

**Fluorine-specific interactions as an  
orthogonal tool in protein folding:  
Laying the groundwork by chemical synthesis of  
small proteins**

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry, Pharmacy  
of Freie Universität Berlin

by

VALENTINA STULBERG

from Pavlodar, Kazakhstan

January 2023



1<sup>st</sup> Reviewer: Prof. Dr. Beate Kokschi (Freie Universität Berlin)  
2<sup>nd</sup> Reviewer: Prof. Dr. Nediljko Budisa (University of Manitoba)

Date of defense: 28.06.2023

---

---

## **Declaration**

The research presented in this doctoral thesis was performed under the supervision of Prof. Dr. Beate Kokschi from April 2015 until December 2022 at the Institute of Chemistry and Biochemistry in the Department of Biology, Chemistry, and Pharmacy of Freie Universität Berlin.

I hereby declare that I have prepared this dissertation entitled “Fluorine-specific interactions as an orthogonal tool in protein folding: Laying the groundwork by chemical synthesis of small proteins” autonomously and without impermissible help. All external sources and resources have been specified and properly cited or acknowledged. This thesis has not been submitted, accepted, rated as insufficient, or rejected in any other doctorate degree procedure.

Berlin, January 2023

Valentina Stulberg

---

## Publications

- (1) K. Ataka, J. Drauschke, **V. Stulberg**, B. Kokschi, J. Heberle, *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2022**, 1864, 183873. (DOI: doi.org/10.1016/j.bbamem.2022.183873)
- (2) J. Moschner, **V. Stulberg**, R. Fernandes, S. Huhmann, J. Leppkes, B. Kokschi, Approaches to Obtaining Fluorinated  $\alpha$ -Amino Acids. *Chem. Rev.* **2019**, 119, 10718-10801. (DOI: doi.org/10.1021/acs.chemrev.9b00024)

## Oral Presentations

- (1) “Effect of fluorine interactions on protein folding”, 2<sup>nd</sup> Network-Meeting of Collaborative Research Center (CRC) 1349 “Fluorine-Specific Interactions”, Berlin, Germany 27/8 - 28/8/**2020**.
- (2) “Synthese fluorierter Ubiquitin-Analoga”, 18<sup>th</sup> German Fluorine Days, Schmitten, Germany, 17/9 - 19/9/**2018**.
- (3) “Synthesis of Fluorinated Ubiquitin Variants”, 68<sup>th</sup> Seminar of Kollegseminar of the research training group 1582/2 “Fluorine as a Key Element”, Berlin, Germany, 27/4/**2018**.
- (4) “Synthesis of Fluorinated Ubiquitin Variants”, 9<sup>th</sup> Ph.D. workshop of the research training group 1582 “Fluorine as key element,” Potsdam, Germany, 28/03 – 29/03/**2018**.

## Poster Presentations

- (1) “Impact of Fluorine-Specific Interactions on Protein Folding”, 20<sup>th</sup> European Symposium on Fluorine Chemistry, Berlin, Germany, 14/08 – 19/08/**2022**.
- (2) “Impact of Fluorine-Specific Interactions on Protein Folding”, 19<sup>th</sup> European Symposium on Fluorine Chemistry, Warsaw, Poland, 25/08 – 31/08/**2019**.
- (3) “Impact of Fluorine-specific on Protein Folding”, 8<sup>th</sup> Chemical Protein Synthesis meeting, Berlin, Germany 16/6 - 19/6/**2019**
- (4) “Synthesis of Fluorinated Ubiquitin Variants”, 8<sup>th</sup> Peptide Engineering Meeting (PEM-2018), Berlin, Germany, 08/11–10/11/**2018**.

- 
- (5) *“Synthesis of Fluorinated Ubiquitin Variants”*, 35<sup>th</sup> European Peptide Symposium, Dublin, Ireland, 26/8 - 31/8/**2018**.
- (6) *“Peptide-based structure-activity studies on G-protein coupled receptor 83 for elucidation of the signaling regulation mechanism”*, 13<sup>th</sup> German Peptide Symposium, Erlangen, Germany 20/03 – 23/03/**2017**.

---

## Acknowledgments

First of all, I want to thank Prof. Dr. Beate Kokschi for her support throughout my doctoral studies. In particular, I thank her for her trust, which enabled me to pursue my own ideas and paths in research. I am grateful to her for encouraging and guiding me whenever necessary and allowing me to grow and develop as a scientist. In addition, I am very grateful to her because she has enabled me to collaborate with other interdisciplinary scientists and share my ideas at national and international conferences and meetings.

I want to thank Prof. Dr. Nediljko Budisa for taking over the second supervision of this dissertation.

I would also like to thank the current and former members of the Kokschi group for the pleasant working atmosphere and the numerous inspiring discussions. In particular, I want to express my gratitude to Kristin and Susanne, who warmly welcomed me into the team and introduced me to the field of peptide synthesis. I want to thank Johann and Allison, who have always been mentors of the group and always had an open ear and advice. I also thank Allison and Tobias for proofreading the present work. I thank Alex, Thomas, and Suvrat for providing some fluorinated amino acids or the starting material thereof.

To my students Nils, Suvrat, Abdul, Yifu, Michael B., Eunice, and Jana, I express my deep gratitude for their participation in my research work during their internships. Thank you for allowing me to work out new research ideas with you and for allowing me to grow with you.

I want to thank Michael K. and Dr. Christian Roth of the Max Planck Institute for facilitating crystallization studies of 1FYN and working out new analytical strategies such as SDS-PAGE and IEC. I want to thank Dr. Kenichi Ataka for performing *in situ* SEIRAS measurements. Thanks to Patricia for conducting AFM measurements and to Dr. Carsten Grötzinger for studying my GPR83 samples in various biological assays. Thanks to Prof. Dr. Daniel Varón Silva for the opportunity to use the microwave synthesizer under the supervision of Dana at Max Planck Institute



---

Potsdam for the initial studies with this technology. I would also like to thank the members of the Core facility of the Freie Universität Berlin for measuring my samples.

I thank the CEM staff Sandeep, Shaina, Monika, and Sören for their support in developing new protocols using MW-SPPS.

For the financial support, I thank the Deutsche Forschungsgemeinschaft within the Collaborative Research Center (CRC) 1349 "Fluorine-Specific Interactions." I would also like to thank the Center for International Cooperation (CIC) of the Freie Universität Berlin for funding my research stay in Matthews (NC, USA) and the Frauenförderung of the Freie Universität Berlin for supporting my participation in the German Fluorine Day in Schmitten.

My deepest gratitude goes to my family and friends, who have supported me all this time and stood by me, especially after my surgical procedures. Without you, none of this would have been possible. I am forever thankful to know you by my side.

---

## Kurzzusammenfassung

Proteine sind an allen lebenswichtigen Prozessen in Zellen beteiligt. Als Werkzeuge der Zelle besitzen sie regulatorische oder strukturgebende Funktionen. Im *Protein und Peptide Engineering* versucht man die Funktionsweise von Proteinen besser zu verstehen oder die physikochemischen Eigenschaften von Peptiden und Proteinen über nicht-natürliche Aminosäuren mit neu-artigen funktionellen Gruppen zu modifizieren. In der vorliegenden Arbeit wurden beide Aspekte untersucht.

Der erste Teil der vorliegenden Arbeit befasst sich mit den unterschiedlichen Anwendungsmöglichkeiten von Fluormodifikationen in Peptiden und Proteinen. Das Element Fluor kommt in natürlichen biologischen Systemen nur sehr selten vor. Doch die einzigartigen Eigenschaften des Fluors, wie die stärkste Elektronegativität bei gleichzeitig äußerst geringer Polarisierbarkeit, führten nicht nur in der pharmazeutischen und Agrarchemie zu einer weitverbreiteten Anwendung. Im *Protein und Peptide Engineering* bringt eine Fluormodifikation nicht selten überraschende Ergebnisse zum Vorschein.

In gezielt zugeschnitten Systemen sollte die Einführung des Elements Fluors Form fluorierter Aminosäuren in der vorliegenden Doktorarbeit zum einen als Sonde in unterschiedlichen kleineren Peptidmodellen mit neuen Methoden (*in situ surface-enhanced infrared absorption spectroscopy (SEIRAS)*,  $^{19}\text{F}$ -MRI (*magnetic resonance imaging*) sowie AFM (*Atomic force microscopy*) untersucht werden sollen und zum anderen als orthogonales Werkzeug zur Steuerung der Proteinfaltungskinetik erforscht werden.

Die peptidischen Systeme umfassten Sequenzlängen von bis zu 90 Aminosäuren (AAs) in linearer Anordnung oder auch Konjugat-Konstrukte aus einem peptidischen Gerüst (26 AAs), welches an der Seitenkette mit einem weiteren Peptid (bis zu 21 AAs) dekoriert worden ist. Für diese komplexen nativen Systeme wurde jeweils erfolgreich ein chemischer Ansatz entwickelt, der aus einem Ensemble von geeigneten Harzen, Kopplungsbedingungen, der Verwendung spezifischer Bausteine und der Auswahl des richtigen TFA-Spaltungscocktails besteht. Darüber hinaus wurde eine universelle Methode erfolgreich ausgearbeitet, mit deren Hilfe fluoridierte Aminosäuren in solche komplexen Systeme zuverlässig eingebaut werden können.

---

Im zweiten Projekt wurde an der Aufklärung der regulatorischen Funktionsweise des G-protein gekoppelten Rezeptors 83 (GPR83) geforscht. Dieser Rezeptor wird in Gehirnregionen exprimiert, die unter anderem für den Energiemetabolismus, räumliches Lernen sowie Stressregulation verantwortlich sind. Zu Beginn der vorliegenden Arbeit galt der Rezeptor als „*orphan*“, da sein endogener Ligand unbekannt war. Die Studienlage wies daraufhin, dass die extrazelluläre N-terminale Domäne (eNDo) von GPR83 die regulatorischen Prozesse als „*inverser agonist*“ beeinflusst. In der vorliegenden Arbeit wurde dieser Bereich mithilfe einer Peptidbibliothek aus 13 Mitgliedern unterschiedlicher Länge (10 bis 56 AAs) mit einem *overlapping* Design nachgebildet. Mithilfe von Circular dichroism-Messungen erfolgte die Bestimmung der Sekundärstruktur der jeweiligen Peptide. Der zwischenzeitlich publizierte endogene Ligand des GPR83, hPEN, wurde ebenfalls im Rahmen der vorliegenden Arbeit synthetisiert und eine Co-Supplementation mit den Mitgliedern der Peptidbibliothek in biologischen Studien geplant.

In den zahlreichen biologischen Assays war es jedoch weder möglich, eine Rezeptoraktivität durch hPEN noch durch die synthetischen Abschnitte der eNDo nachzuweisen. Weitere Forschungsarbeiten an dem Projekt wurden daher abgebrochen.

---

## Abstract

Proteins are involved in all vital processes in cells. As tools of the cell, they have regulatory or structure-giving functions. In protein and peptide engineering, attempts are made to understand how these biological systems function or to modify the physicochemical properties of peptides and proteins via non-natural amino acids with novel functional groups. In the present work, both aspects were investigated.

The present work's first part deals with various fluorine modification applications in peptides and proteins. The element fluorine is very rarely found in natural biological systems. However, the unique properties of fluorine, such as the strongest electronegativity combined with extremely low polarizability, have led to its widespread application not only in pharmaceutical and agricultural chemistry. In protein and peptide engineering, fluorine modification often yields surprising results.

In specifically tailored systems, the introduction of the element fluorine by applying fluorinated amino acids in the present doctoral thesis should be investigated both as a probe in different smaller peptide models using new methods (*in situ* surface-enhanced infrared absorption spectroscopy (SEIRAS),  $^{19}\text{F}$ -MRI (*magnetic resonance imaging*) as well as AFM (*atomic force microscopy*) and on the other hand as an orthogonal tool to control protein folding kinetics.

The peptidic systems investigated here include sequence lengths of up to 90 amino acids (AAs) in a linear arrangement or conjugate constructs consisting of a peptidic scaffold (26 AAs) decorated on the side chain with another peptide (up to 21 AAs). For each of these complex native systems, a chemical approach has been successfully established from an ensemble of the suitable resin, coupling conditions, use of specific building blocks, and selection of the TFA cleavage cocktail. Moreover, a universal method has been successfully elaborated to incorporate fluorinated amino acids into such complex systems reliably.

In the second project, research was conducted to elucidate the regulatory function of G-protein coupled receptor 83 (GPR83). This receptor is expressed in brain regions responsible for energy metabolism, spatial learning, and stress regulation. At the beginning of the present work, the receptor was considered an "orphan" because its endogenous ligand was unknown. Studies indicated that the

---

extracellular N-terminal domain (eNDo) influences the regulatory processes as an "*inverse agonist*". In the present work, this domain was mapped using a peptide library consisting of 13 members of varying lengths (10 to 56 AAs) with an overlapping design. Circular dichroism measurements were used to determine the secondary structure of the respective peptides. The meanwhile published endogenous ligand of GPR83, hPEN, was also synthesized in the present work, and a co-supplementation with the peptide library members was planned. However, through numerous biological assays, it was not possible to detect receptor activity by either hPEN or the synthetic portions of eNDo. Further research on the project was therefore terminated.

---

## Symbols and Abbreviations

<b>Å</b>	Angstrom (1 Å = 10 <sup>-10</sup> m or 0.1 nm)
<b>AA</b>	amino acid(s)
<b>Ac</b>	acetyl
<b>Ac<sub>2</sub>O</b>	acetic anhydride
<b>ACN</b>	acetonitrile
<b>Boc</b>	<i>tert</i> -butyloxycarbonyl
<b>calc.</b>	calculated
<b>CD</b>	circular dichroism
<b>Cl-MPA</b>	<i>N</i> -Fmoc-4'-[poly(oxyethylene)carbamoylmethoxy]-2, 4-dimethoxybenzhydrylamine polymer-bound Poly(oxyethylene)-RAM polymer-bound
<b>Da</b>	Dalton
<b>DBU</b>	1,8-diazabicyclo[5.4.0]undec-7-en
<b>Dbz</b>	3,4-amino-benzoic acid
<b>DCM</b>	dichloromethane
<b>DIC</b>	diisopropylcarbodiimide
<b>DIPEA</b>	<i>N,N</i> -diisopropylethylamine
<b>DMF</b>	dimethyl formamide
<b>DSL</b>	diselenide selenoester peptide ligation
<b>EDT</b>	1,2-ethanediol
<b>eq.</b>	equivalents
<b>ESI</b>	electron spray ionization
<b>Et<sub>2</sub>O</b>	diethylether
<b>EtOH</b>	ethanol
<b>Fmoc</b>	9- <i>N</i> -fluorenylmethyloxycarbonyl
<b>GndHCl</b>	guanidine hydrochloride
<b>h</b>	hour
<b>HATU</b>	1-[bis(dimethylamino)methylene]-1 <i>H</i> -1,2,3-triazolo[4,5- <i>b</i> ]pyridinium 3-oxide hexafluorophosphate
<b>HCTU</b>	0-(1 <i>H</i> -6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
<b>HFIP</b>	1,1,1,3,3,3-hexafluoro-2-propanol
<b>HOAt</b>	1-Hydroxy-7-azabenzotriazole
<b>HOBt</b>	1-hydroxy-benzotriazole
<b>HPLC</b>	High performance liquid chromatography
<b>HR</b>	high resolution
<b>ivDde</b>	4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)-3-methylbutyl
<b>KAHA</b>	ketoacid-hydroxylamine ligation
<b>LL</b>	low loading
<b>M</b>	molar
<b>m/z</b>	mass per charge
<b>min</b>	minutes
<b>Mmt</b>	monomethoxytrityl
<b>Mtt</b>	4-methyltrityl
<b>MW</b>	micro wave

---

<b>Nbz</b>	<i>N</i> -acyl-benzimidazolidone
<b>NCL</b>	native chemical ligation
<b>Nle</b>	Norleucine
<b>nm</b>	nanometer
<b>NMP</b>	1-Methylpyrrolidin-2-one
<b>OEpe</b>	<i>O</i> -3-ethyl-pent-3-yl
<b>OMpe</b>	<i>O</i> -3-methylpent-3-yl
<b>p</b>	para
<b>PBS</b>	phosphate-buffered saline
<b>PDB</b>	Protein databank
<b>PEG</b>	Polyethylene glycol
<b>Pip</b>	piperidine
<b>quant.</b>	quantitative
<b>rpm</b>	rounds per minute
<b>rt</b>	room temperature
<b>SDS-PAGE</b>	sodium dodecylsulfate polyacrylamide gel electrophoresis
<b>SPPS</b>	solid phase peptide synthesis
<b>STL</b>	serine/threonine ligation
<b>TCEP</b>	tris-(2-carboxyethyl)phosphine
<b>TFA</b>	trifluoroacetic acids
<b>TIS</b>	triisopropylsilane
<b>ToF</b>	time of flight
<b>Trt</b>	Trityl-
<b>Ub</b>	ubiquitin

Abbreviations of the 20 canonical amino acids are consistent with the one- or three-letter code recommended by the IUPAC-IUB Joint Commission on Biochemical Nomenclature (*Eur. J. Biochem.* **1984**, 138, 9-37).

Abbreviations for fluorinated amino acids relevant to this thesis are given below.

The abbreviations correspond to the L amino acids.

**Abz (Z)**      *ortho*-aminobenzoic acid

**(3R)-TfV**      (2*S*,3*R*)-4,4,4-trifluorovaline

IUPAC: (2*S*,3*R*)-2-amino-4,4,4-trifluoro-3-methylbutanoic acid

**(3S)-TfV**      (2*S*,3*S*)-4,4,4-trifluorovaline

IUPAC: (2*S*,3*S*)-2-amino-4,4,4-trifluoro-3-methylbutanoic acid

**TfeGly**      Trifluoroethyl glycine,

IUPAC (*S*)-2-amino-4,4,4-trifluorobutanoic acid

---



---

## Content of this doctoral thesis

<b>Part A:</b> Fluorine in Protein Landscapes	1-143
<b>Part B:</b> Structural Elucidation of Signaling Pathways of GPR83	144-178
<b>Part C:</b> Material and Methods	179-258

---

## **Part A**

# **Fluorine in Protein Landscapes**

## Content Part A

<b>1</b>	<b><i>Fluorine in Protein Landscapes</i></b> .....	<b>1</b>
<b>1.1</b>	<b>Fluorine - Unique in its Nature</b> .....	<b>2</b>
1.1.1	Atomic Physical Properties .....	2
1.1.2	Fluorine Isotopes - Important Diagnostic Tools .....	3
1.1.3	The C-F System .....	3
1.1.4	Impact of Fluorine Substitution on Heteroatoms and Functional Groups .....	4
1.1.5	Non-covalent Interactions .....	5
1.1.6	Why is Fluorine a Rising Star in Modifying Molecular Properties?.....	8
<b>1.2</b>	<b>Fluorinated Amino Acids as Powerful Tools in Peptide and Protein Engineering..</b>	<b>10</b>
1.2.1	Incorporation of Fluorine into Peptides and Proteins .....	11
1.2.2	Fluorine Modifications in Peptides and Proteins.....	14
1.2.3	Properties of Fluorinated Amino Acids.....	18
1.2.4	Impact of Fluorinated Amino Acids on the Secondary Structure and Folding of Proteins .	24
1.2.5	Impact of Fluorine Modifications on Kinetics of Peptide and Protein Folding .....	29
<b>2</b>	<b><i>Model Systems Used in the Doctoral Thesis</i></b> .....	<b>32</b>
<b>2.1</b>	<b>Cell-penetrating Peptides: Fluorine-specific Interactions on Lipid Surfaces</b> .....	<b>32</b>
<b>2.2</b>	<b>Coiled-coil Systems</b> .....	<b>33</b>
2.2.1	FF03 Scaffold as Carrier for <i>in vivo</i> <sup>19</sup> F-MRI Examinations.....	33
2.2.2	A4/B4 System a well-defined Coiled-coil Dimer for AMF Measurements.....	35
<b>2.3</b>	<b>Ubiquitin- Oligomerization Ruled by Hydrophobic Patches</b> .....	<b>36</b>
<b>2.4</b>	<b>Small all-<math>\beta</math>-proteins: Folding of Barrel and Sandwich Structures</b> .....	<b>38</b>
<b>3</b>	<b><i>Strategies in Chemical Protein Synthesis</i></b> .....	<b>39</b>
<b>3.1</b>	<b>Ligation Techniques: a Boon in Chemical Protein Synthesis</b> .....	<b>39</b>
<b>3.2</b>	<b>Native Chemical Ligation: New Paths in Protein Landscapes</b> .....	<b>42</b>
3.2.1	Diaminobenzyl (Dbz) Linkers .....	44
3.2.2	Hydrazides .....	45
3.2.3	Bis(2-sulfanylethyl)amido (SEA) group .....	46
<b>4</b>	<b><i>Aim of the Presented Studies</i></b> .....	<b>47</b>
<b>5</b>	<b><i>Results and Discussion</i></b> .....	<b>50</b>

---

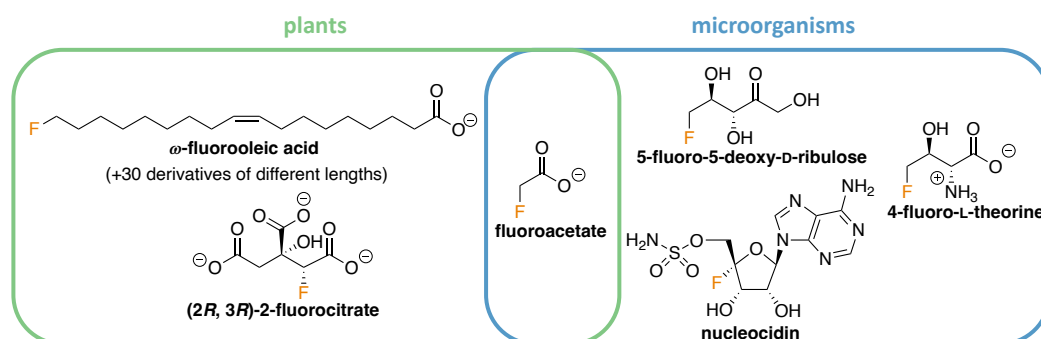
<b>5.1</b>	<b>Synthesis of Fluorinated Amino Acids.....</b>	<b>50</b>
5.1.1	Synthesis of Trifluoroethylglycine (TfeGly) as Substituent for Val .....	50
5.1.2	Synthesis of Trifluorovaline Derivatives (TfVal) as Leu/Ile Substituents .....	51
5.1.3	Chiral Ni-Complexes in the Synthesis of Fluorinated Amino Acids .....	53
5.1.4	Incorporation of Fluorinated Amino Acids into the Peptide Backbone .....	54
<b>5.2</b>	<b>Insertion of pHLIP Peptides into Lipid Bilayers.....</b>	<b>55</b>
5.2.1	Synthesis and Purification of pHLIP Peptides.....	55
5.2.2	Insertion Behavior Investigated by <i>in situ</i> SEIRAS.....	58
<b>5.3</b>	<b>Coiled-coil systems.....</b>	<b>59</b>
5.3.1	Fiber-forming Peptide FF03 .....	59
5.3.2	A4/B4 - a well-defined Heterodimeric System for AFM Experiments.....	69
<b>5.4</b>	<b>Fluorine as Orthogonal Tool to Impact Kinetics of Protein Folding via Product Formation of Ubiquitin Dimers .....</b>	<b>77</b>
5.4.1	Design of the Model System .....	77
5.4.2	Full-length Synthesis of Ubiquitin by MW-assisted SPPS.....	80
5.4.3	Comparison of MW-SPPS and Conventional SPPS in Ub Synthesis .....	85
5.4.4	Thioester Surrogates .....	86
5.4.5	Synthesis of Thiol Fragment of the NCL approach .....	96
5.4.6	Summary of Synthesized Monomers for Native Chemical Ligation .....	97
5.4.7	Native Chemical Ligation for Ubiquitin Dimerization.....	97
<b>5.5</b>	<b>Chemical Synthesis of Small all-<math>\beta</math> Proteins to Study Fluorine's Impact on Protein Folding Kinetics .....</b>	<b>102</b>
5.5.1	1FYN .....	102
5.5.2	Synthesis and Purification of 1FYN Derivatives and Pro-rich Peptides .....	103
5.5.3	Structural Examination of 1FYN Derivatives .....	111
5.5.4	Synthesis and Purification of 1TEN .....	112
5.5.5	Structural Examination of 1TEN Derivatives .....	115
<b>6</b>	<b>Summary and Outlook .....</b>	<b>116</b>
<b>7</b>	<b>References .....</b>	<b>122</b>



## 1 Fluorine in Protein Landscapes

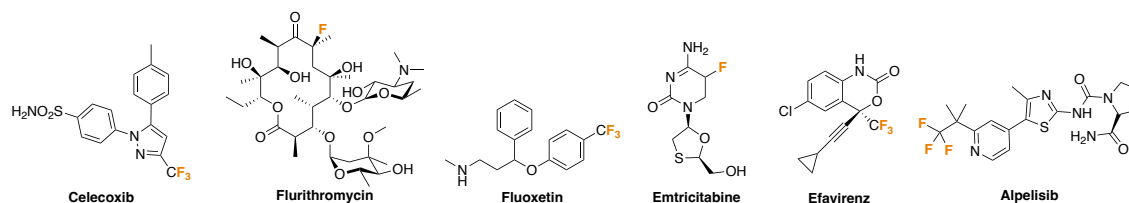
Like no other element, fluorine permits the modification of the properties of small molecules and large systems such as peptides and proteins. As the 24<sup>th</sup> most abundant element, fluorine occurs in nature primarily in the earth's crust in the form of its inorganic salts or complexes such as fluorspar (fluorite, CaF<sub>2</sub>), cryolite (Na<sub>3</sub>[AlF<sub>6</sub>]), and fluorapatite (Ca<sub>5</sub>(PO<sub>4</sub>)<sub>3</sub>F). [1,2]

In the living organism, fluorine appears relatively rarely, mainly because of its poor bioavailability, but it is by no means unknown. There are about 30 identified natural products (a selection is found in Figure 1.1) and even enzymes that perform fluorinations. [3,4]



**Figure 1.1:** Chosen known fluorinated natural products. Adapted from Cheng *et al.* [3] Copyright © 2021, The Author(s), under exclusive license to Springer-Verlag GmbH Germany, part of Springer Nature

In recent decades, interest in synthesizing novel fluorinated components has steadily increased, especially in developing new pharmaceuticals [5-7] or agrochemicals, [8,9] often with beneficial features like improved metabolic stability or biological activity. [10] This is particularly evident from the fact that the number of newly approved drugs bearing at least one fluorine atom has increased to about 20-25% (a selection is given in Figure 1.2). [11,12]



**Figure 1.2:** A selection of fluorine-containing pharmaceuticals. (Left to right): Celecoxib is a COX-2 inhibitor and nonsteroidal anti-inflammatory drug utilized in pain relief; flurithromycin is a broad-spectrum antibiotic that survives acidic environment; fluoxetine is an anti-depressant, emtricitabine and efavirenz are used in HIV-treatment; alpelisib is a PI3K-inhibitor applied in treatment of certain types of breast cancer. [5-7]

In addition, fluorine has become an essential tool in modifying properties, structure, and function in peptide and protein research. The following chapter highlights the unique characteristics of fluorine, its properties, and its effects on the physiological environment of peptides and proteins.

## 1.1 Fluorine - Unique in its Nature

Within this section, the comparison of fluorine and hydrogen is elaborated, and the modifications caused by the substitution of both elements are briefly discussed. A particular emphasis is placed on non-covalent interactions of fluorine, which may impact the secondary-structure formation and protein folding.

### 1.1.1 Atomic Physical Properties

Based on the comparison of the van der Waals (vdW) radii, fluorine is about 20% (1.47 Å, Table 1.1) larger than hydrogen (1.20 Å) and only slightly smaller than oxygen (1.52 Å). It is, therefore, commonly used to mimic those elements. Substituting hydrogen with fluorine is usually considered conservative as it generally causes minimal steric perturbation. With its strongest electron-withdrawing character of all elements, fluorine induces an electronic distribution that can change molecular properties. The fluorine atom has lower polarizability than hydrogen. Its electron affinity deviates remarkably from the trend among the halogens as it is lower than chlorine's. [13]

**Table 1.1:** Key properties of fluorine compared with other halogens and common elements of organic chemistry.

Atom	vdW radius [Å] <sup>[14]</sup>	atom polarizability (x10 <sup>-24</sup> cm <sup>3</sup> ) <sup>[15]</sup>	electronegativity of the element [ $\chi_p$ ] <sup>[15]</sup>	electron affinity [kJ mol <sup>-1</sup> ] <sup>[15]</sup>
H	1.20	0.67	2.20	72.8
C	1.70	1.76	2.55	153
N	1.55	1.10	3.40	-6.3
O	1.52	0.82	3.44	141.4
<b>F</b>	<b>1.47</b>	<b>0.56</b>	<b>3.98</b>	<b>328.2</b>
Cl	1.75	2.18	3.16	348.5
Br	1.85	3.05	2.96	324.5
I	1.98	4.70	2.66	295.4



### 1.1.2 Fluorine Isotopes - Important Diagnostic Tools

$^{19}\text{F}$  has a natural abundance of 100% with a spin of  $\frac{1}{2}$ . Fluorine has a wide chemical shift range ( $> 300$  ppm) and high sensitivity in  $^{19}\text{F}$ -NMR, as it reveals a receptivity of 0.834 relative to  $^1\text{H}$ .<sup>[16-18]</sup> Due to its absence in biological systems, fluorine enables unique opportunities in medical imaging studies (e.g.,  $^{19}\text{F}$ -Magnetic Resonance Imaging, MRI)<sup>[19]</sup> or studies of unlabeled proteins by  $^{19}\text{F}$ -NMR.<sup>[20,21]</sup>

$^{18}\text{F}$  is the most used radioisotope in positron emission tomography (PET) with nuclear characteristics of 97%  $\beta^+$  decay and 109.7 min half-life.  $^{18}\text{F}$ -PET is used for validating drug targets by biodistribution imaging of a radiotracer, e.g., in order to find off-target binding.<sup>[22]</sup>

### 1.1.3 The C-F System

With a dissociation energy of  $105 \text{ kcal mol}^{-1}$ , the C-F bond is the strongest single bond in organic chemistry (Table 1.2). Because of the strongest electronegativity, fluorine causes a highly polarized C-F bond with an ionic character and a high dipole moment. Compared to the C-H bond, the C-F bond has an inverted polarity ( $\text{C}^{\delta-} - \text{H}^{\delta+}$  vs.  $\text{C}^{\delta+} - \text{F}^{\delta-}$ ). The bond has a length of  $1.35 \text{ \AA}$ , between C=O ( $1.23 \text{ \AA}$ ) and C-OH ( $1.43 \text{ \AA}$ ) and is significantly longer than the C-H bond ( $1.09 \text{ \AA}$ ).<sup>[23]</sup>

**Table 1.2:** Comparison of C-F bond.<sup>[5]</sup>

Bond	Bond length [ $\text{\AA}$ ]	Dipole moment $\mu$ [D]	Bond dissociation energy [ $\text{kcal mol}^{-1}$ ] <sup>[24]</sup>
C-H	1.09	$\sim -0.4$	98.8
C-F	1.35	1.41	105.4
C-Cl	1.77	1.87 ( $\text{CH}_3\text{Cl}$ )	78.5
C=O	1.23	2.33 ( $\text{H}_2\text{C}=\text{O}$ )	85 ( $\pi$ -bond)
C-OH	1.43 (MeOH)	2.87 (MeOH)	84.0
	1.48 (EtOH)	1.66 (EtOH)	

Upon increasing fluorination at the  $\text{C}_\alpha$  atom, the steric demand of the alkyl chains also rises. That becomes particularly clear for the trifluoromethyl group, for which a vdW volume of  $38.9 \text{ \AA}^3$  is described, representing rather an ethyl residue ( $38.9 \text{ \AA}^3$ ). But the trifluoromethyl group has a different shape than the ethyl group. Depending on the method employed and the interpretation of the results, different values are assigned to the steric demand of the trifluoromethyl group.<sup>[25]</sup> Thus, the data varies from a simple methyl,<sup>[26,27]</sup> to twice as large as a methyl group,<sup>[28]</sup> to a *sec*-butyl or

a phenyl group.<sup>[29]</sup> In other studies, the CF<sub>3</sub> group is described as slightly larger than the *iso*-propyl group.<sup>[30,31]</sup> By X-ray crystal analysis, however, the size of an *iso*-butyl was determined.<sup>[32,33]</sup>

This research demonstrates that with fluorine substitution, it is not only the purely steric effects that need to be considered but also the electrostatic changes induced by the high electron-withdrawing character of fluorine. Here, two concepts play an active role in the electronegativity of fluorine: on the one hand, the intrinsic ability to attract electrons and, on the other hand, the capacity to stabilize the gained charge in the electron shell, which is determined by the low atomic polarizability of fluorine.<sup>[34]</sup>

These two aspects play an essential role in consideration of the non-covalent bonding of fluorine below.

#### 1.1.4 Impact of Fluorine Substitution on Heteroatoms and Functional Groups

Thanks to the highest electronegativity and the resulting inductive effect, fluorine substitution modulates the electron density of neighboring functional groups and thereby reduces their nucleophilicity, for example. In this way, the pK<sub>a</sub> values of COOH and OH groups are lowered with increasing fluorine content resulting in a higher acidity. Compared with the other halogens, the effect of the electron-withdrawing character of fluorine is significant (Table 1.3). For amines, however, the basicity is reduced.

This inductive effect will therefore influence the binding affinity of receptors, biological activities, and pharmacokinetics by increasing the hydrophobicity, lipophilicity, and membrane permeability.<sup>[35]</sup>

**Table 1.3:** Impact of fluorination on functional groups.

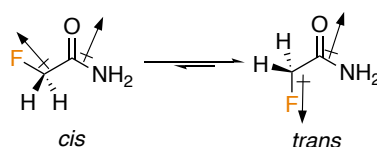
compound <sup>[36]</sup>	pK <sub>a</sub>	compound <sup>[37]</sup>	pK <sub>a</sub>	compound <sup>[38]</sup>	pK <sub>a</sub>
CH <sub>3</sub> CO <sub>2</sub> H	4.76	CH <sub>3</sub> CH <sub>2</sub> OH	15.93	CH <sub>3</sub> CH <sub>2</sub> NH <sub>2</sub>	10.6
CH <sub>2</sub> FCO <sub>2</sub> H	2.59	CF <sub>3</sub> CH <sub>2</sub> OH	12.39	CH <sub>2</sub> FCH <sub>2</sub> NH <sub>2</sub>	9.0
CH <sub>2</sub> ClCO <sub>2</sub> H	2.87	(CH <sub>3</sub> ) <sub>2</sub> CHOH	17.1	CHF <sub>2</sub> CH <sub>2</sub> NH <sub>2</sub>	7.3
CH <sub>2</sub> BrCO <sub>2</sub> H	2.90	(CF <sub>3</sub> ) <sub>2</sub> CHOH	9.3	CF <sub>3</sub> CH <sub>2</sub> NH <sub>2</sub>	5.7
CHF <sub>2</sub> CO <sub>2</sub> H	1.33	(CH <sub>3</sub> ) <sub>3</sub> COH	19.0		
CF <sub>3</sub> CO <sub>2</sub> H	0.50	(CF <sub>3</sub> ) <sub>3</sub> COH	5.4		

### 1.1.5 Non-covalent Interactions

Non-covalent forces play a fundamental role in forming secondary and tertiary structures of peptides and proteins, which also control biological functions and processes. Here the influences of the strong electronegativity of fluorine can lead to far-reaching consequences; hence, the options for fluorine to form non-covalent interactions will be outlined in the following sections.

#### 1.1.5.1 Polar Interactions

The strongly polarized C-F bond can therefore have a significant influence on macromolecules. For example, the strong ionic character of the C-F bond enables a high value in the dipole moment, which can also determine the conformation of organofluorine compounds (Figure 1.3) and allows interactions in immediate proximity. These types of interactions are well known from fluorine-containing drug molecules and their targeted receptors. [39]



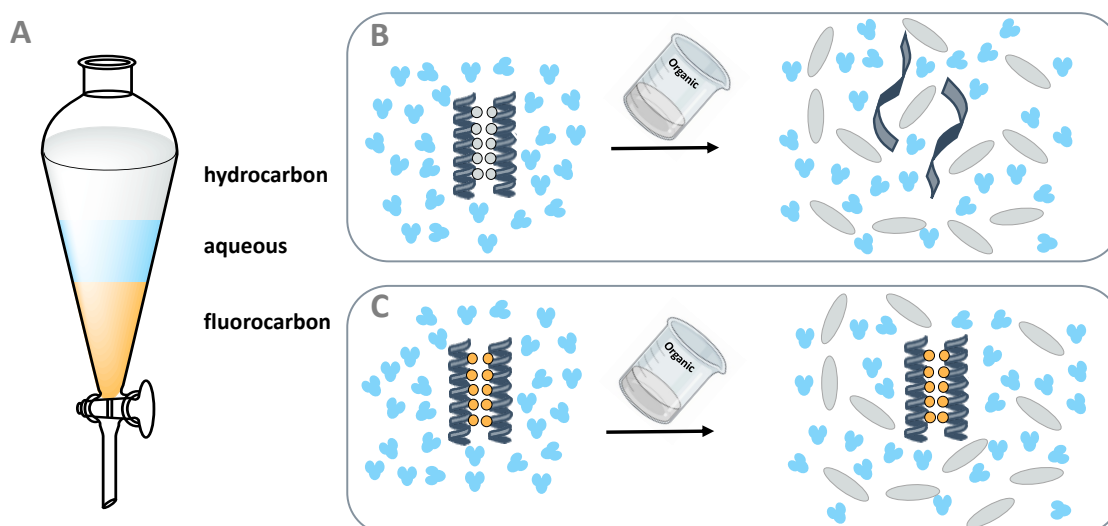
**Figure 1.3:** Conformational preference of  $\alpha$ -fluoroamide, opposing dipoles (indicated with arrows) favor the *trans* conformation. [24]

Fluorine's ability to form H-bonds has long been highly disputed. It has three lone pairs, potentially making it a good H-bond acceptor. However, these lone pairs are tightly bound to the nucleus due to the strong electronegativity, which has led to the assumption that fluorine cannot participate in H-bonds. Recent experimental and theoretical studies demonstrate that X-H...F-C (with X = O, N, S) interactions are possible. [40-50] But these are significantly weaker (F...H with 2.0 to 3.2 kcal mol<sup>-1</sup> instead of 5 to 10 kcal mol<sup>-1</sup> for O...H) than traditional H-bonds. [51,52] The main explanation for this weak interaction is the low polarizability and low capacity of fluorine to absorb the charge gained, making it a poorer H-bond acceptor than its electronegativity would allow. [34] Nevertheless, weak H-bond formation can be important in many processes beyond ligand-macromolecule target recognition, e.g., chemical synthesis and physicochemical properties of molecules. [53]

Moreover, the presence of numerous C-F---H-X (where X = O, N, S) as well as C-F---H-C<sub>α</sub> (e.g., C<sub>α</sub>-carbon of AAs)-contacts described in the Protein database (PDB) indicate that it is advantageous for the C-F dipole to engage in multipolar interactions, especially C-F---H-N (backbone amide) interactions are prominent. [54,55]

### 1.1.5.2 Fluorous Effect and Hydrophobicity

Another vital interaction that organofluorine compounds can undergo, and one that is controversially discussed in the literature, is the so-called *fluorous effect*. This term comes from studies that have shown that perfluorinated aliphatic compounds can form a phase of their own, i.e., they are immiscible with both the aqueous and the organic hydrocarbon phases (Figure 1.4). Derived from the word aqueous, this phase was named *fluorous*. [56,57] This phenomenon opened the door for a new field for research groups but also industries generating a multitude of applications in catalysis [58–62], synthesis [63,64], separation techniques like fluorosolid phase extraction (FSPE) [65,66] or fluorosolid HPLC purification, [67] construction of microarrays [68–78], and supramolecular chemistry in which perfluorinated alkyl chains are also used to design amphiphilic polymers. [79–86] Sometimes, this effect is also described as fluorophilicity. [87–89]



**Figure 1.4:** A) Separation of phases with fluorocarbons. B) Disruption of the inter- and intramolecular forces of peptides by adding organic solvents (grey eclipse). C) Suggested stabilization of a peptide by a fluorosolid core (highlighted in orange circles, grey circles represent a typical hydrocarbon core). Adapted with permission from Biava *et al.* [90] Created with *Biorender.com* © 2013 The Authors. Engineering in Life Sciences published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

Perfluoroalkyl chains differ from their hydrocarbon analogs in several ways, such as their geometry: they are more space-consuming and adopt a fully trans-helical structure instead of the trans-planar zig-zag conformation of hydrocarbons. They are more rigid and have less conformational freedom.<sup>[91]</sup> Compared to hydrogen, fluorine has a denser and less polarizable electron cloud that shields the C-C bond. This leads to weaker intermolecular attraction through vdW interactions, resulting in, for example, lower boiling points (BP) compared to their hydrocarbon counterparts (BP: C<sub>6</sub>H<sub>14</sub> 69°C (M<sub>w</sub>: 86 g · mol<sup>-1</sup>) vs. 56°C (M<sub>w</sub>: 338 g · mol<sup>-1</sup>)).<sup>[92]</sup> These features make perfluoroalkyl chains an advantageous modification in many areas of chemical biology and material science.<sup>[79,93,94]</sup>

To determine possible F---F interactions, numerous predominantly computer-assisted studies have been performed.<sup>[95-97]</sup> It was found that different factors, such as London dispersion forces or polarizability, play an influential role in this interaction to overcome the repulsive electrostatic effects.<sup>[98,99]</sup>

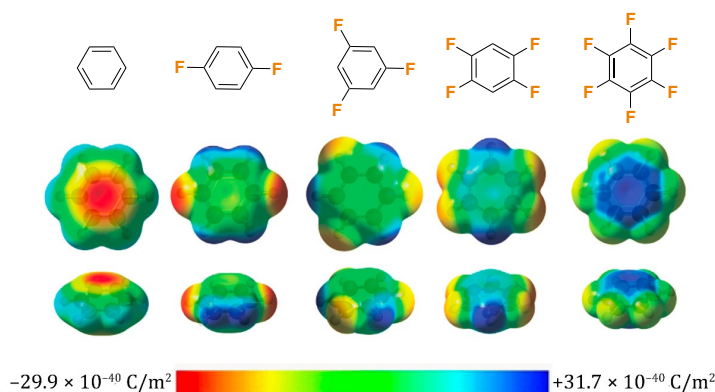
When examining crystal structures, for example, a few compounds are known that exhibit a short distance of F---F, which is shorter than the sum of the vdW radii of the individual fluorine atoms (2.94 Å).<sup>[100]</sup> These effects are highly debated whether they have a stabilizing character or are due to packing in a solid state.<sup>[101-104]</sup>

Investigations by Marsh *et al.*, however, showed that the formation of the fluorous phase is not caused by attractive, enthalpy-favored F---F interactions but rather, as described above, by the absence of intermolecular interactions such as vdW or H-bridges with other molecules.<sup>[105]</sup> In general, highly fluorinated alkyl chains are considered to be more hydrophobic than the parent systems, which is why fluorinated building blocks are also used in protein and peptide design in order to direct self-organization and packing into highly stable and ordered assemblies.<sup>[79,106]</sup>

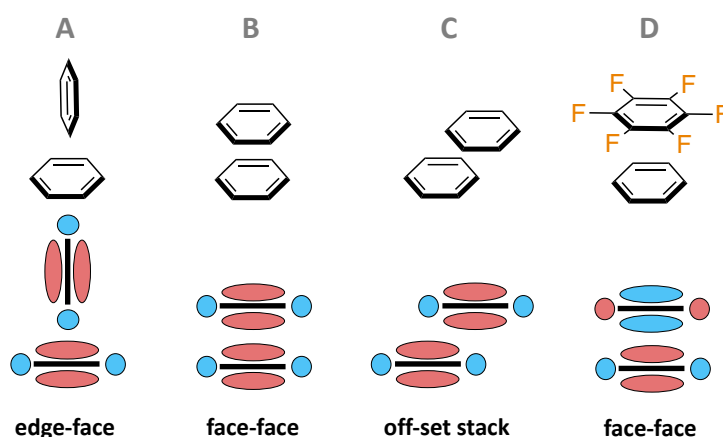
### 1.1.5.3 Interactions of Fluorinated Aromatics

In contrast to perfluorinated aliphatic compounds, perfluorinated aromatics preferentially interact with their hydrocarbon counterparts. After all, due to the strong electronegativity, the electrons are withdrawn from the aromatic system and are no longer as strongly delocalized as in the case of hydrocarbon-aromatics. As a result, the electron-rich  $\pi$ -donor benzene is converted to an electron-deficient

$\pi$ -acceptor in its perfluorinated state (Figure 1.5). This has profound consequences for the interaction capabilities of fluorinated aromatic compounds (Figure 1.6).<sup>[107]</sup>



**Figure 1.5:** Influence of fluorine substitutions on electrostatic potentials in aromatic systems. Adapted from Monkovic *et al.*<sup>[108]</sup>



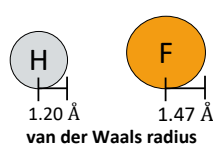
**Figure 1.6:** Geometry of common aromatic interactions and a schematic representation of the corresponding electrostatic potentials. Red indicates electron-rich potentials, while blue represents electron deficiency. Adapted with permission from Waters.<sup>[109]</sup> Copyright © 2002 Elsevier Science Ltd.

### 1.1.6 Why is Fluorine a Rising Star in Modifying Molecular Properties?

The preceding sections have clearly shown that the incorporation of fluorine modifies the electrostatic distribution within molecules, thereby modifying the nucleophilicity, hydrophobicity, or membrane permeability. The C-F bond is the strongest single bond in organic chemistry with a strong ionic character, shielded by the dense electron cloud of fluorine. Although hydrogen exchange by fluorine is usually considered structurally conservative, i.e., associated with low perturbation, the introduction of fluorine is also suitable for mimicking other functional groups (Figure 1.7 B). This is mainly due to the similar orientation of the dipoles (e.g., C=O vs. C-F). Fluorinated compounds are capable of weak non-covalent interactions with

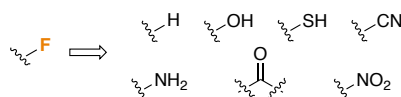
their immediate environment, which can affect the organization of macromolecules. In particular, the ability of perfluorinated aliphatic compounds to form a phase of their own allows fluorine to be incorporated as an orthogonal tool in systems for various applications. The inverted electrostatic potential map of fluorinated aromatics makes them interesting candidates for modulating assembly processes and folding of peptides and proteins. Its low bioavailability is also a great advantage, as its  $^{19}\text{F}$  and  $^{18}\text{F}$  isotopes are widely used in medical diagnostics ( $^{19}\text{F}$ -MRI,  $^{18}\text{F}$ -PET) or in structural elucidations of proteins by  $^{19}\text{F}$ -NMR.

### A Comparison of hydrogen and fluorine nuclei

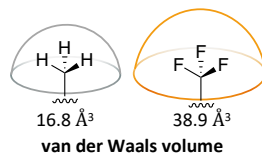


	H	F
Electronegativity (Pauling)	2.20	3.98
polarizability [ $\times 10^{-24} \text{ cm}^3$ ]	0.667	0.557

### B Fluorine as mimic for other functional groups

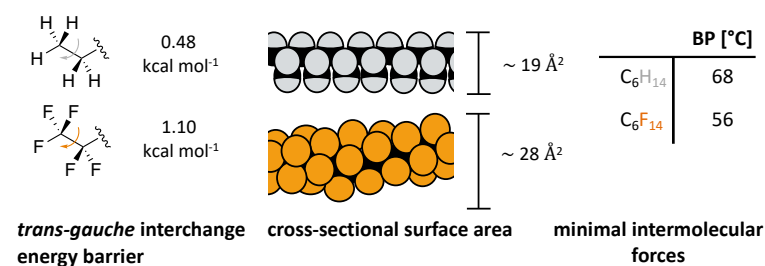


### C Comparison of carbon bound to hydrogen vs. fluorine



	C-H	C-F
dipole	←-+	+→
bond length [Å]	1.09	1.35
bond strength [kcal mol <sup>-1</sup> ]	~ 99	~ 105

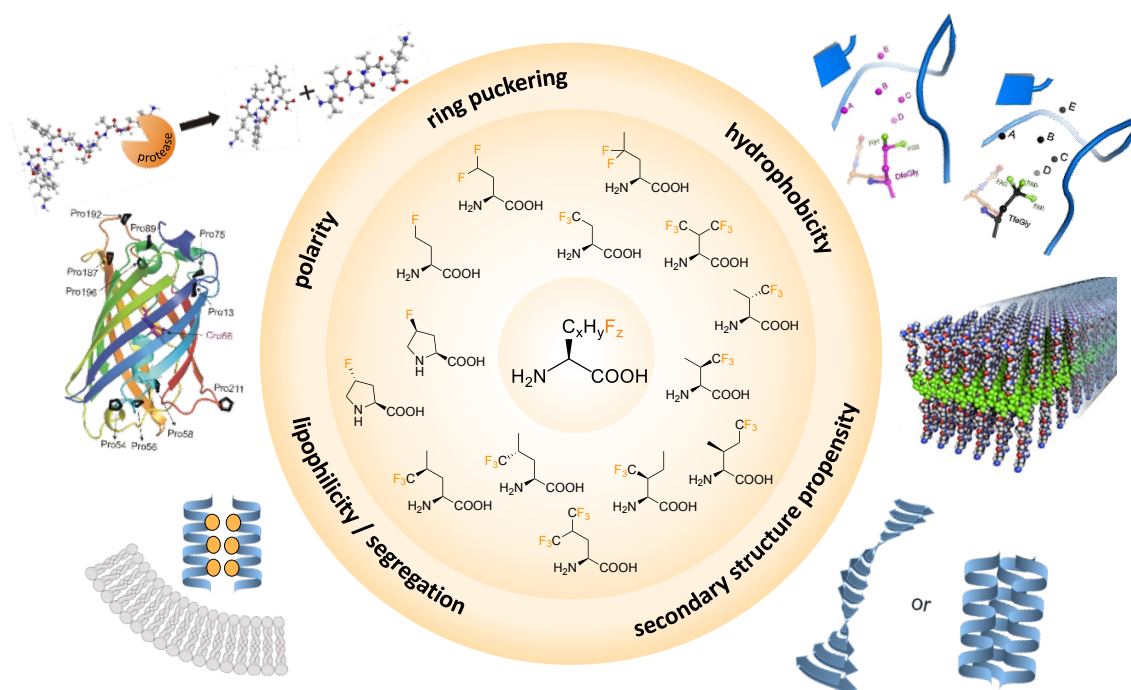
### D Comparison of hydrocarbons vs. perfluorocarbons



**Figure 1.7:** Comprehensive overview of key parameters of hydrogen substitution by fluorine. **A)** Comparison of the physical characteristics of both atoms; **B)** possible bioisosteric applications; **C)** Comparison of the C-H and C-F bond as well as methyl and trifluoromethyl group; **D)** differences between perfluorocarbons and the corresponding hydrocarbons. Adapted with permission from Miller *et al.*<sup>[110]</sup> (© 2020 Wiley-VCH GmbH) and Monkovic *et al.*<sup>[108]</sup>

## 1.2 Fluorinated Amino Acids as Powerful Tools in Peptide and Protein Engineering

As demonstrated in the previous chapter, incorporating fluorine has far-reaching consequences for a molecule. The unique combination of the small size, the very low polarizability, and the strong inductive effect is already widely used to investigate the physicochemical properties of peptides and proteins concerning their stability or to enable novel applications. At the same time, however, there is no direct prediction of the corresponding consequences of fluorine modification possible. Extensive review articles by the Kokschi and Montclare groups deal with the various implications of fluorine substitution in peptide and protein engineering, summarized in Figure 1.8. [108,111,112] Sloand *et al.* report an excellent overview of fluorinated peptidic biomaterials. [106]



**Figure 1.8:** Overview of the potential factors of the introduction of fluorine on the properties of amino acids as well as their consequences in the peptide backbone and application in peptide and protein engineering. Adapted from Berger *et al.* [112]

The following section provides a short overview of aspects related to incorporating fluorine to modify the formation of secondary structures and protein folding. For this purpose, the methods for incorporating fAAs into peptides and proteins are first briefly presented. In addition, changes in properties such as hydrophobicity or

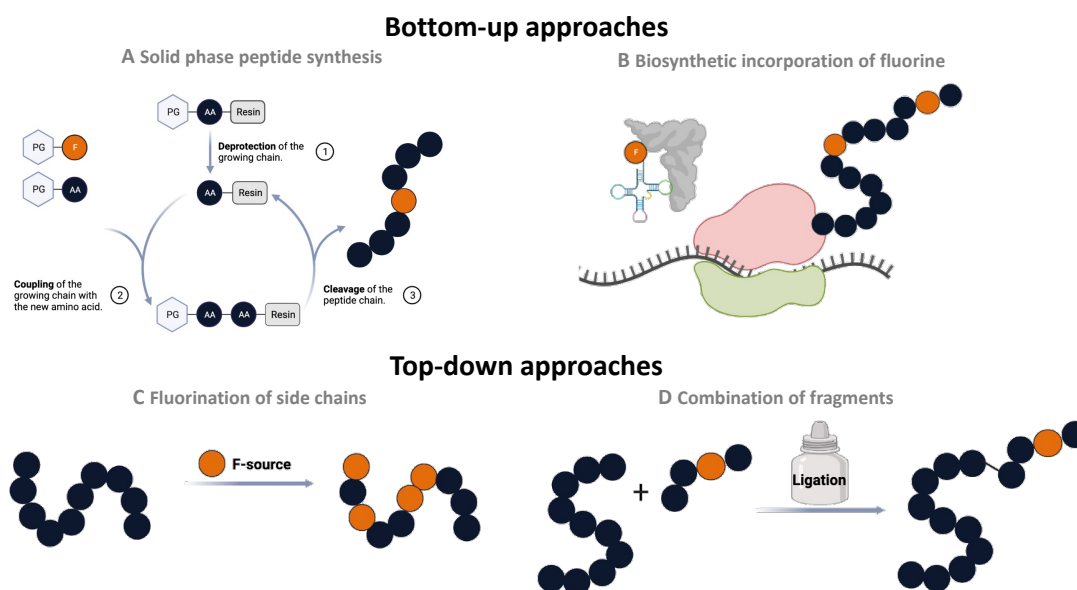


secondary structure propensity of fluorinated hydrophobic aliphatic AAs will be discussed, and their impact on peptide and protein structure will be addressed.

### 1.2.1 Incorporation of Fluorine into Peptides and Proteins

Depending on the desired degree of fluorination and the planned investigation, numerous ways exist to introduce fluorine into the world of peptides and proteins. [113]

In principle, there are two main strategies (Figure 1.9): bottom-up approaches (Figure 1.9 A + B), in which the peptide backbone is built up either chemically via the solid phase peptide synthesis (SPPS) or with the biological machinery, and fluorine is introduced via corresponding fluorinated building blocks. Or top-down approaches (Figure 1.9 C + D) where fluorine modifications are added to already built peptide/protein strands with the help of fluorinating agents or by conjugation reactions. There is also a combination in which, for example, ligation of a fluorinated fragment and a chemically or biologically synthesized peptide/protein is used (in this case, it is called semi-synthesis).



**Figure 1.9:** Methods of introducing fluorine (indicated by orange circles) into peptides and proteins (dark circles). Bottom-up approaches: **A)** solid phase peptide synthesis and **B)** biosynthesis. Top-down approaches: **C)** Fluorination of side chains and **D)** Combination of fluorinated fragments (ligation or fragment condensation). Created with *Biorender.com* and adapted from Monkovic *et al.* [108]

### 1.2.1.1 Solid phase peptide synthesis (SPPS)

Each of these methods has its advantages and disadvantages. For example, SPPS allows the well-defined incorporation of selected fAAs at distinct positions of the primary structure. However, this method is usually limited to a length of 50 AAs; ligation steps must be considered for longer sequences (detailed mechanisms are presented in section 3 on chemical protein synthesis). In addition, the incorporation of fAAs by SPPS is far from being straightforward, as the strong inductive effect leads to a decreased nucleophilicity of the N-term and, therefore, to a reduction of the coupling efficiency. Thus, for each sequence, a specific protocol has to be established for incorporating the fAAs and the subsequent sequence. Devillers *et al.* developed, for example, a di-peptide building block to enable the incorporation of trifluoromethyl-Ala, which cannot be introduced into peptide backbone using SPPS but must instead be synthesized in solution as an alternative. [114]

Another disadvantage that affects all bottom-up approaches is the low commercial availability of fluorinated building blocks. These usually have to be elaborated in complex syntheses. [115] Inserting those particular building blocks using SPPS is challenging, as they are generally coupled with fewer equivalents than the canonical AAs.

### 1.2.1.2 Biosynthetic Approaches

There are many ways to enable residue- or site-specific incorporation when using biological machinery, e.g., auxotroph bacteria or bio-orthogonal translation. [116–126] Note that these methods are limited to unnatural building blocks recognized by the endogenous or bio-engineered aminoacyl tRNA synthases. This, in turn, constrains the application of AAs with similar structural architecture (e.g., L-configuration) as their non-fluorinated counterparts. Another limitation is that 100% incorporation rarely occurs (90-95% is already considered very good). [127]

A new research strategy was published by Agostini *et al.* They have engineered *E. coli* by adaptive laboratory evolution experiments to live on fluorinated indole variants, forcing the bacteria to synthesize fluorinated Trp-variants and thereby synthesizing the whole proteome with fluorinated Trp-analogs. [128]

### 1.2.1.3 Late-stage Fluorination

Top-down methods circumvent the difficulties of incorporating fluorinated building blocks during peptide assembly by subjecting the peptide/protein to conjugation reactions with small, fluorinated molecules or fluorinating agents after synthesis and purification. These methods are often used in  $^{19}\text{F}$ -NMR studies. They take advantage of the fact that the  $^{19}\text{F}$  chemical shifts are extremely sensitive to changes in the immediate environment and provide an opportunity to study otherwise unlabeled proteins' structure or conformational changes. [129]

Generally, this procedure involves global fluorination of suitable free functional groups such as  $-\text{NH}_2$  or  $-\text{SH}$  throughout the protein. Often 2,2,2-trifluoroethanethiol (TFET) [130-133] or 3-bromo-1,1,1-trifluoroacetone (BTFA) [134,135] are chosen for conjugation reactions with thiol groups. BTFA involves the formation of a thioester bond, whereas TFET is introduced via a disulfide linkage and is, therefore, reversible. [129] In addition, there are studies on labeling free amino groups of, for example, Lys residues by trifluoro acetic acid (TFA) acetylation or reaction with S-ethyl trifluorothioacetate. [136-138]

The investigation of a single fluorinated position is only possible to a limited extent with these conjugation reactions. Either a system must be selected with only one functional group required for conjugation, or a targeted mutation must be performed to introduce this functional group. The latter has been exploited by Shekhawat *et al.* to characterize conformationally distinct ubiquitin oligomers via  $^{19}\text{F}$ -NMR by introducing a Cys-mutation at the Gln46 position and then treatment with BTFA. [139]

Efforts to employ fluorinating agents are also intriguing; for example, Hebel *et al.* published acetyl hypofluorite as an electrophilic fluorinating reagent on the aromatic ring of Tyr in a heptapeptide. [140] Imiołek *et al.* reported on a radical trifluoromethylation of Trp-residues, [141] and Ichiishi *et al.* of Phe-residues under mild conditions. [142] However, using fluorination reagents on synthesized peptides or proteins is not widespread due to their complex handling (e.g., the need for dry organic solvents), harsh reaction conditions, and undirected fluorination.

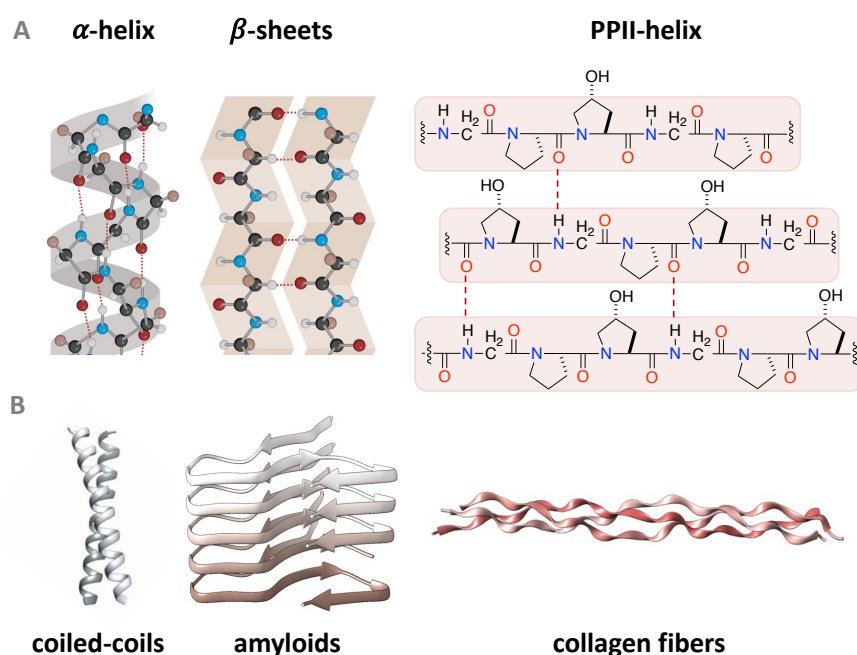
Another application of the top-down approach is the use of  $^{18}\text{F}$ -labeled proteins and peptides, for which the modification is used as late as possible in the synthesis due to the short half-life of the  $^{18}\text{F}$ -isotope. [143]

## 1.2.2 Fluorine Modifications in Peptides and Proteins

In this section, the main secondary structures of peptides and their organization into higher structures will be presented to lay a groundwork for a better understanding of fluorine's influence later on.

### 1.2.2.1 Common Secondary Structures and Folding Motifs of Peptides

The two major folding motifs of peptides include the  $\alpha$ -helix (grey) and  $\beta$ -sheets (brown), whose higher organizations are supercoils in coiled coils or amyloids (Figure 1.10). In proteins, all- $\beta$ -structures can adopt two folded structures, on the one hand, a barrel structure and, on the other hand, a sandwich structure (not depicted). Another secondary structure, although less common, is the poly-L-proline type II helix (PPII), which is the principal segment of fibril-like proteins such as collagen (red).



**Figure 1.10:** Overview of the most important peptide folding motifs: **A**)  $\alpha$ -helix (grey),  $\beta$ -sheets (brown), and poly-L-proline type II helix (PPII-helix, red) (hydrogen bonds are highlighted in dark-red) as well as **B**) their organization to higher order structures (coiled coils (PDB-ID: 1P9I [144]), amyloids (PDB-ID: 2XMU[145]), and collagen fibers (PDB-ID: 1CGD [146]) Created with Biorender.com and <http://www.cgl.ucsf.edu/chimera>.

All arrangements are determined by an ensemble of hydrophobic, vdW interactions or H-bonds, upon which a fluorine substitution is likely to have an effect.

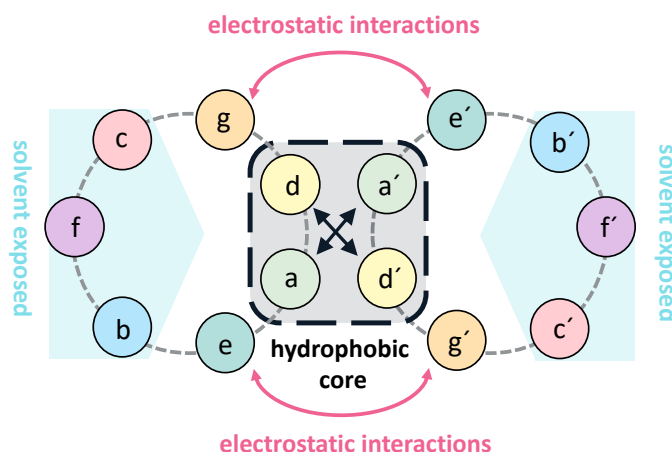
### *$\alpha$ -Helical Structures*

The right-handed  $\alpha$ -helix is the most common secondary structure of peptides and is formed by 3.6 residues per turn. This structure is stabilized by H-bonds of the carbonyl group of the backbone of residues  $i$  and the amide group of residues  $i + 4$ . All side chains point away from the helix axis. [147]

Isolated monomeric  $\alpha$ -helices, however, are rarely observed in the solution phase; usually, higher organized structures, so-called coiled coils, are formed. Here, two or more  $\alpha$ -helices self-assemble and twist around each other, creating an amphipathic left-handed superhelix. [148,149] In this arrangement, the periodicity is reduced from 3.6 AAs per turn to 3.5, or the sequences are repeated after two turns or after seven AAs, respectively (Figure 1.12 A). [150,151]

The motif was developed from research on  $\alpha$ -creatin by Crick, Pauling, and Corey nearly 70 years ago. [152–155] Nowadays, it is a widely used motif, and coiled coils are widely used as scaffolds for various applications. [156–159]

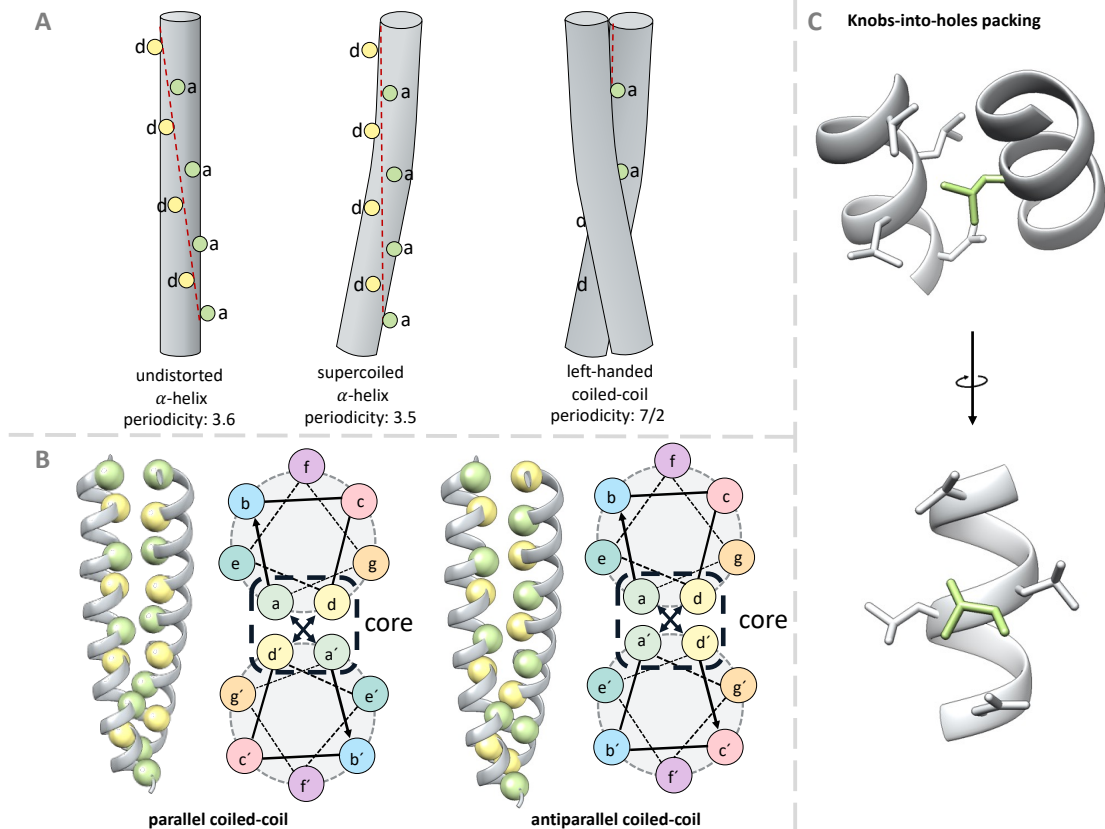
The simple design is characterized by varying repetitions of a heptad pattern denoted as  $(a-b-c-d-f-e-g)_n$ . [160] It is typically depicted in a helical wheel diagram (Figure 1.11). Here, the number  $n$  of heptad repeats varies from, for example, two in *de novo* designs [161] to 200 in naturally occurring fibril-like proteins. [162] Even though dimers, trimers, and tetramers are the most common types of coiled coils, larger designs with up to seven helices are also reported in the literature. [163,164]



**Figure 1.11:** Helical wheel representation of a parallel coiled-coil design. Positions  $a$  (green) and  $d$  (yellow) refer to hydrophobic AAs; dark arrows indicate their interaction. Positions  $e$  (petrol) and  $g$  (orange) are occupied by charged AAs contributing electrostatic interactions (indicated with pink arrows) for folding stabilization. Polar or neutral AAs are used at solvent-exposed positions ( $c$  (red),  $f$  (purple), and  $b$  (blue)) that are underlined in a light-blue box.

In the helical-wheel diagram, hydrophobic AAs (e.g., Val, Leu, Ile) are used in positions *a* and *d*, forming a hydrophobic core enhanced by electrostatic interactions between positions *e* and *g* from complementary charged AAs (e.g., Glu, Lys, Arg). Positions *b*, *c*, and *f* are solvent exposed. Here, additional charged AA are incorporated, e.g., to control the degree of oligomerization or to allow fibrillation.<sup>[158,165]</sup> Other polar AAs, such as Ser, are suitable for increasing the solubility of the coiled coil. Alternatively, Ala can also be used at these positions since it is a small, uncharged AA that favors the  $\alpha$ -helix.<sup>[166]</sup> Conjugations with different modifications (drugs, epitopes, etc.) can occur at the *f* position because this position does not contribute to the stabilization of the coiled coil.<sup>[167,168]</sup> Crucial for the organization of the helices is a special packing: the knobs-into-holes arrangement.<sup>[152]</sup> Here, the one hydrophobic AA of one helix acts as a knob and packs into a diamond-like hole of the other helix, which is formed by four AAs (Figure 1.12 C) on the other helix. The formation of the hydrophobic packing is the driving force, and the motif is further stabilized by intermolecular interactions such as salt bridges.<sup>[169]</sup>

Helices can be either identical (homotypic) or different (heterotypic), and they can adopt different orientations towards each other, parallel on the one hand and antiparallel on the other (Figure 1.12 B).<sup>[166,170,171]</sup> The selected AAs determine the direction in the respective positions. As an example, Ile and Val are used as  $\beta$ -branched AA predominantly in the *a* position of parallel dimers, while Leu is utilized in the *d* position. Similarly, the arrangement of the charged AAs in the *e* and *g* positions can control the orientation.<sup>[172,173]</sup>



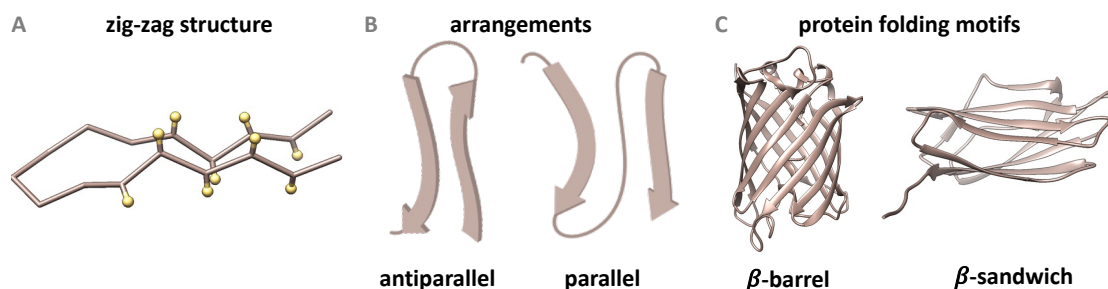
**Figure 1.12:** Packing effects of coiled coils. **A)** Comparison of undistorted  $\alpha$ -helix vs. supercoiled and coiled-coil arrangements. **B)** Parallels and antiparallel coiled-coil dimers with their respective helical wheel diagram and illustration of hydrophobic interactions. The hydrophobic core is build-up by *a* (green sphere) and *d* (yellow sphere). Layers, while mixed *a*-*d* layers, construct the hydrophobic core of an antiparallel dimer. **C)** Schematic representation of knobs-into-holes packing of an Ile-residue (highlighted in green). The ribbon of the second helix was removed to visualize the diamond-shaped hole (grey). Adapted from Truebestein *et al.*<sup>[151]</sup> and Rhys *et al.*<sup>[166]</sup> Created with <http://www.cgl.ucsf.edu/chimera>.

### ***$\beta$* -Structures**

$\beta$ -Sheets are the second most abundant secondary structural element of peptides. They are characterized through an accordion-like folded structure (Figure 1.13 A). In general,  $\beta$ -sheets are designed by alternating a polar and a hydrophobic AA. Val, Thr, and Ile are often used in this motive, as are aromatic AAs such as Phe, Trp, and Tyr.<sup>[174]</sup> The side chains of all AAs point downwards or upwards out of the plane of the sheet.<sup>[175]</sup> The assembly of  $\beta$ -sheets can lead to the formation of amyloid structures. These are usually insoluble aggregates associated with diseases such as Alzheimer's in the human body.<sup>[176]</sup>

Like coiled coils,  $\beta$ -strands can adopt different arrangements: parallel or antiparallel, with the latter having a higher natural occurrence (Figure 1.13 B).<sup>[177]</sup> In each case, stabilization occurs via H-bonds from the opposite  $\beta$ -sheet. A  $\beta$ -hairpin

is the minimum increment, combining two antiparallel  $\beta$ -sheets with a turn. [178–180] In proteins, the  $\beta$ -structures can adopt the folding motif of a barrel or a sandwich (Figure 1.13 C).



**Figure 1.13:** Summary of key elements of  $\beta$ -structures. **A)** common zig-zag structure of  $\beta$ -sheets (PDB-ID: 1U6U [181]). Residues of AAs are indicated with yellow spheres. **B)** Possible arrangements of  $\beta$ -strands and **C)** common folding motifs of all- $\beta$ -proteins (PDB-ID: 4OGS [182]  $\beta$ -barrel, 4JOX [183]  $\beta$ -sandwich). Created with *Biorender.com* and <http://www.cgl.ucsf.edu/chimera>.

### ***PPII-helix and Collagens***

Although PPII is found in only a few proteins, this motif has proven to be valuable for studying protein-protein or protein-DNA interactions and for biomedical applications. [184] Repetitions also construct this design motif; here, the tripeptide  $(X_{AA}-Y_{AA}-Gly)_n$  is used, whereby  $X_{AA}$  and  $Y_{AA}$  can be any AA. By natural occurrence, commonly, 2*S*-Pro for  $X_{AA}$  and (2*S*,4*R*)-4hydroxyproline for  $Y_{AA}$  are implemented. [185] Gly is invariant as it is crucial for tight packing. The basic framework is a left-handed helix that assembles into a right-handed triple helix, which is stabilized by interhelical H-bonds of  $C=O_{(X_{AA})}$  and  $H-N_{(Gly)}$  intrahelical H-bonds do not occur (Figure 1.10). Similar to the coiled coil structures, the individual components of the triple-helix can be either identical or different up to heterotrimeric. [186] Typically, *de novo* designs are composed of six to ten repeats of the tripeptide sequence; in nature, more than 300 repetitions can be achieved. [187] So-called sticky ends are also incorporated in *de novo* designs for improved assembly. [188]

### **1.2.3 Properties of Fluorinated Amino Acids**

The formation of secondary and tertiary structures in the protein world depends on the amino acids' properties and their arrangement. When selecting a fAA, it is essential to consider which properties of the AA are modified by fluorine and what influence these changes have on the overall appearance of a peptide or protein. In



the present doctoral thesis, the replacement of mainly hydrophobic aliphatic AAs by fluorinated variants is scheduled. Therefore, the following section will discuss the hydrophobicity and the secondary structure propensity of aliphatic fAAs. It is important to note that C $\alpha$ -trifluoromethylated AAs represent a specialized class that is not in the context of the present work and will, consequently, not be considered further.

### 1.2.3.1 Hydrophobicity of Fluorinated Amino Acids

There are various methods for determining AAs' hydrophobicity based on mimicking a protein's interior, as hydrophobic interactions are key factors in protein folding. These methods include partitioning methods between two immiscible phases (e.g., water vs. octanol [189], *N*-cyclohexyl-2-pyrrolidone [190] or *N*-methyl-acetamide [191]), chromatography-based studies, [192,193] site-directed mutagenesis in proteins, [194] and computer-based analyses. [195] All of these methods have been applied primarily to canonical AAs, but it has been found that a different hydrophobicity scale can be determined by each method. To date, there are more than 100 different hydrophobic scales. [196,197]

fAAs have also been investigated with respect to their hydrophobicity, but these experiments are usually limited due to low commercial availability.

Partition studies were performed by Lee *et al.* with HfLeu in heptanol and *n*-1,1-H<sub>2</sub>-perfluoroheptanol. [198] Here, HfLeu was found to exhibit a higher coefficient in each case (<sup>H</sup> $\Pi_{\text{Leu}}$  1.58 vs. <sup>H</sup> $\Pi_{\text{HfLeu}}$  1.87 and <sup>F</sup> $\Pi_{\text{Leu}}$  1.56 vs. <sup>F</sup> $\Pi_{\text{HfLeu}}$  1.87), indicating a higher hydrophobicity.

This coefficient was determined with the following equation.

$${}^X\Pi = \log(D_{AA}) - \log(D_{Gly}) \quad (1)$$

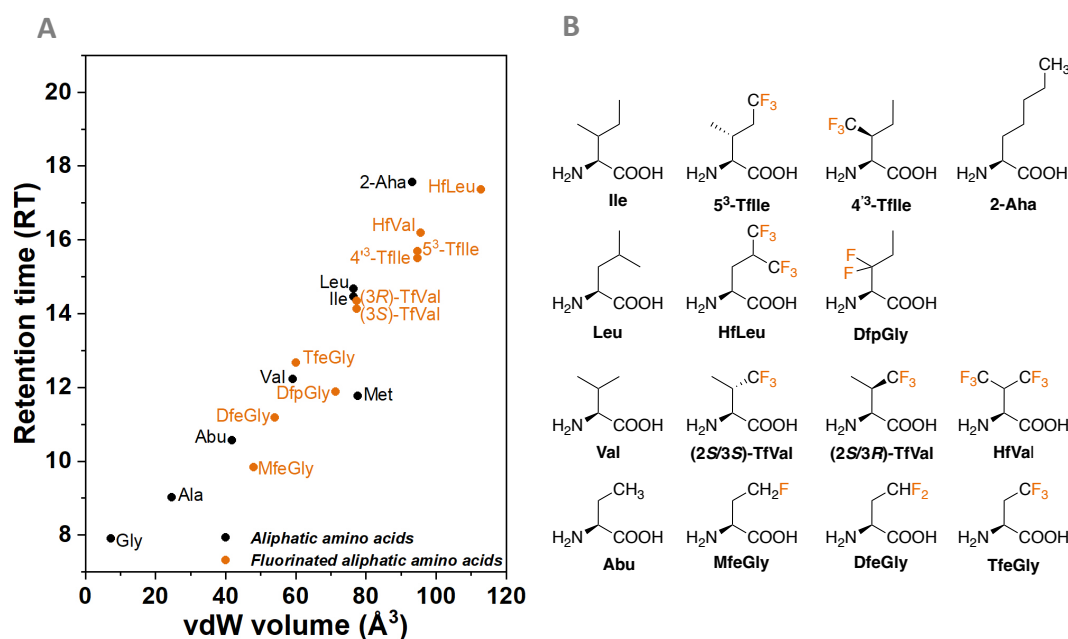
*X* refers to the solvent used in the experiment: H = heptanol, F = *n*-1,1-H<sub>2</sub>-perfluoroheptanol, and D is the distribution coefficient between water and chosen organic solvent.

Interestingly, no higher coefficient for HfLeu was obtained in the fluorinated solvent, indicating the absence of a favorable fluorous effect since it has been hypothesized that HfLeu may favor perfluorinated solvents due to the presence of both CF<sub>3</sub> groups. The free energy of the partition was also investigated, with a difference of -0.4 kcal mol<sup>-1</sup> ( $\Delta G_{\text{Leu}} = -2.1$  kcal mol<sup>-1</sup>,  $\Delta G_{\text{HfLeu}} = -2.5$  kcal mol<sup>-1</sup>).

HPLC-based methods to study the hydrophobicity of fAAs have been reported by Kokschi and Brigaud group. [199–203]

Here it is assumed that the non-polar stationary phase can mimic a hydrophobic environment, whereby the retention time correlates with the hydrophobicity of the corresponding compound as a function of size. [204] An increase in retention time corresponds to a higher hydrophobicity (Figure 1.14). For this purpose, Fmoc-protected derivatives of fAAs were used by our group. The calculation of the vdW volumes was carried out according to the method of Zhao *et al.* [205]

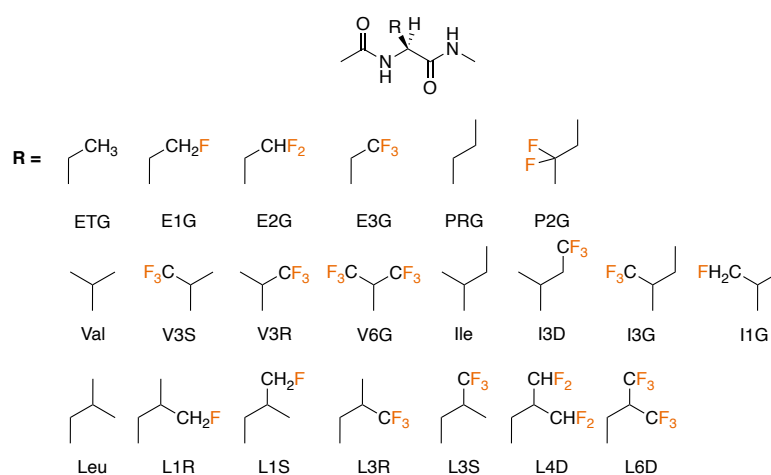
For the non-fluorinated variants, hydrophobicity increases with a higher side chain space requirement (i.e., larger vdW volume). Note that the rise in hydrophobicity is not linear. When comparing fluorinated (TfVal, HfVal, and HfLeu) and non-fluorinated derivatives of Val and Leu, it is evident that fluorine substitution leads to increased hydrophobicity. At the same time, it is noteworthy that the two TfVal diastereomers exhibit slight differences in hydrophobicity due to the different orientations of the CF<sub>3</sub> group. While having a similar vdW volume, TfeGly has a minimal higher hydrophobicity than Val.



**Figure 1.14:** HPLC-based investigation of hydrophobicity. **A)** The retention time of Fmoc-protected fluorinated aliphatic AAs (orange circles) [199–201] with their non-fluorinated counterparts (black circles) plotted against the corresponding vdW volumes. **B)** Structures of selected AAs.

Interestingly, there are opposing trends; MfeGly possesses a lower hydrophobicity than the non-fluorinated AA Abu, although they have a similar spatial demand in the side chain. DfpGly, which has an equal vdW volume to Leu or Ile, also shows lower hydrophobicity. In addition, both trifluorinated analogs of Ile, while sterically similar in size to the 2-Aha, exhibit lower hydrophobicity.

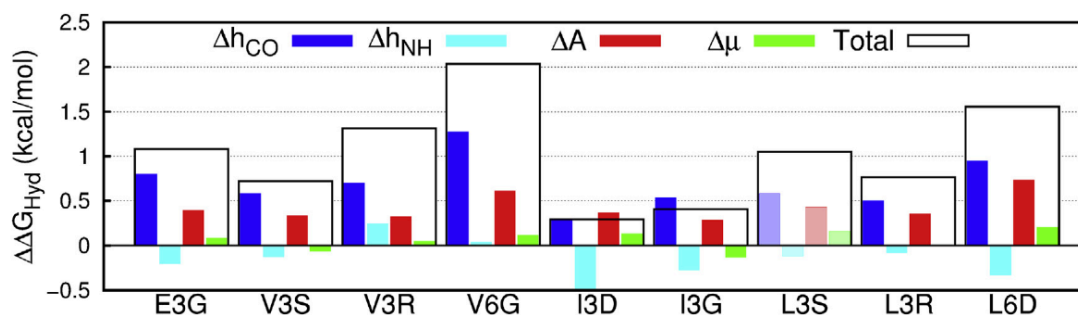
There are different explanations for the observed effects. First, the C-F bond is more polarized than the C-H bond and has an ionic character. This allows it to interact increasingly electrostatically with the environment. As previously discussed, the exchange of C-H to C-F leads to a significant spatial increase since the fluorine atom is 20% larger than the H-atom. Thus, an increase in hydration energy would be expected. Theoretical considerations using molecular dynamic (MD) simulations addressed the aspects of the factors contributing to hydrophobicity. In their first studies, Robalo *et al.* investigated the hydration energy of tri- and hexafluorinated AAs.<sup>[206]</sup> The corresponding AAs were acetylated at the N-term and amidated at the C-Term. Scheme 1.1 provides the structures and nomenclature used in MD simulation studies.



**Scheme 1.1:** Nomenclature of fluorinated structures used in MD-simulation studies by Robalo *et al.* <sup>[206,207]</sup>

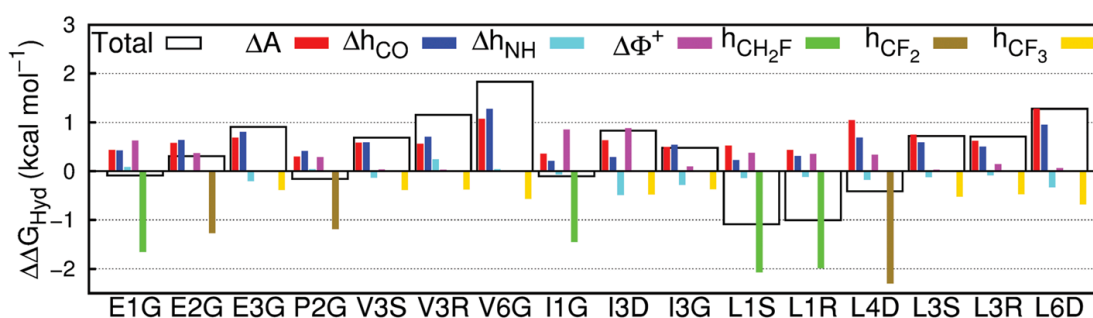
Here only a small contribution to hydrophobicity by an increased surface area ( $\Delta A$  red bar) was determined (Figure 1.15). The decisive factors of hydrophobicity are the H-bonds with the backbone amines and the carbonyl groups in water ( $\Delta h_{CO}$  and  $\Delta h_{NH}$ ). Overall, the number of possible H-bonds decreased for flAA, which increased the hydrophobicity. This factor also explains the slight difference in hydrophobicity

between the two TfVal isomers. On the other hand, the changes in the dipole moment account for a small proportion of the hydrophobicity.



**Figure 1.15:** Contribution of different factors to the overall free hydration energy ( $\Delta \Delta G_{Hyd}$  broad white bars) of fluorinated side-chains: Backbone-water H-bonds ( $\Delta h_{CO}$  dark blue bars and  $\Delta h_{NH}$  light blue bars), surface area ( $\Delta A$ , red bars) and dipole moment ( $\Delta \mu$ , green bars). Reprinted with permission from Robalo *et al.*<sup>[206]</sup> © 2017 Elsevier Inc.

In their second study, mono- and di-fluorinated AA analogs were included. The result showed that fluorination could not only lead to positive hydration energies (i.e., increase hydrophobicity) but also cause negative values (indicating a decrease in hydrophobicity), which is in a great alignment with the HPLC-based investigation.<sup>[207]</sup> It was demonstrated that the polarity of the side chains has a significant impact on the overall hydrophobicity. In the end, it was found that the effects of fluorine substitution remain challenging to predict, regardless of the degree of fluorination. For example, the total free hydration energy of DfeGly (E2G) is positive free hydration energy in contrast to DfpGly (P2G), although both possess a  $CF_2$  group as well.



**Figure 1.16:** further investigation of contributing factors of hydrophobicity measured by the overall free hydration energy ( $\Delta \Delta G_{Hyd}$  broad white bars): Backbone-water H-bonds ( $\Delta h_{CO}$  dark blue bars and  $\Delta h_{NH}$  light blue bars), surface area ( $\Delta A$ , red bars) and positive electrostatic potential ( $\Delta \Phi^+$ ) and hydrogen bonds between water and fluorine in mono- ( $h_{CH_2F}$ ), di- ( $h_{CF_2}$ ) or trifluorinated ( $h_{CF_3}$ ) AA. Reprinted from Robalo *et al.*<sup>[207]</sup>

These studies demonstrate that the hydrophobic character of fAAs depends on various factors, including the larger surrounding context (e.g., other AAs in the environment, possible H-bonds to the peptide backbone) in which a fAA is incorporated plays a pivotal role.

### 1.2.3.2 Secondary Structure Propensity of Fluorinated Amino Acids

Chakrabarty *et al.* established a simple model peptide that can be used to study  $\alpha$ -helix propensity. The AA of interest ( $X_{AA}$ ) is incorporated into an Ala-rich peptide (Ac-YGGKAAA $X_{AA}$ AAKAAA $X_{AA}$ AKAAAA-NH<sub>2</sub>) with no other helix-stabilizing interactions. The design involved an acetylated N-term and an amidated C-term to minimize electrostatic interactions. [208] Circular dichroism (CD) measurements then evaluate the helical content of the peptides. The average ellipticity of the residues at 222 nm ( $\theta_{222\text{nm}}$ ) is used to calculate the helix fraction ( $f_{\text{helix}}$ ). By utilizing the modified Lifson-Roig theory, the  $\alpha$ -helix propensity  $\omega$  of the AA is quantified. [209] Cheng *et al.* studied the  $\alpha$ -helical propensity of fAAs for the first time using this method. [210] Our group also applied this method to investigate various fAAs. [200,201,211] A summary of the results is shown in Table 1.4. From these data, it is clear that most of the fAAs generally have poor to no  $\alpha$ -helical propensity. Of all the fAAs studied, MfGly has the highest  $\alpha$ -helical propensity. Its value is almost twice as high as the  $\beta$ -branched AAs Ile and Val, which are usually found in  $\beta$ -sheet structures.

**Table 1.4:** Determination of  $\alpha$ -helix propensity  $\omega$  of fluorinated and non-fluorinated AA after incorporation into a model peptide. Mean residue ellipticity ( $\theta$ ), fraction helix ( $f_{\text{helix}}$ ), and helix propensity ( $\omega$ ).

$X_{AA}$	$[\theta]_{222\text{ nm}}$	$f_{\text{helix}}$	$\omega$
Abu <sup>[210]</sup>	-18100 ± 200	0.52 ± 0.01	1.22 ± 0.14
Ala <sup>[210]</sup>	-19100 ± 200	0.60 ± 0.01	1.46 ± 0.01
Val <sup>[201]</sup>	-13054 ± 452	0.38 ± 0.01	0.41 ± 0.04
Ile <sup>[211]</sup>	-13876 ± 186	0.40 ± 0.01	0.53 ± 0.05
Leu <sup>[210]</sup>	-17400 ± 200	0.50 ± 0.01	1.06 ± 0.12
MfGly <sup>[201]</sup>	-16814 ± 379	0.48 ± 0.01	0.87 ± 0.07
DfGly <sup>[201]</sup>	-13969 ± 569	0.40 ± 0.02	0.50 ± 0.06
TfGly <sup>[201]</sup>	-7469 ± 475	0.21 ± 0.01	0.06 ± 0.02
(2S,3S)TfVal <sup>[211]</sup>	-2685 ± 526	0.08 ± 0.01	0
(2S,3R)TfVal <sup>[211]</sup>	-3887 ± 547	0.11 ± 0.02	0
5 <sup>3</sup> -TfIle <sup>[200]</sup>	-10776 ± 216	0.31 ± 0.01	0.26 ± 0.03
4 <sup>3</sup> -TfIle <sup>[211]</sup>	-3602 ± 216	0.10 ± 0.01	0
HfLeu <sup>[210]</sup>	-8720 ± 200	0.25 ± 0.01	0.13 ± 0.02

One possible reason for the low  $\alpha$ -helix propensity is the hydrophobic character of the fAAs. As discussed before, the side chains of a fAA can also induce polarization, which might explain why MfeGly exhibits a higher  $\omega$  value. It is also suggested that the increased bulkiness of the fluorinated side chains and the proximity of the  $\text{CF}_3$  group to the peptide backbone negatively affect the formation of  $\alpha$ -helical structures. [211]

#### **1.2.4 Impact of Fluorinated Amino Acids on the Secondary Structure and Folding of Proteins**

As the previous chapters have shown, fluorine substitution can profoundly affect the properties of AAs. It is, therefore, not surprising that numerous studies are available that have investigated these properties and the fascinating interactions of fluorine in larger constructs in peptide and protein engineering. This section presents a selection of these studies and does not claim to be exhaustive. Highlights are introduced here that are necessary for a better understanding of the studies planned for the presented doctoral research projects. The reviews of the Kokschi and Montclare group are recommended for further detailed overviews. [108,111,112]

##### **1.2.4.1 Fluorinated Monomeric Secondary Structures**

Monomeric units of secondary structures are rarely found in nature. Only a few important studies are available on fluorine substitution (Table 1.5). For example, all studies on the  $\alpha$ -helix propensity of fAAs are based on an isolated model  $\alpha$ -helix (YCC, detailed information is provided in section 1.2.3.2). From these reports, fAAs appear to prefer  $\beta$ -structures. Clark *et al.* developed two facile  $\beta$ -hairpin systems, which they modified with HfLeu to investigate cross-strand interactions. [212] In the first scaffold, the fluorinated side chains were directed toward each other without facing any other H-bonds. The cyclization of the peptide archived this through disulfide bonds. The second scaffold was designed to investigate the fluorinated side chain pointing towards the exterior of the peptide.

Their study demonstrated by NMR analysis that HfLeu did not perturb the formation of the secondary structures. Determination of the cross-strand revealed that HfLeu-HfLeu interactions were weaker than Leu-Leu or Leu-HfLeu interactions, indicating that canonical Leu-Leu interactions have a higher propensity to assemble into  $\beta$ -hairpins than their highly fluorinated counterparts. As a possible reason, the

authors suggested that a complex interplay of polarizability, dipolar interactions, and hydrophobicity controls the cross-strand interactions.

