prof. Dr. Beate Koksch für Ihr Vertrauen in mich und Ihre allseits gewährte Unterstützung sowie für die Chance mich in Ihrer Arbeitsgruppe auch über die Forschung hinaus zu entwickeln und weiterzubilden. Ich danke Ihr für die herausfordernden Projekte, die sie mir anvertraut hat und für den Freiraum auch meine eigenen wissenschaftlichen Ideen umzusetzen.

Herrn Jun.-Prof. Dr. Kevin Pagel danke ich für die fruchtbare Kooperation im NFGAIL Projekt und für die Übernahme des Zweitgutachtens dieser Dissertation.

Ohne die hervorragende und kreative Zusammenarbeit mit M.Sc. Waldemar Hoffmann und Dr. Johann Moschner wäre das spannende NFGAIL Projekt nicht zustande gekommen. Gleiches gilt für M.Sc. Susanne Huhmann, ohne die das Protease und auch so manch anderes Projekt ganz anders verlaufen wäre.

Herrn Dr. Andreas Schäfer Danke ich für seinen großen Einsatz beim Messen und seine Hilfe beim Auswerten der aufwändigen ^{31}P NMR Kinetik. Frau Carola Feindt danke ich für die Einführung in die spannende Welt der HPLC Technik. Für die SAXS Messungen vom NFGAIL danke ich Frau M.Sc. Claudia Kästner. Mein Dank gilt ebenfalls Herrn Dr. Hans von Berlepsch und Herrn Dr. Boris Schade für das Anfertigen der elektronenmikroskopischen Aufnahmen. Dem Gerätezentrum BioSupraMol verdanke ich die Verfügbarkeit der guten analytischen Ausstattung.

Dr. Ulla Gerling-Driessen, Dr. Allison Berger und Dr. Johann Moschner möchte ich dafür danken, dass Sie mit mir ihre berufliche und private Lebenserfahrung geteilt haben. Ihnen und M.Sc. Anne-Katrin Stegemann, M.Sc. Maiko Schulze, M.Sc. Damian Klemczak und besonders M.Sc. Susanne Huhmann danke ich für aufmerksames Korrekturlesen dieser Arbeit und die vielen anderen Wege der Unterstützung, die sie für mich gegangen sind. Diesen besonderen Menschen sowie allen jetzigen und bisherigen Mitgliedern der Arbeitsgruppe Koksch, mit denen ich zusammen arbeiten durfte, danke ich für die inspirierende Zeit der Zusammenarbeit, für abwechslungsreiche Kuchenrunden und viele schöne Erinnerungen.

Meinen Studenten, Hans, Mary-Ann, Tobias, Christian, Phu, Maiko, Anne, Damian, Ullrike, Merlin, Michelle, Tinki, Ayse, Alex und Dorian danke ich für die Unterstützung bei meinen Projekten und auch für die vielen Dinge, die ich bei Ihrer Betreuung gelernt habe. Ich hoffe, Ihnen mindestens ebenso viel mitgegeben zu haben, wie sie mir.

Meinem wunderbaren Lebensgefährten Janis, meinen treuen Freunden und meinen zwei Familien Danke ich für Ihren Rückhalt, Ihr Verständnis, Ihren Zuspruch und den Sonnenschein, den sie jederzeit in mein Leben bringen.

Table of Contens

Abbreviations

Abz	aminobenzoic acid
Abu	(*S*)-2-aminobutanoic acid
ACN	acetonitrile
Ac_2O	acetic anhydride
AD	Alzheimer's disease
AFM	atomic force microscopy
AND	8-anilino-1-napththalene-1-sulfonate
APP	amyloid precursor protein
ATP	adenosine triphosphate
Bhc	6-bromo-4-hydroxymethyl-7-hydroxycoumarin
BCN	bicyclo[6.1.0]non-4-yn-9-ylmethanol
Bis-ANS	bis-4,4'-dianilino-1,1'binaphthyl-5,5'-disulfonate
cAMP	cyclic adenosine monophosphate
CD	circular dichroism
c	concentration
conc.	concentrated
coumarin	1-benzopyran-2-one
Cy5	Cyanine 5
d	path length (of a cuvette)
d	day
DBU	1,8-diazabicyclo[5,4,0]undec-7-en
DCM	dichloromethane
DfeGly	(*S*)-2-amino-4,4-difluorobutanoic acid
DIC	*N,N'*-diisocarbodiimide
DIPEA	*N,N*-diisopropylethylamine
DMF	*N,N*-dimethylformamide
DMSO	dimethylsulfoxid
DSC	*N,N'*-succinimidyl carbonate
EDC-HCl	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
Et_2O	diethyl ether
EtOAc	ethyl acetate
ESI Tof	*electronspray ionization time of flight*
ε	extinction coefficient
E_λ	extinction at certain wavelength
FF	diephenylalanine peptide
Fmoc-Ahx-OH	6-(Fmoc-amino)-hexanoic acid
Fmoc-Lys(Mtt)-OH	*N*-α-Fmoc-N-ε-4-methyltrityl-L-lysine
GdmHCl	guianidinium hydrochloride
h	hour/s
HATU	*O*-(7-azabenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium-hexafluorophosphate
HFIP	1,1,1,3,3,3-hexafluoro-2-propanol
HfLeu	5,5,5,5',5',5'-hexafluoroleucine
hIAPP	human islet amyloid polypeptide

HOAt	1-hydroxy-7-azabenzotriazole
HOBt	1-hydroxybenzotriazole
HPLC	*high performance liquid chromatography*
HPMA	N-(2-hydroxypropyl) methacrylamide
I_0	incident light
I_1	transmitted ligth
IAPP	islet amyloid polypeptide, amylin or diabetes-associated peptide
NCL	native chemical ligation
NHS	N-succinimidyl ester
MS	*mass spectrometry*
Mtt	4-methyltrityl
iPr_3SiH / TIPS	triisopropylsilane
PKA	cAMP-dependent protein kinase A
PTA	phosphotungstic acid
θ	quantum yield
ROS	reactive oxygen species
RP	reversed phase
rt	room temperature
SPAAC	strain promoted azide alkyne cycloaddition
TAMRA	5(6)-carboxytetramethylrhodamine
T2D	diabetes mellitus type two
TEA	triethylamine
TEM	*transmission electron microscopy*
TfeGly	(S)-2-amino-4,4,4-trifluorobutanoic acid
ThT	thioflavine T
TIPS	triisopropylsilane
TLC	thin layer chromatography
TFA	2,2,2,-trifluoroacetic acid
Trt	triphenylmethyl-
λ	wavelength
WHO	World Health Organization

Abbreviations of the 20 canonical amino acids are consistent with the biochemical nomenclature proposed by the IUPAC-IUB commission.[1]

Zusammenfassung

Peptide und Proteine stehen im Zusammenhang mit zahlreichen Krankheiten, werden andererseits aber auch als Wirkstoffe oder innovative Biomaterialen hoch geschätzt. Ihre Vielseitigkeit, ihre Spezifität und ihr komplexes Verhalten macht sie zu interessanten Studienobjekten. Eines der bedeutendsten wissenschaftlichen Studienfelder beschäftigt sich beispielsweise mit der Fehlfaltung von Peptiden, die zu sogenannten Amyloidosen führt. Dieses pathologische Phänomen ist eine intrinsische Gemeinsamkeit zahlreicher Krankheiten wie Alzheimer, Parkinson oder Diabetes Typ zwei. Allerdings gibt es auch Beispiele in der Natur, wo der hochgradig stabile und homogene Aufbau der Amyloidfibrillen gezielt genutzt wird. Ein berühmtes Beispiel ist die Spinnenseide, aber auch Hormone des endokrinen Nervensystem werden in Amyloid-artigen Strukturen gelagert. Zahlreiche wissenschaftliche Arbeiten beschäftigen sich damit Amyloide als funktionale Biomaterialien nutzbar zu machen. Alle natürlichen Amyloid-bildenden Peptide teilen jedoch die Gemeinsamkeit, dass sie anspruchsvoll in der Gewinnung und Handhabung sind. Deshalb ist es besonders sinnvoll auf Modellpeptide zurück zu greifen, die gezielt für eine spezifische Fragestellung und für die gewünschte Analytik designt werden können. In dieser Arbeit werden zwei verschiedene Arten von aggregierenden Modellpeptiden vorgestellt. Zum einen wird ein *de novo* designtes Peptid genutzt um den Einfluss enzymatischer Phosphorylierung nachzustellen und gezielt zu steuern. Zum anderen wird die entscheidende Teilsequenz des mit Diabetis Typ zwei in Verbindung gebrachten Peptids IAPP genutzt, um die Struktur von Oligomeren zu Beginn des Aggregationsprozesses und weiterhin den Ablauf der Aggregation und der dabei gebildeten Strukturen zu untersuchen. Auf diese Weise können Aussagen über die bisher kaum experimentell untersuchten Anfänge des pathologischen Verhaltens gewonnen werden. Um zukünftig das durch Laborversuche gewonnene Wissen über Modellpeptide auch *in vivo* anwenden zu können wird ein weiteres *de novo* designtes Modellpeptid durch einen Licht-sensitiven Linker modifiziert. Die optimierte Synthese des auf dem natürlichen Duftstoff Coumarin-basierenden Linkers sowie seine Kupplung an verschiedene Peptide wird beschrieben. Ein großes Problem bei der Verwendung von Peptiden *in vivo* ist die Bioverfügbarkeit, welche durch den Abbau durch Proteasen erheblich reduziert wird. Im vierten Projekt dieser Arbeit wird anhand eines idealisierten Protease Substrat Peptids der möglicherweise stabilisierende Einfluss von fluorierten Seitenketten untersucht. Die hier gezeigten Ergebnisse sind dabei Teil einer Reihe von ähnlichen Studien mit anderen

fluorierten Seitenketten. Das Ziel ist es zu verstehen unter welchen Vorrausetzungen ein fluoriertes Peptid nichtmehr proteolytisch abgebaut wird. Auf diesem Weg wird eine Datenbank mit Wissen generiert, die Zukünftig zum Beispiel bei dem Design von Peptid-basierten Medikamenten helfen soll.

Abstract

Peptide and protein dysfunction are a major aspect for many diseases, but also various peptide based drugs and innovative biomaterials are highly esteemed. Many studies concentrate on peptides due to their versatility, their specificity on the other hand and the complex structural patterns. After cancer, the so called amyloidosis are the second largest research field in the globalized world. The misfolding of functional peptides and proteins is one intrinsic pathologic pattern for many diseases like Alzheimer's and Parkinson's disease or diabetes type two. On the other hand uses nature the highly stable and homogenous morphology of amyloids. One famous example is spider silk, but also hormones in the endocrine system are stored in amyloid-like structures. In addition, a range of studies concentrate on applications which are based on functional amyloids. Nevertheless, the collective disadvantages of natural amyloid prone peptides are the difficult isolation and handling. Hence, model peptides with defined design can be used to address certain questions and analytical approaches. In this manuscript presented are two different kinds of amyloid-forming model peptides. The first one is a *de novo* designed model peptide with a recognition motif for enzymatic phosphorylation. The goal is to understand the impact of the phosphate group and control the phosphorylation process. The second peptide is the critical hydrophobic sequence part of the diabetes type two derived peptide IAPP. With the help of the short peptide sequence, the structure of oligomers during the initial growing phase and the following processes of amyloid formation can be investigated. Thus, insights into so far sparsely examined beginnings of the pathologic event can be described. To transfer the knowledge on model peptides for *in vivo* studies, a light sensitive caging linker was designed. The optimized synthesis method for the, on natural flavour-based, coumarin linker and the caging process of different model peptides is described in detail. One major problem for peptide applications *in vivo* is the low bioavailability due to proteolytic digestion. Thus, aim of the fourth project is an idealized protease substrate peptide to study the impact of fluorinated side chains on proteolytic stability. Here presented results are part of a series of comparable studies regarding different

kinds of side chain fluorination. Objective is to understand which requirements have to be addressed to inhibit proteolytic digestion of a peptide. As a result, a data bank with this knowledge will be generated as a toolbox for peptide design. One example is the design of metabolic resistant peptide drugs.

Introduction

1.   Designing Peptides - Advantages Compared to Native Proteins

Native proteins and peptides perform a wide range of functions in the cell and in all type of living organisms including signaling[2-3], transduction of energy[4-5], transport of ions[6-7], stability of membranes[8-9], as hormones[10-12] and many more[13-15]. The diversity of proteins in nature is mirrored in the modern material and medical sciences, where the high biological specificity and affinity as well as the multifunctional nature of proteins are beneficial. Despite their importance, knowledge about the structural properties and environmental influences are, for most native proteins, relatively sparse. Furthermore, proteins are limited due to their individual natural functions and in terms of their solubility, synthetic availability, practicality, cost effectiveness, immuneogenecity, regulatory considerations, and others.[16] Designed peptides are able to combine the attributes of proteins, especially their recognition specificity, with being highly chemically well defined, accessible, synthetically feasible, simple and thus cost effective. That is why, for example, recent advances in designing peptide ligands for therapeutic targets are making designed peptides an attractive alternative to small molecules and proteins.[17] Small peptides can, for example, address tasks in various scientific investigations by explicitly modeling one structural motif or functional behavior and carrying a label to enable monitoring of the progress of its action. Thus, also unnatural amino acids can be incorporated to enhance the functionality or stability of given peptide.[18-21] Also designed peptides with affinities comparable to antibodies[22] or DNA[23-24] and with specificities much better than small molecules or antibodies, have been developed. In addition, numerous peptide drugs[25-26] like tumor targeting cytotoxic drugs[27] or targeted diagnostics[28], where peptides can be used as a delivery vehicle for drugs[29] or diagnostics[30], advance medical applications.

Small peptides can also be used to understand the characteristics of natural proteins and their functional or pathological behaviors. Thus, designing peptides can break down the more complex topic into one-parameter models. By this means, Holt *et al*. demonstrate how well-characterized model peptides can be used to determine the orientation and tilt angles of transmembrane peptides in lipid bilayers and they were able to illustrate possible consequences of hydrophobic mismatches in molecular detail.[31] Cyclic peptides have been studied to understand the racemization process during solid phase peptide synthesis [SPPS][32] and several research groups

have made progress towards the '*de novo*' design of specific metal sites to understand the structural, electronic, and catalytic properties of metalloproteins[33], learning about the cumulative influences of proteins. Model peptides have been also used to mimic the catalytic centers of enzymes[34-35], to investigate the specifics of substrates for native enzymes like discussed in part D[19, 36] or to modulate self-assembling supramolecular structures like nanotubes[37], nano-helices[38] or tissue-like scaffolds[39]. Beside these examples, one source of utility for designed peptides is the imitation of conformational changes. Particularly self-assembling and aggregation processes are difficult to examine with native proteins, because they are often poorly synthetically accessible and, due to their low solubility, also difficult to manipulate. Additionally, the aggregation process *in vivo* is followed by a cascade of events due to the sometimes toxic form of aggregates. Models enable the simulation of a defined process, in a predictable time frame, under controllable conditions and are easy to monitor with common lab techniques. Peptide models can be, for instance, features with key sequences that are found in natural amyloid forming peptides, in order to show specific properties and characteristics.[40] One example is the NFGAIL sequence, discussed in more detail in part B, which is the shortest amyloid forming part of diabetes mellitus type two [T2D] relevant human islet amyloid polypeptide [hIAPP].[41] Also the insulin derived VEALYL[42-44] and the central hydrophobic amino acid sequence LVFF of Alzheimer's amyloid peptide Aβ[45-46] are other examples which also form amyloid structures.

A second approach is the use of rational or *de novo* designed peptides.[47-50] *De novo* protein structure prediction methods attempt to create tertiary structures from sequences based on general principles that govern protein folding energetics and/or statistical tendencies of conformational features that native structures acquire. Therefore databases like the protein data bank (PDB)[51] or the protein knowledge base (UniProt)[52] and algorithms like TANGO[53-55] with the collected knowledge about interactions and parameters are used. *De novo* design can be used to design a variety of structures like amphiphilic peptides[56-57], dimers[58], bundle systems as collagen-like frameworks[59] or hydrogelating peptides[60-61]. For example, Koksch *et al*. designed coiled-coil-based model peptides that are able to undergo structural transitions. They combine structural characteristics of helical coiled-coil, random coil and β-sheets with the propensity to form amyloids.[62-68] The knowledge about these model systems was later used for a broad range of studies about the influence of metal ions[68], pH[62-63], the interaction with and on nanoparticles[69],

catalytic activity[34], posttranslational modifications[70-72], the incorporation of unnatural amino acids[18, 73-74] and was further used for the projects discussed in part A and C of this thesis. All described coiled-coil peptide models consist of at least two helices which are wrapped around each other with a helical twist. The sequence reiterates a periodicity of seven amino acids, the so-called heptad repeat (**a/b/c/d/e/f/g**) as shown in **Figure 1**.
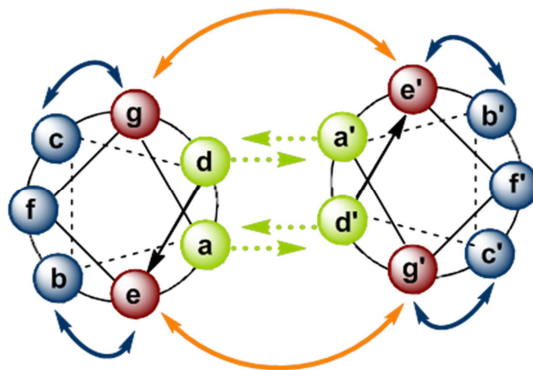
Figure 1: Helical wheel projection of a dimeric coiled-coil based model peptide with arrows to symbolize the intermolecular (orange) and intramolecular (blue) electrostatic interactions and the hydrophobic interactions which direct into the steric zipper motif (green).

One driving force for assembly into coiled-coils is hydrophobic core packing, when hydrophobic amino acids like leucine, isoleucine or valine in positions **a** and **d** interact to form the steric zipper motif. Using larger residues like phenylalanine in these positions can direct the formation of coiled-coils of even higher oligomeric states. The second driving force is the synergy of intermolecular-, electrostatic interactions induced by charged amino acids like lysine, arginine or glutamic acid in positions **e** and **g**. Rational positioning of positively and negatively charged amino acids in **e** and **g** positions can direct the coiled-coil into either parallel or antiparallel orientations or can enable or disable the assembly of heterodimeric coiled-coils.[73, 75] The positions **b**, **c** and **f** are oriented towards the solvent and are thus crucial for solubility. For buffered media serine is favorable, but also charged amino acids which can stabilize the helix *via* intramolecular interactions between amino acids in positions **e** and **g**. To induce β-sheet propensities and the ability to form amyloid fibrils, three residues in solvent exposed positions **b**, **c** and **f** can be substituted with hydrophobic valine side chains.[58, 63]

2. Amyloids

2.1 Amyloid Peptides in Diseases

The complex tertiary structure of native proteins under physiological conditions can be influenced by many environmental conditions[60-61, 75, 78-88], thus a conformational transition to insoluble β-sheet rich amyloid fibrils, sometimes leading to a loss of function[76-79], could take place. In 1853 Virchow coined the term amyloid for extracellular plaques which assemble around cells and disrupt the healthy function of tissues in the humans brain, liver or kidney.[80] The ability to form amyloid fibrils is a common feature of several natural proteins.[76] Many neurodegenerative diseases such as Huntington's disease[81-82] and also T2D[83-85], transmissible spongiform encephalopathies[86-87], hemolytic anemia[88], cystic fibrosis[89], amyotrophic lateral sclerosis[90], secondary amyloidosis[91], dialysis-related amyloidosis[92-93] and more than ten other rare diseases are associated with the conformational transition of a peptide or protein from its native, functional form into insoluble amyloid deposits[76, 94]. In addition to these mostly rare diseases, Alzheimer's disease [AD] is the most common form of amyloidosis. Virtually unknown to the general public four decades ago, AD has risen in prevalence to an estimated 40 million patients worldwide.[95] Since early 1980, when molecular studies of AD did begin, thousands of scientists and healthcare professionals have delved into all aspects of this complex, multifactorial syndrome, hoping to help patients now and shelter others in the future.[95] Selkoe and Hardy reviewed in detail 2016 the findings of the last 25 years of AD amyloid hypothesis and thereby also the fact that numerous clinical trials of antiamyloid agents have not met their prespecified endpoints. They also state that several of these agents are characterized by inadequate preclinical data, poor brain penetration, little human biomarker change, and/or low therapeutic indexes. Therefore, they see herein the need to proceed with more detailed engineering.[95] The pathological behavior of AD is typically defined by the interaction of two proteins, the abnormally hyperphosphorylated tau protein and the amyloid precursor protein (APP).[96-97] The former tends to form intracellular neurofibrillary tangles in neurons[98] and the aggregated amyloid form of the transmembrane protein APP assembles in extracellular plaques[99]. The predominant component of the extracellular plaques and subject of the many studies regarding amyloids is Aβ(1-42) which is generated when APP is digested by the proteases β- and γ-secretase.[100-101] Mutations in the genes encoding either natural APP, which is the substrate for the enzyme

presinilin, or presinilin itself, are sufficient to cause the disease.[102] The early phase of amyloid growth of Aβ(1-42) and other proteins is characterized by different sized and charged β-sheet rich oligomers which later assemble into larger protofilaments.[44, 103-106] Currently, these early state oligomers are presumed to be the cellular toxic species, due to three different neurodegenerative cascades. First is the generation of reactive oxygen species, which directly react with receptors like the ones for tau for example.[107] The second is that early state oligomers form amyloid pore channels and thereby cause an increased concentration of free intracellular Ca^{2+} that leads to neuronal cell death.[102-105] The last hypothesis deals with nonspecific membrane bilayer perturbation.[108] Also the protofilaments of hyperphosphorylated tau are discussed to be more toxic than the resulting fibrils.[109-110] Benilova and De Strooper reviewed the different theories of toxicity of Alzheimer's relevant amyloid species and their highly complex interplay.[108] Nevertheless, the formed amyloid plaques are the major characteristic and found in all necrotic tissues in amyloidosis and most diagnoic techniques concentrate on the resulting fibrils.[111]

One special case of amyloids are the so-called prions. Prions are infectious proteins that can transmit biological information by propagating protein misfolding and induce aggregation of other proteins, which are usually not prone to amyloid formation. Such infection thereby causes protein misfolding diseases in unexpected tissues.[112] In conclusion, the proteins involved in the described diseases do not share sequence or structural identity, but all of them can adopt at least two different conformations without mutations of their amino acid sequence.

2.2   Forms of Amyloid Structures and Fibril Growth Mechanism

Amyloid formation, however, is not a specific feature of disease related proteins, but instead appears to be a general property of peptides and proteins.[113-114] Traditionally amyloids are defined by binding to the dye Congo red[115-116], and this has since been extended to include several more dyes like Thioflavin T [ThT][117-118], NIAD-4[119], Nile red[120], 8-anilino-1napththalenesulfonate [ANS][121] or bis-4,4'-dianilino-1,1'binaphthyl-5,5'-disulfonate [Bis-ANS][121]. The early state oligomers during the growing phase can be bound by the antibody A11, which exclusively binds to amyloid precursor oligomers.[122]

The soluble precursor proteins differ widely in their size, primary and secondary structure, but the insoluble amyloid fibrils share similar highly organized morphology.[78, 123-125] They all appear straight and unbranched with a length in the range of micrometers and a diameter between 8-25 nm depending on the size of the protein and the number of protofilaments.[126] Circular dichroism [CD] spectroscopy reveals a high β-sheet content of the fibrils.[127] X-ray fiber diffraction shows a characteristic anisotropic cross-β-diffraction pattern with two diffuse reflections in which β-strands run perpendicular to the fiber axis. One reflection arises at a spacing of 4.8 Å along the vertical axis from the stacks of extended β-strands and the second at a spacing of ~10 Å along the horizontal axis from the separation of adjacent β-sheets (**Figure 2**, right).[9,11-13] The sheets are held together by N-H⋯O=C hydrogen bonds between the backbone amide groups parallel to the fibril axis.[128] The β-sheet strands can be oriented in pairs parallel or antiparallel to each other, while the amino acid residues are exposed in to alternating fashion along the sides of the sheets. Interactions between the side chains of each pair form a steric zipper motif.[129] Nevertheless, protein monomers with amyloid tendency do not automatically aggregate even when the critical concentration for aggregation is exceeded. The initial nucleation event is thermodynamically unfavored, thus it is the rate-limiting step of the aggregation process.[130] The high-energy barrier of the nucleus formation can be explained by the requirements for a first steric zipper in which at least three, sometimes four molecules are needed to stabilize the β-sheet pairing. For side chain interdigitation between two interacting β-sheets at least one more molecule from the second sheet is required to form the steric zipper.
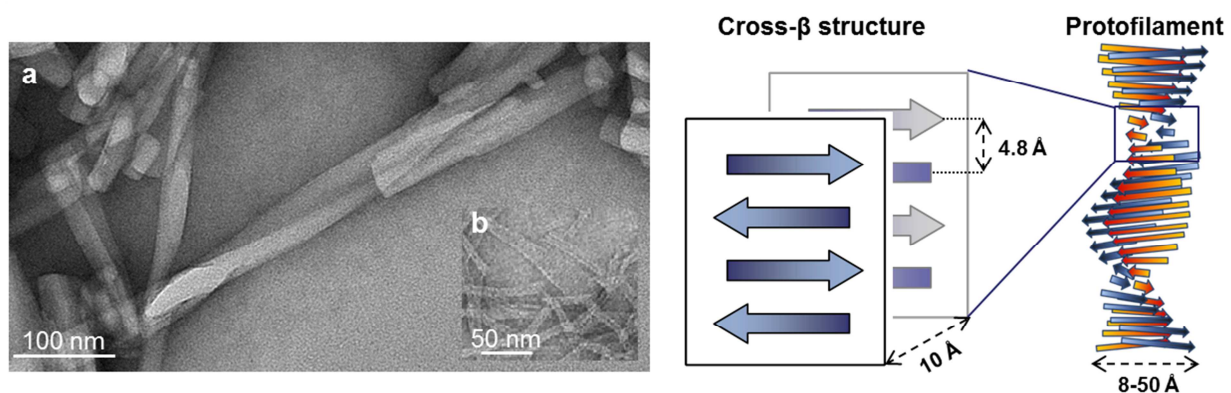
Figure 2: *TEM image of hIAPP related NFGAIL amyloid fibril with a characteristic twist (a), funnel-like pattern of KFM6[70] (b) and a cartoon showing the cross-β structure of amyloid fibrils (right).*

For this first nucleus or seed to form, the three or four participating protein molecules must all be spatially close and have undergone a conformational change to expose their zipper-forming stretch. Once the nucleus is formed, growth of the strand proceeds linearly from both ends, presumably one monomer at a time.[131-133] The elongated strands are called protofilaments. Morphological studies using transmission electron microscopy [TEM] or atomic force microscopy [AFM] show that amyloid fibrils commonly consist of a small number of protofilaments with typical diameters in the range of 20 to 50 Å (**Figure 2**, left).[9,14] Two or more protofilaments twist around each other and are held together by van der Waals and hydrophobic interactions in the amyloid fibril. Data analysis of 77 published amyloid crystal structures resulted in the calculation of an average of 50,000 layers (parallel and perpendicular to the fibril axis) of β-strands in one fiber.[134] Each type of amyloid fibril is a regular aggregate of one single protein or peptide, thus seeding the growth from monomers into fibrils only works with seeds from the same protein or at least one with a similar sequence.[135] Nevertheless, it is possible to receive several different cross-β structured fibrils with the same protein sequence in one experimental batch. This effect is called polymorphism and can be explained by varying pairing between β-sheet regions (homo- or hetero-steric zippers) of the sequence and different folding patterns of the protein segments, which are not part of the β-sheet stacking.[136]
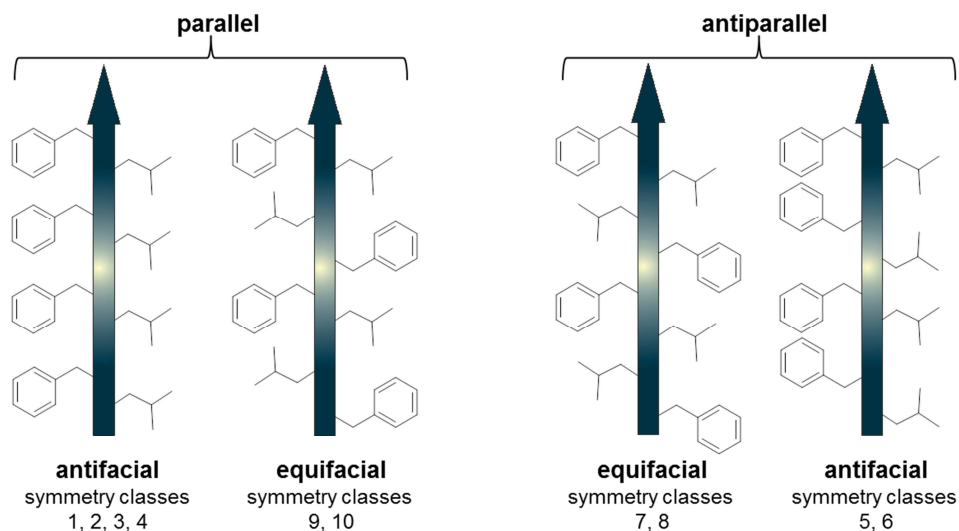
Figure 3: *Schematic representation of the four basic β-sheet symmetry patterns. The protofilament strand is indicated with a blue arrow that symbolizes the direction of growth. Parallel sheets can be either antifacial or equifacial, as can be the antiparallel sheets.[129]*

The two most elementary types of amyloid structures are the assembly of strands in either a single sheet or pairs of sheets which interact to form a spine. These two basic structure elements can be further subdivided into four β-sheet symmetry patterns depicted in **Figure 3** and ten symmetry classes of higher order for cross-β spines, which are depicted in the cartoon in **Figure 4**.[137-138] The four subpattern arise on the one hand from a parallel (C-terminus next to C-terminus) or antiparallel (C-terminus next to N-terminus) orientation of adjacent sheets and on the other hand because of the chirality of peptides which results in two faces of a single peptide strand, also symbolized with two different colored arrows in **Figure 4**. Thus the sheets can also align antifacial and equifacial. Antifacial sheets are defined when strands align in such a way as to yield a face composed of even-numbered residues and opposite face composed of odd-numbered residues. Conversely, equifacial sheets orient both faces with odd- and even-numbered residues.[129] Only the parallel-equifacial pattern was not yet observed by crystallographic measurements.[139] The higher order symmetry shown in **Figure 4** can be used to explain the characteristic morphology of the fibrils as visualized in TEM, solution-based NMR or X-ray crystallography[140-141].
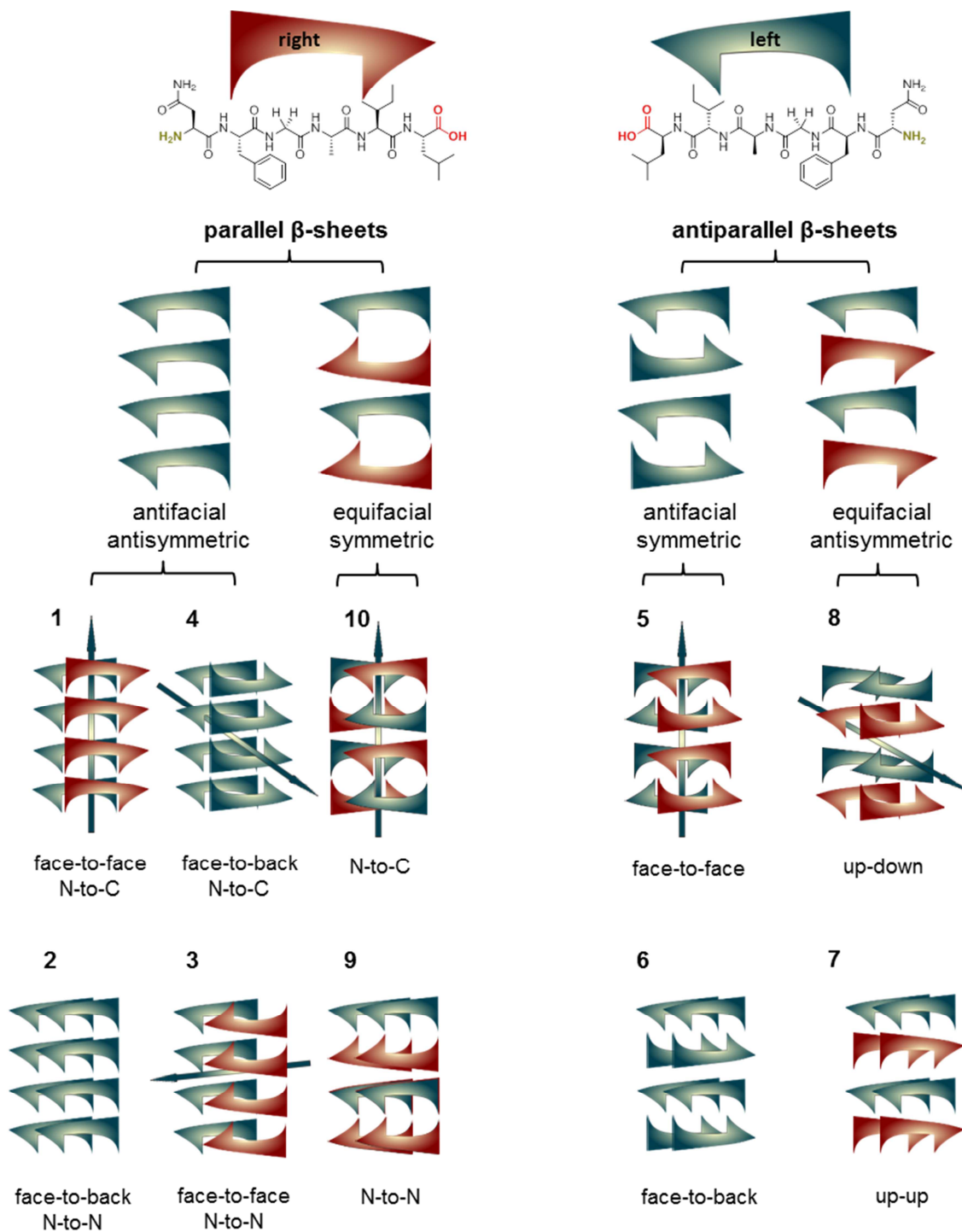
Figure 4: *The ten amyloid symmetry classes of homo-steric zipper amyloid spines, illustrated with arrows. At the top is shown that a β-strand has an N-terminus, a C-terminus, a front face (blue arrows) and a back face (red arrows). The thin blue arrows signify 2_1 symmetry axes, meaning that the two sheets are related by 180° rotation about the axes and a translation along the axes of one-half the distance between the arrows. Cartoon is adapted with permission from Eisenberg and Sawaya.[129]*

If hydrophilic and hydrophobic residues alternate in the peptide sequence, antifacial sheets may display amphipathic character and could as pairs lead to protofilaments through their hydrophobic residues on the inside and with polar surfaces on the outside of the protofilament (class 1, 3 or 5). Equifacial sheets must be symmetrically identical and therefore they cannot be amphiphatic (classes 7-10). Those fibrils usually assemble with no discrete termination point, sometimes with variable ribbon- or funnel-like dimensions as was shown for the model peptide in part A of this thesis[70] (**Figure 2b**). Similarly, spines with face-to-back symmetry may pack with no discrete termination point (classes 2, 4 or 6).[129]

2.3   Amyloid Structures as Scaffolds and Templates for Biomaterials

Recently amyloids also gained attention in the field of bio- and nanotechnological applications since these structures show outstanding mechanical stability and remarkably regular fibrous architecture, which makes them ideal starting points for functional and responsive biomaterials.[142-144] Also their resistance to dissolution under conditions of 5% sodium dodecyl sulfate or 4 M urea, which brake apart most peptides structures, or reagents like 1 M acetic acid, glutamine or ammonium hydroxide are not affecting the amyloids. Only some solvents like hexafluoroisopropanol [HFIP][145], TFA, NaOH, pure formic acid, or 4 M guanidinium hydrochloride can dissolve the fibrils.[146] Their unusual stability results from the dual character of the cross-β pattern of the protofilaments. Compared to single β-sheets only stabilized by backbone hydrogen bonds, the cross-β pattern is additionally stabilized by van der Waals and hydrophobic interactions resulting from the tight packing of the steric zipper interfaces. Furthermore, the amides, running along the strands of the steric zipper, are mutually polarized and have a dipole moment oriented in the same direction, resulting in favorable multipolar stabilization. In addition to both factors, supports the calculated averaged hydrogen bond energy of -9.1 kcal/mol-per-hydrogen-bond the resistance of amyloids.[147] Furthermore, the release of water during the tight packing process generates an increase in entropy and also the side chains can form ladders with high binding affinities in dependence of the sequence. Systematic scanning approaches combined with low resolution techniques have been used to characterize certain functions of residues[148] or secondary structures[58, 149-150] within the sequences of amyloidogenic peptides. All the described factors can be used to design nano-materials[151], 3D

hydrogel scaffolds[152] or in drug design with finely-tuned mechanical properties and assembly architectures. Gazit *et al*. describes for example, the smallest aggregating peptide, a dipeptide composed of two phenylalanines [FF], that self-assembles into regular elongated nano-tubes.[153] By adding additional phenylalanine, plates, spheres and two-dimensional structures can be formed[154] and with only minor modifications the FF hydrogel has also been shown to support cell growth, and it can be used to release small molecules in a time-controlled manner and it has been part of a electrochemical biosensor platform to detect enzyme activity.[155] Two more complex complementary, 11 residue β-sheet prone peptides, attached to an N-(2-hydroxypropyl) methacrylamide [HPMA] copolymer, formed a hydrogel with a pattern perfect for the long term viability and proliferation of cells, thus it was used as a scaffold for bone tissue engineering.[156] In 2017 Nystrom *et al*. used amyloid fibrils generated from freeze-dried and cross-linked β-lactoglobulin as a cell growth scaffold for epithelial cell lines and they could show how to control the pore structure and elastic modulus of the aerogels by simple concentration and freezing gradients. They further proposed to expand the knowledge for wet-state applications such as heterogeneous catalysis or purification membranes.[157] At the same time Al-Garawi *et al*. presented an amyloid fibril that can bind Zn^{2+} ions *via* its three-dimensional, histidine-rich structure and mimic carbonic anhydrase with high esterase activity.[158] Amyloid-like bovine serum albumin was used to prepare nanofibers for controlled drug release and the release of antibodies against *Escherichia coli* and *Staphylococcus aureus*.[159] Rengifo *et al*. reviewed work on amyloid scaffolds which are able to organize light harvesting chromophores and suggest chlorosomes antenna array-based self-assembling architectures to break peptide bilayer symmetry for directional energy and electron transfer, and to incorporate redox active metal ions at high densities for energy storage.[39]

It is important to note, that natural amyloids can also have nonpathological, biological functions. Fowler *et al*. demonstrated for example that the amyloid protein Pmel17 is involved in mammalian skin pigmentation[160] and Maji et. al. showed that some protein hormones in secretory granules of the endocrine system are stored in an amyloid-like cross–β-sheet-rich conformation.[161] A well-known example is the amyloid-like spider silk peptide fibroin, which was modified to release bioactive materials and to operate as a drug delivery system.[162] It was also described that flexible amyloids are an important part of the surface of *Escherichia coli*.[163]

Furthermore, some domains of RNA-binding proteins in RNA-rich stress granules of higher cells does form amyloid-like structures.[164] In the brain, where amyloids are also linked to many diseases, exists a connection to the formation of long term memories, as was described by Si *et al.*.[165]

2.4   Relevance of IAPP derived NFGAIL

In 95% of T2D patients the islets of Langerhans in the pancreas are damaged by amyloid plaques, the so-called islet amyloids.[166-169] T2D is a heterogeneous and multifactorial disorder with several complications and a reduced life expectancy. 422 million diabetic people in all countries of the world were tallied by the World Health Organization [WHO] in 2014 and 3.4 million died in 2004, with increasing mortality rates since then.[170-171] More than 80% of these deaths occurred in low- and middle-income countries.[172] In 1986, the disease was for the first time found to be connected to the islet amyloid polypeptide [IAPP] as an aggregation product of the β-cell regulatory system (**Figure 5**).
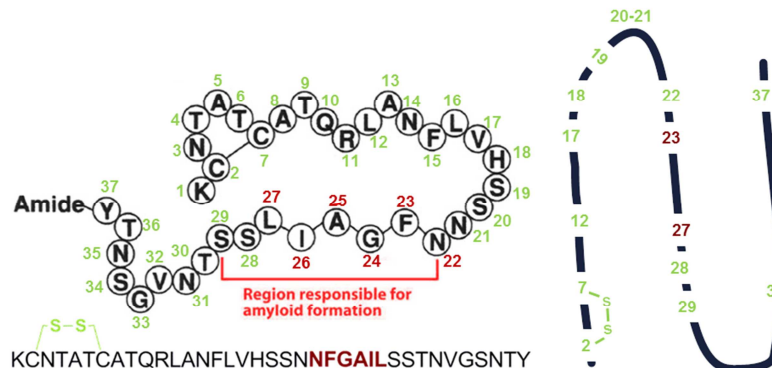
Figure 5: *Schematic representation of the β-sheet and β-turn regions of soluble, monomeric hIAPP as a result of combined calculations proposed from Pillay and Govender[173], derived from three different models[174-176] (right). Primary structure of human IAPP [hIAPP] with disulfide bridge between Cys2 and Cys7 in green and the NFGAIL part highlighted in red (left).*

IAPP, which is also called amylin or diabetes-associated peptide, is processed in the pancreatic β-cells and is secreted together with insulin in a 20-to-1 molar ratio of insulin to IAPP.[170] It is released as an 89 amino acid long precursor protein, referred to as preProIAPP, and is cleaved at

the N-terminus yielding ProIAPP. ProIAPP is subsequently posttranslationally processed by the prohormone convertase to IAPP.[177] IAPP is known to work antagonistically to insulin by inhibiting glycogenesis and promoting glycolysis.[178-179] Further studies suggested that IAPP suppresses the release of glucagon from pancreatic β-cells and thus prevents the release of glucose from the liver.[180-181] In its native, functional state IAPP is a soluble 37 residue monomer with a disulfide bridge between Cys2 and Cys7[182], but it was also found to be the major protein component in islet amyloid plaques.[183] Several studies regarding the monomeric structure of IAPP led to the model presented in **Figure 5**, where the regions between 12-17, 23-27 and 32-37 are prone to β-sheet structures and the region 20-21 was identified as β-turn. The residues 18, 19, 22 and 28-31 could participate in forming either β-sheets or -turns.[173]

hIAPP amyloid formation was described to occur in a stepwise manner, starting with almost the same nucleation mechanism as Alzheimer's peptide Aβ[83], where the monomers assemble, in an on-pathway manner, into oligomeric structures starting with tetramers[184] or even higher oligomer numbers.[185] In the presence of copper ions also globular and toxic off-pathway hIAPP oligomers can be formed.[184-185] Protofilaments with width of 5-8 nm can grow from both ends at a rate of approximately 1.1 nm per minute.[186] Higher order fibrils are formed when two protofilaments coil about each other with a 25 nm axial cross-over repeat, containing 2.6 hIAPP molecules per 1 nm length of the fibril.[187] Three dominant theories about the toxicity of IAPP in pancreatic β-cells are discussed in literature. All of them share the common hypothesis, that oligomers present at early stages of the aggregation process are more toxic than the full-length fibrils. The first theory involves the incorporation of cell membrane lipids into the growing fibrils, and states that the resulting ion voltage-dependent channels within the membrane, leads to a nonphysiological exchange of Ca^{2+}, Na^+ or K^+ ions as the main factor for cell mortality.[188-190] The second theory states that hIAPP induces DNA fragmentation and thus increases the expression of p21 and p53 tumor suppressor genes. Both genes encode proteins that arrest cell proliferation and lead to apoptosis.[173, 191] The third theory deals with reactive oxygen species [ROS] and is in agreement with models for amyloid toxicity in other amyloidoses. On the pathway to amyloid fibrils, the oligomers can bind ions like copper or iron by a mixture of nitrogen and oxygen donor complexation of mostly histidine and thereby generate an imbalance of ROS in the cell which

leads to oxidative stress.[192-194] In this way, the destruction of the pancreatic β-cells results in a decrease in insulin production and thereby manifests as T2D.[195]

As a region with high amyloid propensity the hydrophobic patch region of hIAPP(20-29) SNNFGAILSS was of particular interest, since hIAPP fibril formation is driven by nucleation in this domain[196], while mutations in this region inhibit fibril formation.[169] Several other amyloid forming regions of IAPP have been observed or predicted during the last decades (**Table 1**), but the shortest fragment with high amyloid forming tendency is the NFGAIL, residues 22-27[41, 197].

Table 1: *Observed and predicted amyloid forming regions of hIAPP.*[174-176, 197-206]

amyloidogenic region	predicted or observed	year	reference
20-29	Observed	1990	Westermark et al.
17–34, 24–37, 30–37	Observed	1999	Nilsson et al.
20–29	Observed	2000	Goldsbury et al.
23–27 and 22–27	Observed	2000	Tenidis et al.
22–29	Observed	2001	Azriel and Gazit
8–20	Observed	2001	Jaikaran et al.
14–18, 14–22, 14–20, 15–20, 15–19	Observed	2002	Mazor et al.
22–27	Observed	2002	Scrocchi et al.
12–17, 15–20	Observed	2003	Scrocchi et al.
12–17, 22–27, 31–37	Observed	2005	Kajava et al.
13–18	Predicted	2006	Galzitskaya et al.
8–17, 28–37	Predicted	2007	Luca et al.
12–18, 15–20, 22–28	Predicted	2007	Zhang et al.
8, 13, 17, 25, 27	Predicted	2009	Shim et al.

The isolated NFGAIL segment is known to form aggregates which coil along each other resulting in fibrils, whereas the FGAIL segment forms broad ribbon-like funnels.[200] Tenidis *et al*. have shown that the single fragment NFGAIL has cytotoxicity towards the pancreatic cell line RIN5fm comparable to that of full-length hIAPP.[207] Thus, NFGAIL became a common model for hIAPP aggregation and its characteristics have been extensively studied experimentally[208-210] and theoretically[211-216]. Ionic strength influences NFGAIL assembly[211] and oligomers are formed due to attractive, hydrophobic interactions between phenylalanine residues[210, 213-214, 216]. The Phe23 residue has been found to be crucial for the amyloid formation of NFGAIL and full-length

hIAPP.[217] It probably directs the stacking of β-sheets to form ordered multilayer aggregates during the nucleation process[213]. These findings are complemented by simulations, showing that the side-chain of Phe fits into the concave surface around Gly of the neighboring peptides in a well-ordered cross-β pattern.[218] The amyloid fibrils of NFGAIL were investigated by Nielsen *et al.* conducting solid-state NMR and the structure shows that an antiparallel hetero zipper with interacting Phe and Leu residues between two neighboring molecules exists.[219] The authors further described by X-ray fiber diffraction a 4.65 Å distance between bridging strands and a 9.01 Å distance between ladders of NFGAIL protofilaments.[219] In contrast, Manine *et al.* found a polymorphic structure for hIAPP(20-29), with mixtures of larger populations of parallel and smaller populations of antiparallel strands.[220] The NFGAIL model was further used to simulate the oligomerization process in general. In this way, Wu *et al.* established that only 8% of early oligomers up to octamers are well-ordered and stabilized by β-cross-strand hydrogen bonds and favorable side-chain stacking in a face-to-face pattern with either parallel or antiparallel orientation of the strands. The remaining disordered oligomers are mainly stabilized by nonspecific hydrophobic interactions.[213] Guo *et al.* predicted a clear tendency for an antiparallel and equifacial orientation of β-sheets of hIAPP(22-28) NFGAILS and suggested that the hydrophobicity of the Phe23 residue plays a greater role than the expected π-stacking between phenylalanine residues (**Figure 6**).[221] Other simulations suggest a 67% content of parallel NFGAIL β-sheets during the nucleation process[222] which would be in good agreement with previous models for hIAPP[175]. NFGAIL has also been used to test possible inhibitors for hIAPP aggregation like green tea polyphenols[223] and (-)-epigallocatechin gallate [EGCG][85]. Furthermore, it was demonstrated that seeding with NFGAIL fibrils induces the aggregation of full-length hIAPP.[197] As previously described in section 2.2, the seeded nucleation mechanism only applies to peptides that adopt similar cross-β patterns.[135]
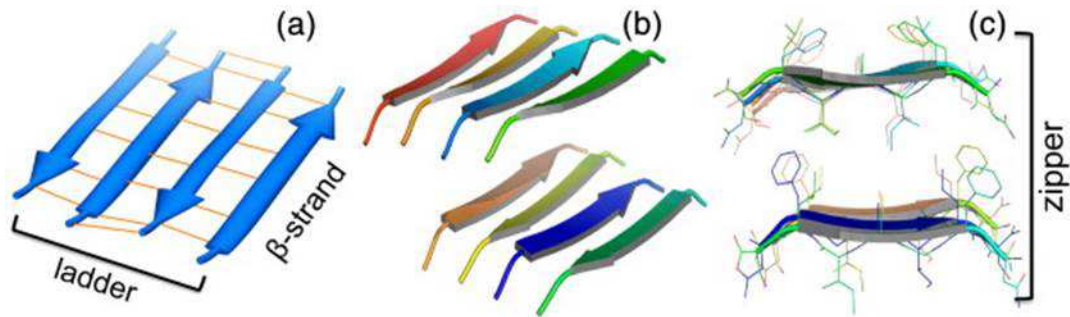
Figure 6: *Schematic representation of the respective hIAPP(22-28) NFGAILS fibril (PDB ID: 2KIB). a) β-strands are arranged by hydrogen bonding (shown as orange lines) into a ladder, which is antiparallel. b) Two ladders combined into an antiparallel steric zipper. Atoms are shown as lines without hydrogen atom. c) Orientation of modeled cross-β-pattern with symmetry class 5. (Figure reproduced with permission from Guo et al.; copyright © Elsevier.)[221]*

2.5 Influence of Phosphorylation

The reversible transfer of phosphate groups is a crucial regulatory mechanism *in vivo*. A popular example is cellular respiration, which is driven by adenosine triphosphate [ATP] as phosphate donor.[224] The enzyme class responsible for the transfer are the so-called phosphotransferases with the two subclasses of kinases and phosphatases with antagonistic function.[225] 1.7% of the human genome encode phosphotransferases, thus this is the second largest class of proteins in humans after collagens.[226] Both kinases and phosphatases, play an important role in cell growth and differentiation[227], for the interaction between proteins[228-229], nucleic acids[230], cofactors[231] and antibodies[232] and they are crucial for almost every process in metabolism[233], biosynthesis[234], gene regulation[235], signal transduction[236] and muscle contraction[237]. Phosphorylation can also modulate enzyme activity, as discovered by the Nobel prize recipients Fisher, Krebs, Sutherland and Wosilait during studies of the break down of glycogen in the liver.[238-239] The first characterized and purified enzyme of this class was the phosphorylase kinase, reported in 1959.[240] Until today, three groups of kinases have been described, serine and/or threonine kinases [PSTK], tyrosine kinases [PTK] or kinases with double specificity, able to phosphorylate all three types of amino acids [DSK].[241] As a consequence of their unique function in nature, imbalances in phosphotransferases commonly result in dysfunction and pathogenesis. Numerous studies have concentrated on such phosphorylation dysfunctions in diseases.[242-245] The pathological disorder of hyperphosphorylated tau protein in Alzheimer's disease is arguably the best described negative effect of phosphorylation.[95-96, 246-248] Furthermore, Seyfried and Shelton

describe that in contrast to normal cells, all cancer cells are dependent on substrate phosphorylation level to meet energy demands and not on oxidative phosphorylation. Thus, they hypothesize cancer could be seen primarily as a metabolic disease with characterized by an overabundance of phosphorylation.[249] It has been further proposed, that an acquired defect in oxidative phosphorylation in mitochondria prevents cells from using molecular oxygen for adenosine triphosphate production, and potentially causes sepsis-induced organ dysfunction.[250] A virus protein of the virus *Rous Sarcoma* is known to have protein kinase activity and thus can induce cancer.[251] This virus protein is almost identical to the cyclic adenosine monophosphate [cAMP]-dependent protein kinase A [PKA] which was used for project A of this thesis.[252] On the other hand, phosphorylation is an important post-translational modification of proteins and a critical quality attribute for protein therapeutics, especially if it is required for protein function or sub-cellular targeting. To quantify phosphorylation rates of amino acid residues on the backbones of proteins, Ketcham *et al.* described in 2017 a new phosphate-specific binding dye.[250] Greengard *et al.* tried to use enzymatic phosphorylation to target APP to prevent amyloidosis and Alzheimer's disease. The authors tested the impact of phosphorylation by protein kinase C, Ca^{2+}/Calmodulin-dependent protein kinase H, cGMP-dependent protein kinase, casein kinase H, insulin receptor protein tyrosine kinase and PKA against several serine or threonine containing segments of APP. The most effective phosphorylation, with a reduction in amyloid propensity of 50-80% under physiological conditions, was realized with Ca^{2+}/Calmodulin-dependent protein kinase H and protein kinase C, all other kinases were less effective.[253] Koksch *et al.* used the 26-residue coiled coil model peptide with amyloid tendency VW18, to investigate the impact of phosphorylation on peptide aggregation under controllable conditions (**Figure 7**). One or two substitutions of serine with phosphoserine were suitable to inhibit peptide aggregation but also decreased the helix propensity of single phosphorylated analogues from 77 to 15%.[72]
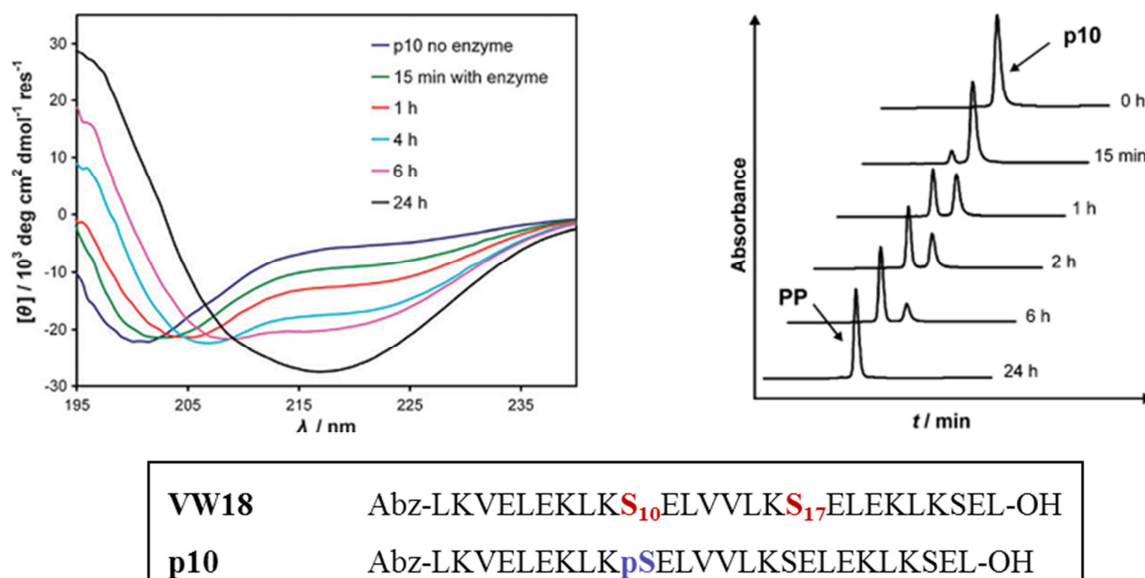
Figure 7: Investigation of the impact of single phosphorylation on the amyloid propensity of a 26-residue model peptide and the kinetics of the enzymatic dephosphorylation induced conformational change from random structure to β-sheet in CD (left) and the dephosphorylation ratio via HPLC (right). Sequences of parent peptide VW18 and monophosphorylated analogue p10 are depicted in the lower panel. Reused with permission, © The Royal Society of Chemistry 2010.[254]

The impact of the phosphate group was described as sterically demanding and the Coulomb repulsion on neighboring amino acid residues was claimed as the main factor hindering the formation of a complex secondary structure pattern.[255] One of the monosubstituted peptides p10 was further used by the Koksch group to demonstrate the induced amyloid formation by enzymatic dephosphorylation with the phosphatase λ-PPase. As depicted in **Figure 7**, the conformational change to β-sheets is directly proportional to the dephosphorylation ratio.[254] Many more studies concentrate on the structural effects, possible applications and in particular the pathological consequences of enzymatic phosphorylation and dephoshorylation.[258-260]

cAMP-dependent Proteinkinase [PKA]

PKA is the best investigated kinase regarding structural and biochemical characteristics. Due to its broad substrate specificity it commonly occurs in nature, also in human organisms. It is a key factor in energy metabolism[256-257], apoptosis[258], gene expression[259], signal transduction[260] between cells[261] and cell growth.[262] The molecular biology of memories is mainly driven by PKA[263] and the important messenger molecule cAMP[264] could not be transmitted without these

enzyme.[265] PKA is a serine/threonine kinase with the capability of transferring phosphate groups to both amino acids.[266] In its inactive form two catalytic subunits are bound to two regulatory units, which are cleaved off when two molecules of cAMP are bound by each regulatory peptide.[267]
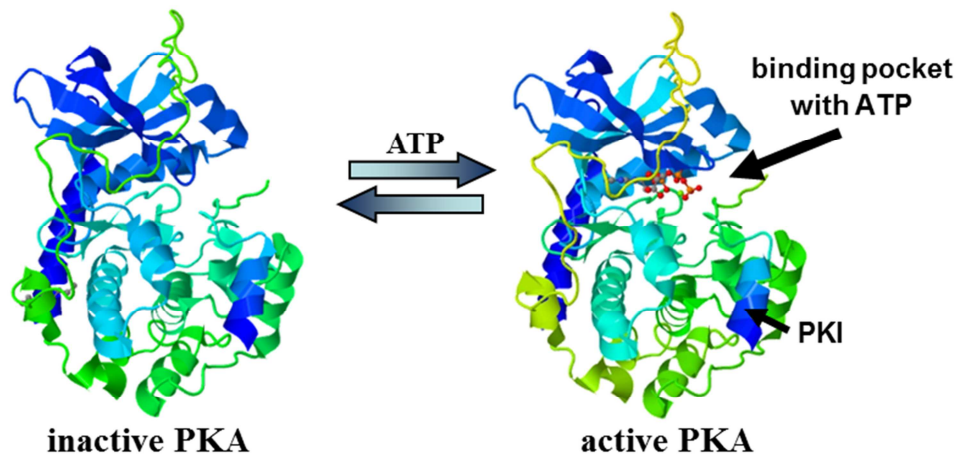
Figure 8: Cartoon image of the PKA activation with ATP. Inactive PKA with protein kinase inhibitor peptide [PKI] (PDB ID: 2GFC)[268] and active form of PKA and MgATP coordinated into the hydrophobic binding pocket (PDB ID: 1Q24)[269].

The hydrophobic binding pocket is one of the specific characteristics of all protein kinases and enables the enzyme the antagonistic interaction with complementary peptides or protein domains (**Figure 8**).[270] The substrate has a recognition motif with up to four, upstream and downstream, amino acids surrounding the phosphorylation site. For high phosphorylation efficiencies the recognition motif is located in solvent exposed and/or flexible regions of the substrate. Krebs and Beavo established in 1979 the shortest, still effective recognition sequence for PKA as **RRXSX**.[271] Three years later the recognition sequence was refined by Zetterquvist and Ragnarsson to increase the enzyme activity. They suggest **RRASL** as most promising substrate recognition site for PKA.[272]

3. Caging Compounds

3.1 Caging Compounds and Functionalities

Caging compounds are light-sensitive molecules that encapsulate and thereby mask functional effector molecules. The uncaging or release and reactivation of the effector molecule can be realized with light irradiation.[273] The advantages of caging in comparison to other orthogonal protecting groups are highly specific release, this is also possible intracellular with rapidly, controlled in time and space as well as quantitatively controlled uncaging of the biomolecule of interest.[274] Caged compounds are used to release effectors at a certain time point to enable an increase of concentration, after longer accumulation, or to start defined kinetics in a challenging surrounding.[275-276] Every kind of effector or messenger from size of proton or Ca^{2+} to proteins can be caged.[277] The first caged biomolecules, cAMP[278] and ATP[279], are still widely used and are available in specialized trade. Today, caged compounds are used in various applications like time-resolved enzyme release[280-282], controlled gene-expression[283-287] or release of signal molecules[288-289], in the brain for example[288]. In medical science caging can be used to produce prodrug conjugates, where the cage masks the drugs activity until the effective concentration at the desired site is accumulated and the drug can be activated *via* optical radiation. Applications with light as trigger enable a high level of pharmacological control.[290-291] Ito *et al.* review UV radiation trigger able 5-fluorouracil-cage conjugates for an anti-cancer pro-drug against solid tumors.[292] In organic chemistry cages are used to induce reactions, like the release of a molecule which can be monitored during the intercalation into a supramolecular host matrix.[293] Epley *et al.* describe a method to slowly release the cell tissue dye Nile Red *via* photo-induced degradation of and cargo release from a nanoscale metal-organic framework.[294]

The light-induced uncaging can be summarized in three states. The reaction starts with the absorption of a photon, thus transferring the light-sensitive caging moiety into an excited state. The next step is the photochemical reaction which results in the release of the effector molecule. The mechanism of this reaction can vary between the caging classes and is also dependent upon the different linkers to the effector (see classes of caging compounds below). During the last step, the caging compound relaxates back to the ground state.[295] If lamps with high energy transfer or lasers are used for the illumination of caged compounds with high quantum yields, the

release can be realized in timescales from pico- to milliseconds.[296] The propensity of a compound to absorb a photon of a certain wavelength [λ] is quantified as extinction coefficient [ε_λ]. The extinction coefficient can be described with the Lambert's Beer law (1) where E_λ is the absorbance of the compound at concentration [c], the path length of the cuvette [d] and the incident [I_0] and transmitted [I_1] light.

$$\varepsilon_\lambda = \frac{E_\lambda}{cd} = \frac{lg\left(\frac{I_0}{I_1}\right)}{cd} \tag{1}$$

A high ε_λ at the illumination wavelength is crucial for a caging compound to realize a good ratio of uncaging per absorbed photons, also quantified as quantum yield [θ]. The efficiency of uncaging can be thereby defined as the product of ε_λ and θ. Thus, a good caging compound is characterized by having both a high ε_λ and a high θ value. In addition, the cage-effector conjugates should be stable and soluble until uncaging. The caging should further completely inhibit the (biological) function of the effector molecule. Moreover, products of the caging compound after uncaging should be non-toxic to the cells or tissue where the reaction is performed and formation of radicals should be avoided.[295] Many biomolecules like peptides absorb light in the range between 240-280 nm. To prevent concurrence in photon absorption or possible damage of the organic tissue or effector molecule, the caging compound should preferentially absorb in the long-wave range upon 300 nm.[273] The covalent bonds between caging compound and effector molecule are related from photosensitive protecting groups in organic chemistry. The fast uncaging process can be followed by a slower degradation step of the remaining linker functionality. One example is the linking of amines with carbamates. After being photoreleased from the caging compound, the carbamate is converted into carbon dioxide and the free amine. A slower release of amine groups can be beneficial for the folding of proteins.[297] Other established linking groups are carboxylic acids, phosphoric esters, alcohol and amine groups, thiols, ketones and aldehydes.[295] The five following classes of caging compounds are frequently described in literature for caging biomolecule effectors (**E**).

o-Alkylated Nitrophenol Cage

The most common caging compounds are the o-nitrobenzyl derivatives. They were originally developed for the use in organic chemistry[298], but with the caging of ATP they became the pioneer caging compound in bioorganic chemistry[279]. The reaction proceeds by flash laser photolysis at a wavelength around 400 nm depending on the functionalization of the R_{1-3} positions.[299] In an attempt to elucidate the underlying mechanism the first two reactions were kinetically investigated at physiological pH. Thus, the uncaging starts with the rate-determining step, a proton release by rapid ionization, followed by the slower formation of the free effector (**E**⁻) and mesomerization of the side product (**Scheme 1**).[300] The advantages are the simple synthetic availability and handling of the compounds. The photochemistry of uncaging of o-alkylated nitrophenol is well understood, thus enabling defined analysis of the effector release. The disadvantages are poor solubility in aqueous solutions, toxicity, significant absorption of the byproducts of uncaging in the range below 300 nm, and low efficiency of the release step.[301-302] In acidic aqueous solutions the reaction is slightly reversible if the effector molecule is not removed from the solution.[295]
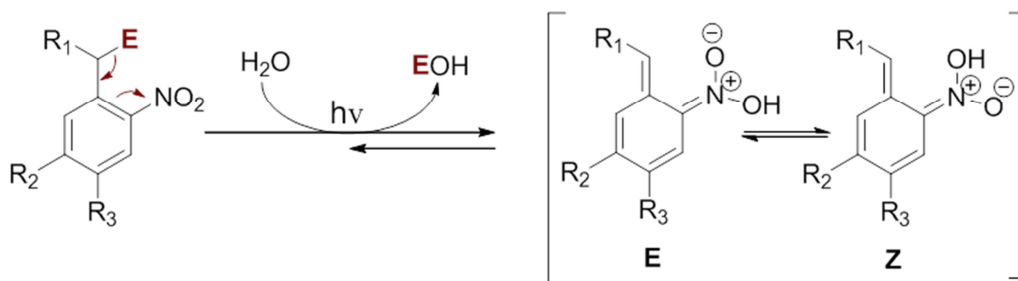
Scheme 1: *Uncaging mechanism of generalized o-alkylated nitrophenol cages. **E** represents the effector molecule.*

p-Hydroxyphenylacyl Cage

Straight forward synthesis and high quantum yields coupled to a high photolytic efficiency make the p-hydroxyphenylacyl compounds a promising alternative to the alkylated nitrobenzyl cages. The cage conjugates are soluble and stable in buffer for at least 24 hours, and the byproducts after uncaging are non-toxic.[303] The byproduct is not chiral and its absorbance is blue shifted and therefore the release is precisely detectable what is a benefit.[304] **Scheme 2** shows the proposed cleaving mechanism from Pelliccioli and Wirz.[295]
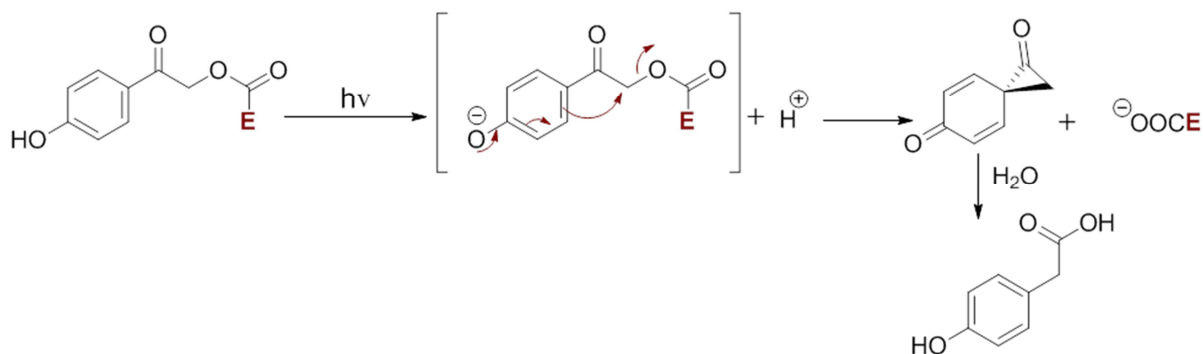
Scheme 2: *Uncaging mechanism of generalized p-hydroxyphenylacyl cages. **E** symbolizes the effector molecule.*

1-Acyl-7-nitroindoline Cage

The proposed mechanism for the uncaging of 1-acyl-7-nitroindoline compounds is characterized by the transfer of an acyl group to the neighboring nitro substituent (**Scheme 3**).[305] The quantum yield and release rate are rather low, but good solubility and no toxicity are advantages of these cages. Excitation occurs in the range of 254 to 365 nm. A carbamate linking group between cage and effector can be used to increase the efficiency of uncaging.[306]
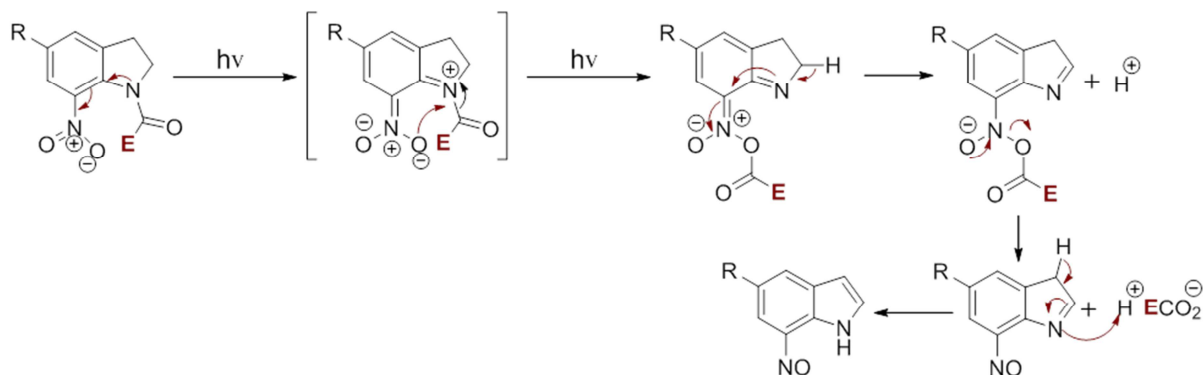
Scheme 3: *Theoretical uncaging mechanism of generalized 1-acyl-7-nitroindoline cages. **E** represents the effector molecule.*

Benzoin Cage

Benzoin cages have high quantum yields, very high efficiencies and uncaging rates (2-7 ns) after excitation in the range of 300 nm.[307] A drawback is their complex synthesis with a chiral product. The two isomers have to be separated which decreases the yield of the synthesis. The uncaging byproduct benzofuran is inert and nontoxic, but absorbs in the same range as benzoin, thus differential detection by UV spectroscopy is not possible. Also, efficiency of photolysis is

23

dependent on the pH in aqueous solutions and is enhanced under acidic pH conditions (**Scheme 4**).[308]
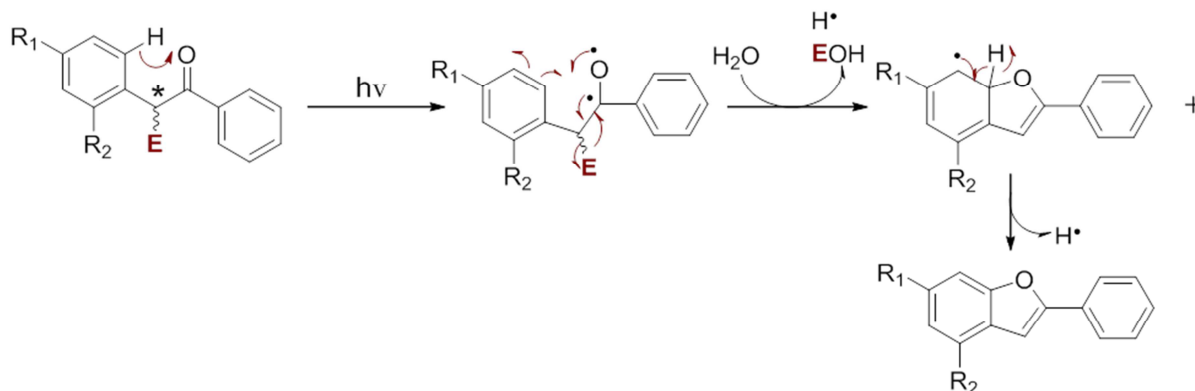
Scheme 4: *Theoretical uncaging mechanism of generalized benzoin cages.* ***E*** *represents the effector molecule.*

Coumarin Cage

1-benzopyran-2-one [coumarin] cages are used in combination with linkers like carbamates (**Scheme 5**) and carbonates, or can be coupled to carboxylic acids and phosphate esters. The syntheses of several coumarin derivatives are described in the literature as feasible procedure.[309-311] The photorelease occurs on a nano- to picosecond timescale[312] and can be monitored (in cells) *via* fluorescence spectroscopy.[313] The byproduct is non-toxic and is generated through an S_N1 mechanism involving heterolysis from the singlet-excited state to form a radical pair intermediate after irradiation with light (**Scheme 5**).[309, 314]
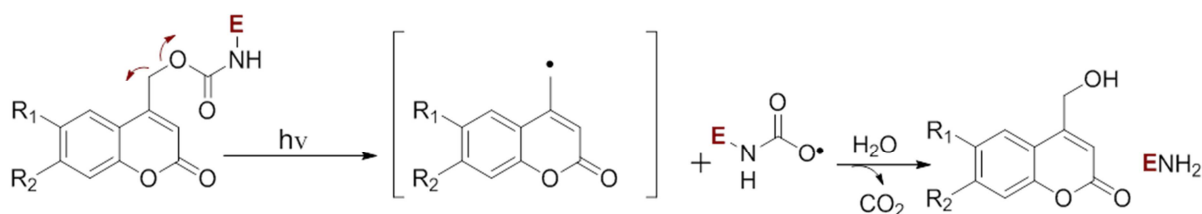
Scheme 5: *Uncaging mechanism of generalized coumarin cages.* ***E*** *represents the effector molecule.*

The subsequent decarboxylation of the carbamate linker is the rate-limiting step and occurs in the subsecond range.[312] The cage conjugates are described to be stable in the dark for more than one week and carbamates are even more stable. Depending on the modifications in positions $R_{1/2}$ the solubility and absorption range of the coumarin can be fine-tuned. The extinction coefficients

are, in contrast to the quantum yields, high in the range above 300 nm and also in the visible range. Brominated coumarin cages are described to be suitable for two-photon absorption.[315]

3.2   Coumarin as Cage

Coumarin was first extracted from tonka beans in 1820 but is also found in vanilla grass, sweet-clover, cherry blossoms and cassia cinnamon.[316] Due to its characteristic smell, reminiscent of dried hey, and its sweet flavor, it is widely used in fragrances, cosmetics, tobacco products, and as a food additive. In the EU and USA the use of coumarin and its analogues is regulated because of its metabolic effects.[317] At higher concentrations coumarin can cause headache, vomiting, dizziness, hypersomnia, paralysis, apnea or coma. In animal experiments also liver- and kidney damage is described.[318] The LD_{50} value for rats was determined to be 293 mg/kg.[319] Nevertheless, only minimal risks for humans have been described.[319-320] In contrast, coumarines are used in medicine for the treatment of lymphedema[321] or as an anticoagulant[322] (Warfarin[323]), and some have antimicrobial[324] while others have antifungal[325-326] activity. The anticoagulant effect is exploited in rat poison.[326-327] Coumarins are extensively used as gain media in blue-green tunable organic dye lasers.[328-329]

Furthermore, coumarins have multiple applications as caging compounds. They have been used to cage DNA strands in hydrogels and neurons[330] and to regulate Ca^{2+} homeostasis *via* ion channels when coupled to membrane lipid mediators and thus elucidate the mechanisms of cellular signaling pathways[331]. A coumarin-fibronil conjugate was successfully tested as an insecticide against mosquito larvae, which is inactive at night and becomes active during the day upon exposure to blue light (420 nm).[332] Also, by means of blue light, caged cGMP can be released from 7-diethylaminocoumarin phosphoester in neurons and thin brain slices with the high quantum yield of 0.18 and an extinction coefficient of 43 000 M^{-1} cm^{-1} at 453 nm.[333] Caged ceramide analogues could be taken up by cells to increase cell viability by using a coumarin modified with a short polyethylene glycol.[334] Nadler *et al.* went one step further, modifying arachidonic acid with sulfonated or neutral, membrane-penetrating coumarin caging groups to locally induce signaling either at the plasma membrane or on internal membranes in beta-cells and thin brain slices derived from mice. With a live-cell microscopy setup they could control the uncaging either at the plasma membrane, where it triggers a strong enhancement of calcium

oscillations in beta-cells and a pronounced potentiation of synaptic transmission, or inside cells, where arachidonic acid blocks calcium oscillations in beta-cells and causes a more transient effect on neuronal transmission.[335] Coumarin caging compounds are also used for various applications in peptide chemistry. For example the dicyanocoumarin called DEAdcCE was used to cage a cyclic RGD peptide to demonstrate that uncaging can be efficiently performed with biologically compatible green light. DEAdcCE was linked with an ester bonding to the side chain carboxyl group of Fmoc-protected aspartic acid and incorporated into the peptide sequence *via* common solid phase peptide synthesis [SPPS]. Caging minimized aspartimide-related side reactions during synthesis.[336] The bioactive tetrapeptide AAPV, which is able to inhibit human neutrophil elastase, was linked to 7-methoxycoumarin-2-ylmethyl and thus successfully tested against the enzyme involved in chronic inflammatory diseases of the skin, like psoriasis.[280] Furthermore, the protein synthesis inhibitor anisomycin was caged with the brominated coumarin derivative 6-bromo-4-hydroxymethyl-7-hydroxycoumarin [Bhc], masking the biological activity that was subsequently made functional by uncaging *in vivo* with high spatial and temporal precision under non-damaging light conditions in one defined cell in a culture of neuronal and kidney cells (**Figure 9**).[337] The brominated coumarin analogues, Bhc esters or carbamates, were first described by Furuta et al. in 1999 as effective cages for amines and carboxylates that undergo efficient photolysis after one- or two-photon absorption.[338] The two-photon cross sections were up to one or two orders of magnitude better than those of other described coumarin cages. The authors further demonstrated these advantages, as well as non-toxicity of the uncaging byproducts, on neurons in thin brain slices. Bhc-Glu conjugates were locally applied to the slices by pressure-ejection from a capillary tube.[338] Eisenhardt *et al.* improved the Bhc-caged protein synthesis inhibitors emetine and anisomycin with a highly water soluble BBhcMOC coumarin derivate (**Figure 9**). It was determined that 99% of emetine and 82% of anisomycin were recovered after two seconds of irradiation with 360 nm light.[297] For the caging of thiols 6-bromo-7-hydroxy-3-methylcoumarin-4-ylmethyl [mBhc] was explored by Distefano *et al.* and tested in modified hydrogels with mBhc-protected cysteamine. The cage was introduced *via* SPPS using an Fmoc-Cys(mBhc)-OH building block (**Figure 9**).[339] Katayama *et al.* described a method to connect two peptides with the use of Bhc caging with an additional Fmoc-protected linker on the hydroxyl group in position 7 (**Figure 9**). During uncaging, acetylated peptide 1 is

tracelessly cleaved off and peptide 2, a cell penetrating peptide for example, remains attached to the Bac linker.[340]
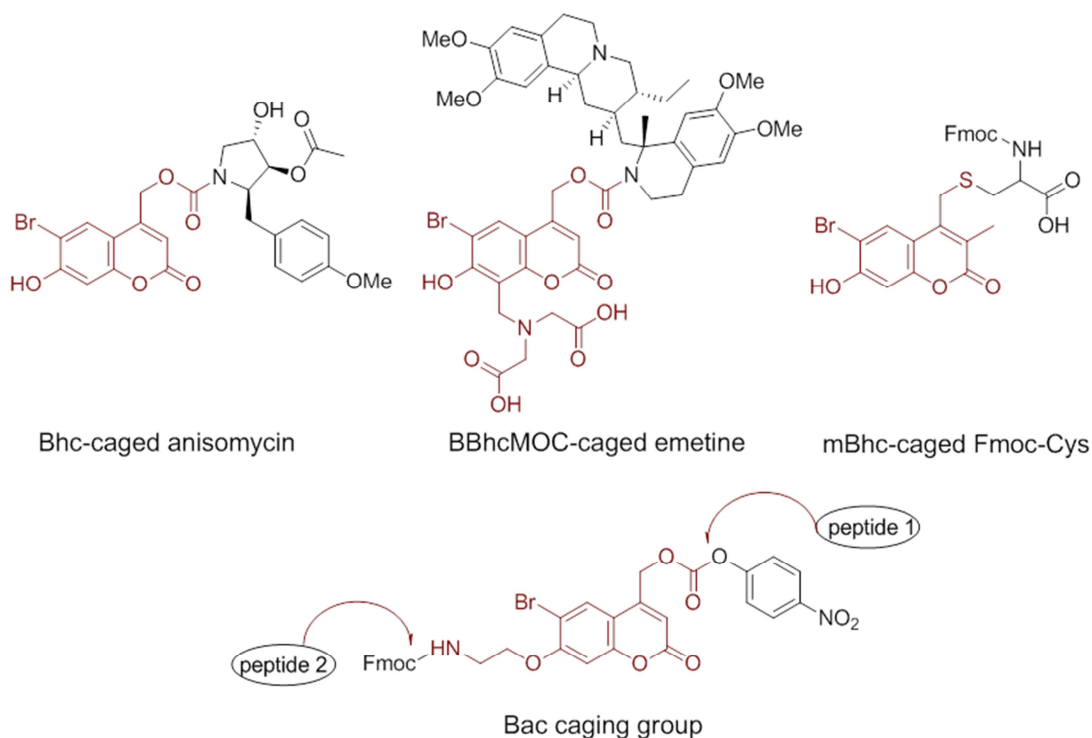
Figure 9: Two examples for caged protein synthesis inhibitors. Brominated 7-hydroxycoumarin-4-ylmethyl [Bhc] caged anisomycin[337] and with improved water-solubility BBhcMOC-caged emetine[297]. In addition the mBhc-caged Fmoc-cysteine building block[339] and the Bac building block[340] for solid phase peptide synthesis.

