



universität  
wien

# MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Auxiliary- Mediated Synthesis of Homogeneously N-glycosylated Tumor Necrosis Factor Ligand Superfamily 6 (sFasL) analogs “

verfasst von / submitted by

Jiahui Huang BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of  
Master of Science (MSc)

Wien, 2016 / Vienna 2016

Studienkennzahl lt. Studienblatt /  
degree programme code as it appears on  
the student record sheet:

A 066 862

Studienrichtung lt. Studienblatt /  
degree programme as it appears on  
the student record sheet:

Masterstudium Chemie

Betreut von / Supervisor:

Univ.-Prof. Dr. Christian Becker

Mitbetreut von / Co-Supervisor:

Dr. Claudia Bello



# Acknowledgement

Finally, today is the day I write this note of thanks as the final part of my thesis. It has been a pleasant time period of learning and growing for me, not only in the scientific areas but also on a personal level. I would like to acknowledge those people who have helped me on the journey of my master thesis and because of whom this experience has become one that I will cherish very much.

First, I would like to express my sincere gratitude to Dr. Claudia Bello for the continuous support of my master thesis. Thank you for always patiently taking the time for me whenever troubles in the Lab or questions about my research/writing have arisen, for leading me in the right direction whenever I have felt confused or frustrated and for your endless optimism and enthusiasm.

I would also like to thank Prof. Dr. Christian Becker for accepting me into his group. I'm deeply grateful for your immense knowledge and your many advices throughout my work, for all of the opportunities I have been given to realize this project and for your revision of my thesis.

In my daily work I'm very thankful for the friendly and cheerful members of the Institute of Biological Chemistry group. Particularly, I would like to acknowledge Dr. Aleksandr Kravchuk for his brilliant ideas and assistance provided during the research project, Mag. Gerhard Niederacher for his help in regard of the Endo-A expression and many other areas, Mag. Manuel Felkl and MSc. Margret Vogt for their support in handling the Liberty Blue system and optimizing the peptide synthesis. Appreciation also goes out to the other members for their comments and advices at different stages of my work, for the exchanges of knowledge and for their understanding and encouragement.

Special thanks goes out to Anna and Alex from the MS-center of the University of Vienna for their countless HPLC-MS analysis throughout my master thesis, Univ.-Prof. Dr. Christopher Gerner and his group for their help and advices with the MS-MS measurements, Prof. Dr. Kaoru Takegawa from Kyushu University for kindly providing us the plasmid of Endo-A enzyme.

I would like to thank my friends for their support and care. I deeply value their friendship and appreciate their belief in me. Also, I'm very grateful for my boyfriend, who has helped me to overcome difficulties and to stay sane during stressful times.

Last but not least, none of this could have been possible without the unconditional love and patience of my family. Therefore, I would like to dedicate this thesis to my mother and my grandparents. From the bottom of my heart, I admire you for all of your accomplishments in life and I am greatly indebted to you for all you have done for me. Thank you for supporting me through my entire life, for your encouragement and motivation at times when I needed them the most, for your knowledge and wisdom.



# Curriculum Vitae

## PERSÖNLICHE DATEN:

BSc. Jiahui (Helen) Huang  
Zur Spinnerin 16/7  
1100 Wien



Geburtsdatum: 23. August 1992  
Staatsangehörigkeit: Österreich  
Familienstand: Ledig  
Tel.: +43 650 8891669  
Email: [jiahui.huang92@yahoo.de](mailto:jiahui.huang92@yahoo.de)

## AUSBILDUNG:

- 10/2013-  
vorauss. 05/2016      **Masterstudium Chemie (Universität Wien)**  
Schwerpunkt: Biologische Chemie, Analytik, Organik  
**Alle erforderlichen Prüfungen mit Auszeichnung bestanden**
- 07/2015- laufend      **Masterarbeit (Universität Wien)**  
**Institut für biologische Chemie**  
*„Auxiliary- Mediated Synthesis of Homogeneously N- glycosylated  
Tumor Necrosis Factor Ligand Superfamily 6 (sFasL)“*  
Supervisor: Dr. Claudia Bello, Univ.-Prof. Dr. Christian Becker  
**Schwerpunkt:** chemische und biologische Proteinchemie,  
Proteinanalytik, Posttranslationale Modifikation, Organische Synthese
- 01-06/2015      **Erasmus Forschungsprojekt (University of Warwick, UK)**  
Department of Chemistry and Life Sciences  
*„Synthesis and testing of RSV Inhibitors“*  
Supervisor: As. Prof. Dr. Andrew Marsh, Prof. Andrew Easton  
**Schwerpunkt:** Organische Synthese, Zellbiologische Assay,  
Entwicklung von High Throughput Screening Method
- 10/2010:      **Bachelorstudium Chemie (Universität Wien)**  
Bachelorarbeit mit ausgezeichnetem Erfolg absolviert
- 07/2010:      **Matura mit ausgezeichnetem Erfolg** (Bundesgym. Klostergasse)

## BERUFSERFAHRUNG:

- 2013-2016 Freie Mitarbeiterin  
**Fakultät für Chemie, Universität Wien**  
▶ Betreuung und Unterstützung der Bachelorstudenten im organisch-chemischen Praktikum  
▶ Projektmitarbeit mit Dr. Walter Weissensteiner (organische Synthese)
- 08-09/2014 Praktikantin  
**Baxter AG und Baxter Innovations GmbH** (Wien, Österreich)  
▶ Qualitätssicherung und Routineanalytik der Produktion (GMP)
- 2013- 2014 Technische Mitarbeiterin (geringfügig)  
**ESW Consulting Wruss ZT GmbH** (Wien, Österreich)  
▶ Analysen der Wasser- und Erdproben mittels Massenspektrometrie (GLP)  
▶ Auswertung, Dokumentation und Präsentation der Messdaten
- 07-08/2013: Praktikantin  
**Nievelt Labor Ges.m.b.H** (Stockerau, Österreich)  
▶ Entnahme und Analysen der Wasser- und Erdproben  
▶ Auswertung, Dokumentation und Präsentation der Messdaten
- 07/2012: Praktikantin  
**Institute of Molecular Genetics** (Prag, Tschechien)  
▶ Zellbiologische Assays  
▶ Entwicklung von High Throughput Screening Methoden (Luminescence assay, Viability assay)
- 08/2010: Praktikantin  
**Ce-M-M Research Center** (Wien, Österreich)  
▶ Zellbiologische Assays, Routineanalytik

## SONSTIGE QUALIFIKATIONEN:

- Sprachen: Deutsch (Muttersprache)  
Chinesisch (Muttersprache)  
Englisch (Verhandlungssicher)  
Französisch (Niveau B1)
- PC-Kenntnisse: MS Office, Adobe, Chemdraw, Endnote, Origin Data analysis, analytisch-chemische Softwares  
(e.g. Waters LC-MS, Dionex HPLC, Bruker Maldi-MS)

# TABLE OF CONTENTS

ABSTRACT.....	4
1. INTRODUCTION.....	6
1.1. Glycoprotein synthesis.....	6
1.1.1. Glycosylation as posttranslational modification.....	6
1.1.2. Peptide synthesis (SPPS) .....	7
1.1.3. Protein synthesis (NCL) .....	8
1.1.4. Protein glycosylation.....	10
1.1.5. Multifunctional PEGylated photocleavable auxiliary.....	12
1.2. sFasL.....	13
1.2.1. The FasR/FasL death initiating pathway .....	13
1.2.2. The structure of FasL.....	14
1.2.3. Soluble Fas Ligand (sFasL) .....	15
2. AIM.....	17
3. EXPERIMENTAL PART .....	19
3.1. Material and methods .....	19
3.2. Reaction monitoring, product analysis and purification by RP-HPLC/MS .....	19
3.3. Isolation and purification of SGP from egg yolk .....	20
3.4. Synthesis of Man <sub>3</sub> GlcNAc <sub>2</sub> -peptide.....	21
3.4.1. Desialylation via neuraminidase .....	21
3.4.2. Neuraminidase activity assay.....	21
3.4.3. Desialylation via acetic acid .....	23
3.4.4. Degalactosylation via $\beta$ -1,4 galactosidase S (NEB).....	24
3.4.5. Galactosidase activity assay .....	24
3.5. Synthesis of the key-sequence (246-281aa) .....	26
3.5.1. Manual synthesis with TentaGel resin (first approach).....	26
3.5.3. Automated Liberty Blue synthesis with TentaGel resin.....	28
3.5.4. Automated Liberty Blue synthesis with ChemMatrix resin .....	28
3.5.5. Automated Liberty Blue synthesis with Wang resin .....	29

3.6.	Attempt of auxilliary attachment.....	29
3.6.1.	PEGylation of the key-sequence .....	29
3.7.	Synthesis of the remaining sequences (141-201aa; 202-245aa) .....	30
3.7.1.	Synthesis of the sequence 141-201aa .....	30
3.7.2.	Synthesis of the sequence 202-245aa .....	30
3.8.	Ligations of the N-terminal and middle sequences .....	31
3.8.1.	Hydrazine cleavage of the sequence 141-201aa .....	31
3.8.2.	Thioester preparation of the sequence 141-201aa .....	32
3.8.3.	Hydrazine cleavage of the sequence 202-245aa .....	32
3.8.4.	Ligation assays for 141-245aa-Hydrazide peptide .....	33
3.9.	Endo-A.....	34
3.9.1.	Endo-A test expression .....	34
3.9.2.	Endo-A activity assay.....	35
4.	RESULTS AND DISCUSSION.....	36
4.1.	Isolation and purification of SGP from egg yolk .....	36
4.1.1.	Isolation of SGP from egg yolk .....	36
4.1.2.	Purification of SGP crude .....	37
4.2.	Synthesis of Man <sub>3</sub> GlcNAc-peptide .....	38
4.2.1.	Desialylation.....	39
4.2.2.	Degalactosylation.....	42
4.3.	Synthesis of the key-sequence (246-281aa) .....	45
4.3.1.	Manual synthesis with TentaGel resin (first approach).....	45
4.3.2.	Manual synthesis with TentaGel resin (second approach).....	46
4.3.3.	Automated Liberty Blue synthesis with TentaGel resin.....	48
4.3.4.	Automated Liberty Blue synthesis with ChemMatrix resin .....	50
4.3.5.	Automated Liberty Blue synthesis with Wang resin .....	50
4.4.	Preliminary solubility/ precipitation assays .....	51
4.4.1.	PEGylation of the key-sequence .....	52
4.5.	Synthesis of the remaining sequences (141-201aa; 202-245aa) .....	55
4.5.1.	Synthesis of the sequence 141-201aa .....	55
4.5.2.	Synthesis of the sequence 202-245aa .....	59
4.6.	Ligations of the N-terminal and middle sequences .....	60
4.6.1.	Hydrazine cleavage of sequence 141-201aa.....	60



4.6.2.	Thioester preparation of sequence 141-201aa .....	62
4.6.3.	Hydrazine cleavage of sequence 202-245aa.....	63
4.6.4.	Ligation assays for 141-245aa-Hydrazine .....	64
4.7.	Endo-A.....	65
4.7.1.	Endo-A test expression .....	66
4.7.2.	Endo- A activity assay.....	68
5.	CONCLUSION & OUTLOOK .....	70
6.	ABBREVIATIONS .....	72
7.	LITERATURE .....	75

# ABSTRACT

The interaction between Fas receptor and Fas ligand represents one of the most important pathway for programmed cell death and is very well investigated. It has been widely accepted that the metalloproteinases cleavage of Fas ligand generates the soluble form of FasL (sFasL). However, the exact mechanism of sFasL has still not been fully understood and its biological function remains unclear. The *N*-glycosylation of sFasL has been proven to be essential for the efficient secretion, but little is known about their role in the interaction with Fas receptor. There are still many unanswered questions and controversy about sFasL, despite the fact that there are numerous disease-related overexpressions of sFasL and therefore tremendous amount of research has already been done in the past. Never before, has been homogenously *N*-glycosylated sFas ligand obtained and only very limited knowledge is available about its properties such as stability or solubility.

To shed light on this issue, this master thesis is dedicated to the synthesis attempt of the homogenously *N*-glycosylated sFasL analogs. Fundamental progress has been made in all three targeted areas of interest, regarding the first steps toward the isolation and transfer of a homogeneous core *N*-glycan to peptides, the optimized synthesis conditions for all three peptide segments via Fmoc based solide phase peptide synthesis and also the *N*- and *C*-terminal peptide modifications with PEG<sub>27</sub>, hydrazide and thioesters. The following data demonstrate that the successful synthesis method of homogeneously *N*-glycosylated sFasL via auxiliary is within one's reach.

# ABSTRACT

Die Interaktion zwischen Fas Rezeptor und Fas Ligand repräsentiert einen der wichtigsten und gründlich erforschten Signalwege, der für den programmierten Zell Tod verantwortlich ist. Es ist allgemein bekannt, dass sFasL (soluble form of FasL) durch die Spaltung des Fas Liganden von Metalloproteasen entstehen. Sowohl der genaue Mechanismus von sFasL als auch seine biologischen Funktionen sind bis heute unklar und wurden noch nicht vollständig verstanden. Es ist erwiesen, dass die *N*-glycosylierte Form des FasL unerlässlich für die effiziente Sekretion ist, jedoch gibt es nur wenige Informationen über ihre Rolle bei der Interaktion mit Fas Rezeptoren. Trotz des enormen geleisteten Forschungsbeitrags in diesem Bereich und zahlreichen krankheitsbezogenen Überexpressionen des sFasL, existieren viele unbeantwortete Fragen und umstrittene Meinungen bezüglich dessen Rolle. Homogen *N*-glycosylierter sFas Ligand wurde nie zuvor synthetisch hergestellt, weswegen große Ungewissheiten über dessen Eigenschaften wie Stabilität oder Löslichkeit herrschen.

Diese Masterarbeit richtet sich an den Syntheserversuch des homogen *N*-glycosylierten sFasL und seines Analogons, um mehr Klarheit in diesem umstrittenen Thema zu schaffen. Fundamentaler Vorschrift wurde in allen drei Interessensgebieten erreicht, die sich wie folgt zusammenfassen lassen: Die ersten Schritte hinsichtlich der Trennung und des Transfers eines homogenen *N*-Glycan Kerns auf einem Peptid, die Optimierung der Fmoc basierten Festphasen Peptid Synthesebedingungen des Proteins, das in drei Peptidfragmenten unterteilt ist, sowie der Vorschrift in Bezug auf *N*- und *C*-terminalen Peptid Modifikationen mit PEG<sub>27</sub>, Hydrazid- und Thioestergruppen. Die nachfolgenden Daten bestätigen, dass die erfolgreiche Synthesemethode des homogen *N*-glycosylierten sFasL mit Hilfe von Auxiliar zum Greifen nah ist.

# 1. INTRODUCTION

## 1.1. Glycoprotein synthesis

### 1.1.1. Glycosylation as posttranslational modification

Glycosylation is one of the most prevalent and ubiquitous co-translational and posttranslational modifications. This enzyme regulated site and substrate specific reaction leads to the addition of a sugar moiety to proteins or lipids, resulting in different kinds of glycoconjugates. It is well known that glycosylation plays not only an essential role in intrinsic properties, e.g. promoting the folding, quality control, solubility and intracellular trafficking of proteins, but also biological recognition processes, such as cell adhesion, signaling, host-pathogen interaction and immune response <sup>[1]</sup>. In contrast to other posttranslational modifications, which only involve one functional group transfer, protein glycosylation is much more diverse. There are more than 40 different types of sugar-amino acid linkages <sup>[2]</sup>, involving 13 different monosaccharides and 8 different amino acid residues, where each single linkage can represent a specific function, while the most abundant carbohydrate-peptide bond is the  $\beta$ -glycosidic linkage of *N*-acetyl glucosamine (GlcNAc) to Asn. It is not surprising that there are still many fundamental questions left unanswered due to the impressive high structural and function complexity of glycans, despite the enormous amount of research already done in this area <sup>[3]</sup>.

All glycans are classified in five categories, regarding the type of residue linked to the oligosaccharide: the *N*-linked glycans, the *O*-linked glycans, the phosphor-glycans, the *C*-linked glycans and glypiation (GPI-anchored). In *N*-glycoproteins, the glycan is attached covalently to the side chain of asparagine or arginine. Therein, *N*-acetylglucosamine binding to asparagine (GlcNAc-Asn) in a consensus (Asn-Xaa-Ser/Thr) sequence is the most common among the five possible linkages between sugar and amino acid. All GlcNAc-Asn *N*-glycans share a common core sugar structure consisting of two *N*-acetyl glucosamines and three mannose residues. There are three types of *N*-glycans: oligomannose, complex and hybrid, classified due to their different terminal structures.

The biosynthesis of *N*-glycans begins on the cytoplasmic face of the ER membrane with the transfer of GlcNAc-phosphate (GlcNAc-P) from uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc) to the membrane bound polyisoprenol lipid precursor dolichol phosphate (Dol-P) to form the dolichol pyrophosphate *N*-acetylglucosamine (Dol-P-P-GlyNAc). Stepwise, a second *N*-acetylglucosamine and five mannose residues are transferred, generating Man<sub>5</sub>GlcNAc<sub>2</sub>-P-P-Dol. Subsequently, the precursor is translocated by flippase to the lumen of the ER and extended by the gradual addition of four mannose residues followed by three glucose residues. The resulting 14-sugar glycan Glc<sub>3</sub>Man<sub>9</sub>GlcNAc<sub>2</sub>-P-P-Dol is bound by the multisubunit protein complex oligosaccharyltransferase, which transfers the glycan “en bloc” onto asparagine of a nascent

protein translocated across the ER membrane. The following specific sequential removal of the three glucose and one mannose residue of the  $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$  structure leaves a  $\text{Man}_8\text{GlcNAc}_2$ -glycoprotein, which is transported to the Golgi apparatus for further sugar addition and modification. During the trimming steps of glucoses and mannose, the precursor undergoes quality control pathways which assure that the proper folding is accomplished <sup>[4]</sup>.

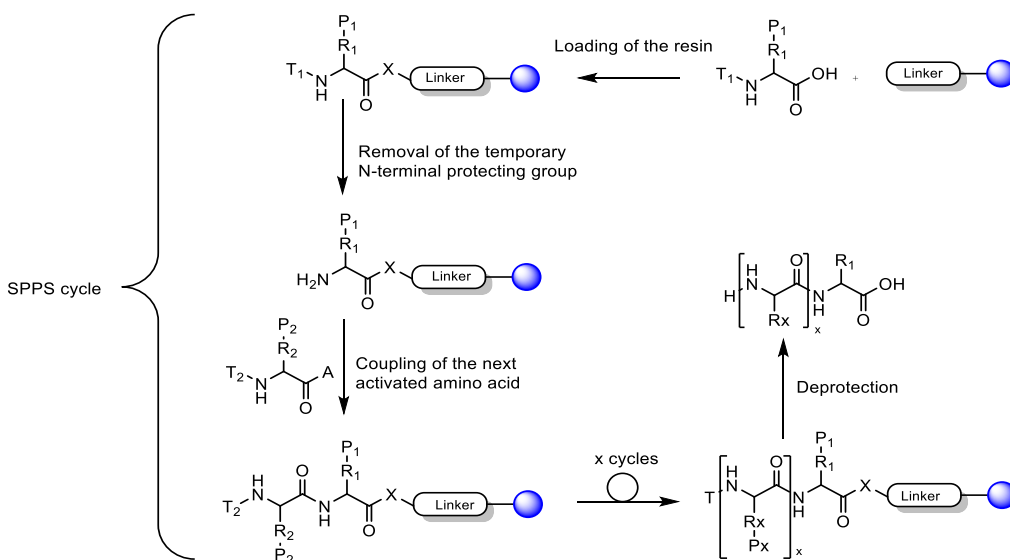
It is no surprise that *N*-glycans belong to one of the most challenging synthetic targets of natural products, due to the large size of peptide and sugar moiety, the complex branching and the heterogeneity of glycosylation. However, for the better understanding of the specific functions and structural patterns of glycopeptides, different strategies of homogeneous glycoprotein synthesis have been developed. The synthesis of the target peptide is essential, subsequently the glycosylation modification is performed either enzymatically or chemically. As for the synthesis of the peptide, two different strategies are normally used: chemical synthesis via liquid or solid phase and recombinant expression via cell based <sup>[5]</sup> or cell free systems <sup>[6]</sup>. Depending on the character of the target peptide and sugar structure, variations and combinations of strategies can be applied and investigated. In this project however, the combination of solid phase peptide synthesis (SPPS) with insertion of the first sugar and the following enzymatic coupling of the isolated *N*-glycosylated core structure are pursued.

### 1.1.2. Peptide synthesis (SPPS)

The preparation of a specific peptide sequence is commonly achieved via recombinant bacterial expression or via chemical synthesis on solid support. Especially for the synthesis of natural peptides, which are difficult to express in bacteria, for peptide/protein backbone modification and for the incorporation of unnatural amino acids within the sequence, the solid phase peptide synthesis (SPPS) method developed by Robert Bruce Merrifield is preferred <sup>[7]</sup>.

The solid phase is made of inert and insoluble carrier beads or resins which can be treated with functional units with the first amino acid attached. Unlike ribosomal protein synthesis, it proceeds from *C*-terminus to *N*-terminus. The principle of SPPS is a repeated cycle of deprotection of the *N*-terminal amine of the solid phase attached side-chain protected amino acid and its coupling with another activated side-chain protected amino acid. The *N*-terminus of the amino acid is either protected with base-sensitive 9-fluorenylmethyloxycarbonyl (Fmoc) or with acid-sensitive tert-butoxycarbonyl (Boc), presenting two different synthesis strategies. At the end of the synthesis, the immobilized and fully protected peptide chain will be cleaved from the solid support (Figure 1). Depending on which one of the two main strategies is used, different cleavage cocktails are applied. During the SPPS, Boc group is deprotected by TFA. The side chain protecting groups and peptide-resin linkages are removed by the much stronger anhydrous hydrofluoric acid (HF). Although this method allows efficient syntheses of many peptides, the necessary application of highly dangerous HF and the need for special polytetrafluoroethylene-lined apparatus have to be put into consideration. In comparison, Fmoc protecting group can be

easily removed by 20-50 % piperidine and TFA is used for the total deprotection from the resin, giving an orthogonal synthesis method. Since no acid is used for the Fmoc removal, the peptide-resin does not need to be neutralized, but the lack of electrostatic repulsions between the peptides can lead to increased aggregation. Nevertheless, the overall milder cleavage condition and the easier handling have established Fmoc strategy as the method of choice for the routine synthesis of peptides.



**FIGURE 1 THE PRINCIPLE OF SPPS**

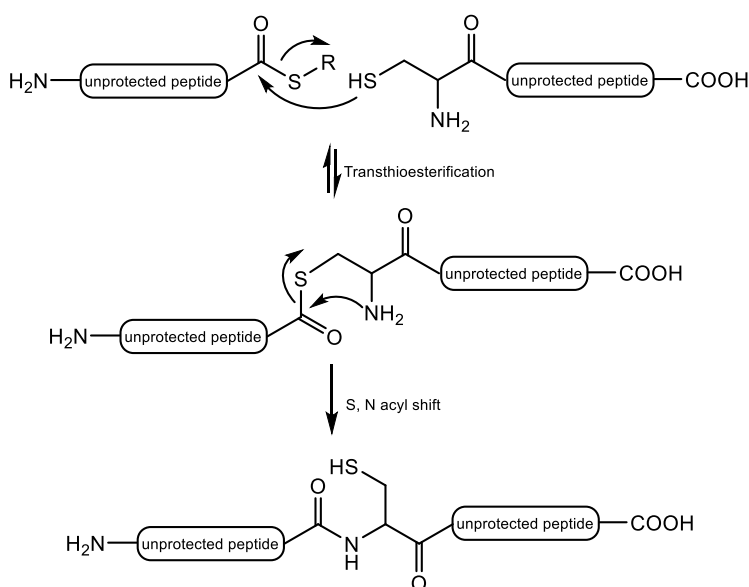
The advantage of SPPS lies in the achieved specificity of sequence elongation by using protecting groups on the side chain and specific deprotection of the *N*-terminus. Also, wash cycles can be performed after each reaction. The peptide remains on the insoluble solid support, which allows the removal of excess reagent and therefore prevents side reactions. Hence, extremely high yield can be generated in each step. Depending on the physical character of the peptide chain, peptides and proteins in the range of 70 amino acids are still successfully achievable [8]. However, ligation methods have to be used to couple two or more peptide fragments with each other, if the target molecule is a large peptide or even a complete protein [9].

### 1.1.3. Protein synthesis (NCL)

After obtaining peptides with either expression in bacteria or SPPS, these fragments can be further conjugated with each other by chemoselective reactions. With several successive ligation steps, proteins up to 312 amino acids can be produced [10]. The principle lies in the two unique reactive terminal functional groups, one on each of the reacting peptide fragment. These two are

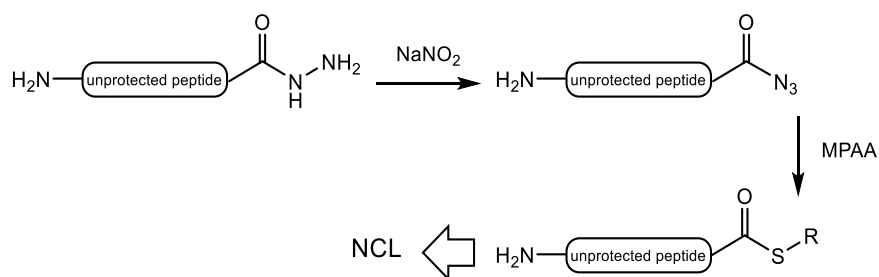
designed to only react with each other and not with any of the other functional groups e.g. the unprotected side chains of the peptides.

The most robust and successful ligation method is the native chemical ligation (NCL) developed by Kent <sup>[11]</sup>. It is based on the selective and reversible thioesterification of one C-terminal thioester peptide with another N-terminal cysteine peptide under mild conditions. After the *S,N*-acyl-shift of the intermediate, a peptide bond is formed (Figure 2).



**FIGURE 2 REACTION SCHEME OF NATIVE CHEMICAL LIGATION**

This technique is not only limited to chemically synthesized peptides. The expressed protein ligation established (EPL) by Muir <sup>[12]</sup> offers the elegant combination of both synthetic and recombinant peptides resulting in the semisynthesis approach of various proteins <sup>[13]</sup>. Despite numerous advantages of this technique, certain limitations still exist. One of the challenges lies in the preparation of the necessary unprotected peptide-thioester fragment, which can be either prepared with Boc chemistry or linker mediated Fmoc chemistry. Also, since cysteine is the least common amino acid found in proteins, the possibility of application is narrowed down enormously. Already shortly after the native chemical ligation, Kent et al. <sup>[14]</sup> has developed the  $N^\alpha$ -oxyethanethiol auxiliary which generates glycine at the ligation site. With selective desulfurization techniques shown by Dawson et al., alanine can be formed instead of cysteine<sup>[15]</sup>. Tremendous efforts <sup>[16]</sup> have been made since then to extend the native chemical ligation method with chemical modifications of the cysteine residue. It has been already successfully proven that for example His <sup>[17]</sup> and Met <sup>[18]</sup> are possible substitutes. In addition, the application of diverse cleavable auxiliaries has been investigated. Later in this work, one novel multifunctional auxiliary will be presented.