**Table 1.5:** Overview of chosen monomeric constructs used to investigate the impact of fluorine substitution. The sequences are provided along with replaced residue (highlighted in orange), the number of fluorine atoms in the construct, and the synthetic method applied to obtain the desired peptides. @ in the single letter code refers to Ati (1,2-dihydro-3(6*H*)-pyridinone). Adapted and modified from Monkovic *et al.*<sup>[108]</sup>

construct	sequence	fIAA	No. of F atoms	synthetic method	secondary structure	reference
YCC	Ac-YGGKAAAAKA- X-AAKAAAAK-NH <sub>2</sub>	divers	divers	SPPS	$\alpha$ helix	section 1.2.3.2
HL1	QRLTNCCTXEG	HfLeu	6	SPPS	$\beta$ -hairpin	[212]
HH1	QRXTNCCTXEG	HfLeu	2 x 6	SPPS	$\beta$ -hairpin	[212]
AH1	QRATNCCTXEG	HfLeu	6	SPPS	$\beta$ -hairpin	[212]
HL2	L@VpAVTX	HfLeu	6	SPPS	$\beta$ -hairpin	[212]
HH2	X@VpAVTX	HfLeu	2 x 6	SPPS	$\beta$ -hairpin	[212]
HG	X@VpAVTG	HfLeu	6	SPPS	$\beta$ -hairpin	[212]
GH	G@VpAVTX	HfLeu	6	SPPS	$\beta$ -hairpin	[212]

#### 1.2.4.2 Fluorine's Role in Higher Organization of Secondary Structures

In the investigation of the higher organization of peptides, a plethora of different systems were used, summarized in the Table 1.6. The number of fluorine atoms was varied from 1 to 144. It was shown that even one fluorine atom is sufficient to influence the structural organization. A different orientation of the pyrrolidine ring of proline (*C $\gamma$ -exo* vs. *C $\gamma$ -endo* conformers) is possible due to the fluorine substitution. In the case of (2*S*,4*R*)-4-fluoroPro (*R*-FlPro) hyper stable collagen structures can be formed, whereas, in the case of (2*S*,4*S*)-4-fluoroPro (*S*-FlPro) no fiber formation is possible.<sup>[213]</sup>

Aromatic interactions were also applied in the construction of collagen fibers, Gottlieb and Cejas *et al.* exploited the preference of perfluoro-Phe for their hydrocarbon counterpart to create sticky ends.<sup>[214,215]</sup>

A major part of the research deals with coiled-coil motifs and the stabilizing effects observed through the global incorporation of fluorinated AA, such as HfLeu, TfLeu, and TfVal. Here the fluorous effect was also considered a possible explanation, but Buer's intensive studies with 4-helix bundles contradicted this theory and showed that hydrophobic interactions are the main stabilizing component.<sup>[216]</sup>

In general, a coiled-coil motif was always formed despite the higher spatial demand and the lower  $\alpha$ -helix propensity of fIAAs. Sometimes structures with higher

stability were obtained, and a higher organization of helices was determined in the fluorinated variants, e.g., tetramers instead of dimers of peptide F were reported by Bilgiçer *et al.* [217] Substitutions at single positions in one helix of heterodimeric systems were also explored. Here it was shown that the packing of fluorinated AA, i.e., its position within the hydrophobic core (*a* or *d* positions), plays a crucial role. [218–220]

**Table 1.6:** Summary of the peptidic systems used in studies with regard to a higher-order organization in coiled-coil and collagen motifs. SC = segment condensation, CRP = collagen-related peptide, CMP = collagen-model peptide, CLP = collagen-like peptide, CC = Coiled coil, RSI = residue-specific incorporation, \* = diastereomeric \*\*\* = shortened sequence, EF = electrophilic fragment

entry	construct	sequence	fIAA	No. of F atoms	synthetic method	secondary structure	reference
1	Dfp-CRP	(PPG) <sub>4</sub> -PXG-	(4,4)DfPro	2	SC + SPPS	CRP	[221]
2	Flp-CRP	(PPG) <sub>5</sub>	2S-FIPPro	1	SC + SPPS	CRP	
3	PfPhe-CMP	X-(GPO) <sub>n</sub> -F n = 5 or 10	PfPhe	5	SPPS	CMP	[214]
4	S/R-Flp-CMP	(X <sub>1</sub> -X <sub>2</sub> -Gly) <sub>7</sub>	3S-FIPPro+ 4R-FIPPro	14x1	SC + SPPS	CMP	[222]
5	PfPhe-CRP	X-(GPO) <sub>4</sub> -GPL- (GPO) <sub>5</sub> -F	PfPhe	5	SPPS	CLP	[215]
6	peptide 2	HNRMKQLEDK <b>V</b> EELLSKNASLE <b>Y</b> E <b>V</b> ARLKKLV <b>G</b> E	TfLeu*+ TfVal*	7x3	SPPS	CC	[223]
7	Q+TFL	***RVKEITFLKN TAPQMLRE <b>L</b> QET NAALQDVRE <b>L</b> LR Q <b>Q</b> SK <b>L</b>	5-TfLeu	7x3	RSI by biosynthesis	CC	[224]
8	C+TFL	***RAPQM <b>X</b> RE <b>X</b> QETNAALQDV RE <b>X</b> XRQ <b>Q</b> VKEIT F <b>X</b> KNTSK <b>X</b>	5-TfLeu	7x3	RSI by biosynthesis	CC	[224]
9	TfLeu-GCN4-p1d	RMKQ <b>X</b> EDK <b>V</b> EE <b>X</b> LSKNYH <b>X</b> ENE <b>V</b> AR <b>X</b> KKLVGER	TffLeu *	4x3	SPPS	CC	[225]
10	tfi-INL	***AARRSRARK LQR <b>X</b> KQLEDK <b>X</b> E ELLSKNYHLE <b>N</b> <b>X</b> ARLKKL <b>X</b> GER	5-TfIle	4x3	RSI by biosynthesis	CC	[226]
11	tfv-VNL	***AARRSRARK LQR <b>X</b> KQLEDK <b>X</b> E ELLSKNYHLE <b>N</b> <b>X</b> ARLKKL <b>X</b> GER	R-TfVal	4x3	RSI by biosynthesis	CC	
12	K8-DfeGly	Ac- RLEELRED <b>X</b> LESL RKKLA	DfeGly	2	SPPS	EF of CC	[218]
13	L9-DfeGly	Ac- RLEELREDK <b>X</b> ESL RKKLA	DfeGly	2	SPPS	EF of CC	



Table 1.6 continuing

14	K8TfeGly	Ac- RLEELREDXLESL RKKLA	TfeGly	3	SPPS	EF of CC	[218]
15	L9TfeGly	Ac- RLEELREDXLESL RKKLA	TfeGly	3	SPPS	EF of CC	
16	VPK-HfLeu	Abz- KVSALKEKVASLK EKVSA <del>X</del> KKEEVASL EEKVSALK	Hfleu	6	SPPS	VPK strand CC	[219]
17	VPK-R- TfVal		S-TfVal	3	SPPS	VPK strand CC	
18	VPK-S- TfVal		R-TfVal	3	SPPS	VPK strand CC	
19	VPK- 16aTfeGly	ZKVSALKEKVASL KEK <del>X</del> SALKEEVASL LEEKVSALK	TfeGly	3	SPPS	VPK strand CC	
20	VPK-16a- DfeGly		DfeGly	2	SPPS	VPK strand CC	[220]
21	VPK-16a- DfpGly		DfpGly	2	SPPS	VPK strand CC	
22	VPK- 19dTfeGly		TfeGly	3	SPPS	VPK strand CC	
23	VPK- 19dDfeGly	ZKVSALKEKVASL KEK <del>X</del> VSA <del>X</del> KKEEVA SLEEKVSALK	DfeGly	2	SPPS	VPK strand CC	
24	VPK- 19dDfpGly		DfpGly	2	SPPS	VPK strand CC	
25	FA1-92	***SGD <del>X</del> ENEYA Q <del>X</del> EREVRS <del>X</del> EDE AAE <del>X</del> EQKVS <del>R</del> <del>X</del> K NEIED <del>X</del> KAEIGD <del>X</del> NNT***	TfLeu *	7x3	RSI by biosynthesis	CC	[227]
26	SR-A1	***SGD <del>X</del> ENEYA Q <del>X</del> EREVRS <del>X</del> EDE	5-TffLeu	7x3	RSI by biosynthesis	CC	[228]
27	SR-A1	AAE <del>X</del> EQKVS <del>R</del> <del>X</del> K NEIED <del>X</del> KAEIGD <del>X</del> NNT***	4-TffLeu	7x3	RSI by biosynthesis	CC	
28	HFL-A1		HfLeu	7x6	RSI by biosynthesis	CC	[229]
29	F-peptide	CGGAQ <del>X</del> KKEXQ A <del>X</del> KKENAQ <del>X</del> KW EXQA <del>X</del> KKEXAQ	HfLeu	7x6 HF 14x6 FF	SPPS	CC	[217]
30	$\alpha_4F_6$		HfLeu	4x6x6	SPPS	4-helix bundle	
31	$\alpha_4F_{2(6, 24)}$		HfLeu	4x2x6	SPPS	4-helix bundle	
32	$\alpha_4F_{2(10, 20)}$		HfLeu	4x2x6	SPPS	4-helix bundle	
33	$\alpha_4F_{2(13, 17)}$	GNADE <del>X</del> YKEXED XQER <del>X</del> RK <del>X</del> RKK <del>X</del> RSG	HfLeu	4x2x6	SPPS	4-helix bundle	[216]
34	$\alpha_4F_{3a}$		HfLeu	4x3x6	SPPS	4-helix bundle	
35	$\alpha_4F_{3d}$		HfLeu	4x3x6	SPPS	4-helix bundle	
36	$\alpha_4F_{3af_3d}$		HfLeu+ TfeGly	4x3x6 + 4x3x3	SPPS	4-helix bundle	

### 1.2.4.3 Fluorine-modified Proteins

Due to limitations in the SPPS, so far, only a few fluorinated proteins have been synthesized chemically. Therefore, a major part of the studies explored fluorinated proteins synthesized by the biological machinery (Table 1.7). These methods do not always enable complete incorporation of the fAAs and require elaborate engineering of the biological processes.

In the proteins studied, it was generally found that no significant structural perturbation occurred. Regarding the effects of fluorine modification on the thermodynamic parameters of proteins, there are reports of an enhanced (entries 1-3, 5+6) as well as a reduced chemical or thermal stability (entries 10, 13) or unchanged thermostability (entry 12).

**Table 1.7:** Summary of synthesized fluorinated proteins. RSI = residue-specific incorporation, SSI = site-specific incorporation, SPI = selective pressure incorporation, QfLeu = (*S*)-5,5,5',5'-tetrafluoroleucine

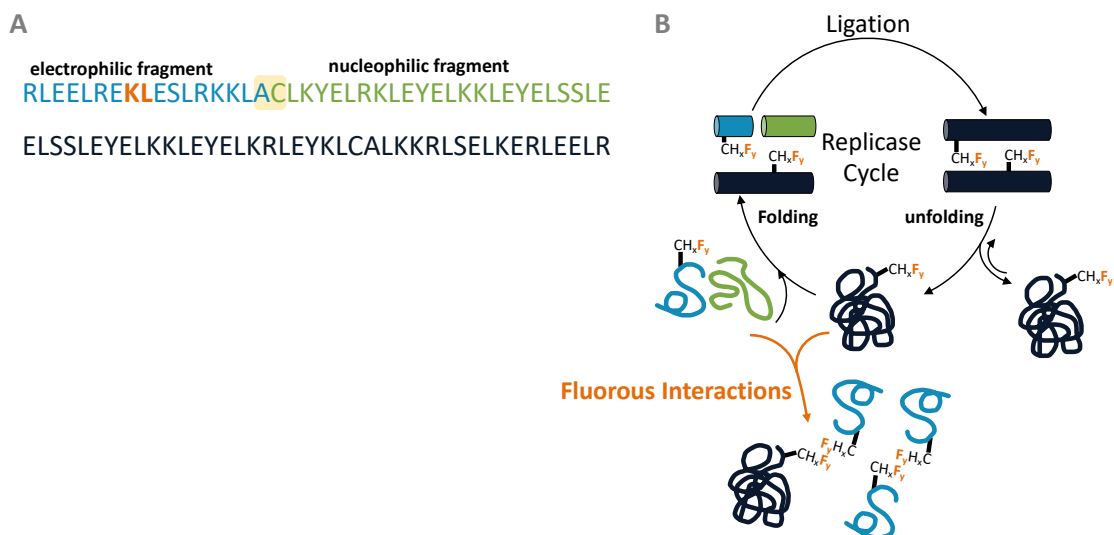
entry	protein	type and size of protein	fIAA	No. of F atoms	region	synthetic method	reference
1	GB1-Qfl	$\alpha+\beta$ (56)	QfLeu	4	$\beta$ -sheet	SPPS	[230]
2	GB1-Hfl	$\alpha+\beta$ (56)	Hfle	6	$\beta$ -sheet	SPPS	[230]
3	GB1-Pff	$\alpha+\beta$ (56)	PfPhe	5	$\beta$ -sheet	SPPS	[230]
4	GB1-Trp	$\alpha+\beta$ (56)	5-Trp	1	$\beta$ -strand	SSI	[231]
5	NTL9-V3	$\alpha+\beta$ (56)	TfVal	3	$\beta$ -sheet	SPPS	[232]
6	NTL9-V21	$\alpha+\beta$ (56)	TfVal	3	$\beta$ -sheet	SPPS	[232]
7	BsCspB-Phe	all- $\beta$ (67)	2/3/4/-FPhe	3x1	$\beta$ -barrel	RSI	[233]
8	BsCspB-Trp	all- $\beta$ (67)	4/5/6-FTrp	3x1	$\beta$ -barrel	RSI	[233]
9	GFP-F-Pro	all- $\beta$ (238)	4R or 4S-Fpro	10x1	$\beta$ -barrel	RSI	[234]
10	Ub-MfLeu	$\alpha+\beta$ (76)	5MfLeu	2x1	hydrophobic core	SPPS	[235]
11	Ub-FPro	$\alpha+\beta$ (76)	4R-Fpro	3x1	turns	SPI	[236]
12	HRAV-FTrp	all- $\alpha$ (320)	4/5/6-FTrp	1	turn	SSI	[237]
13	CAT-TfLeu	$\alpha+\beta$ (219)	5-TfLeu	up to 8x3	global	RSI	[121]

### 1.2.5 Impact of Fluorine Modifications on Kinetics of Peptide and Protein Folding

To date, there are only a few studies on fluorine's influence on folding kinetics. Some studies above examined the folding kinetics through Trp or Tyr stopped-flow fluorescence of proteins. In the case of BsCsp, fluorination only had minimal effects on the kinetics, which validated the thermodynamic stability.<sup>[233]</sup> Due to the different kinetics of isomerization of FPro compared to native Pro, slightly faster refolding kinetics and 2.6-fold slower unfolding kinetics in Ub-FPro were reported.<sup>[234]</sup> Stopped-flow fluorescence measurements are also applied to TfVal-modified NTL9-V3 and NTL9-V21.<sup>[232]</sup> As a result, a slower unfolding rate was detected for both fluorinated variants compared to the native protein indicating a stabilization of the native state by the CF<sub>3</sub> group. The folding rate was increased in both fluorinated derivatives, suggesting a stabilization of the transition state by hydrophobic interactions. In the case of GFP-FPro, the folding rate was determined by measuring the fluorescence recovery after chemical denaturation.<sup>[234]</sup>

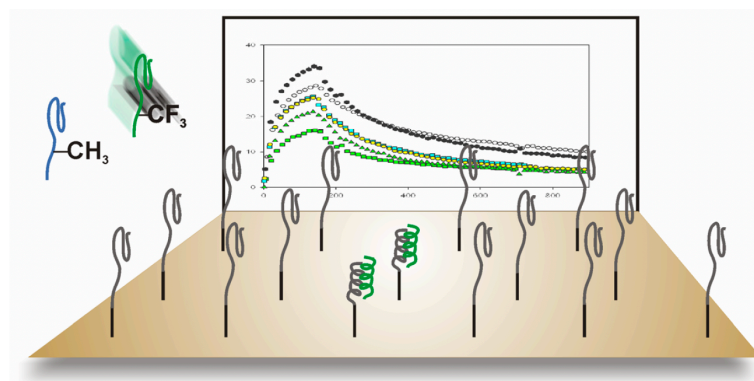
Our group developed two coiled-coil model systems to study the folding kinetics. The first system is based on an antiparallel homodimeric coiled coil in which the positions Lys8 and Lys9 have been substituted with Abu and the fluorinated derivatives DfeGly, TfeGly, and DfpGly, respectively.<sup>[238]</sup> To investigate the folding kinetics, the self-replication properties of coiled coils were exploited by template-directed ligation.<sup>[239,240]</sup> For this purpose, a helix was constructed from an electrophilic and nucleophilic fragment (Figure 1.17). The rate of product formation corresponded to the folding rate.

It was shown that the product formation decreased upon an increase in the degree of fluorination. For this observation, the authors suggested fluorine-induced interactions between fluorinated peptides in the unfolded strands.



**Figure 1.17:** Concept of folding investigations by Jäckel *et al.* **A)** Primary structures of model peptides. Orange-colored positions were substituted with DfeGly, DfpGly, and TfeGly, respectively. The yellow box refers to the ligation position. **B)** Schematic representation of autocatalyzed peptide ligation (replicase): dark blue cylinder = ligation product, green = nucleophilic fragment, blue = electrophilic fragment. Fluorous interaction (orange arrows) of fragments might interfere with the folding process. Adapted from Jäckel *et al.*,<sup>[238]</sup> Copyright © 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

The folding of the heterodimeric coiled coil system of VPE/VPK was examined by surface plasmon resonance.<sup>[241]</sup> For these studies, Abu and the fluorinated derivatives DfeGly, TfeGly, and DfpGly were substituted in the VPK strand at position *d* of the third heptad. The native VPE strand was immobilized on a gold surface by an N-terminal Cys, so the added VPK strands served as binding analytes (Figure 1.18). Dfe and Dfp-modified VPK strands showed a higher association rate than the non-fluorinated Abu derivative. TfeGly, on the other hand, showed low association, which led the authors to suggest possible F---F interactions in the unfolded state.



**Figure 1.18:** Schematic representation of experimental setup surface plasmon resonance. VPE-strand is indicated in blue, and the fluorine-modified VPK strands are colored in green. Coiled-coil formation can be monitored in real-time (the typical sensorgrams are depicted in the background). Reprinted with permission from Salwiczek *et al.* [241] Copyright © 2009 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

These systematic studies showed that influences could be exerted on the thermodynamics and the folding kinetics depending on the fluorinated side chain and the position in the primary structure. Therefore, choosing fIAA is critical concerning the planned study in suitable peptide or protein systems. In this context, properties such as the size, hydrophobicity, or secondary structure propensity of the fIAA, but also the immediate environment (e.g., other functional groups in the vicinity), play a decisive role. Therefore, the characteristics of the peptidic systems and proteins selected in this doctoral thesis will be outlined in the next section.

## 2 Model Systems Used in the Doctoral Thesis

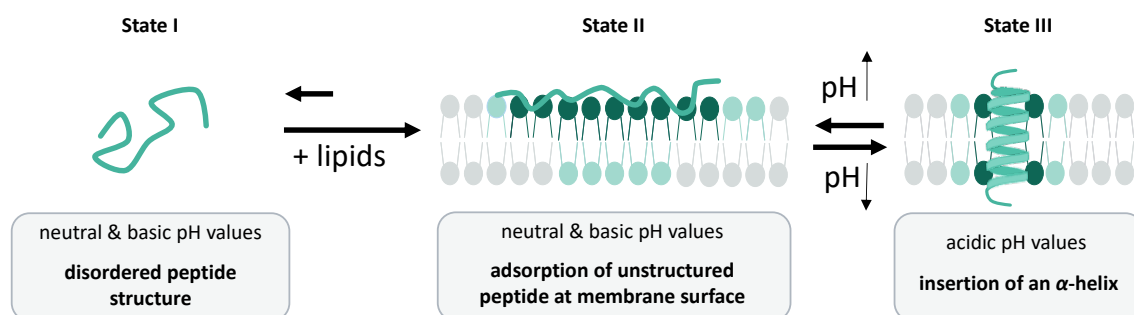
In this section, the model systems for the planned investigations carried out during this doctoral thesis are presented. All these well-tailored systems were used for different applications to study fluorine-specific interactions. A detailed discussion of required design changes or adaptations for the corresponding scheduled studies is given in the respective results sections.

### 2.1 Cell-penetrating Peptides: Fluorine-specific Interactions on Lipid Surfaces

In 1997, Englemann *et al.* developed a system based on the sequence of one of six transmembrane helices of bacteriorhodopsin, which behaves differently in a membrane environment depending on pH: pH (low) lipid insertion peptide (pHLIP).<sup>[242]</sup> The unique feature in the design of the primary structure of the peptide was the incorporation of Asp residues. These allow water solubility at neutral pH, and their degree of protonation influences the insertion into a lipid bilayer. The process is defined by three states (Figure 2.1): At neutral or high pH values, pHLIP is soluble in water at low concentrations of about 7  $\mu\text{M}$ . However, it does not adopt defined folding (**State I**), and even when a lipid membrane is added, initially, only adsorption of unstructured peptide occurs on its surface (**State II**). With a decrease in pH, the Asp residues of pHLIP are protonated. As a result, the peptide inserts into the lipid bilayer as an  $\alpha$ -helix (**State III**). The driving force here is the increased hydrophobicity of the pHLIP, caused by the protonation and, thus, neutralization of the Asp charges. This characteristic makes pHLIP an exciting candidate for drug carrier systems against cancer cells, as these cells typically exhibit an acidic milieu.<sup>[243,244]</sup> But pHLIP is also suitable for studies on interactions of peptides on a membrane surface.

On the one hand, fluorine can be incorporated as a sensor for *in situ* surface-enhanced infrared absorption spectroscopy (SEIRAS) because the C-F vibration is extremely sensitive to microenvironmental changes. In addition, the influence of the increased hydrophobicity and lipophilicity of fAAs on the folding process and, thereby, on the insertion of the  $\alpha$ -helix can be investigated here. In the further course of the study, model membranes carrying a fluorinated lipid tail can also be

harnessed, and therefore possible F---F interactions (between peptides and lipids) might be explored.



**Figure 2.1:** Schematic representation of pHLIP's interactions at a membrane surface. **State I** illustrates pHLIP as an unstructured motif at neutral and high pH values. After adding a lipid bilayer, adsorption of the unstructured peptide occurs at the surface (**State II**). The directly involved lipids are highlighted in dark petrol, while the neighboring lipids are marked in light petrol. As a consequence of the lowering of the pH, the hydrophobicity of the pHLIP is enhanced, leading to the insertion of the peptide into the membrane bilayer in the form of an  $\alpha$ -helix (**State III**). Adapted from Andreev *et al.* [245] Copyright (2010) National Academy of Sciences.

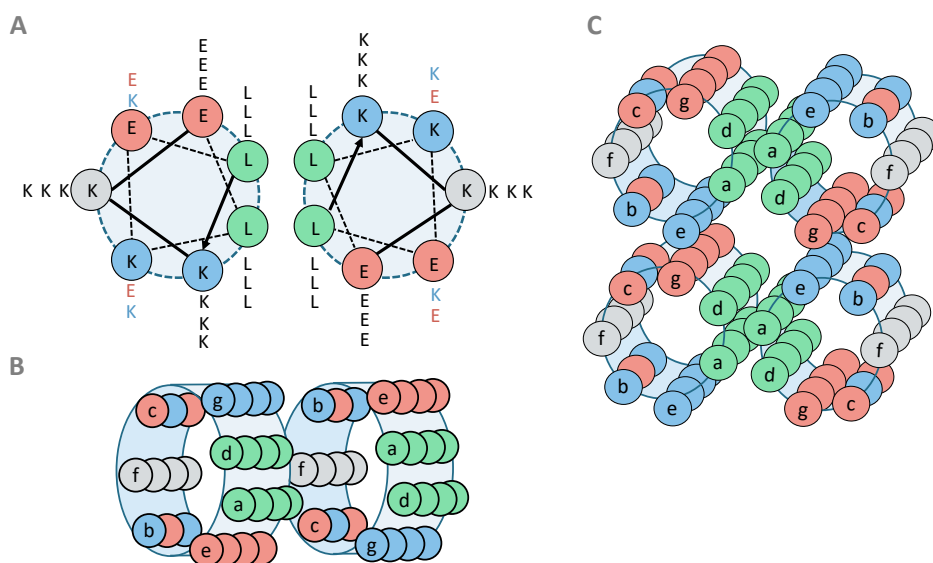
## 2.2 Coiled-coil Systems

As previously presented, coiled-coil systems offer a broad spectrum of applications. The unique features of coiled-coil motifs were explained in detail (in section 1.2.2.1). In the following section, two different models and their specifications will be described, which are employed in two projects in the present work.

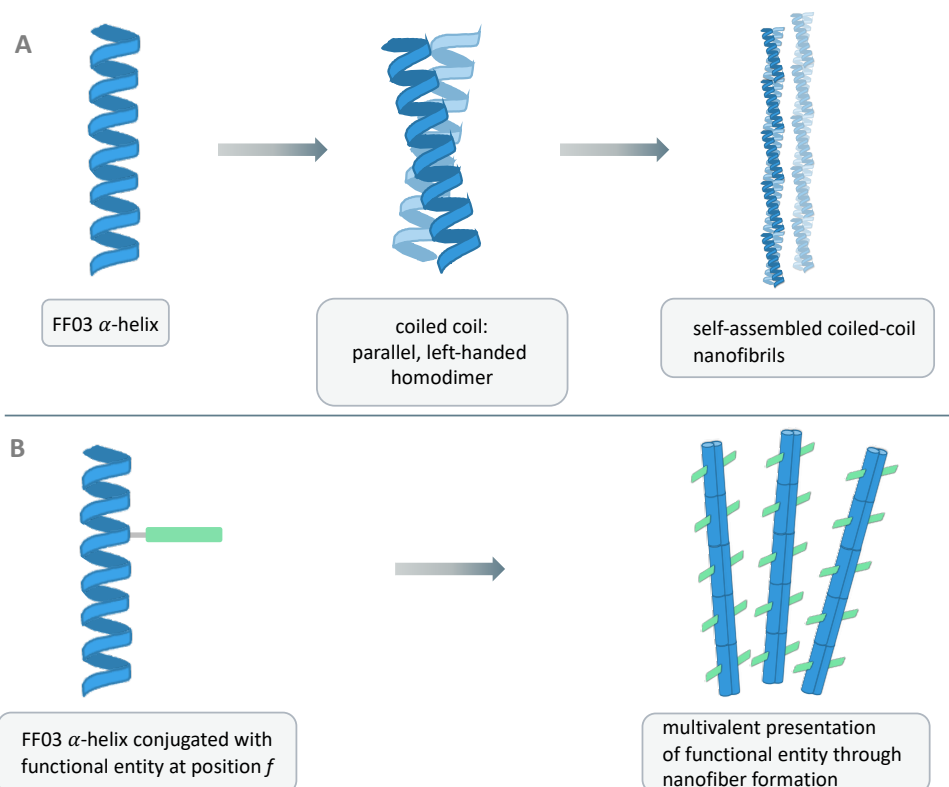
### 2.2.1 FF03 Scaffold as Carrier for *in vivo* $^{19}\text{F}$ -MRI Examinations

The fiber-forming peptide FF03 is a homodimeric coiled coil system. Its unique characteristic is the self-assembly into nanofibrils, enabled by a sticky-end design (Figure 2.2). Designed by Brandenburg in 2012, FF03 consists of 26 AAs with a triple heptad repeat. [246] Leu builds up the hydrophobic core of the coiled coil in positions *a* and *d*. Complementary electrostatic interactions of Glu and Lys in positions *e*, *g*, as well as *c*, and *b*, provide additional stabilization. The latter positions are also directed toward the solvent-exposed side, contributing to higher solubility. Position *f* contains only Lys, which can be substituted at the  $\epsilon$ -NH<sub>2</sub> group of the side chain. Thus, this system is an excellent scaffold for the multivalent presentation of epitopes, drugs, or sugar moieties. [167,168] Besides, the formation of nanofibers provides a high local concentration of the scaffold. This combination makes FF03 an ideal candidate for *in vivo* studies; a fluorinated hydrophobic core would allow FF03

to be explored in  $^{19}\text{F}$ -MRI. Even though  $^{19}\text{F}$  produces a sensitive signal, *in vivo* studies require the highest fluorine concentration possible.



**Figure 2.2:** Introduction of FF03-System: **A)** Helical wheel presentation of FF03 sequence and schematic representation of **B)** sticky ends concept (elongation via longitudinal axis) and **C)** electrostatic interactions between different FF03 strands to form fibers. Light blue cylinders illustrate single FF03 strands. Hydrophobic AAs are highlighted in green, positively charged AA in blue, negatively charged AA in red, and AA in solvent-exposed positions only are grey. Adapted with permission from Zacco *et al.* [168] Copyright © 2015, American Chemical Society.

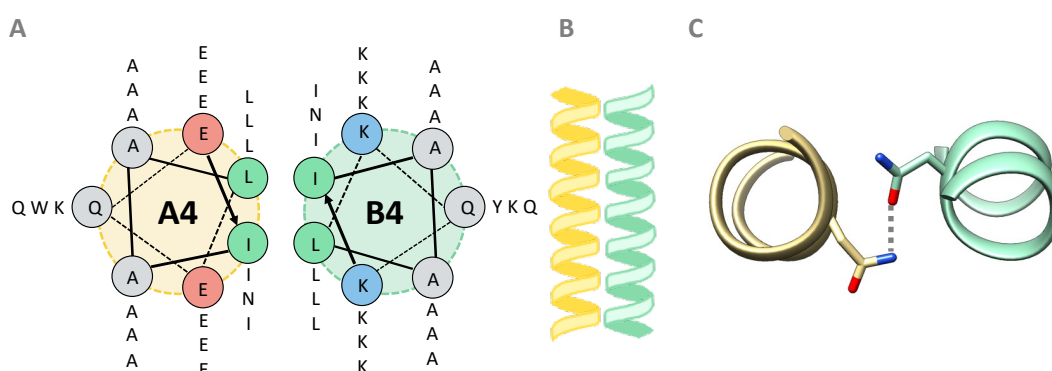


**Figure 2.3:** Illustration of nanofiber formation. **A)** Two FF03 strands form a helical coiled coil that forms nanofibers due to self-assembly. **B)** Conjugated FF03 serves as a scaffold for the multivalent presentation of a functional entity at *f*-position.



### 2.2.2 A4/B4 System a well-defined Coiled-coil Dimer for AMF Measurements

In contrast to the FF03 system, the A4/B4 system developed by Thomas *et al.* in 2013 consists of a parallel heterodimeric arrangement of an "acidic" strand (A4) with a "base" strand (B4) (Figure 2.4).<sup>[247]</sup> Another difference is that no higher-order structures and, consequently, no self-assembly have been described here. The individual strands consist of four heptad repeats with the pattern (EIAALEX) in the acidic strand and (KIAALKX) in the base strand. Inside the hydrophobic core,  $\beta$ -branched AAs are located in the *a* position (Ile), and  $\gamma$ -branched AAs in the *d* position (Leu). According to the Leu zipper design, an Asn was incorporated in the third heptad to allow a specification towards a dimer.<sup>[248]</sup>



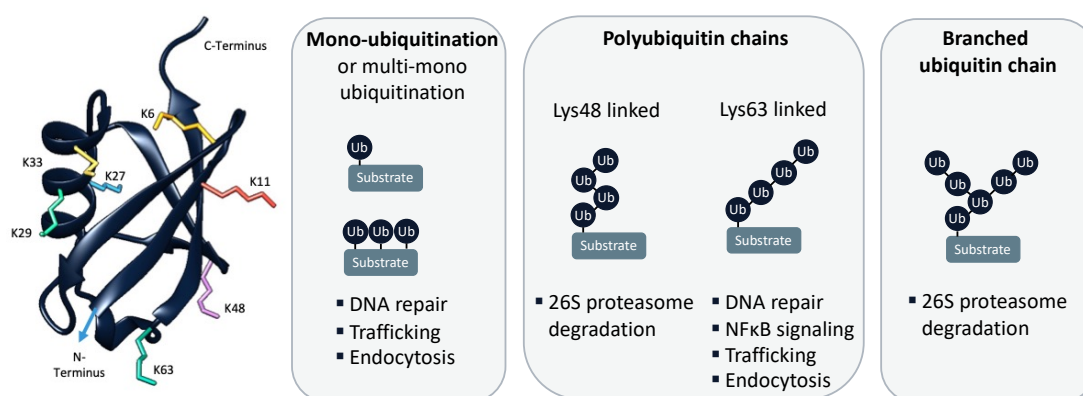
**Figure 2.4:** Features of A4/B4 system. **A)** Helical wheel representation of A4 and B4 sequences. Hydrophobic AAs are highlighted in green, positively charged AAs in blue, and negatively charged AA in red. **B)** Schematic representation of coiled-coil formation. **C)** Hydrophobic core of Asn containing heptad of the GCN4 leucine zipper; inter-strand hydrogen bond is indicated with a dotted line (PDB: 2ZTA).<sup>[248]</sup> Adapted with permission from Thomas *et al.*<sup>[247]</sup> Copyright © 2013, American Chemical Society. Created with <http://www.cgl.ucsf.edu/chimera>.

Although a buried Asn in the hydrophobic core is known to destabilize the coiled-coils overall structure, it is nevertheless a widespread motif in *de novo* designs of heterodimeric coiled-coil systems, as Asn can generally control the oligomerization of the coiled-coils towards dimers.<sup>[249–252]</sup> Equally charged AAs are presented in positions *e* and *g* on each peptide and thus, through oppositely charged peptides, form a complementary system. At positions *c* and *b*, Ala is used as a small, helix-favoring residue without additional electrostatic stabilization. In the *f* position, other charged AAs (Lys and Gln) are introduced to enhance solubility. Besides, an aromatic AA (Trp in A strand and Tyr in B strand) is incorporated at this position to allow concentration measurements. Moreover, the individual peptides are synthesized as amides at the C-terminus with acetylated N-terminus to avoid further

charges and, ultimately, unwanted interactions. Altogether, the design laid the foundation for a well-defined parallel heterodimer. Since this dimeric system has only a few interactions, it is ideally suited for investigating the effect of a single modification on the mechanical stability of the entire structure. In addition, it is also possible to incorporate fAAs facing each other and to investigate the potential effects of fluorine-fluorine interactions.

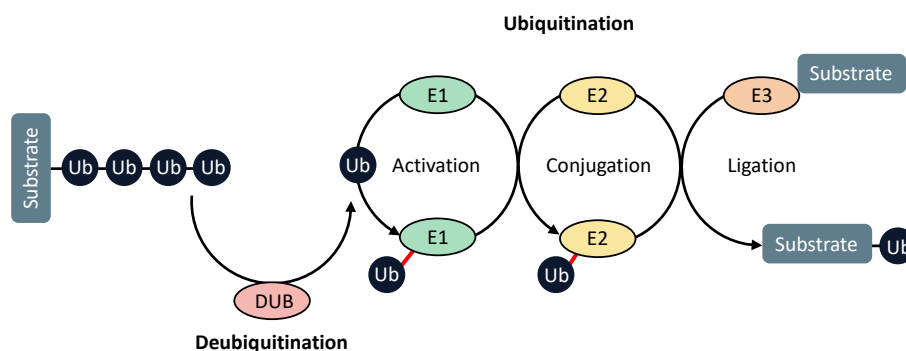
### 2.3 Ubiquitin- Oligomerization Ruled by Hydrophobic Patches

Ub is a small globular  $\alpha+\beta$  protein (76 AAs, 8.5 kDa) found in all eukaryotic cells. [253] It has a highly conserved primary structure across all orthologs. This is evident, for example, when comparing human sequences with those of yeast or oats, where the difference is only three AA. [254] Its structure consists of a five-stranded mixed  $\beta$ -sheet, a short  $3_{10}$  helix, and a 3.5-turn  $\alpha$ -helix (Figure 2.5). [255] The key feature of Ub is its stable, rigid H-bond network, which protects it from pH values between 1.18 and 8.48 and temperature degradation between 23-80°C. Additionally, high stability towards proteolytic digestion by trypsin, chymotrypsin, and carboxypeptidase of Ub is reported despite the presence of Arg or Lys residues. [256-258] A flexible six-residue C-terminal tail is essential for Ub conjugation. Ub has seven Lys residues Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, and Lys63, through which the protein can be bound to other proteins or to itself (oligomerization, Figure 2.5).



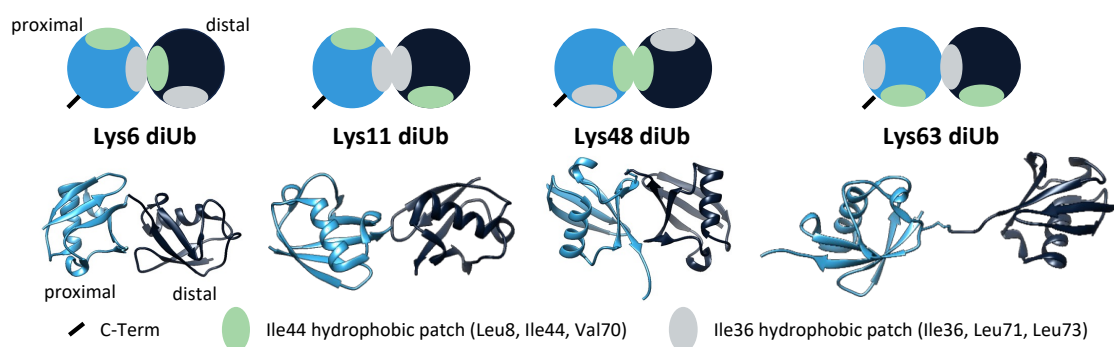
**Figure 2.5:** The Ub-system. Crystal structure of Ub (PDB-ID: 1UBQ) [255] with its seven Lys residues (differentially color-coded). Different patterns of ubiquitination of a substrate and selected physiological responses. Adapted with permission from Buetow *et al.* [259-261] Copyright © 2016, Nature Publishing Group, a division of Macmillan Publishers Limited. Created with <http://www.cgl.ucsf.edu/chimera>.

This process is also called ubiquitination, an essential post-translational modification of proteins (Figure 2.6). In this process, the physiological effect is determined by the linkage of different Ub chains. For instance, the Lys48 conjugation of four Ub residues provides the signal for the digestion of the target protein in the proteasome.



**Figure 2.6:** Schematic representation of the enzymatic machinery involved in the Ub system. Deubiquitinases (DUB) separate Ub units without digesting the respective Ub. There are three different classes of enzymes involved in the ubiquitination process: Ub-activating (E1), Ub-conjugating (E2), and Ub-ligating enzyme (E3). Adapted from Kliza *et al.* [262]

A unique element of Ub is the formation of two solvent-exposed hydrophobic regions, so-called hydrophobic patches (Leu8, Ile44, Val70 and Ile36, Leu71, Leu73). [263,264] These patches are crucial for Ub-binding enzymes and a critical component in the oligomerization behavior of Ub. Depending on the linkage, the orientation of these hydrophobic patches to each other varies (Figure 2.7).



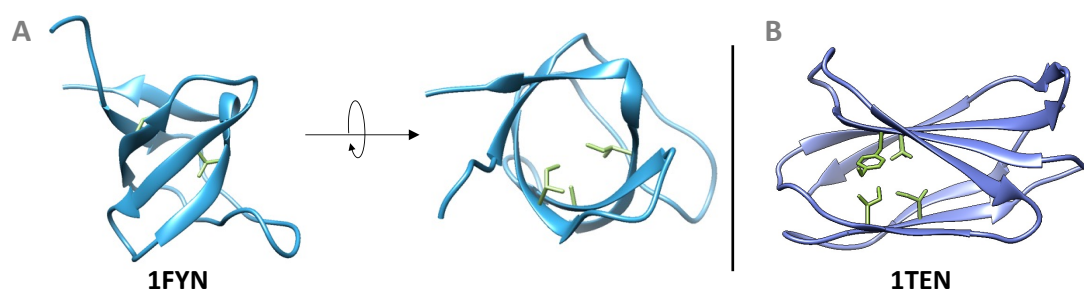
**Figure 2.7:** Orientation of hydrophobic patches in different Ub dimers. The distal Ub unit (linked with C-term) is depicted in dark blue, while the proximal (C-Term + Lys residue bearing) is defined in light blue. The Ile44 hydrophobic patch is represented in green, and the Ile36-hydrophobic patch in grey. PDB-ID: Lys6 diUb (PDB-ID: 2XK5), [265] Lys11 diUb (PDB-ID: 2MBQ), [266] Lys48 diUb (PDB-ID: 1AAR), [267] and Lys63 diUb (PDB-ID: 2JF5). [268] Adapted from Shekhawat *et al.* [139] Created with <http://www.cgl.ucsf.edu/chimera>.

These hydrophobic patches are a suitable target for systematic investigation for fluorine modification. In the Lys48 diUb, the patches are directed towards each

other, which allows the exploration of the oligomerization behavior of fluorinated patches or a possible preorganization through F---F interaction, as observed in Jäckel *et al.*, to control folding kinetics (see 1.2.5).<sup>[238]</sup>

## 2.4 Small all- $\beta$ -proteins: Folding of Barrel and Sandwich Structures

Apart from GFP and BsCsp, no known all- $\beta$  proteins have been studied with respect to protein folding, especially modifications of the hydrophobic core with fluorinated aliphatic AAs. This gap will be addressed in the present doctoral thesis. For this purpose, two "small" protein domains (1FYN and 1TEN, Figure 2.8) were selected that are frequently studied in kinetics studies of protein folding using stopped-flow Trp fluorescence (Figure 2.8). In both systems, a Trp is located in the proximity of the hydrophobic core. Since this AA is very sensitive to changes in the microenvironment, both selected all- $\beta$  structures are ideally suited for systematic studies of fluorine modifications in the hydrophobic core on protein folding kinetics. 1FYN is the phosphotransferase region of the SH3 domain. It is a 62 AAs long all- $\beta$  protein segment with a barrel structure consisting of five anti-parallel  $\beta$  sheets. In the hydrophobic core, three AA are responsible for the correct folding (Ile28, Ala39, Ile50).<sup>[269]</sup> The interaction with Pro-rich peptides (PXXP) is essential for signal transduction of SH3 via the 1FYN domain (transfer of a phosphor).<sup>[270]</sup>



**Figure 2.8:** Structures of selected all- $\beta$  proteins. **A)** 1FYN (PDB-ID)<sup>[270]</sup> and **B)** 1TEN (PDB-ID).<sup>[271]</sup> Essential AAs in the hydrophobic core are highlighted in green. Created with <http://www.cgl.ucsf.edu/chimera>.

1TEN is the fibronectin type III domain of human tenascin, a protein found in the extracellular matrix, consists of 90 AAs, and has a  $\beta$ -sandwich fold. The vital AA of the hydrophobic core are Ile20, Tyr36, Ile59, and Val70. This system is also suitable for investigating the influences of aliphatic and aromatic fAAs on protein folding.<sup>[272]</sup>

Both systems have no disulfide bonds, and protein folding proceeds two-stately.

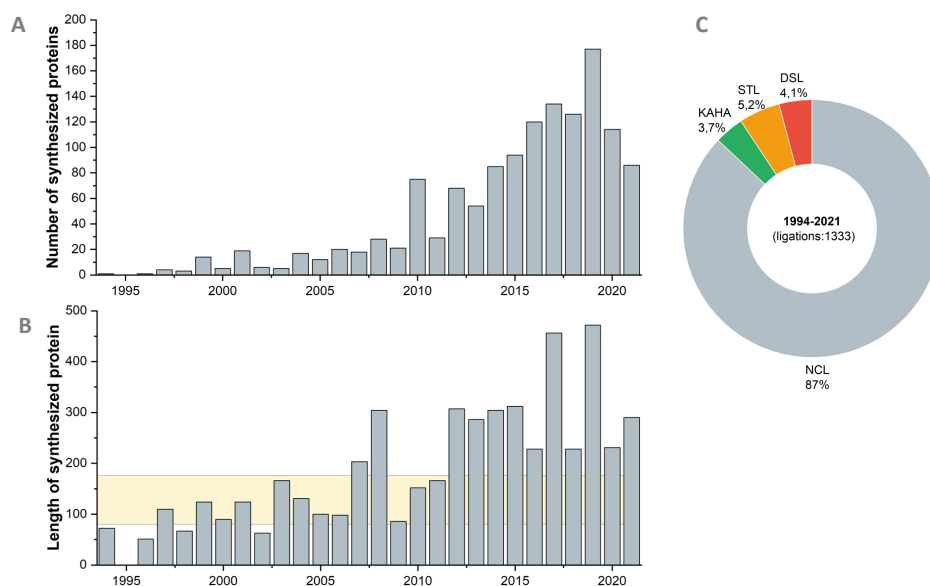
### 3 Strategies in Chemical Protein Synthesis

With a linear sequence length of 90 AAs and 152 AAs in the Ub dimer, the selected model systems represent a significant challenge for stepwise SPPS. Encouragingly, substantial improvements in SPPS have been achieved. Among these is applying elevated temperature in combination with microwave irradiation (MW-SPPS) or automated flow peptide synthesis (AF-SPPS), which can be used to obtain sequences up to 164 AAs.<sup>[273]</sup> However, AF-SPPS is not widely available and requires a custom-made instrument setup. In this dissertation, long linear proteins were prepared using MW-SPPS (details are described in 5.4.2), and their synthesis protocols are discussed in detail in the results section. Further strategies are essential for synthesizing ubiquitin dimers, which will be presented in the following section.

#### 3.1 Ligation Techniques: A Boon in Chemical Protein Synthesis

Research on the total chemical synthesis of proteins began in the Emil Fischer era. Since then, several key milestones have been elaborated, enabling researchers today to synthesize a wide variety of landscapes in biomacromolecules. These milestones include the establishment of SPPS by Merrifield,<sup>[274]</sup> the development of orthogonal protecting group strategies,<sup>[275]</sup> various coupling reagents,<sup>[276]</sup> and the automation of SPPS. Nevertheless, due to inefficient results, chemical access to peptides with a length of 50 AAs and more remained limited.

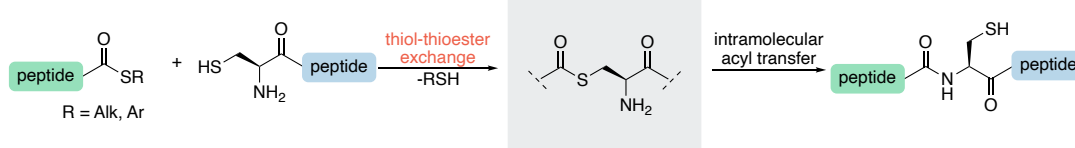
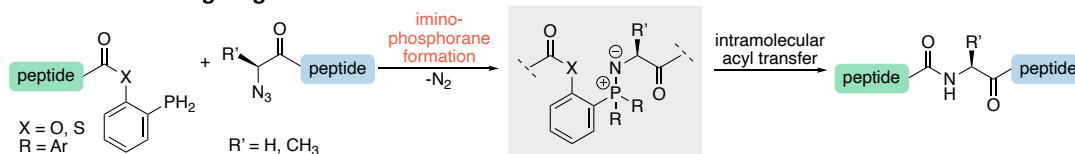
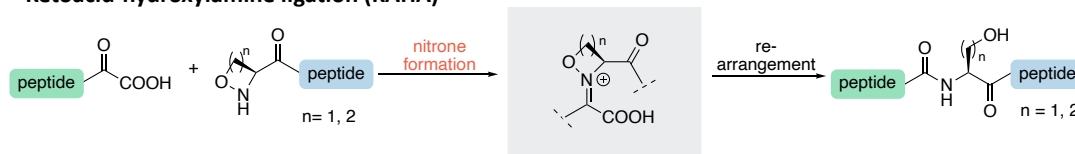
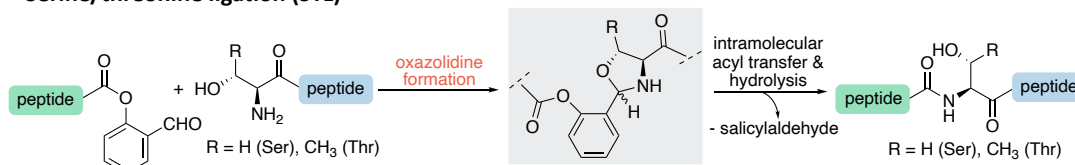
A paradigm shift was brought about by the publication of Kent *et al.* in 1994, who succeeded in conjugating unprotected peptide fragments in an aqueous milieu by a chemoselective thiol-thioester reaction, calling it the native chemical ligation (NCL).<sup>[277]</sup> Since then, the chemical synthesis of proteins has increased steadily and made advancements (Figure 3.1 A), with the length of proteins synthesized using at least one ligation step growing as well (Figure 3.1 B, up to approx. 470 AAs in 2019).<sup>[278]</sup> Besides, the field of unique chemo-selective reactions was also broadened (Figure 3.1 C): ketoacid-hydroxylamine ligation (KAHA),<sup>[279]</sup> diselenide selenoester ligation (DSL)<sup>[280]</sup>, and serine/threonine ligation (STL) were developed.<sup>[281]</sup> Agouridas *et al.* have classified these methods into three categories, depending on the functional group-bearing position required for the corresponding reaction (Figure 3.2).<sup>[282]</sup>



**Figure 3.1:** Chemical protein synthesis per year from 1994 to 2021. **A)** Number of synthesized proteins; **B)** length of synthesized proteins; the yellow box indicates the size region of functional protein domains; <sup>[283]</sup> **C)** Prevalence of applied ligation techniques: NCL, DSL, STL, and KAHA. Illustration adapted with permission and modified from Agouridas *et al.* <sup>[282]</sup> Source for used data extraction: Protein chemical synthesis (PCS) database (<http://pcs-db.fr/>). <sup>[284]</sup> Semi-synthesis was excluded from the dataset. Note: the PCS database provides only data on peptides and proteins of biological relevance; it does not include ligation reactions used to synthesize model peptides, polymers, or hybrid material. Copyright © 2019, American Chemical Society.

In type **I** (native chemical ligation (NCL), Figure 3.2 A), the ligations are performed through thiol functionalities at the side chain of an N-terminal cysteinyl peptide (blue) and a C-terminal thioester (green). The intermediate is generated by a thiol-thioester exchange followed by a spontaneous irreversible intramolecular S-to-N acyl transfer to produce a native peptide bond. In 2015, the DSL was developed by Mitchell *et al.* <sup>[280]</sup> Although the exact mechanism of this ligation is still not fully understood, the DSL is classified in this category (not shown in fig).

The type **II** ligations are characterized by the involvement of the  $\alpha$ -nitrogen (Figure 3.2 B). Traceless Staudinger ligation, introduced in 2000, <sup>[285,286]</sup> and KAHA ligation, developed in 2006, <sup>[279]</sup> belong to this category. In the traceless Staudinger ligation, a phosphine-functionalized peptide fragment (green) reacts with an azide component (blue) under nitrogen elimination to form an imino-phosphorane. Again, an intramolecular acyl transfer occurs, creating the native peptide bond. In KAHA ligation, a nitron is formed by ketoacids (green) and hydroxylamines (blue). The desired peptide backbone is built up by decarboxylation and rearrangement.

**A Type I (side-chain)****Native chemical ligation (NCL)****B Type II ( $\alpha$ -nitrogen)****Traceless Staudinger ligation****Ketoacid-hydroxylamine ligation (KAHA)****C Type III ( $\alpha$ -amino group & side-chain)****Serine/threonine ligation (STL)**

**Figure 3.2:** Overview of amide bond forming ligation methods applied in chemical protein synthesis. Green and blue boxes indicate unprotected peptides, i.e., with free functional groups on the side chains. The methods are categorized according to the functional groups of the amine component (blue peptide) involved in chemo-selective reactions. After intramolecular rearrangements, the desired peptide bond is formed (intermediates are shown in grey boxes). **A**) Type I (side-chain): native chemical ligation (NCL)<sup>[277]</sup> and diselenide selenoester ligation (DSL, not shown); **B**) Type II ( $\alpha$ -nitrogen): traceless Staudinger ligations<sup>[285,286]</sup> and ketoacid-hydroxylamine (KAHA); **C**) Type III ( $\alpha$ -amino group + side-chain): serine/threonine ligation (STL). Adapted with permission from Agouridas *et al.*<sup>[282]</sup> Copyright © 2019, American Chemical Society.

Both positions ( $\alpha$ -amino group and side chain) are used in type III ligations (Figure 3.2 C). This category includes serine/threonine ligation, inspired by the concept of pseudoproline ligation published by Tam *et al.* in the 1990s.<sup>[287–289]</sup> The core element is the chemo-selective formation of an oxazolidine by an aldehyde (green) fragment and the hydroxyl group at the side chain of Ser or Thr (blue). After acyl rearrangement and hydrolysis, the peptide bond is formed.

All these methods show key achievements for accessing synthetic proteins, using a wide variety of chemo-selective reactions under mild conditions while tolerating a broad range of functional groups of the side chains of AAs. Detailed reviews by the Becker,<sup>[290]</sup> Hackenberger,<sup>[291]</sup> Li,<sup>[292]</sup> Melnyk,<sup>[282]</sup> and Payne<sup>[293]</sup> groups show the plethora of possible applications of the presented methods and discuss more

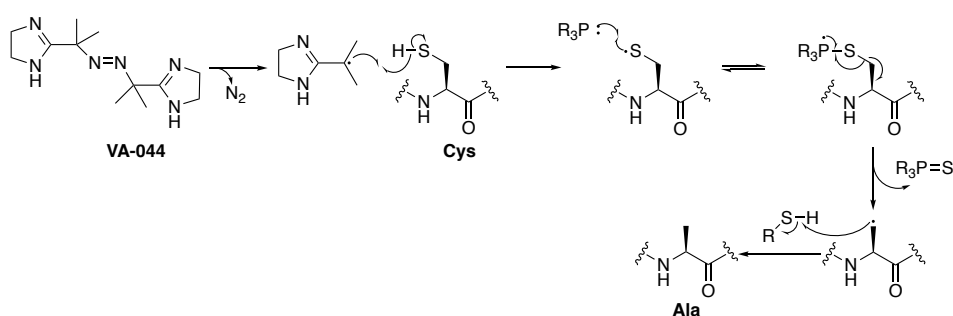
techniques such as enzymatic involvement or the use of different auxiliaries in ligation strategies. Jaradat's review provides an excellent overview of key achievements in peptide synthesis over the past thirteen decades. [294]

The remainder of this section will focus on native chemical ligation, representing the most significant proportion of published strategies for chemical protein synthesis (Figure 3.1 A) and chosen technique of this doctoral thesis.

### 3.2 Native Chemical Ligation: New Paths in Protein Landscapes

The low natural occurrence of Cys in natural proteins (1.3% in nature) [295] initially posed a limitation to native chemical ligation.

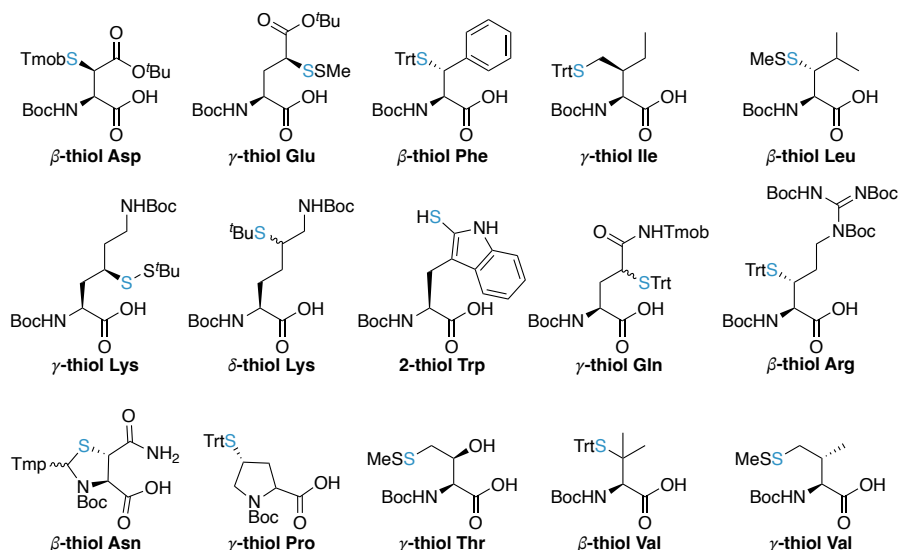
Therefore, research was conducted to find a suitable way to convert the thiol functionality of Cys into an Ala residue (9,0 % natural occurrence). [295] First, studies explored metal-based radical desulfurization reactions with Pd/Al<sub>2</sub>O<sub>3</sub>, Pd/C, Pd/BaSO<sub>4</sub>, and RANEY® nickel. However, the results showed low yields in product formation, which was assumed to be caused by aggregation with the peptide and adsorption processes on the metal surface. [296] In addition, side reactions such as hydrogenation of Trp and demethylthiolation of Met were observed. [297] Finally, Danishefsky *et al.* established a modified protocol of the one initially reported by Valencia *et al.* [298] harnessing VA-044 as a radical initiator with TCEP and *t*BuSH in water at room temperature for desulfurization. [299] The postulated mechanism is shown in Scheme 3.1.



**Scheme 3.1:** Proposed mechanism of radical desulfurization. [299]

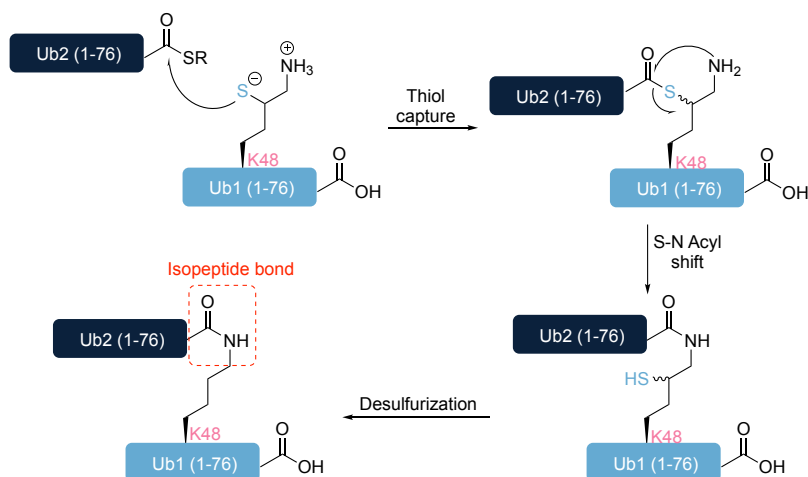
Efficient desulfurization allows a broad linkage field in which a thiol group can be inserted at the desired ligation site and later removed. Various groups have developed synthesis protocols to produce different thiolated AAs (Figure 3.3).





**Figure 3.3:** Toolbox of Fmoc-SPPS compatible thiolated AA. <sup>[300–302]</sup> Note: the amino function is Boc-protected as these building blocks are usually introduced at the N-terminal position:  $\beta$ -thiol Asp, <sup>[303,304]</sup>  $\gamma$ -thiol Glu, <sup>[305]</sup>  $\beta$ -thiol Phe, <sup>[306]</sup>  $\gamma$ -thiol Ile, <sup>[307]</sup>  $\beta$ -thiol Leu, <sup>[308,309]</sup>  $\gamma$ -thiol Lys, <sup>[310,311]</sup>  $\delta$ -thiol Lys, <sup>[312,313]</sup> 2-thiol-Trp, <sup>[314]</sup>  $\gamma$ -thiol Gln, <sup>[315]</sup>  $\beta$ -thiol Arg, <sup>[316]</sup>  $\beta$ -thiol Asn, <sup>[317]</sup>  $\gamma$ -thiol Pro, <sup>[318,319]</sup>  $\gamma$ -thiol Thr, <sup>[320]</sup>  $\beta$ -thiol Val, <sup>[321]</sup>  $\gamma$ -thiol Val. <sup>[322]</sup>

One important compound,  $\delta$ -thiol Lys (also referred to as  $\delta$ -mercapto-Lys) is employed to synthesize *iso*-peptide bond-linked peptides/proteins such as Ub-dimers (Scheme 3.2).



**Scheme 3.2:** Strategy to synthesize native Ub dimers via isopeptide linkage by using  $\delta$ -mercapto Lys. <sup>[313,323]</sup>

Another limitation in NCL is the use of thioesters in the Fmoc strategy of SPPS. These are reactive to oxygen, nitrogen, or sulfur nucleophiles. In particular, thioesters are vulnerable to bases, making it more challenging to produce C-terminal thioesters by the Fmoc strategy, as they would not survive the frequent use of piperidine. <sup>[324,325]</sup> In contrast, thioesters are highly stable to strong acids such as TFA or HF. However,

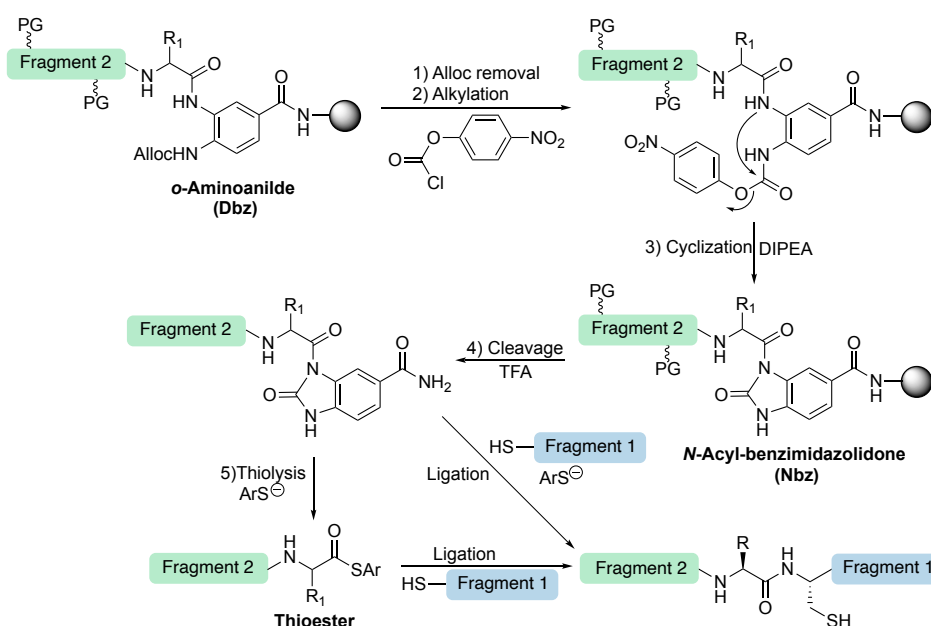
the Boc strategy raises safety concerns due to the need for anhydrous HF.<sup>[296]</sup> Therefore, in the Fmoc strategy, precursor functionalities were introduced to overcome the limitation posed by the reactivity of the thioester. These functionalities, so-called thioester surrogates or latent/masked thioesters, withstand the peptide assembly conditions on the solid phase and are converted into the desired thioester at a later stage. They include *N*-acyl-ureas (Dbz-linkers),<sup>[326]</sup> hydrazides,<sup>[327]</sup> and the bis(2-sulfanylethyl)amido (SEA) group.<sup>[328,329]</sup> All the strategies shown allow an *in situ* transfer into the desired thioester, reducing the need for purification steps. It is noteworthy that the thiolysis can be performed with both aryl and alkyl thiols. However, it was shown that the aryl components have a higher reaction rate.<sup>[330]</sup>

### 3.2.1 Diaminobenzyl (Dbz) Linkers

*N*-Acyl ureas are among the most widely used functionalities for synthesizing thioesters. In 2008, *o*-aminoanilides (also diaminobenzyl (Dbz) linkers) were developed by Blanco-Canosa *et al.* for application on Rink-amide resins.<sup>[326]</sup> In general, the *p*-amino group is protected by an Alloc group to prevent unwanted side reactions during the assembly of the peptide on the solid support. After complete synthesis of the desired peptide fragment, the Alloc protecting group is removed before cleavage of the peptide from the resin under standard conditions (Pd(Ph<sub>3</sub>P)<sub>4</sub> in CHCl<sub>3</sub>/AcOH/ *N*-methyl morpholine). Next, a two-step cyclization procedure to form the *N*-acyl-benzimidazolidone (Nbz) follows (Scheme 3.3). The deprotected amino group is alkylated with *p*-nitrophenyl chloroformate and cyclized in the second step by treatment with DIPEA. At this point, it should be noted that for both stages, the reaction time and concentration of the reagents must be adjusted for each individual peptide fragment. Only after cyclization, the peptide is cleaved from the resin by TFA treatment and purified by HPLC. Subsequently, the desired thioester can be isolated after thiolysis or directly reacted in a ligation reaction with the thiol peptide fragment and, for example, mercapto phenylacetic acid (MPAA), because the Nbz group is labile to thiolysis under neutral conditions.

Proteins that have been produced by using Dbz/Nbz methods include ubiquitin oligomers,<sup>[323]</sup> cyclotides such as sunflower trypsin inhibitor 1,<sup>[331]</sup> the human

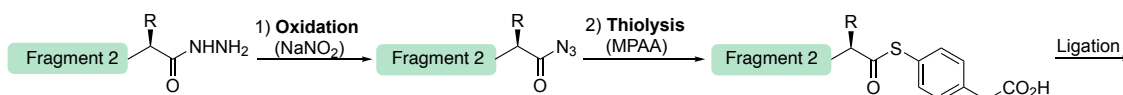
selenoproteins M, W<sup>[332]</sup> and F,<sup>[333]</sup>  $\alpha$ -synuclein,<sup>[334]</sup> and the transactivation domain of p53.<sup>[335]</sup>



**Scheme 3.3:** Diaminobenzyl linkers as masked thioesters. The Dbz-group is transferred to the corresponding resin-bound Nbz moiety by Alloc removal and subsequent two-step cyclization procedure by alkylation with p-nitrophenyl chloroformate and treatment with DIPEA. In the Next step, the desired unprotected peptide-Nbz fragment is generated by TFA cleavage and purification. This compound can, in turn, be either directly ligated by *in situ* thiolyis and the thiol-peptide fragment or the desired thioester is isolated after thiolyis. PG = protecting group. Adapted with permission from Blanco-Canosa *et al.*<sup>[326]</sup> Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

### 3.2.2 Hydrazides

Although protected peptide hydrazides were reported as early as the 1970s,<sup>[336,337]</sup> it was not until 2011 that Liu's group succeeded in exploiting this functionality to ligate unprotected peptide fragments.<sup>[327]</sup> In this method, oxidation with sodium nitrite generates the corresponding azide (Scheme 3.4) in an acidic pH (3.0 to 4.0), which then can be converted to the desired thioester via thiolyis under mildly acidic conditions (pH 5 to 6.2, note: higher pH values might lead to hydrolysis of the thioester). In this reaction sequence, the oxidation must proceed at low temperatures ( $-20$  to  $-10^\circ\text{C}$ ) to avoid the Curtius rearrangement into the corresponding isocyanate.<sup>[338]</sup>

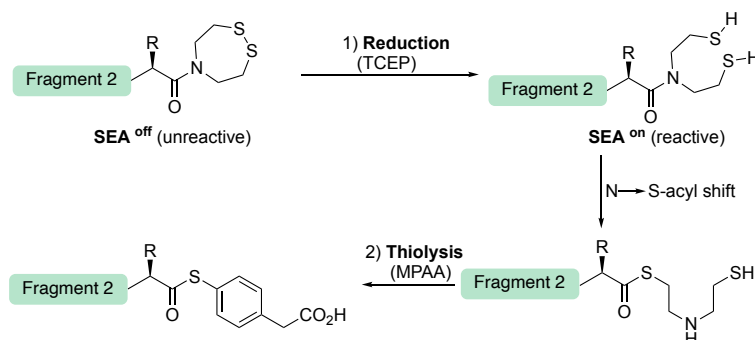


**Scheme 3.4:** Hydrazides as thioester surrogates. Oxidation of the desired peptide fragments provides the corresponding azide under acidic conditions at low temperatures. Following thiolysis with MPAA furnishes the desired thioester for the NCL.

Hydrazides are the second most common form of thioester surrogates. They have been applied, for example, in the synthesis of ubiquitin chains [339–341] or various peptidic toxins such as mambalgin-1 [342] or scorpion toxin AaH-II. [343] More detailed applications of hydrazide intermediates in chemical protein synthesis are described in a review by Lui *et al.* [344]

### 3.2.3 Bis(2-sulfanylethyl)amido (SEA) group

The SEA group is also a latent thioester. Still, the mode of operation differs from the two previously mentioned functionalities because its transfer to the thioester involves a reverse N-to-S-acyl shift system. Such systems have lower reaction rates but allow higher stability concerning possible hydrolysis of the generated thioester. [282,345] Another advantage is that SEA systems can be used at different pH values (3.0–7.2). [346] A unique feature of this functionality is that the SEA group can be present in a reactive  $\text{SEA}^{\text{on}}$  and an inactive state  $\text{SEA}^{\text{off}}$  state (Scheme 3.5). Its inactive state is used, for example, to facilitate purification of the desired SEA fragment or to be specifically "switched on" during sequential ligation steps. Without the N-to-S-acyl shift, conversion to the desired thioester cannot proceed. For this purpose, the disulfide bond must be cleaved by TCEP, as the disulfide bond is stable at neutral pH values with excess MPAA. [347]



**Scheme 3.5:** Concept of  $\text{SEA}^{\text{on/off}}$  in ligation. [348]

Different proteins have been synthesized with this methodology, such as the biotinylated N domain of human hepatocyte growth factor [349] or a small ubiquitin-like modifier. [350]

## 4 Aim of the Presented Studies

Fluorine is perceived as an essential small molecule modifier, especially reflected in the increasing number of fluorine-containing pharmaceuticals and agrochemical compounds. In peptide and protein engineering, fluorine modifications have uniquely opened the door to modify thermodynamic and kinetic factors. Previous studies have shown that their effects depend on the selected fluorinated side chain and its position in the primary structure. Often the studies reveal surprising results since the C-F bond has various interaction possibilities.

In the present work, model systems were selected that are suitable to parameterize fluorine modifications. The designs are roughly divided into two areas: i) smaller peptide models were planned to establish fluorine as a probe in different applications, and ii) small proteins were selected to investigate fluorine as an orthogonal tool in controlling the kinetic of protein folding processes.

### 1. *In situ (SEIRAS) Investigations of Translocation of pHLIP into Lipid Membranes*

With the help of *in situ surface-enhanced infrared absorption spectroscopy (SEIRAS)*, the translocation process of the cell-penetrating pHLIP (pH (low) lipid insertion peptide) will be investigated. First of all, the regulatory area of the native peptide will be localized. Subsequently, specific substitutions with fluorinated amino acids (fIAAs) will be performed. On the one hand, to establish the sensitive vibration of the C-F bond as a probe in *in situ SEIRAS*. Secondly, to investigate the influence of fluorine's higher hydrophobicity and lipophilicity on the insertion process. Finally, possible F-F interactions with fluorinated lipids are to be explored.

### 2. *Project: FF03 carrier scaffold for in vivo <sup>19</sup>F-MRI Studies*

The fiber-forming peptide FF03 has already proven to be a reliable scaffold for the multivalent presentation of epitopes and drugs in our group. For the planned studies, EAE (experimental autoimmune encephalomyelitis) epitopes were selected to be conjugated to the side chain at position Lys17 of FF03. These epitopes are a 21 AAs long segment of myelin oligodendrocyte glycoprotein (MOG<sub>35-55</sub>) and a 13 AAs section of proteolipid protein (PLP<sub>139-151</sub>), which have been shown in *in vivo* studies to induce autoimmune encephalomyelitis (an animal model of multiple sclerosis

(T-cell mediated autoimmune disease)). The first step is synthesizing a fluorescent-labeled peptide scaffold conjugate to demonstrate its cellular uptake. In a second step, the hydrophobic core of the coiled coil will be furnished with fAAs, which would allow the application of *in vivo*  $^{19}\text{F}$ -MRI experiments to investigate the distribution of EAE peptide conjugates in mice.

### *3. Impact of Fluorine on Mechanical Stability of A4/B4 Dimer*

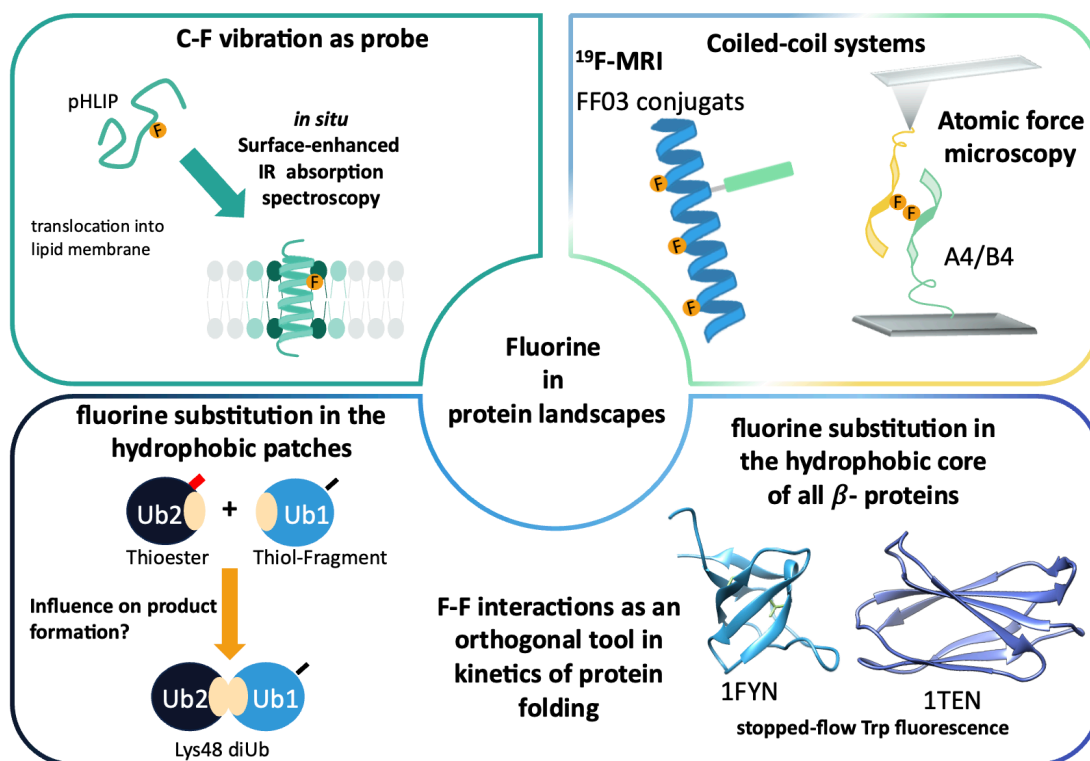
The influence of F-F interactions on the mechanical stability of the coiled-coil system A4/B4 will be investigated by atomic force microscopy (AFM). For this purpose, TfeGly will be introduced at selected positions in the hydrophobic core.

### *4. Influence of F-F Interactions on Protein Folding Kinetics by Investigating Dimerization of Ubiquitin with Fluorinated Hydrophobic Patches*

Ub is a 76 AA long small protein with specific solvent-exposed hydrophobic regions for molecular recognition involving Leu8/Ile44/Val70. In the present work, these patches will be systematically replaced by fAAs with similar hydrophobic character and vdW volume. In the course of the native chemical ligation (NCL), the influence of possible F-F interactions will be determined based on the kinetic product formation.

### *5. Chemical Synthesis of Small all- $\beta$ Proteins to Study Fluorine's Impact on Protein Folding Kinetics with Stopped-flow Trp Fluorescence*

Literature contains only a little information on the influence of fluorinated modifications on the hydrophobic core of all- $\beta$  structures. Therefore, in the present work, access to the chemical synthesis of 1FYN (62 AAs,  $\beta$ -barrel) and 1TEN (90 AAs,  $\beta$ -sandwich) will be established. AAs will be exchanged selected at positions in the hydrophobic core with fAAs of similar hydrophobic character and vdW volume. The folding kinetics shall be investigated using stop-flow Trp fluorescence. This technique exploits the sensitivity of Trp to changes in its microenvironment to provide information about folding kinetics.



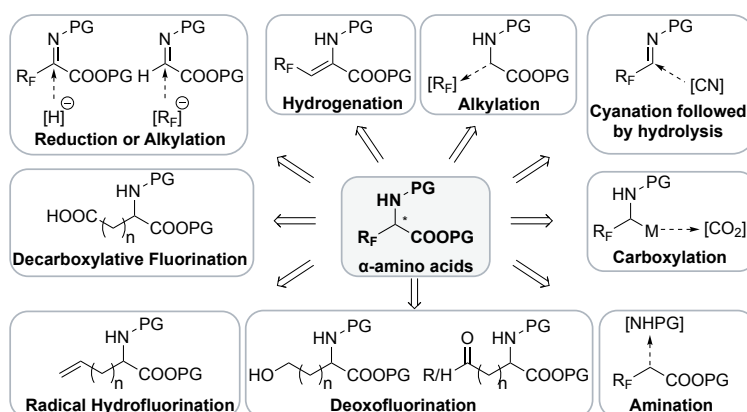
**Figure 4.1:** Illustration of selected protein landscapes to investigate different aspects of fluorine modification.

## 5 Results and Discussion

Some of the selected systems present an extraordinary challenge to chemical protein synthesis. Therefore, establishing an individual protocol for each design is necessary. Furthermore, the fAAs need to be synthesized, and a methodology to incorporate them into large peptidic systems has to be elaborated.

### 5.1 Synthesis of Fluorinated Amino Acids

There are numerous strategies for synthesizing fAAs, which were summarized by Moschner *et al.* (Figure 5.1).<sup>[115]</sup> However, there is limited commercial availability for these unique building blocks, especially for enantiomerically pure compounds. Consequently, the envisaged studies of this thesis require the synthesis of Fmoc-protected TfeGly (**1**) as well as (2*S*,3*S*)-TfVal (**2**) and (2*S*,3*R*)-TfVal (**3**). The following paragraphs will cover different synthetic procedures dealing with the 2*S*-configured fAAs and the achieved results. In the end, a new method of incorporating the synthesized fAAs using MW-assisted SPPS will be introduced.



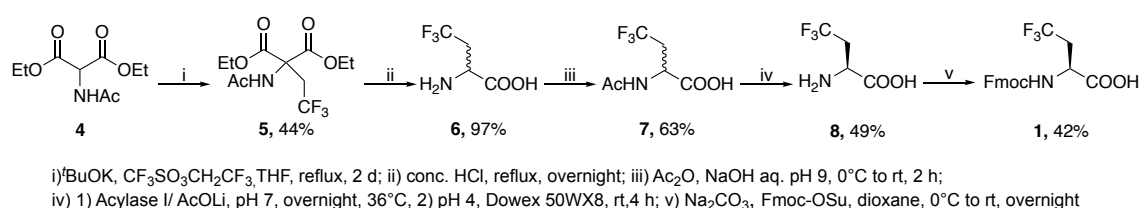
**Figure 5.1:** General synthetic strategies to obtain fluorinated  $\alpha$ -amino acids. Adapted with permission from Moschner *et al.*<sup>[115]</sup> Copyright © 2019, American Chemical Society.

#### 5.1.1 Synthesis of Trifluoroethylglycine (TfeGly) as Substituent for Val

The semi-synthesis followed the protocol of Tsushima *et al.* (Scheme 5.1).<sup>[351]</sup> First, the commercially available diethyl acetamidomalonate (**4**) was deprotonated with *t*BuOK. The resulting carbanion was then reacted with the trifluoroethyl triflate to give the diethyl *N*-acetylamino-(2,2,2-trifluoroethyl)malonate (**5**). At this point, the CF<sub>3</sub> group is introduced. Next, acidic hydrolysis of the ethyl esters and simultaneous decarboxylation of **5** provides (DL)-2-amino-4,4,4-trifluorobutanoic acid (TfeGly, **6**)



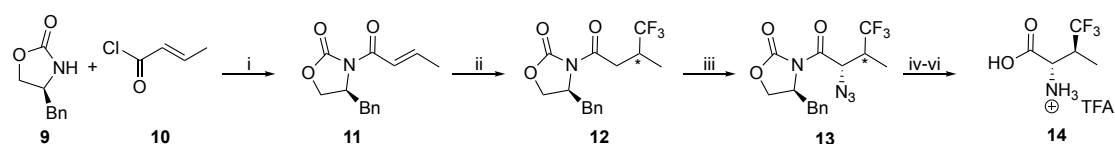
as a racemic mixture. Subsequently, this mixture is acetylated, and an enzymatic kinetic resolution is achieved by acylase I. This metalloenzyme can selectively deacetylate the L-enantiomer (**8**), [352] allowing the desired product to be separated from the D-Ac-TfeGly by ion exchange chromatography. In the final step, the Fmoc protecting group required for the SPPS is introduced with Fmoc-OSu and Na<sub>2</sub>CO<sub>3</sub>. Following this procedure, Fmoc-TfeGly (**1**) was obtained with an overall yield of 5.7%.



**Scheme 5.1:** Synthesis of Fmoc-TfeGly (**1**).

### 5.1.2 Synthesis of Trifluorovaline Derivatives (TfVal) as Leu/Ile Substituents

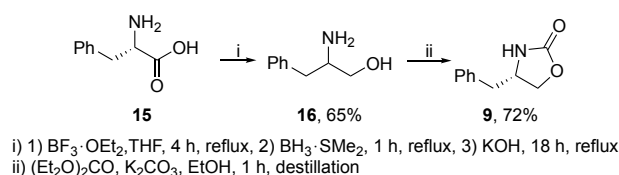
Based on the protocol of Erdbrink *et al.*, the two diastereomers (*S/S*)- and (*S/R*)-TfVal were to be synthesized (Scheme 5.2). [211] The Evans auxiliary (**9**) is used to introduce the desired chirality. The CF<sub>3</sub> group is implemented by treatment with trifluorodomethane, triethyl borane, and ytterbium triflate in an oxygen atmosphere. It is assumed that the CF<sub>3</sub> radicals generated via a reaction of Et<sub>3</sub>B/O<sub>2</sub> form a diethylborone enolate with the Michael system of **11**. The *in situ* hydrolysis of this enolate leads to the trifluoromethylated product. In addition, the Lewis acid Yb(OTf)<sub>3</sub>·nH<sub>2</sub>O is expected to suppress by-product formation. [211] Note that at this point, it should be possible to separate the obtained diastereomers by flash chromatography. Subsequently, the amino group is obtained via α-azidation with trisylazide followed by reduction and protection with Boc<sub>2</sub>O. After hydrolysis of the Evans auxiliary, the desired AA is generated.



i) 1) *n*-BuLi, 2) **10**; ii) CF<sub>3</sub>·, Et<sub>3</sub>B/O<sub>2</sub>, Yb(OTf)<sub>3</sub>·H<sub>2</sub>O; iii) 1) KHMDS, 2) trisyl azide, HOAc; iv) H<sub>2</sub> Pd/C, Boc<sub>2</sub>O; v) H<sub>2</sub>O<sub>2</sub>, LiOH, vi) TFA

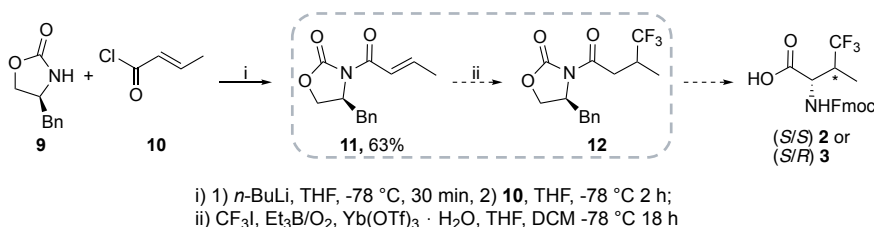
**Scheme 5.2:** Synthesis strategy of TfVal according to Erdbrink *et al.* [211]

The Evans auxiliary (**9**) was prepared in two steps in the first synthesis section. [353] Starting from phenylalanine, (*S*)-phenylalaninol was initially generated with a boron trifluoride etherate borane-dimethyl sulfide complex (Scheme 5.3). This was then reacted with potassium carbonate and diethyl carbonate to produce the Evans auxiliary. Following literature-known procedures, **9** was synthesized with an overall yield of 42%.



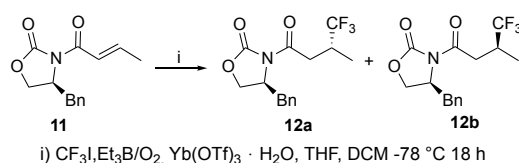
**Scheme 5.3:** Synthesis of Evans auxiliary (**9**).

Next, the procedure of Erdbrink *et al.* was followed. [211] The Evans auxiliary was first deprotonated with *n*-BuLi and then reacted with a crotonic acid chloride. For this reaction, a yield of 63% was obtained (Scheme 5.4). The fluorination step (formation of **12**) proved to be a significant challenge: a low yield of 28%, compared to 70% described in the literature, and a poor separation by flash chromatography were the main difficulties.



**Scheme 5.4:** Synthesis of **11** as starting material for the fluorination step.

Therefore, several attempts to optimize the conditions were examined. First, the amount of  $\text{CF}_3\text{I}$  and  $\text{O}_2$  was increased to optimize the radical formation (Table 5.1, entry 2). Here the impact of a single flash chromatographic purification was also investigated, revealing contamination with starting material (**11**) which was determined at 20% by  $^1\text{H}$ -NMR analysis. This result indicated a poor conversion of the reaction. Nonetheless, it was possible to increase the overall yield to 35% (calc.), which is why the addition of  $\text{CF}_3\text{I}$  and  $\text{O}_2$  was increased even further (entry 3). However, this attempt did not show the desired outcome. On the contrary, a decreased product formation was observed.

**Table 5.1:** Approaches to optimize the trifluoromethylation.

Entry <sup>a</sup>	9 [mg]	$\text{CF}_3\text{I}$ in THF	$\text{O}_2$ (10 mL/1 mmol $\text{Et}_3\text{B}$ )	purification	yield [%] (12a)	yield [%] (12b)
1	200	9 eq. (3x)	41 mL per 15 min	1) DCM: toluene 2:1 2) hexane/EtOAc 4:1	13	15
2	200	9 eq. (4x)	41 mL per 15 min	hexane/EtOAc 4:1	18 <sup>b</sup>	17 <sup>b</sup>
3	500	12 eq. (4x) 9 eq. (1x)	135 or 100 mL per 10 min	1) DCM: toluene 2:1 2) hexane/EtOAc	9 <sup>b</sup>	9 <sup>b</sup>

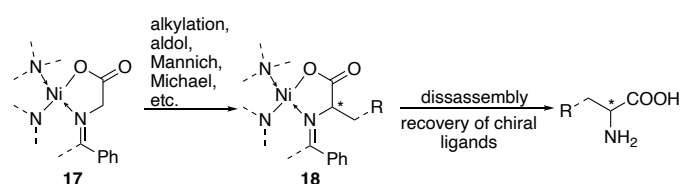
a) Conditions:  $-78^\circ\text{C}$  in THF/DCM 1:1,  $\text{Yb}(\text{OTf})_3 \cdot n\text{H}_2\text{O}$  (2 eq.),  $\text{CF}_3\text{I}$  (addition in intervals of 1.5 h),  $\text{Et}_3\text{B}$  in hexane (5 eq. for each 9 eq.  $\text{CF}_3\text{I}$ ); b) calc. according to the integrals in the  $^1\text{H-NMR}$  spectra.

The handling of the gaseous  $\text{CF}_3\text{I}$  could cause a possible source of error in the experimental procedure. The literature does not specify in which form this is added to the reaction mixture. In our experiments, we condensed the gas at  $-78^\circ\text{C}$  in THF and dosed it via a syringe. This seemed to be the safest way of handling the gas. Therefore, it was decided to explore a new synthesis strategy. One interesting approach was published by the Soloshonok group, which utilized chiral Ni(II)-complexes in the asymmetric synthesis of different types of AAs, including TfeGly. [354]

### 5.1.3 Chiral Ni-Complexes in the Synthesis of Fluorinated Amino Acids

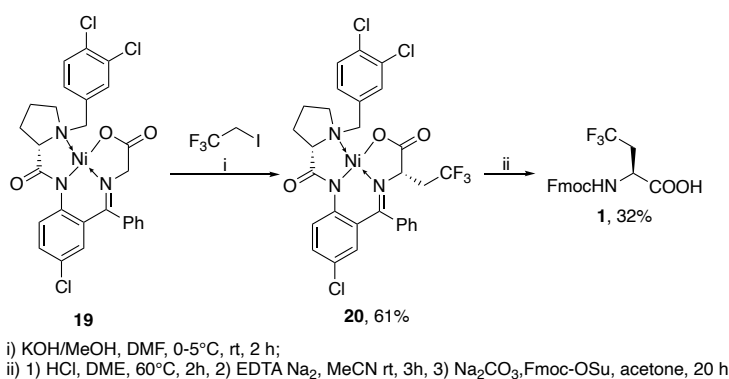
The Belokon group developed a stereoselective synthesis of amino acids mediated by chiral Nickel(II)-complexes in the late 1980s and early 1990s. [355] Since then, this versatile method has been used in a variety of different syntheses of modified aliphatic and aromatic AA. The characteristic feature of this method is the modular structure of the Ni(II)-complex, by which, for example, the chirality of the desired AA can be regulated (Scheme 5.5). During the key process, a prochiral Schiff's base **17** generated by Gly and Ni-complex is reacted with corresponding acceptors. For this step, there is a wide range of possible reactions, such as alkylation, aldol reaction, etc. The Ni(II) complex **18** is hydrolyzed in the second step with conc. HCl and the chiral Ni (II) ligands can be recovered while the free AA becomes accessible.

For more detailed information, the following review articles are recommended. [115,356–359]



**Scheme 5.5:** General Ni(II)-complex-mediated synthesis strategy of enantio-pure AAs. According to Han *et al.* [354]

Following the Strategy of Han *et al.*, Fmoc-TfeGly was synthesized in 20% yield in a two-step procedure. The starting material **19** was kindly provided by Alexander Langhans.



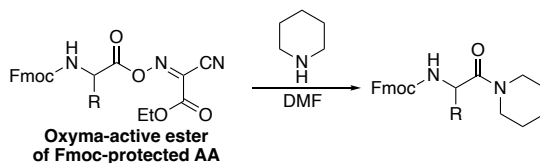
**Scheme 5.6:** Synthesis of Fmoc-TfeGly (**1**) according to Han *et al.* [354]

Hohman *et al.* succeeded in establishing this methodology in the synthesis of both enantio-pure Fmoc-TfVal isomers. Here, it was especially crucial to developing access to the alkyl-iodide. [360] Both fAAs were kindly provided by Thomas Hohmann for the ubiquitin project of the present work.

#### 5.1.4 Incorporation of Fluorinated Amino Acids into the Peptide Backbone

The incorporation of fAAs into the peptide backbone is a significant challenge that was overcome in this work using MW-assisted SPPS. In the present work, a special cycle was developed in collaboration with CEM for building blocks that either require complex synthesis as described above (section 5.1) or are very expensive. With the established cycle, these building blocks can be coupled at 90°C for 10 min with at least two eq. Unique to this cycle are four additional washing steps after Fmoc deprotection to prevent the deactivation of the active ester of the AA/building

blocks from being coupled (Scheme 5.7). In current strategies, this deactivation is usually used to save washing steps after AA coupling. [361]



**Scheme 5.7:** Deactivation of the AA-active ester by Fmoc removal cocktail.

This protocol has been implemented in several projects of our group. [362,363] Before MW-assisted SPPS, these components were manually incorporated with HOAt/DIC activation at rt in very long reaction times (up to 24 h). This approach is also applied in the presented work in the synthesis of **A4F/B4F** (section 5.3.2).

## 5.2 Insertion of pHLIP Peptides into Lipid Bilayers

The project was realized in cooperation with the Heberle group (Experimental Molecular Biophysics, Department of Physics, Freie Universität Berlin). Full detailed results of the non-fluorinated peptides can be found in the publication:

K. Ataka, J. Drauschke, **V. Stulberg**, B. Kokschi, J. Heberle, *BBA-Biomembranes*, **2022**, Volume 1864, Issue 6, 1 June 2022, 183873

Published by Elsevier B.V., Amsterdam, Netherlands [364]

Received 19 August 2021, revised 22 November 2021, accepted 20 January 2022.

Title: *pH-induced insertion of pHLIP into a lipid bilayer: SEIRAS characterization of a folding intermediate at neutral pH*

The published work is available online:

<https://doi.org/10.1016/j.bbamem.2022.183873>

**Contribution of presented doctoral thesis:** Synthesis and purification of pHLIP peptides, which will be briefly presented in the following section. Afterwards, the results of *in situ* SEIRAS measurements (measured by the Heberle group) will be presented.

### 5.2.1 Synthesis and Purification of pHLIP Peptides

A known system, the pH-low insertion peptide (pHLIP), was selected to study interactions on lipid surfaces. This peptide tends to intercalate as an  $\alpha$ -helix at low pH values in the presence of lipids. This process involves three states described above (section 2.1) that were investigated by *in situ* SEIRAS measurements.

For the intended studies, two non-fluorinated peptides were initially synthesized in order to narrow down the regulatory area for the insertion process. The first sequence, **pHLIP1**, is the C-terminal amidated pHLIP known from the literature (Table 5.2).<sup>[242]</sup> In the second sequence (**pHLIP2**), substitution by an Asp block was applied to the N-terminus, and Asp23 was replaced by His in order to investigate the influence of the carboxyl groups on the insertion. These modifications are also literature-known.<sup>[365]</sup>

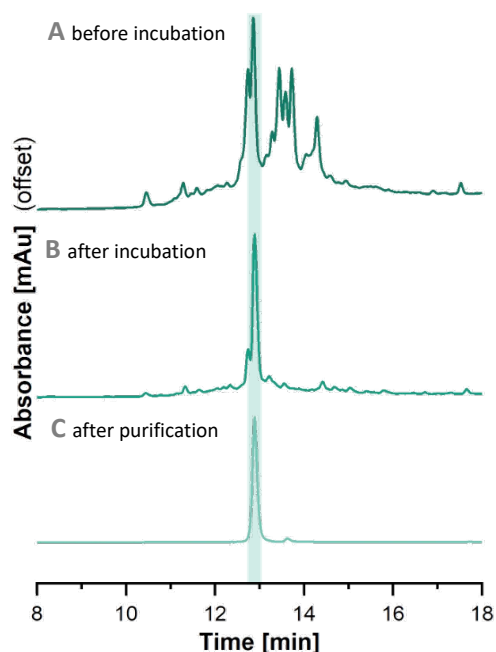
Then a selected position next to an Asp residue was substituted with TfeGly to incorporate the C-F vibration as a tool and also to investigate the influence of flAAs on the insertion process due to its higher hydrophobicity and lipophilicity. For this purpose, position 29 was selected, in which TfeGly replaced the Val residue. This flAA has a similar vdW volume and hydrophobicity.

**Table 5.2:** Planned sequences for *in situ* SEIRAS measurements.

name	sequence
<b>pHLIP1</b>	H <sub>2</sub> N-AEQNPIYWARYADWLFTTPLLLLDLALLVDADEGT-CONH <sub>2</sub>
<b>pHLIP2</b>	H <sub>2</sub> N-DDDEDNPIYWARYADWLFTTPLLLHGHALLVDADECT-CONH <sub>2</sub>
<b>pHLIP1_29TfeGly</b>	H <sub>2</sub> N-AEQNPIYWARYADWLFTTPLLLLDLALLTfeGlyDADEGT-CONH <sub>2</sub>

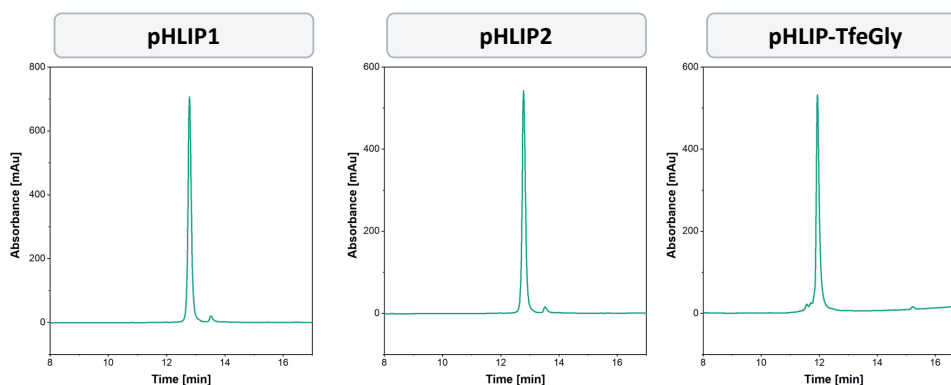
The proposed peptides contain a combination of Asp-Trp (highlighted in dark petrol in Table 5.2), a synthetic challenge (more details are presented in part A). Therefore, the strategies elaborated in Part B were used for synthesis: incorporation of a bulky side-chain protecting group of Asp with Fmoc-Asp(OMpe)-OH at the distinguished position, Fmoc-Asp(O<sup>t</sup>Bu)-OH at the remaining Asp residues. As a base for the Fmoc removal, 20% Pip with 0.1 M HOBT was used. The desired peptides were synthesized as C-functionalized amides by utilizing a Rink amide pro tide resin (0.19 mmol/g), while the N-terminus was kept unprotected. In the synthesis, the first 20 AAs were introduced with a 4 min single coupling at 90°C; the subsequent AAs were incorporated with a double coupling of 2 min at 90°C. According to the previously established protocol, TfeGly was coupled with two eq. at 90°C for 10 min. Cocktail **III** (TFA/H<sub>2</sub>O/EDT/TIS: 94/2.5/2.5/1 % v/v) was applied for cleavage from the solid support. It was found here that, at first glance, many side reactions had occurred (Figure 5.2 A). For example, a carbamylation of Trp was observed (+44 m/z in ESI-ToF-spectra), caused due to incomplete removal of the Boc side-

chain protecting group. Additionally, the peptide structure of the peptide is very pH-dependent by design. It was assumed that the peptide must reach a uniform folding state. To address both issues, the samples were incubated overnight in an acidic environment (30% ACN in H<sub>2</sub>O + 0.1% TFA).