4 Advantages of Fluorine

4.1 Fluorine in Peptides

Many small molecules[341-344], metals[50, 345-346] and other surfaces[344, 347], salts[348-349] and covalent bonds to atoms like selenium[350], iodine[351] or chlorine[352-353] have been studied in context with peptides and proteins and the list of known unnatural amino acid building blocks grows continually. In 2010 Liu and Schultz named already 67 unnatural amino acids that could be incorporated *in vivo via* the expanded genetic code.[354] One particularly interesting modification is the incorporation of fluorine into peptides and proteins. Although fluorine is found in inorganic materials in earth's crust[355] and living organisms, no organic fluorine molecules have been isolated from animals, only a few such compounds from a small number of tropical and subtropical plants[356] and in two actinomycetes thus far.[357] Fluorine is the smallest of the halogens

27

with a van der Waals radius of 1.35 Å, only slightly greater than hydrogen with a radius of 1.20 Å. Furthermore, it is the most electronegative element in the periodic table with an electronegativity of 4.0.[357] The thereby induced strong inductive effect leads to a polarized, ionic character of the C-F bond and supports electrostatic or dipolar interactions with surrounding functional groups. Due to the magnitude of electrostatic attractions between the polarized $C^{\delta+}$ and $F^{\delta-}$ atoms, organic fluorine is typically a poor electron donor for hydrogen bond formation. Thus, the substitution of hydrogen against fluorine in a peptide has a minor steric but potentially major electronic consequence.[358] Although none of the 20 naturally occurring amino acids contains fluorine, fluorinated amino acids have become a powerful tool to modulate the properties of peptides. The incorporation of even a single fluorine atom can alter the hydrophobicity, polarity, acidity/basicity, reactivity or conformation of a peptide or protein. These properties are crucial factors in peptide engineering, thus fluorinated amino acids are frequently used to fine-tune properties such as protein folding, proteolytic stability, and protein-protein interactions.[359-360] Further, biomolecules containing fluorine have been found to gain an increase in biological activity as well as bioavailability.[361-363] This is why as of 2016 already up to 30% of administered drugs contained fluorine atoms or fluoroalkyl groups, and this trend is increasing.[360] Furthermore, the isotopes ^{18}F and ^{19}F are excellent tools in medical imaging by means of NMR, PET or MRI.[364-366] With the use of fluorine the design of peptide and protein based drugs could be improved in terms of reducing their rate of digestion by proteases, which limits their application as therapeutics.[19] Most studies concentrate on fluorinated aliphatic or aromatic hydrophobic amino acid building blocks[367-368], but for example also fluorinated histidines[369], prolines[370], methionines[371] or phenylalanines[372] have been described during the last years. However, careful differentiation between single or multiple fluorination of alkyl chains is necessary. Whereas the single substitution of one hydrogen with fluorine is often discussed to be almost isosteric, the CF_3 group is already twice the van der Waals volume of a CH_3 group and the steric effects are more comparable to an isopropyl group.[373] Furthermore, a certain number of fluorine atoms is necessary to increase the hydrophobicity[374], since single fluorination of an aliphatic side chain in fact slightly reduces the hydrophobicity of the particular amino acid.[375] Berger and Koksch reviewed in 2017 the major findings of fluorine's impacts as follows: "1) small numbers of fluorine atoms introduce polar character into otherwise hydrophobic amino acid side chains, representing a departure from earlier studies that had demonstrated the

hydrophobicity/lipophilicity of heavily fluorinated side chains; 2) fluorine dramatically changes the secondary structure propensity of aliphatic amino acids and, thus, the folding properties of accordingly modified peptides and proteins; 3) fluorine can considerably influence the proteolytic stability of peptides; and 4) the impact of fluorine not only depends on the nature of the fluorinated side chain but also on the immediate environment with which it interacts."[376] Other groups exemplify generalizable paradigms for future design of coiled-coils and collagen mimetics with fluorine substituents[377] and the thermal and chemical stability of those peptides[378]. The VPE/VPK coiled-coil model was used to describe how the packing of the monomers was effected in dependency to the orientation and neighboring side chains and relative to the number of fluorine atoms in (*S*)-2-aminobutyric acid [Abu] analogues.[73] Also the various biosynthetic methods for the introduction of fluorinated amino acids in lager proteins promises to expand the capabilities of protein design.[379]

4.2   5,5,5,5',5',5'-Hexafluoroleucine [HfLeu]

Highly fluorinated amino acids have been used to investigate potential applications in various protein-based biotechnologies. The *L*-enantiomer of HfLeu is an attractive target as fluorinated analogue of leucine, an amino acid that has been demonstrated to play an important role in the folding of many peptides and proteins. The first synthesis of racemic HfLeu was reported in 1968[371] and the first chiral *L*-HfLeu synthesis was reported in 1998[380]. Today there is more than one synthetic route described to achieve *L*-HfLeu depending on the starting molecule.[381-382] Due to its high fluorine content, HfLeu is much more hydrophobic and sterically demanding than comparable aliphatic amino acids, or the unbranched (*S*)-2-aminobutyric acid [Abu] and the even longer (*S*)-2-heptanoic acid [2-Aha], which is comparable in size, but still more polar (**Figure 10**).[380]
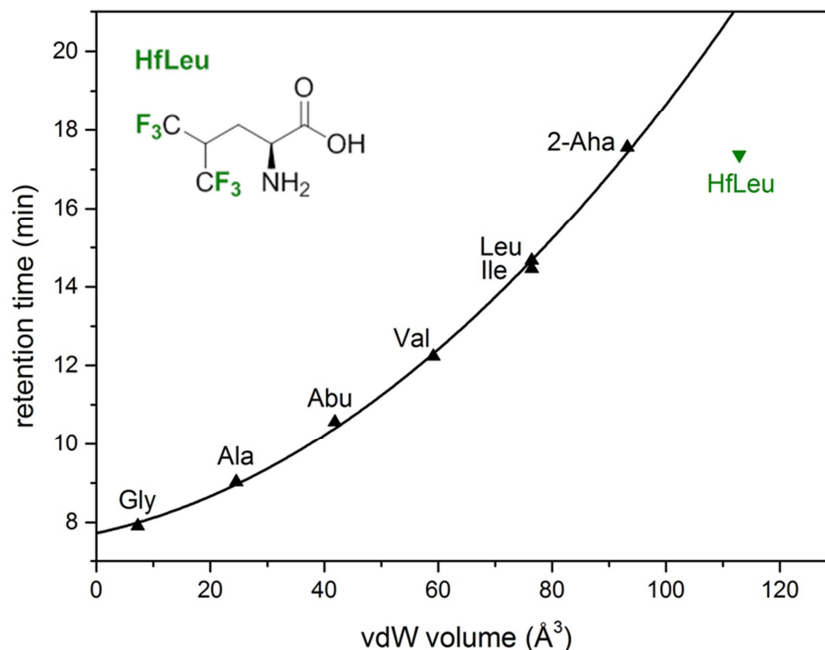
Figure 10: *Retention times of the Fmoc-amino acids plotted against the van der Waals volume of their side chains. The van der Waals volumes were calculated according to Zhao et al..[383] Nonfluorinated amino acids are represented by black triangle, the correlation between them is shown with a black line, and the fluorinated amino acid is represented by a green triangle. The structural formula of L-HfLeu is shown with fluorines highlighted in green.[380, 384-385]*

Chiu *et al.* compared in 2009 fluorinated GB1 mutants regarding their thermal stability and found the HfLeu mutant to be more resistant towards thermal denaturation than the Abu, Leu, (*S*)-5,5,5',5'-tetrafluoroleucine or (*S*)-2-amino-4,4,4-trifluorobutanoic acid analogues. Only the aromatic phenylalanine and pentafluorophenylalanine were even more stabilizing.[386] A conformation stabilizing effect was also reported for helical coiled-coil model peptides where HfLeu was incorporated with a high amount into the hydrophobic core instead of Leu. Buer *et al.* found the most stable coiled-coil structure when HfLeu alternates with Leu residues in the **a** and **d** positions.[387] On the other hand the free energy of unfolding a 27-residue antiparallel 4-α-helix peptide bundle was increased up to 25% when HfLeu was substituted for all Leu side chains in the **a** and **d** positions.[388]

Furthermore, HfLeu containing peptides and proteins were investigated concerning their protease stability and there are diverse discussions about the impact of fluorination. While HfLeu could increase the proteolytic stability of the antimicrobial peptide MSI-78 towards trypsin and chymotrypsin only in the presence of liposomes[389], the resistance of magainin 2 amide peptide

and buforin were enhanced towards proteolytic degradation by trypsin when HfLeu was incorporated[390]. Also the proteolytic stability of GLP-1 peptide towards serine protease dipeptidyl peptidase was increased by HfLeu in the peptide sequence.[391]

4.3 Fluorinated Amino Acids in Protease Studies

The electronic and steric properties of fluorine containing amino acids was shown to have an impact on substrate-enzyme binding, where the presence of fluorine can modify the acidity of neighbouring functional groups and affect the binding affinity of an enzyme for the certain fluorinated substrate.[357] The Koksch group developed a simple 9-residue peptide model as protease substrate named FA (**Figure 11**).
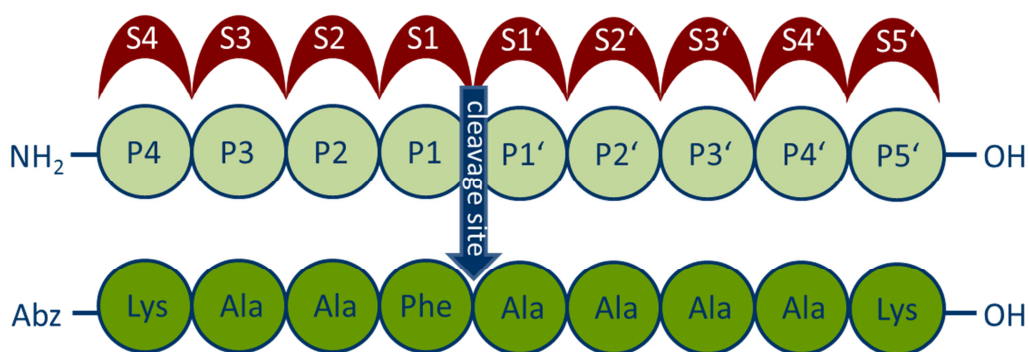
Figure 11: *Schechter and Berger model[392] with the enzyme in red and the substrate in light green with the desired cleavage site, between position P1 and P1', in blue. The sequence of the protease substrate FA is given below the general substrate sequence.[19]*

The peptide was designed according to the Schechter and Berger model[392], with a central Phe residue in the P1 position immediately upstream of the cleavage site to satisfy the known substrate specificities of serine and aspartic endopeptidases.[19, 36, 376] The inexpensive and easy to handle fluorescence label Abz was used to enable the visualization of the digestion process. Abz is a small fluorescent molecule with size and structure comparable to smaller natural amino acids and has been used as an extrinsic probe for peptides thereby having only minor influence on the peptide conformation when coupled to the N-terminus.[393] Because of its high quantum yield, Abz is used as a convenient fluorescence donor group in peptides that are substrates for several proteolytic enzymes, forming a donor-acceptor pair with N-[2,4-dinitrophenyl]-ethylenediamine, as in human tissue kallikrein for example. Here, Abz was coupled to the substrate peptide

bradykinin.[394] The FA peptide contains two lysine residues for better solubility in physiological media. All other positions are occupied by Ala to ensure flexibility of the substrate. In previous studies one Ala of FA in either position P2, P1' or P2' was substituted against Abu or its fluorinated analogues *(S)*-2-amino-4,4-difluorobutanoic acid [DfeGly] or *(S)*-2-amino-4,4,4-trifluorobutanoic acid [TfeGly] (**Figure 12**) to investigate the fluorine effect. The resulting peptides were exposed to human blood plasma, elastase, α-chymotrypsin or pepsin. For α-chymotrypsin and pepsin most fluorinated substrate peptides were cleaved even faster than the parent FA sequence. Only in two particular cases was the observed digestion rate decreased. If TfeGly or DfeGly are incorporated into the P2' position, resistance towards pepsin was enhanced. The same effect for α-chymotrypsin could be achieved only with DfeGly in position P2. The authors interpreted these effects with the fluorine-induced polarity which disfavors the accommodation of the fluorinated residue within the protease active site.[36] Also the P2 variants with DfeGly or Abu were highly stable against elastase and in human blood plasma. Furthermore, a similar effect could be observed for TfeGly in position P1. On the other hand, all other substrate peptides were degraded even faster than FA when incubated with elastase or human blood plasma.[19] Nevertheless, in about 25% of all cases studied in the Koksch group so far, peptide substrates were protected due to the incorporation of a fluorinated amino acid.
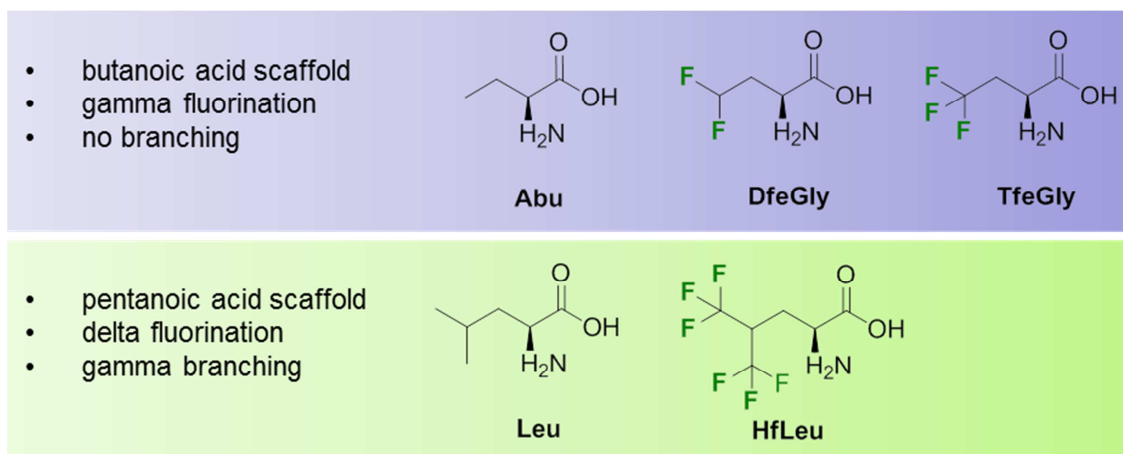
Figure 12: Structures, abbreviations and characteristics of fluorinated Leu and Abu analogue amino acids.

4.4   Elastase, α-Chymotrypsin, Proteinase K and Pepsin

Approximately 1.7% of the human genome encodes proteases, and these enzymes are present in all living organisms.[395] Proteases are structurally diverse and have the capability to hydrolyse

peptide bonds. They are divided into two groups. First the endopeptidases, like all four proteases discussed here, which cleave the substrate peptide at specific amino acids within a protein. Secondly the exopeptidases, including carboxypeptidases that cleave peptide bonds at the C-terminus of proteins. As are all enzymes, proteases are defined by the EC nomenclature, in section 3.4 with 14 subsections.[396] Elastase, α-chymotrypsin and proteinase K are serine proteases with the EC sub-section 3.4.21 and pepsin is a typical aspartic protease with the EC sub-section 3.4.23.[396] Serine proteases form the largest group of known proteases and they do play an important role in many physiological processes like digestion (elastase, α-chymotrypsin), blood coagulation or as granzymes for example.[397-399] The X-ray crystallographic structures of the serine proteases are highly similar, but they have different specificity pockets that allow only selected peptide bonds to be hydrolyzed.[400] To hydrolyze the peptide bond, serine proteases use a "catalytic triad", which was first described in 1960.[401] Typically the catalytic triad is composed of serine, which acts as a nucleophile, a general base like histidine and an aspartic acid.[402] Nevertheless, as Polgár reviewed in 2005, several new families of catalytic triads have been discovered in which the nucleophile-base-acid pattern is generally conserved, but the individual amino acids can differ.[403-404] In contrast to serine proteases, aspartic proteases do not need a covalent acyl-enzyme intermediate on the reaction pathway. During hydrolysis, the substrate is attacked directly by a water molecule which is coordinated *via* hydrogen bridge linkage between two aspartic residues in the catalytic site.[405-406] Aspartic proteases are the smallest class of human proteases with only 15 members. Most aspartic proteases are mammalian enzymes like pepsin, as well as a range of microbial and plant enzymes. Over the past years, they have received considerable attention as potential targets for pharmaceutical intervention since many have been shown to play important roles in physiological and pathological processes.[407]

Pancreatic Elastase (EC 3.4.21.36)

Elastases are a subclass of chymotrypsin-like serine proteases.[408] The name was derived from the fact that most elastases are able to digest the structural protein elastin, a highly insoluble extracellular protein giving many tissues their elastic properties.[409] Pancreatic proelastase is structurally similar to trypsin and other serine proteases in the pancreas and has a signal peptide, activation peptide, and a mature, functional protein formed by a single peptide. Proelastase is activated by trypsin but an excess of trypsin could also serve as inhibitor for elastase

formation.[410] The classic target molecule, elastin is composed mainly of glycine, proline, and other hydrophobic residues. It contains multiple lysine-derived cross-links so that the enzymes that digest elastin are generally powerful proteases able to hydrolyze numerous proteins.[400] Pancreatic elastase prefers small uncharged amino acids like alanine, glycine and serine adjacent to the cleaved bond in the substrate and does not accept isoleucine in its active site.[400, 411] Its binding site is spread over eight subsites, S5 to S1, and S1' to S3'.[412]

α-Chymotrypsin (EC 3.4.21.1)

α-Chymotrypsin is the second most abundant serine protease in pancreatic secretions and makes up 9% of total pancreatic juice protein.[413] It is synthesized in the pancreas with an 18 amino acid signal peptide, a 15 amino acid activation peptide and a 230 amino acid active molecule. Activation of chymotrypsinogen to α-chymotrypsin is accomplished by trypsin acting at the Arg15–Ile16 bond. However, the activation peptide remains attached to the α-chymotrypsin.[414] Retention of the activation peptide through disulfide bonds stabilizes the active enzyme against sensitivity to heat or acid denaturation.[400] α-Chymotrypsin hydrolyses peptide bonds adjacent to large hydrophobic amino acids with aromatic side chains, including phenylalanine, tyrosine, and tryptophan in P1 and P2 positions.[415-416] Secondary hydrolysis also occurs at the carbonyl end of isoleucine, methionine, serine, threonine, valine, histidine, glycine, and alanine.[404, 417]

Proteinase K (EC 3.4.21.64)

Proteinase K is a serine protease produced by the fungus *Tritirachium album Limber*[418] with a broad substrate specificity and high activity. Also proteinase K is specified as an endopeptidase, it is able to cleave bonds in the middle of numerous peptides,[419] where it prefers aromatic or aliphatic hydrophobic amino acids in position P1[418]. Interestingly, Saenger *et al.* even claimed that not the substrate sequence but only the volume is important for the efficacy of proteinase K.[420] Furthermore, Ala is favored in position P2 and enhances cleavage efficiency.[421] The name proteinase K was derived from its ability to digest even keratin in hair. The enzyme is extraordinarily stable in the pH range 4-12 and also at higher temperatures (60°C). Many denaturation agents like guanidinium chloride or sodium dodecyl sulfate help to increase the activity of proteinase K by gaining the flexibility of the substrate but leaving the enzyme intact. Nevertheless, these findings could be only secured for native substrates and complex peptides.[422]

Pepsin (EC 3.4.23.1)

The stomach produces pepsin that is especially useful in digesting muscle, tendons and other components of meat with high collagen contents.[400] It has its optimal activity at pH 1.8-3.5 and is irreversibly inactivated at pH ranges over 7.[400, 423] The extended active site of pepsin can bind at least seven residues[424] and it prefers to cleave peptide bonds N-terminal[425] to hydrophobic especially aromatic residues like Phe, Trp, and Tyr but also Leu at the P1 and P1' positions.[426] His, Lys, Arg, or Pro residues prohibit cleavage when found at the P1 position and Pro is also not tolerated at position P2.[427] The S2 subsite of pepsin preferentially accommodates hydrophobic residues such as Leu, Ala or norleucine as well as Ile and Val, but can also bind charged residues.[428]

General Aim

Four different peptide models are discussed in this thesis. Each model describes a different perspective on a topic of general interest. Thereby four different ways of peptide design are presented to generate a toolbox for peptide engineering. The aim of the project A was to design a peptide that undergoes a time-dependent conformational transition from random-coil to amyloids and additionally includes a recognition motif for phosphorylation by the kinase PKA. If the model peptide was phosphorylated, the aggregation could be inhibited even during the aggregation process. In previous studies the opposite reaction was shown. The aggregation was induced by enzymatic dephosphorylation. Combination of the gained knowledge could result in a dynamic switchable amyloid biomaterial.
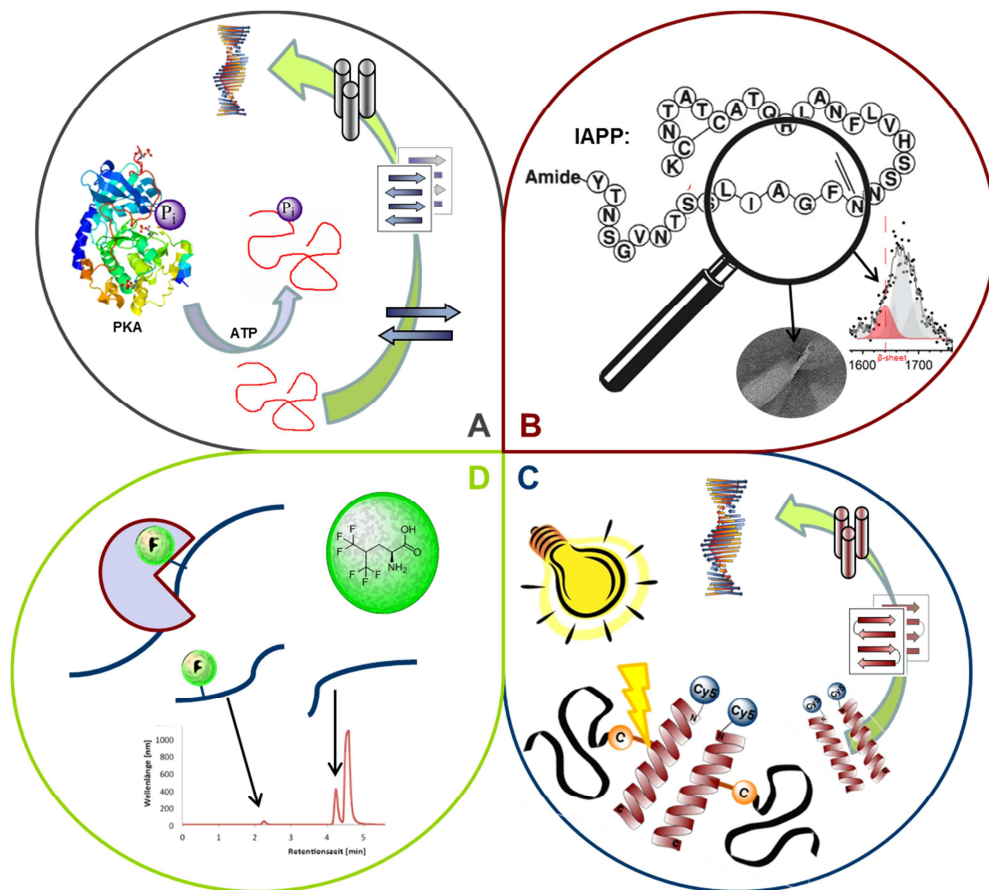
Figure 13: *Inhibition of peptide aggregation by means of enzymatic phosphorylation (project A). Oligomer investigation with NFGAIL (project B). A light-sensitive cage for aggregating peptides (project C). Hexafluoroleucine helps to increase protease stability of peptides (project D).*

In project B the focus is on pathologic aggregation behavior of diabetes type two related hIAPP and its hydrophobic patch region NFGAIL. Previous studies identified the short NFGAIL sequence as ideal model to study the aggregation pattern of full length hIAPP. Because of its short sequence, several molecular modeling simulations did concentrate on NFGAIL and results are discussed controversial. The results presented show some of the first experimental insights into structural identity, form and charge of the small oligomers on the beginning of the aggregation process. These oligomers are known to be the toxic species during amyloid formation. Also new insights about the aggregation pathway and amyloid fibril pattern are discussed. To transfer the knowledge and advantages of model peptides into cell experiments a caging model was enhanced in project C. A coumarin based Bhc caging compound was designed to link a cell penetrating polymer to the side chain of a peptide. The caging should carry the peptide into the cell and prevent it from folding and proteolytic digestion. The peptide can be traceless released by illumination with light. The thereby induced conformational behavior can be monitored inside the cell by fluorescence labeling of the peptide. A model peptide to study proteolytic digestion was modified in project D. The influence of the fluorinated HfLeu amino acid in different positions of the protease substrate peptide FA was evaluated. Fluorinated amino acids could help to increase the bioavailability of peptidic drugs, if they are able to decrease the proteolytic digestion in human metabolism.

Results and Discussion

Applied Methods

Analytical High Performance Liquid Chromatography [HPLC]

Analytical HPLC was used to determine the purity of the synthesized peptides and the retention time is one parameter for the identification of a peptide.[429-431] Using a diode array detector [DAD] or a fluorescence detector the visualization of optical active compounds like Abz, phenylalanine, Fmoc protected amino acids or Bhc becomes possible.[36, 72, 432-433]

Electrospray Ionization Time of Flight [ESI ToF] Mass Spectrometry [MS]

ESI Tof MS gives defined results on the high resolution mass to charge ratio of a compound, if the molecule is ionizable. In combination with analytical HPLC characterization, a high resolution mass qualifies to unequivocal identify peptides. In ESI ToF the mass of a compound is denoted in dependency to the charge in a mass to charge ratio.[44, 434-436]

Nuclear Magnetic Resonance [NMR]

NMR gives defined insights of the molecular structure of a compound. Thereby, it gives defined insights about the identity of a compound and its impurities. With NMR the detailed structure of a (small) compound and its functional groups can be described. Additionally, it can be used to follow reaction kinetics and to monitor different reaction states. NMR is strongly dependent on the probed atoms covalent and noncovalent interactions. NMR data are given with respect to the frequency of the instrument and the signal of an internal standard (tetramethylsilane) as chemical shift of the certain atom and coupling constants to interacting atoms.[437-439]

Circular Dichroism [CD] Spectroscopy

CD is a useful spectroscopic technique for studying the secondary structure, folding and binding properties of peptides. According to the extrema, the spectrum could be interpreted as an average of all present secondary structures of a peptide solution. Two minima at 208 and 222 nm are characteristic for an α-helix, a random structure appears with an minimum around 200 nm and a β-sheet with a minimum at 218 nm and a maximum around 200 nm.[127, 440-442]

Thioflavin T [ThT] Fluorescence Spectroscopy

ThT fluorescence assays are performed to monitor time-dependent amyloid formation. The dye becomes fluorescent when intercalated into amyloid fibers.[117-118, 121, 443]

Transmission Electron Microscopy [TEM]

TEM images show the morphology of structures visible in the nanometer scale. Thus the identification of structures and some pattern like diameters of protofibrils, aggregates or peptide crystals is possible.[444-445]

Small Angle X-ray Spectroscopy [SAXS]

In combination with TEM, SAXS is a technique for the elucidation of the topology of an oligomer or aggregate in the nanoscale range. Morphology pattern like form, length, high and diameter could be calculated from the data resulting from SAXS.[446-451]

Part A: Inhibition of Peptide Aggregation by Means of Enzymatic Phosphorylation

The results presented in this section were originally published as **K. Folmert**, M. Broncel, H. v. Berlepsch, C. H. Ullrich, M. Siegert and B. Koksch, Inhibition of Peptide Aggregation by Means of Enzymatic Phosphorylation*, Beilstein J. Org. Chem.* **2016,** *12,* 2462–2470.

The original paper with supporting information is available at: doi:10.3762/bjoc.12.240

A1   Abstract

The presented studies were inspired by previous results presented from Koksch *et al.* in 2010 in which an amyloid-forming model peptide remained in a random coil conformation if phosphorylated at different positions of the sequence during the synthesis,[72] but spontaneously aggregated if the phosphate group was enzymatically removed[254]. Hence, a new previously unreported 26-residue peptide model KFM6 was designed. KFM6 combines structural characteristics of random coils, α-helical coiled-coils and β-sheets in one sequence and thus undergoes a spontaneous conformational transition to amyloid fibrils over time under physiological conditions. Furthermore, the peptide contains a recognition site for cAMP-dependent protein kinase A (PKA) that enables enzymatic phosphorylation. The study included

an analysis of how enzymatic phosphorylation influences the pathway of peptide aggregation and illustrated a remarkable inhibitory effect. While most inhibition studies concentrate on the later phase of the aggregation process, the results presented in this report suggest a strong directing effect, concomitant with enzymatic phosphorylation, to random structures during the initial phase of the aggregation process in which small oligomers and protofilaments are assumed to prevail. The influence of the electronegative phosphate group on the peptide conformation is discussed with respect to different properties including steric bulk, charge and solubility. Furthermore, it was found that the location of the phosphate group, in the vicinity of the suggested aggregation domain of our model peptide, could enable a "gatekeeping" phenomenon that reduces the aggregation propensity of the peptide. Additionally, the impact of the phosphorylation components was investigated during time-dependent aggregation studies. PKA and the phosphate donor ATP make the peptide more prone to aggregation, but this process can be halted or even reversed by inducing enzymatic phosphorylation.[70]

Figure A1: *Schematic aggregation pathway from random coil to amyloid-fibrils and inhibition of aggregation by phosphorylation.*

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Inhibition of peptide aggregation by means of enzymatic phosphorylation

Kristin Folmert[1], Malgorzata Broncel[2], Hans v. Berlepsch[1], Christopher H. Ullrich[3], Mary-Ann Siegert[4] and Beate Koksch[*1]

Abstract

As is the case in numerous natural processes, enzymatic phosphorylation can be used in the laboratory to influence the conformational populations of proteins. In nature, this information is used for signal transduction or energy transfer, but has also been shown to play an important role in many diseases like tauopathies or diabetes. With the goal of determining the effect of phosphorylation on amyloid fibril formation, we designed a model peptide which combines structural characteristics of α-helical coiled-coils and β-sheets in one sequence. This peptide undergoes a conformational transition from soluble structures into insoluble amyloid fibrils over time and under physiological conditions and contains a recognition motif for PKA (cAMP-dependent protein kinase) that enables enzymatic phosphorylation. We have analyzed the pathway of amyloid formation and the influence of enzymatic phosphorylation on the different states along the conformational transition from random-coil to β-sheet-rich oligomers to protofilaments and on to insoluble amyloid fibrils, and we found a remarkable directing effect from β-sheet-rich structures to unfolded structures in the initial growth phase, in which small oligomers and protofilaments prevail if the peptide is phosphorylated.

Introduction

Amyloid fibrils are one of the most important and studied self-assembled materials in nature. A wide range of peptides and proteins with various primary sequences and functions are able to form these wel organized and highly stable aggregates under multiple conditions [1]. The morphology ranges from liquid crystals to ribbons, rigid nanotubes and funnels [2]. Although amyloid aggregates are mostly known for pathologic events like Alzheimer's and Parkinson's disease or type two diabetes [3-6],

they also combine the potential for applications with biological function with desirable mechanical properties. Many such peptide aggregates display exceptional stability, mechanical strength, stability at high temperatures and they are resistant towards enzymatic degradation [2]. Several studies have demonstrated the utilization of amyloids as functional templates like conductive nanowires [7], water-filled nanotubes [8] or biocompatible hydrogels [9]. Likewise, nature uses amyloid-like cross-β-sheet-rich conformations to store peptide hormones in secretory granules of the endocrine system [10]. As another example it was found that mammalian Pmel17 amyloid templates and accelerates the covalent polymerization of reactive small molecules into melanin, an important biopolymer that protects against a broad range of cytotoxic insults, including UV and oxidative damage [11]. While amyloid morphologies and events which induce amyloid formation are the subject of numerous studies, the pathway of aggregation from soluble α-helical-rich or partially unfolded peptides into insoluble β-sheet-rich amyloid fibrils remains unclear. Several studies discovered parameters of peptide environment, like pH [12,13], oxidative stress [14], presence of organic components [15,16] or metal ions [17,18], to have a crucial influence on the conformational transition. The challenging physicochemical properties, the low solubility and the tendency to aggregate make natural amyloid-forming peptides difficult to synthesize and complicates detailed structural characterization. One possibility to overcome these problems is the use of de novo designed model peptides [19]. The identification of short domains in peptide sequence which make a peptide prone to amyloid formation [20-23] inspired the design of a series of amyloid-forming model peptides with defined characteristics, outstanding stability, regular fibrous architecture and high synthetic accessibility with numerous chemoselective ligation and modification methods [24-26]. Also, various post-translational modifications, like phosphorylation or glycosylation have been studied as aggregation triggers [27-29]. Phosphorylation caused by abnormal activities of phosphatases and/or kinases is associated with known diseases such as cancer [30], multiple sclerosis [31], diabetes [32] and Alzheimer's disease [33]. The accumulation of hyperphosphorylated tau protein as trigger for several dementias was intensively discussed [33-35]. Moreover, the effect of enzymatic phosphorylation on Alzheimer-relevant APP was demonstrated already decades ago [36] and has been followed up in plenty of studies since then [37-40]. Phosphorylation experiments were also transferred to model peptide systems by our group. We explored a 26-residue coiled-coil peptide which undergoes a conformational transition to amyloid fibrils in 24 hours under physiological conditions [41], but remains random coil if one of three serine residues carries a phosphate group [27]. The aggregation process could be restored by addition of Lambda Protein Phosphatase that removes the phos-

phate group [42]. It is widely accepted that this reversible process of phosphorylation directs diverse properties of peptides or proteins in nature, ranging from interactions with other proteins and nucleic acids to subcellular localization and binding [43]. Additionally, phosphorylation may be found to direct desired behaviors in the currently intensively studied area of amyloid-based biomaterials. In this report we show the use of the negatively charged phosphate group to control the aggregation process. Using PKA and ATP (adenosine triphosphate) as phosphate group donor, we studied the impact of the phosphate group on the oligomerization pathway at different conformational states from random coil, over β-sheet monomers to protofilaments and amyloid fibrils. Furthermore, we demonstrate the importance of the influence of phosphorylation components on the peptide conformation during this process. This knowledge could become a useful tool in employing enzymatic phosphorylation and dephosphorylation as triggers for or inhibitors of amyloid formation, as it was previously shown for a self-assembling, supramolecular hydrogel [44].

Results
Peptide model

Peptide KFM6 follows a typical coiled-coil heptad repeat sequence and it includes five amino acids (-R-R-A-S-L-) in positions 20–24, in proximity to the C-terminus, that serve as the recognition motif for PKA. The crucial role of this recognition motif for efficient phosphorylation has long been established [45]. It is also known that enzyme activity depends upon the local flexibility and solvent-exposed position of the target amino acid in the peptide [46]. Scheme 1 depicts the key design

Scheme 1: Helical wheel projection of two parallel helical strands and primary structure of KFM6. The recognition motif KFM6(20–24) for PKA is given in bold and valine residues are bold and underlined.
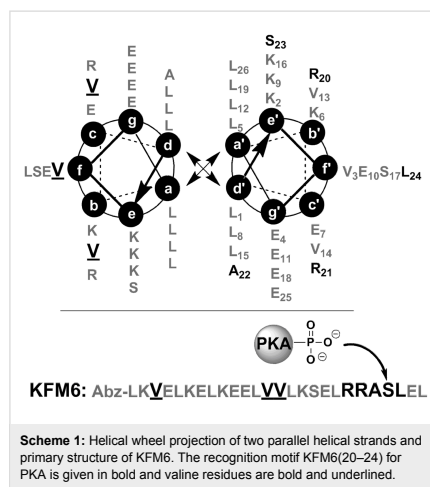
features, which are based on one of our previously published peptides [47], and the primary structure of KFM6. The nonpolar leucine residues at positions a and d contribute to thermal stability by hydrophobic core packing of the leucine zipper motif ("knobs-into-holes principle"). Charged amino acids at positions e, g, b, and c stabilize the coiled-coil by intramolecular electrostatic attractions between glutamates and lysines.

Intermolecular coulombic interactions between e and g direct the monomers into a parallel dimeric orientation. The solvent-exposed position f is occupied by serine for better solubility. One site of each position b, c, and f contains valine, making the system prone to amyloid formation. To ensure the UV-activity of the peptide as analytical tool, Abz (anthranilic acid) was coupled to the N-terminus. The resulting peptide contains elements of α-helices and β-sheets, and a recognition site for PKA. To investigate the structural changes that KFM6 undergoes in the absence of phosphorylation, time-dependent CD (circular dichroism) measurements were performed. To ensure comparable and reproducible starting conditions, a disaggregation and concentration validation assay using HFIP (1,1,1,3,3,3-hexafluoropropan-2-ol) was performed. Thus, all aggregates that may have formed during peptide synthesis and purification were disrupted.

As shown in Figure 1 the peptide KFM6 adopts a random structure upon dissolution in buffer and undergoes a time-resolved conformational transition to β-sheet-rich amyloid fibrils within 24 hours. The CD spectra (Figure 1a) show a typical minimum at 200 nm indicating a random-coil conformation with decreasing intensity over 24 hours, and another increasing minimum at 218 nm, which is characteristic for β-sheets [48]. Absolute values of ellipticity IΘI for 200 or 218 nm are plotted versus the time (Figure 1b) for better visualization.

Although a white precipitate indicates the formation of aggregates, amyloid fibril formation was studied by means of a ThT (thioflavine T) binding assay (Figure 2). ThT yields in an enhanced fluorescence signal at 485 nm when bound to amyloid fibrils, and can be used to determine the kinetics of fibril formation [49]. KFM6 shows a strongly increasing fluorescence signal without lag time, reaching a plateau after ten hours (Figure 2, circles).

$$I_{Fl} = a + \left(\frac{b}{\left(1 + \exp\left(\frac{t_{50} - t}{\tau} \right) \right)} \right)$$
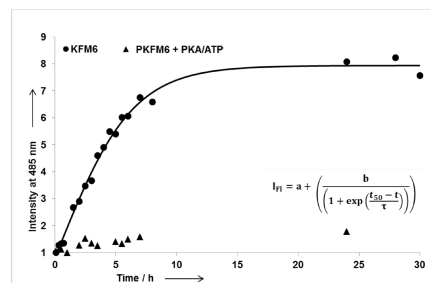
Figure 2: ThT-binding assay of of 15 μM KFM6 in 50 mM Tris/HCl buffer with 10 mM MgCl$_2$ at pH 7.5, 24 °C, and a ThT concentration of 20 μM. The amyloid growth process was monitored for untreated peptide (circles) and peptide in the presence of PKA and ATP (triangles). All values were determined in triplicate and are normalized based on the initial fluorescence intensity (t = 5 min; λ = 485 nm).

The displayed data indicate a fast amyloid growth process that appears to begin upon dissolution in buffer. This observation is in good agreement with the CD spectra, in which no lag time was observed. To discover the morphology of the amyloid fibrils, a TEM (transmission electron micrograph) was obtained at 24 hours of incubation (Figure 3a). The TEM shows typical amyloid morphologies such as extended tubular fibers, some of
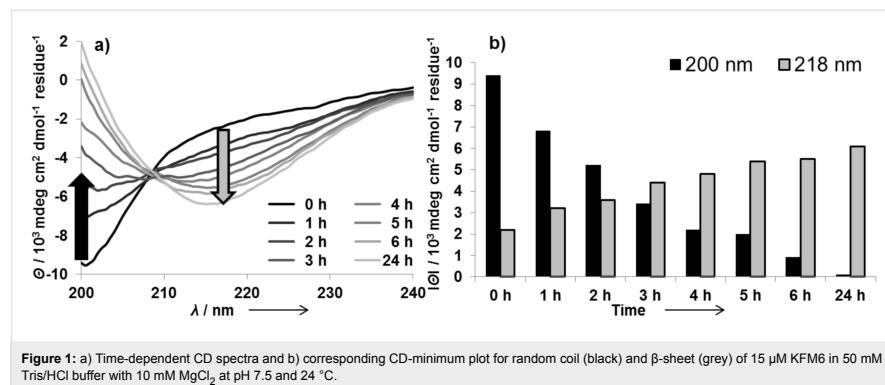

Figure 1: a) Time-dependent CD spectra and b) corresponding CD-minimum plot for random coil (black) and β-sheet (grey) of 15 μM KFM6 in 50 mM Tris/HCl buffer with 10 mM MgCl$_2$ at pH 7.5 and 24 °C.
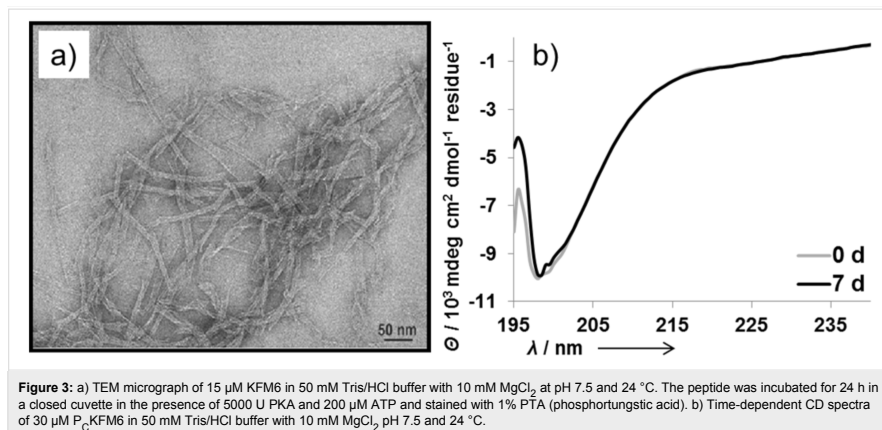
Figure 3: a) TEM micrograph of 15 µM KFM6 in 50 mM Tris/HCl buffer with 10 mM MgCl$_2$ at pH 7.5 and 24 °C. The peptide was incubated for 24 h in a closed cuvette in the presence of 5000 U PKA and 200 µM ATP and stained with 1% PTA (phosphortungstic acid). b) Time-dependent CD spectra of 30 µM P$_C$KFM6 in 50 mM Tris/HCl buffer with 10 mM MgCl$_2$, pH 7.5 and 24 °C.

them ending in funnels. The protofilaments are frequently well resolved. Their spacing is about 3.5 nm, which is in good agreement with estimates for the protofilaments of other related peptides [47].

Enzymatic phosphorylation

To investigate the influence of enzymatic phosphorylation on the aggregation process, buffer containing PKA and ATP was prepared and used to dissolve lyophilized KFM6. PKA is the most studied and structurally well-characterized protein kinase, and it is known to have relatively broad substrate specificity due to its roles in the regulation of energy metabolism, growth, signal transduction, and apoptosis of cells [50]. The serine/threonine kinase PKA uses ATP as phosphate donor. It has to be activated from its natural species with cAMP which leads to dissociation of PKA in two regular and two catalytic subunits [51]. We used the preactivated catalytic subunit of PKA to perform the reactions. The structural behavior of KFM6 during enzymatic phosphorylation was determined by different analytical methods.

The ThT binding assay shows no increase in amyloid formation compared to the unphosphorylated sample over 24 hours (Figure 2, triangles) and the CD-minimum plot indicates a large population of random structures and a constantly low intensity for β-sheets over seven days of incubation (Figure 4a). Together with the TEM micrograph, which shows no sign of amyloid morphologies or other higher oligomers, the results suggest the inhibition of amyloid formation of KMF6 by enzymatic phosphorylation. Also, no precipitation was observed during all enzymatic phosphorylation experiments, which is a further indicator for less aggregate formation.

To learn more about the influence of phosphorylation components on the pathway of aggregation and at which state the growth of oligomers is stopped by phosphorylation, we performed time-dependent enzymatic phosphorylation studies. Therefore, ATP and PKA were added separately at time zero and the other one was added after a certain time assuming that the phosphorylation of KFM6 begins immediately when both phosphorylation components are present in solution. In this way, the influence of ATP or PKA on amyloid fibril formation was monitored as control. In Figure 4b, the persistent minimum at 218 nm for β-sheets and the apparent absence of a random-coil structure suggest that ATP accelerates the aggregation process, which is in a good agreement with previous studies in which organic anions were found to enhance the formation of amyloid fibrils [16]. When, after a five minute incubation time of KFM6 with ATP, PKA is added, the concentration of β-sheets decreases and the minimum at 200 nm for random structures strongly increases (Figure 4b). This conformational transition shows a strong directing effect of enzymatic phosphorylation in the early stages of aggregate growth, during which small oligomers and protofilaments prevail. If the peptide was incubated with PKA for two hours (Figure 4d), the soluble peptide concentration decreases, what could be an effect of aggregation. After two hours and induced phosphorylation, the peptide concentration and structure seems to be stabilized over at least seven days. The inverse experiment, with ATP incubation for two hours, results in an intensive increase of β-sheet structures in KFM6 (Figure 4c). After induced phosphorylation, the concentration of β-sheets decreased in good agreement with the previous observation. To further clarify the conformational shift the 200/218 nm ratio of some CD experiments was depicted in Supporting Information File 1 (Figure S3).

2465

45

Figure 4: a) Enzymatic phosphorylation of 15 µM KFM6 in time-dependent CD-minimum plots. Addition of both 200 µM ATP and 5000 U PKA at time of dissolution of peptide; b) ATP-induced β-sheet formation, reversed by the addition of PKA after five minutes of incubation; c) the influence of ATP during the first two hours of conformational change and effect of two hours delayed phosphorylation; and d) the inverse experiment with PKA.

Chemical phosphorylation

To demonstrate the directing effect of the phosphate group towards soluble random structures, chemically phosphorylated P_CKFM6 was synthesized for control experiments. The peptide was dissolved in HFIP stock solution, as was unphosphorylated KFM6. For CD experiments P_CKFM6 was prepared in two-fold higher concentration (30 µM) compared to KFM6, with the intention to make it more prone to aggregation. However, only random structures were observed after seven days (Figure 3b). A TEM was taken 24 hours after dissolving and incubating the peptide. P_CKFM6 shows no morphologies of amyloid fibrils, its preliminary states or other higher order oligomers. Also no precipitation was observed. These results indicate high stability of the chemically phosphorylated peptide in comparison to the unphosphorylated analog.

Discussion

In the accepted model pathway of amyloid formation, a lag phase populated by functional soluble peptide structures is followed by partial unfolding, then a conformational transition into β-sheet monomers follows and these assemble into oligomers and protofilaments which then form amyloid fibrils [26,52]. While most inhibition studies concentrate on the later phase of this process, the results we present here suggest a strong directing effect to random structures during the initial phase of

the aggregation process concomitant with enzymatic phosphorylation. Phosphorylation experiments at different time points of the aggregation process suggest a time-dependent inhibitory effect, which becomes less efficient with increasing concentration of amyloid fibrils. The impact of phosphorylation must be considered with respect to three properties: steric bulk, charge and hydrophobicity. It was calculated that phosphorylation increases the van der Waals volume of the serine side chain by a factor of two [27] and it seems natural that flexible structures and solvent exposed positions should be preferred for both the interaction of the recognition site with the enzyme and the resulting, bulky phosphate group. Nevertheless, comparative studies between phosphate and β-galactose, incorporated in an amyloid forming coiled-coil peptide, showed no significant structural influence of the three times larger sugar residue, whereas the phosphate group directed the peptide into unfolded structures [27]. Additionally, the phosphate group introduces a formal charge of −2 to the peptide at pH 7.5. This leads to intramolecular Coulombic repulsions between negatively charged glutamate residues and the phosphate, or in electrostatic pairing with positively charged lysine or arginine residues, resulting in the perturbation of higher ordered secondary structures [53]. The impact of electrostatics on peptide and protein folding and self-assembly has been studied extensively [12,53], and it is known that the impact of the phosphate group charge strongly

2466

depends on the pH and the neighboring residues in the peptide [53,54]. In contrast to the accepted model where negatively charged phosphates must be close to the N-terminus to have a stabilizing effect on the original secondary structure [55], we demonstrate that phosphorylation proximal to the C-terminus can also significantly inhibit amyloid formation.

Delayed ATP or PKA addition lead in all cases to a stop of the amyloid formation and to stabilization of remaining soluble structures, as no further precipitation was observed. It remains unclear whether the decreasing effect of the phosphorylation during later stages of the aggregation is due to the lower flexibility of the recognition site, leading to slow reaction kinetics of the enzyme or to the competitive effects of nucleation-induced aggregation and phosphorylation. ^{31}P NMR experiments displaying the enzymatic phosphorylation over 12 hours with measurements every 30 minutes suggest rapid phosphorylation, as no change in the integral ratio of β-phosphate to α-phosphate of ATP was observed subsequent to the first time point (Supporting Information File 1, Figure S4). These observations support the theory that fully formed aggregates in a protofilamental or amyloidal stage cannot be accessed by the enzyme. This could be either a consequence of the low solubility or the tight cross-β-structure of the amyloid fibrils [56]. Furthermore, the presence of amyloid fibrils is known to have a strong promoting effect on the aggregation process [57,58], whereby nucleation, together with the unfavorable structure could conceal the effect of phosphorylation. Thus, a further proof was the chemically phosphorylated peptide P_CKFM6 which completely lost its ability to undergo a conformational transition and remains random coil even at higher concentrations.

The conformational transition to the β-sheet-rich structures is provided by the incorporation of hydrophobic domains, which can also decrease the overall solubility of the peptide. Calculated secondary structure propensities, using the TANGO algorithm, identify a nine amino acid long β-aggregation domain (E-L-V-V-L-K-S-E-L-) for KFM6 (Supporting Information File 1, Figure S5). Those aggregation domains are found in almost every peptide with the propensity to form amyloids [59,60]. The recognition site for PKA is located next to the aggregation domain of KFM6, with just three amino acids distance to the phosphorylated serine residue. The short distance to this hydrophobic aggregation domain could enable an interaction with the polar phosphate group, resulting in the suppression of the β-aggregation propensity. This is in accordance with simulations carried out by Rousseau et al. who proposed that charged amino acid residues flanking aggregating peptide segments could act as gatekeeper residues that reduce the aggregation propensity of the peptide [61-63]. Nevertheless, the change of the serine in position 23 of KFM6 to a glutamic acid does not

influence the calculated propensities to form β-aggregates as the phosphorylation does (Supporting Information File 1, Figure S5). It has also been reported that not only the charge, but also the position of the phosphate group can influence the amyloid formation and even the morphology of the formed fibrils [54,64]. Determining the relative contributions of the above factors to the role of phosphorylation in the amyloid forming pathway is a major challenge, however, crucial for the molecular understanding of the process.

Conclusion

Post-translational modifications like phosphorylation have become the most promising approach to examine and control the pathway of amyloid fibrillization. For a better understanding of the challenges of phosphorylation in this context, we designed a model peptide which undergoes a conformational transition from soluble structures to amyloids under physiological conditions and additionally presents a recognition site for PKA. During the initial growth phase, in which small oligomers and protofilaments are assumed to prevail, our results demonstrate a remarkable directing effect towards unfolded structures, if the peptide was phosphorylated close to the hydrophobic aggregation domain of the peptide. However, the enzymatic phosphorylation had to be realised on flexible secondary structures and couldn't dissipate grown aggregates in our experiments. The ability to identify those influences is a major challenge to use enzymatic phosphorylation as a tool to control the aggregation process and for the development of new functional biomaterials based on amyloid morphologies.

Experimental

Peptide synthesis, purification and characterization

Peptides were synthesized manually according to standard Fmoc chemistry using preloaded Fmoc-Leu-NovaSyn® TGA resin (0.3 mmol g^{-1}, Novabiochem). Standard couplings were performed in DMF with Fmoc-amino acids and TBTU (O-(benzotriazole-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate)/HOBt (1-hydroxybenzotriazole)/DIC (*N,N'*-diisopropylcarbodiimide) in eight-fold and NaClO$_4$ in ten-fold excess with double couplings of one hour coupling time. Fmoc-Ser(PO(OBzl)OH)-OH was activated with HATU (O-(7-azabenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate)/HOBt five-fold and fifteen-fold excess of DIPEA (*N,N*-diisopropylethylamine) with respect to the resin and two hour double couplings. A mixture of DBU (1,8-diazabicyclo[5,4,0]undec-7-ene) and piperidine (2% each) in DMF was used for Fmoc-deprotection (3 × 10 min). The resin was washed between each step with DMF and DCM (3 × 6 mL each). Peptides were cleaved from the resin by treatment with 2 mL TFA (trifluoroacetic acid iPr₃SiH/H$_2$O (90:9:1)) for three

hours. The resin was washed twice with 1 mL TFA and DCM, and excess of solvent was removed by evaporation. The peptides were precipitated with cold Et_2O, pelleted by centrifugation and dried by lyophilization. All peptides were purified with preparative reversed-phase HPLC by using a Knauer Smartline system (Knauer, Berlin, Germany) equipped with a Luna C8 column (10 u, 250 × 21.20 mm, Phenomenex, Torrance, CA, USA) running with a ACN/H_2O + 0.1% TFA gradient at a flow rate of 20 mL min^{-1} flow. The crude peptide was dissolved in 5 mL ACN/H_2O/DMSO (1:1:1). The collected fractions were evaporated and the purified peptide was dissolved in 2 mL of water and lyophilized to give the peptide as a white powder. Pure peptides were characterized by means of analytical HPLC and ESI–ToF mass spectrometry. Analytical HPLC was carried out with a VWR-Hitachi Elite LaCrome system (VWR, Darmstadt, Germany) using a Luna C8 column (5 u, 250 × 4.6 mm, Phenomenex, Torrance, CA, USA). The mass to charge ratios were determined with an Agilent 6210 ESI–ToF (Agilent Technologies, Santa Clara, CA, USA). Peptide solutions were injected by a syringe pump with a flow rate of 20 μL min^{-1}. Spray voltage was set to 4000 V, drying gas flow rate was 5 L min^{-1} and gas temperature was set to 300 °C. Both retention times and peptide masses are given in Supporting Information File 1 (Table S1).

Concentration determination

A stock solution was prepared by dissolving the purified, Abz labeled peptide in HFIP ($\approx$1 mg mL^{-1}) and sonicating for 15 minutes to dissolve all aggregates. 50 μL of this solution were aliquoted and dried under nitrogen flow, before the residue was dissolved in 1 mL 50 mM Tris/HCl buffer containing 10 mM $MgCl_2$ at pH 7.5. UV spectra were recorded in a 1 cm path length cuvette using a Cary 50 UV–vis spectrophotometer (Varian, Palo Alto, CA, USA) and the absorbance maximum at 312 nm was compared to a standard curve of the dipeptide H_2N-Abz-Gly-OH·HCl to calculate the concentration of the peptide stock solution. The stock solution was stored at −20 °C.

Circular dichroism

CD spectra were recorded by using a Jasco J-810 spectropolarimeter (Jasco, Gross-Umstadt, Germany) at 24 °C (Jasco PTC-348W1 peltier thermostat) using 0.2 mm path length Quartz Suprasil cuvettes (Hellma, Müllheim, Germany). After background correction, the spectra were averaged over three scans (λ = 195–240 nm; 0.5 nm intervals; 2 mm bandwidth; 4 s response time, 100 nm min^{-1} scanning speed). Ellipticity was normalized to concentration (c [mol L^{-1}]), number of residues (n = 27, including the N-terminal Abz group) and path length (l [cm]) by using Equation 1 in which Θ_{obs} is the measured ellipticity in millidegrees and Θ is the mean residue ellipticity in 10^3 mdeg cm^2 $dmol^{-1}$ $residue^{-1}$.

$$\Theta = \Theta_{obs} / (1000 \times 1 \times c \times n) \qquad (1)$$

Aliquots of the peptide stock solution were dried under nitrogen flow and, immediately before measurement, dissolved in 350 μL 50 mM Tris/HCl buffer with 10 mM $MgCl_2$, including 5000 U PKA and 200 μM ATP for the enzymatic phosphorylation studies. The pH was adjusted to 7.5 with 1 M NaOH. For time-dependent phosphorylation experiments, the buffer was prepared with just one of the phosphorylation components, while the other one was added at different time points.

Thioflavin T fluorescence spectroscopy

Fluorescence spectra were recorded by using a 1 cm path length quartz cuvette (Hellma, Müllheim, Germany) and a luminescence spectrometer LS50B (Perkin-Elmer, Boston, MA, USA). Spectra were recorded at room temperature from 470−500 nm after excitation at 450 nm (excitation slit width 5 nm; emission slit width 20 nm; scan speed = 500 nm min^{-1}; accumulations = 5). Sample preparation was the same as for CD with a volume of 500 μL buffer including 20 μM ThT. The fluorescence intensity at 485 nm was recorded at different time points over a total time period of 24 hours and normalized to the starting value at t = 5 min for one. The shown plots represent an average of three independent measurements.

Transmission electron microscopy

Peptides, prepared as for CD spectroscopy measurements, were examined after 24 hours. Aliquots (6 μL) of the corresponding solution were placed for 60 seconds onto glow-discharged (60 s plasma treatment at 8 W in BAL-TEC MED 020), carbon-coated collodium films on 400-mesh copper grids (Leica Microsystems, Wetzlar, Germany). After blotting and negative staining with 1% PTA, the grids were left to air-dry. The TEM images were recorded with a Philips CM12 transmission electron microscope (FEI, Oregon, USA) at 100 kV acceleration voltage and at a primary magnification of 58000× on Kodak SO-163 negative film by using a defocus of 900 nm.

Supporting Information

Supporting Information File 1

Description of further methods and additional figures.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-12-240-S1.pdf]

References

1. Chiti, F.; Dobson, C. M. *Annu. Rev. Biochem.* **2006,** *75,* 333–366. doi:10.1146/annurev.biochem.75.101304.123901
2. Cherny, I.; Gazit, E. *Angew. Chem., Int. Ed.* **2008,** *47,* 4062–4069. doi:10.1002/anie.200703133
3. DeToma, A. S.; Salamekh, S.; Ramamoorthy, A.; Lim, M. H. *Chem. Soc. Rev.* **2012,** *41,* 608–621. doi:10.1039/C1CS15112F
4. Seeman, P.; Seeman, N. *Synapse* **2011,** *65,* 1289–1297. doi:10.1002/syn.20957
5. Westermark, P. *Upsala J. Med. Sci.* **2011,** *116,* 81–89. doi:10.3109/03009734.2011.573884
6. Vandenberghe, R. *Curr. Neurol. Neurosci. Rep.* **2014,** *14,* No. 498. doi:10.1007/s11910-014-0498-9
7. Baldwin, A. J.; Bader, R.; Christodoulou, J.; MacPhee, C. E.; Dobson, C. M.; Barker, P. D. *J. Am. Chem. Soc.* **2006,** *128,* 2162–2163. doi:10.1021/ja0565673
8. Perutz, M. F.; Finch, J. T.; Berriman, J.; Lesk, A. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99,* 5591–5595. doi:10.1073/pnas.042681399
9. Yang, F.; Lim, G. P.; Begum, A. N.; Ubeda, O. J.; Simmons, M. R.; Ambegaokar, S. S.; Chen, P. P.; Kayed, R.; Glabe, C. G.; Frautschy, S. A.; Cole, G. M. *J. Biol. Chem.* **2005,** *280,* 5892–5901. doi:10.1074/jbc.M404751200
10. Maji, S. K.; Perrin, M. H.; Sawaya, M. R.; Jessberger, S.; Vadodaria, K.; Rissman, R. A.; Singru, P. S.; Nilsson, K. P. R.; Simon, R.; Schubert, D.; Eisenberg, D.; Rivier, J.; Sawchenko, P.; Vale, W.; Riek, R. *Science* **2009,** *325,* 328–332. doi:10.1126/science.1173155
11. Fowler, D. M.; Koulov, A. V.; Alory-Jost, C.; Marks, M. S.; Balch, W. E.; Kelly, J. W. *PLoS Biol.* **2006,** *4,* e6, 100–107. doi:10.1371/journal.pbio.0040006
12. Pagel, K.; Wagner, S. C.; Araghi, R. R.; von Berlepsch, H.; Böttcher, C.; Koksch, B. *Chem. – Eur. J.* **2008,** *14,* 11442–11451. doi:10.1002/chem.200801206
13. Aggeli, A.; Bell, M.; Carrick, L. M.; Fishwick, C. W. G.; Harding, R.; Mawer, P. J.; Radford, S. E.; Strong, A. E.; Boden, N. *J. Am. Chem. Soc.* **2003,** *125,* 9619–9628. doi:10.1021/ja021047i
14. Bieschke, J.; Zhang, Q.; Powers, E. T.; Lerner, R. A.; Kelly, J. W. *Biochemistry* **2005,** *44,* 4977–4983. doi:10.1021/bi0501030
15. Klajnert, B.; Cladera, J.; Bryszewska, M. *Biomacromolecules* **2006,** *7,* 2186–2191. doi:10.1021/bm060229s
16. Lendel, C.; Bolognesi, B.; Wahlström, A.; Dobson, C. M.; Gräslund, A. *Biochemistry* **2010,** *49,* 1358–1360. doi:10.1021/bi902005t
17. Cerasoli, E.; Sharpe, B. K.; Woolfson, D. N. *J. Am. Chem. Soc.* **2005,** *127,* 15008–15009. doi:10.1021/ja0543604
18. Karr, J. W.; Kaupp, L. J.; Szalai, V. A. *J. Am. Chem. Soc.* **2004,** *126,* 13534–13538. doi:10.1021/ja0488028
19. Pagel, K.; Koksch, B. *Curr. Opin. Chem. Biol.* **2008,** *12,* 730–739. doi:10.1016/j.cbpa.2008.09.005
20. Scrocchi, L. A.; Chen, Y.; Waschuk, S.; Wang, F.; Cheung, S.; Darabie, A. A.; McLaurin, J.; Fraser, P. E. *J. Mol. Biol.* **2002,** *318,* 697–706. doi:10.1016/S0022-2836(02)00164-X
21. Tenidis, K.; Waldner, M.; Bernhagen, J.; Fischle, W.; Bergmann, M.; Weber, M.; Merkle, M.-L.; Voelter, W.; Brunner, H.; Kapurniotu, A. *J. Mol. Biol.* **2000,** *295,* 1055–1071. doi:10.1006/jmbi.1999.3422
22. Balbach, J. J.; Ishii, Y.; Antzutkin, O. N.; Leapman, R. D.; Rizzo, N. W.; Dyda, F.; Reed, J.; Tycko, R. *Biochemistry* **2000,** *39,* 13748–13759. doi:10.1021/bi0011330
23. Melquiond, A.; Gelly, J.-C.; Mousseau, N.; Derreumaux, P. *J. Chem. Phys.* **2007,** *126,* 065101. doi:10.1063/1.2435358
24. Elgersma, R. C.; van Dijk, M.; Dechesne, A. C.; van Nostrum, C. F.; Hennink, W. E.; Rijkersa, D. T. S.; Liskamp, R. M. J. *Org. Biomol. Chem.* **2009,** *7,* 4517–4525. doi:10.1039/b912851d
25. Chaudhary, N.; Nagaraj, R. *J. Pept. Sci.* **2011,** *17,* 115–123. doi:10.1002/psc.1339
26. Hackenberger, C. P. R.; Schwarzer, D. *Angew. Chem., Int. Ed.* **2008,** *47,* 10030–10074. doi:10.1002/anie.200801313
27. Broncel, M.; Falenski, J. A.; Wagner, S. C.; Hackenberger, C. P. R.; Koksch, B. *Chem. – Eur. J.* **2010,** *16,* 7881–7888. doi:10.1002/chem.200902452
28. Artner, L. M.; Merkel, L.; Bohlke, N.; Beceren-Braun, F.; Weise, C.; Dernedde, J.; Budisa, N.; Hackenberger, C. P. R. *Chem. Commun.* **2012,** *48,* 522–524. doi:10.1039/C1CC16039G
29. Shumyantseva, V. V.; Suprun, E. V.; Bulko, T. V.; Archakov, A. I. *Biosens. Bioelectron.* **2014,** *61,* 131–139. doi:10.1016/j.bios.2014.05.001
30. Nakamura, K.; Shima, H.; Watanabe, M.; Haneji, T.; Kikuchi, K. *Biochem. J.* **1999,** *344,* 819–825. doi:10.1042/bj3440819
31. Auch, C. J.; Saha, R. N.; Sheikh, F. G.; Liu, X.; Jacobs, B. L.; Pahan, K. *FEBS Lett.* **2004,** *563,* 223–228. doi:10.1016/S0014-5793(04)00302-3
32. Hutton, J. C.; Eisenbarth, G. S. *Proc. Natl. Acad. Sci. U. S. A.* **2003,** *100,* 8626–8628. doi:10.1073/pnas.1633447100
33. Eidenmüller, J.; Fath, T.; Hellwig, A.; Reed, J.; Sontag, E.; Brandt, R. *Biochemistry* **2000,** *39,* 13166–13175. doi:10.1021/bi001290z
34. Beharry, C.; Cohen, L. S.; Di, J.; Ibrahim, K.; Briffa-Mirabella, S.; Alonso, A. del C. *Neurosci. Bull.* **2014,** *30,* 346–358. doi:10.1007/s12264-013-1414-z
35. Russell, C. L.; Koncarevic, S.; Ward, M. A. *J. Alzheimer's Dis.* **2014,** *41,* 345–364. doi:10.3233/JAD-132312
36. Gandy, S.; Czernik, A. J.; Greengard, P. *Proc. Natl. Acad. Sci. U. S. A.* **1988,** *85,* 6218–6221. doi:10.1073/pnas.85.16.6218
37. Hung, A. Y.; Haass, C.; Nitsch, R. M.; Qiu, W. Q.; Citron, M.; Wurtman, R. J.; Growdon, J. H.; Selkoe, D. J. *J. Biol. Chem.* **1993,** *268,* 22959–22962.
38. Chen, Y.-X.; Du, J.-T.; Zhou, L.-X.; Liu, X.-H.; Zhao, Y.-F.; Nakanishi, H.; Li, Y.-M. *Chem. Biol.* **2006,** *13,* 937–944. doi:10.1016/j.chembiol.2006.06.017
39. Liang, Z.; Liu, F.; Grundke-Iqbal, I.; Iqbal, K.; Gong, C.-X. *J. Neurochem.* **2007,** *103,* 2462–2470. doi:10.1111/j.1471-4159.2007.04942.x
40. Liu, X.-A.; Zhu, L.-Q.; Zhang, Q.; Shi, H.-R.; Wang, S.-H.; Wang, Q.; Wang, J.-Z. *Neurochem. Res.* **2008,** *33,* 1811–1820. doi:10.1007/s11064-008-9638-4
41. Gerling, U. I. M.; Brandenburg, E.; von Berlepsch, H.; Pagel, K.; Koksch, B. *Biomacromolecules* **2011,** *12,* 2988–2996. doi:10.1021/bm200587m
42. Broncel, M.; Wagner, S. C.; Hackenberger, C. P. R.; Koksch, B. *Chem. Commun.* **2010,** *46,* 3080–3082. doi:10.1039/c001460e
43. Hirsch, A. K. H.; Fischer, F. R.; Diederich, F. *Angew. Chem., Int. Ed.* **2007,** *46,* 338–352. doi:10.1002/anie.200603420
44. Yang, Z.; Liang, G.; Wang, L.; Xu, B. *J. Am. Chem. Soc.* **2006,** *128,* 3038–3043. doi:10.1021/ja057412y
45. Zetterqvist, Ö.; Ragnarsson, U. *FEBS Lett.* **1982,** *139,* 287–290. doi:10.1016/0014-5793(82)80872-7

46. Kreegipuu, A.; Blom, N.; Brunak, S.; Järv, J. *FEBS Lett.* **1998,** *430,* 45–50. doi:10.1016/S0014-5793(98)00503-1

47. Pagel, K.; Wagner, S. C.; Samedov, K.; von Berlepsch, H.; Bottcher, C.; Koksch, B. *J. Am. Chem. Soc.* **2006,** *128,* 2196–2197. doi:10.1021/ja057450h

48. Greenfield, N.; Fasman, G. D. *Biochemistry* **1969,** *8,* 4108–4116. doi:10.1021/bi00838a031

49. Levine, H., III. *Protein Sci.* **1993,** *2,* 404–410. doi:10.1002/pro.5560020312

50. Meinkoth, J. L.; Alberts, A. S.; Went, W.; Fantozzi, D.; Taylor, S. S.; Hagiwara, M.; Montminy, M.; Feramisco, J. R. *Mol. Cell. Biochem.* **1993,** *127,* 179–186. doi:10.1007/BF01076769

51. Kim, C.; Cheng, C. Y.; Saldanha, S. A.; Taylor, S. S. *Cell* **2007,** *130,* 1032–1043. doi:10.1016/j.cell.2007.07.018

52. Pagel, K.; Vagt, T.; Koksch, B. *Org. Biomol. Chem.* **2005,** *3,* 3843–3850. doi:10.1039/b510098d

53. Inoue, M.; Hirata, A.; Tainaka, K.; Morii, T.; Konno, T. *Biochemistry* **2008,** *47,* 11847–11857. doi:10.1021/bi8010994

54. Valette, N. M.; Radford, S. E.; Harris, S. A.; Warriner, S. L. *ChemBioChem* **2012,** *13,* 271–281. doi:10.1002/cbic.201100607

55. Inoue, M.; Konno, T.; Tainaka, K.; Nakata, E.; Yoshida, H.; Morii, T. *Biochemistry* **2012,** *51,* 1396–1406. doi:10.1021/bi201451z

56. Hamley, I. W. *Angew. Chem., Int. Ed.* **2007,** *46,* 8128–8147. doi:10.1002/anie.200700861

57. Bernhardt, N. A.; Berhanu, W. M.; Hansmann, U. H. E. *J. Phys. Chem. B* **2013,** *117,* 16076–16085. doi:10.1021/jp409777p

58. O'Nuallain, B.; Williams, A. D.; Westermark, P.; Wetzel, R. *J. Biol. Chem.* **2004,** *279,* 17490–17499. doi:10.1074/jbc.M311300200

59. Fernandez-Escamilla, A.-M.; Rousseau, F.; Schymkowitz, J.; Serrano, L. *Nat. Biotechnol.* **2004,** *22,* 1302–1306. doi:10.1038/nbt1012

60. Azriel, R.; Gazit, E. *J. Biol. Chem.* **2001,** *276,* 34156–34161. doi:10.1074/jbc.M102883200

61. Rousseau, F.; Schymkowitz, J.; Serrano, L. *Curr. Opin. Struct. Biol.* **2006,** *16,* 118–126. doi:10.1016/j.sbi.2006.01.011

62. Rousseau, F.; Serrano, L.; Schymkowitz, J. W. H. *J. Mol. Biol.* **2006,** *355,* 1037–1047. doi:10.1016/j.jmb.2005.11.035

63. Pedersen, J. S.; Christensen, G.; Otzen, D. E. *J. Mol. Biol.* **2004,** *341,* 575–588. doi:10.1016/j.jmb.2004.06.020

64. Schneider, A.; Biernat, J.; von Bergen, M.; Mandelkow, E.; Mandelkow, E.-M. *Biochemistry* **1999,** *38,* 3549–3558. doi:10.1021/bi981874p

Supporting Information

for

Inhibition of peptide aggregation by means of

enzymatic phosphorylation

Kristin Folmert[1], Malgorzata Broncel[2], Hans v. Berlepsch[1], Christopher H. Ullrich[3],

Mary-Ann Siegert[4] and Beate Koksch[1]*

Address: [1]Department of Chemistry and Biochemistry, Freie Universität Berlin, Takustr.

3, 14195 Berlin, Germany, [2]The Institute of Cancer Research, 237 Fulham Road,

London SW3 6JB, UK, [3]Deutsche Bahn, Umweltservice, Bahntechnikerring 74, 14774

Kirchmöser, Germany and [4]Department of Organic Chemistry, Technische Universität

Berlin, Strasse des 17. Juni 124, 10623 Berlin, Germany

Email: Beate Koksch - Beate.Koksch@fu-berlin.de

* Corresponding author

Description of further methods and additional figures

S1

51

1. Materials

Fmoc-L-amino acids were purchased from ORPEGEN Peptide Chemicals GmbH (Heidelberg, Germany). Fmoc-Ser(PO(OBzl)OH)-OH and Boc-Abz-OH were obtained from Bachem (Weil am Rhein, Germany). Fmoc-Leu Nova Syn®TGA-resin with 0.3 mmol/g loading was from the company NovaBiochem (Wolfshagen, Germany). PKA and ATP were purchased from New England BioLabs (Frankfurt am Main, Germany). All solvents and common chemicals were used from Acros Organics (Geel, Germany) without further purification.

2. Peptide characterization

ESI-ToF

Mass-to-charge ratios were determined with an Agilent 6210 ESI-ToF (Agilent Technologies, Santa Clara, CA, USA). Peptides were dissolved in ACN/H_2O (1:1) and the peptide solutions were injected by a syringe pump with a flow rate of 10 µL min^{-1}. Spray voltage was set to 4000 V, drying gas flow rate was 5 L min^{-1} and gas temperature was set to 300 °C.

HPLC

Analytical HPLC was carried out with a VWR-Hitachi Elite LaCrome system (VWR, Darmstadt, Germany) and a Luna C8 column (5 u, 250 × 4.6 mm, Phenomenex, Torrance, CA, USA) was used for data analysis. LaChrom-software (Version 3.0, Merck) was used to analyze data. The gradient was a mixture of H_2O and ACN with 0.1% TFA from 5 to 70% ACN over 30 minutes followed by a washing and calibration time. Peptides were detected at a wavelength of 220 nm. Phosphorylated peptide PcKFM6 is degrading under HPLC conditions. The products of degradation, KFM6 and polyphosphoric acid, appear with the same retention time.

Table S1: Determined and calculated mass to charge ratios and retention times on analytical HPLC of the peptides.

peptide	charge	calc. mass (*m/z*)	det. mass (*m/z*)	t_R (min)
KFM6	0	3980.9592	–	
	+2	1591.9174	1592.4449	
	+3	1061.6142	1061.9654	23.48
	+4	796.4626	796.7276	
P$_C$KFM6	0	4077.9292	–	
	+2	1640.9174	1640.9378	
	+3	1094.2808	1094.4539	27.15
	+4	820.9626	820.9653	

S2

Figure S1: HPLC chromatogram of KFM6 (top) and PcKFM6 (below).