**FIGURE 3 REACTION SCHEME OF HYDRAZIDE NATIVE CHEMICAL LIGATION**

One of the extensions of NCL is the well-known hydrazide ligation method <sup>[19]</sup>, which has been further optimized by Liu et al (Figure 3) <sup>[20]</sup>. This strategy involves the chemoselective reaction between a C-terminal hydrazide peptide with another N-terminal cysteine peptide. Under acidic condition, the hydrazide functionality is rapidly oxidized by NaNO<sub>2</sub> to an azide and converted to a thioester after adding MPAA or mesNa under neutral condition. Thereafter, the native chemical ligation between the C-terminal thioester peptide and the N-terminal cysteine peptide takes place. The advantage of this method in comparison to NCL is that peptide hydrazides can be prepared through either Boc- or Fmoc- based SPPS and are chemically more stable than peptide thioesters, hence easier to work with. However, the modification of the solid phase (Wang resin) with 4-nitrophenyl carbonochloridate before SPPS would be still necessary. To simplify this method even further, direct hydrazinolysis of the peptide loaded TentaGel resin has been performed and the hydrazide peptides has been obtained with success <sup>[21]</sup>. Detailed stereochemical integrity of the resulting peptide hydrazide and the influence of the C-terminal amino acid have been investigated. Although this technique has only been applied to middle length peptide sequences, it has the potential to be universally applicable.

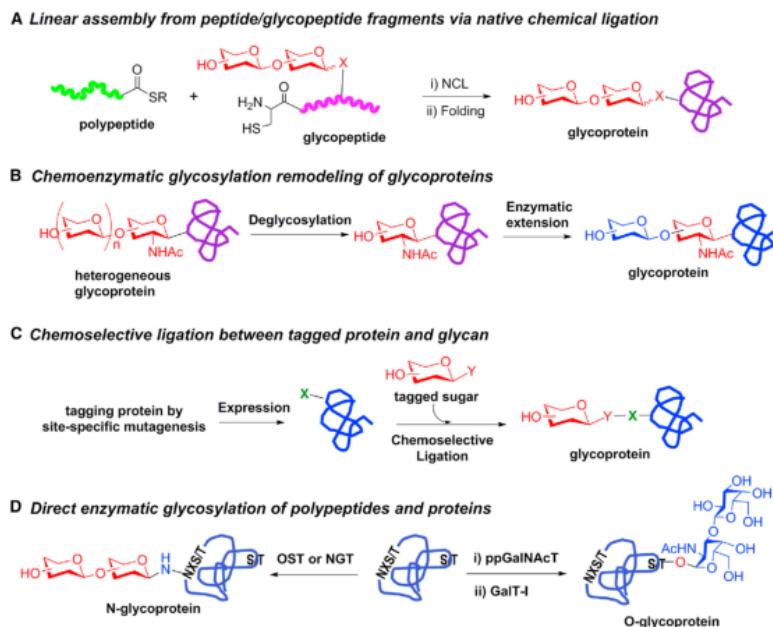
#### 1.1.4. Protein glycosylation

Despite the difficulties mentioned before, several important synthetic glycoprotein strategies have been explored. The currently most successful methods which overall allow the controlled synthesis of full size homogeneous glycoproteins are: native chemical ligation of glycosylated fragments, chemoenzymatic transformation, “tag and modify” method and direct enzymatic glycosylation (Figure 4).

Not long after the discovery of native chemical ligation, the first NCL based total chemical linear assembly of an O-glycoprotein has been performed by Bertozzi et al. <sup>[22]</sup> (Figure 4A). The harsh acidic conditions have been used for the Boc- protected solid phase synthesis would damage the glycosidic linkages. Also, thioester cleavage occurs under the basic condition used for Fmoc deprotection. Therefore, the acid- and base-stable “safety-catch” linker has been used, which is activated and cleaved off after the complete Fmoc-protected glycopeptide synthesis. The GalNAc residues have been incorporated during the solid phase synthesis with glycosylated amino acids



and the two *O*-glycosylated peptide fragments have been ligated via NCL, forming the antibacterial 82mer glycoprotein diptericin. Despite the successful synthesis of the large target molecule, the method by itself still contains several limitations, such as the length of the glycosylation and the dependency of cysteine at the ligation position.



**FIGURE 4 FOUR MAJOR STRATEGIES FOR THE SYNTHESIS OF HOMOGENEOUS GLYCOPROTEINS (FIGURE TAKEN FROM [1])**

Glycosylation remodeling of heterogeneous glycoproteins is particularly useful for *N*-glycoproteins and it involves the specific enzymatic trimming to remove the heterogeneous part and the following sugar chain elongation by glycosyltransferases or endoglycosidases (Figure 4B). The discovery of synthetic sugar oxazoline as donor substrates and the transglycosylation capacity of modified or not modified glycosidases such as Endo-A, Endo-M, Endo-D etc. enables this successful and elegant method. In combination with NCL, it is possible to synthesize full size glycosylated proteins [23].

The “tag and modify” strategy has been applied to recombinant proteins by introducing specific tags at chosen glycosylation sites with mutagenesis. After expression, the tagged protein is chemoselectively conjugated with a functionalized glycan at the tagged site. The drawback of this approach is the introduction of unnatural sugar-amino acid linkages. The direct enzymatic glycosylation of proteins in vitro with oligosaccharyl transferases (OST), *N*-glycosyltransferase (NGT) or other sugar transferases would be highly valuable and it has been shown in several cases (Figure 4D) [24]. However, at the moment the practical application remains unfulfilled caused by the instability of the enzymes themselves and the inability of the enzymes to glycosylate folded proteins [25].

Enormous efforts in the development of sophisticated chemical and biochemical ligation methods have made it possible to synthesize full-length natural glycoproteins, despite the large

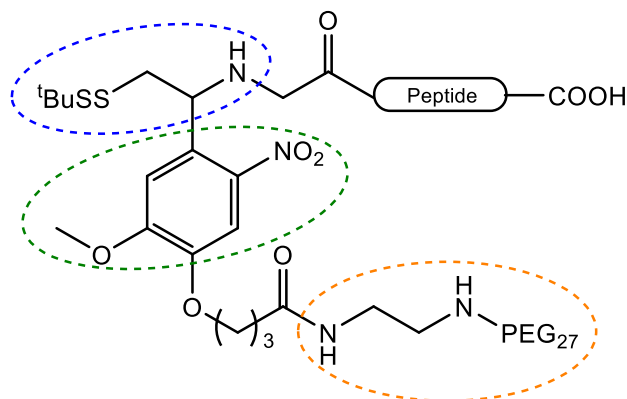
size and enormously complex glycan structures. However, one should consider that the synthesis of each specific target glycoprotein concludes its own special challenges and the careful choice of glycosylation strategy and design is essential. There is still a great need of universally applicable synthetic technologies for forming homogeneous glycoproteins.

#### 1.1.5. Multifunctional PEGylated photocleavable auxiliary

In this work, the combination of two methods mentioned in the previous chapter will be applied for the synthesis of the target molecule sFas ligand. While the native chemical ligation is necessary for the extension of the glycosylated peptide fragment, the stepwise or “en bloc” enzymatic glycosylation of single sugar building block or pre-trimmed sugar oxazoline provides the selective and yet convenient installation of the sugar chain. Several linkers and auxiliary based strategies <sup>[26]</sup> have been reported to mimic the NCL reaction for overcoming the limitation of Fmoc- based SPPS synthesis and the reliance on Cys residues for ligation. At the same time, different attempts have been made to optimize the yield and purity of the enzymatic glycosylation reaction <sup>[27]</sup>. Possible incomplete reaction leads to mixtures of glycopeptides and the purification of the desired target has been a complicated and time consuming task.

A recent development in overcoming the limit of both fields is represented by the multifunctional PEGylated photocleavable auxiliary developed by Bello and Becker et al. <sup>[28]</sup>, which allows the straightforward preparation and isolation of homogeneously glycosylated peptide. Ligation mediated by photocleavable auxiliaries has been previously investigated by Aimoto <sup>[29]</sup> and Dawson <sup>[30]</sup>. The mercaptoethylene amino group of the auxiliary mediates NCL, while the abundant occurring glycine is generated after photolytic removal under mild conditions. The incorporation of a monodisperse PEG polymer in the auxiliary scaffold enables quantitative sequential enzymatic glycosylation and high yielding purification through precipitation, thanks to the solubility of PEG polymer in a wide range of organic solvents and the strong tendency to crystallize. <sup>[31]</sup> The application of soluble polymer supports for the separation of the products from byproducts has been widely used in “liquid phase synthesis” <sup>[32]</sup> for the purification of organic compounds and peptides.

The coupling of the auxiliary to the peptide has to be carried out stepwise while the peptide is still bound to the solid support. First, the iodoacetylation of the deprotected *N*-terminus of the peptide takes place. In the presence of DIEA, the PEG free auxiliary reacts with the iodoacetyl peptide with very high conversion. After this  $S_N2$  reaction, the Fmoc of the primary amine on the auxiliary is deprotected and the subsequent PEGylation is performed quantitatively.



**FIGURE 5 STRUCTURE OF THE MULTIFUNCTIONAL PEGYLATED PHOTOCLEAVABLE AUXILIARY; BLUE CIRCLE: MEDIATING NCL AND GENERATING GLYCINE AT LIGATION POSITION; GREEN CIRCLE: RESPONSIBLE FOR PHOTOLYTIC REMOVAL AFTER LIGATION; ORANGE CIRCLE: PEG RESIDUE, ENHANCING OVERALL SOLUBILITY AND ENABLING HIGH YIELDING PURIFICATION OF GLYCOSYLATED PEPTIDE THROUGH PRECIPITATION**

With this method, the *O*-glycosylated Mucin 1 (MUC1) polypeptide has been successfully synthesized [28]. It has been shown that the subsequent NCL reaction of two glycosylated MUC1 tandem repeats reaches conversion over 65 %. The auxiliary attached MUC1 has been obtained in 34 % yield and after the ligation, the photocleavage can be performed without any complications. The enzymatic glycosylation has been efficiently supported by the auxiliary and there is no need for chromatographic purification. The only limitation of this technique so far is the identification of a suitable glycine residue. Therefore, the full potential still needs to be exploited.

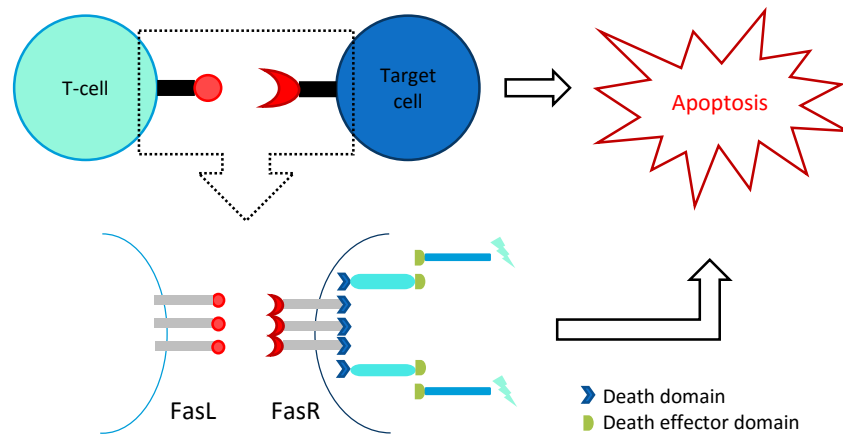
In summary it can be said that this novel PEGylated photocleavable auxiliary combines different necessary features for NCL and enzymatic glycosylation. Although this method has been only used for *O*-glycosylation so far, there is no apparent limitation of its applicability. Therefore, this auxiliary is going to be the key element for the synthesis of the target molecule *N*-glycosylated sFas ligand.

## 1.2. sFasL

### 1.2.1. The FasR/FasL death initiating pathway

The complex process of programmed cell death, apoptosis, is regulated *in vivo* by different pathways [33]. One of them is the direct signal transduction, triggered by the interaction of cell surface protein Fas receptor (FasR), also known as apoptosis antigen 1 (APO-1 or APT), cluster of differentiation 95 (CD95) or tumor necrosis factor receptor superfamily member 6 (TNFRSF6), with its natural ligand FasL or agonistic antibodies [34]. While FasR is expressed ubiquitously with the highest level found in the heart, lungs, liver etc., the expression of FasL is mainly associated

with activated T lymphocytes, natural killer cells <sup>[35]</sup> and in tissues of “immune-privilege sites” such as the testis, eye, lung and placenta <sup>[36]</sup>. However, cell expression of FasL can be induced by cellular stress.



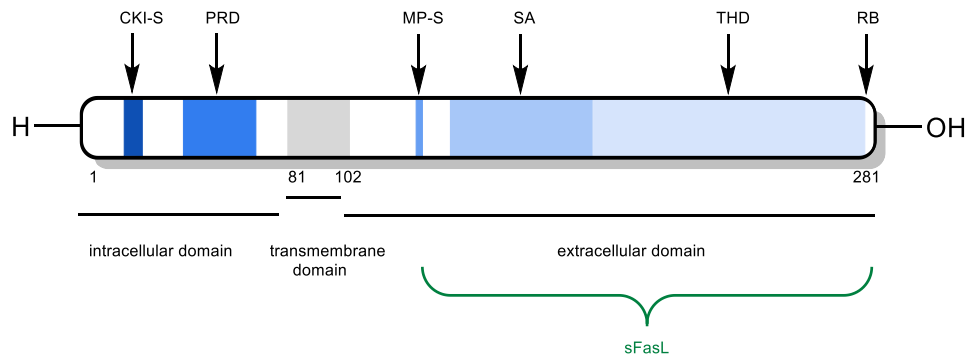
**FIGURE 6 THE FASR/FASL DEATH INITIATING PATHWAY BETWEEN A T-CELL AND THE TARGET CELL; BINDING OF TRIMER FASL WITH TRIMER FASR LEADS TO DISC FORMATION, HENCE CONSEQUENTLY TO APOPTOSIS OF THE TARGET CELL**

The activation of the FasR/FasL death initiating pathway starts with the spontaneous trimer-formation of the membrane-anchored FasL on the surface of an adjacent cell, which causes the trimerization and binding of the FasR on the target cell. This cell-to-cell contact leads to the rearrangement of the death-inducing signaling complex (DISC), which consequently generates the death signal. The protein-protein interaction domain on the cytoplasmic site of FasR known as the death domain (DD) interacts with the DD domain of the adaptor molecule Fas-associated DD containing protein (FADD) through homotypic interaction <sup>[37]</sup>. The death effector domain (DED) on the *N*-terminus of the FADD is crucial for the recruitment of caspases containing the same DED domain, such as caspase 8 or caspase 10 (Figure 6) <sup>[38]</sup>. Caspases are cysteine-dependent aspartate-directed proteases which either work as initiator by activating other caspases or as executor by triggering the apoptosis through protein cleavages <sup>[39]</sup>. The FasR/FasL pathway leads to programmed cell death and is one of two apoptosis pathways, the other being the mitochondrial pathway. Due to their significant function, the involvement within diseases especially cancer has been investigated in detail. The idea of selective control of this pathway, and thus of the cell apoptosis, motivates countless research in this particular field <sup>[40]</sup>.

### 1.2.2. The structure of FasL

The 281 amino acids long membrane-bound human FasL is a type II transmembrane molecule, which contains three different regions: the intracellular domain (1-81aa), the transmembrane domain (81-102aa) and the extracellular domain (102-281aa). The intracellular domain harbors a specific double casein kinase substrate motif (CKI-S) and a proline-rich domain (PRD), which is

responsible for trafficking FasL to secretory lysosomes upon interactions with e.g. SH3 domain proteins [41]. Several metalloproteases substrate sites (MP-S), the self-assembly site (SA) which is necessary for the trimerization, the glycosylated tumor necrosis factor homology domains (THD) and the important C-terminal receptor-binding site (RB) are located in the extracellular domain, causing it to be not only the biggest, but also the most interesting part of the Fas ligand. The soluble FasL, released from the cell surface by metalloproteinases, contains the extracellular domain from the MP-S to RB, therefore allowing possible trimerization and FasR binding. [42]



**FIGURE 7 MAIN STRUCTURE FEATURES OF THE HUMAN FAS LIGAND**

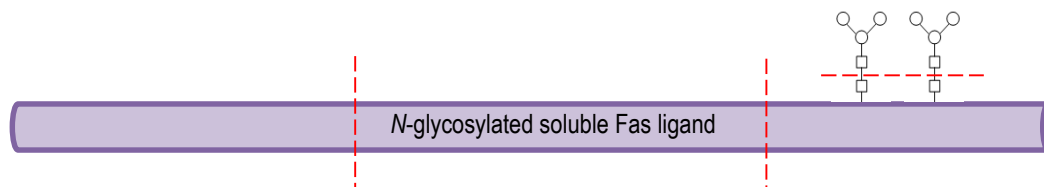
### 1.2.3. Soluble Fas Ligand (sFasL)

In comparison to the membrane-bound FasL, there is little known about the soluble Fas ligand, despite its over-expression correlating to numerous different kinds of diseases, e.g. rheumatic diseases [43], diverse types of cancer [44], malaria [45] and heart diseases [46]. It has been widely accepted that the cleavage of metalloproteinases generates the soluble form of FasL. However, the exact mechanism has still not been fully understood [47]. Its biological function remains unclear, since both apoptosis activating and deactivating responses have been observed in different types of cells [44, 47-48]. It has been discovered that different enzymatic cleavage sites in FasL are present which would generate multiple possible forms of sFasL with individual amino terminal sequences, which possess varying biological activities. This presents one possible explanation for the controversy in the literature. Variations of cell types and enzymes forms disparate sFasL types, which in general have been identified as the same sFasL and compared with each other in regard of biological functionality. Overall it can be said that sFasL definitely plays an essential role in the regulation of apoptosis by influencing the FasL/FasR interaction depending on specific cell type and enzymatic cleavage site. However, the physical role of sFasL remains unclear due to the complex cleavage site - function relation [48a].

The three *N*-linked glycans on the tumor necrosis factor homology domain of sFasL and FasL have been found to be essential for efficient secretion. However, their role in the interaction with Fas

receptor has not been further investigated<sup>[49]</sup>. There are some approaches for the recombinant production of glycosylated and non-glycosylated sFasL, but none of them has been chemically synthesized yet. Especially the homogeneously glycosylated sFasL would be of great interest for further investigating the role/effect of the glycans on the stability, solubility and biological function of the protein. The better understanding of these properties could open the door to eventual therapeutic applications.

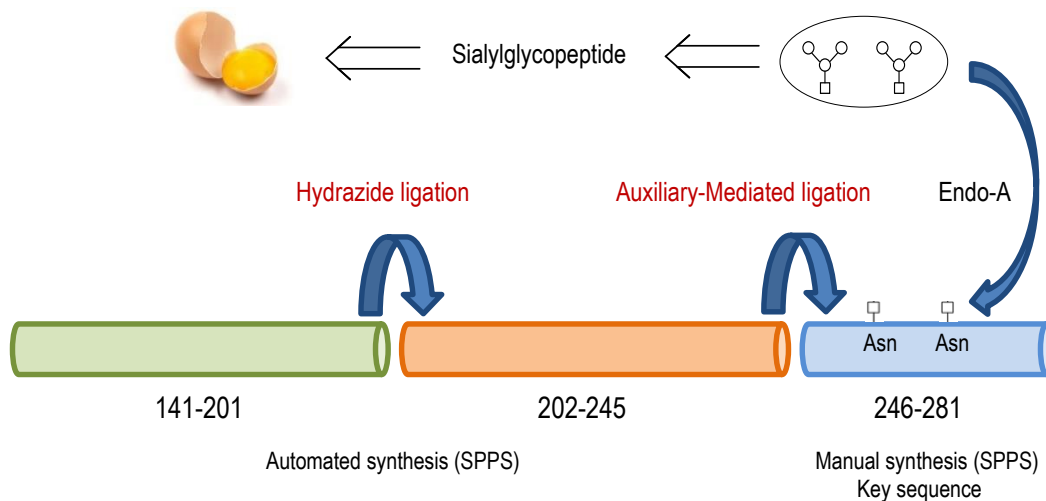
## 2. AIM



**FIGURE 8 STRATEGY FOR ACCESSING HOMOGENOUS *N*-GLYCOSYLATED sFAS LIGAND**

Despite the enormous effort invested in studying the disease related over-expression of sFas ligand, its exact biological function still remains a mystery. The *N*-glycosylations in the tumor necrosis factor homology domains (THD) have been proven to be essential for efficient secretion, however little is known about their role in the interaction with Fas receptor. Never before, has homogeneously *N*-glycosylated sFas ligand been obtained and therefore only very limited knowledge is available about its properties such as stability or solubility. To shed light on this issue, this master thesis aims at the synthesis of homogeneously *N*-glycosylated sFasL analogs for the better understanding of sFas ligand properties and its role in the FasR/FasL death initiating pathway.

In native sFasL, there are three biological *N*-glycosylation sites at positions Asn-184, Asn-250 and Asn-260. In this project however, the focus lies on the two sites closer to the C-terminus due to simplification of the synthesis approach. The strategy is to synthesis the sFasL ligand protein via Fmoc-based solid phase peptide synthesis (SPPS) in three fragments (Figure 8). The *N*-terminal fragment will be linked to the middle fragment via native chemical ligation using a C-terminal hydrazide as thioester precursor. The multifunctional PEGylated auxiliary approach will be applied to the C-terminal fragment (key sequence), which can contain two, one or no asparagine-GlcNAc building block, whereupon the core structures of *N*-glycosylation can be transferred enzymatically with endoglycosidase and a sugar oxazoline isolated from egg yolk. PEGylating of the auxiliary (Figure 5) should enhance the solubility of the key sequence and support the enzymatic glycosylation via *Endo*- $\beta$ -*N*-Acetylglucosaminidase from *Arthrobacter protophormiae* (*Endo*-A). The Man<sub>3</sub>GlcNAc oxazoline can be obtained through enzymatic digestion of sialylglycopeptide, isolated from egg yolk. After glycosylation, the PEGylated residue should allow purification via precipitation. To finalize the synthesis of the target molecule, the *N*-glycosylated key sequence will be ligated via auxiliary-mediated NCL to the protein segment comprising the other two fragments. Subsequently, glycine would be generated at its natural position after removal of the auxiliary from the ligation product by radiation (Figure 5).



**FIGURE 9 DETAILED SYNTHESIS STRATEGY FOR HOMOGENOUS *N*-GLYCOSYLATED sFASL: THREE PEPTIDE SEGMENTS ARE COUPLED VIA MODIFIED VERSION OF NATIVE CHEMICAL LIGATION, GIVE THE COMPLETE PROTEIN SEQUENCE OF sFASL; THE SIALYLGLYCOPEPTIDE ISOLATED FROM EGG YOLK IS ENZYMATICALLY TRIMMED AND CHEMOENZYMATICALLY ATTACHED TO THE PEPTIDE KEY SEQUENCE**

As for the protein part, the intermediate objectives would be the synthesis optimization with respect to total yield and purity, the characterization of the peptides and the investigation of further modification possibilities, such as PEGylation, hydrazinolysis and ligation. Special focus is placed on the key sequence due to the high purity requirement for the quantitative auxiliary conjugation reaction on resin. The aim is to achieve the homogeneous *N*-glycosylation, the isolation and purification of SGP, the testing of the enzymatic trimming steps with several enzymes. Possible process optimization needs to be investigated in detail.



# 3. EXPERIMENTAL PART

## 3.1. Material and methods

### Peptide synthesis and modification:

9-Fluorenylmethoxycarbonyl (Fmoc)-protected amino acids, 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), DIC and Oxyma were purchased from Merck (Darmstadt, Germany), Iris Biotech (Marktredwitz, Germany), ORPEGEN Peptide Chemicals (Heidelberg, Germany) and Sigma-Aldrich (Taufkirchen, Germany). TentaGel™-S-PHB-Lys(Boc)Fmoc and TentaGel™-S-PHB-Ser(t-Bu)Fmoc were purchased from Rapp polymere (Tübingen, Germany). H-Lys(Boc)-HMPB-ChemMatrix® resin was purchased from Sigma-Aldrich, Fmoc-Lys(Boc)-Wang resin was purchased from Iris Biotech. N,N-Dimethylformamide (DMF), methanol, dichloromethane (DCM), acetonitrile (ACN) and trifluoroacetic acid (TFA) were obtained from Biosolve (Valkenswaard, The Netherlands). Fmoc-PEG<sub>27</sub>-COOH was from Polypure (Oslo, Norway).

### SGP isolation and trimming:

Fresh egg yolks were obtained from hen's eggs purchased from local market. Celite and carbon were purchased from Sigma Aldrich. Diethylether (Et<sub>2</sub>O), aceton, acetonitrile (ACN) and trifluoroacetic acid (TFA) were obtained from Biosolve (Valkenswaard, The Netherlands) and Sigma-Aldrich. Neuraminidase and β-1,4-galactosidase were purchased from New England Biolab (Frankfurt am Main, Germany) and Sigma-Aldrich. The unit definitions of enzyme activities were following the description from manufacturers. Fetuin from fetal bovine serum was obtained from Sigma-Aldrich. Acetic acid was purchased from Merck. P-nitrophenyl-α-D-galactosidase and albumin from bovine serum were obtained from Carl Roth (Karlsruhe, Germany). Stain-free Mini-PROTEAN® TGX™ Precast gradient Gels were purchased from Bio-Rad (Munich, Germany)

All other chemicals were obtained from Sigma-Aldrich at the highest purity available and used without any further purification.

## 3.2. Reaction monitoring, product analysis and purification by RP-HPLC/MS

Waters AutoPurification HPLC/MS system (3100 Mass Detector, 2545 Binary Gradient Module, 2767 Sample Manager and 2489 UV/Visible Detector) was used for analytical liquid chromatography-mass spectrometry (LC-MS). Mass spectra were acquired by electrospray

ionization (ESI-MS) operating in positive ion mode. If not stated otherwise, separation was achieved with a Kromasil 300-5-C<sub>4</sub> or 300-5-C<sub>18</sub> column (50 × 4.6 mm, 5 μm particle size) at a flow rate of 1 mL/min running a linear gradient from 5% to 65% of buffer B (ACN + 0.05 % TFA) in buffer A (ddH<sub>2</sub>O + 0.05 % TFA) in 10 min. Crude hydrazide peptides were purified with the same system on Semiprep Kromasil 300-10-C<sub>4</sub> (250 × 10 mm, 5 μm particle size) at a flow rate of 10mL/min with gradient from 5-20% of buffer B in 5 min and 20-50 % of buffer B in 30 min.

For analytical RP-HPLC measurements and purification, Dionex Ultimate 3000 instrument at a flow rate of 1 mL/min with a linear gradient from 5 to 65 % buffer B in buffer A over 30 min (buffer A: 0.1 % (v/v) TFA in ddH<sub>2</sub>O, buffer B: 0.08 % (v/v) TFA in ACN) has been used. Crude peptides and peptide thioesters were analyzed and purified on a Kromasil 300-5-C<sub>4</sub> column (150 × 4.6 mm, 5 μm particle size). SGP and SGP trimming has been analyzed and purified on a Kromasil 300-5-C<sub>18</sub> column at flow rate of 1 mL/min with a gradient (150 × 4.6 mm, 5 μm particle size) with a linear gradient from 2-5 % buffer B in 10 min and 5-90 % buffer B in 25 min. For biological assays, NanoDrop UV-Vis2000c Spectrophotometer from Thermo Scientific has been used.

For SGP and other detailed HPLC-MS mass measurements, Dionex Ultimate 3000 instrument RSLC-nano system coupled with Thermo Scientific LTQ Orbitrap VELOS via nano-Viper C4 column (75 μm × 15 cm, 150 Å) or C18 column (75 μm × 50 cm, 100 Å) with a gradient of 2 % B constant for 5 min, 2-7 % B for 1 min, 7-35 % B for 24 mins, 35-40 % B for 1 min, 40 % B constant for 1 min, 40-80 % B for 1 min and 80 % B constant for 2 mins (buffer A: 98 % ddH<sub>2</sub>O, 2 % ACN, 0.1 % formic acid, buffer B: ACN/ddH<sub>2</sub>O 80/20, 0.1 % formic acid) has been used. Precolumn nano-Viper C18 (100 μm × 2 cm, 100 Å) has been coupled when necessary. The samples have been prepared in buffer A or B before injection.

### 3.3. Isolation and purification of SGP from egg yolk

Fresh 30 egg yolks were separated from egg white carefully and stirred with 150 mL ddH<sub>2</sub>O at room temperature for 1 h. The residue was lyophilized until complete dryness, giving the yellow egg yolk powder 200-250 g.