**Figure 5.2:** Investigation of the influence of incubation under acidic conditions on the pHLIP system by HPLC. **A)** Chromatogram of the crude product directly after treatment with cocktail **III**. **B)** Chromatogram after incubation under acidic conditions (30% ACN in H<sub>2</sub>O + 0.1% TFA, pH 2.0). **C)** Chromatogram of the purified product. HPLC system: Chromaster® Kinetex® C18 column; eluents: A = H<sub>2</sub>O, B = ACN both containing 0.1% (v/v) TFA; gradient: 30% → 100% B over 18 min. The chromatogram is depicted from 8 to 18 min.

The chromatogram revealed a clear improvement after the acidic treatment (Figure 5.2 B), providing the ideal conditions for purification (Figure 5.2 C). The parameters for purification, analysis and the yield (in mg) of the peptides generated are summarized in Table 5.3 and Figure 5.3.



**Figure 5.3:** Chromatograms of purified pHLIP peptides. The gradients and RT are given in table 5.3.

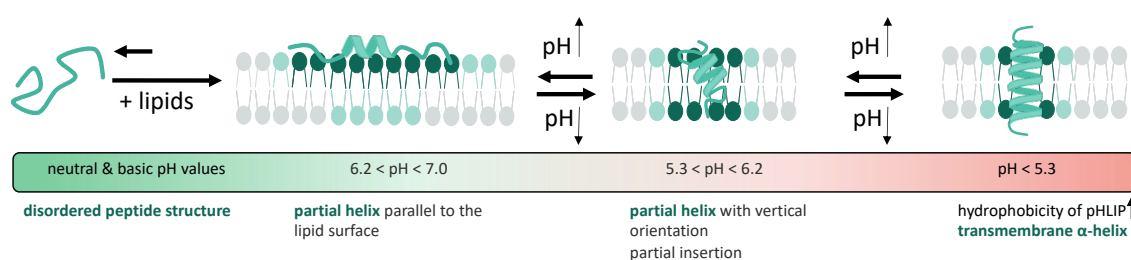
**Table 5.3:** Synthesized and purified pHLIP-peptides.

peptide	ESI-ToF	calc. <sup>a</sup> m/z	calc. obs.	purification	obtained purity [%]	yield [mg]
<b>pHLIP1</b>	[M+2H] <sup>2+</sup>	2003.5359	2003.5458	30% → 100% B over 18 min RT = 12.74 min	> 98	20.3
	[M+3H] <sup>3+</sup>	1336.0265	1336.0265			
	[M+4H] <sup>4+</sup>	1002.2718	1002.2720			
	[M+5H] <sup>5+</sup>	802.0190	802.0207			
<b>pHLIP2</b>	[M+2H] <sup>2+</sup>	2140.0205	2140.0217	30% → 100% B over 18 min RT = 12.81 min	> 98	19.7
	[M+3H] <sup>3+</sup>	1427.0163	1427.0143			
	[M+4H] <sup>4+</sup>	1070.5141	1070.5138			
	[M+5H] <sup>5+</sup>	856.6129	856.6146			
<b>pHLIP1_29TfeGly</b>	[M+2H] <sup>2+</sup>	2024.0167	2024.0491	30% → 100% B over 18 min RT = 11.94 min	> 98	12
	[M+3H] <sup>3+</sup>	1349.6804	1349.6983			
	[M+4H] <sup>4+</sup>	1012.5122	1012.5249			

a) monoisotopic calc.

## 5.2.2 Insertion Behavior Investigated by *in situ* SEIRAS

*In situ* SEIRAS measurements were conducted at the Heberle group. This technique enables a very detailed elucidation of the insertion pathway of the pHLIP1 into the lipid bilayer revealing new intermediate structures between states **II** and **III** (Figure 5.4). We were able to show that partial helix formation occurs at very slight acidic pH values between 7 and 6.2, which is parallel to the membrane surface. Re-orientation into a vertical axis is observed at pH 5.3 to 6.2 with a beginning insertion. Complete insertion is achieved at pH values below 5.3. Critical factors for the individual transitions are the protonation states of the Asp, as well as Glu residues.



**Figure 5.4:** Schematic representation of *in situ* SEIRAS results. New intermediates identified: parallel (pH 6.2 < 7.0) and vertical (pH 5.3 < 6.2) partial helix.

With **pHLIP2**, no insertion behavior was detected, indicating that Asp23 is a critical residue. In the next step, it would be interesting to substitute the hydrophobic residues around Asp23 to examine fluorine's impact on membrane permeability.

The studies with native peptide sequences laid the groundwork for further investigation of the possible impact of fluorine on membrane surfaces. For this purpose, a Val residue close to Asp30 was replaced by a TfeGly, a suitable



substituent for Val due to its similar vdW volume and hydrophobicity. This modification will also enable studies to establish the C-F vibrations as a probe for membrane insertion of peptides. Synthesis and purification were performed with the established protocol for the native sequences without difficulty. The TfeGly modified pHLIP1 (**pHLIP1\_29TfeGly**) peptide was delivered to the collaboration partner and is subject to on ongoing *in situ* SEIRAS investigations.

### 5.3 Coiled-coil systems

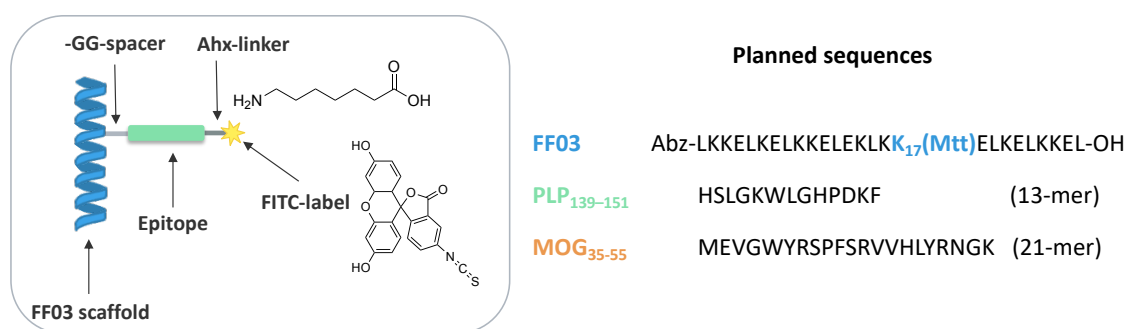
#### 5.3.1 Fiber-forming Peptide FF03

The encephalitogenic peptides (a 21 AA long segment of myelin oligodendrocyte glycoprotein (MOG<sub>35-55</sub>) and a 13-AA section of proteolipid protein (PLP<sub>139-151</sub>), induce experimental autoimmune encephalomyelitis (EAE) - the animal model of multiple sclerosis (T-cell mediated autoimmune disease). In collaboration with S. Waiczies (AG Niendorf, Experimental Ultrahighfield MR at Berlin Institute for Medical Systems Biology of the Max Delbrück Center) *in vivo* <sup>19</sup>F-MRI studies are planned with EAE-inducing epitopes (**MOG** and **PLP**) conjugated to a peptidic scaffold to track the distribution of distinct immune cell populations. In an initial experiment, a cell-penetrating peptide-derived scaffold RF011 was modified with TfVal (replacement of Leu) in the *a* position in three heptads. During *in vivo* studies, however, no clear <sup>19</sup>F signal could be detected. As a possible cause, either a poor uptake of RF011 by the cells or a low local fluorine concentration were hypothesized. [366]

Therefore, in the present work, the FF03 system was selected for a multivalent presentation of EAE epitopes. Due to its self-assembly property of forming nanofibers (section 2.2.1), FF03 provides the opportunity to achieve a high local fluorine concentration once the hydrophobic core is substituted with fluorinated AA, which is essential for *in vivo* <sup>19</sup>F-MRI studies. Prior to the incorporation of fluorinated AA into the system, a FITC-labeled scaffold will be synthesized to investigate its cellular uptake. In the second step, fluorinated AA should be introduced into the hydrophobic core of FF03.

### 5.3.1.1 Peptide-Design

Based on the all-on solid support approach, the following conjugate structure has been proposed: From previous work of our group, it is known that the position Lys17 is best suited for a modification with various functionalities (Figure 5.5).<sup>[167,168]</sup> Therefore, Lys17 was first equipped with the Mtt protecting group at the  $\epsilon$ -NH<sub>2</sub>-group, which can be removed regio-selectively under mild conditions (1% TFA in DCM). The N-term was equipped with an Abz, which can be used for concentration determinations. Before incorporating the respective epitope sequences, a small spacer consisting of two Gly residues was introduced. After the epitope sequence, another spacer has to be inserted to avoid Edman degradation due to the coupling of FITC during TFA cleavage.



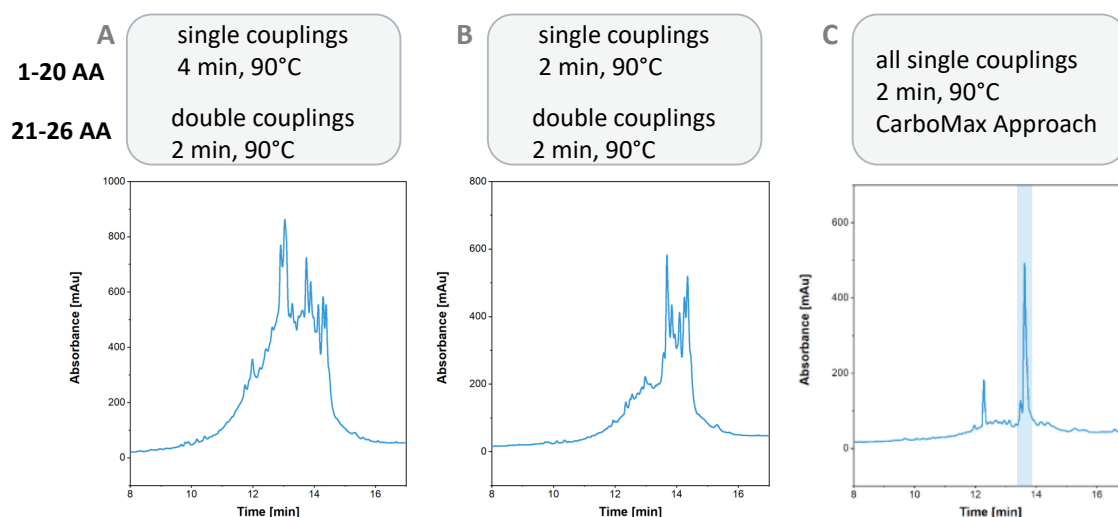
**Figure 5.5:** Design of EAE decorated FF03. Schematic representation of the designed conjugate of EAE epitope and peptidic scaffold FF03 (**left**) with all spacers and labels, sequences of EAE epitopes and FF03, the position for the linkage (Lys17) is marked with blue (**right**). MOG = myelin oligodendrocyte glycoprotein PLP = proteolipid protein.

### 5.3.1.2 Peptide Synthesis and Purification

MW-SPPS was utilized for the synthesis. A low-loaded Fmoc-Leu-TGA resin (0.19 mmol/g loading) was used. Couplings were performed with five eq. AA, Oxyma Pure, and DIC at 90°C. For the first 20 AAs, a 4 min - single coupling while a double coupling was used for the last six AAs. 20% Pip was applied for Fmoc cleavage with a reaction time of 1 min at 90 °C. After the synthesis of the scaffold, a reaction control by a micro-cleavage was performed, which revealed a variety of by-products via HPLC analysis (Figure 5.6 A). Due to the relatively simple primary structure, it was assumed that the TGA resin had low heat resistance, which may have caused the side reactions (more studies on the resins are discussed in section 5.4.2). Alternative resins did not provide suitable candidates for the planned synthesis; on the one

hand, Cl-based resins, for example, are highly acid-labile and, therefore, not ideal for the downstream regio-selective cleavage of the Mtt protecting group. On the other hand, the existing Wang resins exhibited comparatively high loading (0.3 mmol/g), which could lead to low efficiency in the conjugation of epitopes + Gly-spacer (up to 23 AAs). After all, the close spatial proximity of the individual peptide strands at high loading on the solid support could lead to aggregation, or intermolecular interactions between the strands could hinder the assembly of the peptide backbone.

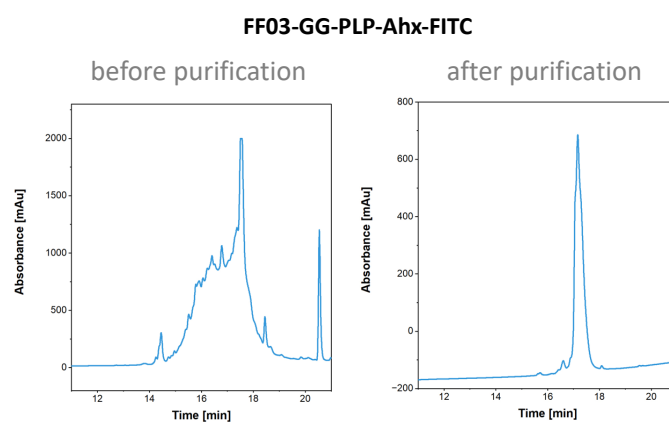
For this reason, in a second approach, the reaction time of the single couplings in the first section was reduced to 2 min to minimize the thermal stress on the TGA resin. Again, the reaction control showed a highly impure outcome (Figure 5.6 B). Subsequently, 2 min single couplings were used exclusively, and the CarboMax approach (CMA see experimental section) was applied. This involves adding 0.1 eq. DIPEA to the activator base (Oxyma Pure) doubling the amount of DIC. With this procedure, the coupling efficiency is expected to increase.<sup>[367]</sup> In addition, DIPEA as an additive is expected to reduce the acidity during AA coupling, yielding improved results, especially after incorporating the highly acid-labile Mtt protecting group. With this protocol, the scaffold was prepared in good purity, which is particularly important for the further synthesis of conjugating epitopes at the side chain (Figure 5.6 C).



**Figure 5.6:** Optimization of the FF03 scaffold synthesis. **A)** In the first experiment, the established steady-state conditions were used. **B)** A decreased coupling reaction time of the first 20 AA resulted in a minor refinement. **C)** Performing all couplings in a 2 min single coupling and with CMA provided the best results. The desired product is marked with a blue box. HPLC system: Chromaster®; Kinetex® C18 column; eluents: A = H<sub>2</sub>O, B = ACN both containing 0.1% (v/v) TFA; gradient: 10% → 80% B over 18 min. The chromatogram is depicted from 8 to 17 min.

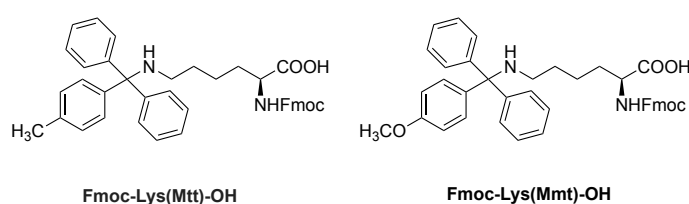
With the scaffold in satisfactory quality in hand, the conjugation of the epitope sequences began by regio-selective cleavage of the Mtt protecting group with 1% TFA in DCM. The reaction was carried out overnight at rt. By-products were observed when monitoring the reaction after the first coupling of the Gly-spacer. Uncoupled FF03 was identified as a significant by-product, assumably caused by insufficient cleavage of the Mtt protecting group or a poor Gly coupling. Repeated coupling of Gly showed no improvement, while prolonged or repeated Mtt cleavage caused further by-products, supposedly by cleavage of the acid-labile side-chain protecting groups at the peptide backbone. This is why the synthesis was continued with the first results of Mtt-cleavage.

Encouragingly, it was possible to build up the conjugate. Here, the reaction control was performed with cocktail **III** for the TFA cleavage since the PLP sequence contains a Trp. Also, 0.1 M HOBt was added to the Fmoc cleavage cocktail to avoid possible aspartimide formation. Synthesis of the epitope sequences was accomplished in several consecutive steps to avoid possible overheating of the resin. A double-coupling likewise implemented the Ahx linker. FITC was incorporated with five eq. of FITC isomer I and ten eq. of DIPEA at rt overnight to furnish the desired peptide-scaffold conjugate **FF03-GG-PLP-Ahx-FITC**. Still, by-products of the inefficient Mtt cleavage had a sustained effect on the quality of the synthesis, which is why purification by HPLC could not be achieved satisfactorily (broad signal in HPLC Figure 5.7 containing truncated peptide fragments determined by ESI-ToF measurements).



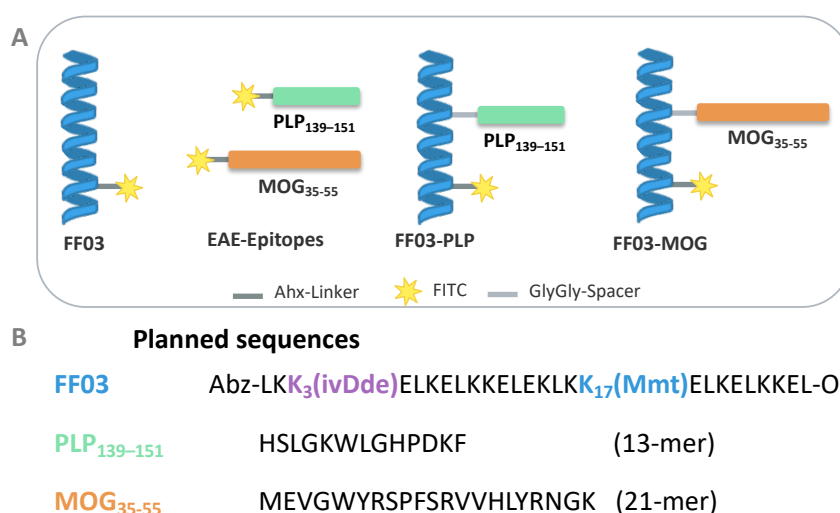
**Figure 5.7:** Chromatograms of the crude product FF03-GG-PLPAhx-FITC and the result of purification. HPLC system: Chromaster® (slow Fritz); Kinetex® C18 column; eluents: A = H<sub>2</sub>O, B = ACN both containing 0.1% (v/v) TFA; gradient: 10% → 80% B over 18 min. The chromatogram is depicted from 10 to 22 min.

As a consequence, the strategy was adjusted. In order to improve the efficiency of conjugation at the side chain, Mtt protecting group was replaced by Mmt (Scheme 5.8), which is easier to remove. For its cleavage, a protocol is available for the MW synthesizer, which would allow continuous automated synthesis. It was furthermore decided not to introduce the fluorescent label directly at the epitope but at a different *f* position of the FF03 backbone, which would allow for a better presentation of the respective epitope.

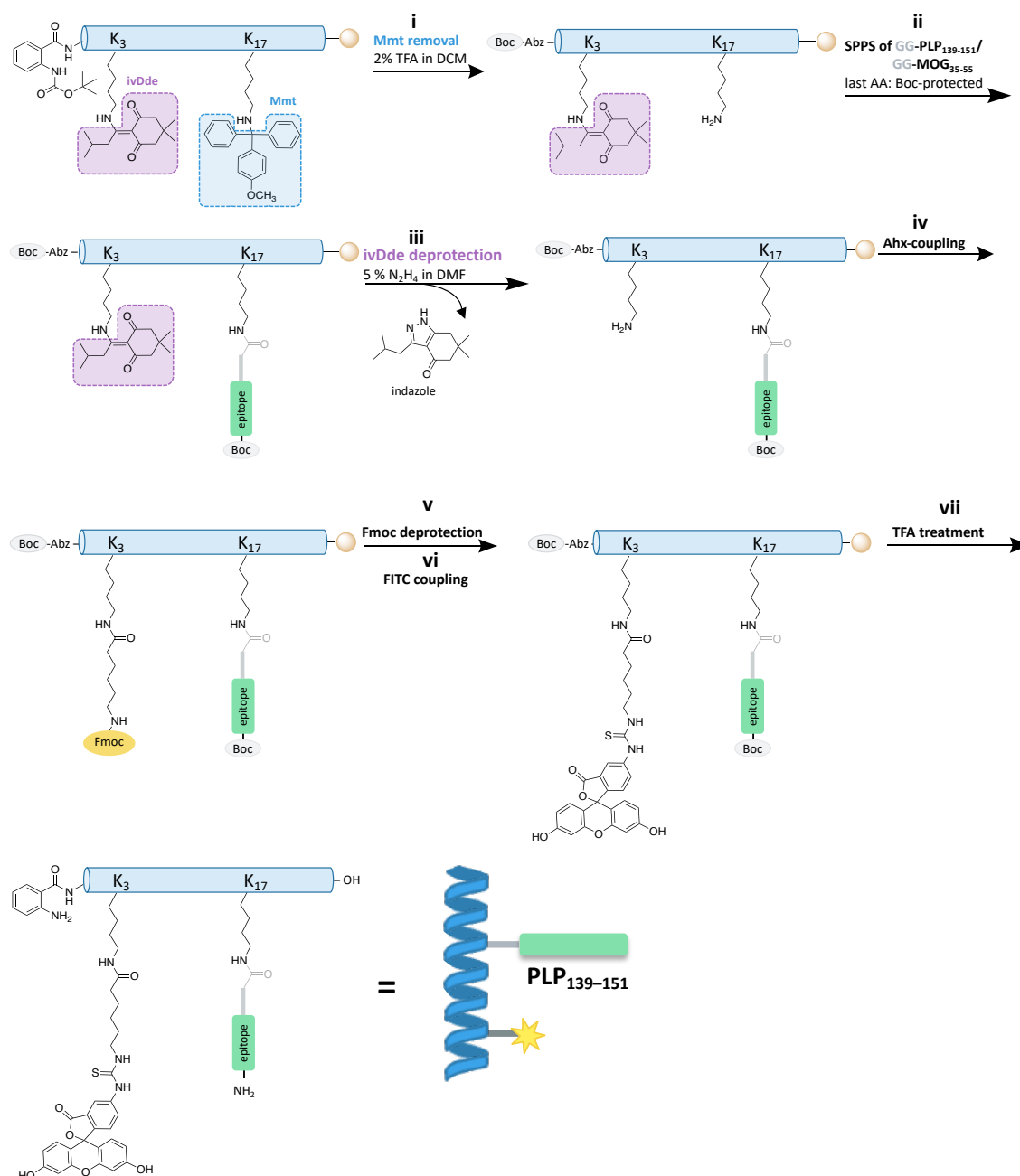


**Scheme 5.8:** Different protecting groups at the  $\epsilon$ -NH<sub>2</sub> group of Fmoc-Lys-OH used at the site of epitope recognition, regio-selectively removable by cleavage under mildly acidic conditions. Mtt: 4-methyltrityl Mmt: monomethoxytrityl.

For this purpose, an ivDde protecting group was introduced at position Lys3 for the conjugation of Ahx and FITC. Orthogonally to the other protecting groups, ivDde can be removed by treatment with 5% hydrazine solution on the MW synthesizer (3 x 3 min at 90°C). It is crucial that all N-termini (FF03-scaffold and epitope on the side chain) are equipped with a Boc-protecting group (see synthesis route in Figure 5.9). Fmoc is not suitable as it also is removed by hydrazine.

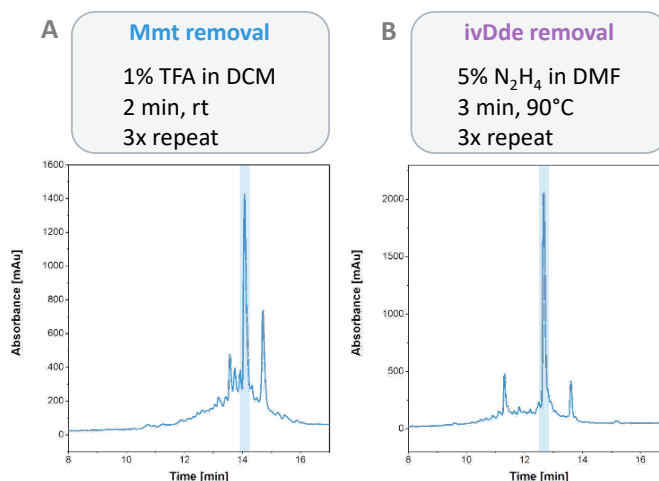


**Figure 5.8:** Outline of planned peptides for the multivalent presentation of EAE epitopes; **A)** a schematic representation of the conjugates. Non-conjugated peptides (PLP, MOG, and FF03) serve as control samples. **B)** Sequences of peptides involved.



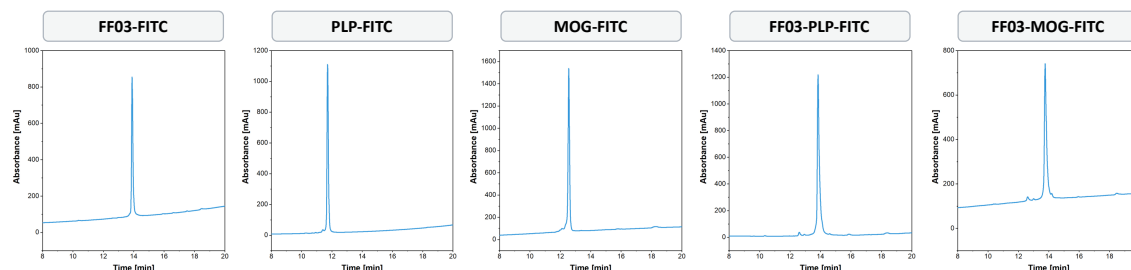
**Figure 5.9:** The route of the all-on solid support approach. MW-assisted SPPS of the scaffold is followed by regio-selective cleavage of the Mmt protecting group with 2% TFA in DCM, also on the MW synthesizer (i). Subsequently, a short Gly-Gly spacer and the epitope sequence are incorporated. Note that the last AA (His or Met) must be Boc-protected at the N-term (ii). With a 5% hydrazine solution, the ivDde protecting group is removed in the next step (iii). A Fmoc-Ahx linker is then introduced to prevent Edman degradation of FITC (iv). Fmoc cleavage (v) is followed by FITC coupling (vi). Finally, all side chain protecting groups are removed, and the peptide is cleaved from the solid support (vii).

The first step was to establish the removal of the new protecting groups at FF03. This proved to generate reliable results in each case with only few by-products (Figure 5.10).



**Figure 5.10:** Regio-selective cleavages of the protecting groups at different conjugation positions. **A)** Mmt removal conditions, **B)** ivDde deprotection. Both protocols were performed by automated means. HPLC system: Chromaster®; Kinetex® C18 column; eluents: A = H<sub>2</sub>O, B = ACN both containing 0.1% (v/v) TFA; gradient: 10% → 80% B over 18 min. The chromatogram is depicted from 8 to 17 min.

Thereupon, both conjugates were incorporated according to the previously described procedure. Since Boc-His(Trt)-OH (N-term of PLP) was not soluble in DMF, it had to be coupled manually by applying HATU coupling chemistry in a 1:1 mixture of DMC/DMF. Overall, both desired conjugates were generated in good quality and purified by HPLC without difficulties (Table 5.4 + Figure 5.11). The non-conjugated epitopes were synthesized on the Cl-MPA (0.16 mmol/g) resin using 2 min single couplings at 90°C. Ahx-linker and FITC label were coupled according to previously established protocols of the conjugate. The following figure (Figure 5.11) and Table 5.4 summarize the results on the synthesized, isolated and purified EAE-conjugated FF03 derivatives and control peptides.



**Figure 5.11:** Chromatograms of synthesized and purified FITC-labeled peptides. Gradients and RT are given in table 5.4. The chromatograms are depicted from 8 to 17 min.

**Table 5.4:** Synthesized and purified members of the proposed FF03-peptide library. **a)** monoisotopic calc. **b)** average calc.

Peptide	ESI-ToF	calc. m/z	obs. m/z	purification	obtained purity [%]	yield [mg]
FF03-FITC <sup>a</sup>	[M+3H] <sup>3+</sup>	1286.6804	1286.7318	10% → 80% B		18.1
	[M+4H] <sup>4+</sup>	965.2622	965.3004	over 18 min	> 99	
	[M+5H] <sup>5+</sup>	772.4114	772.4411	RT = 13.98 min		
PLP-FITC <sup>a</sup>	[M+2H] <sup>2+</sup>	1012.9022	1012.9685	10% → 80% B		15.1
	[M+3H] <sup>3+</sup>	675.6040	675.6492	over 18 min	> 98	
	[M+4H] <sup>4+</sup>	506.9550	506.9894	RT = 11.77 min		
MOG-FITC <sup>a</sup>	[M+3H] <sup>3+</sup>	1028.7802	1028.8129	10% → 80% B		12.5
	[M+4H] <sup>4+</sup>	771.8371	771.8616	over 18 min	98	
	[M+5H] <sup>5+</sup>	617.6712	617.6910	RT = 12.55 min		
FF03-PLP <sup>b</sup>	[M+5H] <sup>5+</sup>	1096.3757	1096.1743	10% → 80% B		10.2
	[M+6H] <sup>6+</sup>	913.8144	913.6442	over 18 min	> 98	
	[M+7H] <sup>7+</sup>	783.4134	783.2749	RT = 13.84 min		
FF03-MOG <sup>b</sup>	[M+4H] <sup>4+</sup>	1635.2790	1635.0259	10% → 80% B		8.5
	[M+5H] <sup>5+</sup>	1308.4248	1308.2272	over 18 min	> 98	
	[M+6H] <sup>6+</sup>	1090.5219	1090.3584			
	[M+7H] <sup>7+</sup>	934.8771	934.7420	RT = 13.75 min		

### 5.3.1.3 Structural Elucidation by Circular Dichroism

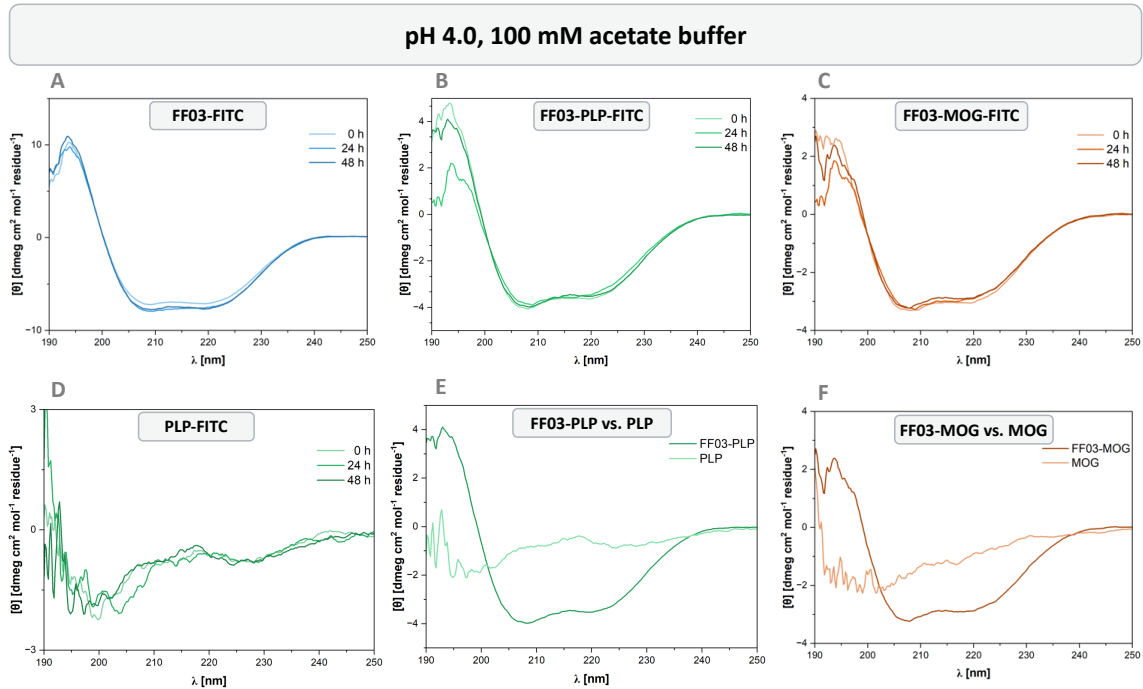
Before a potential *in vivo* application, it is necessary to verify that the conjugation of the epitopes does not disrupt the overall structure of the coiled coil construct. For this purpose, CD measurements at a peptide concentration of 100  $\mu$ M were investigated under three different conditions: pH 4.0 in 100 mM acetate buffer and pH 7.4 in 100 mM phosphate buffer, each at rt. In addition, physiological conditions were examined to ensure that the  $\alpha$ -helix is maintained at 37°C, using a 15 mM HEPES buffer commonly found in commercially available media for cell culture studies. All spectra were also recorded after 24 h and 48 h incubation of the peptides to monitor folding processes and FF03 fibril formation through changes of the minima's intensities.

According to the results, both the native FF03 and the conjugates form an  $\alpha$ -helix at pH 4.0, recognizable by the characteristic minima at 222 nm (corresponding to the  $n \rightarrow \pi^*$  transition) and 208 nm ( $\pi \rightarrow \pi^*$  transition), and a maximum at 193 nm ( $\pi \rightarrow \pi^*$ ) (Figure 5.12 A, B, and C). The non-conjugated epitopes exhibit a random structure (D), as shown by the comparison spectra E and F (spectra were compared after 48 h). Only a marginal time dependence of the coiled-coil formation was observed at pH 4.0.

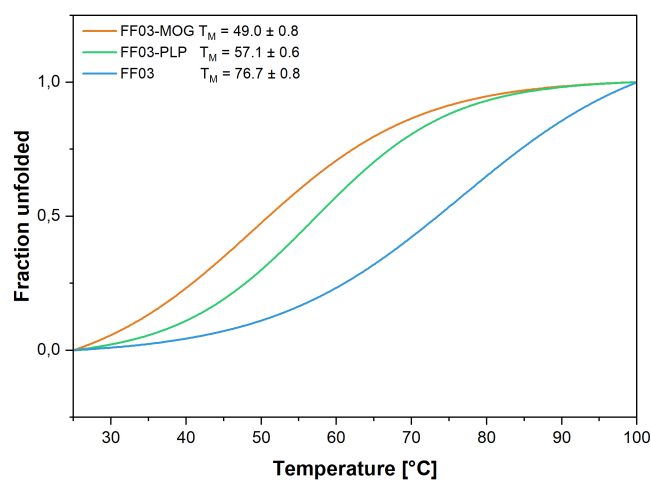
However, it is apparent that the minima's intensity decreases with the conjugate's length. This may indicate a lower  $\alpha$ -helical content (minimum at -10 at 208 nm for



non-conjugated FF03-FITC, -4 for FF03-PLP, and -3 for FF03-MOG) and, thus a weaker interaction in forming the coiled-coil system of the conjugates. A similar finding is reflected by the thermal denaturation study (Figure 5.13). Here, a melting temperature of 49°C was determined for the FF03-MOG, and 57°C for FF03-PLP, while the non-conjugated FF03-FITC had a melting temperature of 76°C.

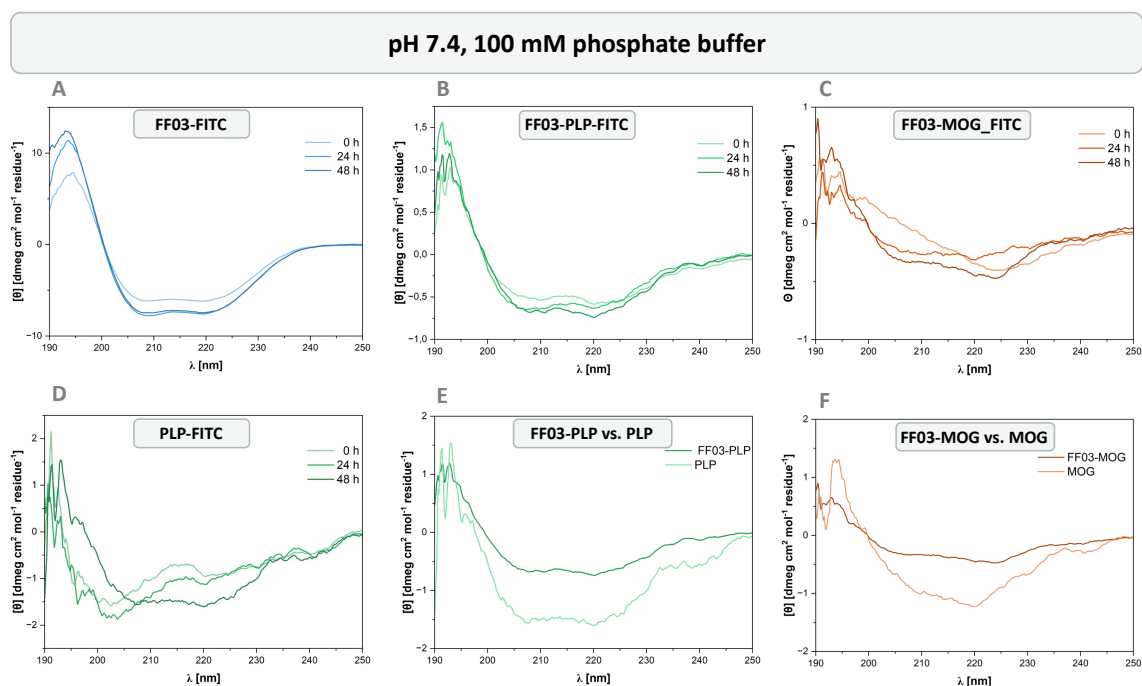


**Figure 5.12:** CD spectra of the FITC-labeled EAE-epitope-conjugates and control peptides. The spectra were recorded after 24 h and 48 h at pH 4.0 at rt in 100 mM acetate buffer with an overall peptide concentration of 100  $\mu$ M. The depicted spectra are normalized and represent the mean of three independent measurements. **A)** FF03 scaffold, **B)** FF03-PLP conjugate, **C)** FF03-MOG conjugate, **D)** non-conjugated PLP; comparison of conjugate vs. non-conjugated epitope sequence after 48 h: **E)** FF03-PLP vs. PLP and **F)** FF03-MOG vs. MOG. The depicted spectra are normalized and represent the mean of three independent measurements.



**Figure 5.13:** Thermal denaturation curves FF03 peptide and EAE-epitope conjugates measured by mean molar ellipticity  $\theta$  at 22 nm. These spectra were recorded at pH 4.0 in 100 mM acetate buffer with a peptide concentration of 100  $\mu$ M. The spectra are depicted in terms of fraction unfolded.

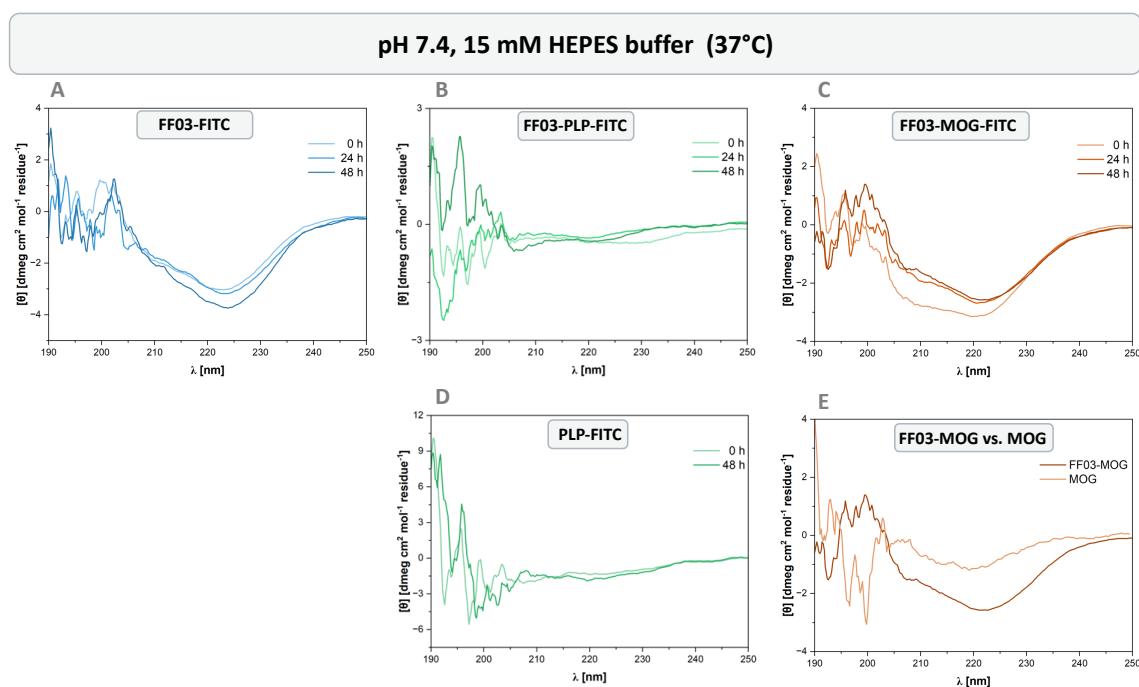
In the 100 mM phosphate buffer (Figure 5.14), the intensity of the typical  $\alpha$ -helical minima decreases significantly for FF03-PLP (to around -1), while the intensity of the non-conjugated FF03 scaffold remains approximately the same. For FF03-MOG, less distinctive characteristic signals are detectable after 48 h. Generally, the structure formation shows a time dependence for phosphate buffer. This is evident for non-conjugated FF03 and the non-conjugated PLP, which exhibits a different pattern after 48 h at pH 7.4 than at pH 4.0.



**Figure 5.14:** CD spectra of the FITC-labeled EAE-epitope-conjugates and control peptides. The spectra were recorded after 24 h and 48 h at pH 7.4 at rt in 100 mM phosphate buffer with an overall peptide concentration of 100  $\mu$ M. The depicted spectra are normalized and represent the mean of three independent measurements. **A)** FF03 scaffold, **B)** FF03-PLP conjugate, **C)** FF03-MOG conjugate; **D)** non-conjugated PLP, comparison of conjugate vs. non-conjugated epitope sequence after 48 h; **E)** FF03-PLP vs. PLP and **F)** FF03-MOG vs. MOG. The depicted spectra are normalized and represent the mean of three independent measurements.

Investigations under physiological conditions (15 mM HEPES buffer, 37°C) proved to be challenging, as no clear signals could be recorded below 210 nm due to the high voltage values (Figure 5.15). The FF03-PLP showed no precise structure formation, while FF03-MOG and the FF03 scaffold showed a displaced minimum at 222 nm. For FF03, this value continues to increase over time, while for FF03-MOG, it decreases again. The pronounced minimum at 222 nm could indicate the formation of higher-order structures.<sup>[368]</sup> Still, since the signal at 208 nm is not

distinctly visible due to the high voltage values, an investigation by cryo-TEM may provide an alternative means of structure elucidation.



**Figure 5.15:** CD spectra of the FITC-labeled EAE-epitope-conjugates and control peptides. The spectra were recorded after 24 h and 48 h at pH 7.4 at 37°C in 15 mM HEPES with an overall peptide concentration of 100  $\mu$ M. The depicted spectra are normalized and represent the mean of three independent measurements. **A)** FF03 scaffold, **B)** FF03-PLP conjugate, **C)** FF03-MOG conjugate; comparison of conjugate vs. non-conjugated epitope sequence after 48 h: **D)** FF03-PLP vs. PLP and **E)** FF03-MOG vs. MOG. The depicted spectra are normalized and represent the mean of three independent measurements.

Next, the samples were submitted to the collaboration partner for studies on cellular uptake. These investigations are ongoing.

### 5.3.2 A4/B4 - a well-defined Heterodimeric System for AFM Experiments

This project was carried out in collaboration with the Blank group (Mechano(bio)chemistry, Max Planck Institute of Colloids and Interfaces (MPI), Potsdam). The collaboration partner defined the design. It is based on the A4/B4 system (section 2.2.2) and was extended by a dipeptide spacer of Gly-Gly at the N- and C-terminus (Table 5.5). A terminal Cys was introduced to use the thiol function of the side chain for immobilization of the peptides for AFM experiments (N-terminal at the A strand, C-terminal at the B strand). The collaborator also provided the native non-fluorinated A4 and B4 sequences. In the second heptad, Leu, in *d* position of the hydrophobic core, was to be replaced by TfeGly.

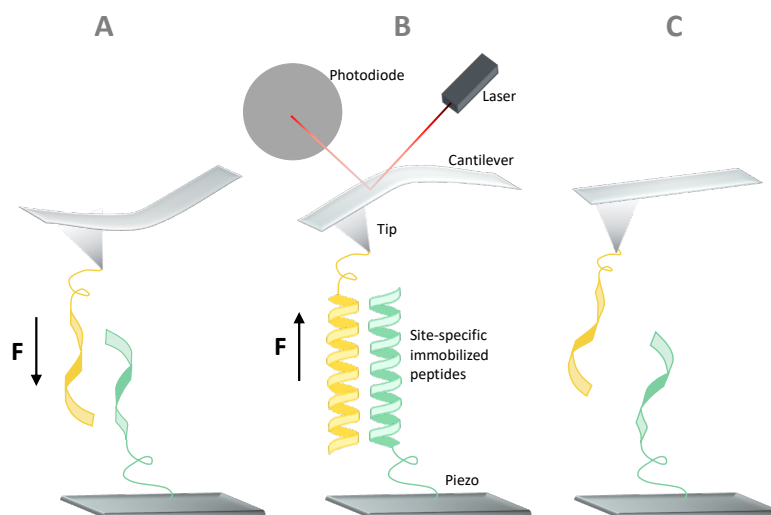
Furthermore, it was considered to substitute Val or TfeGly in the  $\alpha$  position of the heptad since  $\beta$ -branched AA are more advantageous in this position for forming the hydrophobic core in coiled-coil systems. The proposed sequences are shown in the following table.

**Table 5.5:** Planned sequences for AFM examination. **X** = TfeGly

name	N-term	sequence																								C-term				
		g	a	b	c	d	e	f	g	a	b	c	d	e	f	g	a	b	c	d	e	f	g	a	b		c	d	e	f
<b>A4</b>	Ac-CGG	E	I	A	A	L	E	Q	E	I	A	A	L	E	K	E	N	A	A	L	E	W	E	I	A	A	L	E	Q	GG-NH <sub>2</sub>
<b>A4F</b>	Ac-CGG	E	I	A	A	L	E	Q	E	I	A	A	X	E	K	E	N	A	A	L	E	W	E	I	A	A	L	E	Q	GG-NH <sub>2</sub>
<b>B4</b>	Ac-GG	K	I	A	A	L	K	Q	K	I	A	A	L	K	Y	K	N	A	A	L	K	K	K	I	A	A	L	K	Q	GGC-NH <sub>2</sub>
<b>B4F</b>	Ac-GG	K	I	A	A	L	K	Q	K	I	A	A	X	K	Y	K	N	A	A	L	K	K	K	I	A	A	L	K	Q	GGC-NH <sub>2</sub>
<b>A4V2</b>	Ac-CGG	E	I	A	A	L	E	Q	E	V	A	A	L	E	K	E	N	A	A	L	E	W	E	I	A	A	L	E	Q	GG-NH <sub>2</sub>
<b>B4V2</b>	Ac-GG	K	I	A	A	L	K	Q	K	V	A	A	L	K	Y	K	N	A	A	L	K	K	K	I	A	A	L	K	Q	GGC-NH <sub>2</sub>
<b>B4V2F</b>	Ac-GG	K	I	A	A	L	K	Q	K	X	A	A	L	K	Y	K	N	A	A	L	K	K	K	I	A	A	L	K	Q	GGC-NH <sub>2</sub>
<b>A4V2d</b>	Ac-CGG	E	I	A	A	L	E	Q	E	I	A	A	V	E	K	E	N	A	A	L	E	W	E	I	A	A	L	E	Q	GG-NH <sub>2</sub>
<b>B4V2d</b>	Ac-GG	K	I	A	A	L	K	Q	K	I	A	A	V	K	Y	K	N	A	A	L	K	K	K	I	A	A	L	K	Q	GGC-NH <sub>2</sub>

The collaborating partner, respectively, synthesized the native A4 and B4 sequences.

With a well-defined structure of a heterodimer, the A4/B4 system is an excellent candidate for investigations using AFM. Here, the force leading to the rupture of the interactions of a coiled coil is to be measured. The schematic experimental set-up is depicted in the following figure.



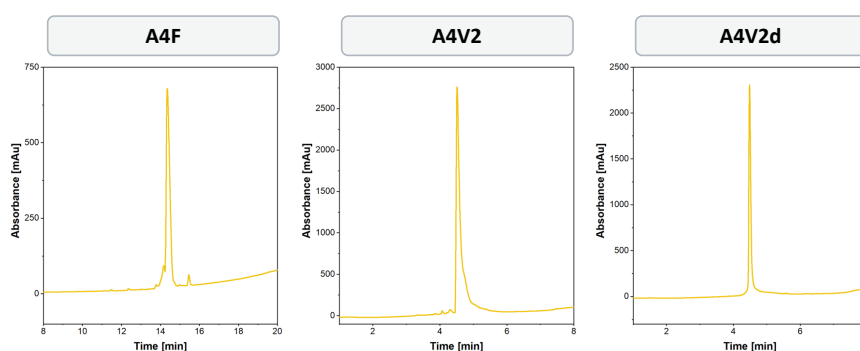
**Figure 5.16:** Illustration of a schematic experimental procedure of an AFM measurement. Using a thiol function, one strand of the coiled-coil system is placed either on the tip of a movable cantilever or on the piezo element. **A)** The flexible cantilever brings the two strands into proximity and allows them to interact via hydrophobic or electrostatic interactions. **B)** Subsequently, the tip is retracted until the interactions between the peptide strands are disrupted. During this process, a focused laser beam, which is directed onto a photodiode, is used for read-out. **C)** At this point, the rupture force is evaluated. Adapted from Garcia. <sup>[369]</sup>

### 5.3.2.1 Synthesis of A4/B4 Derivatives

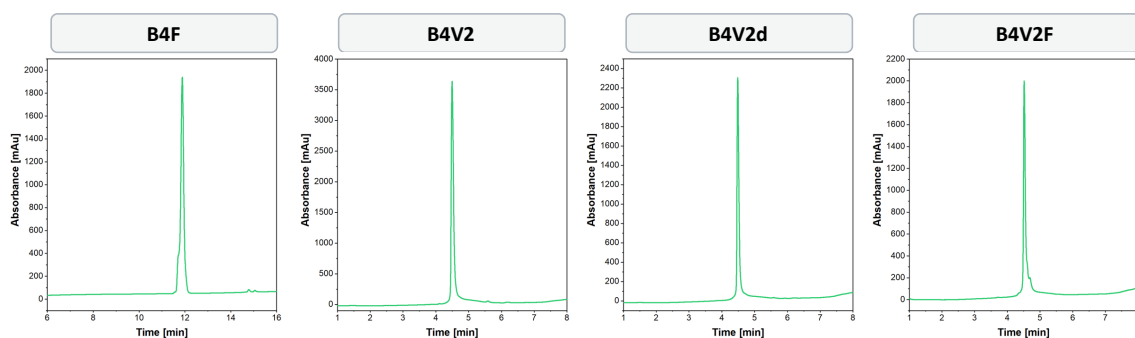
Automated conventional SPPS prepared the synthesis of A4F and B4F on Activotec P11; a TGR-NovaSyn resin was used. After manual coupling of the first AA a loading test was performed. HATU chemistry was applied in general for the AA coupling. Fmoc cleavage was performed in 20% Pip. The TfeGly was incorporated manually, for which 1.5 eq of the AA was first preactivated with HOAt/DIC for 5 min and allowed to react for 6 h at rt. Subsequently, another 0.5 eq. of each was added, and the reaction was allowed to stir overnight. After proving the complete incorporation of TfeGly with a microcleavage, the automated synthesis was continued. Acetylation of the N-term was performed manually with 10% Ac<sub>2</sub>O and 10% DIPEA at rt (3 x 10 min).

MW-SPPS with DIC/Oxyma chemistry provided all V2-variants using a Rink amide pro tide resin. A standard coupling of 2 min at 90°C was conducted for the coupling of the first 20 AAs, and double coupling was performed from the 21st AA onwards. Fmoc-removal was done with 20% Pip. TfeGly was introduced according to the established protocol for MW-SPPS with 2 eq. at 90°C for 10 min. Acetylation was performed with 10% Ac<sub>2</sub>O and 10% DIPEA at 40°C.

All peptides generated were cleaved of the solid support with cocktail **III** because of the presence of Cys. The following figures (Figure 5.17 + Figure 5.18) and Table 5.6 summarize the results on the synthesized, isolated and purified A4 and B4 derivatives.



**Figure 5.17:** HPLC-Chromatograms of synthesized and purified A-strand peptides. Methods are given in table 5.6.



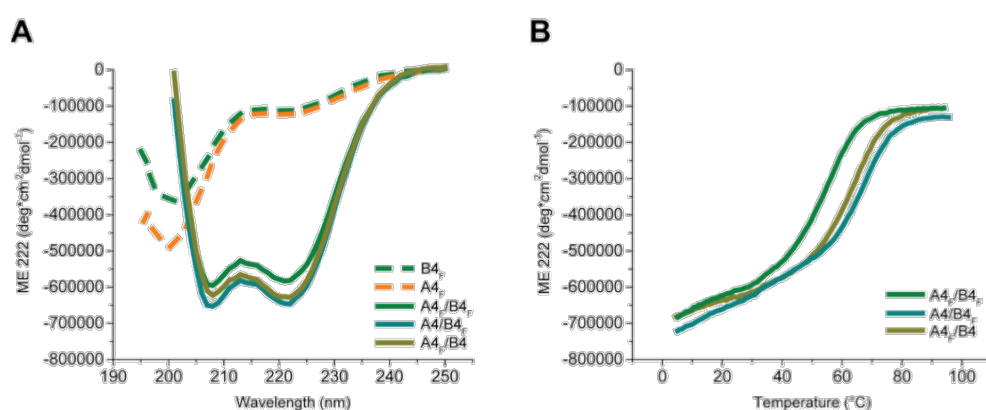
**Figure 5.18:** HPLC-Chromatogram of synthesized and purified B-strand peptides. Methods are given in table 5.6.

**Table 5.6:** Synthesized and A-strand and B-strand peptides.

peptide	ESI-ToF	calc. m/z	obs. m/z	purification	obtained purity [%]	yield [mg]
<b>A4F</b>	[M+3H] <sup>3+</sup>	1165.2106	1165.2231	5% → 100% B over 18 min RT = 14.47 min	96	12.3
	[M+4H] <sup>4+</sup>	874.1559	874.1699			
<b>B4F</b>	[M+2H] <sup>2+</sup>	1155.0116	1155.0228	5% → 100% B over 18 min RT = 11.89 min	98	14.3
	[M+3H] <sup>3+</sup>	866.5106	866.5198			
<b>A4V2</b>	[M+2H] <sup>2+</sup>	1727.3462	1727.3568	10% → 80% B over 7 min RT = 4.51 min	95	15.6
	[M+3H] <sup>3+</sup>	1151.9001	1151.9076			
	[M+4H] <sup>4+</sup>	864.1770	864.1841			
<b>B4V2</b>	[M+3H] <sup>3+</sup>	1141.7011	1141.7092	10% → 60% B over 7 min RT = 4.48 min	> 98	20.8
	[M+4H] <sup>4+</sup>	856.5277	856.5355			
<b>B4V2F</b>	[M+2H] <sup>2+</sup>	1732.0135	1732.0288	10% → 60% B over 7 min RT = 4.51 min	> 98	12.5
	[M+3H] <sup>3+</sup>	1155.0116	1155.0257			
	[M+4H] <sup>4+</sup>	866.5106	866.5224			
	[M+5H] <sup>5+</sup>	693.4101	693.4079			
<b>A4V2d</b>	[M+2H] <sup>2+</sup>	1727.3462	1727.3508	10% → 80% B over 7 min RT = 4.48 min	> 98	15.2
	[M+3H] <sup>3+</sup>	1151.9001	1151.9037			
	[M+4H] <sup>4+</sup>	864.1770	864.1807			
<b>B4V2d</b>	[M+2H] <sup>2+</sup>	1712.0477	1712.065	10% → 60% B over 7 min RT = 4.49 min	98	19.4
	[M+3H] <sup>3+</sup>	1141.7011	1141.7115			
	[M+4H] <sup>4+</sup>	856.5277	856.5369			

### 5.3.2.2 Investigations of Thermal and Mechanical Stability of Fluorinated A4/B4- Derivatives

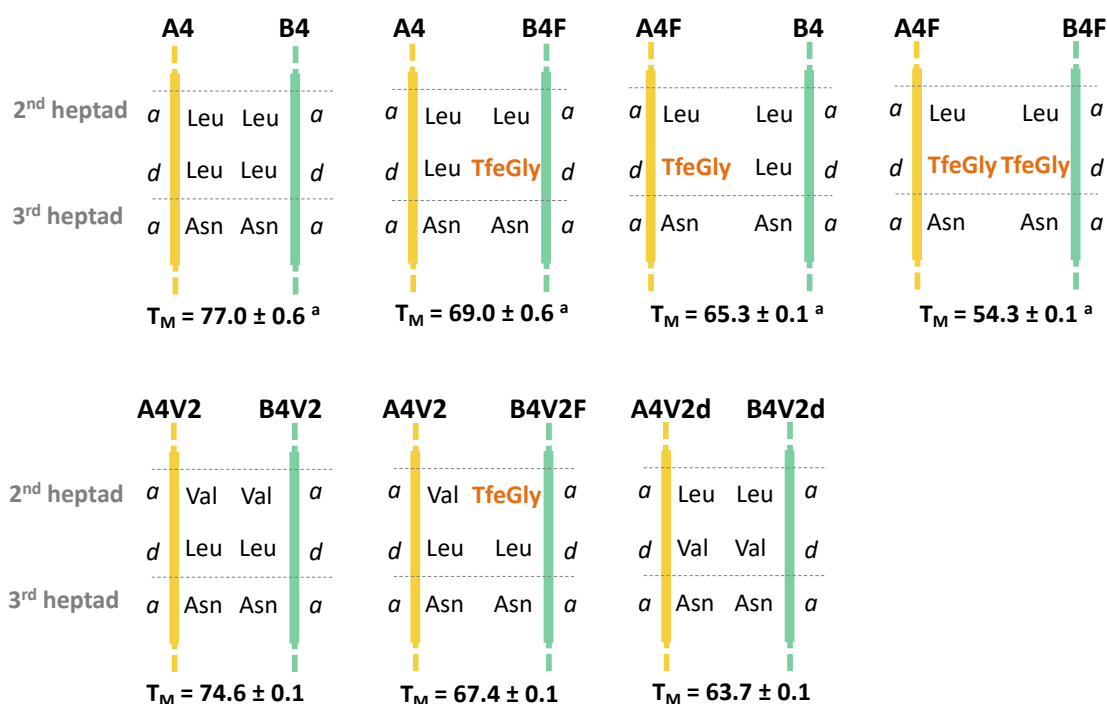
First, the native A4/B4 system was compared with the TfeGly derivatives (A4F/B4F) by CD analysis in 100 mM PBS buffer at rt with an overall peptide concentration of 50  $\mu$ M. All three combinations (A4/B4F, A4F/B4, and A4F/B4F) were observed to form an  $\alpha$ -helical structure according to the characteristic patterns of the recorded CD-spectra (Figure 5.19 A). The monomeric peptides form a random coil structure. When subjected to thermal denaturation, a lower melting temperature ( $T_M$ ) was determined for all combinations that involved a fluorinated strand (Figure 5.19 B + Figure 5.20).



**Figure 5.19:** CD measurements of fluorinated derivatives of A4/B4 system; **A)** CD-spectra for monomeric fluorinated strands (represented by dotted lines) and different dimer combinations (A4F/B4, A4/B4F, and A4F/B4F). The spectra were recorded at pH 7.4 at rt in 100 mM PBS buffer with an overall peptide concentration of 50  $\mu$ M. **B)** Thermal denaturation curves of dimer combinations measured at 222 nm. Data kindly provided by Garcia. [369]

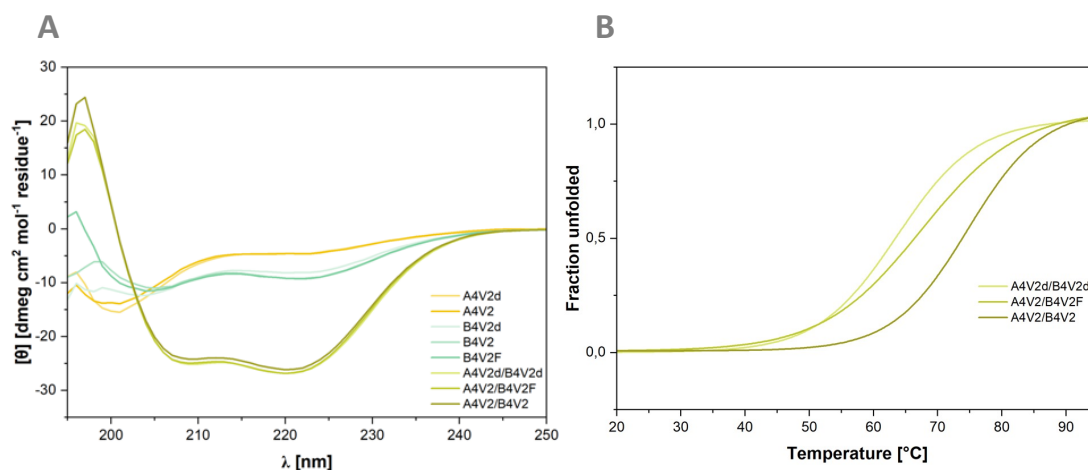
It was assumed that TfeGly, which according to the hydrophobicity plot, is a more suitable substituent for Val, destabilizes the hydrophobic core of the coiled coil by its structure, as TfeGly exhibits a much lower  $\alpha$ -helical propensity. A possible fluorous effect does not compensate for this destabilization. For a more suitable comparison, a dimer combination with Val substitution was synthesized (A4V2d/B4V2d). This dimer also revealed a lower  $T_M$  of 63.7°C, which is significantly lower than the native A4/B4 system ( $T_M = 77$  °C) and lower than the combination of a TfeGly variant and a native strand A4 or B4 (A4F/B4  $T_M = 69.0$ °C or A4/B4F  $T_M = 65.3$ °C), respectively. Furthermore, the direct comparison of Val (A4V2d/B4V2d  $T_M = 63.7$ °C) and TfeGly (A4F/B4F  $T_M = 56.3$ °C) in *d* position shows a difference of 7.4°C. It is necessary to examine further whether an unfavorable

packing effect causes this difference in the hydrophobic core or by other influences of TfeGly, such as polar interactions of the Asn residues in close proximity (3rd heptad) of the hydrophobic core. Studies involving derivatives with substitutions at the *a* position (A4V2/B4V2 or A4V2/B4V2F) were expected to provide clarification here. Since all more stable combinations were achieved using the B strand of the desired modification, a fluorinated A strand (A4V2F) was not synthesized. The recorded CD-spectra demonstrate a typical  $\alpha$ -helical pattern (yellow-green lines Figure 5.21) but also revealed an indication of  $\alpha$ -helical content for the monomeric B strands (green lines) while A strands exhibit a random structure (yellow lines). The Val derivatives (A4V2/B4V2) showed significantly improved thermal stability ( $T_M = 74.4^\circ\text{C}$ ). However, it was also found that a TfeGly mutant (A4V2/B4V2F  $T_M = 67.4^\circ\text{C}$ ) was less favorable, although this showed a more stable combination compared to the A4F/B4F.



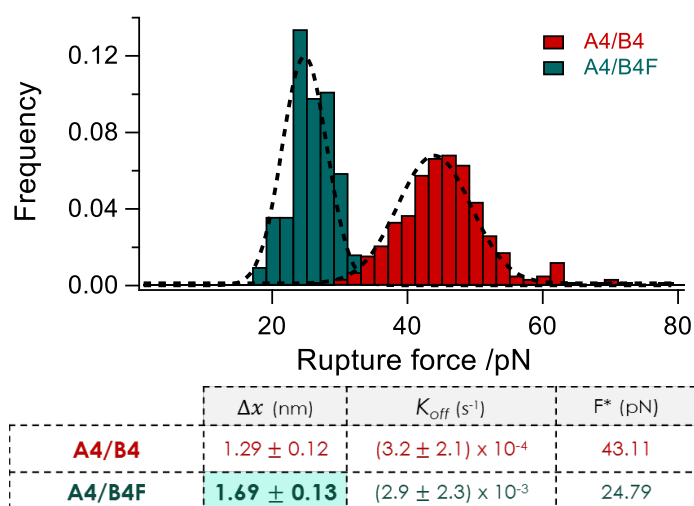
**Figure 5.20:** Summary of measured melting temperatures ( $T_M$  in °C) of different combination of A4 (yellow lines) and B4 (green lines) derivatives by thermal denaturation. Schematic representation of AAs placed in the *a* or *d* position of the hydrophobic core in the heptad of interest. **a:** Measurements by Garcia. [369]





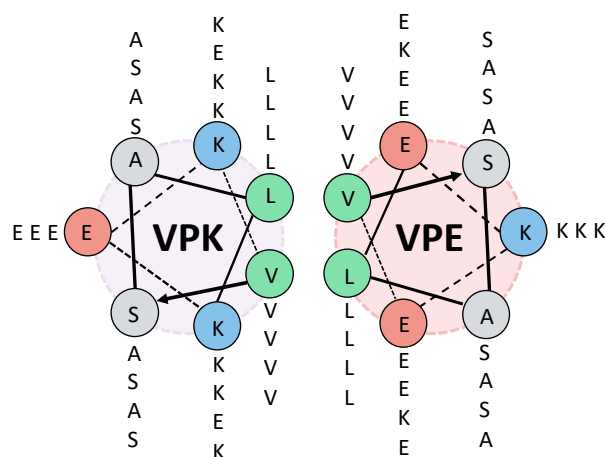
**Figure 5.21:** CD measurements Val substitutions in *a* and *d* Position and TfeGly substitution in a position; **A)** Monomeric A strands a marked in yellow, B strands in green while the different dimer combinations are depicted in yellow-green. The spectra were recorded at pH 7.4 at rt in 100 mM PBS buffer with an overall peptide concentration of 50  $\mu$ M. The depicted spectra are normalized and represent the mean of three independent measurements. **B)** Thermal denaturation of different dimer combinations; curves measured at 222 nm. The spectra are depicted in terms of fraction unfolded.

AFM-based experiments were conducted to investigate the mechanical stability of the most stable combination (A4/B4F) detected by thermal denaturation Garcia carried out the experiments at the MPI. It was shown that destabilization could also be detected by this method (Figure 5.22). The breaking force ( $F^*$ ) decreases from 43 pN to 24 pN. The dissociation constant ( $k_{off}$ ) becomes higher as a consequence of the destabilization. The distance between the native and transition states ( $\Delta x$ ) also becomes more prolonged, mainly and hypothetically, due to a less dense packing.



**Figure 5.22:** Rupture force histograms measured at a retract speed of 400 nm  $s^{-1}$ . The most probable rupture force is determined by a Gaussian fit (black dashed line). Summary of the kinetic and thermodynamic parameters obtained from A4/B4, A4/B4F coiled coils. All values are depicted as the mean  $\pm$  standard error of the mean ( $n = 3$ ). Data kindly provided by Garcia. <sup>[369]</sup>

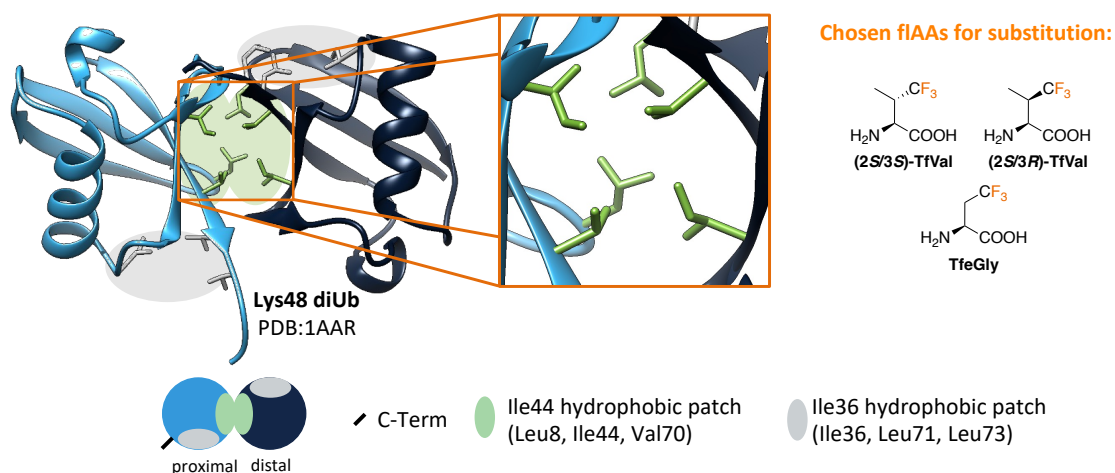
These observations were the reason for synthesizing the Val derivatives and also why only B4V2F, not A4V2F, has been prepared in the additional studies. AFM measurements of the Val/TfeGly-derivatives are part of ongoing investigations. According to the CD studies performed, however, destabilization is most likely to be expected here. Due to the Val substitution in both *d* and *a* position, a lower melting temperature was determined by CD measurements (Figure 5.20). By incorporating TfeGly into the second heptad, destabilization already occurs when it is introduced on one strand. But in addition to potential packing effects, probable interactions with the polar Asn-interaction need to be investigated in more detail in this system. So far, the main focus of AFM investigations has been to examine changes in the mechanical stability of the A4/B4 system by modifications (lactam linkage or metal staples) at the *b* position (solvent exposed) but not in the hydrophobic core. [370] To avoid the question of possible interactions with the Asn H-bond by fluorine-specific interactions in the hydrophobic core and to focus only on hydrophobic or fluorine interactions, it might be advantageous for further studies to select a different system. For example, our group's VPE/VPK model appears to be a good choice. This system also forms a parallel heterodimer consisting of Val and Leu in the hydrophobic core (Figure 5.23). Therefore, one could also consider the influence of different substitutions. In fact, initial studies have demonstrated a stabilizing effect of fluorine substitution through thermal denaturation. [371]



**Figure 5.23:** Helical wheel representation of the parallel VPE/VPK heterodimeric model system. [372]

## 5.4 Fluorine as Orthogonal Tool to Impact Kinetics of Protein Folding via Product Formation of Ubiquitin Dimers

Lys48-linked Ub dimer (diUb) was selected as a model system to investigate the influence of F---F interaction between partially fluorinated side chains of hydrophobic AAs on protein folding ("clustering by fluorous interactions"). It is a well-studied protein dimer in which the Ile44 hydrophobic patches are orientated toward each other (Figure 5.24). First, the rationale of the design and synthesis strategy will be introduced, followed by a discussion of the obtained results.



**Figure 5.24:** Crystal structure of Lys48 diUb (PDB-ID: 1AAR) [267], with schematic representation, zoom into Ile44 hydrophobic patch and selected fIAAs for substitution. Created with <http://www.cgl.ucsf.edu/chimera>.

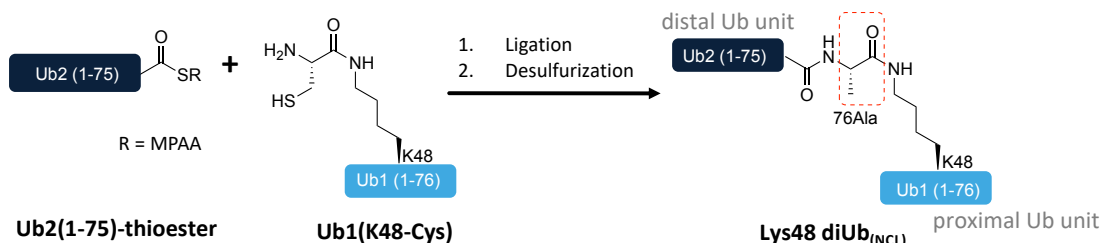
### 5.4.1 Design of the Model System

Two strategies are to be explored to obtain the desired diUb: native chemical ligation (NCL), which leads to an Ala mutation in the diUb (Figure 5.25), and ligation by isopeptide linkage (ICL). In the latter, the native sequence of the diUb is retained (Figure 5.26). However, the required  $\delta$ -mercapto-Lys is very expensive and challenging to synthesize and will be provided by A. Brik (Schulich Faculty of Chemistry in the Technion-Israel Institute of Technology Technion City, Israel), an external collaborator. Therefore, the ICL is mainly scheduled for samples to be studied in terms of the influence of fluorine interactions on the biological properties of diUb. The NCL approach, on the other hand, is ideally suited for structural elucidation of fluorous interactions since the Ala mutation occurs in the C-terminal region and is not involved in protein folding (Figure 5.24). In both strategies, the product formation step of the ligation is the key interest of the present study. Here

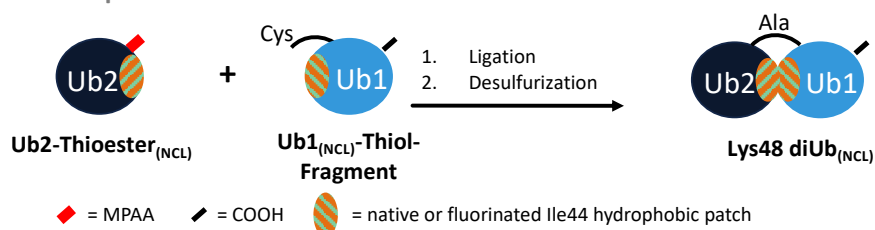
the influence of F-F interactions might impact ligation, as was previously described in our group. [238]

## Native Chemical Ligation (NCL)

### chemical structures



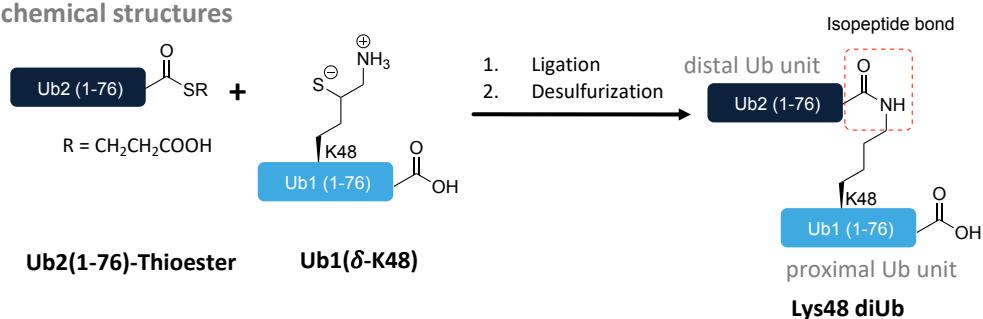
### schematic representation



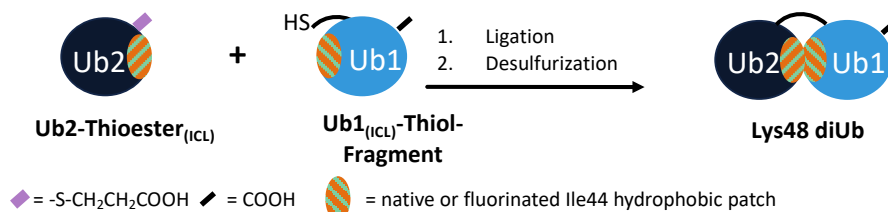
**Figure 5.25:** Schematic representation of planned native chemical ligation (NCL) approach to obtain Lys48 diUb with Ala mutation for structural analysis.

## Ligation via Isopeptide-Bond Formation (ICL)

### chemical structures



### schematic representation



**Figure 5.26:** Illustration of planned ligation via isopeptide bond formation (ICL) for biologically relevant samples.

The ultimate goal is to substitute the canonical AA of the Ile44 hydrophobic patches of Lys48 diUb with fluorinated analogs of similar vdW volume and hydrophobicity.

For this purpose, (S/R)TfVal, (S/S)TfVal, and TfeGly were chosen since our previous studies had demonstrated these desired properties in an HPLC-based hydrophobicity plot (see section 1.2.3.1). We assume this substitution will not perturb Ub's folding, judging from the overall highly conserved nature of the protein. Nevertheless, we will examine the substitution of each position in parallel to the full replacement of the hydrophobic patch.

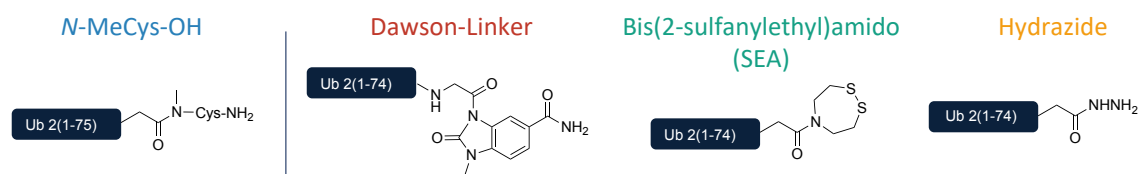
**Table 5.7:** Planned substitution patterns in a Ub monomer.

Name	Sequence
wt	MQIFVKLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQK ESTLHLVLRGG
70TfeGly	MQIFVKLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQK ESTLHLTfeGlyLRLGG
44(S/S)TfVal	MQIFVKLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRL(S/S)TfValFAGKQLEDGRTL SDYNIQKESTLHLVLRGG
8(S/R)TfVal	MQIFVKT(S/R)TfValTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTL SDYNIQKESTLHLVLRGG
8(S/R)TfVal- 44(S/S)TfVal -70TfeGly	MQIFVKT(S/R)TfValTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRL(S/S)TfValFAGKQ LEDGRTLSDYNIQKESTLHLTfeGlyLRLGG

For each ligation approach, two monomers are required: a thiol-functionality (referred to as Ub1 since it is the proximal Ub unit; highlighted in light blue (Figure 5.24) and a thioester (referred to as Ub2 as this is the distal Ub unit; highlighted in dark blue (Figure 5.24)).

$\delta$ -Mercapto-Lys is used to introduce the thiol functionality in Ub1 (Figure 5.26) for the ICL, while an additional Cys is coupled onto the side chain of Lys48 after a regioselective removal of the ivDde side-chain protecting group at  $\epsilon$ -NH<sub>2</sub> (Figure 5.25) Note: this will lead to a shorter sequence of the desired Ub2-thioester in the NCL approach and an extended sequence of Ub1.

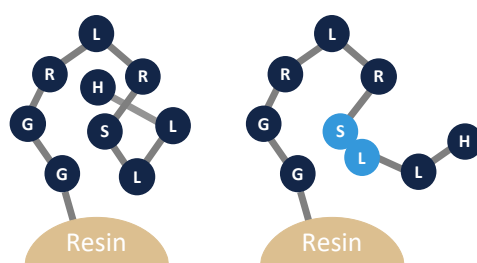
For the synthesis of the respective thioester for each strategy, different precursors, so-called thioester surrogates, are to be examined: on one hand, *N*-Me-Cys for ICL and on the other hand, Dawson, SEA, and Hydrazine for NCL (Scheme 5.9). These functionalities enable the direct transformation into the desired thioester *in situ* after full synthesis on the solid support without the need for additional purification steps.



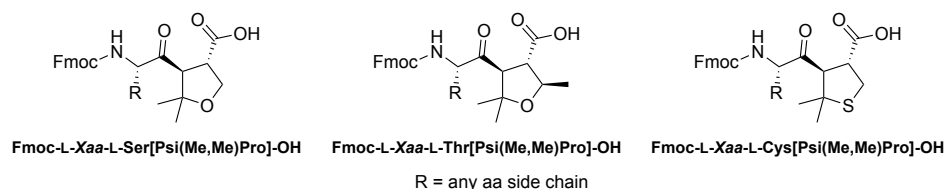
**Scheme 5.9:** Planned thioester surrogates N-Me-Cys for ICL, Dawson-Linker, SEA, and Hydrazides for NCL approach.

### 5.4.2 Full-length Synthesis of Ubiquitin by MW-assisted SPPS

With its full length of 76 AAs, Ub creates a significant challenge for conventional SPPS, which is limited to a length of about 50 AAs. This is mainly reflected by poorer efficiency in coupling AA or Fmoc removal due to aggregations or interactions of the peptide strands at the solid support. As a result, often, a large number of by-products characterized by truncated sequences is obtained. To get peptides of such length, powerful coupling reagents (such as HATU or COMU) are commonly used in high equivalents and in very long reaction times. Alternatively, dipeptide building blocks, so-called pseudoprolines, can be used. These cause a kink in the peptide backbone during the assembly of the peptide sequence, making the N-terminus more accessible to AA couplings or Fmoc cleavages (Figure 5.27).



**Figure 5.27:** An illustration of the impact of a pseudoproline dipeptide building block (highlighted in blue). It causes a "kink" conformation in the polyamide backbone of the peptide (**right**), which disrupts the structure and prevents peptide aggregation (**left**).

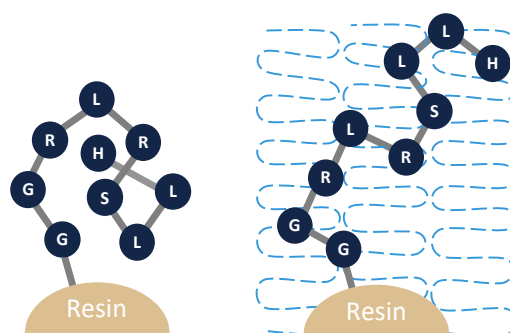


**Scheme 5.10:** Commercially available *N*-Fmoc protected Ser-, Thr-, and Cys-derived cyclic building blocks (pseudoprolines). *Xaa* = canonical amino acid. [373]

A new way to address the mentioned issues is using microwave (MW) radiation during peptide synthesis. [374,375] It is hypothesized that direct interactions with

amide dipoles of the dipole moments in peptides cause direct heating of the peptide leading to the prevention of aggregations and enabling efficient synthesis (Figure 5.28).<sup>[376]</sup> This effect is opposed to indirect heating by contact with the solvent. However, this theory is highly disputed as Bacsa *et al.* showed, that similar synthetic results of three peptides differing in sequence length (9 to 24 AAs) were reached by using conventional thermal heating with an oil bath at 86°C.<sup>[377]</sup>

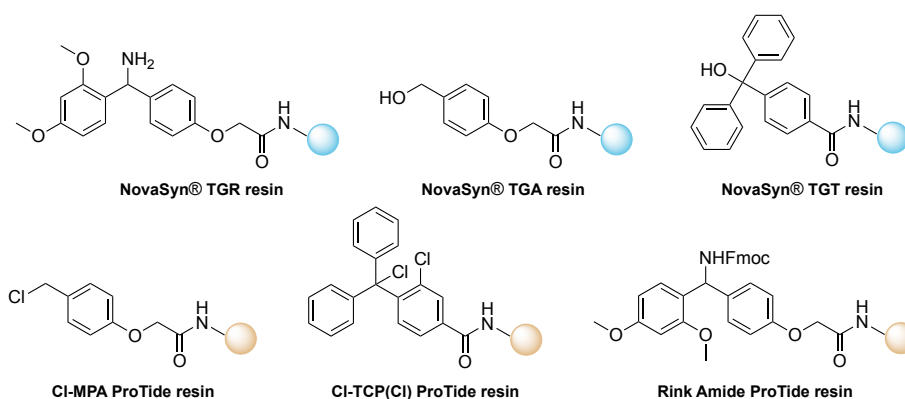
In 2016 the successful synthesis of Ub by MW-assisted SPPS was presented at the International Peptide Symposium in Leipzig.<sup>[378]</sup>



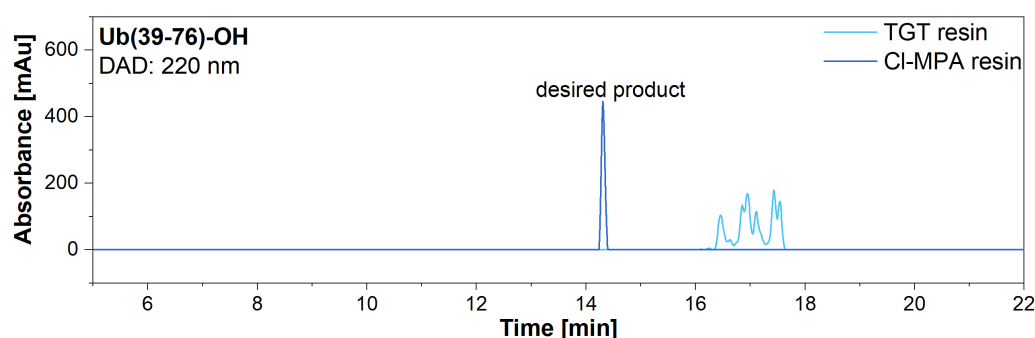
**Figure 5.28:** Peptide assembly on a solid support with microwave irradiation (**right**) and without (**left**). The microwave energy leads to a de-aggregation of the peptide. Adapted with permission from Palasek *et al.*<sup>[379]</sup> Copyright © 2006 European Peptide Society and John Wiley & Sons, Ltd.

#### 5.4.2.1 Challenges in MW-SPPS: Choice of Resin

During the establishment of the protocol in our laboratory, it was found that the high temperature and frequent heating stressed the solid support. This was evident when the NovaSyn TGR resin by Merck (Scheme 5.11 top) was utilized in the synthesis of Ub-amide (precursor for Ub2-*N*-Me-Cys), as it decomposed during the synthesis. While using the NovaSyn Fmoc-Gly-TGT resin in a test reaction for the synthesis of Ub acid (Ub-OH, precursor for thiol fragment), only poor quality of the synthesis was observed (Figure 5.29). The TGT resin has a very labile linker, which is probably why the synthesis could not be achieved in satisfactory yield already for 37 AAs. Similar observations were made in the FF03 synthesis (section 5.3.1.2). Therefore, the pro tide resins from CEM (Scheme 5.11 bottom) were chosen for further optimizations since they were developed for the reaction conditions of the MW-assisted SPPS.



**Scheme 5.11:** Resins used in Ub synthesis: NovaSyn® resins (**top**) and pro tide resins (**bottom**).



**Figure 5.29:** HPLC-Chromatogram of reaction control after 37 AA of Ub utilizing NovaSyn-TGT and Cl-MPA resin. HPLC system: LaChrom®; Kinetex® C18 column; eluents: A = H<sub>2</sub>O, B = ACN both containing 0.1% (v/v) TFA; gradient: 10% → 80% B over 18 min.

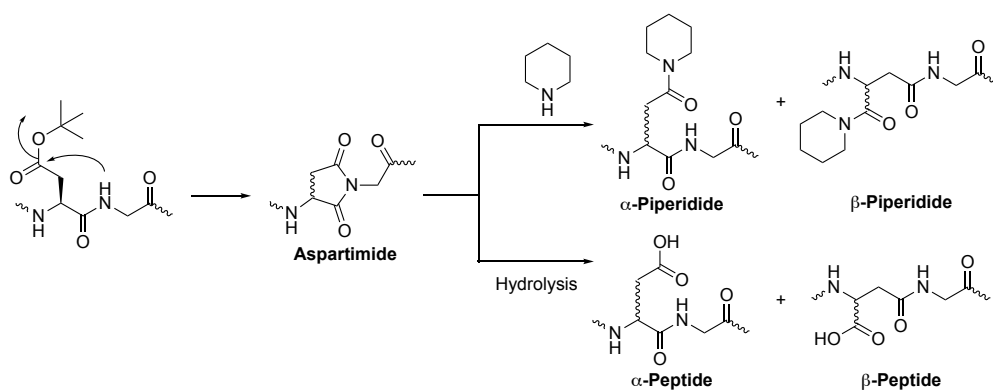
#### 5.4.2.2 Challenges in MW-SPPS: Enhancement of Side-reactions

As used in MW-assisted SPPS, elevated temperature can facilitate side reactions like aspartimide formation or epimerization of histidine.

##### *Aspartimid formation*

Ub has five segments in which aspartimide formation can occur (highlighted in blue in the sequence of Ub-wt in Table 5.7). This is the most common side reaction during the incorporation of Asp into a sequence that produces a cyclic by-product (Scheme 5.12). In general, the carboxyl function of Asp's side chain is protected with a <sup>t</sup>Bu ester during peptide assembly, which is removed during TFA cleavage. However, in a sequence of -Asp-X<sub>AA</sub>-, this ester can be targeted and removed by various nucleophiles during peptide assembly or storage, resulting in a racemic mixture of a five-membered ring.

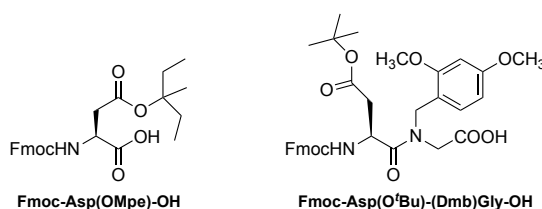




**Scheme 5.12** Mechanism of aspartimid formation and the resulting side products. [380]

This cyclic side product might survive the following reactions and be detected in the final product as an impurity with a  $-18$   $m/z$  difference relative to the desired product. On the other hand, it might undergo a subsequent ring-opening process through nucleophilic attack either by hydrolysis (identical  $m/z$ ) or piperidine addition ( $+67$  in  $m/z$ ). It is difficult to distinguish between the  $\alpha$ - or  $\beta$ -form by conventional mass analysis or HPLC separation in both cases. Because this side reaction can occur during any of the repetitive base-promoted  $N_{\alpha}$ -Fmoc deprotection steps, this effect is accumulative, enhancing the level of aspartimide side product formation.

Therefore, countermeasures were applied to the sequence: Fmoc-Asp(OMpe)-OH and Asp(O<sup>t</sup>Bu)-(Dmb)-Gly-OH were incorporated in parts that are prone to this side reaction. Additionally, 10% piperazine (in NMP/EtOH 9:1 % v/v) instead of 20% Pip (in DMF) was applied with 0.1 M HOBT in Fmoc-removal. Piperazine has a lower  $pK_a$  value (piperazine  $pK_a = 9.73$  vs. piperazine  $pK_a = 10.78$ ), which is why it may reduce the likelihood of deprotonation of the  $N_{\alpha}$ .

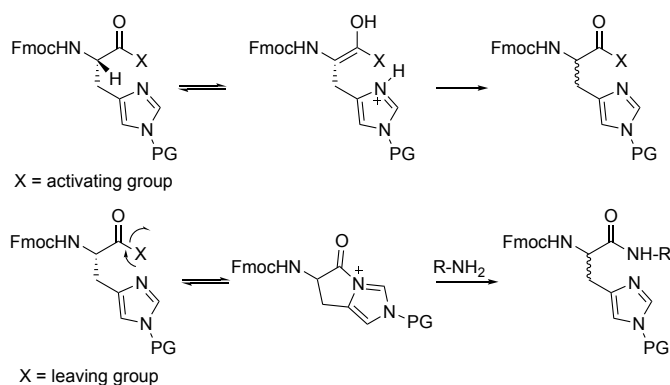


**Scheme 5.13:** Applied sterically hindered Asp building blocks in Ub synthesis to avoid aspartimid formation.

### Histidine epimerization

His tends to epimerize, which is enhanced at high temperatures (Scheme 5.14); This is caused by an intramolecular attack of the unprotected  $N_{\pi}$  of the imidazolyl side chain. <sup>[381]</sup> On the one hand, this nitrogen might abstract the proton of  $C_{\alpha}$ , whose acidity is increased in the active ester formed during AA coupling. Upon re-protonation of the obtained enolate, racemization can occur at this position (Scheme 5.14 top). On the other hand, a nucleophile attack of the  $N_{\pi}$  on the carboxylate might occur, leading to the formation of a liable cyclic carbonyl imidazolide derivative. Ring-opening leads to racemization of the peptide (Scheme 5.14 bottom).

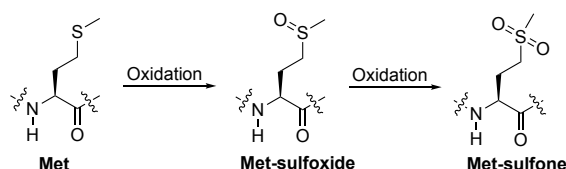
Therefore, His is usually coupled at 50°C for 10 min in MW-assisted SPPS. Nonetheless, the side reaction might still occur. Another option to address the epimerization process is described by applying different  $N_{\tau}$  protecting groups. <sup>[382]</sup> By incorporating Fmoc-His(Boc)-OH instead of Fmoc-His(Trt)-OH, the side-product formation was reduced to a minimum of 2% (instead of 7%) during Ub synthesis. Alternatively, Fmoc-His(Mbom)-OH can be used as it is reported to improve the results more effectively. This compound has not been included in this work for budgetary reasons. <sup>[382]</sup>



**Scheme 5.14:** Possible epimerization reactions of His according to Yang. <sup>[383]</sup>

#### 5.4.2.3 Oxidation of Met

Oxidation of Met is a well-known side reaction in peptide synthesis that occurs independently of the application of MW conditions. The formation of the corresponding sulfoxide or sulfone (Scheme 5.15) can occur during peptide assembly, purification, or storage. <sup>[384]</sup>



**Scheme 5.15.** Oxidation of Met to the corresponding sulfoxide and sulfone, according to Yang. [384]

To avoid this possible side reaction at the N-terminal Met in Ub, this residue was replaced with Nle, which is considered an isostereomer of Met. This replacement often has a low influence on a protein's function or structure. [385–389]

### 5.4.3 Comparison of MW-SPPS and Conventional SPPS in Ub Synthesis

Despite all challenges that arise with high temperatures in SPPS, MW radiation became a meaningful tool in accessing large peptides and difficult sequences with improved quality, which is elucidated in the synthesis of the wt Ub2-N-Me-Cys sequence (Table 5.8). Here it is evident that the desired compound was synthesized in a shorter reaction time while reducing amounts of solvents or reagents and reaction time.

**Table 5.8:** Comparison of Ub2-N-Me-Cys synthetic conditions: conventional [390] vs. MW-assisted SPPS.

Process	Conventional SPPS <sup>a</sup>		MW-assisted SPPS <sup>b</sup>	
	reagents (in DMF)	duration at rt	reagents	duration at 90°C
<b>Fmoc removal</b>	20% piperidine (+ 0.1 M HOBT)	3 x 10 min	10% piperazine (+ 0.1 M HOBT) <sup>c</sup>	1 min
<b>Coupling of canonical AA</b>	10 eq. Fmoc-AA-OH <sup>d</sup> + 20 eq. DIPEA + 10 eq. HCTU	2 x 60 min	5 eq. Fmoc-AA-OH + 5 eq. Oxyma Pure + 5 eq. DIC	1 x 2 min (4-30 AA) <sup>e</sup> 2 x 2 min (31-76 AA)
<b>Coupling of dipeptides</b>	2 eq. HCTU + 4 eq. DIPEA + 2 eq. Fmoc-YAA-XAA-OH <sup>f</sup>	5 h + overnight	5 eq. Fmoc-DG-OH + 5 eq. Oxyma Pure + 5 eq. DIC	1 x 2 min
<b>Coupling of fluorinated AA</b>	1.2 eq. Fmoc-AA-OH + 1.2 eq. DIC + 1.2 eq. HOAt	6 h + overnight	2 eq. Fmoc-AA-OH + 2 eq. Oxyma Pure + 2 eq. DIC	1 x 10 min

**a)** combination of automated (P-11 by Activotec) and manual synthesis; **b)** first 4 AAs were coupled manually, then automated synthesis with Liberty Blue (CEM) was performed; **c)** (10% w/v) in NMP/EtOH 9:1 (% v/v); **d)** Phe, Gln, Arg, Asn, Pro, and the following AA were coupled 2 x 90 min; **e)** His was coupled for 10 min at 50°C, Fmoc-Asp(OMpe)-OH was used, and Fmoc-Arg(Pbf)-OH was implemented with double couplings; **f)** Asp(O<sup>t</sup>Bu)-(Dmb)-Gly-OH, (Fmoc-Ile-Thr(ψ(Me,Me)pro)-OH, Fmoc-Leu-Ser(ψ(Me,Me)pro)-OH.

With a protocol for the full-length protein in hand, the next goal was to establish the desired functionalities for ligation.

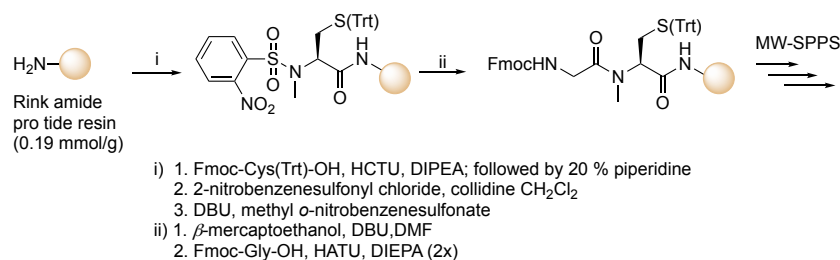
#### 5.4.4 Thioester Surrogates

As described above, different surrogates were screened: on the one hand, the *N*-Me-Cys at the C-term for the ICL approach, and on the other hand, Dawson linker, SEA (bis(2-sulfanylethyl)amido)) as well as hydrazides for the NCL approach. The following sections describe the studies performed and the corresponding results obtained.

