Chemical Synthesis of the Thymocyte Differentiation Antigen 1 (Thy-1) N-glycoprotein

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry, Pharmacy

of Freie Universität Berlin

by

Antonella Rella

from Bari, Italy

June 2021

This work was performed between August 2016 and June 2020 under the supervision of Dr. Daniel Varón Silva in the Department of Biomolecular Systems, Max Planck Institute of Colloids and Interfaces Potsdam, and the Institute of Chemistry and Biochemistry, Freie Universität Berlin.

1st Reviewer: Dr. Daniel Varón Silva

2nd Reviewer: Prof. Dr. Mathias Christmann

Date of defense: 22.10.2021

Declaration

This is to certify that the entire work in this thesis has been carried out by Ms. Antonella Rella, if not stated otherwise. The assistance and help received during the course of investigation have been fully acknowledged.

Berlin, 22.10.2021 ----------

(Date, Place) (Signature)

Acknowledgements

I express my gratitude to Dr. Daniel Varón Silva for let me work in the GPI group and for his guidance during last years.
I also want to thank Prof. Dr. Peter Seeberger for giving me the opportunity to perform my dissertation in the biomolecular systems department at Max Planck Institute of Colloids and Interfaces.
I am thankful to Prof. Dr. Mathias Christmann for kindly agreeing to review this thesis.

I want to thank current and former members of the GPI group for the motivational working environment and support in lab Sandra, Hyunil and Paula.
I have special thanks to Ankita for the amazing time spent in the lab and the outside the work.
I want to thank Renee, for her infinite help and for introducing me to the analytical field. I think I would not do my actual job without you.
A special thanks to Sandra for her huge help and for doing her magic on my peptides.
A special thanks to my MPI colleagues Silvia, Priya, Pietro, Martina, Bart, Mara, Cristian, Susanne, Sebastian, Amira, Mauro, Alonso, Jamal, Alexandra, Vittorio, Mike, CD.
I thank Dorothee Böhme for her help and support and Eva Settels, for outstanding technical and organizational support.

A special thank goes to Maria for helping me, feeding me and everything done during these 4 years.
I will always be grateful to Stefano Piotto and Simona Concilio for their unconditional help and support.
I would like to thank Franco and Angela for their unconditional support and encouragement and for feeding me even from thousand miles away. I am also grateful to my second family Gino and Lucia for their infinite support.

My final thanks go to Roberto. Thank you for your support and patience.

List of Publications

Parts of this work have been or will be published.

Scientific Publications

- Roller R. F., Malik A., Carillo M. A., Garg M., **Rella A.**, Raulf M., Lepenies B., Peter H. Seeberger P.H and Varón Silva D. "Semisynthesis of Functional Glycosylphosphatidylinositol-Anchored Proteins". *Angew. Chemie* **2020**, *132*, 12133–12138.

- Concilio S., Ferrentino I., Sessa L., Massa A., Iannelli P., Diana R., Panunzi B., **Rella A**. and Piotto S. "A novel fluorescent solvatochromic probe for lipid bilayers", *Supramolecular Chemistry*, **2017**, 29:11, 887-895.

Scientific Conference and Symposia

- September 2019, "Chemical Synthesis of Glypiated and Glycosylated Thy-1 Protein", 7th RIKEN-Max Planck Symposium, Ringberg, Germany.
- July 2019, "Chemical Synthesis of Glypiated and Glycosylated Thy-1 Protein" ,10th Eurocarb, Leiden, Netherlands.
- June 2019, "Chemical Synthesis of Glypiated and Glycosylated Thy-1 Protein", 8th Chemical Protein Synthesis Meeting, Berlin, Germany.
- November 2018, " Developments of Strategies for the Chemical Synthesis og Glypiated and Glycosylated Proteins", 8th Peptide Engineering Meeting, Berlin, Germany.
- September 2017, "Synthesis of C-Terminal Modified Peptides", Ringberg Meeting of Max Planck Society, Rottach-Egern, Germany.

Contents

List of abbreviations

a.a.	Amino acid
AcBr	Acetyl bromide
AcCl	Acetyl chloride
AcOH	Acetic acid
Acm	Acetamidomethyl
ACN	Acetonitrile
Ala	Alanine
All	Allyl
Alloc	Allyloxylcarbonyl
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
AP	Anchored protein
Boc	*Tert*-butoxycarbonyl
BOP	(Benzotriazol-1-yloxy)tris(dimethylamino)phosphonium
Bn	Benzyl
Cl-HOBt	6-Chloro-1-hydroxybenzotriazole
COMU	(1-Cyano-2-ethoxy-2-oxoethylidenaminooxy)dimethylamino-morpholino-carbenium hexafluorophosphate
Cys	Cysteine
DCC	N,N'-dicicloesilcarbodiimmide
DCM	Dichloromethane
DIC	N,N′-Diisopropylcarbodiimide
DIPEA	*N*,*N*-diisopropyl-ethylamine
DMC	2-chloro-1,3-dimethylimidazolinium chloride

DMF	*N*,*N*-dimethylformamide
DMSO	Dimethylsulfoxide
DTT	Dithiothreitol
EDC	1-Ethyl-3-(3-dimethylaminopropyl)-carbodiimid
EDT	1,2-Ethanedithiol
ENGases	Endo-β-N-acetylglucosaminidase
Equiv.	Equivalents
ER	Endoplasmic reticulum
ESI-MS	Electrospray ionization mass spectrometry
Et_3N	Triethylamine
EtOAc	Ethyl acetate
FA	Formic acid
Fmoc	9-Fluorenylmethoxycarbonyl
Gal	Galactose
Gdn HCl	Guanidinium hydrochloride
GlcN	Glucosamine
GlcNAc	*N*-acetylglucosamine
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
GPI	Glycosylphosphatidylinositol
h	Hour
H_2	Hydrogen
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate
HBTU	N,N,N′,N′-Tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate
HCl	Hydrochloric acid

Hg	Mercury
$Hg(OAc)_2$	Mercury (II) acetate
His	Histidine
HOAt	3H-[1,2,3]-Triazolo[4,5-b]pyridin-3-ol
HOBt	1-hydroxy benzotriazol
HPLC	High performance liquid chromatography
Hz	Hertz
Ile	Isoleucin
LC-MS	Liquid chromatography mass spectrometry
Fuc	Fucose
Leu	Leucine
Lys	Lysine
MALDI-TOF	Matrix assisted laser desorption ionization time of flight
Man	Mannose
Map	4-(dimethylamino)-phenacyl
Me	Metyl
MeOH	Methanol
Met	Methionine
min	Minutes
MMP	Methyl 3-(methylmercapto)propionate
MPAA	4-mercaptophenylacetic acid
NaN3	Sodium azide
NaNO2	Sodium nitrite
NaOMe	Sodium methoxide
NCL	Native chemical ligation
NMP	*N*-methyl-2-pyrrolidone
Oxyma	Ethyl cyano(hydroxyimino)acetate

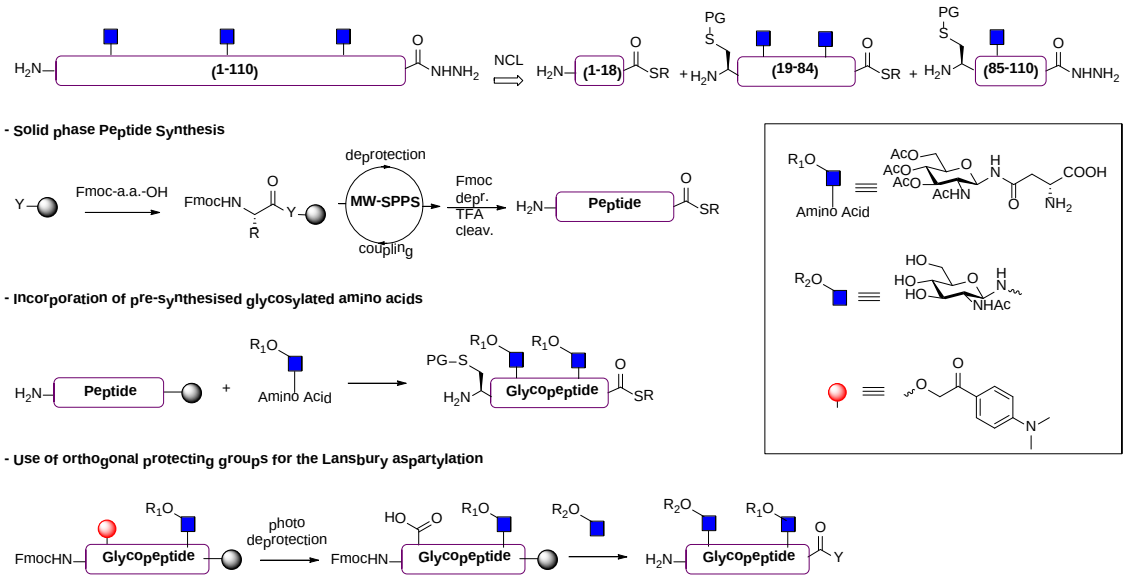
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
Pd	Palladium
PEG	Polyethylene glycol
Pfp	Pentafluorophenyl
PG	Protecting groups
Phe	Phenylalanine
Pro	Proline
PTMs	Post-translational modifications
PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphonium-hexafluorophosphate
Qtof	Quadrupole TOF
RP-HPLC	Reverse phase HPLC
rpm	Rounds per minute
r.t.	Room temperature
Rf	Retention factor
Rt	Retention time
Ser	Serine
SPPS	Solid phase peptide synthesis
TBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate
tBu	Tert-butyl
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
Thr	Threonine
Thz	Thiazolidine
TIPS	Triisopropylsilane
TLC	Thin layer chromatography
Tyr	Tyrosine
Val	Valine

Summary

The glycosylation of the proteins is a common post-translational modification found in eukaryotic cells. Glycosylated proteins play a crucial role in cell adhesion, cell-matrix interaction, and immune response. Several studies suggest that the overexpression of these glycoconjugates is linked to many diseases, inflammation, viral infections, cancer metastasis, and cellular apoptosis among other pathologies. Over the years, the scientific community has tried to answer questions about the influence of the oligosaccharides on the protein structure or the relationship between the oligosaccharide structure and glycoproteins function. The main limitation to study these important aspects is the difficulty to obtain homogeneous isoforms of glycoproteins due to the lack of genetic control in the biosynthesis. The crescent interest in the understanding of structure-activity relationship of glycoproteins has led to the development of new technologies to obtain complex glycans and proteins.

In this work, different methodologies were investigated to synthesize defined isoforms of the ubiquitous glycoprotein called Thy-1 or CD90. This heavily glycosylated protein was first discovered in 1964 in mouse T lymphocytes and later detected in human fibroblast neurons, blood stem cells, and endothelial cells. The 25kDa glycoprotein is translated as a 161 and 160 amino acids length protein in humans and mouse, respectively. The protein is transferred from ribosomes into the ER and receives several post-translational modifications including N-glycosylation at three different asparagine residues and a glypiation at the C-terminal cysteine residue. Several studies showed the association between the expression of Thy-1 and T cell activation, neurite outgrowth, apoptosis, leukocyte and melanoma cell adhesion and migration, tumor suppression, and fibroblast proliferation. The exploration of the structure-activity relationship to determine the role of the glycosylation on this protein required defined glycoprotein and motivated this work to develop a *de novo* synthesis of the Thy-1 pure glycoforms.

The retrosynthesis of the mature glycoprotein was designed considering three key steps: 1) the assembly of the primary sequence of the protein; 2) the introduction of glycans into the protein; 3) the differentiation of the oligosaccharides at the glycosylation sites. The primary sequence of the glycoprotein was obtained using a sequential native chemical ligation (NCL) of three segments: the peptide fragment (1-18), the glycopeptide (19-84) having two N-glycosylations, and the glycopeptide (85-110) containing one N-glycosylation.



Scheme 1 – Strategies for the synthesis of the required peptide and glycopeptide fragments required for the assembly of the glycoprotein Thy-1.

The three peptide fragments were synthesized by solid-phase peptide synthesis (SPPS) and were obtained having a hydrazide moiety on the C-*terminus* that was converted into a thioester prior to the NCL. For this purpose, a Wang and a trityl resin were functionalized with hydrazine and modified manually with the amino acid on the C-*termini*. The protecting acetamidomethyl group was used to mask the thiol of the N-terminal cysteine residues in the fragment (19-84) and (85-110) to allow the stepwise ligation of the fragments.