The yolk powder was stirred vigorously at room temperature with 300-400 mL diethyl ether for 1 hour. The organic solvent was removed by filtration and the washing step was repeated twice. The residue was then mixed at room temperature with 300 mL 70 % acetone for 1 hour, the solvent was removed by filtration and the washing step was repeated once. Finally, the remaining filter cake was stirred with 300 mL 40 % acetone for 1 hour and the process was repeated at least three times. The filtrate was filtrated with Celite and concentrated on the rotary evaporator to remove the acetone. The remaining aqueous solution was lyophilized until complete dryness, giving the light yellow crude SGP product 2-3.5 g, which was analyzed by RP-HPLC and MS.

1.8 g of the egg yolk extract powder was dissolved in 5 mL water and applied to a packed active carbon/Celite (2:1) column (2 x 17 cm). The column was washed by water containing 0.1 % TFA (100 mL), 5 % acetonitrile and 0.1 % TFA (100 mL), 10 % acetonitrile and 0.1 % TFA (100 mL). Finally, the desired product was eluted with 25 % acetonitrile and 0.1 % TFA (200 mL). The elution fractions were lyophilized giving 260 mg yield, which was analyzed by RP-HPLC and MS.

SGP ( $C_{112}H_{189}N_{15}O_{70}$ ) ESI-MS: observed mass: 2865 Da, calculated mass: 2865 Da;  $[C_{112}H_{192}N_{15}O_{70}]^{3+}$  observed: 956.06, calculated: 956.07;  $[C_{112}H_{191}N_{15}O_{70}]^{2+}$  observed: 1433.60, calculated: 1433.60

## 3.4. Synthesis of $Man_3GlcNAc_2$ -peptide

### 3.4.1. Desialylation via neuraminidase

50 mg SGP (0.017 mmol) in a phosphate buffer (50 mM, pH 6.0, 350  $\mu$ L) was incubated with neuraminidase (150 U) at 32°C for 22 hours on the shaker and controlled over time by RP-HPLC and MS. No reaction progress was observed. Therefore, another 300 U neuraminidase were added to the reaction mixture and shaken for altogether up to 48 hours. Still, no reaction at all could be observed.

The reaction was repeated with 5 mg of SGP (1.75  $\mu$ mol) in glycobuffer (from New England Biolab), which consisted of 5 mM  $CaCl_2$  and 50 mM sodium acetate at pH 5.5, with neuraminidase (15 U) at 32°C for up to 24 hours on the shaker. The reaction was monitored by RP-HPLC and MS. Since no reaction was observed after 3 hours, an additional amount of neuraminidase (15 U) was added. The reaction was controlled over 24 hours and no desialylation reaction was detected.

### 3.4.2. Neuraminidase activity assay

#### Approach 1:

In five 1.5 mL Eppendorf, 2  $\mu$ L fetuin was combined with glycoprotein denaturing buffer and 5  $\mu$ L of  $H_2O$  each. The reaction mixture was incubated at 95°C for 10 min. Thereafter, 1  $\mu$ L of Glycobuffer, 1  $\mu$ L of 10 % NP-40 and 1  $\mu$ L of neuraminidase were added subsequently to each tube and softly mixed. Under shaking, two Eppendorf tubes were incubated at 37°C for 1 hour, one for 2 hours and one for 3 hours. All samples and the control (fetuin without enzyme) were visualized on 15 % SDS-PAGE gel (see below).

### Approach 2:

In four 1.5 mL Eppendorf, 2  $\mu$ L fetuin was combined with glycoprotein denaturing buffer and 5  $\mu$ L of H<sub>2</sub>O each. The reaction mixture was incubated at 95°C for 10 min. Thereafter, 1  $\mu$ L of glycobuffer, 1  $\mu$ L of 10 % NP-40 and 1  $\mu$ L of neuraminidase were added subsequently to each tube and mixed gently. Under shaking, one Eppendorf tube was incubated at 37°C for 1 hour, one for 2 hours, one for 3 hours and one for 24 hours.

Four different positive controls were prepared as below (Table 1).

<b>Fetuin</b>	<b>Acid</b>	<b>Reaction time</b>
<b>1mg</b>	2 M acetic acid	2 hours
<b>1mg</b>	2 M acetic acid	3 hours
<b>1mg</b>	2.5 M Formic acid	2 hours
<b>1mg</b>	2.5 M Formic acid	3 hours

**TABLE 1 POSITIVE CONTROLS OF THE FETUIN ACTIVITY ASSAY**

Finally, the samples and the control (fetuin without enzyme) were visualized on 8-16 % SDS-PAGE Mini-protean precast stain-free gel. The bands of the samples look all very similar on the gel and no clear desialylation could be observed, as can be seen in section results and discussion (Figure 17, page 41). However, the bands of the positive control could not be observed on the SDS-PAGE gel.

### General procedure for SDS gels preparation (for 9 gels):

The casting frames were set on the casting stands. The separating gel was mixed thoroughly as described below in a separate beaker. The appropriate amount was pipetted into the gap between the glass plates. Water was filled into the gap until overflow, to ensure the smooth horizontal top formation of the separating gel. After 20-30 min, the separating gel was gelled. The additional water was discarded and the stacking gel pipetted until overflow. The well-forming comb was inserted and after 20-30 min the gel gelled. The comb was taken out after the complete gelation of the stacking gel.

Separating gels consisted of 19.2 mL H<sub>2</sub>O, 40 mL acrylamide, 20.8 mL stacking gel buffer, 837  $\mu$ L 10 % sodium dodecyl sulfate (SDS), 837  $\mu$ L 10 % Ammonium persulfate (APS), 38  $\mu$ L tetramethylethylenediamine (TEMED).

Stacking gels consisted of 29 mL H<sub>2</sub>O, 7.2 mL acrylamide, 5.6 mL separating gel buffer, 426  $\mu$ L 10 % sodium dodecyl sulfate (SDS), 426  $\mu$ L 10 % Ammonium persulfate (APS), 42  $\mu$ L tetramethylethylenediamine (TEMED).

#### General procedure for SDS buffers preparation:

Laemmli buffer (10x concentrated): 250 mM 2-amino-2-(hydroxymethyl)-1,3-propanediol hydrochloride (Tris-HCl), 2 M glycine and 1% (w/v) SDS were dissolved in H<sub>2</sub>O. For the SDS-PAGE use, 10-fold dilution of the stock solution was prepared.

Separating gel buffer (1x concentrated): 1.5 M Tris-HCl, 0.4 % (w/v) SDS were mixed and the pH adjusted to 8.8.

Stacking gel buffer (1x concentrated): 0.5 M Tris-HCl, 0.4 % (w/v) SDS were mixed and the pH adjusted to 6.8.

Sample buffer (loading buffer) (2x concentrated): 500 mM Tris, 6 % (w/v) SDS, 35 % (v/v) glycerin, 3.55 % (v/v) β-mercaptoethanol and 0.05 % (w/v) bromophenolblue were mixed and adjusted to pH 6.8.

Coomassie-Staining solution (1x concentrated): 0.1 % (w/v) Coomassie R250, 10 % (v/v) acetic acid and 45 % (v/v) methanol in water were mixed.

Destaining Solution: 10 % (v/v) acetic acid and 40 % (v/v) methanol in water were mixed.

#### General procedure for SDS-PAGE:

The gels were taken out of the casting frame and set in the cell buffer dam. Laemmli buffer (electrophoresis buffer) was poured into the inner chamber after overflow, until the buffer surface was above the required level in the outer chamber.

The samples were mixed with loading buffer and denatured for 5-10 min at 90°C. The prepared samples and loading protein marker were pipetted into the wells. Finally, the anode and cathode have been connected and the appropriate volt was set.

After the SDS-PAGE, the gels were covered with Coomassie-Staining solution for one hour under shaking and washed twice with H<sub>2</sub>O. After destaining the gels with either destaining solution for one hour or with H<sub>2</sub>O overnight under shaking, the gels were ready for further analysis.

### 3.4.3. Desialylation via acetic acid

5 mg (1.75 μmol) of SGP in 500 μL 2 M acetic acid were shaken for 2 hours at 80°C. After completion of the reaction, the mixture was lyophilized to dryness repeatedly until reaching a neutral pH 3.8 mg of desialylated SGP (95 % yield) have been obtained.

The same reaction was repeated under variation of conditions (Table 2)

SGP [mg/ $\mu$ mol]	2M acetic acid [ $\mu$ L]	Reaction time	temperature
1/0.35	100	2 hours	80°C
1/0.35	200	2 hours	80°C
1/0.35	100	3 hours	80°C
1/0.35	200	3 hours	80°C

TABLE 2 REACTION CONDITIONS FOR DESIALYLATION VIA ACETIC ACID

Desialylated SGP ( $C_{90}H_{155}N_{13}O_{54}$ ) ESI-MS: observed mass: 2283 Da, calculated mass of: 2283 Da;  $[C_{90}H_{158}N_{13}O_{54}]^{3+}$  observed: 762.00, calculated: 762.00;  $[C_{90}H_{157}N_{13}O_{54}]^{2+}$  observed: 1142.50, calculated: 1142.50

#### 3.4.4. Degalactosylation via $\beta$ -1,4 galactosidase S (NEB)

1 mg desialylated SGP (0.44  $\mu$ mol) in phosphate buffer (50 mM, pH 6.0) was incubated with  $\beta$ -1,4 galactosidase S (8 U) at 32°C for 24 hours on the shaker and controlled over time by RP-HPLC and MS. No degalactosylation reaction progress was observed.

The reaction was repeated with 1 mg of desialylated SGP (0.44  $\mu$ mol) in glycobuffer (bought from New England Biolab), which consisted of 5 mM  $CaCl_2$  and 50 mM sodium acetate at pH 5.5, with  $\beta$ -1,4 galactosidase S (8 U) at 32°C for up to 24 hours on the shaker. The reaction was controlled over time by RP-HPLC and MS. No degalactosylation reaction progress was observed.

The reaction with 1 mg of desialylated SGP in phosphate buffer was repeated with new ordered enzymes from New England Biolab and Sigma-Aldrich. With Dionex RP-HPLC system, 0.6 mg desialylated SGP were purified. 0.6 mg of purified desialylated SGP (0.26  $\mu$ mol) in phosphate buffer (50 mM, pH 6.0) was incubated with  $\beta$ -1,4 galactosidase S (3 U) from New England Biolab at 32°C for 24 hours on the shaker and controlled over time by RP-HPLC and MS. No degalactosylation reaction was observed via New England Biolab enzyme and extremely low activity can be seen via Sigma Aldrich enzyme, which does not increase significantly over time.

#### 3.4.5. Galactosidase activity assay

##### Approach 1:

Sodium acetate buffer (400  $\mu$ L) was mixed with 20  $\mu$ L enzyme solution, then 500  $\mu$ L 10 mM p-nitrophenyl- $\beta$ -D-galactopyranoside (PNP-Gal) were added. The mixture was mixed by inversion and incubated under shaking at 25°C for exactly 10 min. Finally, 2 mL borate buffer (200 mM

boric acid, pH 9.8) were added to the solution. 2 mL of the end solution were transferred to a cuvette and the enzyme activity was measured with NanoDrop 2000c Spectrophotometer at  $A_{400\text{nm}}$ .

#### Solution preparation:

Sodium acetate buffer: 200 mM sodium acetate, 0.25 % (w/v) Bovine Serum Albumin, 100 mM sodium chloride, pH 4.4 at 25°C.

$\beta$ -Galactosidase enzyme solution: 1 unit/ml dilution of  $\beta$ -Galactosidase in cold water.

#### Approach 2:

The following reaction was set up in five different falcon tubes and analyzed with NanoDrop 2000c Spectrophotometer at  $A_{405\text{nm}}$ :

Tube	Sodium phosphate [mL]	pNP- $\beta$ Gal [mL]	Enzyme [ $\mu$ L]
1	2.70	0.3	0
2	2.40	0.3	300
3	2.55	0.3	150
4	2.60	0.3	100
5	2.65	0.3	50

**TABLE 3 GALACTOSIDASE ACTIVITY ASSAY (LINEAR APPROACH)**

While sodium phosphate (0.08 M, pH 7.7) and 4-nitrophenyl- $\beta$ -D-galactopyranoside (pNP- $\beta$ Gal, 0.033 mM) solutions were mixed in advance, the enzyme solution was kept on ice and only added just before the UV measurement at  $A_{405\text{nm}}$ . The solution was mixed gently through inversion. The first falcon tube was used as blank. The second falcon tube with the highest enzyme concentration was measured exactly 30 secs after the addition of the enzyme. The measurement continued at 30-sec intervals, until the absorption value did not longer change. The final  $A_{405\text{nm}}$  value was used as  $\Delta A_{405\text{max}}$ . The three tubes were also measured in 30-sec intervals after enzyme addition for three minutes. The absorption values were plotted and the linear gradient  $\Delta A_{405}/\Delta$  time per minute was determined (Figure 19). Instead of the expected 10 units/mL, the results indicated approximately 0.1 unit/mL enzyme activity (Equation 1).

#### Solution preparation:

Enzyme solution: 10 units/mL in sodium phosphate buffer, supplemented with 1 mg/mL bovine serum albumin.

## 3.5. Synthesis of the key-sequence (246-281aa)

### 3.5.1. Manual synthesis with TentaGel resin (first approach)

0.05 mmol scale manual SPPS synthesis was performed on TentaGel-R-PHB- Lys(Boc)Fmoc resin. 2.5 eq Fmoc-aa, 2.38 eq HBTU (0.5 M) and 5 eq DIEA were used for the coupling step of the first 20 amino acids. Afterwards, the double amount was used (5 eq Fmoc-aa, 4.76 eq HBTU (0.5M), 10 eq DIEA). Double coupling and pseudoprolines were used to enhance the quality of the synthesis (Figure 21). Kaiser tests and testcleavages were performed for controlling the reaction progress. The crude product was analyzed via HPLC-MS and RP-HPLC. No purification step was performed due to insufficient purity

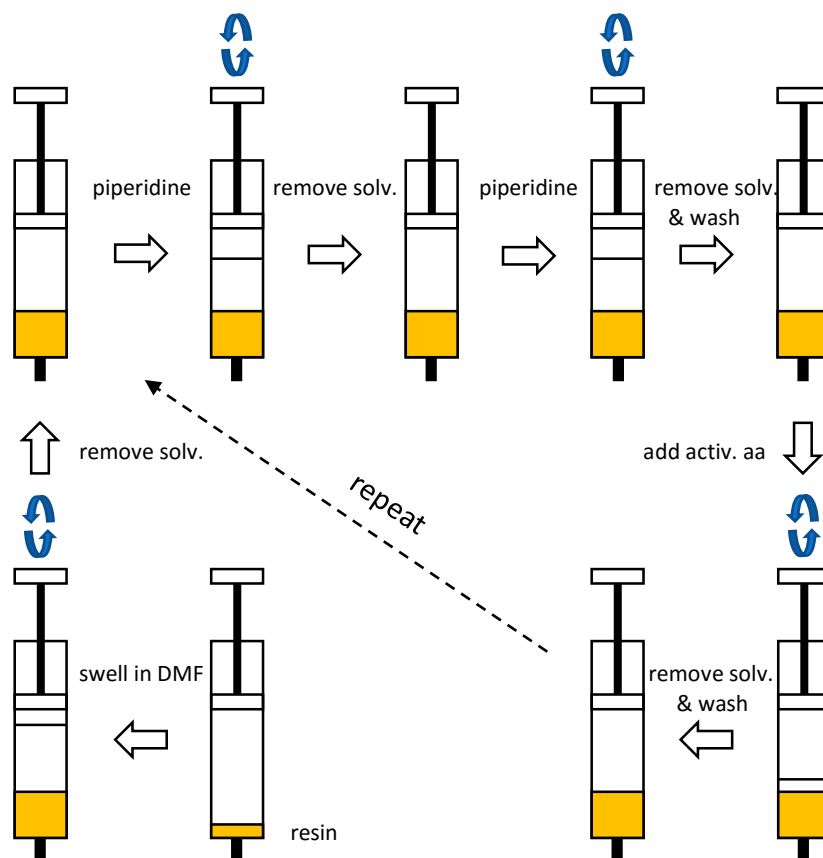
35mer key sequence ( $C_{186}H_{276}N_{42}O_{56}$ ) ESI-MS: observed mass: 3996 Da, calculated mass 3996 Da;  $[C_{186}H_{280}N_{42}O_{56}]^{4+}$  observed: 1000, calculated: 1000;  $[C_{186}H_{279}N_{42}O_{56}]^{3+}$  observed: 1333, calculated: 1333

#### General procedure for manual SPPS:

The appropriate amount of resin was weighted in a syringe and swelled in DMF for at least 30 min. Fmoc from the first amino acid on the solid support was deprotected twice with 20% (v/v) piperidine/DMF solution for 3 min each. After thoroughly washing the resin with DMF, the activated amino acid solution was added and incubated for 30 min. If the amino acid needed to be double coupled, the first solution was removed, a second portion containing the same amount of activated amino acid solution was added and incubated for another 30 min. Kaiser tests could be performed before and after this step. Finally, the resin was washed again with DMF.

Testcleavages and final cleavage could be performed after washing the resin with DCM and drying it in the desiccator. Afterwards, the resin was incubated with a cleavage solution for 2-3 hours. The peptide in the cleavage solution was precipitated by addition of diethylether. The solvents were removed after centrifugation and the peptide pellet was washed again with diethyl ether. Finally, the crude product was dried with nitrogen, dissolved in 50/50 buffer and lyophilized to complete dryness.





**FIGURE 10 GENERAL PROCEDURE FOR MANUAL Fmoc BASED SPPS, THE SYNTHESIS CYCLE IS REPEATED UNTIL THE COMPLETE PEPTIDE SEQUENCE HAS BEEN SYNTHESIZED**

### Solution preparation:

Activated amino acid: 2.5 eq Fmoc-aa were dissolved in 2.38 eq HBTU (0.5M solution) and 5 eq DIEA and vortexed.

Kaiser test: The Kaiser test could be performed before and after adding the activated amino acid. About 5 mg resin were taken out, washed with DCM and dried under vacuum. 40  $\mu$ L KCN, 40  $\mu$ L phenol and 40  $\mu$ L ninhydrin were added and the mixture was heated to 90°C for 5-10 min. Afterwards, depending on the colour of the mixture, the presence of free amino groups could be determined. If the color of the mixture was yellow, the test result was negative and therefore no free amino groups were present. If it was purple, the result was positive which indicated that the protecting group of the N-terminal amino acid was not coupled anymore and the amino group was free.

Standard cleavage solution: 5 % Triisopropylsilane (TIS), 2.5 % H<sub>2</sub>O, 92.5 % TFA were mixed together and applied to the resin

### 3.5.2. Manual synthesis with TentaGel resin (second approach)

0.05 mmol scale manual SPPS synthesis was performed on the TentaGel-R-PHB- Lys(Boc)Fmoc resin. 3.75 eq Fmoc-aa, 3.57 eq HBTU (0.5 M) and 7.5 eq DIEA were used. Double coupling and pseudoprolines were used to enhance the quality of the synthesis (Figure 23). Kaiser tests and test cleavages were performed for controlling the reaction progress. The crude product was analyzed via HPLC-MS and RP-HPLC. No purification step was performed due to insufficient purity.

35mer key sequence ( $C_{186}H_{276}N_{42}O_{56}$ ) ESI-MS: observed mass: 3996 Da, calculated mass 3996 Da;  $[C_{186}H_{280}N_{42}O_{56}]^{4+}$  observed: 1000, calculated: 1000;  $[C_{186}H_{279}N_{42}O_{56}]^{3+}$  observed: 1333, calculated: 1333

### 3.5.3. Automated Liberty Blue synthesis with TentaGel resin

0.05 mmol scale automated SPPS synthesis was performed on the TentaGel-R-PHB- Lys(Boc)Fmoc resin via CEM Liberty Blue Microwave Peptide Synthesizer. The standard programmed condition was used. Double coupling and pseudoprolines were applied to enhance the quality of the synthesis (Figure 23). Kaiser tests and test cleavages were performed for controlling the reaction progress. The crude product was analyzed via HPLC-MS and RP-HPLC. No purification step was performed due to insufficient purity and further development of the PEGylation method.

35mer key sequence ( $C_{186}H_{276}N_{42}O_{56}$ ) ESI-MS: observed mass: 3996 Da, calculated mass 3996 Da;  $[C_{186}H_{280}N_{42}O_{56}]^{4+}$  observed: 1000, calculated: 1000;  $[C_{186}H_{279}N_{42}O_{56}]^{3+}$  observed: 1333, calculated: 1333

### 3.5.4. Automated Liberty Blue synthesis with ChemMatrix resin

0.025 mmol scale automated SPPS synthesis was performed on H-Lys(Boc)-HMPB-ChemMatrix resin via CEM Liberty Blue Microwave Peptide Synthesizer. The standard programmed condition was used. Double coupling and pseudoprolines were applied to enhance the quality of the synthesis (Figure 23). Kaiser tests and test cleavages were performed for controlling the reaction progress. The crude product was analyzed via HPLC-MS and RP-HPLC. No purification step was performed due to insufficient purity.

35mer key sequence ( $C_{186}H_{276}N_{42}O_{56}$ ) ESI-MS: observed mass: 3996 Da, calculated mass 3996 Da;  $[C_{186}H_{280}N_{42}O_{56}]^{4+}$  observed: 1000, calculated: 1000;  $[C_{186}H_{280}N_{42}O_{56}]^{3+}$  observed: 1333, calculated: 1333

### 3.5.5. Automated Liberty Blue synthesis with Wang resin

0.025 mmol scale automated SPPS synthesis was performed on Fmoc-Lys(Boc)-Wang resin via CEM Liberty Blue Microwave Peptide Synthesizer. The standard programmed condition was used. Double coupling and pseudoprolines were applied to enhance the quality of the synthesis (Figure 23). Kaiser tests and test cleavages were performed for controlling the reaction progress. The crude product was analyzed via HPLC-MS and RP-HPLC. No purification step was performed due to further development of the PEGylation method.

35mer key sequence ( $C_{186}H_{276}N_{42}O_{56}$ ) ESI-MS: observed mass: 3996 Da, calculated mass 3996 Da;  $[C_{186}H_{280}N_{42}O_{56}]^{4+}$  observed: 1000, calculated: 1000;  $[C_{186}H_{279}N_{42}O_{56}]^{3+}$  observed: 1333, calculated: 1333

## 3.6. Attempt of auxilliary attachment

### 3.6.1. PEGylation of the key-sequence

14 mg of Fmoc deprotected peptidyl TentaGel resin (appr. 2.3  $\mu$ mol) was swollen in DMF for at least 30 min and transferred into a 0.5 mL Eppendorf tube. 1.38 eq (3.2  $\mu$ mol) PEG<sub>27</sub> was dissolved in 1.19 eq (2.7  $\mu$ mol, 0.5 M solution of 40/60, ACN/DMF) HATU and 2.5 eq (5.8  $\mu$ mol) DIEA solution. The yellow viscous mixture was added to the resin and the Eppendorf tube was gently shaken overnight. Finally, the resin was washed with DMF and DCM, dried in desiccator and cleaved with standard cleavage cocktail. The Fmoc-protection of the PEG<sub>27</sub> was still on. The PEGylation with peptidyl Wang resin was performed accordingly. The crude product was analyzed via HPLC-MS and RP-HPLC. No purification step was performed due to the low purity and conversion rate based on the chromatogram (approx. 50 %).

35mer PEGylated key sequence ( $C_{260}H_{403}N_{43}O_{87}$ ) ESI-MS: observed mass: 5522 Da, calculated mass: 5522 Da;  $[C_{260}H_{409}N_{43}O_{87}]^{6+}$  observed: 921, calculated: 921;  $[C_{260}H_{408}N_{43}O_{87}]^{5+}$  observed: 1105, calculated: 1105,  $[C_{260}H_{407}N_{43}O_{87}]^{4+}$  observed: 1381, calculated: 1382;  $[C_{260}H_{406}N_{43}O_{87}]^{3+}$  observed: 1841, calculated: 1842

## 3.7. Synthesis of the remaining sequences (141-201aa; 202-245aa)

### 3.7.1. Synthesis of the sequence 141-201aa

0.025 and 0.05 mmol scale automated SPPS synthesis were performed on TentaGel-S-PHB-Ser(t-Bu) Fmoc resin via CEM Liberty Blue Microwave Peptide Synthesizer for three different approaches. The standard programmed condition was used. Double coupling and pseudoprolines were applied at different positions to enhance the quality of the synthesis (Figure 32, Figure 35, Figure 37). Kaiser tests and test cleavages were performed for controlling the reaction progress. Cleavage cocktail H was applied for test cleavages. The crude product was analyzed via HPLC-MS and RP-HPLC. No purification step was performed due to further development of the hydrazine cleavage (Section 4.6.1).

61mer N-terminal sequence ( $C_{318}H_{501}N_{83}O_{89}S$ ) ESI-MS: observed mass: 6944 Da, calculated mass: 6943 Da;  $[C_{318}H_{510}N_{83}O_{89}S]^{9+}$  observed: 772, calculated: 772;  $[C_{318}H_{509}N_{83}O_{89}S]^{8+}$  observed: 869, calculated: 869;  $[C_{318}H_{508}N_{83}O_{89}S]^{7+}$  observed: 993, calculated 993;  $[C_{318}H_{507}N_{83}O_{89}S]^{6+}$  observed: 1158, calculated: 1158;  $[C_{318}H_{506}N_{83}O_{89}S]^{5+}$  observed: 1389, calculated 1390,  $[C_{318}H_{505}N_{83}O_{89}S]^{4+}$ : observed: 1737, calculated 1737

#### General procedure for methionine reducing cleavage cocktail (cocktail H):

82.5 % TFA, 2.5 % EDT, 5 % thioanisole, 5 % H<sub>2</sub>O, 5% dimethyl sulfide (Me<sub>2</sub>S) were mixed. Phenol (75 mg), NH<sub>4</sub>I (50 mg) and triisopropylsilane (TIS, 100 μL) were added per mL. Vortex until phenol completely and NH<sub>4</sub>I partially dissolved.

### 3.7.2. Synthesis of the sequence 202-245aa

0.025 mmol scale automated SPPS synthesis was performed on TentaGel-R-PHB- Lys(Boc)Fmoc resin via CEM Liberty Blue Microwave Peptide Synthesizer for three different approaches. The standard programmed condition was used. Double coupling and pseudoprolines were applied to enhance the quality of the synthesis (Figure 45). Kaiser tests and test cleavages were performed for controlling the reaction progress. Cleavage cocktail H was applied for test cleavages. The crude product was analyzed via HPLC-MS and RP-HPLC. No purification step was performed due to further development of the hydrazine cleavage (Section 4.6.3).

44mer middle sequence ( $C_{224}H_{351}N_{61}O_{65}S_8$ ) ESI-MS: observed mass: 5194 Da, calculated mass: 5195 Da;  $[C_{224}H_{358}N_{61}O_{65}S_8]^{7+}$  observed mass: 743, calculated mass: 743;  $[C_{224}H_{357}N_{61}O_{65}S_8]^{6+}$

observed mass: 867, calculated mass: 866;  $[C_{224}H_{356}N_{61}O_{65}S_8]^{5+}$  observed mass: 1040, calculated mass: 1040;  $[C_{224}H_{355}N_{61}O_{65}S_8]^{4+}$  observed mass: 1299, calculated mass: 1300,  $[C_{224}H_{354}N_{61}O_{65}S_8]^{3+}$  observed mass: 1732, calculated mass: 1733

## 3.8. Ligations of the N-terminal and middle sequences

### 3.8.1. Hydrazine cleavage of the sequence 141-201aa

#### Approach 1 (without washing step)

10 mg dry peptidyl resin in 200  $\mu$ L hydrazine (1 M, THF) solution were shaken gently overnight (18-20 hours) at room temperature. The hydrazine solution was removed, neutralized with acetic acid and dried over evaporator. The resin was washed once with DCM and dried in the desiccator. The washing solution was collected and evaporated under vacuum. The resin, the dried hydrazine solution and the washing residue were subject to the methionine reducing cleavage. The crude product was analyzed via HPLC-MS and RP-HPLC. No purification step was performed.