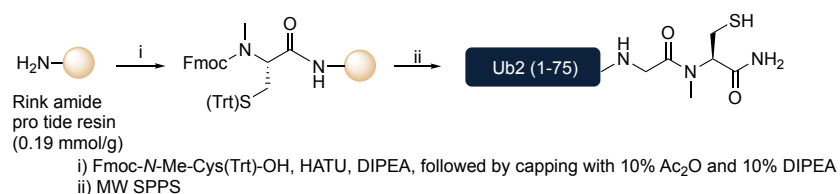
##### 5.4.4.1 Ub2-*N*-Me-Cys as Thioester Surrogate for the ICL Approach

The studies on the Ub2-*N*-Me-Cys provided the basis for the optimizations of the full-length synthesis of Ub. For the desired functionalization, the protocol of Brik's group was followed.<sup>[391]</sup> Initially, the Rink amide pro tide resin was loaded with Fmoc-Cys(Trt)-OH using HCTU/ DIPEA. *N*-methylation was performed thereafter. For this, the Fmoc protecting group was first removed with 20% Pip. In a three-step procedure, the desired functionality was implemented: 1) sulfonylation through treatment with 2-nitrobenzenesulfonyl chloride (*o*-NBS) and collidine, 2) methylation by reaction methyl 4-nitrobenzenesulfonate and DBU; 3) NBS-removal by DBU and  $\beta$ -mercaptoethanol followed.<sup>[392]</sup> Subsequently, the Gly was introduced through a double coupling with HATU chemistry (Scheme 5.16).

Judging from HPLC analysis, the obtained product seemed to be a good purity; however, the mass analysis revealed a contamination of -15 m/z, which might correspond to the absence of the methyl group on the Cys. It was decided for the further procedure to use commercially available Fmoc-*N*-Me-Cys(Trt)-OH. This was incorporated at the C-term with a manual single coupling using HATU, and a capping (10% Ac<sub>2</sub>O + 10% DIPEA in DMF) of the possible unreacted sites on the resin was performed followed by a loading test.

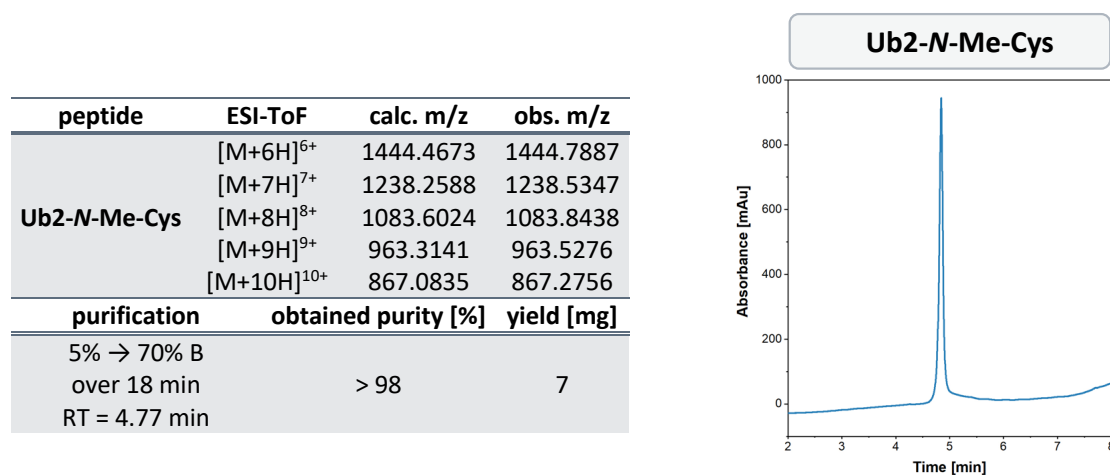


**Scheme 5.16:** Preparation of *N*-Me-Cys as thioester surrogate according to Brik *et al.* [391]



**Scheme 5.17:** Optimized synthesis of Ub2-*N*-MeCys by using commercially available Fmoc-*N*-Me-Cys(Trt)-OH.

With this procedure, the 7 mg of *N*-Me-Cys modified Ub2 was successfully synthesized and purified with no contaminations in mass analysis (Figure 5.30).

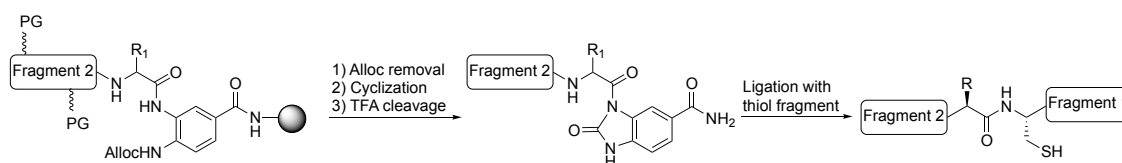


**Figure 5.30:** Analysis of final product Ub2(1-76)-*N*-Me-Cys. **Left:** ESI-ToF analysis and conditions of purification; **right:** Analytical HPLC chromatogram of final product. HPLC system: Chromaster® (FF), method 5-70% MeCN in H<sub>2</sub>O + 0.1% TFA in 10 min, C18 column, DAD 220 nm. The chromatogram is depicted from 2 to 8 min.

### 5.4.4.2 Different C-terminal Functionalizations to Obtain Ubiquitin Thioester Surrogates for the NCL Approach

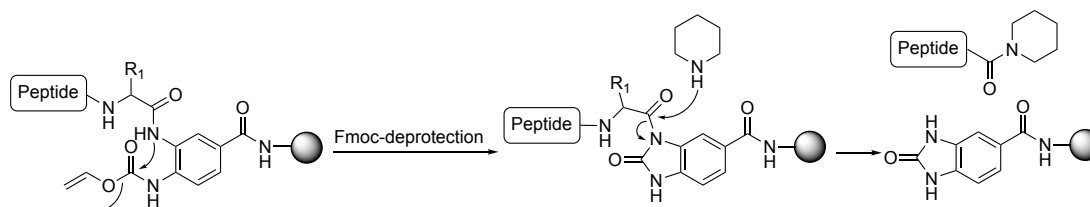
#### *Dawson-linker (Dbz) Strategy*

As described above, the Dawson linker (Dbz) is already known in the literature for synthesizing different diUb by conventional SPPS.<sup>[323]</sup> Here, the Dbz linker is first equipped with the Alloc protecting group.<sup>[393]</sup> The synthesis of the peptide backbone is then carried out at the *o*-position (Scheme 5.18). After complete assembly of the desired peptide, the Alloc protecting group is cleaved off, and Dbz is cyclized with *p*-nitrophenyl chloroformate and DIPEA to give *N*-acylbenzimidazolinone (Nbz) in a two-step procedure. After this treatment, the peptide is removed from the solid phase with TFA and purified by HPLC.

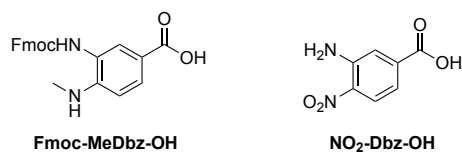


**Scheme 5.18:** Short overview of Dawson linker strategy.<sup>[326]</sup>

When adapting the strategy to MW-assisted SPPS, it is essential to consider that the common Alloc protecting group on the Dbz is unsuitable, as cleavage of this protecting group may occur at the typical high temperatures of MW-SPPS during Fmoc-cleavage (Scheme 5.19). Therefore, Fmoc-MeDbz<sup>[394]</sup> and NO<sub>2</sub>-Dbz<sup>[395]</sup> (Scheme 5.20) are reported in the literature as suitable alternatives; the latter is reduced to an amino group by SnCl<sub>2</sub> after synthesis for cyclization to the Nbz. The methylated Dbz linker has already been described, particularly for C-terminal Gly-units, as is the case with Ub.<sup>[394]</sup>



**Scheme 5.19:** Described side reaction when applying Alloc-protected Dbz linker in MW-assisted SPPS.<sup>[395]</sup>



**Scheme 5.20:** Alternative Dbz-Linkers for MW-assisted SPPS.

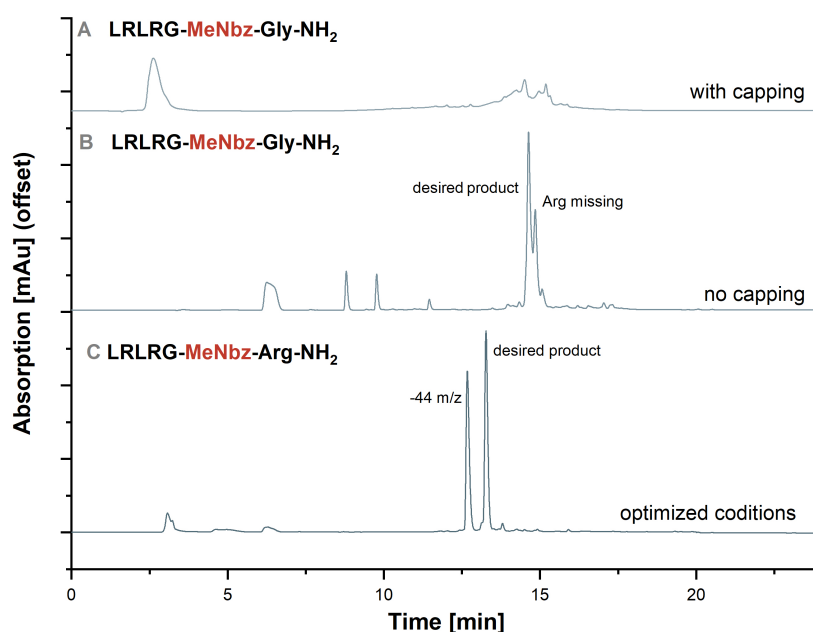
The synthesis was carried out on the Rink-amide pro tide resin (0.19 mmol/g loading). Before incorporating the MeDbz-linker, it is recommended to include a short spacer, e.g., an Arg or Gly unit was described for this purpose.<sup>[394]</sup> Due to its poor solubility in DMF, MeDbz was introduced by manual coupling and could not be added by the synthesis robot. Once the linker was installed, it was not possible to perform capping (acetylation with Ac<sub>2</sub>O and DIPEA) steps as the desired product was not observed in control micro cleavages (Figure 5.31 A). These usually allow better control of the coupling quality, especially in manual synthesis or when difficult building blocks are incorporated. In addition, it was very important to pay attention to the DIPEA eq. Initially, HATU coupling chemistry was used as well as ten eq. of DIPEA; this did not lead to the desired quality of coupling efficiency. When switching to HBTU and using 5.5 eq. DIPEA, much better results were obtained. This was followed by further synthesis under MW conditions. Initially, the coupling of the first AA after the Dbz-linker proved to be challenging. The coupling was finally achieved by MW-SPPS using ten eq. of Fmoc-Gly-OH. In the further course, the installation of an Arg turned out to be complicated, even though a double coupling with ten eq. was carried out (Figure 5.31 B). Therefore, a triple coupling of Arg with ten eq. each fulfilled the desired requirements.

Another challenge with this approach is the conversion from Dbz to Nbz before the peptide is cleaved of the solid phase. Here, different concentrations of DIPEA, as well as different reaction times, were investigated. In the process, entry 8 provided the best results (Table 5.9).

**Table 5.9:** Conditions for the two-step cyclization procedure to convert Dbz to Nbz. The best conditions are highlighted in red.

entry	acetylation with <i>p</i> -nitrophenyl chloroformate duration at rt	cyclization duration and conc. of DIPEA treatment
1	1 h	0.5 h, 0.5 M
2	1 h	1 h, 0.5 M
3	1 h	0.5 h, 1 M
4	1 h	1 h, 1 M
5	2 h	0.5 h, 0.5 M
6	2 h	1 h, 0.5 M
7	2 h	0.5 h, 1 M
8	<b>2 h</b>	<b>1 h, 1 M</b>

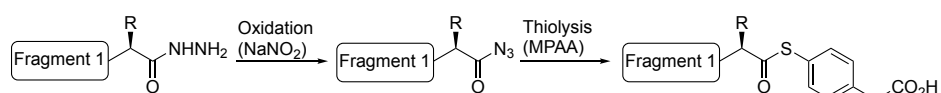
Despite all efforts to optimize the synthesis, establishing the Dawson linker as a suitable thioester surrogate for the NCL approach remains to be achieved. In the test series with the best conditions, another by-product was obtained at RT = 12.58 min (Figure 5.31 C), which exhibits a difference of -44 m/z in the ESI-ToF. This by-product could not be assigned to any known side reaction. Therefore, the Dawson linker approach was not pursued further.

**Figure 5.31:** Summary of reaction controls to establish Me-Dbz as thioester surrogate for NCL approach. HPLC system: Chromaster®, Kinetex® C18 column; eluents: A = H<sub>2</sub>O, B = ACN both containing 0.1% (v/v) TFA; gradient: 10% → 80% B over 18 min. The chromatograms are depicted from 0 to 22 min.



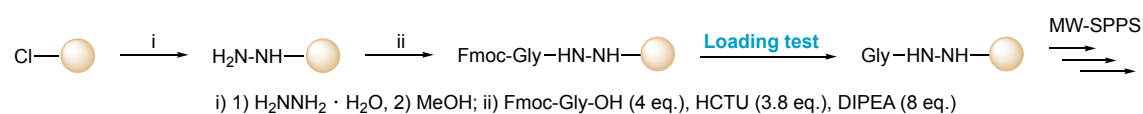
### Hydrazides as Thioester Surrogates for NCL Approach

Hydrazide functionality was used in the synthesis of Ub chains. [339–341] In general, hydrazides can be converted *in situ* to the desired thioester (Scheme 5.21). However, other, shorter fragments were selected here than in the present work. Correspondingly functionalized resins are readily available commercially, but they are typically quite highly loaded (0.3 to 0.8 mmol/g) and thus unsuitable for the planned fragment lengths in the presented work.



**Scheme 5.21:** General procedure to convert hydrazide functionalized peptides into the desired thioester.

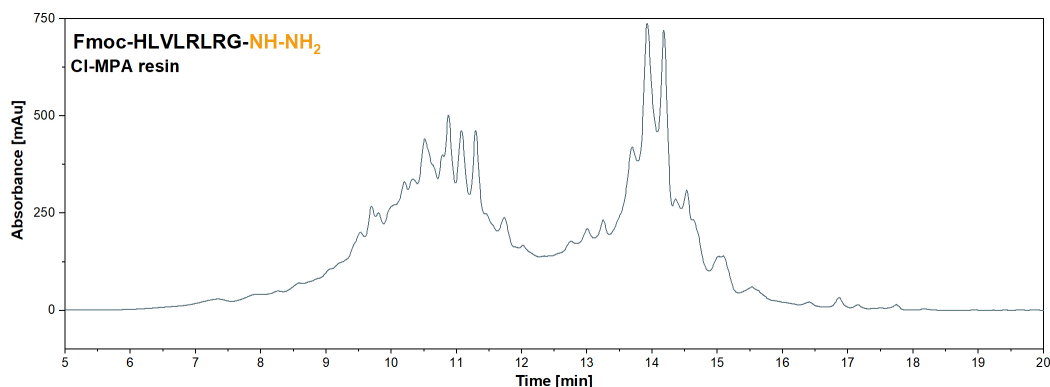
Encouragingly, chlorine-containing resins can be converted into hydrazides by treatment with hydrazine hydrate in DMF (Scheme 5.22). [396] The remaining Cl-functionalities are quenched by using 10% MeOH in DMF. A Kaiser test [397] involving a color reaction of ninhydrin with a primary amino group to monitor the coupling efficiency is unsuitable for detecting hydrazine functionalization. Therefore, an indirect detection was selected; after the coupling of the first AA, Fmoc cleavage was performed, and the supernatant obtained was measured by UV absorption at 301 nm. By the extinction coefficient of dibenzofulvene–piperidine adduct, a conclusion can be drawn about the loading and, thus, the successful functionalization. [398]



**Scheme 5.22:** Preparation of a hydrazine trityl resin and coupling of the first AA according to the procedure published by Zheng *et al.* [399]

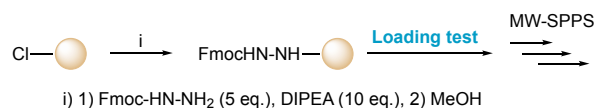
The present work investigated Cl-TCP(Cl) (0.16 mmol/g) and Cl-MPA pro tide resins (0.23 mmol/g loading) (Scheme 5.11 bottom). Cl-TCP(Cl) was successfully functionalized, but an unsatisfactory yield of 0.04 mmol/g of 0.16 mmol/g maximal loading capacity was obtained. It is known that the functionalization of 2-chlorotryl resins is affected by hydrolysis reaction during treatment with hydrazine hydrate leading to lower yields. [400] However, this loading was too low to continue with peptide assembly.

In several approaches, the Cl-MPA functionalization by treatment with hydrazine hydrate indicated a quantitative result, leading to this resin being chosen for further synthetic investigations. However, the micro-cleavages for quality control after eight AAs did not show successful product formation (Figure 5.32), as no confirmation in ESI-ToF analysis was obtained.



**Figure 5.32:** HPLC-Chromatogram of reaction control after 30 AA of Ub utilizing NovaSyn-TGT resin. HPLC system: LaChrom®; Kinetex® C18 column; eluents: A = H<sub>2</sub>O, B = ACN both containing 0.1% (v/v) TFA; gradient: 10% → 80% B over 18 min. The chromatogram is depicted from 5 to 20 min.

Therefore, another functionalization approach was conducted using Fmoc-NHNH<sub>2</sub> and DIPEA in an overnight reaction at rt, allowing the direct confirmation of functionalization by a loading test (Scheme 5.23).<sup>[400,401]</sup> Interestingly, using Cl-MPA resin, no loading functionalization was detected. Several conditions were examined for the desired reaction (Table 5.10). In MW-SPPS, this resin is loaded using KI and DIPEA as activation reagents, which is why MW conditions were also tested for functionalization. Here, no desired functionalization was determined.



**Scheme 5.23:** Route for Fmoc-NHNH<sub>2</sub> functionalization allowing a direct loading test prior to AA coupling.

**Table 5.10:** Functionalization of Cl-MPA resin by treatment with DIPEA and Fmoc-NH-NH<sub>2</sub>.

entry	conditions <sup>a</sup>	loading test (20% Pip) <sup>b</sup>
1	5 eq. Fmoc-NH-NH <sub>2</sub> (0 °C to rt) 14 h in flask	No reaction
2	5 eq. Fmoc-NH-NH <sub>2</sub> (+ KI + MW) <sup>c</sup>	No reaction
3	10 eq. Fmoc-NH-NH <sub>2</sub> (+ KI + MW) <sup>c</sup>	No reaction
4	2 x 10 eq. Fmoc-NH-NH <sub>2</sub> (+ KI + MW) <sup>c</sup>	No reaction

**a:** all reactions were performed with + 10 eq. DIPEA in 1:1 DCM/DMF

**b:** loading procedure according to AAptec, UV/Vis absorption at 301 nm. <sup>[402]</sup>

**c:** DCA-Trityl loading procedure on Liberty Blue procedure: 1 min 80°C, 9 min 90°C.

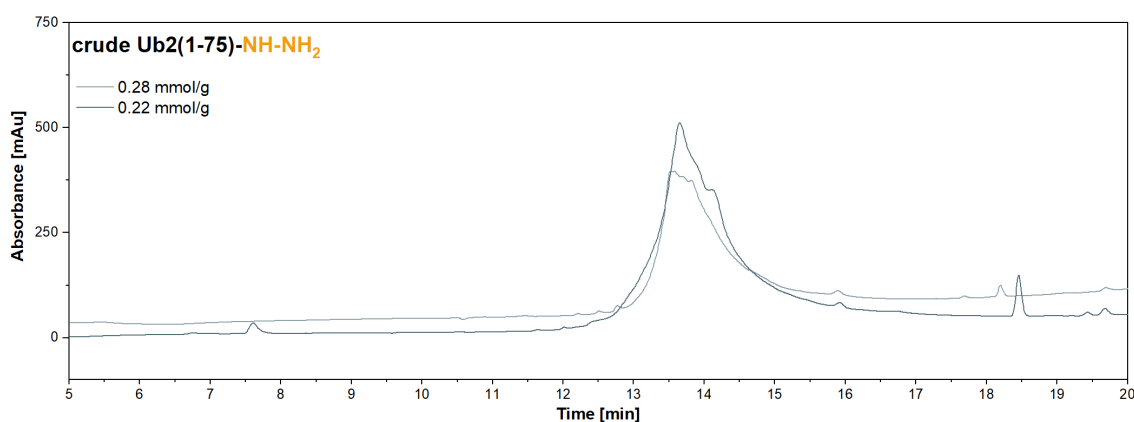
Functionalization of Cl-TCP(Cl) (maximal loading capacity 0.40 or 0.37 mmol/g) achieved loadings of 0.17 to 3.2 mmol/g (Table 5.11). Optimal loading (0.17-0.22 mmol/g) was achieved by reducing the reaction time to 2 h at rt. Nevertheless, some challenges arose using this protocol: First, functionalization in a flask usually had a higher yield but resulted in clogging of the reaction vessel may be due to friction of the solid support by a magnetic stirrer. In addition, high loading led to a less effective synthesis (Figure 5.33).

**Table 5.11:** Functionalization of Cl-TCP(Cl) resin with Fmoc-NH-NH<sub>2</sub>.

entry	conditions <sup>a</sup>	loading test (20% Pip) <sup>b</sup>
1	5 eq. Fmoc-NH-NH <sub>2</sub> (0°C to rt) 14 h in flask	0.32 mmol/g (80%)
2	5 eq. Fmoc-NH-NH <sub>2</sub> (0°C to rt) 14 h in reactor	0.28 mmol/g (70%)
3	5 eq. Fmoc-NH-NH <sub>2</sub> (0°C to rt) 14 h in flask	0.26 mmol/g (65%)
4	5 eq. Fmoc-NH-NH <sub>2</sub> (0°C to rt) 4 h in reactor	0.20 – 0.27 mmol/g (50-68%)
5	5 eq. Fmoc-NH-NH <sub>2</sub> (rt) 2 h in reactor	0.17-0.22 mmol/g (43-55%)

**a:** all reactions were performed with + 10 eq. DIPEA in 1:1 DCM/DMF

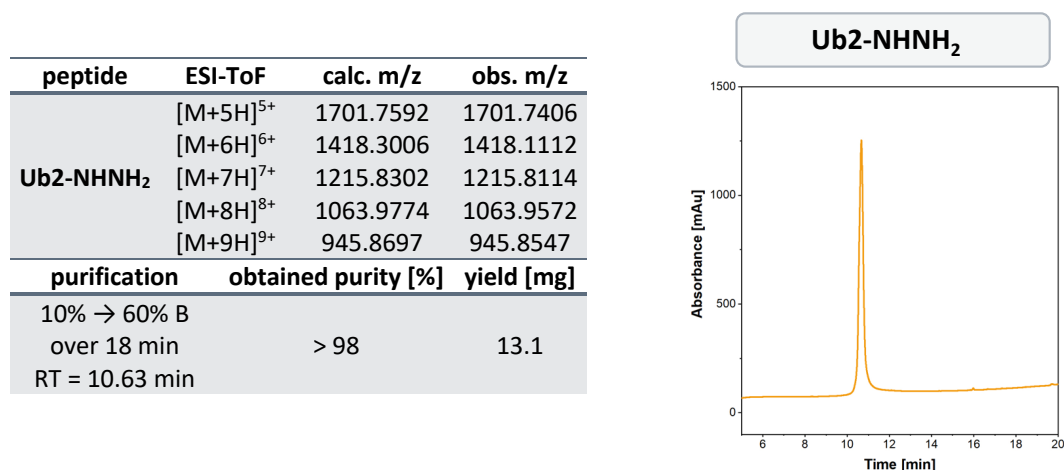
**b:** loading procedure according to AAptec, UV/Vis absorption at 301 nm. <sup>[402]</sup>



**Figure 5.33:** HPLC-Chromatograms depict the influence of loading capacity on product quality. HPLC system: LaChrom®; Kinetex® C18 column; eluents: A = H<sub>2</sub>O, B = ACN both containing 0.1% (v/v) TFA; gradient: 10% → 80% B over 18 min. The chromatogram is depicted from 5 to 20 min.

In general, after the functionalization, the swelling properties of the resin changed, leading to the need for additional washing steps after AA coupling in the protocol for MW-SPPS. Another change in the protocol was established by reducing the single coupling segment from 30 AAs to 20 AAs.

Finally, this conversion protocol was used for the successful synthesis of the desired Ub thioester surrogate. After purification 13.1 mg of **Ub2-NH<sub>2</sub>** were obtained (Figure 5.34).

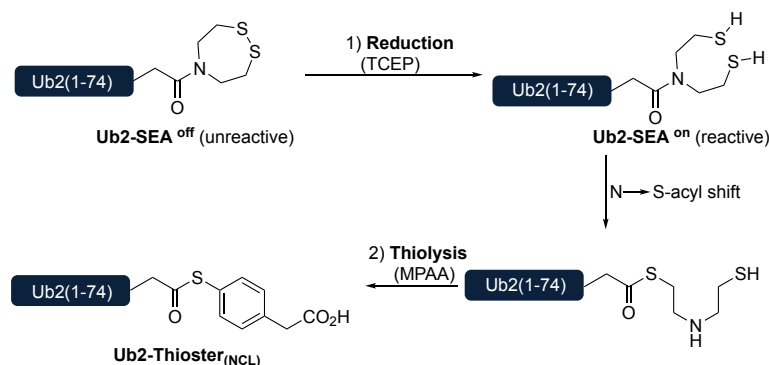


**Figure 5.34:** Analysis of final product Ub2(1-75)-NH<sub>2</sub>. **Left:** ESI-ToF analysis and conditions of purification; **right:** Analytical HPLC chromatogram of final product. HPLC system: Primaide®, Kinetex® C18 column, method 10-60% MeCN in H<sub>2</sub>O + 0.1% TFA in 18 min, DAD 220 nm. The chromatogram is depicted from 5 to 20 min.

### ***N-S shift system bis(2-sulfanylethyl)amido) (SEA) as masked thioester for NCL approach***

In recent years, interest in investigating new functionalities as thioester surrogates has increased rapidly. A promising approach here is the use of SEA. Unlike the Dbz linker, this functionality does not have to be incorporated at the C-term because commercial sources exist with preloaded AA on the solid support. These are polystyrene resins, which are also available in the low-loaded form. This makes them ideally suited for the synthesis of long sequences such as Ub.

Like with the Dawson linker approach, a 7 AA section was first synthesized with the resin. The Val70 was coupled with two eq. to mimic the conditions of flAA incorporation since this resin used has not yet been described for use in MW-SPPS.

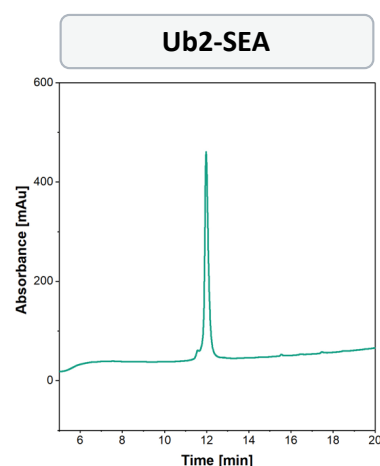


**Figure 5.35:** Planned Ub2-SEA modification for NCL approach.

The first synthesis proceeded without errors, whereupon the complete sequence was assembled on the resin following the established conditions in section 5.4.2. For the cleavage with TFA, a special cocktail (TFA, DMS, TIS, thioanisol, H<sub>2</sub>O: 90/2.5/2.5/2.5/2.5 % v/v) has to be used for SEA at the C-terminus, and the cleavage has to be performed for three h at rt.

Before purification, it is necessary to transfer SEA into the unreactive state SEA<sup>off</sup>. Otherwise, the purification might be complicated since the free thiol groups can always perform N-to-S acyl shifts (Figure 5.35). For this purpose, the crude product is first oxidized with iodine in DMSO; this must not take longer than 30 s. Subsequently, the excess iodine is quenched with DTT. Then the peptide is dissolved in a mixture of AcOH and water (1:4 % v/v) and purified by preparative HPLC to yield 13.4 mg of the desired thioester surrogate **Ub2-SEA** (Figure 5.36).

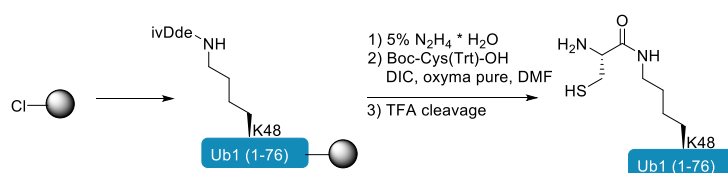
peptide	ESI-ToF	calc. m/z	obs. m/z
<b>Ub2-SEA</b>	[M+6H] <sup>6+</sup>	1435.6253	1435.4552
	[M+7H] <sup>7+</sup>	1230.6799	1230.5166
	[M+8H] <sup>8+</sup>	1076.9709	1076.9600
	[M+9H] <sup>9+</sup>	957.4195	957.3702
	[M+10H] <sup>10+</sup>	861.7783	861.6742
purification	obtained purity [%]	yield [mg]	
10% → 60% B over 18 min RT = 11.97 min	> 96	13.4	



**Figure 5.36:** Analysis of final product Ub2(1-75)-SEA. **Left:** ESI-ToF analysis and conditions of purification; **right:** Analytical HPLC chromatogram of final product. HPLC system: Primaide® Jupiter® C18 column (300Å), method 10-60% MeCN in H<sub>2</sub>O + 0.1% TFA in 18 min, DAD 220 nm. The chromatogram is depicted from 5 to 20 min.

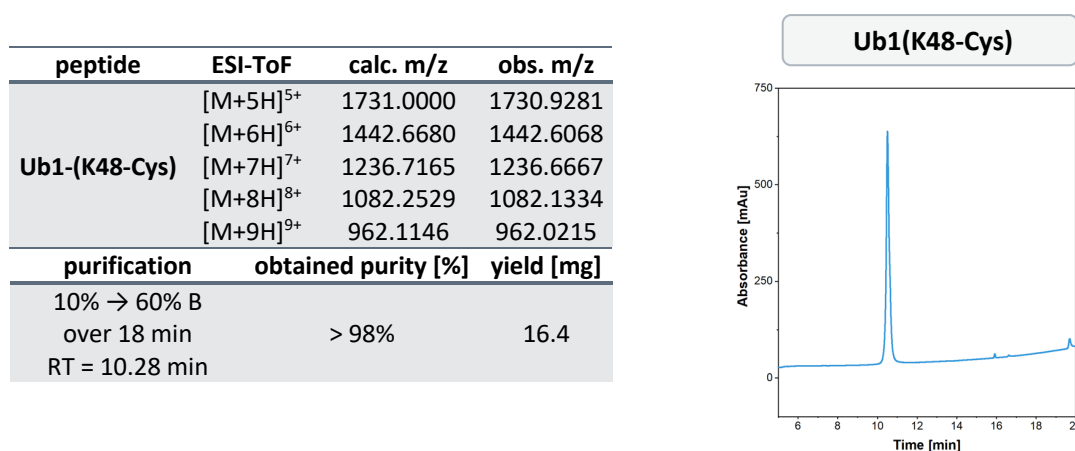
### 5.4.5 Synthesis of Thiol Fragment of the NCL approach

The second monomer for the NCL strategy contains a Cys coupled to the side chain of Lys48. To enable this, a Fmoc-Lys(ivDde)-OH is introduced at the desired position (Figure 5.37). After regioselective cleavage of this protecting group with 5% hydrazine in DMF (3 x 3 min at 90°C), Boc-Cys-OH is implemented in a 4 min double coupling at 90 °C. It is important to note that the last AA (Nle in the present case) at the N-terminus must also be equipped with a Boc-protecting group.



**Figure 5.37:** Strategy to obtain Ub1(K48-Cys).

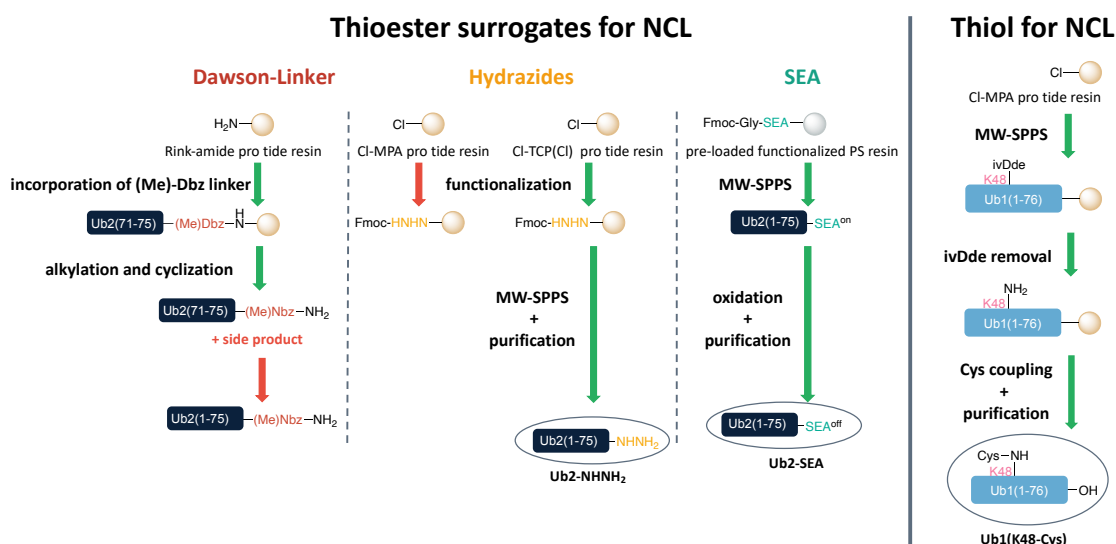
Each crucial step was monitored with micro-cleavages showing the desired outcome. The thiol fragment was successfully synthesized (16.4 mg) and purified (Figure 5.38).



**Figure 5.38:** Analysis of final product Ub1(K48-Cys). **Left:** ESI-ToF analysis and conditions of purification; **right:** Analytical HPLC chromatogram of final product. HPLC system: Chromaster®, Kinetex® C18 column, method 10-60% MeCN in H<sub>2</sub>O + 0.1% TFA in 18 min, DAD 220 nm. The chromatogram is depicted from 5 to 20 min.

### 5.4.6 Summary of Synthesized Monomers for Native Chemical Ligation

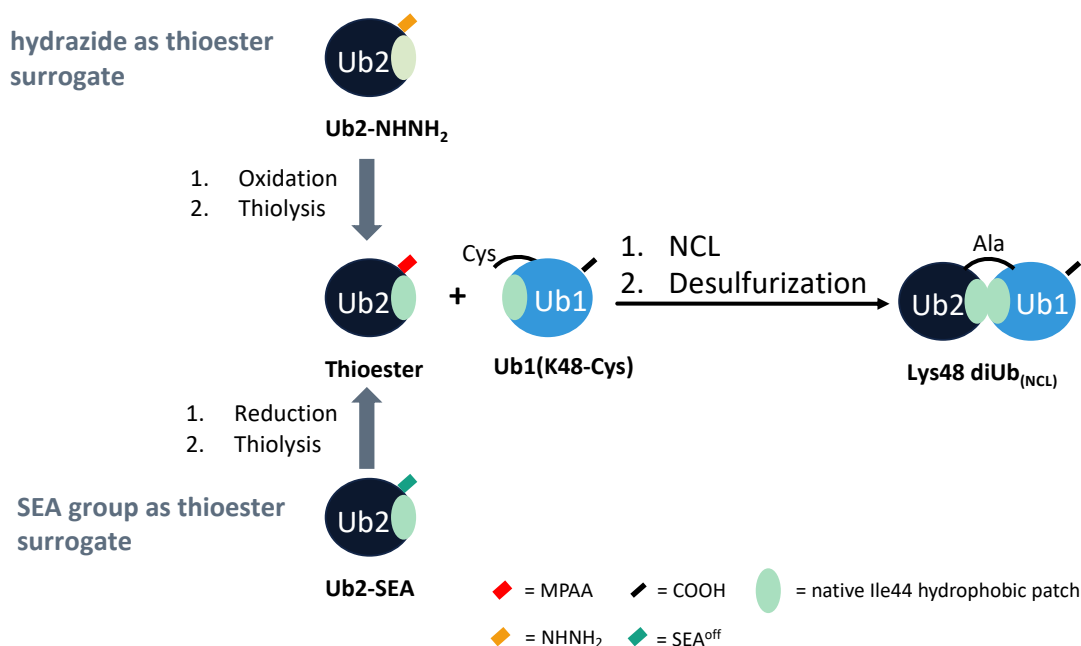
Overall, we succeeded in establishing protocols for the synthesis and purification of the desired thiol fragment (**Ub1(K48-Cys)**) and two thioester surrogates (**Ub2(1-75)-NHNH<sub>2</sub>** and **Ub2-SEA**). Figure 5.39 provides a summary of the elaborated processes. Encircled compounds were next used in native chemical ligation studies.



**Figure 5.39:** Comprehensive overview of investigated processes. Green arrows indicate successful protocols, and red arrows represent unsuccessful trials.

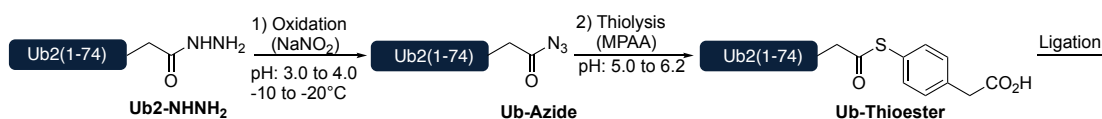
### 5.4.7 Native Chemical Ligation for Ubiquitin Dimerization

The two thioester surrogates, **Ub2-NHNH<sub>2</sub>** and **Ub2-SEA**, were studied with **Ub1(K48-Cys)** in native chemical ligation. For this purpose, the C-terminal functionalities of the thioester surrogates must be first converted into the corresponding thioester. This is usually done *in situ*, i.e., the resulting thioester is not isolated but reacts directly with the thiol fragment (**Ub1(K48-Cys)**).



**Figure 5.40:** Schematic representation of planned ligation studies. First, the thioester surrogates must be transferred into the desired thioester. In the case of Ub2-NHNH<sub>2</sub>, an oxidation to the corresponding Ub2-azide must occur before thiolysis. For Ub2-SEA, a TCEP treatment (reduction) needs to be performed to generate the S-to-N acyl shift system that is reacted in a thiolysis to form the corresponding thioester.

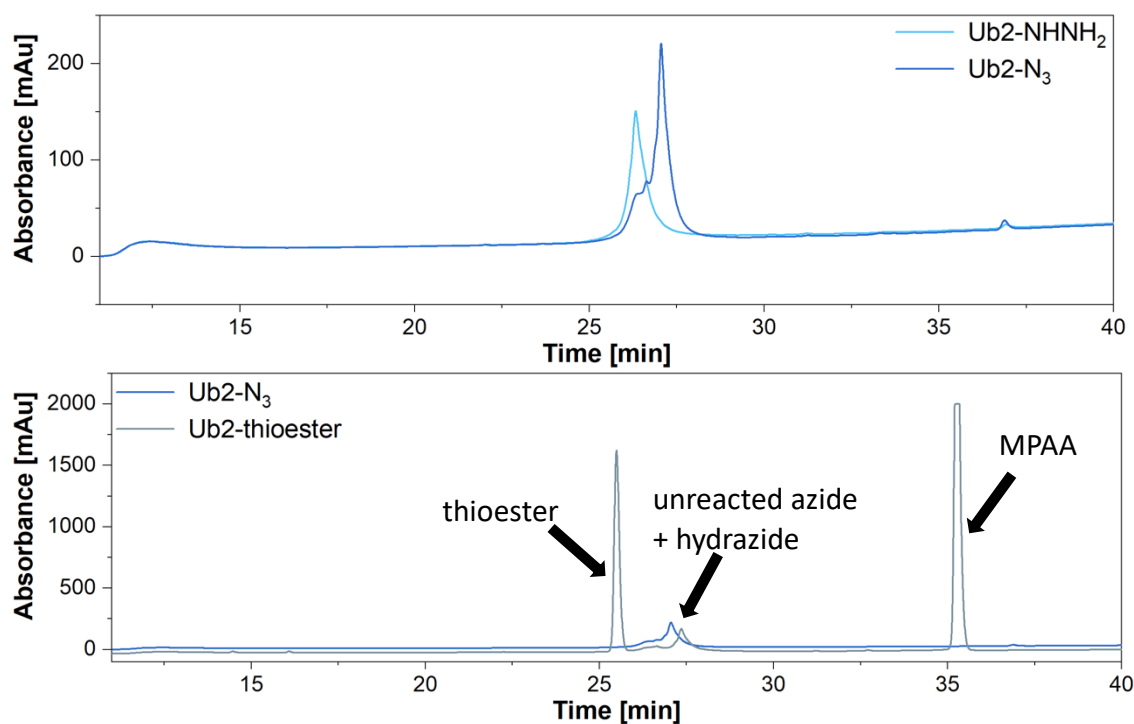
First, following the protocol of Zheng *et al.*, ligation with the hydrazide functionalized thioester surrogate **Ub2-NHNH<sub>2</sub>** was investigated.<sup>[399]</sup> Therefore, the hydrazide is converted to the desired thioester first (Scheme 5.24).



**Scheme 5.24:** Scheduled conversion of Ub2(1-75)-NHNH<sub>2</sub> to the corresponding MPAA-thioester.

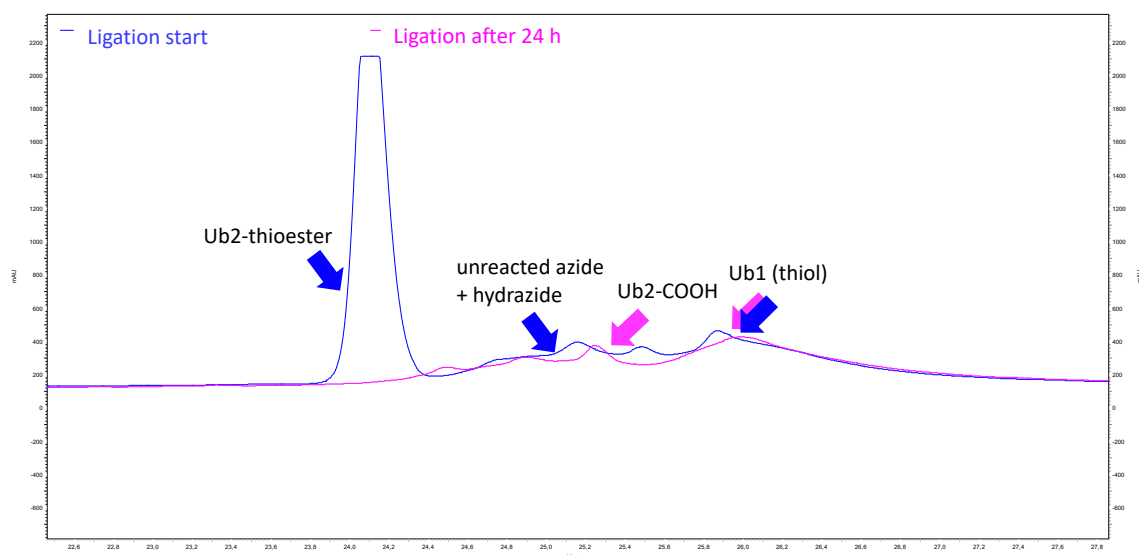
For this purpose, the hydrazide is dissolved at pH 3 in 6 M Gdn-HCl (+ 0.2 M phosphate buffer), cooled to -10°C, treated with 10 eq. NaNO<sub>2</sub> and carefully shaken for 15 min. HPLC analysis showed less than complete conversion of this step (Figure 5.41 top). However, the remaining hydrazide does not affect the next step or ligation. Consequently, thiolysis was conducted. For this, 40 eq. of MPAA were added to the reaction mixture, and the pH was adjusted to 5. The reaction was shaken for 5 min at rt. Reaction monitoring showed only a small amount of unreacted compounds (Figure 5.41 below).





**Figure 5.41:** HPLC-analysis of hydrazide to thioester conversion. **Top:** Oxidation of hydrazide to the corresponding azide. **Below:** thiolysis to furnish the desired thioester. HPLC system: LaChrom®, Jupiter® C18 column (300Å), method 5%(5 min)-20-60% MeCN in H<sub>2</sub>O + 0.1% TFA in 30 min, DAD 220 nm. The chromatogram is depicted from 11 to 40 min.

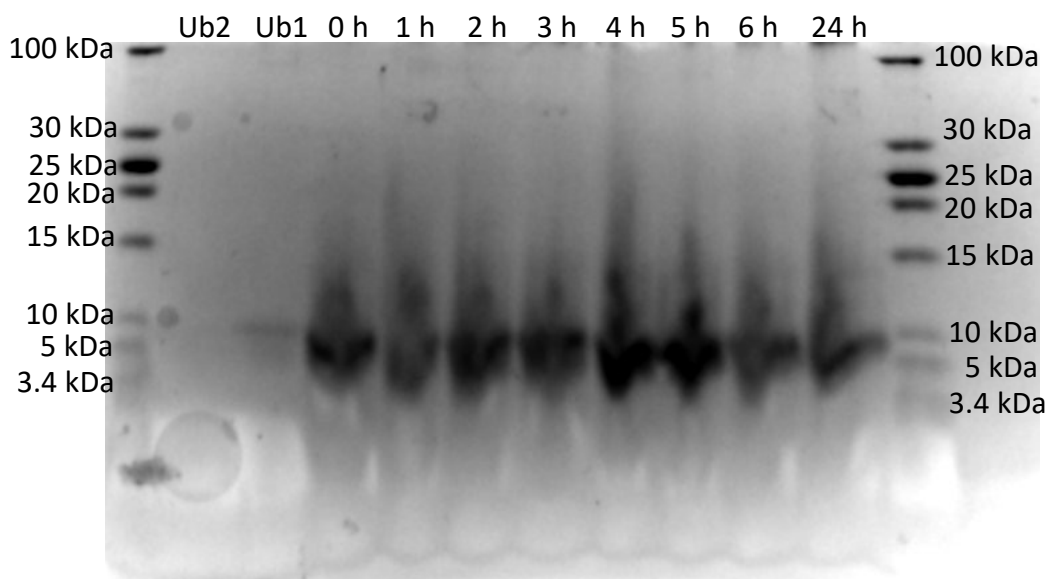
Ligation was continued with the *in situ* generated thioester. The thiol fragment **Ub1(K48-Cys)** (1 eq.) was added to the reaction mixture, and the pH was carefully adjusted to 6.5. After 24 h, the reaction was monitored by HPLC analysis showing a complete consumption of the thioester (Figure 5.42). Still, no formation of the desired Ub-dimer as the new species detected proved to be the hydrolyzed thioester (Ub2-COOH).



**Figure 5.42:** Investigation of ligation HPLC system: LaChrom®, Jupiter® C18 column (300Å), method 5%(5 min)-20-60% MeCN in H<sub>2</sub>O + 0.1% TFA in 30 min, DAD 220 nm. The chromatogram is depicted from 22 to 28 min.

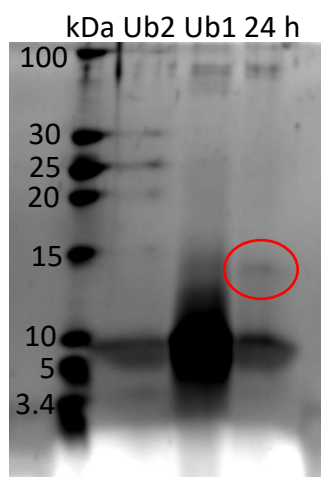
In the further course, the thioester was isolated and purified to guarantee a defined starting point for the planned kinetic studies. Attempts were also made to run the reaction at higher protein concentrations (2 - 4 mM instead of 1 mM). However, these attempts did not provide the desired outcome due to the poor solubility of both synthesized Ub monomers (**Ub2-NH-NH<sub>2</sub>** and **Ub1(K48-Cys)**). Excessive ultrasonic treatment, heat, exchange of Gnd · HCl, and previous desalting of the Ub monomers did not lead to any improvement.

With the help of the SEA group, the hydrolysis susceptibility of the thioester should be controlled since this functionality is based on an S-to-N acyl shift system. The reaction was carried out at a ratio of 1 eq. **Ub2-SEA** to 1.5 eq. **Ub1(K48-Cys)** at 37°C in a nitrogen atmosphere in a 100 mM phosphate buffer containing 6 M Gnd · HCl, 60 eq. TCEP and 60 eq. MPAA. Here, the studies were still impacted by the poor solubility of the monomers (**Ub2-SEA** and **Ub1(K48-Cys)**). Therefore, SDS-PAGE was also resorted to as a possible analytical method. Unlike HPLC, there is no risk of clogging with this method, and the size difference from the Ub monomers to the desired Ub dimer can be well elaborated with Tricine gels. [403]



**Figure 5.43:** Reaction control of SEA-ligation via SDS-PAGE using a 16% Tricine gel. 2  $\mu\text{L}$  of the reaction mixture was diluted with 18  $\mu\text{L}$  100 mM phosphate buffer pH 7.4 and treated with 5  $\mu\text{L}$  of Lämmli dye. A low-range protein ladder was used as an internal standard.

The first analysis showed an insufficient separation of the bands, presumably caused by the high content of Gnd·HCl (Figure 5.43). Therefore, for the second analysis, the 24 h sample was subjected to an extraction procedure with  $\text{Et}_2\text{O}$  to remove excessive MPAA, dialysis was performed to remove Gnd·HCl, and the sample was analyzed again by SDS-PAGE (Figure 5.44).



**Figure 5.44:** Reaction control of SEA-ligation via SDS-PAGE using a 16% Tricine gel after 24h. 2  $\mu\text{L}$  of the reaction mixture was diluted with 18  $\mu\text{L}$  100 mM phosphate buffer pH 7.4 and treated with 5  $\mu\text{L}$  of Lämmli dye. A low range protein ladder was used as an internal standard. The 24 h sample was extracted and dialyzed prior to the analysis.

A new band was detected in size range of the desired Ub dimer, indicating successful ligation. Confirmation by ESI-ToF was not available at the time of submission of the

present work. However, the promising result suggests that further optimization steps are required since, even after 24 h, a complete conversion to the Ub dimer has yet to be achieved.

### **5.5 Chemical Synthesis of Small all- $\beta$ Proteins to Study Fluorine's Impact on Protein Folding Kinetics**

**1FYN**, the phosphotransferase domain of Src-homology 3, and **1TEN**, the fibronectin type III domain of human tenascin, were selected for studies of fluorine-specific interactions on folding kinetics using stopped-flow Trp fluorescence.

#### **5.5.1 1FYN**

1FYN is an all- $\beta$  domain (62 AAs) that forms a  $\beta$ -barrel structure. According to the literature, studies on the folding behavior of this domain revealed three critical positions in the hydrophobic core Ile28, Ala39, and Ile50, which will be first replaced by Val in this presented work. Subsequently, these positions will be substituted by TfeGly, since this flAA has properties similar to Val concerning vdW volume and hydrophobicity as described above (see section 1.2.3.1). It is planned to perform single substitutions and eventually replace all three positions with TfeGly (Table 5.12). It is important to note that there are sequences of different lengths described for this domain, so the positions' numbering might seem not to match at first sight. However, the sequences contain only two Ile in the whole sequence, which are present at the described distance from each other. Also, the essential Ala is located in relation to the Ile residues described in the literature.

**Table 5.12:** Planned sequences for 1FYN derivatives.

Name	Sequence
1FYN-wt	GTGVTLFVALYDYEARTE <b>DDLSFHKGEKFQ</b> ILNSSEG <b>DW</b> WEARSLTTGETGYIPSNYVAPVD
1FYN-V28	GTGVTLFVALYDYEARTE <b>DDLSFHKGEKFQ</b> VILNSSEG <b>DW</b> WEARSLTTGETGYIPSNYVAPVD
1FYN-V39	GTGVTLFVALYDYEARTE <b>DDLSFHKGEKFQ</b> ILNSSEG <b>DW</b> WE <b>V</b> RSLTTGETGYIPSNYVAPVD
1FYN-V50	GTGVTLFVALYDYEARTE <b>DDLSFHKGEKFQ</b> ILNSSEG <b>DW</b> WEARSLTTGETGY <b>V</b> PSNYVAPVD
1FYN-28/39/50V (1FYN 3xV)	GTGVTLFVALYDYEARTE <b>DDLSFHKGEKFQ</b> VILNSSEG <b>DW</b> WE <b>V</b> RSLTTGETGY <b>V</b> PSNYVAPVD
1FYN-TfeGly28	GTGVTLFVALYDYEARTE <b>DDLSFHKGEKFQ</b> TfeGlyILNSSEG <b>DW</b> WEARSLTTGETGYIPSNYVAPVD
1FYN-TfeGly39	GTGVTLFVALYDYEARTE <b>DDLSFHKGEKFQ</b> ILNSSEG <b>DW</b> WE <b>TfeGly</b> RSLTTGETGYIPSNYVAPVD
1FYN-TfeGly50	GTGVTLFVALYDYEARTE <b>DDLSFHKGEKFQ</b> ILNSSEG <b>DW</b> WEARSLTTGETGY <b>TfeGly</b> PSNYVAPVD
1FYN-TfeGly28/39/50	GTGVTLFVALYDYEARTE <b>DDLSFHKGEKFQ</b> TfeGlyILNSSEG <b>DW</b> WE <b>TfeGly</b> RSLTTGETGY <b>TfeGly</b> PSNYVAPVD

## 5.5.2 Synthesis and Purification of 1FYN Derivatives and Pro-rich Peptides

### 5.5.2.1 Synthesis Protocol of 1FYN

Besides the long sequence, there are other challenges in synthesizing 1FYN with MW-SPPS. This protein domain contains the difficult combination of Asp39-Trp40 (highlighted in red in Table 5.12, more details are presented in part B). Here, the Fmoc-Asp(OMpe)-OH was introduced. Fmoc-Asp(O<sup>t</sup>Bu)-OH was used for the other Asp positions (highlighted in blue in Table 5.13). By using 10% piperazine with 0.1 M HOBt during Fmoc cleavage, aspartimide formation was successfully suppressed. TfeGly was incorporated using the established protocol with two eq. The sequence was generally constructed under the previously elaborated conditions: The first 20 AAs were coupled with single 4 min couplings, the rest with 2 min double couplings. Fmoc-Arg(Pbf)-OH was typically introduced with double couplings. Fmoc-His(Boc)-OH was coupled at 50°C for 10 min.

As in the pHLIP project, cocktail **III** (TFA/H<sub>2</sub>O/EDT/TIS: 94/2.5/2.5/1 % v/v) was used for cleavage from the solid support due to the Asp-Trp section. However, ESI-ToF analysis revealed a by-product with an m/z of +44. It was attributed to the incomplete cleavage of the carbonyl group of the Boc protecting group on the side chain of Trp. Incubation of the highly diluted sample under acidic HPLC conditions

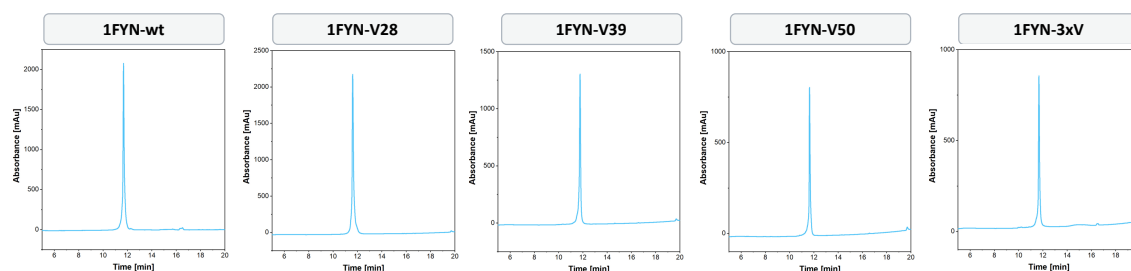
(ACN + H<sub>2</sub>O + 0.1% TFA) significantly improved the quality of the desired crude product in test reactions. In addition, the undesired by-product was also no longer detectable. With these optimizations, all the planned sequences could be successfully synthesized.

### 5.5.2.2 Purification by Preparative HPLC

Problems were experienced during purification by preparative HPLC, presumably caused by aggregation of all  $\beta$ -structures at the low pH of the mobile phase. This was particularly evident in that the samples remained in the stationary phase and could only be removed by extensive washing. With this procedure, **1FYN-wt** and all Val derivatives were purified by injecting 8 mg of peptide dissolved at pH 7 per HPLC-run (Table 5.13 + Figure 5.45). However, this process is ineffective and leads to losses in yield. Various efforts were investigated to improve the purification process.

**Table 5.13:** Synthesized and purified 1FYN derivatives.

peptide	ESI-ToF	calc. m/z	obs. m/z	purification	obtained purity [%]	yield [mg]
<b>1FYN-wt</b>	[M+6H] <sup>6+</sup>	1158.3918	1158.6792	10% → 80% B	> 96	15.6
	[M+7H] <sup>7+</sup>	993.6570	993.3288	over 18 min		
	[M+8H] <sup>8+</sup>	866.5759	867.0158	RT = 11.68 min		
<b>1FYN-V28</b>	[M+6H] <sup>6+</sup>	1156.0559	1156.3267	10% → 80% B	> 96	14.1
	[M+7H] <sup>7+</sup>	867.2939	867.8652	over 18 min		
	[M+8H] <sup>8+</sup>	771.0399	771.6123	RT = 11.60 min		
<b>1FYN-V39</b>	[M+6H] <sup>6+</sup>	1151.9001	1151.9076	10% → 80% B	> 96	9.3
	[M+7H] <sup>7+</sup>	864.1770	864.1841	over 18 min RT = 11.77 min		
<b>1FYN-V50</b>	[M+6H] <sup>6+</sup>	1156.0559	1156.3251	10% → 80% B	> 96	20.8
	[M+7H] <sup>7+</sup>	867.2939	867.8624	over 18 min RT = 11.64 min		
<b>1FYN-3xV</b>	[M+5H] <sup>5+</sup>	1389.8687	1390.0321	10% → 80% B	> 96	17.9
	[M+6H] <sup>6+</sup>	1158.3918	1158.7249	over 18 min		
	[M+7H] <sup>7+</sup>	993.0513	993.2167	RT = 11.67 min		

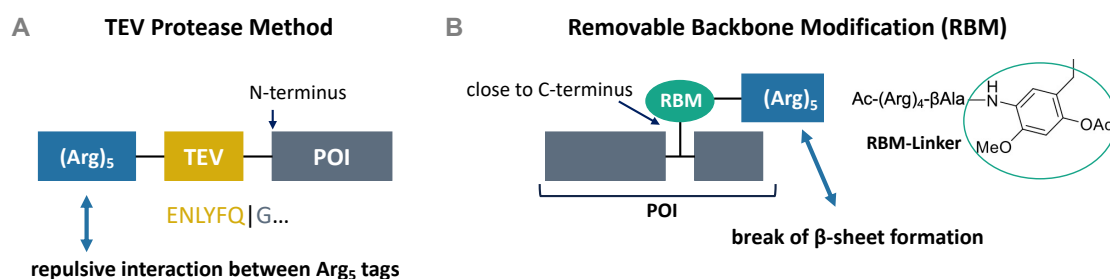


**Figure 5.45:** HPLC-Chromatograms of purified 1FYN derivatives. HPLC system: Chromaster®, Kinetex® C18 column, method 10-80% MeCN in H<sub>2</sub>O + 0.1% TFA in 18 min, DAD 220 nm. Chromatograms are depicted from 3 to 20 min.

The use of 10 mM  $\text{NH}_4\text{HCO}_3$  in the mobile phase of the HPLC resulted in a significant improvement in the purification due to the higher solubility, but the stationary phase was not stable under these conditions. Further studies involved an incorporation of a solubility tag.

### ***Incorporation of Solubility Tag to Improve Purification of 1FYN***

Two methods were investigated: the TEV protease strategy and the removable backbone modification (RBM) (Figure 5.46).

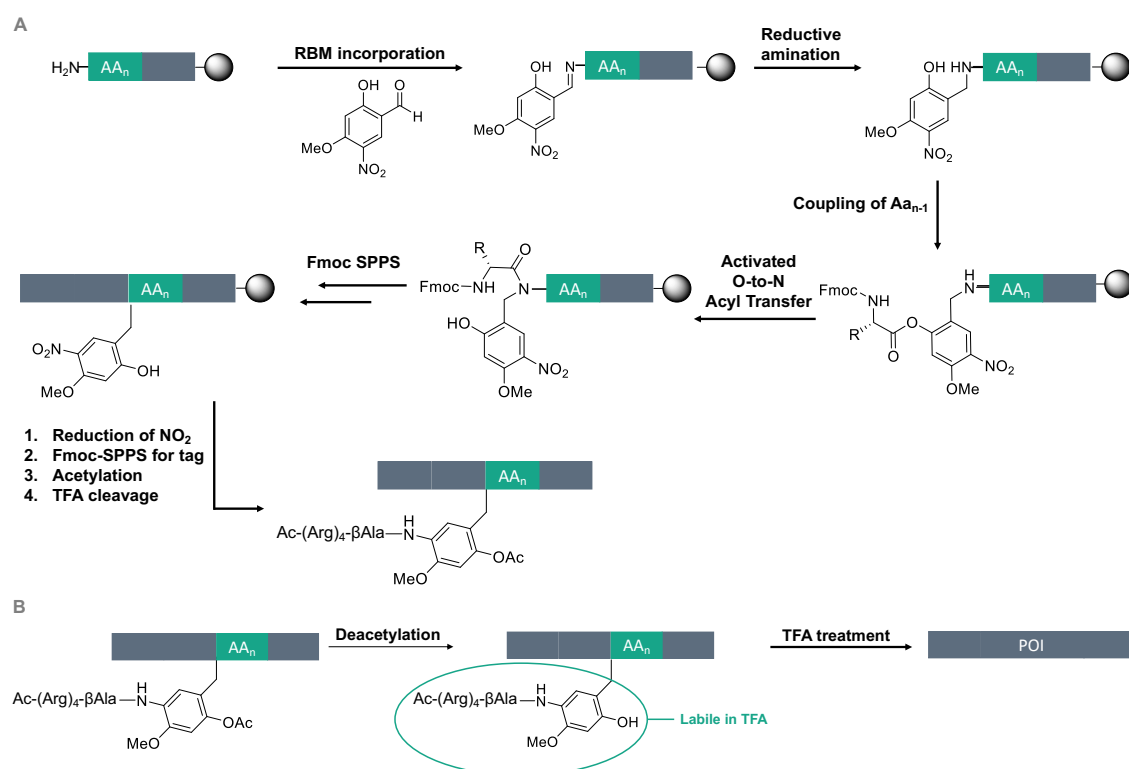


**Figure 5.46:** Strategies to introduce solubility tags into proteins: **A:** TEV protease Method; **B:** removable backbone modification (RBM). POI = protein of interest.

The first method is suitable for the application of N-terminal solubility tags cleaved by an endoprotease. This is done by introducing a binding sequence between the solubility tag and the protein of interest (POI). For the Tobacco Etch Virus (TEV) protease used in the present work, a Gly residue must be located at the N-terminus of the POI since this is the cleavage site of the protease; such is the case for 1FYN (Figure 5.46 left).<sup>[404]</sup> However, it is known in the literature that C-terminal solubility tags significantly improve solubility compared to N-terminal ones.<sup>[405]</sup> For this reason, a removable backbone modification (RBM method) was also investigated, in which a linker is introduced into the peptide backbone near the C-term for the solubility tag. After purification, the corresponding linker is removed along with the solubility tag. Zheng's second-generation linker, 4-methoxy-5-nitrosalicylaldehyde (RMB linker), was chosen for these studies (Figure 5.46 right). This linker is universally applicable by being introduced to any primary AA.<sup>[406,407]</sup> However, it is essential to note that there have been reports of low efficiency of insertion of the linker between sterically hindered AAs such as Val-Ile combinations.

After reductive amination, the subsequent AA can be coupled by incorporating an activated O-to-N-acyl transfer auxiliary (Figure 5.47).

Once the peptide is assembled on the solid support,  $\text{SnCl}_2$  is used to reduce the nitro group on the linker to an amino function. Next, a  $\beta$ -Ala is coupled as a spacer before the solubility tag is implemented at this point. The acetylation of the RBM linker is required before TFA cleavage, making it acid stable. Upon HPLC purification of the desired conjugate, the RBM linker is deacetylated with 0.4 M cysteine, making it acid-labile again. Subsequently, the TFA treatment removes the linker, and the POI is precipitated.

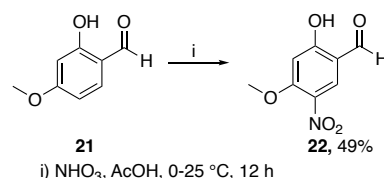


**Figure 5.47:** Removable Backbone Modification according to Zheng *et al.* **A)** Incorporation of the linker is followed by reductive amination and coupling of the next AA. After the O-to-N-acyl transfer, the SPPS continues. For the introduction of the solubility tag, the nitro group is reduced by  $\text{SnCl}_2$ . Before TFA treatment for removal of the assembled peptide, acetylation needs to be performed. **B)** After purification of the POI, a deacetylation is conducted, and the POI is precipitated by TFA treatment. Adapted from Zheng *et al.* [408]

### **RBM method:**

In the present work, this linker was introduced between Val58 and Ala59 of 1FYN. First, the RBM linker (**22**) was successfully prepared according to a literature-known procedure starting from 2-hydroxy-4-methoxy benzaldehyde (**21**) in a 49% yield (Scheme 5.25). [409]





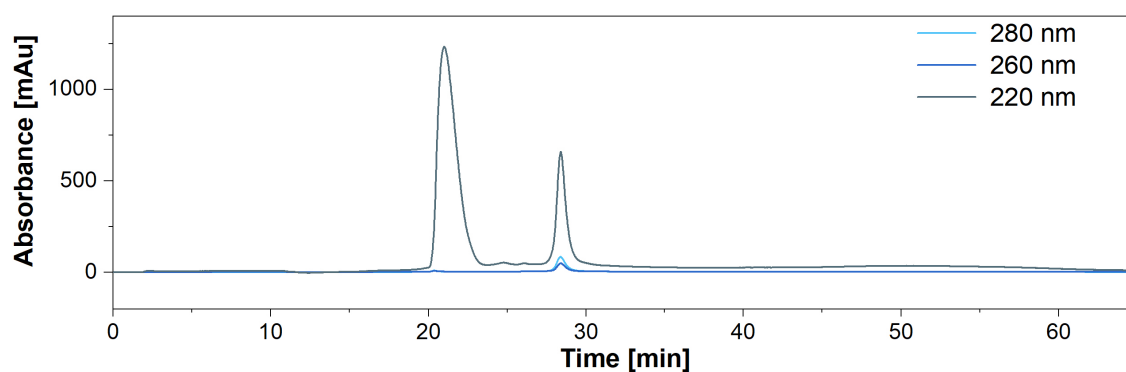
**Scheme 5.25:** Synthesis of RMB linker according to Tang *et al.* [409]

Micro-cleavages monitored each step of the incorporation (data is provided in Part C appendix). Here a capping with  $\text{Ac}_2\text{O}$  and DIPEA was performed before TFA cleavage. The reductive amination was extended to 10 min as the recommended 5 min treatment with 5 eq.  $\text{NaBH}_4$  in DMF at rt was proven to be insufficient. After the coupling of Val58 with HATU chemistry, a side-product was detected. This was assigned by ESI-ToF analysis to the ester formed by Fmoc-Val and the free OH-group of the linker. Additionally, micro-cleavages without prior capping also showed this by-product (with no esterification at the linker, the linker would be removed by TFA treatment). Therefore, the amount of Fmoc-Val was reduced to 2 eq., and mild coupling conditions (Oxyma + DIC at rt) were examined. However, the efforts could not suppress by-product formation entirely. Miranda *et al.* described in their publication that the ester is removed by treatment with piperidine, which was confirmed in our study. [410] Subsequently, the generated construct was subjected to MW-SPPS conditions. Here, no desired compound was determined. Therefore, this method needs further optimizations.

### ***TEV-protease method***

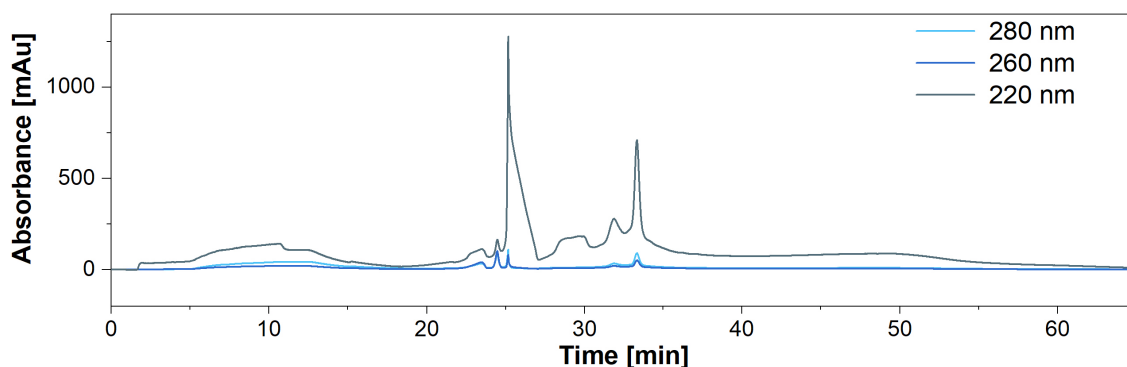
To introduce the solubility tag at the N-terminus of 1FYN, the recognition sequence ENLYFQ was first implemented using MW-SPPS; this was done with double couplings for 4 min each at 90°C. Double-couplings were also performed for the first two Arg during the synthesis of the poly-Arg section (solubility tag). Each of the last three Arg was built-in with a triple coupling. Due to the Pbf protecting group, Arg is already difficult to incorporate into a peptide backbone; therefore, such a concentrated Arg section requires additional couplings. The sequence was successfully extended with these adjustments, and the poly-Arg solubility tag was inserted. For the 1FYN-wt, improved solubility in acidic conditions for preparative HPLC was observed. After successful purification, treatment with TEV protease followed. Here the protein was dissolved in 10 mL reaction buffer (0.5 mM EDTA +

50 mM Tris-HCl + 1 mM DTT, pH: 7.5) 100  $\mu$ L of TEV-solution (7 mg/mL, kindly provided by M. Krummhaar (Roth group, Max Planck Institute of Colloids and Interfaces, Biomolecular Systems, Berlin), were added to the peptide solution. The reaction was carried out overnight at 4°C. Then ion-exchange chromatography (IEC) was performed to separate the cleaved solubility tag, TEV protease, and unreacted tagged protein. (Figure 5.48). As a result, two major peaks were observed (at 23.1 and 28.1 min). The peak at 23.1 min did not exhibit absorption in the aromatic domain of 260 or 260 nm. It was therefore assumed that this compound corresponds to the solubility tag with the TEV-recognition sequence. The second peak was evaluated by SDS-PAGE as well as with HPLC (highly diluted) and ESI-ToF analysis after performing desalting steps. Both analyses confirmed the desired **1FYN-wt**.



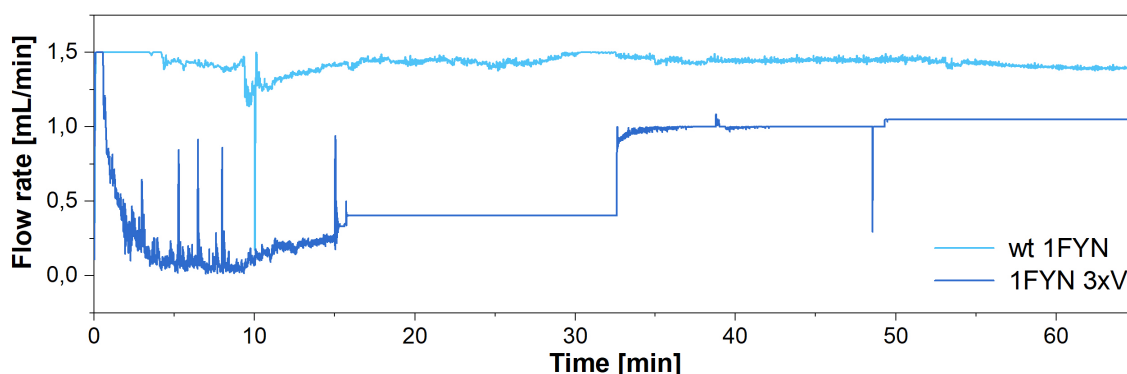
**Figure 5.48:** IEC chromatogram of TEV reaction with 1FYN-wt. Mono-Q<sup>TM</sup> 4.6/100 PE column was used, and a gradient of 0 to 70% B (20 mM Tris buffer) in A (H<sub>2</sub>O) in 25 column volumes was applied.

With this protocol in hand, the 1FYN Val28/39Val39/Val50 (**1FYN 3xV**) derivative was examined. Unfortunately, solubility issues remained in this sample despite the attached solubility tag, causing clogging of the preparative HPLC again. This challenge persisted also in the TEV reaction as here the pH value needed to be adjusted to 8.5 (instead of 7.5), which was still in the range of optimal conditions for the TEV reaction. Purification by IEC revealed several products, indicating an incomplete TEV reaction (Figure 5.49). This finding was confirmed by SDS-PAGE analysis.



**Figure 5.49:** IEC chromatogram of TEV reaction with 1FYN 3xV. Mono-QTM 4.6/100 PE column was used, and a gradient of 0 to 70% B (20 mM Tris buffer) in A (H<sub>2</sub>O) in 25 column volumes was applied.

Interestingly, clogging issues were also encountered also with IEC, as depicted in Figure 5.50. Here the flow rate is stabilized after the products are all eluted. These experiments clearly show that further attempts must be undertaken to develop a suitable purification protocol for the Val derivatives as well. The TfeGly peptides are still on the resin, as an effective purification process has yet to be established.



**Figure 5.50:** Comparison of flow rates during IEC purification. Mono-QTM 4.6/100 PE column was used.

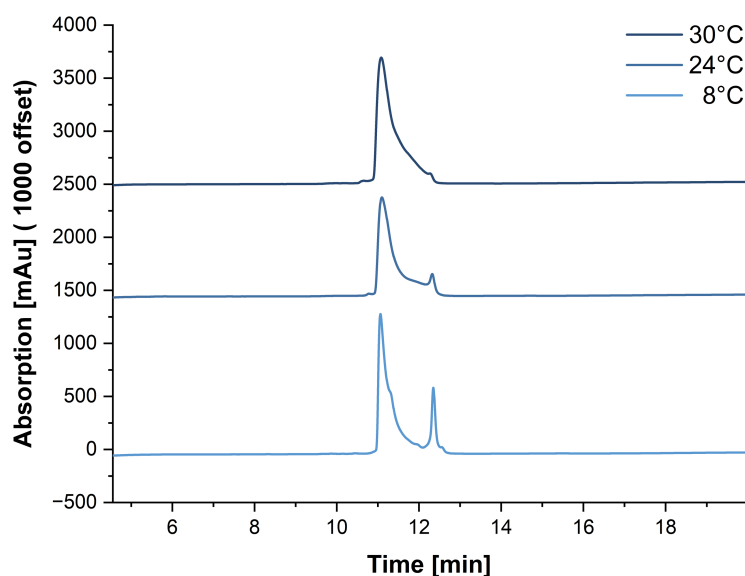
### 5.5.2.3 Pro-rich Peptides for X-tal studies

A majority of the published crystal structures of 1FYN contain different short Pro-rich peptides. These were also synthesized as possible additives to support folding of 1FYN. [270,411–413]

**Table 5.14:** Planned Pro-rich sequences.

name	sequence
<b>N55A</b>	Ac-APPIPPPRRKR-NH <sub>2</sub>
<b>APP12</b>	Ac-APPLPPRNRRL-NH <sub>2</sub>
<b>VSL12</b>	Ac-VSLARRPLPLP-NH <sub>2</sub>

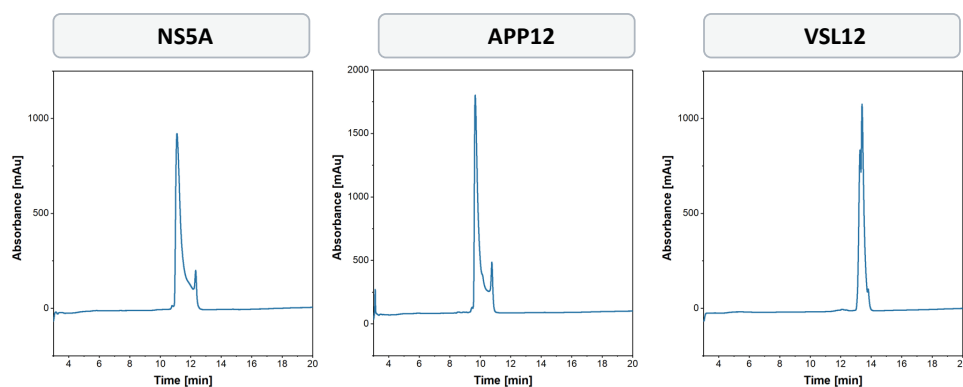
The synthesis was carried out on the Rink amid pro tide (0.19 mmol/g) resin using 2 min single couplings in a 0.1 mmol scale. Since the first cycle is always a "cold cycle", a two-fold Fmoc deprotection was performed at the beginning. For this purpose, 20% Pip was used. The N-term was acetylated on the MW synthesizer (Ac<sub>2</sub>O + DIPEA for 2 min at 60°C). During purification, a second peak was detected for all peptides. It was suspected that this was not an impurity as the ESI-ToF revealed no differences, but that there was an equilibrium between the cis/trans isomerism of the Pro is observed here. This was confirmed by the re-injection of the two separate peaks, as the same profile was revealed in the chromatogram despite the isolation of the samples. Furthermore, HPLC measurements at different temperatures could confirm a clear separation of the peaks, confirming the assumption of a kinetic dependency of product formation (Figure 5.51). Table 5.15 summarizes the obtained results of pro-rich peptides. The chromatograms are given in Figure 5.52.



**Figure 5.51:** Temperature-depended HPLC-measurements of NS5A peptide. HPLC system: Chromaster®, Kinetex® C18 column, method 5-50% MeCN in H<sub>2</sub>O + 0.1% TFA in 18 min, DAD 220 nm. The chromatogram is depicted from 6 to 20 min.

**Table 5.15:** Synthesized and purified Pro-rich peptides.

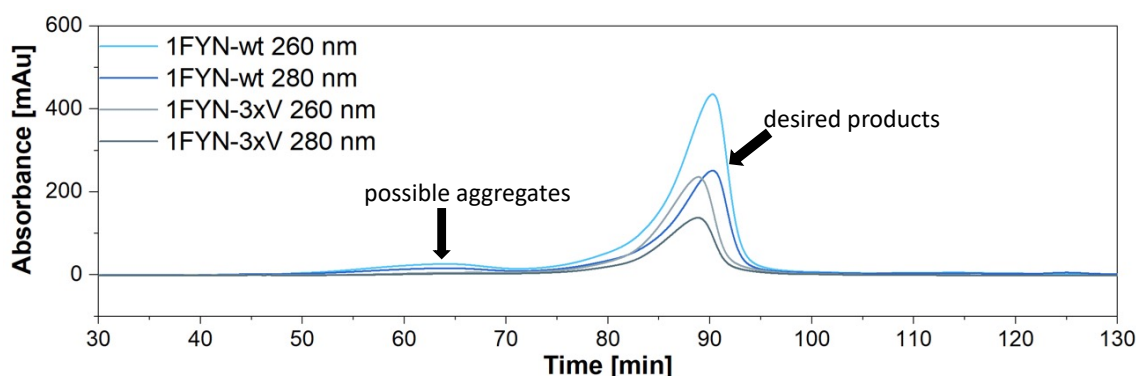
peptide	ESI-ToF	calc. m/z	obs. m/z	purification	yield [mg]
NS5A	[M+1H] <sup>1+</sup>	1325.8281	1325.8216	5% → 50% B over 18 min	15.8
	[M+2H] <sup>2+</sup>	663.4180	663.4171		
APP12	[M+1H] <sup>1+</sup>	1424.8602	1424.8542	5% → 50% B over 18 min	49.5
	[M+2H] <sup>2+</sup>	712.9340	712.9324		
VSL12	[M+1H] <sup>1+</sup>	1356.8479	1356.8428	5% → 50% B over 18 min	26.2
	[M+2H] <sup>2+</sup>	678.9278	678.9268		



**Figure 5.52:** Chromatograms of synthesized and purified Pro-rich peptides. HPLC system: Chromaster®, Kinetex® C18 column, method 5-50% MeCN in H<sub>2</sub>O + 0.1% TFA in 18 min, DAD 220 nm. Chromatograms are depicted from 3 to 20 min.

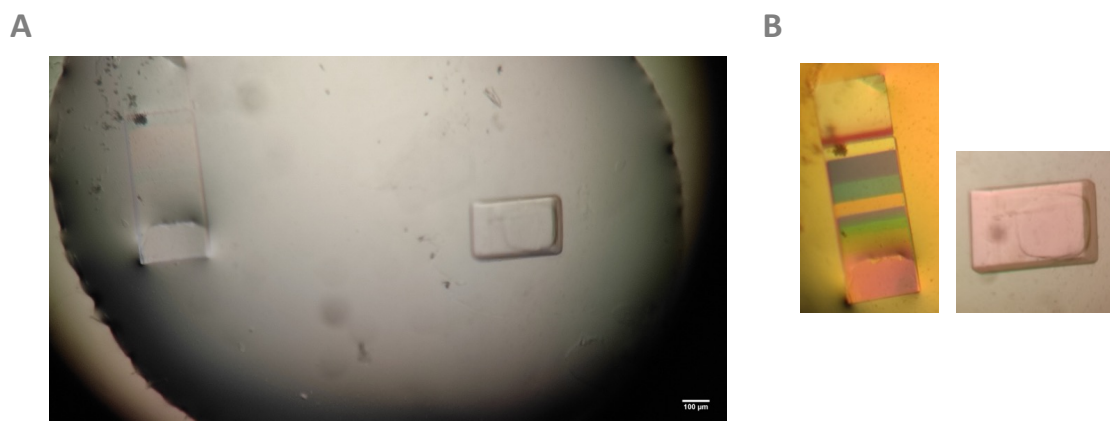
### 5.5.3 Structural Examination of 1FYN Derivatives

In collaboration with M. Krummhaar (Roth group, Max Planck Institute of Colloids and Interfaces, Biomolecular Systems, Berlin), experiments on the crystallization of 1FYN-wt and 1FYN-3xV were performed. Before crystallization, it was essential to conduct gel filtration (size exclusion chromatography, SEC), mainly because of 1FYN's tendency to aggregate. Here, one additional species was detected in both cases that eluted before (50-75 min) the main product (85-95 min). The species, however, exhibit only a low absorption indicating a low concentration. Nevertheless, the desired products were isolated in each case and concentrated by centrifugation.



**Figure 5.53:** Chromatogram of size exclusion chromatography (SEC) in 20 mM Tris buffer (pH 8.2).

Following the published procedure, crystals of 1FYN-wt were successfully obtained without Pro-rich peptides as additives (Figure 5.54 A),<sup>[270]</sup> and a peptidic origin was indicated by polarized filters (Figure 5.54 B). However, analysis by X-ray scattering showed poor diffraction, so further optimization approaches were conducted.



**Figure 5.54:** A) Generated crystals obtained by hanging drop method. reservoir: a solution containing 4M sodium formate (pH 8.3) and 0.6%  $\beta$ -mercaptoethanol) at 4 °C (4 mg/mL peptide conc.) B) Polarized filters hint at a peptidic origin of crystals.

Over 800 different conditions were scanned, but neither **1FYN-wt** nor **1FYN-3xVal** formed a stable crystal. In the next approaches, Pro-rich peptides might induce structural stability. This was investigated in CD studies. Unfortunately, these investigations did not succeed in forming defined  $\beta$ -structures in the given time of the presented doctoral thesis.

#### 5.5.4 Synthesis and Purification of 1TEN

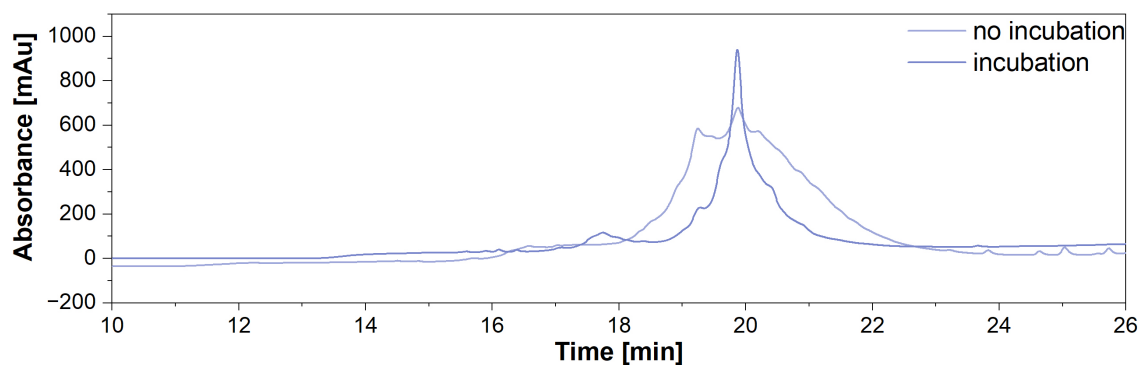
With a total length of 90 AAs, 1TEN is the largest protein in the present work. The positions Ile20, Tyr36, Ile59, and Val70, are essential AAs in the hydrophobic core of the protein domain. In the first step, the Ile residues are to be substituted by Val, as in the case of 1FYN. TfeGly subsequently replaces these positions. In the second series of experiments, the influence of the fluorine content on the folding kinetics can be studied by replacing TfeGly with DfeGly and MfeGly. Tyr36 is also interesting; here, different fluorinated AAs can be studied in this case.

**Table 5.16:** Planned sequences of 1TEN derivatives.

name	sequence
1TEN-wt	RLDAPSQIEVKDVTDTTALITWFKPLAEIDGIELTYGIKDVPGDRRTTIDLTEDE- NQYSIGNLKPDEYEVSLISRRGDMSSNPAKETFTT
1TEN-V20	RLDAPSQIEVKDVTDTTALVITWFKPLAEIDGIELTYGIKDVPGDRRTTIDLTEDE- NQYSIGNLKPDEYEVSLISRRGDMSSNPAKETFTT
1TEN-V59	RLDAPSQIEVKDVTDTTALITWFKPLAEIDGIELTYGIKDVPGDRRTTIDLTEDE- NQYSVGNLKPDEYEVSLISRRGDMSSNPAKETFTT
1TEN-TfeGly20	RLDAPSQIEVKDVTDTTALTfeGlyITWFKPLAEIDGIELTYGIKDVPGDRRTTIDLTEDE- NQYSIGNLKPDEYEVSLISRRGDMSSNPAKETFTT
1TEN-TfeGly59	RLDAPSQIEVKDVTDTTALITWFKPLAEIDGIELTYGIKDVPGDRRTTIDLTEDE- NQYSTfeGlyGNLKPDEYEVSLISRRGDMSSNPAKETFTT
1TEN-TfeGly70	RLDAPSQIEVKDVTDTTALITWFKPLAEIDGIELTYGIKDVPGDRRTTIDLTEDE- NQYSIGNLKPDEYETfeGlySLISRRGDMSSNPAKETFTT

The synthesis was performed using the previously elaborated strategies. For this, a Cl-MPA resin was used; the synthesis was divided into three sections: the 1-20st AA were introduced with 2 min single coupling, 21-60 with double 2 min coupling, and 60-90 in double 4 min. The reaction time of Fmoc cleavage was increased from 1 to 2 min. Because of the many Asp residues, 10% piperazine with 0.1% HOBt and the di-peptide building block were used. TfeGly was incorporated with 2 eq. by applying the 10 min 90°C cycle. The desired peptide was cleaved with cocktail I (TFA, TIS, H<sub>2</sub>O: 95/2.5/2.5) for four h at rt.

During purification, it was found that this large system requires a different column: a wide pore material for achieving effective purification (Phenomenex's C18 Jupiter column was used here.) Unlike the 1FYN, there were no difficulties with solubility under HPLC conditions. But the sample needs a more prolonged incubation with higher TFA content (overnight incubation in 10% ACN in water with 0.2% TFA instead of 0.1% TFA) before purification (Figure 5.55).



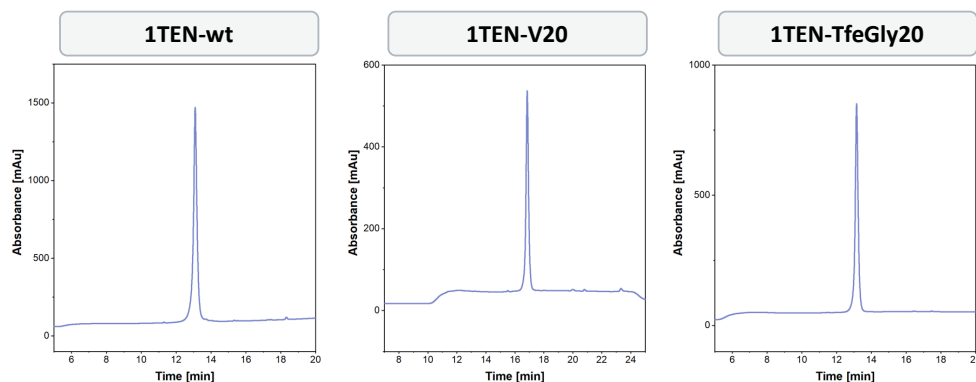
**Figure 5.55:** Impact of incubation prior to purification of 1TEN. HPLC system: Primaid®, Jupiter® C18 column (300Å), method 10-80% MeCN in H<sub>2</sub>O + 0.1% TFA in 20 min (+5 min pre-run), DAD 220 nm. The chromatogram is depicted from 10 to 26 min.

With this optimization, the following peptides have been synthesized and purified:

**Table 5.17:** Synthesized and purified 1TEN derivatives.

peptide	ESI-ToF	calc. m/z	obs. m/z	purification	obtained purity [%]	yield [mg]
<b>1TEN-wt</b>	[M+6H] <sup>6+</sup>	1257.5235	1257.4371	10% → 80% B	> 96	13
	[M+7H] <sup>7+</sup>	1117.9016	1117.8613	over 18 min		
	[M+8H] <sup>8+</sup>	1006.2203	1006.6284	RT = 13.15 min		
<b>1TEN-V20</b>	[M+10H] <sup>10+</sup>	1004.8177	1004.3443	10% → 80% B	> 96	10
	[M+11H] <sup>11+</sup>	913.5622	913.2085	over 18 min RT = 16.85min*		
<b>1TEN-TfeGly20</b>	[M+6H] <sup>6+</sup>	1680.3354	1680.9054	10% → 80% B	> 96	11
	[M+7H] <sup>7+</sup>	1440.4315	1440.9701	over 18 min		
	[M+8H] <sup>8+</sup>	1260.5053	1260.9408	RT = 13.15 min		

\*with 5 min pre-run (5% ACN in H<sub>2</sub>O + 0.1% TFA)

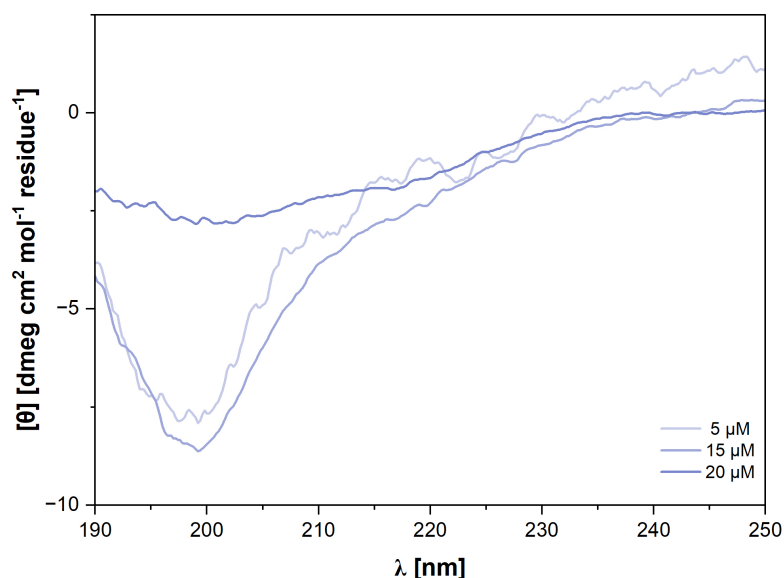


**Figure 5.56:** Analytical runs of final 1TEN products. HPLC system: Primaid®, Jupiter® C18 column (300Å), method 10-80% MeCN in H<sub>2</sub>O + 0.1% TFA in 18 min (+5 min pre-run for 1TEN-V20), DAD 220 nm. The chromatogram is depicted from 6 to 20 min.



### 5.5.5 Structural Examination of 1TEN Derivatives

Structure elucidation by CD was performed first with the final products to ensure that the synthesized samples have a natural  $\beta$ -structure. For this purpose, the native 1TEN wt was first investigated at different concentrations ranging from 5  $\mu\text{M}$  to 20  $\mu\text{M}$ . Different pH values were also analyzed (pH 4 as well as pH 5 in 100 mM acetate buffer and pH 6 in 100 mM phosphate buffer). As a result, however, a disordered structure was repeatedly obtained (Figure 5.57).



**Figure 5.57:** Results of CD studies of **1TEN-wt**. Shown are exemplary results of different concentrations at pH 5 (100 mM acetate buffer), as the most stable structures are reported at this pH.<sup>[414]</sup> The depicted spectra are normalized and represent the mean of three independent measurements.

In summary, exceptionally complex syntheses have been achieved in this project, and the incorporation of fIAAs has been established as a routine procedure. While handling the large systems, it became apparent that not only an individual synthesis protocol is required for each system, but also new purification challenges must be overcome. In the case of 1TEN, this has been realized very successfully. However, in the case of 1FYN, further optimizations must be undertaken, given the aggregation behavior.

Structural examinations of **1TEN** (CD studies) and **1FYN** (CD as well as crystallographic studies) suggest that more than a perfectly elaborated blueprint (primary structure) is needed to guarantee complete construction (folding). Here, additives or other buffer systems should be investigated to establish native folding.

## 6 Summary and Outlook

Fluorine has unique properties and abilities to interact with its environment. The present work deals with five different systems in which the influence of fluorine on the peptide and protein surroundings should be examined.

### 1. Project: Elucidation of Insertion Process of pHLIP

As a small model system, the pH (low) lipid insertion peptide (pHLIP) was investigated concerning its translocation behavior into a lipid bilayer using *in situ* SEIRAS. In this system, fluorinated AA should be incorporated to investigate the sensitive C-F vibration as a probe to study the insertion process. For this purpose, the region responsible for the translocation of pHLIP was to be determined. Based on these studies, fluorinated AA should then be incorporated. First, two non-fluorinated peptides, pHLIP1 (native sequence) and pHLIP2 (Asp23 replaced by His and N-term containing fewer Asp-residues), were synthesized. Standard conditions were used in MW-SPPS to perform the synthesis. Before purification, an overnight incubation in the acidic milieu is required. With *in situ* SEIRAS, we succeeded in restricting the regulatory domain to the Aps23 position. In addition, we succeeded in detecting two new folding intermediates. These results were published in *BBA Biomembranes* by Ataka *et al.* [364] In the second step, Val29 was substituted by TfeGly to establish the C-F vibration as a probe for the insertion process of peptides into the lipid membrane. This work laid the foundation for further studies of the folding behavior of pHLIP in a membrane environment and for the collaboration in the CRC-1349.