3. CD spectroscopic analysis

Time-dependent CD spectra of 15 µM KFM6 to which both 200 µM ATP and 5000 U PKA were added. CD spectra were recorded by using a Jasco J-810 spectropolarimeter (Jasco, Gross-Umstadt, Germany) at 24 °C (Jasco PTC-348W1 peltier thermostat) using 0.2 mm path length Quartz Suprasil cuvettes (Hellma, Müllheim, Germany). After background correction, the spectra were averaged over three scans (λ = 195–240 nm; 0.5 nm intervals; 2 mm bandwidth; 4 s response time, 100 nm min^{-1} scanning speed).
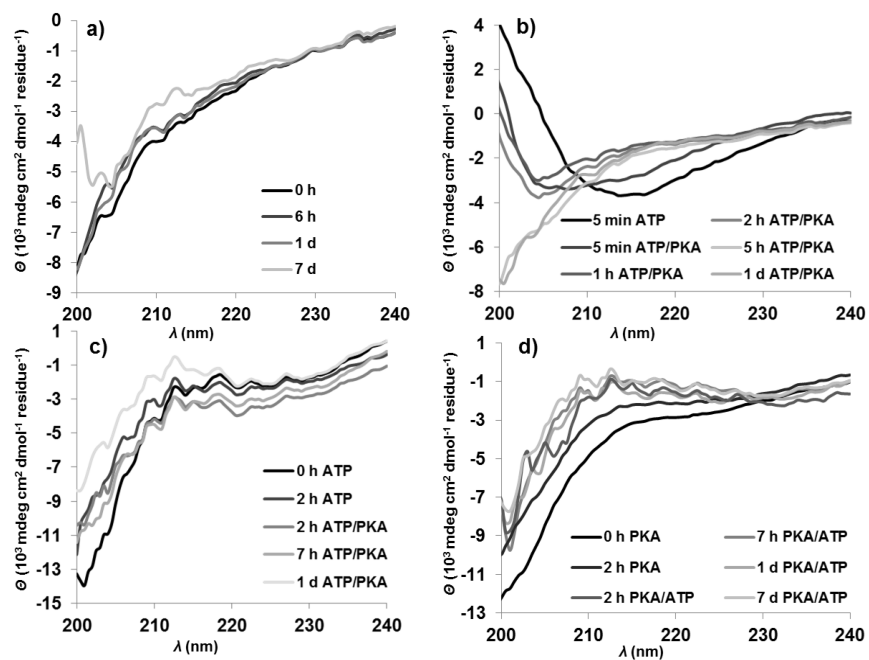
Figure S2: CD spectra of enzymatic phosphorylation of 15 µM KFM6 in time-dependent CD-minimum plots. Addition of both 200 µM ATP and 5000 U PKA at time of dissolution of peptide (a); ATP-induced β-sheet formation, reversed by the addition of PKA after five minutes of incubation (b); the influence of ATP during the first two hours of conformational change and effect of two hours delayed phosphorylation (c); and the inverse experiment with PKA (d).

S3
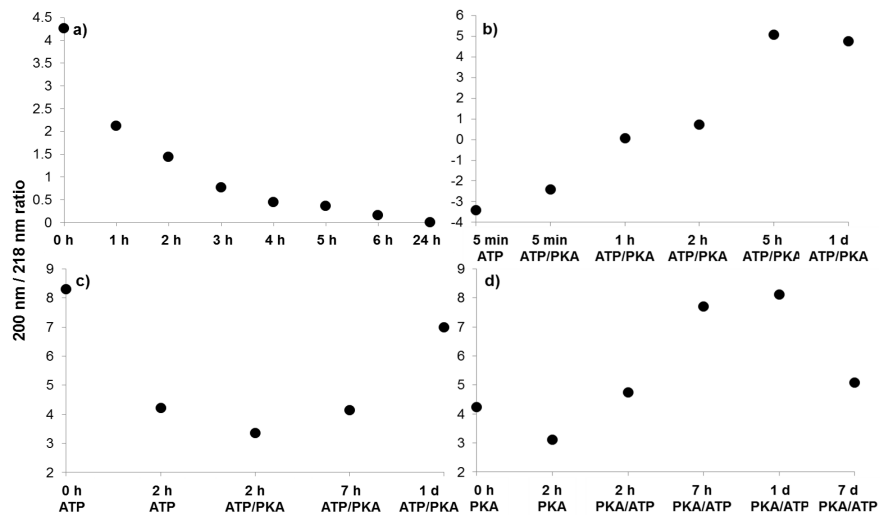
Figure S3: Ratio of 200 nm to 218 nm plotted from CD spectra of enzymatic phosphorylation of 15 µM KFM6 in time-dependent CD-minimum plots. Time-dependent transition from random-coil to β-sheets of plain KFM6 (a); ATP-induced β-sheet formation, reversed by the addition of PKA after five minutes of incubation (b); the influence of ATP during the first two hours of conformational change and effect of two hours delayed phosphorylation (c); and the inverse experiment with PKA (d). The 5 min ATP and 5 min ATP/PKA values of (b) are not divided by 200 nm and plot only 218 nm quantity because 200 nm value was zero.

4. ^{31}P NMR

22 μM of peptide KFM6 was incubated with 200 μM ATP and 5000 U PKA in 50 mM Tris/HCl buffer with 10 mM MgCl$_2$ at pH 7.5 and 27 ± 3 °C. ATP transfers one phosphate during enzymatic phosphorylation to the peptide, leaving an ADP molecule. Compared to ATP, ADP has no β-phosphate with respect to the electronic properties, while the α-phosphate remains constant. The integral ratio from β-phosphate to α-phosphate was followed to determine the reaction ratio to ADP, which is linear proportional to the phosphorylation ratio of KFM6. The pointed data were accumulated for 25 min each. The resulting linear fit curve implies no significant change in the integral ratio during the examined time dimension.

^{31}P NMR (202 MHz, PKA reaction buffer, 27 ± 3 °C, ppm) δ = 6.31 (d, γ-phosphate), 11.45 (d, α-phosphate), 19.66 (dd, β-phosphate).

Figure S4: Structure of ATP (left) and ^{31}P NMR experiment with the integrated ration of β- to α-phosphate as a reaction control for the kinetic of the enzymatic phosphorylation (right).

5. TANGO

The computer algorithm TANGO was used to predict the propensities for α-helices, β-sheets or β-aggregates like amyloids [S1-S3]. The calculations were performed for 293 K, at pH 7.5 and an ionic strength of 0.02. The resulting propensities are plotted over the amino acid sequence of KFM6.
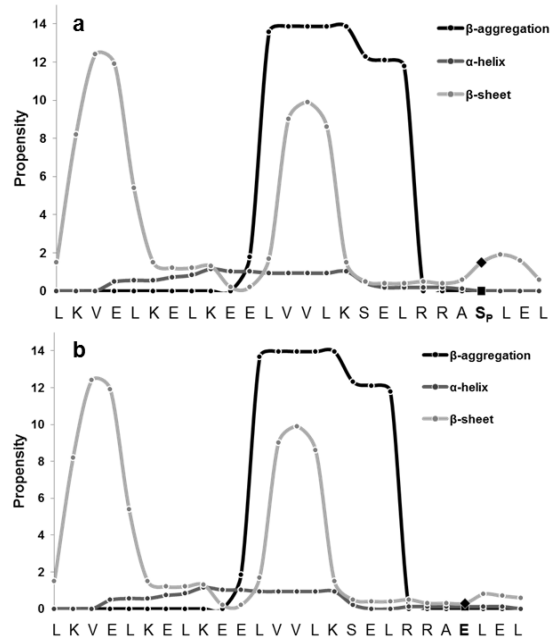
Figure S5: TANGO calculation for the secondary structure propensities of KFM6 with a) serine or b) glutamic acid in position 23.

References

S1. Rousseau, F., Schymkowitz, J., Serrano, L., *Curr. Opin. Struct. Biol.* **2006**, *16*, 118-126.
S2. Fernandez-Escamilla, A.-M., Rousseau, F., Schymkowitz, J., Serrano, L., *Nat. Biotech.* **2004**, *22*, 1302-1306.
S3. Linding, R., Schymkowitz, J., Rousseau, F., Diella, F., Serrano, L., *J. Mol. Biol.* **2004**, *342*, 345-353.

Part B: NFGAIL Amyloid Oligomers: The Onset of β-Sheet Formation and the Mechanism for Fibril Formation

The results presented in this section have been partially published as W. Hoffmann, **K. Folmert**, J. Moschner, X. Huang, H. v. Berlepsch, B. Koksch, M. T. Bowers, G. v. Helden and K. Pagel; NFGAIL Amyloid Oligomers: The Onset of Beta-Sheet Formation and the Mechanism for Fibril Formation, *J. Am. Chem. Soc.*, **2017**, accepted.

Individual contributions: **K. Folmert** synthesized and characterized NFGAIL, performed the ThT experiments, prepared the samples for TEM and SAXS measurements, and interpreted the results in the context of the project aim. **K. Folmert**, W. Hoffmann and J. Moschner had the initial idea for the project concept. W. Hoffmann wrote the paper and performed all experiments, which are not discussed in this thesis. H.v. Berlepsch and X. Huang provided the TEM images. B. Koksch, M. T. Bowers, G. v. Helden and K. Pagel are corresponding authors. SAXS data were provided and interpreted from C. Kästner and A. F. Thünemann.

B1 Abstract

Amyloid plaques, so-called islet amyloids are one major characteristic for T2D and were found in 95% of the islets of Langerhans in concerned patients.[166-169] The 37 residue peptide hIAPP is the main component in islet amyloids and the presumably soluble hIAPP oligomers were identified as one of the cytotoxic species in T2D pathology.[183] Fibril formation of hIAPP is driven from the hydrophobic patch region hIAPP(22-27) NFGAIL where the nucleation process is initiated.[196] Furthermore, the hexapeptide NFGAIL is the shortest segment of hIAPP which is still highly prone to amyloid formation.[41, 197] In addition, the short aggregating sequence has a cytotoxicity towards the pancreatic cell line RIN5fm comparable to hIAPP[207] and thus serves as a suitable model for the full length hIAPP. Several computational modeling studies have concentrated on the NFGAIL sequence and predicted a cross-β-pattern similar to hIAPP

fibrils[211-216] and a crucial impact of the hydrophobic phenylalanine residue was described[217]. Nevertheless, the simulations could not agree on the conformational patterns and the pathway of assembly during the early nucleation and oligomerization events. To investigate the contingent of β-sheets, turns and random structures during these critically oligomerization process ion mobility-mass spectrometry coupled to infrared spectroscopy was used resulting in the conclusion that during the lag phase highly polydisperse, polymorphic, and compact oligomers as well as extended intermediates are present. The compact conformations adopt mainly turn-like and random structures, whereas the extended oligomers include a significant amount of β-sheet content. ThT and SAXS assays and TEM images were used to analyze the aggregation behavior and quaternary morphology of the formed amyloid fibrils. In agreement with molecular dynamic simulations by Wu *et al.*[213] and others[221-222], the results suggest an aggregation mechanism where both species, unordered off-pathway and well-ordered partially β-sheet containing on-pathway oligomers, coexist.
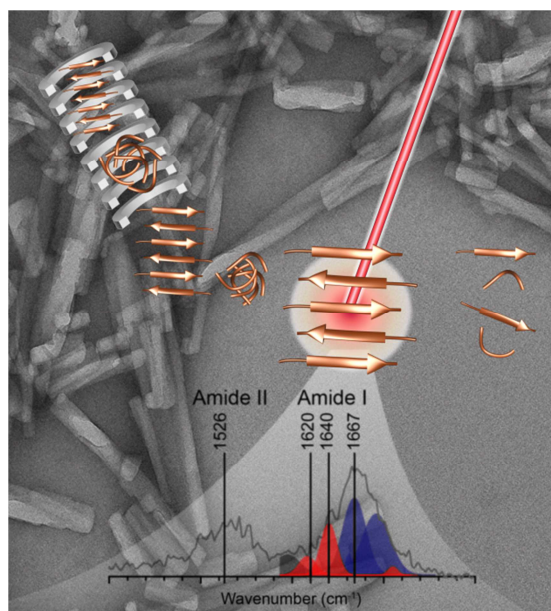
Figure B1: *Cartoon of the instrumental setup for the manuscript attached below. The different species of NFGAIL separated by the drift tube before measuring IM-MS and afterwards analyzed by gas-phase IR spectroscopy.*

B2 Results

B2.1 NFGAIL Synthesis and Characterization

The NFGAIL peptide was synthesized by SPPS using common Fmoc-strategy on a solid support in double hand-couplings. For high yields the full cleavage reaction mixture including TFA/ TIPS/ H_2O (90/9/1) was sonicated for three hours at 30°C. During workup the crude NFGAIL was not precipitated with Et_2O due to its solubility in inorganic solvents. After freeze-drying the crude product, NFGAIL was dissolved in pure methanol, without any sonication, filtration or contact with water to prevent fast aggregation. Before preparative HPLC, the solution of NFGAIL in methanol was centrifuged and decanted to remove debris and not dissolved particles. The peptide was characterized with analytical HPLC and ESI ToF MS and was shown to be at least more than 99% pure. The results presented here and in the manuscript are valid for all batches synthesized.

B2.2 NFGAIL Aggregation in Solution

Sample Preparation

A stock solution of NFGAIL in 1,1,1,3,3,3-hexafluoro-2-propanol [HFIP] (18 mM) was prepared and sonicated for 15 minutes to dissolve all preformed aggregates and to direct peptide monomers into disordered or helical conformations[145]. For aggregation studies aliquots of this stock solution were dried under nitrogen flow and then dissolved to a final concentration of 4 mM in fresh filtered ammonium acetate buffer [10 mM, pH 7.0] and vortexed for 30 seconds. For ThT assays the buffer contained 20 µM ThT. The reaction mixture was free of undissolved particles at the beginning of the assay. Between the measuring points the mixture was kept in an amber colored Eppendorf reactor to prevent bleaching of the dye and additionally incubated at a temperature of 37°C and shaken with 1300 rpm; as this procedures had previously been described to increase the aggregation rate[452].

CD Spectroscopy

CD measurements of NFGAIL were performed in varying concentrations and with different conditions, but due to the short sequence no significant signal could be observed.

ThT Fluorescence Assay

In **Figure B2** two identical ThT assays are depicted. The growth phase into mature fibrils was initiated after 20 hours, whereas insoluble fibrils were present after 22 hours. The increasing fluorescence signal is a strong indicator for the presence of amyloid fibrils.[118, 121] The overall growth process was complete after two hours. The formation of insoluble fibrils, which accumulate at the bottom or at the walls of the cuvette, led to a decrease of the measured fluorescence intensity at 485 nm. A long lag phase followed by a sigmoidal growing phase is an indicator for a nucleation dependent on-pathway aggregation process.[206] During the lag phase early soluble and polydisperse monomers are present until the nucleation event induces fibril growth.[129, 453] The two almost congruent curves emphasize the reproducibility of the behavior.
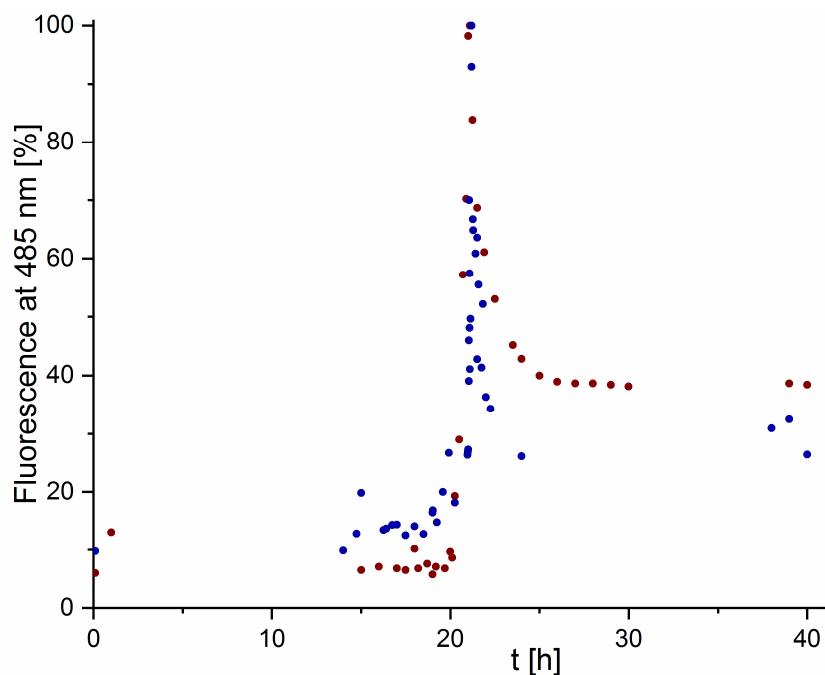
Figure B2: Two individual (blue and red symbols) and normalized Thioflavin T assays of a 4 mM NFGAIL sample incubated at 37°C and 1300 rpm in 10 mM ammonium acetate buffer (10 mM, pH 7.0).

TEM Images

The preparation for the TEM images was performed in an analogous way as the fluorescence samples and images were taken after 24 hours incubation time with 1% phosphotungstic acid [PTA] negative staining. The amyloid fibrils presented in **Figure B3** have a diameter of 20-30 nm and are only a few hundreds of nanometers in length. These findings are consistent with

previous results for NFGAIL from Tenidis *et al.*.[200] The observed discrete termination points of the fibrils indicate an antifacial orientation of the peptide chains with likely amphipathic character. As described for **Figure 4**, the symmetry classes 1, 3 or 5 can adopt a structure in which the twist results from a rotation about a 2_1 symmetry axis.[129] This interpretation is in good agreement with molecular simulations of NFGAILS of Guo *et al.*, who proposed an antiparallel orientation of the β-sheets and a twisted fibril pattern that can be categorized as belonging to mainly symmetry class 5.[221]
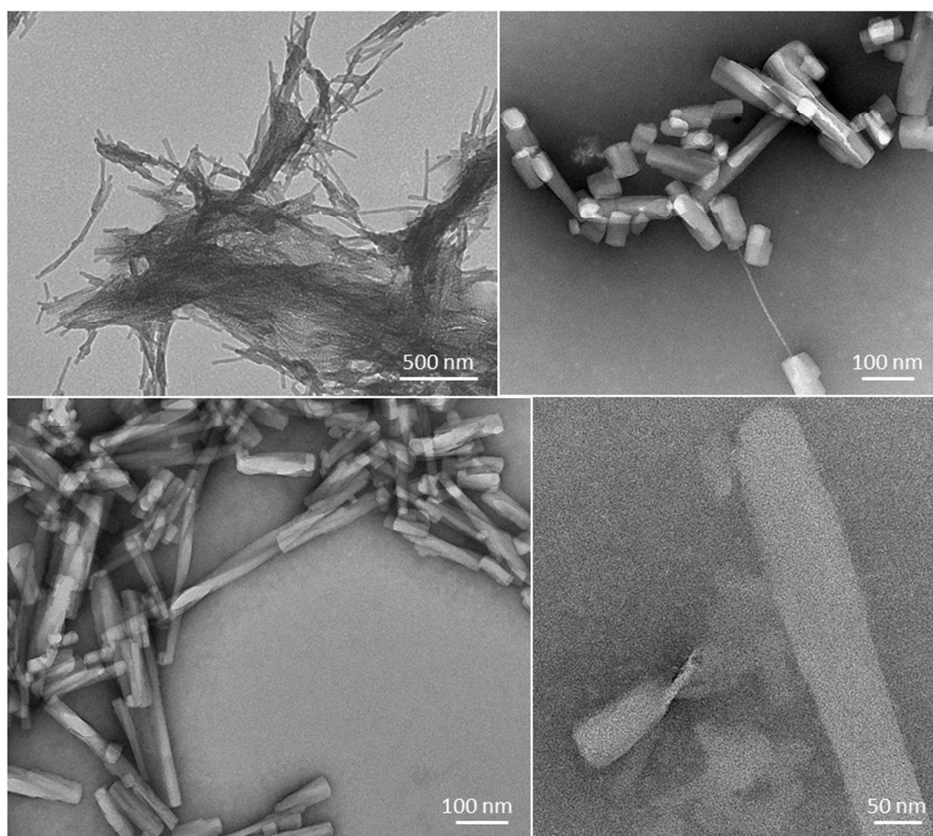
Figure B3: *Four different negative staining (1% PTA) TEM images of 4 mM NFGAIL after 24 h incubation in 10 mM ammonium acetate buffer (pH 7.0) at 37°C and 1300 rpm.*

Small-Angle X-ray Scattering [SAXS]

SAXS data from aggregated NFGAIL suspensions were collected in the size range of π/q_{max} = 0.5 nm to π/q_{min} = 45 nm. SAXS produces intrinsically representative statistical sample averages. The scattering vector is defined in terms of the scattering angle θ and the wavelength λ of the radiation, thus $q = 4\pi / \lambda \sin(\theta)$. The angle between the incident and the scattered beam is

2θ. Valuable information about the filament cross sections is obtainable from the data, whereas determination of the much larger total length of NFGAIL filaments proved not to be possible. Small-angle scattering patterns of fibers with the symmetry of long cylindrical rods of length a and radius R decay in proportion to q^{-1} at low q-values. Characteristic for the rod's radius is the steep bending of the scattering curve in the higher q-range. Fits of the data by using the cylindrical rod model[454] with a filament radius of 6.7 nm are in accordance with the data only for q-values larger than 0.3 nm^{-1} (blue dotted line **Figure B4**). This reveals that it is not possible to fit the SAXS pattern of NFGAIL sufficiently with the simple cylinder model. Instead of the expected q^{-1} scaling of a cylinder, the intensity scales with q^{-2} in the low q-range. This finding indicates a significantly more pronounced ribbon-shaped filament structure for NFGAIL. Ribbon shaped filaments can be modeled as parallelepipeds[455] of length a, width b and thickness c with $a > b > c$. The extended sheet scales with q^{-2}, which indicates that the extension of the ribbons in a lateral direction is larger than 45 nm, forming a sheet-like structure. The best fit of the data of aggregated NFGAIL results in a height of c = 10 nm and a width of b > 30 nm for the cross-section dimensions of the filament. The total fit curve of the parallelepiped is displayed in **Figure B4** as a red solid line. As already mentioned, the filament's length is beyond the maximum size accessible and was therefore held constant at a = 500 nm. Since the value of b is larger than 45 nm the extended sheet model with only the thickness of the sheets as fit parameter could be used alternatively, resulting in the same height of 10 nm (green curve in **Figure B4**). Additionally, model-free data evaluation methods like the cross-section Guinier law[456-457] are useful to check the consistency of the parameters derived from the model curves. Application of the cross-section Guinier law yields an estimate of c = 9.8 ± 0.2 nm. Further insight into the cross-section filament structure yields the pair-distance distribution function of the cross-section, $\rho_t(r)$.[457-459] The $\rho_t(r)$ can be interpreted as the electron density weighted number of all possible connections between points within the filament cross-section.
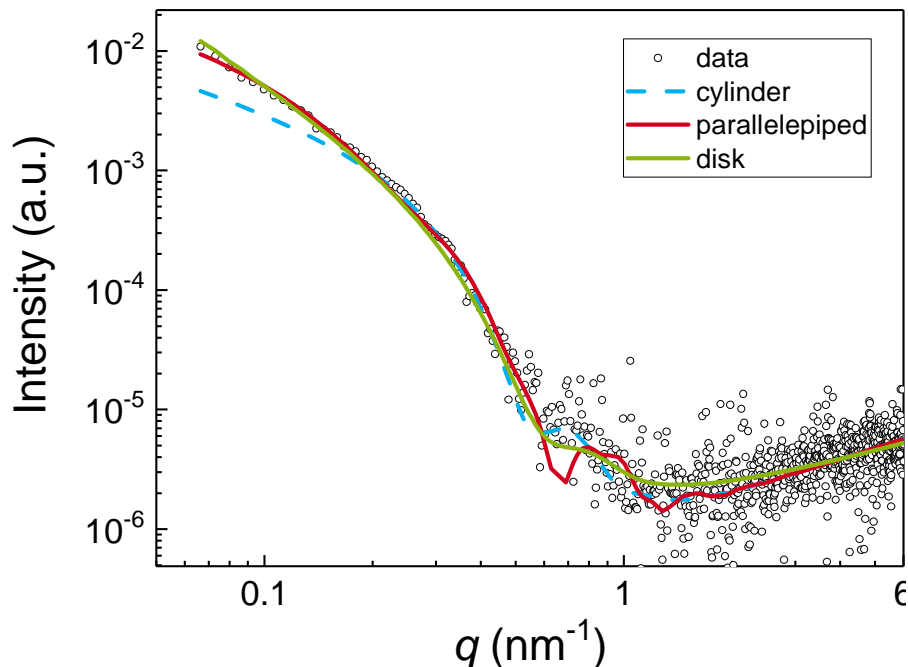
Figure B4: *Small-angle X-ray scattering data of NFGAIL (circles) and curve fits using a cylinder and parallelepiped model (blue dashed and red solid line, respectively). The best fit curves are from a cylinder function for NFGAIL with a radius of R = 6.7 nm and a parallelepiped with c = 10 nm and b > 45 nm. Length of cylinder and parallelepiped where held constant for fitting at 500 nm. The most simple model suitable for curve fitting is a ribbon with a thickness of 10.0 nm with a polydispersity of 2.0 nm (green solid line).*

The result of the experimental $\rho_t(r)$ is given in the inset of **Figure B5**. The $\rho_t(r)$ decreases approximately linearly and the maximum dimension of the filaments is 9 nm, as derived from the approximation with a straight line where $\rho_t(r)$ reaches 0. The model-fit calculations and model-free data evaluation lead to a congruent picture of a ribbon-like structure of aggregated NFGAIL with a height of 10 nm and lateral dimension that exceed the upper limit of size determination which is 45 nm in this study. The height of 10 nm corresponds roughly to the single protofilament diameter of NFGAIL[219] or hIAPP[186], indicating that oligomers of NFGAIL only grow in one dimension (length) and the protofilaments are aligned next to but not above each other.
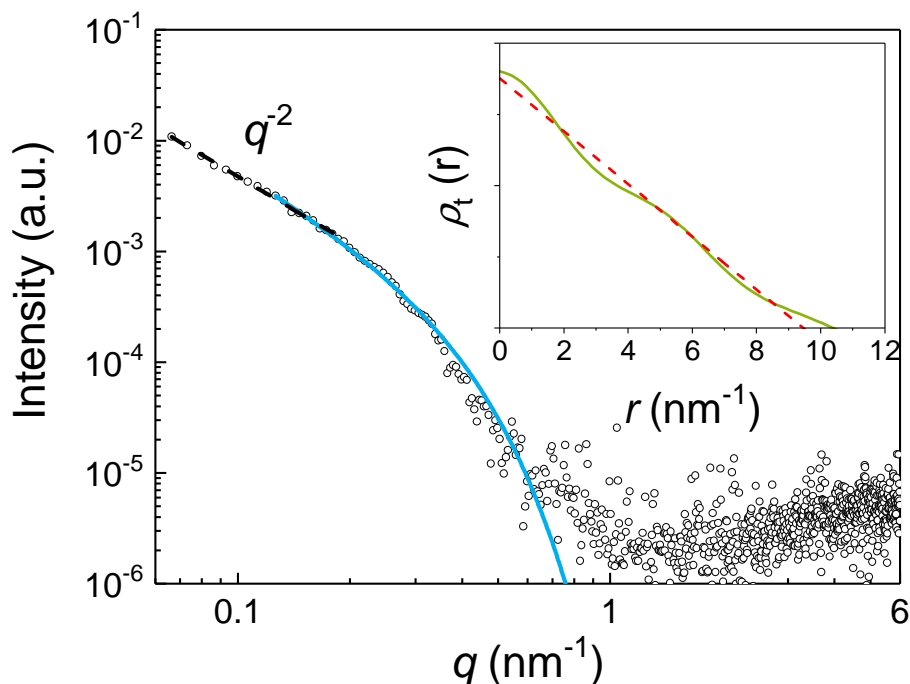
Figure B5: *SAXS data and the Guinier approximation[456-457] (symbols and blue solid line). The intensity scales with q^{-2} at low q-values (black dashed line) Inset: Cross section distance distribution function $\rho_t(r)$ calculated from the scattering curve in **Figure 4** (green solid line) and the best fitting cross section distribution function for a large homogeneous sheet with a height of 9.5 nm (red dashed line).*

B3   Paper

Reproduced with permission from the Journal of the American Chemical Society, accepted for publication. Unpublished work copyright 2017 American Chemical Society.

NFGAIL Amyloid Oligomers: The Onset of Beta-Sheet Formation and the Mechanism for Fibril Formation

Waldemar Hoffmann[†,‡], Kristin Folmert[†], Johann Moschner[†], Xing Huang[‡], Hans von Berlepsch[†], Beate Koksch[†], Michael T. Bowers[§], Gert von Helden[‡] and Kevin Pagel[†]*

[†] Freie Universität Berlin, Institute of Chemistry and Biochemistry - Organic Chemistry, Takustr. 3, 14195 Berlin, Germany

[‡] Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany

[§] Department of Chemistry and Biochemistry, University of California Santa Barbara, Santa Barbara, California 93106, United States

ABSTRACT: The hexapeptide NFGAIL is a highly amyloidogenic peptide, derived from the human islet amyloid polypeptide (hIAPP). Recent investigations indicate that presumably soluble hIAPP oligomers are one of the cytotoxic species in type II diabetes. Here we use Thioflavin T staining, transmission electron microscopy as well as ion mobility-mass spectrometry coupled to infrared (IR) spectroscopy to study the amyloid formation mechanism and the quaternary- and secondary-structure of soluble NFGAIL oligomers. Our data reveal that at neutral pH NFGAIL follows a nucleation-dependent mechanism to form amyloid fibrils. During the lag phase, highly polydisperse, polymorph, and compact oligomers (oligomer number n=2-13) as well as extended intermediates (n=4-11) are present. IR secondary structural analysis reveals that compact conformations adopt turn-like structures, whereas extended oligomers exhibit a significant amount of β-sheet content. This agrees well with previous molecular dynamic simulations and provides direct experimental evidence that unordered off-pathway NFGAIL aggregates up to the size of at least the 13-mer as well as partially folded β-sheet containing oligomers are coexisting.

1. INTRODUCTION

Amyloid formation is a hallmark of various diseases such as type II diabetes (T2D), Parkinson´s, and Alzheimer´s disease. In the case of T2D unstructured human islet amyloid polypeptide (hIAPP) monomers assemble into highly structured, β-sheet rich, insoluble deposits known as amyloid plaques.[1] These plaques are found in more than 95 % of T2D patients, which suggests their direct involvement in β-cell dysfunction in the pancreas.[2-5] Recent studies have shown that especially early hIAPP oligomers represent the toxic species.[1, 5-7] Understanding their detailed structure and the mechanism leading from monomer to fibrils is crucial to selectively modulate the assembly pathway.[1, 7-10]

Aggregation prone domains of hIAPP include the fragments 1-8, 8-20, 20-29 and 30-37, all of which form amyloid fibrils.[11-15] While each of these fragments may play a role in the assembly of the parent hIAPP, the 20-29 region has garnered the most attention since chemical modification[16-17] or mutations[5] in this region disrupt fibril formation. The 22NFGAIL27 fragment within this region is the shortest known hIAPP sequence capable of forming amyloid fibrils at pH ~ 7 and was shown to be cytotoxic towards the pancreatic cell line RIN5fm.[18] In addition, recent evidence suggests the 20-29 domain may play a

central role in fibril formation of full-length hIAPP.[19] Thus, many experimental[20-22] as well as theoretical[23-28] studies have investigated the assembly characteristics of NFGAIL. They show that the ionic strength[23] influences the NFGAIL assembly and oligomers are formed due to attractive, hydrophobic interactions between phenylalanine residues.[22, 25-26, 28] Molecular dynamic simulations further support that hydrophobic interactions lead to the formation of both disordered and structured aggregates.[24, 26] However, the dissociation of those disordered oligomers is the rate-limiting step for the formation of structured aggregates.[27] Those can then further evolve into amyloid fibrils composed of antiparallel β-strands.[22, 28] The characterization of soluble oligomers using solution-based Fourier Transform-infrared (FT-IR), circular dichroism (CD) or nuclear magnetic resonance (NMR) spectroscopy only provides ensemble averaged-information rather than information on individual oligomeric states. Thus, experimental evidence for the proposed assembly is still missing.

Gas-phase techniques can isolate and characterize a single species in the presence of many others without affecting the underlying equilibrium. In particular, ion mobility-mass spectrometry (IM-MS), a technique which separates ions based on their mass, charge, size and shape, has long been applied successfully to study the structure and aggre-

1

gation of reacting systems[29] and more recently has been used to follow the assembly of early amyloid intermediates.[30-33] In a typical IM-MS experiment the analyte is first carefully transferred from solution into the gas phase using soft ionization techniques such as electrospray ionization (ESI). These techniques have been shown to preserve solution backbone structures of peptides and small proteins[34-37] and aggregates of hexameric peptides[38] and Aβ peptides[39]. Transferred ions are then gently pulsed into an IM cell that is filled with an inert buffer gas, through which they move under the influence of a weak electric field. Compact ions undergo fewer collisions with the buffer gas than extended ions and therefore leave the IM cell earlier. The resulting drift time can further be used to calculate a rotationally-averaged collision cross-section (CCS), which is an instrumentally independent value from which information about the quaternary-structure of amyloid oligomers can be deduced.[38]

IR spectroscopy on the other hand is highly depended on intramolecular vibrations, which makes it an ideal tool to obtain detailed structural information.[40] For example, the amide I band, *i.e.* the C=O stretching mode, strongly depends on the secondary structure adopted by peptides and proteins.[41-42] Antiparallel β-sheet rich proteins typically feature an amide I frequency at lower wavenumbers (1610-1640 cm[-1]) and additional weak mode at higher wavenumbers (~ 1700 cm[-1]) compared to helical or turn-like structures (1648-1670 cm[-1], ~1660-1690 cm[-1], respectively).[37, 41-42] Recently, IM-MS was successfully used as a preselection tool to perform gas phase IR spectroscopy on individual amyloid oligomeric states and conformations.[43] The data showed that the tetramers (n=4) of the two amyloid forming peptides VEALYL and YVEALL undergo a characteristic transition from compact and unordered conformations into more extended and β-sheet rich versions.[43] All of the observed higher-order oligomers (n≥4) were shown to exhibit an elevated β-sheet content. Here, we employ IM-MS coupled to gas-phase IR spectroscopy for the structural characterization of NFGAIL oligomers.

2. RESULTS AND DISCUSSION

The hexapeptide NFGAIL is the shortest fragment of hIAPP known to form amyloid fibrils at neutral pH.[18] Those fibrils have been shown to feature antiparallel β-strands and a morphology similar to full length hIAPP fibrils.[21, 28, 44] In order to study the tendency of early oligomers to form β-sheets, the NFGAIL peptide was incubated for two days in ammonium acetate buffer. This procedure lead to the formation of fibrillar aggregates with anisometric cross-section, a mean apparent diameter of 20 to 30 nm and few hundreds of nanometers length (Figure 1a and 1b). Typical protofilaments have not been detected. The kinetics of fibril formation was monitored in real-time using Thioflavin T (ThT) assay.[45] ThT intercalates into the cavity of amyloid fibrils, which leads to an increased fluorescence. Even for high concentrations (4 mM) the hexapeptide NFGAIL follows a nucleation dependent growth mechanism (Figure S2), characteristic of classical on-pathway amyloid formation. The sigmoidal growth behavior can be divided into lag-, growth-, and saturation phases. During
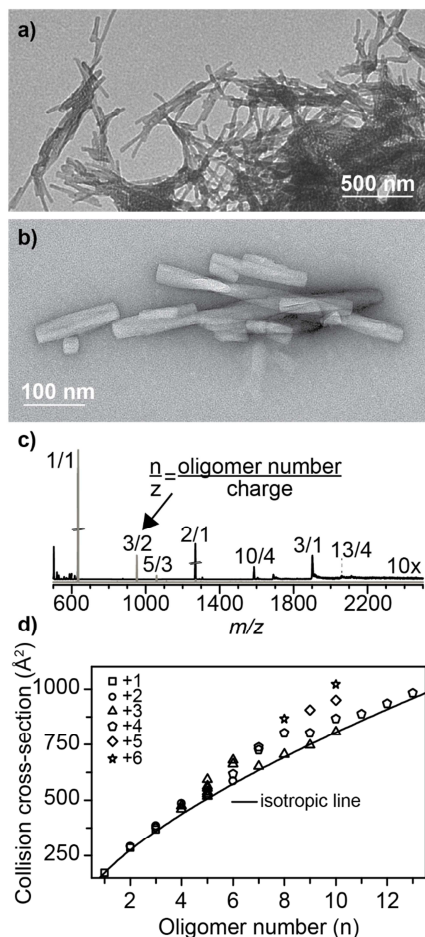
Figure 1. Structural investigation of NFGAIL oligomers. (a,b) Negative staining transmission electron microscopy (TEM) images of a NFGAIL solution (4 mM) in ammonium acetate buffer (10 mM, pH ~ 7) shows the formation of amyloid fibrils. Scale bars: (a) 500 nm, (b) 100 nm. (c) Mass spectra recorded at two different instrumental settings on the iMob[46] instrument (gray and black line) indicate for a freshly electro-sprayed NFGAIL solution the presence of multiple oligomers with *n* being the oligomer number and *z* the charge. (d) Measured collision cross-sections (CCSs) as a function of the oligomer number *n*. The solid line indicates theoretical CCSs expected for an idealized spherical growth. The experimental error of each CCS measurement is less than 1 % and smaller than the size of the symbols.

2

the lag-phase (20 hours), early soluble NFGAIL intermediates are present.[7, 18] These transient oligomers are highly polydisperse, *i.e.* a multitude of oligomeric states coexist. In addition, they most likely are polymorph, *i.e.* they occupy a wide range of different conformations for one oligomeric state. All species can interconvert or undergo subunit exchange, but once a so-called nucleus is formed, the auto-catalytical growth phase is initiated. The nuclei presumably act as a template to assemble monomers into mature fibrils, which are present at the saturation phase (see reference[39] for an example).

The mass spectrum (Figure 1 c) shows that NFGAIL immediately forms a wide distribution of oligomers, spanning from a singly charged monomer $n/z=1/1$ (with n = oligomer number, z = charge) up to a quadruply charged 13-mer n/z = 13/4. All of these oligomers carry much less charge than their number of peptide strands and are therefore most likely not disturbed by Coulomb-repulsion and derived from solution. IM-MS further reveals that multiple conformations and higher oligomeric states of the same m/z value exist (Figure S3, S4, S5). CCSs of all these species (Table S4) are displayed as a function of the oligomer number in Figure 1d. The solid line represents an ideal isotropic (*i.e.* globular) growth behavior, following the equation $\sigma=\sigma_1 \cdot n^{2/3}$, where σ_1 is the monomer´s CCS and n the oligomer number.[38] Oligomers exhibiting an experimental CCS on the isotropic line adopt compact, spherical conformations consisting of turn-like and/or unordered structures. Extended oligomers that are deviating from the isotropic growth model on the other hand are often partially structured and form helical or β-sheet-rich conformations.[38, 43] Figure 1d shows a variety of compact (experimental CCSs close to the isotropic line) NFGAIL oligomers, starting from the singly charged monomer ($n/z=1/1$) up to the quadruply charged 13-mer ($n/z=13/4$). These compact oligomers presumably adopt turn-like or unordered conformations. Interestingly, starting from the tetramer ($n=4$) multiple conformations are present. For example, for the pentamer these range from compact ($n/z=5/2$ with 523 Å^2 and $n/z=5/3$ with 520 Å^2) to highly extended structures ($n/z=5/3$ with 540 and 566 Å^2, respectively). Similar extended conformations are also observed for higher oligomeric states (n=6-10). Thus, the NFGAIL tetramer might represent a structural transition point from unordered/turn-like conformations into at least partially folded, presumably β-sheet rich structures. This observation is consistent with recent results on the 8-20 fragment where a significant β-sheet content begins at the tetramer.[47] In full-length hIAPP, however, a β-hairpin conformer is observed for both the monomer and dimer[30, 48] and is shown to be the conformation that leads to fibrilization. Further, a multidimensional analytical approach for the full-length hIAPP shows that in the presence of copper ions globular and toxic off-pathway hIAPP oligomers are formed, but in neat hIAPP a different assembly pathway starting at tetramer is observed.[49] Another IM-MS study also observed more extended hIAPP versions for n ≥ 4.[50]

Although IM-MS provides information on the stoichiometry and overall size of amyloid oligomers, direct details of the fine structure cannot be deduced. Instead, IM-MS can be used to pre-select individual conformations for a subsequent analysis by orthogonal techniques such as gas phase IR spectroscopy. The combination of IM-MS and IR spectroscopy allows the individual characterization of m/z- and drift-time-selected species on the quaternary- as well as the secondary-structure level. Amide I vibrations (C=O stretching modes) are highly sensitive towards the secondary structure adopted by peptides and proteins.[37, 41-42] However, due to the number of vibrational modes individual amide I features usually overlap. In conventional condensed-phase spectroscopy a deconvolution procedure is typically used to deduce the relative content of each motif.[51-52] In contrast, such relative populations cannot be directly obtained from gas phase infrared multiple photon dissociation (IRMPD) spectra without making assumptions about the oscillator strengths. The IRMPD process is based
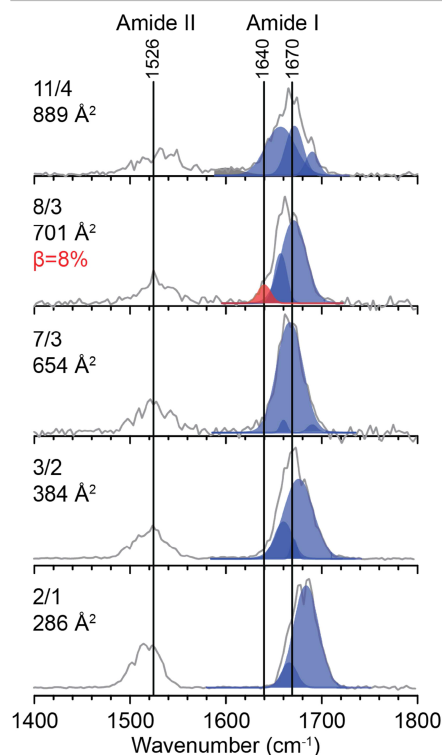
Figure 2. Infrared multiple photon dissociation (IRMPD) spectra of m/z- and drift-time selected compact NFGAIL oligomers, exhibiting experimental CCSs similar to theoretical values predicted by the isotropic growth model. Fractions of the amide I band that are representative for β-sheet structures (1600-1640 cm^{-1}) and turn-like structures (1660-

3

1685 cm^{-1}) are shown in red and blue, respectively. For details about the fitting procedure see Supplementary Information.

on the sequential absorption of multiple photons and the resulting IR intensity does not scale linearly with the number of absorbed photons. A deconvolution of IRMPD spectral features may therefore not be fully quantitative but still represents a reasonably good representation of the different structural features contributing to a complex spectral band. Hence, while we provide "quantitative" percentages of β-sheet content in a given band, these actually represent the relative amounts between spectral bands for differing conformations or oligomeric systems but are less quantitative within a given band.

Figure 2 shows a set of IRMPD spectra of compact NFGAIL oligomers (n=2-11) measured in the wavenumber range from 1400-1800 cm^{-1}. All spectra feature two bands assigned as amide II (1480-1540 cm^{-1}) and amide I (1600-1720 cm^{-1}). The amide I region was fitted with multiple Gaussian curves (for details see Supporting information), where antiparallel β-sheets and turn-like motifs are labelled in red and blue, respectively. The center of the amide I band of the singly charged dimer (n/z=2/1) appears at 1683 cm^{-1}, whereas for higher oligomers (n=3-11) this feature occurs at lower wavenumbers. The extent of hydrogen bonding within the dimer is apparently lower than for larger oligomers and therefore provides less perturbation for the individual C=O oscillators. The amide I band of higher compact oligomers appears around 1670 cm^{-1}, which is indicative for turn-like structures.[41-43] This observation agrees with the fact that the experimental CCSs of these oligomers represent compact, spherical structures. Interestingly, an additional shoulder around 1640 cm^{-1} emerges for the triply charged octamer (n/z = 8/3), indicating a β-sheet content of approximately 8 %. The experimental CCS of this octamer (n/z=8/3 with 701 Å^2), however, fits well with the theoretical value for a spherical, compact conformation (694 Å^2). A small β-sheet content observed via IRMPD may not necessarily be correlated with an extended structural ensemble.

The first NFGAIL oligomers that considerably deviate from the isotropic line are pentamers. They are highly polymorph and range from compact (n/z=5/2 with 523 Å^2 and n/z=5/3 with 520 Å^2) to highly extended (n/z=5/3 with 540 Å^2 and 566 Å^2) conformations. The individual, conformer-selective IR spectra show an amide I band at 1674 cm^{-1}, which indicates a predominantly turn-like structure present for all pentamers (Figure 3). Interestingly, the more extended pentamer isoforms exhibit a broader amide I band than observed for the corresponding compact species and show additional features at around 1620-1640 cm^{-1} and 1700 cm^{-1} (red Gaussians). These features are indicative of antiparallel β-strands. An amide I deconvolution reveals a β-sheet IR fraction of up to 13 %. This observation is in good agreement with previous results, which suggest a higher β-sheet content for more extended conformations.[38, 43] As a result of the non-linear IR absorption process, however, it is not possible to absolutely quantify how many single peptide strands within the NFGAIL pentamer contribute to the β-sheet fraction in the IR spectrum. For a short peptide such as NFGAIL the β-sheet IR

fraction originates from the non-covalent assembly of at least two individual peptide strands. In addition, a recent
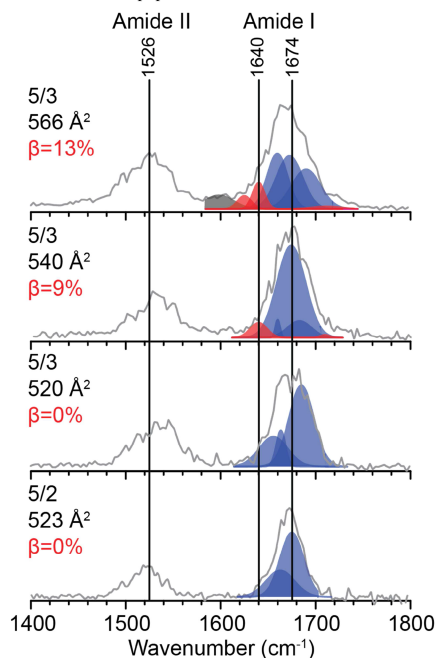
Figure 3. Infrared multiple photon dissociation (IRMPD) spectra of m/z- and drift-time selected NFGAIL pentamers ranging from compact (n/z=5/2 with 523 Å^2 and n/z=5/3 with 520 Å^2) to extended (n/z=5/3 with 540 Å^2 and 566 Å^2, respectively) structures. The amide I band (1600-1700 cm^{-1}) was deconvoluted with multiple Gaussians representing β-sheets (red) or turn-like (blue) conformations. One additional Gaussian between 1560-1600 cm^{-1} (grey) was included to compensate for an overlapping amide I and amide II band. Details of the fitting procedure are given in the Supplementary Information.

2D-IR study on the full-length hIAPP further suggests, that β-sheet rich hIAPP oligomers, which are present in the lag phase, are composed of less than four β-strands.[19] Thus, the here investigated NFGAIL oligomers are likely to be composed of at least two or three β-strands.

The gas phase IR spectra of other extended NFGAIL oligomers (n=4-9) also show a broad amide I band, with elevated intensities at 1617-1640 cm^{-1} (Figure 4). The Gaussian deconvolution analysis reveals a significant β-sheet content of up to 24 %. There is a significant jump in β-sheet content from n=4 (14 %) to n=7 (24 %) but no smooth trend as n increases. Hence factors beyond oligomer size contribute to β-sheet formation. A complex equilibrium between oligomeric states and conformations with different β-sheet content is established in the early stages of

4

NFGAIL aggregation (lag phase). Secondary structural transitions might therefore occur over a multitude of oligomeric states as observed for other amyloid forming sequences such as NNQQNY.[38] Similarly to NFGAIL, also the full-length hIAPP peptide shows a diverse free energy landscape, which is more complicated than a simple transition from a random-coil structure into a perfect fiber-like nucleus.[19] Thus, the short NFGAIL peptide might serve as a good model system and provides critical insights into the full-length hIAPP assembly mechanism. In addition, the coexistence of both turn-like as well as partially β-sheet rich structured aggregates for n=4-10 is in agreement with a previous theoretical study, which showed that unstructured NFGAIL oligomers are initially formed and their subsequent dissociation is the rate-limiting step for the assembly into higher, β-sheet rich oligomers.[27] The extended β-sheet rich NFGAIL oligomers investigated here, therefore, most likely represent on-pathway oligomers to amyloid fibrils.

3. CONCLUSIONS

A combination of condensed-phased methods such as ThT assay and TEM, as well as gas phase methods such as IM-MS coupled to gas-phase IR spectroscopy was used to study the kinetics and oligomeric structural evolution occurring in NFGAIL fibril formation. Under buffered conditions, NFGAIL follows a nucleation dependent growth mechanism into mature fibrils. IM-MS analysis reveals that during the lag phase a variety of oligomeric states n and different conformations ranging from compact (for n=2-13) to extended (for n=4-10) structures are present. Interestingly, also for the full-length hIAPP peptide[49-50] and other amyloid forming systems[38, 43, 48, 53] a similar transition from compact to more extended versions has been observed. In all cases, the conversion typically starts for oligomers as little as two to nine subunits.[49-50] The early transition into more extended, presumably β-sheet rich structures might be therefore a general feature of amyloid forming systems.

The gas phase IR analysis of compact NFGAIL oligomers (n=2-11) shows an amide I band centered at 1670 cm^{-1}, associated with turn-like structures. Extended NFGAIL oligomers (n=4-9), however, exhibit additional IR features at 1617-1640 cm^{-1}, which are representative of β-sheet rich structures. Deconvolution of the amide I band indicates a β-sheet IR content of up to 24 % for extended NFGAIL oligomers.

A previous theoretical study suggests that the formation of unstructured aggregates and their subsequent dissociation is the rate-limiting step to form higher-order, β-sheet rich NFGAIL oligomers.[27] The data presented here support this hypothesis by providing the first direct secondary structure data for individual oligomers. Due to the complex assembly-disassembly-assembly cascade, the investigated oligomers range from largely unordered to significantly folded β-sheet containing species and therefore for every extended version a more compact, unordered counterpart coexists. The conformational complexity for each NFGAIL oligomer observed here is in contrast to our previous study[43] on VEALYL and YVEALL where conformational complexity was limited to only one or two oligomer sizes. This points to coexistence of on- and off-pathway
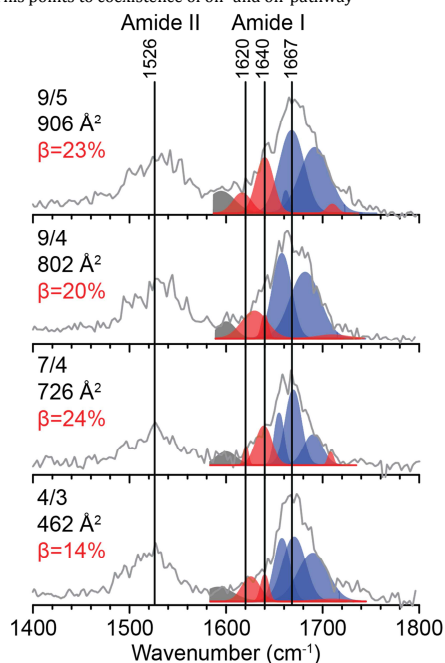
Figure 4. Individual infrared multiple photon dissociation (IRMPD) spectra of extended NFGAIL oligomers. The amide I band (1600-1700 cm^{-1}) was deconvoluted with multiple Gaussians representing β-sheets (red) or turn-like (blue) conformations. Details of the fitting procedure are given in the Supplementary Information.

aggregates in the NFGAIL system and complex assembly/disassembly dynamics that ultimately leads to both an on pathway β-sheet dominated set of oligomers leading to fibrils and an off-pathway more isotropic set of oligomers up to at least the size of an 13-mer.

4. MATERIALS AND METHODS

4.1 Samples. The hexapeptide NFGAIL was synthesized manually according to standard Fmoc-chemistry using a preloaded TGA resin. Peptide purification was performed on a low-pressure HPLC system. Purity was validated using an analytical HPLC system and high-resolution mass spectrometry (Figure S1). Further details can be found in the Supporting information.

4.2 ThT Assay. An NFGAIL stock solution was prepared by dissolving the purified NFGAIL in HFIP (1,1,1,3,3,3-hexafluoro-2-propanol) (~18 mM) and was further sonicated for 15 min to dissolve all preformed aggregates. Aliquots of this stock solution were dried and then redissolved to a final concentration of 4 mM in ammonium acetate buffer (10 mM,

5

pH ~ 7), containing 20 μM ThT. After dissolution, the sample was sonicated for 30 s and then incubated at 37 °C with 1300 rpm.

Fluorescence spectra were recorded using a 1 cm path length quartz cuvette (Hellma, Müllheim, Germany) and a luminescence spectrometer LS50B (Perkin-Elmer, Boston, MA, USA). Spectra were recorded at room temperature from 470-500 nm after excitation at 420 nm (excitation slit width 5 nm; emission slit width 10 nm; scan speed = 300 nm/min; accumalations = 3). The fluorescence intensity at 485 nm was normalized with respect to its maximum value.

4.3 TEM characterization. Sample preparation was performed according to ThT assay. The solution was further incubated at 37 °C and shaken with 1300 rpm for two days. During the incubation time, the formation of insoluble deposition was observed. Samples for staining electron microscopy were prepared by adsorbing 5 μL aliquots of peptide solution to glow-discharged carbon-coated collodium films on 400-mesh copper grids. The grids were blotted, stained with 1 % phosphotungstic acid, and air-dried. TEM micrographs were taken by a FEI Talos L120 TEM operated at 120 kV for morphological characterization.

4.4 Isomer-Selective IR Spectroscopy. The purified peptide was dissolved in ammonium acetate (10 mM, pH ~ 7) to yield a final peptide concentration of 1 mM. For nESI, ~ 8 μL of sample were loaded into in-house-prepared Pd-Pt-coated borosilicate capillaries, and voltages of 0.6-1.0 kV were applied.

The ion mobility method for drift-time-based selection is well established.[54] Ions are drift time preselected prior to irradiation with intense IR light using an in-house constructed drift-tube ion mobility-mass spectrometer similar to one described previously.[40, 55] After ions are generated using a *nano* electrospray ionization (nESI) source they are transferred and stored in an entrance funnel. Subsequently, ions are released by 150 μs long pulses into a drift tube, where they travel under the influence of a weak electric field (10-20 V/cm) through helium buffer gas (~4 mbar). The drift velocity of a particular ion depends on its mobility, which in turn is based on its overall size, shape and charge. After releasing the ion mobility cell ions are mass selected using a quadrupole mass filter and their arrival time distributions (ATDs) can be recorded by measuring the time dependent ion current of the m/z selected species after release of the ion trap. CCSs are measured on two different IMS instruments; 1) HiRes[55] located in Santa Barbara (Resolution, R=100), which allows a high IMS separation of individual conformations. 2) iMob[46] located in Berlin (R=40), which is connected to a Free Electron Laser and therefore allows the recording of conformer- and mass-selected IR spectra.

Gas phase IR spectra are recorded on the iMob[46] instrument by selecting a narrow drift-time window (100 μs width) using electrostatic deflection prior to mass selection. This m/z- and ion mobility-selected ion cloud was further irradiated by an intense (10-40 mJ) 10 μs pulse of IR photons. The photofragmentation is detected by a time-of-flight (ToF) mass analyzer and IR spectra are obtained by plotting the fragmentation yield as a function of the tunable IR wavenumber. The final IR spectrum represents an average of at least two individual scans, where each scan was obtained by scanning in 3 cm^{-1} wavenumber steps and averaging at least 40 spectra for each step.

The tunable mid-IR light is provided by the Fritz-Haber-Institut Free Electron Laser and is transported to the instrument *via* an evacuated beam line, where the last two meters of the beam line are flushed with dry nitrogen to avoid water absorption.

ASSOCIATED CONTENT

Supporting Information

Details of the synthesis protocol, peptide purification, experimental procedures, data from ion mobility-mass spectrometry, collision cross-sections and Gaussian fitting routines are given in the Supporting information. This material is available free of charge via the Internet at http://pubs.acs.org

AUTHOR INFORMATION

Corresponding Author

* kevin.pagel@fu-berlin.de

Author Contributions

W.H., K.F., J.M., M.T.B., B.K., H.v.B., G.v.H., K.P. conceived and designed the experiments; W.H. performed and analyzed the experiments; K.F., J.M., X.H., H.v.B. supported the experiments; all the authors co-wrote the paper.

Notes
The authors declare no competing financial interest.

REFERENCES

(1) Cao, P.; Abedini, A.; Raleigh, D. P. *Curr. Opin. Struct. Biol.* **2013**, *23*, 82.

(2) Clark, A.; Cooper, G. J.; Lewis, C. E.; Morris, J. F.; Willis, A. C.; Reid, K. B.; Turner, R. C. *Lancet* **1987**, *2*, 231.

(3) Luskey, K. L. *Diabetes care* **1992**, *15*, 297.

(4) Kahn, S. E.; Andrikopoulos, S.; Verchere, C. B. *Diabetes* **1999**, *48*, 241.

(5) Cao, P.; Marek, P.; Noor, H.; Patsalo, V.; Tu, L. H.; Wang, H.; Abedini, A.; Raleigh, D. P. *FEBS Lett.* **2013**, *587*, 1106.

(6) Lin, C.-Y.; Gurlo, T.; Kayed, R.; Butler, A. E.; Haataja, L.; Glabe, C. G.; Butler, P. C. *Diabetes* **2007**, *56*, 1324.

(7) Haataja, L.; Gurlo, T.; Huang, C. J.; Butler, P. C. *Endocr. Rev.* **2008**, *29*, 303.

(8) Cooper, G. J.; Willis, A. C.; Clark, A.; Turner, R. C.; Sim, R. B.; Reid, K. B. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8628.

(9) Lorenzo, A.; Razzaboni, B.; Weir, G. C.; Yankner, B. A. *Nature* **1994**, *368*, 756.

(10) Zraika, S.; Hull, R. L.; Verchere, C. B.; Clark, A.; Potter, K. J.; Fraser, P. E.; Raleigh, D. P.; Kahn, S. E. *Diabetologia* **2010**, *53*, 1046.

(11) Gilead, S.; Gazit, E. *Exp. Diabetes Res.* **2008**, *2008*, 256954.

(12) Scrocchi, L. A.; Ha, K.; Chen, Y.; Wu, L.; Wang, F.; Fraser, P. E. *J. Struct. Biol.* **2003**, *141*, 218.

(13) Ehud, G. *Curr. Med. Chem.* **2002**, *9*, 1725.

(14) Nilsson, M. R.; Raleigh, D. P. *J. Mol. Biol.* **1999**, *294*, 1375.

(15) Ilitchev, A. I.; Giammona, M. J.; Do, T. D.; Wong, A. G.; Buratto, S. K.; Shea, J.-E.; Raleigh, D. P.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 1010.

6

(16) Kapurniotu, A.; Schmauder, A.; Tenidis, K. *J. Mol. Biol.* **2002,** *315*, 339.

(17) Tatarek-Nossol, M.; Yan, L. M.; Schmauder, A.; Tenidis, K.; Westermark, G.; Kapurniotu, A. *Chem. Biol.* **2005,** *12*, 797.

(18) Tenidis, K.; Waldner, M.; Bernhagen, J.; Fischle, W.; Bergmann, M.; Weber, M.; Merkle, M. L.; Voelter, W.; Brunner, H.; Kapurniotu, A. *J. Mol. Biol.* **2000,** *295*, 1055.

(19) Shim, S. H.; Gupta, R.; Ling, Y. L.; Strasfeld, D. B.; Raleigh, D. P.; Zanni, M. T. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106*, 6614.

(20) Soriaga, A. B.; Sangwan, S.; Macdonald, R.; Sawaya, M. R.; Eisenberg, D. *J. Phys. Chem. B* **2016,** *120*, 5810.

(21) Deng, W.; Cao, A.; Lai, L. *Protein Sci.* **2008,** *17*, 1102.

(22) Azriel, R.; Gazit, E. *J. Biol. Chem.* **2001,** *276*, 34156.

(23) Cai, Z.; Li, J.; Yin, C.; Yang, Z.; Wu, J.; Zhou, R. *J. Phys. Chem. B* **2014,** *118*, 48.

(24) Melquiond, A.; Gelly, J.-C.; Mousseau, N.; Derreumaux, P. *J. Chem. Phys.* **2007,** *126*, 065101.

(25) Wu, C.; Lei, H.; Duan, Y. *Biophys. J.* **2005,** *88*, 2897.

(26) Wu, C.; Lei, H.; Duan, Y. *J. Am. Chem. Soc.* **2005,** *127*, 13530.

(27) Wu, C.; Lei, H.; Duan, Y. *Biophys. J.* **2004,** *87*, 3000.

(28) Zanuy, D.; Ma, B.; Nussinov, R. *Biophys. J.* **2003,** *84*, 1884.

(29) von Helden, G.; Hsu, M. T.; Gotts, N.; Bowers, M. T. *J. Phys. Chem.* **1993,** *97*, 8182.

(30) Dupuis, N. F.; Wu, C.; Shea, J.-E.; Bowers, M. T. *J. Am. Chem. Soc.* **2009,** *131*, 18283.

(31) Bernstein, S. L.; Dupuis, N. F.; Lazo, N. D.; Wyttenbach, T.; Condron, M. M.; Bitan, G.; Teplow, D. B.; Shea, J.-E.; Ruotolo, B. T.; Robinson, C. V.; Bowers, M. T. *Nat. Chem.* **2009,** *1*, 326.

(32) Cole, H. L.; Kalapothakis, J. M. D.; Bennett, G.; Barran, P. E.; MacPhee, C. E. *Angewandte Chemie, International Edition* **2010,** *49*, 9448.

(33) Hoffmann, W.; von Helden, G.; Pagel, K. *Curr. Opin. Struct. Biol.* **2017,** *46*, 7.

(34) Baumketner, A.; Bernstein, S. L.; Wyttenbach, T.; Bitan, G.; Teplow, D. B.; Bowers, M. T.; Shea, J.-E. *Protein Science : A Publication of the Protein Society* **2006,** *15*, 420.

(35) Wyttenbach, T.; Grabenauer, M.; Thalassinos, K.; Scrivens, J. H.; Bowers, M. T. *The Journal of Physical Chemistry B* **2010,** *114*, 437.

(36) Wyttenbach, T.; Bowers, M. T. *The Journal of Physical Chemistry B* **2011,** *115*, 12266.

(37) Seo, J.; Hoffmann, W.; Warnke, S.; Bowers, M. T.; Pagel, K.; von Helden, G. *Angew. Chem.* **2016,** *55*, 14173.

(38) Bleiholder, C.; Dupuis, N. F.; Wyttenbach, T.; Bowers, M. T. *Nat. Chem.* **2011,** *3*, 172.

(39) Economou, N. J.; Giammona, M. J.; Do, T. D.; Zheng, X.; Teplow, D. B.; Buratto, S. K.; Bowers, M. T. *J. Am. Chem. Soc.* **2016,** *138*, 1772.

(40) Warnke, S.; Seo, J.; Boschmans, J.; Sobott, F.; Scrivens, J. H.; Bleiholder, C.; Bowers, M. T.; Gewinner, S.; Schöllkopf, W.; Pagel, K.; von Helden, G. *J. Am. Chem. Soc.* **2015,** *137*, 4236.

(41) Barth, A. *BBA-Bioenergetics* **2007,** *1767*, 1073.

(42) Jackson, M.; Mantsch, H. H. *Crit. Rev. Biochem. Mol. Biol.* **1995,** *30*, 95.

(43) Seo, J.; Hoffmann, W.; Warnke, S.; Huang, X.; Gewinner, S.; Schöllkopf, W.; Bowers, M. T.; von Helden, G.; Pagel, K. *Nat. Chem.* **2017,** *9*, 39.

(44) Kayed, R.; Bernhagen, J.; Greenfield, N.; Sweimeh, K.; Brunner, H.; Voelter, W.; Kapurniotu, A. *J. Mol. Biol.* **1999,** *287*, 781.

(45) Wolfe, L. S.; Calabrese, M. F.; Nath, A.; Blaho, D. V.; Miranker, A. D.; Xiong, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107*, 16863.

(46) Warnke, S.; von Helden, G.; Pagel, K. *Proteomics* **2015,** *15*, 2804.

(47) Wang, L.; Ilitchev, A. I.; Giammona, M. J.; Li, F.; Buratto, S. K.; Bowers, M. T. *J. Phys. Chem. B* **2016,** *120*, 11905.

(48) Dupuis, N. F.; Wu, C.; Shea, J.-E.; Bowers, M. T. *J. Am. Chem. Soc.* **2011,** *133*, 7240.

(49) Lee, S. J. C.; Choi, T. S.; Lee, J. W.; Lee, H. J.; Mun, D.-G.; Akashi, S.; Lee, S.-W.; Lim, M. H.; Kim, H. I. *Chem. Sci.* **2016,** *7*, 5398.

(50) Young, L. M.; Cao, P.; Raleigh, D. P.; Ashcroft, A. E.; Radford, S. E. *J. Am. Chem. Soc.* **2014,** *136*, 660.

(51) Arrondo, J. L. R.; Goñi, F. M. *Prog. Biophys. Mol. Biol.* **1999,** *72*, 367.

(52) Bramanti, E.; Benedetti, E. *Biopolymers* **1996,** *38*, 639.

(53) Do, T. D.; LaPointe, N. E.; Nelson, R.; Krotee, P.; Hayden, E. Y.; Ulrich, B.; Quan, S.; Feinstein, S. C.; Teplow, D. B.; Eisenberg, D.; Shea, J.-E.; Bowers, M. T. *J. Am. Chem. Soc.* **2016,** *138*, 549.

(54) Bowers, M. T. *Int. J. Mass Spectrom.* **2014,** *370*, 75.

(55) Kemper, P. R.; Dupuis, N. F.; Bowers, M. T. *Int. J. Mass Spectrom.* **2009,** *287*, 46.

7

NFGAIL Amyloid Oligomers: The Onset of Beta-Sheet Formation and the Mechanism for Fibril Formation

Waldemar Hoffmann[†,‡], Kristin Folmert[†], Johann Moschner[†], Xing Huang[‡], Hans von Berlepsch[†], Beate Koksch[†], Michael T. Bowers[§], Gert von Helden[‡] and Kevin Pagel[†]*

[†] Freie Universität Berlin, Institute of Chemistry and Biochemistry - Organic Chemistry, Takustr. 3, 14195 Berlin, Germany

[‡] Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany

[§] Department of Chemistry and Biochemistry, University of California Santa Barbara, Santa Barbara, California 93106, United States

* e-mail: kevin.pagel@fu-berlin.de (K.P.)

SUPPLEMENTARY INFORMATION

S1

Table of Contents

1   Supplementary Materials and Methods

1.1 Solid Phase Peptide Synthesis (SPPS)

Fmoc-L-amino acids were purchased from ORPEGEN Peptide Chemicals GmbH (Heidelberg, Germany). Fmoc-Leu Nova Syn®TGA-resin with 0.3 mmol/g loading was obtained from NovaBiochem (Merck Chemicals GmbH, Darmstadt, Germany). All solvents and common chemicals were used from VWR (Darmstadt, Germany) without further purification.

The peptide was synthesized according to standard Fmoc-chemistry using preloaded Fmoc-Leu-NovaSyn®TGA resin (0.3 mmol g^{-1}, Novabiochem). Standard couplings were performed in DMF with Fmoc-amino acids and HOBt [1-Hydroxybenzotriazole]/ DIC [N,N'-Diisocarbodiimide] in eight-fold excess with respect to the resin amount and with double couplings of one hour coupling time. A mixture of DBU [1,8-diazabicyclo[5,4,0]undec-7-en] and piperidine (2% each) in DMF was used for Fmoc-deprotection (3x10 min). The resin was washed between each step with DMF and DCM (3x6 mL each). Peptide was cleaved from the resin by treatment with 2 mL TFA [trifluoroacetic acid]/iPr3SiH/H_2O (90/5/5) for three hours using sonication at 30 °C. The resin was washed twice with 1 mL TFA and DCM, and excess of solvent was removed by evaporation. The peptide was dried by lyophilization before purification with preparative reversed phase HPLC. After dissolving the peptide in MeOH the purification was performed by using a LaPrepΣ low-pressure HPLC system (VWR, Darmstadt, Germany). A Kinetex RP-C18 endcapped (5 µM, 100 Å, 250x21.2 mm, Phenomenex®, USA) HPLC-column was used. A Security GuardTM PREP Cartridge Holder Kit (21.20 mm, ID, Phenomenex®, USA) served as pre-column. Deionized water (Milli-Q Advantage® A10 Ultrapure Water Purification System, Millipore®, Billerica, MA, USA) and ACN, both containing 0.1% (v/v) TFA were applied as eluents. HPLC runs were performed using an isocratic gradient over five minutes, 5% ACN, flow rate: 10 mL/min, then over 25 minutes, 5-70% ACN, flow rate: 20.0 mL/min. UV-detection was carried out at 220 nm. Data analysis was performed with an EZChrom Elite-Software (Version 3.3.2 SP2, Agilent Technologies, Santa Clara, CA, USA). The collected fractions were evaporated and lyophilized to give the peptide as a white powder.

Analytical HPLC was carried out on a Chromaster 600 bar DAD-System with CSM software (VWR/Hitachi, Darmstadt, Germany). The system works with a low-pressure gradient containing a HPLC-pump (5160) with a 6-channel solvent degasser, an organizer, an autosampler (5260) with a 100 µL sample loop, a column oven (5310) and a diode array flow detector (5430). A Kinetex C18 column (5 µm,250 Å~ 4.6 mm, Phenomenex®, Torrance, CA, USA) was used. Deionized water and ACN, both containing 0.1 % (v/v) TFA served as eluents. A flow rate of 1 mL/min was used and the column was heated to 24℃. The UV-detection of the peptides was performed at 220 nm and a linear gradient of 5–70% ACN + 0.1% TFA in 18 min was applied (Figure S1). The data were analyzed with EZ Chrom ELITE software (version 3.3.2, Agilent Technologies, Santa Clara, CA, USA).

Mass-to-charge ratios were determined with an Agilent 6220 ESI-ToF MS instrument (Agilent Technologies, Santa Clara, CA, USA) (Figure S1). Peptide was dissolved in MeOH and the peptide solution was injected by a syringe pump with a flow rate of 10 µL min^{-1}. Spray voltage was set to 4000 V, drying gas flow rate was 5 L min^{-1} and gas temperature was set to 300 ℃.

S3
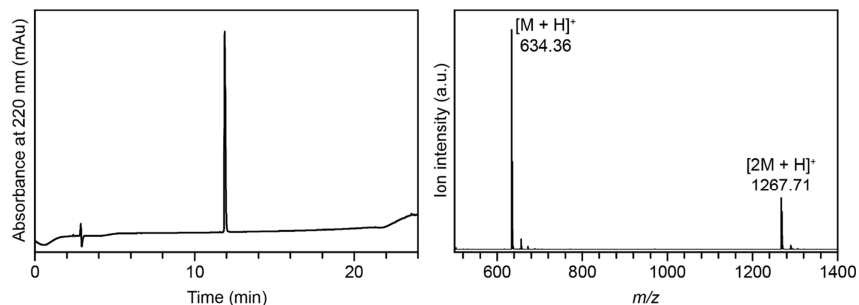
Figure S1. RP-HPLC chromatogram and ESI mass spectrum of purified NFGAIL peptide. According to the chromatogram and mass spectrum, the purity of the synthesized peptide is more than 99%.

1.2 Collision Cross-Section Determination

Ions of different sizes but identical *m/z* ratios can be separated based on their drift velocities when they travel under the influence of a weak electric field through a drift tube, which is filled with inert buffer gas. The drift velocity of an ion depends on different instrumental parameters and it is therefore often useful to convert drift times from ion mobility experiments into a rotationally averaged collision cross-section Ω (CCS). The CCS is a molecular property, which does not depend on instrumental parameters and which describes the overall shape of an ion. When a homogeneous electric field is applied inside of the drift region, the drift time can be converted using the following Mason-Schamp equation (equation S1).[1]

$$\Omega = \frac{3ze}{16N}\sqrt{\frac{2\pi}{\mu k_B T}}\frac{t_D E}{L}\frac{1013T}{273.16P} \tag{S1}$$

where z is the number of charges on the ion, e the elementary charge, μ the reduced mass, k_B the Boltzmann constant, T the temperature, t_D the drift time, E the electric field, L the drift tube length, N the number density of the buffer gas molecules and P the buffer gas pressure (in mbar) inside the drift cell.

1.3 Gaussian fitting procedure for amide I region

Infrared spectroscopy is an established tool for the characterization of the protein secondary structure.[2-3] The protein backbone exhibits two major, strong vibrational modes, namely the amide I and II band. The amide I absorption predominantly originates from C=O stretching vibrations, whereas the amide II absorption is arising from a combination of N-H bending and C-N stretching modes. Especially, the amide I band is very sensitive towards the secondary structure of proteins and peptides and typical amide I frequencies are given in Table S1. A recent study on prototypical examples for β-sheet rich and helical proteins shows a very good agreement between condensed-phase and gas-phase IR positions.[4]

Table S1. *Correlations between Common Protein Structures and Amide I Frequency*[2-3]

Secondary structure	Amide I band (cm^{-1})
aggregated strands	1610–1628
β-sheet	1610–1640
random coil	1640–1648
α-Helix	1648–1660
3$_{10}$-Helix, type II β-turn	1660–1685
antiparallel β-sheet / aggregated strands	1675–1695

Both the amide I and II regions are, however, broad bands and it is not possible to resolve individual bands corresponding to different secondary structure elements. Therefore, a deconvolution procedure of the respective bands is commonly applied in the condensed-phase to extract the secondary structure content of each motif. Recently, a similar approach was also performed for the deconvolution of the amide I band of gas phase IR spectra of peptides.[5] IR features in the 1600 – 1710 cm^{-1} region were fit with multiple Gaussian functions as shown in equation S2. However, the IR absorption process is more complicated in the gas phase (IRMPD process) and a linear correlation of IR intensities and actual secondary structure content does not exist. Instead the fraction of β-sheet IR signature (F(β)) in the amide I region is calculated according to equation S3, where A(n) denotes the area of *n*-th Gaussian.

$$I = \sum_n \frac{A(n)}{\sigma(n) \cdot \sqrt{\pi/2}} \cdot exp\left[-\frac{2\{x - x_C(n)\}^2}{\sigma(n)^2}\right] \tag{S2}$$

$$F(\beta) = \frac{A(1) + A(2) + A(6)}{\sum_{n=1}^{6} A(n)} \tag{S3}$$

The amide I region was fitted with six Gaussian (1-6) using constraints in the peak center (X_C) and width (σ) as shown in Table S2. The constraints were set based on the known IR band positions for the individual secondary structure elements as shown in Table S1. One additional Gaussian at 1560 – 1600 cm^{-1} (0) is used to correct the baseline below 1600 cm^{-1}. For the calculation of β-sheet related IR frequencies we also used a narrow IR window as provided by reported condensed-phase experiments. For other secondary structural elements, we slightly broaden the window to allow a better fit to the experimental data and to avoid underestimation of those elements. Thus, windows related to turns, helices or unordered structures were chosen to slightly overlap with each other, whereas β-sheet related IR windows were set adjacent to those. This further reduces an overestimation of the β-sheet IR content. R-square values (R^2) for all fits are larger than 0.94.

S5

Table S2. *Constraints in the peak center and width for the fitting amide I with multiple Gaussians*

n	Peak center (cm^{-1})	Peak width σ (cm^{-1})
0	1560–1600	0–30
1	1600–1630	0–30
2	1620–1640	0–30
3	1640–1660	0–30
4	1650–1680	0–30
5	1660–1690	0–30
6	1680–1710	0–30

2 Additional Experimental Data of NFGAIL
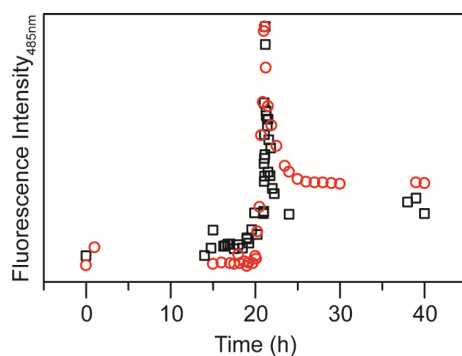
Figure S2. Two individual (black and red symbols) Thioflavin T assays of an incubated NFGAIL (4 mM) sample in ammonium acetate (10 mM) buffer. The growth phase into mature fibrils is initiated after 20 hours, whereas insoluble fibrils are present after 22 hours. The formation of insoluble fibrils, which accumulate at the bottom or at the walls of the cuvette, leads to a decrease of the measured fluorescence intensity at 485 nm.