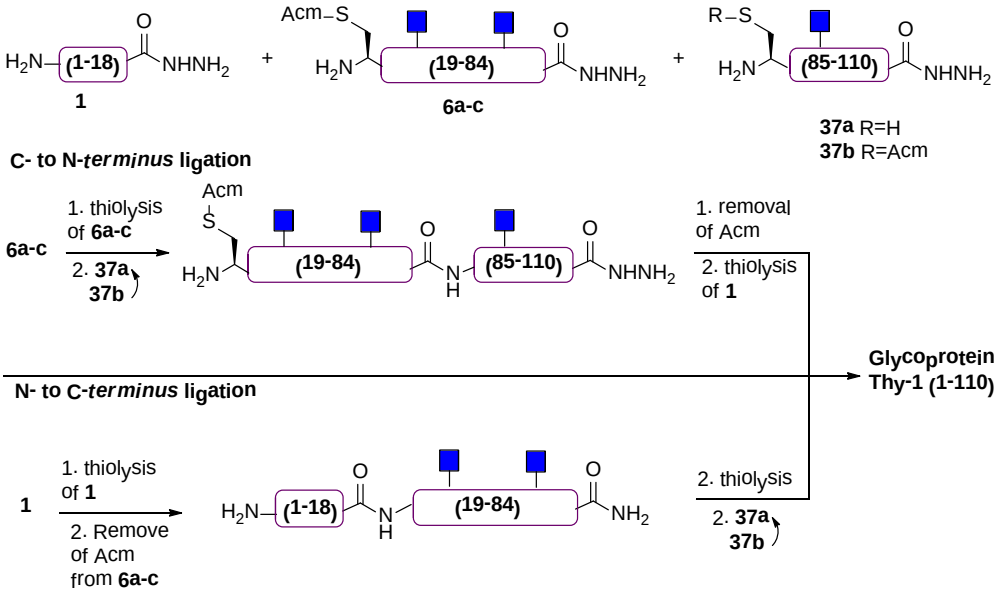
The elongation of the polypeptide chains was carried out in the microwave-assisted synthesizer and careful optimization of deprotection and coupling cycles was executed for the three fragments. The incorporation of amino acids was performed considering the common problems that affect the peptide synthesis including the cyclization reaction on asparagine and aspartic acid, and the racemization of cysteine and histidine at temperatures higher than 50°C. These precautions were not enough to obtain the Thy-1 glycopeptide fragments in high yield and with desired quality. The synthesis of the designed peptides required multiple optimizations to avoid side reactions. Some general requirements were established. Hindered amino acids (glutamic acid, isoleucine, and phenylalanine) required double coupling to assure complete attachment to the growing chain. Vicinal identical amino acids (e.g. serine 25-serine 26) were introduced using single coupling for the first residue in the sequence and a double coupling for the following residue. One of the most crucial point for the optimization of the synthesis was the minimization of the aspartimide

formation involving the cyclization of the aspartic acid. Among different cocktails for the removal of the Fmoc group promising the suppression of this side reaction, the best results were obtained by using 20% piperidine and 0.7% of formic acid in dimethylformamide, which reduced the aspartimide formation up to 70%.

The glycosylations were introduced following two approaches to allow differentiation between the glycan at each glycosylation site. In the cassette-method, a glycosylated asparagine was synthesized and incorporated in the corresponding glycosylation site during the assembly of the peptides. Different procedures were evaluated and optimized for the synthesis of the N-acetyl-O-per-acetylated glucosaminyl asparagine **2** and its introduction into SPPS to get the fragments **6a-c** (19-84) and **37a-b** (85-110) with a peracetylated glucosamine on each glycosylation site.

The convergent approach was used to synthesize fragment **6b** (19-84) having the peracetylated N-acetylglucosamine **2** in one glycosylation site and an unprotected N-acetylglucosamine, installed via Lansbury aspartylation on the second site. The Lansbury strategy required the formation of the amide bond between the free carboxylate at an aspartic acid and the N-acetylglucosamine amine. The reaction was performed on the fully protected peptide fragment and required the synthesis of two building blocks, an orthogonal protected aspartic acid **3**, and a pseudoproline dipeptide **4**. The aspartic acid was synthesized with a photolabile protecting group on the side chain, which was selectively removed on resin to give a free carboxylic acid. To avoid the undesired rearrangement of the amino acid, the protected pseudoproline dipeptide Thr-Ser($\Psi^{Me,Me}$Pro) was incorporated before the aspartic acid.

The three (glyco)peptides were efficiently synthesized by a combination of manual and automated processes and were characterized by their challenging purifications. The difficulties in obtaining the isoforms from the Thy-1 fragment (19-84) glycopeptides **6a-b**, were related to the low solubility of the generated glycopeptides and the loss of acetyl groups on the peracetylated glucosamine. The analysis of these problems led to the design of the optimal strategy to synthesize in high yield the fragment **6c** having a protected and non-protected N-acetylglucosamine. The synthesis of glycoform **6c** involved the coupling of a per-acetylated N-acetylglucosaminyl asparagine building block in the peptide sequence that was de-acetylated before the introduction of the second glycosylated asparagine residue.



Scheme 2 – C- to N-*terminus* ligation and N- to C-*terminus* ligation explored for the assembly of the glycoprotein Thy-1.

The assembly of the Thy-1 glycoprotein involved the ligation of the synthesized fragments by the chemoselective reaction between the C-terminal peptide thioester in one peptide fragment and the free N-terminal cysteine residue of another fragment. The process required the conversion of the synthesized peptide hydrazides into the corresponding thioesters prior to each ligation and the removal of the Acm-group of the fragments having the N-terminal cysteine. Two combinations were studied to ligate the three fragments: from C- to N-*terminus* and from N- to the C-*terminus*. The first approach consisted of the initial ligation of the glycopeptides **6a-c** (19-84) and **37a-b** (85-110) and the subsequent ligation with the peptide **1** (1-18). The second strategy involved the reaction of peptide **1** (1-18) with the glycopeptide **6a-c** (19-84) and the following ligation with the fragment **37a-b** (85-110). The ligation conditions were investigated, and the strategy was selected considering factors such as the availability and solubility of the fragments, number of steps and the efficiency of the processes.

The N-to C-strategy was the first approach adopted. The hydrazide precursor **6a** was converted into the thioester fragment **38** by formation of the peptide azide at low pH and thiolysis with the corresponding thiol. The Acm group from the cysteine in **37b** was removed with mercury (II) acetate and reduction of the sulfur-mercury complex with mercaptoethanol. The ligation of the fragments delivered the glycoprotein **40** and the hydrolysed fragment **38** as by-product. The low solubility of the glycoprotein **40** hampered the isolation of the product and suggested the need of more polar fragment. For this purpose, glycoform **6b** was converted into the thioester **41** and

reacted with fragment **37a** having a free N-terminal cysteine. Surprisingly, this ligation did not proceed suggesting that the treatment with mercury (II) acetate could be beneficial for the ligation. A N- to C- ligation of fragment **6c** treated with the mercury (II) acetate with the thioester of fragment **1** proceeded successfully and the isolation of the ligation product yielded the glycoprotein **45**, confirming the better behaviour of polar fragments and reactivity of the cysteine-containing fragments treated with mercury. A ligation of the MPAA thioester of Thy-1 fragment (1-84) with **37b** to obtain the full glycoprotein Thy-1 (1-120) was hindered by the poor solubility of thioester **46** in the ligation media. A new C- to N- ligation involving the polar glycoform **6c** with the fragment **37b** delivered the glycoprotein **49** that was treated in one-pot with $PdCl_2$ to release the N-terminal cysteine and get **50**. Finally, a methyl 3-mercaptopropionate MMP thioester of fragment 1-18 (**51)** was ligated with glycoprotein fragment **50** to obtain the desired Thy-1 glycoprotein (1-120) having three glycosylations.

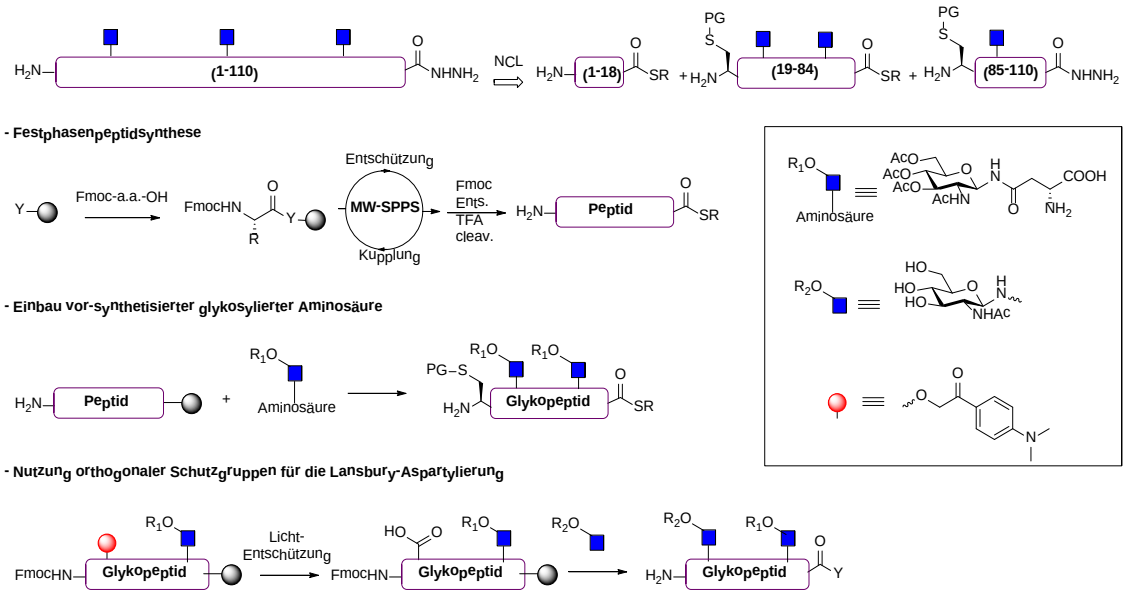
In this work was presented the design and evaluation of different strategies for the synthesis of the glycoprotein Thy-1. The assembly of the 13kDa glycoprotein required different steps and the optimized synthesis of amino acid building blocks for the solid-phase assembly of glycopeptides. Various methods were applied for the generation of glycopeptides, the exploration of the chemical properties of the obtained fragments, and the modulation of the chemical conditions for the ligation of the peptide fragments to get the target glycoprotein. This work focused on the production of homogeneous glycoforms of Thy-1. However, the synthetic methods and protocols established in this work are applicable for the synthesis of any peptides and glycopeptides and contribute to the chemical synthesis of other important glycoproteins.

Zusammenfassung

Glykosylierung ist eine häufige post-translationale Modifikation in eukaryotischen Zellen. Glykosylierte Proteine spielen eine entscheidende Rolle in Zell-Ädhäsion, Zell-Matrix-Adhäsion und Immunantwort. Mehrere Studien deuten darauf hin, dass die Überexpression dieser Glykokonjugate mit vielen Erkrankungen wie etwa Entzündungen, Metastasen bei Krebs, zelluläre Apoptose und weiteren Pathologien im Zusammenhang steht. Über die Jahre hat die wissenschaftliche Gemeinschaft versucht, die offenen Fragen über den Einfluss von Oligosacchariden auf die Proteinstruktur sowie die Beziehung zwischen Oligosaccharidstruktur und Glykoprotein-Funktion zu beantworten. Die größte Limitierung der Untersuchungen dieser wichtigen Aspekte sind bislang die Schwierigkeiten, homogene Isoformen von Glykoproteinen zu erhalten, die in mangelnder genetischer Kontrolle über die Biosynthese begründet sind. Das steigende Interesse am Verständnis der Struktur-Aktivität-Beziehung von Glykoproteinen hat zur Entwicklung neuer Technologien für die Generierung komplexer Glykane und Proteine geführt.

In dieser Arbeit wurden unterschiedliche Methoden untersucht, um definierte Isoformen des ubiquitären Glykoproteins Thy-1, oder CD90 zu synthetisieren. Dieses stark glykosylierte Protein wurde zum ersten Mal 1964 in T-Lymphozyten der Maus entdeckt und später auch in humanen Fibroblast-Neuronen, Blut-Stammzellen sowie Endothelzellen. Das 25kDa-große Glykoprotein wird als 161 bzw. 160 Aminosäuren langes Protein im Menschen bzw. in der Maus translatiert. Das Protein wird von den Ribosomen in das ER transferiert und erhält mehrere post-translationale Modifikationen, darunter eine N-Glykosylierung an drei verschiedenen Asparagin-Resten und eine Glypiation am C-terminalen Cystein. Mehrere Studien zeigten den Zusammenhang zwischen der Expression von Thy-1 und T-Zell-Aktivierung, Neuritenwachstum, Leukozyten- und Melanom-Zell-Adhäsion sowie Migration, Tumorsuppression und Fibroblasten-Proliferation. Die Erforschung dieser Struktur-Aktivitäts-Beziehung für die Bestimmung der Rolle der Glykosylierung dieses Proteins erforderte definierte Glykoproteine und begründete die Entwicklung einer *de novo* Synthese reiner Thy-1 Glykoformen.

Die Retrosynthese des reifen Glykoproteins wurde unter Berücksichtigung der folgenden drei Schritte entwickelt: 1) die Konstruktion der primären Sequenz des Proteins; 2) die Einführung von Glykanen in das Protein; 3) die Differenzierung der Olgisaccharide an den Glykosylierungsstellen. Die primäre Sequenz des Glykoproteins wurde durch eine sequenzielle Native Chemische Ligation (NCL) erreicht: das Peptid-Fragment (1-18), das Glykopeptid (19-84) mit zwei N-Glykosylierungen und das Glykopeptid (85-110) mit einer N-Glykosylierung.



Schema 1 – Strategien für die Synthese der benötigten Peptid- und Glykopeptidfragmente für die Konstruktion des Thy-1 Glykoproteins.

Die drei Peptidfragmente wurden mittels Festphasen-Peptidsynthese (SPPS) synthetisiert und wurden mit einem Hydrazid-Rest am C-Terminus hergestellt, der vor der NCL in einen Thioester umgewandelt wurde. Zu diesem Zweck wurde ein Wang- und ein Trityl-Resin mit Hydrazin funktionalisiert und manuell mit den Aminosäuren des C-Terminus modifiziert. Die Schutzgruppe Acetamidomethylgruppe wurde genutzt, um die Thiole der N-terminalen Cystein-Reste in den Fragmenten (19-84) und (85-110) zu maskieren und eine schrittweise Ligation der Fragmente zu ermöglichen.