### 2. Project: FF03 as Carrier for EAE-inducing Peptides for <sup>19</sup>F-MRI Investigations

Using <sup>19</sup>F-MRI, *in vivo* studies were designed to clarify the distribution of two different EAE (experimental autoimmune encephalomyelitis, animal model of multiple sclerosis) peptides. For this purpose, the fiber-forming peptide FF03 was selected as a scaffold for the transport and multivalent presentation of EAE peptides (MOG<sub>35-55</sub> (myelin oligodendrocyte glycoprotein) as well as PLP<sub>139-151</sub> (proteolipid protein)) in animal models. In the first series of studies, the cellular uptake of the synthesized peptide conjugates was to be determined first; for this purpose, FITC-labeled constructs were to be assembled. In the second step, the hydrophobic core

of FF03 was to be furnished by fAAs to be used in *in vivo*  $^{19}\text{F}$ -MRI studies. Here, the fiber-forming properties of FF03 play a crucial role by allowing a very high local concentration of fluorine.

The synthesis and purification of both conjugates and unconjugated peptides as control samples were established by exploiting orthogonal protecting groups in an all-on solid-support approach. All synthesis steps were established using MW-SPPS, which had to be paused in intervals, as the selected TGA resin did not seem heat stable. For the second study section with fluorinated AA, it should be considered to choose a different resin as the TGA resin would probably not withstand the coupling conditions of 10 min at 90°C for fAAs. The MW-SPPS typical pro tide Cl resins are not compatible with the protective group strategy due to their high acid stability. In this respect, low-loaded Wang resins may offer a good alternative. Studies on cellular uptake of the FITC-labeled peptide-conjugates are ongoing in collaboration with S. Waiczies (AG Niendorf, Experimental Ultrahigh field MR at Berlin Institute for Medical Systems Biology of the Max Delbrück Center).

### *3. Project: Examining the Impact of Possible F-F Interactions on Mechanical Stability*

The heterodimeric coiled-coil system A4/B4 was chosen to study the influence of possible F-F interactions on the mechanical stability of coiled coils. Here, a TfeGly was substituted in place of a Leu in the *a* or *d* position of the second heptad, and a Cys linker extended the sequences at either the N- or C-term so that they could be immobilized for the measurements. The results determined that the incorporation of TfeGly in one strand of the coiled coil already leads to a destabilization. The cause here is thought to be unfavorable intermolecular multipolar interactions of the fluorinated side chain with the side chain of the Asn in the third heptad so that the intermolecular H-bond of the Asn-Asn interaction is compromised. On the other hand, unfavorable packing effects of TfeGly and its lower  $\alpha$ -helical propensity might also lead to destabilization. For further studies on mechanical stability, it would be recommended to investigate the position of fAAs in systems that do not exhibit other polar interactions within the hydrophobic core. One example would be our group's VPE/VPK model; this system has already been investigated concerning the influence of fAAs in different regions of the hydrophobic core by thermal stability.

Most of the work is concerned with studying possible F-F interactions as an orthogonal tool to control protein folding kinetics. The selected systems included the small protein ubiquitin (76 AAs) and two small all- $\beta$  protein domains, 1FYN (62 AAs) and the 1TEN (90 AAs), which pose a significant challenge for chemical protein synthesis. With the help of MW-SPPS, we have succeeded in synthesizing all the proposed native sequences with tailored synthesis protocols and established a protocol for efficiently incorporating the valuable fluorinated AA with low equivalents (1.5 to 2 eq.).

#### *4. Project: Influence of Fluorine-specific Interactions on Protein Folding by Studying Ubiquitin Dimerization*

Ubiquitin (Ub) was to be systematically substituted at its solvent-exposed hydrophobic domains, the so-called hydrophobic patches consisting of Leu8/Ile44/Val70 with fAAs that have similar hydrophobic character and vdW volume. The influence of a fully fluorinated patch (all three AAs substituted) and a non-fully fluorinated patch (substitution of a single position) were also to be investigated. The influence on protein folding kinetics was to be determined via product formation from the ligation of two ubiquitin monomers monitored by HPLC. For this, two approaches were pursued, both native chemical ligation (leading to an Ala mutation at the ligation site) and linkage to the native Ub-dimer via isopeptide binding (ICL). For both ligations, preparing a thioester (Ub2) and a thiol (Ub1) monomer was essential.

The present work explored C-terminal modifications for precursor functionalities that can be converted into the desired thioester after synthesis and purification: the *N*-Me-Cys modification for the ICL approach, and Dbz-linkers, hydrazides, as well as SEA-group for the NCL strategy. In total, we succeeded in establishing the synthesis of the desired Ub2-*N*-Me-Cys modified monomer for the ICL approach and Ub2-NHNH<sub>2</sub>, as well as Ub2-SEA functionalization for the NCL approach. In addition, we succeeded in preparing the thiol Ub1(K48-Cys) monomer required for NCL.

In the case of using Ub2-NHNH<sub>2</sub> for the NCL, the conversion to the desired thioester needed to be established first. Encouragingly, the conversion of the hydrazide to the thioester via the corresponding azide proceeded well. During the ligation

examination, however, no dimerization was observed as only the hydrolyzed Ub2 thioester (Ub2-COOH) was determined.

In further procedures, the evaluation by HPLC became impractical due to precipitations and poor solubility of all synthesized Ub monomers. Treatment with heat or ultrasonication, or desalting did not lead to any improvement. Therefore, on the one hand, SDS-PAGE was explored as a possible analysis method, and an extraction procedure with Et<sub>2</sub>O was also performed to remove the excess of MPAA. Examination of SEA-ligation with SDS-PAGE revealed promising results on Ub dimer formation. However, these results also suggest that further optimizations must be undertaken, as, after 24 h, only an incomplete conversion to the Ub dimer could be determined. The main challenge remains the poor solubility of the Ub monomers. Here, other detergents, as well as buffer systems, should be investigated. In addition, the synthesis of the Ub monomers should be optimized on a larger scale so that more starting product is available (this could facilitate handling the experimental procedure since larger volumes than 100  $\mu$ L could be used.). To prevent thioester hydrolysis in the hydrazide strategy, it would be possible to switch to MesNa; this thioester is less reactive, which could lead to slower ligation times.

##### *5. Project: Investigations of Fluorine-specific Interactions on Protein Folding Kinetics of Small All- $\beta$ Protein Domains*

In the context of this thesis, the data on the influence of fAAs on protein folding kinetics employing stopped-flow Trp fluorescence should be expanded since only little data regarding all- $\beta$  proteins exists in the literature, according to our knowledge. For this purpose, 1FYN, and 1TEN, both small all- $\beta$  protein domains without disulfide bonds and folding in a two-step manner, were selected. The fAAs were to be substituted in positions that are crucial in the hydrophobic core. Ile was first to be exchanged by Val to generate a non-fluorinated counterpart to TfeGly-substitutions.

In the first step, the synthesis of the two models was worked out based on the work on the Ub project. However, difficulties arose during purification. For 1TEN, it was found that this protein requires incubation with higher levels of TFA before purification. In addition, 300 Å column material must be used during purification, because a 100 Å material leads to poor separation results.

1FYN cannot be dissolved in the acidic milieu; this always causes it to stick to the stationary material. Preparative HPLC purification at higher pH values (pH = 8) using  $\text{NH}_4\text{HCO}_3$  as a modifier instead of TFA led to the decomposition of the column material. Therefore, two different methods for introducing a solubility tag were investigated. One was the removable backbone modification (RBM), according to Zheng *et al.*, and the other was the extension of the sequence by the solubility tag at the N-term in combination with a recognition sequence of the TEV protease for the removal of the tag after HPLC purification. The TEV method proved to be successful with the 1FYN-wt. Applying this method to the 3xVal mutant, however, encountered solubility problems even at pH 8, which also resulted in low flow rates when the TEV reaction was purified by ion-exchange chromatography and indicated clogging of this material.

Synthetically, all the desired fluorinated 1FYN variants were generated, but their purification was not successful by the submission date of the present work. To enable the purification, it would be possible in the next step to consider the purification at elevated temperatures or the addition of detergents. These methods are often used in synthesizing and purifying membrane peptides, which are also very difficult to purify. Likewise, different backbone modifications could be explored that, unlike the TEV method, can interrupt the structure in the middle rather than at the N-term and thus prevent aggregation. It would be necessary here that this modification is stable for MW-based SPPS. RBM was not a success in the present work.

Nonetheless, structural elucidation studies using CD were performed with the obtained material at 1FYN wt and all Val mutants. It was found that a  $\beta$ -structure could neither be detected at different concentrations and pH values nor by adding the Pro-rich peptides as additives.

Crystallization experiments did yield crystals suggesting a peptide origin. However, it was not possible to obtain an accurate structure elucidation due to poor diffraction.

In the case of 1TEN, the wt, Va20, and TfeGly20 mutants were examined by CD. Again, the spectra did not show defined  $\beta$  structures. Additive scans (e.g., by Hampton research) can serve as optimization tools for both 1FYN and 1TEN. In

these experiments, a wide variety of compositions are investigated, which can lead to a stabilization of the protein structure.

## 7 References

- [1] K. T. Koga, E. F. Rose-Koga, in *Encyclopedia of Geochemistry. Encyclopedia of Earth Sciences Series* (Ed.: William M. White), Springer, Cham, **2018**, pp. 495–498.
- [2] N. Budisa, V. Kubyshkin, D. Schulze-Makuch, *Life* **2014**, *4*, 374–385.
- [3] X. Cheng, L. Ma, *Appl Microbiol Biotechnol* **2021**, *105*, 8033–8058.
- [4] D. O'Hagan, H. Deng, *Chem Rev* **2015**, *115*, 634–649.
- [5] N. A. Meanwell, *J Med Chem* **2018**, *61*, 5822–5880.
- [6] E. Hammel, T. F. Webster, R. Gurney, W. Heiger-Bernays, *iScience* **2022**, *25*, 104020.
- [7] M. Inoue, Y. Sumii, N. Shibata, *ACS Omega* **2020**, *5*, 10633–10640.
- [8] T. Fujiwara, D. O'Hagan, *J Fluor Chem* **2014**, *167*, 16–29.
- [9] Y. Ogawa, E. Tokunaga, O. Kobayashi, K. Hirai, N. Shibata, *iScience* **2020**, *23*, 101467.
- [10] Y. Zhou, J. Wang, Z. Gu, S. Wang, W. Zhu, J. L. Aceña, V. A. Soloshonok, K. Izawa, H. Liu, *Chem Rev* **2016**, *116*, 422–518.
- [11] J. Wang, M. Sánchez-Roselló, J. L. Aceña, C. del Pozo, A. E. Sorochinsky, S. Fustero, V. A. Soloshonok, H. Liu, *Chem Rev* **2014**, *114*, 2432–2506.
- [12] S. Purser, P. R. Moore, S. Swallow, V. Gouverneur, *Chem. Soc. Rev.* **2008**, *37*, 320–330.
- [13] P. Politzer, *J Am Chem Soc* **1969**, *91*, 6235–6237.
- [14] A. Bondi, *J Phys Chem* **1964**, *68*, 441–451.
- [15] B. E. Smart, in *Organofluorine Chemistry*, Springer US, Boston, MA, **1994**, pp. 57–88.
- [16] J. C. Knight, P. G. Edwards, S. J. Paisey, *RSC Adv* **2011**, *1*, 1415.
- [17] R. K. Harris, E. D. Becker, S. M. Cabral de Menezes, R. Goodfellow, P. Granger, *Solid State Nucl Magn Reson* **2002**, *22*, 458–483.
- [18] I. Tirotta, V. Dichiarante, C. Pigliacelli, G. Cavallo, G. Terraneo, F. B. Bombelli, P. Metrangolo, G. Resnati, *Chem Rev* **2015**, *115*, 1106–1129.
- [19] M. Srinivas, A. Heerschap, E. T. Ahrens, C. G. Figdor, I. J. M. de Vries, *Trends Biotechnol* **2010**, *28*, 363–370.



- [20] D. Gimenez, A. Phelan, C. D. Murphy, S. L. Cobb, *Beilstein Journal of Organic Chemistry* **2021**, *17*, 293–318.
- [21] H. Chen, S. Viel, F. Ziarelli, L. Peng, *Chem Soc Rev* **2013**, *42*, 7971.
- [22] M. M. Alauddin, *Am J Nucl Med Mol Imaging* **2012**, *2*, 55–76.
- [23] L. Hunter, *Beilstein Journal of Organic Chemistry* **2010**, *6*, DOI 10.3762/bjoc.6.38.
- [24] D. O'Hagan, *Chem. Soc. Rev.* **2008**, *37*, 308–319.
- [25] F. Leroux, *ChemBioChem* **2004**, *5*, 644–649.
- [26] Y. Tang, G. Ghirlanda, W. A. Petka, T. Nakajima, W. F. DeGrado, D. A. Tirrell, *Angewandte Chemie International Edition* **2001**, *40*, 1494–1496.
- [27] V. P. Kukhar, *Fluorine-Containing Amino Acids – Synthesis & Properties*, John Wiley & Sons Inc, Hoboken (NJ, USA), **1995**.
- [28] M. Schlosser, D. Michel, *Tetrahedron* **1996**, *52*, 99–108.
- [29] T. Nagai, G. Nishioka, M. Koyama, A. Ando, T. Miki, I. Kumadaki, *J Fluor Chem* **1992**, *57*, 229–237.
- [30] G. Bott, L. D. Field, S. Sternhell, *J Am Chem Soc* **1980**, *102*, 5618–5626.
- [31] C. Wolf, W. A. König, C. Roussel, *Liebigs Annalen* **1995**, *1995*, 781–786.
- [32] M. Zanda, *New Journal of Chemistry* **2004**, *28*, 1401.
- [33] M. Molteni, C. Pesenti, M. Sani, A. Volonterio, M. Zanda, *J Fluor Chem* **2004**, *125*, 1735–1743.
- [34] J. S. Murray, P. G. Seybold, P. Politzer, *J Chem Thermodyn* **2021**, *156*, 106382.
- [35] W. K. Hagmann, *J Med Chem* **2008**, *51*, 4359–4369.
- [36] M. Schlosser, *Angewandte Chemie International Edition* **1998**, *37*, 1496–1513.
- [37] M. H. Abraham, P. L. Grellier, D. V. Prior, P. P. Duce, J. J. Morris, P. J. Taylor, *Journal of the Chemical Society, Perkin Transactions 2* **1989**, 699.
- [38] M. Morgenthaler, E. Schweizer, A. Hoffmann-Röder, F. Benini, R. E. Martin, G. Jaeschke, B. Wagner, H. Fischer, S. Bendels, D. Zimmerli, J. Schneider, F. Diederich, M. Kansy, K. Müller, *ChemMedChem* **2007**, *2*, 1100–1115.
- [39] R. Paulini, K. Müller, F. Diederich, *Angewandte Chemie International Edition* **2005**, *44*, 1788–1805.
- [40] C. Dalvit, C. Invernizzi, A. Vulpetti, *Chemistry - A European Journal* **2014**, *20*, 11058–11068.
- [41] C. Dalvit, A. Vulpetti, *Chemistry – A European Journal* **2016**, *22*, 7592–7601.

- [42] B. Biswas, P. C. Singh, *J Fluor Chem* **2020**, *235*, 109414.
- [43] I. Hyla-Kryspin, G. Haufe, S. Grimme, *Chemistry - A European Journal* **2004**, *10*, 3411–3422.
- [44] L. Shimoni, J. P. Glusker, *Struct Chem* **1994**, *5*, 383–397.
- [45] A. Vulpetti, C. Dalvit, *Chemistry – A European Journal* **2021**, *27*, 8764–8773.
- [46] H.-J. Schneider, *Chem Sci* **2012**, *3*, 1381.
- [47] G. T. Giuffredi, V. Gouverneur, B. Bernet, *Angewandte Chemie International Edition* **2013**, *52*, 10524–10528.
- [48] F. R. Fischer, W. B. Schweizer, F. Diederich, *Angewandte Chemie International Edition* **2007**, *46*, 8270–8273.
- [49] A. J. Peloquin, D. A. Kure, A. R. Jennings, C. D. McMillen, S. T. Iacono, W. T. Pennington, *Cryst Growth Des* **2020**, *20*, 5484–5492.
- [50] C. Li, Y. Cao, B. Hu, Y. Li, *J Phys Org Chem* **2021**, *34*, DOI 10.1002/poc.4151.
- [51] R. Taylor, *Acta Crystallogr B Struct Sci Cryst Eng Mater* **2017**, *73*, 474–488.
- [52] J. D. Dunitz, *ChemBioChem* **2004**, *5*, 614–621.
- [53] W. Pietruś, R. Kafel, A. J. Bojarski, R. Kurczab, *Molecules* **2022**, *27*, 1005.
- [54] K. Müller, C. Faeh, F. Diederich, *Science (1979)* **2007**, *317*, 1881–1886.
- [55] G. Zhang, W. He, D. Chen, *Mol Phys* **2014**, *112*, 1736–1744.
- [56] J. A. Gladysz, D. P. Curran, *Tetrahedron* **2002**, *58*, 3823–3825.
- [57] A. P. Dobbs, M. R. Kimberley, *J Fluor Chem* **2002**, *118*, 3–17.
- [58] J.-M. Vincent, in (Ed.: I.T. Horváth), Springer, Berlin, Heidelberg, **2011**, pp. 153–174.
- [59] C. Cai, W.-B. Yi, W. Zhang, M.-G. Shen, M. Hong, L.-Y. Zeng, *Mol Divers* **2009**, *13*, 209–239.
- [60] M. Wende, J. A. Gladysz, *J Am Chem Soc* **2003**, *125*, 5861–5872.
- [61] G. Pozzi, *Coord Chem Rev* **2003**, *242*, 115–124.
- [62] M. Wende, R. Meier, J. A. Gladysz, *J Am Chem Soc* **2001**, *123*, 11490–11491.
- [63] W. Zhang, *Chem Rev* **2004**, *104*, 2531–2556.
- [64] W. Zhang, *Chem Rev* **2009**, *109*, 749–795.
- [65] W. Zhang, D. P. Curran, *Tetrahedron* **2006**, *62*, 11837–11865.
- [66] V. Montanari, K. Kumar, *J Am Chem Soc* **2004**, *126*, 9528–9529.
- [67] T. Tono, A. Nishikawa, T. Yajima, H. Nagano, K. Mikami, *European J Org Chem* **2008**, *2008*, 1331–1335.

- [68] A. J. Vegas, A. N. Koehler, in (Eds.: M. Uttamchandani, S. Yao), Humana Press, Totowa, NJ, **2010**, pp. 43–55.
- [69] B.-Y. Li, D. S. Juang, A. K. Adak, K.-C. Hwang, C.-C. Lin, *Sci Rep* **2017**, *7*, 7053.
- [70] N. Pohl, *Angewandte Chemie International Edition* **2008**, *47*, 3868–3870.
- [71] B. Y. M. Collet, T. Nagashima, M. S. Yu, N. L. B. Pohl, *J Fluor Chem* **2009**, *130*, 1042–1048.
- [72] A. J. Vegas, J. E. Bradner, W. Tang, O. M. McPherson, E. F. Greenberg, A. N. Koehler, S. L. Schreiber, *Angewandte Chemie International Edition* **2007**, *46*, 7960–7964.
- [73] R. L. Nicholson, M. L. Ladlow, D. R. Spring, *Chemical Communications* **2007**, 3906.
- [74] F. A. Jaipuri, B. Y. M. Collet, N. L. Pohl, *Angewandte Chemie International Edition* **2008**, *47*, 1707–1710.
- [75] G.-S. Chen, N. L. Pohl, *Org Lett* **2008**, *10*, 785–788.
- [76] S. K. Mamidyala, K.-S. Ko, F. A. Jaipuri, G. Park, N. L. Pohl, *J Fluor Chem* **2006**, *127*, 571–579.
- [77] S. Maza, J. L. de Paz, P. M. Nieto, *Molecules* **2019**, *24*, 1591.
- [78] K.-S. Ko, F. A. Jaipuri, N. L. Pohl, *J Am Chem Soc* **2005**, *127*, 13162–13163.
- [79] J.-M. Vincent, *Chemical Communications* **2012**, *48*, 11382.
- [80] S. N. Chavan, A. K. Padhan, D. Mandal, *Polym Chem* **2018**, *9*, 2258–2270.
- [81] A. K. Singh, B. Schade, M. Rosati, R. Rashmi, V. Dichiarante, G. Cavallo, P. Metrangolo, R. Haag, *Macromol Biosci* **2022**, 2200108.
- [82] D. Verma, Rashmi, D. Rathore, K. Achazi, B. Schade, R. Haag, S. K. Sharma, *ACS Appl Polym Mater* **2022**, DOI 10.1021/acsapm.2c01176.
- [83] Z. Zhang, W. Shen, J. Ling, Y. Yan, J. Hu, Y. Cheng, *Nat Commun* **2018**, *9*, 1377.
- [84] H. Omorodion, B. Twamley, J. A. Platts, R. J. Baker, *Cryst Growth Des* **2015**, *15*, 2835–2841.
- [85] J.-M. Vincent, *J Fluor Chem* **2008**, *129*, 903–909.
- [86] J. Lv, H. Wang, G. Rong, Y. Cheng, *Acc Chem Res* **2022**, *55*, 722–733.
- [87] M. Skiba, M. Skiba-Lahiani, P. Arnaud, *J Incl Phenom Macrocycl Chem* **2002**, *44*, 151–154.
- [88] F. T. T. Huque, K. Jones, R. A. Saunders, J. A. Platts, *J Fluor Chem* **2002**, *115*, 119–128.

- [89] M. Jbeily, J. Kressler, *J Fluor Chem* **2017**, *193*, 67–72.
- [90] H. Biava, N. Budisa, *Eng Life Sci* **2014**, *14*, 340–351.
- [91] M. P. Krafft, J. G. Riess, *Chemosphere* **2015**, *129*, 4–19.
- [92] J. G. Riess, *Artificial Cells, Blood Substitutes, and Biotechnology* **2005**, *33*, 47–63.
- [93] M. A. Miller, E. M. Sletten, *ChemBioChem* **2020**, *21*, 3451–3462.
- [94] M. Cametti, B. Crousse, P. Metrangolo, R. Milani, G. Resnati, *Chem. Soc. Rev.* **2012**, *41*, 31–42.
- [95] S. Kawai, A. Sadeghi, F. Xu, L. Peng, A. Orita, J. Otera, S. Goedecker, E. Meyer, *ACS Nano* **2015**, *9*, 2574–2583.
- [96] P. Varadwaj, A. Varadwaj, H. Marques, K. Yamashita, *Computation* **2018**, *6*, 51.
- [97] C. F. Matta, N. Castillo, R. J. Boyd, *J Phys Chem A* **2005**, *109*, 3669–3681.
- [98] A. Varadwaj, P. R. Varadwaj, H. M. Marques, K. Yamashita, *ChemPhysChem* **2018**, *19*, 1486–1499.
- [99] C. Esterhuysen, A. Heßelmann, T. Clark, *ChemPhysChem* **2017**, *18*, 772–784.
- [100] G. V. Janjić, S. T. Jelić, N. P. Trišović, D. M. Popović, I. S. Đorđević, M. K. Milčić, *Cryst Growth Des* **2020**, *20*, 2943–2951.
- [101] M. Yano, T. Taketsugu, K. Hori, H. Okamoto, S. Takenaka, *Chemistry - A European Journal* **2004**, *10*, 3991–3999.
- [102] A. Schwarzer, W. Seichter, E. Weber, H. Stoeckli-Evans, M. Losada, J. Hulliger, *CrystEngComm* **2004**, *6*, 567.
- [103] G. Asensio, M. Medio-Simon, P. Alemán, C. R. de Arellano, *Cryst Growth Des* **2006**, *6*, 2769–2778.
- [104] A. Schwarzer, P. Bombicz, E. Weber, *J Fluor Chem* **2010**, *131*, 345–356.
- [105] E. N. G. Marsh, *Acc Chem Res* **2014**, *47*, 2878–2886.
- [106] J. N. Sloand, M. A. Miller, S. H. Medina, *Peptide Science* **2021**, *113*, DOI 10.1002/pep2.24184.
- [107] C. A. Hunter, J. K. M. Sanders, *J Am Chem Soc* **1990**, *112*, 5525–5534.
- [108] J. M. Monkovic, H. Gibson, J. W. Sun, J. K. Montclare, *Pharmaceuticals* **2022**, *15*, 1201.
- [109] M. L. Waters, *Curr Opin Chem Biol* **2002**, *6*, 736–741.
- [110] M. A. Miller, E. M. Sletten, *ChemBioChem* **2020**, *21*, 3451–3462.

- [111] M. Salwiczek, E. K. Nyakatura, U. I. M. Gerling, S. Ye, B. Kokscha, *Chem. Soc. Rev.* **2012**, *41*, 2135–2171.
- [112] A. A. Berger, J.-S. Völler, N. Budisa, B. Kokscha, *Acc Chem Res* **2017**, *50*, 2093–2103.
- [113] B. C. Buer, E. N. G. Marsh, *Protein Science* **2012**, *21*, 453–462.
- [114] E. Devillers, J. Pytkowicz, E. Chelain, T. Brigaud, *Amino Acids* **2016**, *48*, 1457–1468.
- [115] J. Moschner, V. Stulberg, R. Fernandes, S. Huhmann, J. Leppkes, B. Kokscha, *Chem Rev* **2019**, *119*, 10718–10801.
- [116] L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science (1979)* **2001**, *292*, 498–500.
- [117] P. Wang, Y. Tang, D. A. Tirrell, *J Am Chem Soc* **2003**, *125*, 6900–6906.
- [118] S. Son, I. C. Tanrikulu, D. A. Tirrell, *ChemBioChem* **2006**, *7*, 1251–1257.
- [119] P. Wang, A. Fichera, K. Kumar, D. A. Tirrell, *Angewandte Chemie International Edition* **2004**, *43*, 3664–3666.
- [120] N. Budisa, W. Wenger, B. Wiltschi, *Mol Biosyst* **2010**, *6*, 1630.
- [121] T. Panchenko, W. W. Zhu, J. K. Montclare, *Biotechnol Bioeng* **2006**, *94*, 921–930.
- [122] N. Budisa, *Angewandte Chemie International Edition* **2004**, *43*, 6426–6463.
- [123] S. A. Fraser, C. J. Easton, *Aust J Chem* **2015**, *68*, 9.
- [124] L. Merkel, N. Budisa, *Org Biomol Chem* **2012**, *10*, 7241.
- [125] F. Tobola, M. Lepšík, S. R. Zia, H. Leffler, U. J. Nilsson, O. Blixt, A. Imberty, B. Wiltschi, *ChemBioChem* **2022**, *23*, DOI 10.1002/cbic.202100593.
- [126] H. Qianzhu, E. H. Abdelkader, I. D. Herath, G. Otting, T. Huber, *ACS Sens* **2022**, *7*, 44–49.
- [127] N. C. Yoder, K. Kumar, *Chem Soc Rev* **2002**, *31*, 335–341.
- [128] F. Agostini, L. Sinn, D. Petras, C. J. Schipp, V. Kubyshkin, A. A. Berger, P. C. Dorrestein, J. Rappsilber, N. Budisa, B. Kokscha, *ACS Cent Sci* **2021**, *7*, 81–92.
- [129] T. Didenko, J. J. Liu, R. Horst, R. C. Stevens, K. Wüthrich, *Curr Opin Struct Biol* **2013**, *23*, 740–747.
- [130] R. Horst, J. J. Liu, R. C. Stevens, K. Wüthrich, *Angewandte Chemie International Edition* **2013**, *52*, 10762–10765.

- [131] J. J. Liu, R. Horst, V. Katritch, R. C. Stevens, K. Wüthrich, *Science (1979)* **2012**, 335, 1106–1110.
- [132] J. Klein-Seetharaman, E. V. Getmanova, M. C. Loewen, P. J. Reeves, H. G. Khorana, *Proceedings of the National Academy of Sciences* **1999**, 96, 13744–13749.
- [133] M. C. Loewen, J. Klein-Seetharaman, E. V. Getmanova, P. J. Reeves, H. Schwalbe, H. G. Khorana, *Proceedings of the National Academy of Sciences* **2001**, 98, 4888–4892.
- [134] K. Y. Chung, T. H. Kim, A. Manglik, R. Alvares, B. K. Kobilka, R. S. Prosser, *Journal of Biological Chemistry* **2012**, 287, 36305–36311.
- [135] T. H. Kim, K. Y. Chung, A. Manglik, A. L. Hansen, R. O. Dror, T. J. Mildorf, D. E. Shaw, B. K. Kobilka, R. S. Prosser, *J Am Chem Soc* **2013**, 135, 9465–9474.
- [136] V. D. Mehta, P. V. Kulkarni, R. P. Mason, A. Constantinescu, P. P. Antich, *Bioconjug Chem* **1994**, 5, 257–261.
- [137] A. S. Chubarov, M. M. Shakirov, I. V. Koptuyug, R. Z. Sagdeev, D. G. Knorre, T. S. Godovikova, *Bioorg Med Chem Lett* **2011**, 21, 4050–4053.
- [138] A. S. Chubarov, O. D. Zakharova, O. A. Koval, A. V. Romaschenko, A. E. Akulov, E. L. Zavjalov, I. A. Razumov, I. V. Koptuyug, D. G. Knorre, T. S. Godovikova, *Bioorg Med Chem* **2015**, 23, 6943–6954.
- [139] S. S. Shekhawat, G. H. Pham, J. Prabakaran, E. R. Strieter, *ACS Chem Biol* **2014**, 9, 2229–2236.
- [140] D. Hebel, K. L. Kirk, L. A. Cohen, V. M. Labroo, *Tetrahedron Lett* **1990**, 31, 619–622.
- [141] M. Imiołek, G. Karunanithy, W.-L. Ng, A. J. Baldwin, V. Gouverneur, B. G. Davis, *J Am Chem Soc* **2018**, 140, 1568–1571.
- [142] N. Ichiishi, J. P. Caldwell, M. Lin, W. Zhong, X. Zhu, E. Streckfuss, H.-Y. Kim, C. A. Parish, S. W. Krska, *Chem Sci* **2018**, 9, 4168–4175.
- [143] Z. Wu, F. Kandeel, *Curr Pharm Biotechnol* **2010**, 11, 572–580.
- [144] D. L. Lee, S. Ivaninskii, P. Burkhard, R. S. Hodges, *Protein Science* **2003**, 12, 1395–1405.
- [145] Y. Xiao, B. Ma, D. McElheny, S. Parthasarathy, F. Long, M. Hoshi, R. Nussinov, Y. Ishii, *Nat Struct Mol Biol* **2015**, 22, 499–505.
- [146] J. Bella, B. Brodsky, H. M. Berman, *Structure* **1995**, 3, 893–906.

- [147] A. Acharyya, Y. Ge, H. Wu, W. F. DeGrado, V. A. Voelz, F. Gai, *J Phys Chem B* **2019**, *123*, 1797–1807.
- [148] D. N. Woolfson, in *Adv Protein Chem*, Elsevier, **2005**, pp. 79–112.
- [149] A. N. Lupas, M. Gruber, **2005**, pp. 37–38.
- [150] D. N. Woolfson, in *Fibrous Proteins: Structures and Mechanisms* (Eds.: D. Parry, J. Squire), Springer, Cham., Basel (Switzerland), **2017**, pp. 35–61.
- [151] L. Truebestein, T. A. Leonard, *BioEssays* **2016**, *38*, 903–916.
- [152] F. H. C. Crick, *Acta Crystallogr* **1953**, *6*, 689–697.
- [153] F. H. C. Crick, *Acta Crystallogr* **1953**, *6*, 685–689.
- [154] L. PAULING, R. B. COREY, *Nature* **1953**, *171*, 59–61.
- [155] F. H. C. CRICK, *Nature* **1952**, *170*, 882–883.
- [156] F. Lapenta, J. Aupič, Ž. Strmšek, R. Jerala, *Chem Soc Rev* **2018**, *47*, 3530–3542.
- [157] Y. Wu, J. H. Collier, *WIREs Nanomedicine and Nanobiotechnology* **2017**, *9*, DOI 10.1002/wnan.1424.
- [158] J. M. Mason, K. M. Arndt, *ChemBioChem* **2004**, *5*, 170–176.
- [159] J. Utterström, S. Naeimipour, R. Selegård, D. Aili, *Adv Drug Deliv Rev* **2021**, *170*, 26–43.
- [160] A. Lupas, *Trends Biochem Sci* **1996**, *21*, 375–82.
- [161] P. Burkhard, M. Meier, A. Lustig, *Protein Science* **2000**, *9*, 2294–2301.
- [162] W. D. Kohn, C. T. Mant, R. S. Hodges, *Journal of Biological Chemistry* **1997**, *272*, 2583–2586.
- [163] A. R. Thomson, C. W. Wood, A. J. Burton, G. J. Bartlett, R. B. Sessions, R. L. Brady, D. N. Woolfson, *Science (1979)* **2014**, *346*, 485–488.
- [164] J. Liu, Q. Zheng, Y. Deng, C.-S. Cheng, N. R. Kallenbach, M. Lu, *Proceedings of the National Academy of Sciences* **2006**, *103*, 15457–15462.
- [165] J. F. Conway, D. A. D. Parry, *Int J Biol Macromol* **1990**, *12*, 328–334.
- [166] G. G. Rhys, C. W. Wood, E. J. M. Lang, A. J. Mulholland, R. L. Brady, A. R. Thomson, D. N. Woolfson, *Nat Commun* **2018**, *9*, 4132.
- [167] C. K. Thota, D. J. Mikolajczak, C. Roth, B. Kokscha, *ACS Med Chem Lett* **2021**, *12*, 67–73.
- [168] E. Zacco, C. Anish, C. E. Martin, H. v. Berlepsch, E. Brandenburg, P. H. Seeberger, B. Kokscha, *Biomacromolecules* **2015**, *16*, 2188–2197.
- [169] P. Burkhard, S. Ivaninskii, A. Lustig, *J Mol Biol* **2002**, *318*, 901–910.

- [170] A. Acharya, V. Rishi, C. Vinson, *Biochemistry* **2006**, *45*, 11324–11332.
- [171] G. GRIGORYAN, A. KEATING, *Curr Opin Struct Biol* **2008**, *18*, 477–483.
- [172] N. E. Zhou, C. M. Kay, R. S. Hodges, *Journal of Biological Chemistry* **1992**, *267*, 2664–2670.
- [173] N. E. Zhou, C. M. Kay, R. S. Hodges, *Biochemistry* **1992**, *31*, 5739–5746.
- [174] P. Y. Chou, G. D. Fasman, *Biochemistry* **1974**, *13*, 222–245.
- [175] C. K. Smith, L. Regan, *Acc Chem Res* **1997**, *30*, 153–161.
- [176] F. Chiti, C. M. Dobson, *Annu Rev Biochem* **2017**, *86*, 27–68.
- [177] M. Tsutsumi, J. M. Otaki, *J Chem Inf Model* **2011**, *51*, 1457–1464.
- [178] H. E. Stanger, S. H. Gellman, *J Am Chem Soc* **1998**, *120*, 4236–4237.
- [179] K. Y. Tsang, H. Diaz, N. Graciani, J. W. Kelly, *J Am Chem Soc* **1994**, *116*, 3988–4005.
- [180] F. J. Blanco, M. A. Jimenez, J. Herranz, M. Rico, J. Santoro, J. L. Nieto, *J Am Chem Soc* **1993**, *115*, 5887–5888.
- [181] O. Rosen, J. Chill, M. Sharon, N. Kessler, B. Mester, S. Zolla-Pazner, J. Anglister, *Biochemistry* **2005**, *44*, 7250–7258.
- [182] V. Wineman-Fisher, R. Simkovitch, S. Shomer, R. Gepshtein, D. Huppert, M. Saif, K. Kallio, S. J. Remington, Y. Miller, *Phys. Chem. Chem. Phys.* **2014**, *16*, 11211–11223.
- [183] M. S. Kelker, C. Berry, S. L. Evans, R. Pai, D. G. McCaskill, N. X. Wang, J. C. Russell, M. D. Baker, C. Yang, J. W. Pflugrath, M. Wade, T. J. Wess, K. E. Narva, *PLoS One* **2014**, *9*, e112555.
- [184] A. A. Adzhubei, M. J. E. Sternberg, A. A. Makarov, *J Mol Biol* **2013**, *425*, 2100–2132.
- [185] A. A. Makarov, N. G. Esipova, Yu. A. Pankov, V. M. Lobachev, B. A. Grishkovsky, *Biochem Biophys Res Commun* **1975**, *67*, 1378–1383.
- [186] G. J. Cameron, I. L. Alberts, J. H. Laing, T. J. Wess, *J Struct Biol* **2002**, *137*, 15–22.
- [187] J. Hartmann, M. Zacharias, *PLoS Comput Biol* **2021**, *17*, e1009079.
- [188] S. A. H. Hulgán, J. D. Hartgerink, *Biomacromolecules* **2022**, *23*, 1475–1489.
- [189] W. C. Wimley, T. P. Creamer, S. H. White, *Biochemistry* **1996**, *35*, 5109–5124.
- [190] E. Q. Lawson, A. J. Sadler, D. Harmatz, D. T. Brandau, R. Micanovic, R. D. MacElroy, C. R. Middaugh, *J Biol Chem* **1984**, *259*, 2910–2.



- [191] S. Damodaran, K. B. Song, *J Biol Chem* **1986**, *261*, 7220–2.
- [192] R. S. Hodges, B.-Y. Zhu, N. E. Zhou, C. T. Mant, *J Chromatogr A* **1994**, *676*, 3–15.
- [193] C. T. Mant, J. M. Kovacs, H.-M. Kim, D. D. Pollock, R. S. Hodges, *Biopolymers* **2009**, *92*, 573–595.
- [194] K. Yutani, K. Ogasahara, T. Tsujita, Y. Sugino, *Proceedings of the National Academy of Sciences* **1987**, *84*, 4441–4444.
- [195] C. Zhu, Y. Gao, H. Li, S. Meng, L. Li, J. S. Francisco, X. C. Zeng, *Proceedings of the National Academy of Sciences* **2016**, *113*, 12946–12951.
- [196] C. C. Palliser, D. A. D. Parry, *Proteins: Structure, Function, and Genetics* **2001**, *42*, 243–255.
- [197] K. M. Biswas, D. R. DeVido, J. G. Dorsey, *J Chromatogr A* **2003**, *1000*, 637–655.
- [198] K.-H. Lee, H.-Y. Lee, M. M. Slutsky, J. T. Anderson, E. N. G. Marsh, *Biochemistry* **2004**, *43*, 16277–16284.
- [199] S. A. Samsonov, M. Salwiczek, G. Anders, B. Koksich, M. T. Pisabarro, *J Phys Chem B* **2009**, *113*, 16400–16408.
- [200] H. Erdbrink, E. K. Nyakatura, S. Huhmann, U. I. M. Gerling, D. Lentz, B. Koksich, C. Czekelius, *Beilstein Journal of Organic Chemistry* **2013**, *9*, 2009–2014.
- [201] U. I. M. Gerling, M. Salwiczek, C. D. Cadicamo, H. Erdbrink, C. Czekelius, S. L. Grage, P. Wadhvani, A. S. Ulrich, M. Behrends, G. Haufe, B. Koksich, *Chem. Sci.* **2014**, *5*, 819–830.
- [202] S. Chowdhary, J. Moschner, D. J. Mikolajczak, M. Becker, A. F. Thünemann, C. Kästner, D. Klemczak, A. Stegemann, C. Böttcher, P. Metrangolo, R. R. Netz, B. Koksich, *ChemBioChem* **2020**, *21*, 3544–3554.
- [203] C. Gadais, E. Devillers, V. Gasparik, E. Chelain, J. Pytkowicz, T. Brigaud, *ChemBioChem* **2018**, *19*, 1026–1030.
- [204] T. Bączek, R. Kaliszan, *Proteomics* **2009**, *9*, 835–847.
- [205] Y. H. Zhao, M. H. Abraham, A. M. Zissimos, *J Org Chem* **2003**, *68*, 7368–7373.
- [206] J. R. Robalo, S. Huhmann, B. Koksich, A. Vila Verde, *Chem* **2017**, *3*, 881–897.
- [207] J. R. Robalo, A. Vila Verde, *Physical Chemistry Chemical Physics* **2019**, *21*, 2029–2038.
- [208] A. Chakrabartty, T. Kortemme, R. L. Baldwin, *Protein Science* **1994**, *3*, 843–852.

- [209] A. J. Doig, A. Chakrabartty, T. M. Klingler, R. L. Baldwin, *Biochemistry* **1994**, *33*, 3396–3403.
- [210] H.-P. Chiu, Y. Suzuki, D. Gullickson, R. Ahmad, B. Kokona, R. Fairman, R. P. Cheng, *J Am Chem Soc* **2006**, *128*, 15556–15557.
- [211] H. Erdbrink, I. Peuser, U. I. M. Gerling, D. Lentz, B. Kokschi, C. Czekelius, *Org Biomol Chem* **2012**, *10*, 8583.
- [212] G. A. Clark, J. D. Baleja, K. Kumar, *J Am Chem Soc* **2012**, *134*, 17912–17921.
- [213] M. D. Shoulders, R. T. Raines, *Annu Rev Biochem* **2009**, *78*, 929–958.
- [214] M. A. Cejas, W. A. Kinney, C. Chen, J. G. Vinter, H. R. Almond, K. M. Balss, C. A. Maryanoff, U. Schmidt, M. Breslav, A. Mahan, E. Lacy, B. E. Maryanoff, *Proceedings of the National Academy of Sciences* **2008**, *105*, 8513–8518.
- [215] D. Gottlieb, S. A. Morin, S. Jin, R. T. Raines, *J Mater Chem* **2008**, *18*, 3865.
- [216] B. C. Buer, J. L. Meagher, J. A. Stuckey, E. N. G. Marsh, *Proceedings of the National Academy of Sciences* **2012**, *109*, 4810–4815.
- [217] B. Bilgiçer, X. Xing, K. Kumar, *J Am Chem Soc* **2001**, *123*, 11815–11816.
- [218] C. Jäckel, W. Seufert, S. Thust, B. Kokschi, *ChemBioChem* **2004**, *5*, 717–720.
- [219] S. Huhmann, E. K. Nyakatura, H. Erdbrink, U. I. M. Gerling, C. Czekelius, B. Kokschi, *J Fluor Chem* **2015**, *175*, 32–35.
- [220] M. Salwiczek, S. Samsonov, T. Vagt, E. Nyakatura, E. Fleige, J. Numata, H. Cölfen, M. T. Pisabarro, B. Kokschi, *Chemistry - A European Journal* **2009**, *15*, 7628–7636.
- [221] M. D. Shoulders, K. J. Kamer, R. T. Raines, *Bioorg Med Chem Lett* **2009**, *19*, 3859–3862.
- [222] J. A. Hodges, R. T. Raines, *J Am Chem Soc* **2005**, *127*, 15923–15932.
- [223] B. Bilgiçer, A. Fichera, K. Kumar, *J Am Chem Soc* **2001**, *123*, 4393–4399.
- [224] H. T. More, K. S. Zhang, N. Srivastava, J. A. Frezzo, J. K. Montclare, *Biomacromolecules* **2015**, *16*, 1210–1217.
- [225] Y. Tang, G. Ghirlanda, N. Vaidehi, J. Kua, D. T. Mainz, W. A. Goddard, W. F. DeGrado, D. A. Tirrell, *Biochemistry* **2001**, *40*, 2790–2796.
- [226] S. Son, I. C. Tanrikulu, D. A. Tirrell, *ChemBioChem* **2006**, *7*, 1251–1257.
- [227] Y. Tang, G. Ghirlanda, W. A. Petka, T. Nakajima, W. F. DeGrado, D. A. Tirrell, *Angewandte Chemie International Edition* **2001**, *40*, 1494–1496.

- [228] J. K. Montclare, S. Son, G. A. Clark, K. Kumar, D. A. Tirrell, *ChemBioChem* **2009**, *10*, 84–86.
- [229] Y. Tang, D. A. Tirrell, *J Am Chem Soc* **2001**, *123*, 11089–11090.
- [230] H.-P. Chiu, B. Kokona, R. Fairman, R. P. Cheng, *J Am Chem Soc* **2009**, *131*, 13192–13193.
- [231] R. Campos-Olivas, R. Aziz, G. L. Helms, J. N. S. Evans, A. M. Gronenborn, *FEBS Lett* **2002**, *517*, 55–60.
- [232] J.-C. Horng, D. P. Raleigh, *J Am Chem Soc* **2003**, *125*, 9286–9287.
- [233] H. Welte, T. Zhou, X. Mihajlenko, O. Mayans, M. Kovermann, *Sci Rep* **2020**, *10*, 2640.
- [234] T. Steiner, P. Hess, J. H. Bae, B. Wiltschi, L. Moroder, N. Budisa, *PLoS One* **2008**, *3*, e1680.
- [235] D. Alexeev, P. N. Barlow, S. M. Bury, J.-D. Charrier, A. Cooper, D. Hadfield, C. Jamieson, S. M. Kelly, R. Layfield, R. J. Mayer, H. McSparron, N. C. Price, R. Ramage, L. Sawyer, B. A. Starkmann, D. Uhrin, J. Wilken, D. W. Young, *ChemBioChem* **2003**, *4*, 894–896.
- [236] M. D. Crespo, M. Rubini, *PLoS One* **2011**, *6*, e19425.
- [237] C. Minks, R. Huber, L. Moroder, N. Budisa, *Biochemistry* **1999**, *38*, 10649–10659.
- [238] C. Jäckel, M. Salwiczek, B. Kokschi, *Angewandte Chemie International Edition* **2006**, *45*, 4198–4203.
- [239] A. Saghatelian, Y. Yokobayashi, K. Soltani, M. R. Ghadiri, *Nature* **2001**, *409*, 797–801.
- [240] K. Severin, D. H. Lee, J. A. Martinez, M. R. Ghadiri, *Chemistry - A European Journal* **1997**, *3*, 1017–1024.
- [241] M. Salwiczek, B. Kokschi, *ChemBioChem* **2009**, *10*, 2867–2870.
- [242] J. F. Hunt, P. Rath, K. J. Rothschild, D. M. Engelman, *Biochemistry* **1997**, *36*, 15177–15192.
- [243] O. A. Andreev, A. D. Dupuy, M. Segala, S. Sandugu, D. A. Serra, C. O. Chichester, D. M. Engelman, Y. K. Reshetnyak, *Proceedings of the National Academy of Sciences* **2007**, *104*, 7893–7898.
- [244] L. C. Wyatt, J. S. Lewis, O. A. Andreev, Y. K. Reshetnyak, D. M. Engelman, *Trends Biotechnol* **2017**, *35*, 653–664.

- [245] O. A. Andreev, A. G. Karabadzak, D. Weerakkody, G. O. Andreev, D. M. Engelman, Y. K. Reshetnyak, *Proceedings of the National Academy of Sciences* **2010**, *107*, 4081–4086.
- [246] Enrico Brandenburg, Inhibition Der Bildung Amyloider Aggregate, Freie Universität, **2012**.
- [247] F. Thomas, A. L. Boyle, A. J. Burton, D. N. Woolfson, *J Am Chem Soc* **2013**, *135*, 5161–5166.
- [248] E. K. O'Shea, J. D. Klemm, P. S. Kim, T. Alber, *Science (1979)* **1991**, *254*, 539–544.
- [249] F. Thomas, A. Niitsu, A. Oregioni, G. J. Bartlett, D. N. Woolfson, *Biochemistry* **2017**, *56*, 6544–6554.
- [250] L. Gonzalez, D. N. Woolfson, T. Alber, *Nat Struct Biol* **1996**, *3*, 1011–1018.
- [251] L. Gonzalez, J. J. Plecs, T. Alber, *Nat Struct Mol Biol* **1996**, *3*, 510–515.
- [252] P. B. Harbury, T. Zhang, P. S. Kim, T. Alber, *Science (1979)* **1993**, *262*, 1401–1407.
- [253] G. Goldstein, M. Scheid, U. Hammerling, D. H. Schlesinger, H. D. Niall, E. A. Boyse, *Proceedings of the National Academy of Sciences* **1975**, *72*, 11–15.
- [254] S. Vijay-Kumar, C. E. Bugg, K. D. Wilkinson, R. D. Vierstra, P. M. Hatfield, W. J. Cook, *J Biol Chem* **1987**, *262*, 6396–9.
- [255] S. Vijay-kumar, C. E. Bugg, W. J. Cook, *J Mol Biol* **1987**, *194*, 531–544.
- [256] M. S. Briggs, H. Roder, *Proceedings of the National Academy of Sciences* **1992**, *89*, 2017–2021.
- [257] D. H. Schlesinger, G. Goldstein, H. D. Niall, *Biochemistry* **1975**, *14*, 2214–2218.
- [258] S. Vijay-Kumar, C. E. Bugg, K. D. Wilkinson, W. J. Cook, *Proceedings of the National Academy of Sciences* **1985**, *82*, 3582–3585.
- [259] K. Husnjak, I. Dikic, *Annu Rev Biochem* **2012**, *81*, 291–322.
- [260] L. Buetow, D. T. Huang, *Nat Rev Mol Cell Biol* **2016**, *17*, 626–642.
- [261] M. Rape, *Nat Rev Mol Cell Biol* **2018**, *19*, 59–70.
- [262] K. Kliza, K. Husnjak, *Front Mol Biosci* **2020**, *7*, DOI 10.3389/fmolb.2020.00021.
- [263] R. Beal, Q. Deveraux, G. Xia, M. Rechsteiner, C. Pickart, *Proceedings of the National Academy of Sciences* **1996**, *93*, 861–866.
- [264] K. E. Sloper-Mould, J. C. Jemc, C. M. Pickart, L. Hicke, *Journal of Biological Chemistry* **2001**, *276*, 30483–30489.

- [265] S. Virdee, Y. Ye, D. P. Nguyen, D. Komander, J. W. Chin, *Nat Chem Biol* **2010**, *6*, 750–757.
- [266] C. A. Castañeda, T. R. Kashyap, M. A. Nakasone, S. Krueger, D. Fushman, *Structure* **2013**, *21*, 1168–1181.
- [267] W. J. Cook, L. C. Jeffrey, M. Carson, Z. Chen, C. M. Pickart, *Journal of Biological Chemistry* **1992**, *267*, 16467–16471.
- [268] D. Komander, F. Reyes-Turcu, J. D. F. Licchesi, P. Odenwaelder, K. D. Wilkinson, D. Barford, *EMBO Rep* **2009**, *10*, 466–473.
- [269] J. G. B. Northey, A. A. Di Nardo, A. R. Davidson, *Nat Struct Biol* **2002**, *9*, 126–130.
- [270] A. Musacchio, M. Saraste, M. Wilmanns, *Nat Struct Biol* **1994**, *1*, 546–551.
- [271] D. J. Leahy, W. A. Hendrickson, Ikramuddin Aukhil, H. P. Erickson, *Science (1979)* **1992**, *258*, 987–991.
- [272] S. J. Hamill, A. Steward, J. Clarke, *J Mol Biol* **2000**, *297*, 165–178.
- [273] N. Hartrampf, A. Saebi, M. Poskus, Z. P. Gates, A. J. Callahan, A. E. Cowfer, S. Hanna, S. Antilla, C. K. Schissel, A. J. Quartararo, X. Ye, A. J. Mijalis, M. D. Simon, A. Loas, S. Liu, C. Jessen, T. E. Nielsen, B. L. Pentelute, *Science (1979)* **2020**, *368*, 980–987.
- [274] R. B. Merrifield, *J Am Chem Soc* **1963**, *85*, 2149–2154.
- [275] A. Isidro-Llobet, M. Álvarez, F. Albericio, *Chem Rev* **2009**, *109*, 2455–2504.
- [276] A. El-Faham, F. Albericio, *Chem Rev* **2011**, *111*, 6557–6602.
- [277] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science (1979)* **1994**, *266*, 776–779.
- [278] H. Sun, S. M. Mali, S. K. Singh, R. Meledin, A. Brik, Y. T. Kwon, Y. Kravtsova-Ivantsiv, B. Bercovich, A. Ciechanover, *Proceedings of the National Academy of Sciences* **2019**, *116*, 7805–7812.
- [279] J. W. Bode, R. M. Fox, K. D. Baucom, *Angewandte Chemie International Edition* **2006**, *45*, 1248–1252.
- [280] N. J. Mitchell, L. R. Malins, X. Liu, R. E. Thompson, B. Chan, L. Radom, R. J. Payne, *J Am Chem Soc* **2015**, *137*, 14011–14014.
- [281] Y. Zhang, C. Xu, H. Y. Lam, C. L. Lee, X. Li, *Proceedings of the National Academy of Sciences* **2013**, *110*, 6657–6662.

- [282] V. Agouridas, O. El Mahdi, V. Diemer, M. Cargoët, J.-C. M. Monbaliu, O. Melnyk, *Chem Rev* **2019**, *119*, 7328–7443.
- [283] D. Xu, R. Nussinov, *Fold Des* **1998**, *3*, 11–17.
- [284] “protein chemical synthesis database,” can be found under <http://pcs-db.fr/>, **n.d.**
- [285] B. L. Nilsson, L. L. Kiessling, R. T. Raines, *Org Lett* **2000**, *2*, 1939–1941.
- [286] E. Saxon, J. I. Armstrong, C. R. Bertozzi, *Org Lett* **2000**, *2*, 2141–2143.
- [287] J. P. Tam, Z. Miao, *J Am Chem Soc* **1999**, *121*, 9013–9022.
- [288] J. P. TAM, C. RAO, C.-F. LIU, J. SHAO, *Int J Pept Protein Res* **2009**, *45*, 209–216.
- [289] C. F. Liu, J. P. Tam, *Proceedings of the National Academy of Sciences* **1994**, *91*, 6584–6588.
- [290] A. C. Conibear, E. E. Watson, R. J. Payne, C. F. W. Becker, *Chem Soc Rev* **2018**, *47*, 9046–9068.
- [291] A. L. Baumann, C. P. R. Hackenberger, *Chimia (Aarau)* **2018**, *72*, 802.
- [292] Y. Tan, H. Wu, T. Wei, X. Li, *J Am Chem Soc* **2020**, *142*, 20288–20298.
- [293] S. S. Kulkarni, J. Sayers, B. Premdjee, R. J. Payne, *Nat Rev Chem* **2018**, *2*, 0122.
- [294] D. M. M. Jaradat, *Amino Acids* **2018**, *50*, 39–68.
- [295] <http://www.ebi.ac.uk/uniprot/TrEMBLstats/>, “UniProtKB/TrEMBL,” **n.d.**
- [296] H. Rohde, O. Seitz, *Biopolymers* **2010**, *94*, 551–559.
- [297] L. Z. Yan, P. E. Dawson, *J Am Chem Soc* **2001**, *123*, 526–533.
- [298] G. Arsequell, A. González, G. Valencia, *Tetrahedron Lett* **2001**, *42*, 2685–2687.
- [299] Q. Wan, S. J. Danishefsky, *Angewandte Chemie International Edition* **2007**, *46*, 9248–9252.
- [300] L. R. Malins, R. J. Payne, *Aust J Chem* **2015**, *68*, 521.
- [301] B. Premdjee, R. J. Payne, in *Chemical Ligation: Tools for Biomolecule Synthesis and Modification* (Eds.: L.D. D’Andrea, A. Romanelli), Wiley, Hoboken (NJ, USA), **2017**, pp. 161–218.
- [302] I. Guan, K. Williams, J. S. T. Liu, X. Liu, *Front Chem* **2022**, *9*, DOI 10.3389/fchem.2021.826764.
- [303] R. E. Thompson, B. Chan, L. Radom, K. A. Jolliffe, R. J. Payne, *Angewandte Chemie International Edition* **2013**, *52*, 9723–9727.
- [304] X. Guan, M. R. Drake, Z. Tan, *Org Lett* **2013**, *15*, 6128–6131.

- [305] K. M. Cergol, R. E. Thompson, L. R. Malins, P. Turner, R. J. Payne, *Org Lett* **2014**, *16*, 290–293.
- [306] D. Crich, A. Banerjee, *J Am Chem Soc* **2007**, *129*, 10064–10065.
- [307] K. K. Pasunooti, R. Yang, B. Banerjee, T. Yap, C.-F. Liu, *Org Lett* **2016**, *18*, 2696–2699.
- [308] Z. Harpaz, P. Siman, K. S. A. Kumar, A. Brik, *ChemBioChem* **2010**, *11*, 1232–1235.
- [309] Z. Tan, S. Shang, S. J. Danishefsky, *Angewandte Chemie International Edition* **2010**, *49*, 9500–9503.
- [310] R. Merkx, G. de Bruin, A. Kruithof, T. van den Bergh, E. Snip, M. Lutz, F. El Oualid, H. Ovaa, *Chem Sci* **2013**, *4*, 4494.
- [311] R. Yang, K. K. Pasunooti, F. Li, X.-W. Liu, C.-F. Liu, *J Am Chem Soc* **2009**, *131*, 13592–13593.
- [312] F. El Oualid, R. Merkx, R. Ekkebus, D. S. Hameed, J. J. Smit, A. de Jong, H. Hilkmann, T. K. Sixma, H. Ovaa, *Angewandte Chemie International Edition* **2010**, *49*, 10149–10153.
- [313] K. S. Ajish Kumar, M. Haj-Yahya, D. Olschewski, H. A. Lashuel, A. Brik, *Angewandte Chemie International Edition* **2009**, *48*, 8090–8094.
- [314] L. R. Malins, K. M. Cergol, R. J. Payne, *Chem. Sci.* **2014**, *5*, 260–266.
- [315] P. Siman, S. V. Karthikeyan, A. Brik, *Org Lett* **2012**, *14*, 1520–1523.
- [316] L. R. Malins, K. M. Cergol, R. J. Payne, *ChemBioChem* **2013**, *14*, 559–563.
- [317] J. Sayers, R. E. Thompson, K. J. Perry, L. R. Malins, R. J. Payne, *Org Lett* **2015**, *17*, 4902–4905.
- [318] S. Shang, Z. Tan, S. Dong, S. J. Danishefsky, *J Am Chem Soc* **2011**, *133*, 10784–10786.
- [319] H. Ding, A. Shigenaga, K. Sato, K. Morishita, A. Otaka, *Org Lett* **2011**, *13*, 5588–5591.
- [320] J. Chen, P. Wang, J. Zhu, Q. Wan, S. J. Danishefsky, *Tetrahedron* **2010**, *66*, 2277–2283.
- [321] C. Haase, H. Rohde, O. Seitz, *Angewandte Chemie International Edition* **2008**, *47*, 6807–6810.
- [322] J. Chen, Q. Wan, Y. Yuan, J. Zhu, S. J. Danishefsky, *Angewandte Chemie International Edition* **2008**, *47*, 8521–8524.

- [323] K. S. A. Kumar, L. Spasser, L. A. Erlich, S. N. Bavikar, A. Brik, *Angewandte Chemie International Edition* **2010**, *49*, 9126–9131.
- [324] A. B. Clippingdale, C. J. Barrow, J. D. Wade, *Journal of Peptide Science* **2000**, *6*, 225–234.
- [325] X. Li, T. Kawakami, S. Aimoto, *Tetrahedron Lett* **1998**, *39*, 8669–8672.
- [326] J. B. Blanco-Canosa, P. E. Dawson, *Angewandte Chemie International Edition* **2008**, *47*, 6851–6855.
- [327] G.-M. Fang, Y.-M. Li, F. Shen, Y.-C. Huang, J.-B. Li, Y. Lin, H.-K. Cui, L. Liu, *Angewandte Chemie International Edition* **2011**, *50*, 7645–7649.
- [328] J. Dheur, N. Ollivier, O. Melnyk, *Org Lett* **2011**, *13*, 1560–1563.
- [329] N. Ollivier, J. Dheur, R. Mhidia, A. Blanpain, O. Melnyk, *Org Lett* **2010**, *12*, 5238–5241.
- [330] E. C. B. Johnson, S. B. H. Kent, *J Am Chem Soc* **2006**, *128*, 6640–6646.
- [331] S. Gunasekera, T. L. Aboye, W. A. Madian, H. R. El-Seedi, U. Göransson, *Int J Pept Res Ther* **2013**, *19*, 43–54.
- [332] L. Dery, P. S. Reddy, S. Dery, R. Mousa, O. Ktorza, A. Talhami, N. Metanis, *Chem Sci* **2017**, *8*, 1922–1926.
- [333] Z. Zhao, R. Mousa, N. Metanis, *Chemistry – A European Journal* **2022**, *28*, DOI 10.1002/chem.202200279.
- [334] B. Fauvet, S. M. Butterfield, J. Fuks, A. Brik, H. A. Lashuel, *Chemical Communications* **2013**, *49*, 9254.
- [335] A. Baral, A. Asokan, V. Bauer, B. Kieffer, V. Torbeev, *Tetrahedron* **2019**, *75*, 703–708.
- [336] S.-Sun. Wang, *J Am Chem Soc* **1973**, *95*, 1328–1333.
- [337] J. K. Chang, M. Shimizu, S.-S. Wang, *J Org Chem* **1976**, *41*, 3255–3258.
- [338] A. A. Vinogradov, M. D. Simon, B. L. Pentelute, *Org Lett* **2016**, *18*, 1222–1225.
- [339] S. Gao, M. Pan, Y. Zheng, Y. Huang, Q. Zheng, D. Sun, L. Lu, X. Tan, X. Tan, H. Lan, J. Wang, T. Wang, J. Wang, L. Liu, *J Am Chem Soc* **2016**, *138*, 14497–14502.
- [340] S. Tang, L.-J. Liang, Y.-Y. Si, S. Gao, J.-X. Wang, J. Liang, Z. Mei, J.-S. Zheng, L. Liu, *Angewandte Chemie* **2017**, *129*, 13518–13522.
- [341] M. Pan, S. Gao, Y. Zheng, X. Tan, H. Lan, X. Tan, D. Sun, L. Lu, T. Wang, Q. Zheng, Y. Huang, J. Wang, L. Liu, *J Am Chem Soc* **2016**, *138*, 7429–7435.



- [342] M. Pan, Y. He, M. Wen, F. Wu, D. Sun, S. Li, L. Zhang, Y. Li, C. Tian, *Chem. Commun.* **2014**, *50*, 5837–5839.
- [343] Y. Li, Q. Qu, Y. Qi, L. Liu, K. W. Wang, Y. Liu, G. Fang, *Journal of Peptide Science* **2022**, *28*, DOI 10.1002/psc.3365.
- [344] Y.-C. Huang, G.-M. Fang, L. Liu, *Natl Sci Rev* **2016**, *3*, 107–116.
- [345] M. Cargoët, V. Diemer, B. Snella, R. Desmet, A. Blanpain, H. Drobecq, V. Agouridas, O. Melnyk, *J Org Chem* **2018**, *83*, 12584–12594.
- [346] M. Cargoët, V. Diemer, L. Raibaut, E. Lissy, B. Snella, V. Agouridas, O. Melnyk, in *Peptide and Protein Engineering. Springer Protocols Handbooks*. (Eds.: O. Iranzo, A. Roque), Humana, New York, NY, **2020**, pp. 1–12.
- [347] L. Raibaut, N. Ollivier, O. Melnyk, *Chem Soc Rev* **2012**, *41*, 7001.
- [348] N. Ollivier, J. Vicogne, A. Vallin, H. Drobecq, R. Desmet, O. El Mahdi, B. Leclercq, G. Goormachtigh, V. Fafeur, O. Melnyk, *Angewandte Chemie International Edition* **2012**, *51*, 209–213.
- [349] L. Raibaut, J. Vicogne, B. Leclercq, H. Drobecq, R. Desmet, O. Melnyk, *Bioorg Med Chem* **2013**, *21*, 3486–3494.
- [350] E. Boll, H. Drobecq, N. Ollivier, A. Blanpain, L. Raibaut, R. Desmet, J. Vicogne, O. Melnyk, *Nat Protoc* **2015**, *10*, 269–292.
- [351] T. Tsushima, K. Kawada, S. Ishihara, N. Uchida, O. Shiratori, J. Higaki, M. Hirata, *Tetrahedron* **1988**, *44*, 5375–5387.
- [352] H. A. Lindner, A. Alary, L. I. Boju, T. Sulea, R. Ménard, *Biochemistry* **2005**, *44*, 15645–15651.
- [353] Gage James R., Evans David A., *Organic Syntheses* **1990**, *68*, 77.
- [354] J. Han, R. Takeda, X. Liu, H. Konno, H. Abe, T. Hiramatsu, H. Moriwaki, V. A. Soloshonok, *Molecules* **2019**, *24*, DOI 10.3390/molecules24244521.
- [355] Yu. N. Belokon, A. G. Bulychyev, S. V. Vitt, Yu. T. Struchkov, A. S. Batsanov, T. V. Timofeeva, V. A. Tsyryapkin, M. G. Ryzhov, L. A. Lysova, *J Am Chem Soc* **1985**, *107*, 4252–4259.
- [356] A. E. Sorochinsky, J. L. Aceña, H. Moriwaki, T. Sato, V. Soloshonok, *Amino Acids* **2013**, *45*, 1017–1033.
- [357] A. E. Sorochinsky, J. L. Aceña, H. Moriwaki, T. Sato, V. A. Soloshonok, *Amino Acids* **2013**, *45*, 691–718.

- [358] Y. Wang, X. Song, J. Wang, H. Moriwaki, V. A. Soloshonok, H. Liu, *Amino Acids* **2017**, *49*, 1487–1520.
- [359] J. L. Aceña, A. E. Sorochinsky, V. Soloshonok, *Amino Acids* **2014**, *46*, 2047–2073.
- [360] T. Hohmann, M. Dyrks, S. Chowdhary, M. Weber, D. Nguyen, J. Moschner, B. Kokschi, *J Org Chem* **2022**, *87*, 10592–10604.
- [361] A. Kumar, A. Sharma, B. G. de la Torre, F. Albericio, *Green Chemistry* **2022**, *24*, 4887–4896.
- [362] S. Chowdhary, R. F. Schmidt, A. K. Sahoo, T. tom Dieck, T. Hohmann, B. Schade, K. Brademann-Jock, A. F. Thünemann, R. R. Netz, M. Gradzielski, B. Kokschi, *Nanoscale* **2022**, *14*, 10176–10189.
- [363] J. Leppkes, N. Dimos, B. Loll, T. Hohmann, M. Dyrks, A. Wieseke, B. G. Keller, B. Kokschi, *RSC Chem Biol* **2022**, *3*, 773–782.
- [364] K. Ataka, J. Drauschke, V. Stulberg, B. Kokschi, J. Heberle, *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2022**, *1864*, 183873.
- [365] F. N. Barrera, D. Weerakkody, M. Anderson, O. A. Andreev, Y. K. Reshetnyak, D. M. Engelman, *J Mol Biol* **2011**, *413*, 359–371.
- [366] Fernandes Ana Rita de Lima, Investigation of Supramolecular Assemblies Based on de Novo Coiled Coil Peptidic Scaffolds, Freie Universität, **2019**.
- [367] CEM, “CarboMax approach,” can be found under <https://cem.com/en/carbomax-enhanced-peptide-coupling-at-elevated-temperatures>, **n.d.**
- [368] M. J. Pandya, G. M. Spooner, M. Sunde, J. R. Thorpe, A. Rodger, D. N. Woolfson, *Biochemistry* **2000**, *39*, 8728–8734.
- [369] Patricia López García, Coiled Coils as Mechanical Building Blocks, **2018**.
- [370] P. López-García, A. D. Araujo, A. E. Bergues-Pupo, I. Tunn, D. P. Fairlie, K. G. Blank, *Angewandte Chemie International Edition* **2021**, *60*, 232–236.
- [371] S. Huhmann, E. K. Nyakatura, H. Erdbrink, U. I. M. Gerling, C. Czekelius, B. Kokschi, *J Fluor Chem* **2015**, *175*, 32–35.
- [372] M. Salwiczek, S. Samsonov, T. Vagt, E. Nyakatura, E. Fleige, J. Numata, H. Cölfen, M. T. Pisabarro, B. Kokschi, *Chemistry - A European Journal* **2009**, *15*, 7628–7636.

- [373] T. Wöhr, F. Wahl, A. Nefzi, B. Rohwedder, T. Sato, X. Sun, M. Mutter, *J Am Chem Soc* **1996**, *118*, 9218–9227.
- [374] J. M. Collins, K. A. Porter, S. K. Singh, G. S. Vanier, *Org Lett* **2014**, *16*, 940–943.
- [375] S. L. Pedersen, A. P. Tofteng, L. Malik, K. J. Jensen, *Chem. Soc. Rev.* **2012**, *41*, 1826–1844.
- [376] J. M. Collins, N. E. Leadbeater, *Org Biomol Chem* **2007**, *5*, 1141.
- [377] B. Bacsá, K. Horváti, S. Bősze, F. Andreae, C. O. Kappe, *J Org Chem* **2008**, *73*, 7532–7542.
- [378] S. K. Singh, J. M. Collins, K. A. Porter, M. J. Karney, “Large Scale Stepwise Synthesis of Ubiquitin,” can be found under <https://cem.sharefile.com/share/view/s369a202817074df083cb56789213797b>, **2016**.
- [379] S. A. Palasek, Z. J. Cox, J. M. Collins, *Journal of Peptide Science* **2007**, *13*, 143–148.
- [380] J. L. Lauer, C. G. Fields, G. B. Fields, *Letters in Peptide Science* **1995**, *1*, 197–205.
- [381] J. H. JONES, W. I. RAMAGE, M. J. WITTY, *Int J Pept Protein Res* **2009**, *15*, 301–303.
- [382] H. Hibino, Y. Miki, Y. Nishiuchi, *Journal of Peptide Science* **2012**, *18*, 763–769.
- [383] Y. Yang, in *Side Reactions in Peptide Synthesis*, Elsevier, **2016**, pp. 257–292.
- [384] Y. Yang, in *Side Reactions in Peptide Synthesis*, Elsevier, **2016**, pp. 217–233.
- [385] R. Rocchi, A. Scatturin, L. Moroder, F. Marchiori, A. M. Tamburro, E. Scoffone, *J Am Chem Soc* **1969**, *91*, 492–496.
- [386] Z. I. Randhawa, H. E. Witkowska, J. Cone, J. A. Wilkins, P. Hughes, K. Yamanishi, S. Yasuda, Y. Masui, P. Arthur, *Biochemistry* **1994**, *33*, 4352–4362.
- [387] F. Naider, Z. Bohak, J. Yariv, *Biochemistry* **1972**, *11*, 3202–3208.
- [388] A. M. Gilles, P. Marlière, T. Rose, R. Sarfati, R. Longin, A. Meier, S. Femandjian, M. Monnot, G. N. Cohen, O. Bâzru, *Journal of Biological Chemistry* **1988**, *263*, 8204–8209.
- [389] C. B. Anfinsen, L. G. Corley, *Journal of Biological Chemistry* **1969**, *244*, 5149–5152.
- [390] V. Radtke, Fluorine Modification of K 48 - Linked Diubiquitin, Freie Universität, **2014**.

- [391] L. A. Erlich, K. S. A. Kumar, M. Haj-Yahya, P. E. Dawson, A. Brik, *Org Biomol Chem* **2010**, *8*, 2392.
- [392] S. C. Miller, T. S. Scanlan, *J Am Chem Soc* **1997**, *119*, 2301–2302.
- [393] S. K. Mahto, C. J. Howard, J. C. Shimko, J. J. Ottesen, *ChemBioChem* **2011**, *12*, 2488–2494.
- [394] J. B. Blanco-Canosa, B. Nardone, F. Albericio, P. E. Dawson, *J Am Chem Soc* **2015**, *137*, 7197–7209.
- [395] S. Tsuda, T. Uemura, M. Mochizuki, H. Nishio, T. Yoshiya, *Synlett* **2017**, *28*, 1956–1960.
- [396] G. Stavropoulos, D. Gatos, V. Magafa, K. Barlos, *Letters in Peptide Science* **1996**, *2*, 315–318.
- [397] E. Kaiser, R. L. Colescott, C. D. Bossinger, P. I. Cook, *Anal Biochem* **1970**, *34*, 595–598.
- [398] S. Eissler, M. Kley, D. Bächle, G. Loidl, T. Meier, D. Samson, *Journal of Peptide Science* **2017**, *23*, 757–762.
- [399] J.-S. Zheng, S. Tang, Y.-K. Qi, Z.-P. Wang, L. Liu, *Nat Protoc* **2013**, *8*, 2483–2495.
- [400] Y.-C. Huang, C.-C. Chen, S.-J. Li, S. Gao, J. Shi, Y.-M. Li, *Tetrahedron* **2014**, *70*, 2951–2955.
- [401] M. J. Bird, P. E. Dawson, *Peptide Science* **2022**, *114*, DOI 10.1002/pep2.24268.
- [402] aapptec, “Technical Support Information Bulletin 1198,” can be found under [www.peptide.com/custdocs/1198.pdf](http://www.peptide.com/custdocs/1198.pdf), **n.d.**
- [403] H. Schägger, *Nat Protoc* **2006**, *1*, 16–22.
- [404] R. B. Kapust, J. Tözsér, T. D. Copeland, D. S. Waugh, *Biochem Biophys Res Commun* **2002**, *294*, 949–955.
- [405] V. Paraskevopoulou, F. Falcone, *Microorganisms* **2018**, *6*, 47.
- [406] J.-B. Li, S. Tang, J.-S. Zheng, C.-L. Tian, L. Liu, *Acc Chem Res* **2017**, *50*, 1143–1153.
- [407] J.-S. Zheng, M. Yu, Y.-K. Qi, S. Tang, F. Shen, Z.-P. Wang, L. Xiao, L. Zhang, C.-L. Tian, L. Liu, *J Am Chem Soc* **2014**, *136*, 3695–3704.
- [408] J.-S. Zheng, Y. He, C. Zuo, X.-Y. Cai, S. Tang, Z. A. Wang, L.-H. Zhang, C.-L. Tian, L. Liu, *J Am Chem Soc* **2016**, *138*, 3553–3561.
- [409] S. Tang, C. Zuo, D.-L. Huang, X.-Y. Cai, L.-H. Zhang, C.-L. Tian, J.-S. Zheng, L. Liu, *Nat Protoc* **2017**, *12*, 2554–2569.

- [410] L. P. Miranda, W. D. Meutermans, M. L. Smythe, P. F. Alewood, *J Org Chem* **2000**, *65*, 5460–8.
- [411] S. Arold, P. Franken, M.-P. Strub, F. Hoh, S. Benichou, R. Benarous, C. Dumas, *Structure* **1997**, *5*, 1361–1372.
- [412] J. M. Martin-Garcia, I. Luque, J. Ruiz-Sanz, A. Camara-Artigas, *Acta Crystallogr D Biol Crystallogr* **2012**, *68*, 1030–1040.
- [413] A. Camara-Artigas, E. Ortiz-Salmeron, M. Andujar-Sánchez, J. Bacarizo, J. M. Martin-Garcia, *Acta Crystallogr F Struct Biol Commun* **2016**, *72*, 707–712.
- [414] J. Clarke, S. J. Hamill, C. M. Johnson, *J Mol Biol* **1997**, *270*, 771–778.



## **Part B**

# **Structural Elucidation of Signaling Pathways of GPR83**





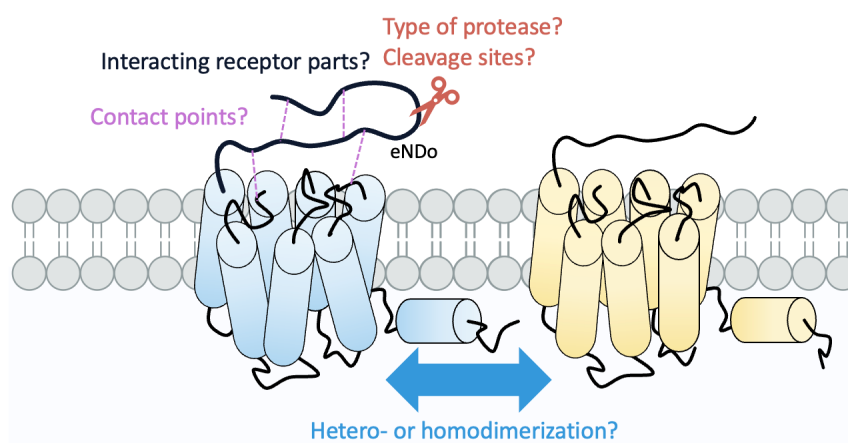
## **Content Part B**

<b>8</b>	<b><i>Motivation</i></b> .....	<b>145</b>
<b>9</b>	<b><i>GPCRs - Fascinating Molecular Switches in Signaling Pathways</i></b> .....	<b>146</b>
9.1	GPCR Signaling .....	147
9.2	Road to <i>De-orphanize</i> GPR83 .....	149
<b>10</b>	<b><i>Peptide Libraries</i></b> .....	<b>155</b>
<b>11</b>	<b><i>Aim of the Study</i></b> .....	<b>157</b>
<b>12</b>	<b><i>Results and Discussion</i></b> .....	<b>158</b>
12.1	Design of Peptide Library .....	158
12.2	Challenges in Peptide Synthesis.....	159
12.2.1	Short Sequences (Peptides 7 to 13) .....	159
12.2.2	Difficult Segment Gpr83_50-52 (Ser-Asp-Trp, SDW).....	160
12.2.3	Medium-length Sequences (Peptides 4 to 6) and hPEN.....	166
12.2.4	Synthesis of eNDo (GPR83_18-74) and Long Segments (Peptides 2 & 3) .....	166
12.3	Structural Studies of Synthesized Peptides by CD measurements .....	167
12.4	Biological assays on G <sub>q</sub> and G <sub>i</sub> activation .....	168
<b>13</b>	<b><i>Summary and Outlook</i></b> .....	<b>173</b>
<b>14</b>	<b><i>References</i></b> .....	<b>175</b>



## 8 Motivation

As a multi-key player in various signaling pathways, the G-protein coupled receptors (GPCRs) are important targets in pharmacological research. <sup>[1]</sup> This fact leads to an extraordinary interest in finding natural ligands of GPCRs whose activation and signaling pathways have to be identified, so-called “*orphan receptors*”. An attractive candidate from this group is GPCR 83 (GPR83), which is expressed in areas that regulate neurological and immune processes. This receptor's investigation suggests a diverse influence on the reward system, stress regulation, learning processes, and energy metabolism. <sup>[2]</sup> In preliminary studies, our collaborators Müller *et al.* designed and characterized several artificial constitutively activating mutations in murine GPR83. As a result, they showed that the GPR83 has a G<sub>q</sub> basal signaling. Additional examinations with deletion models indicated a crucial involvement of the extracellular N-terminal domain (eNDo) in activating the GPR83, which might act as an “*inverse agonist*.” <sup>[3]</sup> To further elucidate the role of eNDo in GPR83 activation, the present work will perform structure-activity studies using a systematic peptide library. This peptide library consists of 13 members of varying chain lengths (15 to 56 AAs) with a partial overlap in sequence to pinpoint potentially significant structural regions for the activation of GPR83.

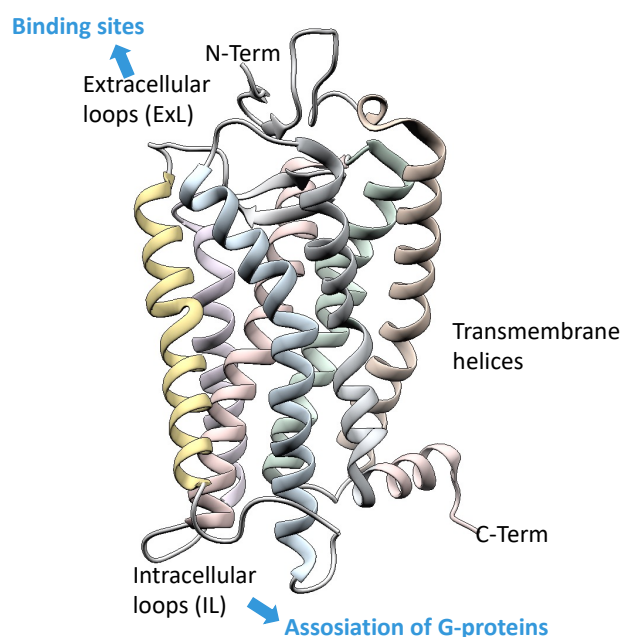


**Figure 8.1:** Schematic representation of the open questions of GPR83 activation. A peptide library will be used to investigate the contact points.

## 9 GPCRs - Fascinating Molecular Switches in Signaling Pathways

G protein-coupled receptors (GPCRs) form the most prominent family of membrane receptors that transmit extracellular stimuli into the cell by various signal transduction pathways and elicit a specific cellular response. Their activators range from ions or small organic molecules to peptides and proteins.<sup>[4]</sup> Given their widespread expression and diverse modes of action, it is hardly surprising that around 40% of marketed drugs target GPCRs as they play a crucial role in different disorders like cancer, diabetes, and Alzheimer's disease.<sup>[5-8]</sup>

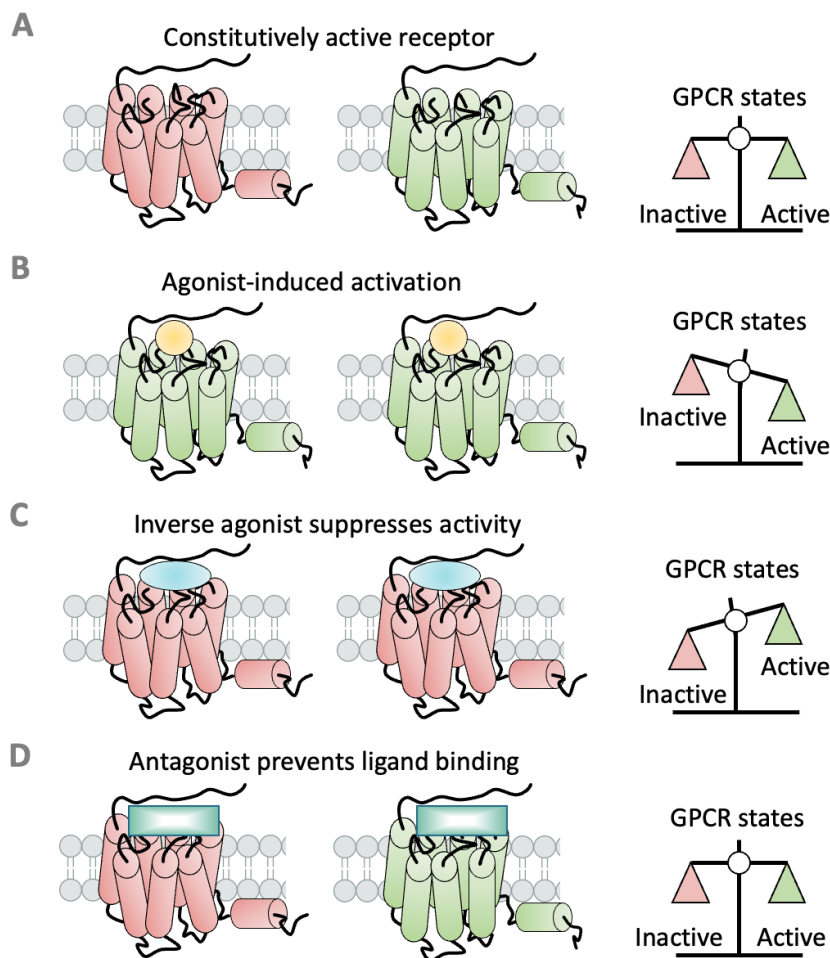
Despite diverse physiologic functions, all GPCRs have a similar architecture of seven transmembrane helices (7TM) combined by three extra- and intracellular loops (Figure 9.1). The extracellular loops shape the ligand-binding pockets, whereas the biggest intracellular loop is associated with a heterotrimeric G-protein.<sup>[9]</sup> Both extra- and intracellular loops and the C- and N-terminal tails can differ considerably, generating a broad receptor recognition for varied ligands. Individual GPCRs may also bind multiple ligands; thus, they have more than one extracellular ligand-binding site. Each of these ligands has a unique effect on the receptor, which changes its conformation in a particular way due to the interaction. Hence these receptors are also referred to as “molecular switches.”<sup>[10]</sup>



**Figure 9.1:** Example of a GPCR. PDB-ID: 1F88.<sup>[11]</sup> Created with <http://www.cgl.ucsf.edu/chimera>.

## 9.1 GPCR Signaling

GPCRs reveal a multiple signaling repertoire. [12] On the one hand, the receptor activity is influenced by the type of ligand (agonist, antagonist, inverse agonist, Figure 9.2) [13,14]; on the other hand, it is also determined by membrane effects (such as the thickness of the membrane) [15,16] or other co-located receptors (e.g., oligomerization behavior). [17,18]

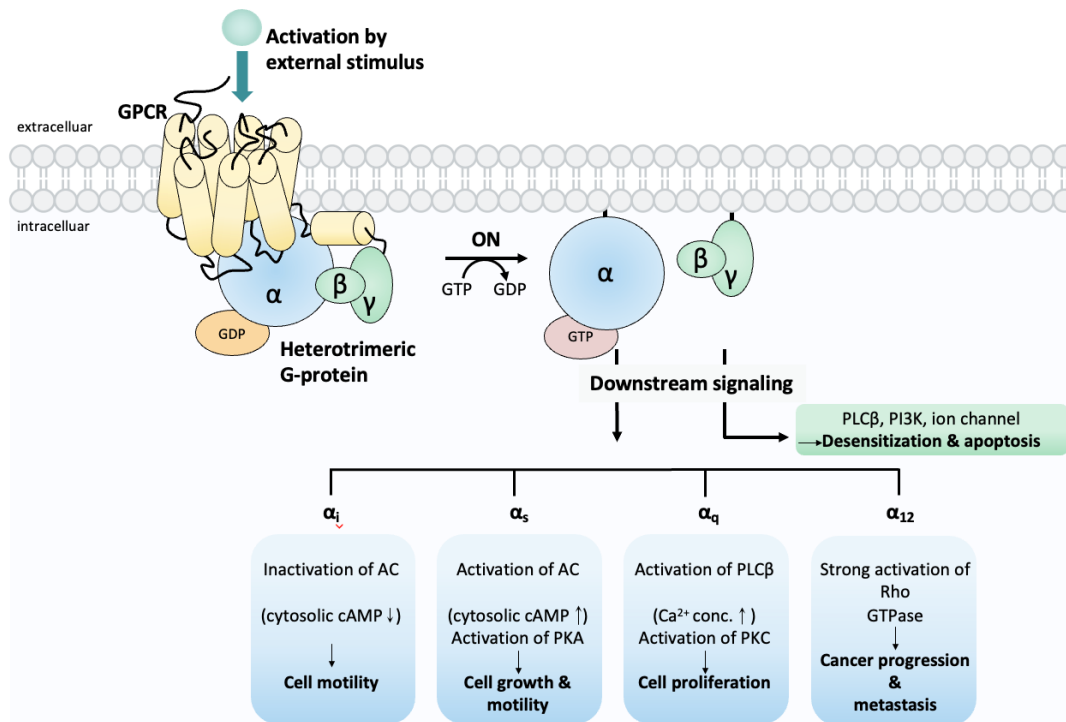


**Figure 9.2:** Modes of activation of GPCRs: **A)** In the classical two-state model, there is an equilibrium between an active (green) and an inactive (red) state. An agonist-independent isomerization can achieve the active state. **B)** Binding an agonist (yellow circle) shifts the balance towards the active state while an inverse agonist (blue ellipse) shifts the equilibrium towards an inactive state **(C).** **D)** An antagonist (dark green square) prohibits the binding of endogenous ligands without interfering with the equilibrium. Adapted from Meye *et al.* and Sato *et al.* [19,20]

Generally, a G-protein is composed of a  $G_{\alpha}$ - and a  $G_{\beta\gamma}$ -dimer subunit. In its inactive form, the G-protein is bound to guanosine diphosphate (GDP). Upon binding an extracellular agonist, a GPCR is activated; followed by the conformational shift, GDP is released and replaced by a GTP resulting in a dissociation of the heterotrimeric G-

protein into the GTP- $G_{\alpha}$ - and the  $G_{\beta\gamma}$ -dimer subunit (Figure 9.3). Both subunits are reported to regulate the activity of varied downstream effectors in different pathways. [6,21] In the following section, the pathways of the  $G_{\alpha}$  subunits will be briefly introduced:  $G_{q/11}$ ,  $G_{12}/G_{13}$ , and  $G_s/G_i$  signaling routes. The  $G_{\beta\gamma}$ -dimer subunit of the activated G protein can regulate diverse signaling pathways via ion channels or phospholipase C $\beta$  (PLC $\beta$ ). [22-24] However, this field will not be the subject of the present work.

Activated GTP- $G_{\alpha}$  subunits subsequently bind to effectors proteins such as adenylyl cyclase (AC), RhoGEF, and phospholipase C $\beta$  (PLC $\beta$ ), affecting their activity. [25,26] These, in turn, can impact downstream effectors either directly or through the formation of *secondary messengers*, e.g., cAMP, diacylglycerol (DAG), and inositol-1,4,5-trisphosphate (IP3). The signal transduction cascade continues then with further downstream effectors such as protein kinase A (PKA) and protein kinase C (PKC), resulting in an influence on different cell processes.



**Figure 9.3:** GPCR signaling pathways. [27] A conformational change of the receptor caused by ligand binding initiates GPCR signaling. In the inactive state, GPCRs interact with a heterotrimeric G protein formed by a GDP-bound  $G_{\alpha}$ -subunit and a  $G_{\beta\gamma}$ -subunit. An external stimulus initiates an exchange of GDP with GTP and dissociation of the  $G_{\beta\gamma}$  subunit, leading to a variety of biological regulations. There are four  $G_{\alpha}$  subclasses. Each of them targets different downstream effectors such as protein kinase A (PKA), phospholipase C $\beta$  (PLC $\beta$ ), protein kinase C (PKC), and Rho-GTPase influencing different physiological processes, e.g., cell growth or cell proliferation. Desensitization and apoptosis are associated with the signaling through the  $G_{\beta\gamma}$  subunit. Adapted with permission from Jo *et al.* [27] Copyright © 2016, The Author(s).

If a  $G_{\alpha}$ -protein belongs to the  $G_s$  (stimulatory) or the  $G_i$  family (inhibitory), the first effector molecule of the signaling cascade is AC. Anchored in the inner cell membrane, this enzyme facilitates the formation of cyclic AMP (cAMP) from abundant ATP in the cytosol. Its activity is regulated by binding the  $G_{\alpha}$ -subunit of the activated G-protein; while  $G_i$ -proteins inhibit the catalytic ability of AC,  $G_s$ -proteins promote enzyme activity. Respectively, this leads to either a reduction or an increase in the intracellular cAMP concentration. [10]

Propagation of a stimulus through the  $G_{q/11}$  signaling pathway involves the activation of phospholipase C (PLC), a phosphodiesterase that catalyzes the cleavage of unsaturated fatty acids of phospholipids. Activated PLC hydrolyzes the phosphoinositol 4,5 bisphosphate anchored in the cell membrane into *second messengers* 1,2-diacylglycerol (DAG), which stays at the membrane and into inositol 1,4,5-triphosphate (IP3) that is released into the cytosol. These molecules trigger different signaling cascades onward in parallel, which will be introduced in the designed biological assays (see experimental section). Another result of PLC activation is an increased cytosolic  $Ca^{2+}$  concentration that plays a vital role in multiple cellular processes.

The  $G_{12/13}$  activation is not extensively characterized as the previously described mechanisms. It is known that RhoGTPase and nucleotide exchange factors (RhoGEFs) are among the effector molecules of this activation. These, in turn, affect RhoA, a key player in the regulation of actin cytoskeleton, cell shape, cell polarity, microtubule dynamics, membrane transport pathways, gene transcription, cell adhesion, cell migration, neurite extension/retraction, and cell growth. [28,29]

## 9.2 Road to *De-orphanize* GPR83

For about 100 members of GPCRs, there is no disaggregated activation mechanism or ligand known; they are referred to as so-called *orphan* receptors. [30] This group is the focus of pharmaceutical research, as it holds considerable promise for drug discovery programs. [6,7,31] GPR83 belonged to this group when the present studies began. [32,33]

GPR83 (also known as GPR72, JP05, or glucocorticoid-induced receptor (GIR)) is a 423 amino acid-long protein (48 kDa) belonging to the rhodopsin-like family but also reveals a high sequence similarity with members of the Neuropeptide Y

receptor family. [34,35] Complete sequence elucidation was performed in 1991 by Harrigan *et al.* [36] So far, four isoforms have been identified in mice. Two of them exhibit significant changes around the second intracellular loop, suggesting that this receptor may target different G-proteins and thus various signaling pathways. In the human genome, only isoform-1 is found. [36,37] The receptor is expressed in significant brain regions involved in food intake, stress regulation, learning, and energy metabolisms, such as the hippocampus, amygdala, and hypothalamic region in humans and mice. To a lesser extent, it is also present in the spleen, thymus and immune cells like CD4+ T-cells. [35,38-40]

Due to the absence of ligands, GPR83 was initially investigated using mRNA expressions. Expression in the hippocampus suggests modulation of memory-related processes such as spatial orientation. [41,42] Interestingly, not only was this receptor identified in different brain regions, but it was also discovered that it plays a crucial role in which sub-neuronal cells it is located. For example, GPR83 is present in the arcuate nucleus and preoptic region of the hypothalamus in neurons expressing AgRP and GSHR. Both receptors are, in turn, responsible for the release of NPY and GABA, regulating hormones such as leptin, insulin, and ghrelin which affect food intake. An interaction with these receptors suggests GPR83 may be a new player in body weight regulation. On the other hand, GPR83 is also present in warm-sensitive neurons in the nucleus arcuatus and thus could influence core body temperature and fever. [42] [43]

Investigations on the expression of the GPR83 also showed its dependence on the availability of nutrients. Under fasting conditions, expression was reduced but increased after resuming food intake. In knock-out (ko) mice, Tschöp *et al.* demonstrated that diet-induced obesity is prevented, identifying GPR83 as a potential new element in body weight regulation. [44] In additional examinations of GPR83 ko mice, Vollmer and co-workers described a reduction in stress-induced anxiety and delayed spatial learning in the Morris water maze. [45]

Research on the oligomerization behavior of GPR83 was also performed. Here, both homodimerization and heterodimerization with the GHSR receptor were observed. The latter showed an influence on the regulatory capacity of the GHSR through reduced uptake of its endogenous ligand acyl-ghrelin. However, these experiments



also found evidence that GPR83 also has a GHSR-independent activation mode as GPR83 is collocated with agouti-related protein (AgRP). [44,46]

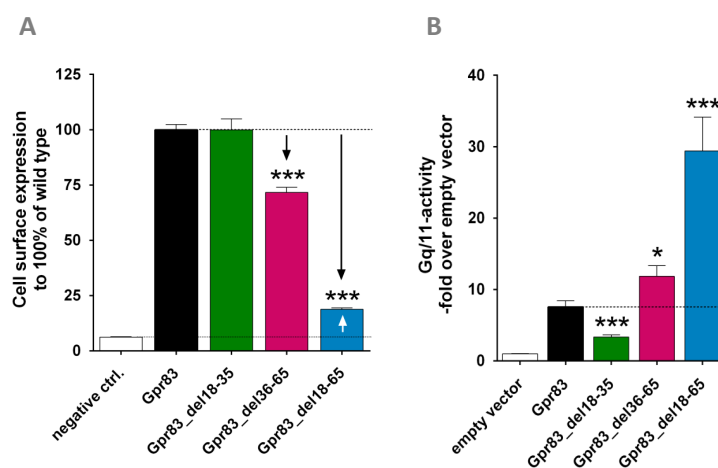
Several investigations on the signaling properties of GPR83 indicate various modes of activation for GPR83. Preliminary experiments focused on the basal activity of GPR83, which is the signaling in the absence of an endogenous ligand. This activity was particularly evident in the G<sub>q</sub> signaling pathway as increased basal levels of IP3 accumulation in combination with the presence of GPR83 were observed. An enhancement of basal activity was demonstrated by treatment with zinc(II), indicating zinc(II) may play an important role in receptor activation, even referred to as an endogenous "molecule" by the authors, since neither Ca<sup>2+</sup> nor Mg<sup>2+</sup> caused the activation. [47] These studies revealed only activity in the G<sub>q</sub> pathway, not in G<sub>i</sub> or G<sub>s</sub> as no altered cAMP levels were observed.

Martin *et al.* reported high constitutive receptor activity for GPR83, an activity in the absence of a ligand caused by a spontaneous isomerization of a GPCR. [34,35] It was hypothesized that the receptor could be activated by proteolytic cleavage of an inhibitory domain. Therefore, eNDo (consisting of about 70 AAs) became the focus of further examinations.