S6

77
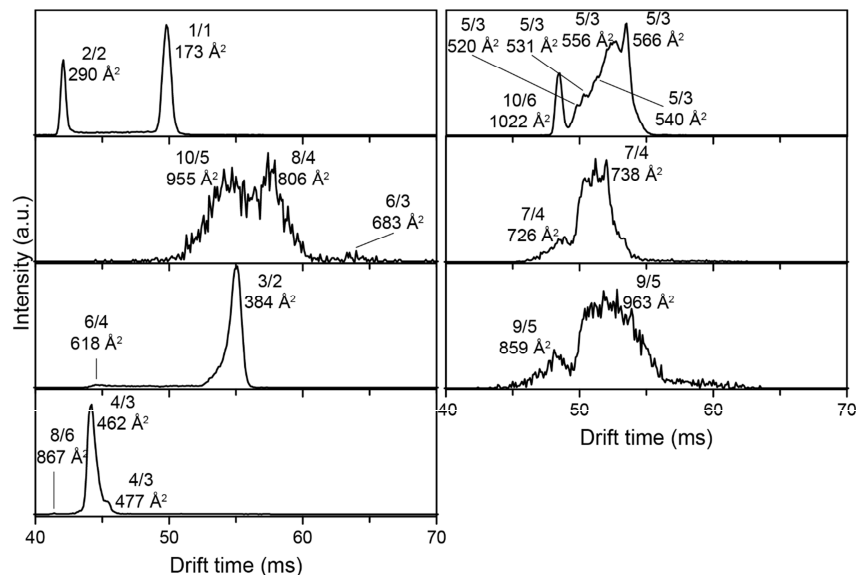
Figure S3. Arrival time distributions (ATDs) of NFGAIL (1 mM) in ammonium acetate buffer (10 mM, pH~7) recorded at a home-built high-resolution IM-MS instrument[6] (HiRes, Resolution, R=100) located in Santa Barbara. Each ATD feature is labelled by its *n/z* notation with *n* being the oligomer number and *z* the overall charge. The respective collision cross-sections are given in Å^2.
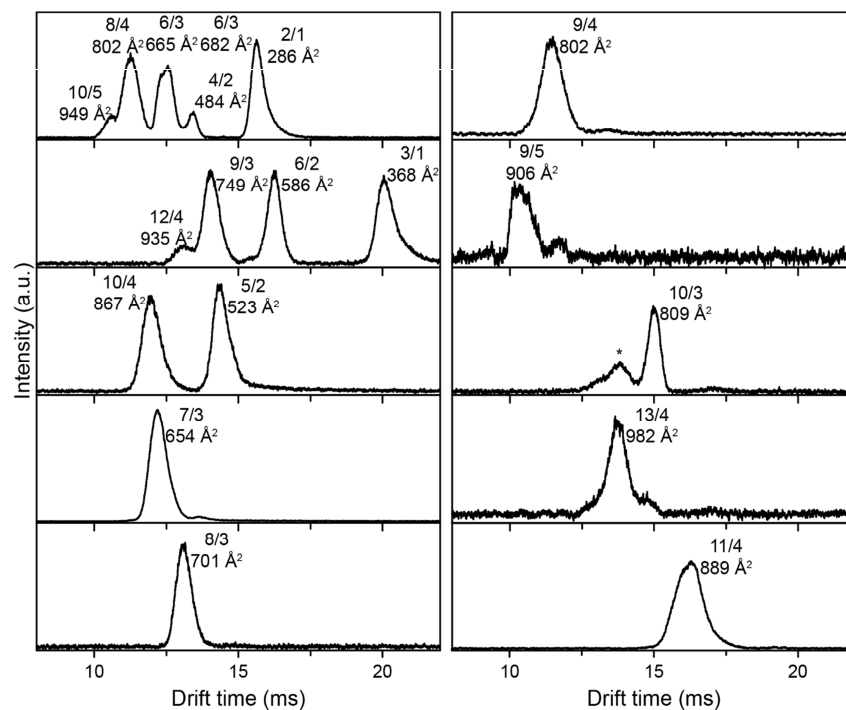
Figure S4. Arrival time distributions (ATDs) of NFGAIL (1 mM) in ammonium acetate buffer (10 mM, pH~7) recorded at a home-built high-resolution IM-MS instrument[7] (iMob, R=40) located in Berlin. Each ATD feature is labelled by its *n/z* notation with *n* being the oligomer number and *z* the overall charge. The respective collision cross-sections are given. Overlapping noise features are shown with an asterisk.

S8

79
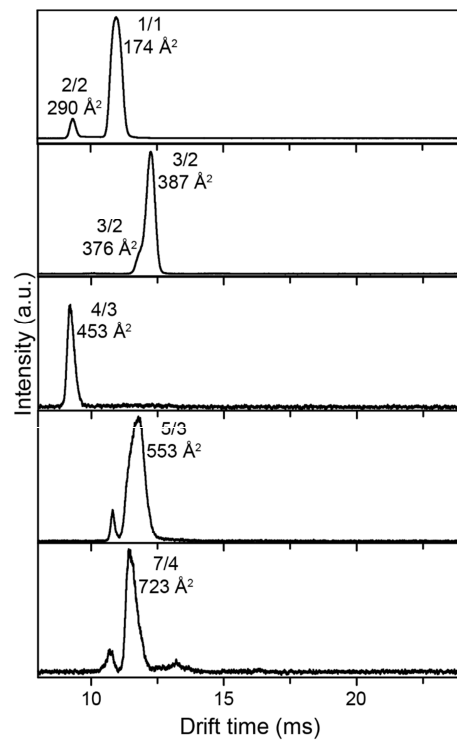
Figure S5. Arrival time distributions (ATDs) of NFGAIL (1 mM) in ammonium acetate buffer (10 mM, pH~7) recorded at a home-built high-resolution IM-MS instrument[7] (iMob, R=40) located in Berlin. Each ATD feature is labelled by its *n/z* notation with *n* being the oligomer number and *z* the overall charge. The respective collision cross-sections are given.

S9

Table S3. *Gaussian fitting parameters for amide I bands of NFGAIL oligomers*

	Fitted parameters of *n*-th Gaussian:			Peak center (FWHM) Relative area				R^2
	n = 0	*n* = 1	*n* = 2	*n* = 3	*n* = 4	*n* = 5	*n* = 6	
2/1	-	-	-	-	1666 (15) 15	1684 (25) 100	-	0.99
3/2	-	-	-	1660 (18) 27	1671 (6) 5	1676 (30) 100	-	0.99
4/3	1592 (30) 28	1626 (18) 30	1640 (8) 14	1658 (15) 66	1670 (23) 100	1690 (30) 100	1710 (30) 4	0.95
5/2	-	-	-	-	1663 (30) 50	1675 (25) 100	-	0.99
5/3 520 Å^2	-	-	-	1656 (30) 42	1663 (10) 17	1685 (26) 100	-	0.95
5/3 540 Å^2	-	-	1640 (18) 10	1660 (5) 3	1674 (30) 100	1682 (25) 15	1710 (10) 1	0.96
5/3 556 Å^2	1598 (30) 28	1625 (14) 14	1640 (13) 23	1660 (22) 82	1672 (27) 100	1690 (30) 82	1710 (30) 5	0.99
7/3	-	-	-	1660 (6) 3	1667 (26) 100	1690 (9) 3	-	0.99
7/4	1600 (24) 21	1621 (5) 8	1639 (19) 54	1655 (11) 41	1670 (19) 100	1690 (22) 49	1708 (6) 5	0.94
8/3	-	-	1640 (13) 11	1657 (12) 28	1671 (25) 100	-	-	0.99
9/4	1600 (18) 15	1629 (26) 35	1640 (6) 8	-	1658 (19) 80	1682 (30) 100	1710 (30) 5	0.94
9/5	1595 (27) 29	1617 (18) 18	1640 (17) 47	1660 (6) 7	1666 (25) 100	1690 (30) 97	1710 (10) 5	0.94
11/4	1600 (30) 14	-	-	1657 (30) 100	1671 (17) 57	1690 (12) 20	-	0.99

S10

81

Table S4. *Collision cross-sections Ω (measured on the iMob[7] and HiRes[6] instruments) and β-sheet IR intensity ratio in the amide I feature of NFGAIL oligomers. Bold collision cross-sections have been used to constructed Figure 1c.*

N	z	$\Omega_{measured}$ (Å²) iMob[7]	$\Omega_{measured}$ (Å²) HiRes[6]	β-sheet ratio in amide I
1	1	**174**	**173**	0
2	1	**286**		0
2	2	**290**	**290**	
3	1	**368**		
3	2	**376**		
		387	**384**	0
4	2	**484**		
4	3	**453**	**462**	0.14
			477	
5	2	**523**		0
5	3		**520**	0
			531	
			540	0.09
		553	**556**	
			566	0.13
6	2	**586**		
6	3	**665**		
		682	**683**	
6	4		618	
7	3	**654**		0
7	4	**723**	**726**	0.24
			738	
8	3	**701**		0
8	4	**802**	**806**	
8	6		**867**	
9	3	**749**		
9	4	**802**		0.20
9	5	**906**	859	0.23
			963	
10	3	**809**		
10	4	**867**		
10	5	**949**	**955**	
10	6		**1022**	
11	4	**889**		0
12	4	**935**		
13	4	**982**		

3 References

(1) Mason, E. A.; McDaniel, E. W. *Transport Properties of Ions in Gases*. John Wiley & Sons: New York, United States, 1988.
(2) Barth, A. *BBA-Bioenergetics* **2007,** *1767*, 1073.
(3) Jackson, M.; Mantsch, H. H. *Crit. Rev. Biochem. Mol. Biol.* **1995,** *30*, 95.
(4) Seo, J.; Hoffmann, W.; Warnke, S.; Bowers, M. T.; Pagel, K.; von Helden, G. *Angew. Chem.* **2016,** *55*, 14173.
(5) Seo, J.; Hoffmann, W.; Warnke, S.; Huang, X.; Gewinner, S.; Schöllkopf, W.; Bowers, M. T.; von Helden, G.; Pagel, K. *Nat. Chem.* **2017,** *9*, 39.
(6) Kemper, P. R.; Dupuis, N. F.; Bowers, M. T. *Int. J. Mass Spectrom.* **2009,** *287*, 46.
(7) Warnke, S.; von Helden, G.; Pagel, K. *Proteomics* **2015,** *15*, 2804.

B4 Outlook

The NFGAIL peptide was the object of several molecular modeling simulations regarding the amyloid formation process and several theories about structural pattern and the impact of single amino acid side chains are based on the NFGAIL sequence.[211-216] We did generate some of the first insights based on experimental data about the structural identity and the charge contribution of the early oligomers during the growing phase of amyloids and could thereby approve some of the modeling theories. Prospectively, the NFGAIL model could be used to answer further, more detailed questions like the controversial discussed influence of Phe23 on hIAPP and the orientation of the Ile side chain within the cross-β-pattern.[212, 215-216, 218] Therefore the impact of fluorination could be used. For example a pentafluorophenylalanine could substitute the phenylalanine residue and thereby switch the π-stacking preference of the aromatic ring. If π-stacking is one driving force for the amyloid formation, would the substitution result in a different pattern. If only the hydrophobicity of the phenylalanine in interaction with the isoleucine residue causes the formation of β-sheets, fluorination should not influence the general aggregation behavior. In NFGAIL mixtures with $NX_{aa}GAIL$ where Phe is substituted by *p*-monofluorophenylalanine or *2,3,5,6*-tetrafluorophenylalanine the π-stacking would be directed into an edge-to-face direction and thus could probably result in an inhibition of aggregation or in generation of completely new structures. Also the isoleucine residue could be fluorinated to increase the hydrophobicity and enable further analytical methods like solid state ^{19}F NMR spectroscopy.

Part C: A Light-sensitive Cage for Aggregating Peptides

C1   Abstract

Model peptides have great potential for biochemical and medical applications *in vivo*, but their utility is limited due to their inefficient penetration of cells.[460] Furthermore, peptides have low metabolic resistance against proteases for example.[461] Conformational switching of peptides has been extensively studied by the use of model peptides and many insights have been gained, but *in vivo* most investigations have concentrated on more complex natural peptide sequences, which are difficult to analyze and synthesize. The 26-residue model peptide VW18 was developed by Koksch *et al*. and undergoes a conformational transition from an α-helical coiled-coil structure to amyloid fibrils. VW18 was already used for various types of studies[18, 58, 72, 462], thus it is an promising candidate for the engineering of a new cell delivery system. Koschek, Rademann, Brandenburg and Koksch designed a methodology based on VW18 for a cell delivery system for model peptides based on a light-sensitive Bhc cage with a linker towards a cell penetrating polymer as the carrier. The polymer should transport the peptide inside the cell and thereby inhibit the conformational change to amyloid fibrils or completely hinder the folding of the peptide. The polymer could be shown to protect the peptide against proteases.[461, 463-464] To link the polymer to the peptide and enable a traceless release of VW18, the coumarin derived caging compound Bhc was modified with a cysteine residue as thiol donor for a native chemical ligation [NCL] or maleimide coupling. The methylhydroxy group of Bhc was activated as a N-hydroxysuccinimidyl [NHS] ester and different methods for coupling to a selectively unprotected lysine side chain of VW18 on the solid support or to a short reference peptide were tested. Koschek described the preliminary results in her doctoral thesis[465], but many issues remained unresolved. Thus, all organic and bioorganic synthetic steps were optimized. Furthermore, the impact of the Bhc coupling on two different fluorescent dyes was tested. Different from the

pioneer studies, the fluorescent dye cyanine 5 [Cy5] was introduced, to enable ThT assays and confocal laser scanning microscopy with the modified peptides *in vivo*.
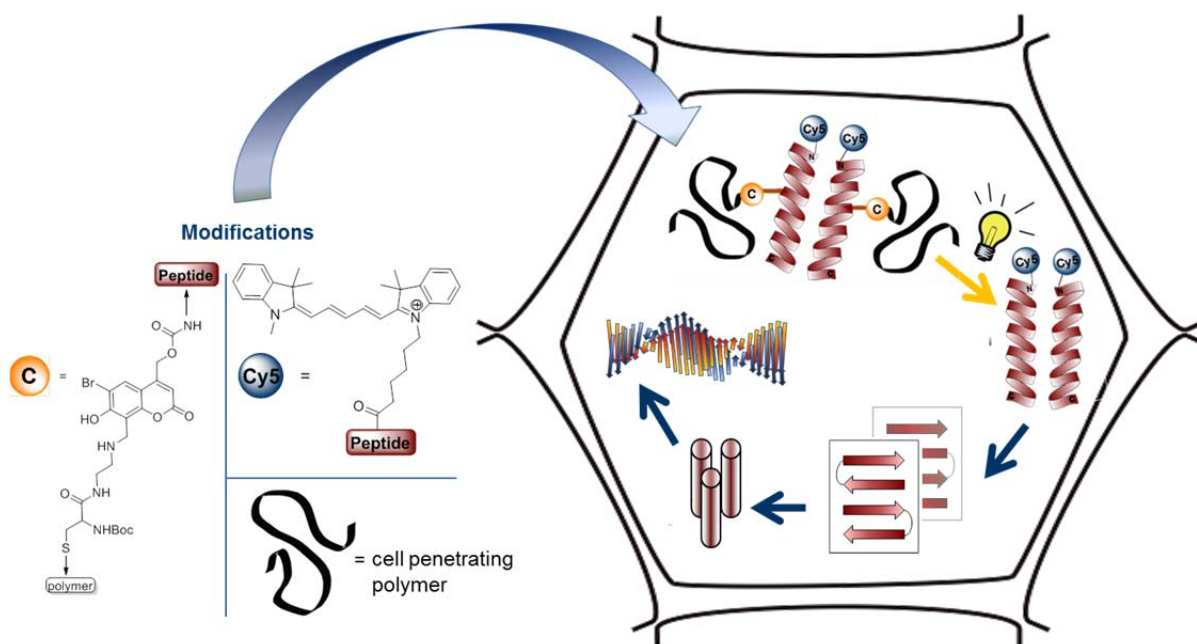
Figure C1: Project design for a cell delivery system for peptides like the amyloid forming coiled-coil peptide VW18, with a light responsive coumarin-based Bhc linker to connect the peptide with a cell penetrating polymer.

C2   Project Design

Peptides

The coiled-coil model peptide VW18K9 can be stabilized by hydrophobic core packing of a leucine zipper *via* heptad positions **a** and **d** (**Figure C2**). Further, intermolecular electrostatic coulomb interactions between oppositely charged lysine and glutamic acid residues in positions **g** and **e** direct the coiled-coil strands into parallel orientation. Compensating charged amino acids in positions **b** and **c** balance the net charge and additionally stabilize the helix strand intramolecular. Serine residues in position **f** tend to increase the solubility of the peptide. To make the peptide prone to amyloid formation, three valines in sequence positions 3, 13 and 14 are introduced. In addition, in position 9 one lysine with 4-methyltrityl [Mtt] protecting group is incorporated to enable site specific deprotection on the solid support.[466] To increase the structural influence of caging, position 9 is located directly in the middle of the sequence and next to the hydrophobic domain LVVL of VW18, which is the driving force for β-sheet formation.
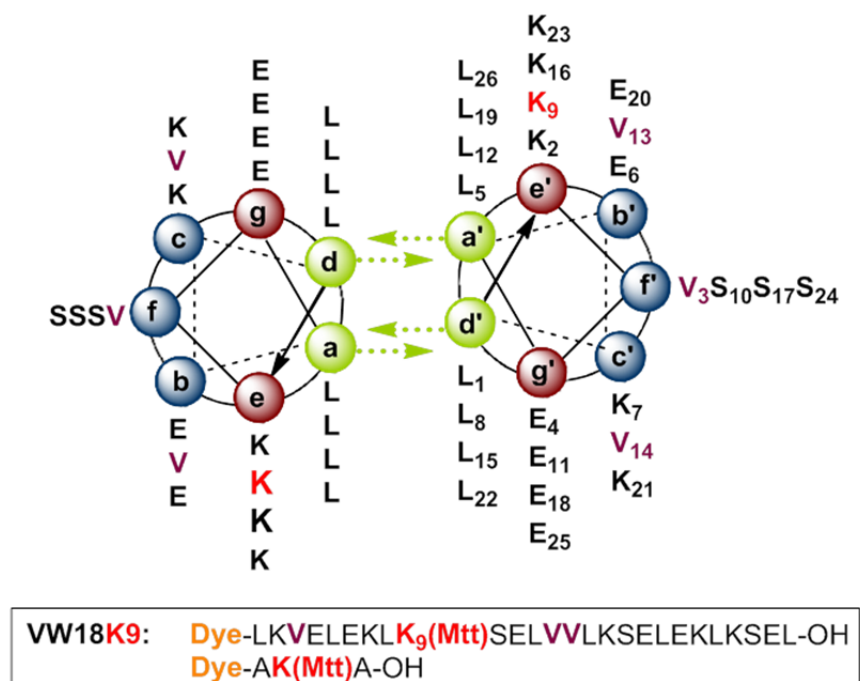
Figure C2: *Helical wheel projection and amino acid sequence of the 26-residue model peptide VW18K9 and sequence of the test peptide. Incorporated valine residues are highlighted in purple and Mtt protected lysine in position 9 is highlighted in red.*

The fluorescent dye is linked *via* a flexible aminohexyl spacer to the N-terminus of the peptide facilitating its purification and visualization inside the cell.[467] The spacer is necessary to decrease the interference of the large hydrophobic dyes with the peptide backbone, thus hindering the conformational behavior. A second short sequence constists of two alanines that surround the Mtt protected lysine and the dyes, which are coupled to the N-terminus without a spacer. The short peptide was designed as control for the cell studies and cost efficient target for test reactions.

Cys-Bhc-NHS Linker

Coumarins can interfere with peptide conformation, hindering their assembly by intercalating into the hydrophobic core. For example, they have been identified to inhibit amyloid formation of Alzheimer's disease related Aβ peptide.[468] Hence, coumarin was chosen as the caging compound to link the cell penetrating polymer to the peptide. The linker was modified starting from the Bhc caging molecule because of its high photolytic efficiency, the non-toxic byproducts released upon uncaging and the relatively long excitation wavelength (λ_{max}= 365 nm).[297, 337-340]

The alkylhydroxy group of Bhc was activated as NHS ester[469] (**Scheme C1**) to form a carbamate bond between the unprotected amine group of the side chain of a lysine[466].
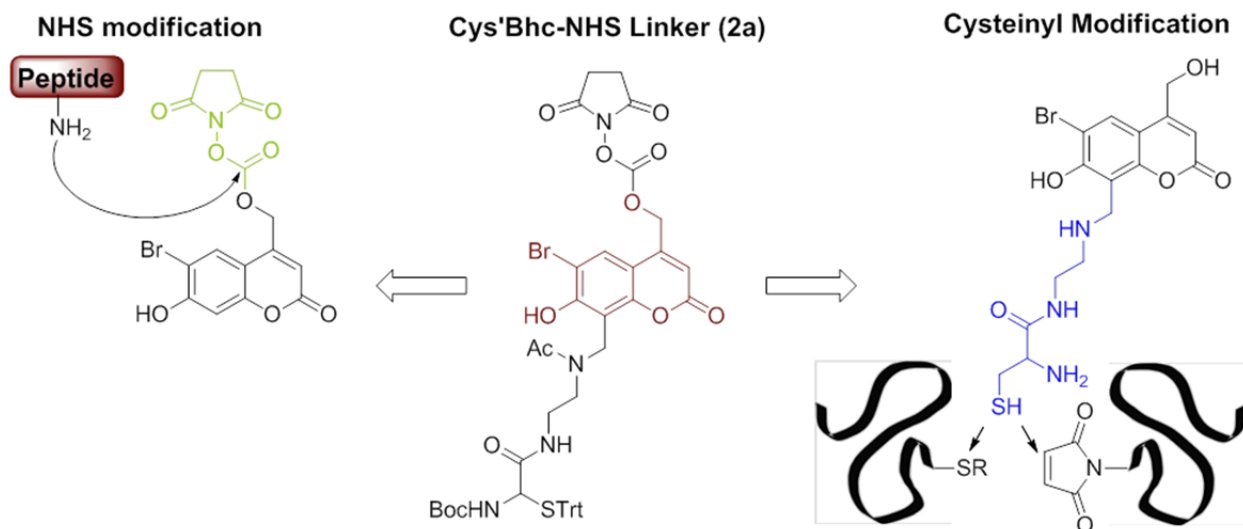
Scheme C1: *Overview of the linker design for the light sensitive Bhc-based linker between peptide and polymer. The Bhc caging part (red) is supplemented by a carbamate that links to the peptide through NHS activation (green) and a cysteinyl linking part (blue) in position 8 of Bhc as donor for thiol coupling to a maleimide or to an activated thioester via NCL.*

Furuta *et al*. described the uncaging in carbamate linked Bhc cages with the decarboxylation of the carbamate as the rate-limiting step, which occurs in the range of subseconds and the photorelease of the Bhc in nanoseconds (**Scheme 5**, section 3.1).[338] The uncaging kinetics can be monitored *via* fluorescence spectroscopy and the decarboxylation using IR.[312] To realize the transport of the caged peptide into cells and to protect the peptide against degradation, including proteolytic digestion, a cell penetrating polymer like *N*-(2-hydroxypropyl) methacrylamide [HPMA][464-465] or sugar-based dextrans[470] can be linked to the Bhc cage. Therefore a cysteinyl linker presenting a thiol group is incorporated into position 8 of the Bhc ring system. The thiol provides the possibility for orthogonal coupling *via* NCL[471-473] or maleimide-thiol reaction[474]. Furthermore, alkylation of positon 8 was shown by Adamczyk to lower the pK_a for the uncaging of coumarins to physiological conditions (pH ~7).[475]

C3   Results

C3.1   Comparison of TAMRA and Cy5

Koschek *et al.* described the labeling of VW18K9 with the rhodamine based fluorescent dye TAMRA. TAMRA has an approximate maximum excitation wavelength of 546 nm and an emission maximum at 579 nm.[476] The authors used an aminohexyl spacer to reduce hydrophobic or steric interactions with the peptide. Both spacer and TAMRA were coupled with Fmoc-mediated SPPS strategy under harsh conditions (8 eq. of *O*-(7-azabenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium-hexafluorophosphate [HATU] and *N,N*-diisopropyl-ethylamine [DIPEA]). The folding of TAMRA-VW18-OH was monitored with CD spectroscopy and TEM. The peptide's conformation changed from α-helical coiled-coil to β-sheet structures during 24 hours, a behavior that had been previously reported[58]. TEM images showed amyloid-like twisted fibers, also comparable to VW18[58]. Nevertheless, a ThT fluorescence assay of the aggregation pathway was not possible, because TAMRA quenched the fluorescence of ThT at 485 nm. Thus, the aggregates visible in TEM could not be reliably identified as amyloids.[465, 467] This thesis describes how the coupling of the dye onto the peptide was optimized, thereby minimizing side reactions and increasing the coupling efficiency (**Figure C3**). Therefore, a NHS activated form of the dye was coupled with only 1.2 eq. with respect to the peptide and without coupling reagents in a single coupling step and a yield above 95%. The reaction had to be performed in the absence of light and water and under mild basic conditions (DIPEA). The presence of coupling reagents like HATU, 1-hydroxybenzotriazole [HOBt] or 1-hydroxy-7-azabenzotriazole [HOAt] decreased the coupling efficiency. Nevertheless, several attempts to couple Cys'Bhc-NHS (**12a**) to the TAMRA labeled VW18 failed, resulting in complete or partwise dissociation of TAMRA or decomposition of the Bhc linker. A test coupling of Cys'Bhc-NHS (**12a**) to TAMRA-AKA* (**14**) peptide on the solid support resulted in 73% caged TAMRA-AK(Cys'Bhc)A-OH (**19**), but the product was not stable for even 12 hours.
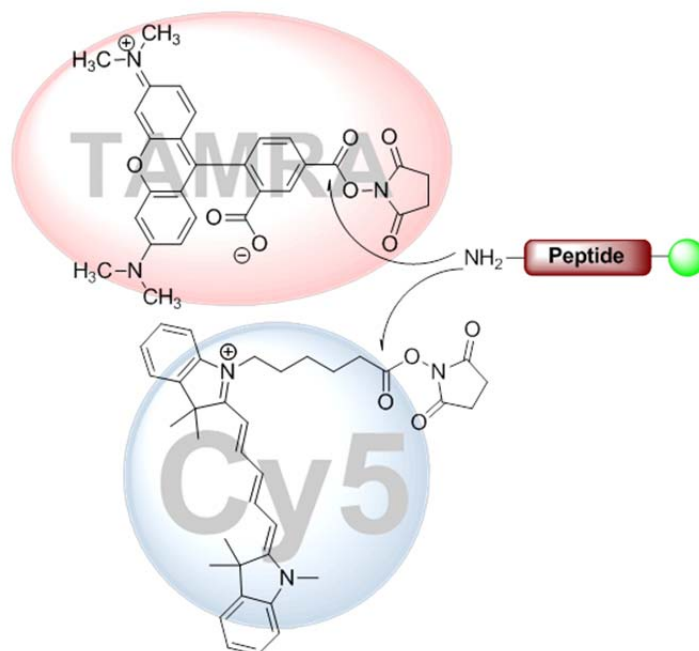
Figure C3: *Two fluorescence dyes in NHS ester form for N-terminal peptide labeling. 5(6)-carboxytetramethylrhodamine [TAMRA] depicted on top and cyanine 5 [Cy5] below.*

It could be hypothesized that a small degree of spectral overlap between Bhc and TAMRA could lead to Förster resonance energy transfer and the subsequent quenching of TAMRA. Furthermore, it became apparent that TAMRA was interfering with the peptide itself when stored on the solid support. Several TAMRA labeled, freeze-dried peptides, stored at -20°C on solid support or purified after full cleavage, where decomposed after only some weeks of storage. This could probably be derived from the sensitivity against basic pH of TAMRA.[477] Because of its hydrophobicity TAMRA is further prone to dye-dye aggregation in aqueous solutions.[478] An alternative option for peptide labeling is the Cy5 dye which is also available as a NHS ester. It has an approximate excitation maximum at 646-650 nm and an emission maximum at 662-670 nm.[479] Cy5 is widely used as fluorescence marker of peptides and other biomolecules *in vivo*. It is commonly detected with confocal laser scanning microscopy [CLSM], but also in FRET experiments.[480-482] Cy5 conjugates are relatively stable if stored in the dark at -20°C. The red shift of the spectral pattern, in comparison to TAMRA, should minimize the interaction between Bhc or ThT and Cy5, thus facilitating coupling of the Bhc linker and the monitoring of amyloid formation with a ThT assay. The coupling of Cy5-NHS onto to the N-terminus of the peptide including the spacer group was successful under the same reaction conditions as earlier

described, using 1.2 eq. dye with respect to the peptide and basic conditions (DIPEA) in either DMF or DCM to yield 99% of the expected product.

For both dyes the usage of hydride ion donors like scavengers such as TIPS to support the full cleavage of the peptide from the solid support or for site specific Mtt cleavage is not possible.

C3.2 Cy5-VW18K9(Ac)-OH

To test the influence of Cy5 labeling on the aggregation behavior of VW18, a small amount was capped with acetic anhydride to protect the side chain amine of lysine 9. After Fmoc-deprotection of the N-terminus, the coupling of Cy5-NHS was performed with high yield as described before. The peptide was fully cleaved with 98% TFA and 2% water, before purification by preparative HPLC. To test the structural behavior of Cy5-VW18K9(Ac)-OH it was dissolved in 1,1,1,3,3,3-hexafluoro-2-propanol [HFIP] to prepare a stock solution and to fully denature the sample. The concentration of the stock solution was determined using the concentration curves for Cy5 (section E4.24). For CD and ThT assays (**Figure C4**), aliquots of the stock solution were evaporated and dissolved in 10 mM phosphate buffer to obtain 100 µM peptide solutions, containing 20 µM ThT for fluorescence spectroscopy (detailed procedure in sections E1.8 for the ThT assay and E1.10 for CD spectroscopy). The CD or ThT measurements began directly after pH adjustment of the buffer to 7.4.
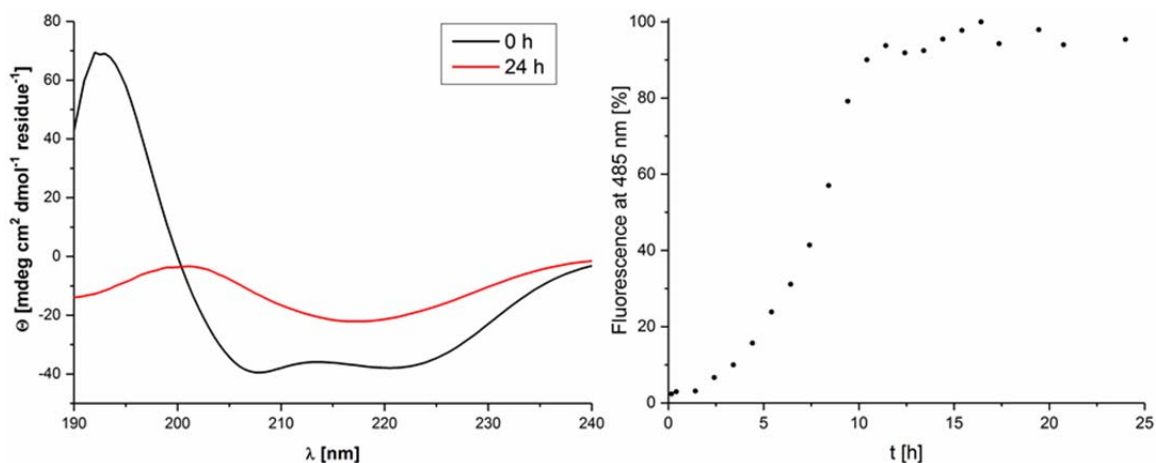
Figure C4: *CD spectra (left) and ThT fluorescence assay (right) of 100 µM Cy5-VW18K9(Ac)-OH peptide in 10 mM phosphate buffer pH 7.4.*

The CD spectrum shows a maximum at 195 nm and two minima at 208 and 222 nm with equivalent intensities, which suggests an α-helical coiled-coil conformation.[127] After 24 hours the minimum is changed to 218 mm and a maximum at 200 nm, which is characteristic for a β-sheet conformation[127]. The lower intensity of the signal at 24 hours can be explained by precipitation of the aggregates that leads to a reduction in peptide concentration in solution. The ThT fluorescence assay shows a short lag time of 1.5 hours and a relatively slow growth phase of 8.5 hours with a characteristic sigmoidal curve resulting in a saturation phase after 10 hours. The increase in ThT fluorescence is a strong hint for the presence of amyloid fibrils.[117-118, 121] Both, CD and ThT spectroscopy of Cy5-VW18K9(Ac)-OH are in good agreement with the conformational behavior of VW18[58, 483]. That implies Cy5 labeling does not affect the conformational behavior of VW18 to a significant extent and therefore could be used as an alternative to TAMRA.

C3.3   Synthesis of the Alkylated Bhc Cage

Outline for the Synthesis of the Alkylated Bhc Cage

The synthesis of the Bhc-Cl **3** caging moiety starts with a Pechmann condensation proceeding from 4-bromresorcinole **1** and 4-chloroacetoacetate **4** (**Scheme C2**). Hydrolysis of the chlorine leads to the hydroxymethyl functionality of Bhc-OH **4**.
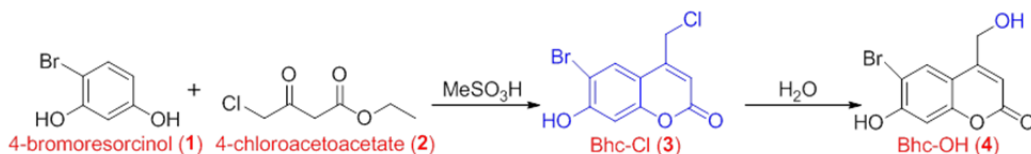
Scheme C2: *Synthesis the coumarin derived Bhc cage Bhc-Cl **3** and its hydrolysation to Bhc-OH 4.*

Bhc-Boc (**5a**) was achieved in a Mannich-like reaction of Boc-ethylenediamine and paraformaldehyde with Bhc-OH (**4**). Afterwards, the TFA salt **6** is generated by Boc deprotection of **5a** (**Scheme C3**).
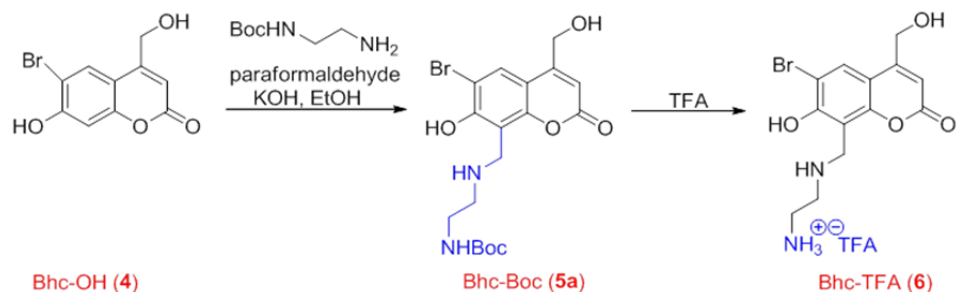
*Scheme C3: Mannich-like alkyltaion of Bhc-OH **4** and Boc deprotection to gain Bhc-TFA **6**.*

Synthesis of Bhc-Cl (3)

4-bromresorcinol (**1**) and 4-chloroacetoactetate (**2**) were combined in a Pechmann-condensation in presence of methanesolfonic acid (**Scheme C2**).[314-315] The reaction was performed with quantitative conversion following an improved protocol from Hagen *et al.*.[125]

Synthesis of Bhc-OH (4)

Furuta el al. performed the hydrolysis to Bhc-OH (**4**) while they did suspend **9** in water and refluxed it for 14 hours to yield 98% **4** without further purification (**Scheme C2**).[315] The repetition of these procedure resulted in 4% conversion to the desired product (HPLC area percent). The reaction was tested at different pH values (5, 7, 9, 10). With respect to Le Chatelier's principle[484], could the formation of HCl during the synthesis and the thereby generated acidic pH shift the reaction equilibrium to the starting reagents. To increase the yield, the reaction mixture was three times per day titrated to a pH 9.5-10 using 1 M NaOH. The reaction mixture could be hot filtrated, because only compound **4** was soluble in hot water. The starting material could be reused. Before workup, it was crucial to titrate the aqueous solution of **4** to pH 7 to prevent side reactions. Following freeze-dying, the crude mixture was dissolved in pure acetone and purified with flash chromatography to gain 19% isolated yield.

Synthesis of Bhc-Boc (5a)

To modify Bhc derivative **4** with an aminoalkyl side chain in position 8, like described by Adamczyk[475] and Furuta[340] and optimized by Hagen *et al.*[314], **4** was dropwise added to a cold and stirring mixture of paraformaldehyde and N-Boc-ethylenediamine (**Scheme C3**). The aminoalkylation mechanism starts with the formation of an imin in a Mannich-like reaction.[485]

The added Bhc-OH reacts with the iminium ion to the desired aminomethylated Bhc-Boc (**5a**). The product could be isolated after precipitation at -20°C and recrystallization with 40% yield.

Synthesis of Bhc-TFA (6)

For Boc removal **5a** was dissolved in a mixture of TFA and DCM and stirred (**Scheme C3**). After removing the solvents, the TFA salt was mixed with water and freeze-dried. In contact with air or traces of water, the Bhc-TFA salt becomes a viscous oil which is difficult to divide. Thus it is recommended to aliquot and store **5a** not **6**. Nevertheless, both compounds were stable for years when stored as solid under exclusion of light at -20°C. The isolated product **6** was achieved with quantitative yield.

C3.4 Synthesis of the Cysteinyl Functionality

Outline for the Synthesis of the Cysteinyl Functionality

A cysteine **7** was activated with an ester bond to a pentafluorophenol **8** as leaving group to gain **9**. The cysteinyl functionality in **10** was realized with an amide coupling between **9** and **6** (**Scheme C4**).
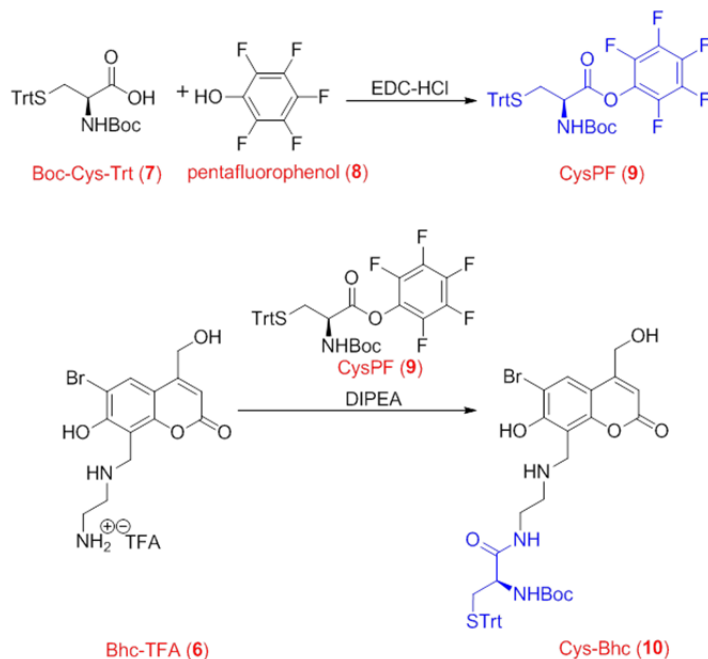
Scheme C4: *Synthesis of the pentafluorophenol ester activated cysteine CysPF **9** and its reaction to Cys-Bhc (**10**) by an amide formation between CysPF **9** and Bhc-TFA **6**.*

Synthesis of CysPF (9)

The esterification of a *N*- and S-protected cysteine (**7**) with a pentafluorophenol (**8**) was supported from 1-ethyl-3-(3-dimethylaminopropyl)carbodiimid [EDC] hydrochloride to gain a pentafluorophenol ester of the cysteine (**Scheme C4**). EDC preactivates the carboxylic acid and could be easily seperated by washing with water. After the reaction was completed, the solvent was evaporated. The difficulty during this procedure was the voluminous foam which suddenly conglomerates when only small amounts of solvent were left. The crude product was purified with flash chromatography using DCM as eluent. The product **9** could be obtained in 56% isolated yield.

Synthesis of Cys-Bhc (10)

The amide condensation between CysPF **9** and Bhc-TFA **6** was realized with Schlenk chemistry under basic conditions (DIPEA) in dry DMF (**Scheme C4**). The reaction had to be stirred at 50°C for several hours to push the reaction equilibrium to the product site. Therefore, HPLC reaction control was useful. If the reaction was heated to more than 70°C the Bhc was rapidly decomposing. Thus, after the reaction was completed, the solvent was removed in vacuum without further heating. Purification was realized with preparative HPLC to gain highly pure **10** with 21% yield. In addition, both left over starting compounds could be regained.

C3.5 NHS Activation of the Cysteinyl Cage

Outline for the NHS Activation of the Cysteinyl Cage

The NHS activation is realized proceeding from the fully protected cysteinyl linker **11** which was previously N-protected with an acetyl group in the diethylamine functionality of **10** (**Scheme C5**). The activation was first optimized with Bhc-Boc **5a** with different solvents and bases to achieve a 99% yield of **5b**. The obtained reaction conditions were later transferred to the activation reaction of **12a**. Furthermore, the pH sensitivity of **12a** was tested in pure TFA.
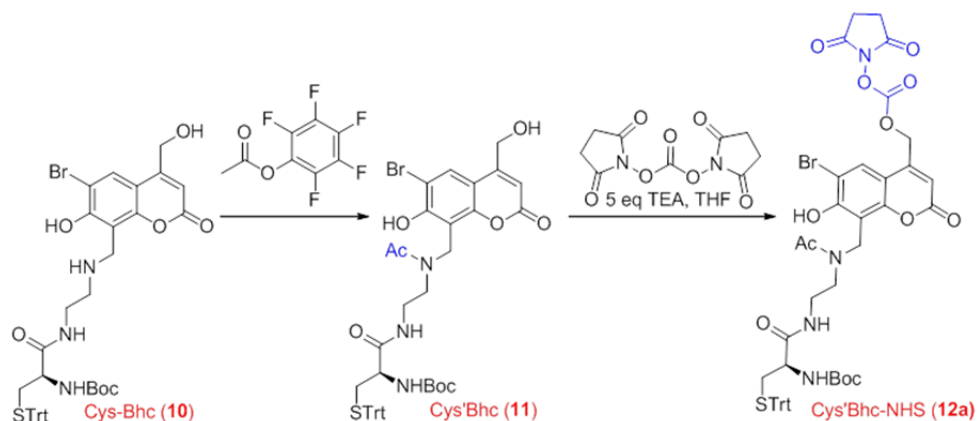
Scheme C5: *Synthesis of the cysteinyl Bhc activated as NHS ester (**12a**) and previously capped (**11**), starting from Cys-Bhc (**10**).*

Synthesis of Cys'Bhc (11)

Adamczyk *et al*. described the necessity to protect the secondary amine of the diethylamine part in the side chain of the linker before the aminoalkylated Bhc can be used as cage under physiological pH. They mentioned a method with pentafluorophenyl acetate to selectively protect primary and secondary amines (**Scheme C5**).[475] The reaction was performed in the absence of a base to prevent the acetylation of the hydroxyl groups and again with Schlenk chemistry in dry DMF with heating to 50°C to push the reaction equilibrium into the desired product direction. The crude was purified with preparative HPLC resulting in 70% isolated yield of **11**. Cys-Bhc **10** could be recovered and reused.

Optimization for NHS Activation with Bhc-Boc (5b)

NHS activation of **11** as described by Koschek[465] with an excess of *N,N'*-succinimidyl carbonate [DSC] with Schlenk chemistry in dry tetrahydrofuran [THF] and with triethylamine [TEA] in the absence of light, did not result in creation of the desired product **12a**. To develop a synthetic method for the activation, derivative **5a** was used because of its easy handling (**Scheme C6**).
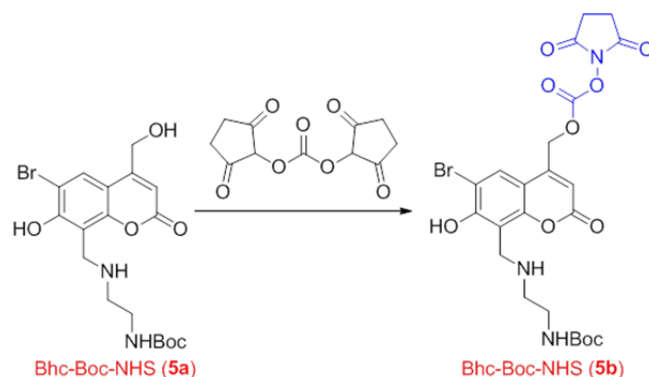
Scheme C6: *Testreaction for the NHS activation using Bhc-Boc (5a).*

Several solvents and bases were tested in small batches to evaluate the number of byproducts and the yield of **5b** after 24 hours (**Table C1**). The NHS activation of **5a** with DSC in THF resulted in 55% product formation. Solubility was identified as one major challenge for the reaction. The chaotropic base guanidinium hydrochloride [GdmHCL] leads to relatively high yields, but its low solubility in organic solvents reduces its utility. Piperidine is a liquid base, but it led to the formation of byproducts and only small amounts of **5b**. It was observed that the presence of acetonitrile [ACN] supports NHS ester production, but neither **5a** nor **12a** were soluble in pure ACN. Thus, the reaction could be optimized with a 1:1 mixture of ACN and THF in combination with freshly distilled TEA to yield 99% conversion of **5a** to **5b**.

Table C1: *Optimization series for NHS activation with **5a** and different solvents and bases.*

solvent	base	comments	product [%]
THF	TEA	no byproducts	55
DMF	TEA	no byproducts	3
ACN	TEA	low solubility/ no byproducts	78
TEA	TEA	low solubility/ no byproducts	26
H_2O/THF	TEA	suspension	79
H_2O/THF	GdmHCl	suspension/ 1 byproduct	20
THF	GdmHCl	no byproducts	80
ACN	GdmHCl	low solubility/ no byproducts	99
MeOH	GdmHCl	low solubility/ 1 byproduct	24
THF	piperidine	no educt left/ 1 byproduct	8
ACN	piperidine	low solubility/ 1 byproduct	7
ACN/THF (1:1)	TEA	THF was added after 5 h/ 3 byproducts	80
ACN/THF (1:1)	TEA	no byproducts	99
ACN/THF (1:1)	-	no reaction	0

Synthesis of Cys'Bhc-NHS (12a)

All kind of handling and synthesis was realized without or in orange light to prevent uncaging. According to the results of NHS activation towards Bhc-Boc (**5b**), Cys'Bhc **11** was dissolved in a 1:1 mixture of ACN and THF. DSC was added in the presence of TEA. The mixture was heated to 50°C and the reaction was monitored by analytical HPLC. If the reaction failed to progress, additional DSC was supplemented. The maximum conversion ratio observed was 3:7 (**11**:**12a**). It turned out to be crucial to separate the product from residual DSC before continuing with the coupling **12a** to the peptides. Thus, the crude product was purified with preparative HPLC. **12a** could be obtained in 16% isolated yield. Observations indicated that **12a** was decomposing and in addition cleaving of the NHS ester under preparative HPLC conditions or while stored in solution afterwards. If stored as a freeze-dried solid under exclusion of light at -20°C, **12a** was stable for several weeks.

Deprotection and Stability Test of Cys'Bhc-NHS (12b)

2a was stirred for 2 h in pure TFA. The acid was removed in vacuum and the crude was purified with analytical HPLC to gain 68% isolated yield of **2b**. NHS ester and the Boc- and trityl-protecting groups were removed during acidic treatment. The cysteinyl Bhc part was resistant against the conditions.

C3.6 Caged TAMRA-Peptides

Side Chain Deprotection of Dye-AKA*

Mtt deprotection of the lysine side chain of Dye-AK(Mtt)A* on the solid support was first tested as described in the literature with 2% TFA and 2-5% TIPS[104, 107, 128], but the bleaching of the dye or complete release of the peptide was observed. It could be assumed, that a hydrogen ion from the scavenger TIPS was undergoing a nucleophile reaction with the double bonds of the dye, resulting in a loss of color (**Scheme C7**). The hydrogen ion could further react in a benzyl ether reductive cleavage reaction with the linker between peptide and the solid support.[486] To prevent these kinds of side reactions, no scavenger was used subsequent to the deprotection of the lysine side chain. Thus, the usually very fast reaction had to be repeated several times until no yellow color was visible, since Mtt cleavage product is yellow. The resulting side chain free Dye-AKA* was further caged with the coumarin linker. Therefore the reactor was covered and light

exposure was minimized by working in the dark or in the presence of orange light. This procedure was repeated for all described peptides.
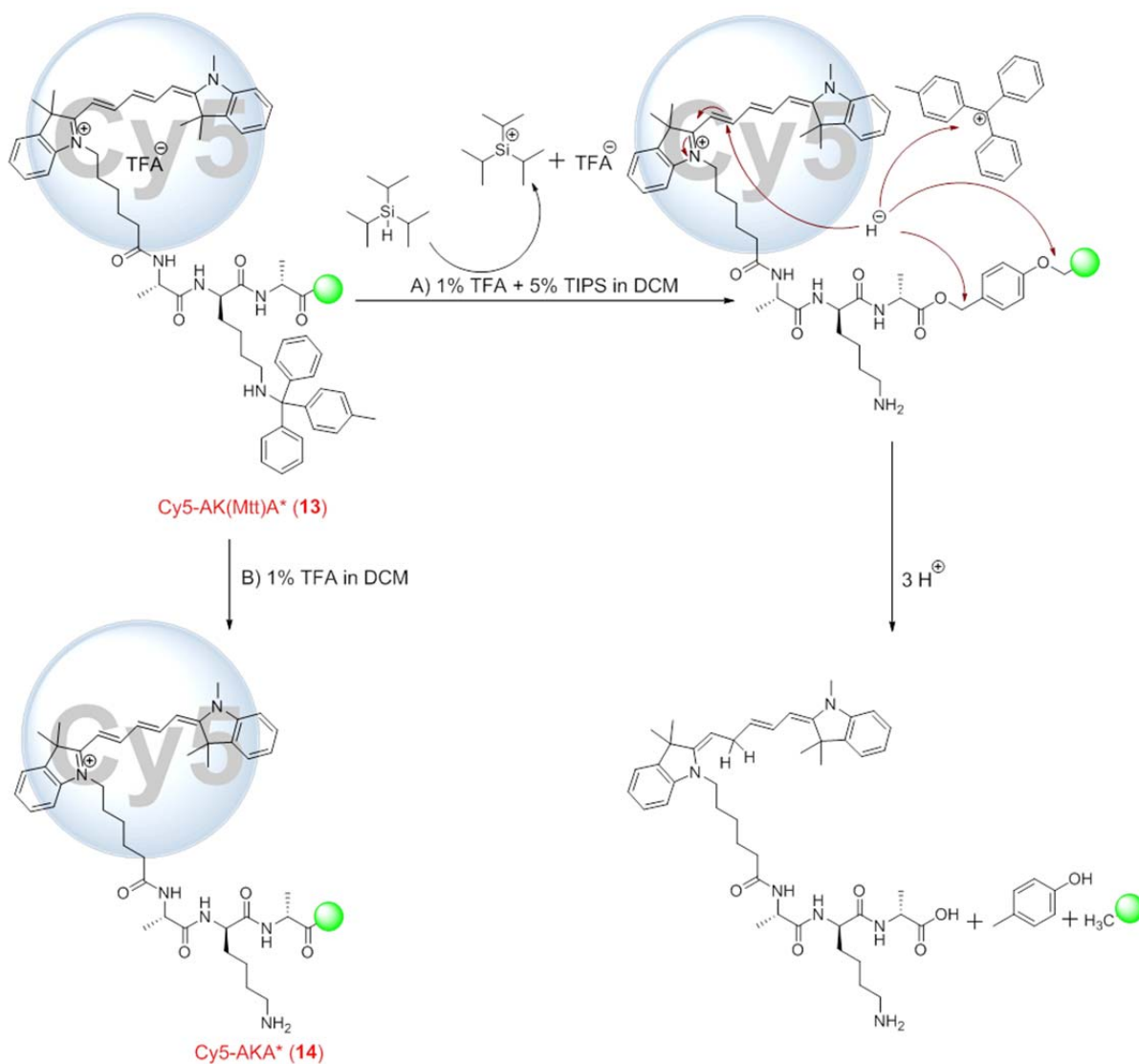
Scheme C7: *Mechanisms for the lysine side chain deprotection with possible side reactions if TIPS is used.*

Synthesis of the caged TAMRA-AK(Bhc)A-OH (15)

Again the reaction for caging was tested with the less sensitive **5a** and **5b** coumarins (**Scheme C8**). The reaction was first performed in a one-pot synthesis like described from Koschek.[465] Therefore Bhc-Boc **5a** was mixed with DSC and TEA in ACN/DCM. This reaction cocktail was added to the resin with side chain free TAMRA-AKA* **14** peptide. After 12 hours a complete conversion from TAMRA-AKA* (**14**) to TAMRA-AK(NHS)A* (**16**) occurred. The excess of

strong activation reagent DSC led to a preferred reaction with the free amine to the N-succinimidyl carbamate.
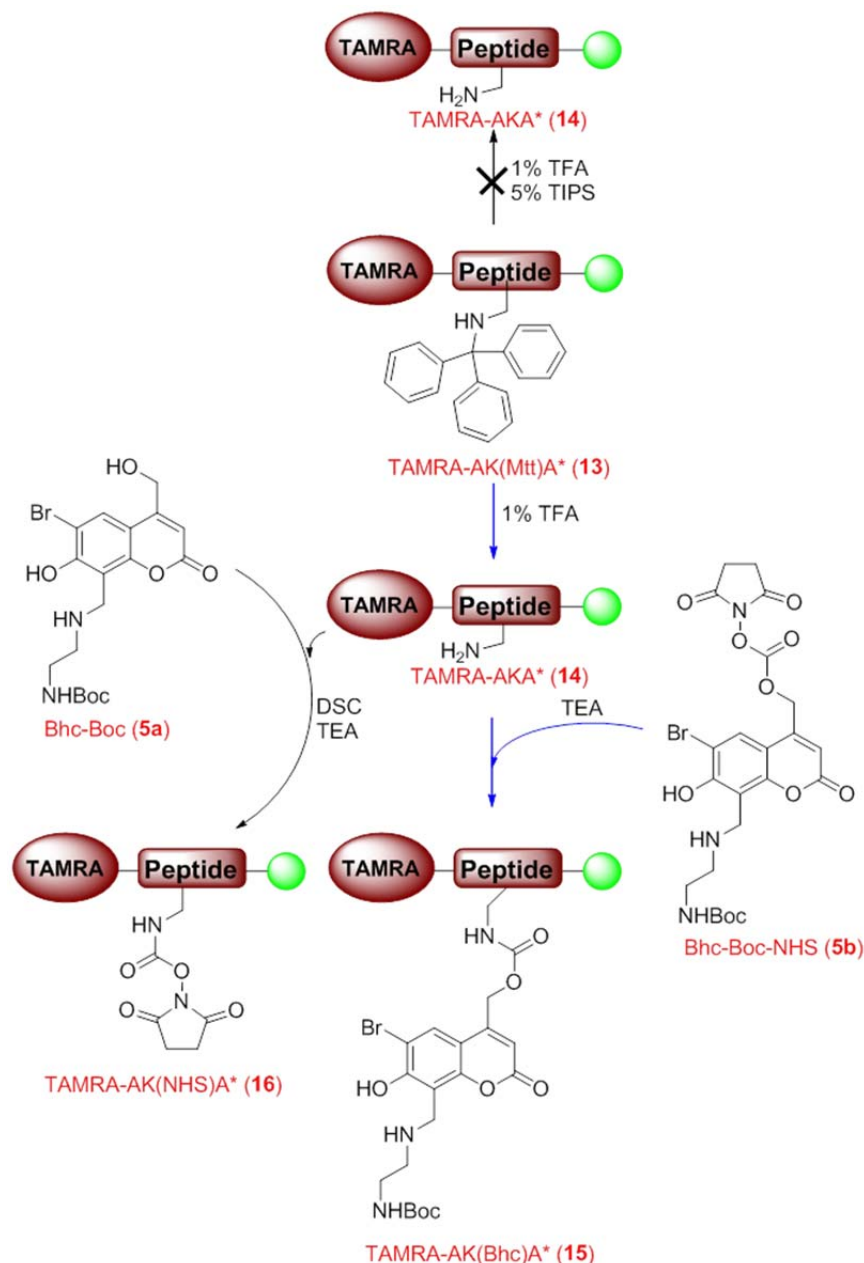
Scheme C8: *Synthesis of caged TAMRA-AK(Bhc)A* (15) on the solid support. The reaction starts with the deprotection of the lysine side chain, which was tested with different reaction mixtures. Afterwards the side chain unprotected TAMRA-AKA* 14 was in a one-pot reaction reacted with Bhc-Boc (5a), resulting in the NHS carbamate of the lysine side chain. In a second reaction, the previous activated Bhc-Boc-NHS (5b) was reacted with the peptide to yield the desired TAMRA-AK(Bhc)A* 15. The successful reaction pathway is highlighted in blue.*

In an alternative reaction pathway **14** was reacted with the previously NHS activated Bhc-Boc-NHS (**5b**). Remaining DSC was separated by preparative HPLC before **5b** was dissolved in a

mixture of ACN and DCM and added to the peptide on the solid support. After 12 hours reaction time, the resin was washed and a test cleavage with 100% TFA, in the absence of TIPS scavenger, revealed 68% conversion to the desired TAMRA-AK(Bhc)A-OH **15**. The product of the test reaction was not refurbished.

Synthesis of TAMRA-AK(Cys'Bhc)A-OH (19)

13 was side chain deprotected as described before with only 1% TFA. Afterwards the caging of **14** with Cys'Bhc-NHS **12a** was tested in the presence of DIPEA and only traces of DSC, resulting in a complete conversion to TAMRA-AK(NHS)A* (**16**) (**Scheme C9**).

In a second approach **12a** was dissolved in a mixture of ACN and DCM before being added to side chain unprotected Fmoc-AKA* (**17**) (**Scheme C9**). A test cleavage with 100% TFA revealed 87% conversion to Fmoc-AK(Cys'Bhc)A-OH (**18**). After Fmoc-cleavage TAMRA-NHS ester was dissolved in DMF with DIPEA and added to the peptide on the solid support. After test cleavage the desired product TAMRA-AK(Cys'Bhc)A-OH (**19**) could be verified in a 41% yield. These observations indicate an interaction between both dyes resulting for example in dye-dye aggregation like described for TAMRA in the literature[478] or uncaging of the Bhc linker. The close proximity between both dyes without a spacer beside alanine could further facilitate these interactions.
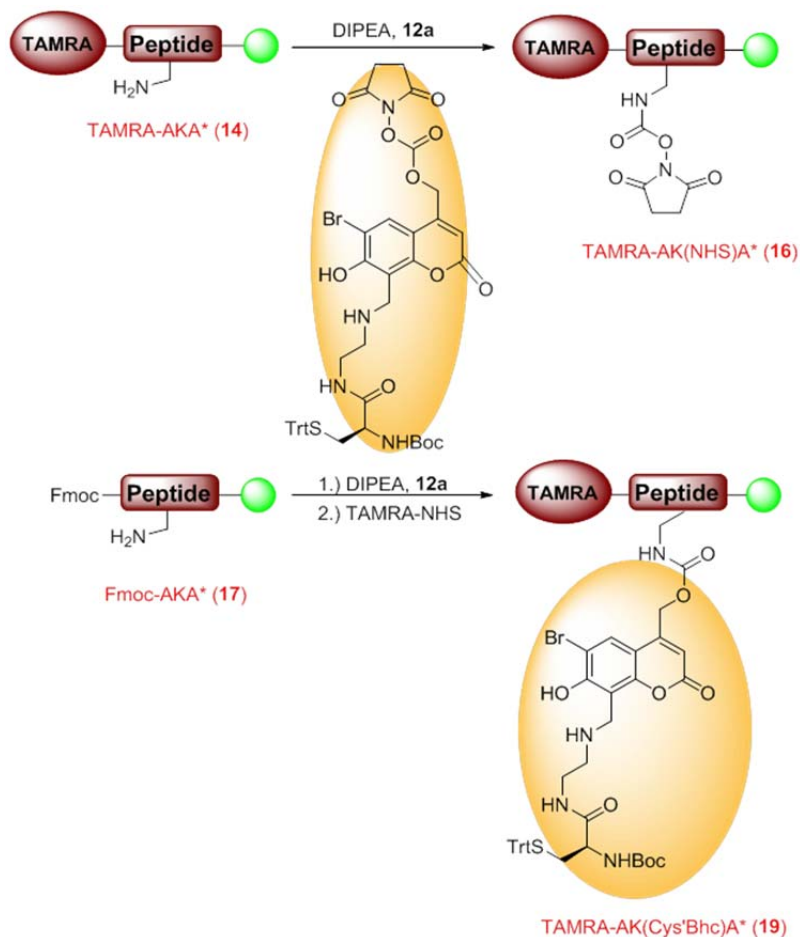
Scheme C9: *Synthesis of caged TAMRA-AK(Cys'Bhc)A-OH **19**. Reaction of **12a** with TAMRA-AKA **14** resulted in TAMRA-AK(NHS)A* **16** side product. Caging of Fmoc-AKA* and subsequent labeling with TAMRA-NHS yields **19**.*

Synthesis of TAMRA-Ahx-VW18K9(Cys'Bhc)-OH (23)

In parallel to the synthesis of **19** also the caging of the longer TAMRA-Ahx-VW18K9* (**20**) was revised. Mtt deprotection was performed as described above. Cys'Bhc-NHS **12a** was dissolved in a mixture of ACN and DCM including DIPEA and was directly added to the resin. After 12 hours the reaction mixture contained only compound **11** without an NHS ester. In a test cleavage of the peptide in 100% TFA neither the desired product **23** nor TAMRA-Ahx-VW18K9-OH (**20**) could be detected. Again, interaction between the dyes could be the reason for this outcome. The reaction was tested with different kinds of solvents (DMF, THF, DCM) to investigate the impact on the peptide or the resin, but only the number and HPLC retention times of the unidentified side products changed.

Based on the successful synthesis of **19** (section above) Fmoc-Ahx-VW18K9* (**21**) was reacted, by analogy, with **12a** resulting in a 75% conversion to Fmoc-Ahx-VW18K9(Cys'Bhc)* (**22**). Nevertheless, two trials of labeling the N-terminus with TAMRA-NHS as described for TAMRA-AK(Cys'Bhc)A-OH **19** resulted in complete decarboxylation of **12a** to **11** and destruction of the peptide. Several mixtures for test cleavage were tested, all resulting in the same outcome. Not even peptidic cleavage products could be identified in MS analyses. Koschek described problems with caging as well, stating only a low yield of 2.7% for **23** and a lack of reproducibility.[465] As described in section C3.1 TAMRA has certain disadvantages. That is why it was decided to change the fluorescent dye to Cy5, as described in detail below.

C3.7 Caged Cy5-Peptides

Synthesis of Cy5-AK(Cys'Bhc)A-OH (26)

By analogy to **19**, Cy5-AK(Mtt)A* (**24**) was Mtt deprotected with TFA in many repetitions. The washing solution turned green, not yellow as expected for Mtt, suggesting that some blue colored Cy5-peptide is removed during this procedure (**Scheme C7** section C3.6).

Cy5-AK(Mtt)A* (**25**)   Cy5-AK(Cys'Bhc)A-OH (**26**)
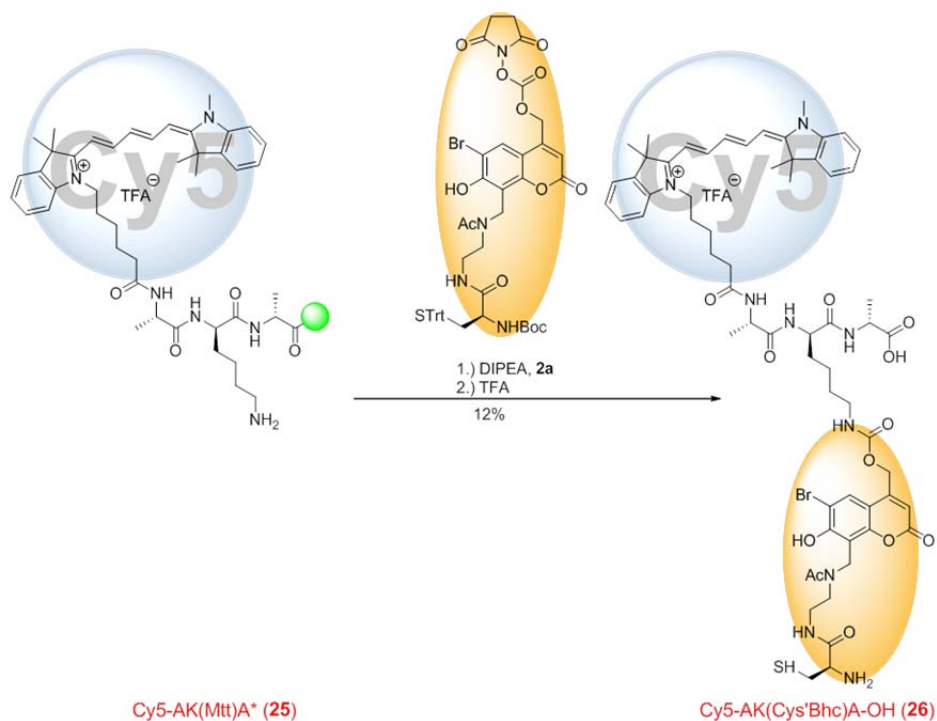
Scheme C10: *Synthesis of caged Cy5-AK(Cys'Bhc)A-OH **26**. The NHS activated **12a** was directly coupled to the side chain unprotected lysine of the peptide on the solid support.*

12a was dissolved in ACN/DCM with DIPEA and added to the resin with lysine side chain unprotected Cy5-AKA* (**25**) (**Scheme C10**). After 12 hours the reaction mixture contained only traces of **12a** and mostly **11**, as confirmed by analytical HPLC. After full cleavage and purification with preparative HPLC the desired product Cy5-AK(Cys'Bhc)A-OH (**26**) could be obtained in 12% isolated yield. After 1 week of storage as a freeze-dried solid in the dark at -20°C the purity was only 90%.

Synthesis of Cy5-VW18K9(Cys'Bhc)-OH (28)

Like described for **24** the side chain deprotection of lysine in sequence position 9 led to a green instead of yellow colored washing solution. In an analytical HPLC of the test cleavage only Cy5 labeled peptide could be obtained. Certainly, an analytical HPLC of the green washing solution did contain a high concentration of Cy5-VW18K9-OH (**27**). Thus the peptide was cleaved from the resin with only 1% TFA, resulting in a decrease of peptide loading on the resin. Nevertheless, **12a** was dissolved as common and added to the resin. After 12 hours no desired product **1a** or uncaged peptide could be identified in analytical HPLC and ESI-ToF analyzes of varying test cleavages.

In a second approach **12a** was coupled to side chain unprotected Fmoc-VW18K9* (**21**) as described before in section C3.6. After Fmoc-removal Cy5-NHS, dissolved in DMF and DIPEA, was added to the resin for 12 hours. After full cleavage the crude was purified with analytical HPLC to gain 1% isolated yield of **28**. The product was not stable when dissolved in HFIP as a stock solution.

Synthesis of Cy5-AK(Cys'Bhc)A-OH (26) in Solution

To test whether the problem during caging could be a consequence of side chain interactions of Cys'Bhc during full cleavage, the reaction was performed in solution with fully protected Cys'Bhc-NHS **12a** and purified, unprotected Cy5-AKA-OH (**25**). The reagents were dissolved in ACN with DIPEA under Schlenk conditions and stirred. After 12 hours almost no product **26** formation could be observed. Thus the reaction mixture was heated to 50°C for 4 days with regular reaction control *via* analytical HPLC. After 4 days the reaction equilibrium was shifted to a 1:1 ratio of product to Cy5-AKA-OH (**25**:**26**) and no **12a** was left. With this knowledge it

could be assumed, that the reaction kinetic was very slow and the reaction equilibrium was probably more on the side of the starting materials. Furthermore, an in situ deactivation of **12a** as competitive reaction could decrease the reaction efficiency.

C4 Discussion and Conclusion

The design for a photo cleavable linker based on the caging coumarin compound Bhc was described in detail. The linker should combine a caging binding site for a connection to a lysine side chain of a peptide and an appropriate binding site for a cell penetrating polymer which should serve as carrier for the peptide into the cell and protect it against proteolytic digestion. Thus Bhc was modified with a cysteinyl linker in position 8 to present a thiol group for the binding *via* NCL or maleimid-thiol coupling to the polymer. For the linkage to the peptide the methylhydroxy group was transferred to the corresponding carbonate, which was then reacted to the NHS ester resulting in molecule **12a**. A detailed synthesis plan for **12a** was presented in this thesis and its coupling to different peptides was discussed. The synthesis of Bhc-Cl **3** and Bhc-OH **4** was described by Furuta[340] and optimized by Hagen *et al.*[314]. The hydrolysis of **3** could not be reproduced and was further optimized with regards of the optimal pH range to shift the reaction equilibrium towards the product **4**. The yield was further increased with regular hot filtration steps to extract the product from the equilibrium. The following synthesis of the light sensitive cysteinyl Bhc linker **11** was described with several improvements in comparison to Koschek[465]. The influence of different bases and solvents on the activation as NHS ester was evaluated using the intermediate **5a**. The presence of ACN and TEA increased the ester formation to a ratio of 99% conversion to **5b**. Using these new methods the desired NHS activated linker Cys'Bhc-NHS (**12a**) could be successful synthesized. During first trials to link **12a** to peptides as described from Koschek[465] it turned out to be crucial to carefully purify **12a** from the activation reagent DSC before added to the peptide on the solid support. It could be impeccable shown that the presence of catalytic amounts of DSC led to a preferred and stable *N*-succinimidyl carbamate formation with the amine group of the peptide side chain. In addition, the procedure for the side chain deprotection of the Mtt protected lysine in the peptide sequence was changed. As previously discussed led the presence of the scavenger TIPS, which is a common reagent to improve TFA mediated full or side chain cleavages[466], to a release of the fluorescence labeled peptides or a bleaching of the dyes (**Scheme C7** in section C3.6). Two

different fluorescence dyes TAMRA and Cy5 were compared and the coupling to the peptides was improved by the use of NHS activated dyes. Thus no aggressive coupling reagents had to be used and the efficiency of the coupling was increased to 99% conversion. Cy5 was tested as alternative fluorescence label to enable ThT aggregation assays of the labeled peptides and to monitor the distribution inside cells. TAMRA quenches the fluorescence emission of ThT, led to a decomposing of the peptides stored on the solid support and is furthermore known to support dye-dye aggregation in aqueous solutions. Hence, Cy5 was shown to not disturb the conformational behavior of the model peptide VW18 and its aggregation to amyloid fibrils.

The coupling of the linker **12a** to two different peptide sequences, VW18K9 and AKA both tested with both dyes, was evaluated. Again the reaction conditions were optimized with the use of the less sensitive linker **5a**. Thus conversion to the desired TAMRA-AK(Bhc)A-OH (**15**) could be realized with high efficiency. It was further possible to synthesize the caged variants of the short sequences Cy5-AK(Cys'Bhc)A-OH **26** and TAMRA-AK(Cys'Bhc)A-OH **19**. Therefore, a strategy, to first cage the peptide on the solid support, and afterwards add the fluorescence label turned out to be favorable. Nevertheless, both caged products were rapidly decomposing. Caging of the VW18K9 sequence was more challenging. Koschek already described problems during the synthesis of TAMRA-Ahx-VW18K9(Cys'Bhc)-OH (**28**) resulting in low yields and a not reproducible synthetic strategy.[465] Again the successful strategy was to first cage the peptide side chain to gain Fmoc-Ahx-VW18K9(Cys'Bhc)* (**22**). Although, following Fmoc-deprotection did perform with high product formation ratio, the finalizing dye coupling was not possible for the TAMRA analogue **23** and in only 1% isolated yield for the Cy5-Ahx-VW18K9(Cys'Bhc)-OH **28**. The successful caging of both AKA and Fmoc-VW18K9 peptides is a strong indicator for the general practicableness of the NHS coupling strategy.

The major difference to previously described Bhc caged compounds with fluorescence labels and to **15** was the cysteinyl linker side chain of **12a** and its thiol functionality. Thiols are known to react with π-electron systems in fluorescence dyes, thus an intramolecular rearrangement or intermolecular dimerization could be an explanation for the instability of caged fluorescence labeled peptides after acidic deprotection (**Scheme 11**). Furthermore, a major side reaction for Bhc cages with a thiol ether bond to the peptide was described. Mahmoodi *et al*. proposed a mechanism for a thiol rearrangement to position 2 of the coumarin ring (**Scheme 11**).
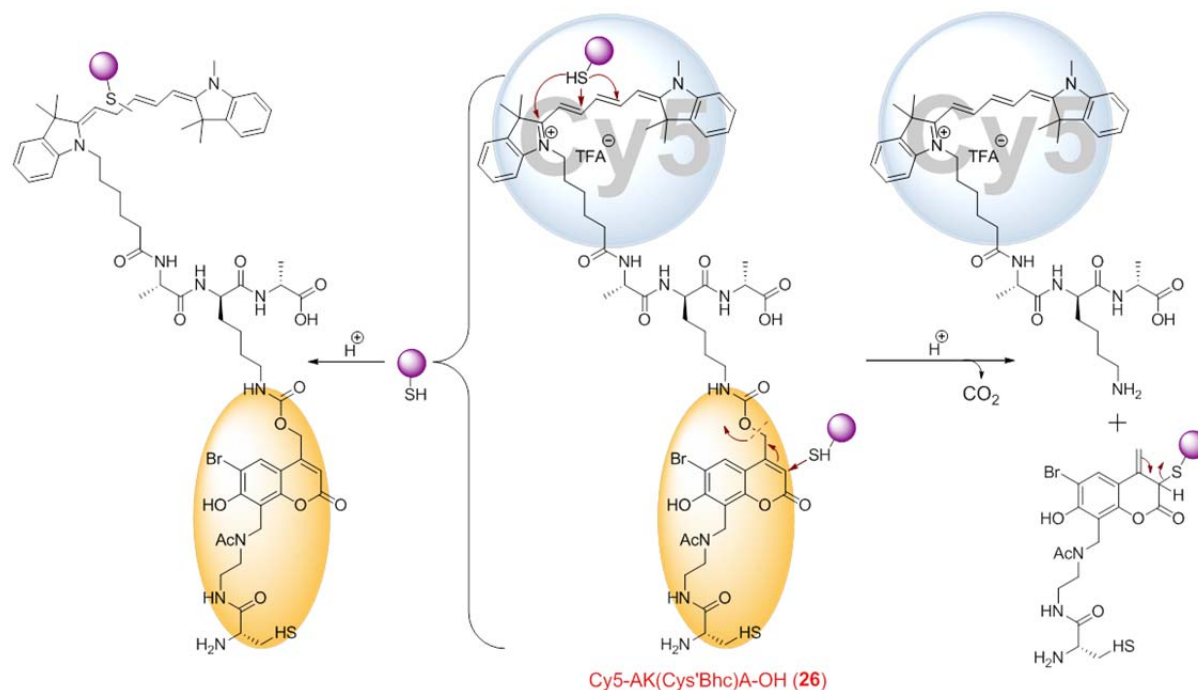
Scheme 11: *Possible intra- or intermolecular side reaction of the cysteinyl thiol functionality with either the π-electron system of the dyes, resulting in dimerization and bleaching of the dye, or at the Bhc position 2 resulting in uncaging of the Bhc from the peptide.*

They suggest a high affinity for this reaction and recommend alternatives for Bhc in presence of thiols.[339, 487] To easy prevent this side reaction the Bhc could be modified with a methyl group in the position 2 (**Scheme C12**). The resulting mBhc is resistant against the thiol exchange reaction.[339] To prevent bleaching of the dyes via a thiol side reaction a change to other functional groups for polymer binding like an azide or alkyne for a copper free click reaction could be used (**Scheme 13**).[488-489] Other known intrinsic side reactions of Bhc like the Photo-Claisen rearangement can be excluded for the tested **12a** due to the fact, that a carbamate binding to the caged compound is recommended to inhibit this side reaction.[285] The fact, that the short peptides with close distance were successfully synthesized could be a hint for an impact of flexibility of the longer VW18K9 sequence. Thus dynamic dye-dye aggregation and photon exchange already on the resin could be a reason for side reactions and decrease of coupling efficiency. To prevent problems with the caging reactions, an alternative setup with Bhc loaded Fmoc-protected amino acids could be used. For example a straight forward method for a Fmoc-cystein-mBhc building block for SPPS was described which could be an alternative setup for a alkylated linker (**Figure 9** section 3.2).[339]
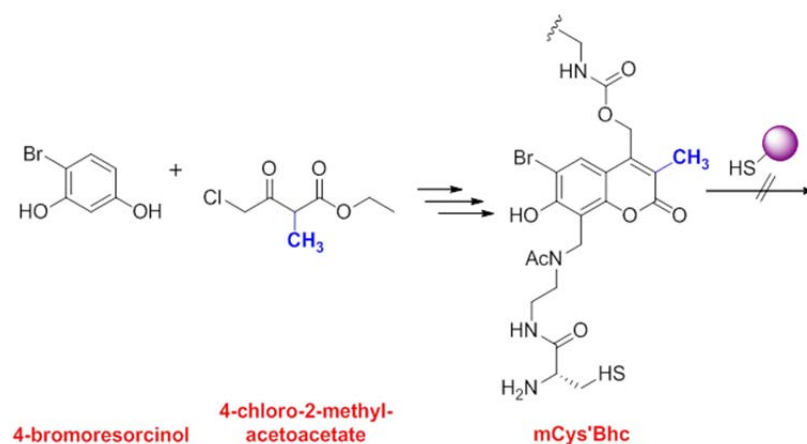
Scheme C12: *A methyl group in position 2 of the Bhc protects from thiol side reactions. The resulting mCys'Bhc can be synthesized starting from 4-bromresorcinol and 4-chloro-2-methyl-acetoacetate.*

Also the Bac caging group (**Figure 9** section 3.2) could be used to link a cell penetrating peptide or polymer to the caged VW18K9. Gandioso recently described a Fmoc-asparagine coumarin derivative for SPPS where the lactone function was replaced by a cyano(4-nitrophenyl)methylene moiety leading to a shift in absorbance to the green and red region.[490] Thus an excitation of the cage with common room light would be reduced, maybe resulting in less sensitivity of the resulting products.
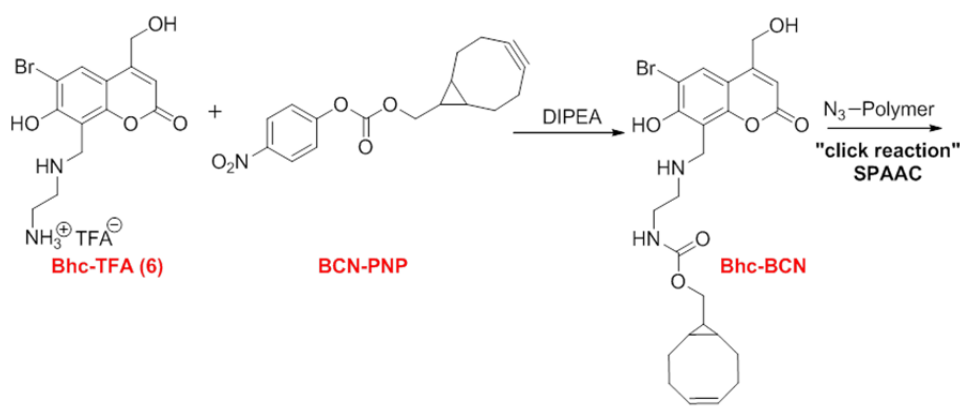
Scheme C13: *Alternative side chain functionalization with a bicyclo[6.1.0]non-4-yn-9-ylmethanol [BCN] as reaction site for a strain promoted azide alkyne cycloaddition [SPAAC].[488-489]*

Part D: HfLeu Helps to Increase Protease Stability of Peptides

The results presented in this section have been originally published as S. Huhmann, A.-K. Stegemann, **K. Folmert**, D. Klemczak, J. Moschner, M. Kube and B. Koksch, Hexafluoroleucine and trifluoroisoleucine help to increase protease stability of peptides, *Beilstein J. Org. Chem., 2017, accepted*.