Die Verlängerung der Polypeptidketten wurde in einem Mikrowellen-assistierten Syntheziser mit sorgfältiger Optimierung der Entschützungs- und Kupplungszyklen für die drei Peptide durchgeführt. Der Einbau der Aminosäuren wurde unter Berücksichtigung der üblichen Probleme bei der Peptidsynthese durchgeführt, inklusive der Zyklisierungsreaktion an Asparagin- und Aspartat-Resten sowie die Racemisierung von Cystein- und Histidin bei Temperaturen über 50° C. Diese Vorsichtsmaßnahmen reichten nicht aus, um die Thy-1 Glykopeptidfragmente in hoher Ausbeute und der gewünschten Qualität herzustellen. Die Synthese der designten Peptide erforderte mehrere Optimierungen um Nebenreaktionen zu vermeiden und die automatisierte Verlängerung zu optimieren. Einige allgemeine Anforderungen wurden daher aufgestellt. Gehinderte Aminosäuren (Glutamat, Isoleucin und Phenylalanin) erforderten doppelte Kupplung, um eine vollständige Kupplung zur wachsenden Peptidkette sicherzustellen. Benachbarte identische Aminosäuren (z.B. Serin-25-Serin-26) wurden mit einfacher Kupplung für die erste Aminosäure und doppelter Kupplung für die zweite Aminosäure eingeführt. Einer der kritischsten

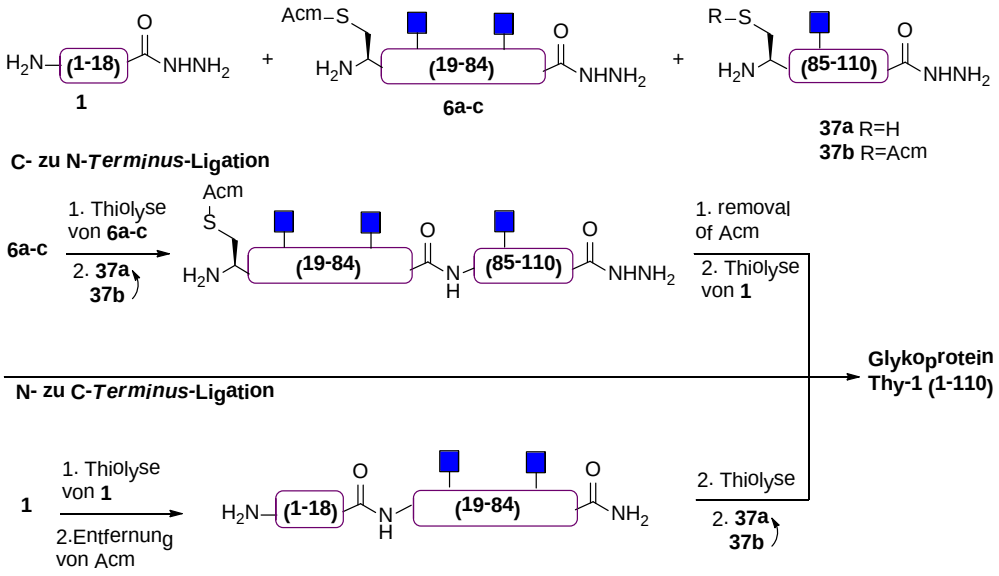
Punkte für die Optimierung der Synthese war die Minimierung der Aspartimid-Bildung unter Zyklisierung des Aspartat-Restes. Aus unterschiedlichen Cocktails für die Entfernung der Fmoc-Gruppe für die Unterdrückung dieser Nebenreaktion wurden die besten Ergebnisse erzielt unter Nutzung von 20 % Piperidin und 0,7 % Ameisensäure in Dimethylformamid. So konnte die Aspartimid-Bildung um bis zu 70 % reduziert werden.

Die Glykosylierungen wurden mittels zwei Ansätzen eingeführt, um die Differenzierung zwischen den Glykanen an jeder Glykosylierungsstelle zu ermöglichen. In der Kassetten-Methode wurde ein glykosyliertes Asparagin synthetisiert und während des Zusammenbaus des Peptids an der entsprechenden Stelle eingebaut. Unterschiedliche Verfahren wurden für die Synthese des N-Acetyl-O-peracetylierten Glucosaminyl-Asparagins **2** und dessen Einführung in die SPPS untersucht und optimiert, um die Fragmente **6a-c** (19-84) und **37a-b** (85-110) mit einem peracetylierten Glucosamin an jeder Glykosylierungsstelle zu erhalten.

Der konvergente Ansatz wurde eingesetzt, um Fragment **6b** (19-84) mit dem peracetylierten N-Acetylglucosamin **2** an einer Glykosylierungsstelle und einem ungeschützten N-Acetlyglucosamin (eingeführt via Lansbury-Aspartylierung) an der zweiten Stelle, zu synthetisieren. Die Lansnbury-Strategie erforderte die Bildung einer Amidbindung zwischen dem freien Carboxylat an einem Aspartat-Rest unddem N-Acetylglucosamin-Amin. Die Reaktion wurde mit dem vollständig geschützten Peptid-Fragment durchgeführt und erforderte die Synthese von zwei Building Blocks, einem orthogonal geschützten Aspartat **3** sowie einem Pseudoprolin-Dipeptid **4**. Das Aspartat wurde mit einer photolabilen Schutzgruppe an der Seitenkette synthetisiert, die selektiv auf dem Resin entfernt wurde, um eine freie Carboxylsäure zu erhalten. Um unerwünschte Umlagerungen der Aminosäure zu vermeiden, wurde das geschützte Pseudoprolin-Dipeptid Thr-Ser($\Psi^{Me,Me}$Pro) vor dem Asparat eingebaut.

Die drei (Glyko-)Peptide konnten mittels einer Kombination von manuellen und automatisierten Prozessen effizient synthetisiert werden und werden durch ihre herausfordernde Aufreinigung charakterisiert. Die Schweirigkeiten, die Isoformen des Thy-1-Fragmentes (19-84) sowie der Glykopeptide **6a-b** herzustellen hingen mit der niedrigen Löslichkeit der generierten Glykopeptide und dem Verlust der Acetylgruppen auf den peracetylierten Glucosaminen zusammen. Die Analyse dieser Probleme führte zum Design der optimalen Strategie für die Synthese des Fragments **6c** mit einem geschützten und einem ungeschützten N-Acetylglucosamin in hoher Ausbeute. Die Synthese der Glycoform **6c** beinhaltete die Kupplung eines peracetylierten N-

Acetylglucosaminyl-Asparagin-Building Blocks in der Peptidsequenz, welches vor der Einführung des zweiten glykosylierten Asparagin-Restes de-acetyliert wurde.



Schema 2 – Untersuchung von C- zu N-Terminus-Ligation und N- zu C-Terminus-Ligation für den Zusammenbau des Glykoproteins Thy-1

Der Zusammenbau des Thy-1 Glykoproteins beinhaltete die Ligation der synthetisierten Fragmente mit Hilfe der chemoselektiven Reaktion zwischen dem C-terminalen Peptid-Thioester eines Fragmentes und dem freien N-terminalen Cystein-Rest eines anderen Fragmentes. Der Prozess erforderte die Umwandlung des synthetisierten Peptid-Hydrazids in den korrespondierenden Thioester vor jeder Ligation und die Entfernung der Acm-Gruppe des Fragmentes mit dem freien N-terminalen Cystein. Zwei Kombinationen für die Ligation der drei Fragmente wurden untersucht: vom C- zum N-Terminus und vom N- zum C-Terminus. Der erste Ansatz bestand aus der initialen Ligation des Glykopeptids **6a-c** (19-84) und **37a-b** (85-110) sowie der nachfolgenden Ligation mit dem Peptid **1** (1-18). Die zweite Strategie beinhaltete die Reaktion von Peptid **1** (1-18) mit dem Glykopeptid **6a-c** (19-84) und der darauffolgenden Ligation mit dem Fragment **37a-b** (85-110). Die Ligationsbedingungen wurden untersucht und die Strategie wurde unter Berücksichtigung von Faktoren wie der Verfügbarkeit und Löslichkeit der Fragmente, der Anzahl der Schritte und der Effizienz der Prozesse ausgewählt.

Die N- zu C-Strategie war der zuerst umgesetzte Ansatz. Der Hydrazid-Vorläufer **6a** wurde in das Thioester-Fragment **38** umgewandelt über die Bildung des Peptid-Azids bei niedrigem pH-Wert und Thiolyse mit dem korrespondierenden Thiol. Die Acm-Gruppe des Cysteins in **37b** wurde mit Quecksilber(II)-Azetat und Reduktion des Schwefel-Quecksilber-Komplexes mit

Mercaptoethanol entfernt. Die Ligation des Fragmentes lieferte das Glykoprotein **40** und das hydrolysierte Fragment **38** als Nebenprodukt. Die niedrige Löslichkeit des Glykoproteins **40** beeinträchtigte die Isolation des Produktes und legte den Bedarf eines polareren Fragmentes nahe. Für diesen Zweck wurde die Glykoform **6b** in den Thioester **41** umgewandelt und reagierte mit dem Fragment **37a**, das ein freies N-terminales Cystein besaß. Überraschenderweise lief diese Ligation nicht erfolgreich ab, was nahelegte, dass die Behandlung mit Quecksilber(II)-Azetat vorteilhaft für die Ligation sein könnte.

Eine N- zu C-Ligation des mit Quecksilber(II)-Azetat behandelten Fragmentes **6c** mit dem Thioester von Fragment **1** lief erfolgreich ab und die Isolation des Ligationsproduktes lieferte das Glykoprotein **45**, was das bessere Verhalten von polaren Fragmenten und die bessere Reaktivität des Cystein-beinhaltenden Fragmentes nach Quecksilber-Behandlung. Eine Ligation des MPAA-Thioesters des Thy-1-Fragmentes 1-84 mit **37b** für den Erhalt des vollständigen Glykoproteins Thy-1 (1-120) wurde durch die schlechte Löslichkeit des Thioesters **46** im Ligationsmedium behindert. Eine erneute C-zu-N-Ligation mit der polaren Glykoform **6c** mit dem Fragment **37b** ergab das Glykoprotein **49**, das in einer One-Pot-Reaktion mit $PdCl_2$ behandelt wurde, um das N-terminale Cystein freizusetzen und **50** zu erhalten. Schlussendlich wurde ein Methyl 3-Mercaptopropionat MMP Thioester von Fragment 1-18 (**51**) mit dem Glykoprotein-Fragment **50** ligiert, um das erwünschte Thy-1 Glykoprotein (1-120) mit drei Glykosylierungen zu erhalten.

In dieser Arbeit wurde das Design und die Untersuchung verschiedener Strategien für die Synthese des Glykoproteins Thy-1 gezeigt. Die Konstruktion des 13kDa Glykoproteins erforderte unterschiedliche Schritte und die optimierte Synthese von Aminosäure-Building Blocks für die Festphasensynthese von Glykopeptiden. Mehrere verschiedene Strategien wurden für die Generierung von Glykopeptiden angewendet, darunter die Untersuchung der chemischen Eigenschaften der erhaltenen Fragmente, die Modulation der chemischen Bedingungen für die Ligation der Peptidfragmente für die Erlangung des Ziel-Glykoproteins. Diese Arbeit fokussierte sich auf die Produktion homogener Glykoformen von Thy-1. Die synthetischen Methoden und Protokolle, die im Rahmen dieser Arbeit etabliert wurden, sind jedoch anwendbar auf die Synthese jeglicher Peptide und Glykopeptide und tragen zur chemischen Synthese anderer wichtiger Glykoproteine bei.

1 Introduction

Proteins are a class of macromolecules assembled by a lineal chain of amino acids that play an essential role in living cells. Two aspects determine the role and function of a protein, the sequence of the polypeptide chain and the folding. Thus, it is possible to understand and predict the process the protein is involved in by analysing the amino acid composition and the interactions taking place among peptidyl chains. Thanks to the implementation of new powerful tools, such as molecular dynamics[1] and electronical microscopy,[2] the prediction and characterization of proteins progressed enormously over the past two decades. However, it is not always easy to identify or isolate proteins due to their structural complexity.[3]

The biosynthesis of proteins is genetically controlled by translation.[4] After the assembly of proteins, they can undergo further trimmers called post-translational modifications (PTMs). [5,6] The modifications include the anchor of small and large structures at protein's *N*- and *C-terminus* or the modification of functional groups on the side chains of amino acids.[7] One of the most common PTM is the glycosylation. It involves the covalent attachment of oligosaccharides to the side chains of specific amino acids in the protein forming glycoproteins. The large number and variety of glycoproteins expressed by a single cell reflects the complexity of this class of macromolecules and their involvement in different biological processes.[8,9]

1.1 Glycoproteins: functions and structure

Glycoproteins participate in cell-cell communication, cell growth and differentiation, viral and parasitic infections, immune defence and inflammation, and many other processes.[10–13] Depending on the amino side chain involved in the protein-sugar bond, most natural glycoproteins are classified as O- and N-glycoproteins (figure 1). Although, other linkages are also possible. In eukaryotic cells, proteins can also have a glycophosphatidylinositol (GPI) glycolipid covalently bound to the C-*terminus* of the protein forming the so-called GPI-anchored glycoproteins (GPI-APs).[14,15] GPI-APs have been of interest in different areas and may participate in signal transduction, cell adhesion and immune recognition among other processes.[16]

Despite the importance of these glycoproteins, there is a big gap between the knowledge about the protein and oligosaccharide moieties. While proteins are linear polymers at molecular level, carbohydrates have highly variable structures for the ring size, the type of branching, the anomeric conformation, and chemical modifications (phosphorylation, methylation, acylation, etc.).[17] The unique biosynthetic pathway of glycoproteins defines an additional limit to the fully understanding

of these macromolecules.[18] Unlikely proteins, oligosaccharides are not under the genetic control and are only regulated by the presence and level of expression of the enzymes involved in the process. This lack of control results in a mixture of glycoproteins having heterogeneous glycans, called microheterogeneity of glycoproteins.[19] Isolation of homogeneous glycoforms (proteins having the same backbone and different oligosaccharides) and their structure-activity relationship (SAR) remain one of the most challenging topics in glycobiology.

N-linked glycan **O-linked glycan**

Figure 1 – General structure of O-linked and N-linked glycans.