Comparing the sequence across multiple species showed that the eNDo could be divided into two sections (Figure 9.4). While section 36-67 is highly conserved and contains nearly identical regions among orthologs, a wide range of different amino acids is present in section 18-35. To understand the influence of the large eNDo on GPR83's activity, three eNDo deletion constructs were generated and analyzed (Figure 9.5). In all deletion model designs, the signal peptide was preserved, and an HA-Tag was introduced to examine the cell surface expression by enzyme-linked immunosorbent assay (ELISA). [3]

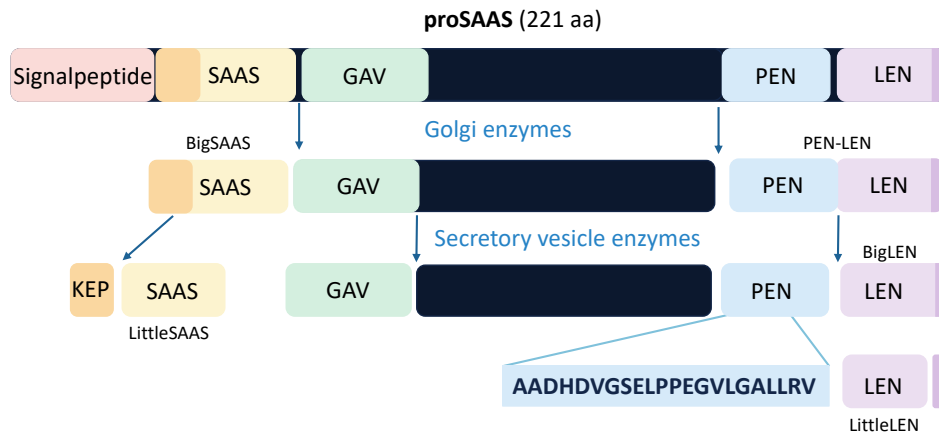


Removal of both the less and highly conserved regions of eNDo (GPR83\_del18-65) resulted in a fourfold increase in basal  $G_q$  activity (Figure 9.6 B). In contrast, removing the highly conserved portion (GPR83\_del36-65) resulted in a 1.4-fold increase in activity. Yet, receptor activity decreased to 60% when the 18-65 region was removed, indicating an inverse agonistic effect of eNDo. This also highlights a possible link between extracellular activation and proteolytic cleavage.  $G_s$ ,  $G_i$ ,  $G_{12/13}$ , or MAPK signaling were also examined, but no activation was detected by the influence of eNDo. However, single-point mutations in the TMH7 region showed increased basal MAPK activity. These findings support the hypothesis that GPR83 has multiple modes of activation.



**Figure 9.6:** Results of deletion model studies. **A)** Cell expression levels determined by ELISA and **B)**  $G_q$  activity investigated by IP3-assay. Reprinted with permission from Müller *et al.*<sup>[3]</sup> Copyright © 2014, Müller et al.; licensee BioMed Central.

In the further course, the human sequence became the center of the investigations, as this is where the therapeutic benefits are targeted. In 2016, Gomes *et al.* identified PEN as the putative endogenous ligand for human and mouse GPR83 by investigating neuropeptides derived from proSAAS (Figure 9.7), a 221 AA long protein that is very prominent in several brain regions associated with GPR83.<sup>[46]</sup> These studies discovered a new receptor in co-localization with GPR83: GPR171. It is noteworthy that both receptors are activated by different peptides derived from proSAAS: BigLEN, not PEN, activates GPR171. Interestingly, the addition of protease inhibitors played a crucial role in the study by Gomes *et al.*, indicating that receptor's activation depends on proteases.



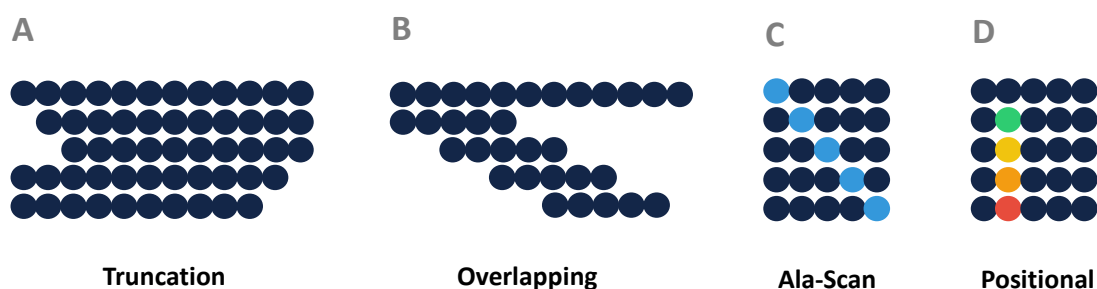
**Figure 9.7:** proSAAS and its neuropeptides. Adapted with permission from Mack *et al.*<sup>[48]</sup> Copyright © 2019, American Chemical Society

Despite these findings, the GPR83 was still considered an *orphan* when our study began. Our main focus was to pinpoint the regions of eNDo that are responsible of GPR83's activation. We decided to include PEN in our investigation as we were looking to find a better elucidation of the activation mechanism of GPR83 using a peptide library.

## 10 Peptide Libraries

With the discovery of Merrifield's solid-phase peptide synthesis (SPPS), peptide libraries have become an indispensable tool for protein-related studies. Peptides are used as mimetics for protein binding sites, structural studies of protein-protein interactions, epitope mapping, and drug discovery. [49–52] The latter plays a vital role in the *de-orphanization* of GPCRs as peptidic endogenous ligands were found for some *orphan* GPCRs. In addition, about 50 GPCR peptide-based drugs have been approved to date. [53] Thus, a well-defined peptide library provides a reasonable basis for exploring new ligands. There are several ways to design a peptide library. An overview is shown in Figure 10.1. Typically, a peptide library contains an extensive collection of peptides with a systematic AA combination. In contrast to the phage display approach, peptides synthesized on solid support offer the advantage of incorporating unnatural AA.

A truncated library is characterized by the systematic removal of AA at the N- or C-terminus. In this way, the minimum peptide length for optimal activity can be identified.



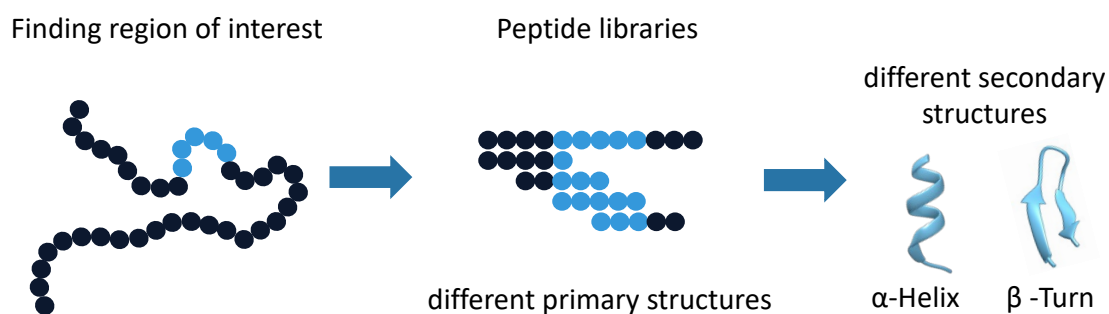
**Figure 10.1:** Peptide library strategies. Adapted from [54–56]

Overlapping peptide libraries (Figure 10.1B) cover the entire sequence of a protein of interest in short overlapping fragments. These peptides usually have the same sequence length but different primary structures resulting in different secondary structures (Figure 10.2). An optimal length for the fragment design is between 6 and 20 AA with offsets of about one-third of the peptide length. These libraries are mainly used to analyze linear binding segments, e.g., for epitope mapping.

In an Ala-Scan (Figure 10.1 C), one position is systematically substituted with alanine. As a result, the contribution of a single AA to the function, activity, or stability of a peptide or protein structure can be determined. Additionally, conclusions can be drawn about its role in the overall structure of a protein. Ala is

preferably chosen due to its chemically inert nature, and it is the smallest chiral AA with almost no interactions within the peptide backbone. [57]

Replacing one AA with different AA at a specific position is achieved using a positional peptide library. Here, the effects of various amino acid residues on the peptide properties, such as activity, binding, and stability, can be studied. In this way, the leader sequence of a peptide can be optimized.

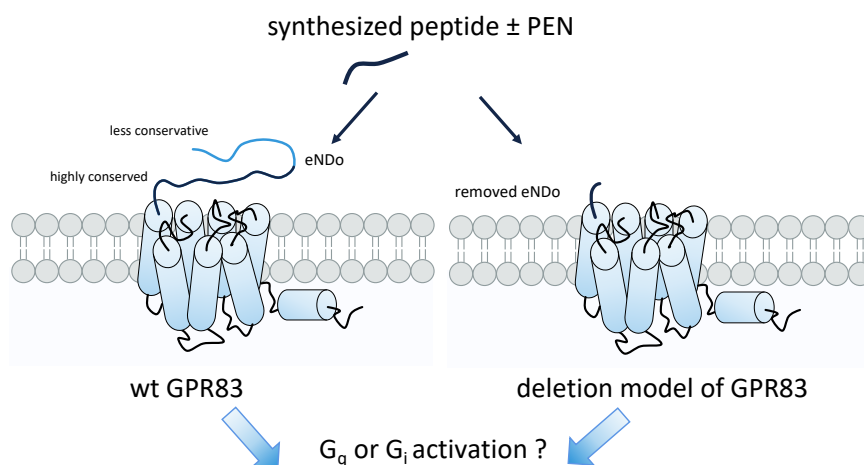


**Figure 10.2:** Localization for linear binding sites by scanning with peptide libraries. The different primary structures of synthesized peptides may reveal different secondary structures. Secondary structures generated with <http://www.cgl.ucsf.edu/chimera>.

## 11 Aim of the Study

GPR83 was reported to be involved in different regulations like stress, anxiety, reward, learning as well as energy metabolism and food intake. However, its biological role remains to be elucidated due to a poor understanding of its activation. So far, different modes of action of the receptor have been published.<sup>[42]</sup> In particular, the discovery by Gomes *et al.* of neuropeptide PEN as a putative endogenous ligand highlights the need to decipher the detailed mechanism of the receptor.<sup>[46]</sup> Our collaborator's studies investigated the very long extracellular N-terminal domain (eNDo) role for the first time.<sup>[3]</sup> Here, deletion models in the eNDo region showed different influences on the receptor expression levels and activity. In the present work, the impact of eNDo will be investigated in more detail using a peptide library. The peptide library is constructed of 13 members. In short fragments, the whole eNDo will be mapped using an overlapping strategy. The fragments consist of about 15 AA with an offset of 5 AA. Special focus will be given to the highly conserved region (36-67), so several peptides containing parts of the first TM helix section will be produced here. Longer peptide sequences will provide insight into the possible influence of the secondary structure of the eNDo. Finally, the previously deleted part of the eNDo (GPR83\_18-74) will be synthesized using SPPS.

The individual members of the peptide library will each be studied in terms of their structure and activity of the receptor when analyzed in bioassays in the  $G_q$  and  $G_i$  signaling pathways. Furthermore, the peptides' co-supplementation with PEN will be investigated for their role as possible competitors or inhibitors.



**Figure 11.1:** Schematic representation of planned biological investigations.

## 12 Results and Discussion

First, the preparation of the peptide library will be discussed, followed by an analysis of the secondary structure of the individual peptides by circular dichroism (CD). Finally, their influence on the activation of GPR83 is investigated in bioassays.

### 12.1 Design of Peptide Library

In the present work, the overlapping peptide library strategy was adapted. For this purpose, the complete sequence of eNDo was divided into fragments of about 15 AA, with an overlap of 5 AA. Particular attention was given to the highly conserved region (36-67), which is why parts of the first TM helix were also included. These short fragments (peptides **7** to **13**) are intended to narrow down the part responsible for receptor activity. In order to also examine the effects that a secondary structure of the peptides might cause, the eNDo was partitioned into larger fragments (peptides **2** to **6**). In addition, the entire deleted part of eNDo (peptide **1**) was to be reconstructed using SPPS. The respective sequences are shown in Table 12.1.

**Table 12.1:** Sequences of the planned peptide library.

peptide	segment in eNDo	sequence
<b>Peptide 1</b>	18-74	EPHEGRADEQSAEAALAVPNASHFFSWNNYTFSDWQNFVGRRRRYGAESQNP TVKALL
<b>Peptide 2</b>	18-45	Abz-EPHEGRADEQSAEAALAVPNASHFFSWN
<b>Peptide 3</b>	46-74	Abz-NYTFSDWQNFVGRRRRYGAESQNPTVKALL
<b>Peptide 4</b>	18-36	Abz-EPHEGRADEQSAEAALAVP
<b>Peptide 5</b>	37-55	Abz-NASHFFSWNNYTFSDWQNFV
<b>Peptide 6</b>	56-74	Abz-VGRRRRYGAESQNPTVKALL
<b>Peptide 7</b>	18-32	Abz-EPHEGRADEQSAEAA
<b>Peptide 8</b>	28-42	Abz-SAEAALAVPNASHFF
<b>Peptide 9</b>	38-53	Abz-ASHFFSWNNYTFSDW
<b>Peptide 10</b>	48-62	Abz-TFSDWQNFVGRRRRYG
<b>Peptide 11</b>	53-66	Abz-QNFVGRRRRYGAESQ
<b>Peptide 12</b>	57-69	Abz-GRRRYGAESQNPT
<b>Peptide 13</b>	58-74	Abz-RRRYGAESQNPTVKALL
<b>hPEN<sup>a</sup></b>	---	AADHDVGSELPPEGVLGALLRV

a) hPEN was described as the putative ligand for GPR83 activation. <sup>[46]</sup>

During our studies, Gomes *et al.* identified neuropeptide hPEN as the endogenous ligand for GPR83. <sup>[46]</sup> It was decided to include this peptide in our studies to



investigate our synthesized peptides for their role as possible competitors or inhibitors of this hPEN.

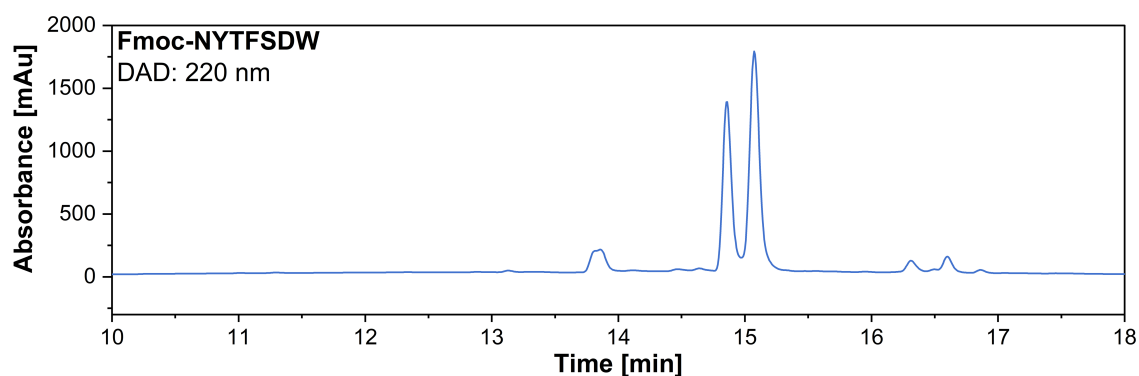
## 12.2 Challenges in Peptide Synthesis

All peptides were synthesized according to the Fmoc-strategy of SPPS. Shorter sequences were obtained by manual amino acid coupling, while the longer sequences were generated with automated synthesizers. After synthesis, the peptides were cleaved from the solid support by TFA treatment, purified by high-performance liquid chromatography (HPLC), and characterized by ESI-ToF mass spectrometry.

### 12.2.1 Short Sequences (Peptides 7 to 13)

These peptides focus on mapping the entire eNDo with an overlapping strategy. Except for peptide **10**, all peptides were synthesized at the 0.05 mmol scale using a NovaSyn TGA resin preloaded with the first Fmoc-protected amino acid (AA). For peptide **10**, a Fmoc-Gly-TGT resin was applied. 2-Amino benzoic acid (Abz, Z) was incorporated into the sequence of the short fragments as a chromophore for concentration determination with UV/Vis spectroscopy for structural analysis.

The coupling of AA was performed with eight equivalents and the use of HOBT and DIC as coupling reagents at room temperature. For each synthesis, a micro-TFA cleavage was performed after the seventh AA to monitor quality during peptide assembly. Peptides **7**, **8**, **11**, **12**, and **13** were successfully synthesized. However, the synthesis of peptide **9** (Gpr83\_38-52, ZASHFFSWNNYTFSDW) was difficult. The first challenge was observed in the micro-cleavage of peptide **9** at position Asn10. A double peak at 15 min was revealed in the HPLC chromatogram (Figure 12.1). Studies with ESI-ToF confirmed an identical m/z for both samples.

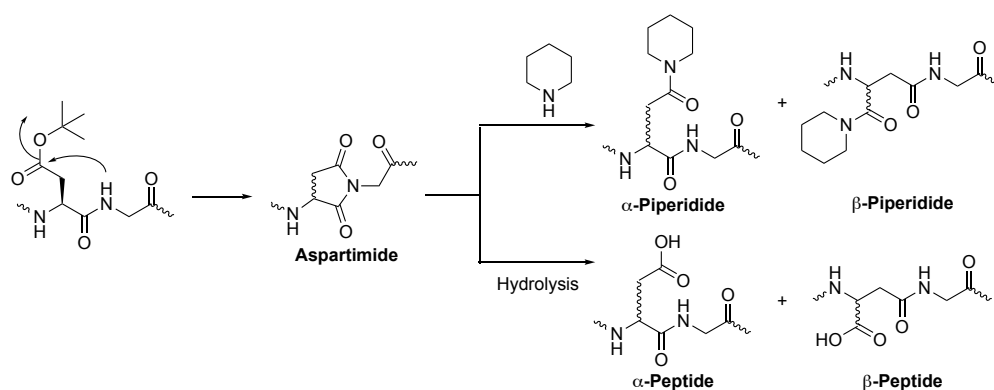


**Figure 12.1:** HPLC chromatogram of micro cleavage of peptide **9** (Fmoc-NYTFSDW) after seven AA. HPLC system: LaChrom ELITE®; column: Kinetex® C18 column, eluent: A = H<sub>2</sub>O, B = ACN, both containing 0.1% (v/v) TFA; gradient: 5% → 70% B over 18 min. The chromatogram is depicted from 12 to 20 min.

### 12.2.2 Difficult Segment Gpr83\_50-52 (Ser-Asp-Trp, SDW)

#### *Aspartimide Formation: Influence of Fmoc-removal Conditions*

In the absence of HOBt in the Fmoc-removal solution, it was assumed that aspartimide formation had occurred. This is the most common side reaction during the incorporation of Asp into a sequence that produces a cyclic by-product (Scheme 12.1). Detailed information on this side reaction is provided in section 5.4.2 in part A.



**Scheme 12.1:** Reminder: mechanism of aspartimide formation and the resulting side products.<sup>[58]</sup>

Therefore, it is required to implement some countermeasures to suppress this side reaction during peptide synthesis. An additive for Fmoc-removal is crucial to alleviate aspartimide formation.<sup>[59,60]</sup>

Peptide **9** was synthesized again to Asn10 with 0.1 M HOBt in the Fmoc-cocktail. Nevertheless, a double peak was observed in the HPLC analysis. For further studies (Table 12.2), a tripeptide (SDW) was used as a model system to find the best conditions for Fmoc removal on a 0.005 mmol scale.

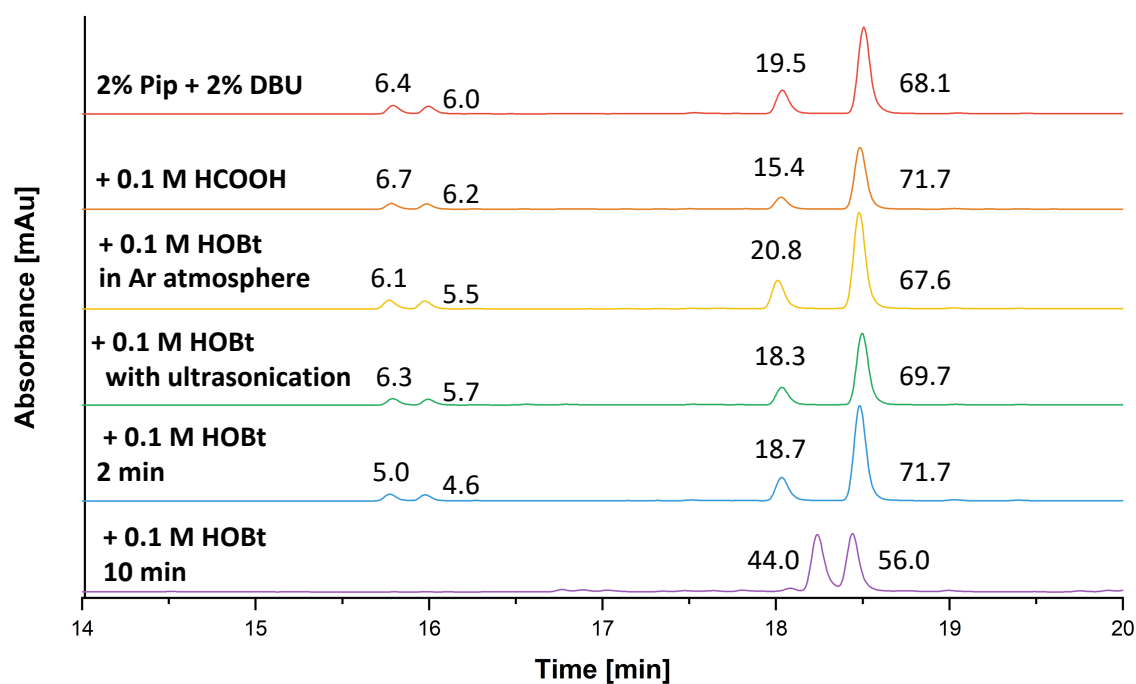
**Table 12.2:** Examined conditions for Fmoc removal.

entry	Fmoc-removal conditions <sup>a</sup>
<b>1</b> <sup>b</sup>	20% Pip, 3 x 10 min
<b>2</b>	20% Pip + 0.1 M HOBT, 3 x 10 min
<b>3</b>	20% Pip + 0.1 M HOBT, 3 x 2 min
<b>4</b>	20% Pip + 0.1 M HOBT, 3 x 10 min (ultra-sonication)
<b>5</b>	20% Pip + 0.1 M HOBT 3 x 10 min (Ar atmosphere)
<b>6</b>	20% Pip + 0.1 M formic acid, 3 x 10 min
<b>7</b>	2% Pip + 2% DBU, 3 x 10 min

a) all reactions were performed in 1 mL DMF at rt, b) sequence: NYTFSDW rest is with SDW.

First, different aspects of that might affect the kinetics during Fmoc-cleavage were examined. Therefore, the reaction time was shortened to two instead of ten minutes (**entry 3**), or ultrasonication was used (**entry 4**). Inert conditions were also investigated to exclude any racemization that might have occurred under an oxygen atmosphere (**entry 5**). Since base-mediated aspartimide formation is initiated by deprotonation of peptide amide backbone, the choice of base and the application of 0.1 M formic acid as an additive to the Fmoc-removal cocktail were explored. Formic acid was used following Mier *et al.* report on positive effects in reducing aspartimide formation (**entry 6**). DBU as a base for Fmoc cleavage was tested since DBU is a non-nucleophilic base, which is why possible cyclization might be avoided (**entry 7**).

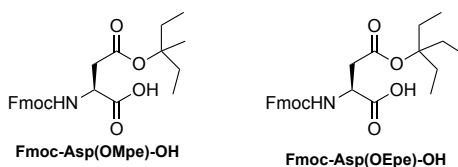
After synthesis, a micro TFA cleavage was performed, and the obtained tripeptides were characterized through HPLC and ESI-ToF analysis. However, all examined procedures had the previously described outcome of a double peak at 18 min (Figure 12.2), but with a shifted ratio that was determined by calculation of the area%. All chromatograms displayed another set of double peaks at 16 min, which was assigned to the corresponding piperidinyll adducts. The best results were achieved using shorter reaction times and formic acid as an additive (**entry 3** and **6**). Despite a clear improvement in product formation, it was still impossible to distinguish between the two formed species. Moreover, the detected piperidinyll adducts in these model studies suggest that further measures needed to be investigated to suppress the side reaction.



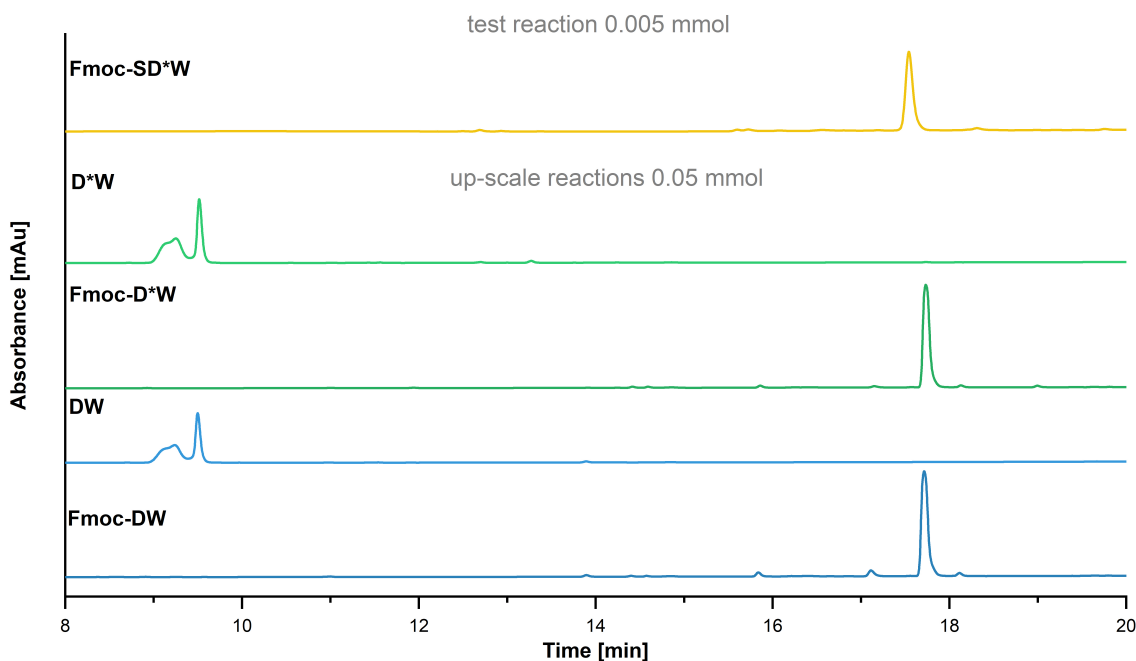
**Figure 12.2:** Results of modifying Fmoc-removal procedure. HPLC chromatograms show traces of crude peptides after a micro cleavage of Fmoc-NYTFSDW (purple line) and Fmoc-SDW under different Fmoc cleavage conditions. The shown numbers represent the calculated area% of the peaks. HPLC system: LaChrom ELITE®; Kinetex® C18 column; eluents: A = H<sub>2</sub>O, B = ACN both containing 0.1% (v/v) TFA; gradient: 5% → 70% B over 18 min. The chromatogram is depicted from 14 to 20 min.

### ***Aspartimide formation: Impact of Side-chain Protecting Groups***

Another way to address aspartimide formation is the bulkiness of the Asp side-chain protecting group so that the five-membered ring closure is avoided by sterically demanding groups at the Asp position. Most commonly, Asp(OMpe) or Asp(OEpe) are used for this purpose (Scheme 12.2). Again, a tripeptide sequence (SDW) was used on a 0.005 mmol scale for test reactions. Fmoc-Asp(OEpe)-OH was introduced in a 2 h coupling using 2 eq. of AA, DIC, and HOAt since this AA is expensive. Fmoc-removal was performed with 0.1 M HOBt as an additive for 2 min at rt (as previously established). This approach yielded a significant improvement (Figure 12.3 yellow line). However, this success was not reproducible in a scaled-up reaction as different side products were detected after Fmoc-removal of both Asp(OEpe) (green lines) and Asp(O<sup>t</sup>Bu) (blue lines).



**Scheme 12.2:** Commercially available bulky side-chain protected Asp derivatives that are used to avoid aspartimide formation.

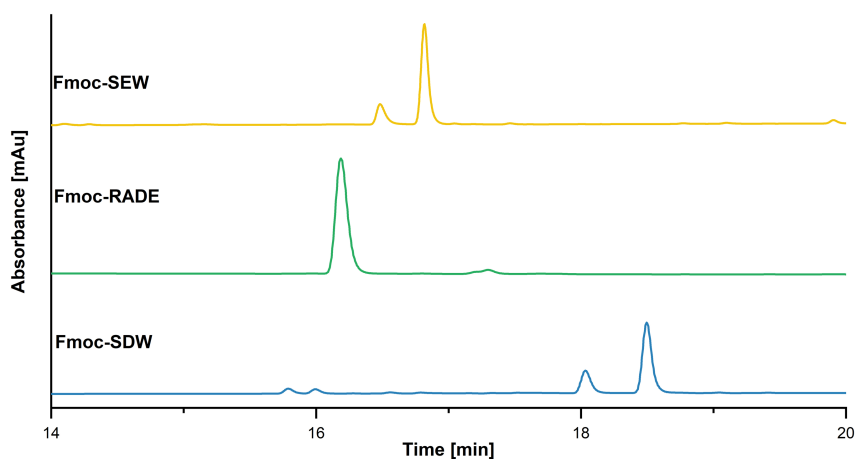


**Figure 12.3:** HPLC chromatograms of micro cleavage of Fmoc-SDW synthesized with  $D^* = \text{Asp}(\text{OEpe})$  in a 0.005 mmol scale (yellow line) and up-scale reactions with (green lines) and without (blue lines) bulky side-chain protecting group. Note: Due to the lack of the Fmoc-protecting group, the absorbance of DW dipeptides is significantly lower.

HPLC system: LaChrom ELITE®; Kinetex® C18 column; eluents: A =  $\text{H}_2\text{O}$ , B = ACN both containing 0.1% (v/v) TFA; gradient: 5%  $\rightarrow$  70% B over 18 min. The chromatogram is depicted from 8 to 20 min.

### ***Aspartimide Formation: Sequence Dependency***

It has been described that the sequence of AA significantly influences the formation of cyclic by-products.<sup>[58]</sup> A negative influence in the sequence Asp-Trp is not yet described in the literature. This aspect was investigated using the GPR83\_24-28 segment consisting of Arg-Ala-Asp-Glu (RADE). Here, no undesired by-product formation was observed with HOBt as an additive and short reaction time in the Fmoc-cleavage cocktail (Figure 12.4 green line). Additionally, Asp was replaced by Glu (yellow line), which also tends to form a cyclic by-product, but the resulting six-membered ring is kinetically less favorable, leading to a smaller amount of aspartimide formation in theory.



**Figure 12.4:** HPLC chromatograms of sequence-dependency studies. The blue line represents the original GPR83\_50-52 segment (SDW), the green line shows GPR83\_24-28 (RADE), and the yellow line corresponds to the replacement of Asp with Glu.

HPLC system: LaChrom ELITE®; Kinetex® C18 column; eluents: A = H<sub>2</sub>O, B = ACN both containing 0.1% (v/v) TFA; gradient: 5% → 70% B over 18 min. The chromatogram is depicted from 14 to 20 min.

All investigations concluded that the by-product is already formed after the coupling of Asp when this is located at the C-term. Moreover, the double peak of Fmoc-SEW (yellow line) indicates that Trp might impact the side product formation. Studies with different coupling reagents, such as HOAt or HATU or shorter coupling duration, did not provide any improvements. The best results for peptide **9** were obtained with the incorporation of Asp at the C-term.

### ***TFA-cleavage***

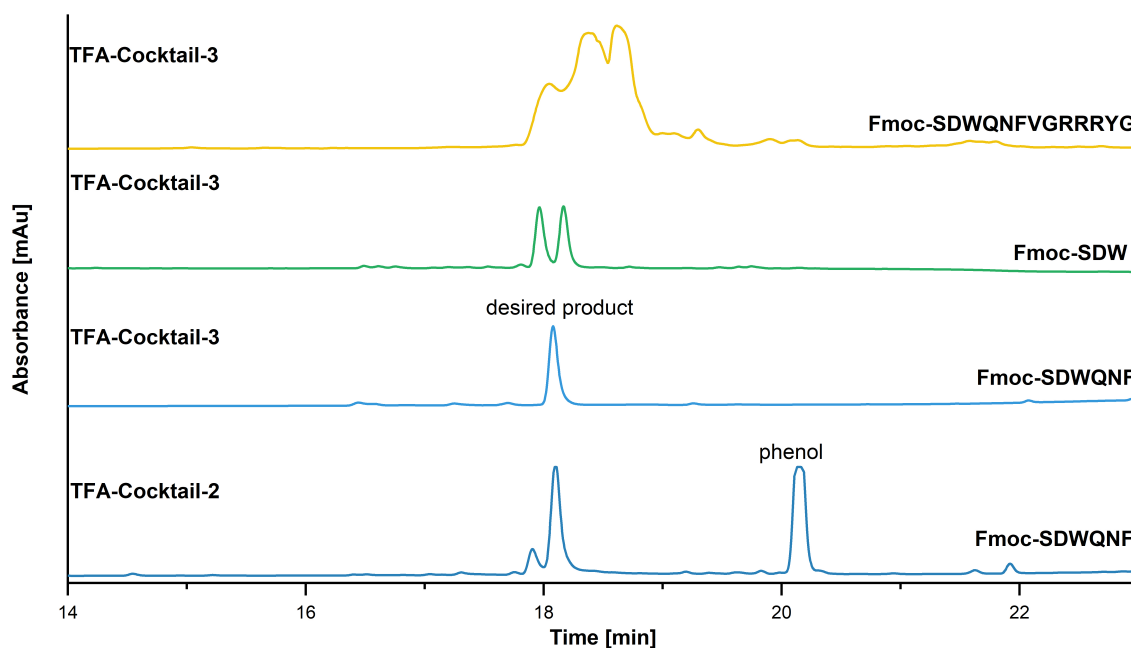
Having thoroughly explored all possibilities of influence during the assembly of the peptide backbone, the focus was put on the conditions of TFA cleavage. Here, three different cleavage cocktails were investigated (Table 12.3).

The first cocktail represents the standard in TFA cleavage. Besides triisopropylsilane (TIS), the other two cocktails contain additional scavengers such as phenol, thioanisole, or 1,2-ethanedithiol (EDT). These are used for complicated sequences, such as Cys or Met-containing peptides. Trp is also partly assigned to this category in the literature and the described cocktails **II** and **III** are often recommended. Interestingly, it was shown that the use of reagent K (cocktail **II**) does not lead to any improvement, while cocktail **III** completely suppresses the formation of the undesired by-product when using segment GPR83\_50-55 (Figure 12.5 blue lines).

**Table 12.3:** TFA cleavage cocktails.

cocktail	composition	ratio in % v/v
I	TFA/TIS/water 95/5/5	95/5/5
II	TFA/EDT/thioanisol/water/phenol 82.4/2.5/5/5/5	82.4/2.5/5/5/5
III	TFA/EDT/water/TIS	94/2.5/2.5/1

A total volume of 200  $\mu$ L was prepared, and the TFA cleavage was performed on a micro-scale for 2.5 h at rt each.



**Figure 12.5:** HPLC chromatograms of different GPR83 segments with different TFA cocktails. The blue lines represent the original GPR83<sub>50-55</sub> segment (SDWQNF) with cocktail 2 (dark blue) and cocktail 3 (light blue), the green line shows GPR83<sub>50-52</sub> (SDW) in which the Asp-Trp section is close to the C-terminus and the yellow line corresponds to a segment (GPR83<sub>50-62</sub>) that is close to the N-terminus.

HPLC system: LaChrom ELITE®; Kinetex® C18 column; eluents: A = H<sub>2</sub>O, B = ACN both containing 0.1% (v/v) TFA; gradient: 5% → 70% B over 18 min. The chromatogram is depicted from 14 to 20 min.

However, no beneficial effect of scavengers was observed for N- or C-terminal Ser-Asp-Trp regions. Based on the studies performed, it was decided to truncate peptides **9** and **10** to avoid the Asp-Trp segment at the C-terminus and in proximity to the N-terminus, resulting in the sequences of peptides **14** and **15** (Table 12.4). Although this decision means the disruption of the overlapping library, this area is covered by peptide **5**, as shown above in Table 12.1. Therefore, the original desire to map the eNDo with the new sequences is still granted.

With this modification, all desired short members of the peptide library could be synthesized in the following yield.

**Table 12.4:** Synthesized short peptides.

peptide	segment in eNDo	sequence	yield [mg]
<b>Peptide 7</b>	18-32	Abz-EPHEGRADEQSAEAA	5.2
<b>Peptide 8</b>	28-42	Abz-SAEAALAVPNASHFF	9.6
<b>Peptide 14</b>	38-51	Abz-ASHFFSWNNYTFSD	7.9
<b>Peptide 15</b>	52-62	Abz-WQNFVGRRRYG	5.6
<b>Peptide 11</b>	53-66	Abz-QNFVGRRRYGAESQ	23.9
<b>Peptide 12</b>	57-69	Abz-GRRRYGAESQNPT	33.5
<b>Peptide 13</b>	58-74	Abz-RRRYGAESQNPTVKALL	17.0

### 12.2.3 Medium-length Sequences (Peptides 4 to 6) and hPEN

These peptides were prepared on a 0.05 mmol scale using automated synthesis with the Activo-P11 synthesizer. Again, predominantly preloaded NovaSyn-TGA resins were used. For a Pro at the C-term in the sequence of peptide **4**, a NovaSyn-TGT resin was chosen. Since no structural investigations were planned with hPEN, no Abz label was introduced in that sequence.

Due to the previous establishments in AA coupling, Fmoc cleavage, and the appropriate composition of the TFA cleavage cocktail, the desired peptides could be produced in the following yields.

**Table 12.5:** Synthesized medium-length peptides and hPEN.

peptide	segment in eNDo	sequence	yield [mg]
<b>Peptide 4</b>	18-36	Abz-EPHEGRADEQSAEAALAVP	12.1
<b>Peptide 5</b>	37-56	Abz-NASHFFSWNNYTFSDWQNFV	10.3
<b>Peptide 6</b>	56-74	Abz-VGRRRYGAESQNPTVKALL	14.6
<b>hPEN</b>	----	AADHDVGSSELPPEGLVLLRV	20.6

### 12.2.4 Synthesis of eNDo (GPR83\_18-74) and Long Segments (Peptides 2 & 3)

The sequence of the eNDo consists of 56 AAs without the signal peptide. It was the first system synthesized using MW-SPPS. This method is introduced in section of Part A. We have examined it to synthesize peptides **2**, **3** (both 28 AA long) and the longest sequence of eNDo (peptide **1**, 56 AAs) (Table 12.6).

**Table 12.6:** Sequences of long peptides obtained by MW-SPPS.

Peptide	Segment in eNDo	Sequence	Yield [mg]
<b>Peptide 1</b>	18-74	EPHEGRADEQSAEAALAVPNASHFFSWNNYTFSDWQNFV GRRRYGAESQNPTVKALL	14.1
<b>Peptide 2</b>	18-45	Abz-EPHEGRADEQSAEAALAVPNASHFFSWN	11.4
<b>Peptide 3</b>	46-74	Abz-NYTFSDWQNFVGRRRRYGAESQNPTVKALL	9.3



Since the synthesis under MW radiation is performed at 90°C. The reaction time was reduced to 2 min for AA coupling and 1 min for Fmoc removal. High temperature, as in MW-assisted SPPS, might enhance aspartimide formation. To reduce this side reaction to a minimum of 10% piperazine in NMP/EtOH (9:1 v/v) with 0.1 M HOBt was used as it is a milder base in comparison with piperidine. Additionally, Fmoc-Asp(OMpe)-OH was used in section GPR83\_52-53 (Asp-Trp). First, both peptides **2** and **3** were synthesized with no difficulties detected. Afterward, the full-length eNDo was successfully prepared.

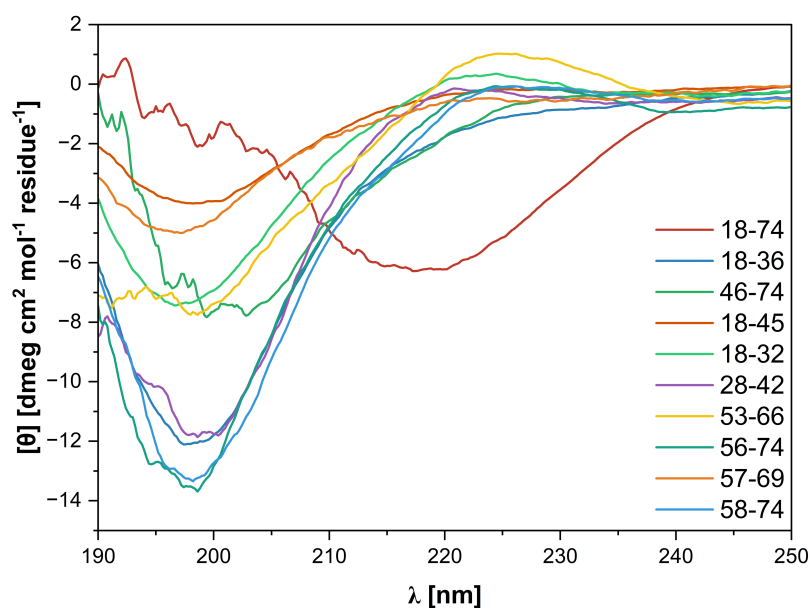
### 12.3 Structural Studies of Synthesized Peptides by CD measurements

Unlike small molecule ligands, peptides can adopt various 3D structures in the form of secondary structures, such as the  $\alpha$ -helix or  $\beta$ -sheet. Such structures are essential for the interactions of peptides with proteins, as in recognizing ligands in the binding pockets of GPCRs.

There are 118 peptide-binding GPCRs described in the literature. Structural elucidations by X-ray crystal structure analysis of 21 GPRCs were used to determine the selectivity as well as similarities in ligand-binding, revealing that predominantly a turn motif is preferred. [4,61]

In the present work, all synthesized peptides mapping the eNDo of GPR83 in different sections and primary structures were analyzed concerning their secondary structure by circular dichroism. A 10 mM phosphate buffer at pH 7.4 and a peptide concentration of 100  $\mu$ M was used. For the determination of the peptide concentration, a calibration curve was first determined using Abz-Gly·HCl. The Abz label was then used to determine the desired peptide concentration. In the case of GPR83\_18-67, the protocol Pace *et al.* was applied as no Abz-label was implemented for this peptide. [62] Detailed information is given in experimental section.

In Figure 12.6 all CD spectra are summarized. Most synthesized peptides have random structure features, except for the 56 AA long eNDo, which shows a  $\beta$ -structure. This observation is in good correlation with described literature for rhodopsin-like receptors, as these receptors have a  $\beta$ -sheet at the N-terminus (Figure 9.1).



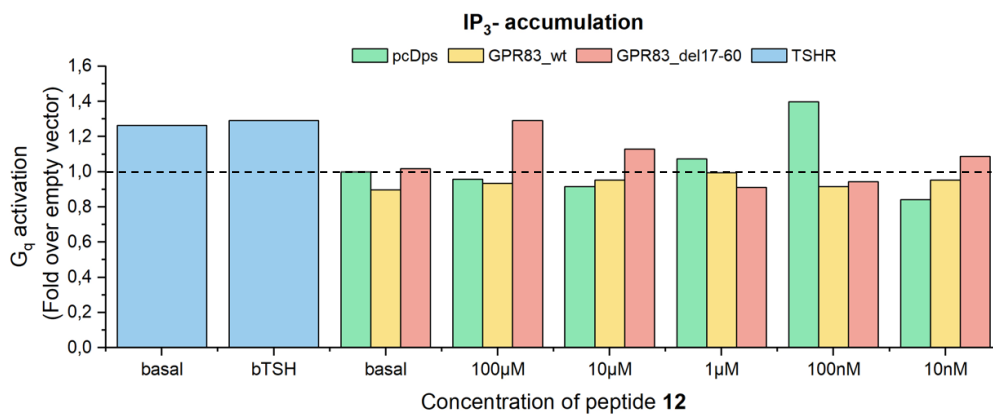
**Figure 12.6:** Measured CD-spectra of peptide library members. The spectra were recorded at pH 7.4 and 37°C in 10 mM phosphate buffer with an overall peptide concentration of 100  $\mu$ M. The depicted spectra are normalized and represent the mean of three independent measurements.

#### 12.4 Biological assays on $G_q$ and $G_i$ activation

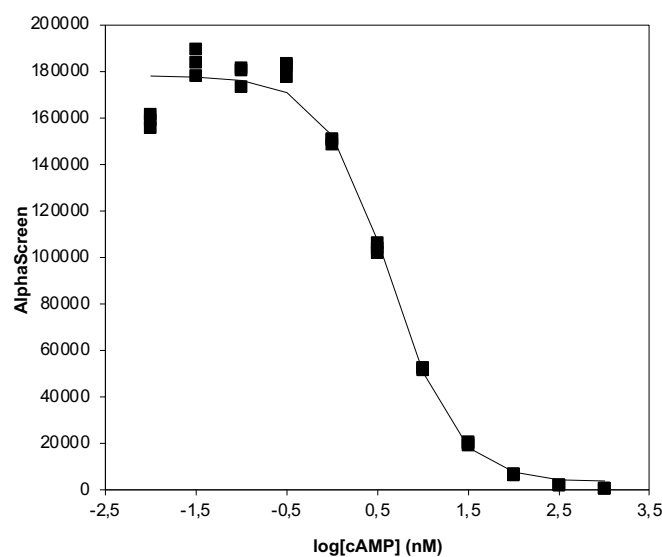
In biological assays, the influence of the synthesized peptides on GPR83's activity was investigated. Here we focused on  $G_q$  and  $G_i$  activation since Gomes *et al.* described a basal activity of GPR83 in these signaling pathways. These experiments were performed at the Institute of Experimental Pediatric Endocrinology (IEPE, Charité Universitätsmedizin Berlin) in the group of Prof. Dr. Heike Biebermann institute (detailed protocols are described in Part C) In our investigation, the embryonic mouse hypothalamus cell line N41 was chosen, representing a neurological cell line that might be more suitable for GPR83 than HEK (human embryonic kidney) or CHO (Chinese hamster ovary) cells. Before performing the experiments, the desired plasmid DNA of the receptors (GPR83-wt and GPR83\_del18-67) was purified following the protocol by the supplier. Then two different assays were conducted with peptide **12** (GPR83\_57-69): a chemiluminescent-based cAMP-assay for  $G_i$  activation, and a reporter-gen (luciferase) assay for  $IP_3$ -accumulation ( $G_q$  activation). Peptide **12** was selected because it covered part of the highly conserved region of eNDo and had already shown an impact on GPR83's activity in previous experiments in HEK cells.

Unfortunately, the biological investigations did not provide the desired outcome. First, it turned out that the initial receptor activity with peptide **12** in HEK cells was

caused by forskolin contamination and was, therefore, not caused by peptide **12**. Second, no precise results were obtained in the IP<sub>3</sub> assay when the N41 cell line was used. The difference between measured activity and positive control was too low to detect an apparent activation by hPEN or peptide **12** (Figure 12.7). For a meaningful evaluation of this assay, the basal values of the positive control (TSHR) would have to be about 20 to 30 times higher than those of the empty vector (pcDps). That is not the case in the present results, as the difference is about 1.4-fold high. Additionally, the cAMP assay revealed no specific receptor activity. Only the standard curve of cAMP was observed (Figure 12.8).

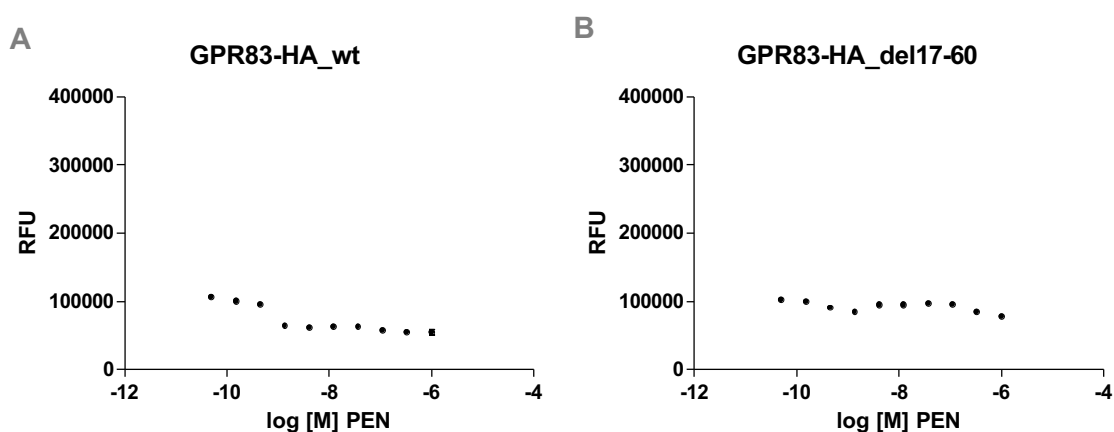


**Figure 12.7:** G<sub>q</sub>-activation measured by IP<sub>3</sub>-accumulation. TSHR with its endogenous ligand bTSH served as a positive control (blue bars) while pcDps was used as an empty vector (green bars). Both wt GPR83 (yellow bars) and a deletion model Grp83\_del17-60 (red bars) were stimulated with peptide **12**.

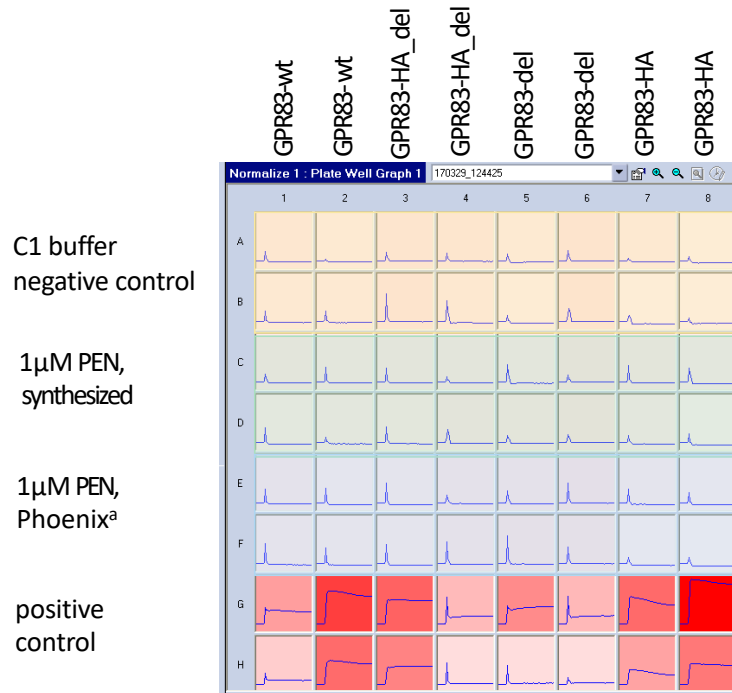


**Figure 12.8:** Results of alpha screen cAMP-assay (G<sub>i</sub> activation).

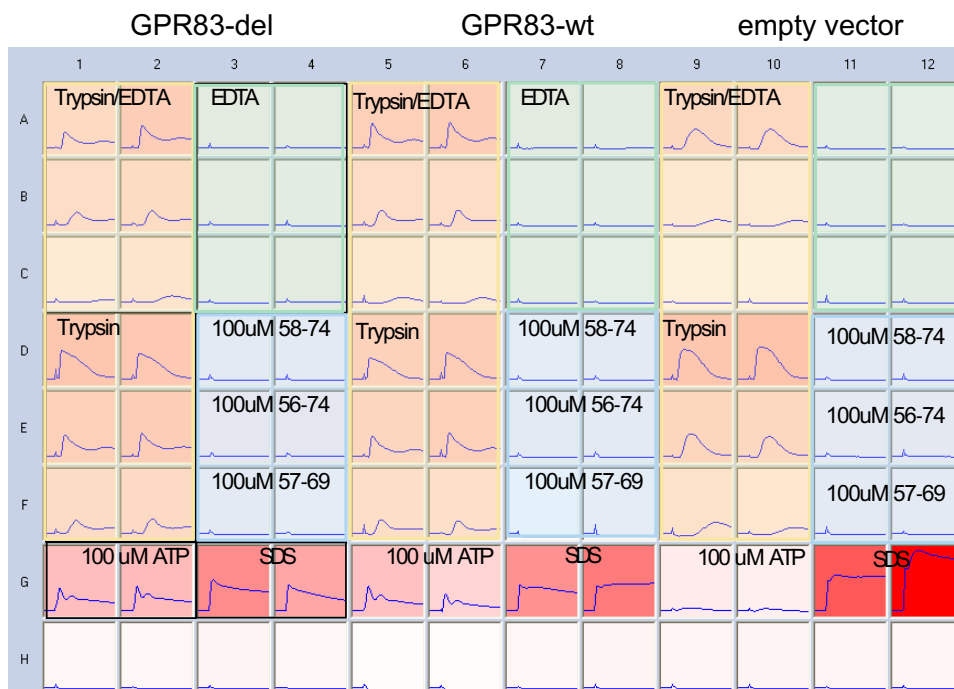
These were the first studies with the N41 cell line, suggesting that individual steps of the assay protocols, such as the duration of transfection, needed further optimization. At this point, however, the project was terminated by the collaborator. A new collaboration partner (Carsten Grötzinger, Research Group Tumor Targeting, Med. Klinik m. S. Hepatology and Gastroenterology, Charité - Universitätsmedizin Berlin) investigated the receptor in HEK293 cells with a variety of literature known assays, different synthesized peptides, and different hPEN sources (both synthesized in the presented work and commercially available). First, a cAMP assay ( $G_i$  activation) was performed with hPEN in both the wt and deletion model of GPR83 (Figure 12.9). Here, no specific activation was found (a positive result would depict a sigmoidal behavior like Figure 12.8). Second,  $G_q$  activation was explored using Calcium imaging assays with different hPEN sources (Figure 12.10 highlighted in yellow and green) and GPR83 receptor mutants with HA-tag and without, various members of the synthesized peptides (Figure 12.11, highlighted in blue: **6** (GPR83\_58-74), **12** (GP83\_56-69) and **13** (GPR83\_56-74), and Co-factors like Trypsin or EDTA (highlighted in yellow or green). A positive control shows a steady signal after treatment with Fluo-4, which is correlated to increased Calcium-concentration caused by  $G_q$ -activation of a GPCR. Treatment with Trypsin or Trypsin/ETDA demonstrated an activation of the receptor (highlighted in yellow Figure 12.11). However, this behavior was also observed in the case of an empty vector indicating that this signal is not a result of a specific response by GPR83.



**Figure 12.9:** Results of cAMP-assays with hPEN stimulation. A) wt-receptor, B) deletion model GPR83-del17-60. RFU = relative fluorescence units. Data kindly provided by the Grötzinger group.

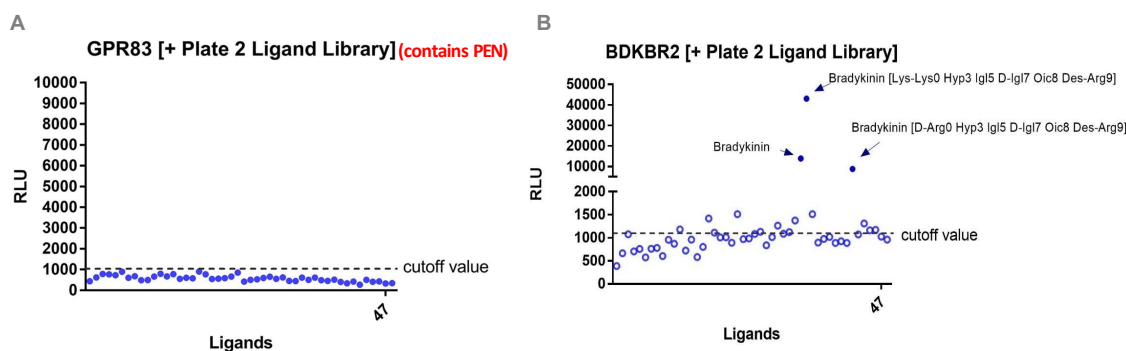


**Figure 12.10:** Calcium-imaging assay. Different GPR83 mutants (wild-type and 18-67deletion model (here referred to as del) with and without HA-tag) were investigated with regard to their  $G_q$ -activity. C1-buffer (130 mM NaCl, 5 mM KCl, 10 mM HEPES, 2 mM  $CaCl_2$ , 10 mM glucose) was used as a negative control (highlighted in yellow), and our synthesized PEN (highlighted in green) in comparison with commercially available PEN (Phoenix Pharmaceuticals Inc.) (highlighted in blue). A positive control (red colored) shows significant receptor activity. a) commercial source Phoenix, USA. Data kindly provided by the Grötzinger group.



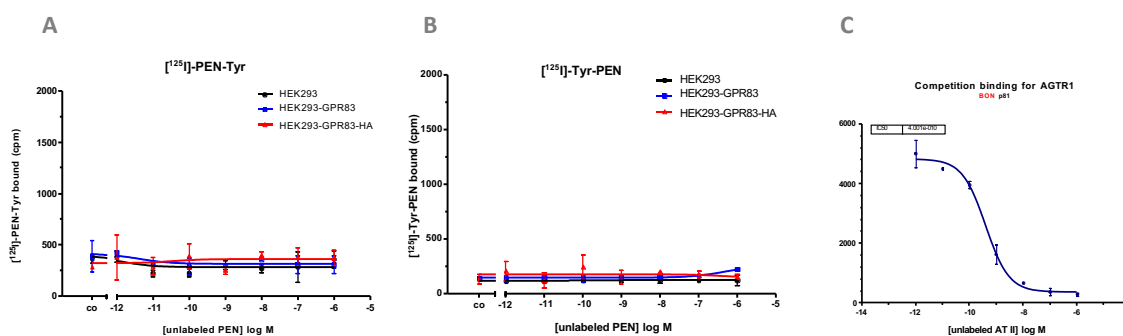
**Figure 12.11:** Calcium imaging assays with different Co-factors. Trypsin/EDTA treatment shows an activation (highlighted in yellow). But no activation with EDTA (highlighted in green) and different synthesized members 6, 12, and 13 of the peptide library (highlighted in blue). Data kindly provided by the Grötzinger group.

Additionally, a commercially available peptide screen (Tango assay) was examined with PEN included as a putative endogenous ligand of GPR83. As a comparison, the same assay was conducted with Bradykinin and its known ligands (Figure 12.12). This assay revealed no significant response by GPR83 towards hPEN or other commercially available ligands.



**Figure 12.12:** Tango-Assay (screening of different commercially available ligands). **A)** GPR83 wt was tested including PEN; **B)** positive control showing activation of Bradykinin-agonists in BRDKR2-cells. Data kindly provided by the Grötzinger group.

Finally, a binding assay was performed (Figure 12.13). For this purpose, hPEN was labeled with  $^{125}\text{I}$ -Tyr at the N- and C-terminus following the Chloramine-T- method. Again, no specific reaction was detected. A positive control is depicted in Figure 12.13 C.



**Figure 12.13:**  $^{125}\text{I}$ -Tyr binding assay. **A)** Treatment with C-terminal labeled hPEN, **B)** N-terminal labeled hPEN, and **C)** positive control of AGTR1 and its ligand. Data kindly provided by the Grötzinger group.

This extensive investigation could not confirm hPEN as an endogenous ligand for GPR83 nor demonstrate any activation by the synthesized peptides. At this point the project was terminated entirely.

### 13 Summary and Outlook

GPR83 is an interesting drug discovery candidate because it is involved in numerous physiological processes. During the present studies, this receptor was *de-orphanized* by Gomes *et al.*, and hPEN was released as a putative ligand. The influence of the extracellular N-terminal domain (eNDo) was nevertheless still of great interest, so we intended to map eNDo with an overlapping peptide library. We extended our studies to the co-supplementation of the synthesized peptides with hPEN as competitors or inhibitors in biological studies.

When chemically synthesizing peptides based on biological templates, it is imperative to keep in mind that, unlike model systems, they may contain complicated sequences that cannot be easily synthesized. This was also the case for some proposed peptides containing the GRP83\_50-52 region. We have succeeded in synthesizing the Asp-Trp region in intermediate-length peptides through extensive studies. However, in the case of a position at the C- or N-term in the primary structure, the synthesis of this section was not successful. Therefore, the overlap region of the peptide library had to be interrupted at this position. In total, we succeeded in synthesizing 13 peptides with different lengths from 10 to 56 AAs, and the putative ligand hPEN, using different well-tailored SPPS protocols according to the Fmoc strategy. The obtained peptides were analyzed for their secondary structure by CD measurements. Most of the peptides showed a disordered structure. Only the eNDo showed a typical tendency toward a  $\beta$ -structure of rhodopsin-like GPCRs.

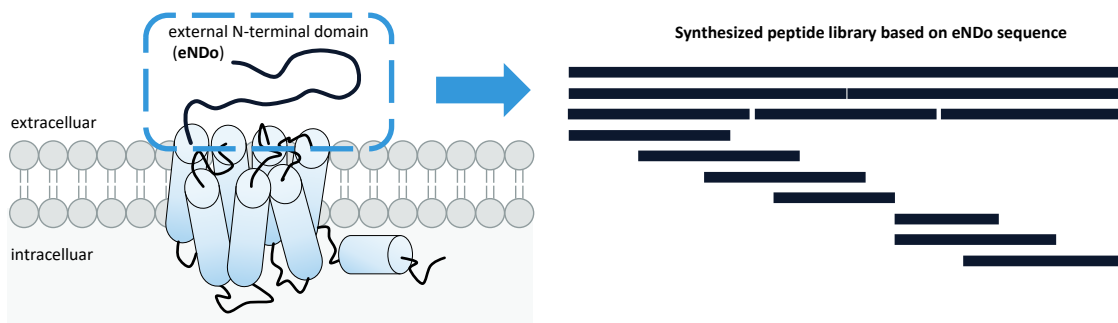
Biological studies of receptor activity were also performed. As a result, difficulties were identified in using the N41 cell line. Although the N41 cell line is better suited for the natural occurrence of GPR83 in neuronal cell lines, the expression of the receptor in this cell line needs to be further optimized, e.g., by establishing a successful transfection protocol.

With the help of Grötzinger group hPEN and the synthesized peptides were investigated in HEK cells. However, here it was not possible to reproduce the results of Gomes *et al.*, neither with the hPEN we provided nor with the commercially available hPEN. At this point, the project was terminated by all collaborators without successfully investigating the synthesized members of the peptide library.

Encouragingly, during the course of the studies, some open questions regarding GPR83 were clarified. For example, both homo- and hetero-oligomerization (MC4R) were detected and demonstrated not to be affected by eNDo. The region shown to be most critical for receptor activity was narrowed down using an Ala scan. [63]

We also hypothesized that different proteases might affect receptor activity because a protease inhibitor was used in all biological studies. It remains an open question for future investigations if the eNDo can affect receptor activity by proteolytic removal.

In conclusion, GPR83 remains an interesting receptor, the *de-orphaning* of which could enable various applications.



**Figure 13.1:** Schematic representation of GPR83 (left) and synthesized library (right).



## 14 References

- [1] B. K. Kobilka, *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2007**, 1768, 794–807.
- [2] M. T. Harrigan, G. Baughman, N. F. Campbell, S. Bourgeois, *Mol Cell Biol* **1989**, 9, 3438–3446.
- [3] A. Müller, B. Leinweber, J. Fischer, T. D. Müller, A. Grüters, M. H. Tschöp, V. Knäuper, H. Biebermann, G. Kleinau, *BMC Res Notes* **2014**, 7, 913.
- [4] F. Wu, G. Song, C. de Graaf, R. C. Stevens, *J Mol Biol* **2017**, 429, 2726–2745.
- [5] T. M. Bridges, C. W. Lindsley, *ACS Chem Biol* **2008**, 3, 530–541.
- [6] A. S. Hauser, M. M. Attwood, M. Rask-Andersen, H. B. Schiöth, D. E. Gloriam, *Nat Rev Drug Discov* **2017**, 16, 829–842.
- [7] M. Rask-Andersen, S. Masuram, H. B. Schiöth, *Annu Rev Pharmacol Toxicol* **2014**, 54, 9–26.
- [8] K. Sriram, P. A. Insel, *Mol Pharmacol* **2018**, 93, 251–258.
- [9] M. Lu, B. Wu, *IUBMB Life* **2016**, 68, 894–903.
- [10] E. Ringel, G. Riemekasten, *Trillium Immunologie* **2021**, 5, 39–47.
- [11] K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. le Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto, M. Miyano, *Science (1979)* **2000**, 289, 739–745.
- [12] D. Hilger, M. Masureel, B. K. Kobilka, *Nat Struct Mol Biol* **2018**, 25, 4–12.
- [13] R. Seifert, K. Wenzel-Seifert, *Naunyn Schmiedebergs Arch Pharmacol* **2002**, 366, 381–416.
- [14] A. F. M. Gavriilidou, H. Hunziker, D. Mayer, Z. Vuckovic, D. B. Veprintsev, R. Zenobi, *J Am Soc Mass Spectrom* **2019**, 30, 529–537.
- [15] A. J. Y. Jones, F. Gabriel, A. Tandale, D. Nietlispach, *Molecules* **2020**, 25, 4729.
- [16] J. Oates, A. Watts, *Curr Opin Struct Biol* **2011**, 21, 802–807.
- [17] K. Palczewski, *Trends Biochem Sci* **2010**, 35, 595–600.
- [18] R. Sleno, T. E. Hébert, **2018**, pp. 141–171.
- [19] F. J. Meye, G. M. J. Ramakers, R. A. H. Adan, *Transl Psychiatry* **2014**, 4, e361–e361.
- [20] J. Sato, N. Makita, T. Iiri, *Endocr J* **2016**, 63, 507–514.
- [21] N. Okashah, Q. Wan, S. Ghosh, M. Sandhu, A. Inoue, N. Vaidehi, N. A. Lambert, *Proceedings of the National Academy of Sciences* **2019**, 201905993.

- [22] D. J. Dupré, M. Robitaille, R. V. Rebois, T. E. Hébert, *Annu Rev Pharmacol Toxicol* **2009**, *49*, 31–56.
- [23] A. v. Smrcka, *Cellular and Molecular Life Sciences* **2008**, *65*, 2191–2214.
- [24] S. M. Khan, R. Sleno, S. Gora, P. Zylbergold, J.-P. Laverdure, J.-C. Labbé, G. J. Miller, T. E. Hébert, *Pharmacol Rev* **2013**, *65*, 545–577.
- [25] G. Milligan, E. Kostenis, *Br J Pharmacol* **2006**, *147*, S46–S55.
- [26] K. Kristiansen, *Pharmacol Ther* **2004**, *103*, 21–80.
- [27] M. Jo, S. T. Jung, *Exp Mol Med* **2016**, *48*, e207–e207.
- [28] S. Siehler, *Br J Pharmacol* **2009**, *158*, 41–49.
- [29] C. K. Billington, R. B. Penn, *Respir Res* **2003**, *4*, 4.
- [30] S. Chung, T. Funakoshi, O. Civelli, *Br J Pharmacol* **2008**, *153*, S339–S346.
- [31] I. Shimada, T. Ueda, Y. Kofuku, M. T. Eddy, K. Wüthrich, in *NMR with Biological Macromolecules in Solution*, WORLD SCIENTIFIC, **2021**, pp. 197–220.
- [32] A. S. Hauser, M. M. Attwood, M. Rask-Andersen, H. B. Schiöth, D. E. Gloriam, *Nat Rev Drug Discov* **2017**, *16*, 829–842.
- [33] “<https://www.guidetopharmacology.org/GRAC/ObjectDisplayForward?objectId=119>,” **n.d.**
- [34] A. L. Martin, M. A. Steurer, R. S. Aronstam, *PLoS One* **2015**, *10*, e0138463.
- [35] R. Sah, L. M. Pritchard, N. M. Richtand, R. Ahlbrand, K. Eaton, F. R. Sallee, J. P. Herman, *Neuroscience* **2005**, *133*, 281–292.
- [36] M. T. Harrigan, N. F. Campbell, S. Bourgeois, *Molecular Endocrinology* **1991**, *5*, 1331–1338.
- [37] W. Hansen, A. M. Westendorf, T. Toepfer, S. Mauel, R. Geffers, A. D. Gruber, J. Buer, *Genes Immun* **2010**, *11*, 357–361.
- [38] S. Brézillon, M. Detheux, M. Parmentier, T. Hökfelt, Y. L. Hurd, *Brain Res* **2001**, *921*, 21–30.
- [39] L. de Moerlooze, J. Williamson, F. Liners, J. Perret, M. Parmentier, *Cytogenet Genome Res* **2000**, *90*, 146–150.
- [40] J. Eberwine, T. Bartfai, *Pharmacol Ther* **2011**, *129*, 241–259.
- [41] P. Pesini, M. Detheux, M. Parmentier, T. Hökfelt, *Molecular Brain Research* **1998**, *57*, 281–300.
- [42] L. M. Lueptow, L. A. Devi, A. K. Fakira, **2018**, pp. 1–25.

- [43] J. S. Dubins, M. Sanchez-Alavez, V. Zhukov, A. Sanchez-Gonzalez, G. Moroncini, S. Carvajal-Gonzalez, J. R. Hadcock, T. Bartfai, B. Conti, *Metabolism* **2012**, *61*, 1486–1493.
- [44] T. D. Müller, A. Müller, C.-X. Yi, K. M Habegger, C. W. Meyer, B. D. Gaylinn, B. Finan, K. Heppner, C. Trivedi, M. Bielohuby, W. Abplanalp, F. Meyer, C. L. Piechowski, J. Pratzka, K. Stemmer, J. Holland, J. Hembree, N. Bhardwaj, C. Raver, N. Ottaway, R. Krishna, R. Sah, F. R. Sallee, S. C. Woods, D. Perez-Tilve, M. Bidlingmaier, M. O. Thorner, H. Krude, D. Smiley, R. DiMarchi, S. Hofmann, P. T. Pfluger, G. Kleinau, H. Biebermann, M. H. Tschöp, *Nat Commun* **2013**, *4*, 1968.
- [45] L. E. Vollmer, S. Ghosal, J. A. Rush, F. R. Sallee, J. P. Herman, M. Weinert, R. Sah, *Genes Brain Behav* **2013**, *12*, 241–249.
- [46] I. Gomes, E. N. Bobeck, E. B. Margolis, A. Gupta, S. Sierra, A. K. Fakira, W. Fujita, T. D. Müller, A. Müller, M. H. Tschöp, G. Kleinau, L. D. Fricker, L. A. Devi, *Sci Signal* **2016**, *9*, DOI 10.1126/scisignal.aad0694.
- [47] A. Müller, G. Kleinau, C. L. Piechowski, T. D. Müller, B. Finan, J. Pratzka, A. Grüters, H. Krude, M. Tschöp, H. Biebermann, *PLoS One* **2013**, *8*, e53347.
- [48] S. M. Mack, I. Gomes, L. A. Devi, *ACS Chem Neurosci* **2019**, *10*, 1884–1891.
- [49] N. C. Wrighton, F. X. Farrell, R. Chang, A. K. Kashyap, F. P. Barbone, L. S. Mulcahy, D. L. Johnson, R. W. Barrett, L. K. Jolliffe, W. J. Dower, *Science (1979)* **1996**, *273*, 458–463.
- [50] T. Beissbarth, J. A. Tye-Din, G. K. Smyth, T. P. Speed, R. P. Anderson, *Bioinformatics* **2005**, *21*, i29–i37.
- [51] S. E. Cwirla, P. Balasubramanian, D. J. Duffin, C. R. Wagstrom, C. M. Gates, S. C. Singer, A. M. Davis, R. L. Tansik, L. C. Mattheakis, C. M. Boytos, P. J. Schatz, D. P. Baccanari, N. C. Wrighton, R. W. Barrett, W. J. Dower, *Science (1979)* **1997**, *276*, 1696–1699.
- [52] S. Mabe, T. Nagamune, M. Kawahara, *Sci Rep* **2015**, *4*, 6127.
- [53] A. P. Davenport, C. C. G. Scully, C. de Graaf, A. J. H. Brown, J. J. Maguire, *Nat Rev Drug Discov* **2020**, *19*, 389–413.
- [54] <https://www.biosyn.com/peptidescreeningtools.aspx>, “peptide libraries,” **n.d.**
- [55] <https://www.proimmune.com/custom-peptide-libraries/>, “,” **n.d.**

- [56] [https://www.genscript.com/peptide\\_screening\\_tools.html](https://www.genscript.com/peptide_screening_tools.html),  
“[https://www.genscript.com/peptide\\_screening\\_tools.html](https://www.genscript.com/peptide_screening_tools.html),” **n.d.**
- [57] K. L. Morrison, G. A. Weiss, *Curr Opin Chem Biol* **2001**, *5*, 302–307.
- [58] J. L. Lauer, C. G. Fields, G. B. Fields, *Letters in Peptide Science* **1995**, *1*, 197–205.
- [59] Y. Yang, in *Side Reactions in Peptide Synthesis*, Elsevier, **2016**, pp. 119–161.
- [60] T. Michels, R. Dölling, U. Haberkorn, W. Mier, *Org Lett* **2012**, *14*, 5218–5221.
- [61] C. W. Gruber, M. Muttenthaler, M. Freissmuth, *Curr Pharm Des* **2010**, *16*, 3071–3088.
- [62] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, T. Gray, *Protein Science* **1995**, *4*, 2411–2423.
- [63] A. Müller, J. C. Berkmann, P. Scheerer, H. Biebermann, G. Kleinau, *PLoS One* **2016**, *11*, e0168260.

## **Part C**

### **Material and Methods**



---

## Content Part C

<b>14. Materials and Methods</b> .....	<b>179</b>
<b>14.1. General Experimental Conditions: Instrumental Setups and Procedures</b> .....	<b>179</b>
14.1.1 Reagents and Solvents .....	179
14.1.2 NMR Spectroscopy .....	180
14.1.3 Solid Phase Peptide Synthesis (SPPS) .....	181
14.1.4 Microwave-assisted Solid Phase Peptide Synthesis (MW-SPPS) .....	185
14.1.5 Analysis, Isolation and Purification of Synthesized Peptides .....	189
14.1.6 Lyophilization .....	197
14.1.7 Mass spectrometry.....	197
14.1.8 SDS-PAGE.....	198
14.1.9 UV/ Vis-spectroscopy .....	199
14.1.10 CD spectroscopy .....	202
14.1.11 Biological Studies: Instrumental Set-up and Procedures .....	204
14.1.12 Activation Studies of GPR83.....	207
<b>14.2. Organic Synthesis</b> .....	<b>210</b>
14.2.1 Synthesis of (S)-2-amino-4,4,4-trifluorobutanoic acid (TfeGly).....	210
14.2.2 Synthesis of Fmoc-TfeGly (1) Following the Ni(II)-complex Strategy .....	213
14.2.3 Synthesis of Fmoc-TfVal Derivatives (2+ 3), according to Erdbrink.....	215
14.2.4 Synthesis of RBM-linker 22.....	218
14.2.5 4-Methoxy-5-nitrosalicylaldehyde synthesis (22).....	218
<b>14.3. Synthesized and Purified Peptides</b> .....	<b>219</b>
14.3.1 Part A: pH (Low) Lipid Insertion Peptides and its Derivatives .....	219
14.3.2 Part A: Multivalent Presentation of EAE-Peptides on FF03 Scaffold .....	221
14.3.3 Part A: Acid and Base Strands of A4/B4 Dimeric Coiled-Coil.....	225
14.3.4 Part A: Functionalized Ubiquitin Monomers .....	230
14.3.5 Ub1(K48-Cys).....	234
14.3.6 Part A: 1FYN and its Derivatives .....	235
14.3.7 Part A: 1TEN and its Derivatives.....	242
14.3.8 Part B: GPR83-Project.....	245
14.3.9 Performed Reactions with synthesized Peptides .....	253
<b>15. References:</b> .....	<b>255</b>
<b>16. Appendix</b> .....	<b>256</b>
<b>16.1 RMB-Optimizations</b> .....	<b>256</b>

---

---



## 14. Materials and Methods

### 14.1. General Experimental Conditions: Instrumental Setups and Procedures

#### 14.1.1 Reagents and Solvents

All syntheses involving air- and moisture-sensitive compounds were carried out using the standard Schlenk technique under an atmosphere of argon. Before use the glassware was flame-dried under vacuum and cooled under flushing with argon. Air- and moisture-sensitive compounds were added under argon counterflow or with the help of septa using an oven-dried syringe or cannula. Solvents were dried with the solvent purification system MB-SPS 800 from M. Braun (M. Braun Inertgas-Systeme GmbH, Garching, Germany). Room temperature refers to 22°C.

Reactions were monitored by thin-layer chromatography (TLC) using Merck KGaA silica gel 60 F254 TLC aluminum sheets (0.25 mm thickness, Merck KGaA, Darmstadt, Germany) and visualized with ceric ammonium molybdate, anis aldehyde staining solution, potassium permanganate staining solution, ninhydrin solution staining, or UV light. Flash chromatography was carried out using Silica 60 M (0.040–0.063 mm) from Macherey-Nagel GmbH & Co. KG (Düren, Germany), using a forced flow of eluent (method of Still)<sup>[356]</sup> 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. Resins (preloaded TGA and TGT resins and not preloaded TGR and Rink amide resins) for conventional solid-phase peptide synthesis were purchased from Novabiochem® (Merck KGaA, Darmstadt, Germany). CEM was used as a commercial source for pro tide resins used in MW-SPPS (Rink Amide, Cl-MPA, and Cl-TCP(Cl) resin), Fmoc-Asp(OMpe), Fmoc-His(Boc)-OH, Fmoc-Lys(Mmt)-OH, Fmoc-Asp(O<sup>t</sup>Bu)-(Dmb)-Gly-OH, Oxyma pure.

Canonical Fmoc- and orthogonal side-chain protected amino acids were obtained from Carbolution. Pseudoproline dipeptide building blocks and Fmoc-Lys(ivDde)-OH were bought from Iris Biotech GmbH (Marktredwitz, Germany) and Boc-2-Abz-OH from Sigma. Coupling reagents came from Novabiochem® (Merck KGaA, Darmstadt, Germany), Carbolution (Carbolution Chemicals GmbH, St. Ingbert,

Germany) or Sigma-Aldrich® (Merck KGaA, Darmstadt, Germany). Solvents for peptide synthesis were purchased as synthesis grade from Fisher Scientific (Schwerte, Germany), Acros Organics (Thermo Fisher Scientific, Geel, Belgium), VWR (VWR International GmbH, Darmstadt, Germany), or Merck (Merck Chemicals GmbH, Darmstadt, Germany). Solvents for peptide purification were obtained as HPLC or spectroscopy grade from Fisher Scientific (Schwerte, Germany) or Merck (Merck Chemicals GmbH, Darmstadt, Germany).

Deionized water used for HPLC or to prepare buffers was obtained using a Milli-Q® Advantage A10 Ultrapure Water Purification System (Merck KGaA, Darmstadt, Germany).

All other chemicals and solvents were purchased from Acros Organics (Thermo Fisher Scientific, Geel, Belgium), abcr GmbH (Karlsruhe, Germany), chemPUR (Karlsruhe, Germany), Fluorochem (Hadfield, United Kingdom), VWR (Darmstadt, Germany), Merck (Darmstadt, Germany), Sigma-Aldrich® (Merck KGaA, Darmstadt, Germany), Alfa Aesar (Thermo Fisher (Kandel) GmbH, Karlsruhe, Germany), Novabiochem® (Merck KGaA, Darmstadt, Germany), Roth (Carl Roth GmbH + Co. KG, Karlsruhe, Germany), or TCI chemicals (Zwijndrecht, Belgium) at highest commercially available purity. Deutero GmbH (Kastellaun, Germany) was used a commercial source for solvents applied in NMR-Measurements.

All reagents and solvents were used without further purification. Acetic anhydride (Grüssing GmbH Analytika, Filsum, Germany) was freshly distilled prior to use.

### 14.1.2 NMR Spectroscopy

<sup>1</sup>H-, <sup>13</sup>C- and <sup>19</sup>F-NMR spectra were recorded at rt using a JEOL ECX400 (<sup>1</sup>H-NMR: 400 MHz, <sup>13</sup>C-NMR:101 MHz, <sup>19</sup>F-NMR: 376 MHz, JEOL, Tokyo, Japan), a JEOL ECP500 (<sup>1</sup>H-NMR: 500 MHz, <sup>13</sup>C-NMR: 126 MHz, JEOL, Tokyo, Japan) or a JOEL ECZ600 S (<sup>1</sup>H-NMR: 600 MHz, <sup>13</sup>C-NMR: 126 MHz, JEOL, Tokyo, Japan). The depiction and analysis of the spectra occurred with the program MestReNova version 14.2.2-28739 (Mestrelab Research S. L., Santiago de Compostela, Spain).

The chemical shifts ( $\delta$ ) are given in parts per million (ppm). The <sup>1</sup>H- and <sup>13</sup>C-NMR chemical shifts were referenced against the internal solvent (CDCl<sub>3</sub>  $\delta$  (<sup>1</sup>H) = 7.26 ppm,  $\delta$  (<sup>13</sup>C) = 77.16 ppm; deuterium oxide D<sub>2</sub>O (<sup>1</sup>H) = 4.79 ppm,  $\delta$  (<sup>13</sup>C) = 29.84 ppm; DMSO-D<sub>6</sub>  $\delta$  (<sup>1</sup>H) = 2.50 ppm,  $\delta$  (<sup>13</sup>C) = 39.52 ppm) and given to trimethylsilane

as internal standard ( $\delta = 0.00$  ppm). The  $^{19}\text{F}$ -NMR chemical shifts are given relative to  $\text{CFCl}_3$  as an external reference or TFA as an internal standard.  $^{13}\text{C}$ -NMR-spectra were recorded  $^1\text{H}$  decoupled. The order of citation in parentheses is a) multiplicity (s = singlet, d = doublet, t = triplet, m = multiplet, br = broad, and combinations thereof), b) coupling constant  $J$ , c) number of nuclei.

### 14.1.3 Solid Phase Peptide Synthesis (SPPS)

Unless otherwise stated, peptide syntheses were performed fully automatically using the Activo-P11 Automated Peptide Synthesizer (Activotec, Cambridge, United Kingdom) according to the protocols in 14.1.3.1. Manual SPPS was also applied for short sequences in the GPR83 project. The conditions for this are listed in Table 14.1 and Table 14.2. Microwave-assisted peptide synthesis was performed exclusively on the CEM Liberty Blue™ automated microwave peptide synthesizer (CEM Corporation, Matthews, NC, USA) according to parameters described in section 14.1.4.

Peptides were assembled from the C to the N term using the Fmoc/ $t$ Bu protecting group strategy. The following standard building blocks were implemented unless stated otherwise:

Fmoc-Ala-OH  $\cdot$  H<sub>2</sub>O, Fmoc-Arg(Pbf)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Asp(O $t$ Bu)-OH, Fmoc-Cys(Trt)-OH, Fmoc-L-Gln(Trt)-OH, Fmoc-L-Glu(O $t$ Bu)-OH  $\cdot$  H<sub>2</sub>O, Fmoc-Gly-OH, Fmoc-His(Trt)-OH, Fmoc-Ile-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Met-OH, Fmoc-Phe-OH, Fmoc-Ser( $t$ Bu)-OH, Fmoc-Thr( $t$ Bu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Tyr( $t$ Bu)-OH, Fmoc-Val-OH.

In addition, the following special Fmoc-protected building blocks were utilized: Fmoc-Ahx-OH, Fmoc-Asp(OEpe)-OH, Fmoc-Asp(OMpe)-OH, Fmoc-His(Boc)-OH, Fmoc-Lys(ivDde)-OH, Fmoc-Lys(Mmt)-OH, Fmoc-Lys(Mtt)-OH, Fmoc-Nle-OH. Furthermore, Boc-protected building blocks were also employed for N-terminal positions in some projects. Boc-Cys(Trt)-OH, Boc-His(Trt)-OH, Boc-Met-OH and Boc-Nle-OH. All amino acids were used in the L configuration.

Moreover, the following dipeptide building blocks were used: Fmoc-Asp(O $t$ Bu)-(Dmb)-Gly-OH, Fmoc-Leu-Ser[PSI(Me,Me)Pro]-OH and Fmoc-Leu-Thr[PSI(Me,Me)-Pro]-OH.

### 14.1.3.1 Reaction Conditions for Conventional SPPS

Conventional SPPS was accomplished either by manual couplings or automated synthesis robots. All reaction steps were carried out at rt. Moreover, the automated synthesis was performed in an inert gas atmosphere, whereas the manual synthesis was conducted in an oxygen atmosphere. Generally, a 0.05 mmol approach was used. The individual protocols are listed in Table 14.1 to Table 14.4.

### 14.1.3.2 Manual Coupling Protocols

For conventional SPPS, resins preloaded with the corresponding first amino acid were generally chosen. For the **A4F/B4F**, a NovaSyn<sup>®</sup> TGR resin (0.25 mmol/g) was used. For this purpose, the resin had to be loaded manually. After the reaction, the loading was determined.

#### ***General procedure: Resin Loading and Loading Test***

First, 500 mg resin was swollen twice with 10 mL DMF for 15 min each at rt. Subsequently, the respective first amino acid was coupled. For this, 5 eq. of the Fmoc-X<sub>AA</sub>-OH (Fmoc-Gly-OH for **A4F** and Fmoc-Cys(Trt)-OH for **B4F**) as well as 4.9 eq. HATU and 10 eq. DIPEA were used. The reaction was carried out for 3 h at rt. Subsequently, the resin was washed with DMF, DCM, and MeOH (each 3 x 5 mL), and the resin was dried *in vacuo* for at least 2 h.

The loading test was performed according to the technical note from aapptec.<sup>[1]</sup> For this purpose, a triplicate determination was consistently performed:

First, approximately 20 mg was weighed out in a 2 mL Eppendorf tube. Then the resin was swollen for 15 min at rt in 1.6 mL DMF. Next, 400 µL of piperidine was added, and the reaction mixture was shaken for an additional 15 min. 200 µL of this mixture was transferred to a new tube and diluted with 1800 µL DMF. For the blank, the corresponding procedure was conducted without resin. Subsequently, the absorbance was measured in a quartz cuvette using UV/Vis spectrometry. The following equation was used to determine the loading of the resin:

$$L = (A_{301} \cdot V \cdot d) / (E_c \cdot w \cdot M) \quad (1)$$

Where  $L$  = resin loading,  $A_{301}$  = absorbance at 301 nm,  $V$  = volume of cleavage cocktail = 2 mL,  $d$  = dilution = 10,  $E_c$  = extinction coefficient = 7800 mL/mmol · cm,  $w$  = width of the cuvette = 1 cm,  $M$  = weight of the resin sample.

Any unreacted NH<sub>2</sub> groups were acetylated on the resin. For this purpose, capping was performed with a solution of 10% (v/v) DIPEA and 10% (v/v) Ac<sub>2</sub>O in DMF. The reaction was carried out three times for 10 min at rt. The resin was then washed with DMF and DCM (3 x 5 mL each).

### ***General Coupling Procedures***

Two different coupling strategies (HOAt/HOBt + DIC or HATU chemistry) were employed in the manual synthesis of the peptides. These are described in the following tables (Table 14.1+ Table 14.2).

At the beginning of each synthesis, the resin was swollen in 2 mL DMF for 30 min in a 10 mL polypropylene syringe reaction vessel. This was followed by Fmoc cleavage with 20% (v/v) piperidine in DMF for 10 min. If an Asp-residue was present in the sequence, 0.1 M HOBt was dosed to the Fmoc cocktail as an additive to suppress aspartimide formation. The procedure was repeated three times, and the resin was subsequently washed three times with 5 mL DMF. Coupling was performed in a solution of 0.4 M NaClO<sub>4</sub> in DMF to avoid aggregation during peptide assembly. The coupling of the AA followed this. Depending on the chosen coupling strategy, either 8 eq. (+ 8eq. DIC + 8 eq. HOAt or HOBt) or 5 eq. (+ 4.9 eq. HATU + 10 eq. DIPEA) of the AA was used. In HOAt/HOBt + DIC, pre-activation of the AA to be coupled is necessary before adding it to the resin. In HATU chemistry, this is not required. In addition, the reaction time for HOAt/HOBt is 60 min, while for HATU, it is 30 min. Double coupling was usually performed, i.e., the AA to be coupled was activated twice and added to the resin. The resin was washed three times with 5 mL DMF and DCM following the coupling step.

After the complete sequence, the N-terminus was either acetylated (capping) as described above or Boc-2-Abz-OH was attached. The latter was coupled in the same way as the canonical amino acids.

**Table 14.1:** Protocol manual SPPS synthesis using a 0.05 mmol scale utilizing HOAt/HOBt + DIC activation.

process	reagents	duration
resin swelling	2 mL of DMF	30 min
Fmoc removal	2 mL of 20% Pip (+ 0.1 M HOBt) in DMF (v/v)	3 x 10 min
washing	3 x 5 mL of DCM; 3 x 5 mL of DMF	1 min
amino acid activation	2 x 2 mL 0.4 M NaClO <sub>4</sub> in DMF + 8 eq. each Fmoc-X <sub>AA</sub> -OH, HOAt or HOBt, DIC	5 min (HOAt) 10 min (HOBt)
amino acid coupling	2 x 2 mL of 0.4 M NaClO <sub>4</sub> in DMF + 8 eq. each Fmoc-X <sub>AA</sub> -OH, HOAt or HOBt, DIC	60 min
washing	3 x 5 mL of DCM; 3 x 5 mL of DMF	1 min
capping	2 mL of 10% DIPEA+10% Ac <sub>2</sub> O in DMF (v/v)	3 x 10 min

**Table 14.2:** The manual SPPS protocol using a 0.05 mmol scale utilizing HATU+ DIPEA activation.

process	reagents	duration
resin swelling	2 mL of DMF	30 min
Fmoc removal	2 mL of 20% Pip (0.1 M HOBt) in DMF (v/v)	3 x 10 min
washing	3 x 5 mL of DCM; 3 x 5 mL of DMF	1 min
amino acid activation	no pre-activation	0 min
amino acid coupling	2 x 2 mL 0.4 M NaClO <sub>4</sub> in DMF + 5 eq. Fmoc-X <sub>AA</sub> -OH, 4.9 eq. HATU, 10 eq. DIPEA	30 min
washing	3 x 5 mL of DCM; 3 x 5 mL of DMF	1 min
capping	2 mL of 10% DIPEA/10% Ac <sub>2</sub> O in DMF (v/v)	3 x 10 min

### *Coupling of Fluorinated Amino Acids*

For fluorine-containing residues, a single coupling step was carried out with a molar excess of the amino acid and coupling reagents of 1.2-fold relative to the resin loading in 2.0 mL DMF containing 0.4 M NaClO<sub>4</sub>. These couplings were performed for 6 h at rt. Another 0.5 eq. of AA and coupling reagents (DIC/HOAt) were added to the resin and shaken overnight. Prior to the Fmoc deprotection of the fluorinated AA, unreacted NH<sub>2</sub> groups on the resin were acetylated in a capping procedure as previously described. The resin was washed with DMF and DCM (3 × 5.0 mL each). The amino acid that is immediately following the fluorinated one was also coupled manually according to the protocol in Table 7.10 using 8-fold excess of amino acid and coupling reagents (DIC/HOAt).

### 14.1.3.3 Automated Conventional Solid Phase Peptide Synthesis

In analogy to the steps described for manual synthesis, peptides were synthesized using Activotec's (Cambridge, United Kingdom) Synthesizer P-11 using the following protocols. All reactions were carried out at rt under nitrogen atmosphere.

**Table 14.3:** The automated SPPS protocol using a 0.05 mmol scale on the Activo-P11 Automated Peptide Synthesizer utilizing HOAt/HOBt + DIC activation.

protocol	process	reagents <sup>a</sup>	reaction time
<b>start</b>	swelling	2 mL DMF	2 x 15 min
	Fmoc deprotection	2 mL 20% Pip <sup>b</sup> in DMF	2 x 5 min + 2 x 10 min
	washing	2 mL DMF	5 x 1 min
<b>double coupling</b>	amino acid coupling	8 eq. each Fmoc-X <sub>AA</sub> -OH, HOAt or HOBt and DIC in 1.6 mL DMF <sup>c</sup>	2 x 60 min
	washing	2 mL DMF	5 x 1 min
	Fmoc-deprotection	2 mL 20% Pip <sup>b</sup> in DMF	2 x 5 min + 2 x 10 min

**a:** After the synthesis, a final wash was performed with DCM, **b:** 0.1 M HOBt was used as an additive in Fmoc removal when synthesizing peptide sequences containing Asp, **c:** Addition of HOAt/HOBt and DIC was combined before Fmoc-X<sub>AA</sub>-OH was dissolved.

**Table 14.4:** Protocol for the automated solid phase peptide synthesis using a 0.05 mmol scale on the Activo-P11 Automated Peptide Synthesizer utilizing HATU+ DIPEA activation.

protocol	process	reagents <sup>a</sup>	reaction time
<b>start</b>	swelling	2 mL DMF	2 x 15 min
	Fmoc deprotection	2 mL 20% Pip <sup>b</sup> in DMF (v/v)	4 x 5 min
	washing	2 mL DMF	5 x 1 min
<b>double coupling</b>	amino acid coupling	5 eq. Fmoc-X <sub>AA</sub> -OH, 4.9 eq. HATU and 10 eq. DIPEA in 1.6 mL DMF <sup>c</sup>	2 x 30 min
	washing	2 mL DMF	5 x 1 min
	Fmoc-deprotection	2 mL 20% Pip <sup>b</sup> in DMF (v/v)	4 x 5 min
	washing	2 mL DMF	6 x 1 min

**a:** After the synthesis, a final wash was performed with DCM, **b:** 0.1 M HOBt was used as an additive in Fmoc removal when synthesizing peptide sequences containing Asp, **c:** Fmoc-X<sub>AA</sub>-OH and HATU were dissolved prior to DIPEA addition.

### 14.1.4 Microwave-assisted Solid Phase Peptide Synthesis (MW-SPPS)

The MW-SPPS was exclusively performed with the CEM Liberty Blue™ Automated Microwave Peptide Synthesizer (CEM Corporation, Matthews, NC, USA). The following sessions describe the methods, reagents set-ups, and programmed cycles used in this doctoral thesis in detail.

#### 14.1.4.1 Microwave Methods

These methods were used for both 0.05 and 0.1 mmol scale synthesis. Except for the acetylation (capping), all microwave methods include two phases. The first phase usually needs a high power for a short period of time to ensure a rapid increase in temperature. To avoid an overshoot in temperature, a second phase is included with lower microwave power for the desired reaction time.

**Table 14.5:** Microwave parameters for different processes.

process	temperature (°C)	power (W)	hold time (s)	delta T (°C)
capping	65	45	35	5
	65	0	50	5
	65	50	35	5
	65	0	45	5
DCA loading	80	75	60	2
	90	20	540	1
coupling 10 min at 50°C	25	0	120	2
	50	35	480	1
coupling 10 min at 90°C	75	217	15	1
	90	40	585	1
coupling 2 min	75	175	20	2
	90	35	120	1
coupling 3 min	75	217	15	2
	90	43	165	2
coupling 4 min	75	217	15	2
	90	43	225	2
deprotection	75	155	15	2
	90	35	120	1
Mmt removal	25	0	120	1

#### 14.1.4.2 Reagent Setup

The overall concentration of Fmoc-X<sub>AA</sub>-OH was 0.2 M in DMF unless stated otherwise.

AA coupling was performed with DIC (activator) and Oxyma Pure (activator base) chemistry. Their concentration depends on the reaction scale of the conducted synthesis (Table 14.6). When using a Chloride-based resin (Cl-MPA pro tide or Cl(TCP) pro tide resin), the AA couplings were performed utilizing the CarboMax-approach (CMA), leading to increased eq. of DIC and an additional 0.1 eq. of DIPEA in the activator base solution. These resins were not pre-swollen and activated using a Potassium iodide/DIPEA solution with the DCA loading procedure on the device. In contrast, Rink-amide pro tide resins were swollen at rt for 30 min prior to usage.