#### Approach 2 (small scale)

30 mg dry peptidyl resin in 600  $\mu$ L hydrazine (1 M, THF) solution were shaken gently overnight (20-24 hours) at room temperature. The hydrazine solution was removed, neutralized with acetic acid and dried over evaporator. The resin was washed with 600  $\mu$ L DCM for 30 min and twice with 1 mL TFE for each 2-3 hours. Every washing step was performed in a syringe under shaking at room temperature. Finally, the resin was dried over desiccator and the washing fractions were evaporated under reduced pressure. The resin, the dried hydrazine and the washing residue were separately incubated with cocktail H to remove possible peptide remained on the resin (in the case of the resin) and the side chain protecting group of the peptide hydrazide. The crude product was analyzed via HPLC-MS and RP-HPLC and purified via SemiPrep column. 2.6 mg (0.37 mmol, 13 % based on the synthesis scale) purified product was obtained.

#### Approach 3 (big scale)

600 mg dry peptidyl resin were separated into three syringes, each 200 mg. 4 mL hydrazine (1 M, THF) solution was added to each of the syringes and gently shaken overnight for 24 hours at room temperature. The hydrazine solution was removed, neutralized with acetic acid and dried over evaporator. The resin was washed with 5 mL DCM for 30 min and twice with 5 mL TFE for each 2-3 hours. Every washing step was performed in a syringe under shaking at room temperature. Finally, the resin was dried over desiccator and the washing fractions were dried under vacuum. The resin, the dried hydrazine and the washing residue were subject to the cocktail H incubation. The crude product was analyzed via HPLC-MS and RP-HPLC and purified via SemiPrep column. 9.7 mg (1.4 mmol, 3 % based on the synthesis scale) purified product was obtained.

*N*-terminal hydrazide peptide (C<sub>318</sub>H<sub>503</sub>N<sub>85</sub>O<sub>88</sub>S) ESI-MS: observed mass: 6958 Da, calculated mass 6957 Da; [C<sub>318</sub>H<sub>512</sub>N<sub>85</sub>O<sub>88</sub>S]<sup>9+</sup> observed: 774, calculated 774; [C<sub>318</sub>H<sub>511</sub>N<sub>85</sub>O<sub>88</sub>S]<sup>8+</sup> observed: 871, calculated 871; [C<sub>318</sub>H<sub>510</sub>N<sub>85</sub>O<sub>88</sub>S]<sup>7+</sup> observed: 996, calculated 995; [C<sub>318</sub>H<sub>509</sub>N<sub>85</sub>O<sub>88</sub>S]<sup>6+</sup> observed: 1160, calculated 1160; [C<sub>318</sub>H<sub>508</sub>N<sub>85</sub>O<sub>88</sub>S]<sup>5+</sup> observed: 1392, calculated 1392; [C<sub>318</sub>H<sub>507</sub>N<sub>85</sub>O<sub>88</sub>S]<sup>4+</sup> observed: 1740, calculated 1740

### 3.8.2. Thioester preparation of the sequence 141-201aa

0.093 μmol (0.65 mg) purified hydrazine peptide was dissolved in 15 μL buffer (pH 3). 7.5 eq. NaNO<sub>2</sub> solution (3.5 μL) was added at -20°C and the mixture was stirred for 15 mins. Afterwards, 40.4 eq. mesNa solution (9.3 μL) was added and after stirring for 15 mins, 1 μL ascorbic acid was added to quench the reaction. The solution was diluted with water and directly purified with HPLC for the further ligation reaction. 0.52 mg (73 μmol, 79 %) yield was obtained and analyzed via HPLC-MS and RP-HPLC.

*N*-terminal thioester peptide (C<sub>320</sub>H<sub>505</sub>N<sub>83</sub>O<sub>91</sub>S<sub>3</sub>) ESI-MS: calculated mass: 7066 Da, observed mass: 7066 Da; [C<sub>320</sub>H<sub>514</sub>N<sub>83</sub>O<sub>91</sub>S<sub>3</sub>]<sup>9+</sup> observed: 786.20, calculated 786.20; [C<sub>320</sub>H<sub>513</sub>N<sub>83</sub>O<sub>91</sub>S<sub>3</sub>]<sup>8+</sup> observed: 884.35, calculated 884.35; [C<sub>320</sub>H<sub>512</sub>N<sub>83</sub>O<sub>91</sub>S<sub>3</sub>]<sup>7+</sup> observed: 1010.39, calculated 1010.39; [C<sub>320</sub>H<sub>511</sub>N<sub>83</sub>O<sub>91</sub>S<sub>3</sub>]<sup>6+</sup> observed: 1178,79, calculated 1178,79; [C<sub>320</sub>H<sub>510</sub>N<sub>83</sub>O<sub>91</sub>S<sub>3</sub>]<sup>5+</sup> observed: 1414.35, calculated 1414.35

#### Solution preparation:

Buffer at pH 3: 6 M guanidine hydrochloride and 0.2 M Na<sub>2</sub>HPO<sub>4</sub> were dissolved and the pH was adjusted to 3 with HCl.

Buffer at pH 7.5: 6 M guanidine hydrochloride and 0.2 M Na<sub>2</sub>HPO<sub>4</sub> were dissolved and the pH was adjusted to 7.5 with NaOH.

Ascorbic acid solution: 2.5 mg ascorbic acid dissolved in 7 μL buffer at pH 7.5.

MesNa solution: 16.4 mg of mesNa were dissolved in 250 μL buffer at pH 7.5 (0.4 M).

NaNO<sub>2</sub> solution: 10.4 mg of NaNO<sub>2</sub> were dissolved in 750 μL H<sub>2</sub>O (0.2 M).

### 3.8.3. Hydrazine cleavage of the sequence 202-245aa

100 mg dry peptidyl resin in 2 mL hydrazine (1 M, THF) solution were gently shaken overnight for 24 hours at room temperature. The hydrazine solution was removed, neutralized with acetic acid and dried over evaporator. The resin was washed with 2 mL DCM for 30 min and twice with 2 mL

TFE for each 2-3 hours. Every washing step was performed in a syringe under shaking at room temperature. Finally, the resin was dried over desiccator and the washing fractions were dried under vacuum. The resin, the dried hydrazine and the washing residue were subject to cleavage cocktail H. 50 mg crude product were analyzed via HPLC-MS and RP-HPLC and purified via SemiPrep column. 2.5 mg (0.48 mmol, 5 % based on the synthesis scale) purified product was obtained.

44mer middle sequence hydrazide peptide ( $C_{224}H_{351}N_{61}O_{65}S_8$ ) ESI-MS: calculated mass: 5209 Da, observed mass: 5208 Da;  $[C_{224}H_{358}N_{61}O_{65}S_8]^{7+}$  observed: 745, calculated 745;  $[C_{224}H_{357}N_{61}O_{65}S_8]^{6+}$  observed: 869, calculated 869;  $[C_{224}H_{356}N_{61}O_{65}S_8]^{5+}$  observed: 1043, calculated 1043;  $[C_{224}H_{355}N_{61}O_{65}S_8]^{4+}$  observed: 1303, calculated 1303;  $[C_{224}H_{354}N_{61}O_{65}S_8]^{3+}$  observed: 1737, calculated 1737

### 3.8.4. Ligation assays for 141-245aa-Hydrazide peptide

1.1 mg (0.156  $\mu$ mol) of thioester peptide and 1.1 eq of hydrazide peptide (0.89 mg, 0.171  $\mu$ mol) were dissolved in 40  $\mu$ L buffer with mesNa at pH 7.5 and stirred at room temperature. 4  $\mu$ L of TCEP solution were added after 5 mins and the reaction was controlled via HPLC-MS and RP-HPLC over time. The desired product can already be observed immediately after adding the hydrazide peptide, however, the conversion of reaction seems to stop after 5 hours. Due to the extremely low reaction conversion and product concentration, the UV peak of the product cannot be clearly determined.

141-245aa hydrazide peptide ( $C_{542}H_{852}N_{146}O_{152}S_9$ ) ESI-MS: calculated mass: 12133 Da, observed mass: 12133 Da;  $[C_{542}H_{863}N_{146}O_{152}S_9]^{11+}$  observed: 1104.12, calculated 1104.12;  $[C_{542}H_{862}N_{146}O_{152}S_9]^{10+}$  observed: 1214.33, calculated 1214.33;  $[C_{542}H_{861}N_{146}O_{152}S_9]^{9+}$  observed: 1349.14, calculated 1349.14

#### Solution preparation:

Buffer with mesNa at pH 7.5: 3.3 mg of mesNa (0.4 M) were dissolved in 50  $\mu$ L of buffer 7.5 (Section 3.8.2) and degassed with  $N_2$  for 5 mins.

TCEP solution: 2 mg were dissolved in 200  $\mu$ L phosphate buffer at pH 8.

## 3.9. Endo-A

### 3.9.1. Endo-A test expression

DNA obtained from professor Takegawa was transferred to E. Coli BL-21 DE3 competent cell agar gel plate and one of the colonies was grown overnight under shaking at 36°C with 100 mL 50/50 (v/v) LB and 2YT medium mixture. Four different cell flasks with each 250 mL media were prepared and 15 mL of the cell mixture were added. To control the specific cell expression, ampicillin was added. The four flasks were shaken at 36°C and the cell density was measured via Nanodrop UV spectrometer. At cell density 0.6, to the flasks with LB and 2YT media 250 µL IPTG were added each and samples (1 mL each) were taken out over time to visualize the protein growth with SDS-PAGE. The same was repeated at cell density 1 (Table 4).

Medium	LB	2YT	LB	2YT
Cell density at IPTG activation	0.6	0.6	1	1

TABLE 4 ENDO-A TEST EXPRESSION REACTION SETUP

The samples for SDS-PAGE were centrifuged for 1.5 min. The supernatant was removed, 5 µL H<sub>2</sub>O and 5 µL SDS PAGE loading buffer were added to each protein pellet. Before applying the prepared samples to SDS-PAGE, the mixture was vortexed and heated to 90°C. The molecular weight of the soluble Endo-A has approximately 80 kDa and its expression already starts before IPTG activation (Figure 53) and therefore no conclusion can be drawn between the two cell activation steps at different cell density.

Therefore, the Endo-A DNA was transferred to E. Coli BL-21 DE3 RIL, BL-21 DE3 Rosetta and BL-21 DE3 pLySs cell agar gel plates. One of the colonies on each agar gel plate was transferred and grown overnight under shaking at 36°C with 50 mL LB medium. Three cell flasks with each 125 mL LB medium were prepared and 7 mL of each cell mixture were added. To control the specific cell expression, ampicillin and chloramphenicol were added. The three flasks were shaken at 36°C and the cell density was measured via Nanodrop UV spectrometer. At cell density 0.6, 250 µL IPTG were added to each flask and samples (1 mL each) were taken out over time to visualize the protein growth with SDS-PAGE. All three variations gave satisfying cell growth signals and the target protein was successfully induced upon IPTG addition.

#### Solution preparation:

2YT medium: 16 g tryptone, 10 g yeast extract and 5 g NaCl were dissolved in 1-liter H<sub>2</sub>O and the solution was autoclaved.

LB medium: 10 g tryptone, 5 g yeast extract and 10 g NaCl were dissolved in 1-liter H<sub>2</sub>O and the solution was autoclaved.



Ampicillin: 100 mg ampicillin were dissolved in 1 mL H<sub>2</sub>O and sterile filtered.

Chloramphenicol: 30 mg chloramphenicol was dissolved in 1 mL ethanol and filtered sterile.

### 3.9.2. Endo-A activity assay

#### Approach 1 (without HPLC SGP purification)

In three separate Eppis, 1.5 mg of SGP each were dissolved in 600  $\mu$ L phosphate buffer (50 mM, pH 6.0). 10  $\mu$ L, 30  $\mu$ L and 50  $\mu$ L enzyme solutions (~2 mg/mL; 10 mM potassium phosphate, 0.5 M NaCl, pH 7.0) were added and gently mixed. Finally, the total volume of each Eppi was increased to 1 mL with phosphate buffer. The three mixtures were shaken at 37°C and the reaction progress was controlled via HPLC-MS and RP-HPLC. Unfortunately, the SGP peak did not decrease overtime due to the reaction of another impurity peak with Endo-A (Figure 56).

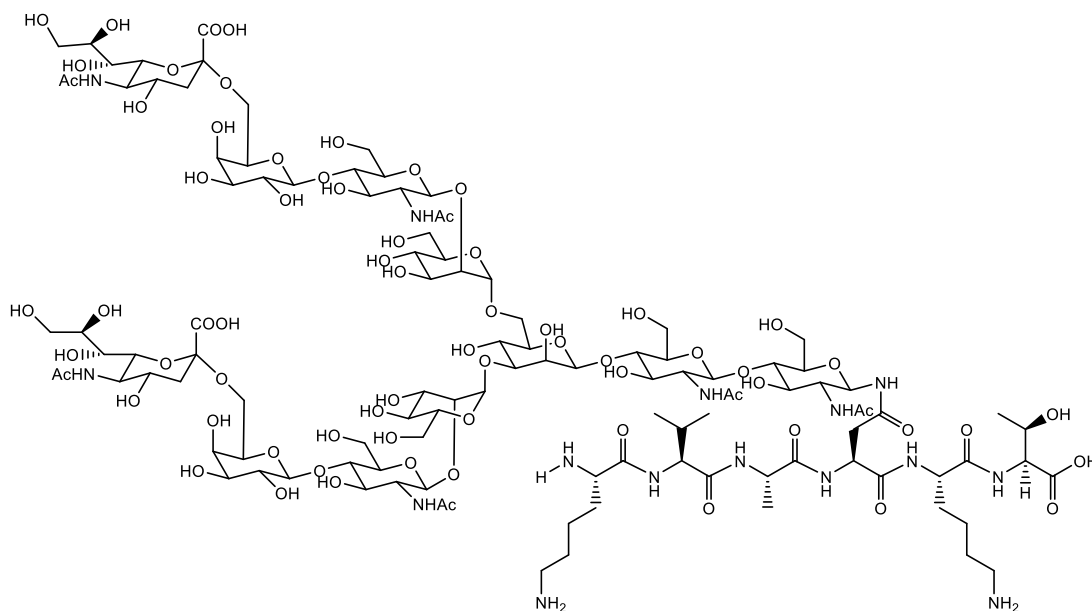
#### Approach 2 (with HPLC SGP purification)

1 mg of purified SGP was dissolved in 300  $\mu$ L phosphate buffer (50 mM, pH 6.0) and 50  $\mu$ L enzyme solution (~4 mg/mL; 10 mM potassium phosphate) were added and gently mixed. Finally, the total volume was increased to 600  $\mu$ L with phosphate buffer. The mixture was shaken at 37°C and the reaction progress was controlled via HPLC-MS and RP-HPLC. The decrease of the SGP peak could be observed over time via HPLC- UV chromatogram, which indicated the successful trimming step. However, the desired mass of the two fragments could not be found with MS.

# 4. RESULTS AND DISCUSSION

## 4.1. Isolation and purification of SGP from egg yolk

For the synthesis of homogenously *N*-glycosylated sFas Ligand, it is necessary to obtain the homogenous sugar core structure. Recently a simplified procedure for gram scale production of sialylglycopeptide (Figure 11) from egg yolk has been reported by Sun et al. : the *N*-glycan core Man3GlcNAc can be obtained <sup>[50]</sup> via glycan-trimming and successive oxazoline formation.

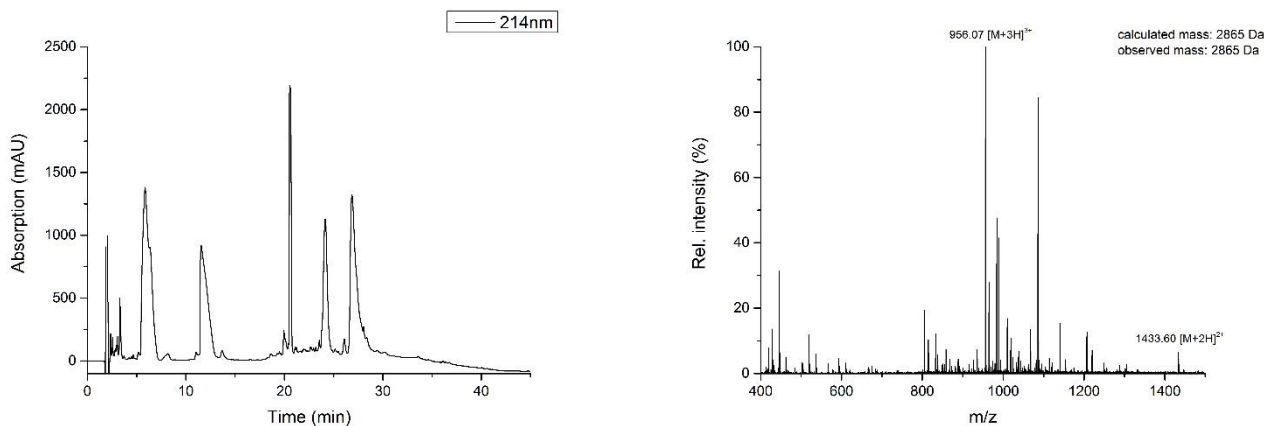


**FIGURE 11 SIALYLGlyCOPEPTIDE STRUCTURE**

### 4.1.1. Isolation of SGP from egg yolk

The modified procedure from Sun et al. <sup>[50]</sup> has been used for the isolation and purification of SGP. Three batches of 30 egg yolks each have been used altogether. The egg yolks are stirred with water for one hour and lyophilized to yolk powder. Thereafter, the powder is washed thoroughly first with diethyl ether and then 70 % acetone to get rid of lipids and unwanted substances without generating emulsion. It becomes apparent that especially the diethyl ether washing step need to be repeated several times. The remaining residue is mixed vigorously with 40 % acetone for a few hours and filtered first with folded filter, then with Celite which prevents

any possible emulsification trouble. 2-3.5 g of crude product have been obtained from each batch.



**FIGURE 12 LEFT: UV CHROMATOGRAM OF THE SGP CRUDE PRODUCT (SGP PEAK: TR = 12 MIN)**  
**RIGHT: MASS SPECTRUM OF CRUDE SGP, TR = 12 MIN; CALCULATED MASS: 2865 DA, OBSERVED MASS: 2865 DA;**  
**SEVERAL IMPURITIES CAN BE OBSERVED UNDER THE SAME PEAK, WHICH SHOWS UV-INACTIVE MOLECULES WITH SIMILAR**  
**RETENTION TIME AND INDICATES THE NECESSITY OF THE PURIFICATION STEP**

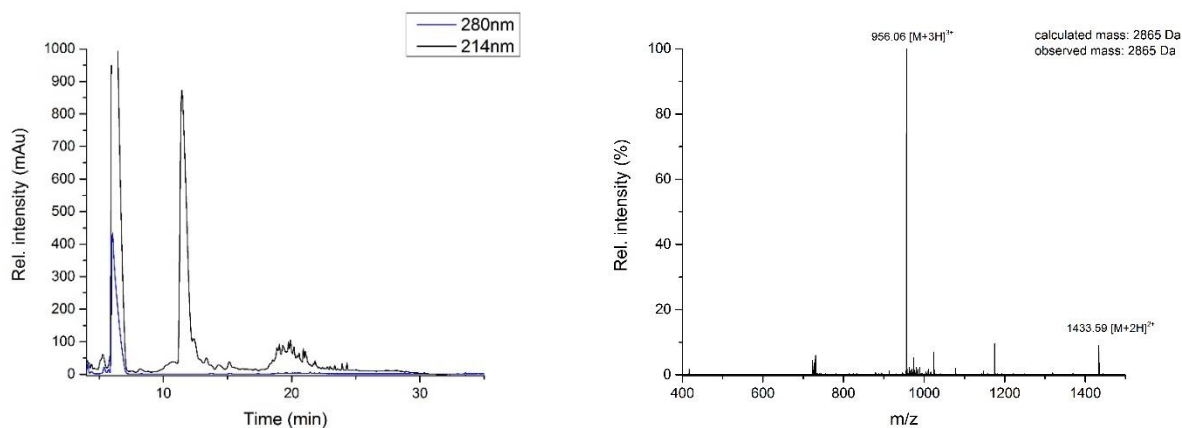
The assumed SGP peak on the UV chromatogram (Figure 12) has been collected and analyzed with HPLC-MS, which shows clearly that the obtained crude product indeed contained the desired SGP and that the assignment of the peak was correct. For further confirmation, the simulation of the SGP mass isotopic pattern has been used as comparison. Since several unidentified substances can also be observed under the same SGP peak on the chromatogram, the necessary purification has been performed. The SGP is very hydrophilic due to the many sugar structures, therefore a gradient of 5-15 % for 5 min, 15-65 % for 30 min on the C18 kromasil column has been selected and used both for analysis.

#### 4.1.2. Purification of SGP crude

The crude is purified via solid-phase extraction with active carbon and Celite. This convenient method has been reported by Sun et al. <sup>[50]</sup> to have the highest separation capability in comparison to other purification methods. The SGP extract is dissolved in water and added onto an active carbon/Celite (2:1 mixture) column, which is prepared with water containing 0.1 % TFA. Thereafter, the column is eluted with continues rising amount of acetonitrile (0-25 %) in water containing 0.1 % TFA. The last elution fraction with 25 % acetonitrile contains the desired SGP

(Figure 13). Altogether, 650 mg purified SGP of good quality (shown as below) have been obtained.

Another 960 mg of SGP side fraction have also been collected and can be used after a small scale carbon/Celite purification.



**FIGURE 13 LEFT: UV CHROMATOGRAM OF THE SGP PURIFIED PRODUCT (SGP PEAK: TR = 12 MIN) RIGHT: MASS SPECTRUM OF PURIFIED SGP, TR = 12 MIN ; CALCULATED MASS: 2865 DA, OBSERVED MASS: 2865 DA**

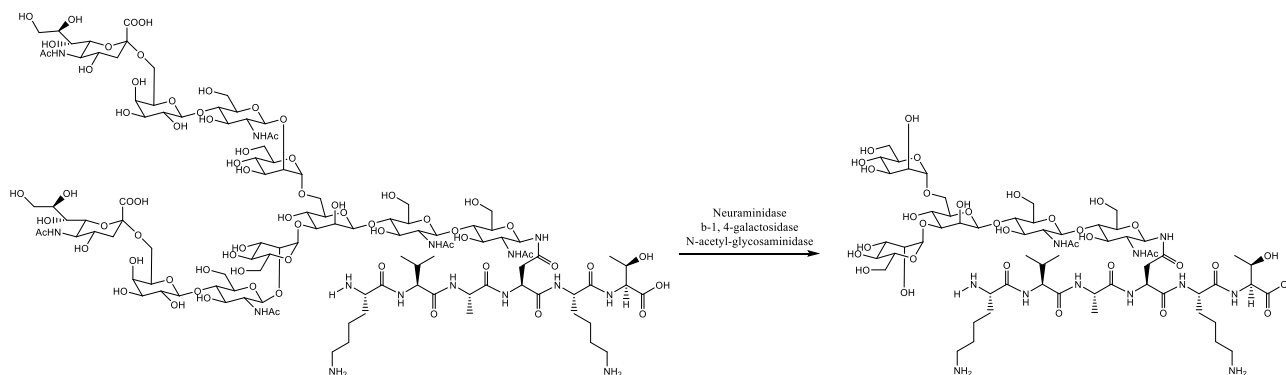
Attempts have been made to increase the acetonitrile percentage up to 35 % in the hope of higher yield. However, the quality of the SGP decreases drastically after the 25 % fraction and therefore it can be assumed that only a small amount, if at all, remains on the column after the purification. Also, no SGP can be observed in the chromatograms of the washing fractions.

In comparison to the crude SGP chromatogram (Figure 12), most of the unwanted substances have been removed during the purification (Figure 13). However, the chromatogram and also the mass spectrum still show traces of other molecules. One of them has a retention time of 5 mins and absorbs at 280 nm, which suggests that it might be another glycoprotein with aromatic amino acids. Unfortunately, due to extremely low ionization character of this molecule in positive and negative mode, an indication of its molecular weight cannot be provided. It is presumed that the purified SGP has sufficient purity for the further modifications.

## 4.2. Synthesis of Man<sub>3</sub>GlcNAc-peptide

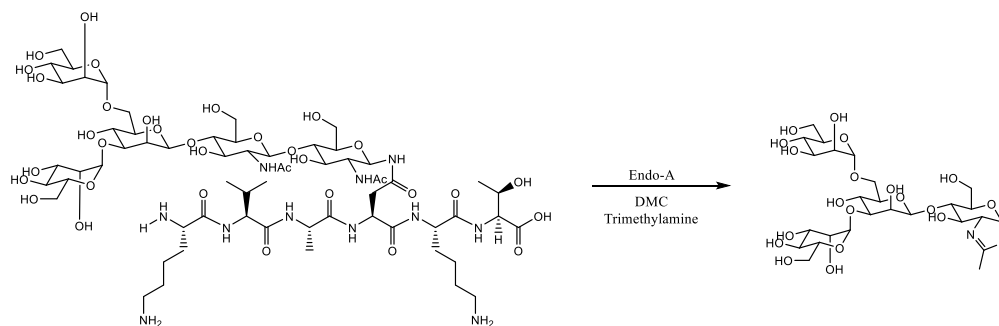
The primary strategy is to apply the described approach <sup>[50]</sup> using exo-glycosidases to trim non-reducing end carbohydrate moieties of *N*-glycans. According to the literature, SGP should be first treated with neuraminidase for the cleavage of sialic acids, then with  $\beta$ -1,4-galactosidase for the

removal of galactose moieties and finally with *N*-acetyl-glycosaminidase to obtain the Man<sub>3</sub>GlcNac-peptide (Figure 14). All enzymatic reactions should be performed sequentially in one pot (pH 6.0) and monitored step after step with HPLC and ESI-MS.



**FIGURE 14 ENZYMATIC TRIMMING FOR OBTAINING MAN<sub>3</sub>GLCNAC-PEPTIDE**

After the catalysis with glycosidase Endo-A, the released sugar core structure Man<sub>3</sub>GlcNac should be treated with 2-chloro-1,3-dimethylimidazolinium chloride (DMC) and trimethylamine to form the desired Man<sub>3</sub>GlcNac oxazoline as in the reported procedure (Figure 15).



**FIGURE 15 ENDO-A MAN<sub>3</sub>GLCNAC-OXAZOLINE FORMATION**

#### 4.2.1. Desialylation

For the further enzymatic and chemical trimming of SGP, the first two terminal sialic acid residues have to be cleaved. Therefore, first the approach with neuraminidase, an exo-glycoside hydrolase enzyme, has been attempted. Finally, the more convenient and inexpensive chemical desialylation method with acetic acid has led to the desired product.

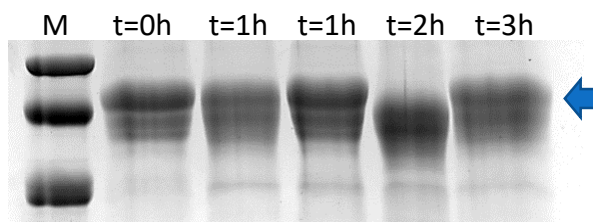
#### 4.2.1.1. Desialylation via Neuraminidase

The approach of  $\alpha$  2-3,6,8 neuraminidase (8.8 unit/ $\mu$ mol) in phosphate buffer (50 mM, pH 6.0) at 32°C has been tested with 50 and 5 mg of SGP [50]. Both reactions have been monitored with HPLC-MS. After 24 hours, no difference can be observed and an additional amount of enzyme has been added to the solution, giving the double enzyme concentration. However, no product formation has been observed after up to 42 hours. Thereafter, the suggested reaction glycosylation buffer (5 mM CaCl<sub>2</sub> and 50 mM sodium acetate at pH 5.5) of the supplier company (New England Biolab) have been tested with 5 mg of SGP. Again, no trace of the trimming reaction can be detected after 30 hours.