D1   Abstract

Peptides and proteins are promising biomaterials and medicinal components with a broad range of possible applications, but limited due to their low bioavailability and high sensitivity towards proteases. One strategy to improve the proteolytic stability of peptides is the modification with fluorinated amino acids. The impact of fluorination on the behavior and stability of peptides depends on the position of the fluorine-substituent inside the amino acid and on interactions with neighboring amino acids and the enzyme binding pocket. This study discusses the systematic screening of the protease substrate FA, which was designed to combine specificities of different proteases.[19, 36] FA was further varied with leucine and its highly fluorinated analogue HfLeu in three different positions adjacent to the scissile bond. The seven resulting substrate peptides were tested regarding their resistance towards proteolysis by α-chymotrypsin, pepsin, proteinase K,

and elastase. The process was monitored *via* HPLC using a fluorescence detector and ESI ToF MS to identify the cleavage products. The enzymes did vary in their preferences regarding the substrate and also in their reaction times. Thus, the results give an overview of the position dependent impact of fluorine.

D2   Results

D2.1   Project Design

To elucidate the impact of fluorination on proteolytic stability six different peptides were designed with respect to the Schechter and Berger model[392] and using our previously published model peptide FA[19, 36] (**Table D1**).

Table D1: *Names and sequences of the used peptides synthesized via SPPS. Abz = o-aminobenzoic acid; Xaa= Leu or HfLeu*

Name	Sequence									
	P5	P4	P3	P2	P1	P1'	P2'	P3'	P4'	P5'
FA	Abz	K	A	A	F	A	A	A	A	K
P2X$_{aa}$FA	Abz	K	A	X$_{aa}$	F	A	A	A	A	K
P1'X$_{aa}$FA	Abz	K	A	A	F	X$_{aa}$	A	A	A	K
P2'X$_{aa}$FA	Abz	K	A	A	F	A	X$_{aa}$	A	A	K

As in our earlier studies the enzymes α-chymotrypsin, elastase, pepsin and proteinase K were investigated to explore different binding pockets.[19, 36] In all peptides position P1 is occupied by phenylalanine to constitute the cleavage site. In positions P4 and P5'lysines are introduced to increase the solubility of the peptides. For all other positions alanine is used, resulting in a disordered conformation to enable a flexible interaction of the peptides with the binding pockets of the enzymes. In order to achieve good substrate specificity position P1 remained unchanged while the positions P2, P1' and P2' were systematically substituted with Leu or HfLeu. As mentioned above, one single CF_3 group of the kind present in 5^3-TfIle has twice the van der Waals volume of a CH_3 group and the steric effects are more comparable to those of an isopropyl group.[373] Thus, due to its high fluorination content HfLeu is much more hydrophobic and also sterically more demanding than the TfIle with only one CF_3 group or Leu/Ile, respectively (**Figure D1**).
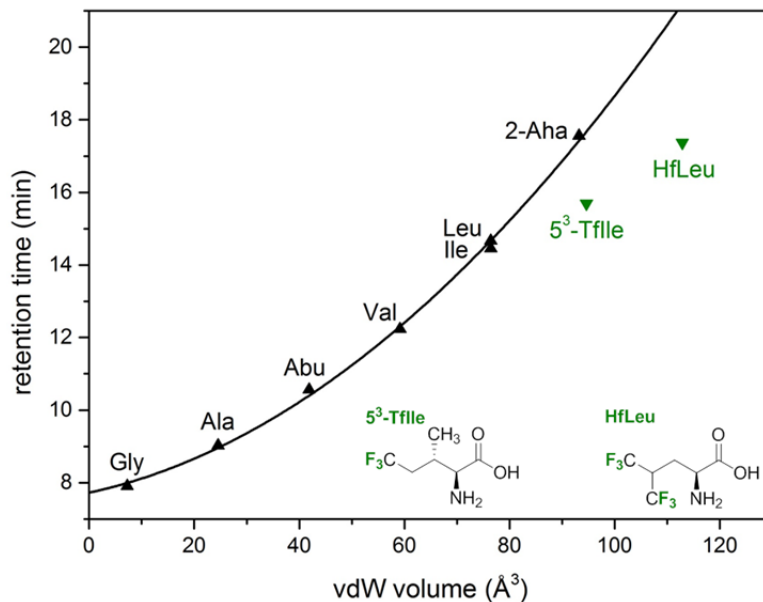
Figure D1: *Retention times of the Fmoc-protected amino acids plotted against the van der Waals volumes of the side chains. Nonfluorinated amino acids are represented by black triangles, the correlation between them is shown with a black line, and fluorinated amino acids are represented by green triangles. The structural formulae of 5³-TfIle and HfLeu are given with fluorine atoms highlighted in green.[374, 380, 385]*

D2.2 α-Chymotrypsin

Based on the hydrophobicity plot[374, 380] (**Figure D1**) all substituted peptides are more hydrophobic than the control peptide FA. Due to α-chymotrypsin's preference for aromatic and bulky hydrophobic side chains in position P1, a decrease in proteolytic stability was expected for the substituted peptides. Nevertheless, after 120 min of incubation all model peptides showed a higher proteolytic stability towards α-chymotrypsin than did the control peptide FA (**Figure D2**). This result indicates that hydrophobicity alone is not the crucial factor for the enzymatic degradation kinetics. After 120 min there was no significant difference for the modified peptides between fluorinated and unfluorinated variants and to the starting values. After an incubation time of 24 h the peptides FA, P2-LeuFA and P2'-LeuFA were almost completely cleaved. On the contrary, substitution with HfLeu led to a remarkable increase in proteolytic stability in comparison to its natural counterparts, except for P1'-HfLeu, where fluorine incorporation reduces proteolytic stability. A previous study with DfeGly in comparison with its hydrocarbon analogue Abu already investigated this influence of fluorination in position P2. It was explained by the fluorine-induced polarity of DfeGly which is induced by the withdrawing of electron

110

density from the neighboring C-H bonds.[36] Introducing the fluorinated, non-polar amino acid HfLeu into position 2 now shows that polarity is not the main argument to discuss the influence on proteolytic stability. A more precise explanation needs to take the complex influence of fluorine and potential fluorine interactions on the physical, chemical and structural properties of the peptides into account.
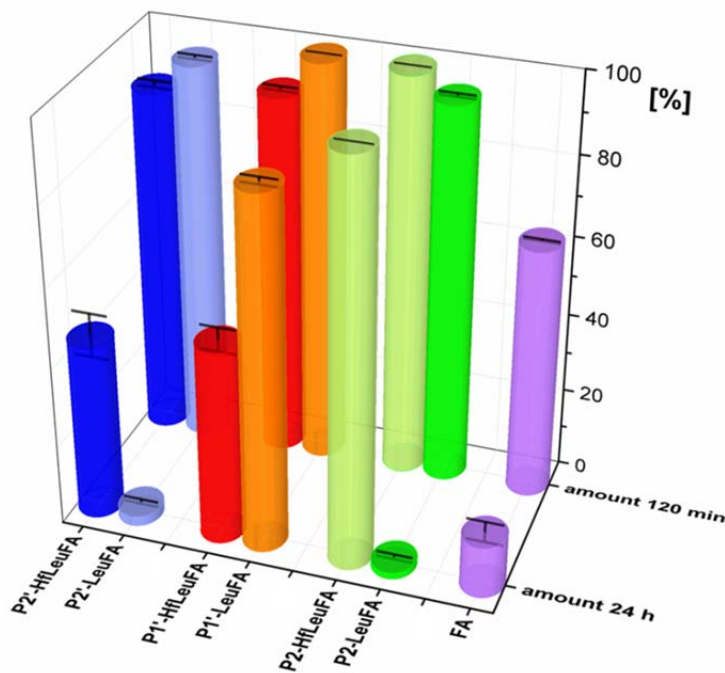
Figure D2: *Percentage of substrate remaining after incubation for 120 min and 24 h with α-chymotrypsin in 10 mM phosphate buffer, pH 7.4, at 30°C. The data shown represent the average of three independent measurements. Errors are derived from the standard deviation.*

ESI ToF MS analysis confirmed that position P1 (Phe) is the main cleavage site for the enzyme α-chymotrypsin; however, further cleavage after P1' position was observed for P1'-LeuFA and P1'-HfLeuFA (**Table S1** appendix). These additional cleavage sites could explain the overall low proteolytic stability of P1'-HfLeuFA

D2.3 Elastase

Elastase prefers to cleave bonds next to neutral and small side chains, that of Ala in particular, and thus a significant effect of introducing large hydrophobic amino acids into the substrate would be expected.[491] This may explain why incubating the peptides with elastase led to almost no proteolytic degradation during the first 120 min. Only the control peptide FA was somewhat

digested by elastase (**Figure D3**). After 24 h of incubation, all model peptides except for P1'-LeuFA persisted to the same extent as the control peptide FA. HfLeu containing peptides showed significantly higher stability than those with Leu for every position P2', P1' and P2. Notably P1'-HfLeuFA increases the resistance against elastase compared to the completely degraded P1'-LeuFA. In summary, for the case of elastase, fluorination had a large impact on the proteolytic stability of the evaluated peptides in all positions.
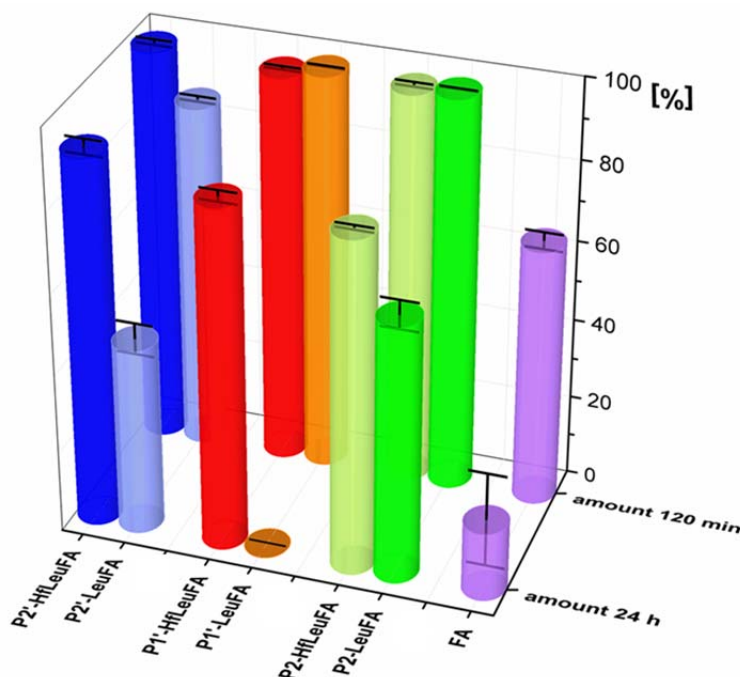
Figure D3: *Percentage of substrate remaining after incubation for 120 min and 24 h with elastase in 100 mM Tris/HCl buffer, pH 8.4, at 37°C. The data shown represent the average of three independent measurements. Errors are derived from the standard deviation.*

Since Ala is the main component of the model peptides, a variety of cleavage products in ESI ToF MS spectrometry was observed (**Table S2** appendix). Nevertheless, the estimated cleavage site between P1' and P1 was not hydrolyzed by elastase.

D2.4 Pepsin

The preference of pepsin is to cleave bonds with aromatic and bulky hydrophobic side chains in position P1 and P1', thus substitutions at the latter position should have a dramatic impact.[492] After incubation with pepsin for 120 min the P2 peptides P2-LeuFA and P2-HfLeuFA were almost entirely degraded (**Figure D4**). P1'HfLeuFA had the same stability as FA and P2'-

HfLeuFA was in this case the most resistant peptide. Nevertheless, after 24 h of incubation most peptides were completely digested, only the fluorinated peptide P2'-HfLeu remains largely full length in the presence of pepsin. Previous studies already investigated the effect of side chain fluorination on position P2' [19, 36] but thus far such an effect for position P2 has not been reported.
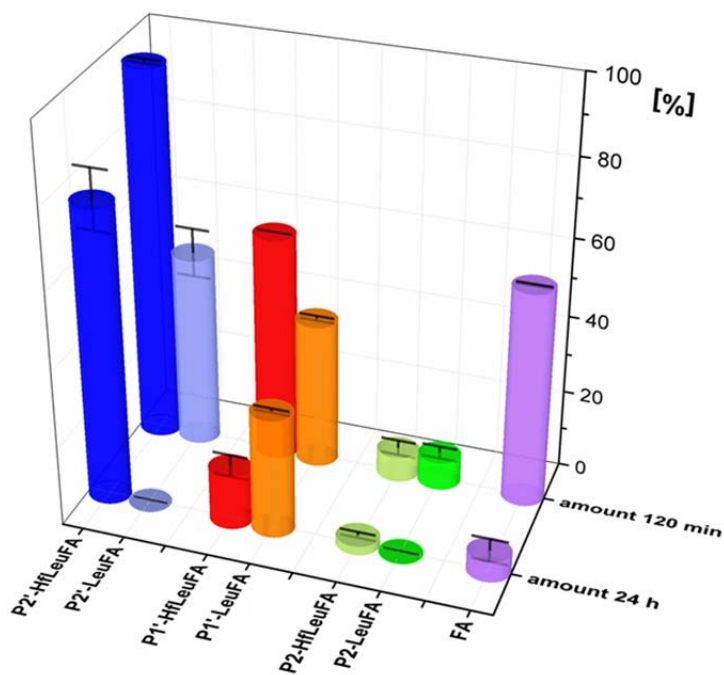
Figure D4: Percentage of substrate remaining after incubation for 120 min and 24 h with pepsin in 10 mM acetate buffer, pH 4.0, at 30°C. The data shown represent the average of three independent measurements. Errors are derived from the standard deviation.

The expected mass fragments for cleavage between positions P1 and P1' were observed (**Table S3** appendix). In addition, the peptides with substitutions in the P1' position showed also shorter fragments after amino acids like Leu and HfLeu which is in good agreement with the results for α-chymotrypsin. Furthermore, fragments of the P2'-HLeu peptide produced by enzyme cleavage between the P1' and P2' positions were observed.

D2.6 Proteinase K

Proteinase K offers a broad substrate specificity, but aliphatic, especially Ala, and aromatic side chains in position P1 and P1' are preferred.[493] Already after a reaction time of 120 min a high degradation rate for all peptides was observed (**Figure D5**). All unfluorinated peptides were digested in a similar way as the control peptide FA, whereas the fluorinated P2-HfLeuFA peptide was more resistant against proteinase K. Nevertheless, P1'-HfLeuFA and P2'-HfleuFA peptides were less stable than their unfluorinated analogues. After 24 h incubation the degradation was even more significant and none of the peptides remained uncleaved.
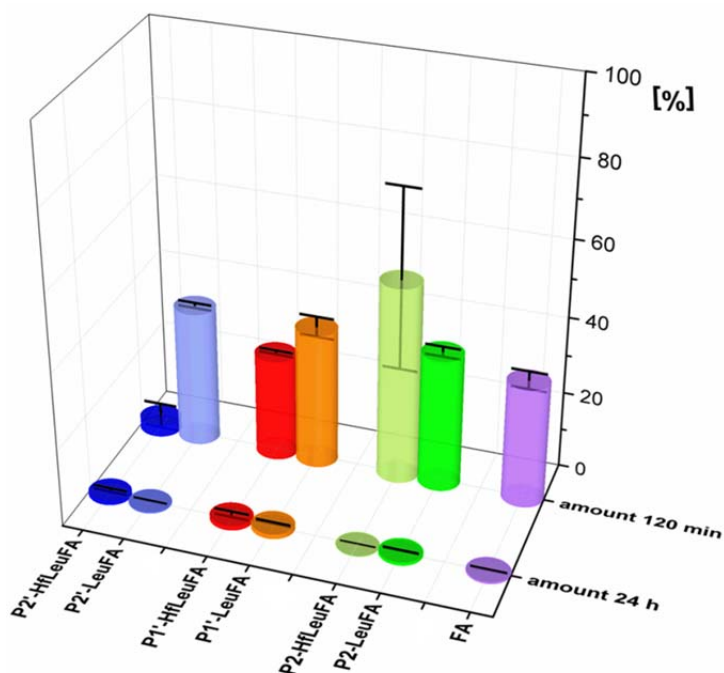
Figure D5: *Percentage of substrate remaining after incubation for 120 min and 24 h with proteinase K in 50 mM Tris/HCl buffer, containing 10 mM CaCl₂ pH 7.5, at 37°C. The data shown represent the mean of three independent measurements. Errors are derived from the standard deviation.*

Due to the preferences of pepsin, mainly fragments corresponding to cleavage N-terminal to each Ala were observed in ESI ToF MS (**Table S4** appendix). For the peptides P2'-HfLeuFA and P2-HfLeuFA some additional fragments where proteinase K hydrolyzed before and after the HfLeu amino acid were identified.

D2.7   Effect of Fluorination on Time-dependent Degradation

Regarding the reaction velocity with the model peptides, the enzymes α-chymotrypsin and elastase were much slower on the one hand and pepsin and proteinase K were comparably fast on the other hand (**Figure D6**).
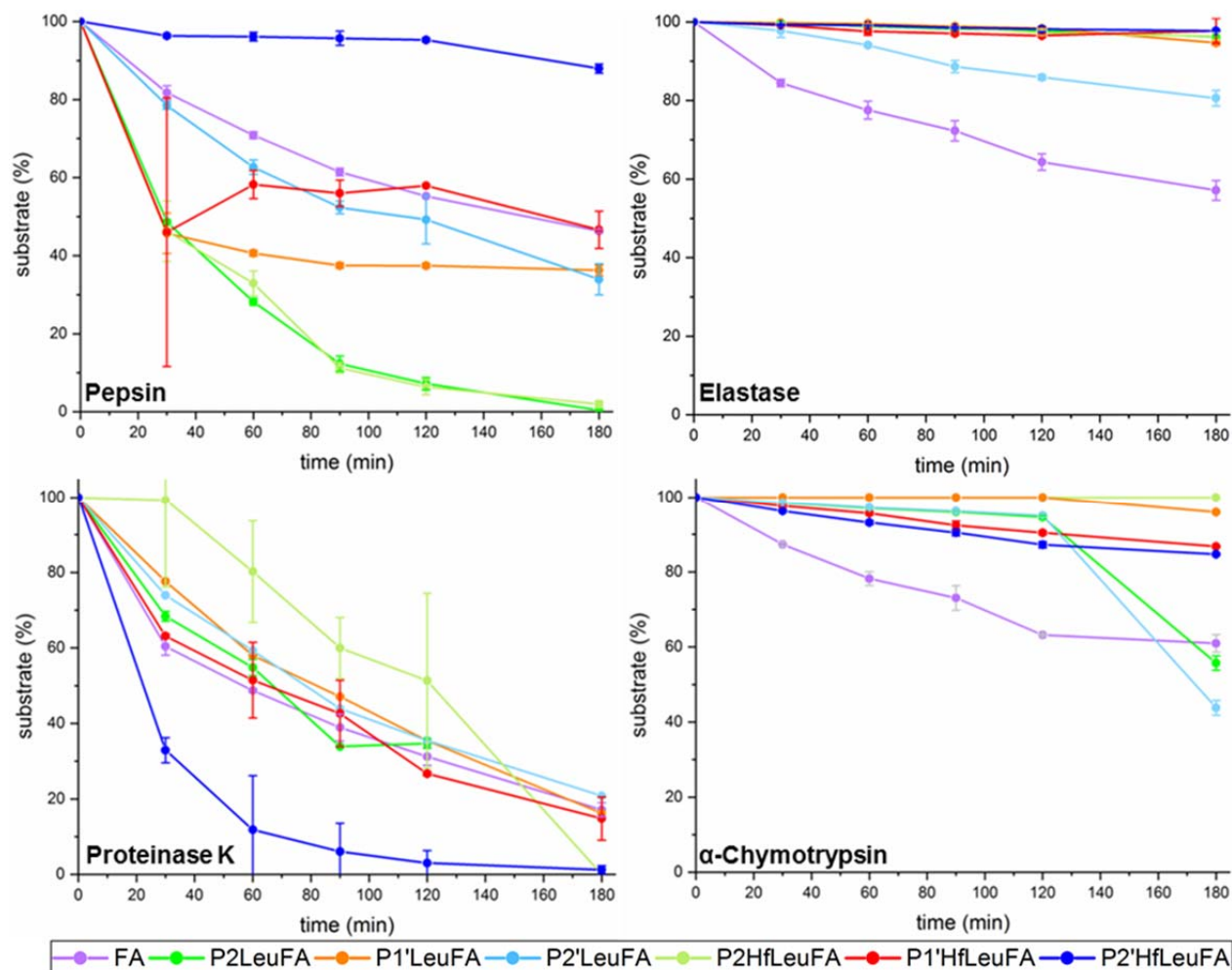
Figure D6: *Substrate amount [%] versus incubation time, up of 180 min, with α-chymotrypsin, pepsin, elastase and proteinase K. The depicted values represent the mean of three independent measurements. Error bars that are not visible are too small for proper display.*

In addition, the stability of the peptides differed from enzyme to enzyme. The most stable peptide was P2'HfLeuFA, that was remarkably resistant towards all other investigated proteases compared to the non-fluorinated analogues, except for proteinase K. In contrast, P2'HfLeuFA was easily digested by proteinase K. The reason for the high stability of P2'-HfLeuFA could be

the impact of the number and conformation of the six fluorine atoms of HfLeu in addition to the localization to the expected cleavage site. Upon comparing the fluorinated peptides with their leucine analogues the dependency of the position of the fluorination becomes even more distinct (**Figure D7**). While the effect ranges from destabilization to a strong increase in stabilization in position P1' next to the expected cleavage site, the incorporation of six fluorine atoms affected only α-chymotrypsin in position P2 and drastically increased the stability of the substrate. All three other protease differentiated either slightly or not at all between fluorinated and non-fluorinated leucine in the P2 position. When HfLeu is incorporated into position P2', digestion by three of the enzymes, elastase, pepsin and α-chymotrypsin, was hindered.
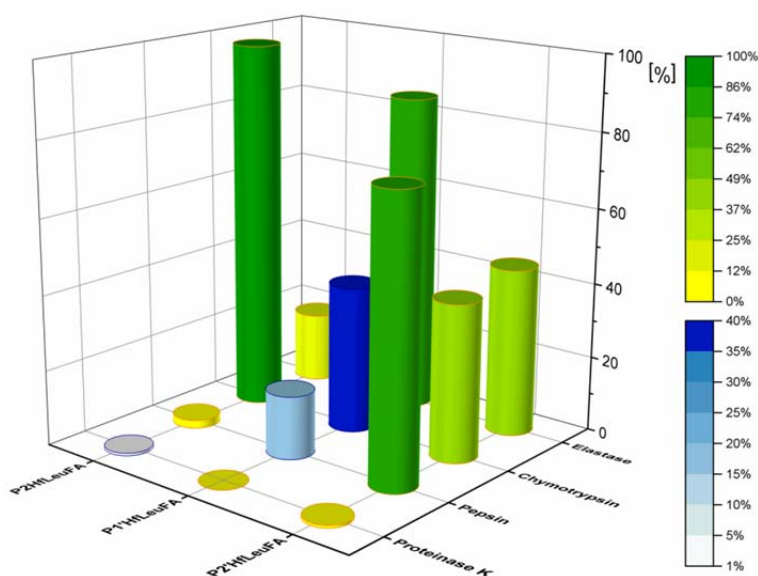
Figure D7: *Dimension of stabilization or destabilization upon HfLeu incorporation compared to the non-fluorinated analogue containing Leu, respectively, for all four different enzymes studied here and measured after 24 h of incubation with stabilization in green and destabilization in blue.*

The properties of fluorine, mainly its high electronegativity, are known to influence the acidity, polarization, and polar interactions of the building blocks. Therefore they may also have an impact on the intra- and intermolecular interactions engaged in by the peptides.[494-502] The very short and strong C-F-bond[503] has a large dipole moment[504] and is highly hydrophobic[505]. There are controversial discussions of the property of hydrogen bond acceptance by fluorine and its inductive effect may enhance Lewis acidity of neighboring group protons and their H-bond donor ability. Nevertheless, there are already investigated examples of OH⋯FC-hydrogen

bonds, for example in fluorinated benzyl alcohol.[506] These different and strong impacts of fluorine may result in a conformational change, in a small molecule as well as in a larger peptide molecule. Therefore, binding to the enzyme's binding pockets S2 through S2' is probably hindered. These results demonstrate that in contrast to some published proteolytic studies[507], the expectation of a general and consistent increase in the proteolytic stability through incorporation of fluorine atoms is not necessarily correct. Thus, proteolytic stability is not always predictable and additional systematic studies are required to in order to begin to develop rules for peptide engineering with fluorine.

The results show an impact of side-chain fluorination on proteolytic stability by influencing the interactions between the peptide and the enzyme binding pocket depending on the side chain proximal to the cleavage site and the preferences of the particular enzyme. Also, in contrast to proteolytic studies published before[390, 508] the expectation of a general increase in proteolytic stability as a result of steric occlusion of the peptide from the active site upon incorporation of sterically demanding fluorinated amino acids could not be verified based on the results of our current study. Nevertheless, the impact of proteolytic stability based on unnatural-fluorinated amino acids has great potential for the design of more bioavailable peptides and peptidomimetics. It holds promise for many applications in medicinal chemistry; especially the unique stability feature of the P2-HfLeuFA peptide provides outstanding opportunities for future therapeutic design.[360, 383]

Position-Dependent Impact of Hexafluoroleucine and Trifluoroisoleucine on Protease Digestion

Susanne Huhmann[1], Anne-Katrin Stegemann[1], Kristin Folmert[1], Damian Klemczak[1], Johann Moschner[1], Michelle Kube[1] and Beate Koksch*[1]

Address: [1]Department of Chemistry and Biochemistry, Freie Universität Berlin, Takustraße 3, 14195 Berlin, Germany

Email: Beate Koksch - beate.koksch@fu-berlin.de

* Corresponding author

Abstract

Rapid digestion by proteases limits the application of peptides as therapeutics. One strategy to increase the proteolytic stability of peptides is modification with fluorinated amino acids. This study presents a systematic investigation of the effects of fluorinated leucine and isoleucine derivatives on the proteolytic stability of a peptide that was designed to comprise substrate specificities of different proteases. Therefore, leucine, isoleucine, and their side-chain fluorinated variants were site-specifically incorporated at different positions of this peptide resulting in a library of 13 distinct peptides. The stability of these peptides towards proteolysis by α-chymotrypsin, pepsin, proteinase K, and elastase was studied, and this process was followed by an FL-RP-HPLC assay in combination with mass spectrometry. In a few cases, we observed an exceptional increase in proteolytic stability upon introduction of the fluorine substituents. The opposite phenomenon was observed in

1

other cases, and this may be explained by specific interactions of fluorinated residues with the respective enzyme binding sites. Noteworthy is that 5,5,5-trifluoroisoleucine is able to significantly protect peptides from proteolysis by all enzymes included in this study when positioned N-terminal to the cleavage site. These results provide valuable information for the application of fluorinated amino acids in the design of proteolytically stable peptide-based pharmaceuticals.

Keywords

fluorinated amino acids, hexafluoroleucine, peptide drugs, protease stability, trifluoroisoleucine

Introduction

Peptide-based drugs are promising pharmaceuticals since they offer several advantages including high selectivity, specificity, and efficacy for recognizing and binding to their targets [1-6]. However, their application as drugs is often limited due to low oral bioavailability and a short half-life attributable in part to proteases of the digestive system and blood plasma [1-8]. Efficient approaches to overcome these limitations have been developed including the incorporation of non-natural amino acids, such as *D*-amino acids, backbone-extended or chemically modified amino acids [1]. In this regard, the incorporation of fluorine into amino acids has become a promising strategy. Fluorine's unique properties, namely low polarizability, a strong inductive effect, and high electronegativity, as well as its small size, result in strong, short C-F bonds and perturb the acidity and basicity of adjacent functional groups. Moreover, these changes may strongly influence hydrogen bonding and electrostatic interactions that are crucial for binding to receptors or, in context of protease stability,

2

enzymes. Thus, when introduced in the form of fluorinated amino acids, this unique element can alter the biophysical, chemical and pharmaceutical properties of proteins and peptides including their interaction with proteases [9-10].

Several laboratories have focused on introducing highly fluorinated analogues of hydrophobic amino acids and have studied the effects on stability of the resulting proteins towards thermal and chemical denaturation [9, 11-22]. These studies prompted further investigation into the extent to which fluorinated amino acids stabilize peptides and proteins against proteolytic degradation in particular. Meng and Kumar reported that the incorporation of 5,5,5,5',5',5'-hexafluoroleucine (HfLeu) into the antimicrobial peptides magainin 2 amide and buforin enhanced their resistance towards proteolytic degradation by trypsin [23]. They also introduced HfLeu into the glucagon-like-peptide-1 (GLP-1), which is an attractive lead compound for the treatment of diabetes mellitus type 2. Unfortunately, the clinical use of the wild-type peptide is severely hampered due to rapid digestion (~ 2 min) by the serine protease dipeptidyl peptidase [24-26]. Satisfyingly, the fluorinated GLP-1 analogues displayed higher proteolytic stability against this enzyme [27].

Usually, the enhanced proteolytic stability of fluorinated peptides is explained by their greater hydrophobicity and altered secondary structure compared to the parent, non-fluorinated peptide. A further reason is the increased steric bulk of the fluorinated amino acid, meaning protection from protease degradation is a result of the steric occlusion of the peptide from the active site [23, 28]. In contrast, the Marsh lab found that the introduction of HfLeu into the antimicrobial peptide MSI-78 only renders it more stable towards proteolysis by trypsin and chymotrypsin in the presence of liposomes [29]. In the absence of liposomes, the fluorinated variants were as rapidly degraded as the non-fluorinated control, suggesting that the incorporation of HfLeu is not the only factor that prevents the peptide from being digested by proteases [29].

3

Fluorinated aromatic amino acids were also investigated regarding their impact on peptide proteolysis. For instance, incorporation of monofluorinated phenylalanine variants into the histone acetyltransferase protein tGN5 resulted in destabilization in a chymotrypsin digestion assay [30]. Substitution of tryptophan, tyrosine, and phenylalanine residues in a glycosylation-deficient mutant of *Candida antarctica* lipase B, CalB N74D, by their monofluorinated analogues, left resistance to proteolytic degradation by proteinase K unchanged [31]. Incorporation of α-fluoroalkyl substituted amino acids can also lead to proteolytically stable peptides, and proteases can even be used to synthesize α-fluoroalkyl substituted peptides [32-38].

These results indicate that the influence of fluorinated amino acids on the proteolytic stability of peptides and proteins remains difficult to predict. In an attempt to systematically study the influence of fluorinated amino acids on the proteolytic stability of peptides, a 10-amino acid peptide (FA) was previously designed in our group, comprising the substrate specificities of the proteases α-chymotrypsin and pepsin [39-40]. 2-Aminobutanoic acid (Abu) and its fluorinated analogues 2-amino-4,4-difluorobutanoic acid (DfeGly) and 2-amino-4,4,4-trifluorobutanoic acid (TfeGly) were individually incorporated at either the P2, the P1' or the P2' position [41] to give nine different analogues of FA. In prior studies, we observed that the introduction of fluorine atoms into the Abu side chain can significantly improve or dramatically reduce resistance to hydrolysis by different enzymes and human blood plasma, depending upon the fluorine content of the side chain, the position of the substitution relative to the cleavage site and the particular protease [39-40].

Here, we extend these studies to include highly fluorinated, sterically demanding HfLeu, and 5,5,5-trifluoroisoleucine (TfIle) and to investigate their effects on proteolytic stability towards the serine proteases α-chymotrypsin, elastase, and proteinase K, and the aspartate protease pepsin.

4

Results and Discussion

Peptide design and structure

To elucidate the impact of fluorination on proteolytic stability we previously designed the peptide FA (Figure 1(b)) that comprises the substrate specificities of α-chymotrypsin and pepsin [39-40]. Consequently, the P1 position is occupied by a phenylalanine residue. Lysine residues were introduced at both ends of the peptide sequence to enhance solubility, and *o*-aminobenzoic acid (Abz) at the N-terminus serves as a fluorescence label. Alanine residues in positions P3, P3', and P4' act as spacers as the peptide binds in an extended conformation to the enzyme's active site [42]. The positions P2, P1' and P2' at or adjacent to the cleavage site [41] carry the key residues for the recognition of the substrate by the protease and serve as substitutions sites.

(a)

1
(76,4 Å^3)

2
(76,4 Å^3)

3
(94,65 Å^3)

4
(112,86 Å^3)

hydrophobicity & spatial demand

(b)

Name		P3	P2	P1	P1'	P2'	P3'	P4'
FA	Abz K	A	A	F	A	A	A	K
P2-XaaFA	Abz K	A	**Xaa**	F	A	A	A	K
P1'-XaaFA	Abz K	A	A	F	**Xaa**	A	A	K
P2'-XaaFA	Abz K	A	A	F	A	**Xaa**	A	K
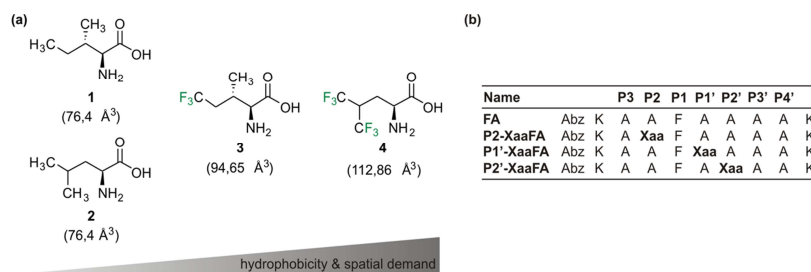
Figure 1: (a) Structures of isoleucine **1**, leucine **2**, and their fluorinated analogues 5,5,5-trifluoroisoleucine **3** (TfIle) and 5,5,5,5',5',5'-hexafluoroleucine **4** (HfLeu). The van der Waals volumes given in parentheses correspond to the side chains (starting at α-carbon), and were calculated according to Zhao et al. [43]; **(b)** Names and amino acid sequences of the studied peptides; the substitution positions are marked

5

as Xaa; Xaa = Leu, HfLeu, Ile or TfIle. Positions are named according to Schechter and Berger nomenclature [41].

The alanines at P2, P1' or P2' positions were substituted individually with either TfIle or HfLeu (Figure 1(a)). Ile and Leu variants were also included in this study as non-fluorinated controls. This led to a library of 12 FA variants (Figure 1(b)).

Leu and Ile are larger and more hydrophobic than Ala. The fluorinated amino acids are even larger and more hydrophobic than their hydrocarbon analogues [44-45]. Furthermore, fluorine substitution has been shown to polarize neighboring C-H bonds (here the γ-hydrogens) that could affect noncovalent interactions [9, 11]. Since the amino acids used here (Figure 1(a)) differ in their degree of fluorination, spatial demand and hydrophobicity, it is expected that they will have different impacts on the enzyme's binding pocket, reflected by different behavior in the proteolysis assay.

Determination of proteolytic stability

All peptides were incubated with the four different proteases and their proteolytic degradation was followed over a period of 24 h. Both, α-chymotrypsin [46-49] and pepsin [50-54] are well characterized digestive proteases. They are, together with trypsin, the main enzymes of the human digestive system. Elastase possesses a wide substrate specificity for non-aromatic, neutral side chains [55-56] and is found in the human pancreas and in blood serum. Proteinase K, an enzyme widely used for inactivation and degradation studies of proteins, was included here since it shows a broad substrate specificity and high activity and, thus, is able to digest numerous native proteins, even in the presence of detergents [57]. These four enzymes have different preferences at their subsites, thus providing a broad scope for our investigations.

6

The course of proteolytic digestion was characterized by an analytical HPLC-assay with fluorescence detection [39-40]. For quantification of substrate degradation, integration of the corresponding HPLC peak was conducted. Cleavage products were identified by ESI-ToF mass spectrometry (see Supporting Information, Table S4-S7). Figure S2 shows the time course of all of the digestion experiments. A detailed description of the results for the individual enzymes is given in the following sections.

Proteolytic stability towards α-chymotrypsin (EC 3.4.21.1)

α-Chymotrypsin is a serine endopeptidase with broad specificity. It preferably cleaves peptide bonds C-terminal to large hydrophobic residues such as phenylalanine, tyrosine, tryptophan, and leucine in the P1 position. Secondary hydrolysis also occurs at the carbonyl end of isoleucine, methionine, serine, threonine, valine, histidine, glycine, and alanine [47, 58-60].

The S2 subsite of α-chymotrypsin generally prefers to accommodate hydrophobic residues [59, 61]. We observed that the fluorinated P2-variants show a smaller amount of digestion after 120 min compared to their hydrocarbon analogues, while all variants are more stable than the control FA (Figure 2). After 24 h, all P2 peptides except for P2-LeuFA are still more stable than FA. Incorporation of Leu into P2 leads to complete proteolysis compared to FA, in which Ala occupies this position. However, incorporation of six fluorine substituents into Leu (resulting in HfLeu) results in an almost 100% gain in proteolytic stability. Ile is not as highly preferred in P2 as Leu, but also here the introduction of three fluorine substituents leads to a 50% gain in stability. P2-HfLeuFA and P2-TfIleFA are not digested at all, suggesting that HfLeu and TfIle are not favored within the S2 pocket of α-chymotrypsin.

7

Figure 2: Percentage of substrate remaining after incubation for 120 min (left) and 24 h (right) with α‑chymotrypsin in 10 mM phosphate buffer, pH 7.4, at 30℃. The data shown represent the average of three independent measurements. Errors are derived from the standard deviation.

P1' substituted peptides are all more stable towards digestion than the control peptide FA, while Leu seems to provide the best protection from proteolysis. Here, introduction of fluorine makes the peptide prone to degradation. The opposite is true for Ile as TfIle leads to less efficient degradation. The S1' subsite of α-chymotrypsin usually accommodates basic residues with long side chains [59, 62-63]. Ile, as a branched amino acid, is obviously not well accommodated in this position for steric reasons. A further increase in side chain volume with TfIle exacerbates this effect. In the case of HfLeu, however, fluorine substituents seem to engage in favorable interactions with amino acid residues of the binding site, thus making P1'-HfLeuFA a better substrate than the non-fluorinated Leu peptide. Several such interactions are possible, as described in our previous work [39-40, 64].

The S2' subsite of α-chymotrypsin exhibits hydrophobic character and thus prefers to accommodate hydrophobic residues [59, 65]. However, the more hydrophobic peptides P2'-LeuFA, P2'-HfLeuFA, P2'-IleFA, P2'-TfIleFA are more stable against digestion by α-chymotrypsin compared to FA after 120 min of incubation. After 24 h,

8

125

only the fluorinated analogues are less degraded than the control FA. Full length P2'-HfLeuFA and P2'-TfIleFA are present at percentages up to 44% and 24%, respectively, while substitution by Leu and Ile in P2' position leads to accelerated proteolysis compared to FA. Thus, both HfLeu and TfIle have a protective effect towards proteolysis in this position.

ESI-ToF mass analysis confirms that the position P1 bearing Phe is the main cleavage site for α-chymotrypsin (Figure 3). Cleavage C-terminal to Leu and HfLeu in P1'-LeuFA and P1'-HfLeuFA is also observed (Figure 3, see Supporting Information, Table S4), which means that the cleavage site was shifted towards the C-terminus by one residue. This is likely a consequence of α-chymotrypsin's preference for not only aromatic residues but also bulky hydrophobic residues in the S1 pocket, thus, HfLeu is accepted by the P1 binding pocket of α-chymotrypsin.
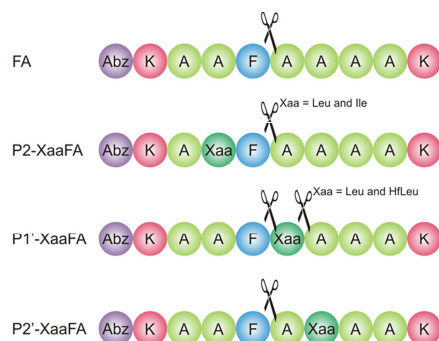
Figure 3: Cleavage positions observed in the digestion of library peptides with α-chymotrypsin.

In summary, the introduction of fluorinated Leu and Ile analogues into a α-chymotrypsin specific peptide sequence can improve proteolytic stability mainly at the P2 and P2' positions, with the strongest effects observed for the P2 position.

9

Proteolytic stability towards pepsin (EC 3.4.23.1)

Pepsin is an aspartic endopeptidase and one of the main digestive enzymes in humans. It exhibits specificity for hydrophobic, especially aromatic residues like Phe, Trp, and Tyr at the P1 and P1' positions [50-54]. It has an extended active site that can bind at least seven residues [66-67], and peptide bond cleavage occurs N-terminal to the residue at position P1. Cleavage efficiency heavily depends upon the identity of this amino acid, with Phe and Leu being the most favored residues. At the P1' position aromatic amino acid residues are preferred, however the influence of the P1' position on proteolytic cleavage is not as significant [68]. Pepsin typically does not cleave at Val, Ala, or Gly linkages [60].

The S2 subsite of pepsin preferentially accommodates hydrophobic residues such as Leu, Ala or norleucine as well as the β-branched species Ile and Val, but can also bind charged residues [69-70]. Except for P2-TfIleFA, we observed that the P2-modified peptides are degraded more rapidly than the control peptide FA and that these peptides are almost or completely degraded after 120 min (Figure 4). For example, whereas after 24 h FA is also almost completely degraded, P2-TfIleFA is still detected at a level of 100%. Incorporation of Leu or Ile leads to complete proteolysis. Remarkably, the introduction of six fluorine atoms into Leu doesn't change this behavior. In sharp contrast yet equally remarkable, the incorporation of three fluorine substituents into Ile results in a 100% gain in proteolytic stability. These results indicate that Leu, HfLeu, as well as Ile are well accommodated in the S2 subsite of pepsin. In contrast, TfIle, although smaller than HfLeu [44], doesn't appear to fit well into the S2 pocket of pepsin.

10
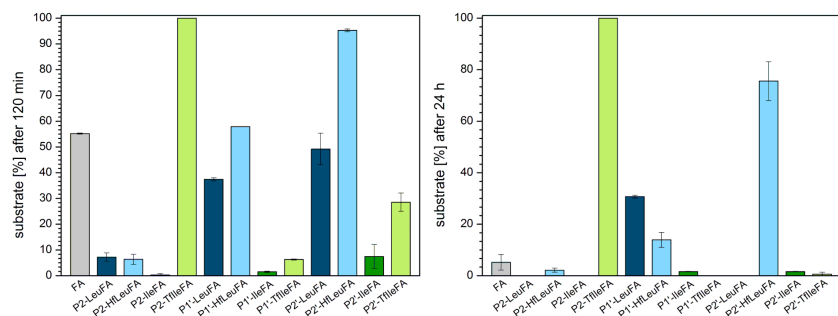
Figure 4: Percentage of substrate remaining after incubation for 120 min (left) and 24 h (right) with pepsin in 10 mM acetate buffer, pH 4.0, at 30℃. The data shown represent the average of three independent measurements. Errors are derived from the standard deviation.

To compare the P1' substituted peptides, only P1'-HfLeu shows the same persistence after 120 min as the control peptide FA, while all other sequences are digested faster. Here, the introduction of fluorine into Leu seems to stabilize the peptide by about 20%. Interestingly, after 24 h this trend is reversed, and the P1'-HfLeuFA peptide is destabilized to an amount of 17% compared to the hydrocarbon analogue, but both peptides are somewhat more stable than the control FA. The incorporation of TfIle into this position doesn't show a significant impact. Although the fluorine substituents slow down the digestion process (see Supporting Information, Figure S2(b)), the TfIle containing peptide as well as its hydrocarbon analogue are fully digested after 24 h. The S1' subsite has hydrophobic character and thus prefers to accommodate hydrophobic or aromatic residues [71]. Ile and TfIle are obviously well accommodated in this position, while Leu and HfLeu are not.

The S2' subsite of pepsin favors hydrophobic amino acids, but also accepts charged polar amino acids like Glu and Thr [52, 72]. After 120 min peptides P2'-LeuFA, P2'-IleFA and P2'-TfIleFA are degraded faster than the control FA. Instead,

11

P2'-HfLeuFA is only digested up to 5%. This effect is even more pronounced after 24 h. While all other P1' substituted peptides, along with the control peptide FA, are almost or completely digested, P2'-HfLeuFA is still present to about 76%. In this case fluorination leads to protection against proteolysis by pepsin. Hfleu is obviously not well accommodated in this position. As already observed for position P1', the introduction of three fluorine atoms into Ile slows down proteolysis, although both peptides are completely digested after 24 h.

For almost all peptides of our library, we observed the expected cleavage pattern with Phe in the P1 position (Figure 5, see Supporting Information, Table S5). Only P2'-HfLeuFA is not hydrolyzed at the designed cleavage site, instead cleavage occurs exclusively N-terminal to the HfLeu residue, thus demonstrating that HfLeu occupies the P1' position. In the case of P2'-TfIle we found two further peptide bonds that are cleaved by pepsin, namely N-terminal cleavage to TfIle and to Phe. These findings indicate that the S1' subsite accommodates bulky hydrophobic residues more readily than does the S2' site of pepsin. For P1'-LeuFA and P1'-HfLeuFA we found an additional cleavage site at which the peptide bonds Leu$^{P1'}$-Ala$^{P2'}$, and HfLeu$^{P1'}$-Ala$^{P2'}$ are hydrolyzed, respectively, which means that the cleavage site was shifted towards the C-terminus by one residue. This cleavage pattern was also detected for α-chymotrypsin before, and indicates that HfLeu is well accepted by pepsin in its S1 binding site. Furthermore, we identified a second cleavage site for P2-HfLeuFA at which the peptide bond N-terminal to Phe is proteolytically cleaved as well. This means that the cleavage site is shifted such that HfLeu occupies the P1 position, and Phe the P1'. However, this perfectly matches the specificity of pepsin that prefers bulky hydrophobic and aromatic amino acids both up- and downstream of the scissile bond.

12
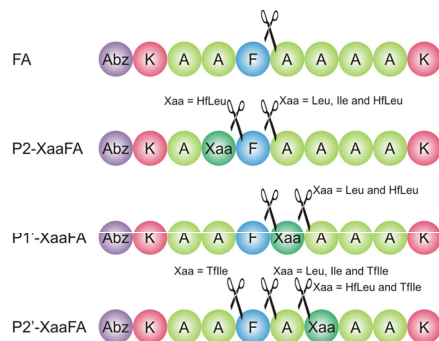
Figure 5: Cleavage positions for digestion of the different peptides with pepsin.

In summary, the introduction of fluorinated Leu into a pepsin specific peptide sequence can improve proteolytic stability at the P2' position, whereas the incorporation of a fluorinated Ile into the P2 position shows the strongest effect in protection from proteolysis.

Proteolytic stability towards elastase (EC 3.4.21.36)

Elastase is a serine endopeptidase, and has a wide specificity for non-aromatic uncharged side chains. It preferentially cleaves peptide bonds C-terminal to small uncharged non-aromatic amino acid residues such as glycine, alanine and serine, but also valine, leucine, isoleucine [56, 73]. Its binding site extends over eight subsites (S5 to S1, and S1' to S3') [74].

The fact that in this study larger and more hydrophobic amino acids [44-45] were introduced may explain why degradation of most of the variants during the first 120 min of incubation with elastase is hardly observed (Figure 6). Only P2'-LeuFA, P2'-IleFA, and P2'-TfIleFA were somewhat digested during this time, however, all of the modified peptides are more stable than the control FA.
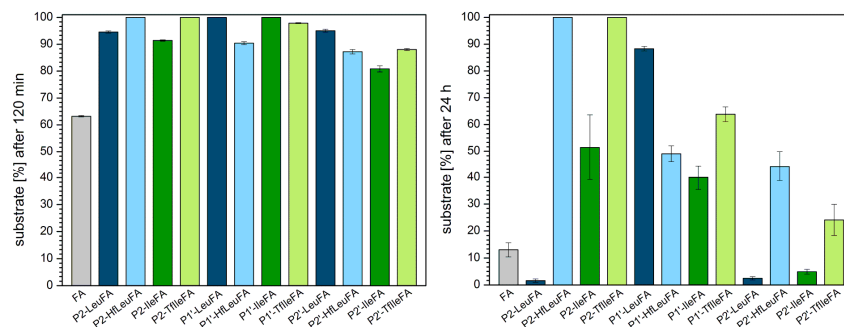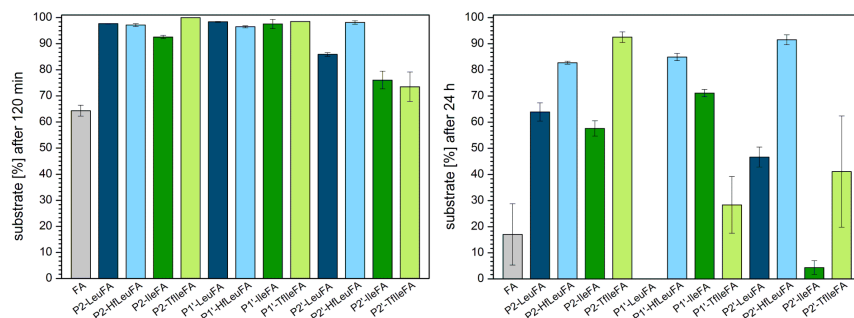
Figure 6: Percentage of substrate remaining after incubation for 120 min (left) and 24 h (right) with elastase in 100 mM Tris/HCl buffer, pH 8.4, at 37℃. The data shown represent the mean of three independent measurements. Errors are derived from the standard deviation.

After 24 h all P2 peptide variants are more stable than the control FA (Figure 6), while TfIle provides the best protection from elastase digestion. Leu and Ile are not quite as preferred in P2 position as is Ala. Fluorination of Leu leads to an increase in stability of around 19%. With 35% this effect is even higher when three fluorine atoms are introduced into Ile.

Modification of the P1'-postion renders P1'-HfLeuFA and P1'-IleFA more stable than the control peptide FA, while P1'-TfIleFA is comparably stable. Incorporation of Leu into P1' leads to complete digestion. However, introducing six fluorine atoms into Leu results in an 85% gain in stability. The opposite is observed for Ile, where TfIle accelerates enzymatic degradation.

Except for P2'-IleFA, all P2' modified variants are more stable compared to the control peptide FA after 24 h. Leu is not as preferred in this position as Ala. Introduction of fluorine strengthens this effect and effectively doubles the stability. In contrast, introduction of Ile leads to almost complete proteolysis. However, substitution by TfIle slows down the degradation rate and results in a stabilization of

14

around 37%. In P2', fluorination shows in both cases a protective effect towards hydrolysis by elastase.

Elastase preferably hydrolyses peptide bonds C-terminal to uncharged non-aromatic amino acids and mainly between Ala-Ala and Ala-Gly bonds [56, 73]. Since Ala is the main residue present in the peptides studied here, we observed various cleavage products in the ESI-ToF analysis (Figure 7, see Supporting Information, Table S6). For none of the peptides were fragments with Phe in the P1 position observed. Since elastase has a constricted S1 pocket, the binding of aromatic amino acids at P1 is deleterious [75]. Here, we also observed that Leu appears to never occupy the P1 position, although it is known to occupy this position in other substrates [73]. Interestingly, the larger fluorinated variant was found in the P1 position in one case, while Ile and its fluorinated analogue occupy this position in two of the three peptide analogues.

The S2' subsite of elastase has a marked specificity for Ala, and can accommodate bulkier residues only with some difficulty [74]. Thus, we did not find the fluorinated amino acids HfLeu and TfIle binding to the S2' subsite of the enzyme as expected, whereas for the Leu and Ile variants this was observed only in one case each.

The S1' subsite usually prefers Lys residues, and to a lesser extent Ala or Glu [74, 76]. Indeed, we found a fragment cleaved off corresponding to a Lys in P1', but primarily detected fragments with Ala in P1' and also Phe that was even more favored than Lys. We observed that Ile as well as its fluorinated analogue TfIle are not accommodated in this subsite, probably due to their β-branched topology.

The S3' pocket in elastase is known to have a high aromatic specificity [74]. Interestingly, in our cases Phe in P3' was less favored. Instead, mainly Lys occupied this position.

15

Ala is favored in P2. Its carboxyl group can form a hydrogen bond with the amide nitrogen of Gly193 in the S2 pocket, and Ala's methyl group faces the solvent [76]. Occupation of the S4 subsite is important for efficient catalysis [76-77]. Thus elastase might not easily split the first three bonds at the amino terminus of a peptide chain, since interactions of a residue with S4 is necessary [77]. Indeed, we only observed a low amount of cleavage proximal to the N-terminus, while most of the hydrolysis occurred at the C-terminal end of the peptides. The S3 subsite seems to accommodate bulkier hydrophobic amino acids well, as we observed cleavage products containing Ile and Leu in P3 position for all the peptides modified with these residues, as well as their fluorinated analogues for two of the three substituted peptides each.
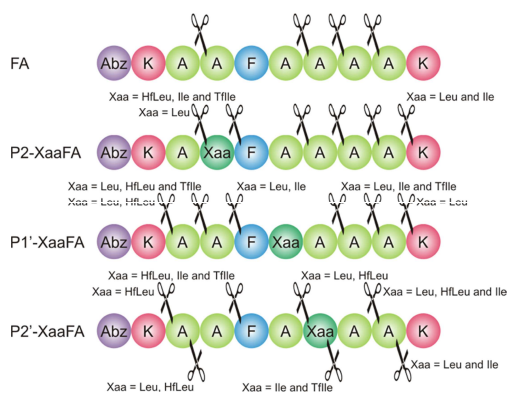
Figure 7: Cleavage positions for digestion of the different peptides with elastase.

In summary, introduction of HfLeu in different positions of a peptide can enhance the proteolytic stability up to 85% compared to the corresponding Leu analogues. Replacing Ile with TfIle can increase the stability against elastase as well, although not as efficiently as HfLeu.

16

Proteolytic stability towards proteinase K (EC 3.4.21.64)

Proteinase K is a non-specific serine endopeptidase and the main proteolytic enzyme produced by the fungus *Tritirachium album Limber* [78]. It has a broad substrate specificity, cleaving peptide bonds C-terminal to a number of amino acids, however prefers aromatic or aliphatic hydrophobic amino acids in position P1 [57, 78]. Furthermore, Ala is favored in position P2 and enhances cleavage efficiency [79-80]. Proteinase K possesses a very high proteolytic activity [79]. Its active center contains an extended binding region consisting of several subsites, at least four or five subsites on the N-terminal side of the scissile bond (S1 to S4/S5) and three subsites C-terminal to the scissile bond (S1' to S3') [81-83]. The "bottom" of substrate recognition site is predominantly hydrophobic and there is evidence that not the sequence of the substrate is of importance in the recognition but only the volume of the side chains [84].

Substitution of Ala in position P2 with Ile and Leu leads to a greater or comparable amount of degradation after 120 min. Introducing fluorine atoms in both cases slows down the digestion process, most pronounced for P2-TfIleFA with a gain of 60% in stability compared to its non-fluorinated analogue. Ile is not preferred to the extent that Leu is, and the introduction of fluorine enhances this effect. While all other peptides are almost completely or entirely degraded after 24 h, P2-TfIleFA is the only peptide that is still left after 24 h of incubation (Figure 8).

17
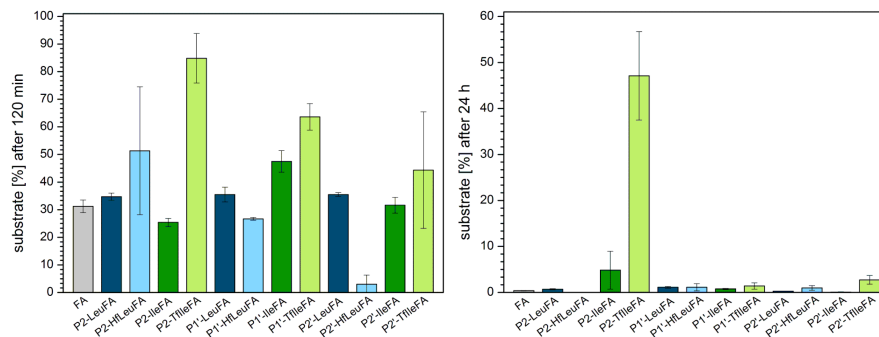
Figure 8: Percentage of substrate remaining after incubation for 120 min (left) and 24 h (right) with proteinase K in 50 mM Tris/HCl buffer, containing 10 mM CaCl$_2$ pH 7.5, at 37℃. The data shown represent the mean of three independent measurements. Errors are derived from the standard deviation.

Introduction of Leu at the P1' position leads to an amount of digestion comparable to FA after 120 min. Fluorination of the Leu side chain leads to a small acceleration in digestion. Ile at this position is not as preferred as is Leu and this enhances the stability to a small extent compared to FA. Introducing three fluorine atoms at the Ile side chains strengthens the stability against proteinase K even further.

As already observed for the other two Leu containing peptides, also substitution of Ala at position P2' with Leu does not change the resistance against proteinase K significantly. Interestingly, when six fluorine atoms are introduced, the digestion process is faster and P2'-HfLeuFA is almost completely degraded after 120 min. The opposite is observed for the fluorination of Ile. While P2'-IleFA is as stable as FA, P2'-TfIleFA shows a small gain in stability of around 12%.

Thus, in the case of P1' as well as P2' peptide variants, only fluorination of Ile leads to a slower digestion by proteinase K, while introducing even more fluorine atoms into the Leu side chain leads to more rapid hydrolysis compared to the non-fluorinated analogues. Ile seems in all investigated cases not as preferred as Leu, since less

18

efficient digestion is observed. Introduction of three fluorine atoms even enhances this protective effect.

Based on the wide substrate specificity of proteinase K and its preference for alanine, and since our studied peptides have a high number of alanine residues present, there are multiple cleavage sites possible in addition to the designed site between Phe[P1] and Xaa[P1']. Indeed, multiple cleavage patterns are observed, especially cleavage C-terminal to Ala (Figure 9, see Supporting Information Table S7). Thus, Ala mainly occupies the S1 subsite, but is also found to bind to the S2 site to a greater extent. Ala is most effective in P2 [80] as the S2 subsite is a small and narrow cleft, which limits the possibilities for effective side chain substitutions [79]. However, Ile and TfIle are well accepted here. A negative or positive charge at S2 is not preferred and hampers the formation of the enzyme substrate complex [82]. Thus, it can be concluded that Lys is poorly accepted at this position. HfLeu, the most sterically bulky amino acid investigated here, is not observed to occupy the S2 subsite. Instead, HfLeu is mainly found to bind to the S1 pocket. Leu is also found to occupy the S1 site of Proteinase K, which is large and has mainly hydrophobic character [82-83, 85]. It does not impose too strong steric limitations on the amino acid side chain but prefers hydrophobic and aromatic residues, with a specificity for Ala [78, 81, 83, 86]. Charged side chains of Glu and Lys are very poorly accepted, as are β-branched functional groups, because the entrance to the S1 subsite is too narrow to allow their passage [79]. Thus, Lys is not observed to occupy the S1 subsite. Neither Ile nor TfIle can be accommodated by the S1 pocket due to their β-branching. Phe is found in S1 in only two cases in our study, and mainly occupies the S3 and S1' pockets. The S3 pocket has a wide specificity due to its location at the protein surface, but exhibits preference for aromatic side chains in P3 (Trp, Phe) [79]. S1' shows a slight preference for smaller residues like Ala and Gly, but also bulkier

19

136

residues such as Phe and Leu are hydrolyzed to a significant extent [81]. In this study Leu apparently does not bind to the S1' site at all, and this is also true of TfIle. Additionally, Lys is not well accommodated here. Phe is also found to occupy S4 to a great extent, and this subsite is known to have an affinity for aromatic groups, especially a marked preference for Phe [79]. S4-P4 interactions are primarily hydrophobic in nature [79], which might explain why we observed that Lys is only poorly accepted in this position. The S3 subsite cannot be defined as a "cleft" or "pocket" [79]. The P3 residue of the peptide substrate lies on the protein surface and the side chain of P3 should be directed toward the solvent [79]. This arrangement might explain the broad specificity of S3 [79]. We observed that all the residues used in this study can occupy the P3 position, mainly Phe and Lys. Leu, Ile and TfIle are also found to a great extent in P3.
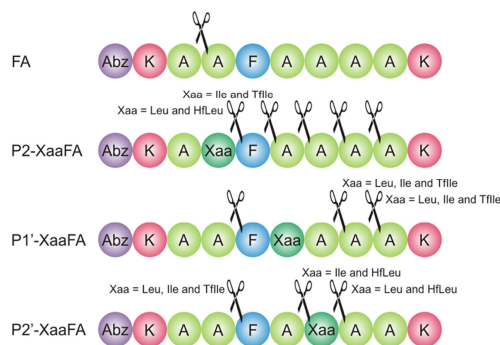
Figure 9: Cleavage positions found with proteinase K.

In summary, fluorination of an Ile residue N-terminal to the cleavage site can help to protect a peptide against proteolysis by proteinase K. Due to its broad specificity and high activity, proteinase K typically digests peptides quickly [57]. This was also observed in this work in experiments in which all peptides, except for P2-TfIleFA, a remarkably stable species, were completely degraded after an incubation time of 24 h.

20

Conclusion

The bulky side-chain fluorinated amino acids HfLeu and TfIle have the power to significantly stabilize peptides against proteolytic degradation. The impact of their incorporation on the proteolytic stability of peptides does not follow a general trend but rather depends on a combination of factors including the nature of the fluorinated amino acid, the substitution position relative to the cleavage site and the studied protease. Also, in contrast to proteolytic studies published before [23, 27-28], the expectation of a general increase in proteolytic stability as a result of steric occlusion of the peptide from the active site upon incorporation of sterically demanding fluorinated amino acids could not be verified based on the results of our current study. We found a significant stabilization towards proteolysis in 13 of a total of 24 peptides of the library studied here upon introduction of either HfLeu or TfIle (Figure 10). However, we observed that even these sterically demanding fluorinated amino acids show in some cases favorable interactions with the enzymes binding sites resulting in a more rapid digestion as the non-fluorinated control.

21
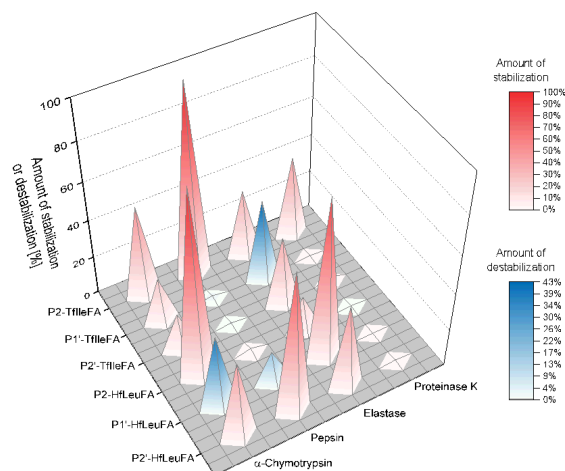
Figure 10: Dimension of stabilization or destabilization upon TfIle or HfLeu incorporation compared to the non-fluorinated analogues containing Ile or Leu, respectively, for all four different enzymes studied here and measured after 24 h of incubation.

Introduction of fluorinated Leu and Ile analogues into P2 and P2' position improved the proteolytic stability towards α-chymotrypsin. When introduced in the P1' position a stabilization was still observed for TfIle, while incorporation of HfLeu made the peptide more prone to proteolytic digestion compared to the non-fluorinated control. Incorporation of HfLeu had a significantly stabilizing effect towards hydrolysis by pepsin only in P2' position, while TfIle develops a protective effect only when incorporated in P2 position.

As both, elastase and proteinase K possess a broad specificity, preferring C-terminal cleavage to Ala, we observed here a rather unspecific cleavage pattern for both enzymes with multiple cleavage products, in which the intended designed cleavage site with Phe in P1 position wasn't affected. However, we observed that the introduction of HfLeu has a general protective effect against degradation by elastase,

22

whereas the effect of TfIle depends on the substitution position. Although the introduction of fluorine substituents generally affected the rate of hydrolysis by proteinase K, only fluorination of an Ile residue N-terminal to the cleavage site effectively protected the peptide from digestion. Particularly noteworthy is the effect of fluorination of the Ile side chain in P2 position. The P2-TfIleFA peptide was the most resistant substrate towards proteolysis by all four proteases applied in this study. In contrast, destabilization due to fluorination was only observed when TfIle and HfLeu were incorporated into the P1' position.

In future studies, we will focus on a more precise characterization of the interaction of fluorinated substrates with proteolytic enzymes to which multiple factors contribute. The steric demand or conformation of the side chain, hydrophobicity, fluorine induced polarity and significant pK_a-value changes of neighboring groups [9-10] can lead to fluorine-specific interactions between substrate and enzyme binding sites as well as to an exclusion of the cleavage-relevant peptide bonds from the active site.

Furthermore, our investigations show that fluorine's impact on proteolytic stability needs to be investigated always case-by-case as there is no general trend to be concluded. Nevertheless, the results of this current study provide valuable knowledge on how bulky fluorinated amino acids can help to increase the proteolytic stability of peptides, and show that upon smart design, these fluorinated amino acids can be used to engineer peptide drug candidates.

Experimental

Materials

Fmoc-L-amino acids were purchased from ORPEGEN Peptide Chemicals GmbH (Heidelberg, Germany). Fmoc-Lys-Nova-Syn®TGA-resin was from Novabiochem

23

(Merck Chemicals GmbH, Darmstadt, Germany). All solvents were used from VWR (Darmstadt, Germany) without further purification. All other chemicals were bought from ACROS (Geel, Belgium), abcr GmbH (Karlsruhe, Germany), fluorochem (Hadfield, United Kingdom), VWR (Darmstadt, Germany) or Merck (Darmstadt, Germany) at highest commercially available purity and used as such. Detailed synthetic strategy for Fmoc-TfIle-OH is described in literature [44]. For the synthesis of Fmoc-HfLeu-OH see supporting information.

Peptide synthesis, purification and characterization

Peptides were synthesized manually in a 0.05 mmol scale on a solid support by means of an Fmoc/tert-butyl protecting group strategy on a preloaded Fmoc-Lys(Boc)Wang resin (0.57 mmol/g loading) using 10 mL polypropylene reactors. HfLeu containing peptides were synthesized with an Activo-P11 Automated Peptide Synthesizer (Activotec, Cambridge, United Kingdom). Couplings of non-fluorinated amino acids were performed in dimethylformamide (DMF) with the Fmoc-L-amino acid, 1-hydroxybenzotriazole (HOBt) and N,N'-diisocarbodiimide (DIC) in eight-fold excess with respect to the resin amount. In order to ensure completion of the reaction the couplings were performed twice for 1 h each. The fluorinated amino acids and coupling reagents 1-Hydroxy-7-azabenzotriazole (HOAt)/DIC were used in 1.2-fold excess, and the coupling was carried out manually one time over night. In case of an insufficient coupling, the coupling was repeated for 3 h with 0.5 equivalents. Prior to the Fmoc deprotection of the fluorinated amino acids, free N-termini were capped by adding a mixture of acetic anhydride (Ac$_2$O) (10% (v/v)) and N,N-diisopropylethylamine (DIPEA) (10% (v/v)) in DMF (3 x 10 min). Fmoc deprotection was achieved by treatment with 20% (v/v) piperidine in DMF (3 x 10 min). All peptides were N-terminally labeled with o-aminobenzoic acid (Abz) to enable photometric detection. The resin was washed between each step with DMF and

24

dicholoromethane (DCM) (3 x 2 mL each). After the synthesis, the peptides were cleaved from the resin by treatment with a solution (2 mL) containing triisopropylsilane (TIS) (10% (v/v)), water (1% (v/v)), and trifluoroacetic acid (TFA) (89% (v/v)) for 3 h. The resin was washed twice with TFA (1 mL) and DCM (1 mL) and excess solvent was removed by evaporation. The crude peptide was precipitated with ice-cold diethyl ether (80 mL), and after centrifugation dried by lyophilization. Purification of the synthesized peptides was performed on a LaPrepΣ low-pressure HPLC system (VWR, Darmstadt, Germany) using a Kinetex RP-C18 endcapped HPLC-column (5 μM, 100 Å, 250 x 21.2 mm, Phenomenex®, USA). A Security GuardTM PREP Cartridge Holder Kit (21.20 mm, ID, Phenomenex®, USA) served as pre-column. As eluents deionized water (Milli-Q Advantage® A10 Ultrapure Water Purification System, Millipore®, Billerica, MA, USA) and acetonitrile (ACN), both containing 0.1% (v/v) TFA were used. HPLC runs were performed starting with an isocratic gradient of 5% ACN over 5 min, flow rate: 10 mL/min, continuing with a linear gradient of 5 – 70% ACN over 25 min, flow rate: 20.0 mL/min. UV-detection occurred at 220 nm. Data analysis was performed with an EZChrom Elite-Software (Version 3.3.2 SP2, Agilent Technologies, Santa Clara, CA, USA). The fractions containing pure peptide were combined, reduced *in vacuo* and lyophilized to give the peptides as a white powder. The purity of the peptides was controlled by analytical HPLC (LUNA™ C8 (2) column, 5 μm, 250 x 4.6 mm, Phenomenex®, Torrance, CA, USA), and the products were identified by high resolution ESI-ToF-MS (see supporting information).

Protease digestion assay

All peptides employed in the degradation studies were used as the TFA salts obtained after lyophilization. Stock solutions of α-chymotrypsin (from bovine pancreas, EC 3.4.21.1, ≥ 40.0 units/mg of protein, Sigma Aldrich, St. Louis, MO,

25

USA), and pepsin (from porcine stomach mucosa, EC 3.4.23.1, ≥ 250 units/mg of protein, Sigma Aldrich, St. Louis, MO, USA) were prepared at concentrations of 1 mg/mL in phosphate buffer (10 mM, pH 7.4), or in acetate buffer (10 mM, pH 4.0), respectively. For proteinase K (from tritirachium album, EC 3.4.21.64, ≥ 30 units/mg of protein, Sigma Aldrich, St. Louis, MO, USA) and elastase (from porcine pancreas, EC 3.4.21.36, 6.2 units/mg of protein, Sigma Aldrich, St. Louis, MO, USA) stock solutions were prepared also at concentrations of 1 mg/mL in tris/HCl (50 mM) + $CaCl_2$ (10 mM) buffer (pH 7.5), or in tris/HCl buffer (100 mM, pH 8.4), respectively. Peptides (0.002 mmol) were prepared as stocks in DMSO (100 µL) and incubated with the respective enzyme at 30℃ (for α-chymotrypsin and pepsin) or 37℃ (for proteinase K and elastase) with shaking at 300 rpm in a thermomixer over a period of 24 h. The reaction mixture consisted of DMSO (15 µL), corresponding buffer (25 µL), peptide solution (5 µL) and the corresponding enzyme solution (5 µL). The concentration of enzyme was optimized so that the hydrolysis of the control peptide FA was about 40% after 120 min. Aliquots of 5 µL were taken at fixed time points (0, 15, 30, 60, 90, 120 min as well as 3 h and 24 h) and either quenched with ACN containing 0.1% (v/v) TFA (95 µL), in the case of α-chymotrypsin, proteinase K and elastase, or 2% aqueous ammonia (95 µL), in the case of pepsin. All samples were subjected to analytical HPLC on a LaChrom-ELITE-HPLC-System equipped with a fluorescence detector (VWR International Hitachi, Darmstadt, Germany). A monolithic reversed-phase C8 Chromolith® Performance HPLC column (100 x 4.6 mm, Merck KGaA, Darmstadt, Germany) was used to resolve and quantify the products of digestion. The used system and gradients are described in detail in the supporting information. Detection based on the Abz label was carried out using a fluorescence detector with λ_{ex} = 320 nm and λ_{em} = 420 nm. In all cases, the peaks corresponding to the starting materials (full-length peptides) or the N-terminal fragments (products)

26

were integrated and used to determine the velocity of the reaction (see supporting information). The FA peptide was used as a reference. Each fragment cleaved from the full-length peptide was identified by ESI–ToF mass analysis on an Agilent 6220 ESI-ToF-MS spectrometer (Agilent Technologies, Santa Clara, CA, USA) (see supporting information). All experiments were performed in triplicate.

Supporting Information

Characterization and identification of synthesized peptides. Characterization of the enzymatic digestion reactions, and identification of proteolytic cleavage products. HPLC methods, and synthesis protocol for Fmoc-HfLeu-OH.

Supporting Information File 1:

File Name: Supporting-Information_revised

File Format: docx

Title: Supporting Information - Hexafluoroleucine and trifluoroisoleucine help to increase protease stability of peptides

Acknowledgements

27

144

References

1. Vlieghe, P.; Lisowski, V.; Martinez, J.; Khrestchatisky, M. *Drug Discov. Today* **2010,** *15* (1–2), 40-56.
2. Sato, A. K.; Viswanathan, M.; Kent, R. B.; Wood, C. R. *Curr. Opin. Biotechnol.* **2006,** *17* (6), 638-642.
3. Albericio, F.; Kruger, H. G. *Future Med. Chem.* **2012,** *4* (12), 1527-1531.
4. Uhlig, T.; Kyprianou, T.; Martinelli, F. G.; Oppici, C. A.; Heiligers, D.; Hills, D.; Calvo, X. R.; Verhaert, P. *EuPA Open Proteom.* **2014,** *4*, 58-69.
5. Fosgerau, K.; Hoffmann, T. *Drug Discov. Today* **2015,** *20* (1), 122-128.
6. Santos, G. B.; Ganesan, A.; Emery, F. S. *ChemMedChem* **2016,** *11* (20), 2245-2251.
7. Jäckel, C.; Koksch, B. *Eur. J. Org. Chem.* **2005,** *2005* (21), 4483-4503.
8. Latham, P. W. *Nat. Biotechnol.* **1999,** *17* (8), 755-757.
9. Salwiczek, M.; Nyakatura, E. K.; Gerling, U. I.; Ye, S.; Koksch, B. *Chem. Soc. Rev.* **2012,** *41* (6), 2135-2171.
10. Berger, A. A.; Völler, J.-S.; Budisa, N.; Koksch, B. *Acc. Chem. Res.* **2017,** *50* (9), 2093-2103.
11. Salwiczek, M.; Samsonov, S.; Vagt, T.; Nyakatura, E.; Fleige, E.; Numata, J.; Cölfen, H.; Pisabarro, M. T.; Koksch, B. *Chem. Eur. J.* **2009,** *15* (31), 7628-7636.
12. Huhmann, S.; Nyakatura, E. K.; Erdbrink, H.; Gerling, U. I. M.; Czekelius, C.; Koksch, B. *J. Fluorine Chem.* **2015,** *175*, 32-35.
13. Marsh, E. N. G. Designing Fluorinated Proteins. In *Methods Enzymol.,* Vincent, L. P., Ed. Academic Press: 2016; Vol. 580, pp 251-278.
14. Vukelić, S.; Moschner, J.; Huhmann, S.; Fernandes, R.; Berger, A. A.; Koksch, B. Synthesis of Side Chain Fluorinated Amino Acids and Their Effects on the Properties of Peptides and Proteins. In *Modern Synthesis Processes and Reactivity of Fluorinated Compounds,* Leroux, F. R.; Tressaud, A., Eds. Elsevier: 2017; pp 427-464.
15. Buer, B. C.; Meagher, J. L.; Stuckey, J. A.; Marsh, E. N. *Protein Sci.* **2012,** *21* (11), 1705-1715.
16. Buer, B. C.; Meagher, J. L.; Stuckey, J. A.; Marsh, E. N. G. *Proc. Natl. Acad. Sci.* **2012,** *109* (13), 4810-4815.
17. Jäckel, C.; Salwiczek, M.; Koksch, B. *Angew. Chem. Int. Ed.* **2006,** *45* (25), 4198-4203.
18. Tang, Y.; Tirrell, D. A. *J. Am. Chem. Soc.* **2001,** *123* (44), 11089-11090.
19. Tang, Y.; Ghirlanda, G.; Petka, W. A.; Nakajima, T.; DeGrado, W. F.; Tirrell, D. A. *Angew. Chem. Int. Ed.* **2001,** *40* (8), 1494-1496.
20. Tang, Y.; Ghirlanda, G.; Vaidehi, N.; Kua, J.; Mainz, D. T.; Goddard, W. A.; DeGrado, W. F.; Tirrell, D. A. *Biochemistry* **2001,** *40* (9), 2790-2796.
21. Yoder, N. C.; Kumar, K. *Chem. Soc. Rev.* **2002,** *31* (6), 335-341.
22. Bilgiçer, B.; Xing, X.; Kumar, K. *J. Am. Chem. Soc.* **2001,** *123* (47), 11815-11816.
23. Meng, H.; Kumar, K. *J. Am. Chem. Soc.* **2007,** *129* (50), 15615-15622.
24. Deacon, C. F.; Nauck, M. A.; Toft-Nielsen, M.; Pridal, L.; Willms, B.; Holst, J. J. *Diabetes* **1995,** *44* (9), 1126.
25. Holst, J. J.; Deacon, C. F. *Curr. Opin. Pharmacol.* **2004,** *4* (6), 589-596.
26. Drucker, D. J. *Curr. Pharm. Des.* **2001,** *7* (14), 1399-1412.
27. Meng, H.; Krishnaji, S. T.; Beinborn, M.; Kumar, K. *J. Med. Chem.* **2008,** *51* (22), 7303-7307.
28. Akcay, G.; Kumar, K. *J. Fluorine Chem.* **2009,** *130* (12), 1178-1182.
29. Gottler, L. M.; Lee, H. Y.; Shelburne, C. E.; Ramamoorthy, A.; Marsh, E. N. *ChemBioChem* **2008,** *9* (3), 370-373.
30. Voloshchuk, N.; Zhu, A. Y.; Snydacker, D.; Montclare, J. K. *Biorg. Med. Chem. Lett.* **2009,** *19* (18), 5449-5451.