Deprotection was performed either with 10% piperazine (%w in NMP/EtOH 9:1 (% v/v)) or 20% piperidine (Pip in DMF); the flow rate of the device was calibrated accordingly.

**Table 14.6:** Reagents for MW-SPPS. Unless stated otherwise, all reagents were dissolved in DMF in the desired concentration. CMA = CarboMax-Approach

reagents	concentration <sup>a</sup>	amount
<b>Fmoc-X<sub>AA</sub>-OH</b>		
0.05 mmol	0.2 M	1.25
0.1 mmol	0.2 M	2.5
Fmoc-flAA-OH	2 eq.	1.25
<b>activator (DIC)</b>		
0.05 mmol scale	0.25	1
CMA:	0.5	1
0.1 mmol scale	0.5	1
CMA	1	1
<b>activator base (oxyma pure)</b>		
0.05 mmol scale	0.5	0.5
CMA: + 0.1 eq. DIPEA	0.5	0.5
0.1 mmol	1	0.5
CMA: + 0.1 eq. DIPEA	1	0.5
<b>chloride loading</b>		2 (0.05 mmol)
KI + DIPEA	0.125 M + 1 M	4 (0.10 mmol)
<b>ivDde removal</b>		
H <sub>2</sub> N-NH <sub>2</sub>	5% (v/v)	4
<b>acetylation (capping)</b>		
Ac <sub>2</sub> O+ DIPEA	10% + 10% (v/v)	4
<b>Mmt removal</b>		
TFA	2% in DCM	2
<b>deprotection</b>		
piperidine <sup>b</sup>	20%	3
piperazine <sup>b</sup>	10 w% in NMP/EtOH (9/1, % v/v)	3

**a:** DMF was used as the solvent if not stated otherwise; **b:** 0.1 M HOBT was added for peptide sequences containing Asp.

### **CarboMax Approach**<sup>[2]</sup>

The CarboMax approach is the method of achieving increased coupling efficiency by adding 0.1 eq. DIPEA in the activator base solution (oxyma pure) and increasing the addition of activator (10 eq. DIC). The addition of DIPEA is beneficial when using acid-labile resins such as Cl-MPA and Cl-TCP(Cl), as the acidity of the coupling solution is reduced. In addition, the formation of by-products should be suppressed.

### 14.1.4.3 Programmed Cycles for the Automated MW-assisted SPPS

The following table shows the different programmed processes during synthesis. Different MW settings can be selected for the couplings.

**Table 14.7:** Programmed cycles for MW-assisted SPPS. For deprotection and coupling, different MW settings, as shown in **Table 14.5**, can be used.

	cycle ID		cycle ID
<b>chloride-loading</b>	DCA (Trityl loading)	<b>single coupling x-tra wash</b>	Deprotection
	WTM <sup>a</sup> 4 x Wash		7 x Wash Coupling Wash
<b>single coupling</b>	Deprotection 4 x Wash Coupling Wash	<b>double deprotection, triple coupling</b>	2 x Deprotection 4 x Wash 3 x Coupling Wash
<b>double coupling</b>	Deprotection 4 x Wash 2 x Coupling Wash	<b>double deprotection, double coupling, capping</b>	2 x Deprotection 4 x Wash 2 x Coupling Wash Acetylation 3 x WTM <sup>a</sup>
<b>double deprotection, double coupling</b>	2 x Deprotection 4 x Wash 2 x Coupling Wash		

a: WTM = wash trough manifold (cleaning process for other different reagents used than AA + coupling reagents).

Workflows were also programmed for the removal of specific protecting groups (Table 14.8). The ivDde protecting group was removed with a solution of 5% (v/v) hydrazine in DMF. The reaction was carried out at 90 °C for 3 min. The procedure was repeated three times altogether. After adding the hydrazine solution, the WTM washes were performed, and the resin was washed five times. The Mmt protecting group was cleaved with 1% TFA in DCM. The reaction was carried out at rt. For this purpose, the conventional coupling was set to 2 min. Before the cleavage, a fivefold wash run with DCM was performed. The procedure was repeated five times in total. Subsequently, the device was cleaned with WTM, and the resin was washed with DMF to remove the residual DCM.

**Table 14.8:** Cycles for the removal of special protecting groups.

ivDde removal	Mmt removal
3 x Wash	5 x Wash DCM
3 x (Add reagent MW method <sup>a</sup>	5x (Add reagent MW method <sup>b</sup>
3 x WTM	Wash DCM)
5x Wash)	5x WTM
	5x Wash DMF

a: MW-method for ivDde removal: coupling 3 min; b: Mmt removal conventional coupling 2 min.

### 14.1.5 Analysis, Isolation and Purification of Synthesized Peptides

Micro-cleavage was performed to validate individual synthesis sections. A combination of analytical HPLC runs and ESI-ToF measurements performed the monitoring. After a complete synthesis, the desired peptide was removed from the solid support in full cleavage and subsequently purified by preparative HPLC. Analytical HPLC characterized the fractions obtained from the preparative HPLC runs before unification.

The following cleavage cocktails were used throughout the present doctoral thesis:

**Table 14.9:** TFA cleavage cocktails for removal of the desired peptide from the solid support.

cocktail	composition	ratio in % v/v
I	TFA/TIS/water 95/5/5	95/5/5
II	TFA/EDT/thioanisol/water/phenol	82.4/2.5/5/5/5
III	TFA/EDT/water/TIS	94/2.5/2.5/1
IV	TFA, DMS, TIS, thioanisol, H <sub>2</sub> O	90/2.5/2.5/2.5/2.5

#### 14.1.5.1 General Procedure of Micro-cleavages

A small amount (0.5 - 1 mg) of resin was transferred into an Eppendorf tube and treated with 200  $\mu$ L of the selected TFA cleavage cocktail. The reaction was stirred between 2 to 4 h at rt. Then, TFA was gently evaporated, and the desired peptide precipitated in 1.5 mL ice-cold diethyl ether. After centrifugation, the ether was decanted, and the isolated peptide was dissolved in a 1:1 mixture of MeCN/water + 0.1 % TFA. Before injection into HPLC, the sample was filtered with 13 mm syringe filters w/0.2  $\mu$ m PTFE membrane (VWR International GmbH, Darmstadt, Germany).

#### **14.1.5.2 General Procedure of Full Cleavages**

The dried resin was divided into portions of 150 mg in a polypropylene syringe reaction vessel and was treated with 3-5 mL of the selected TFA cleavage cocktail. The reaction was allowed to stir for 2 to 4 h at rt. Afterwards, the reaction mixtures were filtered and combined in a round bottom flask. The resin portions were each washed with 3 mL TFA and 3 mL DCM. The organic phase was concentrated *in vacuo*. Subsequently, the crude peptide was precipitated by 50 mL ice-cold diethyl ether. After centrifugation (4°C, 4.4 rpm, 4 to 6 min), the diethyl ether was decanted, and the residue was washed with an additional 30 mL ice-cold diethyl ether and dried *in vacuo* before purification.

#### **14.1.5.3 Preparative HPLC**

Reversed-phase preparative HPLC performed the purification of the isolated crude peptides by applying different linear gradients of MeCN in water, each containing 0.1% TFA. All crude samples were dissolved in starting conditions of selected gradient and filtered prior to injection with 0.45 µm Acrodisc® syringe filters with GHP membrane (Pall Corporation, Port Washington, NY, USA).

#### ***Knauer high-pressure-gradient System***

Several samples of the GPR83 peptide library were purified using an HPLC from Knauer GmbH (D-14167, Berlin) consisting of a smartline manager 5000 interface module with solvent degasser, two smartline pumps 1000 HPLC pumps, one 6-port-3-channel injection valve with a 5 mL sample loop, a variable UV flow detector (UV Detector 2500) and one high-pressure gradient mixer. The chromatogram was recorded with a Knauer GmbH x-y-plotter. An RP-HPLC column Luna™ C8(2) (10 µm, 250 x 21.20 mm, Phenomenex®, USA) was used for the chromatographic separation. The samples were detected at a wavelength of 320 nm. The runs were performed with a flow rate of 20 mL/min applying linear eluent gradients summarized in table Table 14.10.

**Table 14.10:** Applied linear gradients for Knauer high-pressure-gradient system. Eluents: A = water + 0.1% (v/v) TFA, B = MeCN + 0.1% (v/v) TFA.

name	time [min]	A [%]	B [%]	name	time [min]	A [%]	B [%]
<b>T1</b> <b>5-70</b>	0	95	5	<b>T2</b> <b>5-100</b>	0	95	5
	30	30	70		30	0	100
	31	0	100		32	0	100
	33	0	100		33	95	5
	34	95	10		36	95	5
	37	95	10				

### LaPrepΣ HPLC

All other synthesized peptides were purified with a LaPrepΣ HPLC system (VWR International GmbH, Darmstadt, Germany), comprising a LaPrepΣ LP 1200 preparative solvent pump with a 100 mL titanium pump head, 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 -LP 3101 1-channel UV-detector, a LaPrep \_semi-preparative flow cell with 0.5 mm path length and a LaPrep \_LP2016 17-port/1-channel fractionation valve. A Kinetex® C18 RP-HPLC-column with TMS end-capping (5 μm, 100 Å, 250 × 21.2 mm, Phenomenex®, Torrance, CA, USA) or a Luna® C8(2) column (10 μm, 100 Å, 250 × 21.2 mm, Phenomenex®, Torrance, CA, USA) was used. A SecurityGuard™ PREP Cartridge Holder Kit (21.20 mm ID, Ea, Phenomenex®, Torrance, CA, USA) either holding a C18 or C8 cartridge (15 × 21.2mm, Phenomenex®, Torrance, CA, USA) served as pre-column. HPLC runs were performed according to the methods given in Table 14.11. 220 nm were chosen as detection wavelength. Data analysis occurred with EZChrom *Elite* software (Version 3.3.2 SP2, Agilent Technologies, Santa Clara, CA, USA).

**Table 14.11:** Gradients used in purification procedures by LaPrepΣ HPLC. Eluents: A = water + 0.1% (v/v) TFA, B = MeCN + 0.1% (v/v) TFA.

name	time [min]	A [%]	B [%]	name	time [min]	A [%]	B [%]
<b>P1</b> <b>10-80</b> <b>18 min</b>	0	90	10	<b>P2</b> <b>10-80</b> <b>22 min</b>	0	90	10
	18	20	80		22	20	80
	19	0	100		23	0	100
	22	0	100		26	0	100
	23	10	10		27	90	10
	26	10	10	30	90	10	
<b>P3</b> <b>10-80</b> <b>30 min</b>	0	90	10	<b>P4</b> <b>5-100</b> <b>18 min</b>	0	95	5
	30	20	80		18	0	100
	31	0	100		23	0	100
	34	0	100		24	95	100
	36	90	10		27	95	5
	40	90	10				
<b>P5</b> <b>5-70 18 MIN</b>	0	95	5	<b>P6</b> <b>5-70 30 min</b>	0	95	5
	18	30	70		30	30	70
	19	0	100		32	0	100
	22	0	100		35	0	100
	23	95	5		36	95	5
	26	95	5	40	95	5	
<b>P7</b> <b>30-100 18 min</b>	0	70	30	<b>P8</b> <b>30-100 30 min</b>	0	70	30
	18	0	100		30	0	100
	21	0	100		32	0	100
	22	70	30		35	0	100
	25	70	30		36	70	30
				40	70	30	
<b>P9</b> <b>5-70 w PR</b>	0	95	5 <sup>a</sup>	<b>P10</b> <b>10-80 w PR</b>	0	95	5 <sup>a</sup>
	5	95	5 <sup>a</sup>		5	95	5 <sup>a</sup>
	20	30	70		20	20	80
	22	0	100		22	0	100
	25	0	100		25	0	100
	26	95	5	26	95	5	
	28	95	5	30	95	5	
<b>P11</b> <b>5-60 w PR</b>	0	90	10	<b>P12<sup>b</sup></b> <b>5 in 30 min</b>	0	95	5
	15	90	10		30	0	100
	17	80	20		33	0	100
	39	40	60		34	95	5
	40	0	100		37	95	5
	42	0	100				
	43	90	10				
	51	90	10				
<b>P13</b>	0	95	5				
	18	50	50				
	21	0	100				
	22	0	100				
	25	95	5				

PR= pre run, different flowrate: a: 10 mL/min, b: 20 mL/min

After separation, the purity of the collected fractions was determined by analytical HPLC. The used gradient methods are shown in following sections. In general, the UV-detection occurred at 220 nm (unlabeled peptides and peptide fragments, as

well as purity control of the final peptides), 280 nm (peptides containing Tyr and/or Trp residues), or 320 nm (Abz-labeled peptides). Fractions with sufficient purity were combined, and ACN was removed by rotary evaporation. Lyophilization of the remaining aqueous solution yielded the pure peptide. The data were analyzed with EZChrom *Elite* software (version 3.3.2, Agilent Technologies, Santa Clara, CA, USA).

#### 14.1.5.4 Analytical HPLC

##### *Semi-micro Chromaster HPLC System (FF)*

Analytical HPLC was carried out on a VWR-Hitachi Chromaster HPLC 600 bar system (VWR International GmbH, Darmstadt, Germany) that works with a low-pressure gradient. The system contains a 5160 pump with a 6-channel solvent degasser, an organizer, a 5260 autosampler with a 20  $\mu$ L sample loop, a 5310 column oven, and a 5430 diode array detector with a high-pressure semi-micro flow cell (5 mm). A Purospher® STAR RP-C18 end-capped UHPLC column (2  $\mu$ M, 120 Å, 50  $\times$  2.1 mm, Merck, Deutschland) was used. As eluents, deionized water and MeCN, each 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 14.12:** Linear gradients used in analytical runs by Chromaster (FF). Eluents: A = water + 0.1% (v/v) TFA, B = MeCN + 0.1% (v/v) TFA.

name	time [min]	A [%]	B [%]	name	time [min]	A [%]	B [%]
A1	0	95	5	A2	0	95	5
	9	0	100		6	30	70
	10.5	0	100		6.5	0	100
	11	95	5		7	0	100
	14	95	5		7.1	95	5
A3	0	90	10	A4	10	95	5
	6	20	80		0	90	10
	6.1	0	100		6	40	60
	7	0	100		6.5	0	100
	7.1	90	10		7	0	100
10	90	10	7.1	95	95		
				10	95	95	

##### *Chromaster HPLC System*

The VWR-Hitachi Chromaster HPLC 600 bar system (VWR International GmbH, Darmstadt, Germany) works with a low-pressure gradient. It comprises a 5160 pump with a 6-channel solvent degasser, an organizer, a 5260 autosampler with a 100  $\mu$ L sample loop, a 5310 column oven, and a 5430 diode array detector with a

standard flow cell (10 mm optical path length). A Luna<sup>®</sup> C8(2) column (5 μm, 100 Å, 250 × 4.6 mm, Phenomenex<sup>®</sup>, Torrance, CA, USA) or a Kinetex<sup>®</sup> C18 column (5 μm, 100 Å, 250 × 4.6 mm, Phenomenex<sup>®</sup>, Torrance, CA, USA), was used. A SecurityGuard<sup>™</sup> Cartridge Kit (Ea, Phenomenex<sup>®</sup>, Torrance, CA, USA) with either a C8 or C18 SecurityGuard<sup>™</sup> cartridge (4 × 3.0 mm, Phenomenex<sup>®</sup>, Torrance, CA, USA) served as a pre-column. The flowrate was adjusted to 1 mL/min, and the column was heated to 24°C.

**Table 14.13:** Linear gradients for analytical runs on the Chromaster System. Eluents: A = water + 0.1% (v/v) TFA, B = MeCN + 0.1% (v/v) TFA.

name	time [min]	A [%]	B [%]	name	time [min]	A [%]	B [%]
A5	0	90	10	A6	0	70	30
	18	20	80		18	0	100
	19	0	100		20	0	100
	21	0	100		21	30	30
	22	10	10		26	30	30
A7	27	10	10	A8	0	90	10
	0	80	20		18	20	80
	15	20	80		19	0	100
	16	0	100		21	0	100
	17	0	100		23	90	10
A9	17.5	20	20	25	90	10	
	29	20	20	A10	0	95	5
	0	95	5		18	30	70
	18	0	100		19	0	100
	20	0	100		21	0	100
21	95	5	21.5		95	5	
A6	24	95	5	24	95	5	
	0	95	5	A6	0	95	5
	18	50	50		18	50	50
	20	0	100		20	0	100
	21	0	100		21	0	100
26	95	5	26		95	5	

### LaChrom ELITE<sup>®</sup> HPLC System

Analytical HPLC was carried out on a LaChrom ELITE<sup>®</sup>-HPLC-System from VWR-Hitachi (VWR International GmbH, Darmstadt, Germany). The system contains an organizer, two HPLC-pumps (L-2130) with solvent degaser, an autosampler (L-2200) with a 100 μL sample loop, and a diode array flow detector (L-2455), and a high-pressure gradient mixer. A Luna<sup>®</sup> C8(2) column (5 μm, 100 Å, 250 × 4.6 mm, Phenomenex<sup>®</sup>, Torrance, CA, USA) or a Kinetex<sup>®</sup> C18 column (5 μm, 100 Å, 250 × 4.6 mm, Phenomenex<sup>®</sup>, Torrance, CA, USA) a Jupiter<sup>®</sup> C18 column (5 μm, 300 Å, 250 × 4.6 mm, Phenomenex<sup>®</sup>, Torrance, CA, USA) was used. A SecurityGuard<sup>™</sup> Cartridge



Kit (Ea, Phenomenex<sup>®</sup>, Torrance, CA, USA) holding a C8 or C18 SecurityGuard™ cartridge (4 × 3.0 mm, Phenomenex<sup>®</sup>, Torrance, CA, USA), respectively, served as pre-column. A flow rate of 1.0 mL/min was applied with deionized water and MeCN, each containing 0.1% (v/v) TFA, as eluents. The used gradient methods are shown in Table 14.14.

**Table 14.14:** Gradients used in analytical runs by LaChrom ELITE<sup>®</sup> HPLC system. Eluents: A = water + 0.1% (v/v) TFA, B = MeCN + 0.1% (v/v) TFA.

name	time[min]	A [%]	B [%]	name	time [min]	A [%]	B [%]
A11	0	90	10	A12	0	70	30
	18	20	80		18	0	100
	19	0	100		20	0	100
	21	0	100		21	30	30
	22	10	10		26	30	30
	27	10	10				
A13	0	80	20	A14	0	90	10
	15	20	80		18	20	80
	16	0	100		19	0	100
	17	0	100		21	0	100
	17.5	20	20		23	90	10
	29	20	20		25	90	10
A15	0	95	5	A16	0	95	5
	18	0	100		18	30	70
	20	0	100		19	0	100
	21	95	5		21	0	100
	24	95	5		21.5	95	5
					24	95	5
A17	0	95	5				
	5	95	5				
	6	80	20				
	36	50	60				
	38	0	100				
	40	0	100				
	41	95	5				
	47	95	5				

### **Primaide™ DAD system**

Analytical HPLC was carried out on a Primaide™ DAD system (VWR/Hitachi, Germany). The system works with a low-pressure gradient containing an HPLC pump (1110) with a 6-channel solvent degasser, an organizer, an autosampler (1210) with a 100 µL sample loop, a column oven (1310) and a diode array detector (1430). A Luna<sup>®</sup> C8(2) column (5 µm, 100 Å, 250 × 4.6 mm, Phenomenex<sup>®</sup>, Torrance, CA, USA) or a Kinetex<sup>®</sup> C18 column (5 µm, 100 Å, 250 × 4.6 mm, Phenomenex<sup>®</sup>, Torrance, CA, USA) or a Jupiter<sup>®</sup> C18 column (5 µm, 300 Å, 250 × 4.6 mm, Phenomenex<sup>®</sup>, Torrance, CA, USA). A SecurityGuard™ Cartridge Kit (Ea,

Phenomenex<sup>®</sup>, Torrance, CA, USA) holding a C8 or C18 SecurityGuard<sup>™</sup> cartridge (4 × 3.0 mm, Phenomenex<sup>®</sup>, Torrance, CA, USA), respectively, served as pre-column. A flow rate of 1.0 mL/min was applied with deionized water and MeCN, each containing 0.1% (v/v) TFA, as eluents. The used gradient methods are shown in **Table 14.15**.

**Table 14.15:** Linear gradients applied in analytical runs on the Primaide system.

name	time [min]	A [%]	B [%]	name	time [min]	A [%]	B [%]
<b>A18</b>	0	90	10	<b>A19</b>	0	70	30
	18	20	80		18	0	100
	19	0	100		20	0	100
	21	0	100		21	30	30
	22	10	10		26	30	30
	27	10	10				
<b>A20</b>	0	95	5	<b>A21</b>	0	95	5
	5	95	5		5	95	5
	6	80	20		6	90	10
	36	50	60		36	20	80
	38	0	100		38	0	100
	40	0	100		40	0	100
	41	95	5		41	95	5
47	95	5	47	95	5		

#### 14.1.5.5 Ion Exchange (IEC) and Size Exclusion Chromatography (SEC)

For the 1FYN samples, derivatives with a solubility tag were also produced in this doctoral thesis. These were purified by ion exchange chromatography after a reaction with the endoprotease TEV. In addition, the purified untagged FYN samples were purified by gel filtration (size exclusion chromatography) prior to possible crystallization experiments. Both procedures were carried out in the C. Roth group (Max Planck Institute of Colloids and Interfaces, Biomolecular Systems, Berlin).

The following equipment and parameters were used:

Cation exchange HPLC ÄKTA pure (Cytiva Europe GmbH, Freiburg, Germany) was used to analyze the products of the TEV protease cleavage reaction on the Polyarginine tag and ENLYFQ attached 1FYN. The Mono Q 4.6/100 PE HPLC column was used. A summary of prepared buffers and the employed gradient is provided in Table 14.16.

**Table 14.16:** Composition of applied buffers (wash and eluent) for the IEC.

wash (A)	eluent (B)	gradient
20 mM Tris base pH 8.0	20 mM Tris base 1 M NaCl pH 7.5	0% → 70% B in A 25 column volumes 70% → 100% B in A 3 column volumes 100% B in A 3 column volumes

For the size exclusion chromatography, HPLC ÄKTA pure (Cytiva Europe GmbH, Freiburg, Germany) pure was equipped with a HiLoad™ 16/600 Superdex™ 75 µg column. The prepared buffers are summarized in Table 14.17.

**Table 14.17:** Composition of buffers used in SEC.

eluent (A)	wash (B)
20 mM Tris base 8 pH.0	1 M NaOH pH 7.5

#### 14.1.6 Lyophilization

Cleaved and purified peptides were lyophilized with a laboratory freeze dryer ALPHA 1-2 LD (Christ Gefriertrocknungsanlagen GmbH, Osterode am Harz, Germany) was used. A chemistry hybrid pump RC 6 (Vacuubrand GmbH + Co KG, Wertheim, Germany) was connected.

#### 14.1.7 Mass spectrometry

High-resolution mass spectrometry was conducted on an Agilent 6220 ESI-ToF LC-MS spectrometer (Agilent Technologies Inc., Santa Clara, CA, USA). Mass-to-charge ratios were also determined on an Agilent Technologies 6230 TOF LC/MS instrument with an ESI injector (Agilent Technologies Inc., Santa Clara, CA, USA).

To determine the mass of the synthesized peptides, high-resolution mass spectrometry (HRMS) was conducted on an Agilent ESI-ToF 6230 LC-MS spectrometer (Agilent Technologies Inc., Santa Clara, CA, USA). Parameters were optimized for the measurements: spray voltage was set to 3.5 kV, nebulizer to 20 psi, gas temperature to 325°C, fragmentor voltage to 250 V, and drying gas flow rate to 5 L/min. All samples were dissolved in a mixture of water and acetonitrile (1/1, v/v) containing 0.1 % TFA (v/v). The infusion was carried out at a flow rate of

10  $\mu\text{L}/\text{min}$  via a Harvard Apparatus 11 plus syringe pump (Harvard Apparatus, Holliston, MA, USA).

For the depiction and analysis of the data, the MassHunter Workstation software version B.02.00 or version B.08.00 (Agilent Technologies, Santa Clara, CA, USA) or MestReNova version 14.2.2-28739 (Mestrelab Research S. L., Santiago de Compostela, Spain) Mestrenova was used. Calculated molar masses and mass to charge residues were obtained from the peptide mass calculator v3.2(<http://rna.rega.kuleuven.be/masspec/pepcalc.htm>).

### **14.1.8 SDS-PAGE**

These experiments were conducted at Roth Group at MPI. For all measurements, a vertical electrophoresis apparatus without special cooling was used with continuous current (initial 10-15 mA/gel for the stacking gel and 30 mA/gel for the separation gel).

#### **14.1.8.1 Sample Preparation**

The samples were prepared in a total amount of 20  $\mu\text{L}$  in the desired buffer, and the concentration was determined ( $\sim 0.2$  mg/mL) by nanodrop measurement on the Spektralfotometer/Fluorometer DS-11 FX+ (“Nanodrop”, DeNovix Inc. Wilmington, Delaware, USA ). The final peptide concentration was adjusted to 0.2 mg/mL by measurements on the “Nano-drop”. The final samples were treated with 5  $\mu\text{L}$  loading dye, incubated at 95 °C for 5 min, and centrifuged before being transferred into the gel. The Page Ruler Low Range Unstained Protein (Fischer Scientific) Ladder with a range from 3.4-100 kDa was used as an internal standard with eight defined markers and a highlighted reference at 25 kDa.

#### **14.1.8.2 Composition of Gels**

In course of this presented doctoral thesis, two different gel setups were used: On one hand, BIS-TRIS gels and on the other hand, Tricine-gels. <sup>[3]</sup> The gel compositions and the applied running buffers are summarized in Table 14.18 and Table 14.19.

**Table 14.18:** Gel compositions for the SDS-PAGE with BIS-TRIS gels.

<b>BIS-TRIS</b>		<b>Running Buffer</b>
<b>12% Separating gel</b>	<b>4% Stacking Gel</b>	
1.5 M Bis-Tris buffer (pH 6.8) <b>3.4 mL</b> H <sub>2</sub> O deionized <b>8.66 mL</b> Bis-/Acrylamide (40 %, 39:1) <b>4 mL</b> TEMED <b>40 μL</b> APS (10%) <b>80 μL</b>	1.5 M Bis-Tris buffer (pH 6.8) <b>1.7 mL</b> H <sub>2</sub> O deionized <b>6 mL</b> Bis-/Acrylamide (40 %, 39:1) <b>1.5 mL</b> TEMED <b>20 μL</b> APS (10 %) <b>40 μL</b>	MES

**Table 14.19:** Gel compositions for the SDS-PAGE with Tricine gels.

<b>Tricine</b>	<b>4% Stacking Gel</b>	<b>Gel Buffer</b>	<b>Anode Buffer (10x)</b>	<b>Cathode Buffer (10x)</b>
<b>16% Separating Gel</b>				
Gel buffer <b>6.6 mL</b> H <sub>2</sub> O deionized <b>3.4 mL</b> Glycerol <b>3 mL</b> Urea <b>7.2 g</b> Bis-/Acrylamide (40 %, 19:1) <b>8 mL</b> TEMED <b>20 μL</b> APS (10 %) <b>200 μL</b>	Gel buffer <b>3 mL</b> H <sub>2</sub> O deionized <b>8 mL</b>  Bis-/Acrylamide (40 %, 19:1) <b>1 mL</b> TEMED <b>10 μL</b> APS (10 %) <b>100 μL</b>	3 M Tris base  0.3 % SDS pH 8.45	1 M Tris base  pH 8.9	1 M Tris base 1 M Tricine 1 % SDS pH not adjusted

### 14.1.8.3 Staining of the Gels

Tricine gels were cross-linked with 15 mL 5% (v/v) glutaric aldehyde in water (incubate 60 min, wash three times with water, and keep the gel in water overnight) to avoid migration of the small peptides.

The finished gel was dyed with “magic dye “ containing Coomassie Brilliant Blue G250 10.3791/1350 (2009). The gel was incubated with the dye for 5 min at 95 °C.

### 14.1.9 UV/ Vis-spectroscopy

UV-spectra were measured using a Varian Cary 50 spectrophotometer (Varian Medical Systems, Palo Alto, CA, USA) and quartz cuvettes (10 mm path length, 1.5 mL, Hellma Analytics, Müllheim, Germany) for wavelengths < 300 nm.

All obtained spectra are baseline corrected with a sample measurement of the respective buffer used as a blank for the subsequent triplicate measurements.

Alternatively, the absorbance was measured on the Eppendorf *BioPhotometer plus* using plastic cuvettes (UVette® routine pack, Eppendorf SE, Hamburg, Germany) or on the Spektrofotometer/Fluorometer DS-11 FX+ (“Nanodrop”, DeNovix Inc. Wilmington, Delaware, USA).

### 14.1.9.1 Peptide Concentration Determination Procedure

Peptide involved in both coiled coil projects were stored in 1 mL HFIP stock solution (with an approximate concentration of 2 mg/mL). A defined aliquot (10 µL) of each solution was taken, and HFIP was gently removed in a stream of nitrogen. The remaining peptide films were dissolved in the buffer for absorption measurements. The proteins Ub, 1FYN, 1TEN, and their synthesized derivatives were stored in buffer solutions.

UV/Vis measurements were generally performed to determine the absorbance at different wavelengths. Using the Lambert-Beer law, the peptide concentration was determined.

$$E_{\lambda} = \varepsilon_{\lambda} \cdot c \cdot d \quad (2)$$

Where  $E_{\lambda}$  = measured absorbance at the wavelength,  $\varepsilon_{\lambda}$  = extinction coefficient of the peptide at the wavelength  $\lambda$  [ $M^{-1} \text{ cm}^{-1}$ ],  $c$  = concentration of the peptide [ $\text{mol L}^{-1}$ ],  $d$  = path length of the cell [cm].

Two different calculation methods were used based on the primary structure of the peptides to be investigated.

### **General Procedure for Peptides Containing Cys, Trp or Tyr**

Pace *et al.* suggested a method to calculate the extinction coefficient of peptides in buffer solution without 6 M Gnd·HCl:<sup>[4]</sup>

$$\varepsilon_{280\text{nm}} = \#Trp \cdot 5500 + \#Tyr \cdot 1490 + \#Cys \cdot 125 \quad (3)$$

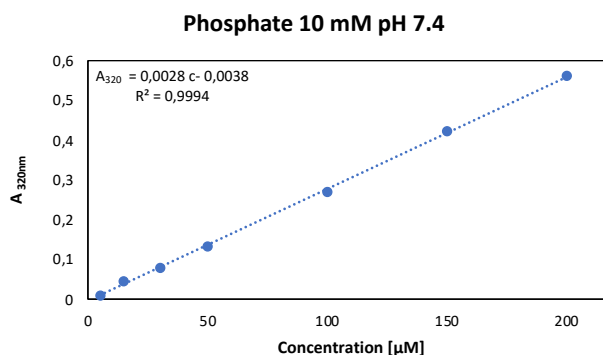
With  $\varepsilon_{280\text{nm}}$  = extinction coefficient at 280 nm, # = number of corresponding AA in the primary structure of the peptide.

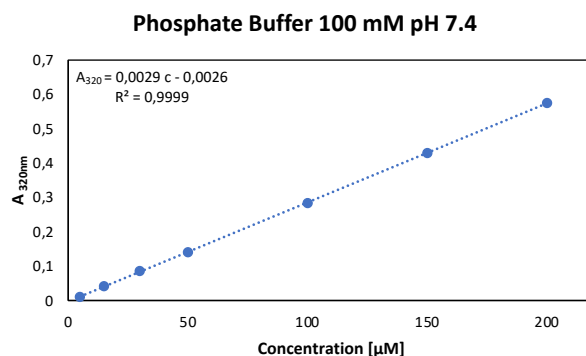
**Table 14.20:** Calculated extinction coefficient of used peptides in this doctoral thesis according to equation (3).

peptide	calc. $\epsilon_{280\text{nm}}$ [ $\text{M}^{-1}\text{cm}^{-1}$ ]
GPR83-18-74	15470
MOG-FITC	8480
PLP-FITC	5500
A4 peptides	5625
B4 peptides	1615
Ub and derivatives	1490
1FYN and derivatives	16960
1TEN and derivatives	9970

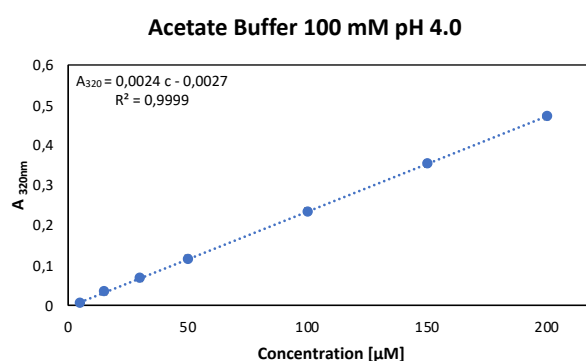
### ***General Procedure for Peptides Containing Abz***

UV/Vis measurements determined concentrations of Abz-containing peptides at 320 nm in a buffer solution containing 6 M Gdn·HCl. Calibration curves for different buffers were recorded with a concentration series of Abz-Gly-COOH · HCl (Bachem AG, Bubendorf, Switzerland) in the selected buffer. The concentration of the peptide stock solutions can be calculated using the measured absorbance, rearranging the formula of the linear regression line of the calibration curve for the used buffer. For each buffer used for CD measurements, a calibration curve was recorded.

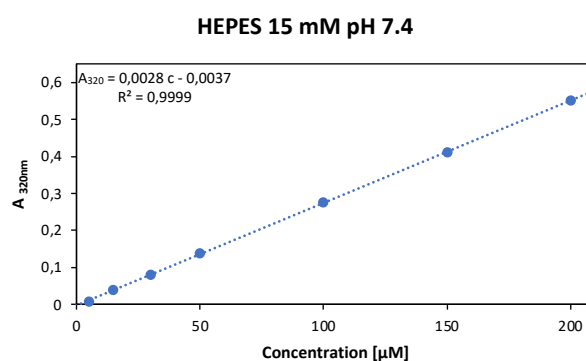
**Figure 14.1:** Calibration line for peptide concentration determination; recorded in 10 mM phosphate buffer with pH 7.4 at 20°C containing 6 M Gdn · HCl.



**Figure 14.2:** Calibration line for peptide concentration determination; recorded in 100 mM phosphate buffer with pH 7.4 at 20°C containing 6 M Gdn · HCl.



**Figure 14.3:** Calibration line for peptide concentration determination; recorded in 100 mM acetate buffer with pH 4.0 at 20°C containing 6 M Gdn · HCl.



**Figure 14.4:** Calibration curve for peptide concentration determination; recorded in 15 mM HEPES buffer with pH 7.4 at 20°C containing 6 M Gdn·HCl.

### 14.1.10 CD spectroscopy

All measurements were conducted in a quartz cuvette of 0.1 cm path length (Hellma Analytics QS High Precision Cell) using a JASCO-810 spectropolarimeter (JASCO Deutschland GmbH, Pfungstadt, Germany) equipped with a Jasco PTC-423S Peltier temperature control system and a HAAKE WKL water recirculator (Thermo Electron GmbH, Karlsruhe, Germany) for temperature control. Data analysis was



performed with the software J-700 (JASCO Deutschland GmbH, Pfungstadt, Germany).

The measurements were done in the UV-CD range of 190 nm to 250 nm, with 3 scans and duplicate measurements for each spectrum. The measured spectra were baseline-corrected using the respective buffer as a blank. All further parameters are summarized in Table 14.21.

**Table 14.21:** Parameters of CD spectroscopy used in this doctoral thesis.

parameter	value
sensitivity	standard (100 mdeg)
data pitch	0.2 nm
scanning mode	continuous
scanning speed	100 nm/min
response	2 s
bandwidth	2 nm
accumulation	3
N2 flow	3 L/min

#### 14.1.10.1 Conversion of the Spectra to Mean Residue Molar Ellipticity [ $\theta$ ]

The measured spectra were normalized using equation (4)

$$[\theta] = \frac{\theta}{n \times l \times c \times 1000} \text{ [mdeg cm}^2 \text{ dmol}^{-1} \text{ residue}^{-1}] \quad (4)$$

with  $\theta$  = measured value [mdeg],  $n$  = # of amide bonds in peptide,  $l$  = light path (0.1 cm)  
 $c$  =: concentration of peptide solution [mol/L]

#### 14.1.10.2 Thermal denaturation

For thermal denaturation curves, the CD signal at 222 nm was recorded by employing the following parameters described in Table 14.22. The obtained melting curves were normalized according to equation (4) and fitted by adapting a two-state unfolding to determine the melting temperature.

Next, the fraction unfolded was calculated:

$$f = \frac{\theta_{222} - \theta(\text{folded})_{222}}{\theta(\text{unfolded})_{222} - \theta(\text{folded})_{222}} \quad (5)$$

where  $f$  = unfolded fraction,  $\theta_{222}$  = measured ellipticity at 222 nm,

$\theta(\text{folded})_{222}$  = ellipticity at 222 nm of folded peptide (linear fitting line of the lower plateau of Boltzmann fit)

$\theta(\text{unfolded})_{222}$ : ellipticity at 222 nm of unfolded peptide (linear fitting line of the higher plateau of Boltzmann fit)

**Table 14.22:** Parameters of thermal denaturation by CD measurements.

FF03-project		A4/B4 system	
parameter	value	parameter	value
peptide conc.	100 $\mu\text{M}$	peptide conc	50 $\mu\text{M}$ (total, i.e. 25 $\mu\text{M}$ per strand)
sensitivity	standard (100 mdeg)	sensitivity	standard (100 mdeg)
starting temperature	25°C	starting temperature	20°C
ending temperature	100°C	ending temperature	100°C
data pitch	0.1°C	data pitch	0.1°C
delay time	0 s	delay time	0 s
temperature slope	3°C/min	temperature slope	1°C/min
response	1 s	response	1 s
bandwidth	1 nm	bandwidth	1 nm
N2 flow	3 L/min	N2 flow	3 L/min

#### 14.1.11 Biological Studies: Instrumental Set-up and Procedures

Cell culture studies were performed at the Institute of Experimental Pediatric Endocrinology (IEPE, Charité Universitätsmedizin Berlin) with the following experimental setup:

##### Equipment:

**Table 14.23:** Devices employed in Bio-Assay.

Instrument	Company
Cell culture hood LaminAir HBB 2448	Heraeus Instruments, Hanau, DE
Centrifuge (SORVALL RC 6 Plus, 5417R / 5417C, Labofuge 200)	Thermo Scientific, Karlsruhe, DE; Eppendorf; Heraeus Instruments
Lab water supply Millipore Milli-Q® Biocel	Merck Millipore, Darmstadt, DE
Incubator (CO <sub>2</sub> incubator type BB 6220 O2, Incubator type B 6030)	Heraeus Instruments
LUNA™ Automated Cell Counter	Logos biosystems, Annandale, USA
Multitechnology microplate reader Mithras LB 940	Berthold Technologies, Regensdorf, CH
Scale (MC1 Laboratory LC 2200 P, CPA223S)	Sartorius, Goettingen, DE

Varioklav steam sterilizer

HP Medizintechnik GmbH, Oberschleissheim,  
DE**Cell line:** Embryonic Mouse Hypothalamus Cell LineN41Incubation with Peptide **12** ( GPR83\_57-69) (100 mM stock solution in 1 mM PBS buffer +0.1% BSA) + protease inhibitor.**Plasmids:**

Plasmids were provided by collaborators as DMSO-glycerol stock solutions and were purified following the manufacturer's protocol for purification kits.

**Table 14.24:** Origin of plasmids used in BioAssays.

number:	insert	plasmid backbone	origin	meaning in experiment
<b>P1469</b>	NHA-TSHR	pcDps; Amp <sup>r</sup>	cloned from human DNA of the thyroid	Other GPCR Positive control
<b>P1596</b>	GPR83 wt (human)	pcDps; Amp <sup>r</sup>	Cloned from human hypothalamic cDNA	Probe
<b>P1337</b>	luc2P/NFAT-RE/Hygro	pGL4.30; Hygro <sup>r</sup>	Promega	Transcriptionsfactor
<b>P798</b>	Empty vector	pcDps; Amp <sup>r</sup>	cloned from human cDNA (UMR cDNA Resource Center, Rolla, MO, USA)	Negative control
<b>P1621</b>	GPR83/_del17-60/-67 (human)	pcDps; Amp <sup>r</sup>	Cloned from human hypothalamic cDNA	Probe

**Kits:**

The PureYield™ Plasmid Midiprep System was used to purify all control samples while the GPR83 containing were purified with NucleoBond® Xtra Midi.

Kit	Supplier
NucleoBond® Xtra Midi	Macherey-Nagel, Düren, DE
PureYield™ Plasmid Midiprep System	Promega

### Buffers, reagents, and stimulation agents

Buffer / solution	Components/ Company/ Comments
Ampicillin stock solution	50 mg/mL Ampicillin
cAMP assay medium	138 mM NaCl, 6 mM KCl, 1 mM MgCl <sub>2</sub> · 6 H <sub>2</sub> O, 5.5 mM Glucose, 20 mM HEPES, 1 mM CaCl <sub>2</sub> * 2 H <sub>2</sub> O, 0.1% BSA pH 7.4
70 % Ethanol	70 % (v/v) (96% Ethanol diluted in Milli-Q-water)
LO buffer (basic solution for LI buffer)	0.1% BSA (Bovine Serum Albumin); 0.3% Tween <sup>®</sup> 20; 5 nM HEPES pH 7.4
LI buffer	1 mM IBMX in LO buffer
3-Isobutyl-1-methylxanthine (IBMX) stock solution	500 mM IBMX (Sigma-Aldrich), dissolved in DMSO

### Commercially purchased reagents and solutions:

#### Buffers:

PBS Dulbecco (w/o Ca <sup>2+</sup> , w/o Mg <sup>2+</sup> ; instamed 9.55 g/L; 1 x: diluted in 5 L Milli-Q water)	Biochrom AG, Berlin, DE
---	-------------------------

#### Stimulating reagents:

<b>Forskolin</b>	10 mM, in 100% DMSO	BioChemica - AppliChem, Darmstadt, DE
<b>Thyrotropic Hormone, from bovine pituitary (bTSH)</b>	10 IU/mL (international Units)	Sigma-Aldrich

#### Cell culture reagents

Dulbecco's MEM (w 3.7 g/L NaHCO <sub>3</sub> , w 4.5 g/L d-Glucose, w/o L-Glutamine, w/o Na-Pyruvate, low endotoxin)	Biochrom AG
Fetal Bovine Serum (FBS)	Biochrom AG
Metafectene <sup>®</sup>	Biontex Laboratories GmbH, Martinsried, DE
Phosphate buffered saline (PBS), 1 x, Dulbecco (w/o Ca <sup>2+</sup> , w/o Mg <sup>2+</sup> , low endotoxin)	Biochrom AG
Penicillin/ Streptomycin (10,000 U/mL/10,000 µg/mL)	Biochrom AG

#### 14.1.11.1 Cell Culture Methods

##### *Cultivation of cell line*

N41 cells were cultivated in polystyrene 75 cm<sup>2</sup> culture flasks (Sarstedt) under standard cell culture conditions.

The cells were washed with Dulbecco's PBS and trypsinized using 1x Trypsin/EDTA solution in PBS until detachment for passaging. Light tapping facilitated the harvesting of the cells, which were re-suspended in a serum-containing medium to stop trypsinization. Then, the cells were split and seeded into 96-well plates.

##### *Transfection*

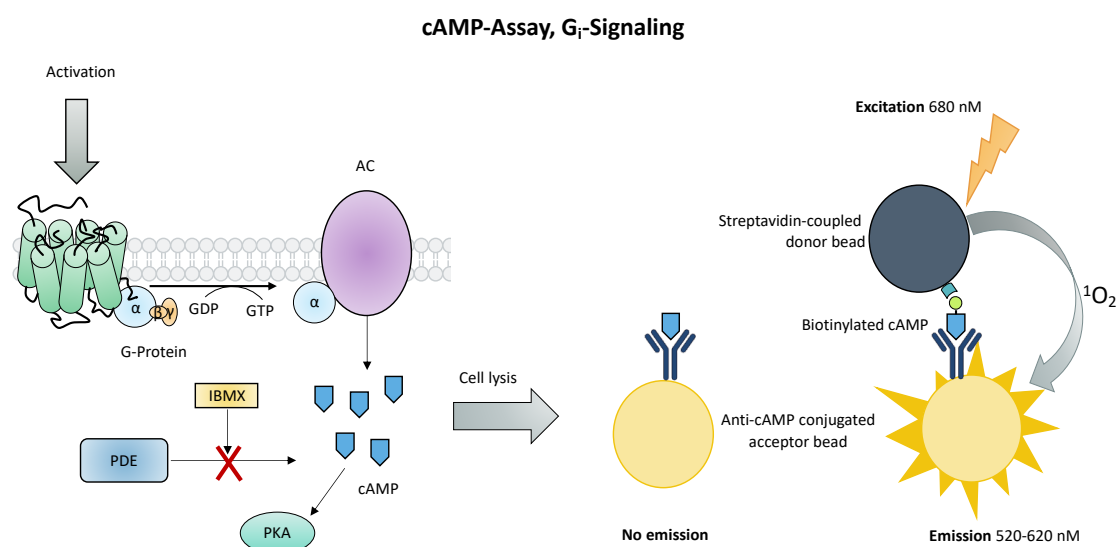
**Table 14.25:** Transfection conditions for scheduled assays.

Assay	Cell line	Cells per well/dish	Format	Amount of DNA per well/dish	Metaftec-tene <sup>®</sup> per well/dish
cAMP accumulation assay (G <sub>i</sub> )	N41	1.5 * 10 <sup>5</sup>	96-well	62.4 ng (R-A/+R-B)	0.47 µL
IP3 accumulation assay (G <sub>q/11</sub> )	N41	1.5 * 10 <sup>5</sup>	96-well	62.4 ng (R-A/+R-B) 62.4 ng ( <i>luc2P/NFAT</i> )	0.47 µL

#### 14.1.12 Activation Studies of GPR83

The GPR83 variants were characterized concerning their capacities to activate G<sub>i</sub>- and G<sub>q</sub>- pathways. Here, two assays investigated the accumulation of the second messengers IP3 (G<sub>q</sub>) and cAMP (cyclic adenosine monophosphate).

### 14.1.12.1 cAMP accumulation assay for measuring G<sub>i</sub>-activity



**Figure 14.5:** G<sub>i</sub>-activity assay. AC = adenosyl cyclase, PKA = protein kinase A, PDE = phosphodiesterase, IBMX = 3-isobutyl-1-methylxanthine.

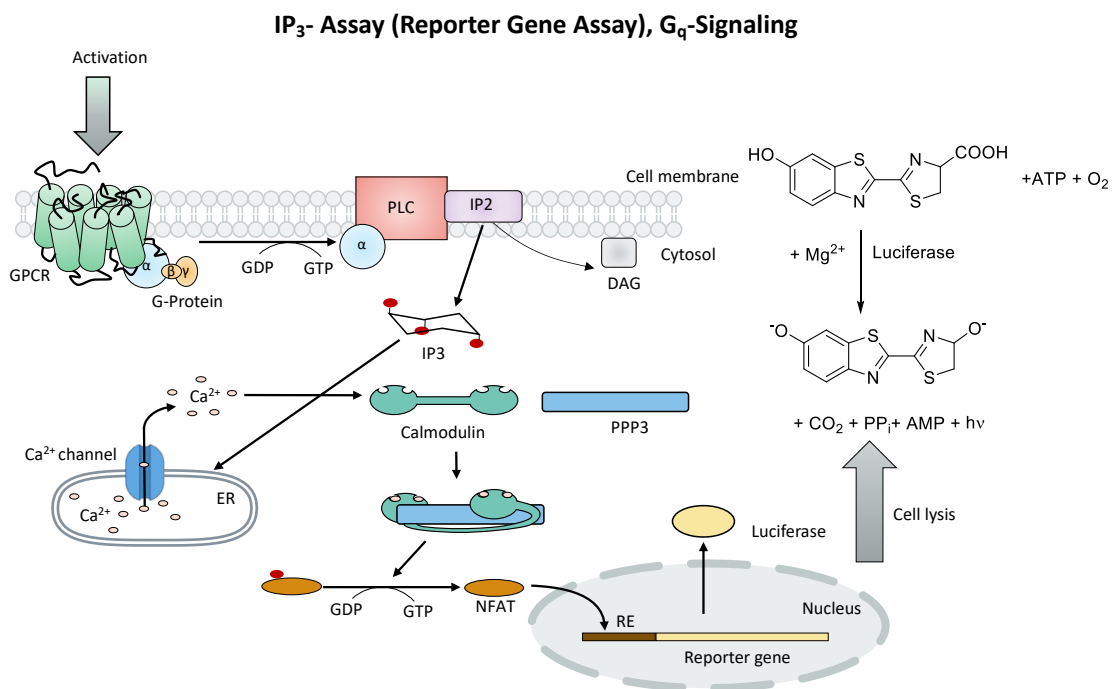
The AlphaLISA (Alpha: amplified luminescent proximity homogeneous assay) technology of PerkinElmer was applied to measure the intracellular cAMP accumulation and, thus, to determine the G<sub>i</sub>-activity of GPR83. Cyclic adenosine monophosphate (cAMP) is synthesized from ATP. G<sub>i</sub>-signaling inhibits this enzyme and, therefore, results in reduced cAMP levels. In order to measure the suppressive effect on the intracellular cAMP concentration of G<sub>i</sub>-active GPCRs, a pre-stimulation with forskolin, a known adenylyl cyclase enhancer, is necessary. IBMX blocks the consumption of produced cAMP by PDE. The assay is based on a competitive principal of endogenous cAMP and exogenous biotinylated cAMP for the binding sites of the anti-cAMP antibody. Due to the conjugated biotin, the exogenous cAMP can be detected by streptavidin-coupled donor beads. When donor and acceptor beads are in close proximity (anti-CAMP antibody-conjugated), the induced excitation of the donor beads leads to measurable light emission from the acceptor beads. The measured light signal is thus inversely correlated with the concentration of endogenous cAMP, since a low signal indicates that many acceptor beads are occupied by intracellular cAMP. Therefore, only a few donor and acceptor beads come very close to each other.

This assay was conducted using N41 cells as described in **Table 14.25**. After 48 h post-transfection, the cells were pre-stimulated with 50  $\mu\text{L}/\text{well}$  1 mM IBMX/cAMP assay medium for 5 min at 37  $^{\circ}\text{C}$ .

If the GPCR should be activated by ligand interaction, 50  $\mu\text{L}$  of ligand dilution was added to the cells and incubated for 45 min at 37  $^{\circ}\text{C}$ . Then, the mixture was sucked off and the cells were lysed at 4  $^{\circ}\text{C}$  for 1.5 h on a shaker by adding 100  $\mu\text{L}/\text{well}$  LI buffer. For the actual measurement, 5  $\mu\text{L}$  of lysed cell suspension was transferred to a white 384-well plate.

A cAMP standard curve, ranging from 0 to 10  $\mu\text{M}$ , was recorded for means of quantification. Initially, anti-cAMP antibody-conjugated acceptor beads in a concentration of 1 U beads in 10  $\mu\text{L}$  ALPALisa buffer/ well were incubated with the suspension for 30 min at rt in the dark. This step was followed by the addition of 10  $\mu\text{L}/\text{well}$  of a mixture of streptavidin-coupled donor beads and biotinylated cAMP (1 U beads and 1 U biotinylated cAMP in 10  $\mu\text{L}$  ALPHALisa buffer) with an incubation time of 1 h at rt in the dark. The emission at 520- 620 nm was measured by the Mithras LB 940 (Berthold technologies) after excitation with 680 nm.

#### 14.1.12.2 Reporter gene assays for measuring $G_q$ activity



**Figure 14.6:**  $G_q$  activity assay. IP<sub>3</sub> = Inositol Trisphosphate, DAG = diacylglycerol, RE = responsive element.

The Gq activity was investigated by a reporter gene assay. This system is based on the co-transfection of the gene of interest with a reporter gene construct encoding the *Photinus pyralis* (firefly) luciferase. Luciferase transfection is controlled by a responsive element (RE), which tailors the synthesis of the luciferase protein to the different signaling pathways under investigation. Signaling either generates a responsive factor for each RE or forms a protein fundamental to complex formation, which in turn can bind to the RE and trigger luciferase transcription. Therefore, a high luciferase signal corresponds to a high activity of the signaling pathway.

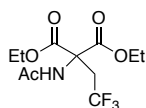
48 h post-transfection, the cells were stimulated with an adequate ligand for 6 h diluted in a serum-lacking medium at 37 °C and afterwards, lysed with 1x Passive Lysis Buffer (PLB, 50 µl/well) for 15 min at rt on a shaker. 10 µl of the lysed solution were transferred to a black 96-well plate. For the measurement, 40 µl of luciferase substrate solution were added by the Mithras LB 940 and the light outcome was determined. The substrate luciferol is converted by ATP hydrolysis and enzymatic catalysis (luciferase) to oxyluciferol, AMP, PPI (pyrophosphate), CO<sub>2</sub> and light. Coenzyme A is additionally used to stabilize the light signal.

## 14.2. Organic Synthesis

### 14.2.1 Synthesis of (S)-2-amino-4,4,4-trifluorobutanoic acid (TfeGly)

The synthesis scheme was followed by Tsushima *et al.* [5]

#### 14.2.1.1 Diethyl 2-acetamido-(2,2,2-trifluoroethyl) malonate (5)



In a two-neck flask, diethyl 2-acetamidomaalonate (9.34 g, 43.0 mmol, 1.0 eq.) and potassium *tert*-butoxide (4.83 g, 43.0 mmol, 1.0 eq.) were dissolved in 80 mL of dry THF under argon atmosphere. The reaction mixture was heated to 70 °C and refluxed for two hours. Afterward, the resulting white suspension was cooled to RT and 2,2,2-trifluoroethyl trifluoromethanesulfonate (6.81 mL, 47.3 mmol, 1.1 eq.) was added by a syringe. The resulting reaction mixture was refluxed at 70 °C for two days. THF was removed under reduced pressure, and the residue was treated with a 1 M HCl solution (50 mL). The reaction mixture was extracted with ethyl acetate

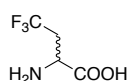


(4 × 50 mL), washed with water (2 × 20 mL), dried over anhydrous sodium sulfate, and concentrated *in vacuo*. The crude product was purified by silica column chromatography using n-hexane/EtoAc (3:1 %v/v) as eluent to yield **5** as a colorless solid (5.7 g, 19 mmol, 44% yield).

**<sup>1</sup>H-NMR** (400 MHz, CDCl<sub>3</sub>): δ [ppm] = 6.93 (s, 1H), 4.32-4.20 (m, 4H, (2x2H)), 3.37-3.25 (m, 2H), 2.11-2.04 (m, 3H), 1.25 (t, *J* = 7.1 Hz, 6H (2x3H)).

**<sup>19</sup>F-NMR** (376 MHz, CDCl<sub>3</sub>): δ [ppm] = -62.02 (t, *J*<sub>HF</sub> = 10.5 Hz, 3F).

#### 14.2.1.2 2-Amino-4,4,4-trifluorobutanoic acid (**6**)

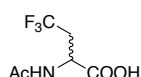


Diethyl 2-acetamido-(2,2,2-trifluoroethyl) malonate (**5**) (5.70 g, 19.6 mmol) was dissolved in conc. HCl (20 mL) and refluxed overnight. The reaction mixture was concentrated *in vacuo* and the product was dried by lyophilization (2.9 g, 18.4 mmol, 97 %). The product **6** was used in the next step without further purification.

**<sup>1</sup>H-NMR** (400 MHz, D<sub>2</sub>O): δ [ppm] = 4.39 (dd, *J* = 8.1, 4.0 Hz, 1H), 2.89-3.13 (m, 2H).

**<sup>19</sup>F-NMR** (376 MHz, D<sub>2</sub>O): δ [ppm] = -64.06 (t, *J*<sub>HF</sub> = 10.2 Hz, 3F).

#### 14.2.1.3 2-Acetamido-4,4,4-trifluorobutanoic acid (**7**)



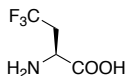
Racemic 2-amino-4,4,4-trifluorobutanoic acid (**6**) (2.0 g, 12.7 mmol, 1.0 eq.) was dissolved in 30 mL of water (*Milli-Q*) and the pH was adjusted to 9 with 1 M NaOH. Then, the reaction mixture was cooled to 0 °C with an ice bath. Acetic anhydride (1.34 mL, 14 mmol, 1.1 eq.) was carefully added to the reaction mixture and the pH was adjusted to 9. Subsequently the mixture was stirred for 30 min at 0°C, warmed to rt, and stirred for additional 2 h at the same temperature. Next, conc. HCl was added slowly until a pH of 2 was reached. The reaction mixture was extracted with ethyl acetate (5 × 30 mL). The organic layers were combined, dried over sodium sulfate and the solvent was removed *in vacuo* to give the desired product as a colorless solid (1.5 g, 7.7 mmol, 61% yield).

**<sup>1</sup>H NMR** (400 MHz, D<sub>2</sub>O): δ [ppm] = 4.42(dd, *J* = 7.7, 4.3 Hz, 1H), 2.69-2.92 (m, 2H), 2.01 (s, 3H).

**<sup>19</sup>F NMR** (376 MHz, D<sub>2</sub>O): δ [ppm] = -64.30 (t, *J*<sub>HF</sub> = 10.3 Hz, 3F).

#### 14.2.1.4 (S)-2-Amino-4,4,4-trifluorobutanoic acid (8)

The synthesis was performed according to describe procedures by Lindner *et al.* [6]

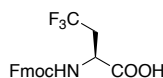


2-acetamido-4,4,4-trifluorobutanoic acid (**7**) (1.5 g, 7.7 mmol) was dissolved in 20 mL water (*Milli-Q*, generating a 50 mM solution). With 1 M NaOH the pH was adjusted to 11. Then, a 10% AcOH solution was added until pH 7 was reached, creating a buffer system. Next, Acylase I from the porcine kidney (553 units/mg) was dissolved in water (*Milli-Q*) (5 mg/100 mg AA). The solution was added to the AA solution, and the pH was adjusted to 7.0 using a 0.1% AcOH and a 10 mM NaOH solution. Subsequently, the reaction mixture was stirred slowly for 4 h at 36 °C and overnight at rt. To quench the reaction, the pH was adjusted to 4. Ion-exchange beads (Dowex 50WX8, 100-200 mesh) were added, and the reaction mixture was stirred slowly for 4 h. Completion of the reaction was verified by Ninhydrin reaction. Finally, the resin was packed into a column and washed with water until a neutral pH was reached. The L-AA was eluted from the resin with a 1 M NH<sub>3</sub> solution, and all fractions with a positive Ninhydrin staining were combined and dried by lyophilization. The product was obtained as a colorless solid (583 mg, 3.7 mmol, 48% yield).

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O): δ [ppm] = 3.75 (dd, *J* = 8.1, 4.4 Hz, 1H), 2.58-2.89 (m, 2H).

<sup>19</sup>F NMR (376 MHz, D<sub>2</sub>O): δ [ppm] = -64.06 (t, *J*<sub>HF</sub> = 10.2 Hz, 3F).

#### 14.2.1.5 (S)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-4,4,4-trifluorobutanoic acid (1)



L-TfeGly (0.5 g, 3.1 mmol, 1.0 eq.) was dissolved in 10% Na<sub>2</sub>CO<sub>3</sub> (5 mL) and cooled to 0 °C in an ice bath. Afterwards, 1,4-dioxane (1 mL) was added dropwise to the reaction mixture, and it was stirred for 15 min. Next, Fmoc-succinimide (1.15 g, 3.4 mmol, 1.1 eq.) was added. The mixture was stirred initially for three hours on ice and subsequently at room temperature overnight. The reaction was quenched with water (50 mL) and extracted with diethyl ether (50 mL). The aqueous phase was cooled to 0 °C in an ice bath, and the pH was adjusted to 2, using conc. HCl. This lead

to the precipitation of a white solid. Then, the suspension was extracted with DCM (4 × 50 mL). The combined organic phases were washed with water (50 mL), dried over anhydrous sodium sulfate, and the solvent was removed *in vacuo*. Purification of the crude product was achieved using preparative HPLC with a Kinetex RP-C18 end-capped (5 μM, 100 Å, 250 × 21.2 mm, Phenomenex®, USA) reverse phase HPLC column. For this purpose, gradient **P7** was applied with a UV-detection at  $\lambda_{\text{abs}} = 280$  nm. The fractions containing pure Fmoc-amino acid were combined, MeCN was removed *in vacuo*. The pure product was obtained as a colorless solid after lyophilization (0.260 mg, 0.6 mmol, 22% yield).

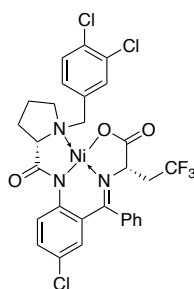
<sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  [ppm] = 7.77 (d,  $J = 7.5$  Hz, 2H), 7.58 (d,  $J = 7.5$  Hz, 2H), 7.42 (d,  $J = 7.5$  Hz, 2H), 7.38 (d,  $J = 7.5$  Hz, 2H), 5.53 (d,  $J = 9.4$  Hz, 1H), 4.64 (dd,  $J = 8.1, 4.4$  Hz, 1H), 4.40-4.25 (m, 2H), 4.23 (t,  $J = 6.1$  Hz, 1H), 2.85-2.72 (m, 2H).

<sup>19</sup>F-NMR (376 MHz, CDCl<sub>3</sub>):  $\delta$  [ppm] = -62.92 (t,  $J_{\text{HF}} = 10.2$  Hz, 3F).

#### 14.2.2 Synthesis of Fmoc-TfeGly (**1**) Following the Ni(II)-complex Strategy

Following the procedure of Han *et al.*, Fmoc-TfeGly was prepared with the chiral Ni(II) complex strategy.<sup>[7]</sup> The starting material **19** was kindly provided by Langhans.

##### 14.2.2.1 Alkylation of **19** with 1,1,1-Trifluoro-2-iodoethane to form **20**



In a nitrogen atmosphere, **19** (2.5 g, 4.1 mmol, 1eq.) and 1,1,1-trifluoro-2-iodoethane (430 μL, 4.3 mmol, 1.05 eq.) were dissolved in 25 mL DMF (dry degassed), and the mixture was cooled to 0-5°C. Subsequently, 2.25 mL of 10% KOH in MeOH was added, and the reaction mixture was stirred for 2 h at 0-5°C. The reaction was quenched with 8 mL of water and warmed up to rt. After 1 h, another 4.5 mL of water was added, and the resulting solution was again stirred for 1 h at rt. The precipitate obtained was separated and washed with a 1:1 mixture of DMF and water. The isolated product was dried overnight at 60°C. The product was obtained

as a dark-red solid, which was used in the next step without further purification (1.70 g, 2,5 mmol, 61% yield).

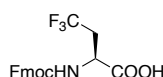
**<sup>1</sup>H-NMR** (600 MHz, CDCl<sub>3</sub>): δ [ppm] = 8.95 (s, 1H), 8.18 (d, *J*=9.3, 1H), 7.72 (d, *J*=8.1, 1H), 7.62 – 7.57 (m, 2H), 7.54 (d, *J*=7.5, 1H), 7.36 (d, *J*=8.1, 1H), 7.29 (d, *J*=6.5, 1H), 7.13 (dd, *J*=9.3, 2.5, 1H), 6.91 (d, *J*=7.6, 1H), 6.57 (d, *J*=2.6, 1H), 4.29 (d, *J*=12.6, 1H), 4.19 (dd, *J*=7.0, 3.0, 1H), 3.58 (dd, *J*=10.7, 6.7, 1H), 3.44 (q, *J*=10.5, 1H), 3.34 (dd, *J*=10.7, 6.3, 1H), 3.23 (d, *J*=12.7, 1H), 2.78 (dt, *J*=15.3, 8.1, 1H), 2.65 – 2.51 (m, 2H), 2.17 – 2.02 (m, 3H).

**<sup>13</sup>C-NMR** (151 MHz, CDCl<sub>3</sub>): δ [ppm] = 180.23, 177.37, 172.54, 141.50, 135.08, 133.87, 133.67, 133.49, 133.20, 133.08, 132.40, 131.22, 130.85, 130.01, 129.84, 129.74, 127.61, 127.16, 126.78, 125.84, 124.33, 71.82, 64.71, 63.45, 58.66, 36.07, 35.88, 31.11, 22.90.

**<sup>19</sup>F-NMR** (565 MHz, CDCl<sub>3</sub>): δ [ppm] = -59.70 (t, *J*<sub>HF</sub> = 10.8 Hz, 3F, CF<sub>3</sub>).

#### 14.2.2.2 Removal of Chiral Ligand and Fmoc-protection: Synthesis of Fmoc-TfeGly (1)

The synthesis followed the protocol of Yin *et al.* [8]



To a solution of **20** (1.4 g, 2 mmol 1 eq.) in 7 mL of dimethoxyethane (DME) was added 3N HCl (3.3 mL, 10 mmol, 5 eq.) and the reaction mixture was stirred for 2 h at 60°C. During this time, a white suspension was formed, which was cooled to rt after 2 h. The obtained precipitate (chiral ligand) was separated, washed 2x with 6 mL of water, and the filtrate was concentrated *in vacuo*. The procedure was repeated once.

EDTA-Na<sub>2</sub> (584.5 mg, 2 mmol, 1eq.) and 4.2 mL MeCN were then added to the concentrated filtrate. The reaction mixture was stirred for 3 h at rt. Subsequently, the solution was brought to a pH of 8.4 with 48% NaOH (4.1 eq.) and 6 M HCl (0.06 eq.). NaCO<sub>3</sub> (273 mg, 2.6 mmol, 1.3 eq.) and Fmoc-OSu (674 mg, 2 mmol, 1 eq.) were slowly added. This reaction mixture was stirred at rt for 20 h, concentrated *in vacuo* (MeCN removed), and the phases separated with EtOAc and water. The organic phase was dried with anhydrous sodium sulfate, and the solvent was removed *in vacuo*. Purification of the crude product was achieved using preparative HPLC with a Kinetex RP-C18 end-capped (5 μM, 100 Å, 250 x 21.2 mm, Phenomenex®, USA) reverse phase HPLC column. For this purpose, gradient **P7** was applied with a UV-detection at λ<sub>abs</sub> = 280 nm. The fractions containing pure Fmoc-amino acid were

combined, MeCN was removed *in vacuo*. The pure product was obtained as a colorless solid (250 mg, 0.64 mmol, 32 % yield) after lyophilization.

**<sup>1</sup>H-NMR** (400 MHz, CDCl<sub>3</sub>): δ [ppm] = 7.76 (d, *J* = 7.6 Hz, 2H), 7.57 (d, *J*<sub>HH</sub> = 7.5 Hz, 2H), 7.40 (d, *J* = 7.5 Hz, 2H), 7.38 (d, *J* = 7.5 Hz, 2H), 5.53 (d, *J* = 9.4 Hz, 1H), 4.64 (dd, *J* = 8.1, 4.4 Hz, 1H), 4.50-4.36 (m, 2H), 4.23 (t, *J* = 6.8 Hz, 1H), 2.85-2.72 (m, 2H).

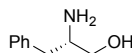
**<sup>19</sup>F-NMR** (376 MHz, CDCl<sub>3</sub>): δ [ppm] = -62.93 (t, *J*<sub>HF</sub> = 10.3 Hz, 3F).

### 14.2.3 Synthesis of Fmoc-TfVal Derivatives (2+ 3), according to Erdbrink

At the beginning of the synthesis sequence, the Evans auxiliary **9** was first prepared in a 2-step procedure,<sup>[9]</sup> followed by alkylation and fluorination according to the protocol of Erdbrink *et al.*<sup>[10]</sup>

#### 14.2.3.1 (S)-Phenylalanol (**8**)

The synthesis was performed by described procedures of Gage *et al.*<sup>[9]</sup>



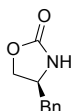
Phe (27.25 g, 165 mmol, 1 eq.) was stored overnight in a vacuum in a three-neck flask. In an Ar atmosphere, 60 mL dry THF was added. To the resulting white suspension, BF<sub>3</sub>·OEt<sub>2</sub> (20.5 mL, 1.5 mmol, 1 eq.) was carefully added via a dropping funnel. The reaction mixture was heated under reflux until a clear solution was obtained (approx. 1 h). Then, BH<sub>2</sub>·SMe<sub>2</sub> (19.4 mL, 190 mmol, 1.15 eq.) was carefully added via a dropping funnel under strong reflux. This reaction mixture was heated under reflux until Phe was consumed (approx. 1 h). Then, the obtained reaction solution was cooled to rt, and 20 mL of a mixture of THF/water (1:1 %v/v) was added. Subsequently, 125 mL of 5 M NaOH solution was added, and the obtained two-phase mixture was heated overnight under reflux. After the reaction time, the mixture was cooled to rt, and the solid was filtered off and washed two times with 50 mL THF. The filtrate was concentrated *in vacuo* and extracted with DCM (2 x 100 mL). The combined organic phases were dried over anhydrous sodium sulfate, and the solvent was removed under reduced pressure. The crude product was purified with crystallization in EtOAc (approx. 180 mL). The desired product was isolated as a colorless solid (16.25 g, 107 mmol, 65%).

<sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>) δ = 7.34 – 7.14 (m, 5H), 3.64-3.61 (dd, *J*=10.6, 3.9, 1H), 3.40-3.36 (dd, *J*=10.6, 7.2, 1H), 3.14-3.09 (m, 1H), 2.78 (dd, *J*=13.5, 5.2, 1H), 2.51 (dd, *J*=13.5, 8.7, 1H), 2.05-1.95 (bs, 3H)

[TLC] R<sub>f</sub>: 0.41 (*n*-hexane: EtOAc 4:1)

#### 14.2.3.2 (S)-4-(Phenylmethyl)-2-oxazolidinone - Evan's auxiliary (**9**)

The synthesis was performed by described procedures of Gage *et al.*<sup>[9]</sup>



The reaction was carried out in a distillation apparatus with a Vigreux column. The K<sub>2</sub>CO<sub>3</sub> used was dried overnight *in vacuo*.

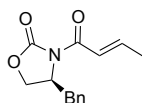
In an Ar atmosphere, **8** (16.25 g, 107 mmol, 1 eq.) of K<sub>2</sub>CO<sub>3</sub> (1.5 g, 10.7 mmol, 0.1 eq.) were combined, and (EtO)<sub>2</sub>CO (26 mL, 221 mmol, 2.06 eq.) was added. The reaction mixture was heated in the distillation apparatus until no more EtOH was formed (approx. 1 h). The resulting white suspension was cooled to rt and taken up in 300 mL DCM. The aqueous phase was extracted with DCM (2x 100 mL). The combined organic phases were dried over anhydrous sodium sulfate, and the solvent was removed under reduced pressure.

The residue was dissolved in a hot mixture of EtOAc/hexane (2:1 %v/v, 150 mL), hot filtered, and the filtrate stored overnight at -18 °C. The crystallized solid was isolated and dried. The desired product was obtained as a colorless solid (13.71 g, 77 mmol, 72 %).

<sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>) δ = 7.30 – 7.14 (m, 5H), 4.72 (bs, 1H), 3.32 (dd, *J*=13.4, 3.3, 3H), 2.79 (dd, *J*=13.3, 9.5, 2H).

#### 14.2.3.3 (S,E)-4-benzyl-3-(but-2-enoyl)oxazolidin-2-one (**11**)

Alkylation (synthesis of **11**) and fluorination (synthesis of **12a/12b**) were conducted according to the protocol of Erdbrink *et al.*<sup>[10]</sup>

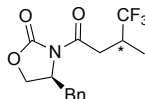


**9** (2.3 g, 13 mmol, 1 eq.) was dissolved in 26 mL dry THF (2 mL/mmol) and cooled to -78°C. Then, *n*-BuLi (2.5 M in hexane, 5.2 mL, 13 mmol, 1 eq.) was carefully added

via syringe, and the reaction mixture was stirred for 30 min at  $-78^{\circ}\text{C}$ . Subsequently, freshly distilled crotonic acid chloride (1.51 mL, 14.2 mmol, 1.1 eq.) was dissolved in 14 mL THF (1 mL/mmol) and carefully added to the reaction mixture for 25 min. The reaction mixture was stirred for an additional 2 h at  $-78^{\circ}\text{C}$ . The reaction was then warmed to rt and quenched with saturated  $\text{NH}_4\text{Cl}$  solution. The aqueous phase was extracted with EtOAc (3x 50 mL). The combined organic phases were washed with saturated NaCl solution and dried over anhydrous sodium sulfate. The solvent was removed under reduced pressure, and the crude product was purified by flash column chromatography (EtOAc/n-hexane, 1:4) to yield **11** as a white solid (2g, 8.2 mmol, 63%).

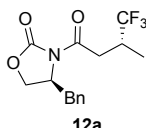
$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  = 7.31 – 7.22 (m, 3H), 7.25 – 7.17 (m, 2H), 6.87 (m, 1H), 6.17-6.13 (m, 1H), 4.39 – 4.30 (m, 2H), 4.15 – 4.08 (m, 1H), 3.18 – 3.10 (m, 1H), 2.93 – 2.84 (m, 1H), 1.87 (dd,  $J=6.0$ , 1.6, 3H).

#### 14.2.3.4 (S)-4-Benzyl-3-((R)-4,4,4-trifluoro-3-methyl-butanoyl)oxazolidin-2-one (**12a**) and (S)-4-benzyl-3-((S)-4,4,4-trifluoro-3-methylbutanoyl)oxazolidin-2-one (**12b**)



In a Schlenk vessel, dry THF (4.3 mL) was added and cooled to  $-78^{\circ}\text{C}$ . Then trifluoroiodomethane  $\text{CF}_3\text{I}$  (1.7 mL, 22.01 mmol) was condensed into THF at  $-78^{\circ}\text{C}$  and chilled at same temperature for further use. Meanwhile, (*S,E*)-4-benzyl-3-(but-2-enoyl)oxazolidin-2-one (**11**) (200 mg, 0.81 mmol) and ytterbium(III) trifluoromethanesulfonate  $\text{Yb}(\text{OTf})_3 \cdot n\text{H}_2\text{O}$  (1.01 g, 1.63 mmol, 2 eq.) were dissolved in DCM (4.1 mL) and THF (4.1 mL) and stirred for 30 min at rt. Then the reaction mixture was cooled to  $-78^{\circ}\text{C}$  and stirred for an additional 20 min. Thereupon  $\text{CF}_3\text{I}$  condensed in THF (2 mL) and triethylborane  $\text{Et}_3\text{B}$  (1.0 M in hexane, 4.1 mL, 4.1 mmol, 5 eq.) were added by syringes. Finally,  $\text{O}_2$  (41 mL) was added into the now yellow-colored solution via syringe over 1.5 h in intervals of 15 min, then the reaction mixture was stirred at  $-78^{\circ}\text{C}$  for additional 15 min. The addition of  $\text{CF}_3\text{I}$  condensed in THF,  $\text{Et}_3\text{B}$  and finally  $\text{O}_2$  was repeated twice. Next, the now intensely yellow-colored solution was stirred overnight at  $-78^{\circ}\text{C}$  before it was quenched by addition of 0.1 M HCl solution (130 mL). After stirring for additional 2 h at rt, the

reaction mixture was diluted with H<sub>2</sub>O (100 mL) and the aqueous phase was extracted with DCM (3 x 300 mL). The combined organic phases were washed with brine, dried over anhydrous sodium sulfate, filtered, and concentrated *in vacuo*. For separation of educt residues, the crude product was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>:toluene 2:1). The diastereomers were afterwards separated by column chromatography (hexane : EtOAc 4:1) to afford **12a** (34.6 mg, 0.11 mmol, 13% yield) and **12b** (37.7 mg, 0.12 mmol, 15% yield) as colorless oils.

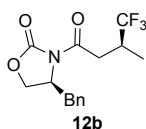


**<sup>1</sup>H-NMR** (400 MHz, CDCl<sub>3</sub>): δ = 7.63 – 7.18 (m, 5H), 4.73 – 4.65 (m, 1H), 4.26 – 4.16 (m, 2H), 3.37 – 3.24 (m, 2H), 3.02 – 2.90 (m, 2H), 2.76 (dd, J = 13.4 Hz, 9.7 Hz, 1H), 1.20 (d, J = 6.4 Hz, 3H).

**<sup>19</sup>F-NMR** (376 MHz, CDCl<sub>3</sub>): δ = -73.27 (d, J = 7.4 Hz, 3F).

**[TLC]** R<sub>f</sub>: 0.35 (*n*-hexane: EtOAc 4:1)

**[Entry 2]** <sup>1</sup>H-NMR ratio: [1.00 (**12a**) : 0.72 (**11**)]



**<sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>)**: δ = 7.37 – 7.18 (m, 5H), 4.77 – 4.66 (m, 1H), 4.29 – 4.17 (m, 2H), 3.30 (dd, J = 13.4, 3.4 Hz, 1H), 3.25 (dd, J = 17.7, 4.3 Hz, 1H), 3.05 (dd, J = 17.6, 8.4 Hz, 1H, -COCH<sub>2</sub>), 3.01 – 2.90 (m, 1H), 2.79 (dd, J = 13.3, 9.6 Hz, 1H), 1.23 (d, J = 6.8 Hz, 3H).

**<sup>19</sup>F-NMR (400 MHz, CDCl<sub>3</sub>)**: δ = -73.23 (d, J = 9.5 Hz, 3F).

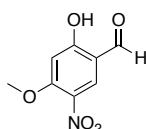
**[TLC]** R<sub>f</sub>: 0.4 (hexane: EtOAc 4:1)

**[Entry 2]** <sup>1</sup>H-NMR ratio: [1.00 (**12a**) : 0.15 (**11**)]

#### 14.2.4 Synthesis of removable backbone modification (RBM) - linker 22

**22** was synthesized following the procedure by Tang *et al.* [11]

#### 14.2.5 4-Methoxy-5-nitrosalicylaldehyde synthesis (**22**)



4 mL of glacial acetic acid and 4 mL of nitric acid were mixed in a 50 mL round-bottomed flask. The mixture was then chilled to 0 °C in an ice bath. 2-hydroxy-4-methoxy benzaldehyde (**21**, 1 g, 6.6 mmol) was slowly added to the solution. Once



all the reactants were added, the mixture turned pink. The reaction was allowed to warm up to rt overnight and was then quenched by slow addition of 30 mL of ice water. A yellow solid was formed. The solid was isolated from a yellow transparent solvent by vacuum filtration. The crude product was purified through crystallization in EtOH (approx. 85 mL) with 1 mL of water. The crystallized product was isolated through vacuum filtration and was washed with cold ethanol and water (1:1 %v/v) and dried *in vacuo* to yield a pale yellow powder (0.64 g, 3.2 mmol, 49% yield). The final product was stored under N<sub>2</sub>.

<sup>1</sup>H-NMR (500 MHz, DMSO-d<sub>6</sub>): δ = 9.95 (s, 1H), 8.18 (s, 1H), 6.59 (s, 1H) 3.86 (s, 3H).

<sup>13</sup>C-NMR (151 MHz, DMSO-d<sub>6</sub>): δ = 189.8, 166.3, 159.4, 132.2, 129.1, 115.8, 101.4, 57.6.

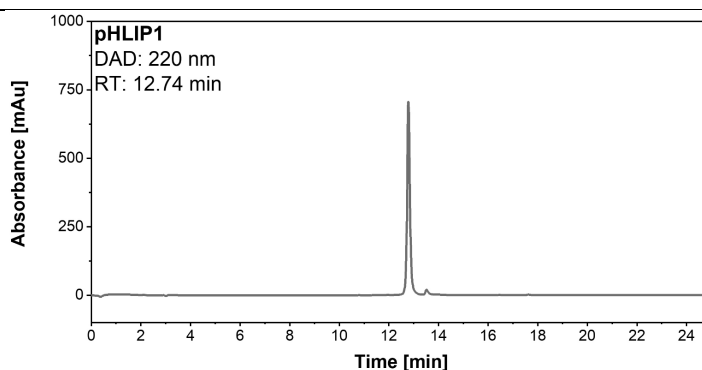
### 14.3. Synthesized and Purified Peptides

#### 14.3.1 Part A: pH (Low) Lipid Insertion Peptides and its Derivatives

All sequences contain the difficult combination Asp-Trp; therefore, Fmoc-Asp(OMpe)-OH was included in this position (highlighted in blue). Before purification, all samples were incubated overnight in an acidic environment (30% MeCN in H<sub>2</sub>O + 0.1% TFA). The first AA was introduced after a double Fmoc-deprotection. The loading was not examined each time as this procedure was proven to yield quantitative loading.

**Table 14.26:** Conditions for the synthesis, isolation, and purification of **pHLIP1**.

Name	Sequence	
<b>pHLIP1</b>	H <sub>2</sub> N-AEQNPIYWARYAD <sup>W</sup> WLFTTPLLLLDLALLVDADEGT-CONH <sub>2</sub>	
Scale	Resin	Amount of Resin
0.05 mmol	Rink amide pro tide (0.19 mmol/g)	262.7 mg
Synthesis	Coupling conditions	Fmoc-removal
MW-SPPS	1 AA: double deprotection, 4 min single coupling 2-20 AA: 4 min single coupling 21-36 AA: 2 min double coupling	20 % Pip +0.1 M HOBT
TFA-Cleavage:	Purification gradient: P7	Obtained purity:
Cocktail III 2.5 h	Kinetex <sup>®</sup> C18 column	> 98%

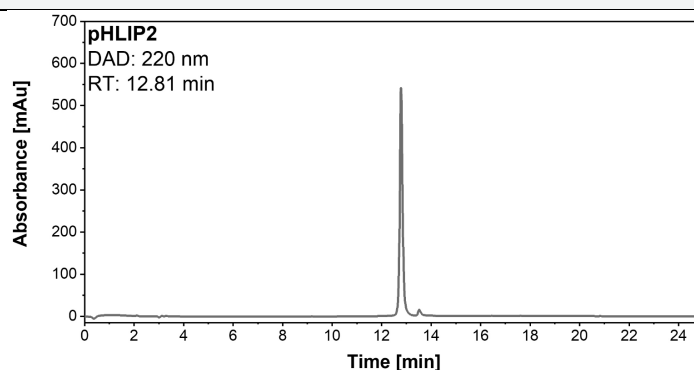


HPLC system: Chromaster® Kinetex® C18 column;  
 eluents: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: A6

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+2H] <sup>2+</sup>	2003.5359	2003.5458	20.3 mg
[M+3H] <sup>3+</sup>	1336.0265	1336.0265	
[M+4H] <sup>4+</sup>	1002.2718	1002.2720	
[M+5H] <sup>5+</sup>	802.0190	802.0207	

**Table 14.27:** Conditions for the synthesis, isolation, and purification of pHLIP2.

Name	Sequence	
pHLIP2	H <sub>2</sub> N-DDEEDNPIYWARYADWLF <del>TTPLLLLHGALLVDA</del> ECT-CONH <sub>2</sub>	
Scale	Resin	Amount of Resin
0.05 mmol	Rink amide pro tide (0.19 mmol/g)	263.5 mg
Synthesis	Coupling conditions	Fmoc-removal
MW-SPPS	1 AA: double deprotection, 4 min single coupling 2-20 AA: 4 min single coupling 21-36 AA: 2 min double coupling	20 % Pip +0.1 M HOBt
TFA-Cleavage:	Purification gradient: P7	Obtained purity:
Cocktail III 2.5 h	Kinetex® C18 column	> 98%

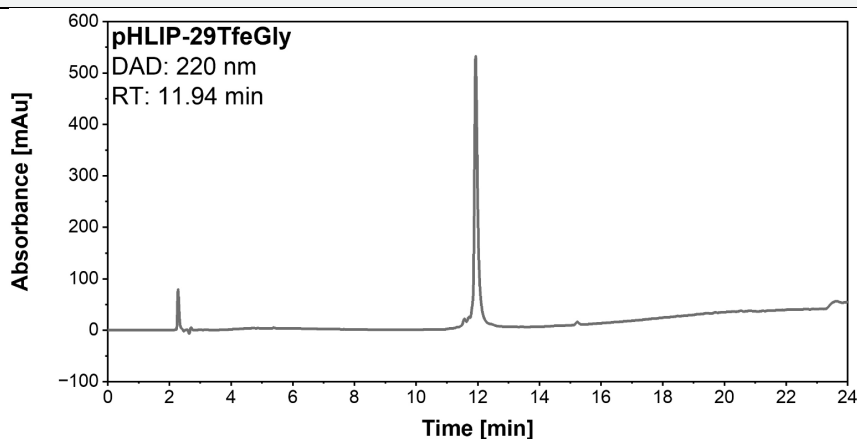


HPLC system: Chromaster® Kinetex® C18 column;  
 eluents: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: A6

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+2H] <sup>2+</sup>	2140.0205	2140.0217	19.7 mg
[M+3H] <sup>3+</sup>	1427.0163	1427.0143	
[M+4H] <sup>4+</sup>	1070.5141	1070.5138	
[M+5H] <sup>5+</sup>	856.6129	856.6146	

**Table 14.28:** Conditions for the synthesis, isolation, and purification of **pHLIP-29-TfeGly**.

Name	Sequence	
<b>pHLIP-29TfeGly</b>	H <sub>2</sub> N-DDDEDNPIYWARYADWLFTTPLLLLHGALL <b>TfeGly</b> DADECT-CONH <sub>2</sub>	
Scale	Resin	Amount of Resin
0.05 mmol	Rink amide pro tide (0.19 mmol/g)	262.6 mg
Synthesis	Coupling conditions	Fmoc-removal
MW-SPPS	1 AA: double deprotection, 4 min single coupling 2-20 AA: 4 min single coupling 21-36 AA: 2 min double coupling	20 % Pip +0.1 M HOBT
TFA-Cleavage:	Purification gradient: P7	Obtained purity:
Cocktail III 2.5 h	Kinetex <sup>®</sup> C18 column	> 98%



HPLC system: Primaide <sup>®</sup> Kinetex <sup>®</sup> C18 column;			
eluent: A = H <sub>2</sub> O, B = MeCN both containing 0.1% (v/v) TFA; gradient: <b>A19</b>			
ESI-ToF	calc. m/z	obs. m/z	Yield
[M+2H] <sup>2+</sup>	2024.0167	2024.0491	12 mg
[M+3H] <sup>3+</sup>	1349.6804	1349.6983	
[M+4H] <sup>4+</sup>	1012.5122	1012.5249	

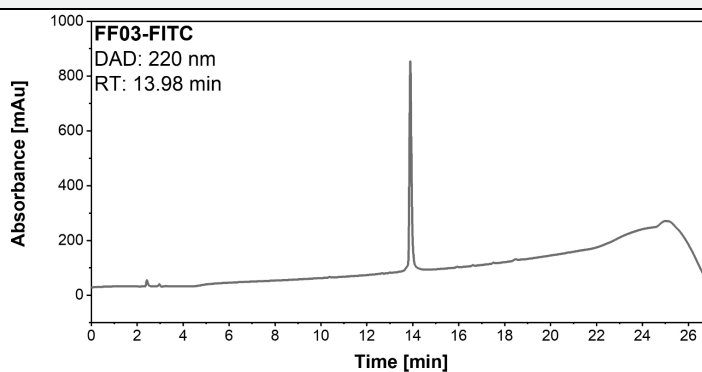
### 14.3.2 Part A: Multivalent Presentation of EAE-Peptides on FF03 Scaffold

The syntheses of the conjugates were carried out on a TGA resin. This resin is not suitable for prolonged coupling conditions under MW radiation; hence the synthesis was carried out in sections to allow the solid support to cool down.

At position Lys3, Fmoc-Lys(ivDde)-OH was incorporated for coupling of the fluorescent label, and at position Lys17, Fmoc-Lys(Mmt)-OH was incorporated for conjugation with the EAE-inducing peptides, MOG or PLP.

**Table 14.29:** Conditions for the synthesis, isolation, and purification of **FF03-FITC**.

Name	Sequence	
<b>FF03-FITC</b>	Abz-LKK <sub>3</sub> (FITC-Ahx-)ELKELKKELEKLK <sub>17</sub> ELKELKKELOH	
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Leu-TGA resin (0.19 mmol/g)	260.3 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: double deprotection, 2 min single coupling 2-27 AA 2 min single coupling Fmoc-Ahx-OH: 2 min double coupling FITC: 5 eq. + 10 eq. DIPEA, rt, overnight	20 % Pip in DMF
TFA-Cleavage:	Purification gradient: P2	Obtained purity:
Cocktail I 2.5 h	Kinetex® C18 column	> 99%



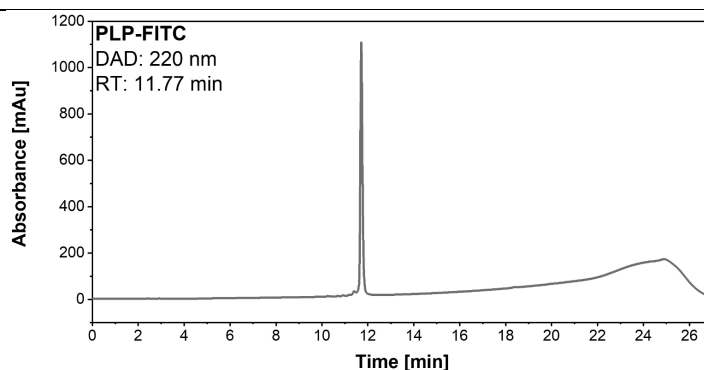
HPLC system: Primaide® Kinetex® C18 column;

eluents: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: gradient: **A18**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+3H] <sup>3+</sup>	1286.6804	1286.7318	18.1 mg
[M+4H] <sup>4+</sup>	965.2622	965.3004	
[M+5H] <sup>5+</sup>	772.4114	772.4411	
[M+6H] <sup>6+</sup>	643.8441	643.8701	

**Table 14.30:** Conditions for the synthesis, isolation, and purification of **PLP-FITC**.

Name	Sequence	
<b>PLP-FITC</b>	FTIC-Ahx-HSLGKWLGHDPKF-OH	
Scale	Resin	Amount of Resin
0.05 mmol	Cl-MPA pro tide (0.23 mmol/g)	217.2 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: DCA loading 2-13 AA: 2 min single coupling Fmoc-Ahx-OH: 2 min double coupling FITC: 5 eq. + 10 eq. DIPEA, rt, overnight	20 % Pip in DMF + 0.1 M HOBt
TFA-Cleavage:	Purification gradient: P2	Obtained purity:
Cocktail I 2.5 h	Kinetex® C18 column	> 98%



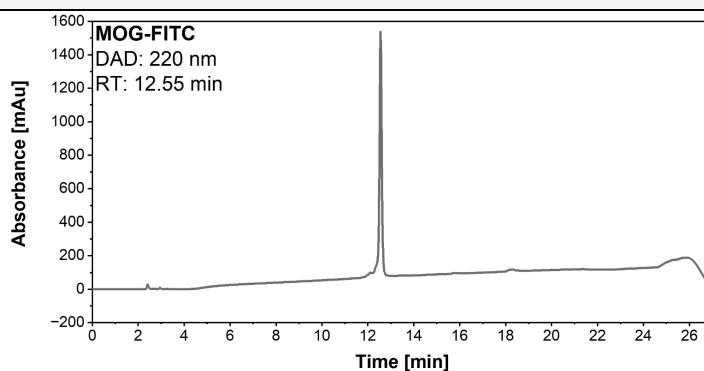
HPLC system: Primaide® Kinetex® C18 column;

eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: gradient: **A18**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+2H] <sup>2+</sup>	1012.9022	1012.9685	
[M+3H] <sup>3+</sup>	675.6040	675.6492	15.1 mg
[M+4H] <sup>4+</sup>	506.9550	506.9894	

**Table 14.31:** Conditions for the synthesis, isolation, and purification of **MOG-FITC**.

Name	Sequence	
<b>MOG-FITC</b>	FTIC-Ahx-MEVGWYRSPFSRVVHLYRNGK-OH	
Scale	Resin	Amount of Resin
0.05 mmol	Cl-MPA pro tide (0.23 mmol/g)	217 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: DCA loading 2-21 AA: 2 min single coupling Fmoc-Ahx-OH: 2 min double coupling FITC: 5 eq. + 10 eq. DIPEA, rt, overnight	20 % Pip in DMF
TFA-Cleavage:	Purification gradient: P2	Obtained purity:
Cocktail I 2.5 h	Kinetex® C18 column	> 98%



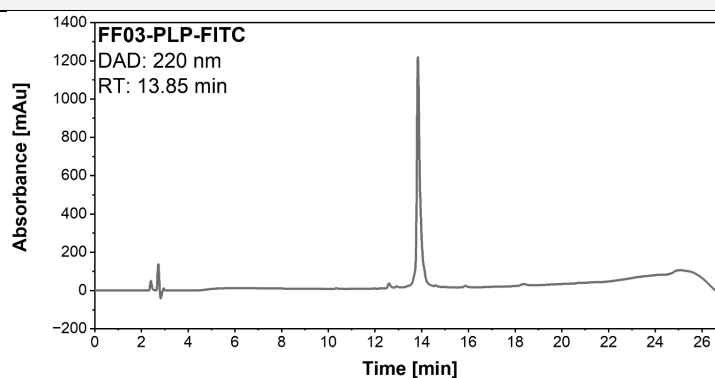
HPLC system: Primaide® Kinetex® C18 column;

eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: gradient: **A18**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+3H] <sup>3+</sup>	1028.7802	1028.8129	
[M+4H] <sup>4+</sup>	771.8371	771.8616	12.5 mg
[M+5H] <sup>5+</sup>	617.6712	617.6910	

**Table 14.32:** Conditions for the synthesis, isolation and purification of **FF03-PLP-FITC**.

Name	Sequence	
<b>FF03-PLP-FITC</b>	Abz-LKK <sub>3</sub> (FITC-Ahx-)ELKELKKELEKLLK <sub>17</sub> (HSLGKWLGHDPKF-GG-)JELKELK KEL-OH	
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Leu-TGA resin (0.19 mmol/g)	264.2 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: double deprotection, 2 min single coupling	20 % Pip in DMF (+ 0.1 M HOBt with PLP sequence)
	2-27: AA 2 min single coupling	
	Side-chain coupling: 2 min double coupling	
	Fmoc-Ahx-OH: 2 min double coupling	
	FITC: 5 eq. + 10 eq. DIPEA, rt, overnight	
TFA-Cleavage:	Purification gradient: P2	Obtained purity:
Cocktail I 2.5 h	Kinetex® C18 column	> 98%



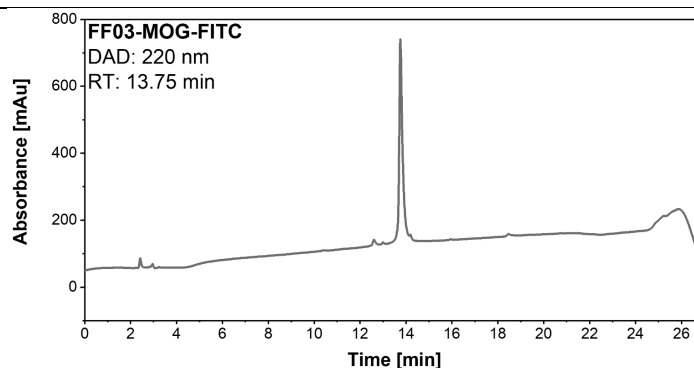
HPLC system: Primaide® Kinetex® C18 column;

eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: gradient: **A18**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+5H] <sup>5+</sup>	1096.3757	1096.1743	10.2 mg
[M+6H] <sup>6+</sup>	913.8144	913.6442	
[M+7H] <sup>7+</sup>	783.4134	783.2749	
[M+8H] <sup>8+</sup>	685.6127	685.4897	

**Table 14.33:** Conditions for the synthesis, isolation and purification of **FF03-MOG-FITC**.

Name	Sequence	
<b>FF03-MOG-FITC</b>	Abz-LKK <sub>3</sub> (FITC-Ahx-)ELKELKKELEKLLK <sub>17</sub> (MEVGWYRSPFSRVVHLYRNGK-GG-)JELKELK KEL-OH	
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Leu-TGA resin (0.19 mmol/g)	264.2 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: double deprotection, 2 min single coupling	20 % Pip in DMF
	2-27 AA 2 min single coupling	
	Side-chain coupling: 2 min double coupling	
	Fmoc-Ahx-OH: 2 min double coupling	
	FITC: 5 eq. + 10 eq. DIPEA, rt, overnight	
TFA-Cleavage:	Purification gradient: P2	Obtained purity:
Cocktail I 2.5 h	Kinetex® C18 column	> 98%



HPLC system: Primaide® Kinetex® C18 column;

eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: gradient: **A18**

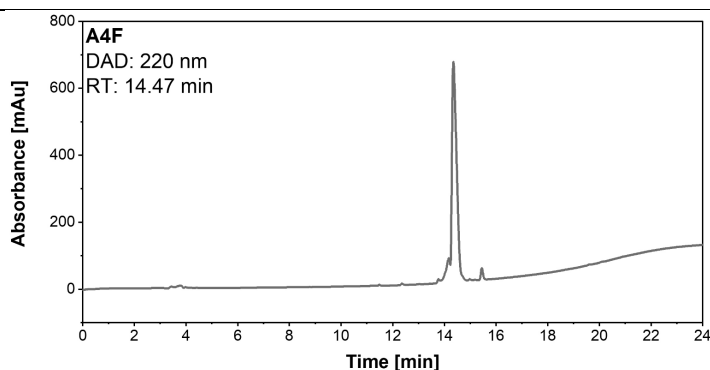
ESI-ToF	calc. m/z	obs. m/z	Yield
[M+4H] <sup>4+</sup>	1635.2790	1635.0259	8.5 mg
[M+5H] <sup>5+</sup>	1308.4248	1308.2272	
[M+6H] <sup>6+</sup>	1090.5219	1090.3584	
[M+7H] <sup>7+</sup>	934.8771	934.7420	

### 14.3.3 Part A: Acid and Base Strands of A4/B4 Dimeric Coiled-Coil

First, A4F and B4F peptides were prepared. For this purpose, a TGR resin was preloaded with Fmoc-Gly-OH or Fmoc-Cys(Trt)-OH, respectively, according to the described procedure (14.1.3.2). The synthesis was performed using automated conventional SPPS on the Activo-P11 synthesizer.