It is rather unlikely that the purity of SGP is not sufficient enough for the enzymatic reaction, since the following desialylation via acetic acid worked successfully as described below. Hence, some hypothesis can be formulated to explain this result. The possibility cannot be ruled out that the sample contains some kind of inhibitor for the enzyme, although nothing has been reported in the literature. To verify the enzyme activity and to bring some clarity into this matter, the suggested neuraminidase activity assay with fetuin from fetal bovine serum has been performed.

#### 4.2.1.2. Activity assay of Neuraminidase

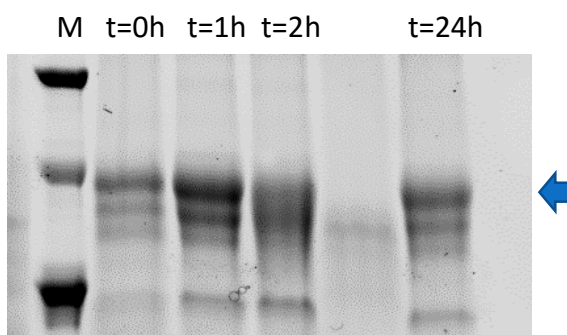
The recommended neuraminidase activity assay conditions according to New England Biolab have been used on fetuin. The first approach is visualized with 15 % SDS-PAGE gel (Figure 16) and the fetuin band appears around 70 kDa. At the first glance the band at t = 2h seems to be a bit lower than the rest, indicating possible enzymatic reaction. But since there is almost no difference between fetuin without enzyme (t = 0h) and fetuin after three hours of enzymatic incubation (t = 3h), it is more likely that there has been an execution error. However, there might be a small hint of activity around 45 kDa (blue arrow).



**FIGURE 16 ACTIVITY ASSAY OF NEURAMINIDASE: 15 % SDS PAGE GEL (M= MARKER, T= TIME)**

The band of fetuin is rather thick due to different branching and other heterogeneity of glycosylation. The assay has been repeated under the same reaction condition as in the first

approach and the reaction time is extended to 24 hours. In order to enhance the visualization, gradient SDS-PAGE gel is used (Figure 17).



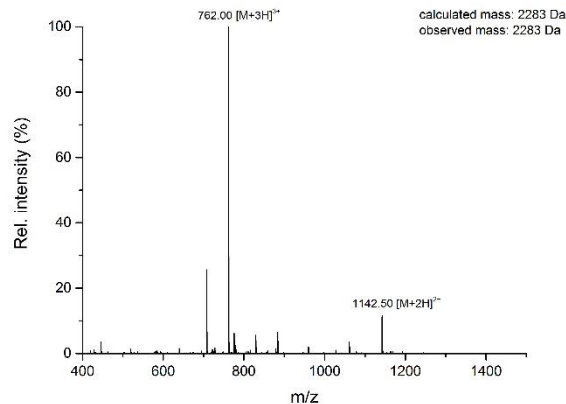
**FIGURE 17 ACTIVITY ASSAY OF NEURAMINIDASE: 8-16 % GRAD. SDS-PAGE GEL (M=MARKER, T=TIME)**

Fetuin without enzyme at  $t = 0h$  shows three strips around 50-66 kDa (Figure 17). Again, one can observe a trace of possible desialylated fetuin around 45 kDa (blue arrow), but no obvious new signal is present. Attempts have been made to generate the positive control of desialylated fetuin, unfortunately only limited informations about chemical desialylation of fetuin are reported in the literature <sup>[51]</sup> and no optimal condition has been found.

In summary it can be stated that the neuraminidase does not show any significant activity with fetuin as substrate after up to 24 hours, which also explains the failed enzymatic trimming reactions with SGP.

#### 4.2.1.3. Desialylation via Acetic Acid

After the unsuccessful enzymatic trimming with neuraminidase, an alternative strategy described by Shigeta et al. <sup>[52]</sup> has been applied. The sialic acid residues have been removed from the purified SGP by hydrolysis with 2 M acetic acid at 80°C for 2 hours. HPLC and MS analysis have confirmed that the desired product has been obtained (Figure 18). Furthermore, the simulation of the desialylated SGP mass isotopic pattern has been used as comparison to the measured signal.



**FIGURE 18 MASS SPECTRUM OF DESIALYLATED SGP, INTEGRATED FROM GLYCOPEPTIDE PEAK AT TR = 12 MIN;  
CALCULATED MASS: 2283 DA, OBSERVED MASS: 2283 DA**

This method has proven to be very stable, since small variations of the reaction condition regarding volume of acetic acid, reaction time or temperature apparently do not influence the desialylation process. For the following enzymatic reaction, the desialylated SGP is dissolved in water and lyophilized several times until a neutral pH is reached.

#### 4.2.2. Degalactosylation

The desialylated SGP needs to be further trimmed by  $\beta$  1-4 galactosidase, to remove two endo galactose residues. It has not been reported to be a reversible enzymatic reaction and the specific cleavage of this position is desired.

##### 4.2.2.1. Desialylation via Galactosidase

The described approach of  $\beta$  1-4 galactosidase S (4.7 unit/ $\mu$ mol) in phosphate buffer (50 mM, pH 6.0) at 32°C [50] and the suggested reaction conditions of the supplier company (New England Biolab) with their glycosylation cocktail have been tested with 1 mg of desialylated SGP each. Both reactions have been monitored with HPLC-MS. After 24 hours an additional amount of enzyme has been added to the solution. None of the two reactions could give the desired desgalactosyl product.

After confirming the low activity of the galactosidase from New England Biolab, a new batch of enzyme has been ordered from Sigma Aldrich. The approach of  $\beta$  1-4 galactosidase S (4.7 unit/ $\mu$ mol) in phosphate buffer (50 mM, pH 6.0) at 32°C has been repeated for 48 hours and the degalactosyl product can be observed with HPLC-MS after 6 hours. The simulation of the



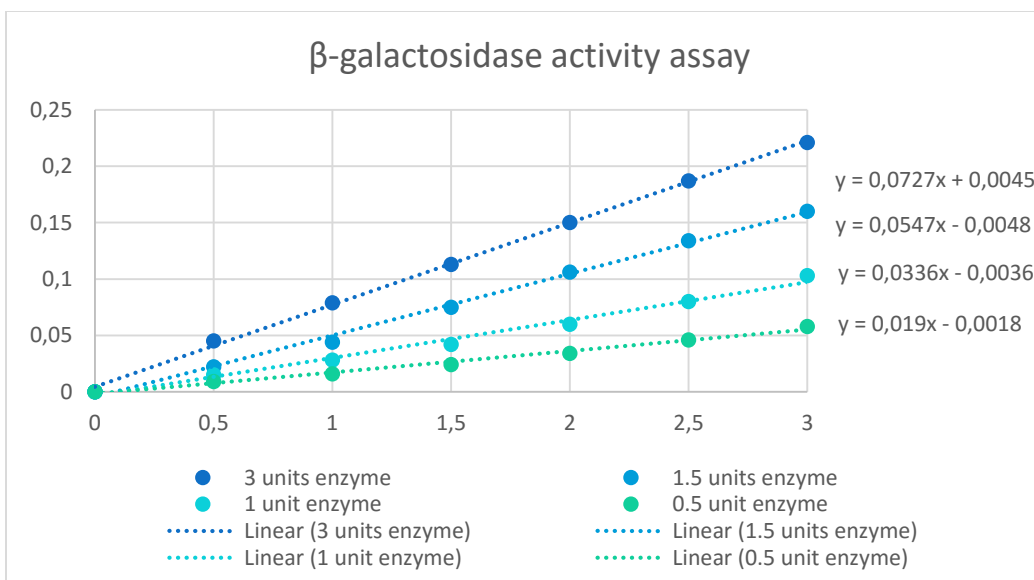
degalactosylated SGP mass isotopic pattern has been used as comparison to the measured signal. However, the mass spectrum signal of the product is extremely low and no significant UV change can be observed. In addition, the signal does not increase significantly over time.

Although the pH of the desialylated SGP has been neutral before the reaction, it is still possible that the amount of existing acetic acid salts influences the activity of the galactosidase enzyme. Also, it cannot be ruled out that there could be an inhibition caused by other kind of small inhibitor or substances in the sample, although nothing has been reported in the literature. The reaction condition and the concentration of used enzyme needs to be further enhanced. To verify the enzyme activity, the recommended galactosidase activity assay with 4-nitrophenyl- $\beta$ -D-galactopyranoside (pNP- $\beta$ Gal) has been performed <sup>[53]</sup>.

#### 4.2.2.2. Activity assay of Galactosidase

The reaction conditions described by Distler et al. <sup>[53-54]</sup> have been used for the activity assay. In sodium acetate buffer (pH 4.4) with bovine serum albumin, pNP- $\beta$ Gal is added to galactosidase and incubated for 10 minutes at 25°C. Thereafter, the reaction is quenched with borate buffer (pH 9.8) and the UV absorption at 400nm is recorded. This procedure has been repeated twice and in both cases no absorption can be measured. This indicates that the enzyme  $\beta$  1-4 galactosidase S from supplier NEB is inactive, which also explains the failed enzymatic reactions with desialyted SGP. The same activity reaction has been applied to the  $\beta$  1-4 galactosidase from Sigma Aldrich, weak absorption signals can be detected.

To determine the activity of  $\beta$  1-4 galactosidase from Sigma Aldrich more in detail and to establish an optimal  $\beta$  1-4 galactosidase activity assay for future applications, different amounts of enzyme have been used for the same concentration of pNP- $\beta$ Gal and the increase of absorption at 405nm over time has been measured. The highest amount of enzyme contains three units, with which the  $A_{405nm}$  value has been determined in 30 seconds intervals until it no longer changes. The final  $A_{405nm}$  value (0.442) indicates the completion of the reaction and has been used as  $\Delta A_{405max}$ . The other reactions with lower enzyme concentrations have been also determined in 30 seconds intervals for three mins, so that sets of  $A_{405nm}$  values that are linear with respect to time are obtained. Finally, the data have been plotted in the same graph and the slope of each linear fitted line has been determined as  $\Delta A_{405nm} / \Delta \text{time}$  (Figure 19).



**FIGURE 19 BETA- GALACTOSIDASE ENZYME ACTIVITY ASSAY; PLOT OF  $A_{405nm}$  VERSUS TIME, THREE DIFFERENT CONCENTRATIONS OF ENZYME HAVE BEEN USED**

For each line, the activity of enzyme has been calculated via the following equation (Equation 1). Since the concentration of the used stock enzyme solution is 10 units/mL according to the definition of the supplier, the average value should be in the magnitude of 10.

$$Activity = \frac{\left(\frac{\Delta A_{405}}{\Delta time}\right) \left(\frac{0.099}{\Delta A_{405Max}}\right)}{V_p}$$

$\frac{\Delta A_{405}}{\Delta time}$ : slope of the fitted linear function

$\Delta A_{405Max}$ : final  $A_{405nm}$  value at the end of the enzymatic reaction

$V_p$ : total volume of  $\beta$ - galactosidase solution used in each reaction in mL

0.099: amount of pNP- $\beta$ Gal in  $\mu$ mol

**EQUATION 1 EQUATION OF BETA-GALACTOSIDASE ACTIVITY; MICROMOLES OF SUBSTRATE HYDROLYZED PER MINUTE/PER MILLILITER OF ENZYME**

However, the average of the calculated activity is 0.075 unit/mL, which is around 100 times weaker than expected. The same reaction has been repeated with more concentrated enzyme and substrate solution. In all cases, similar results have been obtained.

### 4.3. Synthesis of the key-sequence (246-281aa)

**H-G AVFN LTSADHLYVN VSELSLVNFE ESQTFFGLYK L-OH**

**FIGURE 20 KEY SEQUENCE (BLACK); GLYCINE (RED) WILL BE GENERATED AFTER REMOVAL OF THE AUXILIARY FROM THE LIGATION PRODUCT**

The key sequence of this project is a 35 amino acids long peptide, which contains two naturally occurring *N*-glycosylation sites (Figure 20). First, it is crucial to improve the synthesis of the sequence with Fmoc-based SPPS technique in regard of purity and yield, since the following auxiliary attachment and PEGylation reaction have to be performed without purification step in between. In order to obtain the sequence with one and more glycosylated asparagine, the manual synthesis is necessary. Finally, the photocleavable auxiliary will be attached onto the peptide on resin, to support the following Man<sub>3</sub>GlcNAc oxazoline coupling and to mediate glycine generating ligation. Due to the application of highly valuable building blocks (monoglycosylated asparagine) and complex self-synthesized auxiliary, the maximal yield of this sequence is required.

#### 4.3.1. Manual synthesis with TentaGel resin (first approach)

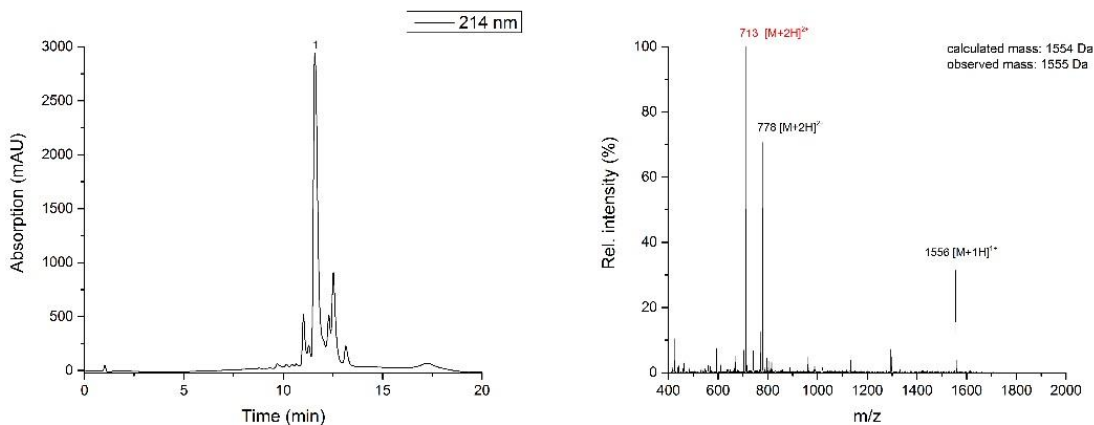
The first approach is the manual synthesis of the key sequence on 0.05 mmol scale TentaGel resin and 2.5 eq Fmoc-amino acids. Fmoc-Val-Ser( $\psi^{\text{Me,Me}}\text{pro}$ )-OH and Fmoc-Glu-Ser( $\psi^{\text{Me,Me}}\text{pro}$ ) pseudo-prolines are used to enhance the stability of the sequence. Besides, several amino acids are double coupled. For a better control of the synthesis, Kaiser tests after each pseudo-proline and test cleavages are performed.

**H-G AVFN LTSADHLYVN VSELSLVNFE ESQTFFGLYK L-OH**

**FIGURE 21 FIRST MANUAL SYNTHESIS OF THE KEY SEQUENCE ON TENTAGEL; PSEUDOPROLINES (UNDERLINED) AND DOUBLE COUPLINGS (GREEN) HAVE BEEN APPLIED**

The first test cleavage is performed after 11 amino acids without Fmoc-deprotection of the last amino acid (Figure 21). As shown below, one can see the lysine deletion product with the mass of 1424 Da ( $\Delta$  -130 Da) together with the desired peptide mass under the product peak (1) at tR = 12 min (Figure 22). Several side products such as dehydrated peptide ( $\Delta$  -18 Da), double dehydrated peptide ( $\Delta$  -36 Da) and trifluoroacetylation product ( $\Delta$  +96 Da) can be observed around the peptide peak (1). In addition, the bad ionization properties and solubility of the

peptide have been noticed. Therefore, a second synthesis of the key sequence has been performed.



**FIGURE 22 LEFT: UV CHROMATOGRAM OF THE 11MER KEY SEQUENCE AT 214 nm; PEPTIDE PEAK (1) AT TR = 12 MIN; RIGHT: MASS SPECTRUM OF THE 11MER KEY SEQUENCE INTEGRATED AT PEPTIDE PEAK (1); CALCULATED MASS: 1554 DA, OBSERVED MASS: 1555 DA; RED: LYSINE DELETION PRODUCT**

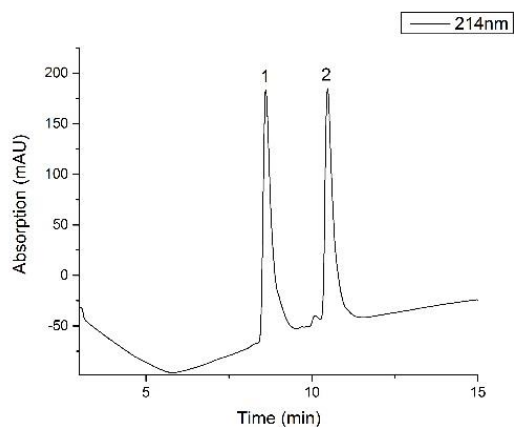
#### 4.3.2. Manual synthesis with TentaGel resin (second approach)

This second approach has been performed under the same conditions as the first one (0.05 mmol scale TentaGel and 2.5 eq Fmoc-amino acids), Fmoc-Val-Ser( $\psi^{\text{Me,Me}}$ pro)-OH and Fmoc-Glu-Ser( $\psi^{\text{Me,Me}}$ pro)-OH have been used, except the additional double coupling of the second amino acid lysine. For a better control of the synthesis, Kaiser tests after each pseudo-proline and test cleavages have been performed (Figure 23).

**H-G AVFN LTSADHLYVN VSELSLVNFE ESQTFFGLYK L-OH**

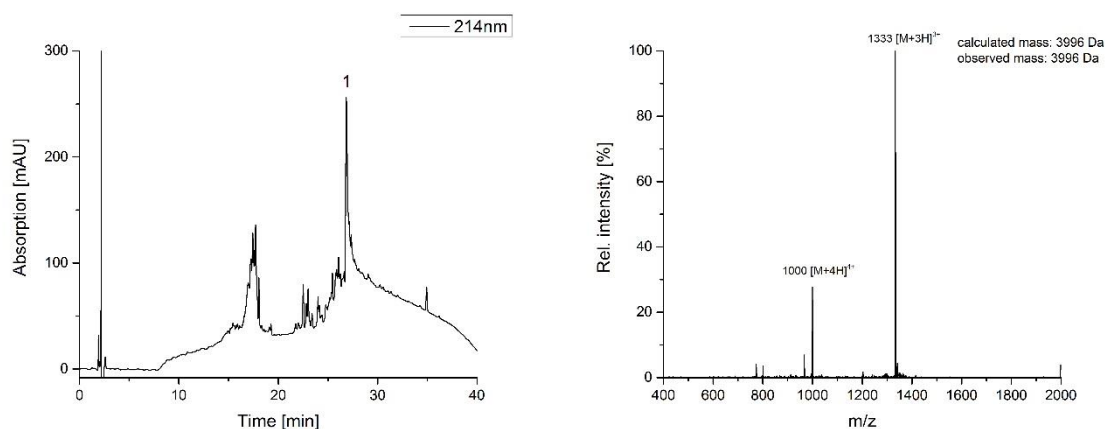
**FIGURE 23 SECOND MANUAL SYNTHESIS OF THE KEY SEQUENCE ON TENTAHEL; PSEUDOPROLINES (UNDERLINED) AND DOUBLE COUPLINGS (GREEN) HAVE BEEN APPLIED**

The first test cleavage has been performed after 9 amino acids without Fmoc-deprotection of the *N*-terminal amino acid. Surprisingly, we have observed the peptide with (peak 2) and without (peak 1) Fmoc protecting group, which might indicate that the Fmoc group is not stable during the cleavage (Figure 24).



**FIGURE 24 UV CHROMATOGRAM OF THE 9MER KEY SEQUENCE WITH (2) AND WITHOUT (1) Fmoc PROTECTING GROUP AT 214nm**

After the complete synthesis, the 35mer has been deprotected and cleaved from the TentaGel resin with the standard cleavage cocktail (Section 3.5). The crude peptide has shown extremely low solubility in 50/50 buffer and analog solutions, therefore it has been dissolved in the strong chaotrope 6 M guanidinium buffer and analyzed with LC-MS (Figure 25). As shown below, the desired product (peak 1) can be observed. However, impurity such as trifluoroacetylation has also been discovered. The peptide tends to aggregate during the analysis, therefore it is challenging to quantify the amount of impurities since a part of the peptide remains on the column. Extra column washing steps are necessary after each analysis run.

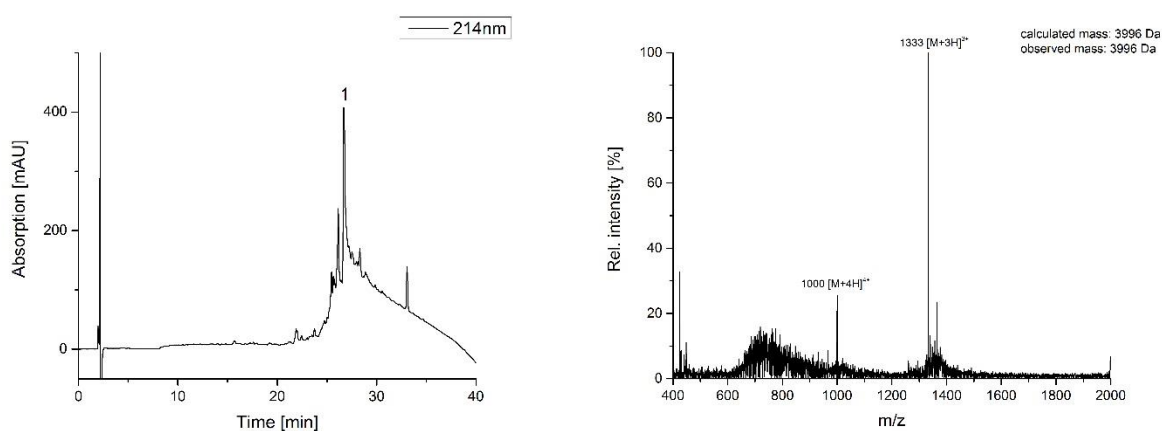


**FIGURE 25 LEFT: UV CHROMATOGRAM OF THE 35MER COMPLETE KEY SEQUENCE AT 214 nm; PEPTIDE PEAK (1) AT TR = 28 MIN; RIGHT: MASS SPECTRUM OF THE 35MER COMPLETE KEY SEQUENCE INTEGRATED AT PEPTIDE PEAK (1); CALCULATED MASS: 3996 DA, OBSERVED MASS: 3996 DA**

To verify that the solubility of the peptide can be improved by PEGylation and that the conjugate can be purified via precipitation by addition of organic solvents, polyethylene glycole<sub>27</sub> (PEG<sub>27</sub>) has been coupled to the key sequence (Section 4.4.1).

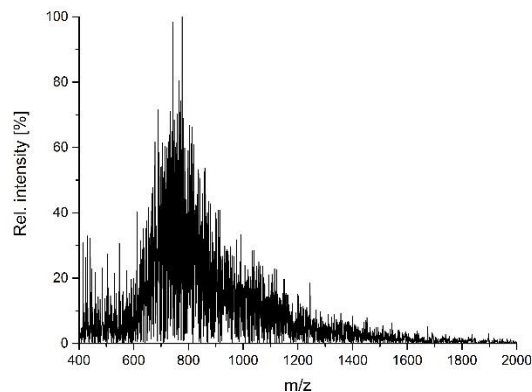
### 4.3.3. Automated Liberty Blue synthesis with TentaGel resin

Since no optimal purity of the key sequence could be achieved yet, automated SPPS using a Liberty Blue microwave synthesizer on TentaGel resin has been performed. On 0.05 mmol scale of resin, 0.1 mmol amino acid, the same pseudoprolines and double coupling positions (Section 4.3.2) are used and the peptide is analyzed after the cleavage. The desired product (1) and some unidentified impurities can be observed (Figure 26). Again, due to strong aggregation tendency, part of the desired peptide remains on the column and the analysis of the synthesized peptide has been challenging.



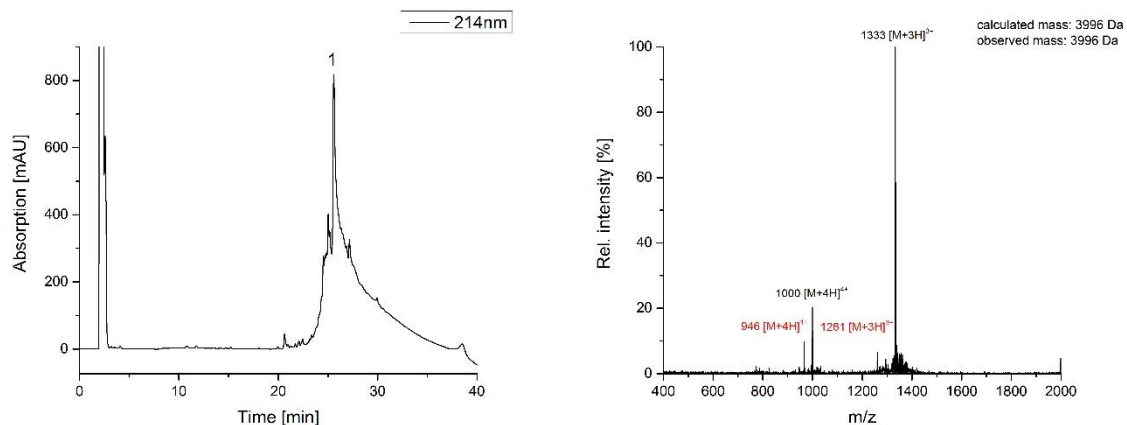
**FIGURE 26 LEFT: UV CHROMATOGRAM OF THE 35MER COMPLETE KEY SEQUENCE AT 214nm; PEPTIDE PEAK (1) AT TR = 28 MIN; RIGHT: MASS SPECTRUM OF THE 35MER COMPLETE KEY SEQUENCE INTEGRATED AT PEPTIDE PEAK (1); CALCULATED MASS 3996 DA, OBSERVED MASS: 3996 DA**

Small masses have been observed in the retention time of 30-40 min, which could be caused by PEG leakage from TentaGel resin. To further investigate this matter, solubility assay with diverse buffers and organic solvents have been performed. In summary it can be stated that both key sequence and resin PEG residues dissolve in 50/50 buffer, ethanol and methanol. While the key sequence is mainly soluble in acetonitrile, 6 M guanidinium at pH 6, 4 M tris buffer at pH 4, the undesired PEG residues are very well soluble in water, isopropanol, urea and phosphate buffer (Figure 27). Interestingly, an UV active impurity has also been found in isopropanol. It is possible that the signal is caused by the 4-hydroxybenzyl alcohol (PHB) Wang linker, which is UV active but too small to be detected by mass. The amount of impurities found in the crude explains the poor solubility. Therefore, it is difficult to draw a clear conclusion on whether the synthesis itself is successful or not, since PEG residues ionize better and thus may suppress the ionization of possible side products.



**FIGURE 27 MASS SPECTRUM OF THE EXTRACTED PEG RESIDUES OF THE CRUDE PEPTIDE WITH H<sub>2</sub>O, DUE TO PEG LEAKAGE OF TENTA GEL RESIN DURING SPPS**

To avoid this complication, the discovered impurities within the crude have been extracted with water and isopropanol. The remaining crude peptide has been dissolved in guanidinium buffer and analyzed (Figure 28).