1.2 N-glycoproteins

N-glycosylation is the most prevalent form of glycosylation in eukaryotic cells and it has been found in proteins from *Archaea* and *Bacteria*.[20] N-glycoproteins contain a glycan bound via amide bond to the side chain of asparagine within the consensus sequence Asn-Xa.a.-Thr/Ser (where Xa.a. is any amino acid different from Pro). All N-glycans share the pentasaccharide core structure Manα1-6(Manα1-3)Manβ1-4GlcNAcα1-4GlcNAc-Asn and most of them contain other monosaccharides added to this core. Based on the structure and composition, N-glycans are classified into oligomannose, complex, and hybrid type (figure 2a).[21]

The high-mannose type glycans contain between two and six α-mannoses bound to the pentasaccharide core. Complex-type glycans accommodate from two (biantennary) to four (tetraantennary) branches at the mannoses having a lactosamine unit (Gal β1-4 GlcNAc) that can terminate with sialic acid residues (figure 2b). Variations on these glycans include the fucosylation of the N-acetylglucosamine bound to the asparagine (GlcNAc-Asn) and/or from the lactosamine unit (GlcNAcβ1-4).[22] The different combination of the glycan core and the branching determines the high structural degree of the complex-type glycans (figure 2b). The hybrid-type N-glycans have mannose residues on the Man α1-6 of the sugar core (as in the high-mannose type glycans), and one or two *antennae* bound to the Man α1-3 are similar to the complex-type glycans.

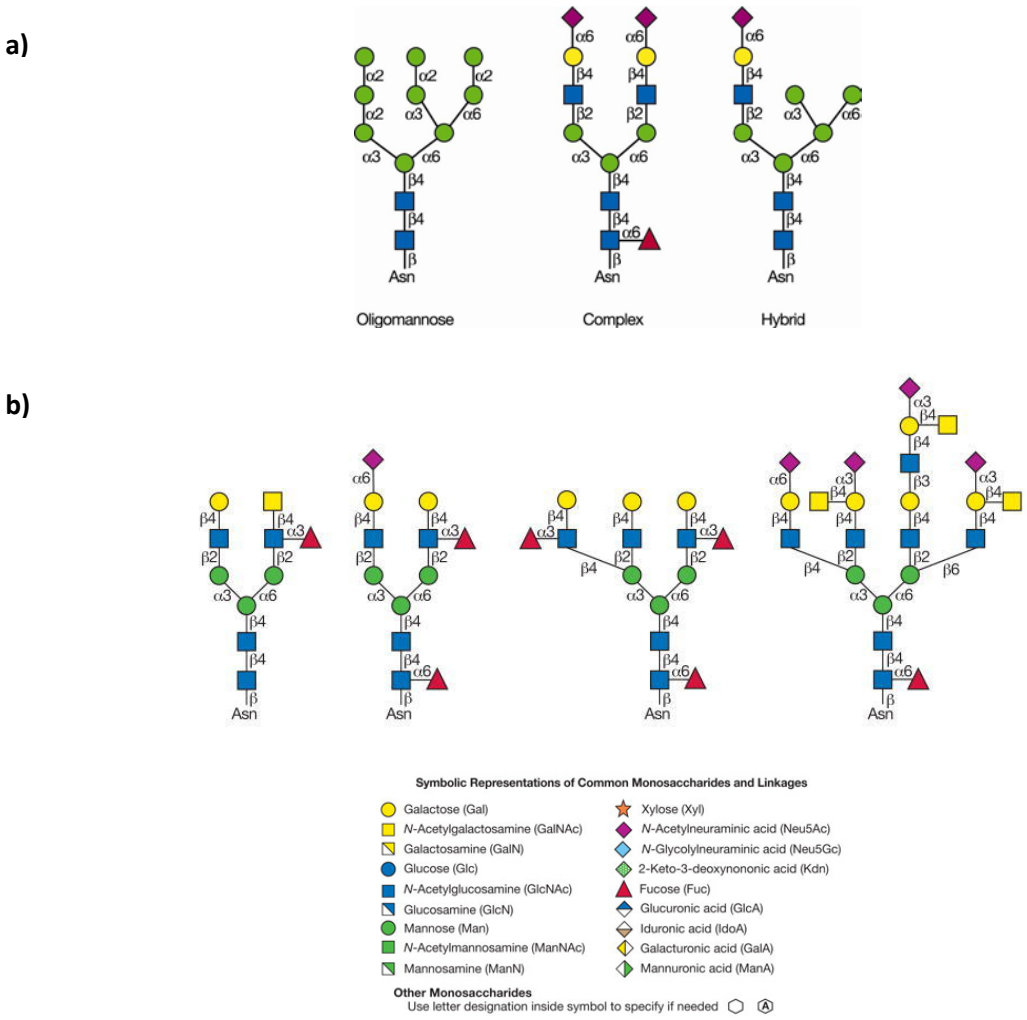


Figure 2 - a) General structures of N-linked glycan types. b) Examples of complex-type glycans found on glycoproteins [23]

The structural diversity of carbohydrates not only reflects the different localization of N-glycoproteins, it also affects the conformation and the biological functions of the glycoprotein.[24] It has been suggested that glycans of intracellular N-glycoproteins control the correct folding of proteins involving the chaperone system called calnexin-calreticulin cycle.[25] The glycans of extracellular N-glycoproteins play a role in the binding of these molecules as ligands for receptors: e.g. the binding of the glycoprotein erythropoietin (EPO) to the erythropoietin receptor (EpoR) is the initial step for the red blood cell production.[26] Many progresses have been done for deciphering the participation of the glycans of N-glycoproteins in different processes. For instance, the elucidation of the structure of carbohydrates decorating the S2 subunit of the coronavirus SARS-CoV-2 may provide insights for the design and synthesis of a vaccine against COVID-19.[27] Beyond the recent technological development, the information available about the role of carbohydrates is still limited due to the difficult access to homogeneous N-glycoproteins.

1.3 Biosynthesis of N-glycoproteins

The biosynthesis of N-glycoproteins takes place in the ER and Golgi compartments and involves different enzymes (figure 3). The process starts in the ER with the attachment of the tetradecasaccharide $Glc_3Man_9GlcNAc_2$ to a dolichylpyrophosphate.[28] This glycolipid is the precursor for all N-glycoproteins and contains the pentasaccharide core for the assembly of high-mannose, complex and hybrid type glycans.[21] In the next step, an oligosaccharyltransferase (OST) recognizes the specific sequon Asn-Xa.a.-Thr/Ser on the newly synthesised polypeptide and transfers the oligosaccharide from the dolichol glycolipid to the protein.[29,30] The addition of the N-linked glycan occurs co-translationally, meaning before the folding of the protein.[31] The terminal glucose and mannose of the tetradecasaccharide on the nascent N-glycoprotein are removed by a glucosidase and a mannosidase that are in the ER.[32] The temporary re-addition of the glucose mediated by an N-acetylglucosamine-1-phosphotransferase and the subsequent removal of the mannose on the central arm by a mannosidase contribute to the protein folding. If the glycoprotein is correctly folded, it leaves the ER and is transported by vesicle to the Golgi apparatus.[33] By misfolding of the glycoprotein, an enzyme called EDEM (ER Degradation Enhancing α-Mannosidase I-like protein) recognizes the mistake and activates the degradation of the glycoprotein.[34] In the cis Golgi *cisternae* α-mannosidase removes further mannose residues of the glycoprotein to produce the main intermediate for the assembly of hybrid and complex-type N-glycans. Some glycoproteins skip this process giving the precursors for the oligomannose type N-glycans. Glycoproteins moving to the medial Golgi *cisternae* with N-glycans targeted as N-complex type receive an N-acetylglucosamine, get removal of two mannoses and the addition of a second N-acetylglucosamine.[35] A fucose can be added on the N-acetylglucosamine bound directly to the protein. At this point, N-acetylglucosaminetransferases catalyses the addition of N-acetylglucosamine to the (α1-3)- and (α1-6)-mannoses producing tri- and tetra-antennary N-glycans.[36] If the α-mannosidase responsible for the hydrolysis of the mannose from the octasaccharide intermediate leaves this glycan unmodified, the precursor for the hybrid type N-glycan is generated. Bisecting N-glycan structures on complex and hybrid type N-glycans are the product of the catalysed addition of a GlcNAc to the β-mannose of the core.[21] Final oligosaccharide modifications, including the addition of fucose and sialic acid residues as capping sugars, take place in the trans Golgi *cisternae*. Hybrid and complex type N-glycans can undergo the addition of N-acetylglucosamine and galactose residues to yield terminal lactosamine oligosaccharides.[23]

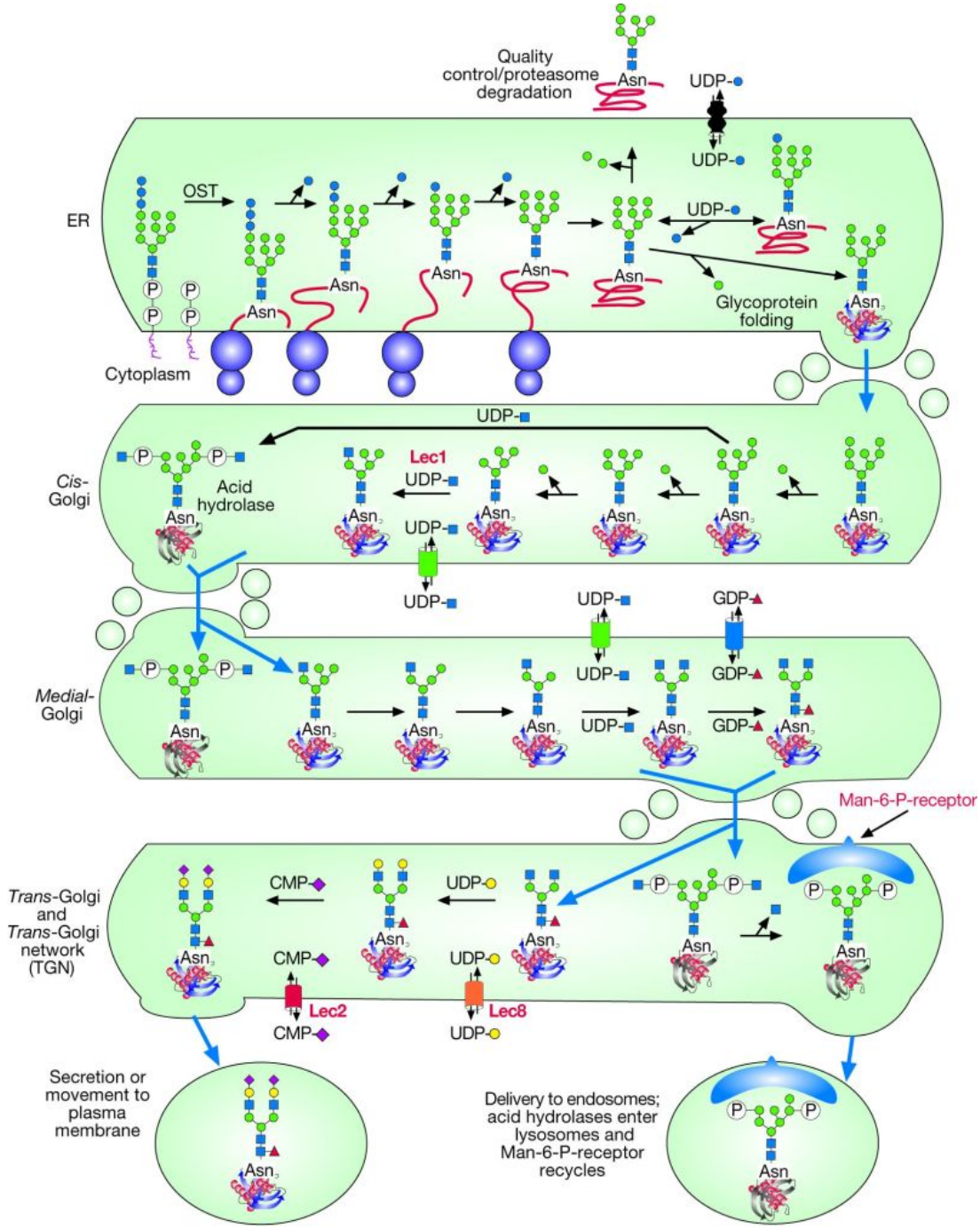


Figure 3 - Biosynthesis of N-glycoproteins in eukaryotic cells.[23]

The availability of enzymes in the Golgi apparatus, the residence time of the glycoprotein in the ER, and the different activity of the enzymes acting on the same substrate are responsible for the heterogeneity of expressed N-glycoprotein.[37] Moreover, it is nearly impossible to control each single step of the biosynthesis. Thus, alternatives need to be investigated and used to obtain homogeneous isoforms of N-glycoproteins.