**Table 14.34:** Conditions for the synthesis, isolation, and purification of **A4F**.

Name	Sequence	
<b>A4F</b>	Ac-GGKIAALKQKIAA <b>TfeGly</b> KYKNAALKKKIAALKQGGC-CONH <sub>2</sub>	
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Gly-TGR-NovaSyn (0.22 mmol/g)	227.5 mg
Synthesis	Coupling conditions	Fmoc-removal
Conventional SPPS Manual + P-11	HATU chemistry according to Table 14.4 Fmoc-TfeGly-OH (2 eq.) was coupled manually	20% Pip in DMF
TFA-Cleavage:	Purification gradient: <b>P4</b>	Obtained purity:
Cocktail III 2.5 h	Kinetex® C18 column	96%



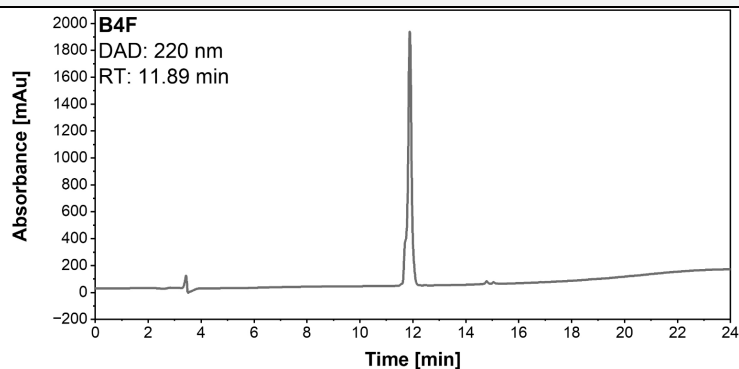
HPLC system: LaChrom® Kinetex® C18 column;

eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A15**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+3H] <sup>3+</sup>	1165.2106	1165.2231	12.3 mg
[M+4H] <sup>4+</sup>	874.1559	874.1699	

**Table 14.35:** Conditions for the synthesis, isolation, and purification of **B4F**.

Name	Sequence	
<b>B4F</b>	Ac-GGKIAALKQKIAA <b>TfeGly</b> KYKNAALKKKIAALKQGGC-CONH <sub>2</sub>	
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Cys(Trt)-TGR-NovaSyn (0.23 mmol/g)	218.2 mg
Synthesis	Coupling conditions	Fmoc-removal
Conventional SPPS Manual + P-11	HATU chemistry according to Table 14.4 Fmoc-TfeGly-OH (2 eq.) was coupled manually	20% Pip in DMF
TFA-Cleavage:	Purification gradient: <b>P4</b>	Obtained purity:
Cocktail III 2.5 h	Kinetex® C18 column	98%



HPLC system: LaChrom® Kinetex® C18 column;

eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A15**

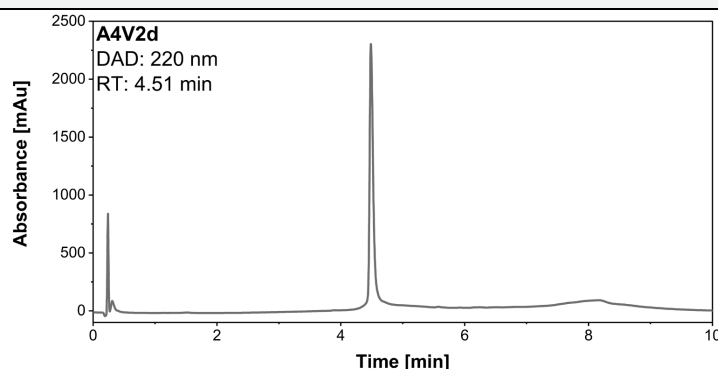
ESI-ToF	calc. m/z	obs. m/z	Yield
[M+3H] <sup>3+</sup>	1155.0116	1155.0228	14.3 mg
[M+4H] <sup>4+</sup>	866.5106	866.5198	

All Val derivatives and B4V2F were synthesized using MW-SPPS. The rink amide resin was pre-loaded after a double deprotection.



**Table 14.36:** Conditions for the synthesis, isolation, and purification of **A4V2d**.

Name		Sequence
<b>A4V2d</b>		Ac-GGKIAALKQKIAAVKYKNAALKKKIAALKQGGC-CONH <sub>2</sub>
Scale	Resin	Amount of Resin
0.05 mmol	Rink amide pro tide (0.19 mmol/g)	263.7 mg
Synthesis	Coupling conditions	Fmoc-removal
MW-SPPS	1 AA: double deprotection, 2 min double coupling 2-20 AA: 2 min single coupling 21-33 AA: 2 min double coupling Capping at N-term	20% Pip in DMF
TFA-Cleavage:	Purification gradient: P4	Obtained purity:
Cocktail III 2.5 h	Kinetex® C18 column	95%

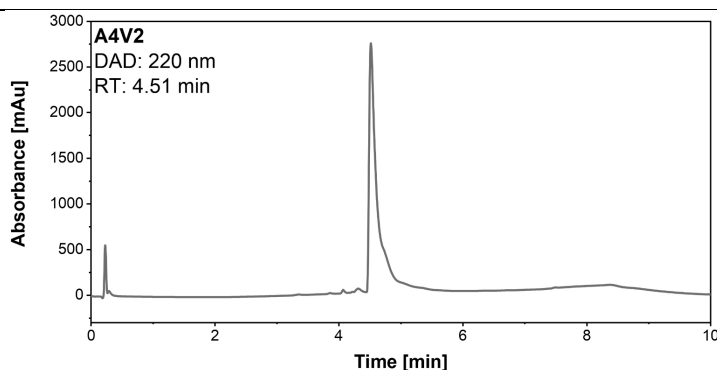


HPLC system: Chromaster®(FF) Purospher® STAR RP-C18 end-capped UHPLC column;  
eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A4**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+2H] <sup>2+</sup>	1727.3462	1727.3508	15.2 mg
[M+3H] <sup>3+</sup>	1151.9001	1151.9037	
[M+4H] <sup>4+</sup>	864.1770	864.1807	

**Table 14.37:** Conditions for the synthesis, isolation, and purification of **A4V2**.

Name		Sequence
<b>A4V2</b>		Ac-GGKIAALKQKVAALKYKNAALKKKIAALKQGGC-CONH <sub>2</sub>
Scale	Resin	Amount of Resin
0.05 mmol	Rink amide pro tide (0.19 mmol/g)	263 mg
Synthesis	Coupling conditions	Fmoc-removal
MW-SPPS	1 AA: double deprotection, 2 min double coupling 2-20 AA: 2 min single coupling 21-33 AA: 2 min double coupling Capping at N-term	20% Pip in DMF
TFA-Cleavage:	Purification gradient: P4	Obtained purity:
Cocktail III 2.5 h	Kinetex® C18 column	95%

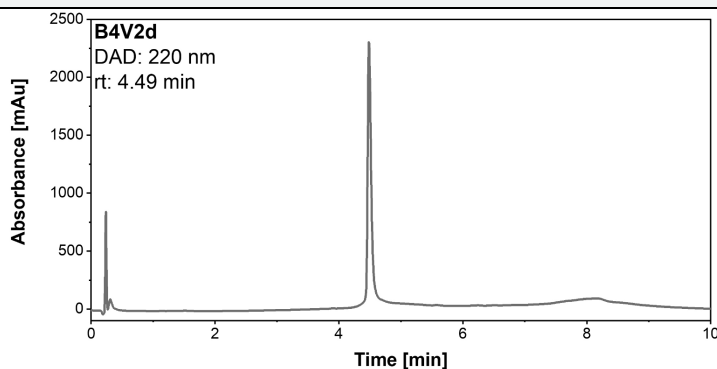


HPLC system: Chromaster®(FF) Purospher® STAR RP-C18 end-capped UHPLC column;  
 eluents: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A3**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+2H] <sup>2+</sup>	1727.3462	1727.3568	15.6 mg
[M+3H] <sup>3+</sup>	1151.9001	1151.9076	
[M+4H] <sup>4+</sup>	864.1770	864.1841	

Table 14.38: Conditions for the synthesis, isolation, and purification of **B4V2d**.

Name	Sequence	
<b>B4V2d</b>	Ac-GGKIAALKQKIAAVKYKNAALKKKIAALKQGQC-CONH <sub>2</sub>	
Scale	Resin	Amount of Resin
0.05 mmol	Rink amide pro tide (0.19 mmol/g)	263.3 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: double deprotection, 2 min double coupling	20% Pip in DMF
	2-20 AA: 2 min single coupling	
	21-33 AA :2 min double coupling	
	Capping at N-term	
TFA-Cleavage:	Purification gradient: P4	Obtained purity:
Cocktail III 2.5 h	Kinetex® C18 column	98%

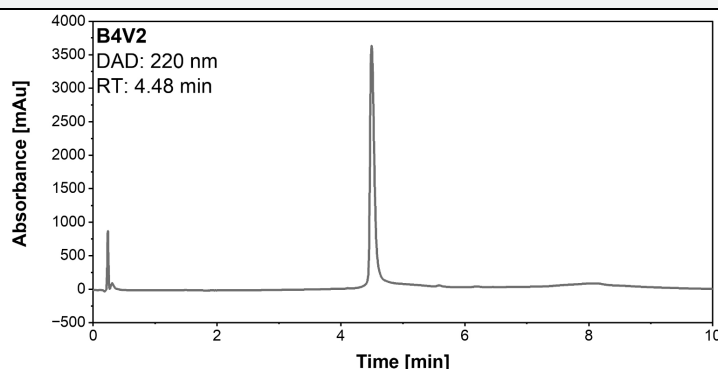


HPLC system: Chromaster®(FF) Purospher® STAR RP-C18 endcapped UHPLC column;  
 eluents: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A4**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+2H] <sup>2+</sup>	1712.0477	1712.0657	19.4mg
[M+3H] <sup>3+</sup>	1141.7011	1141.7115	
[M+4H] <sup>4+</sup>	856.5277	856.5369	

**Table 14.39:** Conditions for the synthesis, isolation, and purification of **B4V2**.

Name	Sequence	
<b>B4V2</b>	Ac-GGKIAALKQKVAALKYKNAALKKKIAALKQGGC-CONH <sub>2</sub>	
Scale	Resin	Amount of Resin
0.05 mmol	Rink amide pro tide (0.19 mmol/g)	263.1 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: double deprotection, 2 min double coupling 2-20 AA: 2 min single coupling 21-33 AA: 2 min double coupling Capping at N-term	20% Pip in DMF
TFA-Cleavage:	Purification gradient: P4	Obtained purity:
Cocktail III 2.5 h	Kinetex® C18 column	98%

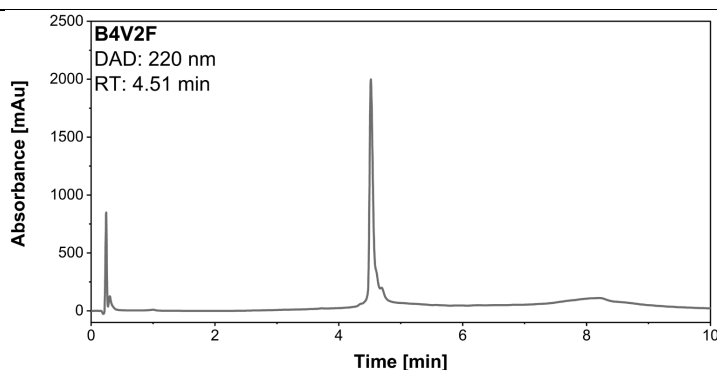


HPLC system: Chromaster®(FF) Purospher® STAR RP-C18 endcapped UHPLC column;  
eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A3**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+2H] <sup>2+</sup>	1727.3462	1727.3568	15.6 mg
[M+3H] <sup>3+</sup>	1151.9001	1151.9076	
[M+4H] <sup>4+</sup>	864.1770	864.1841	

**Table 14.40:** Conditions for the synthesis, isolation, and purification of **B4V2F**.

Name	Sequence	
<b>B4V2F</b>	Ac-GGKIAALKQK <b>TfeGly</b> AALKYKNAALKKKIAALKQGGC-CONH <sub>2</sub>	
Scale	Resin	Amount of Resin
0.05 mmol	Rink amide pro tide (0.19 mmol/g)	263.3 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: double deprotection, 2 min double coupling 2-20 AA: 2 min single coupling 21-33 AA: 2 min double coupling Fmoc-TfeGly-OH: 10 min x-tra wash Capping at N-term	20% Pip in DMF
TFA-Cleavage:	Purification gradient: P4	Obtained purity:
Cocktail III 2.5 h	Kinetex® C18 column	98%



HPLC system: Chromaster®(FF) Purospher® STAR RP-C18 end-capped UHPLC column;  
eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A4**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+2H] <sup>2+</sup>	1732.0135	1732.0288	12.5 mg
[M+3H] <sup>3+</sup>	1155.0116	1155.0257	
[M+4H] <sup>4+</sup>	866.5106	866.5224	
[M+5H] <sup>5+</sup>	693.4101	693.4079	

### 14.3.4 Part A: Functionalized Ubiquitin Monomers

#### 14.3.4.1 Ub<sub>2</sub>-N-Me-Cys

The synthesis of the *N*-Me-Cys modified Ub<sub>2</sub> for the ICL approach was performed according to the protocol of Kumar *et al.* Since a difference of -15 in *m/z* was observed in the synthesis by MALDI, commercially available Fmoc-*N*-Me-Cys(Trt)-OH (Iris Biotec) was utilized. This was manually incorporated at rt. For this, the Rink-amide per tide resin was first swelled in DMF for 30 min at rt, then treated with 3 mL of 20% Pip in DMF three times for 10 min each at rt. This was followed by the coupling of 5 eq. Fmoc-*N*-Me-Cys-OH with 4.9 eq. HATU and 10 eq. DIPEA. The reaction mixture was shaken for three h at rt. Subsequently, a loading test was performed according to aapptec, and all unreacted NH<sub>2</sub> groups on the resin were acetylated using a capping procedure (10% Ac<sub>2</sub>O + 10% DIPEA). The following four AA were also manually incorporated for the improved synthesis, according to Table 10.3. Only after that was the synthesis continued with the MW-SPPS. Here, the dipeptide Fmoc-Asp(O<sup>t</sup>Bu)-Gly(Dmb)-OH was used for the Asp<sub>52</sub>-Gly<sub>53</sub> section. Since the Ub sequence has four other Asp areas potentially prone to aspartimide formation, Fmoc-Asp(OMpe)-OH was generally applied for all other Asp residues. In addition, 10% (%w) piperazine in NMP/EtOH (9:1% v/v) was employed in the Fmoc cleavage. Fmoc-Arg(Pbf)-OH was generally implemented with a double coupling and Fmoc-His(Trt)-OH with a 10 min coupling at 50°C. Furthermore, Fmoc-

**Nle**-OH was incorporated instead of Met at the N-terminus to circumvent potential oxidation of Met.

**Table 14.41:** Conditions for the synthesis, isolation, and purification of **Ub2-N-Me-Cys**.

Name	Sequence	
<b>Ub2-N-Me-Cys</b>	H <sub>2</sub> N- <b>M</b> QIFVKLTGKTTITLEVEPS <b>D</b> TIENVKAKIQ <b>D</b> KEGIPP <b>D</b> QQRLI-FAGKQLE <b>D</b> GR <b>T</b> LS <b>D</b> YNIQKESTLHLVLRRLGG- <b>N-Me-Cys</b> -CONH <sub>2</sub>	
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-N-Me-Cys-Rink amide pro tide (0.18 mmol/g)	277.9 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
Manual + MW-SPPS	1-4 AA: manual coupling HATU chemistry 4-30 AA :2 min single coupling 31-76 AA: 2 min double coupling	20% Pip in DMF 10% piperazine in NMP/EtOH (9:1 % v/v) + 0.1 M HOBt
TFA-Cleavage:	Purification gradient: P6	Obtained purity:
Cocktail III 4 h	Luna® C8(2) column	> 98%

**Ub2-N-Me-Cys**  
DAD: 220 nm  
RT: 4.77 min

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+6H] <sup>6+</sup>	1444.4673	1444.7887	7 mg
[M+7H] <sup>7+</sup>	1238.2588	1238.5347	
[M+8H] <sup>8+</sup>	1083.6024	1083.8438	
[M+9H] <sup>9+</sup>	963.3141	963.5276	

HPLC system: Chromaster® (FF) Purospher® STAR RP-C18 end-capped UHPLC column;  
eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A2**

#### 14.3.4.2 Ub2-NHNH<sub>2</sub>

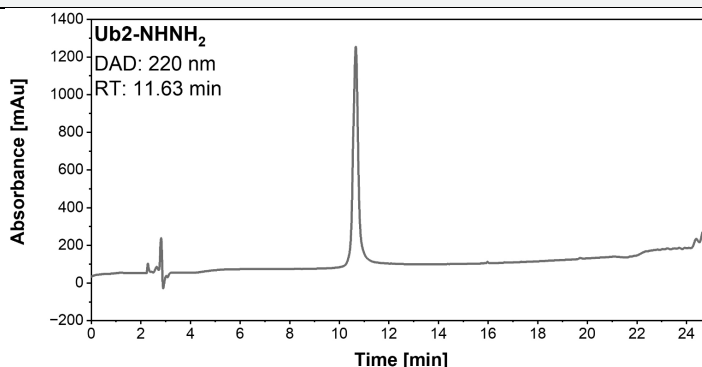
The Cl-TCP(Cl) pro tide resin was functionalized with 5 eq. Fmoc-NHNH<sub>2</sub> and 10 eq. DIPEA in DCM/DMF 1:1 (v/v) at for two h at rt. All unreacted Cl groups were quenched by treatment with 10% MeOH (v/v) in DMF for 30 min at rt. Loading determination was carried out according to aapptec. The functionalized resin was swelled before MW-SPPS in DCM/DMF 1:1 v/v for 30 min at rt. Subsequently, the synthesis was carried out under MW conditions. For this purpose, the previously elaborated synthesis protocol was further optimized. Thus, Fmoc-**His(Boc)**-OH was used instead of Fmoc-**His(Trt)**-OH under previous coupling conditions. In addition, the single-coupling section was shortened from 30 AA to 20 AA. In addition, a **LS-**

dipeptide building block (Fmoc-Leu-Ser[PSI(Me,Me)Pro]-OH) was incorporated with 2 eq. for 10 min at 90°C. The first AA was introduced after double Fmoc-deprotection using double coupling. All cycles must have an **additional washing step** programmed after the coupling, as the swelling properties are altered after functionalization! All couplings were performed in the CarboMax approach (CMA) as the TCP(Cl) is a highly acid-sensitive resin.

Optimizations with regard to suppressing aspartimid formation (Fmoc-Asp(OMpe)-OH, Fmoc-Asp(O<sup>t</sup>Bu)-Gly(Dmb)-OH) remain in this protocol. Likewise, is a double coupling for each Fmoc-Arg(Pbf)-OH and Fmoc-Nle-OH at the N-terminus. Overnight incubation in an acidic medium (starting conditions for HPLC gradient) prior to preparative HPLC purification leads to improved results.

**Table 14.42:** Conditions for the synthesis, isolation, and purification of Ub2-NHNH<sub>2</sub>.

Name	Sequence	
Ub2-NHNH <sub>2</sub>	H <sub>2</sub> N-MQIFVKLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI-FAGKQLEDGRTLSDYNIQKESTLHLVLRRLG-NHNH <sub>2</sub>	
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-NHNH-TCP(Cl) pro tide (0.21 mmol/g)	238.1 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: double deprotection, double coupling 2min	10% piperazine in NMP/EtOH
	2-20 AA: 2 min single coupling	(9:1 % v/v)
	21-75 AA: 2 min double coupling	+ 0.1 M HOBT
	LS dipeptide: 10 min xtra wash	
TFA-Cleavage:	Purification gradient: P11	Obtained purity:
Cocktail III 4 h	Kinetex® C18 column	> 98%



**HPLC system:** Chromaster® Kinetix® C18 column;  
eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: A14

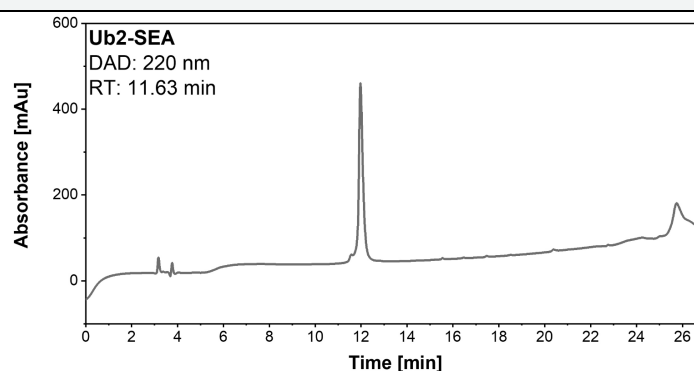
ESI-ToF	calc. m/z	obs. m/z	Yield
[M+5H] <sup>5+</sup>	1701.7592	1701.7406	13.1 mg
[M+6H] <sup>6+</sup>	1418.3006	1418.1112	
[M+7H] <sup>7+</sup>	1215.8302	1215.8114	
[M+8H] <sup>8+</sup>	1063.9774	1063.9572	
[M+9H] <sup>9+</sup>	945.8697	945.8547	

### 14.3.4.3 Ub2-SEA

The required resin Fmoc-Gly-SEA is commercially available. Since this resin has not yet been described for MW-SPPS, Val70 was coupled with 2 eq. for 10 min at 90°C to test the conditions for fluorinated AA. For the MW-SPPS, all sections established for Ub2-NHNH<sub>2</sub> were preserved. The resin must also be cleaved under particular conditions (TFA cocktail **IV**). In addition, the crude product must be brought to the SEAoff state before purification by HPLC. For this purpose, the peptide is dissolved in (AcOH/water 1:4 %v/v) and oxidized with max 2 eq. iodine in DMSO (150 mM) for max. 30 s. Subsequently, excess iodine is quenched with DTT (10 mg/mL).

**Table 14.43:** Conditions for the synthesis, isolation, and purification of **Ub2-SEA**.

Name	Sequence	
<b>Ub2-SEA</b>	H <sub>2</sub> N-MQIFVKLTGKITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLE <span style="color:red">D</span> GRT <span style="color:red">L</span> S <span style="color:red">D</span> YNIQKESTLHLVLRRLRG-SEA	
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Gly-SEA PS resin (0.14 mmol/g)	350 mg
Synthesis	Coupling conditions	Fmoc-removal
MW-SPPS	1 AA: double deprotection + 2 min double coupling 2-20 AA: 2 min single coupling 21-75 AA: 2 min double coupling <span style="color:red">LS</span> dipeptide: 10 min xtra wash	10% piperazine in NMP/EtOH (9:1 % v/v) + 0.1 M HOBT
TFA-Cleavage:	Purification gradient: P11	Obtained purity:
Cocktail <b>IV</b> 3.5 h	Kinetex® C18 column	> 96%



HPLC system: Primaide® Jupiter® C18 column (300Å); eluent: A = H <sub>2</sub> O, B = MeCN both containing 0.1% (v/v) TFA; gradient: <b>A18</b>			
ESI-ToF	calc. m/z	obs. m/z	Yield
[M+6H] <sup>6+</sup>	1435.6253	1435.4552	13.4 mg
[M+7H] <sup>7+</sup>	1230.6799	1230.5166	
[M+8H] <sup>8+</sup>	1076.9709	1076.9600	
[M+9H] <sup>9+</sup>	957.4195	957.3702	
[M+10H] <sup>10+</sup>	861.7783	861.6742	

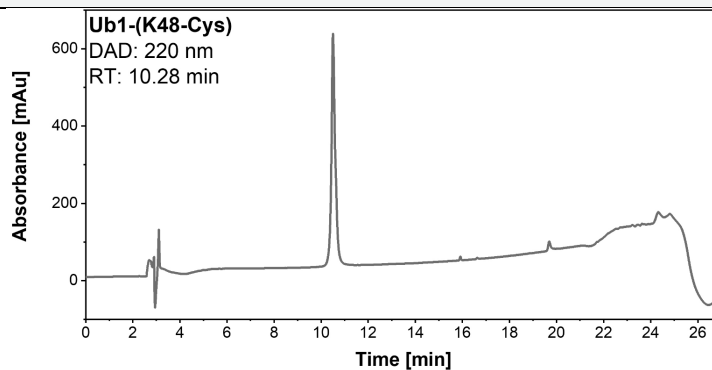
### 14.3.5 Ub1(K48-Cys)

For the thiol fragment of the NCL strategy, Ub1(K48-Cys) was synthesized. Here, the thiol function is introduced by a Cys on the side chain of Lys48. The synthesis was carried out on the Cl-MPA resin. Since this resin is susceptible to acidity, the CarboMax approach (CMA) was used for all couplings. The resin was not swelled and was loaded on the instrument with the first AA through a DCA loading procedure. All other optimizations were adopted from Ub2-NH<sub>2</sub>. It is important here to replace the Fmoc-Nle-OH with a **Boc-Nle**-OH. However, this AA is poorly soluble and requires ultrasonication for complete dissolution.

Special features in this synthesis include the incorporation of Fmoc-**Lys(ivDde)**-OH in position 48, the protecting group of which was removed according to the conditions in Table 10.7. The Boc-Cys-OH was then introduced in a 4 min double coupling.

**Table 14.44:** Conditions for the synthesis, isolation, and purification of **Ub1 (K48-Cys)**.

Name	Sequence	
<b>Ub1(K48-Cys)</b>	H <sub>2</sub> N- <b>M</b> QIFVKTLTGKTITLEVEPS <b>D</b> TIENVKAKIQ <b>D</b> KEGIPP <b>D</b> QQRLI-FAGK <b>(C)</b> QLE <b>DGR</b> TL <b>LS</b> DYNIQKESTL <b>L</b> LVLRRLRGG-OH	
Scale	Resin	Amount of Resin
0.05 mmol	Cl-MPA pro tide resin (0.23 mmol/g)	217.7 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: DCA loading	10% piperazine in NMP/EtOH (9:1 % v/v) + 0.1 M HOBT
	2-20 AA: 2 min single coupling	
	21-76 AA: 2 min double coupling	
	<b>LS</b> dipeptide: 10 min xtra wash	
	Boc-Cys-OH: 4 min double coupling	
TFA-Cleavage:	Purification gradient: P11	Obtained purity:
Cocktai III 4 h	Kinetex® C18 column	> 98%



**HPLC system:** LaChrom® Kinetix® C18 column;  
 eluents: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A11**



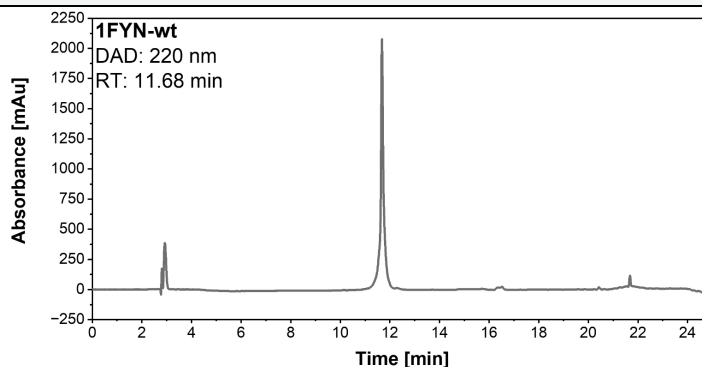
ESI-ToF	calc. m/z	obs. m/z	Yield
[M+5H] <sup>5+</sup>	1071.0000	1070.9281	16.4 mg
[M+6H] <sup>6+</sup>	1442.6680	1442.6068	
[M+7H] <sup>7+</sup>	1236.7165	1236.6667	
[M+8H] <sup>8+</sup>	1082.2529	1082.1334	
[M+9H] <sup>9+</sup>	962.1146	962.0215	

### 14.3.6 Part A: 1FYN and its Derivatives

The synthesis of 1FYN and the corresponding derivatives was carried out on Cl-MPA resin. The synthesis conditions followed the CarboMax approach. All sequences contain the difficult series Asp-Trp. Therefore, a Fmoc-Asp(OMpe) was used in this position. The cleavage was carried out with Cocktail III for 4 h. The product yielded is poorly soluble in the acidic milieu for purification by HPLC. The samples were dissolved at neutral pH and injected. The stationary phase is clogged with an injection of > 8 mg peptide to the preparative HPLC.

**Table 14.45:** Conditions for the synthesis, isolation, and purification of **1FYN-wt**.

Name	Sequence	
<b>1FYN-wt</b>	H <sub>2</sub> N-GTGVTLFVALYDYEARTEDDLSF <b>HK</b> GKEKFQILNSSE- GD <b>W</b> WEARSLTTGETGYIPSNYVAPVD -OH	
Scale	Resin	Amount of Resin
0.05 mmol	Cl-MPA pro tide resin (0.23 mmol/g)	217.4 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: DCA loading	10% piperazine in
	2-20 AA: 4 min single coupling	NMP/EtOH
	21-62 AA: 4 min double coupling	(9:1 % v/v)
		+0.1 M HOBt
TFA-Cleavage:	Purification gradient: P2	Obtained purity:
Cocktail III 4 h	Kinetex® C18 column	> 96%

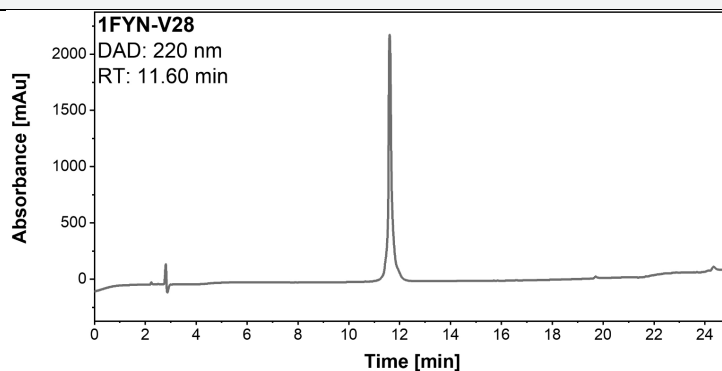


**HPLC system:** Primaide® Kinetix® C18 column;  
 eluents: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A18**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+6H] <sup>6+</sup>	1158.3918	1158.6792	15.6 mg
[M+7H] <sup>7+</sup>	993.6570	993.3288	
[M+8H] <sup>8+</sup>	866.5759	867.0158	

**Table 14.46:** Conditions for the synthesis, isolation, purification of **1FYN-V28**.

Name	Sequence	
<b>1FYN-V28</b>	H <sub>2</sub> N- GTGVTLFVALYDYEARTEDDLSF <b>H</b> KGEKFQVLNSS-EG <b>DW</b> WEARSLTTGETGYIPSNYVAPVD -OH	
Scale	Resin	Amount of Resin
0.05 mmol	Cl-MPA pro tide resin (0.23 mmol/g)	217.6 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: DCA loading 2-20 AA: 4 min single coupling 21-62 AA: 4 min double coupling	10% piperazine in NMP/EtOH (9:1 % v/v) + 0.1 M HOBT
TFA-Cleavage:	Purification gradient: P2	Obtained purity:
Cocktail III 4 h	Kinetex® C18 column	> 96%



**HPLC system:** Primaide® Kinetix® C18 column;

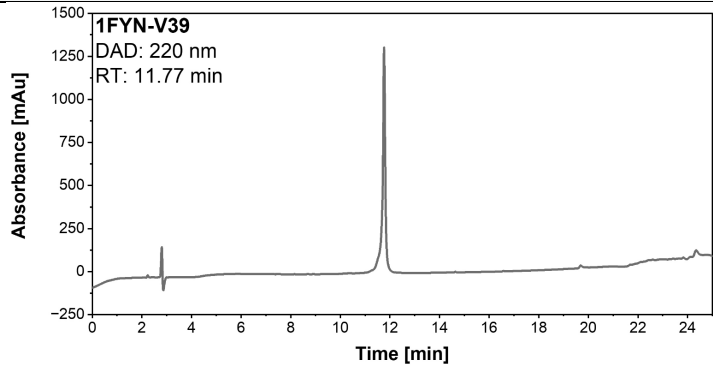
eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A18**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+8H] <sup>8+</sup>	867.2939	867.8652	14.1 mg
[M+9H] <sup>9+</sup>	771.0399	771.6123	

**Table 14.47:** Conditions for the synthesis, isolation, and purification of **1FYN-V39**.

Name	Sequence	
<b>1FYN-V39</b>	H <sub>2</sub> N-GTGVTLFVALYDYEARTEDDLSF <b>H</b> KGEKFQILNSSE-G <b>DW</b> WEVRSLTTGETGYIPSNYVAPVD -OH	
Scale	Resin	Amount of Resin
0.05 mmol	Cl-MPA pro tide resin (0.23 mmol/g)	217.4 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA :DCA loading 2-20 AA: 4 min single coupling 21-62 AA: 4 min double coupling	10% piperazine in NMP/EtOH (9:1 % v/v) + 0.1 M HOBT

TFA-Cleavage:	Purification gradient: P2	Obtained purity:
Cocktail III 4 h	Kinetex® C18 column	> 96%

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+6H] <sup>6+</sup>	1163.7742	1163.4315	9.3 mg
[M+7H] <sup>7+</sup>	997.6647	997.3117	

**Table 14.48:** Conditions for the synthesis, isolation, and purification of **1FYN-V50**.

Name	Sequence
<b>1FYN-V50</b>	H <sub>2</sub> N-GTGVTLFVALYDYEARTEDDLSEKGEKFKQLNSSE-GDWWEARSLTTGETGYVPSNYVAPVD-OH

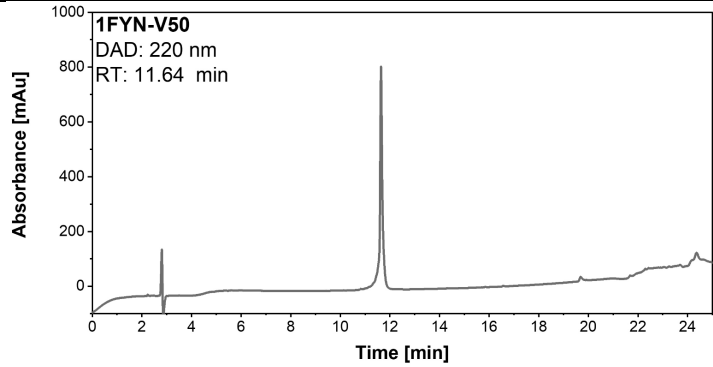
Scale	Resin	Amount of Resin
0.05 mmol	Cl-MPA pro tide resin (0.23 mmol/g)	217.7 mg

Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: DCA loading 2-20 AA: 4 min single coupling 21-62 AA: 4 min double coupling	10% piperazine in NMP/EtOH (9:1 % v/v) + 0.1 M HOBt

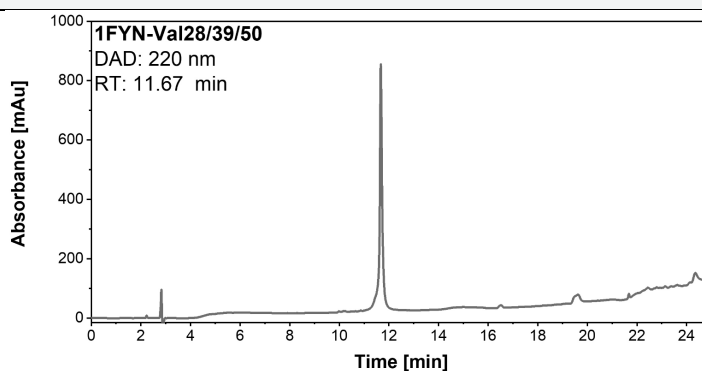
TFA-Cleavage:	Purification gradient: P2	Obtained purity:
Cocktail III 4 h	Kinetex® C18 column	> 96%

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+6H] <sup>6+</sup>	1156.0559	1156.3251	20.8 mg
[M+7H] <sup>7+</sup>	991.6532	991.4563	

**Table 14.49:** Conditions for the synthesis, isolation, and purification of **1FYN-V28/39/50**.

Name		Sequence
<b>1FYN-V28/39/50</b>		H <sub>2</sub> N-GTGVTLFVALYDYEARTEDDLSEFKGEKFQVLNSSE-GDWWEVRSLTTGETGYVPSNYVAPVD -OH
Scale	Resin	Amount of Resin
0.05 mmol	Cl-MPA pro tide resin (0.23 mmol/g)	218.1 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: DCA loading 2-20 AA: 4 min single coupling 21-62 AA: 4 min double coupling	10% piperazine in NMP/EtOH (9:1 % v/v) + 0.1 M HOBt
TFA-Cleavage:	Purification gradient: P2	Obtained purity:
Cocktail III 4 h	Kinetex® C18 column	> 96%



**HPLC system:** Primaide® Kinetix® C18 column;  
eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A18**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+5H] <sup>5+</sup>	1389.8687	1390.0321	
[M+6H] <sup>6+</sup>	1158.3918	1158.7249	17.9 mg
[M+7H] <sup>7+</sup>	993.0513	993.2167	

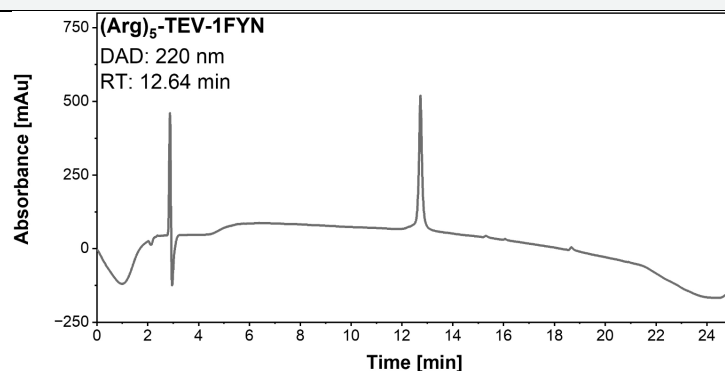
### Solubility-tagged 1FYN Derivatives

Two derivatives were prepared with a poly-Arg solubility tag due to the aggregation challenges of the 1FYN samples obtained. This tag was introduced after the recognition sequence for the TEV protease. Since Fmoc-Arg(Pbf)-OH is fundamentally difficult to introduce into the peptide backbone, the last three Arg residues were each implemented with a triple coupling for 4 min at 90°C.

The tagged 1FYN-wt variant showed significantly improved solubility and purification at acidic pH, but the tagged 1FYN-3xV derivative was only slightly more soluble.

**Table 14.50:** Conditions for the synthesis, isolation, and purification of **(Arg)<sub>5</sub>-TEV-1FYN-wt**.

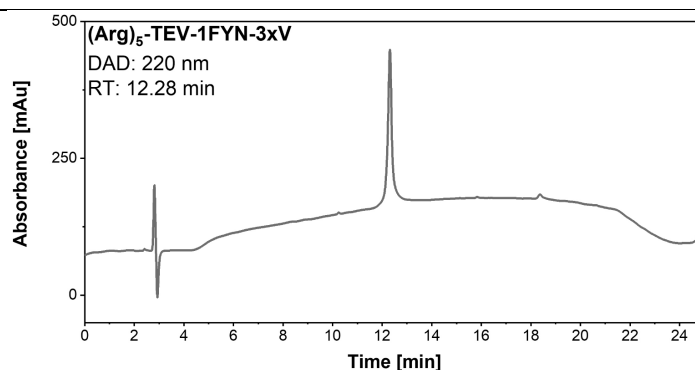
Name	Sequence	
<b>(Arg)<sub>5</sub>-TEV-1FYN-wt</b>	H <sub>2</sub> N-(Arg) <sub>5</sub> -ENLYFQ-GTGVTLFVALYDYEARTEDDLSFHKGEKFQILNSSE-GDWWEARSLTTGETGYIPSNYVAPVD -OH	
Scale	Resin	Amount of Resin
0.05 mmol	Cl-MPA pro tide resin (0.23 mmol/g)	217.9 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: DCA loading	10% piperazine in
	2-20 AA: 4 min single coupling	NMP/EtOH
	21-70 AA: 4 min double coupling	(9:1 % v/v)
	71-73 AA: 4 min triple coupling	+ 0.1 M HOBT
TFA-Cleavage:	Purification gradient: P2	Obtained purity:
Cocktail III 4 h	Kinetex® C18 column	> 96%



HPLC system: Primaide® Kinetix® C18 column;			
eluents: A = H <sub>2</sub> O, B = MeCN both containing 0.1% (v/v) TFA; gradient: A18			
ESI-ToF	calc. m/z	obs. m/z	Yield
[M+4H] <sup>4+</sup>	2132.0939	2132.0070	6 mg
[M+5H] <sup>5+</sup>	1705.8767	1705.8020	
[M+6H] <sup>6+</sup>	1421.7319	1421.7193	

**Table 14.51:** Conditions for the synthesis, isolation, and purification of **(Arg)<sub>5</sub>-TEV-1FYN-3xV**.

Name	Sequence	
<b>(Arg)<sub>5</sub>-TEV-1FYN-3xV</b>	H <sub>2</sub> N-(Arg) <sub>5</sub> -ENLYFQ-GTGVTLFVALYDYEARTEDDLSFHKGEKFQVLNSSE-GDWWEVRSLTTGETGYVPSNYVAPVD -OH	
Scale	Resin	Amount of Resin
0.05 mmol	Cl-MPA pro tide resin (0.23 mmol/g)	217.8 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: DCA loading	10% piperazine in
	2-20 AA: 4 min single coupling	NMP/EtOH
	21-70 AA: 4 min double coupling	(9:1 % v/v)
	71-73AA: 4 min triple coupling	+ 0.1 M HOBT
TFA-Cleavage:	Purification gradient: P2	Obtained purity:
Cocktail III 4 h	Kinetex® C18 column	> 98%



**HPLC system:** Primaide<sup>®</sup> Kinetix<sup>®</sup> C18 column;

eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A18**

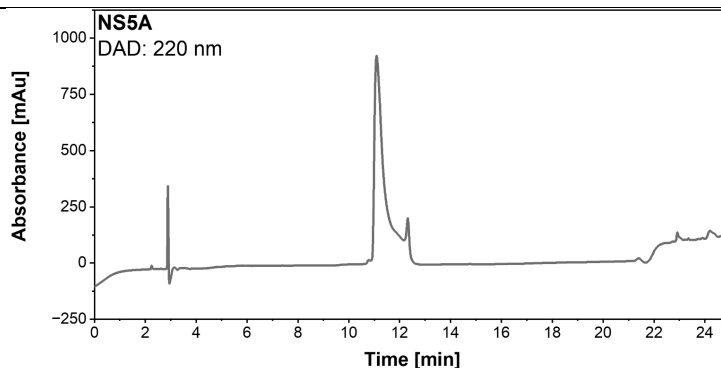
ESI-ToF	calc. m/z	obs. m/z	Yield
[M+4H] <sup>4+</sup>	2132.0939	2132.0063	
[M+5H] <sup>5+</sup>	1705.8767	1705.8053	5.4 mg
[M+6H] <sup>6+</sup>	1421.7319	1421.7193	

### *Pro-rich Peptides*

The pro-rich peptides were prepared as additives for crystallization studies. To avoid influences of charges, the peptides were prepared as C-terminal amides with acetylated N-term. A rink-amide pro tide resin was used for this purpose. The cleavage was performed with cocktail I. Upon purification, a dynamic equilibrium between two peaks was found to be due to a cis/trans isomerism of Pro.

**Table 14.52:** Conditions for the synthesis, isolation, and purification of **NS5A**.

Name	Sequence	
<b>NS5A</b>	Ac-APPIPPRRKR-CONH <sub>2</sub>	
Scale	Resin	Amount of Resin
0.1 mmol	Rink amide pro tide (0.19 mmol/g)	526.5 mg
Synthesis	Coupling conditions	Fmoc-removal
MW-SPPS	1 AA: double deprotection, 4 min single coupling 2-11 AA: 2 min single coupling Capping at N-term	20% Pip in DMF
TFA-Cleavage:	Purification gradient: P13	Obtained purity:
Cocktail I 2 h	Kinetex <sup>®</sup> C18 column	

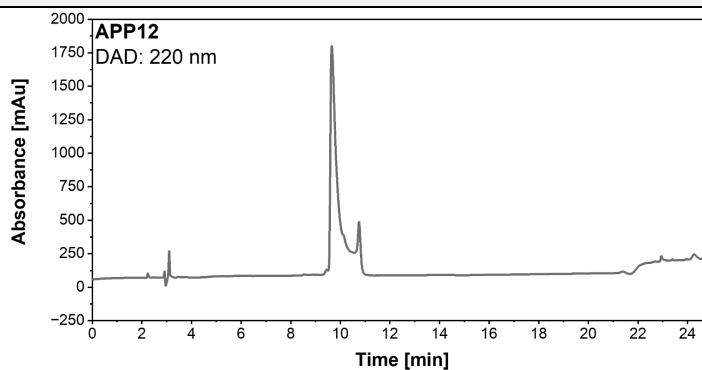


**HPLC system:** Chromaster® Kinetix® C18 column;  
 eluents: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A6**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+1H] <sup>1+</sup>	1325.8281	1325.8216	15.8 mg
[M+2H] <sup>2+</sup>	663.4180	663.4171	

**Table 14.53:** Conditions for the synthesis, isolation, and purification of **APP12**.

Name	Sequence	
<b>APP12</b>	Ac-APPLPPRNRRL-CONH <sub>2</sub>	
Scale	Resin	Amount of Resin
0.1 mmol	Rink amide pro tide (0.19 mmol/g)	526.3 mg
Synthesis	Coupling conditions	Fmoc-removal
MW-SPPS	1 AA: double deprotection, 4 min single coupling 2-12 AA: 2 min single coupling Capping at N-term	20% Pip in DMF
TFA-Cleavage:	Purification gradient: P13	Obtained purity:
Cocktail I 2 h	Kinetex® C18 column	



**HPLC system:** Chromaster® Kinetix® C18 column;  
 eluents: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A6**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+1H] <sup>1+</sup>	1424.8602	1424.8542	49.5
[M+2H] <sup>2+</sup>	712.9340	712.9324	

**Table 14.54:** Conditions for the synthesis, isolation, and purification of **VSL12**.

Name	Sequence		
<b>VSL12</b>	Ac-VSLARRPLPPLP- CONH <sub>2</sub>		
Scale	Resin	Amount of Resin	
0.1 mmol	Rink amide pro tide (0.19 mmol/g)	526.2 mg	
Synthesis	Coupling conditions	Fmoc-removal	
MW-SPPS	1 AA: double deprotection, 4 min single coupling 2-12 AA: 2 min single coupling Capping at N-term	20% Pip in DMF	
TFA-Cleavage:	Purification gradient: P13	Obtained purity:	
Cocktail I 2 h	Kinetex® C18 column	> 98%	

<b>HPLC system:</b> Chromaster® Kinetix® C18 column; eluent: A = H <sub>2</sub> O, B = MeCN both containing 0.1% (v/v) TFA; gradient: <b>A6</b>			
ESI-ToF	calc. m/z	obs. m/z	Yield
[M+1H] <sup>1+</sup>	1356.8479	1356.8428	26.2 mg
[M+2H] <sup>2+</sup>	678.9278	678.9268	

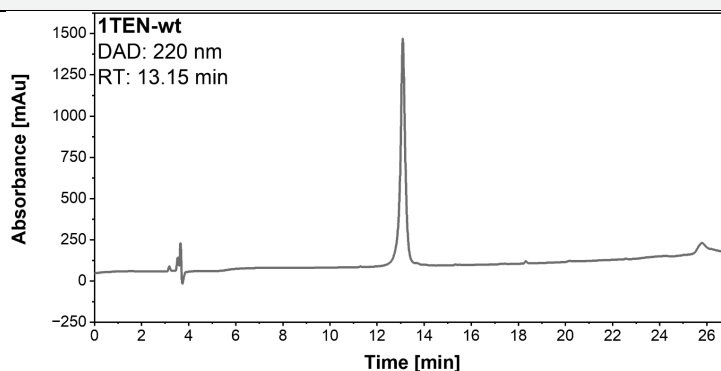
### 14.3.7 Part A: 1TEN and its Derivatives

The synthesis of 1TEN and the derivatives was performed in two steps, first coupling AA 1-60. Subsequently, the last 30 AA were introduced in a second synthesis. For the last section, a 4 min double coupling was used. Fmoc-TfeGly was coupled according to the established protocol with 2 eq. The cleavage was performed with Cocktail III. Before purification, the sample was incubated overnight in the starting conditions for preparative HPLC. Since the wide pore preparative Jupiter column was defective during the doctoral thesis, purification was performed on the Kinetex column. It is recommended to use a wide-pored material to purify such large systems since an improved separation is achieved.



**Table 14.55:** Conditions for the synthesis, isolation, and purification of **1TEN-wt**.

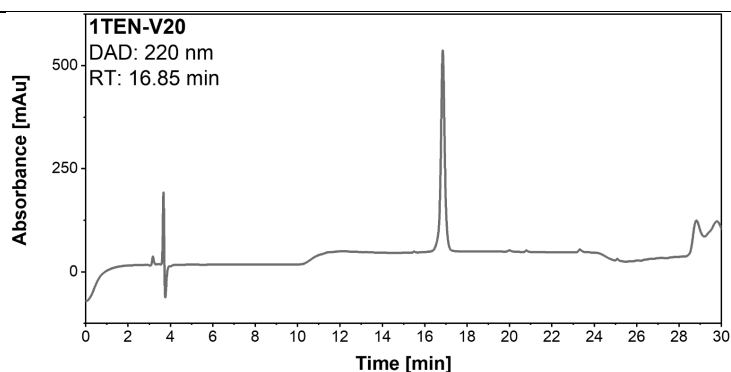
Name	Sequence	
<b>1TEN-wt</b>	H <sub>2</sub> N- RLDAPSQIEVKDVTDTTALITWFKPLAEIDGIELTYGIKDVPGDRTTIDLTEDE-NQYSIGNLKPDTYEYVSLISRRGDMSSNPAKETFTT-OH	
Scale	Resin	Amount of Resin
0.05 mmol	Cl-MPA pro tide resin (0.23 mmol/g)	217.7 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: DCA loading 2-20 AA: 2 min single coupling 21-60 AA: 2 min double coupling 61-90 AA: 4 min double coupling	10% piperazine in NMP/EtOH (9:1 % v/v) + 0.1 M HOBt
TFA-Cleavage:	Purification gradient: P3	Obtained purity:
Cocktail III 4 h	Kinetex® C18 column	> 98%



<b>HPLC system:</b> Primaide® Jupiter® C18 column (300 Å)			
eluents: A = H <sub>2</sub> O, B = MeCN both containing 0.1% (v/v) TFA; gradient: <b>A18</b>			
ESI-ToF	calc. m/z	obs. m/z	Yield
[M+8H] <sup>8+</sup>	1257.5235	1257.4371	13 mg
[M+9H] <sup>9+</sup>	1117.9016	1117.8613	
[M+10H] <sup>10+</sup>	1006.2203	1006.6284	

**Table 14.56:** Conditions for the synthesis, isolation, and purification of **1TEN-V20**.

Name	Sequence	
<b>1TEN-V20</b>	H <sub>2</sub> N- RLDAPSQIEVKDVTDTTALVTWFKPLAEIDGIELTYGIKDVPGDRTTIDLTEDE-NQYSIGNLKPDTYEYVSLISRRGDMSSNPAKETFTT-OH	
Scale	Resin	Amount of Resin
0.05 mmol	Cl-MPA pro tide resin (0.23 mmol/g)	218.1 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: DCA loading 2-20 AA: 2 min single coupling 21-60 AA: 2 min double coupling 61-90 AA: 4 min double coupling	10% piperazine in NMP/EtOH (9:1 % v/v) + 0.1 M HOBt
TFA-Cleavage:	Purification gradient: P3	Obtained purity:
Cocktail III 4 h	Kinetex® C18 column	> 96%

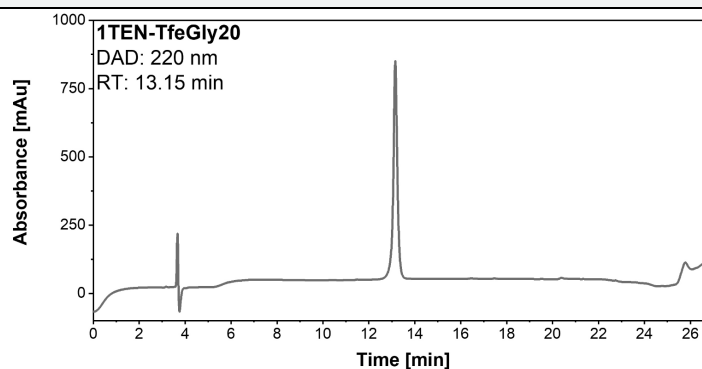


**HPLC system:** Primaide® Jupiter® C18 column (300 Å);  
 eluents: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A21**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+10H] <sup>10+</sup>	1004.8177	1004.3443	10 mg
[M+11H] <sup>11+</sup>	913.5622	913.2085	

**Table 14.57:** Conditions for the synthesis, isolation, and purification of **1TEN-TfeGly20**.

Name	Sequence	
<b>1TEN-TfeGly20</b>	H <sub>2</sub> N- RLDAPSQIEVKDVTDTTALT <b>TfeGly</b> TWFKPLAEIDGIELTYGIKDVPGDRTTIDLTEDE-NQYSIGNLKPDTHEYVSLISRRGDMSSNPAKETFTT-OH	
Scale	Resin	Amount of Resin
0.05 mmol	Cl-MPA pro tide resin (0.23 mmol/g)	217.8 mg
Synthesis	Coupling conditions CMA	Fmoc-removal
MW-SPPS	1 AA: DCA loading 2-20 AA: 2 min single coupling 21-60 AA: 2 min double coupling 61-90 AA: 4 min double coupling	10% piperazine in NMP/EtOH (9:1 % v/v) + 0.1 M HOBT
TFA-Cleavage:	Purification gradient: P3	Obtained purity:
Cocktail III 4 h	Kinetex® C18 column	> 96%



**HPLC system:** Primaide® Jupiter® C18 column (300 Å);  
 eluents: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient **A18**

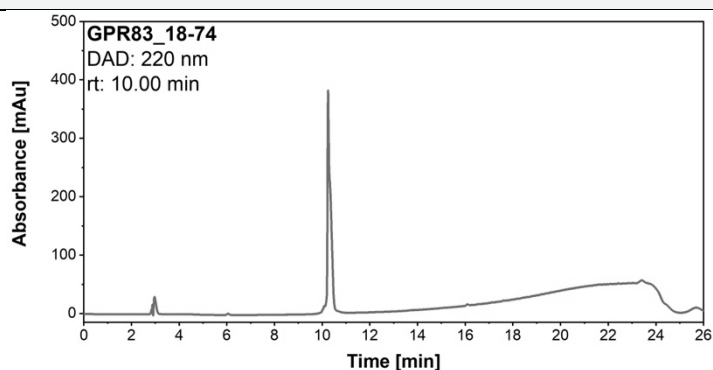
ESI-ToF	calc. m/z	obs. m/z	Yield
[M+6H] <sup>6+</sup>	1680.6688	1680.9054	11 mg
[M+7H] <sup>7+</sup>	1440.7172	1440.9701	
[M+8H] <sup>8+</sup>	1260.7535	1260.9408	

### 14.3.8 Part B: GPR83-Project

In this project, various methods were developed for synthesizing the peptides. These can be found in the tables for the respective peptides.

**Table 14.58:** Conditions for the synthesis, isolation, and purification of **GPR83\_18-74**.

Name	Sequence	
<b>GPR83_18-74</b>	H <sub>2</sub> N-EPHEGRADEQSAEAALAVPNASHFFSWNNYTFSDWQNFVGRRRYGAESQ-NPTVKALL-OH	
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Leu-TGA-NovaSyn (0.19 mmol/g)	2262.8 mg
Synthesis	Coupling conditions	Fmoc-removal
MW-SPPS	1-30: AA 2 min single coupling 31-52AA: 2 min double coupling	20% Pip in DMF +0.1 HOBt
TFA-Cleavage:	Purification gradient: P4	Obtained purity:
Cocktail III 2.5 h	Kinetex® C18 column	98%



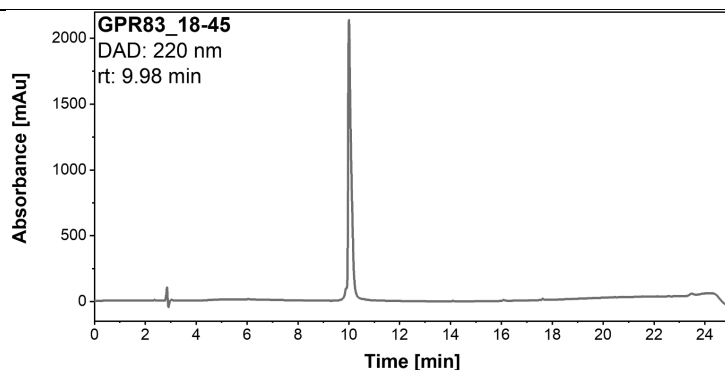
HPLC system: Chromaster® Kinetex® C18 column;

eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A9**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+4H] <sup>4+</sup>	1618.2548	1618.3102	
[M+5H] <sup>5+</sup>	1294.8054	1294.8487	14.1 mg
[M+6H] <sup>6+</sup>	1079.1725	1079.2069	

**Table 14.59:** Conditions for the synthesis, isolation, and purification of **GPR83\_18-45**.

Name	Sequence	
<b>GPR83_18-45</b>	Abz-EPHEGRADEQSAEAALAVPNASHFFSWN-OH	
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Asn(Trt) TGA-NovaSyn (0.17 mmol/g)	294.3 mg
Synthesis	Coupling conditions	Fmoc-removal
MW-SPPS	1-28 AA: 2 min single coupling	20% Pip in DMF + 0.1M HOBt
TFA-Cleavage:	Purification gradient: P4	Obtained purity:
Cocktail I 2.5 h	Kinetex® C18 column	98%

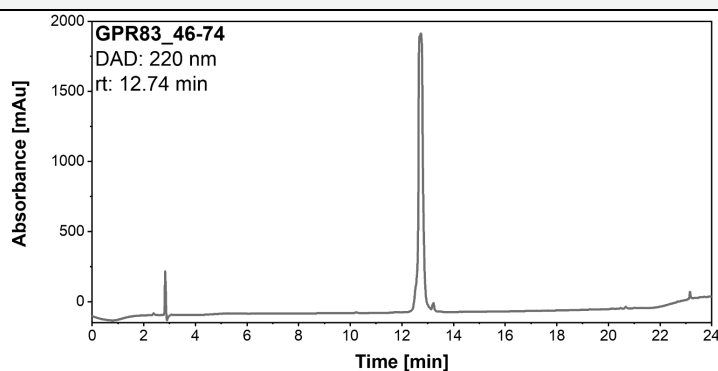


HPLC system: Chromaster® Kinetex® C18 column;  
 eluents: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A9**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+2H] <sup>2+</sup>	1593.6966	1593.7379	11.4 mg
[M+3H] <sup>3+</sup>	1062.8004	1062.5479	

**Table 14.60:** Conditions for the synthesis, isolation, and purification of **GPR83\_46-74**.

Name	Sequence	
GPR83_46-74	Abz-NYTFSDWQNFVGRRRRYGAESQNPTVKALL-COOH	
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Leu-TGA NovaSyn (0.19 mmol/g)	263.5 mg
Synthesis	Coupling conditions	Fmoc-removal
MW-SPPS	1-28 AA: 2 min single coupling	20% Pip in DMF + 0.1 M HOBt
TFA-Cleavage:	Purification gradient: P9	Obtained purity:
Cocktail III 2.5 h	Kinetex® C18 column	97%

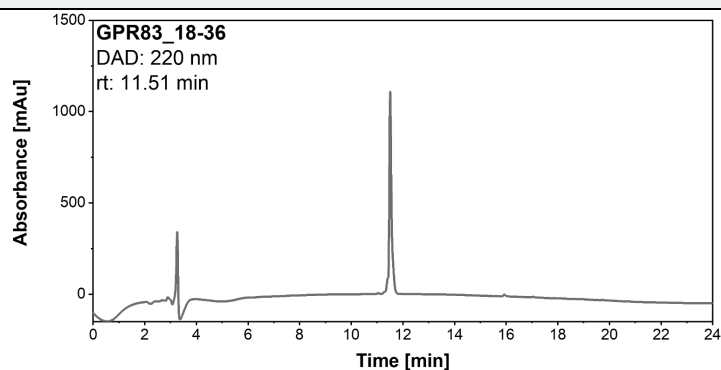


HPLC system: Chromaster® Kinetex® C18 column;  
 eluents: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A10**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+3H] <sup>3+</sup>	1179.5723	1179.5631	9.3 mg
[M+4H] <sup>4+</sup>	884.9312	884.8973	

**Table 14.61:** Conditions for the synthesis, isolation, and purification of **GPR83\_18-36**.

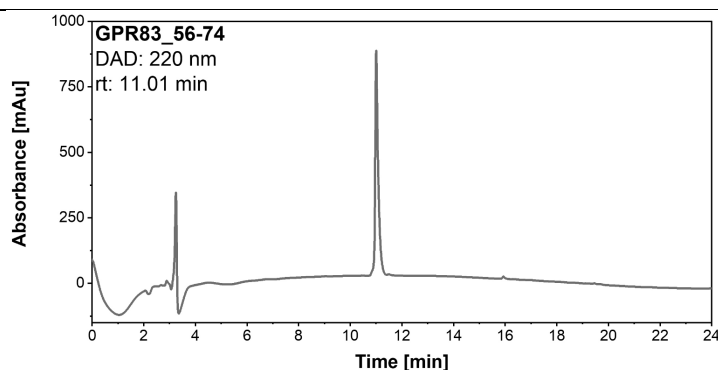
Name		Sequence
GPR83_18-36		Abz-EPHEGRADEQSAEAALAVP-COOH
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Pro-TGT NovaSyn (0.23 mmol/g)	263.4 mg
Synthesis	Coupling conditions	Fmoc-removal
Conventional SPPS P-11	HOAt chemistry according to Table 14.1	20% Pip in DMF + 0.1 M HOBT
TFA-Cleavage:	Purification gradient: P9	Obtained purity:
Cocktail I 2.5 h	Kinetex® C18 column	> 98%



HPLC system: LaChrom® Kinetex® C18 column; eluent: A = H <sub>2</sub> O, B = MeCN both containing 0.1% (v/v) TFA; gradient: <b>A16</b>			
ESI-ToF	calc. m/z	obs. m/z	Yield
[M+1H] <sup>1+</sup>	2095.9234	2095.9599	12.1 mg
[M+2H] <sup>2+</sup>	1048.4656	1048.4853	

**Table 14.62:** Conditions for the synthesis, isolation, and purification of **GPR83\_56-74**.

Name		Sequence
GPR83_56-74		Abz-VGRRRYGAESQNPTVKALL-COOH
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Leu-TGA NovaSyn (0.19 mmol/g)	263.1 mg
Synthesis	Coupling conditions	Fmoc-removal
Conventional SPPS P-11	HOAt chemistry according to Table 14.1	20% Pip in DMF
TFA-Cleavage:	Purification gradient: P9	Obtained purity:
Cocktail I 2.5 h	Kinetex® C18 column	> 98%



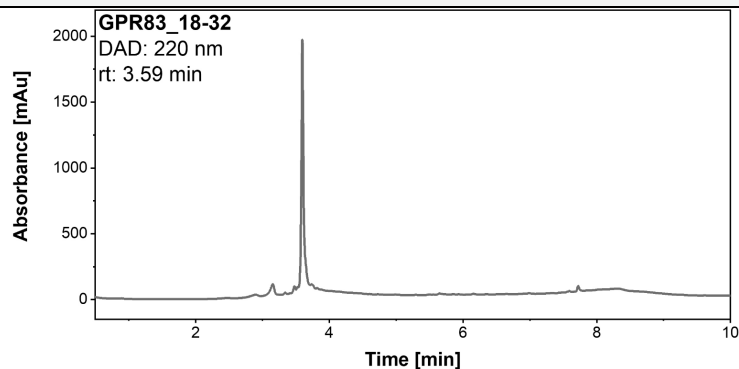
HPLC system: LaChrom® Kinetex® C18 column;

eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A15**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+2H] <sup>2+</sup>	1117.5893	1117.5489	14.6 mg
[M+3H] <sup>3+</sup>	745.3954	745.3785	

**Table 14.63:** Conditions for the synthesis, isolation, and purification of **GPR83\_18-32**.

Name	Sequence	
GPR83_18-32	Abz-EPHEGRADEQSAEAA-COOH	
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Ala-Wang LL (0.32 mmol/g)	157.5 mg
Synthesis	Coupling conditions	Fmoc-removal
Conventional	HOBt chemistry according to Table 14.1	20% Pip in DMF
SPPS		
Manual		
TFA-Cleavage:	Purification gradient: T2	Obtained purity:
Cocktail I 2.5 h	Luna® C8(2) column	95%



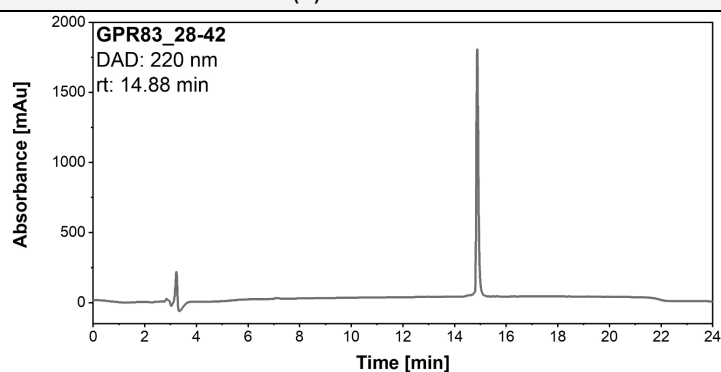
HPLC system: Chromaster® (FF) Purospher® STAR RP-C18 end-capped UHPLC column

eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A1**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+3H] <sup>3+</sup>	1715.6811	1715.7355	5.2 mg
[M+4H] <sup>4+</sup>	858.3344	858.3727	

**Table 14.64:** Conditions for the synthesis, isolation, and purification of **GPR83\_28-42**.

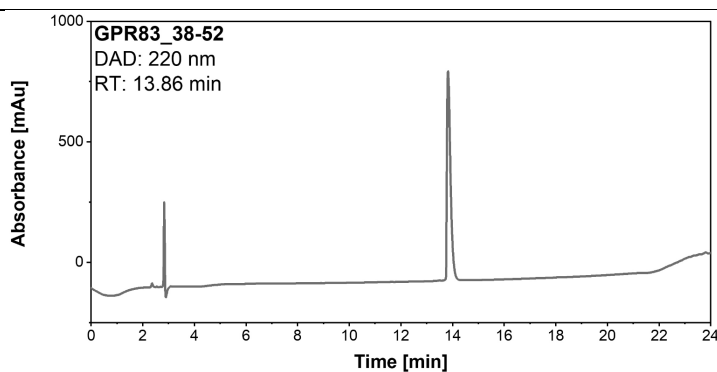
Name		Sequence
GPR83_28-42		Abz-SAEAALAVPNASHFF-COOH
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Phe-TGA NovaSyn (0.19 mmol/g)	266.4 mg
Synthesis	Coupling conditions	Fmoc-removal
Conventional SPPS Manual	HOBt chemistry according to Table 14.1	20% Pip in DMF
TFA-Cleavage:	Purification gradient: T1	Obtained purity:
Cocktail I 2.5 h	Luna® C8(2) column	> 98%



HPLC system: LaChrom® Luna® C8(2) column; eluent: A = H <sub>2</sub> O, B = MeCN both containing 0.1% (v/v) TFA; gradient: <b>A16</b>			
ESI-ToF	calc. m/z	obs. m/z	Yield
[M+1H] <sup>1+</sup>	1650.7466	1650.7842	9.6 mg
[M+2H] <sup>2+</sup>	825.8772	825.8973	

**Table 14.65:** Conditions for the synthesis, isolation, and purification of **GPR83\_38-52**.

Name		Sequence
GPR83_38-52		Abz-ASHFFSWNNYTFSD-COOH
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Asp(O <sup>t</sup> Bu)-TGA-NovaSyn (0.24 mmol/g)	210 mg
Synthesis	Coupling conditions	Fmoc-removal
Conventional SPPS Manual	HOAt chemistry according to Table 14.1	20% Pip in DMF
TFA-Cleavage:	Purification gradient: P9	Obtained purity:
Cocktail III 2.5 h	Kinetex® C18 column	> 98%



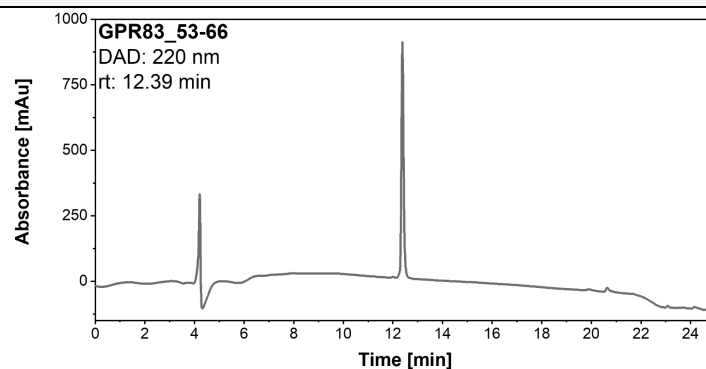
HPLC system: Chromaster® Kinetex® C18 column;

eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A10**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+1H] <sup>1+</sup>	1842.7904	1842.7261	7.9 mg
[M+2H] <sup>2+</sup>	921.3594	921.3878	

**Table 14.66:** Conditions for the synthesis, isolation, and purification of **GPR83\_53-66**.

Name		Sequence
GPR83_53-66		Abz-QNFVGRRRRYGAESQ-COOH
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Gln(Trt)-TGA NovaSyn (0.21 mmol/g)	238.3 mg
Synthesis	Coupling conditions	Fmoc-removal
Conventional		
SPPS	HOBt chemistry according to Table 14.1	20% Pip in DMF
Manual		
TFA-Cleavage:	Purification gradient: T2	Obtained purity:
Cocktail I 2.5 h	Luna® C8(2) column	98%



HPLC system: LaChrom® Luna® C8(2) column

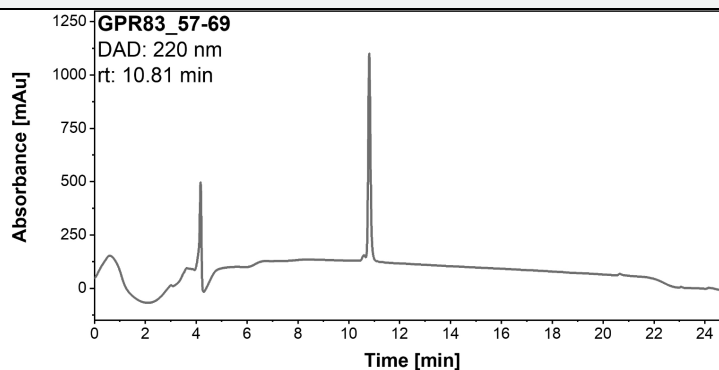
eluent: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A16**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+1H] <sup>1+</sup>	1786.8287	1789.8195	23.9 mg
[M+2H] <sup>2+</sup>	893.9183	893.9172	



**Table 14.67:** Conditions for the synthesis, isolation, and purification of **GPR83\_57-69**.

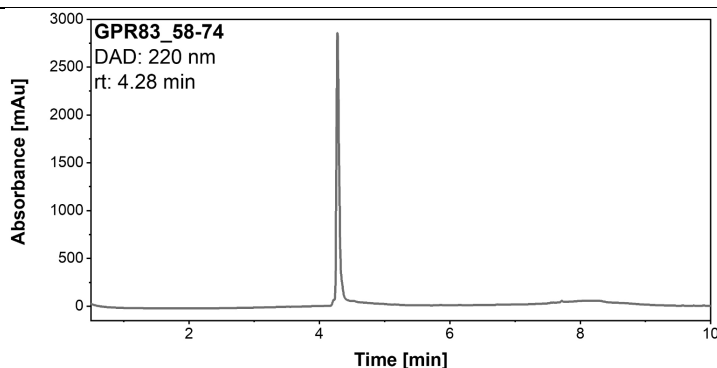
Name		Sequence	
GPR83_57-69		Abz-GRRRYGAESQNPT-COOH	
Scale	Resin	Amount of Resin	
0.05 mmol	Fmoc-Thr( <sup>t</sup> Bu)-TGA NovaSyn (0.18 mmol/g)	278 mg	
Synthesis	Coupling conditions	Fmoc-removal	
Conventional SPPS Manual	HOBt chemistry according to Table 14.1	20% Pip in DMF	
TFA-Cleavage:	Purification gradient: T1	Obtained purity:	
Cocktail III 2.5 h	Luna <sup>®</sup> C8 column	98%	



HPLC system: LaChrom <sup>®</sup> Luna <sup>®</sup> C8(2) column; eluent: A = H <sub>2</sub> O, B = MeCN both containing 0.1% (v/v) TFA; gradient: <b>A16</b>			
ESI-ToF	calc. m/z	obs. m/z	Yield
[M+1H] <sup>1+</sup>	1610.7338	1610.7259	33.5 mg
[M+2H] <sup>2+</sup>	805.8708	805.9003	

**Table 14.68:** Conditions for the synthesis, isolation, and purification of **GPR83\_58-74**.

Name		Sequence	
GPR83_58-74		Abz-RRRYGAESQNPTVKALL-COOH	
Scale	Resin	Amount of Resin	
0.05 mmol	Fmoc-Leu- TGA NovaSyn (0.19 mmol/g)	264.5 mg	
Synthesis	Coupling conditions	Fmoc-removal	
Conventional SPPS Manual	HOBt chemistry See Table 10.3	20% Pip in DMF	
TFA-Cleavage:	Purification gradient: T2	Obtained purity:	
Cocktail I 2.5 h	Luna <sup>®</sup> C8(2) column	96%	

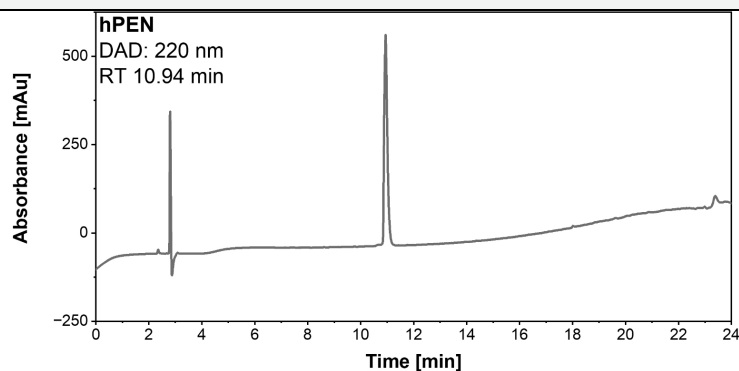


HPLC system: Chromaster® Purospher® STAR RP-C18 end-capped UHPLC column;  
 eluents: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A10**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+2H] <sup>2+</sup>	1040.1237	1040.0631	17 mg
[M+3H] <sup>3+</sup>	693.3655	693.3786	
[M+4H] <sup>4+</sup>	520.2761	520.2858	

**Table 14.69:** Conditions for the synthesis, isolation, and purification of **hPEN**.

Name	Sequence	
<b>hPEN</b>	H <sub>2</sub> N-AADHDVGSPELPEGVLGALLRV-COOH	
Scale	Resin	Amount of Resin
0.05 mmol	Fmoc-Val-Wang NovaSyn (0.32 mmol/g)	156.4 mg
Synthesis	Coupling conditions	Fmoc-removal
Conventional	HATU chemistry	20% Pip in DMF
SPPS	See Table 10.4	
P-11	capping after LLR	
	Fmoc-Asp(OEpe)-OH	
TFA-Cleavage:	Purification gradient: P4	Obtained purity:
Cocktail I 2.5 h	Kinetex® C18 column	96%



HPLC system: Chromaster® Kinetex® C18 column;  
 eluents: A = H<sub>2</sub>O, B = MeCN both containing 0.1% (v/v) TFA; gradient: **A10**

ESI-ToF	calc. m/z	obs. m/z	Yield
[M+1H] <sup>1+</sup>	2215.1721	2215.1834	20 mg
[M+2H] <sup>2+</sup>	1108.0900	1108.0985	

### 14.3.9 Performed Reactions with Synthesized Peptides

#### 14.3.9.1 Native Chemical Ligation (NCL)

In each case, the performance of the reaction was challenging because the synthesized Ub monomers are only very poorly soluble in the corresponding buffers. Extensive ultrasonic treatment in combination with centrifugation is essential for the dissolution of the samples.

##### ***NCL with NHNH<sub>2</sub> as thioester surrogates***

The synthesis followed the procedures by Zheng *et al.* and Pan *et al.* [12,13]

1 mg Ub2 (1 eq.) was dissolved at rt in 0.2 M phosphate solution containing 6 M Gnd-HCl (pH 3.0-3.1) (total conc. 2 mM). For this purpose, the sample was treated for a long time in an ultrasonic bath and repeatedly centrifuged to suppress foam formation. The solution was then cooled to -15 to -10°C and treated with 0.5 M NaNO<sub>2</sub> (10 eq.) solution. The resulting reaction mixture was gently stirred for 20 min at min -10°C. Subsequently, the reaction mixture was warmed to rt, and MPAA (40 eq.) was added. The pH was adjusted to 5.0, and the reaction mixture was stirred for 5 min. 1 mg Ub1(K48-Cys) (1 eq.) was added, and the pH was carefully adjusted to 6.5 for ligation (avoid over titrating otherwise hydrolysis of the thioester will occur). After 24 h, a reaction control was performed by HPLC or SDS-PAGE.

##### ***NCL with SEA-group as thioester surrogates***

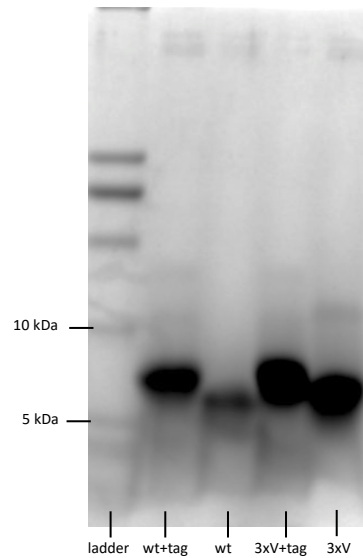
The synthesis followed the procedures by Melnyk *et al.* [14]

In a nitrogen atmosphere 1.5 mg Ub2-SEA (1 eq.) were dissolved at rt in 0.1 M phosphate buffer containing 6 M Gnd-HCl (pH 6), 60 eq. TCEP and 60 eq. MPAA (final conc. of peptide 2 mM). 2.25 mg (1.5 eq) of Ub1(K48-Cys) were added and the reaction mixture was stirred for 24 h at 37°C in a nitrogen atmosphere. After 24 h, a reaction control was performed by SDS-PAGE.

#### 14.3.9.2 Digestion with TEV

5-6 mg of purified tagged peptide were dissolved in 10 mL reaction buffer (0.5 mM EDTA + 50 mM Tris-HCl + 1 mM DTT, initial pH: 7.5). Due to poor solubility in acidic and neutral pH, the pH of the buffer solution was slowly increased to 8.4 for the

1FYN-3xV mutant (optimal TEV activity between pH 5-9). Then 100  $\mu$ L of TEV-solution (7 mg/mL, kindly provided by M. Krummhaar, MPI) were added to the peptide solution. The reaction was carried out overnight at 4°C. The reaction mixture was purified by IEC. The isolated fractions were analyzed by SDS-PAGE and combined.



**Figure 14.7:** Difference of tagged and untagged 1FYN derivatives in SDS-PAGE using 16% Tricine gels.

---

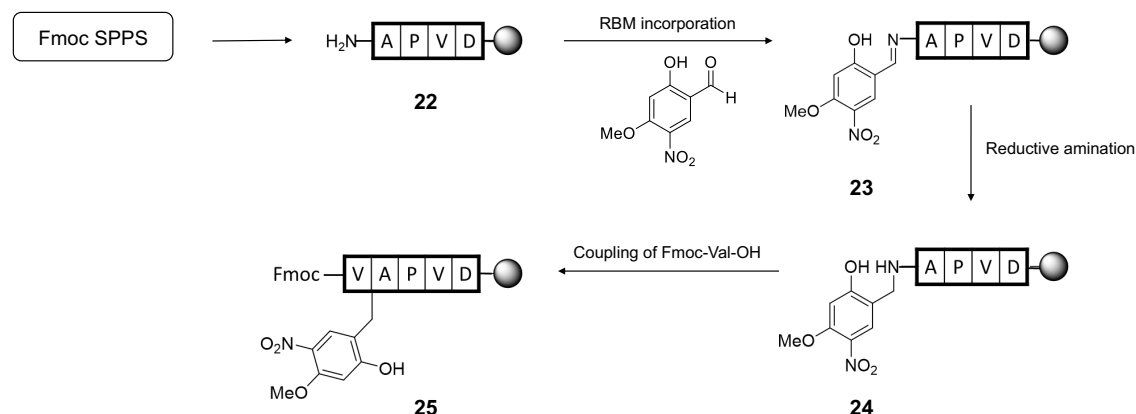
**15. References:**

- [1] aapptec, "Technical Support Information Bulletin 1198," **last access 15/01/2023**.
- [2] CEM, "CarboMax approach," **last access 15/01/2023**
- [3] H. Schägger, *Nat Protoc* **2006**, *1*, 16–22.
- [4] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, T. Gray, *Protein Science* **1995**, *4*, 2411–2423.
- [5] T. Tsushima, K. Kawada, S. Ishihara, N. Uchida, O. Shiratori, J. Higaki, M. Hirata, *Tetrahedron* **1988**, *44*, 5375–5387.
- [6] H. A. Lindner, A. Alary, L. I. Boju, T. Sulea, R. Ménard, *Biochemistry* **2005**, *44*, 15645–15651.
- [7] J. Han, R. Takeda, X. Liu, H. Konno, H. Abe, T. Hiramatsu, H. Moriwaki, V. A. Soloshonok, *Molecules* **2019**, *24*, DOI 10.3390/molecules24244521.
- [8] Z. Yin, H. Moriwaki, H. Abe, T. Miwa, J. Han, V. A. Soloshonok, *ChemistryOpen* **2019**, *8*, 701–704.
- [9] Gage James R., Evans David A., *Organic Syntheses* **1990**, *68*, 77.
- [10] H. Erdbrink, I. Peuser, U. I. M. Gerling, D. Lentz, B. Koksich, C. Czekelius, *Org Biomol Chem* **2012**, *10*, 8583.
- [11] S. Tang, C. Zuo, D.-L. Huang, X.-Y. Cai, L.-H. Zhang, C.-L. Tian, J.-S. Zheng, L. Liu, *Nat Protoc* **2017**, *12*, 2554–2569.
- [12] M. Pan, S. Gao, Y. Zheng, X. Tan, H. Lan, X. Tan, D. Sun, L. Lu, T. Wang, Q. Zheng, Y. Huang, J. Wang, L. Liu, *J Am Chem Soc* **2016**, *138*, 7429–7435.
- [13] J.-S. Zheng, S. Tang, Y.-K. Qi, Z.-P. Wang, L. Liu, *Nat Protoc* **2013**, *8*, 2483–2495.
- [14] M. Cargoët, V. Diemer, B. Snella, R. Desmet, A. Blanpain, H. Drobecq, V. Agouridas, O. Melnyk, *J Org Chem* **2018**, *83*, 12584–12594.

## 16. Appendix

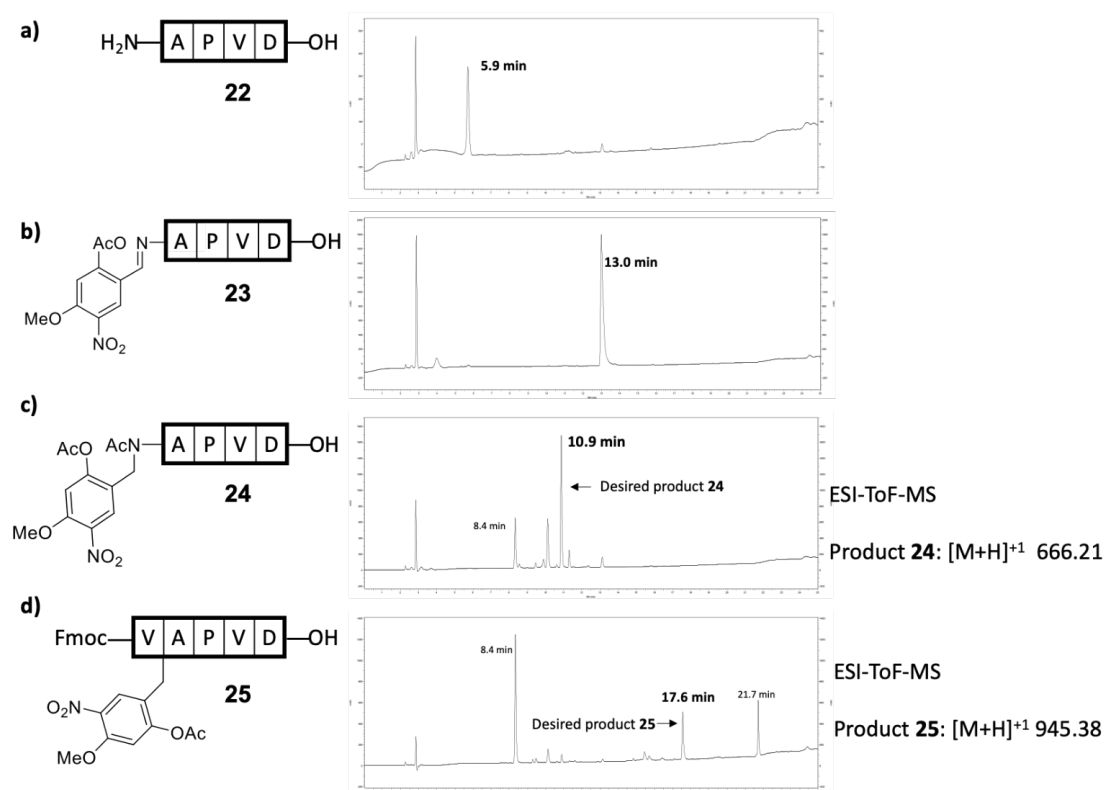
### 16.1 RMB-Optimizations

The following processes were investigated for the introduction of the RMB linker:



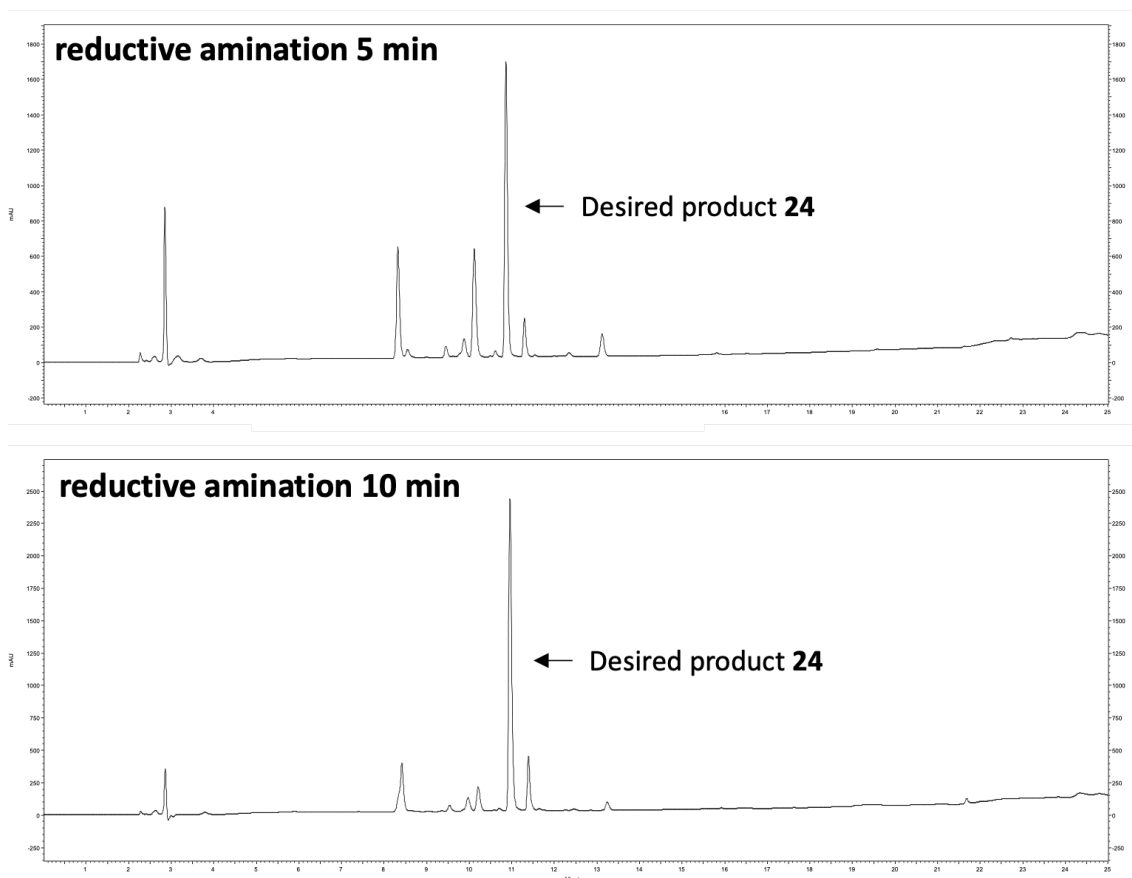
**Figure 16.1:** Processes involved in the incorporation of RMB-linker into 1FYN sequence.

HPLC analysis revealed a side-product by 8.4 min in c) and additional side-product at 21.7 min.



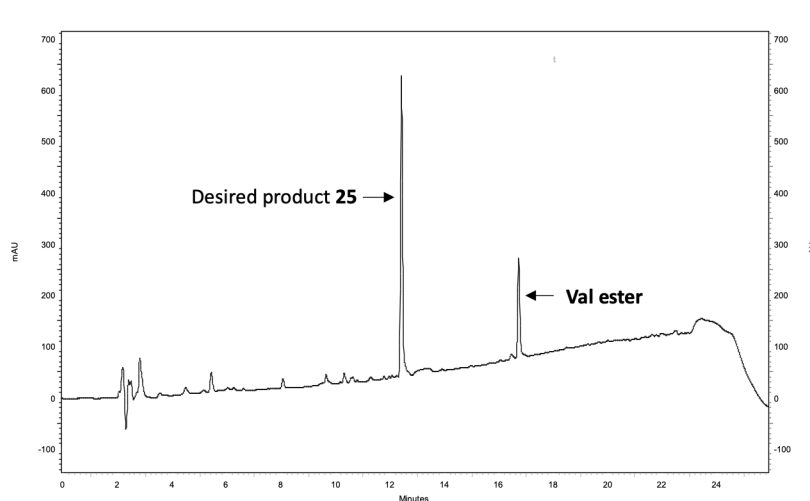
**Figure 16.2:** Incorporation of RMB linker into 1FYN-wt sequence. HPLC-Chromatograms of reaction monitoring. HPLC system: Chromaster®, Kinetex® C18 column, method 10-80% MeCN in H<sub>2</sub>O + 0.1% TFA in 18 min, DAD 220 nm.

The side-product at 8.4 min was identified as incomplete product formation of the reductive amination. Therefore, increasing the time of reductive amination (10 min instead of 5 min) led to an improvement.

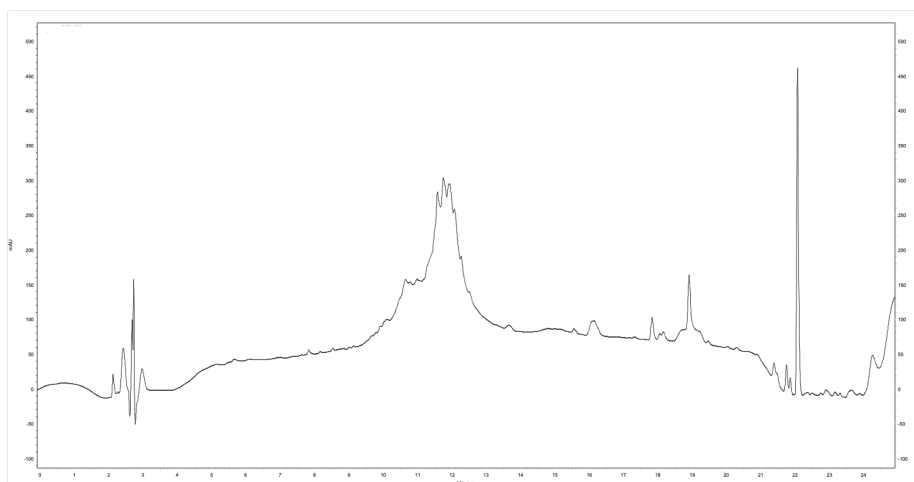


**Figure 16.3:** HPLC-Chromatograms of optimization of reductive amination. HPLC system: Chromaster®, Kinetex® C18 column, method 10-80% MeCN in H<sub>2</sub>O + 0.1% TFA in 18 min, DAD 220 nm.

The side-product at 21.7 min was identified as a Val-ester at the RMB-linker of **24**. Several coupling conditions were screened. Showing, that a decrease in coupling eq. of Val or reduced coupling time leads to an improvement, however the side-product formation is not fully suppressed. Only by treatment with 20% piperidine (as necessary for Fmoc-removal) the Val-ester is cleaved. With this procedure in hand the peptide was subjected to MW-SPPS for the full sequence, but the desired RMB-modified 1FYN-wt was not obtained. The synthesis requires further optimizations.



**Figure 16.4:** HPLC-Chromatograms of optimized Val coupling: 2 eq. 45 min coupling time with 2 eq. Oxyma pure and 2 eq. DIC at rt. HPLC system: Chromaster®, Kinetex® C18 column, method 30-100% MeCN in H<sub>2</sub>O + 0.1% TFA in 18 min, DAD 220 nm.



**Figure 16.5:** Incorporation of RMB linker into 1FYN-wt sequence. HPLC-Chromatograms of reaction monitoring after MW-SPPS. HPLC system: Chromaster®, Kinetex® C18 column, method 10-80% MeCN in H<sub>2</sub>O + 0.1% TFA in 18 min, DAD 220 nm.