**FIGURE 28 LEFT: UV CHROMATOGRAM OF THE 35MER COMPLETE KEY SEQUENCE AFTER EXTRACTION; PEPTIDE PEAK (1) RIGHT: MASS SPECTRUM OF THE 35MER COMPLETE KEY SEQUENCE INTEGRATE AT PEAK PEAK (1); CALCULATED MASS 3996 DA, OBSERVED MASS: 3996 DA; RED: GLU-SER( $\Psi^{Me,Me}PRO$ )-OH DELETION PRODUCT**

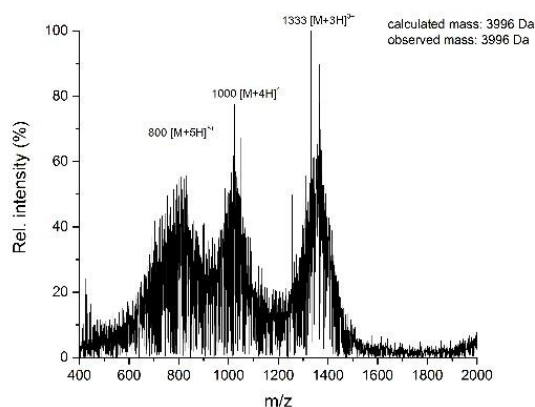
In comparison to the analysis before (Figure 25, Figure 26), the UV chromatogram (Figure 28) shows a much sharper peptide peak with less impurity peaks. However, the optimal peak width has still not been achieved. The peptide does not remain on the column at all after the analysis and the mass spectrum showed the Glu-Ser( $\Psi^{Me,Me}pro$ )-OH deletion product ( $\Delta$  -216 Da), which has not been detected before due to ionization repression caused by PEG impurities. However, the solubility does not enhance significantly.

To test the optimal conditions for the following auxiliary attachment reaction and also to improve the character of the peptide, such as solubility and aggregation, PEG<sub>27</sub> has been coupled to the key sequence (Section 4.4.1).

#### 4.3.4. Automated Liberty Blue synthesis with ChemMatrix resin

To improve the synthesis of the key sequence in regard of purity and yield, ChemMatrix- HPMB resin has been applied. In comparison to TentaGel resin, which contains polymerized polystyrene with grafted PEG, ChemMatrix is a totally PEG-based resin. Due to the primary ether bonds and its high cross-linked matrix, it has been proved to possess a higher chemical and mechanical stability. Moreover, it has been reported to support the synthesis of bigger aggregating peptides, which are more difficult to synthesize on PEG-PS resins. The same reacting conditions (Section 4.3.3) have been applied for the automated Liberty Blue synthesis.

Unfortunately, the use of this resin has caused a very high degree of side product (Figure 29). It has been assumed that the C-terminal PEG has not been cleaved correctly, since side product masses corresponding to peptide carrying different length of PEG fragments have been observed.

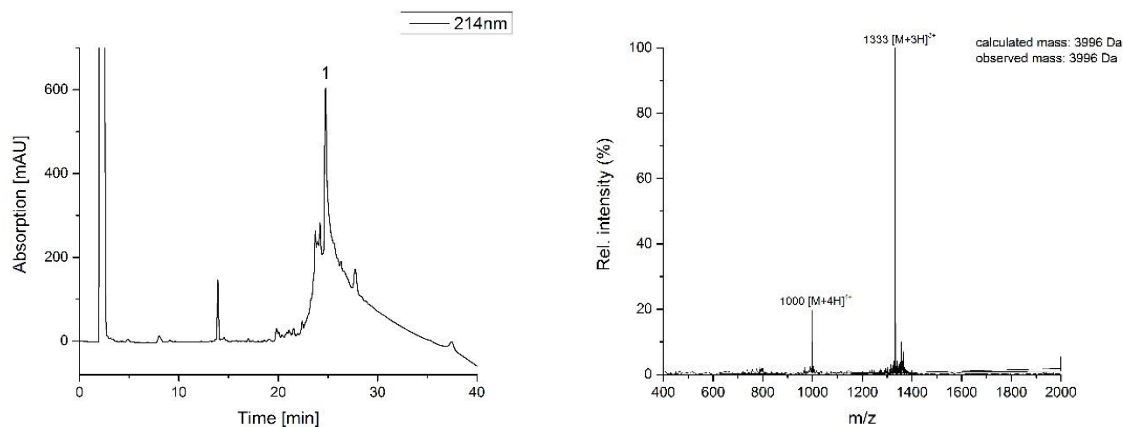


**FIGURE 29 MASS SPECTRUM OF THE 35MER COMPLETE KEY SEQUENCE SYNTHESIZED VIA CHEMMATRIX RESIN**

#### 4.3.5. Automated Liberty Blue synthesis with Wang resin

Despite the excellent swelling property and high flexibility of the high loading TentaGel resin, the PEG free and polystyrene-based solid support low loading Wang resin has been chosen to avoid any further resin leakage problems. The same reaction conditions (Section 4.3.3) have been used and as shown below, the desired product has been synthesized. Overall, the use of Wang resin seems to be most promising (Figure 30), despite small impurities cause by trifluoroacetylation ( $\Delta +96$  Da,  $\Delta +192$  Da etc.). The analysis remains challenging, due to the low solubility of the peptide and the broad peak during the analysis run. However, the strong aggregation behavior observed previously, such as peptide remaining on the column after analysis run, does not occur anymore.





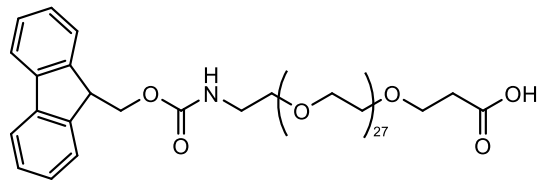
**FIGURE 30 LEFT: UV CHROMATOGRAM OF THE 35MER COMPLETE KEY SEQUENCE AFTER EXTRACTION; PEPTIDE PEAK (1) AT  $t_R = 28$  MIN; RIGHT: MASS SPECTRUM OF THE 35MER KEY SEQUENCE INTEGRATE AT PEPTIDE PEAK (1); CALCULATED MASS 3996 DA, OBSERVED MASS 3996 DA**

Overall it can be said that the SPPS method of the key sequence gives acceptable results. However, for the following reactions such as auxiliary attachment and precipitation assays, the quality of the synthesis needs to be enhanced. Especially because those critical reactions will be done while the peptide is still on the solid support, it is essential to achieve the best possible purity to avoid side reactions. It would also be beneficial to increase the solubility, improve the ionization and prevent the tendency of the peptide to remain on the column, so that more conclusive analysis can be performed.

#### 4.4. Preliminary solubility/ precipitation assays

The attachment of the photocleavable multifunctional auxiliary to the key sequence should enhance the enzymatic  $\text{Man}_3\text{GlcNAc}$  oxazoline coupling by improving the solubility of the peptide and enabling the more convenient purification via precipitation. Thereafter, the auxiliary will be used for the ligation step and generates the abundant natural amino acid glycine after photocleavage.

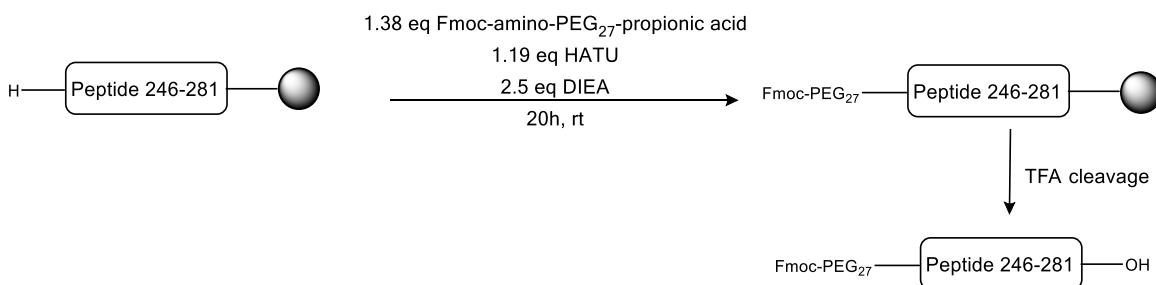
Since the attachment has to be done while the peptide is still on its solid support, an optimized peptide synthesis is necessary. To perform solubility/precipitation assays, a peptide conjugate carrying the PEG polymer (Figure 31) directly at the N-terminus of the peptide has been prepared. This will give us the possibility to check and eventually optimize the conditions for the glycosylation/precipitation procedure that will be subsequently applied in the reaction with the more precious peptide-auxiliary conjugate. Indeed, it has been previously shown using MUC1 peptide that such a conjugate can be efficiently used in the glycosylation/precipitation procedure.



**FIGURE 31 FMOC-AMINO-PEG<sub>27</sub>-PROPIONIC ACID**

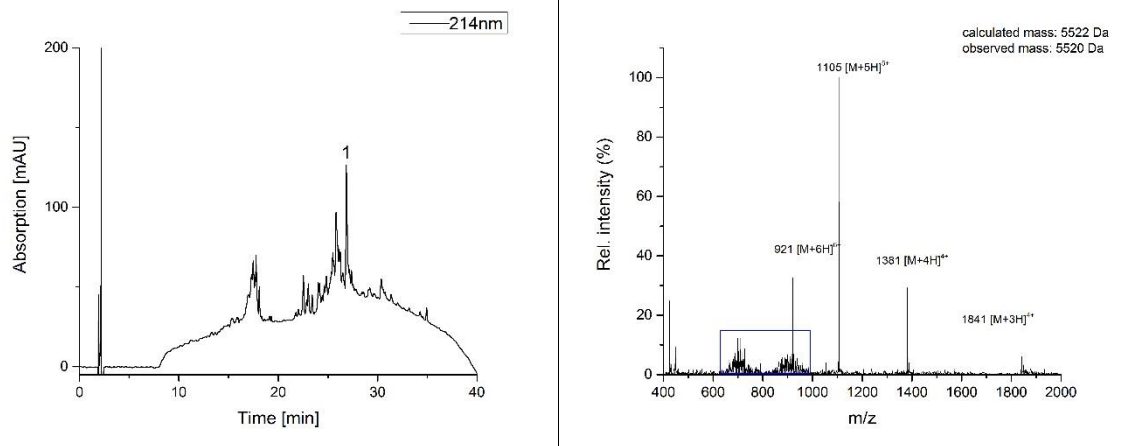
#### 4.4.1. PEGylation of the key-sequence

For the first PEGylation the loaded resin from the second approach of manual synthesis (Section 4.3.2) has been used. After applying the standard procedure (Figure 32) for the PEG coupling, the PEGylated peptide has been cleaved from TentaGel resin and analyzed with LC-MS.



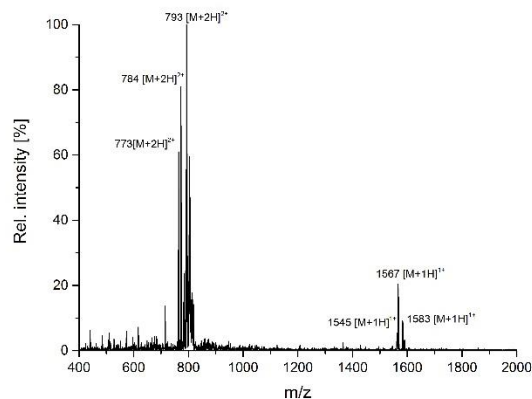
**FIGURE 32 PEGYLATION REACTION OF THE KEY SEQUENCE: AFTER THE PEGYLATION OVERNIGHT AT ROOM TEMPERATURE, THE PEGYLATED PEPTIDE IS REMOVED FROM THE RESIN VIA ACIDIC CLEAVAGE**

As shown below (Figure 33), the desired PEGylated peptide has been obtained. Surprisingly, we observed PEG impurities which does not belong to the PEG<sub>27</sub> reagent (blue square). We suspect that during the PEGylation and cleavage, some of the PEG residue of the solid support TentaGel has also been removed. This and the other impurities from the peptide synthesis itself lead to no precise conclusion in regard of solubility or ionization.

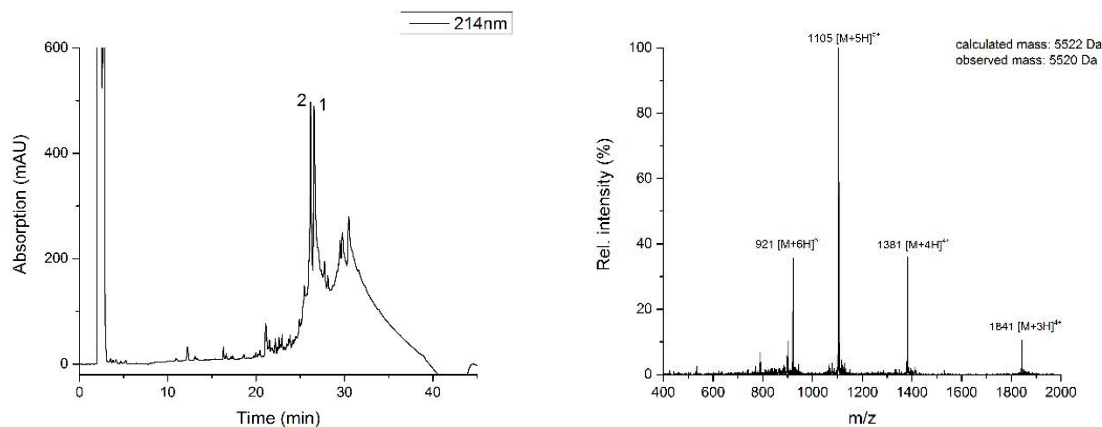


**FIGURE 33 LEFT: UV CHROMATOGRAM OF THE PEGYLATED 35MER KEY SEQUENCE; PEPTIDE PEAK (1) RIGHT: MASS SPECTRUM OF THE PEGYLATED 35MER KEY SEQUENCE AT PEPTIDE PEAK (1); PEG IMPURITIES (BLUE SQUARE); CALCULATED MASS: 5522 DA, OBSERVED MASS: 5520 DA**

The second assay has been performed on the loaded TentaGel resin of automated Liberty Blue synthesis (Section 4.3.3). As shown below at peak 1 (Figure 35) the desired PEGylated peptide has been obtained. A rather big amount of unreacted PEG<sub>27</sub> can be found under peak 2 (Figure 34), which indicates an incomplete reaction or an impure key peptide. Similar to the UV chromatogram of the key sequence before the PEGylation (Figure 26), some impurity peaks can be observed. In comparison, the impurities seem to be less than before. However, depending on which solvent is used for the analysis, the signals vary strongly due to PEG leakage from TentaGel resin. Therefore, no conclusion can be drawn after the PEGylation.

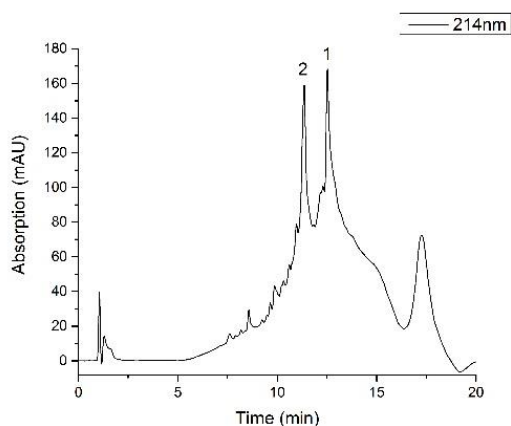


**FIGURE 34 MASS SPECTRUM OF THE UNREACTED PEG<sub>27</sub>- RESIDUE, INTEGRATED AT TR = 25MIN**



**FIGURE 35 LEFT: UV CHROMATOGRAM OF THE PEGYLATED 35MER KEY SEQUENCE; PEPTIDE PEAK (1) AT  $t_R = 26$  MIN RIGHT: MASS SPECTRUM OF THE PEGYLATED 35MER KEY SEQUENCE AT PEPTIDE PEAK (1); CALCULATED MASS: 5522 Da, OBSERVED MASS: 5520 Da**

After using Wang resin, the PEG impurity problem has been finally solved. The third PEGylating assay was done on the loaded Wang resin of automated Liberty Blue synthesis (Section 4.3.5). As shown below (Figure 36) the desired PEGylated peptide has been obtained (peak 1). However, the ionization of the PEGylated peptide has not been enhanced and the low solubility seems to remain the same. In addition, the PEGylation apparently causes the peptide to remain on the column, which makes the analysis and following purification step much more difficult. Under peak 2, the not PEGylated trifluoroacetylated key sequence has been found. Besides, several unidentified small impurities have also been detected, which have not been observed before the PEGylation.



**FIGURE 36 UV CHROMATOGRAM OF THE PEGYLATED 35MER KEY SEQUENCE, PEPTIDE PEAK AT  $t_R = 13$  MIN**

Therefore, for the further investigation of PEGylation reactions with the key sequence, not only the synthesis itself should be further modified to give a more soluble and less aggregating key sequence, but also the PEGylation procedure should be optimized to avoid any possible side

reactions. For a better analysis, different gradient and solvent mixtures could be tested to prevent aggregation behaviors on the column.

## 4.5. Synthesis of the remaining sequences (141-201aa; 202-245aa)

The remaining sequence 141-245 should be synthesized with Liberty blue in two fragments (141-201aa; 202-245aa) and coupled with each other via hydrazide ligation. It has been reported that the hydrazide functionality can be obtained via hydrazinolysis, which has proved to be an easy and convenient method<sup>[55]</sup>. Depending on the C-terminal amino acid and the solubility behavior of the peptide chain itself, yield up to 98 % can be obtained. Therefore, after the successful synthesis of both sequences, the peptidyl resin has been incubated with a solution of hydrazine in THF which leads to its hydrazinolysis (Section 4.6.1, Section 4.6.3).

### 4.5.1. Synthesis of the sequence 141-201aa

The synthesis of the *N*-terminal peptide seems to present a challenge due to its long sequence (61aa), oxidation tendency caused by the presence of several methionine residues and bad ionization character. However, in comparison to the key sequence, the synthesis itself has proved to be less problematic than at the first glance, nevertheless, during the purification step after the hydrazineolysis, the low solubility and aggregation tendency of the peptide have caused several complications (Section 4.6.1).

First, the *N*-terminal sequence is synthesized on Liberty Blue synthesizer with 0.025 mmol scale of TentaGel resin. Five different pseudoprolines have been used to give the synthesis more stability. All of the amino acids are double coupled to prevent possible deletion products and to give an overall good synthesis (Figure 37). For a better control, three test cleavages have been performed.

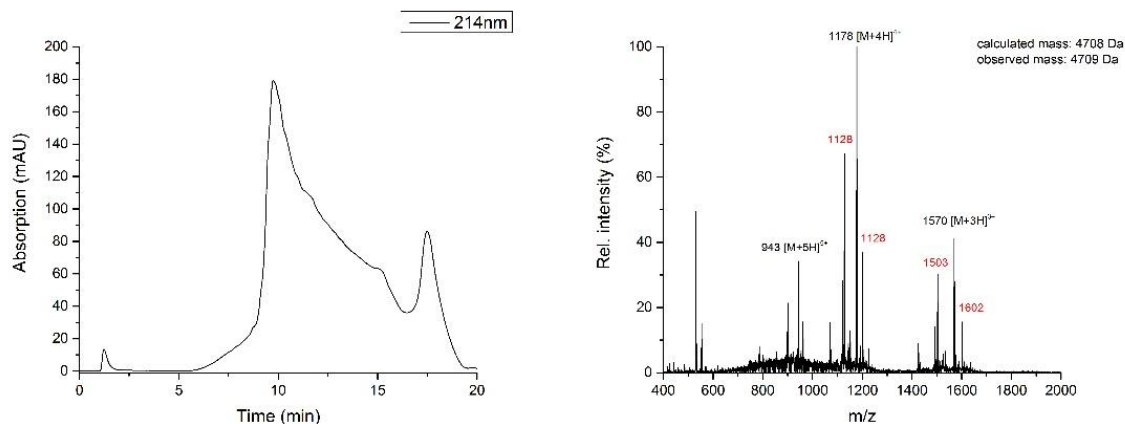


H-KELRKVAHLT GKSNSRSMPL EWEDTYGIVL  
LSGVKYKKGG LVINETGLYF VYSKVYFRGQ S-OH

**FIGURE 37** FIRST AUTOMATED LIBERTY BLUE SYNTHESIS OF THE *N*-TERMINAL SEQUENCE ON TENTA GEL RESIN; PSEUDOPROLINES (UNDERLINED), ALL AA HAVE BEEN DOUBLE COUPLED

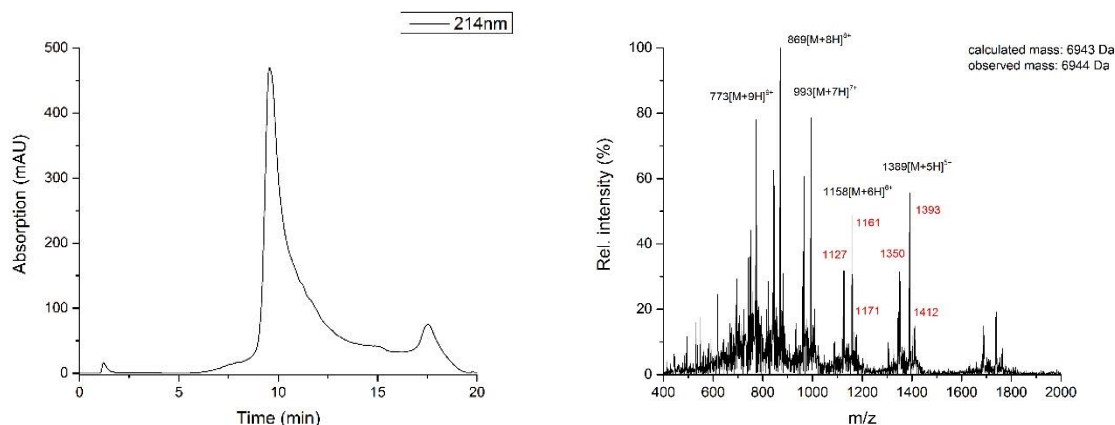
Already after the second test cleavage (1-41aa), several impurities can be found besides the desired product (Figure 38), such as the trifluoroacetylation ( $\Delta$  +96 Da) and the Leu-

Ser( $\psi^{\text{Me,Me}}\text{pro}$ )-OH deletion product ( $\Delta$  -200 Da). It can be assumed that the pseudoproline does not couple well at that specific position. This issue can be prevented by using normal amino acids instead. In addition, it has been observed that the peptide tends to remain on the column and gives a very broad UV chromatogram peak (over 10 min). It is possible that the broad peak overlaps with the peaks of other impurities present in the crude, which explains the observed impurities under the peptide peak.



**FIGURE 38 LEFT: UV CHROMATOGRAM OF THE 41MER *N*-TERMINAL SEQUENCE; PEPTIDE PEAK AT TR = 10 MIN; RIGHT: MASS SPECTRUM OF THE 41MER *N*-TERMINAL SEQUENCE INTEGRATED AT PEPTIDE PEAK; CALCULATED MASS: 4708 DA, OBSERVED MASS: 4709 DA; RED: LEU-SER( $\psi^{\text{ME,ME}}\text{PRO}$ )-OH DELETION PRODUCT**

Almost the same aggregation properties have been noticed after the final test cleavage of the peptide (Figure 39). Also, due to the bad ionization character of the peptide, the mass spectrum appears messy. In addition, the methionine oxidation side product can be observed alone ( $\Delta$  +18 Da) and with trifluoroacetylation ( $\Delta$  +113 Da). To solve the problem of methionine oxidation, one possibility is the use of reducing cleavage cocktail by Hackenberger (cocktail H) [56].



**FIGURE 39 LEFT: UV CHROMATOGRAM OF THE 61MER *N*-TERMINAL SEQUENCE; PEPTIDE PEAK AT TR = 10 MIN; RIGHT: MASS SPECTRUM OF THE 61MER *N*-TERMINAL SEQUENCE INTEGRATED AT PEPTIDE PEAK; CALCULATED MASS: 6943 DA, OBSERVED MASS: 6944 DA; RED: TRIFLUOROACETYLATION AND OXIDATION IMPURITIES**

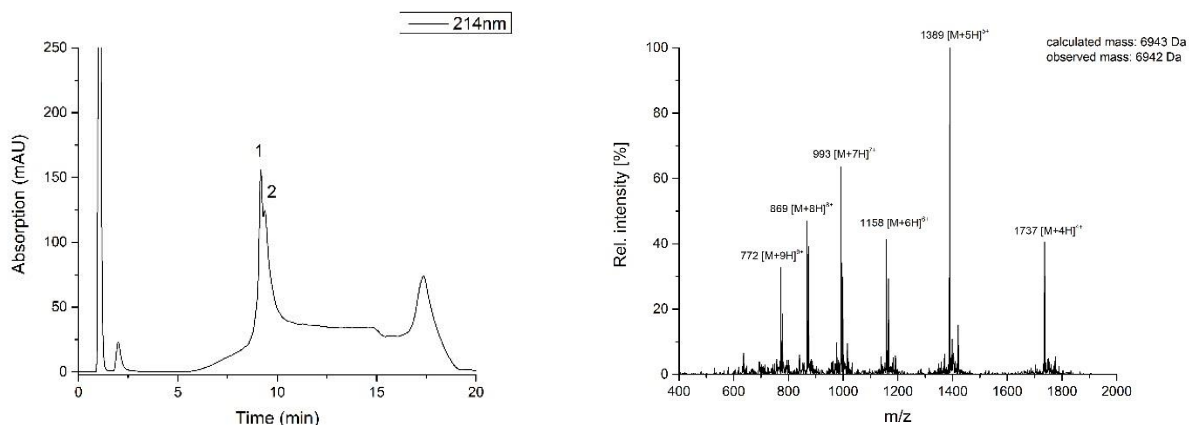
Several hydrazine cleavage assays have been performed with this peptide (Section 4.6.1). However, due to the messy spectrum and the numerous side products, no conclusive results can be obtained. Therefore, the synthesis condition has been further optimized.

The second approach is similar to the first one in regard of scale and double coupling, the only difference is that normal lysine and serine have been used instead of pseudoproline Leu-Ser( $\psi^{\text{Me,Me}}\text{pro}$ )-OH.

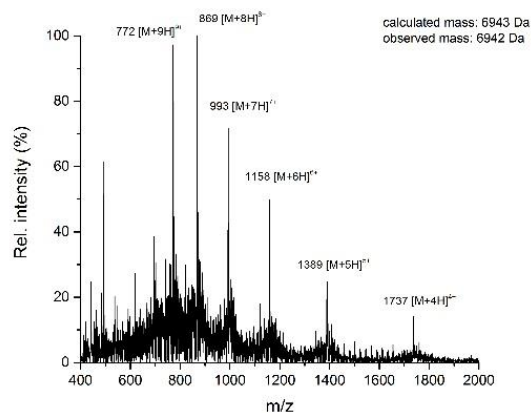
**H-KELRKVAHLT GKSNSRSMPL EWEDTYGIVL  
LSGVKYKGG LVINETGLYF VYSKVYFRGQ S-OH**

**FIGURE 40 SECOND AUTOMATED LIBERTY BLUE SYNTHESIS OF THE N-TERMINAL SEQUENCE ON TENTAGEL RESIN; PSEUDOPROLINES (UNDERLINED), ALL AA HAVE BEEN DOUBLE COUPLED**

After the final test cleavage, the direct injection analysis gives a rather clean spectrum (Figure 41). The HPLC-MS spectrum shows the desired product and the protected pseudoproline ( $\Delta +40$  Da) under peak 1 and the trifluoroacetylation ( $\Delta +96$  Da) under peak 2. The incomplete pseudoproline deprotection can be solved with longer incubation time during cleavage. Since both peaks can be well separated and synthesis is successful, the following hydrazine cleavage investigation has been performed. Traces of the missing pseudoproline Glu-Thr( $\psi^{\text{Me,Me}}\text{pro}$ )-OH ( $\Delta -230$  Da) have been observed.