1.4 Chemical and chemoenzymatic synthesis of N-glycoproteins

Different strategies have been reported for the N-glycoproteins synthesis, including total chemical synthesis,[38–40] chemoenzymatic protocols,[41,42] and recombinant methods.[43,44]

a) Recombinant expression of N-glycoproteins from mutant cell lines

b) N-linked glycoprotein remodelling

c) Total chemical synthesis and semi-synthesis of N-glycoproteins

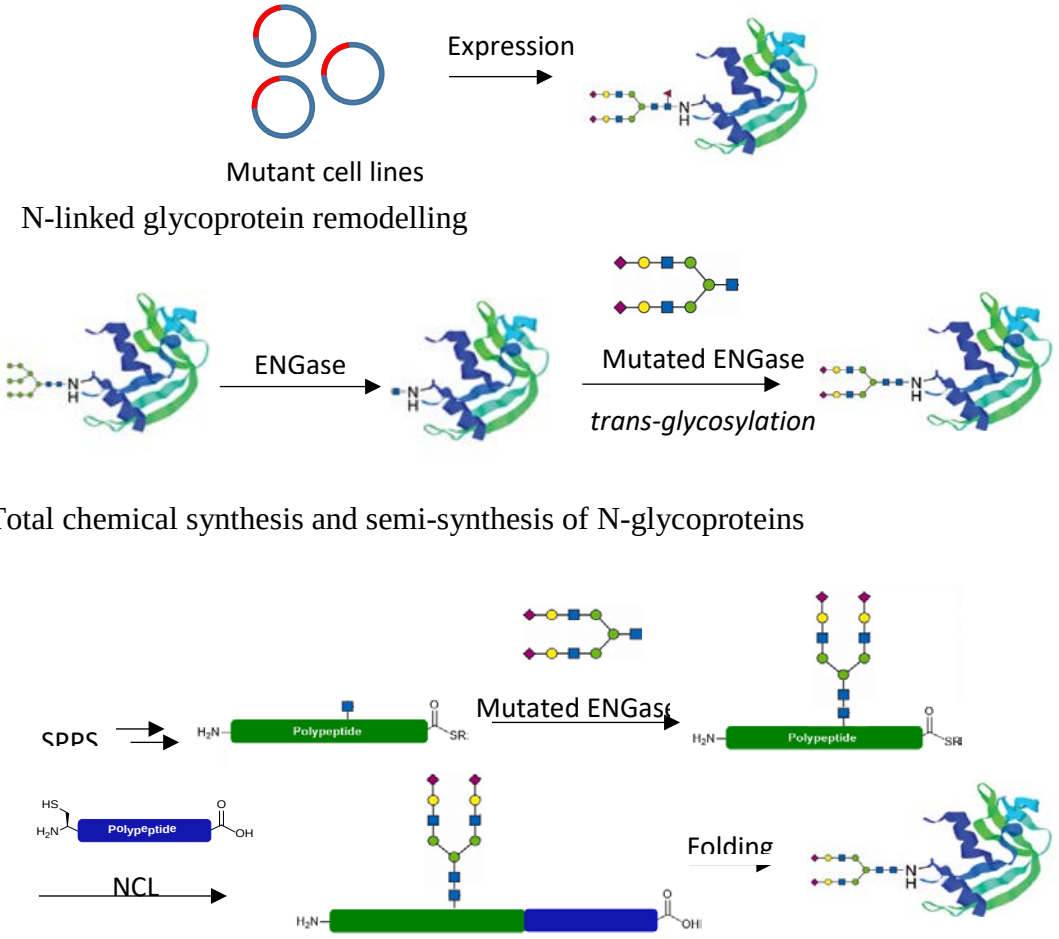
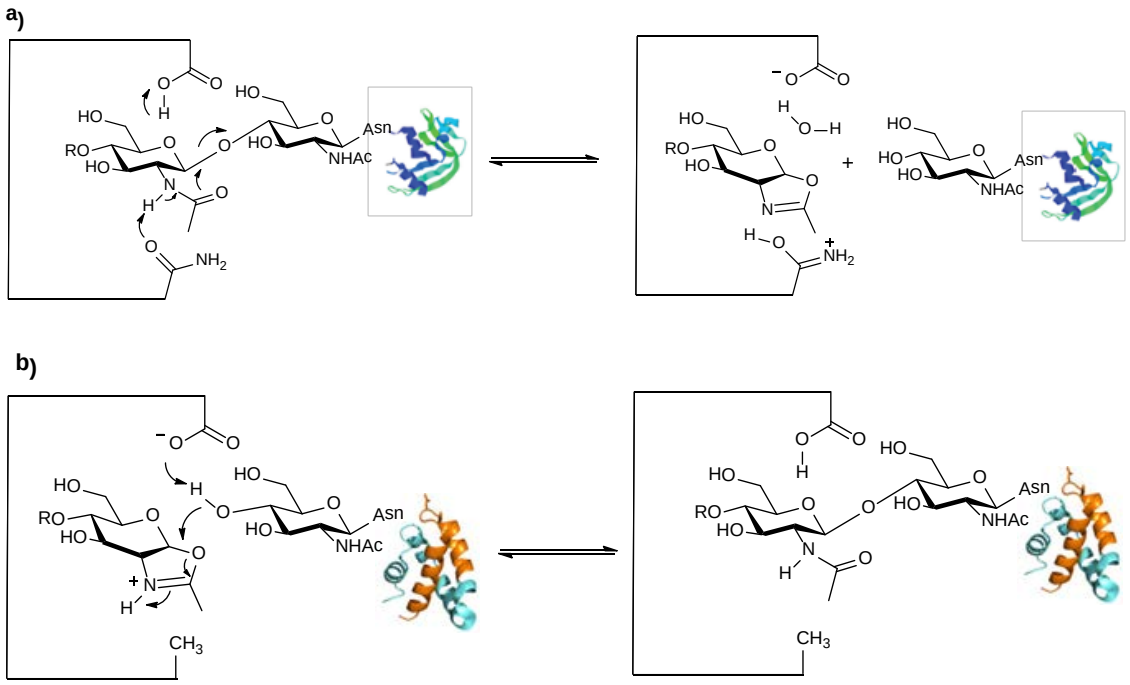


Figure 4 – Strategies for the generation of homogeneous isoforms of N-glycoproteins. a) Recombinant method; b) chemoenzymatic method; c) total chemical synthesis. SPPS = Solid Phase Peptide Synthesis, ENGase = Endo-β-N-acetylglucosaminidase, NCL = Native Chemical Ligation.

Recombinant methods involve the genetic modification of the glycosylation pathway by removing, changing or adding enzymes to increase the expression of a single N-glycoprotein. An example of this technology is the engineering of α2-3-sialyltransferase in Chinese Hamster Ovary (CHO) cells performed by the Krummen group to increase the amount of N-glycoproteins with a large fraction of fully sialylated products.[45]

Another method to obtain homogeneous N-glycoproteins uses a class of enzymes called endo-β-N-acetylglucosaminidases (ENGases). These enzymes are able to hydrolyse N-glycan between the

two N-acetylglucosamine residues yielding the protein with an N-acetylglucosamine in each glycosylation site and N-glycan oxazolines (scheme 3a).[46] The reaction involves an acidic and a basic residue in the active site of the enzyme and is pH-depending. By mutation of one of the amino acids on the active site (*e.g.* glutamine is replaced by alanine), the enzyme loses its hydrolase activity and catalyses the reverse reaction, a transfer of glycans (scheme 3b).[47] The generated N-glycan oxazolines are the substrate in the reverse-hydrolysis reaction or "trans-glycosylation" of proteins and they have been used in chemoenzymatic remodelling approaches for the generation of homogeneous glycoforms of glycoproteins.[48]

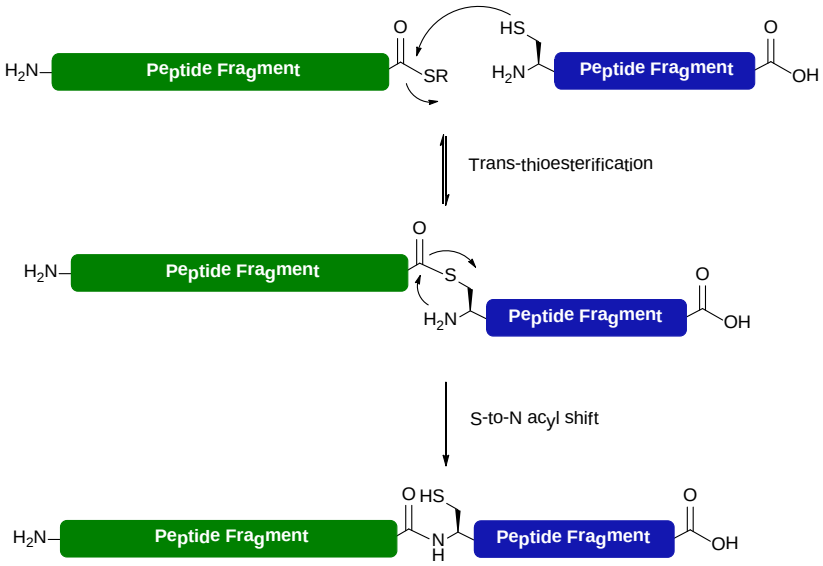


Scheme 3 - Mechanism of action of ENGases for a) hydrolysis, and b)transglycosylation

This technique is receiving special attention for remodelling of monoclonal antibodies (mAbs) due to their clinical applications. In this context, Wang *et. al.* reported the application of mutated ENGases S2 to catalyse the trans-glycosylation of mAbs rituximab and trastuzumab, two antibodies for the treatment of follicular lymphoma and breast cancer.[49,50] Although these mutated enzymes appear to be a promising tool for the production of N-linked glycoproteins, there are some limitations such as the glycan substrate specificity of the enzymes, the reaction of N-glycan–oxazolines involved in the trans-glycosylation process with amino groups present on the protein and the different reactivity of N-acetylglucosamine residues at different glycosylation sites on the protein.

1.5 Ligation methods for the assembly of N-glycoproteins

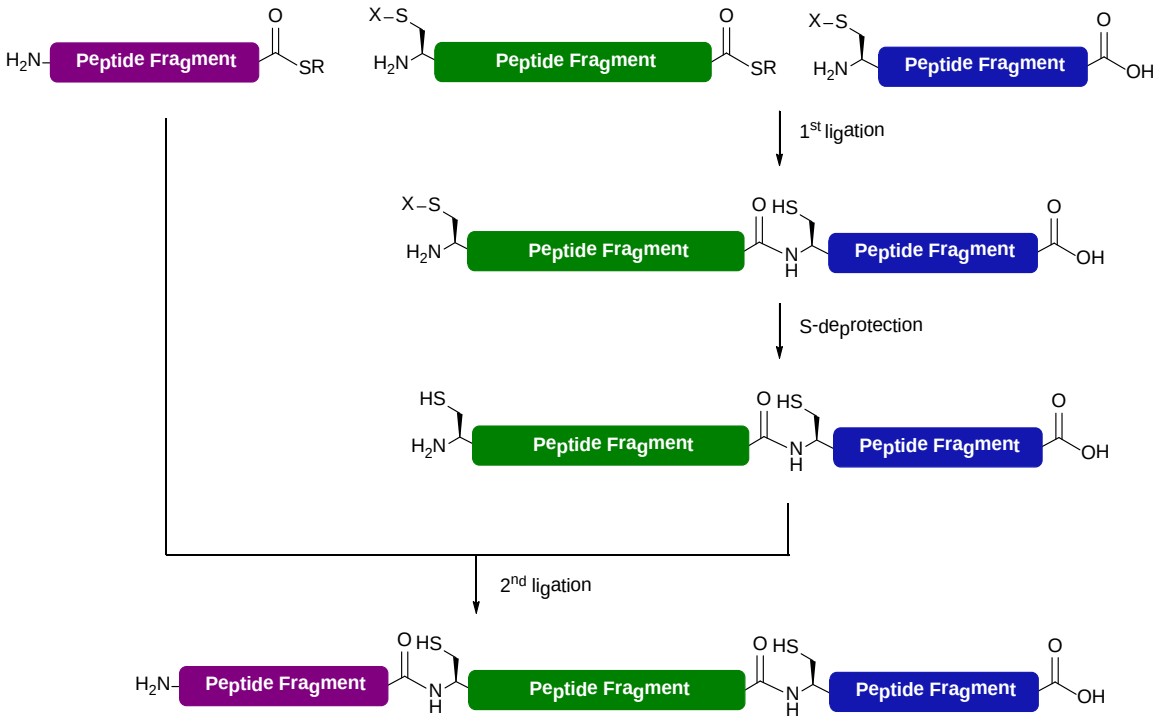
The chemical synthesis of polypeptides and glycopeptides and their subsequent conjugation is a reliable alternative for generating homogeneous N-glycoproteins (figure 4c). Since native chemical ligation (NCL) was developed by Kent in 1994, the synthesis of large proteins and glycoproteins have been reported in the literature.[51–54] The NCL involves a chemo-selective reaction between two unprotected peptide fragments, one having a cysteine at the N-*terminus* and the other having a thioester at the C-*terminus.* The traditional N-to-C ligation comprises two steps, a reversible trans-thioesterification between the thiol of the N-terminal cysteine and the thioester group of the other fragment, and the following intramolecular and irreversible S-to-N acyl shift to generate the new amide bond (scheme 4).



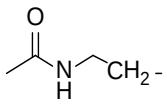
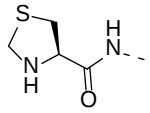
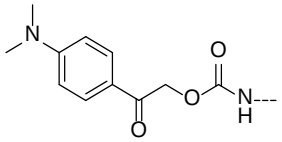
Scheme 4 - Native chemical ligation of two unprotected peptide fragments.

The introduction of temporary protecting groups on the terminal cysteine residues facilitates the ligation of the peptides and has found solid application in synthesis involving sequential ligation reactions, and in one-pot ligation of three or more fragments.[55] Some of these groups includes the acetoamidomethyl (Acm) group, the 1,3-thiazolidine-4-carboxo (Thz) group and others. The Acm-group was introduced by Veber, it mask the thiol of the terminal cysteine during the ligations and can be removed by mercury(II) acetate.[56] The Thz-group can be removed with hydroxylamine and was first reported by the Kent's group in the assembly of the protein crambin.[57] In general, temporary groups can be distinguished depending on the conditions employed for their removal (change of the pH, reaction with specific reagents, UV light irradiation), and if one of these conditions interfere with the subsequent ligation, a different group should be selected. Moreover,

these temporary groups are used to protect internal cysteines. Table 1 lists some protecting groups for the protection of cysteine residues in the synthesis of proteins using ligations.



Scheme 5 – Iterative native chemical ligation of three peptide fragments.