31. Budisa, N.; Wenger, W.; Wiltschi, B. *Mol. BioSyst.* **2010,** *6* (9), 1630-1639.
32. Bordusa, F.; Dahl, C.; Jakubke, H.-D.; Burger, K.; Koksch, B. *Tetrahedron: Asymmetry* **1999,** *10* (2), 307-313.
33. Koksch, B.; Sewald, N.; Burger, K.; Jakubke, H. D. *Amino Acids* **1996,** *11* (3), 425-34.
34. Koksch, B.; Sewald, N.; Hofmann, H.-J.; Burger, K.; Jakubke, H.-D. *J. Pept. Sci.* **1997,** *3* (3), 157-167.
35. Smits, R.; Koksch, B. *Curr. Top. Med. Chem.* **2006,** *6* (14), 1483-1498.
36. Thust, S.; Koksch, B. *J. Org. Chem.* **2003,** *68* (6), 2290-2296.
37. Thust, S.; Koksch, B. *Tetrahedron Lett.* **2004,** *45* (6), 1163-1165.
38. Sewald, N.; Hollweck, W.; Mütze, K.; Schierlinger, C.; Seymour, L. C.; Gaa, K.; Burger, K.; Koksch, B.; Jakubke, H. D. *Amino Acids* **1995,** *8* (2), 187-194.
39. Asante, V.; Mortier, J.; Schlüter, H.; Koksch, B. *Biorg. Med. Chem.* **2013,** *21* (12), 3542-3546.
40. Asante, V.; Mortier, J.; Wolber, G.; Koksch, B. *Amino Acids* **2014,** *46* (12), 2733-744.
41. Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967,** *27* (2), 157-162.
42. Davies, D. R. *Annu. Rev. Biophys. Biophys. Chem.* **1990,** *19* (1), 189-215.
43. Zhao, Y. H.; Abraham, M. H.; Zissimos, A. M. *J. Org. Chem.* **2003,** *68* (19), 7368-373.
44. Erdbrink, H.; Nyakatura, E. K.; Huhmann, S.; Gerling, U. I. M.; Lentz, D.; Koksch, B.; Czekelius, C. *Beilstein J. Org. Chem.* **2013,** *9*, 2009-2014.
45. Samsonov, S. A.; Salwiczek, M.; Anders, G.; Koksch, B.; Pisabarro, M. T. *J. Phys. Chem. B* **2009,** *113* (51), 16400-16408.
46. Blow, D. M.; Birktoft, J. J.; Hartley, B. S. *Nature* **1969,** *221* (5178), 337-340.
47. Blow, D. M. The Structure of Chymotrypsin. In *The Enzymes,* Paul, D. B., Ed. Academic Press: 1971; Vol. 3, pp 185-212.
48. Derewenda, Z. S.; Derewenda, U.; Kobos, P. M. *J. Mol. Biol.* **1994,** *241* (1), 83-93.
49. Polgár, L. *Cell. Mol. Life Sci.* **2005,** *62* (19), 2161-2172.
50. Fruton, J. S. The Specificity and Mechanism of Pepsin Action. In *Adv. Enzymol. Relat. Areas Mol. Biol.,* John Wiley & Sons, Inc.: 1970; pp 401-443.
51. Fruton, J. S. Pepsin. In *The Enzymes,* Paul, D. B., Ed. Academic Press: 1971; Vol. 3, pp 119-164.
52. Dunn, B. M. *Chem. Rev.* **2002,** *102* (12), 4431-4458.
53. Antonov, V. K. New Data on Pepsin Mechanism and Specificity. In *Acid Proteases:Structure, Function, and Biology,* Tang, J., Ed. Springer US: Boston, MA, 1977; pp 179-198.
54. Powers, J. C.; Harley, A. D.; Myers, D. V. Subsite Specificity of Porcine Pepsin. In *Acid Proteases:Structure, Function, and Biology,* Tang, J., Ed. Springer US: Boston, MA, 1977; Vol. 92, pp 141-157.
55. Largman, C.; Brodrick, J. W.; Geokas, M. C. *Biochemistry* **1976,** *15* (11), 2491-2500.
56. Shotton, D. M. Elastase. In *Methods Enzymol.,* Academic Press: 1970; Vol. 19, pp 113-140.
57. Sweeney, P. J.; Walker, J. M. Proteinase K (EC 3.4.21.14). In *Enzymes of Molecular Biology,* Burrell, M. M., Ed. Humana Press: Totowa, NJ, 1993; pp 305-311.
58. Czapinska, H.; Otlewski, J. *Eur. J. Biochem.* **1999,** *260* (3), 571-595.
59. Hedstrom, L. *Chem. Rev.* **2002,** *102* (12), 4501-4524.
60. Sweeney, P. J.; Walker, J. M. Proteolytic Enzymes for Peptide Production. In *Enzymes of Molecular Biology,* Burrell, M. M., Ed. Humana Press: Totowa, NJ, 1993; pp 277-303.
61. Brady, K.; Abeles, R. H. *Biochemistry* **1990,** *29* (33), 7608-7617.
62. Schellenberger, V.; Jakubke, H.-D. *Biochim. Biophys. Acta, Protein Struct. M.* **1986,** *869* (1), 54-60.
63. Schellenberger, V.; Turck, C. W.; Rutter, W. J. *Biochemistry* **1994,** *33* (14), 4251-257.

29

64. Ye, S.; Loll, B.; Berger, A. A.; Mülow, U.; Alings, C.; Wahl, M. C.; Koksch, B. *Chem. Sci.* **2015,** *6* (9), 5246-5254.

65. Antal, J.; Pál, G.; Asbóth, B.; Buzás, Z.; Patthy, A.; Gráf, L. *Anal. Biochem.* **2001,** *288* (2), 156-167.

66. Sampath-Kumar, P. S.; Fruton, J. S. *Proc. Natl. Acad. Sci.* **1974,** *71* (4), 1070-1072.

67. Fruton, J. S. *Acc. Chem. Res.* **1974,** *7* (8), 241-246.

68. Hamuro, Y.; Coales, S. J.; Molnar, K. S.; Tuske, S. J.; Morrow, J. A. *Rapid Commun. Mass Spectrom.* **2008,** *22* (7), 1041-1046.

69. Rao, C.; Dunn, B. M. Evidence for Electrostatic Interactions in the S2 Subsite of Porcine Pepsin. In *Aspartic Proteinases: Structure, Function, Biology, and Biomedical Implications,* Takahashi, K., Ed. Springer US: Boston, MA, 1995; pp 91-94.

70. Fujinaga, M.; Chernaia, M. M.; Tarasova, N. I.; Mosimann, S. C.; James, M. N. *Protein Sci.* **1995,** *4* (5), 960-972.

71. Kageyama, T. *Cell. Mol. Life Sci.* **2002,** *59* (2), 288-306.

72. Dunn, B. M.; Hung, S.-H. *Biochim. Biophys. Acta, Protein Struct. M.* **2000,** *1477* (1–2), 231-240.

73. Naughton, M. A.; Sanger, F. *Biochem. J* **1961,** *78* (1), 156-163.

74. Renaud, A.; Lestienne, P.; Hughes, D. L.; Bieth, J. G.; Dimicoli, J. L. *J. Biol. Chem.* **1983,** *258* (13), 8312-8316.

75. Qasim, M. A. *Protein Pept. Lett.* **2014,** *21* (2), 164-170.

76. Atlas, D. *J. Mol. Biol.* **1975,** *93* (1), 39-53.

77. Atlas, D.; Levit, S.; Schechter, I.; Berger, A. *FEBS Lett.* **1970,** *11* (4), 281-283.

78. Ebeling, W.; Hennrich, N.; Klockow, M.; Metz, H.; Orth, H. D.; Lang, H. *Eur. J. Biochem.* **1974,** *47* (1), 91-97.

79. Georgieva, D.; Genov, N.; Voelter, W.; Betzel, C. *Z. Naturforsch., C: Biosci.* **2006,** *61* (5-6), 445-452.

80. Morihara, K.; Tsuzuki, H. *Agric. Biol. Chem.* **1975,** *39* (7), 1489-1492.

81. Brömme, D.; Peters, K.; Fink, S.; Fittkau, S. *Arch. Biochem. Biophys.* **1986,** *244* (2), 439-446.

82. Kraus, E.; Femfert , U. *Hoppe-Seyler's Z. Physiol. Chem.* **1976,** *357* (2), 937-947.

83. Betzel, C.; Singh, T. P.; Visanji, M.; Peters, K.; Fittkau, S.; Saenger, W.; Wilson, K. S. *J. Biol. Chem.* **1993,** *268* (21), 15854-15858.

84. Wolf, W. M.; Bajorath, J.; Müller, A.; Raghunathan, S.; Singh, T. P.; Hinrichs, W.; Saenger, W. *J. Biol. Chem.* **1991,** *266* (26), 17695-17699.

85. Betzel, C.; Bellemann, M.; Pal, G. P.; Bajorath, J.; Saenger, W.; Wilson, K. S. *Proteins: Struct., Funct., Bioinf.* **1988,** *4* (3), 157-164.

86. Kraus, E.; Kiltz, H.-H.; Femfert , U. F. *Hoppe-Seyler's Z. Physiol. Chem.* **1976,** *357* (1), 233-237.

Supporting Information

Position-Dependent Impact of Hexafluoroleucine and

Trifluoroisoleucine on Protease Digestion

Susanne Huhmann[a], Anne-Katrin Stegemann[a], Kristin Folmert[a], Damian Klemczak[a], Johann Moschner[a], Michelle Kube[a] and Beate Koksch[a,*]

[a]*Department of Chemistry and Biochemistry, Freie Universität Berlin, Takustraße 3, 14195 Berlin, Germany*

* Corresponding author. Tel.: +49 30 838-55344; fax: +49 30 838-55644.
  E-mail address: beate.koksch@fu-berlin.de

Synthesis of Fmoc-HfLeu-OH

General information

All reactions were run under an argon atmosphere unless otherwise indicated. Room temperature refers to 22°C. Reagents and anhydrous solvents were transferred *via* oven-dried syringe or cannula. Flasks were flame-dried under vacuum and cooled under a constant stream of argon. Reactions were monitored by thin layer chromatography using Merck KGaA silica gel 60 F_{254} TLC aluminium sheets and visualized with ceric ammonium molybdate, vanillin staining solution or potassium permanganate staining solution. Chromatographic purification was performed as flash chromatography on Macherey-Nagel GmbH & Co. KG silica gel 60 M, 0.04 – 0.063 mm, using a forced flow of eluent (method of Still). Concentration under reduced pressure was performed by rotary evaporation at 40°C at the appropriate pressure. Yields refer to chromatographically purified and spectroscopically pure compounds. NMR measurements were recorded on a JEOL-ECX400 (operating at 400 MHz for ^{1}H-NMR, 101 MHz for ^{13}C-NMR and 376 MHz for ^{19}F-NMR). Chemical shifts δ are reported in ppm with the solvent resonance as the internal standard. Coupling constants J are given in Hertz (Hz). Multiplicities are classified by the following abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, br = broad or m = multiplet and combinations thereof. High resolution mass spectra were obtained on an Agilent ESI-ToF 6220 (Agilent Technologies, Santa Clara, CA, USA).

Scheme S1: Synthesis of Fmoc-HfLeu-OH **9**.
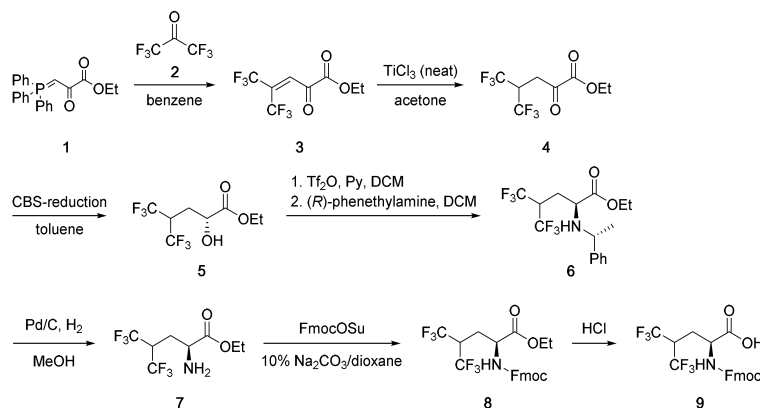
Compounds **3** to **7** were synthesized according to literature [1-2]. Obtained NMR data (^{1}H, ^{13}C, and ^{19}F) are consistent with literature [1-2].

Synthesis of Fmoc-HfLeu-OEt (**8**)

(**S**)-**7** (1.04 g, 3.89 mmol) was dissolved in 10% Na_2CO_3, aq (4 mL) and cooled to 0°C. Dioxane (1 mL) was added and the suspension wa s stirred for 15 min at 0°C after which FmocOSu (1.44 g, 4.28 mmol) was added. The mixture was stirred for 3 h at 0°C and at room temperature overnight. The react ion was diluted with H_2O (50 mL) and extracted with Et_2O (4 x 25 mL). The combined organic layers were concentrated *in vacuo* and the residue was subjected to column chromatography (*n*-hexane/Et_2O, 3:1) to give (**S**)-**8** (1.12 g, 2.29 mmol, 59%) as a waxy solid.

TLC: R_f = 0.45 (*n*-hexane/Et_2O, 5:1).

^{1}H-NMR (400 MHz, $CDCl_3$): δ = 7.77 (d, *J* = 7.5, 2H); 7.58 (d, *J* = 7.2, 2H); 7.40 (t, *J* = 7.5, 2H); 7.32 (t, *J* = 7.0, 2H); 5.40 (d, *J* = 7.50, 1H); 4.47 (dt, *J* = 20.0; 13.40, 3H); 4.32 – 4.18 (m, 3H); 3.18 (s, 1H); 2.40 (d, *J* = 14.5, 1H); 2.05 (d, *J* = 10.0, 1H); 1.30 (t, *J* = 7.1, 3H).

^{13}C-NMR (101 MHz, $CDCl_3$) δ = 170.68, 143.75, 143.53, 141.45, 141.44, 140.84, 130.32, 127.92, 127.90, 127.19, 127.17, 125.07, 124.98, 120.32, 120.15, 120.11, 67.29, 62.53, 51.96, 47.21, 37.15, 27.27, 14.13.

^{19}F-NMR (376 MHz, $CDCl_3$): δ = -67.27 – -67.44 (m), -67.63 – -67.79 (m).

HRMS calculated for $C_{23}H_{31}F_6NNaO_4$ [M+Na]$^+$: 512.1267; observed: 512.1294.

Synthesis of Fmoc-HfLeu-OH (**9**)

A solution of (**S**)-**8** (55.0 mg, 11.2 mmol) in HCl_{conc} (2 mL) was stirred at room temperature for 24 h. The crude product was lyophilized and purified *via* a LaPrepΣ low-pressure HPLC system (VWR, Darmstadt, Germany) using a Kinetex RP-C18 endcapped HPLC-column (5 µM, 100 Å, 250 x 21.2 mm, Phenomenex$^®$, USA). Deionized water and acetonitrile (ACN), both containing 0.1% (v/v) TFA served as eluents. A linear gradient of 30 – 100% ACN + 0.1% (v/v) TFA over 18 min with a flow rate of 20.0 mL/min was applied. UV-detection occurred at 280 nm. This gave (**S**)-**9** (36.3 mg, 7.87 mmol, 70%) as a white powder.

^{1}H-NMR (400 MHz, DMSO-D$_6$): δ = 7.82 (d, *J* = 7.6, 2H); 7.77 (d, *J* = 8.7, 1H); 7.63 (d, *J* = 7.5, 2H); 7.36 (t, *J* = 7.4, 2H); 7.26 (t, *J* = 7.4, 2H); 4.35 – 4.23 (m, 2H); 4.17 (t, *J* = 6.7, 1H); 4.04 (br, 1H); 2.30 – 2.17 (m, 1H); 2.13 – 2.01 (m, 1H).

^{13}C-NMR (101 MHz, DMSO-D$_6$): δ = 175.95 (s); 158.87 (s); 144.17 (s); 144.10 (s); 141.23 (s); 141.22 (s); 128.27 (s); 128.25 (s); 127.62 (s); 127.61 (s); 125.69 (s); 125.69 (s); 125.63 (s); 125.61 (s); 120.65 (s); 120.61 (s); 66.23 (s); 51.54 (s); 47.11 (s); 29.52 (s); 26.34 (s).

^{19}F-NMR (376 MHz, DMSO-D6): δ = -65.91 – -66.13 (m); -66.38 – -66.62 (m).

HRMS calculated for C$_{21}$H$_{17}$F$_6$NO$_4$ [M+Na]$^+$: 484.0954; observed: 484.0942.

Peptide synthesis, purification and characterization

Peptide synthesis

Peptides containing HfLeu were synthesized on an Activo P11 Automated Peptide Synthesizer (Activotec, Cambridge, United Kingdom) working under nitrogen atmosphere. All other peptides, either non-fluorinated or TfIle containing, were synthesized manually under standard conditions.

Peptide characterization

High resolution mass spectra were recorded on an Agilent 6220 ESI-ToF LC–MS spectrometer (Agilent Technologies Inc., Santa Clara, CA, USA) to identify the pure peptide products. The samples were dissolved in a 1:1 mixture of water and acetonitrile containing 0.1% (v/v) TFA and injected directly into the spray chamber by a syringe pump using a flow rate of 10 µL min^{-1}. A spray voltage of 3.5 kV was used, the drying gas glow rate was set to 5 L min^{-1} and the nebulizer to 30 psi. The gas temperature was 300℃.

To verify purity of the synthesized peptides analytical HPLC was carried out on a Chromaster 600 bar DAD-System with CSM software (VWR/Hitachi, Darmstadt, Germany). The system works with a low-pressure gradient containing a HPLC-pump (5160) with a 6-channel solvent degasser, an organizer, an autosampler (5260) with a 100 µL sample loop, a column oven (5310) and a diode array flow detector (5430). A LUNA^{TM} C8 (2) column (5 µm, 250 x 4.6 mm, Phenomenex$^®$, Torrance, CA, USA) was used. As eluents water and ACN, both containing 0.1% (v/v) TFA were used, the flow rate was adjusted to 1 mL/min and the column was heated to 24℃. The used gradient method is shown in table S1. The UV-detection of the peptides occurred at 220 nm. The data were analyzed with EZChrom Elite software (version 3.3.2, Agilent Technologies, Santa Clara, CA, USA).

Table S1: Used linear gradient for the purity determination of the synthesized peptides.

Time [min]	Water + 0.1% (v/v) TFA [%]	ACN + 0.1% (v/v) TFA [%]
0	95	5
18	30	70
19	0	100
21	0	100
21.5	95	5
24	95	5

Table S2: Identification of the synthesized peptides by ESI-ToF mass spectrometry and analytical RP-HPLC.

Peptide	Retention time [min]	Charge	m/z calculated	m/z observed
FA	10.597	+1	967.5364	967.5396
		+2	484.2721	484.2736
P2-LeuFA	12.500	+1	1009.5463	1009.5849
		+2	505.2956	505.2970
P2-HfLeuFA	12.393	+1	1117.4622	1117.5306
		+2	559.2573	559.2691
P2-IleFA	12.137	+1	1009.5463	1009.5849
		+2	505.2956	505.2971
P2-TfIleFA	12.493	+1	1063.4622	1063.5576
		+2	532.2814	532.2845
P1'-LeuFA	11.773	+1	1009.5463	1009.5863
		+2	505.2956	505.7982
P1'-HfLeuFA	11.917	+1	1117.4622	1117.5272
		+2	559.2573	559.2684
P1'-IleFA	11.370	+1	1009.5463	1009.5858
		+2	505.2956	505.2975
P1'-TfIleFA	11.870	+1	1063.4622	1063.5556
		+2	532.2814	532.2816
P2'-LeuFA	11.847	+1	1009.5463	1009.5866
		+2	505.2956	505.2981
P2'-HfLeuFA	12.197	+1	1117.4622	1117.5305
		+2	559.2573	559.2693
P2'-IleFA	11.557	+1	1009.5463	1009.5864
		+2	505.2956	505.2980
P2'-TfIleFA	12.283	+1	1063.4622	1063.5576
		+2	532.2814	532.2835
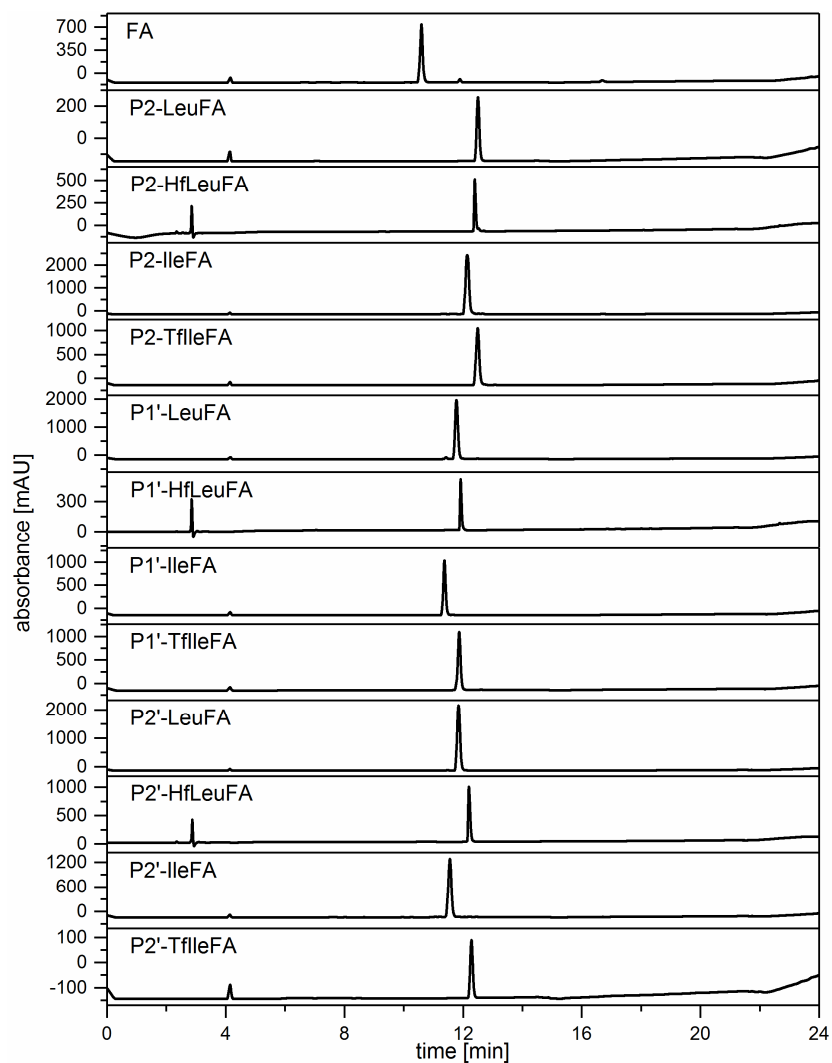
Figure S1: Analytical HPLC chromatograms of purified peptides; column: Luna^{TM}C8 (5 μM, 250 x 4.6 mm, Phenomenex®); Solvent A was H$_2$O, solvent B was acetonitrile, both containing 0.1% (v/v) TFA. The flow rate was 1 mL/min; linear gradient from 5% B to 70% B over 18 min (see table S1).

Enzymatic digestion studies

Characterization of the enzymatic digestion reactions was carried out *via* analytical HPLC on a LaChrom-ELITE-HPLC-System from VWR International Hitachi (Darmstadt, Germany). The system contains an organizer, two HPLC-pumps (L-

2130) with solvent degasser, an autosampler (L-2200) with a 100 µL sample loop, a diode array flow detector (L-2455), a fluorescence detector (L-2485) and a high pressure gradient mixer. As eluents water and ACN, both containing 0.1% (v/v) TFA were used, and a flow rate of 3 mL/min was applied. The used linear gradients are shown in table S3. For the non-fluorinated peptides method A was used to follow the digestion process, and for the fluorinated peptides method B was applied. For chromatograms where an insufficient baseline separation was observed, measurements were repeated using methods C [FA (pepsin), P2-LeuFA (proteinase K), P2-IleFA (pepsin), P2-IleFA (proteinase K), P1'-LeuFA (elastase), P1'-LeuFA (proteinase K), P1'-IleFA (proteinase K)] or D [P2-HfleuFA (proteinase K), P2-TflleFA (pepsin), P2-TflleFA (proteinase K), P1'-TflleFA (elastase), P2'-TflleFA (proteinase K)]. The obtained data were analyzed with EZChrom Elite software (version 3.3.2, Agilent Technologies, Santa Clara, CA, USA).

Table S3: Used linear gradients to follow the digestion process by FL-RP-HPLC.

Method	Time [min]	Water + 0.1% (v/v) TFA [%]	ACN + 0.1% (v/v) TFA [%]
A	0	95	5
	5	70	30
	5.5	70	30
	6	95	5
	9	95	5
B	0	95	5
	5	60	40
	5.5	60	40
	6	95	5
	9	95	5
C	0	95	5
	15	70	30
	15.5	70	30
	16	95	5
	17	95	5
D	0	95	5
	15	55	45
	15,5	55	45
	16	95	5
	17	95	5
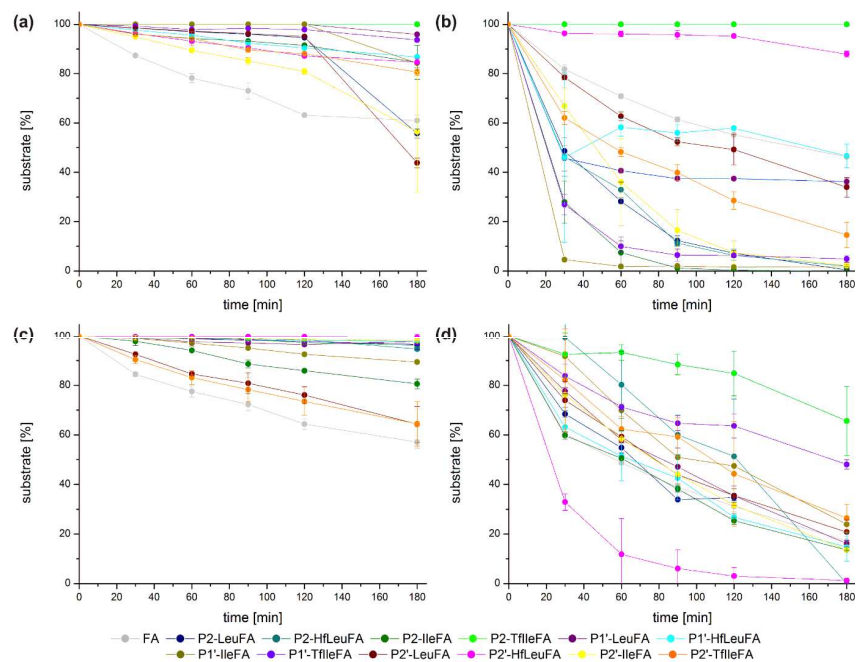
Figure S2: Chronological sequence of the substrate amount [%] over an incubation time of 180 min with **(a)** α-chymotrypsin, **(b)** pepsin, **(c)** elastase, and **(d)** proteinase K. The depicted values represent the mean of three independent measurements.

Identification of the proteolytic cleavage products (table S4 – S7) occurred according to the mass-to-charge ratios determined with an Agilent 6220 ESI-ToF MS instrument (Agilent Technologies, Santa Clara, CA, USA). For this, the quenched peptide-enzyme-solutions after 120 min and 24 h incubation were analyzed. The solutions were injected directly into the spray chamber using a syringe pump with a flow rate of 10 µL min^{-1}. Spray voltage was set to 3.5 kV, a drying gas flow rate of 5 L min^{-1} was used, the nebulizer was set to 30 psi, and the gas temperature to 300℃. The fragmentor voltage was 200 V. Not all corresponding fragments could be detected.

Table S4: Identification of the cleavage products of the different peptides by ESI-ToF mass spectrometry after digestion with α–chymotrypsin.

Peptide	Fragment	[M + H]$^{1+}$ calculated	[M + H]$^{1+}$ observed
FA	Abz-KAAFAAAAK	967.5364	967.5376
	Abz-KAAF	555.2559	555.2938
	AAAAK	431.2617	431.2627
P2-LeuFA	Abz-KALeuFAAAAK	1009.5463	1009.5883
	Abz-KALeuF	597.3029	597.2609
	AAAAK	431.2617	431.2617
P2-HfLeuFA	Abz-KAHfLeuFAAAAK	1117.4622	1117.5298

155

Peptide	Fragment	[M + H]$^{1+}$ calculated	[M + H]$^{1+}$ observed
P2-IleFA	Abz-KAIleFAAAAK	1009.5463	1009.5851
	Abz-KAIleF	597.3029	597.3435
	AAAAK	431.2617	431.2647
P2-TfIleFA	Abz-KATfIleFAAAAK	1063.4622	1063.5577
P1'-LeuFA	Abz-KAAFLeuAAAK	1009.5463	1009.5866
	Abz-KAAFLeu	668.3400	668.3760
	Abz-KAAF	555.2559	555.2907
	LeuAAAK	473.3087	473.3087
	AAAK	360.2246	360.2239
P1'-HfLeuFA	Abz-KAAFHfLeuAAAK	1117.4622	1117.5280
	Abz-KAAFHfLeu	776.2559	776.3214
	HfLeuAAAK	581.2246	581.2246
	Abz-KAAF	555.2559	555.2934
	AAAK	360.2246	360.3630
P1'-IleFA	Abz-KAAFIleAAAK	1009.5463	1009.5825
	Abz-KAAF	555.2559	555.2951
	IleAAAK	473.3087	473.3104
P1'-TfIleFA	Abz-KAAFTfIleAAAK	1063.4622	1063.5604
	Abz-KAAF	555.2559	555.2954
	TfIleAAAK	527.2246	527.2827
P2'-LeuFA	Abz-KAAFALeuAAK	1009.5463	1009.5872
	Abz-KAAF	555.2559	555.2922
	ALeuAAK	473.3087	473.3087
P2'-HfLeuFA	Abz-KAAFAHfLeuAAK	1117.4622	1117.5331
	AHfLeuAAK	581.2246	581.2550
	Abz-KAAF	555.2559	555.2965
P2'-IleFA	Abz-KAAFAIleAAK	1009.5463	1009.5875
	Abz-KAAF	555.2559	555.2943
	AIleAAK	473.3087	473.3112
P2'-TfIleFA	Abz-KAAFATfIleAAK	1063.4622	1063.5575
	Abz-KAAF	555.2559	555.2945
	ATfIleAAK	527.2246	527.2822

Table S5: Identification of the cleavage products of the different peptides by ESI-ToF mass spectrometry after digestion with pepsin.

Peptide	Fragment	[M + H]$^{1+}$ calculated	[M + H]$^{1+}$ observed
FA	Abz-KAAFAAAAK	96.5364	967.5434
	Abz-KAAF	555.2559	555.2967
	AAAAK	431.2617	431.2617
P2-LeuFA	Abz-KALeuFAAAAK	1009.5463	1009.5895
	Abz-KALeuF	597.3029	597.3438
	AAAAK	431.2617	431.2642
P2-HfLeuFA	Abz-KAHfLeuFAAAAK	1117.4622	1117.5302
	Abz-KAHfLeuF	705.2188	705.2870
	FAAAAK	578.3301	578.3327
	AAAAK	431.2617	431.2636
P2-IleFA	Abz-KAIleFAAAAK	1009.5463	1009.5916
	Abz-KAIleF	597.3029	597.3442
	AAAAK	431.2617	431.2647
P2-TfIleFA	Abz-KATfIleFAAAAK	1063.4622	1063.5639
P1'-LeuFA	Abz-KAAFLeuAAAK	1009.5463	1009.5926
	Abz-KAAFLeu	668.3400	668.3820
	Abz-KAAF	555.2559	555.2971

	LeuAAAK	473.3087	473.3126
	AAAK	360.2246	360.2271
P1'-HfLeuFA	Abz-KAAFHfLeuAAAK	1117.4622	1117.5325
	Abz-KAAFHfLeu	776.2559	776.3236
	HfLeuAAAK	581.2246	581.2553
	Abz-KAAF	555.2559	555.2956
	AAAK	360.2246	360.2273
P1'-IleFA	Abz-KAAFIleAAAK	1009.5463	1009.5908
	Abz-KAAF	555.2559	555.2969
	IleAAAK	473.3087	437.3087
P1'-TfIleFA	Abz-KAAFTfIleAAAK	1063.4622	1063.5634
	Abz-KAAF	555.2559	555.2969
	TfIleAAAK	527.2246	527.2843
P2'-LeuFA	Abz-KAAFALeuAAK	1009.5463	1009.5905
	Abz-KAAF	555.2559	555.2963
	ALeuAAK	473.3087	473.3117
P2'-HfLeuFA	Abz-KAAFAHfLeuAAK	1117.4622	1117.5307
	Abz-KAAFA	626.2930	626.3344
	HfLeuAAK	510.1875	510.2170
P2'-IleFA	Abz-KAAFAIleAAK	1009.5463	1009.5889
	Abz-KAAF	555.2559	555.2970
	AIleAAK	473.3087	473.3121
P2'-TfIleFA	Abz-KAAFATfIleAAK	1063.4622	1063.5627
	FATfIleAAK	674.2930	674.3530
	Abz-KAAFA	626.2930	626.3333
	Abz-KAAF	555.2559	555.2969
	ATfIleAAK	527.2246	527.2845
	TfIleAAK	456.1875	456.2462

Table S6: Identification of the cleavage products of the different peptides by ESI-ToF mass spectrometry after digestion with elastase.

Peptide	Fragment	[M + H]$^{1+}$ calculated	[M + H]$^{1+}$ observed
FA	Abz-KAAFAAAAK	967.5364	967.5352
	Abz-KAAFAAA	768.3672	768.4080
	Abz-KAAFAA	697.3301	697.3690
	AFAAAAK	649.3673	648.2935
	Abz-KAAFA	626.2930	626.3308
	AAAK	360.2246	360.2238
	AAK	289.1875	289.1872
P2-LeuFA	Abz-KALeuFAAAAK	1009.5463	1009.5871
	Abz-KALeuFAAAA	881.4513	881.4831
	Abz-KALeuFAAA	810.4142	810.4831
	Abz-KALeuFAA	739.4116	739.4116
	LeuFAAAAK	691.4142	691.4116
	Abz-KALeuFA	668.3404	668.3745
	AAAAK	431.2617	430.0496
	AAAK	360.2246	360.2205
	AAK	289.1875	289.1858
P2-HfLeuFA	Abz-KAHfLeuFAAAAK	1117.4622	1117.5325
	Abz-KAHfLeuFAAA	918.3301	918.3980
	Abz-KAHfLeuFAA	847.2930	847.3610
	Abz-KAHfLeuFA	776.2559	776.2559
	FAAAAK	578.3301	578.2440

P2-IleFA	Abz-KAIleFAAAAK	1009.5463	1009.5884
	Abz-KAIleFAAAA	881.4513	881.4960
	Abz-KAIleFAAA	810.4142	810.4960
	Abz-KAIleFAA	739.3771	739.4176
	Abz-KAIleFA	668.3400	668.3787
	FAAAAK	578.3301	578.3322
	AAAK	360.2246	360.2256
	AAK	289.1875	289.1880
	AK	218.1504	218.1505
P2-TfIleFA	Abz-KATfIleFAAAAK	1063.4622	1063.5576
	Abz-KATfIleFAAA	864.3301	864.4210
	Abz-KATfIleFAA	793.2930	793.3844
	Abz-KATfIleFA	722.2559	722.3470
	FAAAAK	578.3301	578.3271
	AAAK	360.2246	360.224
P1'-LeuFA	Abz-KAAFLeuAAAK	1009.5463	1009.5887
	Abz-KAAFLeuAAA	881.4513	881.4815
	Abz-KAAFLeuAA	810.4142	810.4472
	AAFLeuAAAK	762.4513	762.4458
	Abz-KAAFLeuA	739.3771	7394120
	AFLeuAAAK	691.4142	691.4142
	FLeuAAAK	620.3771	620.3783
	Abz-KAA	408.1875	408.1875
	Abz-KA	337.1504	337.1504
	AAK	289.1504	289.1858
P1'-HfLeuFA	Abz-KAAFHfLeuAAAK	1117.4622	1117.5330
	Abz-KAAFHfLeuAA	918.3301	918.3982
	AAFHfLeuAAAK	870.3673	870.3975
	AFHfLeuAAAK	799.3301	799.3600
	Abz-KA	337.1504	337.1864
P1'-IleFA	Abz-KAAFIleAAAK	1009.5463	1009.5853
	Abz-KAAFIleAA	810.4142	810.4500
	Abz-KAAFIleA	739.3771	739.4144
	FIleAAAK	620.3771	620.3747
P1'-TfIleFA	Abz-KAAFTfIleAAAK	1063.4622	1063.5578
	Abz-KAAFTfIleAA	864.3301	864.4205
	Abz-KAAFTfIleA	793.2930	793.3844
	AFTfIleAAAK	745.3301	745.3814
	AAK	289.1875	289.1857
	AK	218.1504	218.1483
P2'-LeuFA	Abz-KAAFALeuAAK	1009.5463	1009.5894
	Abz-KAAFALeuAA	881.4513	881.4938
	Abz-KAAFALeuA	810.4142	810.4938
	AFALeuAAK	691.4142	691.4168
	Abz-KAAFA	626.2930	626.3324
	LeuAAK	402.2716	402.2730
	Abz-KA	337.1504	337.1872
P2'-HfLeuFA	Abz-KAAFAHfLeuAAK	1117.6622	1117.5330
	Abz-KAAFAHfLeuA	918.3301	918.4349
	AAFAHfLeuAAK	870.3673	870.4001
	AFAHfLeuAAK	799.3301	799.3612
	FAHfLeuAAK	728.2930	728.3229
	Abz-KAAFA	626.2930	626.3230
	Abz-KAAFA	408.1875	408.2262

P2'-IleFA	Abz-KAAFAIleAAK	1009.5463	1009.5866
	Abz-KAAFAIleAA	881.4513	881.4910
	Abz-KAAFAIleA	810.4142	810.4539
	Abz-KAAFAIle	739.3771	739.4171
	FAIleAAK	620.3771	620.3687
	Abz-KAA	408.1875	408.2252
P2'-TfIleFA	Abz-KAAFATfIleAAK	1063.4622	1063.5530
	Abz-KAAFATfIle	793.2930	793.3851
	FATfIleAAK	674.2930	674.3441
	Abz-KAA	408.1875	408.2233
	AAK	289.1875	289.1855

Table S7: Identification of the cleavage products of the different peptides by ESI-ToF mass spectrometry after digestion with proteinase K.

Peptide	Fragment	[M + H]$^{1+}$ calculated	[M + H]$^{1+}$ observed
FA	Abz-KAAFAAAAK	967.5364	967.5376
	AFAAAAK	649.3673	649.2762
P2-LeuFA	Abz-KALeuFAAAAK	1009.5463	1009.5863
	Abz-KALeuFAAA	810.4142	810.4536
	Abz-KALeuFAA	739.4116	739.4179
	Abz-KALeuFA	668.3404	668.3816
	Abz-KALeu	450.2345	450.2719
	AAK	289.1504	289.1880
P2-HfLeuFA	Abz-KAHfLeuFAAAAK	1117.4622	1117.5298
	Abz-KAHfLeuFAAA	918.3301	918.3970
	Abz-KAHfLeuFAA	847.2930	847.3608
	Abz-KAHfLeuFA	776.2559	776.3233
	Abz-KAHfLeu	558.1504	558.2168
	AAK	289.1875	289.1884
P2-IleFA	Abz-KAIleFAAAAK	1009.5463	1009.5872
	Abz-KAIleFAAA	810.4142	810.4551
	Abz-KAIleFAA	739.3771	739.4192
	Abz-KAIleFA	668.3400	889.3806
	Abz-KAIleF	597.3029	597.3422
	AAK	289.1875	289.1891
P2-TfIleFA	Abz-KATfIleFAAAAK	1063.4622	1063.5604
	Abz-KATfIleFAAA	864.3301	864.4204
	Abz-KATfIleFAA	793.2930	793.3895
	Abz-KATfIleFA	722.2559	722.3512
	Abz-KATfIleF	651.2188	651.3140
	AAAAK	431.2617	430.0513
	AAK	289.1875	289.1888
P1'-LeuFA	Abz-KAAFLeuAAAK	1009.5463	1009.5877
	Abz-KAAFLeuAA	810.4142	810.4552
	Abz-KAAFLeuA	739.3771	739.4182
	FLeuAAAK	620.3771	620.3813
	Abz-KAA	408.1875	408.2276
	AAK	289.1504	2891885
P1'-HfLeuFA	Abz-KAAFHfLeuAAAK	1117.4622	1117.5271
	FHfLeuAAAK	728.2930	728.3226
	Abz-KAA	408.1875	408.2261
P1'-IleFA	Abz-KAAFIleAAAK	1009.5463	1009.5878
	Abz-KAAFIleAA	810.4142	810.4545

	Abz-KAAFIleA	739.3771	739.4169
	FIleAAAK	620.3771	620.3795
	Abz-KAA	408.1875	408.2265
	AAK	289.1504	289.1891
P1'-TfIleFA	Abz-KAAFTfIleAAAK	1063.4622	1063.5580
	Abz-KAAFTfIleAA	864.3301	864.4253
	Abz-KAAFTfIleA	793.2930	793.3853
	FTfIleAAAK	647.2930	674.3504
	Abz-KAA	408.1875	408.2260
	AAK	289.1875	289.1880
	AK	218.1504	218.1508
P2'-LeuFA	Abz-KAAFALeuAAK	1009.5463	1009.5845
	Abz-KAAFALeu	739.3771	739.4166
	FALeuAAK	620.3771	620.3782
	Abz-KAA	408.1875	408.2252
	AAK	289.1875	289.1883
P2'-HfLeuFA	Abz-KAAFAHfLeuAAK	1117.4622	1117.5304
	Abz-KAAFAHfLeu	847.2930	847.3603
	HfLeuAAK	510.1875	510.2172
	AAK	289.1875	268.1885
P2'-IleFA	Abz-KAAFAIleAAK	1009.5463	1009.5854
	Abz-KAAFA	626.2930	626.338
	FAIleAAK	620.3771	620.3791
	Abz-KAA	408.1875	408.2260
P2'-TfIleFA	Abz-KAAFATfIleAAK	1063.4622	1063.5563
	FATfIleAAK	674.2930	674.3496
	Abz-KAA	408.1875	408.2239

References

1.   Chiu, H.-P.; Cheng, R. P. *Org. Lett.* **2007,** *9* (26), 5517-5520.
2.   Zhang, C.; Ludin, C.; Eberle, M. K.; Stoeckli-Evans, H.; Keese, R. *Helv. Chim. Acta* **1998,** *81* (1), 174-181.

General Conclusion

Four models for peptide engineering have been described as possible approach for four different questions. Project **A** raises the use of enzymatic phosphorylation as a tool to control amyloid formation. Project **B** used a sequence part of a natural peptide to gain insights into a pathologic dysfunction. With project **C** a light-sensitive cage for the transport of a model peptide into cells was described and project **D** did concentrate on the question, how fluorinated amino acids could help to increase the proteolytic stability of a peptide.

Project A

- A model peptide was designed with characteristic of α-helical coiled-coil, random-coil and β-sheet structures and with the propensity to form amyloid fibrils with time.
- The peptide includes a recognition motif for PKA.
- Enzymatic phosphorylation can inhibit amyloid formation in dependency from its position to the hydrophobic core domain of the peptide.
- The Inhibition occured in the early fibril growing phase where small oligomers and protofilaments prevail.
- Premature influence of phosphorylation can reverse the conformational transition.

Project B

- The short hydrophobic patch domain NFGAIL can be used to study the aggregation process and structural pattern of T2D derived hIAPP.
- The contingent of β-sheets, turns and random structures during the critical oligomerization process was described in a cooperation project.
- The aggregation pathway was coined by a coexistence of unordered off-pathway and well-ordered partially β-sheet containing on-pathway oligomers.
- The morphology and diameters of grown fibrils were described by TEM and SAXS, concluding in a flat 2-dimensional expansion with the high of roughly one protofilament, a diameter of 2-5 protofilaments slightly coiled around a symmetry axis and some hundreds of nanometers length.

- A suggestion for the symmetry classes 1, 2 and 5 for the alignment of β-sheets in the protofilament was discussed.

Project C

- The design of a cysteinyl Bhc caging linker and its optimized synthesis was described in detail.
- Caging methods of four different peptides were described and alternative strategies to overcome recorded problems are presented.
- Suitability of two different fluorescence dyes, TAMRA and Cy5, their coupling to the model peptides, with and without linker, and all possible (dis)advantages were discussed.
- The influence of Cy5 labeling on the peptide aggregation of a known model peptide was evaluated.

Project D

- Idealized protease substrate FA was modified with leucine and the highly fluorinated analogue HfLeu in three different positions next to the desired cleavage site.
- The seven resulting substrate peptides were tested regarding their resistance towards proteolysis by α-chymotrypsin, pepsin, proteinase K, and elastase.
- Most resistant peptide compared to FA and leucine analogues was P2'HfLeuFA, except for proteinase K.
- The effect on proteolytic digestion depends on the intrinsic position of the fluorinated residue in the substrate and the preference of the enzyme.
- The expectation of a general increase in proteolytic stability, as a result of steric occlusion of the peptide from the active site upon incorporation of sterically demanding fluorinated amino acids, could not be verified based on the results.

All four models are highly specialized for their certain application but the gained knowledge can be transferred into a generalized toolbox for the design of more functional biomaterials, or drugs or in comprehension of pathologic mechanisms. Taking together all information from model peptide studies could led to a completely new scientific age of peptide engineering.

Experimental Part

E1 General Proceedings

E1.1 Chemicals and other Materials

Fmoc-L-amino acids were purchased from ORPEGEN Peptide Chemicals GmbH (Heidelberg, Germany). Resins were obtained from NovaBiochem (Merck Chemicals GmbH, Darmstadt, Germany). All solvents and common chemicals were used from VWR (Darmstadt, Germany) without further purification.

E1.2 Solid Phase Peptide Synthesis

Unless otherwise stated, the peptides were synthesized in 0.05 mmol batches according to standard Fmoc-chemistry using preloaded Fmoc-AA-NovaSyn®TGA resin. Standard couplings were performed in DMF with Fmoc-amino acids and 1-hydroxybenzotriazole [HOBt] / *N,N'*-diisocarbodiimide [DIC] in eight-fold excess with respect to the resin amount and with double couplings of one hour coupling time. A 2 L mixture of 1,8-diazabicyclo[5,4,0]undec-7-en [DBU] and piperidine (2% each) in DMF was used for Fmoc-deprotection (3x10 min). The resin was washed between each step with DMF and DCM (3x6 mL each). Peptides were cleaved from the resin by treatment with 2 mL TFA /TIPS /H_2O (90/5/5) for three hours. The resin was washed twice with 1 mL TFA and DCM, and excess of solvent was removed by evaporation. The peptides were precipitated with -20°C cold Et_2O, pelleted by centrifugation and dried by lyophilization before purification with preparative RP-HPLC.

E1.3 Test Cleavage

Test cleavages during peptide synthesis were performed with small amounts of the resin separated from the synthesis batch. For analysis with HPLC, test cleavage is recommended before Fmoc-deprotection. Therefore, 200 µL TFA, 20 µL TIPS and 5 µL H_2O were used for common peptides and pure TFA for peptides with Cy5 or TAMRA labeling. In any case, the reaction mixture was shaken for 2 h and evaporated, before precipitating the peptide with -20°C Et_2O. After decanting the Et_2O, peptides have been dissolved for analysis with HPLC and ESI-ToF MS. Only the small hydrophobic peptides NFGAIL, TAMRA-Leu-OH, TAMRA-AKA-OH,

Fmoc-AKA-OH, Cy5-Ala-OH and Cy5-AKA-OH were not treated with Et_2O after evaporating the TFA mixture.

E1.4 Lyophilization

For lyophilization the compound was dissolved in H_2O and frozen by the use of liquid nitrogen, before sublimation of the H_2O by using an Alpha 1-2 LD lyophilizator (Martin Christ GmbH, Osterode, Germany) and a vacuum pump RZ 2 from Vacuumbrand GmbH und CoKG (Wertheim, Germany).

E1.5 Preparative HPLC

Table E1: *Gradients for preparative HPLC.*

Gradient P1

time [min]	A [%]	B [%]	flow [mL/min]
0	95	5	10
5	95	5	10
30	30	70	20
32	0	100	20
35	0	100	20
36	95	5	20
40	95	5	20

Gradient P2

time [min]	A [%]	B [%]	flow [mL/min]
0	95	5	10
5	95	5	10
30	0	100	20
35	0	100	20
36	95	5	20
40	95	5	20

Gradient P3

time [min]	A [%]	B [%]	flow [mL/min]
0	70	30	20
30	0	100	20
32	0	100	20
33	70	30	20
36	70	30	20

Preparative HPLC was performed on a Knauer low-pressure HPLC system (Knauer GmbH, Berlin, Germany) sold by VWR (Darmstadt, Germany), comprising a LaPrep Sigma preparative pump (LP1200), a ternary low-pressure gradient, a dynamic mixing chamber, a 6-port-3-channel injection valve with an automated preparative 10 mL sample loop, a LaPrep Sigma standard 1-channel-UV-detector (LP3101), a flow cell with 0.5 mm thickness and a 16-port LaPrep Sigma fractionation valve (LP2016). A Kinetex RP-C18 endcapped (5 µM, 100 Å, 250 x 21.2 mm, Phenomenex®, USA) HPLC-column was used. A Security GuardTM PREP Cartridge Holder Kit (21.20 mm, ID, Phenomenex®, USA) served as pre-column. As eluents water and ACN,

both containing 0.1% (v/v) TFA were applied. UV-detection occurred at 220 nm. Data analysis occurred with an EZChrom Elite-Software (Version 3.3.2 SP2, Agilent).

E1.6 Analytical HPLC

LaChrom-Elite HPLC (Rosi)

The analytical LaChrom-ELITE-HPLC-System from VWR International Hitachi (Darmstadt, Germany) contains an organizer, two HPLC-pumps (L-2130) with solvent degaser, an autosampler (L-2200) with a 100 μL sample loop, a diode array flow detector (L-2455), a fluorescence detector (L-2485) and a high pressure gradient mixer. The data were analysed with EZ Chrom ELITE software (version 3.3.2, Agilent). A Kinetex C18 column (5 μm, 250 Å, 4.6 mm, Phenomenex®, Torrance, CA, USA) was used. A flow rate of 1.0 mL/min was used, with H_2O and ACN, both containing 0.1% (v/v) TFA as eluents.

Chromaster (slow Fritz)

The analytical Chromaster 600 bar DAD-System uses EZ Chrom ELITE software (version 3.3.2, Agilent) and works with a low-pressure gradient containing a HPLC-pump (5160) with a 6-channel solvent degaser, an organizer, an autosampler (5260) with a 100 μL sample loop, a column oven (5310) and a diode array flow detector (5430). A Kinetex C18 column (5 μm, 250 Å, 4.6 mm, Phenomenex®, Torrance, CA, USA) was used. As eluents water and ACN, both containing 0.1% (v/v) TFA were applied. A flow rate of 1 mL/min was used and the column was heated to 24°C.

Semi-micro Chromaster (fixer Fritz)

Analytical HPLC was carried out on a Chromaster 600 bar DAD-System with EZ Chrom ELITE software (version 3.3.2, Agilent). The system works with a low-pressure gradient containing a HPLC-pump (5160) with a 6-channel solvent degaser, an organizer, an autosampler (5260) with a 20 μL sample loop, a column oven (5310) and a diode array flow detector (5430) with a high pressure semi-micro flow cell. A Purospher®STAR RP-C18 endcapped (2 μM, 50 x 2.1 mm, Merk, Deutschland) UHPLC column was used. As eluents water and ACN, both containing 0.1% (v/v) TFA were applied. A flow rate of 0.6 mL/min was used and the column was heated to 24°C.

Table E2: Gradients for analytical HPLC.

Gradient A4

time [min]	A [%]	B [%]	flow [mL/min]
0	95	5	1
18	30	70	1
18.5	0	100	1
20	0	100	1
21	95	5	1
24	95	5	1

Gradient A5

time [min]	A [%]	B [%]	flow [mL/min]
0	95	5	1
30	30	70	1
32	0	100	1
35	0	100	1
36	95	5	1
40	95	5	1

Gradient A6

time [min]	A [%]	B [%]	flow [mL/min]
0	95	5	1
18	30	70	1
19	0	100	1
21	0	100	1
21.5	95	5	1
30	95	5	1

Gradient A7

time [min]	A [%]	B [%]	flow [mL/min]
0	95	5	1
6	0	70	1
6.5	0	100	1
7	0	100	1
7.1	95	5	1
10	95	5	1

Gradient A8

time [min]	A [%]	B [%]	flow [mL/min]
0	95	5	3
5	70	30	3
5.5	70	30	3
6	95	5	3
9	95	5	3

Gradient A9

time [min]	A [%]	B [%]	flow [mL/min]
0	95	5	3
5	60	40	3
5.5	60	40	3
6	95	5	3
9	95	5	3

Gradient A10

time [min]	A [%]	B [%]	flow [mL/min]
0	95	5	3
15	70	30	3
15.5	70	30	3
16	95	5	3
17	95	5	3

Gradient A11

time [min]	A [%]	B [%]	flow [mL/min]
0	95	5	3
15	55	45	3
15.5	55	45	3
16	95	5	3
17	95	5	3

E1.7 ESI-ToF Mass Spectrometry

Mass-to-charge ratios were determined with an Agilent 6220 ESI-ToF MS instrument (Agilent Technologies, Santa Clara, CA, USA). Analyte was injected by a syringe pump with a flow rate of 10 µL min^{-1}. Spray voltage was set to 4000 V, drying gas flow rate was 5 L min^{-1} and gas temperature was set to 300°C.

E1.8 Fluorescence Spectroscopy

Fluorescence spectra were recorded by using a 1 cm path length quartz cuvette (Hellma, Müllheim, Germany) and a luminescence spectrometer LS50B (Perkin-Elmer, Boston, MA, USA).

E1.9 UV/Vis Spectroscopy

UV/Vis measurements were performed with a Cary 50 UV/Vis spectrophotometer from Fa. Varian (Darmstadt, Germany).

E1.10 Circular Dichroism Spectroscopy

CD spectra were recorded by using a Jasco J-810 spectropolarimeter (Jasco, Gross-Umstadt, Germany) at 24°C (Jasco PTC-348W1 peltier thermostat) using 2 mm path length Quartz Suprasil cuvettes (Hellma, Müllheim, Germany). After background correction, the spectra were averaged over three scans (λ=195-240 nm; 0.5 nm intervals; 2 mm bandwidth; 4 s response time, 100 nm min^{-1} scanning speed). Ellipticity was normalized to concentration (c [mol L^{-1}]), number of residues (n=27, including the N-terminal labeling) and path length (l [cm]) by using equation (1) in which Θ_{obs} is the measured ellipticity in millidegrees and Θ is the mean residue ellipticity in 10^3 deg cm^2 dmol^{-1} residue^{-1}.

E1.11 Nuclear Magnetic Resonace Spectroscopy

NMR measurements were recorded on a JEOL-ECX400 instrument. The NMR instrument was operating at 400 MHz for ^{1}H-NMR, 101 MHz for ^{13}C-NMR and 376 MHz for ^{19}F-NMR. Chemical shifts δ are reported in ppm with the solvent resonance as the internal standard. Coupling constants J are given in Hertz (Hz). Multiplicities are classified by the following abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, br = broad or m = multiplet and combinations thereof.

E2   Project A

E2.1   Peptide Synthesis

Peptides were synthesized manually according to standard Fmoc-chemistry using preloaded Fmoc-Leu-NovaSyn®TGA resin (0.3 mmol g^{-1}, Novabiochem). Fmoc-Ser(PO(OBzl)OH)-OH (Bachem, Weil am Rhein, Germany) was activated with O-(7-azabenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium-hexafluorophosphate [HATU] / HOBt 5-fold and 15-fold excess of *N,N*-diisopropylethylamine [DIPEA] with respect to the resin and two hour double couplings. All amino acids coupled after Fmoc-Ser(PO(OBzl)OH)-OH, including the fluorescence dye Abz were coupled with common Fmoc-coupling reagents and two hour double couplings. After cleavage the crude peptides were dissolved in 5 mL ACN/ H$_2$O/ DMSO (1:1:1) and purified with preparative HPLC and gradient P2. Characterization of both peptides was performed with the LaChrom-ELITE-HPLC-System using gradient A4 and ESI-ToF mass spectrometry.

Table E3: *Measured and calculated mass to charge ratios and retention times on analytical HPLC of the peptides.*

peptide	charge	calc. mass [m/z]	obs. mass [m/z]	t$_R$ [min]
KFM6	0	3980.9592	-	23.48
	2	1591.9174	1591.9449	
	3	1061.6142	1061.9654	
	4	7964.6260	796.7276	
P$_C$KFM6	0	4077.9292	-	27.15
	2	1640.9174	1640.9378	
	3	1094.2808	1094.4539	
	4	820.9626	820.9653	

E2.2   Concentration Determination and Sample Preparation

A stock solution was prepared by dissolving the purified, Abz labeled peptide in HFIP ($\approx$1 mg/ mL^{-1}) and sonicating for 15 minutes to dissolve all aggregates. 50 µL of this solution was aliquoted and dried under nitrogen flow, before the residue was dissolved in 1 mL 50 mM Tris/HCl buffer containing 10 mM MgCl$_2$ at pH 7.5. UV spectra were recorded in a 1 cm path length cuvette using a Cary 50 UV/Vis spectrophotometer and the absorbance maximum at

312 nm was compared to a standard curve of the dipeptide H$_2$N-Abz-Gly-OHxHCl to calculate the concentration of the peptide stock solution where **y** is used for the UV intensity and **x** for the concentration. The stock solution was stored at 20°C.

Abz concentration standard curve[70]: $y = 0.00323 \, x + 0.00261$ with $R^2 = 0.99993$ (X)

Calculated aliquots of the peptide stock solution were dried under nitrogen flow to generate 15 µM peptide concentration in 350 µL sample solution for CD or 500 µL for fluorescence assays. The dried peptide was, immediately before measurement, dissolved in 50 mM Tris/HCl buffer with 10 mM MgCl$_2$, including 5000 U PKA and 200 µM ATP for the enzymatic phosphorylation studies. For ThT measurements 20 µM ThT were added to the buffer solution. The pH was adjusted to 7.5 with 1 M NaOH. For time-dependent phosphorylation experiments, the buffer was prepared with just one of the phosphorylation components, while the other one was added at different time points.

E2.3 Thioflavin T Assay

For ThT assays the fluorescence intensity at 485 nm was recorded at different time points over a total time period of 24 hours and normalized to the starting value at t=5 min of one. The shown plots represent an average of three independent measurements. Spectra were recorded at room temperature from 470-500 nm after excitation at 450 nm (excitation slit width 5 nm; emission slit width 20 nm; scan speed = 500 nm min^{-1}; accumulations = 5).

E2.4 Transmission Electron Microscopy

Peptides, prepared as for CD spectroscopy measurements, were examined after 24 hours. Aliquots (6 µL) of the corresponding solution were placed for 60 seconds onto glow-discharged (60 sec plasma treatment at 8 W in BAL-TEC MED 020), carbon-coated collodium films on 400-mesh copper grids (Leica Microsystems, Wetzlar, Germany). After blotting and negative staining with 1% PTA, the grids were left to air-dry. The TEM images were recorded with a Philips CM12 transmission electron microscope (FEI, Oregon, USA) at 100 kV acceleration

voltage and at a primary magnification of 58000x on Kodak SO-163 negative film by using a defocus of 900 nm.

E2.5 ^{31}P Nuclear Magnetic Resonance

KFM6 (22 µM) was incubated with 200 µM ATP and 5000 U PKA in 50 mM Tris/HCl buffer with 10 mM $MgCl_2$ at pH 7.5 and 27±3°C. One phosphate of ATP was transferred during enzymatic phosphorylation to the peptide, leaving an ADP molecule. Compared to ATP, ADP has no β-phosphate with respect to the electronic properties, while the α-phosphate remains constant. The integral ratio from β-phosphate to α-phosphate was followed to determine the reaction ratio to ADP, which is linear proportional to the phosphorylation ratio of KFM6. The pointed data were accumulated for 25 min each. The resulting linear fit curve implies no significant change in the integral ratio during the examined time dimension.

^{31}P NMR (202 MHz, PKA reaction buffer, 27±3°C, ppm) δ= 6.31 (d, γ-phosphate), 11.45 (d, α-phosphate), 19.66 (dd, β-phosphate).

E3   Project B

E3.1   Peptide Synthesis

The peptide was synthesized manually according to standard Fmoc-chemistry using preloaded Fmoc-Leu-NovaSyn®TGA resin (0.3 mmol g^{-1}, Novabiochem). Different from the standard protocol, the peptide was cleaved from the resin by treatment with 2 mL TFA/ TIPS/ H$_2$O (90/9/1) for three hours using sonication at 30°C. For purification the lyophilized peptide was dissolved in pure MeOH. The purification was performed by using gradient P1and UV-detection at 220 nm.

Table E4: Observed and calculated mass to charge ratios and retention times on LaChrom ELITE analytical HPLC with gradient A4.

peptide	charge	calc. mass [m/z]	obs. mass [m/z]	t$_R$ [min]
NFGAIL	[M]	633.3485	-	
	[M+H]$^{1+}$	634.3564	634.3653	11.88
	[2M+H]$^{1+}$	1268.7306	1268.7227	

E3.2   Sample Preparation

A stock solution was prepared by dissolving the purified NFGAIL in HFIP (~18 mM) and was further sonicated for 15 min to dissolve all preformed aggregates. Aliquots of this stock solution were dried and then redissolved to a final concentration of 4 mM in ammonium acetate buffer (10 mM), containing 20 µM Thioflavin T. After dissolution, the sample was sonicated for 30 sec and then incubated at 37°C with 1300 rpm.

E3.3   Thioflavin T Fluorescence Assay

Fluorescence spectra were recorded using a 1 cm path length quartz cuvette and spectra were recorded at room temperature from 470-500 nm after excitation at 420 nm (excitation slit width 5 nm; emission slit width 10 nm; scan speed = 300 nm/min; accumalations = 3). The fluorescence intensity at 485 nm was normalized with respect to its maximum value. After each measurement the sample was retransferred into the incubation at 37°C with 1300 rpm.

E3.4 Small-Angle X-ray Scattering

The small-angle X-ray scattering measurements [SAXS] of the samples were performed with a SAXSess camera (Anton Paar, Austria). This Kratky type of camera is attached to a laboratory X-ray generator (PW3830, PANalytical), and is operated with a fine focus glass X-ray tube at 40 kV and 40 mA ($Cu_{K\alpha}$, λ = 0.1542 nm). A focusing multi-layer optic and a block collimator provide a monochromatic primary beam with low background. Samples were filled in a reusable vacuum tight flow cell sample holder. SAXS data (intensity as a function of the scattering vector) was recorded for 43200 s (4320 x 10 s) with a Mythen detection system (Dectris Inc.) in a q-range of 0.07 to 7.0 nm^{-1} (Anton Paar). The scattering vector is defined in terms of the scattering angle, θ and the wavelength, λ of the radiation, thus $q = 4\pi / \lambda \sin(\theta)$. The angle between incident and scattered beam is 2 θ. The two-dimensional intensity data was converted to one-dimensional data with SAXSQuant software (Anton Paar). The temperature of 21°C was controlled with a TCS 120 sample holder (Anton Paar) with an accuracy of ± 0.2°C. A reusable capillary was used for all measurements to attain the same scattering volume and background contribution. The resulting scattering curves were corrected for the contribution of the suspension medium (water) and the capillary. Furthermore, the data was desmeared using the length profile of the primary beam[509] with SAXSQuant (Anton Paar AG, Graz). As sample a 4 mM NFGAIL sample incubated 24 h at 37°C and 1300 rpm in 10 mM ammonium acetate buffer pH 7 was used. The buffer treated with same conditions preserved as blank.

E4. Project C

E4.1 Peptide Synthesis

The parts with natural amino acids of the peptides were synthesized manually according to standard Fmoc-chemistry using preloaded Fmoc-Leu-NovaSyn®TGA resin (0.15 mmol g^{-1}, Novabiochem) or Fmoc-Ala-Wang LL resin (0.32 mmol g^{-1}, Novabiochem). Orthogonally protected N-α-Fmoc-N-ε-4-methyltrityl-L-lysine [Fmoc-Lys(Mtt)-OH] and 6-(Fmoc-amino)-hexanoic acid [Fmoc-Ahx-OH] were both coupled twice with 4 eq. with respect to the resin and two hours coupling time. 1.2 eq. of 5(6)-Carboxytetramethylrhodamine succinimidyl ester [TAMRA-NHS] or Cyanine 5 succinimidyl ester [Cy5-NHS] were used for N-terminal peptide labeling. Therefore, the dye was dissolved in 1 mL DMF in an amber colored Eppendorf reactor and 4 eq. DIPEA were added directly before transferring the reaction mixture to the resin. After 12 h coupling time, under exclusion of light, the reaction mixture was removed and a test cleavage was performed to ensure the coupling achievement. Labeled peptides were kept in dark during all steps of synthesis, purification and analysis and peptides on resin were stored freeze-dried at -20°C.

Unless specified otherwise, the peptides were treated with conc. TFA for h for full cleavage. Afterwards the reaction mixture was caught in a dark flask and evaporated under reduced pressure before adding 5 mL water and freeze-dry the crude peptides. Purification was achieved with preparative HPLC and gradient P1 for TAMRA-Leu-OH, TAMRA-AKA and Cy5-Ala-OH. Gradient P3 was used for TAMRA-Ahx-VW18-OH.

Table E5: *Full cleaved and test cleaved* peptides with labeling.*

peptide	charge	calc. mass [m/z]	obs. mass [m/z]	t_R [min]	HPLC System	method
TAMRA-Leu-OH	1	544.2461	544.2465	4.073	Semi-micro Chromaster	A12
Fmoc-AK(Mtt)A*	1	511.5891	511.2561	13.917	Chromaster	A4
TAMRA-AK(Mtt)A*	1 2	701.3477 351.6738	701.3482 351.6859	13.733	LaChrom-ELITE	A4
TAMRA-Ahx-VW18-OH	3 4 5	1184.6084 888.7082 711.1681	1184.6878 888.7683 711.0197	25.147	LaChrom-ELITE	A5
Cy5-Ala-OH	1	554.34	554.3415	21.213	LaChrom-ELITE	A4
Cy5-AKA*	1	753.47	753.4759	18.633	LaChrom-ELITE	A6
Cy5-Ahx-VW18K9(Ac)-OH	3 4 5	1215.9417 912.2082 729.9681	1215.7300 912.5493 729.8418	19.953	LaChrom-ELITE	A4
Cy5-Ahx-VW18K9*	3 4 5	1201.9417 901.7082 721.5681	1202.9661 901.8023 721.8431	19,860	LaChrom-ELITE	A6

E4.2 6-Bromo-4-chloromethyl-7-hydroxycoumarin (Bhc-Cl, 3)
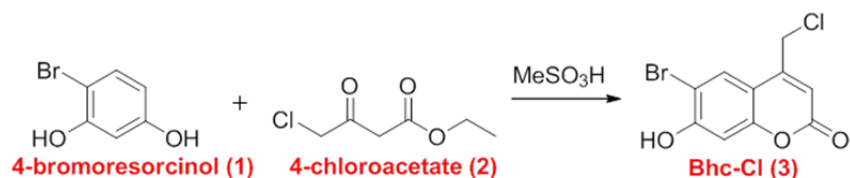
2.84 g 4-bromresorcinol **1** (15.0 mmol, 1.0 eq.) and 3 mL ethyl-4-chloracetoacetate **2** (22.5 mmol, 1.5 eq.) were dissolved in 20 mL methanesulfonic acid and stirred for 2 h at rt. The reaction progress was monitored by HPLC. After completion, the reaction was quenched by subjecting it into an ice/water bath. Successful quenching was characterized by a color change from black to white. The mixture was further stirred for 2 h at 0°C. Filtration and lyophilization led to 4 g brown solid as product **3** (15.0 mmol, 100%).

LaChrom-ELITE HPLC (gradient A5, 312 nm): t_R = 24.014 min
ESI(-)-ToF MS: calculated $[C_{10}H_5BrClO_3]^-$ = 286.9289; obtained $[M-H]^-$ = 286.9285
^{1}H NMR (Acetone-d$_6$, 400 MHz, ppm) δ 8.00 (s, 1H), 6.96 (s, 1H), 6.46 (s, 1H), 4.96 (s, 2H).

^{13}C NMR (Acetone-d$_6$, 100 MHz, ppm) δ 159.6, 157.4, 154.8, 149.8, 129.1, 112.8, 106.1, 103.7.

E4.3 6-Bromo-4-hydroxymethyl-7-hydroxycoumarin (Bhc-OH, 4)
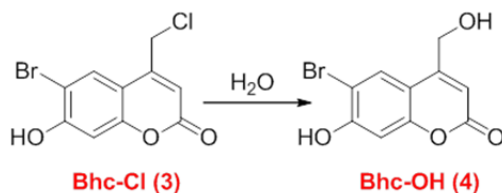
4.9 g Bhc-Cl **3** (18.14 mmol, 1 eq.) were suspended in 1 L H$_2$O and refluxed for 2 d. Three times per day the pH was titrated to 9.5-10 using 1 M NaOH. The reaction progress was monitored by HPLC. After the reaction was finished, the pH was set to 7 and the water lyophilized. The crude product was dissolved in acetone and purified with flash chromatography. The eluent was changed from EtOAc:Hex (1:1) to acetone:MeOH (9:1) after **8** was regained. After evaporating the solvents, 0.93 g **4** (3.44 mmol, 19%) was obtained as slightly yellow powder.

LaChrom-ELITE HPLC (gradient A5, 312 nm): t$_R$ = 15.927 min

ESI(-)-ToF MS: calculated [C$_{10}$H$_6$BrO$_4$]$^-$ = 270.9528; obtained [M-H]$^-$ = 270.9524

^{1}H NMR (Acetone-d$_6$, 400 MHz, ppm) δ 8.01 (s, 1H), 6.96 (s, 1H), 6.47 (s, 1H), 4.97 (s, 2H).
^{13}C NMR (Acetone-d$_6$, 100 MHz, ppm) δ 160.1, 157.1, 155.1, 128.2, 111.8, 108.5, 105.9, 103.5, 59.7, 55.8.

E4.4 N-Boc-N'-(Bhc)-ethylendiamin (Bhc-Boc, 5a)
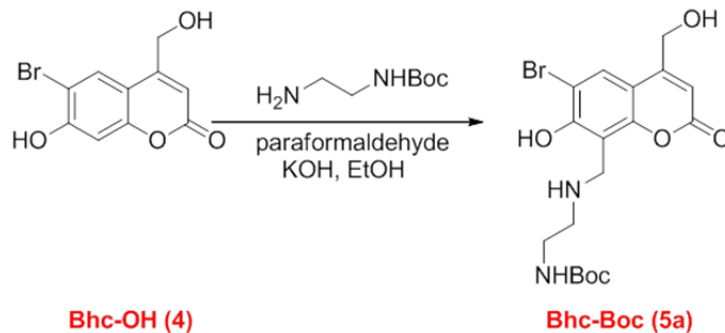
0.3 g paraformaldehyde (9.9 mmol, 3.3 eq.) and 16.4 mg KOH (0.29 mmol, 0.12 eq.) were suspended in 2 mL EtOH. 1.53 mL Boc-ethylenediamine (9.9 mmol, 3.3 eq.) was added dropwise at 0°C. The reaction mixture was stirred for 1 h at rt and was then slowly transferred into a boiling solution of 0.8 g Bhc-OH **4** (3.0 mmol, 1.0 eq.) in 50 mL EtOH. Combined mixture

was refluxed for 3 h. The reaction progress was monitored by HPLC. The solution was stored at -20°C until the product precipitated as yellow solid. The solid was filtrated and washed from cold EtOH. Recrystallization with EtOH and vacuum drying led to 0.52 g yellow powdered **5a** (1.17 mmol, 40%).

LaChrom-ELITE HPLC (gradient A5, 312 nm): t_R = 18.053 min

ESI(-)-ToF MS: calculated $[C_{18}H_{22}BrN_2O_8]^-$ = 442.0739; obtained $[M-H]^-$ = 442.0738

Note: Due to very low solubility in all solvents beside acids, no NMR measurement in appropriate quality was possible.

E4.5 (Bhc)-ethylendiamin TFA salt (Bhc-TFA, 6)
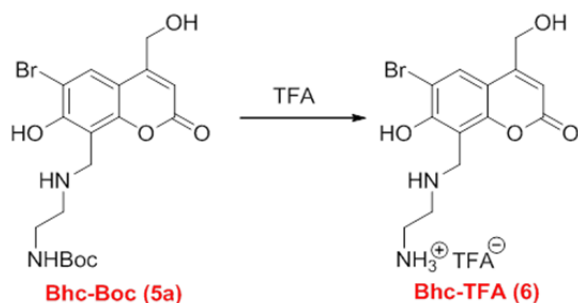
10 mL TFA was added to a solution of 160 mg Bhc-Boc **5a** (0.36 mmol, 1 eq.) in 10 mL DCM. Under slightly nitrogen pressure the reaction was stirred for 5 h. Subsequently, the solvents were removed leaving a brown oil. The oil was mixed with 10 mL water and lyophilized to yield Bhc-TFA **6** as amber colored oil (0.36 mmol, 100%).

LaChrom-ELITE HPLC (gradient A5, 312 nm): t_R = 11.333 min

ESI(-)-ToF MS: calculated $[C_{15}H_{12}BrF_3N_2O_6]^-$ = 453.0288; obtained $[M-H]^-$ = 453.0283

^{1}H NMR (DMSO-d_6, 400 MHz, ppm) δ 7.97 (s, 1H), 5.64 (s, 1H), 4.70 (s, 2H), 4.37 (s, 2H), 3.23 - 3.16 (m, 2H), 3.14 - 3.09 (m, 2H).

Note: The concentration was too low for an appropriate ^{13}C spectrum due to low solubility of the product.

E4.6 N-Boc-S-Trt-L-Cysteinylpentafluorophenylester (CysPF, 9)

Boc-Cys-Trt (7)   **pentafluorophenol (8)**   **CysPF (9)**
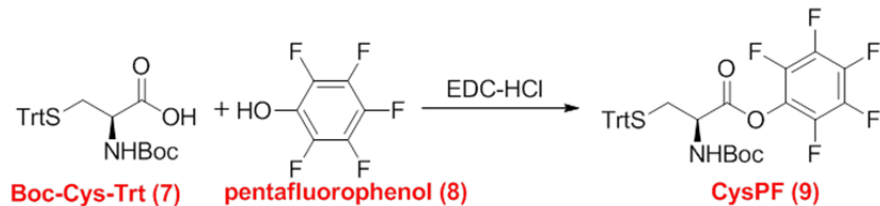
0.93 g Boc-Cys-Trt **7** (2.0 mmol, 1.0 eq.), 0.55 g pentafluorophenol **8** (3.0 mmol, 1.5 eq.) and 0.57 g EDC x HCl (3.0 mmol, 1.5 eq.) were dissolved in 10 mL DCM and stirred for 1 d. The reaction progress was monitored by HPLC and TLC (100% DCM). Afterwards the solvents were removed under reduced pressure. Almost dry product was foaming to a large volume. Residue was dissolved in 50 mL EtOAc and washed twice with brine, conc. Na_2CO_3 solution and water. The combined aqueous phases were re-extracted with EtOAc and the combined organic layers were dried with Na_2SO_4. After removing the solvent (foaming) the crude was dissolved in DCM and purified with flash chromatography with DCM as eluent. Dried and crushed **9** was achieved as 0.703 g white solid (1.12 mmol, 56%).

LaChrom-ELITE HPLC (gradient A5, 237 nm): t_R = 37.540 min

ESI(+)-ToF MS: calculated $[NaC_{33}H_{28}F_5NO_4S]^+$= 653.1479;

obtained $[M+Na+H]^+$= 653.1586

^{1}H-NMR (400 MHz, $CDCl_3$): δ = 1.44 (s, 9H, H_3-1), 2.65-2.82 (m, 2H, H_2-2), 4.30 (dd, 1H, H-3), 4.99 (d, 1H, NH), 7.21-7.45 (m, 15H, H-5) ppm.

^{19}F-NMR (400 MHz, $CDCl_3$): δ = -161.92 (t, 2F,F-1), -157.32 (t, 1F, F-2), - 151.45 (d, 2F, F-3) ppm.