**FIGURE 41 LEFT: UV CHROMATOGRAM OF THE 61MER N-TERMINAL SEQUENCE, PEPTIDE PEAK (1) AT tR = 9 MIN; RIGHT: MASS SPECTRUM OF THE 61MER N-TERMINAL SEQUENCE INTEGRATED AT PEPTIDE PEAK (1); CALCULATED MASS: 6943 DA, OBSERVED MASS: 6942 DA**

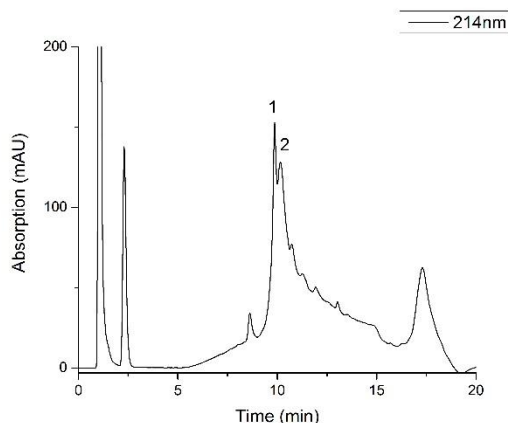


**FIGURE 42 DIRECT INJECTION MASS SPECTRUM OF THE 61MER N-TERMINAL SEQUENCE**

The last synthesis is performed with 0.05 mmol scale of resin, three different pseudoprolines and several double couplings (Figure 43). It is possible that the pseudoproline Glu-Thr( $\psi^{\text{Me,Me}}\text{pro}$ )-OH does not couple well due to the short distance to the nearby pseudoproline Tyr-Ser( $\psi^{\text{Me,Me}}\text{pro}$ )-OH. Therefore, normal amino acids have been used instead. Also, to further investigate the quality difference of the synthesis, only some amino acids at certain critical positions are double coupled this time.

H-KELRKVAHLT GKSNSRSMPL EWEDTYGIVL  
 LSGVKYKKGG LVINETGLYF VYSKVYFRGQ S-OH

**FIGURE 43 THIRD AUTOMATED LIBERTY BLUE SYNTHESIS OF THE N-TERMINAL SEQUENCE ON TENTA GEL RESIN; PSEUDOPROLINES (UNDERLINED) AND DOUBLE COUPLING (GREEN) HAVE BEEN APPLIED**



**FIGURE 44 UV CHROMATOGRAM OF THE 61MER N-TERMINAL SEQUENCE, PEPTIDE PEAK (1) AT TR = 10 MIN**

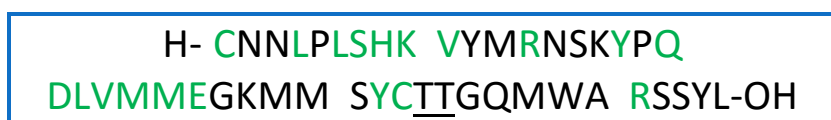
Although there is no apparent local hindrance, the pseudoproline Glu-Thr( $\psi^{\text{Me,Me}}\text{pro}$ )-OH is still not well coupled. This can be observed after the first test cleavage. Protecting group of the



pseudoproline Glu-Thr( $\psi^{\text{Me,Me}}\text{pro}$ )-OH does not come off after 2 hours of cleavage time. The according protected pseudoproline mass ( $\Delta +40$  Da) and the desired peptide mass are both found under peak 1 (Figure 44). Even after extending the cleavage time to almost 3 hours, the protecting group can still not be completely removed. However, this issue luckily does not occur during the side chain deprotection step following the hydrazine cleavage. It is possible that, due to the long sequence and the lower percentage of TFA within the cleavage cocktail H, the deprotection step needs longer to complete. Under peak 2, trifluoroacetylation with ( $\Delta +141$  Da) and without ( $\Delta +96$  Da) protected pseudoproline has been observed. This can possibly be reduced by using Boc-protected lysine as the last amino acid at the end of peptide sequence. However, this method cannot prevent the trifluoroacetylation as an esterification on the side chain. Some other protected peptides can be observed between retention times 11-13 min. Although not all of the amino acids are double coupled, the quality of the synthesis has remained the same. This indicates that despite the fact that the sequence is long, not all amino acids need to be double coupled for the SPPS. After establishing this optimized synthesis, the following hydrazine cleavage of the peptide has been performed (Section 4.6.1).

#### 4.5.2. Synthesis of the sequence 202-245aa

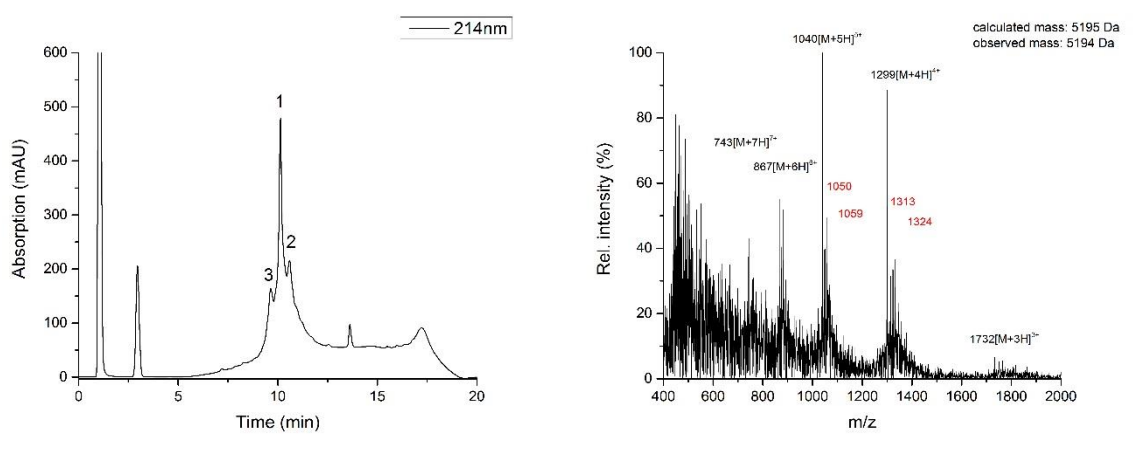
The methionine rich middle fragment of the sFas ligand should not be challenging to synthesize. However, the reducing cleavage condition is necessary to avoid any possible oxidations. Due to this reason, contact with air should be limited for the loaded resin. The automated Liberty Blue synthesis has been performed on 0.025 mmol scale. One pseudoproline and several double coupling have been applied (Figure 45).



**FIGURE 45 FIRST AUTOMATED LIBERTY BLUE SYNTHESIS OF THE MIDDLE SEQUENCE WITH TENTAGEL RESIN; PSEUDOPROLINES (UNDERLINED) AND DOUBLE COUPLING (GREEN) HAVE BEEN APPLIED**

The desired product can be observed in high quality under the peak 1 (Figure 46). Trifluoroacetylation impurity ( $\Delta +96$  Da; peak 2) and peptide with tert-butyl protecting group ( $\Delta +57$  Da; peak 3) can also be found. All of the impurities can be also shown in the direct injection measurement. Similar as described before at the synthesis of the *N*-terminal peptide sequence, the pseudoproline protecting group does not completely come off after two hours cleavage reaction. Due to the high quality of the peptide and the time limitation, no further synthesis of this sequence has been performed. The following hydrazine cleavage and purification gives

successful results without any complication (Section 4.6.3). For the future synthesis, it can be considered using Boc instead of Fmoc protected *N*-terminal cysteine to reduce trifluoroacetylation reaction. Again, extending the cleavage time should give the completely deprotected product.



**FIGURE 46 LEFT: UV CHROMATOGRAM OF THE 44MER MIDDLE SEQUENCE, PEPTIDE PEAK (1) AT TR = 10 MIN; RIGHT: DIRECT INJECTION MEASUREMENT; MASS SPECTRUM OF THE 44MER MIDDLE SEQUENCE; CALCULATED MASS: 5195 DA, OBSERVED MASS: 5194 DA; RED: TRIFLUOROACETYLATION IMPURITY AND PEPTIDE WITH TERT-BUTYL PROTECTING GROUP**

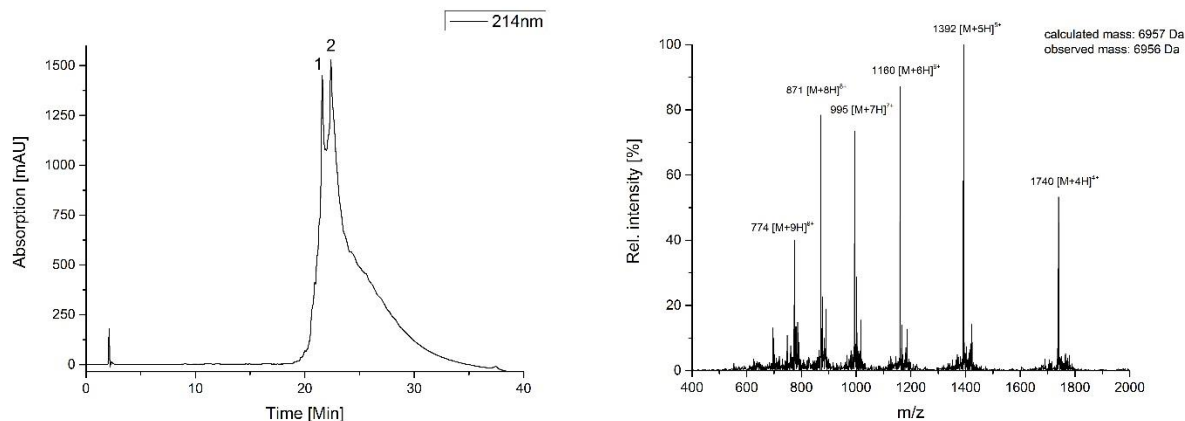
## 4.6. Ligations of the *N*-terminal and middle sequences

After the successful synthesis of the *N*-terminal and middle sequences, the hydrazine cleavage has been applied to both peptidyl resins for further ligation steps. As described in the aim (page 17), first the *N*-terminal and middle sequences will be ligated via hydrazine ligation, then the key sequence will be coupled via auxiliary mediated ligation. In both cases, the formation of hydrazide peptide and subsequently the thioester is essential.

### 4.6.1. Hydrazine cleavage of sequence 141-201aa

Hydrazine cleavages performed on Wang-type and TentaGel resin have been reported to give hydrazide peptides with high efficiency and good quality [55]. Several different *C*-terminal amino acids have been tested to verify the universal application of this method. However, no peptide longer than 20 amino acids has been converted in the past. For further investigation of this convenient technique and its size limit, the hydrazine cleavage has been performed on the 61mer peptide.

Different methods (hydrazine in THF and hydrazine hydrate in MeOH) and various hydrazine concentrations and reaction times have been tested. While the variation of reaction time does not strongly influence the reaction, the use of hydrazine in THF works significantly better than hydrazine hydrate in different organic solvents. The cleaved and protected hydrazide peptide however, does not dissolve in THF and precipitates on the resin. The length of the peptide and the presence of many hydrophobic protecting groups are responsible for this behavior. After drying the resin, cleavage cocktail H has been applied to give the desired product (Figure 47).



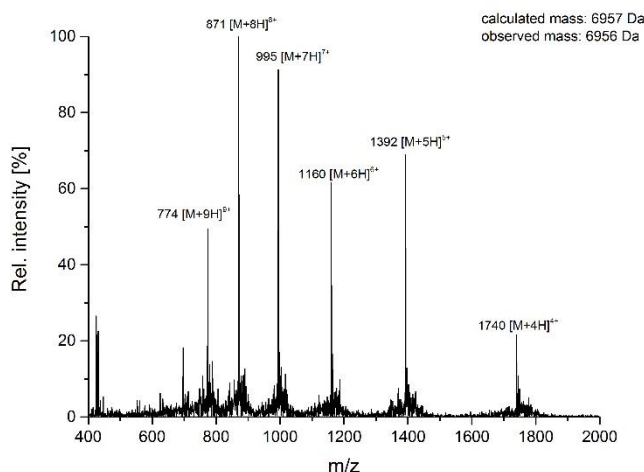
**FIGURE 47 LEFT: UV CHROMATOGRAM OF THE 61MER N-TERMINAL HYDRAZIDE PEPTIDE, PEPTIDE PEAK (1) AT TR = 20 MIN ; RIGHT: MASS SPECTRUM OF THE 61MER N-TERMINAL HYDRAZIDE PEPTIDE INTEGRATED AT PEPTIDE PEAK (1); CALCULATED MASS: 6957 DA, OBSERVED MASS: 6958 DA**

To prove that the hydrazineolysis indeed occurs, thus the hydrazide functionality is on the C-terminal position and not on the side chain of an aspartic or glutamic acid (the two residues are protected as t-butyl esters, they can undergo ester-amide exchange), the fully protected hydrazide peptide has been isolated from TentaGel resin after applying the hydrazine cleavage overnight. Among several solvents, trifluoroethanol has been proven to be able to dissolve the protected peptide. While the dried extraction residue gives the desired product, the resin does not contain any acid peptide or hydrazide peptide at all. This has been proven by the subsequent acidic cleavage, at which no protein product at all can be obtained. Thus, it shows that the reaction conversion is complete and the correct hydrazide peptide has been obtained. MS-MS analysis (Table 5) has also been performed and further verifies this matter.

Sequences	Fragments found	Modification
VYFRGQS	32	C-terminal hydrazide
GGLVINETGLYFVYSKVYFRGQS	2	C-terminal acid
SMPLEWEDTYGIVLLSGVKYKK	5	Sidechain monohydrazide
GGLVINETGLYFVYSK	1	Sidechain monohydrazide
KGGLVINETGLYFVYSK	2	Sidechain monohydrazide

**TABLE 5 MS-MS ANALYSIS OF THE HYDRAZIDE PEPTIDE, ONLY MONOHYDRAZIDE AND ACID PEPTIDES HAVE BEEN OBSERVED; RED: POSSIBLE HYDRAZIDE POSITION**

These results clearly prove that the hydrazine cleavage method is not necessarily limited by size of the peptide. With the right solvent or solvent mixture, it is even possible to isolate the cleaved hydrazide peptide from the resin. Without any kind of resin modification, it facilitates the synthesis of hydrazide peptides and hence the hydrazine ligation enormously. Also, the double hydrazide peptide has not been observed with Waters HPLC-MS and the MS-MS data confirms the presence of side-chain hydrazide modification only in amounts lower than 15-20% (Table 5).

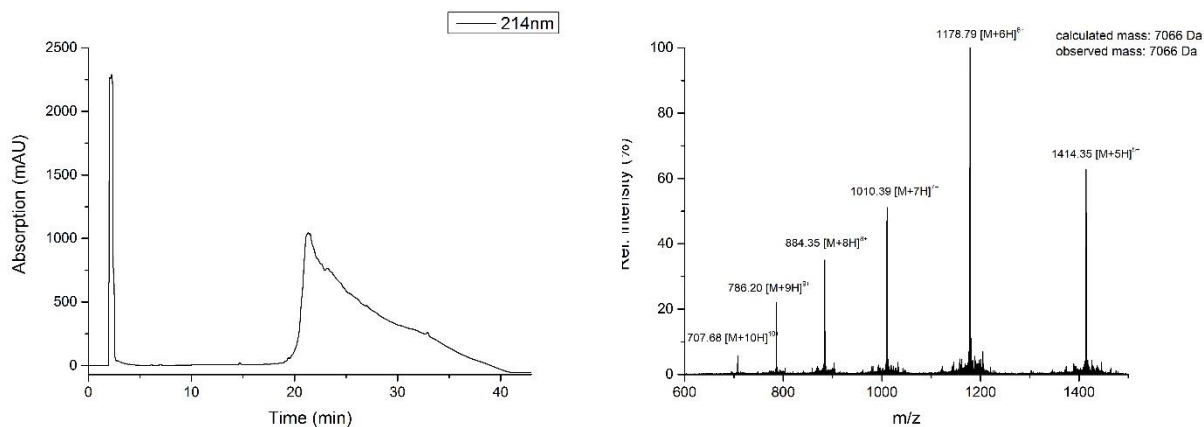


**FIGURE 48 DIRECT INJECTION: MASS SPECTRUM OF THE PURIFIED *N*-TERMINAL HYDRAZIDE PEPTIDE; CALCULATED MASS: 6957 DA, OBSERVED MASS: 6958 DA**

The purification of the hydrazide peptide however, turns out to be very challenging due to the low ionization character and the hydrophobicity of the peptide. Due to possible hydrazine salts and the high salt concentration of the cleavage cocktail H, it has not been possible to estimate the real amount of peptide in the crude product. The difficult detection and the aggregation behavior have led to great yield loss, especially by scaling up the reaction. While only 3 % of purified hydrazide peptide (Figure 48) have been obtained with 600 mg of loaded resin, 13 % yield have been achieved with 30 mg. For this reason, more investigation should be done in enhancing the purification and analysis. Obviously, optimization of the peptide synthesis, e.g. preventing TFA adduct formation by using Boc protected *N*-terminal amino acid, will also increase the yield.

#### 4.6.2. Thioester preparation of sequence 141-201aa

A modified procedure from Bello et al. has been applied for the thioester preparation<sup>[55]</sup>. Different reaction times have been tested and after 15 min the reaction has been proved to be complete.

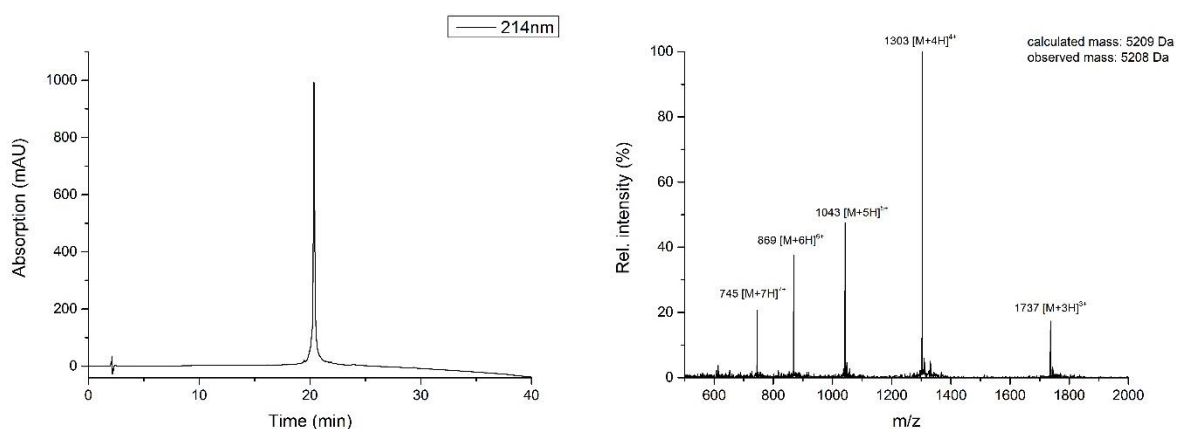


**FIGURE 49 LEFT: UV CHROMATOGRAM OF THE 61MER *N*-TERMINAL THIOESTER PEPTIDE, PEPTIDE PEAK AT TR = 22 MIN; RIGHT: MASS SPECTRUM OF THE *N*-TERMINAL THIOESTER PEPTIDE INTEGRATED AT PEPTIDE PEAK; CALCULATED MASS: 7066 DA, OBSERVED MASS: 7066 DA**

It has been noticed that under acidic conditions, the thioester peptide tends to hydrolyze. Similar to the hydrazide peptide, the thioester product remains on the column and gives very broad UV peaks, which complicate the purification step considerably.

#### 4.6.3. Hydrazine cleavage of sequence 202-245aa

The hydrazine cleavage of the middle sequence has been performed the same way as before (Section 4.6.1). A solution of hydrazine in THF has been used and 5 % purified product have been obtained.

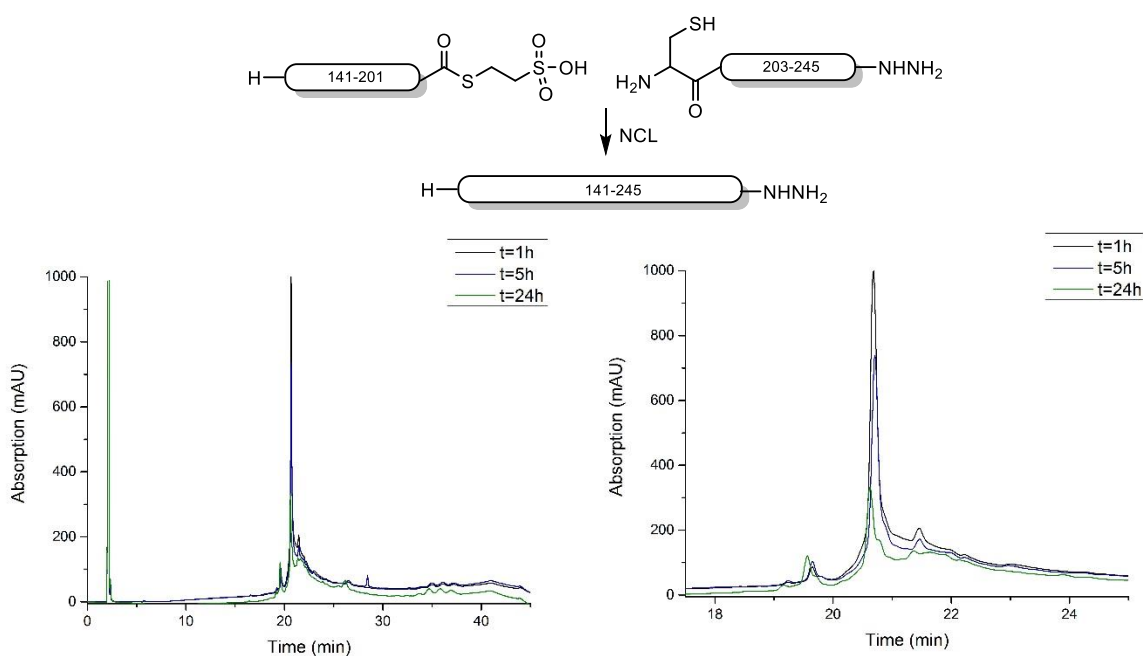


**FIGURE 50 LEFT: UV CHROMATOGRAM OF THE 44MER MIDDLE SEQUENCE HYDRAZIDE PEPTIDE, PEPTIDE PEAK AT TR = 20 MIN; RIGHT: MASS SPECTRUM OF THE MIDDLE SEQUENCE HYDRAZIDE PEPTIDE INTEGRATED AT PEPTIDE PEAK; CALCULATED MASS: 5209 DA, OBSERVED MASS: 5208 DA**

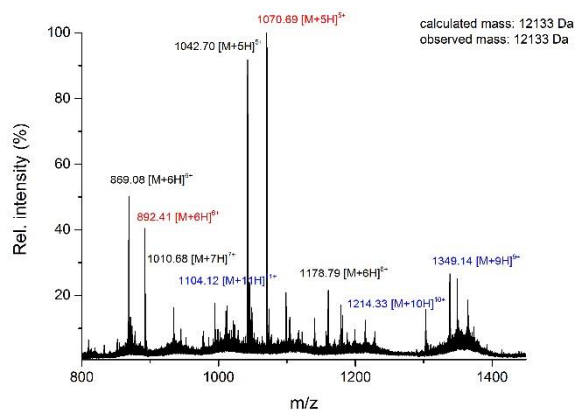
Again, the fully protected hydrazide peptide can be isolated with TFE. As shown in the clean UV chromatogram and mass spectrum, the product has been obtained in high quality after the purification. In comparison to the previous purification, this peptide is easy to handle with due to the high ionization character.

#### 4.6.4. Ligation assays for 141-245aa-Hydrazine

The modified procedure from Bello et al. has been applied for the ligation assays. Different equivalences of peptides have been tested and the best result has been obtained with 1 eq of 141-201aa thioester and 1.1 eq of 202-245aa hydrazide peptide so far (Figure 51).



**FIGURE 51 TOP: REACTION MECHANISM OF THE LIGATION; LEFT: UV CHROMATOGRAMS OF THE LIGATION REACTION AT T=1H, 5H AND 24H; RIGHT: ZOOM OF THE CHROMATOGRAM AT PEAKS OF INTEREST (202-245AA HYDRAZIDE PEPTIDE AT TR = 20.8 MIN, 141-201AA THIOESTER PEPTIDE AT TR = 21.5 MIN)**



**FIGURE 52 MASS SPECTRUM OF THE LIGATION REACTION AFTER 5 HOURS, INTEGRATED OVER PEPTIDE PEAKS; BLACK: THIOESTER PEPTIDE AND HYDRAZIDE PEPTIDE; RED: HYDRAZIDE PEPTIDE WITH MESNA DISULFIDE; BLUE: LIGATION TARGET PROTEIN FORMATION; CALCULATED MASS: 12133 Da, OBSERVED MASS 12133 Da**

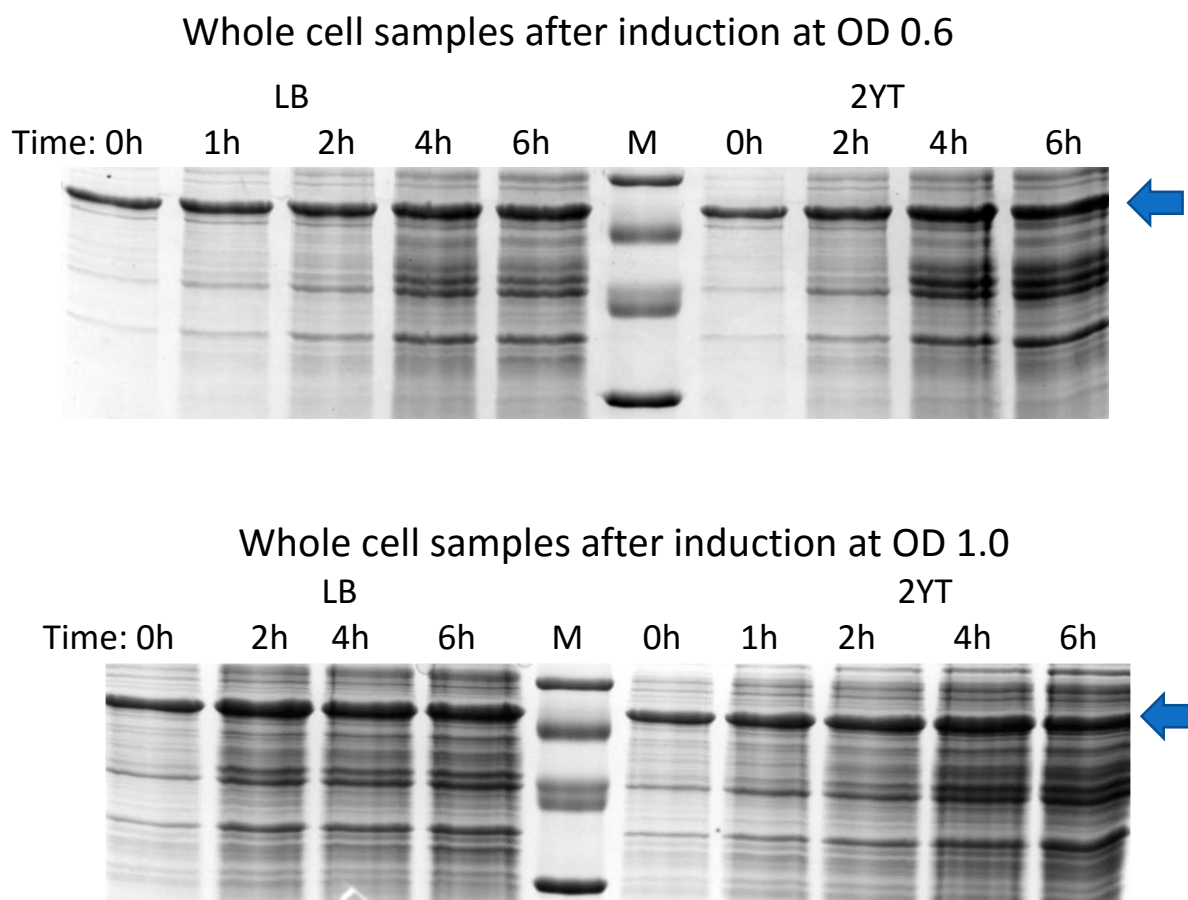
Due to the different ionization character, a significant excess of hydrazide peptide would probably cause ionization repression of the ligated protein and the thioester peptide. The addition of a small amount of TCEP at the beginning is necessary for the reduction of cysteine to avoid disulfide binding to mesNa. The reaction has been controlled with HPLC-MS over time. The desired product can already be observed immediately after adding the hydrazide peptide, however, the conversion of reaction seems to stop after 5 hours. Both educts are still present after up to 48 hours. Hence the reaction needs to be further optimized and different other ligation methods should be investigated.