Protecting group	Structure	Conditions
S-Amidomethyl	 Acm	• 1) $Hg(OAc)_2$, 2) thiol excess[58] • 1) $PdCl_2$, 2) DTT[59] • 1) Ag(OAc), 2) DTT[60,61]
Thiazolidines	 Thz	• 0.2-0.4 M $MeONH_2$•HCl, pH 4.0[60,62] • 1) $PdCl_2$, 2) DTT[63] • 1) $[Pd(allyl)Cl]_2$, 2) DTT[63,64]
Disulphide	 S^tBu	Thiol or TCEP[65–67]
Carbamates	 Mapoc	Photolysis at $\lambda > 300$ nm[68]

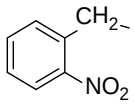
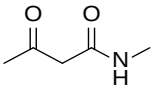
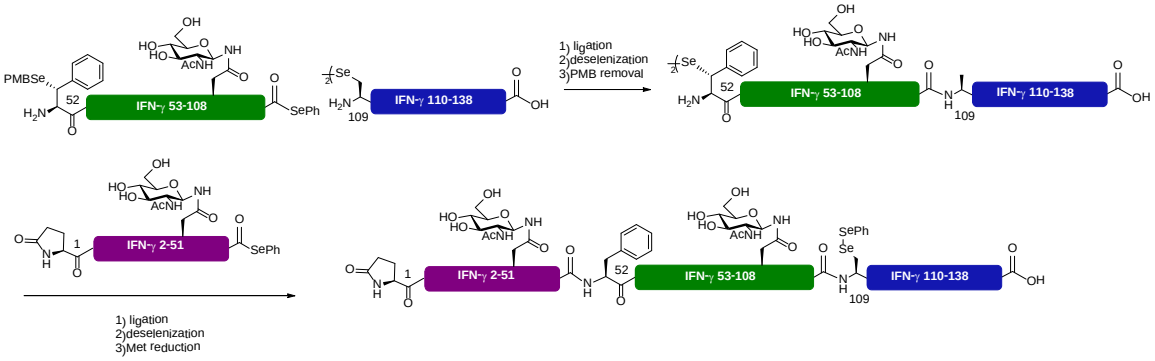
Thioethers	 S-(2-nitro)benzyl	Photolysis at λ = 365 nm[69,70]
Amides	 AcA	NH_2OH/aq., AcOH[71,72]

Table 1 – Protecting groups of cysteine

The main limitation of the native chemical ligation is the low abundance of cysteine in natural proteins (1-2%). Thus, impressive progresses have been made in the development of techniques that bypass the use of cysteine residues in this process. Different β-mercapto amino acids have been synthesised and employed as cysteine surrogates at N-*terminus* of peptides.[73,74] These surrogates undergo the ligation reaction similar to cysteine and they are submitted in the following step to a desulfurization process that restores the corresponding residue at the junction site. [75,76] Similarly, β-seleno amino acids have been used to mimic cysteine residues at the N-*terminus* for the assembly of N-glycoproteins (figure 5a).[77] The Payne group has recently reported the synthesis of glycosylated interferon-γ (IFN-γ) using iterative diselenide-selenoester ligation-deselenization of two N-glycosylated peptide fragments and one peptide segment.[78]

Another method to circumvent the need of cysteine at the junction point in ligations involves a chemoselective reaction between a peptide fragment having a salicylaldehyde at the C-*terminus* and a fragment ending with a serine or threonine residue at the N-*terminus* (figure 5b).[79] The N,O-benzylidene intermediate is hydrolysed in acidic conditions to yield the amide bond. Li and co-workers used this approach for the convergent synthesis of the N-linked glycoprotein interleukin-25.

a) Selenoester ligation[78]



b) Ser/Thr ligation (STL)[81]

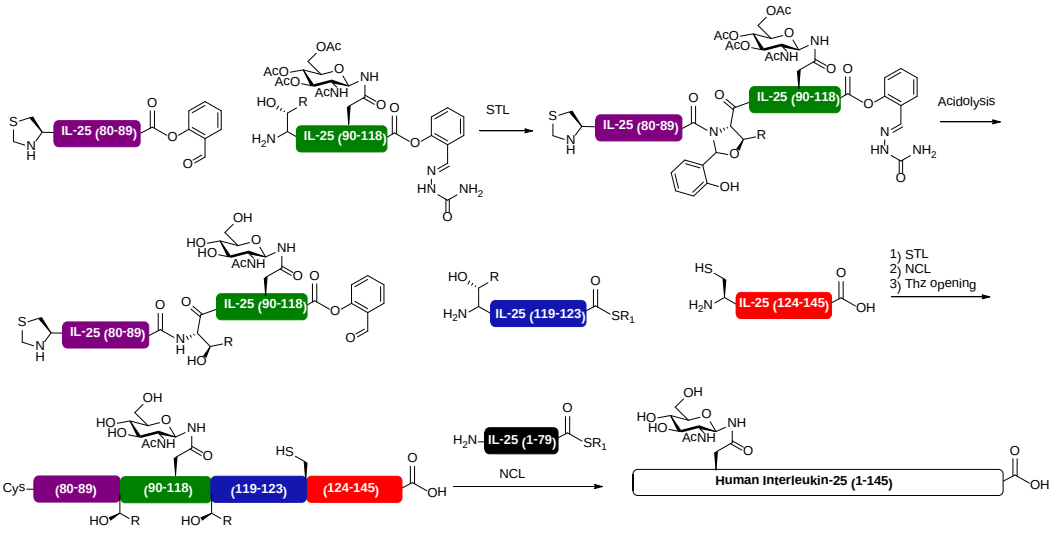


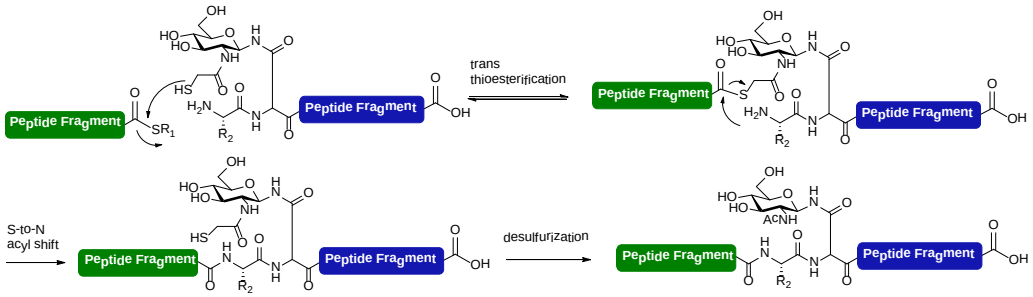
Figure 5 - Strategies for a) the synthesis of Interferon-γ (IFN-γ); b) the synthesis of interleukin-25.

Structures with auxiliary thiol groups have been also used in Sugar-Assisted Ligation (SAL). In this process, the sugar contains a thiol function that participates in the trans-thioesterification reaction with the thioester (figure 6a).[80] The process involved a peptide fragment carrying an N-acetylglucosamine modified with a thiol group at the C2 of the acetyl group. After the rearrangement to the ligation product, this group is removed under reductive conditions.

Conceptually similar to SAL, the dual native chemical ligation is another approach for NCL (figure 6b). This strategy uses a thiol group that is part of an aromatic moiety attached to the anomeric amine of the GlcNAc.[81] The process relies on two ligation steps. The first step is the ligation of GlcNAc having an anomeric trimethoxybenzyl group attached to 3-nitro-2-pyridinesulfenyl with a peptide fragment having an N-terminal aspartate with a thioester at the β-carboxylate. After

formation of the peptide-sugar amide bond, a second ligation with another peptide thioester fragment delivers the ligated and rearranged N-glycopeptide product. The trimethoxybenzyl group is cleaved from the sugar-peptide amide under acidic conditions and the glycan is deacetylated to deliver the desired glycopeptide.

a) Sugar-assisted ligation (SAL)[80]



b) Dual native chemical ligation[81]

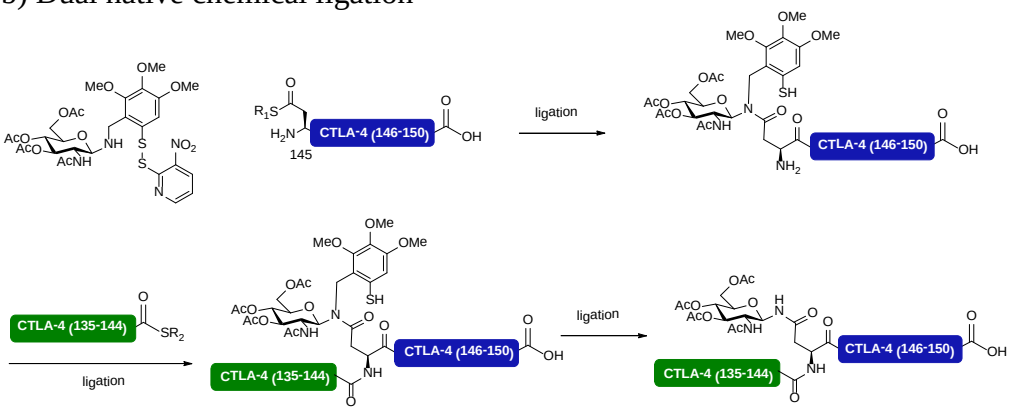
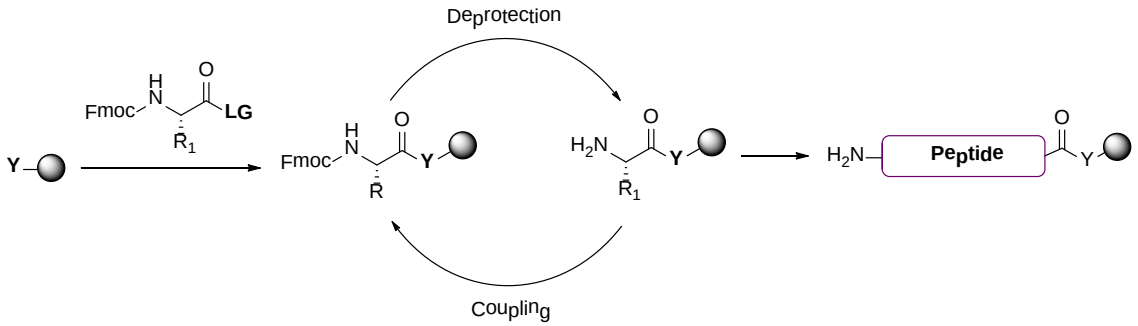


Figure 6 – a) Sugar-assisted ligation (SAL); b) Dual native chemical ligation.

1.6 Solid Phase Peptide Synthesis (SPPS)

In the past decades, peptides gained crescent interest and have been widely employed in pharmaceutical research[82,83], biomedicine[84], and drug discovery[85]. As a consequence, peptide synthesis improved enormously giving access to long and difficult sequences with high yields and in short times. After the first synthesis of a 18-mer peptide by Fisher,[86] a second important milestone in this field is the introduction of the solid phase peptide synthesis (SPPS) by Merrifield in 1963.[87,88] This revolutionary method, comprising the sequential coupling of N-protected amino acids to an inert solid support, removed the need of intermediates purification and simplify the

process by an easy elimination of the reagents excess and soluble side products, facilitating the assembly of long peptides (scheme 6).



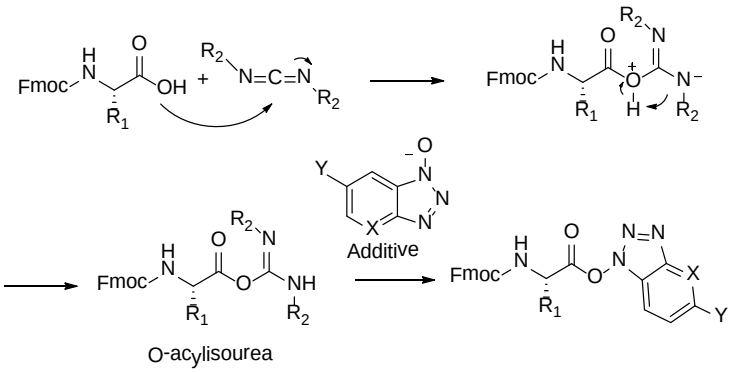
Scheme 6- Solid phase peptide synthesis.

The SPPS requires two main steps that are repeated for each amino acid cycle. The removal of a temporary protecting group from the previous amino acid (or from the resin) and the coupling of an amino acid. Each coupling cycle requires the activation of the amino acid, generally by formation of "stable" active esters. Carbodimide-based reagents have been widely used to activate amino acids[89]; however, they are generally employed in combination of auxiliaries such as hydroxybenzotriazole (HOBt)[90] or ethyl 2-cyan-2-(hydroxymino)acetate (Oxyma)[91] to avoid the rearrangement of the formed O-acylisourea into the unreactive N-acylurea.

Carbodiimides		
DCC[89]	DIC[92,93]	EDC[92]
Auxiliary additives		
HOBt (X = H)[92] Cl-HOBt (X = Cl)[94]	HOAt[95]	Oxyma[91]

Table 2 – Exemples of carbodiimides coupling reagents and additives

The activation with carbodiimides comprises a deprotonation of the carboxylic acid by a base and the concomitant formation of the O-acyl urea. In the following step, a nucleophilic addition of the auxiliary and the elimination of the urea deliver the corresponding active esters (scheme 7).



Scheme 7 – General mechanism for the activation of amino acid C-*terminus* with carbodiimides and auxiliary additives.

A second big family of reagents used in SPPS are the so-called *in situ* coupling reagents. These compounds are mainly HOBt derivatives and include phosphonium salts, aminium salts, and uranium salts.