E4.7 N-(N-Boc-S-Trt-L-Cysteinyl)-N'-(Bhc)-ethylendiamin (Cys-Bhc, 10)

Bhc-TFA (6)   CysPF (9)   Cys-Bhc (10)
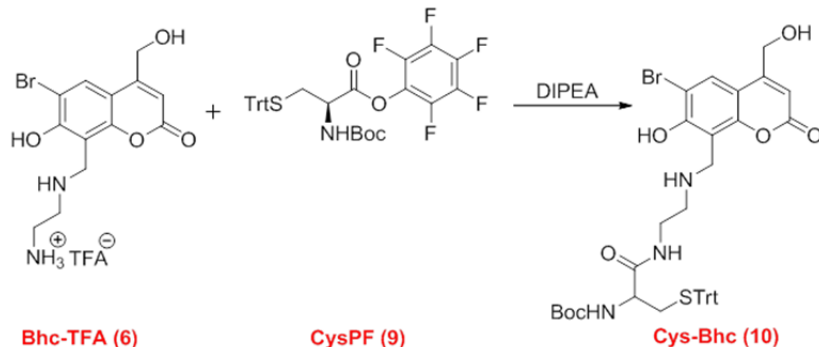
The reaction was performed under exclusion of light (instead of orange light) and dry conditions and in an amber colored flask. 306 mg Bhc-TFA **6** (0.67 mmol, 1 eq.) and 506 mg CysPF **9** (0.80 mmol, 1.2 eq.) were dissolved in 6 mL dry DMF. 242 µL DIPEA (1.3 mmol, 2 eq.) and was added to the reaction. The reaction was stirred for 4 h at 50°C. A nitrogen filled balloon and a septum were used to generate a constant pressure while the reaction was heated to 50°C and stirred for 4 h. The reaction progress was monitored by HPLC. Afterwards DMF and pentafluorophenol were removed at 50°C and 10^{-2} mbar. The crude was dissolved in DCM/MeOH (4:1) and purified with preparative HPLC and gradient P2 at 220 nm (t_R = 23 min). Both starting reagents **6** (t_R = 14 min) and **9** (t_R = 32 min) have been collected separately. 108 mg Cys-Bhc **10** (0.14 mmol, 21%) was obtained as colorless crystals.

$T_{decomposing} \approx 70°C$

Chromaster HPLC (gradient A5, 312 nm): t_R = 27.790 min

ESI(+)-ToF MS: calculated $[C_{40}H_{43}BrN_3O_7S]^+$ = 788.1927; obtained $[M+H]^+$ = 788.1933

^{1}H NMR (CDCl$_3$, 400 MHz, ppm) δ 7.42 - 7.16 (m, 16H), 6.38 (s, 1H), 5.01 (s, 1H), 4.73 (s, 2H), 4.42 (s, 2H), 3.82 - 3.62 (m, 2H), 3.59 - 3.49 (m, 2H), 3.36 - 3.25 (m, 2H), 1.30 (s, 9H).
Note: The concentration was too low for an appropriate ^{13}C spectrum due to low solubility of the product.

E4.8   N-(N-Boc-S-Trt-L-Cysteinyl)-N'-Ac-N'-(Bhc)-ethylendiamin (Cys'Bhc, 11)
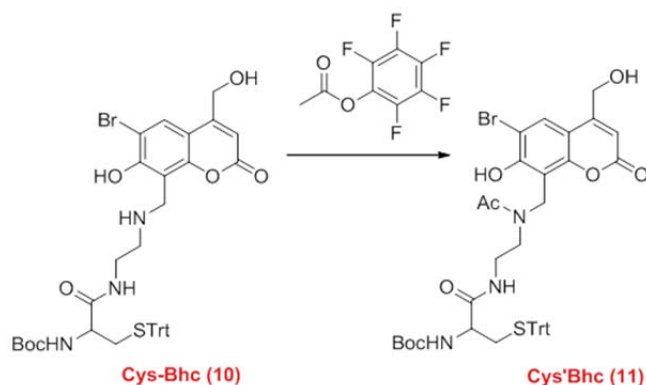
The reaction was performed under exclusion of light or with orange light if necessary and in an amber colored flask. 194 mg Cys-Bhc **10** (0.246 mmol, 1 eq.) was dissolved in 6 mL dry DMF. A solution of 262 mg pentafluorophenol acetate (0.984 mmol, 4 eq.) in 2 mL dry DMF was added dropwise. A nitrogen filled balloon and a septum were used to effect a constant pressure while the reaction was stirred at 50°C for 4 h. The reaction progress was monitored by HPLC by conveying 5 µL reaction solution into 50 µL ACN. Afterwards DMF and pentafluorophenol were removed at 50°C and 10^{-2} mbar. The crude was dissolved in ACN and purified with preparative HPLC and gradient P2 (t_R= 27 min) to yield 143 mg Cys'Bhc **11** (0.172 mmol, 70%) as a colorless solid.

$T_{decomposing} \approx 70°C$

Chromaster HPLC (gradient A6, 312 nm): t_R = 22.570 min

ESI(-)-ToF MS: calculated $[C_{42}H_{45}BrN_3O_8S]^- = 828.2032$; obtained $[M-H]^- = 828.1930$

^{1}H NMR (CDCl$_3$, 400 MHz, ppm) δ 7.43 - 7.17 (m, 16H), 6.42 (s, 1H), 5.16 (s, 2H), 4.89 (s, 1H), 4.67 (s, 2H), 3.60 - 3.46 (m, 4H), 2.78 - 2.71 (m, 4H), 2.54 - 2.48 (m, 1H), 1.40 (s, 9H).
Note: The concentration was too low for an appropriate ^{13}C spectrum due to poor solubility of the product.

E4.9 N-(N-Boc-S-Trt-L-Cysteinyl)-N'-Ac-N'-(Bhc-4-Succinimidylcarbamate)-ethylendiamin (Cys'Bhc-NHS, 12a)

Cys'Bhc (11)   Cys'Bhc-NHS (12a)
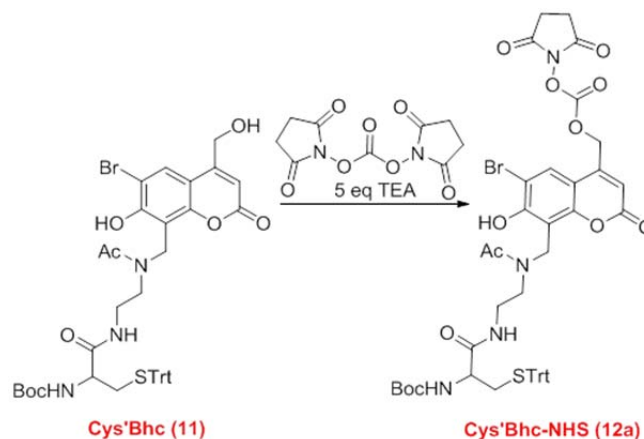
The reaction was performed under exclusion of light except orange light if necessary and in an amber colored flask. 204 mg Cys'Bhc **11** (0.246 mmol, 1 eq.) and 189 mg N,N'-succinimidyl carbonate [DSC] (0.738 mmol, 3 eq.) was dissolved in 6 mL ACN/THF (1:1). Afterwards, 188 μL fresh distilled triethylamine [TEA] (1.35 mmol, 5.5 eq.) was added. A nitrogen filled balloon and a septum were used to effect a constant pressure while the reaction was heated to 50°C and stirred for 3 h. The reaction progress was monitored by HPLC by conveying 5 μL reaction solution into 50 μL ACN. Afterwards the solvents were removed under reduced pressure and the crude mixed with H_2O and lyophilized. For purification the crude was dissolved in 6 mL ACN and purified with preparative HPLC and gradient P2 (t_R= min) to yield 38 mg Cys'Bhc-NHS **12a** (0.039 mmol, 16%) as colorless solid.

$T_{decomposing} \approx 70°C$

Chromaster HPLC (gradient A6, 312 nm): $t_R = 22.697$ min

ESI(+)-ToF MS: calculated $[C_{47}H_{48}BrN_4O_{12}S]^+ = 971.2095$; obtained $[M+H]^+ = 971.2089$

^{1}H NMR (CDCl$_3$, 400 MHz, ppm) δ 7.41 - 7.16 (m, 16H), 6.41 (s, 1H), 5.15 (s, 2H), 4.93 - 4.88 (m, 1H), 4.67 (s, 2H), 3.59 - 3.50 (m, 4H), 2.79 - 2.69 (m, 4H), 2.54 - 2.47 (m, 1H), 1.40 (s, 9H). Note: The concentration was too low for an appropriate ^{13}C spectrum due to poor solubility of the product. The measurement was quickly performed, in an amber colored NMR tube with CDCl$_3$ filtered over Al$_2$O$_3$, nevertheless the sensitive NHS-ester was not detectable in all 5 trials of different batches.

E4.10 Optimization Series for NHS Activation with Bhc-Boc (Bhc-Boc-NHS, 5b)

Bhc-Boc (5a) → Bhc-Boc-NHS (5b)
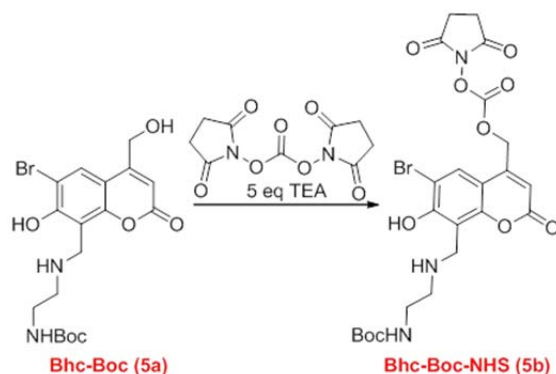
2.66 mg Bhc-Boc **5a** (0.006 mmol, 1 eq.) and 2.50 mg DSC were dissolved in 1 mL of variuos solvents (**Table C1** in section C3.3) and treated with different bases (0.033 mmol, 5.5 eq.). The reaction was stirred for 1 d at rt. The reaction progress was monitored with analytical HPLC Chromaster System and gradient A4 at 312 nm (t_R = 15.100 min). Therefore, 5 µL reaction solution were mixed with 50 µL ACN and injected. (conditions and results: **Table C1** in section C3.3)

E4.11 TAMRA-AK(Bhc)A-OH 1 (TAMRA-AK(NHS)A-OH, 16)
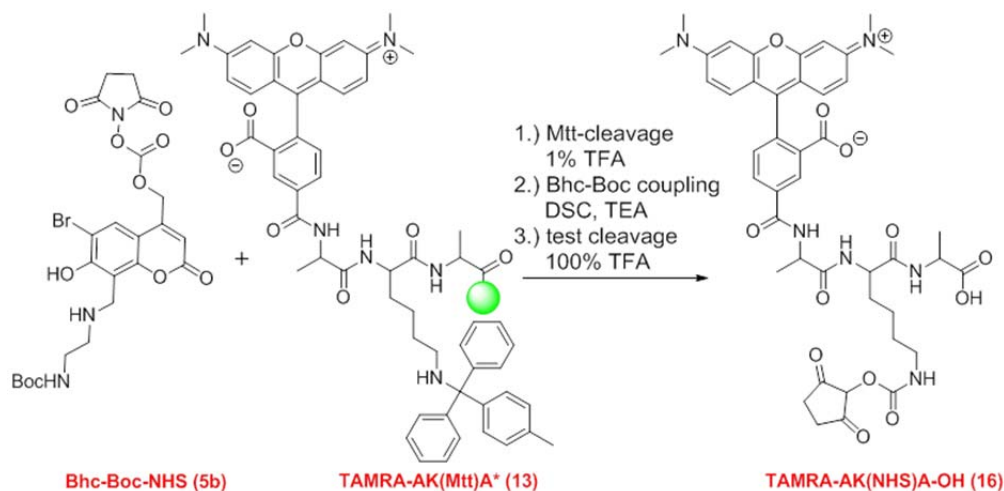
1.) Mtt-cleavage 1% TFA
2.) Bhc-Boc coupling DSC, TEA
3.) test cleavage 100% TFA

Bhc-Boc-NHS (5b) + TAMRA-AK(Mtt)A* (13) → TAMRA-AK(NHS)A-OH (16)

The reaction was performed under exclusion of light. For Mtt-cleavage TAMRA-AK(Mtt)A* **13** (0.005 mmol, 1 eq.) was alternately treated 3 x 1 min with 2 mL of 1% TFA in DCM and 1 x 1 min with 2 mL DCM. The cleavage cocktail turned yellow until no cleaved Mtt was left. Afterwards, the TAMRA-AKA* resin (**14**) was washed intensively with DCM. 3.85 mg Bhc-Boc **5b** (0.008 mmol, 1.7 eq.) and 4.29 mg DSC (0.009 mmol, 1.8 eq.) were dissolved in 1 mL

ACN/DCM (1:1), mixed with 5.9 µL TEA (0.042 mmol, 8.4 eq.) and added to the resin. After 12 h shaking at rt, the reaction mixture was removed and the resin washed with DCM. A test cleavage with 100% TFA for 2 h showed the NHS-activated peptide TAMRA-AK(NHS)A-OH **16** (100% HPLC area percent).

LaChrom ELITE HPLC (gradient A4, 240 nm): t_R = 15.647 min

ESI(+)-ToF MS: calculated $[C_{43}H_{49}N_6O_{12}]^+$ = 844.3330; obtained $[M+H]^+$ = 844.3378

E4.12 TAMRA-AK(Bhc)A-OH 2 (15)

1.) Mtt-cleavage 1% TFA
2.) Bhc-Boc-NHS coupling TEA
3.) test cleavage 100% TFA

Bhc-Boc-NHS (5b)   TAMRA-AK(Mtt)A* (13)   TAMRA-AK(Bhc)A-OH (15)

The reaction was performed under exclusion of light. For Mtt-cleavage TAMRA-AK(Mtt)A* **13** (0.005 mmol, 1 eq.) was alternately treated 3 x 1 min with 2 mL of 1% TFA in DCM and 1 x 1 min with 2 mL DCM. The cleavage cocktail turned yellow until no cleaved Mtt was left. Afterwards, the TAMRA-AKA* resin **14** was washed intensively with DCM. 4.89 mg Bhc-Boc-NHS **5b** (0.008 mmol, 1.7 eq.) was dissolved in 1 mL ACN/DCM (1:1), mixed with 5.9 µL TEA (0.042 mmol, 8.4 eq.) and added to the resin. After 12 h shaking at rt, the reaction mixture was removed and the resin washed with DCM. A test cleavage with 100% TFA for 2 h showed the peptide TAMRA-AK(Bhc)A-OH **15** (68% HPLC area percent) with some byproducts.

LaChrom ELITE HPLC (gradient A4, 240 nm): t_R = 15.213 min

ESI(+)-ToF MS: calculated $[C_{51}H_{58}BrN_8O_{13}]^+$ = 1069.3228.; obtained $[M+H]^+$ = 1069.3325

E4.13 Cys'Bhc-NHS in TFA (12b)

The test was performed under exclusion of light. 1 mg Cys'Bhc-NHS **12a** was stirred for 2 h in 0.5 mL 100% TFA. Afterwards, the TFA was removed in nitrogen flow and vacuum. The crude was dissolved in 100 μL ACN. HPLC showed the Cys'Bhc **11** without protection groups or NHS ester (68% HPLC area percent) and byproducts like leaving groups.

LaChrom ELITE HPLC (gradient A4, 240 nm): t_R = 18.960 min

ESI(-)-ToF MS: calculated $[C_{16}H_{19}BrN_3O_5S]^-$ = 445.0307; obtained $[M-H]^-$ = 445.0319

E4.14 TAMRA-AK(Cys'Bhc)A-OH 1 (TAMRA-AK(NHS)A-OH, 16)
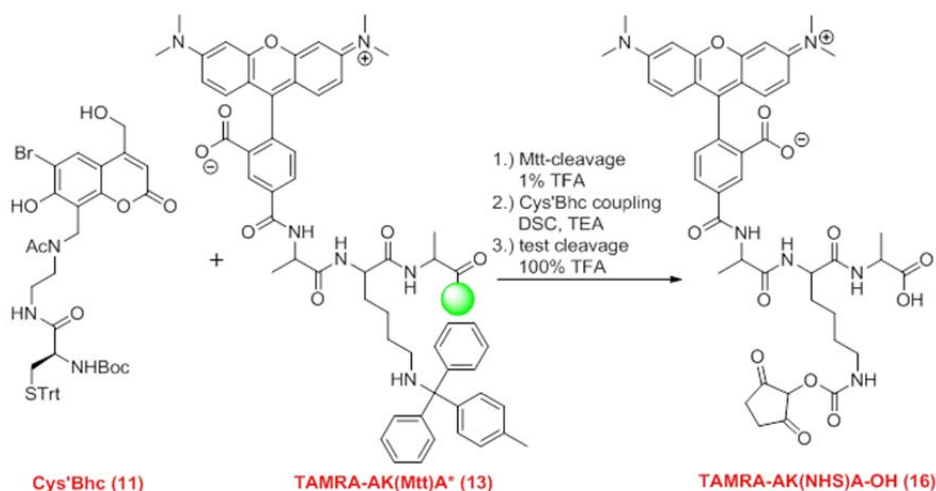
Cys'Bhc (11)   TAMRA-AK(Mtt)A* (13)   TAMRA-AK(NHS)A-OH (16)

The reaction was performed under exclusion of light. For Mtt-cleavage TAMRA-AK(Mtt)A* **13** (0.005 mmol, 1 eq.) was alternately treated 3 x 1 min with 2 mL of 1% TFA in DCM and 1 x 1 min with 2 mL DCM. The cleavage cocktail turned yellow until no cleaved Mtt was left. Afterwards, the TAMRA-AKA* resin **14** was washed intensively with DCM. 5.0 mg Cys'Bhc **11** (0.006 mmol, 1.2 eq.) and 4.8 mg DSC (0.008 mmol, 1.7 eq.) were dissolved in 1 mL ACN/DCM (1:1), mixed with 4.8 μL TEA (0.035 mmol, 7.0 eq.) and added to the resin. After 12 h shaking at rt, the reaction mixture was removed and the resin washed with DCM. A test cleavage with 100% TFA for 2 h showed the NHS-activated peptide TAMRA-AK(NHS)A-OH **16** (100% HPLC area percent) no product or byproducts.

LaChrom ELITE HPLC (gradient A4, 240 nm): t_R = 15.648 min

ESI(+)-ToF MS: calculated $[C_{43}H_{49}N_6O_{12}]^+$ = 844.3330; obtained $[M+H]^+$ = 844.3388

E4.15 Fmoc-AK(Cys'Bhc)A-OH 18

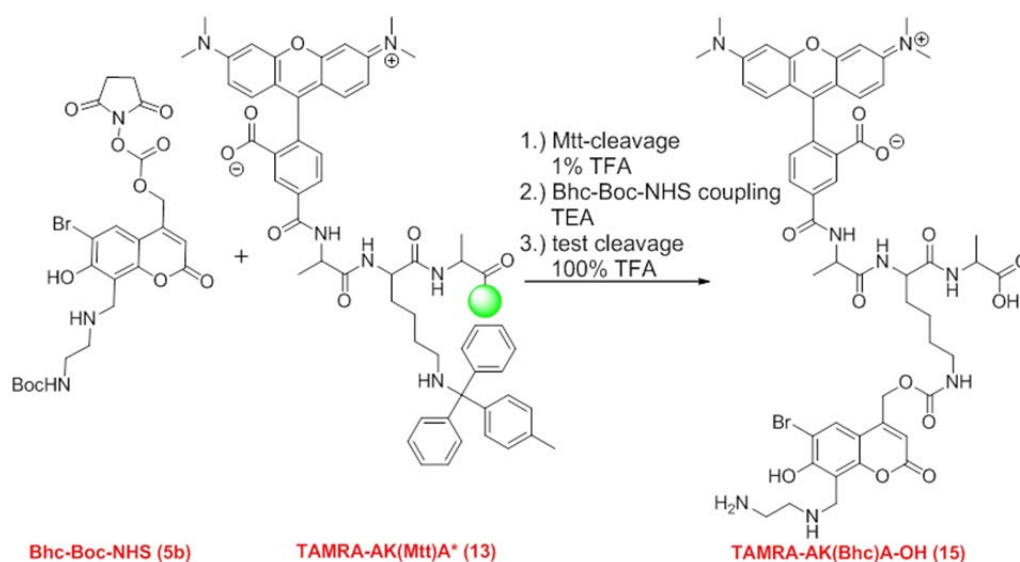Cys'Bhc-NHS (12a)   Fmoc-AK(Mtt)A* (17)   Fmoc-AK(Cys'Bhc)A-OH (18)

The reaction was performed under exclusion of light. For Mtt-cleavage Fmoc-AK(Mtt)A* **17** (0.005 mmol, 1 eq.) was alternately treated 3 x 1 min with 2 mL of 1% TFA in DCM and 1 x 1 min with 2 mL DCM. Cleavage cocktail turned yellow until no cleaved Mtt was left. Afterwards, the Fmoc-AKA* resin was washed intensively with DCM. 5.0 mg Cys'Bhc-NHS **12a** (0.016 mmol, 3.2 eq.) was dissolved in 1 mL ACN/DCM (1:1), mixed with 5.0 µL TEA (0.036 mmol, 7.2 eq.) and added to the resin. After 12 h shaking at rt, the reaction mixture was removed and the resin washed with DCM. A test cleavage with 96% TFA, 2% TIPS and 2% H_2O for 2 h showed the Fmoc-AK(Cys'Bhc)A-OH **18** (81% HPLC area percent) with a few byproducts .

LaChrom ELITE HPLC (gradient A4, 240 nm): t_R = 17.783 min

ESI(+)-ToF MS: calculated $[C_{45}H_{52}BrN_8O_{13}S]^+$ = 1025,2636; obtained $[M+H]^+$ = 1025.2648

E4.16 TAMRA-AK(Cys'Bhc)A-OH 2 (19)

TAMRA-NHS   **Fmoc-AK(Cys'Bhc)A* (18)**   **TAMRA-AK(Cys'Bhc)A-OH (19)**

The reaction was performed under exclusion of light. For Fmoc-removal Fmoc-AK(Cys'Bhc)A* **18** (~0.020 mmol, 1 eq.) was treated 3 x 7 min with 20% piperidine in DMF and afterwards washed with 3 x 2 mL DMF and DCM respectively. 12.7 mg TAMRA-NHS (0.024 mmol, 1.2 eq.) was dissolved in 1 mL DMF and 14 µL DIPEA (0.080 mmol, 4 eq.) was added directly before the mixture was added to the H_2N-AK(Cys'Bhc)A* resin (LaChrom ELITE HPLC, gradient A4, t_R = 112.507 min). After 1 h shaking at rt, the reaction mixture was removed and the resin intensively washed with DMF until the washing solvent remained colorless. A test cleavage with 100% TFA for 2 h yielded the TAMRA-AK(Cys'Bhc)A-OH **19** with 41% (HPLC area percent). The product was not purified.

LaChrom ELITE HPLC (gradient A4, 240 nm): t_R = 15.480 min
ESI(+)-ToF MS: calculated $[C_{56}H_{65}BrN_9O_{15}S]^+$= 1214.3426; $[C_{56}H_{66}BrN_9O_{15}S]^{2+}$= 607.6713 obtained $[M+H]^+$= 1214.3429; $[M+2H]^{2+}$= 607.6714

E4.17 TAMRA-Ahx-VW18K9(Cys'Bhc)-OH 1 (23)

Cys'Bhc-NHS (12a)   TAMRA-Ahx-VW18* (20)   TAMRA-Ahx-VW18K9(Cys'Bhc)-OH (23)

The reaction was performed under exclusion of light. For Mtt-cleavage TAMRA-Ahx-VW18K9* **20** (0.025 mmol, 1 eq.) was alternately treated 3 x 1 min with 2 mL of 1% TFA in DCM and 1 x 1 min with 2 mL DCM. The cleavage cocktail turned yellow until no cleaved Mtt was left. Afterwards, the TAMRA-Ahx-VW18K9* resin was washed intensively with DCM. 34 mg Cys'Bhc-NHS **12a** (0.035 mmol, 1.4 eq.) was dissolved in 1 mL ACN/DCM and 14 µL DIPEA (0.080 mmol, 3.2 eq.) was added directly before the mixture was transferred to the TAMRA-Ahx-VW18* resin (LaChrom ELITE HPLC, gradient A5, t_R = 25.147 min). After 12 h shaking at rt, the reaction mixture was removed and the resin washed with DMF. In HPLC and ESI-ToF of a test cleavage with 100% TFA for 2 h, neither starting reagents nor TAMRA-Ahx-VW18K9(Cys'Bhc)-OH **23** could be observed.

E4.18 Fmoc-Ahx-VW18K9(Cys'Bhc)* (22)

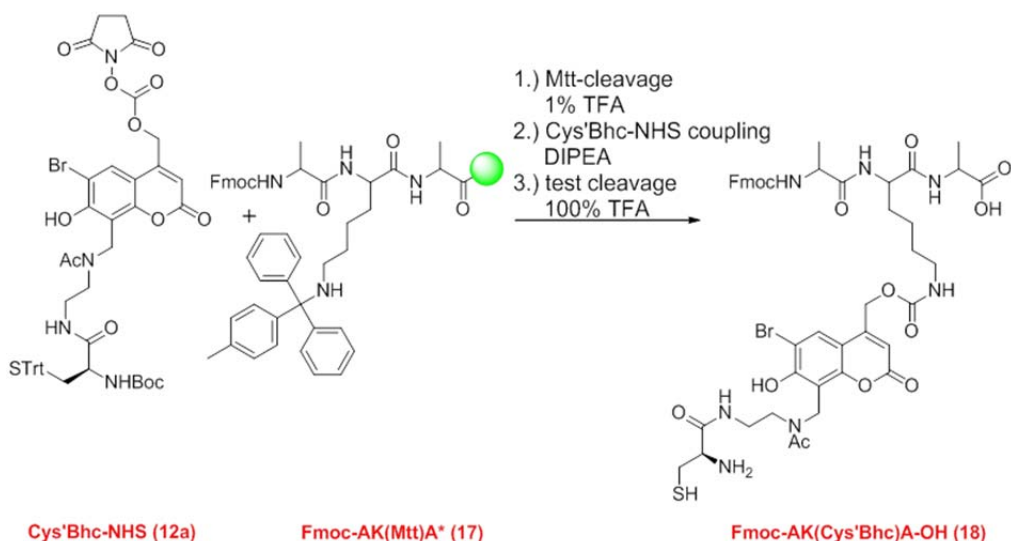Cys'Bhc-NHS (12a) + Fmoc-Ahx-VW18* (20) → Fmoc-Ahx-VW18K9(Cys'Bhc)-OH (22)
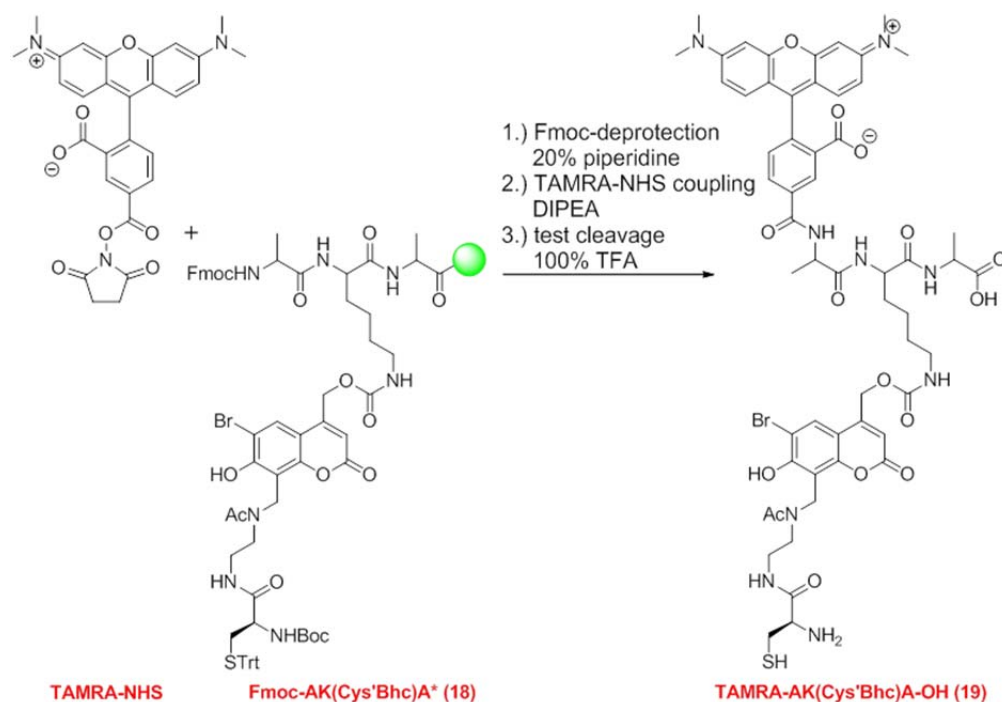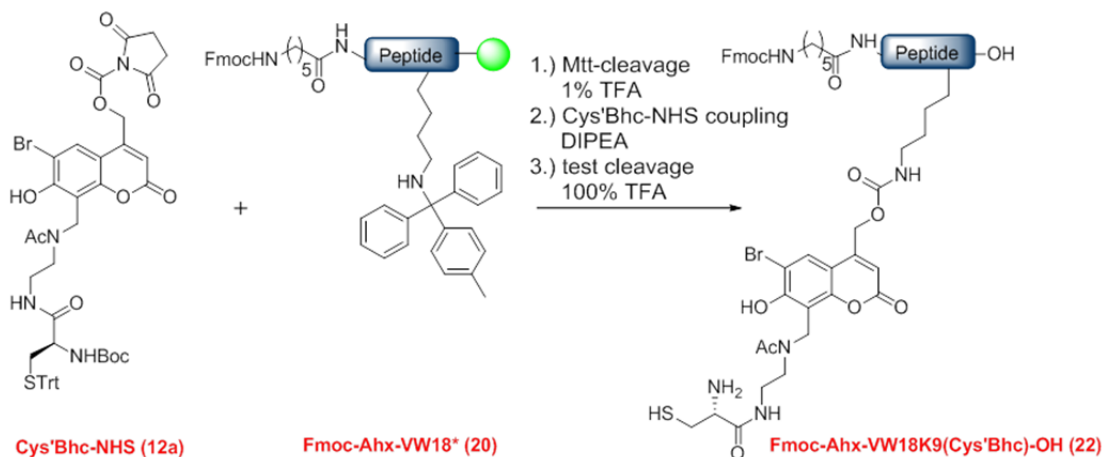
The reaction was performed under exclusion of light. For Mtt-cleavage Fmoc-Ahx-VW18K9* **20** (0.025 mmol, 1 eq.) was alternately treated 3 x 1 min with 2 mL of 1% TFA in DCM and 1 x 1 min with 2 mL DCM. Cleavage cocktail turned yellow until no cleaved Mtt was left. Afterwards, the Fmoc-Ahx-VW18K9* resin (LaChrom ELITE HPLC, gradient A4, t_R = 16.113 min) was washed intensively with DCM. 34 mg Cys'Bhc-NHS **12a** (0.035 mmol, 1.4 eq.) was dissolved in 2 mL ACN/DCM (1:1) and 17 µL DIPEA (0.097 mmol, 3.8 eq.) was added directly before the mixture was transferred to the resin. After 12 h shaking at rt, the reaction mixture was removed and the resin washed with DMF and DCM. A test cleavage with 98% TFA and 2% H_2O for 2 h yielded the Fmoc-Ahx-VW18K9(Cys'Bhc)-OH **22** with 75% (HPLC area percent).

LaChrom ELITE HPLC (gradient A4, 240 nm): t_R= 19.580 min

ESI(+)-ToF MS: calculated [M]= 3872.8018; [M+3H]$^{3+}$= 1291.9417; [M+4H]$^{4+}$= 969.2082; obtained [M+3H]$^{3+}$= 1291.9944; [M+4H]$^{4+}$= 969.2364

E4.19 TAMRA-Ahx-VW18K9(Cys'Bhc)-OH 2 (23)

TAMRA-NHS + **Fmoc-Ahx-VW18* (22)** → **TAMRA-Ahx-VW18K9(Cys'Bhc)-OH (23)**

Reagents: 1.) Fmoc-deprotection 20% piperidine; 2.) TAMRA-NHS coupling DIPEA; 3.) test cleavage 100% TFA
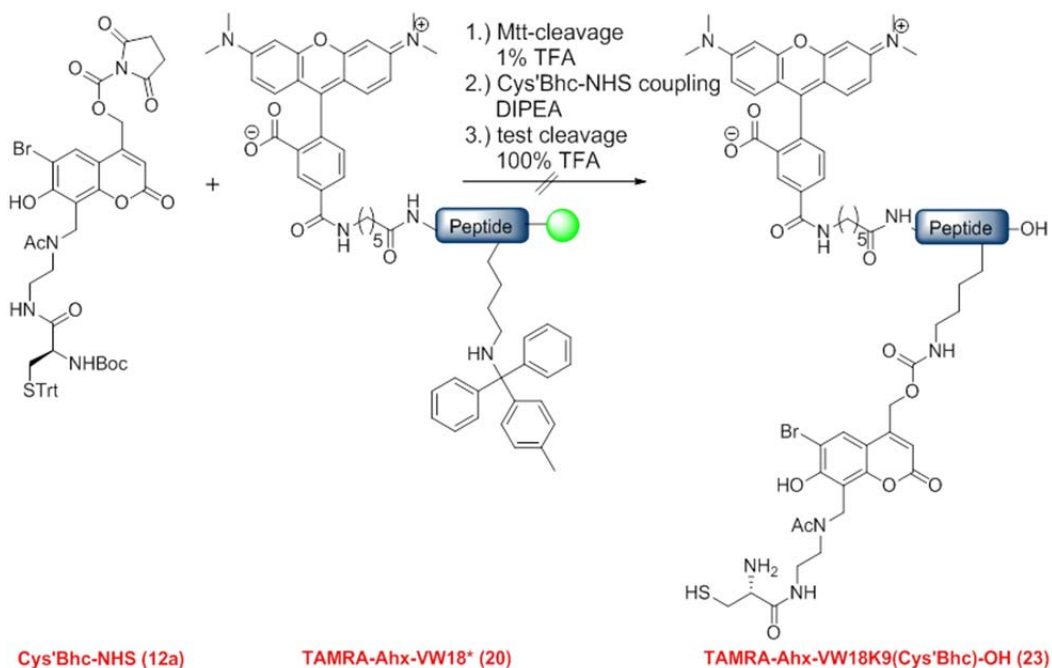
The reaction was performed under exclusion of light. For Fmoc-removal Fmoc-Ahx-VW18-Cys'Bhc* **22** (~0.025 mmol, 1 eq.) was treated 3 x 7 min with 20% piperidine in DMF and afterwards washed with 3 x 2 mL DMF and DCM respectively. 12.7 mg TAMRA-NHS (0.024 mmol, 1.2 eq.) was dissolved in 1 mL DMF and 14 µL DIPEA (0.080 mmol, 4 eq.) was added directly before the mixture was added to the H$_2$N-Ahx-VW18-Cys'Bhc* resin. After 12 h shaking at rt, the reaction mixture was removed and the resin intensively washed with DMF until the washing solvent remained colorless. In several test cleavages with 100% TFA (or other mixtures) for 2 h neither one of the starting compounds nor the product **23** could be observed.

E4.20 Cy5-AK(Cys'Bhc)A-OH (26)

Cys'Bhc-NHS (12a) **Cy5-AK(Mtt)A* (24)** **Cy5-AK(Cys'Bhc)A-OH (26)**
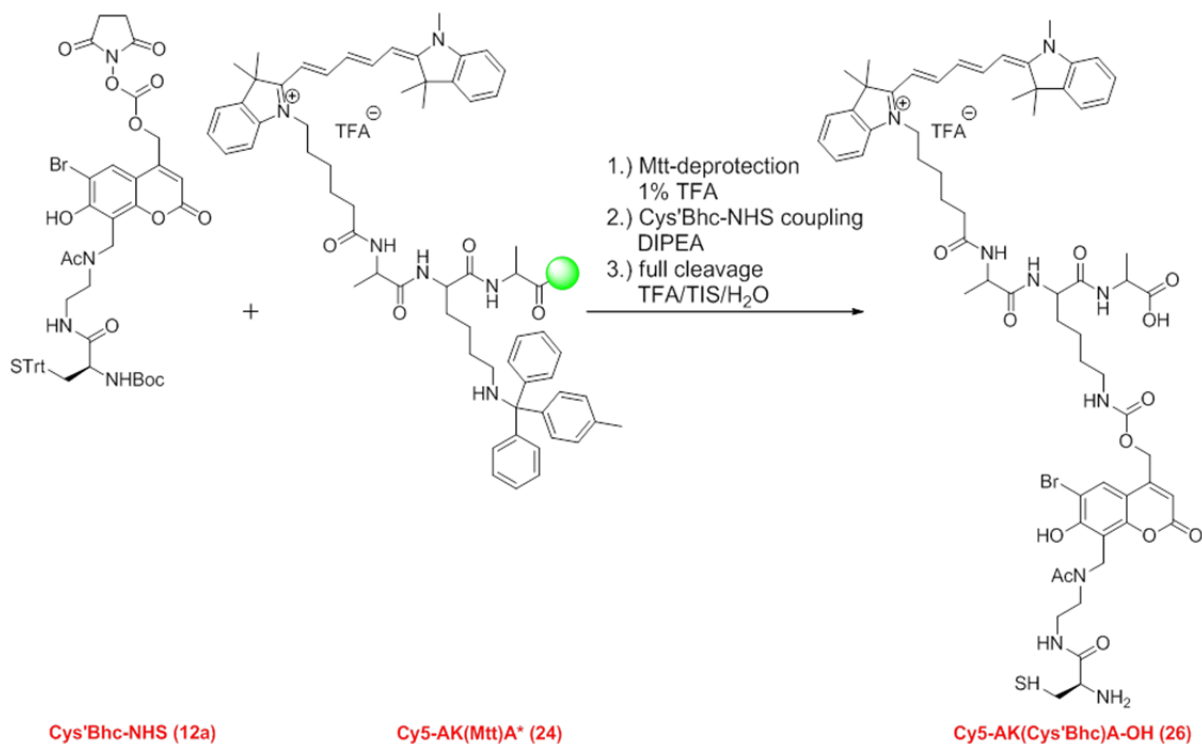
The reaction was performed under exclusion of light. For Mtt-cleavage Cy5-AK(Mtt)A* **24** (0.025 mmol, 1 eq.) was alternately treated 3 x 1 min with 2 mL of 1% TFA in DCM and 1 x 1 min with 2 mL DCM. The cleavage cocktail turned yellow until no cleaved Mtt was left. Afterwards, the Cy5-AKA* resin **25** (Chromaster HPLC, gradient A4, t_R = 16.370 min) was washed intensively with DCM. 29.0 mg Cys'Bhc-NHS **12a** (0.030 mmol, 1.2 eq.) was dissolved in 1 mL DCM/ACN (1:1) and 20 µL DIPEA (0.114 mmol, 4.5 eq.) was added directly before the mixture was added to the H_2N-Ahx-VW18(Cys'Bhc)* resin. After 12 h shaking at rt, the reaction mixture was removed and the resin intensively washed with DMF until the washing solvent remained colorless. For full cleavage the resin was treated with 1.8 mL TFA, 20 µL TIS and 20 µL H_2O for 1 h. Afterwards the reaction mixture was subjected to an amber colored flask and the solvents were removed under reduced pressure before adding 5 mL water and freeze-dry the crude peptide. Purification was achieved with preparative HPLC, gradient P1 and detection at 220 nm. 3.7 mg of Cy5-AK(Cys'Bhc)A-OH **26** (3 µmol, 12%) was obtained with >99% purity (HPLC area percent) and further 15 mg **26** with 90% purity (HPLC area percent).

LaChrom ELITE HPLC (gradient A4, 240 nm): t_R = 17.500 min

ESI(-)-ToF MS: calculated $[C_{63}H_{80}BrN_9O_{12}S]^-$ = 1266.4903; obtained $[M-H]^-$ = 1266.4994

E4.21 Cy5-Ahx-VW18K9(Cys'Bhc)-OH 1 (28)

Cys'Bhc-NHS (12a)   **Cy5-Ahx-VW18K9* (27)**   **Cy5-Ahx-VW18K9(Cys'Bhc)-OH (28)**
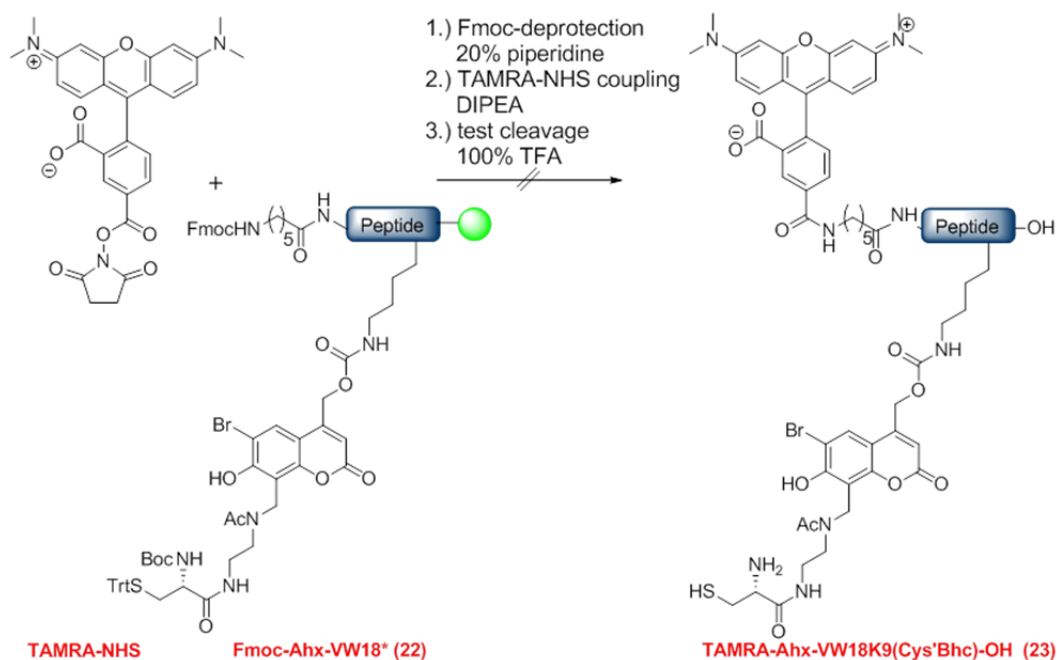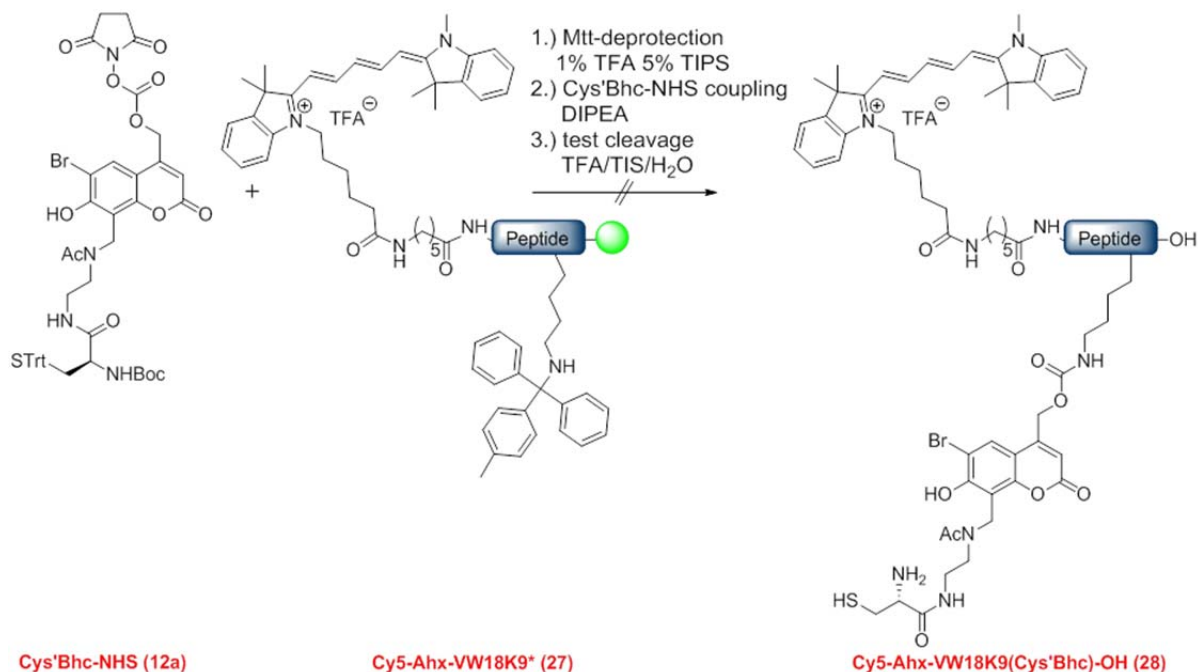
The reaction was performed under exclusion of light. For Mtt-cleavage Cy5-Ahx-VW18K9* **27** (0.0125 mmol, 1 eq.) was alternately treated 3 x 1 min with 2 mL of 1% TFA and 5% TIPS in DCM and 1 x 1 min with 2 mL DCM. The cleavage cocktail turned green until no cleaved Mtt was left. Afterwards, the Cy5-Ahx-VW18K9* resin was washed intensively with DCM. The cleavage solution was examined with HPLC and the chromatogram did contain a high concentration of Cy5-Ahx-VW18K9-OH peptide.

15 mg Cys'Bhc-NHS **12a** (0.015 mmol, 1.2 eq.) was dissolved in 1 mL ACN/DCM and 14 µL DIPEA (0.015 mmol, 1.2 eq.) was added directly before the mixture was transferred to the Cy5-Ahx-VW18K9* resin. After 12 h shaking at rt, the reaction mixture was removed and the resin washed with DMF and DCM. In HPLC and ESI-ToF of a test cleavage with 200 µL TFA and 20 µL TIPS for 2 h, neither the starting reagents nor product **28** could be observed.

E4.22 Cy5-Ahx-VW18K9(Cys'Bhc)-OH 2 (28)
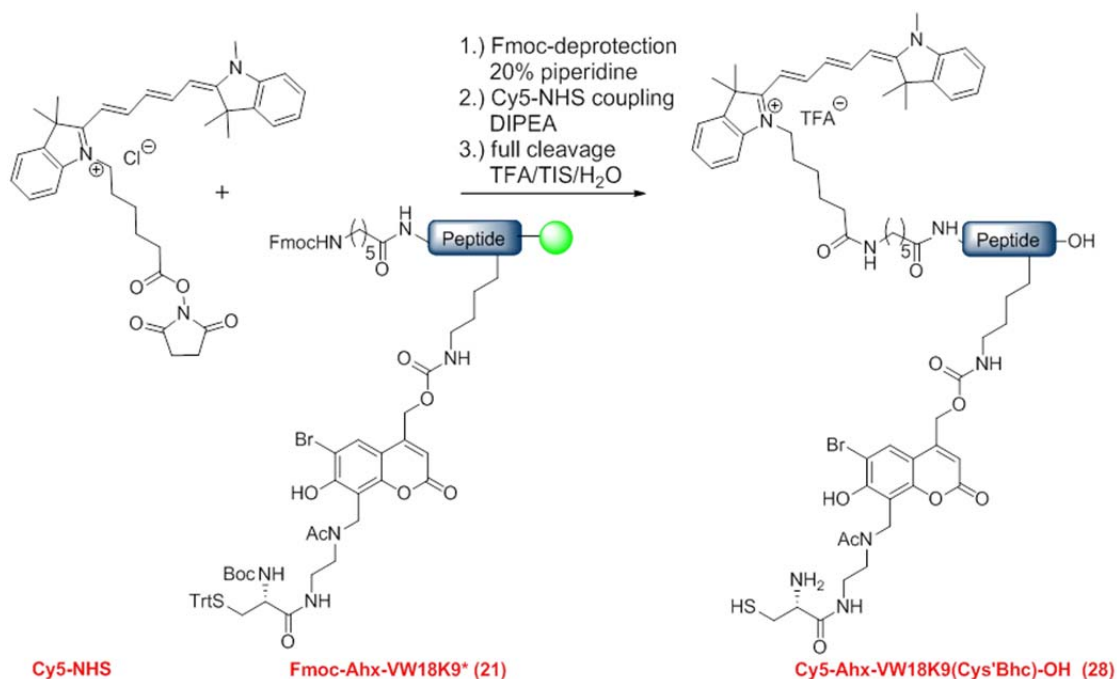
The reaction was performed under exclusion of light. For Fmoc-removal Fmoc-Ahx-VW18K9(Cys'Bhc)* **21** (~0.025 mmol, 1 eq.) was treated 3 x 7 min with 2% piperidine and 2% DBU in DMF and afterwards washed with 3 x 2 mL DMF and DCM respectively. 10.0 mg Cy5-NHS (0.030 mmol, 1.2 eq.) was dissolved in 1 mL DMF and 20 µL DIPEA (0.114 mmol, 4.5 eq.) were added directly before the mixture was added to the H_2N-Ahx-VW18K9(Cys'Bhc)* resin. After 12 h shaking at rt, the reaction mixture was removed and the resin intensively washed with DMF until the washing solvent remained colorless. For full cleavage the resin was treated with 1 mL TFA, 20 µL TIS and 20 µL H_2O for 3 h. Afterwards, the reaction mixture was subjected to an amber colored flask and the solvents were removed under reduced pressure before adding 5 mL water and freeze-drying the crude peptide. Purification was achieved with analytical HPLC, gradient A4 and detection at 240 nm. 1 mg of Cy5-Ahx-VW18K9(Cys'Bhc)-OH **28** (0.2 µmol, 1%) was obtained with >99% purity (HPLC area percent).

LaChrom ELITE HPLC (gradient A4, 240 nm): t_R = 19.927 min
ESI(+)-ToF MS: calculated [M] = 4116.2118; $[M+3H]^{3+}$ = 1373.0784; $[M+4H]^{4+}$ = 1030.0607; $[M+5H]^{5+}$ = 824.2501; obtained $[M+3H]^{3+}$ = 1373.0790; $[M+4H]^{4+}$ = 1030.0576; $[M+5H]^{5+}$ = 824.2474

E4.23 Cy5-AK(Cys'Bhc)-OH (26) in Solution

Cys'Bhc-NHS (12a)   Cy5-AKA-OH (25)   Cy5-AK(Cys'Bhc)A-OH (26)
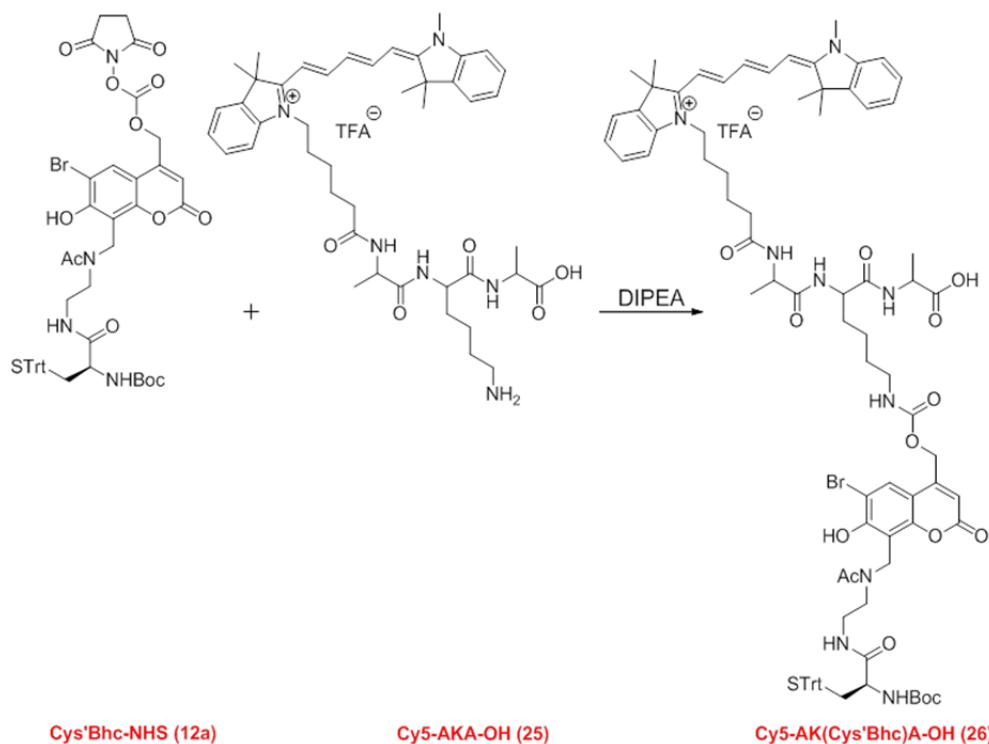
The reaction was performed under exclusion of light and inert argon conditions. 2.3 mg Cy5-AKA-OH **25** (3.07 µmol, 1 eq.) and 4.9 mg Cys'Bhc-NHS **12a** (5.04 µmol, 1.4 eq.) were dissolved in 1 mL ACN and 20 µL DIPEA (0.114 mmol, 38 eq.) was added directly afterwards. After 12 h stirring at rt, the reaction progress was monitored by analytical HPLC by transferring 5 µL reaction solution into 50 µL ACN. With gradient A6 on Chromaster HPLC (312 nm) only few product could be observed. The reaction mixture was further heated to 50°C for 4 d with frequent reaction control on HPLC to achieve 50% **26/25** ratio (HPLC area percent) and 0% of Cys'Bhc-NHS **12a**. The product was not isolated.

Chromaster HPLC (gradient A6, 312 nm): t_R = 22.917 min

E4.24 Concentration curves with Cyanine 5 and Cy5-AKA-OH

For later concentration determination with Cy5 two curves with different methods were realized. Therefore, three stock solutions with 200 µM of Cy5 labeled peptide were prepared. Cy5-Ala-OH was dissolved in ACN/H_2O (2:8) and 0.1% formic acid was added. All three stock solutions were diluted to 100 µM, 50 µM, 25 µM, 10 µM and 5 µM concentrations. With each concentration UV/Vis measurements in a 1 mm QS cuvette and Semi-micro Chromaster HPLC using gradient A7 were performed. The UV absorption at 600 nm and the absolute area of the HPLC peak t_R= 6.683 min at 600 nm wavelength were plotted as an average of three measurments against the concentration of Cy5-Ala-OH to obtain two concentration curves.
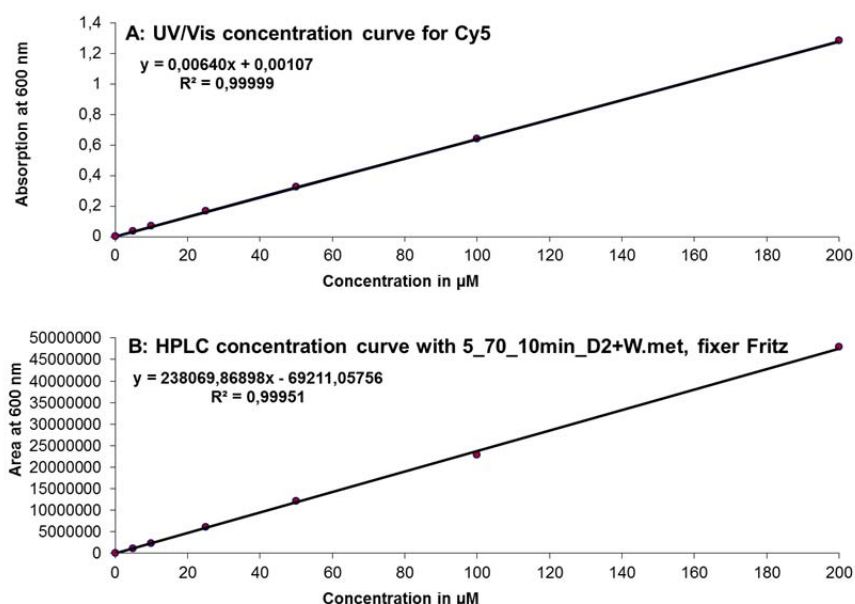
Figure E1: *Concentration curves of Cy5-Ala-OH in ACN/H_2O (2:8) + 0.1% formic acid. The curve contains the average of three measurements per concentration.*

E5 Project D

E5.1 Peptide Synthesis

Peptides were synthesized manually on a preloaded Fmoc-Lys(Boc)-Wang resin (0.57 mmol/g loading), using standard couplings for common amino acids. HfLeu containing peptides were synthesized with an Activo-P11 Automated Peptide Synthesizer (Activotec, Cambridge, United Kingdom). Only for fluorinated amino acids the coupling reagents 1-Hydroxy-7-azabenzotriazole [HOAt]/DIC were used in 1.2-fold excess, and the coupling was carried out manually one time over night. In case of an insufficient coupling, the coupling was repeated for 3 h with 0.5 equivalents. Prior to the Fmoc deprotection of the fluorinated amino acids, free N-termini were capped by adding a mixture of acetic anhydride [Ac_2O] (10% (v/v)) and DIPEA (10% (v/v)) in DMF (3x10 min). All peptides were N-terminally labeled with Abz to enable photometric detection. Purification of the synthesized peptides was performed with preparative HPLC and gradient P1. The fractions containing pure peptide were combined, reduced *in vacuo* and lyophilized to give the peptides as a white powder. The purity of the peptides was controlled by analytical HPLC (LUNA^{TM} C8 (2) column, 5 μm, 250 x 4.6 mm, Phenomenex®, Torrance, CA, USA) and gradient A4. The products were identified by high resolution ESI-ToF.

Table E7: *Observed and calculated mass to charge ratios and retention times on analytical HPLC of the peptides.*

peptide	charge	calc. mass [m/z]	obs. mass [m/z]	t_R [min]
FA	1	967.5364	967.5396	10.597
	2	484.2721	484.2736	
P2-LeuFA	1	1009.5463	1009.5849	12.500
	2	505.2956	505.2970	
P2-HfLeuFA	1	1117.4622	1117.5306	12.393
	2	559.2573	559.2691	
P1'-LeuFA	1	1009.5463	1009.5863	11.773
	2	505.2956	505.7982	
P1'-HfLeuFA	1	1117.4622	1117.5272	11.917
	2	559.2573	559.2684	
	2	532.2814	532.2816	
P2'-LeuFA	1	1009.5463	1009.5866	11.847
	2	505.2956	505.2981	
P2'-HfLeuFA	1	1117.4622	1117.5305	12.197
	2	559.2573	559.2693	

E5.2 Enzyme Digestion Assay

1 mg/mL stock solutions of α-chymotrypsin (from bovine pancreas, EC 3.4.21.1, ≥ 40.0 units/mg of protein, Sigma Aldrich, St. Louis, MO, USA), and pepsin (from porcine stomach mucosa, EC 3.4.23.1, ≥ 250 units/mg of protein, Sigma Aldrich, St. Louis, MO, USA) were prepared in 10 mM phosphate buffer pH 7.4, or in 10 mM acetate buffer pH 4.0. For proteinase K (from tritirachium album, EC 3.4.21.64, ≥ 30 units/mg of protein, Sigma Aldrich, St. Louis, MO, USA) and elastase (from porcine pancreas, EC 3.4.21.36, 6.2 units/mg of protein, Sigma Aldrich, St. Louis, MO, USA) 1 mg/mL stock solutions were prepared in 50 mM tris/HCl buffer including 10 mM CaCl$_2$ at pH 7.5, or in 100 mM tris/HCl buffer pH 8.4, respectively. 0.002 mmol peptide stock solutions in 100 μL DMSO (100 μL) were prepared and incubated with the respective enzyme at 30°C (for α-chymotrypsin and pepsin) or 37°C (for proteinase K and elastase) with shaking at 300 rpm in a thermomixer over a period of 24 h. The reaction mixture consisted of 15 μL DMSO, 25 μL corresponding buffer, 5 μL peptide solution and 5 μL of the corresponding enzyme solution. The concentration of enzyme was optimized so that the hydrolysis of the control peptide FA was about 40% after 120 min. At fixed time points, aliquots of 5 μL were

taken (0, 15, 30, 60, 90, 120 min as well as 3 h and 24 h) and either quenched with 95 µL ACN containing 0.1% (v/v) TFA, in the case of α-chymotrypsin, proteinase K and elastase, or 95 µL of 2% aqueous ammonia, in the case of pepsin. All samples were subjected to analytical HPLC on a LaChrom-ELITE-HPLC-System equipped with a fluorescence detector. A monolithic reversed-phase C8 Chromolith® Performance HPLC column (100 x 4.6 mm, Merck, Darmstadt, Germany) was used to identify the products of digestion. For the non-fluorinated peptides gradient A7 was used to follow the digestion process, and for the fluorinated peptides gradient A8 was applied. For chromatograms where an insufficient baseline separation was observed, the measurements were repeated using either gradient A9 for (FA/ pepsin), (P2-LeuFA/ proteinase K), (P1'-LeuFA/ elastase) and (P1'-LeuFA/ proteinase K), or gradient A10 for the experiment with (P2-HfLeuFA/ proteinase K). Fluorescence detection with $\lambda_{ex} = 320$ nm and $\lambda_{em} = 420$ nm was based on the Abz label. In all cases, the peaks corresponding to the starting materials or the N-terminal fragments were integrated and used to determine the velocity of the degradation. The FA peptide was used as a reference. All experiments were performed in triplicate. Each fragment cleaved from the full-length peptide was identified by ESI–ToF mass experiments according to the mass-to-charge ratios determined with an Agilent 6220 ESI-ToF MS instrument (Agilent Technologies, Santa Clara, CA, USA). For this, the quenched peptide-enzyme-solutions after 120 min and 24 h incubation were analyzed. The solutions were injected directly into the spray chamber using a syringe pump with a flow rate of 10 µL min^{-1}. Spray voltage was set to 3.5 kV, a drying gas flow rate of 5 L min^{-1} was used, the nebulizer was set to 30 psi, and the gas temperature to 300°C. The fragmentor voltage was 200 V. Not all corresponding fragments could be detected.

Appendix

Table S1: *Identification of the cleavage products of the different peptides by ESI-ToF mass spectrometry after digestion with α-chymotrypsin.*

peptide	fragment	calc. [M + H]$^{1+}$	opt. [M + H]$^{1+}$
FA	Abz-KAAFAAAAK	967.5364	967.5376
	Abz-KAAF	555.2559	555.2938
	AAAAK	431.2617	431.2627
P2-LeuFA	Abz-KALeuFAAAAK	1009.5463	1009.5883
	Abz-KALeuF	597.3029	597.2609
	AAAAK	431.2617	431.2617
P2-HfLeuFA	Abz-KAHfLeuFAAAAK	1117.4622	1117.5298
P1'-LeuFA	Abz-KAAFLeuAAAK	1009.5463	1009.5866
	Abz-KAAFLeu	668.3400	668.3760
	Abz-KAAF	555.2559	555.2907
	LeuAAAK	473.3087	473.3087
	AAAK	360.2246	360.2239
P1'-HfLeuFA	Abz-KAAFHfLeuAAAK	1117.4622	1117.5280
	Abz-KAAFHfLeu	776.2559	776.3214
	HfLeuAAAK	581.2246	581.2246
	Abz-KAAF	555.2559	555.2934
	AAAK	360.2246	360.3630
P2'-LeuFA	Abz-KAAFALeuAAK	1009.5463	1009.5872
	Abz-KAAF	555.2559	555.2922
	ALeuAAK	473.3087	473.3087
P2'-HfLeuFA	Abz-KAAFAHfLeuAAK	1117.4622	1117.5331
	AHfLeuAAK	581.2246	581.2550
	Abz-KAAF	555.2559	555.2965

Table S2: *Cleavage products of the different peptides by ESI-ToF mass spectrometry after digestion with elastase.*

peptide	fragment	calc. [M + H]$^{1+}$	opt. [M + H]$^{1+}$
FA	Abz-KAAFAAAAK	967.5364	967.5352
	Abz-KAAFAAA	768.3672	768.4080
	Abz-KAAFAA	697.3301	697.3690
	AFAAAAK	649.3673	648.2935
	Abz-KAAFA	626.2930	626.3308
	AAAK	360.2246	360.2238
	AAK	289.1875	289.1872
P2-LeuFA	Abz-KALeuFAAAAK	1009.5463	1009.5871
	Abz-KALeuFAAAA	881.4513	881.4831
	Abz-KALeuFAAA	810.4142	810.4831
	Abz-KALeuFAA	739.4116	739.4116
	LeuFAAAAK	691.4142	691.4116
	Abz-KALeuFA	668.3404	668.3745
	AAAAK	431.2617	430.0496
	AAAK	360.2246	360.2205
	AAK	289.1875	289.1858
P2-HfLeuFA	Abz-KAHfLeuFAAAAK	1117.4622	1117.5325
	Abz-KAHfLeuFAAA	918.3301	918.3980
	Abz-KAHfLeuFAA	847.2930	847.3610
	Abz-KAHfLeuFA	776.2559	776.2559
	FAAAAK	578.3301	578.2440
P1'-LeuFA	Abz-KAAFLeuAAAK	1009.5463	1009.5887
	Abz-KAAFLeuAAA	881.4513	881.4815
	Abz-KAAFLeuAA	810.4142	810.4472
	AAFLeuAAAK	762.4513	762.4458
	Abz-KAAFLeuA	739.3771	739.4120
	AFLeuAAAK	691.4142	691.4142
	FLeuAAAK	620.3771	620.3783
	Abz-KAA	408.1875	408.1875
	Abz-KA	337.1504	337.1504
	AAK	289.1504	289.1858
P1'-HfLeuFA	Abz-KAAFHfLeuAAAK	1117.4622	1117.5330
	Abz-KAAFHfLeuAA	918.3301	918.3982
	AAFHfLeuAAAK	870.3673	870.3975
	AFHfLeuAAAK	799.3301	799.3600
	Abz-KA	337.1504	337.1864

peptide	fragment	calc. [M + H]$^{1+}$	opt. [M + H]$^{1+}$
P2'-LeuFA	Abz-KAAFALeuAAK	1009.5463	1009.5894
	Abz-KAAFALeuAA	881.4513	881.4938
	Abz-KAAFALeuA	810.4142	810.4938
	AFALeuAAK	691.4142	691.4168
	Abz-KAAFA	626.2930	626.3324
	LeuAAK	402.2716	402.2730
	Abz-KA	337.1504	337.1872
P2'-HfLeuFA	Abz-KAAFAHfLeuAAK	1117.6622	1117.5330
	Abz-KAAFAHfLeuA	918.3301	918.4349
	AAFAHfLeuAAK	870.3673	870.4001
	AFAHfLeuAAK	799.3301	799.3612
	FAHfLeuAAK	728.2930	728.3229
	Abz-KAAFA	626.2930	626.3230
	Abz-KAAFA	408.1875	408.2262

Table S3: Identification of the cleavage products of the different peptides by ESI-ToF mass spectrometry after digestion with pepsin.

peptide	fragment	calc. [M + H]$^{1+}$	opt. [M + H]$^{1+}$
FA	Abz-KAAFAAAAK	965.3640	967.5434
	Abz-KAAF	555.2559	555.2967
	AAAAK	431.2617	431.2617
P2-LeuFA	Abz-KALeuFAAAAK	1009.5463	1009.5895
	Abz-KALeuF	597.3029	597.3438
	AAAAK	431.2617	431.2642
P2-HfLeuFA	Abz-KAHfLeuFAAAAK	1117.4622	1117.5302
	Abz-KAHfLeuF	705.2188	705.2870
	FAAAAK	578.3301	578.3327
	AAAAK	431.2617	431.2636
P1'-LeuFA	Abz-KAAFLeuAAAK	1009.5463	1009.5926
	Abz-KAAFLeu	668.3400	668.3820
	Abz-KAAF	555.2559	555.2971
	LeuAAAK	473.3087	473.3126
	AAAK	360.2246	360.2271
P1'-HfLeuFA	Abz-KAAFHfLeuAAAK	1117.4622	1117.5325
	Abz-KAAFHfLeu	776.2559	776.3236
	HfLeuAAAK	581.2246	581.2553
	Abz-KAAF	555.2559	555.2956
	AAAK	360.2246	360.2273
P2'-LeuFA	Abz-KAAFALeuAAK	1009.5463	1009.5905
	Abz-KAAF	555.2559	555.2963
	ALeuAAK	473.3087	473.3117
P2'-HfLeuFA	Abz-KAAFAHfLeuAAK	1117.4622	1117.5307
	Abz-KAAFA	626.2930	626.3344
	HfLeuAAK	510.1875	510.2170

Table S4: *Identification of the cleavage products of the different peptides by ESI-ToF mass spectrometry after digestion with proteinase K.*

peptide	fragment	calc. [M + H]$^{1+}$	opt. [M + H]$^{1+}$
FA	Abz-KAAFAAAAK	967.5364	967.5376
	AFAAAAK	649.3673	649.2762
P2-LeuFA	Abz-KALeuFAAAAK	1009.5463	1009.5863
	Abz-KALeuFAAA	810.4142	810.4536
	Abz-KALeuFAA	739.4116	739.4179
	Abz-KALeuFA	668.3404	668.3816
	Abz-KALeu	450.2345	450.2719
	AAK	289.1504	289.1880
P2-HfLeuFA	Abz-KAHfLeuFAAAAK	1117.4622	1117.5298
	Abz-KAHfLeuFAAA	918.3301	918.3970
	Abz-KAHfLeuFAA	847.2930	847.3608
	Abz-KAHfLeuFA	776.2559	776.3233
	Abz-KAHfLeu	558.1504	558.2168
	AAK	289.1875	289.1884
P1'-LeuFA	Abz-KAAFLeuAAAK	1009.5463	1009.5877
	Abz-KAAFLeuAA	810.4142	810.4552
	Abz-KAAFLeuA	739.3771	739.4182
	FLeuAAAK	620.3771	620.3813
	Abz-KAA	408.1875	408.2276
	AAK	289.1504	289.1885
P1'-HfLeuFA	Abz-KAAFHfLeuAAAK	1117.4622	1117.5271
	FHfLeuAAAK	728.2930	728.3226
	Abz-KAA	408.1875	408.2261
P2'-LeuFA	Abz-KAAFALeuAAK	1009.5463	1009.5845
	Abz-KAAFALeu	739.3771	739.4166
	FALeuAAK	620.3771	620.3782
	Abz-KAA	408.1875	408.2252
	AAK	289.1875	289.1883
P2'-HfLeuFA	Abz-KAAFAHfLeuAAK	1117.4622	1117.5304
	Abz-KAAFAHfLeu	847.2930	847.3603
	HfLeuAAK	510.1875	510.2172
	AAK	289.1875	268.1885

References

1. *Eur. J. Biochem.* **1984**, *138* (1), 9-37.
2. M. J. Lee; M. B. Yaffe, *Cold Spring Harb. Perspect. Biol.* **2016**, *8* (6), 156-163.
3. L. Birnbaumer, *Annu. Rev. Pharmacol. Toxicol.* **1990**, *30*, 675-705.
4. W. A. Cramer, *Faseb J.* **1994**, *8* (2), 155-163.
5. T. Y. Tsong, *Annu. Rev. Biophys. Biomo.l Struct.* **1990**, *19*, 83-106.
6. M. Blank, *J. Electrochem. Soc.* **1979**, *126* (3), 143-156.
7. D. D. Fanestil; R. J. Kessler; C. S. Park, *Cur. Top. Membr. Transp.* **1984**, *20*, 259-270.
8. J. K. Lanyi, *Abstr. Pap. Am. Chem. Soc.* **1972**, *164* (1), 218-225.
9. M. J. Nadolski; M. E. Linder, *Febs J.* **2007**, *274* (20), 5202-5210.
10. A. Kassenaar; G. A. Smart; S. Brody; R. Luft; R. Schwyzer; W. R. Trotter; A. Prader, *Acta Endocrinol.* **1963**, *42*, 13-26.
11. G. Rosselin; D. Bataille; M. Laburthe; S. Durangarcia, *Pathol. Biol.* **1975**, *23* (10), 793-803.
12. F. M. P. Tonelli; S. Lacerda; F. C. P. Tonelli; G. M. J. Costa; L. R. de Franca; R. R. Resende, *Biotechnol. Adv.* **2017**, *35* (6), 832-844.
13. D. G. Smyth, *Annu. Rep. Prog. Chem.* **1963**, *60*, 468-479.
14. T. Wieland; H. Determann, *Annu. Rev. Biochem.* **1966**, *35*, 651-666.
15. A. Brik, *Bioorg. Med. Chem.* **2013**, *21* (12), 3398-3399.
16. J. H. Collier; T. Segura, *Biomaterials* **2011**, *32* (18), 4198-4204.
17. A. Shrivastava; A. D. Nunn; M. F. Tweedle, *Curr. Pharm. Design* **2009**, *15* (6), 675-681.
18. U. I. M. Gerling; M. Salwiczek; C. D. Cadicamo; H. Erdbrink; C. Czekelius; S. L. Grage; P. Wadhwani; A. S. Ulrich; M. Behrends; G. Haufe; B. Koksch, *Chem. Sci.* **2014**, *5* (2), 819-830.
19. V. Asante; J. Mortier; H. Schlüter; B. Koksch, *Bioorg. Med. Chem.* **2013**, *21* (12), 3542-3546.
20. R. Saladino; G. Botta; M. Crucianelli, *Mini-Rev. Med. Chem.* **2012**, *12* (4), 277-300.
21. H. Ishida; Y. Inoue, *Rev. Heteroatom Chem.* **1999**, *19*, 79-142.
22. S. Borman, *Chem. Eng. News* **2015**, *93* (14), 9-18.
23. Z. Szekeres; A. Isaak; J. Prechl; A. Erdei, *Febs J.* **2005**, *272*, 529-529.
24. R. Gambari, *Curr. Pharm. Design* **2001**, *7* (17), 1839-1862.
25. C. L. Orthner, *Chem. Eng. News* **2006**, *84* (43), 8-16.
26. K. P. Devi; B. Shanmuganathan; A. Manayi; S. F. Nabavi; S. M. Nabavi, *Mol. Neurobiol.* **2017**, *54* (9), 7028-7041.
27. J. D. Sadowsky; T. H. Pillow; J. H. Chen; F. Fan; C. R. He; Y. L. Wang; G. Yan; H. Yao; Z. J. Xu; S. Martin; D. L. Zhang; P. Chu; J. dela Cruz-Chuh; A. O'Donohue; G. M. Li; G. Del Rosario; J. T. He; L. N. Liu; C. Ng; D. A. Su; G. D. L. Phillips; K. R. Kozak; S. F. Yu; K. Y. Xu; D. Leipold; J. Wai, *Bioconjugate Chem.* **2017**, *28* (8), 2086-2098.
28. D. B. Guo; S. T. Xu; N. Wang; H. Y. Jiang; Y. Huang; X. Jin; B. Xue; C. Zhang; X. Y. Zhu, *Biomaterials* **2017**, *144*, 188-198.
29. G. Verdine, *Biopolymers* **2009**, *92* (4), 296-296.
30. A. Gupta; A. Mishra; N. Puri, *J. Biotechnol.* **2017**, *259*, 148-159.
31. A. Holt; J. A. Killian, *Eur. Biophys. J. Biophys. Lett.* **2010**, *39* (4), 609-621.

32. M. J. O. Anteunis, *Bull. Soc. Chim. Belg.* **1978**, *87* (8), 627-650.

33. G. Xing; V. J. DeRose, *Cur. Opin. Chem. Biol.* **2001**, *5* (2), 196-200.

34. J. L. Heier; D. J. Mikolajczak; C. Bottcher; B. Koksch, *Biopolymers* **2017**, *108* (1), 1-7.

35. K. R. Lee; E. S. Kang; Y. T. Kim; N. H. Kim; D. Youn; Y. D. Kim; J. Lee; Y. H. Kim, *Catal. Today* **2017**, *295*, 95-101.

36. V. Asante; J. Mortier; G. Wolber; B. Koksch, *Amino Acids* **2014**, *46* (12), 2733-2744.

37. M. L. Lian; X. Chen; X. J. Liu; Z. C. Yi; W. S. Yang, *Sens. Actuator B-Chem.* **2017**, *251*, 86-92.

38. Y. F. Wang; W. Qi; R. Z. Xing; Q. G. Xing; R. X. Su; Z. M. He, *Adv. Mater. Interfaces* **2017**, *4* (19), 1-23.

39. R. F. Rengifo; N. X. Li; A. Sementilli; D. G. Lynn, *Org. Biomol. Chem.* **2017**, *15* (34), 7063-7071.

40. Jason M. Kalapothakis; Ryan J. Morris; J. Szavits-Nossan; K. Eden; S. Covill; S. Tabor; J. Gillam; Perdita E. Barran; Rosalind J. Allen; Cait E. MacPhee, *Biophys. J.* **2015**, *108* (9), 2300-2311.

41. K. Tenidis; M. Waldner; J. Bernhagen; W. Fischle; M. Bergmann; M. Weber; M.-L. Merkle; W. Voelter; H. Brunner; A. Kapurniotu, *J. Mol. Biol.* **2000**, *295* (4), 1055-1071.

42. W. Ye; Y. Chen; W. Wang; Q. Yu; Y. Li; J. Zhang; H.-F. Chen, *PLoS ONE* **2012**, *7* (5), 36382-36395.

43. D. Matthes; V. Daebel; K. Meyenberg; D. Riedel; G. Heim; U. Diederichsen; A. Lange; B. L. de Groot, *J. Mol. Biol.* **2014**, *426* (2), 362-376.

44. J. Seo; W. Hoffmann; S. Warnke; X. Huang; S. Gewinner; W. Schöllkopf; M. T. Bowers; G. von Helden; K. Pagel, *Nat. Chem.* **2017**, *9* (1), 39-44.

45. T. Takahashi; H. Mihara, *Acc. Chem. Res.* **2008**, *41* (10), 1309-1318.

46. V. Castelletto; P. Ryumin; R. Cramer; I. W. Hamley; M. Taylor; D. Allsop; M. Reza; J. Ruokolainen; T. Arnold; D. Hermida-Merino; C. I. Garcia; M. C. Leal; E. Castaño, *Sci. Rep.* **2017**, *7*, 43637-43649.