## 4.7. Endo-A

Endo-A is necessary for the formation of the tetrasaccharide from SGP and also for transferring Man<sub>3</sub>GlcNAc oxazoline to the GlcNAc present on the asparagine side chain. Several test expressions with the plasmid kindly provided by Prof. K. Takegawa (Department of Bioscience and Biotechnology, Kyushu University) [57] have been performed. Further scale up expression and purification steps are not part of this thesis due to time limitation and have been performed by Mag. Gerhard Niederacher with the help of Dr. Aleksandr Kravchuk. However, the activity of the purified enzyme has been investigated.

### 4.7.1. Endo-A test expression

First, *E. coli* BL-21 DE3 competent cells have been transformed with the Endo-A plasmid and used to test the cell growth in LB and 2YT cell medium. Two different cell densities (0.6 and 1) have been chosen for IPTG activation. Samples have been taken over time, centrifuged and the protein pellets have been dissolved in ddH<sub>2</sub>O and loading buffer for SDS-PAGE analysis.



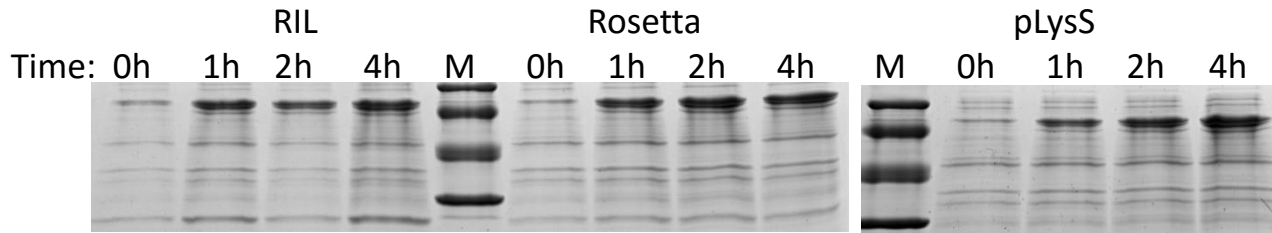
**FIGURE 53 ENDO-A TEST EXPRESSION WITH *E. COLI* BL-21 DE3 CELLS, AFTER INDUCTION AT CELL DENSITIES OF 0.6 AND 1.0; BLUE ARROW: ENDO-A ENZYME SIGNAL**

As shown above (Figure 53), the soluble endo A has approximately a molecular weight of 80 k Da<sup>[57]</sup> and its expression already starts before IPTG induction (at time 0h) and therefore no conclusion can be drawn from induction at different cell densities. Moreover, no clear difference in cell growth between LB and 2YT buffers can be observed. Although the protein concentration seems to increase slightly over time, a controlled induction would be preferred.

Three variations of BL-21 DE3 have been tested in LB medium: BL-21 DE3 RIL, which contains extra copies of the *argU*, *ileY* and *leuW* tRNA genes and is suitable for AT-rich genomes; BL-21

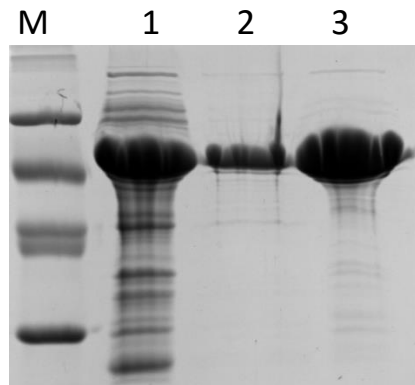


DE3 Rosetta which hosts strains enhancing the expression of proteins that contain codons rarely used in *E. coli*; BL-21 DE3 pLySs, which allows highly efficient protein expression under a T7 promoter also due to the additionally encoded T7 lysozyme that lowers the background expression level. IPTG has been added at a cell density of 0.6. As shown below (Figure 54), all three variations give satisfying cell growth signals and the target protein is successfully induced upon IPTG addition.



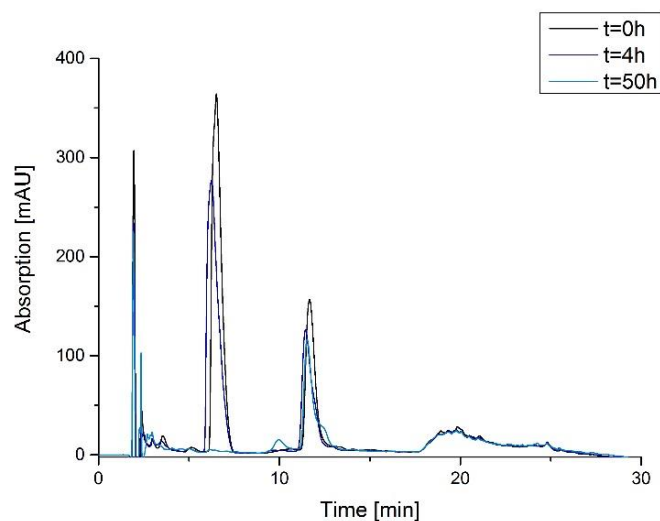
**FIGURE 54 BL-21 DE3 RIL, ROSETTA AND PLYS CELL ENDO-A TEST EXPRESSION**

In comparison, the pLysS strain has proven to be the most promising for the Endo-A expression, due to constant and linear cell growth over time and an increasing amount of Endo-A. Therefore, further expressions have been carried out in pLysS cells with IPTG activation at a cell density of 0.6 and purified via ion exchange and phenyl-sepharose chromatography by colleague.



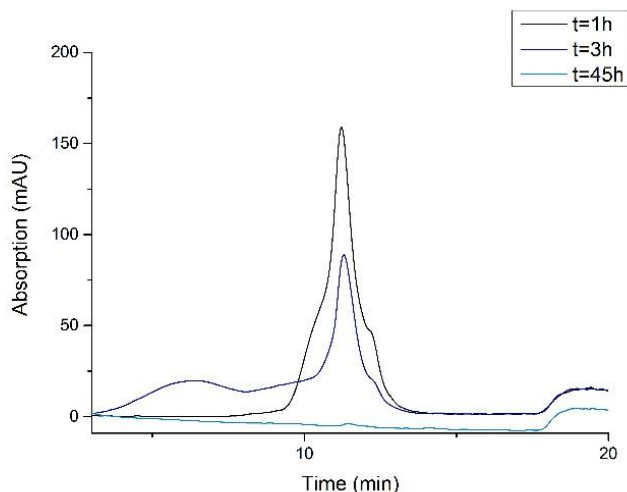
**FIGURE 55 PLYS CELL EXPRESSION; 1: CRUDE ENDO-A ENZYME BEFORE PURIFICATION; 2: ENDO-A ENZYME AFTER THE FIRST PURIFICATION (ION EXCHANGE CHROMATOGRAPHY); 3: ENDO-A ENZYME AFTER THE SECOND PURIFICATION (PHENYL-SEPHAROSE CHROMATOGRAPHY)**

#### 4.7.2. Endo- A activity assay



**FIGURE 56 UV CHROMATOGRAMS OF THE ENDO-A ACTIVITY ASSAY WITH SGP, CONTROLLED OVER TIME WITH HPLC; DECREASING IMPURITY PEAK AT TR = 7 MIN AND SGP PEPTIDE TR = 12 MIN**

After the first purification step of the desired enzyme Endo-A, the activity has been tested on SGP and the reaction progress has been controlled over time with HPLC. It can be observed that the impurity peak at 7 min and the SGP peak at 12 min are decreased after 4 hours. No new peak has been formed, which indicates the trimmed fragments are not concentrated enough to be detected or elute at the very beginning of the HPLC run. After 50 hours, the impurity peak has completely disappeared and the SGP peak has remained. This indicates that the enzyme activity is present, but the conversion of the impurity is faster than of the SGP. Moreover, the impurity probably contains linkages, which can be cleaved with Endo-A enzyme. Hence, some hypothesis can be formulated for the impurity peak. It is possible that the molecular mass is much smaller than SGP and therefore easier to be accessed by Endo-A, which would also explain the lack of detection in the mass spectrum.



**FIGURE 57 UV CHROMATOGRAMS (ZOOM) OF THE ENDO-A ACTIVITY ASSAY WITH RP-HPLC PURIFIED SGP; DECREASING SGP PEPTIDE TR=12MIN OVER TIME**

To further investigate this matter, the SGP has been purified with HPLC and the Endo-A catalyzed reaction was performed under the same condition described above. A higher concentration of Endo- A has been used due to the rather slow enzymatic reaction and the reaction progress has been controlled over time with HPLC. Already after 3 hours, the amount of SGP has decreased enormously. After 45 hours, no SGP can be detected with HPLC (Figure 57). It can be assumed that the purified Endo-A showed activity towards the isolated SGP substrate. However, better analysis of the trimmed fragments and a more constant activity assay condition still needs to be developed.

# 5. CONCLUSION & OUTLOOK

In this work, the basis towards the synthesis of homogeneously glycosylated sFasL has been laid. Fundamental progress has been made in all three targeted areas of interest, regarding the first steps toward the isolation and transfer of a homogeneous core *N*-glycan to peptides, the optimized synthesis conditions for all three peptide segments via Fmoc based SPPS and also regarding the *N*- and *C*-terminal peptide modifications with PEG27, hydrazide and thioesters.

The process of SGP isolation from egg yolk and the subsequent purification has been successfully performed with yields comparable to those found in the literature (lit. value: 1.9g SGP isolated from 300 egg yolks). Attempts have been made for the trimming of SGP, whereby desialylation with acetic acid has been found to be most useful. Assay conditions have been established for testing the activity of neuraminidase and  $\beta$ -1,4 galactosidase. However, the conditions for complete enzymatic trimming of SGP still need to be optimized and the subsequent Man<sub>3</sub>GlcNAc oxazoline reaction to be tested.

Different strategies have been considered for the synthesis of the *C*-terminal fragment (key sequence). High quality and yield of the crude product are desired due to the planned insertion of highly valuable GlcNAc-asparagine building blocks and coupling of the multifunctional auxiliary at the *N*-terminal position. Both manual and automated synthesis of the key sequence have been performed several times. It has been noticed that the high swelling TentaGel shows PEG leakage upon cleavage, therefore the low loading, PEG-free Wang resin has been used instead. Despite the low solubility and the elution problems during the analysis, acceptable results have been obtained. However, the strategy of the synthesis still requires additional modifications in order to improve the solubility properties.

The crude peptides have been PEGylated with Fmoc-amino-PEG<sub>27</sub>-propionic acid for preliminary solubility and precipitation assays. The results indicate that the quality of the key sequence needs to be further enhanced with respect to purity, solubility and aggregation behavior. Other reaction conditions also need to be tested to avoid incomplete PEGylation. Based on these findings no conclusive statement can be made regarding the solubility behavior of the PEGylated peptide.

The remaining two peptide sequences (*N*-terminal and middle fragment of sFasL) have been synthesized successfully. To avoid methionine oxidation, the cleavage cocktail H has been used. Trifluoroacetylated and incomplete deprotected products can be removed in the purification step. However, the use of Boc-protected *N*-terminal amino acid and extending the cleavage time would reduce these impurities and therefore increase the yield for future synthesis. Due to the difficult elution profile and the low ionization of the *N*-terminal peptide fragment, the purification step is challenging and needs to be improved. After incubation of peptidyl TentaGel resin with hydrazine in THF overnight, *C*-terminal hydrazide peptides have been obtained. This is the first

time that such large hydrazide peptides been synthesized via direct hydrazinolysis from Wang linkers. The correct position of the hydrazide group has been confirmed with TFE extraction of the fully protected hydrazide peptide from TentaGel and subsequent MS-MS measurements.

The C-terminal hydrazide of the 141-201aa fragment has been successfully converted into a thioester. After purification, the ligation reaction of the 141-201aa thioester and 202-245aa hydrazide peptide has been performed and product was observed immediately after adding the hydrazide peptide. However, the conversion of the reaction seems to stop after 5 hours although both educts are still intact and available. Unfortunately, it has not been possible to estimate the exact amount of product, since the UV trace of the product overlays these of the educts. This observation indicates that the reaction conditions need to be further optimized.

Furthermore, the first test expression tests of the Endo-A enzyme have been performed using a plasmid kindly provided by Prof. K. Takegawa (Department of Bioscience and Biotechnology, Kyushu University) and suitable expression conditions have been found. The purified Endo-A showed activity towards the isolated SGP substrate. However, better analysis of the trimmed fragments and a more robust activity assay condition still need to be developed.

Taken together, these data clearly show that the synthesis of homogeneously N-glycosylated Tumor Necrosis Factor Ligand Superfamily 6 (sFasL) via auxiliary mediated native chemical ligation is feasible. Upon addressing the points mentioned above, an optimized strategy to generate glycosylated sFasL will become available for the first time.

# 6. ABBREVIATIONS

A, Ala	Alanine
Aa	Amino acid
ACN	Acetonitrile
Boc	Tert-butoxycarbonyl
C, Cys	Cysteine
CaCl <sub>2</sub>	Calcium chloride
D, Asp	Aspartate
DCM	Dichloromethane
DMF	N,N-Dimethylformamide
DNA	Deoxyribonucleic acid
E, Glu	Glutamate
E. Coli	Escherichia coli
Endo-A	<i>Endo-β-N-Acetylglucosaminidase from Arthrobacter protophormiae</i>
ER	Endoplasmic reticulum
Et <sub>2</sub> O	Diethylether
F, Phe	Phenylalanine
FasL	Fas ligand
FasR	Fas receptor
Fmoc	9-Fluorenylmethyloxycarbonyl
G, Gly	Glycine
GlcNAc	N-Acetylglucosamin
Gua	Guanidiniumhydrochlorid
H, His	Histidine
HCl	Hydrochloric acid
HPLC	High-performance liquid chromatography
I, Ile	Isoleucine
IPTG	Isopropyl-β-D-thiogalactopyranosid

K, Lys	Lysine
L, Leu	Leucine
LC	Liquid chromatography
M, Met	Methionine
Man	Mannose
MeOH	Methanol
mesNa	2-Mercaptoethanesulfonic acid
MPAA	4-Mercaptophenylacetic acid
MS	mass spectrometry
N, Asn	Asparagine
N <sub>2</sub>	Nitrogen gas
NaCl	Sodium chloride
NaNO <sub>2</sub>	Natriumnitrit
NaOH	Natriumhydroxid
NCL	Native chemical ligation
Nonoxinol 40	NP-40
P, Pro	Proline
PAGE	Polyacrylamide gel electrophoresis
PEG	Polyethylene glycol
pNP-βGal	4-Nitrophenyl-β-D-galactopyranoside
Q, Gln	Glutamine
R, Arg	Arginine
S, Ser	Serine
SDS	Sodium dodecyl sulfate
sFasL	Soluble Fas ligand
SGP	Sialylglycopeptide
SPPS	Solid phase peptide synthesis
T, Thr	Threonine
t-Bu	Tert-Butyl
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Trifluoroacetic acid

THD	Tumor necrosis factor homology domains
THF	Tetrahydrofuran
Tris-HCl	2-Amino-2-(hydroxymethyl)-1,3-propanediol hydrochloride
W, Trp	Tryptophan
Y, Tyr	Tyrosine



# 7. LITERATURE

- [1] L. X. Wang, M. N. Amin, *Chemistry & Biology* **2014**, *21*, 51-66.
- [2] R. G. Spiro, *Glycobiology* **2002**, *12*, 43-56.
- [3] A. Varki, *Glycobiology* **1993**, *3(2)*, 97-130.
- [4] a) H. Lodish, A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore, J. Darnell, *W.H. Freeman* **2000**;  
b) J. Roth, *Chemical Reviews* **2002**, *102*, 285-304.
- [5] a) F. R. Schmidt, *Applied Microbiology and Biotechnology* **2004**, *65*, 363-372; b) H. P. Sørensen, K. K. Mortensen, *Journal of Biotechnology* **2005**, *115*, 113-128.
- [6] G. Rosenblum, B. S. Cooperman, *FEBS letters* **2014**, *588*, 261-268.
- [7] a) R. B. Merrifield, *The Chemistry of Polypeptides* **1973**, *16*, 335-361; b) R. B. Merrifield, *Journal of the American Chemical Society* **1963**, *85*, 2149-2154.
- [8] S. B. H. Kent, *Annual Review of Biochemistry* **1988**, *57*, 957-989.
- [9] M. Amblard, J. A. Fehrentz, J. Martinez, G. Subra, *Molecular Biotechnology* **2006**, *33*, 239-254.
- [10] a) M. T. Weinstock, M. T. Jacobsen, M. S. Kay, *Proceedings of the National Academy of Sciences of the United States of America* **2014**, *111*, 11679-11684; b) S. B. H. Kent, *Chemical Society Reviews* **2009**, *38*, 338-351.
- [11] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776-779.
- [12] T. W. Muir, D. Sondhi, P. A. Cole, *The National Academy of Sciences* **1998**, *95*, 6705-6710.
- [13] a) Y. Sohma, Q. X. Hua, J. Whittaker, M. A. Weiss, S. B. H. Kent, *Angewandte Chemie International Edition* **2010**, *49 (32)*, 5489-5493; b) Y. W. Wu, L. K. Oesterlin, K. T. Tan, H. Waldmann, K. Alexandrov, R. S. Goody, *Nature Chemical Biology* **2010**, *6*, 534-540.
- [14] L. E. Canne, S. J. Bark, S. B. H. Kent, *Journal of the American Chemical Society* **1996**, *118*, 5891-5896.
- [15] L. Z. Yan, P. E. Dawson, *Journal of the American Chemical Society* **2001**, *123*, 526-533.
- [16] L. R. Malins, R. J. Payne, *Current Opinion in Chemical Biology* **2014**, *22*, 70-78.
- [17] L. Zhang, J. P. Tam, *Tetrahedron Letters* **1996**, *38*, 4375-4378.
- [18] J. P. Tam, Q. Yu, *Biopolymers* **1998**, *46*, 319-327.
- [19] a) J. A. Camarero, B. J. Hackel, J. J. d. Yoreo, A. R. Mitchell, *Journal of the American Chemical Society* **2004**, *69*, 4145-4151; b) M. Arthur, Z. Felix, R. B. Merrifield, *Journal of the American Chemical Society* **1970**, *92*, 1385-1391; c) H. Romovacek, S. R. Dowd, K. Kawasaki, N. Nishi, K. Hofmann, *Journal of the American Chemical Society* **1979**, *101*, 6081-6091.
- [20] G. M. Fang, Y. M. Li, F. Shen, Y. C. Huang, J. B. Li, Y. Lin, H. K. Cui, L. Liu, *Angewandte Chemie* **2011**, *123*, 7787-7791.
- [21] C. Bello, F. Kikul, C. F. W. Becker, *Journal of Peptide Science* **2015**, *21*.
- [22] Y. Shin, K. A. Winans, B. J. Backes, S. B. H. Kent, J. A. Ellman, C. R. Bertozzi, *Journal of the American Chemical Society* **1999**, *121*, 11684-11689.
- [23] K. S. Asahina Y, Takao T, Nishi N, Hojo H., *Angewandte Chemie International Edition* **2013**, *52*, 9733-9737.
- [24] a) J. V. Lomino, A. Naegeli, J. Orwenyo, M. N. Amin, M. Aebi, L. X. Wang, *Bioorganic & Medicinal Chemistry* **2013**, *21*, 2262-2270; b) E. W. S. N. Kerney Jebrell Glover, Barbara Imeriali, *Chemistry and Biology* **2005**, *12*, 1311-1316; c) B. Imperiali, T. L. Hendrickson, *Bioorganic & Medicinal Chemistry* **1995**, *3*, 1565-1578; d) M. M. Chen, E. Weerapana, E. Ciepichal, J. Stupak, C. W. Reid, E. Swiezewska, B. Imperiali, *Biochemistry* **2007**, *46*, 14342-14348; e) R. E. Dempiski, B. Imperiali, *Current Opinion in Chemical Biology* **2002**, *6*, 844-850.

- [25] M. Kowarik, S. Numao, M. F. Feldman, B. L. Schulz, N. Callewaert, E. Kiermaier, I. Catrein, M. Aebi, *Science* **2006**, *314*, 1148-1150.
- [26] a) Y. Yuan, J. Chen, Q. Wan, R. M. Wilson, S. J. Danishefsky, *Biopolymers Peptide Science* **2010**, *94*, 373-384; b) C. S. Bennett, C. H. Wong, *Chemical Society Reviews* **2007**, *36*, 1227-1238; c) G. Chen, Q. Wan, Z. Tan, C. Kan, Z. Hua, K. Ranganathan, S. J. Danishefsky, *Angewandte Chemie International Edition* **2007**, *46*, 7383 -7387.
- [27] a) R. S. Haltiwanger, G. D. Holt, G. W. Hart, *The Journal of Biological Chemistry* **1990** *265*, 2563-2568; b) E. Bause, L. Lehle, *European Journal of Biochemistry* **1979**, *101*, 531-540; c) W. Huang, S. Groothuys, A. Heredia, B. H. M. Kuijpers, F. P. J. T. Rutjes, F. L. v. Delft, L. X. Wang, *ChemBioChem* **2009**, *10*, 1234-1242.
- [28] C. Bello, S. Wang, L. Meng, K. W. Moremen, C. F. Becker, *Angewandte Chemie International Edition* **2015**, *54*, 7711-7715.
- [29] T. Kawakami, S. Aimoto, *Tetrahedron Letters* **2003**, *44*, 6059-6061.
- [30] C. Marinzi, J. Offer, R. Longhi, P. E. Dawson, *Bioorganic and Medical Chemistry* **2004**, *12*, 2749-2757.
- [31] C. Bello, K. Farbiarz, J. F. Möller, C. F. W. Becker, T. Schwientek, *Chemical Science* **2014**, *5*, 1634.
- [32] E. Bayer, M. Mutter, *Nature* **1972**, *237*, 512-513.
- [33] G. Cooper, *The Cell: A Molecular Approach 2nd Edition* **2000**.
- [34] a) M. E. Peter, A. Hadji, A. E. Murmann, S. Brockway, W. Putzbach, A. Pattanayak, P. Ceppi, *Cell Death and Differentiation* **2015**, *22*, 549-559; b) S. Nagata, *Cell* **1997**, *88*, 355-365.
- [35] E. Rouvier, M.-F. Luciani, P. Golstein, *The Journal of experimental medicine* **1993**, *177*, 195-200.
- [36] H. J. Kaplan, M. A. Leibole, T. Tezel, T. A. Ferguson, *Nature Medicine* **1999**, *5*, 292-297.
- [37] S. Nagata, *Annual Review of Genetics* **1999**, *33*, 29-55.
- [38] M. Muzio, A. M. Chinnaiyan, F. C. Kischkel, K. O'Rourke, A. Shevchenko, J. Ni, C. Scaffidi, J. D. Bretz, M. Zhang, R. Gentz, M. Mann, P. H. Kramme, M. E. Peter, V. M. Dixit, *Cell* **1996**, *85*, 817-827.
- [39] W. C. Earnshaw, L. M. Martins, S. H. Kaufmann, *Annual Review of Biochemistry* **1999**, *68*, 383-424.
- [40] a) J. O'Connell, G. C. O'Sullivan, J. K. Collins, F. Shanahan, *The Journal of Experimental Medicine* **1996**, *184*, 1075-1082; b) K. Miwa, M. Asano, R. Horai, Y. Iwakura, S. Nagata, T. Suda, *Nature Medicine* **1998**, *4*, 1287-1292; c) S. Strand, W. J. Hofmann, H. Hug, M. Müller, G. Otto, D. Strand, S. M. Mariani, W. Stremmel, P. H. Krammer, P. R. Galle, *Nature Medicine* **1996**, *2*, 1361-1366; d) T. S. Griffith, T. Brunner, S. M. Fletcher, D. R. Green, T. A. Ferguson, *Science* **1995**, *270*, 1189-1192; e) A. M. Hunter, E. C. LaCasse, R. G. Korneluk, *Apoptosis* **2007**, *12*, 1543-1568.
- [41] E. J. Blott, G. Bossi, R. Clark, M. Zvelebil, G. M. Griffiths, *Journal of Cell Science* **2001**, *114*, 2405-2416.
- [42] a) M. Tanaka, T. Suda, K. Haze, N. Nakamura, K. Sato, F. Kimura, K. Motoyoshi, M. Mizuki, S. Tagawa, S. Ohga, K. Hatake, A. H. Drummond, S. Nagata, *Nature Medicine* **1996**, *2*, 317-322; b) O. Janssen, J. Qian, A. Linkermann, D. Kabelitz, *Cell Death & Differentiation* **2003**, *11*, 1215-1225.
- [43] K. Nozawa, N. Kayagaki, Y. Tokano, H. Yagita, K. Okumura, H. Hasimoto, *Arthritis & Rheumatism* **2005**, *40*, 1126-1129.
- [44] E. Song, J. Chen, N. Ouyang, F. Su, M. Wang, U. Heemann, *British Journal of Cancer* **2001**, *85*, 1047-1054.
- [45] P. Kern, M. Dietrich, C. Hemmer, N. Wellinghausen, *Infection and Immunity* **1999**, *68*, 3061-3063.
- [46] S. Adamopoulos, J. T. Parissis, M. Georgiadi, D. Karatzas, J. Paraskevaidis, C. Kroupis, G. Karavolias, K. Koniavitou, D. T. Kremastinos, *American Heart Journal* **2002**, *144*, 359-364.

- [47] W. C. Powell, B. Fingleton, C. L. Wilson, M. Boothby, L. M. Matrisian, *Current Biology* **1999**, *9*, 1441-1447.
- [48] a) T. Vargo-Gogola, H. C. Crawford, B. Fingleton, L. M. Matrisian, *Archives of Biochemistry and Biophysics* **2002**, *408*, 155-161; b) T. Suda, H. Hashimoto, M. Tanaka, T. Ochi, S. Nagata, *The Journal of experimental medicine* **1997**, *186*, 2045-2050.
- [49] P. Schneider, J. L. Bodmer, N. Holler, C. Mattmann, Patricia Scuderi, A. Terskikh, M. C. Peitsch, J. Tschopp, *the Journal of Biological Chemistry* **1997**, *272*, 18827-18833.
- [50] B. Sun, W. Bao, X. Tian, M. Li, H. Liu, J. Dong, W. Huang, *Carbohydrate Research* **2014**, *396*, 62-69.
- [51] R. G. Spiro, *The Journal of Biological Chemistry* **1961**, *237*, 646-652.
- [52] M. Shigeta, Y. Shibukawa, H. Ihara, E. Miyoshi, N. Taniguchi, J. Gu, *Glycobiology* **2006**, *16*, 564-571.
- [53] J. J. Distler, G. W. Jourdian, *Journal of Biological Chemistry* **1973**, *248*, 6772-6780.
- [54] G. G. Sahagian, J. Distler, G. W. Jourdian, *Proceedings of the National Academy of Sciences* **1981**, *78*, 4289-4293.
- [55] C. Bello, F. Kikul, C. F. W. Becker, *Journal of Peptide Science* **2014**, *21*, 201-207.
- [56] C. P. Hackenberger, *Organic & Biomolecular Chemistry* **2006**, *4*, 2291-2295.
- [57] K. Takegawa, M. Nakoshi, S. Iwahara, K. Yamamoto, T. Tochikura, *Applied and Environmental Microbiology* **1989**, *55*(12), 3107-3112.