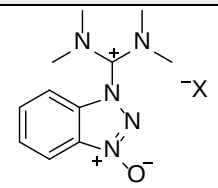
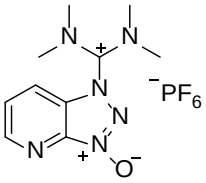
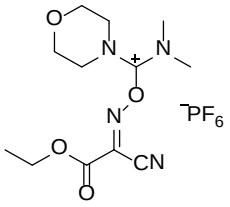
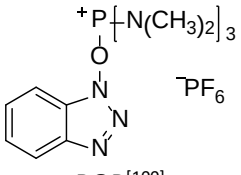
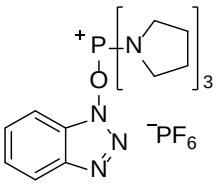
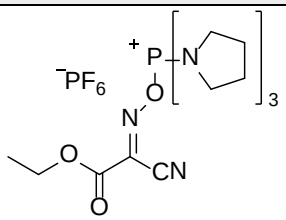
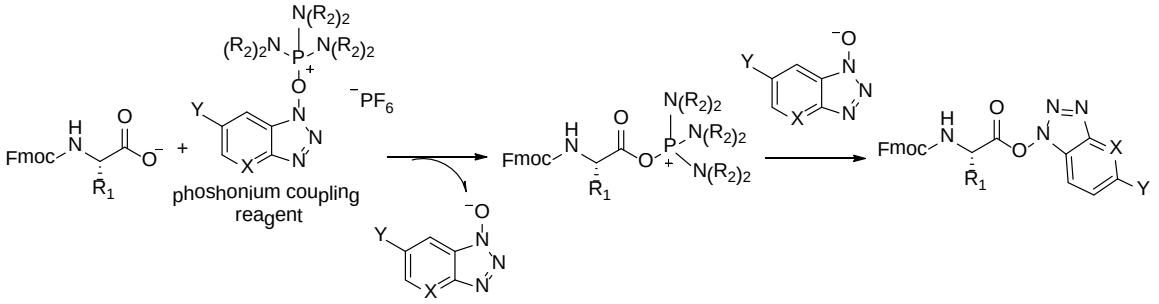
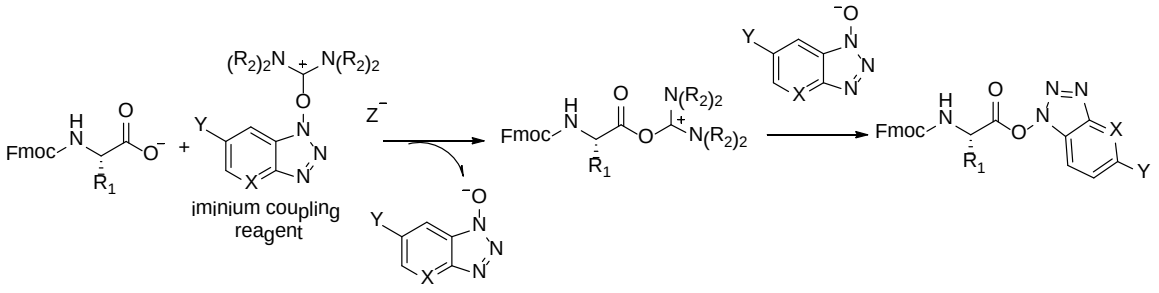
Aminium reagents		**Uronium reagent**
 TBTU (X = BF_4^-)[96] HBTU (X = PF_6^-)[97]	 HATU[98]	 COMU[99]
Phosphonium reagents		
 BOP[100]	 PyBOP[101]	 PyOxim[102]

Table 3 – Exemples of aminium, uranium, and phosphonium reagents

The activation process with these compounds requires the deprotonation of the carboxylic by a base, and the following reaction of the carboxylate to form active species that can either react directly with an amino group or form an activated ester with the HOBt-derivative.[103]



Scheme 8 - Mechanism for the activation of amino acids using phosphonium reagents.



Scheme 9- Mechanism for the activation of amino acids using phosphonium and iminium reagents.

The solid support used in SPPS is generally a polymeric resin carrying a spacer or linker with a functional group. The most common supports are resins of polystyrene cross-linked with 1-2% divinylbenzene[104] and polyethylene glycol-based resins (PEG).[105] The material differs in the physico-chemical properties of the beads such as the size and swelling in different solvents.[106] These resins are commercially available and they are accessible with different linkers for anchoring the first amino acid and obtain peptide acids, amide, thioester or hydrazide.[107] The type of linker-resin combination requires a careful selection prior to each peptide synthesis and it depends on the type and amount of peptide to be synthesized, the properties of the polymer in the solvents used in the synthesis, the type and substitution degree of the linker, and the chemical conditions for the release of the peptide from the solid support.

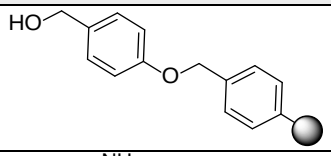
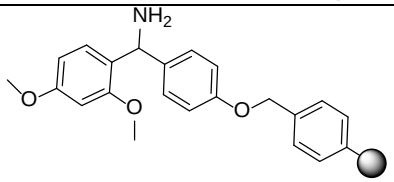
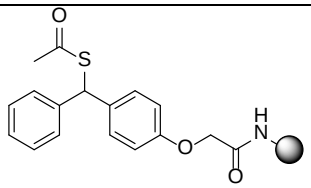
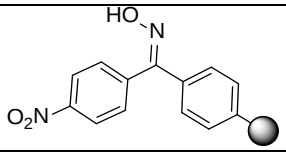
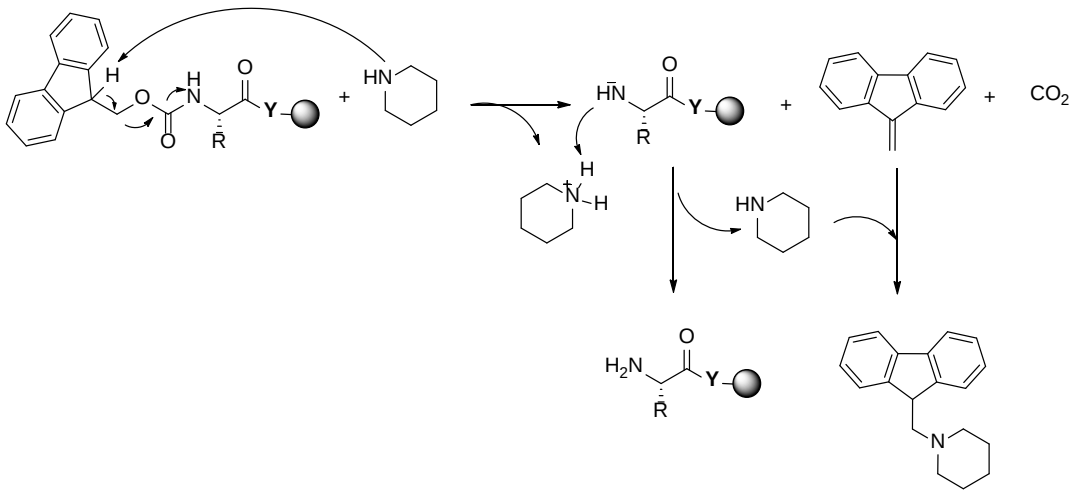
Name	Structure	C-terminal after cleavage
Wang resin[108]		Carboxylic acid
Rink amide resin[109]		Amide
MBPA resin[110]		Thioacid
Kaiser oxime resin[111]		Hydrazide

Table 4 – Examples of linker-resins used in SPPS.

The type and combination of protecting groups on the amino acid side chains are another important aspect to consider in the synthesis design. Two groups generally used for protection of the N-*terminus* of α-amino acids are the fluoren-9-ylmethyloxycarbonyl (Fmoc)[112] and the tert-butoxycarbonyl (Boc).[113] The Fmoc is the most used due to the mild basic conditions required for its removal (scheme 10)[114]. By using the Fmoc-strategy, the side chains of amino acids are orthogonally protected with acid-labile groups that are mostly removed simultaneously with the cleavage of the peptide from the solid support.[115] Alternatively, other groups such as acetoamidomethyl (Acm), allyl or allyloxycarbonyl (Alloc) are introduced on specific amino acids and removed under reductive conditions before or after the peptide is cleaved from the resin (table 1). The removal of protecting groups is generally associated with active species, such as carbocations, that can react with some amino acids and functional groups on the peptides. Thus, different scavengers are employed in the cleavage cocktail to capture these species from the side chain protecting groups.[116] These scavengers are mostly nucleophilic compounds such as water, silanes, thiols and phenol-derivatives.



Scheme 10 - Mechanism of the removal of Fmoc protecting group under basic conditions.

Solid phase synthesis has several advantages: 1) the unreacted building blocks are removed by washing and filtering the solid support; 2) coupling or deprotection steps can be repeated to improve the yields; 3) the desired peptide can be easily released from the support using different chemical conditions; 4) SPPS is suitable for automation. On the other hand, limitations of solid phase synthesis include the aggregation of growing peptides[117,118] and side reactions during the coupling-deprotection steps.[119–121] Inter- and intra-molecular interaction can be disrupted by increasing the reaction temperature or selecting a resin with a low degree substation.[122,123] Different progresses have been made to overcome the occurring side reactions such as the δ-lactame formation affecting the arginine or the epimerization of cysteine and histidine residues.[124,125] However, there are still sequences that are difficult to synthesize: unsolved side reactions represent the major problem in the synthesis of peptides.[126] One of them is the formation of aspartimide, a side reaction concerning the cyclization of aspartic acid.

1.7 Thy-1- A Glycosylated and GPI-anchored Protein

Many proteins contain multiple and different PTMs. Thy-1 or CD90 is a membrane GPI-anchored glycoprotein belonging to the Ig V-type superfamily domain located in the outer leaflet of lipid raft.[127] Discovered in the earlies 1960 in mouse thymocyte cells, Thy-1 was also detected in rodent and human species over the past decades.[128–130] One third of the molecular mass of Thy-1 corresponds to the glycans of the (two and three) N-glycosylation sites reported in human and mouse cells.[131,132] The composition of these glycans differs between cells and tissues.[133] The Thy-1 precursor is translated in humans as a 161 amino acids protein, 160 in rat, and it contains a signal peptide (residues 1-19) and a transmembrane sequence (residues 131-162) that are removed

during the attachment of the GPI anchor.[134] Two disulphide bonds between the cysteine residues at positions 28-13 and 38-104 characterizes this V-like immunoglobulin domain.[135] Several studies showed the involvement of Thy-1 in processes like cell-cell and cell-matrix interactions, nerve regeneration, cancer, apoptosis, inflammation, and fibrosis.[136] It has been speculated that the expression of Thy-1 has been linked to different cancer types such as liver, myeloid, skin and brain. Despite its physiological and pathological relevance, the biological function of Thy-1 remains unclear, mainly due to the difficulties to obtain homogeneous glycoforms of the glycoprotein.

To investigate the biological activity of the protein and determine the role of the glycans in the structure-activity relationship of Thy-1, there is a need of obtaining variants of this protein in homogeneous form and having different modifications.

1.8 Aim of the thesis

The aim of this thesis was to establish a strategy and synthesize a glycosylated human Thy-1 N-glycoprotein using sequential native chemical ligation. Three peptide fragments were designed to obtain the 110 amino acids protein ready for the attachment of a glycosylphosphatidylinositol glycolipid and installation of up to three different large N-glycans using enzymatic trans-glycosylation (figure 7).

The fragments were designed based on the position of the cysteine residues on the natural glycoprotein, fragment 1 corresponding to the amino acids (1-18), without glycosylation; fragment 2 corresponding to the residues (19-84) with two glycosylations, and fragment 3 corresponding to residues (85-110) with one glycosylation site. To elongate the glycans using enzymatic trans-glycosylation, the synthetic fragments included a protected or unprotected N-acetylglucosamine that can serve as acceptor for mutated endo-β-N-acetylglucosaminidases (Endo X enzymes, some of them are listed in the table 5). To allow the attachment of a synthetic GPI-anchor via native chemical ligation, the protein fragments were designed to deliver a glycoprotein containing a thioester precursor at the C-*terminus*.

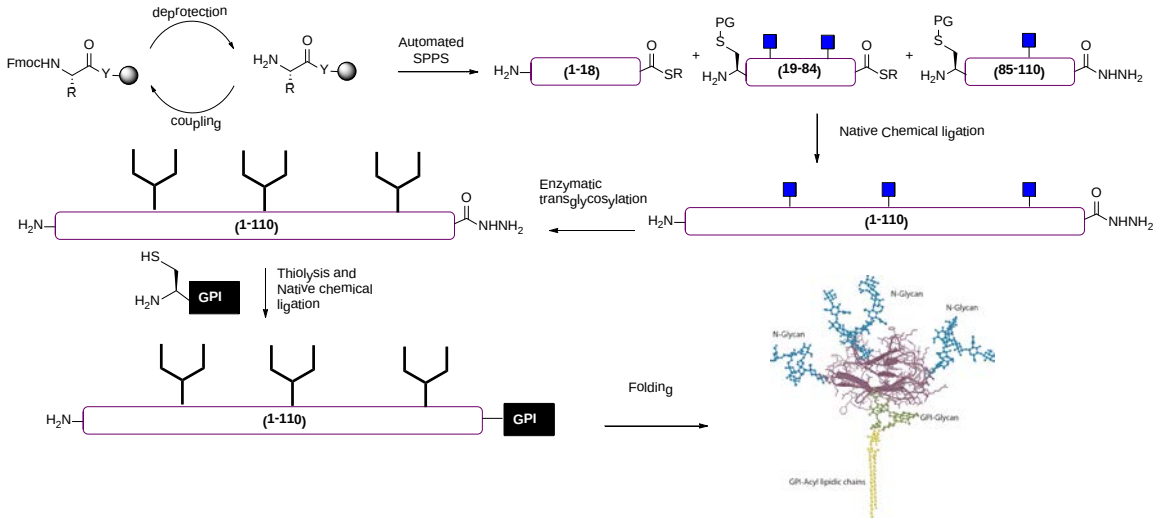


Figure 7 – General strategy for the total chemical Synthesis of GPI-anchored N-glycoprotein Thy-1.

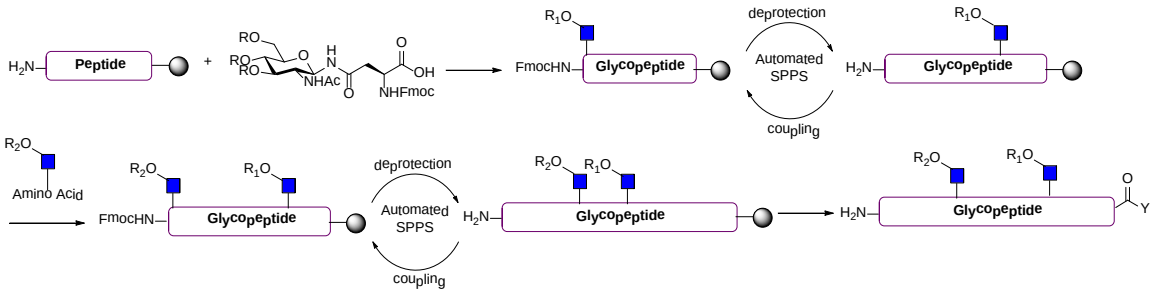
Mutant enzyme	Donor Oxazoline structure
Endo A N171A	High mannose/hybrid
Endo A E173H	
Endo M N175Q	High mannose/hybrid Biantennary complex
Endo M N175Q/W251N	
Endo D N322Q	Core structure only
Endo F3 D126A	Biantennary complex Triantennary complex
Endo S D233A	Core structure Biantennary complex
Endo S D233Q	
Endo S2 D184M Endo	High mannose/hybrid Biantennary complex
Endo CC1 N180H	Biantennary complex

Table 5 – List of the most used mutant ENGases[48].