47. R. S. Signarvic; W. F. DeGrado, *J. Mol. Biol.* **2003**, *334* (1), 1-12.

48. H. Gradisar; R. Jerala, *J. Pep. Sci.* **2011**, *17* (2), 100-106.

49. A. L. Boyle; D. N. Woolfson, *Chem. Soc. Rev.* **2011**, *40* (8), 4295-4306.

50. G. Xing; V. J. DeRose, *Cur. Opin. Chem. Biol.* **2001**, *5* (2), 196-200.

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

52. UniProtDatabase, *Nucleic Acids Res.* **2017**, *45* (D1), 158-169.

53. A.-M. Fernandez-Escamilla; F. Rousseau; J. Schymkowitz; L. Serrano, *Nat. Biotech.* **2004**, *22* (10), 1302-1306.

54. R. Linding; J. Schymkowitz; F. Rousseau; F. Diella; L. Serrano, *J. Mol. Biol.* **2004**, *342* (1), 345-353.

55. F. Rousseau; J. Schymkowitz; L. Serrano, *Cur. Opin. Struct. Biol.* **2006**, *16* (1), 118-126.

56. L. Aulisa; N. Forraz; C. McGuckin; J. D. Hartgerink, *Acta Biomater.* **2009**, *5* (3), 842-853.

57. G. A. Silva; C. Czeisler; K. L. Niece; E. Beniash; D. A. Harrington; J. A. Kessler; S. I. Stupp, *Science* **2004**, *303* (5662), 1352-1355.

58.   U. I. M. Gerling; E. Brandenburg; H. v. Berlepsch; K. Pagel; B. Koksch, *Biomacromolecules* **2011**, *12* (8), 2988-2996.

59.   A. Yoshizumi; J. M. Fletcher; Z. X. Yu; A. V. Persikov; G. J. Bartlett; A. L. Boyle; T. L. Vincent; D. N. Woolfson; B. Brodsky, *J. Biol. Chem.* **2011**, *286* (20), 17512-17520.

60.   B. Ozbas; J. Kretsinger; K. Rajagopal; J. P. Schneider; D. J. Pochan, *Macromolecules* **2004**, *37* (19), 7331-7337.

61.   J. P. Schneider; D. J. Pochan; B. Ozbas; K. Rajagopal; L. Pakstis; J. Kretsinger, *J. Am. Chem. Soc.* **2002**, *124* (50), 15030-15037.

62.   K. Pagel; S. C. Wagner; R. R. Araghi; H. von Berlepsch; C. Böttcher; B. Koksch, *Chem. Eur. J.* **2008**, *14* (36), 11442-11451.

63.   K. Pagel; S. C. Wagner; K. Samedov; H. von Berlepsch; C. Böttcher; B. Koksch, *J. Am. Chem. Soc.* **2006**, *128* (7), 2196-2197.

64.   K. Pagel; T. Vagt; B. Koksch, *Org. Biomol. Chem.* **2005**, *3* (21), 3843-3850.

65.   K. Pagel; K. Seeger; B. Seiwert; A. Villa; A. E. Mark; S. Berger; B. Koksch, *Org. Biomol. Chem.* **2005**, *3* (7), 1189-1194.

66.   J. M. Mason; K. M. Arndt, *ChemBioChem* **2004**, *5* (2), 170-176.

67.   P. Burkhard; S. Ivaninskii; A. Lustig, *J. Mol. Biol.* **2002**, *318* (3), 901-910.

68.   K. Pagel; T. Vagt; T. Kohajda; B. Koksch, *Org. Biomol. Chem.* **2005**, *3* (14), 2500-2502.

69.   D. Schöne; B. Schade; C. Bottcher; B. Koksch, *Beilstein J. Org. Chem.* **2015**, *11*, 792-803.

70.   K. Folmert; M. Broncel; H. v. Berlepsch; C. H. Ullrich; M.-A. Siegert; B. Koksch, *Beilstein J. Org. Chem.* **2016**, *12*, 2462-2470.

71.   E. Zacco; J. Hutter; J. L. Heier; J. Mortier; P. H. Seeberger; B. Lepenies; B. Koksch, *ACS Chem. Biol.* **2015**, *10* (9), 2065-2072.

72.   M. Broncel; J. A. Falenski; S. C. Wagner; C. P. R. Hackenberger; B. Koksch, *Chem. Eur. J.* **2010**, *16* (26), 7881-7888.

73.   M. Salwiczek; S. Samsonov; T. Vagt; E. Nyakatura; E. Fleige; J. Numata; H. Cölfen; M. T. Pisabarro; B. Koksch, *Chem. Eur. J.* **2009**, *15* (31), 7628-7636.

74.   M. Salwiczek; B. Koksch, *ChemBioChem* **2009**, *10* (18), 2867-2870.

75.   C. Jackel; W. Seufert; S. Thust; B. Koksch, *ChemBioChem* **2004**, *5* (5), 717-720.

76.   C. M. Dobson, *Nature* **2003**, *426* (6968), 884-890.

77.   C. M. Dobson, *Nature* **2002**, *418* (6899), 729-730.

78.   F. Chiti; C. M. Dobson, *Annu. Rev. Biochem.* **2006**, *75*, 333-366.

79.   C. Hilbich; B. Kisterswoike; J. Reed; C. L. Masters; K. Beyreuther, *J. Mol. Biol.* **1991**, *218* (1), 149-163.

80.   R. Virchow, *Archiv. f. Pathol. Anat.* **1853**, *6* (4), 135-138.

81.   D. E. Ehrnhoefer; B. K. Y. Wong; M. R. Hayden, *Nat. Rev. Drug Discov.* **2011**, *10* (11), 853-867.

82.   N. S. Moily; A. R. Ormsby; A. Stojilovic; Y. M. Ramdzan; J. Diesch; R. D. Hannan; M. S. Zajac; A. J. Hannan; A. Oshlack; D. M. Hatters, *Mol. Cell. Neurosci.* **2017**, *83*, 103-112.

83.   A. Abedini; A. M. Schmidt, *FEBS Lett.* **2013**, *587* (8), 1119-1127.

84.   A. B. Soriaga; S. Sangwan; R. Macdonald; M. R. Sawaya; D. Eisenberg, *J. Phys. Chem. B* **2016**, *120* (26), 5810-5816.

85. J. J. W. Wiltzius; S. A. Sievers; M. R. Sawaya; D. Eisenberg, *Prot. Sci.* **2009**, *18* (7), 1521-1530.

86. A. Mata; L. Urrea; S. Vilches; F. Llorens; K. Thune; J. C. Espinosa; O. Andreoletti; A. M. Sevillano; J. M. Torres; J. R. Requena; I. Zerr; I. Ferrer; R. Gavin; J. A. del Rio, *Mol. Neurobiol.* **2017**, *54* (8), 6412-6425.

87. E. Katorcha; N. Makarava; Y. J. Lee; I. Lindberg; M. J. Monteiro; G. G. Kovacs; I. V. Baskakov, *PLoS Pathog.* **2017**, *13* (8), 36-42.

88. C. Y. Li; R. Kisilevsky, *Amyloid-J. Protein Fold. Disord.* **2008**, *15* (4), 246-254.

89. G. Tetz; V. Tetz, *Pediatr. Pulmonol.* **2017**, *52*, 300-312.

90. S. Ono; S. Matsuno; N. Shimizu; S. Shoji; A. Tamaoka, *Lancet* **1998**, *351* (9107), 956-957.

91. E. Sinagra; M. Ciofalo; G. Tomasello; F. Cappello; G. C. Morreale; G. Amvrosiadis; P. Damiani; F. Damiani; G. Pompei; A. G. Rizzo; C. Canale; G. Mastrocinque; F. Carini; D. Raimondo, *Prog. Nutr.* **2017**, *19* (1), 5-13.

92. M. Stoppini; V. Bellotti, *J. Biol. Chem.* **2015**, *290* (16), 9951-9958.

93. G. B. Piccoli; M. Hachemi; I. Molfino; J. P. Coindre; C. Boursot, *BMC Nephrol.* **2017**, *18*, 1-6.

94. D. J. Selkoe, *Nature* **2003**, *426* (6968), 900-904.

95. D. J. Selkoe; J. Hardy, *EMBO Mol. Med.* **2016**, *8* (6), 595-608.

96. Z. Y. Lv; C. C. Tan; J. T. Yu; L. Tan, *Neurotox. Res.* **2017**, *32* (4), 707-722.

97. B. H. Meier; R. Riek; A. Bockmann, *Trends Biochem.Sci.* **2017**, *42* (10), 777-787.

98. P. Verwilst; H. R. Kim; J. Seo; N. W. Sohn; S. Y. Cha; Y. Kim; S. Maeng; J. W. Shin; J. H. Kwak; C. Kang; J. S. Kim, *J. Am. Chem. Soc.* **2017**, *139* (38), 13393-13403.

99. T. C. W. Julia; A. M. Goate, *Cold Spring Harb. Perspect. Med.* **2017**, *7* (6), 11169-11187.

100. D. A. Butterfield; D. Boyd-Kimball, *BBA-Proteins Proteomics* **2005**, *1703* (2), 149-156.

101. A. Y. Hung; C. Haass; R. M. Nitsch; W. Q. Qiu; M. Citron; R. J. Wurtman; J. H. Growdon; D. J. Selkoe, *J. Biol. Chem.* **1993**, *268* (31), 22959-22962.

102. B. De Strooper, *Physiol. Rev.* **2010**, *90* (2), 465-494.

103. S. T. Ferreira; M. N. N. Vieira; F. G. De Felice, *IUBMB Life* **2007**, *59* (4-5), 332-345.

104. K. M. Psonka-Antonczyk; P. Hammarström; L. B. G. Johansson; M. Lindgren; B. T. Stokke; K. P. R. Nilsson; S. Nyström, *Front. Chem.* **2016**, *4-31*, 44.

105. G. Bitan; M. D. Kirkitadze; A. Lomakin; S. S. Vollers; G. B. Benedek; D. B. Teplow, *Proc. Natl. Acad. Sci. USA* **2003**, *100* (1), 330-335.

106. C. G. Glabe, *J. Biol. Chem.* **2008**, *283* (44), 29639-29643.

107. H. Kadowaki; H. Nishitoh; F. Urano; C. Sadamitsu; A. Matsuzawa; K. Takeda; H. Masutani; J. Yodoi; Y. Urano; T. Nagano; H. Ichijo, *Cell Death Differ* **2005**, *12* (1), 19-24.

108. I. Benilova; E. Karran; B. De Strooper, *Nat. Neurosci.* **2012**, *15* (3), 349-357.

109. M. Sakono; T. Zako, *Febs J.* **2010**, *277* (6), 1348-1358.

110. A. d. C. Alonso; I. Grundke-Iqbal; H. S. Barra; K. Iqbal, *Proc. Natl. Acad. Sci. USA* **1997**, *94* (1), 298-303.

111. P. Scheltens; K. Blennow; M. M. B. Breteler; B. de Strooper; G. B. Frisoni; S. Salloway; W. M. Van der Flier, *Lancet* **2016**, *388* (10043), 505-517.

112. C. Soto; L. Estrada; J. Castilla, *Trends Biochem.Sci.* **2006**, *31* (3), 150-155.

113. C. M. Dobson, *Nature* **2002**, *418* (6899), 729-730.

114. M. Stefani; C. M. Dobson, *J. Mol. Med.* **2003**, *81*, 678-699.
115. J. D. Sipe; M. D. Benson; J. N. Buxbaum; S. Ikeda; G. Merlini; M. J. M. Saraiva; P. Westermark, *Amyloid-J. Protein Fold. Disord.* **2012**, *19* (4), 167-170.
116. M. T. Elghetany; A. Saleem; K. Barr, *Ann. Clin. Lab. Sci.* **1989**, *19* (3), 190-195.
117. H. Levine, *Prot. Sci.* **1993**, *2* (3), 404-410.
118. M. Biancalana; S. Koide, *BBA-Proteins Proteomics* **2010**, *1804* (7), 1405-1412.
119. E. Brandenburg; H. von Berlepschz; B. Koksch, *Mol. Biosyst.* **2012**, *8* (2), 557-564.
120. A. Hawe; M. Sutter; W. Jiskoot, *Pharm. Res.* **2008**, *25* (7), 1487-1499.
121. N. D. Younan; J. H. Viles, *Biochemistry* **2015**, *54* (28), 4297-4306.
122. R. Kayed; E. Head; J. L. Thompson; T. M. McIntire; S. C. Milton; C. W. Cotman; C. G. Glabe, *Science* **2003**, *300* (5618), 486-489.
123. C. M. Dobson; M. Karplus, *Cur. Opin. Struct. Biol.* **1999**, *9* (1), 92-101.
124. O. S. Makin; L. C. Serpell, *Febs J.* **2005**, *272*, 5950-5961.
125. O. S. Makin; E. Atkins; P. Sikorski; J. Johansson; L. C. Serpell, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 315.
126. A. S. Cohen; E. Calkins, *Nature* **1959**, *183* (4669), 1202-1203.
127. N. Greenfield; G. D. Fasman, *Biochemistry* **1969**, *8* (10), 4108-4116.
128. L. Pauling; R. B. Corey, *Proc. Natl. Acad. Sci. USA* **1951**, *37* (11), 729-740.
129. D. S. Eisenberg; M. R. Sawaya, Structural Studies of Amyloid Proteins at the Molecular Level. In *Annual Review of Biochemistry, Vol 86*, R. D. Kornberg, Ed. Annual Reviews: Palo Alto, **2017**; Vol. 86, pp 69-95.
130. D. Thirumalai; D. K. Klimov; R. I. Dima, *Cur. Opin. Struct. Biol.* **2003**, *13* (2), 146-159.
131. S. B. Prusiner, *Science* **2012**, *336* (6088), 1511-1513.
132. A. Aguzzi; L. Rajendran, *Neuron* **2009**, *64* (6), 783-790.
133. K. C. Luk; V. Kehm; J. Carroll; B. Zhang; P. O'Brien; J. Q. Trojanowski; V. M. Y. Lee, *Science* **2012**, *338* (6109), 949-953.
134. R. Nelson; M. R. Sawaya; M. Balbirnie; A. O. Madsen; C. Riekel; R. Grothe; D. Eisenberg, *Nature* **2005**, *435* (7043), 773-778.
135. M. I. Ivanova; S. A. Sievers; E. L. Guenther; L. M. Johnson; D. D. Winkler; A. Galaleldeen; M. R. Sawaya; P. J. Hart; D. S. Eisenberg, *Proc. Natl. Acad. Sci. USA* **2014**, *111* (1), 197-201.
136. J. J. W. Wiltzius; M. Landau; R. Nelson; M. R. Sawaya; M. I. Apostol; L. Goldschmidt; A. B. Soriaga; D. Cascio; K. Rajashankar; D. Eisenberg, *Nat. Struct. Mol. Biol.* **2009**, *16* (9), 973-998.
137. J. C. Stroud, *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **2013**, *69*, 540-545.
138. M. R. Sawaya; S. Sambashivan; R. Nelson; M. I. Ivanova; S. A. Sievers; M. I. Apostol; M. J. Thompson; M. Balbirnie; J. J. W. Wiltzius; H. T. McFarlane; A. O. Madsen; C. Riekel; D. Eisenberg, *Nature* **2007**, *447* (7143), 453-457.
139. K. C. Chou; M. Pottle; G. Nemethy; Y. Ueda; H. A. Scheraga, *J. Mol. Biol.* **1982**, *162* (1), 89-112.
140. M. I. Apostol; M. R. Sawaya; D. Cascio; D. Eisenberg, *J. Biol. Chem.* **2010**, *285* (39), 29671-29675.
141. R. Nelson; M. R. Sawaya; M. Balbirnie; A. O. Madsen; C. Riekel; R. Grothe; D. Eisenberg, *Nature* **2005**, *435* (7043), 773-778.
142. I. Cherny; E. Gazit, *Angew. Chem.-Int. Ed.* **2008**, *47* (22), 4062-4069.

143. H. A. Lashuel; S. R. LaBrenz; L. Woo; L. C. Serpell; J. W. Kelly, *J. Am. Chem. Soc.* **2000**, *122* (22), 5262-5277.
144. C. E. MacPhee; D. N. Woolfson, *Curr. Opin. Solid State Mater. Sci.* **2004**, *8* (2), 141-149.
145. T. Konno; S. Oiki; K. Hasegawa; H. Naiki, *Biochemistry* **2004**, *43* (42), 13613-13620.
146. M. Balbirnie; R. Grothe; D. S. Eisenberg, *Proc. Natl. Acad. Sci. USA* **2001**, *98* (5), 2375-2380.
147. K. Tsemekhman; L. Goldschmidt; D. Eisenberg; D. Baker, *Prot. Sci.* **2007**, *16* (4), 761-764.
148. A. D. Williams; S. Shivaprasad; R. Wetzel, *J. Mol. Biol.* **2006**, *357* (4), 1283-1294.
149. A. D. Williams; E. Portelius; I. Kheterpal; J.-T. Guo; K. D. Cook; Y. Xu; R. Wetzel, *J. Mol. Biol.* **2004**, *335* (3), 833-842.
150. A. Morimoto; K. Irie; K. Murakami; Y. Masuda; H. Ohigashi; M. Nagao; H. Fukuda; T. Shimizu; T. Shirasawa, *J. Biol. Chem.* **2004**, *279* (50), 52781-52788.
151. L. Adler-Abramovich; E. Gazit, *Chem. Soc. Rev.* **2014**, *43* (20), 6881-6893.
152. J. Kopecek; J. Y. Yang, *Angew. Chem.-Int. Ed.* **2012**, *51* (30), 7396-7417.
153. M. Reches; E. Gazit, *Nat. Nanotechnol.* **2006**, *1* (3), 195-200.
154. M. Reches; E. Gazit, *Phys. Biol.* **2006**, *3* (1), S10-S19.
155. X. H. Yan; P. L. Zhu; J. B. Li, *Chem. Soc. Rev.* **2010**, *39* (6), 1877-1890.
156. L. C. Wu; J. Y. Yang; J. Kopecek, *Biomaterials* **2011**, *32* (23), 5341-5353.
157. G. Nystrom; W. K. Fong; R. Mezzenga, *Biomacromolecules* **2017**, *18* (9), 2858-2865.
158. Z. S. Al-Garawi; B. A. McIntosh; D. Neill-Hall; A. A. Hatimy; S. M. Sweet; M. C. Bagley; L. C. Serpell, *Nanoscale* **2017**, *9* (30), 10773-10783.
159. G. Kabay; A. E. Meydan; G. K. Can; C. Demirci; M. Mutlu, *Mater. Sci. Eng. C- Mater. Biol. Appl.* **2017**, *81*, 271-279.
160. D. M. Fowler; A. V. Koulov; C. Alory-Jost; M. S. Marks; W. E. Balch; J. W. Kelly, *Plos Biol.* **2006**, *4* (1), 100-107.
161. S. K. Maji; M. H. Perrin; M. R. Sawaya; S. Jessberger; K. Vadodaria; R. A. Rissman; P. S. Singru; K. P. R. Nilsson; R. Simon; D. Schubert; D. Eisenberg; J. Rivier; P. Sawchenko; W. Vale; R. Riek, *Science* **2009**, *325* (5938), 328.
162. H. Ghandehari; J. Cappello, *Pharm. Res.* **1998**, *15* (6), 813-815.
163. D. R. Booth; M. Sunde; V. Bellotti; C. V. Robinson; W. L. Hutchinson; P. E. Fraser; P. N. Hawkins; C. M. Dobson; S. E. Radford; C. C. F. Blake; M. B. Pepys, *Nature* **1997**, *385* (6619), 787-793.
164. M. Kato; T. N. W. Han; S. H. Xie; K. Shi; X. L. Du; L. C. Wu; H. Mirzaei; E. J. Goldsmith; J. Longgood; J. M. Pei; N. V. Grishin; D. E. Frantz; J. W. Schneider; S. Chen; L. Li; M. R. Sawaya; D. Eisenberg; R. Tycko; S. L. McKnight, *Cell* **2012**, *149* (4), 753-767.
165. K. Si; S. Lindquist; E. R. Kandel, *Cell* **2003**, *115* (7), 879-891.
166. A. Clark; G. J. Cooper; C. E. Lewis; J. F. Morris; A. C. Willis; K. B. Reid; R. C. Turner, *Lancet* **1987**, *2* (8553), 231-234.
167. K. L. Luskey, *Diabetes care* **1992**, *15* (2), 297-299.
168. S. E. Kahn; S. Andrikopoulos; C. B. Verchere, *Diabetes* **1999**, *48* (2), 241-253.
169. P. Cao; P. Marek; H. Noor; V. Patsalo; L. H. Tu; H. Wang; A. Abedini; D. P. Raleigh, *FEBS Lett.* **2013**, *587* (8), 1106-1118.

170. WHO, **download:** **30.10.2017**, http://www.who.int/mediacentre/factsheets/fs312/en/index.html.

171. C. D. Mathers; D. Loncar, *Plos Med.* **2006**, *3* (11), 2011-2030.

172. WHO, *Geneva, Global status report on noncommunicable diseases 2010* **download: 30.10.2017**.

173. K. Pillay; P. Govender, *Biomed. Res. Int.* **2013**.

174. E. T. A. S. Jaikaran; C. E. Higham; L. C. Serpell; J. Zurdo; M. Gross; A. Clark; P. E. Fraser, *J. Mol. Biol.* **2001**, *308* (3), 515-525.

175. A. V. Kajava; U. Aebi; A. C. Steven, *J. Mol. Biol.* **2005**, *348* (2), 247-252.

176. S. Luca; W.-M. Yau; R. Leapman; R. Tycko, *Biochemistry* **2007**, *46* (47), 13505- 13522.

177. M. Krampert; J. Bernhagen; J. Schmucker; A. Horn; A. Schmauder; H. Brunner; W. Voelter; A. Kapurniotu, *Chem. Biol.* **2000**, *7* (11), 855-871.

178. T. H. Zhu; Y. X. Wang; B. J. He; J. J. Zang; Q. He; W. H. Zhang, *Diabetes Metab. Res. Rev.* **2011**, *27* (1), 28-34.

179. P. Degano; R. A. Silvestre; M. Salas; E. Peiro; J. Marco, *Regul. Pept. Suppl.* **1993**, *43* (1-2), 91-96.

180. B. Akesson; G. Panagiotidis; P. Westermark; I. Lundquist, *Regul. Pept. Suppl.* **2003**, *111* (1-3), 55-60.

181. C. Martin, *Diabetes Educ.* **2006**, *32*, 101S-104S.

182. G. J. S. Cooper; A. C. Willis; A. Clark; R. C. Turner; R. B. Sim; K. B. M. Reid, *Proc. Natl. Acad. Sci. USA* **1987**, *84* (23), 8628-8632.

183. P. Cao; P. Marek; H. Noor; V. Patsalo; L.-H. Tu; H. Wang; A. Abedini; D. P. Raleigh, *FEBS Lett.* **2013**, *587* (8), 1106-1118.

184. S. J. C. Lee; T. S. Choi; J. W. Lee; H. J. Lee; D.-G. Mun; S. Akashi; S.-W. Lee; M. H. Lim; H. I. Kim, *Chem. Sci.* **2016**, *7* (8), 5398-5406.

185. L. M. Young; P. Cao; D. P. Raleigh; A. E. Ashcroft; S. E. Radford, *J. Am. Chem. Soc.* **2014**, *136* (2), 660-670.

186. C. Goldsbury; J. Kistler; U. Aebi; T. Arvinte; G. J. S. Cooper, *J. Mol. Biol.* **1999**, *285* (1), 33-39.

187. C. S. Goldsbury; G. J. S. Cooper; K. N. Goldie; S. A. Muller; E. L. Saafi; W. T. M. Gruijters; M. P. Misur, *J. Struct. Biol.* **1997**, *119* (1), 17-27.

188. A. Lorenzo; B. Razzaboni; G. C. Weir; B. A. Yankner, *Nature* **1994**, *368* (6473), 756-760.

189. M. F. M. Engel; L. Khemtemourian; C. C. Kleijer; H. J. D. Meeldijk; J. Jacobs; A. J. Verkleij; B. de Kruijff; J. A. Killian; J. W. M. Hoppener, *Proc. Natl. Acad. Sci. USA* **2008**, *105* (16), 6033-6038.

190. T. A. Mirzabekov; M. C. Lin; B. L. Kagan, *J. Biol. Chem.* **1996**, *271* (4), 1988-1992.

191. S. P. Zhang; J. X. Liu; G. MacGibbon; M. Dragunow; G. J. S. Cooper, *J. Mol. Biol.* **2002**, *324* (2), 271-285.

192. K. Hensley; J. M. Carney; M. P. Mattson; M. Aksenova; M. Harris; J. F. Wu; R. A. Floyd; D. A. Butterfield, *Proc. Natl. Acad. Sci. USA* **1994**, *91* (8), 3270-3274.

193. M. P. Mattson; Y. Goodman, *Brain Res.* **1995**, *676* (1), 219-224.

194. A. S. DeToma; S. Salamekh; A. Ramamoorthy; M. H. Lim, *Chem. Soc. Rev.* **2012**.

195. R. Kodali; R. Wetzel, *Cur. Opin. Struct. Biol.* **2007**, *17* (1), 48-57.

196. S. H. Shim; R. Gupta; Y. L. Ling; D. B. Strasfeld; D. P. Raleigh; M. T. Zanni, *Proc. Natl. Acad. Sci. USA* **2009**, *106* (16), 6614-6619.

197. L. A. Scrocchi; Y. Chen; S. Waschuk; F. Wang; S. Cheung; A. A. Darabie; J. McLaurin; P. E. Fraser, *J. Mol. Biol.* **2002**, *318* (3), 697-706.

198. R. Azriel; E. Gazit, *J. Biol. Chem.* **2001**, *276* (36), 34156-34161.

199. L. A. Scrocchi; K. Ha; Y. Chen; L. Wu; F. Wang; P. E. Fraser, *J. Struct. Biol.* **2003**, *141* (3), 218-227.

200. K. Tenidis; M. Waldner; J. Bernhagen; W. Fischle; M. Bergmann; M. Weber; M.-L. Merkle; W. Voelter; H. Brunner; A. Kapurniotu, *J. Mol. Biol.* **2000**, *295* (4), 1055-1071.

201. P. Westermark; U. Engström; K. H. Johnson; G. T. Westermark; C. Betsholtz, *Proc. Natl. Acad. Sci. USA* **1990**, *87* (13), 5036-5040.

202. M. R. Nilsson; D. P. Raleigh, *J. Mol. Biol.* **1999**, *294* (5), 1375-1385.

203. C. Goldsbury; K. Goldie; J. Pellaud; J. Seelig; P. Frey; S. A. Müller; J. Kistler; G. J. S. Cooper; U. Aebi, *J. Struct. Biol.* **2000**, *130* (2), 352-362.

204. Y. Mazor; S. Gilead; I. Benhar; E. Gazit, *J. Mol. Biol.* **2002**, *322* (5), 1013-1024.

205. Z. Zhang; H. Chen; L. Lai, *Bioinformatics* **2007**, *23* (17), 2218-2225.

206. S.-H. Shim; R. Gupta; Y. L. Ling; D. B. Strasfeld; D. P. Raleigh; M. T. Zanni, *Proc. Natl. Acad. Sci. USA* **2009**, *106* (16), 6614-6619.

207. K. Tenidis; M. Waldner; J. Bernhagen; W. Fischle; M. Bergmann; M. Weber; M. L. Merkle; W. Voelter; H. Brunner; A. Kapurniotu, *J. Mol. Biol.* **2000**, *295* (4), 1055-1071.

208. A. B. Soriaga; S. Sangwan; R. Macdonald; M. R. Sawaya; D. Eisenberg, *J. Phys. Chem. B* **2016**, *120* (26), 5810-5816.

209. W. Deng; A. Cao; L. Lai, *Prot. Sci.* **2008**, *17* (6), 1102-1105.

210. R. Azriel; E. Gazit, *J. Biol. Chem.* **2001**, *276* (36), 34156-34161.

211. Z. Cai; J. Li; C. Yin; Z. Yang; J. Wu; R. Zhou, *J. Phys. Chem. B* **2014**, *118* (1), 48-57.

212. A. Melquiond; J.-C. Gelly; N. Mousseau; P. Derreumaux, *J. Chem. Phys.* **2007**, *126* (6), 065101.

213. C. Wu; H. Lei; Y. Duan, *Biophys. J.* **2005**, *88* (4), 2897-2906.

214. C. Wu; H. Lei; Y. Duan, *J. Am. Chem. Soc.* **2005**, *127* (39), 13530-13537.

215. C. Wu; H. Lei; Y. Duan, *Biophys. J.* **2004**, *87* (5), 3000-3009.

216. D. Zanuy; B. Ma; R. Nussinov, *Biophys. J.* **2003**, *84* (3), 1884-1894.

217. R. Azriel; E. Gazit, *J. Biol. Chem.* **2001**, *276* (36), 34156-34161.

218. H. X. Lei; C. Wu; Z. X. Wang; Y. Duan, *J. Mol. Biol.* **2006**, *356* (4), 1049-1063.

219. J. T. Nielsen; M. Bjerring; M. D. Jeppesen; R. O. Pedersen; J. M. Pedersen; K. L. Hein; T. Vosegaard; T. Skrydstrup; D. E. Otzen; N. C. Nielsen, *Angew. Chem.-Int. Ed.* **2009**, *48* (12), 2118-2121.

220. J. Madine; E. Jack; P. G. Stockley; S. E. Radford; L. C. Serpell; D. A. Middleton, *J. Am. Chem. Soc.* **2008**, *130* (45), 14990-15001.

221. J. J. Guo; Y. Zhang; L. L. Ning; P. Z. Jiao; H. X. Liu; X. J. Yao, *Biochim. Biophys. Acta-Gen. Subj.* **2014**, *1840* (1), 357-366.

222. A. Melquiond; J. C. Gelly; N. Mousseau; P. Derreumaux, *J. Chem. Phys.* **2007**, *126* (6).

223. G. Pathuri; H. B. Agashe; V. Awasthi; H. Gali, *J. Label. Compd. Radiopharm.* **2010**, *53* (3-4), 186-191.

224. B. J. M. Stryer L., Tymoczko J. L. , *Stryer Biochemie*. Spectrum Akademischer   Verlag: **2007**; Vol. 6.

225. J. Deutscher; C. Francke; P. W. Postma, *Microbiol. Mol. Biol. Rev.* **2006**, *70* (4), 939-1031.

226. G. Manning; D. B. Whyte; R. Martinez; T. Hunter; S. Sudarsanam, *Science* **2002**, *298* (5600), 1912-+.

227. W. Vogel; R. Lammers; J. T. Huang; A. Ullrich, *Science* **1993**, *259* (5101), 1611- 1614.

228. G. T. Robillard; M. Blaauw, *Biochemistry* **1987**, *26* (18), 5796-5803.

229. X. Soberon; M. H. Saier, *J. Mol. Microbiol. Biotechnol.* **2006**, *11* (6), 302-307.

230. K. Kollmann; S. Pohl; K. Marschner; M. Encarnacao; I. Sakwa; S. Tiede; B. J. Poorthuis; T. Lubke; S. Muller-Loennies; S. Storch; T. Braulke, *Eur. J. Cell Biol.* **2010**, *89* (1), 117-123.

231. H. M. Zhao; W. A. van der Donk, *Curr. Opin. Biotechnol.* **2003**, *14* (6), 583-589.

232. P. R. Thompson; J. Schwartzenhauer; D. W. Hughes; A. M. Berghuis; G. D. Wright, *J. Biol. Chem.* **1999**, *274* (43), 30697-30706.

233. A. K. H. Hirsch; F. R. Fischer; F. Diederich, *Angew. Chem.-Int. Ed.* **2007**, *46* (3), 338-352.

234. J. Deutscher; F. M. D. Ake; M. Derkaoui; A. C. Zebre; T. N. Cao; H. Bouraoui; T. Kentache; A. Mokhtari; E. Milohanic; P. Joyet, *Microbiol. Mol. Biol. Rev.* **2014**, *78* (2), 231-256.

235. S. Scholtes; E. Kramer; M. Weisser; W. Roth; R. Luginbuhl; T. Grossner; W. Richter, *J. Cell. Physiol.* **2018**, *233* (1), 699-711.

236. T. H. Chi; G. R. Crabtree, *Science* **2000**, *287* (5460), 1937-+.

237. G. M. Yang; X. Y. Peng; Y. Wu; T. Li; L. M. Liu, *Am. J. Physiol. Cell. Physiol.*   **2017**, *313* (4), C362-C370.

238. E. G. Krebs; E. H. Fischer, *J. Biol. Chem.* **1955**, *216* (1), 113-120.

239. E. W. Sutherland; W. D. Wosilait, *Nature* **1955**, *175* (4447), 169-170.

240. E. G. Krebs; D. J. Graves; E. H. Fischer, *J. Biol. Chem.* **1959**, *234* (11), 2867-2873.

241. R. A. Lindberg; A. M. Quinn; T. Hunter, *Trends Biochem.Sci.* **1992**, *17* (3), 114-119.

242. M. Morris; G. M. Knudsen; S. Maeda; J. C. Trinidad; A. Ioanoviciu; A. L. Burlingame; L. Mucke, *Nat. Neurosci.* **2015**, *18* (8), 1183-+.

243. F. Ahmad; K. Singh; D. Das; R. Gowaikar; E. Shaw; A. Ramachandran; K. V. Rupanagudi; R. P. Kommaddi; D. A. Bennett; V. Ravindranath, *Antioxid. Redox Signal.* **2017**, *27* (16), 1269-1280.

244. T. Ma; M. A. Trinh; A. J. Wexler; C. Bourbon; E. Gatti; P. Pierre; D. R. Cavener; E. Klann, *Nat. Neurosci.* **2013**, *16* (9), 1299-U1185.

245. R. A. DeFronzo; D. Tripathy, *Diabetes care* **2009**, *32*, S157-S163.

246. C. Beharry; L. S. Cohen; J. Di; K. Ibrahim; S. Briffa-Mirabella; A. D. Alonso, *Neurosci. Bull.* **2014**, *30* (2), 346-358.

247. W. T. Lee; C. C. L. Hsian; Y. A. Lim, *Neuroreport* **2017**, *28* (16), 1043-1048.

248. A. Lathuiliere; P. Valdes; S. Papin; M. Cacquevel; C. Maclachlan; G. W. Knott; A. Muhs; P. Paganetti; B. L. Schneider, *Sci. Rep.* **2017**, *7*.

249. T. N. Seyfried; L. M. Shelton, *Nutr. Metab.* **2010**, *7*.

250. R. J. Levy, *Shock* **2007**, *28* (1), 24-28.

251. M. S. Collett; R. L. Erikson, *Proc. Natl. Acad. Sci. USA* **1978**, *75* (4), 2021-2024.

252. W. C. Barker; M. O. Dayhoff, *Proc. Natl. Acad. Sci. USA* **1982**, *79* (9), 2836-2839.

253. J. D. Buxbaum; E. H. Koo; P. Greengard, *Proc. Natl. Acad. Sci. USA* **1993**, *90* (19), 9195-9198.
254. M. Broncel; S. C. Wagner; C. P. R. Hackenberger; B. Koksch, *Chem. Commun.* **2010**, *46* (18), 3080-3082.
255. M. Inoue; A. Hirata; K. Tainaka; T. Morii; T. Konno, *Biochemistry* **2008**, *47* (45), 11847-11857.
256. A. Feliciello; M. E. Gottesman; E. V. Avvedimento, *Cell. Signal.* **2005**, *17* (3), 279-287.
257. E. London; M. Nesterova; C. A. Stratakis, *J. Mol. Endocrinol.* **2017**, *59* (1), 1-12.
258. P. A. Insel; A. Wilderman; L. Zhang; M. M. Keshwani; A. C. Zambon, *Horm. Metab. Res.* **2014**, *46* (12), 854-862.
259. M. Haidar; G. Ramdani; E. J. Kennedy; G. Langsley, *Horm. Metab. Res.* **2017**, *49* (4), 296-300.
260. S. S. Taylor; P. Banky; D. Johnson; L. Burns, *Biochemistry* **2001**, *40* (29), 8610- 8611.
261. R. Lang; A. L. Gundlach; F. E. Holmes; S. A. Hobson; D. Wynick; T. Hokfelt; B. Kofler, *Pharmacol. Rev.* **2015**, *67* (1), 118-175.
262. J. L. Meinkoth; A. S. Alberts; W. Went; D. Fantozzi; S. S. Taylor; M. Hagiwara; M. Montminy; J. R. Feramisco, *Mol. Cell. Biochem.* **1993**, *127/128*, 179-186.
263. E. R. Kandel, *Mol. Brain* **2012**, *5*.
264. J. Berthet; T. W. Rall; E. W. Sutherland, *J. Biol. Chem.* **1957**, *224* (1), 463-475.
265. D. A. Walsh; J. P. Perkins; E. G. Krebs, *Fed. Proc.* **1968**, *27* (2), 339-&.
266. B. Zimmermann; J. A. Chiorini; Y. L. Ma; R. M. Kotin; F. W. Herberg, *J. Biol. Chem.* **1999**, *274* (9), 5370-5378.
267. C. Kim; C. Y. Cheng; S. A. Saldanha; S. S. Taylor, *Cell* **2007**, *130* (6), 1032-1043.
268. S. Bonn; S. Herrero; C. B. Breitenlechner; A. Erlbruch; W. Lehmann; R. A. Engh; M. Gassel; D. Bossemeyer, *J. Biol. Chem.* **2006**, *281* (34), 24818-24830.
269. T. G. Davies; M. L. Verdonk; B. Graham; S. Saalau-Bethell; C. C. Hamlett; T. McHardy; I. Collins; M. D. Garrett; P. Workman; S. J. Woodhead; H. Jhoti; D. Barford, *J. Mol. Biol.* **2007**, *367* (3), 882-894.
270. G. S. Anand; C. A. Hughes; J. M. Jones; S. S. Taylor; E. A. Komives, *J. Mol. Biol.* **2002**, *323* (2), 377-386.
271. E. G. Krebs; J. A. Beavo, *Annu. Rev. Biochem.* **1979**, *48*, 923-959.
272. Ö. Zetterqvist; U. Ragnarsson, *FEBS Lett.* **1982**, *139* (2), 287-290.
273. G. C. R. Ellis-Davies, *Nat. methods* **2007**, *4* (8), 619-628.
274. S. R. Adams; R. Y. Tsien, *Annu. Rev. Physiol.* **1993**, *55* (1), 755-784.
275. A. Morckel; D. Young; A. Deiters; N. Nascone-Yoder, *Dev. Biol* **2009**, *331* (2), 482-482.
276. A. Deiters, *Cur. Opin. Chem. Biol.* **2009**, *13* (5-6), 678-686.
277. G. Mayer; A. Heckel, *Angew. Chem.-Int. Ed.* **2006**, *45* (30), 4900-4921.
278. J. Engels; E. J. Schlaeger, *J. med. Chem.* **1977**, *20* (7), 907-911.
279. J. H. Kaplan; B. Forbush; J. F. Hoffman, *Biochemistry* **1978**, *17* (10), 1929-1935.
280. F. C. Santos; A. M. S. Soares; M. S. T. Goncalves; S. P. G. Costa, *Amino Acids* **2017**, *49* (6), 1077-1088.
281. J. Luo; J. Torres-Kolbus; J. Liu; A. Deiters, *ChemBioChem* **2017**, *18* (14), 1442- 1447.
282. C. Zscherp; A. Barth, *Biochemistry* **2001**, *40* (7), 1875-1883.
283. L. Kröck; A. Heckel, *Angew. Chem.* **2005**, *117* (3), 475-477.
284. R. M. Zadegan; W. L. Hughes, *ACS Synth. Biol.* **2017**, *6* (10), 1800-1806.

285. P. T. Wong; E. W. Roberts; S. Z. Tang; J. Mukherjee; J. Cannon; A. J. Nip; K. Corbin; M. F. Krummel; S. K. Choi, *ACS Chem. Biol.* **2017**, *12* (4), 1001-1010.

286. T. Lucas; F. Schafer; P. Muller; S. A. Eming; A. Heckel; S. Dimmeler, *Nat. Commun.* **2017**, *8*.

287. P. M. Kusen; G. Wandrey; V. Krewald; M. Holz; S. M. zu Berstenhorst; J. Buchs; J. Pietruszka, *J. Biotechnol.* **2017**, *258*, 117-125.

288. M. T. Richers; J. M. Amatrudo; J. P. Olson; G. C. R. Ellis-Davies, *Angew. Chem.-Int. Ed.* **2017**, *56* (1), 193-197.

289. R. Cabrera; O. Filevich; B. Garcia-Acosta; J. Athilingam; K. J. Bender; K. E. Poskanzer; R. Etchenique, *ACS Chem. Neurosci.* **2017**, *8* (5), 1036-1042.

290. C. P. McCoy; C. Brady; J. F. Cowley; S. M. McGlinchey; N. McGoldrick; D. J. Kinnear; G. P. Andrews; D. S. Jones, *Expert Opin. Drug. Deliv.* **2010**, *7* (5), 605-616.

291. C. Q. Liu; Y. Zhang; M. Liu; Z. W. Chen; Y. H. Lin; W. Li; F. F. Cao; Z. Liu; J. S. Ren; X. G. Qu, *Biomaterials* **2017**, *139*, 151-162.

292. T. Ito; K. Tanabe; H. Yamada; H. Hatta; S. Nishimoto, *Molecules* **2008**, *13* (10), 2370-2384.

293. M. A. Romero; N. Basilio; A. J. Moro; M. Domingues; J. A. Gonzalez-Delgado; J. F. Arteaga; U. Pischel, *Chem. Eur. J.* **2017**, *23* (53), 13105-13111.

294. C. C. Epley; K. L. Roth; S. Y. Lin; S. R. Ahrenholtz; T. Z. Grove; A. J. Morris, *Dalton Trans.* **2017**, *46* (15), 4917-4922.

295. A. P. Pelliccioli; J. Wirz, *Photochem. Photobiol. Sci.* **2002**, *1* (7), 441-458.

296. A. Dinache; A. Smarandache; A. Simon; V. Nastasa; T. Tozar; A. Pascu; M. Enescu; A. Khatyr; F. Sima; M. L. Pascu; A. Staicu, *Appl. Surf. Sci.* **2017**, *417*, 136-142.

297. K. Marter; J. Wetzel; J. Eichhorst; N. Eremina; G. Leboulle; A. Barth; B. Wiesner; D. Eisenhardt, *ChemistrySelect* **2017**, *2* (22), 6212-6217.

298. V. N. R. Pillai, *Synthesis Stuttg.* **1980**, (1), 1-26.

299. Q. Q. Zhu; W. Schnabel; H. Schupp, *J. Photochem.* **1987**, *39* (2), 317-332.

300. J. W. Walker; G. P. Reid; J. A. McCray; D. R. Trentham, *J. Am. Chem. Soc.* **1988**, *110* (21), 7170-7177.

301. J. A. McCray; D. R. Trentham, *Annu. Rev. Biophys. Biomo.l Struct.* **1989**, *18* (1), 239-270.

302. K. Grygoryeva; J. Kubecka; A. Pysanenko; J. Lengyel; P. Slavicek; M. Farnik, *J. Phys. Chem. A* **2016**, *120* (24), 4139-4146.

303. P. G. Conrad; R. S. Givens; J. F. W. Weber; K. Kandler, *Organic Letters* **2000**, *2* (11), 1545-1547.

304. R. S. Givens; J. F. W. Weber; P. G. Conrad; G. Orosz; S. L. Donahue; S. A. Thayer, *J. Am. Chem. Soc.* **2000**, *122* (12), 2687-2697.

305. G. Papageorgiou; D. C. Ogden; A. Barth; J. E. T. Corrie, *J. Am. Chem. Soc.* **1999**, *121* (27), 6503-6504.

306. S. Loudwig; M. Goeldner, *Tetrahedron Lett.* **2001**, *42* (45), 7957-7959.

307. H. Boudebous; B. Košmrlj; B. Šket; J. Wirz, *J. Phys. Chem. A* **2007**, *111* (15), 2811-2813.

308. J. C. Sheehan; R. M. Wilson; A. W. Oxford, *J. Am. Chem. Soc.* **1971**, *93* (26), 7222-7232.

309. B. Schade; V. Hagen; R. Schmidt; R. Herbrich; E. Krause; T. Eckardt; J. Bendig, *J. Org. Chem.* **1999**, *64* (25), 9109-9117.

310. V. Hagen; J. Bendig; S. Frings; B. Wiesner; B. Schade; S. Helm; D. Lorenz; U. B. Kaupp, *J. Photochem. Photobiol.* **1999**, *53* (1-3), 91-102.
311. A. Z. Suzuki; T. Watanabe; M. Kawamoto; K. Nishiyama; H. Yamashita; M. Ishii; M. Iwamura; T. Furuta, *Org. Lett.* **2003**, *5* (25), 4867-4870.
312. L. M. Herzig; I. Elamri; H. Schwalbe; J. Wachtveitl, *Phys. Chem. Chem. Phys.* **2017**, *19* (22), 14835-14844.
313. R. O. Schonleber; J. Bendig; V. Hagen; B. Giese, *Bioorg. Med. Chem.* **2002**, *10* (1), 97-101.
314. V. Hagen; F. Kilic; J. Schaal; B. Dekowski; R. Schmidt; N. Kotzur, *J. Org. Chem.* **2010**, *75* (9), 2790-2797.
315. T. Furuta; S. S. H. Wang; J. L. Dantzker; T. M. Dore; W. J. Bybee; E. M. Callaway; W. Denk; R. Y. Tsien, *Proc. Natl. Acad. Sci. USA* **1999**, *96* (4), 1193-1200.
316. L. K. A. M. Leal; A. A. G. Ferreira; G. A. Bezerra; F. J. A. Matos; G. S. B. Viana, *J. Ethnopharmacol.* **2000**, *70* (2), 151-159.
317. N. Z. Ballin; A. T. Sørensen, *Food control.* **2014**, *38* (C), 198-203.
318. E. d. S. Marques; D. B. Salles; E. L. Maistro, *Toxicol. Rep.* **2015**, *2*, 268-274.
319. B. G. Lake, *Food Chem. Toxicol.* **1999**, *37* (4), 423-453.
320. K. P. Link, *Circulation.* **1959**, *19* (1), 97-107.
321. N. Farinola; N. Piller, *Lymphat. Res. Biol.* **2005**, *3* (2), 81-86.
322. M. Akram; A. Rashid, *J. Thromb. Thrombolysis* **2017**, *44* (3), 406-411.
323. T. Lomonaco; S. Ghimenti; I. Piga; D. Biagini; M. Onor; R. Fuoco; A. Paolicchi; L. Ruocco; G. Pellegrini; M. G. Trivella; F. Di Francesco, *Microchem. J.* **2018**, *136*, 170-176.
324. D. Popp; C. M. Plugge; S. Kleinsteuber; H. Harms; H. Strauber, *Appl. Environ. Microbiol.* **2017**, *83* (13).
325. Q. Xie; S. X. Li; D. F. Liao; W. Wang; B. Tekwani; H. Y. Huang; A. Ali; J. U. Rehman; K. K. Schrader; S. O. Duke; C. L. Cantrell; D. E. Wedge, *Rec. Nat. Prod.* **2016**, *10* (3), 294-306.
326. M. Lacova; J. Chovancova; V. Konecny, *Chem. Pap.-Chem. Zvesti* **1986**, *40* (1), 121-126.
327. M. Chand; A. Gupta; S. C. Jain, *Heterocycl. Lett.* **2017**, *7* (1), 215-230.
328. S. S. Anufrik; V. V. Tarkovsky; G. G. Sazonko; M. M. Asimov, *Appl. Spectrosc.* **2012**, *79* (1), 46-52.
329. S. S. Anufrik; V. V. Tarkovsky, *Appl. Spectrosc.* **2010**, *77* (5), 640-647.
330. M. A. H. Fichte; X. M. M. Weyel; S. Junek; F. Schafer; C. Herbivo; M. Goeldner; A. Specht; J. Wachtveitl; A. Heckel, *Angew. Chem.-Int. Ed.* **2016**, *55* (31), 8948-8952.
331. O. Tiapko; B. Bacsa; G. G. de la Cruz; T. Glasnov; K. Groschner, *Sci. China-Life Sci.* **2016**, *59* (8), 802-810.
332. Z. H. Gao; P. T. Yuan; D. H. Wang; Z. P. Xu; Z. Li; X. S. Shao, *Bioorg. Med. Chem. Lett.* **2017**, *27* (11), 2528-2535.
333. H. K. Agarwal; S. Y. Zhai; J. Surmeier; G. C. R. Ellis-Davies, *ACS Chem. Neurosci.* **2017**, *8* (10), 2139-2144.
334. Y. A. Kim; J. Day; C. A. Lirette; W. J. Costain; L. J. Johnston; R. Bittman, *Chem. Phys. Lipids* **2016**, *194*, 117-124.
335. A. Nadler; D. A. Yushchenko; R. Muller; F. Stein; S. H. Feng; C. Mulle; M. Carta; C. Schultz, *Nat. Commun.* **2015**, *6*, 123-134.

336. A. Gandioso; M. Cano; A. Massaguer; V. Marchan, *J. Org. Chem.* **2016**, *81* (23), 11556-11564.

337. M. Goard; G. Aakalu; O. D. Fedoryak; C. Quinonez; J. St Julien; S. J. Poteet; E. M. Schuman; T. M. Dore, *Chem. Biol.* **2005**, *12* (6), 685-693.

338. T. Furuta; S. S.-H. Wang; J. L. Dantzker; T. M. Dore; W. J. Bybee; E. M. Callaway; W. Denk; R. Y. Tsien, *Proc. Natl. Acad. Sci. USA* **1999**, *96* (4), 1193-1200.

339. M. M. Mahmoodi; S. A. Fisher; R. Y. Tam; P. C. Goff; R. B. Anderson; J. E. Wissinger; D. A. Blank; M. S. Shoichet; M. D. Distefano, *Org. Biomol. Chem.* **2016**, *14* (35), 8289-8300.

340. K. Katayama; S. Tsukiji; T. Furuta; T. Nagamune, *Chem. Commun.* **2008**, (42), 5399-5401.

341. Q. Nie; X. G. Du; M. Y. Geng, *Acta Pharmacol. Sin.* **2011**, *32* (5), 545-551.

342. T. K. Chakraborty; P. Srinivasu; S. Tapadar; B. K. Mohan, *Glycoconjugate J.* **2005**, *22* (3), 83-93.

343. A. McFedries; A. Schwaid; A. Saghatelian, *Chem. Biol.* **2013**, *20* (5), 667-673.

344. Y. L. Yang; C. Wang, *Philos. Trans. R. Soc. A-Math. Phys. Eng. Sci.* **2013**, *371* (2000), 1-20.

345. J. C. Steffens, *Annu. Rev. Plant Physiol. Plant Molec. Biol.* **1990**, *41*, 553-575.

346. H. B. Kraatz; M. Galka, *F. Matter* **2001**, *38*, 385-409.

347. A. Vallee; V. Humblot; C. M. Pradier, *Acc. Chem. Res.* **2010**, *43* (10), 1297-1306.

348. P. K. Nandi; D. R. Robinson, *J. Am. Chem. Soc.* **1972**, *94* (4), 1299-1312.

349. T. J. Gaborek; J. D. Madura, *Biophys. J.* **2012**, *102* (3), 44A-44A.

350. M. Muttenthaler; P. F. Alewood, *J. Pep. Sci.* **2008**, *14* (12), 1223-1239.

351. L. T. Hung; S. Fermandjian; J. L. Morgat; P. Fromageot, *J. Label. Compd.* **1974**, *10* (1), 1-21.

352. A. J. Pearson; P. Zhang; G. Bignan, *J. Org. Chem.* **1997**, *62* (13), 4536-4538.

353. S. Venkatraman; F. G. Njoroge; V. M. Girijavallabhan; V. S. Madison; N. H. Yao; A. J. Prongay; N. Butkiewicz; J. Pichardo, *J. med. Chem.* **2005**, *48* (16), 5088-5091.

354. C. C. Liu; P. G. Schultz, *Annu. Rev. Biochem.* **2010**, *79* (1), 413-444.

355. D. O'Hagan; C. Schaffrath; S. L. Cobb; J. T. G. Hamilton; C. D. Murphy, *Nature* **2002**, *416* (6878), 279-279.

356. C. J. Dong; F. L. Huang; H. Deng; C. Schaffrath; J. B. Spencer; D. O'Hagan; J. H. Naismith, *Nature* **2004**, *427* (6974), 561-565.

357. D. B. Harper; D. O'Hagan, *Nat. Prod. Rep.* **1994**, *11* (0), 123-133.

358. D. O'Hagan, *Chem. Soc. Rev.* **2008**, *37* (2), 308-319.

359. P. Y. Chou; G. D. Fasman, *Biochemistry* **1974**, *13* (2), 211-222.

360. Y. Zhou; J. Wang; Z. N. Gu; S. N. Wang; W. Zhu; J. L. Acena; V. A. Soloshonok; K. Izawa; H. Liu, *Chem. Rev.* **2016**, *116* (2), 422-518.

361. M. B. van Niel; I. Collins; M. S. Beer; H. B. Broughton; S. K. F. Cheng; S. C. Goodacre; A. Heald; K. L. Locker; A. M. MacLeod; D. Morrison; C. R. Moyes; D. O'Connor; A. Pike; M. Rowley; M. G. N. Russell; B. Sohal; J. A. Stanton; S. Thomas; H. Verrier; A. P. Watt; J. L. Castro, *J. med. Chem.* **1999**, *42* (12), 2087-2104.

362. C. W. Thornber, *Chem. Soc. Rev.* **1979**, *8* (4), 563-580.

363. L. M. A. Lima; E. J. Barreiro, *Curr. Med. Chem.* **2005**, *12* (1), 23-49.

364. A. F. Brooks; J. J. Topczewski; N. Ichiishi; M. S. Sanford; P. J. H. Scott, *Chem. Sci.* **2014**, *5* (12), 4545-4553.

365.	I. Tirotta; V. Dichiarante; C. Pigliacelli; G. Cavallo; G. Terraneo; F. B. Bombelli; P. Metrangolo; G. Resnati, *Chem. Rev.* **2015**, *115* (2), 1106-1129.
366.	J. N. Dahanayake; C. Kasireddy; J. M. Ellis; D. Hildebrandt; O. A. Hull; J. P. Karnes; D. Morlan; K. R. Mitchell-Koch, *J. Comput. Chem.* **2017**, *38* (30), 2605-2617.
367.	C. Odar; M. Winkler; B. Wiltschi, *Biotechnol. J.* **2015**, *10* (3), 427-446.
368.	A. H. Harkiss; A. Sutherland, *Org. Biomol. Chem.* **2016**, *14* (38), 8911-8921.
369.	C. Kasireddy; J. M. Ellis; J. G. Bann; K. R. Mitchell-Koch, *Sci. Rep.* **2017**, *7*, 42651.
370.	J. Ichikawa; R. Nadano; N. Ito, *Chem. Commun.* **2006**, (42), 4425-4427.
371.	J. Lazar; W. A. Sheppard, *J. med. Chem.* **1968**, *11* (1), 138-140.
372.	M. G. Woll; E. B. Hadley; S. Mecozzi; S. H. Gellman, *J. Am. Chem. Soc.* **2006**, *128* (50), 15932-15933.
373.	G. Bott; L. D. Field; S. Sternhell, *J. Am. Chem. Soc.* **1980**, *102* (17), 5618-5626.
374.	S. A. Samsonov; M. Salwiczek; G. Anders; B. Koksch; M. T. Pisabarro, *J. Phys. Chem. B* **2009**, *113* (51), 16400-16408.
375.	B. E. Smart, *J. Fluor. Chem.* **2001**, *109* (1), 3-11.
376.	A. A. Berger; J.-S. Völler; N. Budisa; B. Koksch, *Acc. Chem. Res.* **2017**, *50* (9), 2093-2103.
377.	N. C. Yoder; K. Kumar, *Chem. Soc. Rev.* **2002**, *31* (6), 335-341.
378.	Y. Tang; G. Ghirlanda; W. A. Petka; T. Nakajima; W. F. DeGrado; D. A. Tirrell, *Angew. Chem.-Int. Ed.* **2001**, *40* (8), 1494-1496.
379.	B. C. Buer; E. N. G. Marsh, *Prot. Sci.* **2012**, *21* (4), 453-462.
380.	H. Erdbrink; E. K. Nyakatura; S. Huhmann; U. I. M. Gerling; D. Lentz; B. Koksch; C. Czekelius, *Beilstein J. Org. Chem.* **2013**, *9*, 2009-2014.
381.	J. T. Anderson; P. L. Toogood; E. N. G. Marsh, *Org. Lett.* **2002**, *4* (24), 4281-4283.
382.	X. Xing; A. Fichera; K. Kumar, *Org. Lett.* **2001**, *3* (9), 1285-1286.
383.	Y. H. Zhao; M. H. Abraham; A. M. Zissimos, *J. Org. Chem.* **2003**, *68* (19), 7368- 7373.
384.	M. Salwiczek; S. Samsonov; T. Vagt; M. T. Pisabarro; B. Koksch, *J. Pep. Sci.* **2008**, *14* (8), 4-4.
385.	D. Klemczak. The Role of 4-Itf-Phe in Amyloid Formation: A Model Peptide Study Master Thesis, Freie Universität Berlin, **2017**.
386.	H.-P. Chiu; B. Kokona; R. Fairman; R. P. Cheng, *J. Am. Chem. Soc.* **2009**, *131* (37), 13192-13193.
387.	B. C. Buer; R. de la Salud-Bea; H. M. Al Hashimi; E. N. G. Marsh, *Biochemistry* **2009**, *48* (45), 10810-10817.
388.	H.-Y. Lee; K.-H. Lee; H. M. Al-Hashimi; E. N. G. Marsh, *J. Am. Chem. Soc.* **2006**, *128* (1), 337-343.
389.	L. M. Gottler; H.-Y. Lee; C. E. Shelburne; A. Ramamoorthy; E. N. G. Marsh, *ChemBioChem* **2008**, *9* (3), 370-373.
390.	H. Meng; K. Kumar, *J. Am. Chem. Soc.* **2007**, *129* (50), 15615-15622.
391.	H. Meng; S. T. Krishnaji; M. Beinborn; K. Kumar, *J. med. Chem.* **2008**, *51* (22), 7303-7307.
392.	I. Schechter; A. Berger, *Biochem. Biophys. Res. Commun.* **1967**, *27* (2), 157-162.
393.	A. S. Ito; R. D. F. Turchiello; I. Y. Hirata; M. H. S. Cezari; M. Meldal; L. Juliano, *Biospectroscopy* **1998**, *4* (6), 395-402.
394.	A. Siuiti Ito; E. Sérgio de Souza; S. dos Reis Barbosa; C. Ryuichi Nakaie, *Biophys. J.* **2001**, *81* (2), 1180-1189.

395. X. S. Puente; L. M. Sanchez; C. M. Overall; C. Lopez-Otin, *Nat. Rev. Genet.* **2003**, *4* (7), 544-558.

396. A. G. McDonald; S. Boyce; K. F. Tipton, *Nucleic Acids Res.* **2009**, *37* (1), 593-597.

397. N. D. Rawlings; A. J. Barrett, *Biochem. J.* **1993**, *290* (Pt 1), 205-218.

398. N. D. Rawlings; F. R. Morton; C. Y. Kok; J. Kong; A. J. Barrett, *Nucleic Acids Res.* **2008**, *36*, 320-325.

399. Y. A. DeClerck; A. M. Mercurio; M. S. Stack; H. A. Chapman; M. M. Zutter; R. J. Muschel; A. Raz; L. M. Matrisian; B. F. Sloane; A. Noel; M. J. Hendrix; L. Coussens; M. Padarathsingh, *Am. J. Pathol.* **2004**, *164* (4), 1131-1139.

400. D. C. Whitcomb; M. E. Lowe, *Dig. Dis. Sci.* **2007**, *52* (1), 1-17.

401. M. L. Bender; F. J. Kezdy, *Annu. Rev. Biochem.* **1965**, *34* (1), 49-76.

402. S. Scheiner; W. N. Lipscomb, *Proc. Natl. Acad. Sci. USA* **1976**, *73* (2), 432-436.

403. L. Polgár, *Cell Mol Life Sci.* **2005**, *62* (19), 2161-2172.

404. L. Hedstrom, *Chem. Rev.* **2002**, *102* (12), 4501-4524.

405. I. V. Kashparov; M. E. Popov; E. M. Popov, Mechanism of Action of Aspartic Proteases. In *Aspartic Proteinases: Retroviral and Cellular Enzymes*, M. N. G. James, Ed. Springer US: Boston, MA, **1998**; pp 115-121.

406. L. Polgár, *FEBS Lett.* **1987**, *219* (1), 1-4.

407. J. Eder; U. Hommel; F. Cumin; B. Martoglio; B. Gerhartz, *Curr. Pharm. Design* **2007**, *13* (3), 271-285.

408. I. Mandl, [92] Pancreatic elastase. In *Methods in Enzymology*, Academic Press: **1962**; Vol. 5, pp 665-673.

409. U. J. Lewis; D. E. Williams; N. G. Brink, *J. Biol. Chem.* **1956**, *222* (2), 705-720.

410. C. Ghelis; M. Tempete-Gaillourdet; J. M. Yon, *Biochem. Biophys. Res. Commun.* **1978**, *84* (1), 31-36.

411. M. Zimmerman; B. M. Ashe, *Biochim. Biophys. Acta Enzym.* **1977**, *480* (1), 241-245.

412. A. Renaud; P. Lestienne; D. L. Hughes; J. G. Bieth; J. L. Dimicoli, *J. Biol. Chem.* **1983**, *258* (13), 8312-8316.

413. J. Carrere; C. Figarella; O. Guy; J. P. Thouvenot, *Biochim. Biophys. Acta* **1986**, *883* (1), 46-53.

414. W. H. Cruickshank; H. Kaplan, *Biochem. J.* **1975**, *147* (3), 411-416.

415. D. M. Blow; J. J. Birktoft; B. S. Hartley, *Nature* **1969**, *221* (5178), 337-340.

416. H. Czapinska; J. Otlewski, *Eur. J. Biochem.* **1999**, *260* (3), 571-595.

417. K. Brady; A. Z. Wei; D. Ringe; R. H. Abeles, *Biochemistry* **1990**, *29* (33), 7600-7607.

418. W. Ebeling; N. Hennrich; M. Klockow; H. Metz; H. D. Orth; H. Lang, *Eur. J. Biochem.* **1974**, *47* (1), 91-97.

419. C. Betzel; T. P. Singh; M. Visanji; K. Peters; S. Fittkau; W. Saenger; K. S. Wilson, *J. Biol. Chem.* **1993**, *268* (21), 15854-15858.

420. A. Müller; W. Hinrichs; W. M. Wolf; W. Saenger, *J. Biol. Chem.* **1994**, *269* (37), 23108-23111.

421. D. Georgieva; N. Genov; W. Voelter; C. Betzel, *Z. Naturforsch.(C)* **2006**, *61* (5-6), 445-452.

422. H. Hilz; U. Wiegers; P. Adamietz, *Eur. J. Biochem.* **1975**, *56* (1), 103-108.

423. Z. Bohak, *J. Biol. Chem.* **1969**, *244* (17), 4638-4648.

424. G. P. Sachdev; J. S. Fruton, *Proc. Natl. Acad. Sci. USA* **1975**, *72* (9), 3424-3427.

425. L. E. Orgel; R. Lohrmann, *Acc. Chem. Res.* **1974**, *7* (11), 368-377.

426. B. M. Dunn, *Chem. Rev.* **2002**, *102* (12), 4431-4458.

427. Y. Hamuro; S. J. Coales; K. S. Molnar; S. J. Tuske; J. A. Morrow, *Rapid Commun. Mass Spectrom.* **2008**, *22* (7), 1041-1046.

428. M. Fujinaga; M. M. Chernaia; S. C. Mosimann; M. N. G. James; N. I. Tarasova, *Prot. Sci.* **1995**, *4* (5), 960-972.

429. J. L. Glajch; J. J. Kirkland; J. Köhler, *J. Chrom. A* **1987**, *384* (Supplement C), 81-90.

430. K. Kalghatgi; C. Horváth, *J. Chrom. A* **1987**, *398* (Supplement C), 335-339.

431. T. E. Young; S. T. Ecker; R. E. Synovec; N. T. Hawley; J. P. Lomber; C. M. Wai, *Talanta* **1998**, *45* (6), 1189-1199.

432. H. Engelhardt; S. Kromidas, *Naturwissenschaften* **1980**, *67* (7), 353-354.

433. U. D. Neue; M. Kele; B. Bunner; A. Kromidas; T. Dourdeville; J. R. Mazzeo; E. S. Grumbach; S. Serpa; T. E. Wheat; P. Hong; M. Gilar, Ultra-Performance Liquid Chromatography Technology and Applications. In *Advances in Chromatography, Vol 48*, E. Grushka; N. Grinberg, Eds. **2010**; Vol. 48, pp 99-143.

434. P. Hernandez; M. Müller; R. D. Appel, *Mass Spectrom. Rev.* **2006**, *25* (2), 235-254.

435. R. L. Gundry; M. Y. White; C. I. Murray; L. A. Kane; Q. Fu; B. A. Stanley; J. E. Van Eyk, *Preparation of Proteins and Peptides for Mass Spectrometry Analysis in a Bottom-Up Proteomics Workflow*. Current protocols in molecular biology **2009**; p 10.25-10.25.

436. S. A. Trauger; W. Webb; G. Siuzdak, *Spectroscopy* **2002**, *16* (1).

437. C. S. Johnson, *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34* (3), 203-256.

438. R. Neufeld; D. Stalke, *Chem. Sci.* **2015**, *6* (6), 3354-3364.

439. B. Fürtig; C. Richter; J. Wöhnert; H. Schwalbe, *ChemBioChem* **2003**, *4* (10), 936- 962.

440. N. J. Greenfield, *Nat. Protoc.* **2006**, *1* (6), 2876-2890.

441. S. R. Martin; M. J. Schilstra, Circular Dichroism and Its Application to the Study of Biomolecules. In *Methods in Cell Biology*, Academic Press: **2008**; Vol. 84, pp 263-293.

442. N. J. Greenfield, *Nat. Protoc.* **2007**, *1*, 2891.

443. N. Darghal; A. Garnier-Suillerot; M. Salerno, *Biochem. Biophys. Res. Commun.* **2006**, *343* (2), 623-629.

444. S. L. Gras; L. J. Waddington; K. N. Goldie, Transmission Electron Microscopy of Amyloid Fibrils. In *Protein Folding, Misfolding, and Disease: Methods and Protocols*, A. F. Hill; K. J. Barnham; S. P. Bottomley; R. Cappai, Eds. Humana Press: Totowa, NJ, **2011**; pp 197-214.

445. M. R. Nilsson, *Methods* **2004**, *34* (1), 151-160.

446. C. L. P. Oliveira; J. S. Pedersen, Structures of Aggregating Species by Small-Angle X-Ray Scattering. In *Amyloid Fibrils and Prefibrillar Aggregates*, Wiley-VCH Verlag GmbH & Co. KGaA: **2013**; pp 85-102.

447. A. E. Langkilde; B. Vestergaard, *FEBS Lett.* **2009**, *583* (16), 2600-2609.

448. A. E. Langkilde; K. L. Morris; L. C. Serpell; D. I. Svergun; B. Vestergaard, *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **2015**, *71* (Pt 4), 882-895.

449. D. Eshan; C. Mina; A. Nadia; A. B. Ashwinkumar; L. B. Serge; B. Aldo, *Phys. Biol.* **2017**, *14* (4), 046001.

450. A. R. von Gundlach; V. M. Garamus; T. Gorniak; H. A. Davies; M. Reischl; R. Mikut; K. Hilpert; A. Rosenhahn, *Biochim. Biophys. Acta Biomem.* **2016**, *1858* (5), 918-925.

451. L. Malik; J. Nygaard; N. J. Christensen; W. W. Streicher; P. W. Thulstrup; L. Arleth; K. J. Jensen, *J. Pep. Sci.* **2013**, *19* (5), 283-292.
452. J. M. D. Kalapothakis; R. J. Morris; J. Szavits-Nossan; K. Eden; S. Covill; S. Tabor; J. Gillam; P. E. Barran; R. J. Allen; C. E. MacPhee, *Biophys. J.* **2015**, *108* (9), 2300-2311.
453. N. J. Economou; M. J. Giammona; T. D. Do; X. Y. Zheng; D. B. Teplow; S. K. Buratto; M. T. Bowers, *J. Am. Chem. Soc.* **2016**, *138* (6), 1772-1775.
454. J. S. Pedersen, *J. Colloid Interface Sci.* **1997**, *70*, 171-210.
455. A. Kuchibhatla; A. S. A. Rasheed; J. Narayanan; J. Bellare; D. Panda, *Langmuir* **2009**, *25* (6), 3775-3785.
456. O. Glatter; O. Kratky, *Small Angle X-ray Scattering*. Academic Press: London, **1982**.
457. O. Glatter, *J. Appl. Crystallogr.* **1980**, *13* (DEC), 577-584.
458. H. Schnablegger; M. Antonietti; C. Goltner; J. Hartmann; H. Colfen; P. Samori; J. P. Rabe; H. Hager; W. Heitz, *J. Colloid Interface Sci.* **1999**, *212* (1), 24-32.
459. H. Schnablegger; M. Antonietti; C. Goltner; I. H. Stapff; F. Brink-Spalink; A. Greiner, *Acta Polym.* **1999**, *50* (11-12), 391-398.
460. R. Sawant; V. Torchilin, *Mol. Biosyst.* **2010**, *6* (4), 628-640.
461. M. C. Morris; J. Depollier; J. Mery; F. Heitz; G. Divita, *Nat. Biotechnol.* **2001**, *19* (12), 1173-1176.
462. M. Salwiczek; E. K. Nyakatura; U. I. M. Gerling; S. J. Ye; B. Koksch, *Chem. Soc. Rev.* **2012**, *41* (6), 2135-2171.
463. K. Koschek; M. Dathe; J. Rademann, *ChemBioChem* **2013**, *14* (15), 1982-1990.
464. I. R. Ruttekolk; F. Duchardt; R. Fischer; K. H. Wiesmuller; J. Rademann; R. Brock, *Bioconjugate Chem.* **2008**, *19* (10), 2081-2087.
465. K. Koschek. *Darstellung multivalenter Peptid-Polymerkonjugate ausgehend von Thioester-HPMA-Copolymeren zur Beantwortung biologischer Fragestellungen in vitro und in vivo*. Dissertation, Freie Universität Berlin, **2012**.
466. L. Bourel; O. Carion; H. Gras-Masse; O. Melnyk, *J. Pep. Sci.* **2000**, *6* (6), 264-270.
467. E. Brandenburg. Inhibition der Bildung amyloider Aggregate. Dissertation, Freie Universität Berlin, **2012**.
468. D. D. Soto-Ortega; B. P. Murphy; F. J. Gonzalez-Velasquez; K. A. Wilson; F. Xie; Q. A. Wang; M. A. Moss, *Bioorg. Med. Chem.* **2011**, *19* (8), 2596-2602.
469. G. W. Anderson; J. E. Zimmerman; F. M. Callahan, *J. Am. Chem. Soc.* **1964**, *86* (9), 1839-1842.
470. C. A. Ventura; C. Cannavà; R. Stancanelli; D. Paolino; D. Cosco; A. La Mantia; R. Pignatello; S. Tommasini, *Biomed. Microdevices* **2011**, *13* (5), 799-807.
471. M. Dittmann; M. Sadek; R. Seidel; M. Engelhard, *J. Pep. Sci.* **2012**, *18* (5), 312-316.
472. M. Dittmann; J. Sauermann; R. Seidel; W. Zimmermann; M. Engelhard, *J. Pep. Sci.* **2010**, *16* (10), 558-562.
473. A. B. Clippingdale; C. J. Barrow; J. D. Wade, *J. Pep. Sci.* **2000**, *6* (5), 225-234.
474. B. H. Northrop; S. H. Frayne; U. Choudhary, *Polym. Chem.* **2015**, *6* (18), 3415-3430.
475. M. Adamczyk; M. Cornwell; J. Huff; S. Rege; T. V. S. Rao, *Bioorg. Med. Chem. Lett.* **1997**, *7* (15), 1985-1988.
476. P. L. Khanna; E. F. Ullman, *Anal. Biochem.* **1980**, *108* (1), 156-161.
477. L. J. Nielsen; S. Eyley; W. Thielemans; J. W. Aylott, *Chem. Commun.* **2010**, *46* (47), 8929-8931.

478. M. Ogawa; N. Kosaka; P. L. Choyke; H. Kobayashi, *ACS Chem. Biol.* **2009**, *4* (7), 535-546.

479. K. Umezawa; A. Matsui; Y. Nakamura; D. Citterio; K. Suzuki, *Chem. Eur. J.* **2009**, *15* (5), 1096-1106.

480. J. Y. Rho; J. C. Brendel; L. R. MacFarlane; E. D. H. Mansfield; R. Peltier; S. Rogers; M. Hartlieb; S. Perrier, *Adv. Funct. Mater* **2017**, 1704569-1704580.

481. L. Li; J. Yang; J. Wang; J. Kopeček, *Macromol. Biosci.* **2017**, 1700196-1700209.

482. C. M. Madl; B. L. LeSavage; R. E. Dewi; C. B. Dinh; R. S. Stowers; M. Khariton; K. J. Lampe; D. Nguyen; O. Chaudhuri; A. Enejder; S. C. Heilshorn, *Nat. Mater.* **2017**, *16*, 1233-1242.

483. E. Brandenburg; H. von Berlepsch; U. I. M. Gerling; C. Boettcher; B. Koksch, *Chem. Eur. J.* **2011**, *17* (38), 10650-10660.

484. H. Le Chatelier, O. Boudouard, *Bull. Soc. Chim. Fr* **1898**, *19*, 483–488.

485. M. Miocque, *Ann. Pharm. Fr.* **1969**, *27* (5), 381-&.

486. M. Ramanathan; D.-R. Hou, *Tetrahedron Lett.* **2010**, *51* (47), 6143-6145.

487. M. M. Mahmoodi; D. Abate-Pella; T. J. Pundsack; C. C. Palsuledesai; P. C. Goff; D. A. Blank; M. D. Distefano, *J. Am. Chem. Soc.* **2016**, *138* (18), 5848-5859.

488. P. Dey; S. Hemmati-Sadeghi; R. Haag, *Polym. Chem.* **2016**, *7* (2), 375-383.

489. E. H. P. Leunissen; M. H. L. Meuleners; J. M. M. Verkade; J. Dommerholt; J. G. J. Hoenderop; F. L. van Delft, *ChemBioChem* **2014**, *15* (10), 1446-1451.

490. A. Gandioso; S. Contreras; I. Melnyk; J. Oliva; S. Nonell; D. Velasco; J. Garcia-Amoros; V. Marchan, *J. Org. Chem.* **2017**, *82* (10), 5398-5408.

491. C. Largman; J. W. Brodrick; M. C. Geokas, *Biochemistry* **1976**, *15* (11), 2491-2500.

492. V. K. Antonov; L. M. Ginodman; Y. V. Kapitannikov; T. N. Barshevskaya; A. G. Gurova; L. D. Rumsh, *FEBS Lett.* **1978**, *88* (1), 87-90.

493. P. J. Sweeney; J. M. Walker, Proteinase K (EC 3.4.21.14). In *Enzymes of Molecular Biology*, M. M. Burrell, Ed. Humana Press: Totowa, NJ, **1993**; pp 305-311.

494. H.-J. Böhm; D. Banner; S. Bendels; M. Kansy; B. Kuhn; K. Müller; U. Obst-Sander; M. Stahl, *ChemBioChem* **2004**, *5* (5), 637-643.

495. J. A. Olsen; D. W. Banner; P. Seiler; B. Wagner; T. Tschopp; U. Obst-Sander; M. Kansy; K. Müller; F. Diederich, *ChemBioChem* **2004**, *5* (5), 666-675.

496. M. Rowley; D. J. Hallett; S. Goodacre; C. Moyes; J. Crawforth; T. J. Sparey; S. Patel; R. Marwood; S. Patel; S. Thomas; L. Hitzel; D. O'Connor; N. Szeto; J. L. Castro; P. H. Hutson; A. M. MacLeod, *J. med. Chem.* **2001**, *44* (10), 1603-1614.

497. J. D. Dunitz; R. Taylor, *Chem. Eur. J.* **1997**, *3* (1), 89-98.

498. J. A. K. Howard; V. J. Hoy; D. O'Hagan; G. T. Smith, *Tetrahedron* **1996**, *52* (38), 12613-12622.

499. V. R. Thalladi; H.-C. Weiss; D. Bläser; R. Boese; A. Nangia; G. R. Desiraju, *J. Am. Chem. Soc.* **1998**, *120* (34), 8702-8710.

500. G. Gerebtzoff; X. Li-Blatter; H. Fischer; A. Frentzel; A. Seelig, *ChemBioChem* **2004**, *5* (5), 676-684.

501. G. Rose; A. Geselowitz; G. Lesser; R. Lee; M. Zehfus, *Science* **1985**, *229* (4716), 834-838.

502. N. T. Southall; K. A. Dill; A. D. J. Haymet, *J. Phys. Chem. B* **2002**, *106* (3), 521-533.

503. K. B. Park; N. R. Kitteringham; P. M. O'Neill, *Annu. Rev. Pharmacol. Toxicol.* **2001**, *41* (1), 443-470.

504. J. D. Dunitz, *ChemBioChem* **2004**, *5* (5), 614-621.

505. J. C. Biffinger; H. W. Kim; S. G. DiMagno, *ChemBioChem* **2004**, *5* (5), 622-627.

506. T. J. Barbarich; C. D. Rithner; S. M. Miller; O. P. Anderson; S. H. Strauss, *J. Am. Chem. Soc.* **1999**, *121* (17), 4280-4281.

507. K.-H. Lee; H.-Y. Lee; M. M. Slutsky; J. T. Anderson; E. N. G. Marsh, *Biochemistry* **2004**, *43* (51), 16277-16284.

508. M. Bruncko; T. K. Oost; B. A. Belli; H. Ding; M. K. Joseph; A. Kunzer; D. Martineau; W. J. McClellan; M. Mitten; S.-C. Ng; P. M. Nimmer; T. Oltersdorf; C.-M. Park; A. M. Petros; A. R. Shoemaker; X. Song; X. Wang; M. D. Wendt; H. Zhang; S. W. Fesik; S. H. Rosenberg; S. W. Elmore, *J. med. Chem.* **2007**, *50* (4), 641-662.

509. D. Orthaber; A. Bergmann; O. Glatter, *J. Appl. Crystallogr.* **2000**, *33*, 218-225.