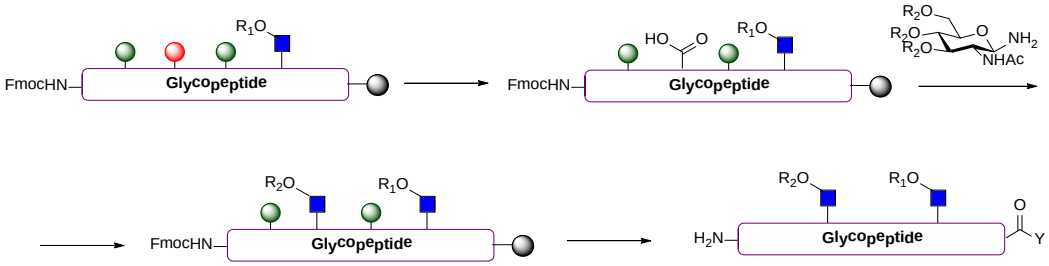
To investigate and optimize the assembly of the glycoprotein either from N- to C-*terminus* or from C- to N-*terminus*, all peptides were designed as hydrazides precursors. To synthesize the complex glycosylated fragment (19-84) different strategies were designed and evaluated. They included the linear synthesis on the cassette-based and a convergent method[137]. The cassette-based method involves the incorporation of a pre-synthesised glycosylated asparagine unit and required the

optimization of the coupling conditions for the elongation of the designed peptide. The convergent method involves the site-specific glycosylation of the pre-assembled peptide fragment.

a) Cassette-based method



b) Convergent method



Scheme 11 - Strategies for the synthesis of N-glycopeptide fragments.

2 Results and discussion

The synthesis of different glycoproteins has been reported over the past decades.[67,138,139] However, it is still a challenging task to overcome the numerous difficulties that have plagued their general access by chemical synthesis. One of the main problem is the generation of the carbohydrate part and the complex chemistry involved in the synthesis of oligosaccharides.[17] The presence of many functional groups and the high degree of stereo- and regiocontrol have made this branch of organic chemistry very challenging.[140] Further, the synthesis of proteins can be equally challenging and problematic as that of oligosaccharides.[123] Although improvements have been made in automated SPPS, it is quite difficult to synthesize peptides that contain more than 40 amino acids due to the number of truncated sequences and other by-products as the sequence increases in length.[103,141]

In 1984, the seminal work by Kent et *al*. facilitated the assembly of long peptides and proteins by a technique they coined Native Chemical Ligation (NCL).[104] Another significant advancement was the exploitation of enzymes for the construction of oligosaccharide structures.[48] Considering these state-of-the-art, the synthesis of a GPI-anchored glycoprotein Thy-1 was designed considering enzymatic methodologies and chemical synthesis (scheme 1).

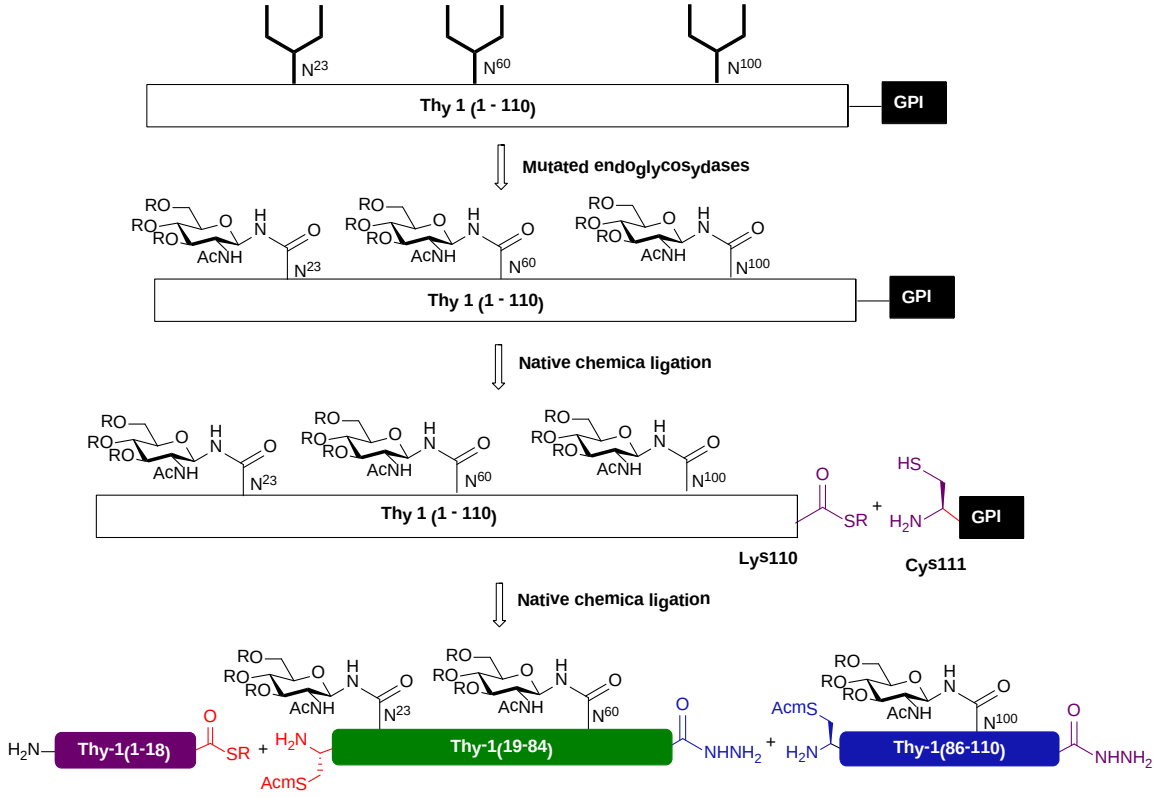
2.1 Retrosynthesis analysis and strategy design

Mutated ENGases were considered for the trans-glycosylation of N-acetylglucosamine residues at each glycosylation site (Asn23, Asn60, and Asn100) of the glycoprotein. The glypiation of the Thy-1 was outlined using a native chemical ligation between a C-terminal thioester of the Thy-1 glycoprotein (1-110) and a cysteine attached to the glycolipid. A strategy involving sequential native chemical ligation using Cys19 and Cys85 as junction points was selected for the assembly of the full-length glycoprotein Thy-1 (1-110).

Three peptide fragments were necessary for the assembly of the protein: a non-glycosylated fragment (1-18), the glycopeptide (19-84) having two N-glycosylated residues, and the glycopeptide (85-110) containing an N-glycosylated residue at Asn100. One important consideration in the design of the strategy comprised the use of different building blocks at the glycosylated positions that will allow differentiation between the three glycosylation sites and the synthesis of diverse Thy-1 glycoforms. Installation of functional groups on the glycosylated residues in the sequence that are orthogonal removable after the assembly of the protein may allow

specific elongation of the N-acetylglucosamine units using trans-glycosylation with enzymes. For this purpose, two synthetic routes were considered to obtain the fragment (19-84) having either an O-*per*-acetylated glucosamine or unprotected glucosamine on the Asn23 and Asn60.

The success of the ligation required the three fragments in high yield and purity. For this purpose, different protocols, and conditions were envisaged for the synthesis of the designed fragments.



Scheme 12 - Retrosynthesis of glycoprotein Thy-1 (R = -H or –Ac).

2.2 Solid-Phase Peptide Synthesis of the designed peptide and glycopeptide fragments

Solid-phase peptide synthesis (SPPS) has been improved since its introduction by Merrifield in the earlies 1960s.[87,88] The automation of SPPS and the use of microwave heating led to shorter reaction times and increase the purity and yield of the products.[141] The major efforts to improve the synthesis of the peptide are focused on minimizing or eliminating the formation of unwanted by-products and the occurring of common side reactions.[142] Despite the advances in the past decades, many peptides are still not accessible by the general protocols and every sequence has inherent synthetic challenges. Moreover, the problems associated with the growth of the peptide chain are still a serious obstacle in obtaining biologically active proteins. For these reasons,

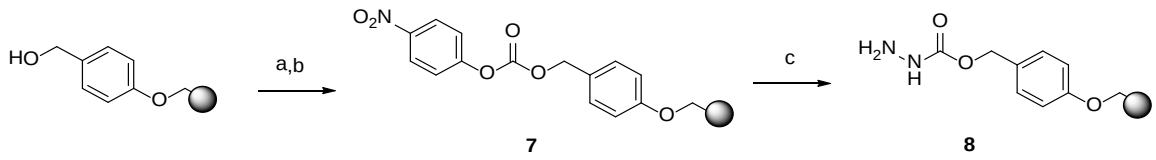
different aspects were considered to synthesize the three targeted peptide and glycopeptide fragments for downstream assembly of the Thy-1 glycoprotein.

The first parameter considered in the synthesis of the fragments was the type of linker for the functionalization and loading of the resin. This selection was directed by the amino acid at the C-*terminus* of the peptide, the length of the fragment, and the physicochemical properties of the resin. By amino acids that are prone to side reactions, the selection of the linker plays a crucial role in the synthesis of the desired peptide. For instance, a hindered linker is required to reduce aspartimide formation when aspartic acid is the C-terminal amino acid.[119] The other important aspect that was considered for the synthesis of the fragments was the length of the peptide. A resin with a high swelling is preferred for the synthesis of long sequences to ensure enough exposition of the reacting groups of the growing chain to the amino acids and reagents required in couplings and deprotections. Different linker-resins were employed for the synthesis of glycosylated or non-glycosylated fragments. The functionalization and the synthesis of different linkers that were non-commercially available or had poor stability under operational conditions were performed. All these processes are described for each peptide in the following sub-chapters.

2.3 Synthesis of the Thy-1 peptide fragment 1 (1-18)

The synthesis of the Thy-1 fragments started with the preparation of the non-glycosylated N-terminal peptide **1** corresponding to the residues 1-18 of the protein. To obtain the fragment having a hydrazide as a thioester precursor at the C-*terminus*, the initial step was the functionalization of the solid support with the required hydrazine linker.

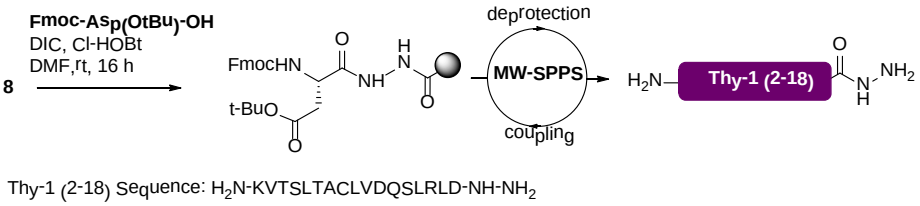
A Wang Nova PEG resin was selected as solid support for the synthesis of the fragment. The first step involved the reaction of the hydroxyl group on the resin with *p*-nitrophenyl chloroformate in the presence of N-methyl morpholine as a base. The reaction was performed twice, 1 hour at 0°C and 12 hours at r.t., using five molar excess of *p*-nitrophenyl chloroformate. The carbonate-resin was treated with hydrazine monohydrate to form the required hydrazide **8** for the peptide synthesis (scheme 13). The nucleophilic attack of the hydrazine to the carbonyl group led to the elimination of the *p*-nitrophenyl group, observed by the formation of an intense yellow colour.[143]



Scheme 13 - Functionalization of NovaPEG Wang resin **8**. Conditions: a) *p*-nitrophenyl chloroformate (5 equiv.), N-methyl morpholine (5 equiv.), DCM, 0°C to r.t., 16 h; b) hydrazine monohydrate (5 equiv.), DCM/DMF (1:1), 0°C to r.t., 16 h.

The first amino acid, Fmoc-Asp(OtBu)-OH was coupled manually to the hydrazine resin **8** using five molar excess of DIC as activator and HOBt as an additive (ratio 1:1:1) in DMF. The success of the functionalization of **7** with hydrazine and the coupling of the amino acid was determined by spectrophotometric (UV) quantification of the resin loading determined by Fmoc group removal from the aspartic acid residue. The new loading of the resin was 0.49 mmol/g and corresponded to a 98% conversion over the two steps based on the theoretical loading.

The peptide fragment was elongated in the MW-assisted peptide synthesizer (scheme 14). The coupling and deprotections were performed at 90°C for 180 and 90 sec, respectively. The arginine residue required different conditions and it was coupled twice at r.t for 25 minutes to avoid the δ-lactam formation.[124] Similarly, the cysteine residue was coupled twice at 50°C for 10 minutes to prevent the epimerization of the amino acid.[126]



Thy-1 (2-18) Sequence: H_2N-KVTSLTACLVDQSLRLD-NH-NH_2

Scheme 14- Synthesis of Thy-1 (2-18) using the functionalized Wang resin **8**

The product was released from the solid support using the cocktail TFA/H_2O/TIPS (95:2.5:2.5) for 2 hours. A LC-MS analysis of the resulting mixture showed different truncated sequences and included the presence of a peptide product having a mass corresponding to the expected mass minus 18 a.m.u (figure 8). This mass difference is characteristic of the elimination of a molecule of water, generally observed by the formation of an aspartimide in sequences having aspartic acid.

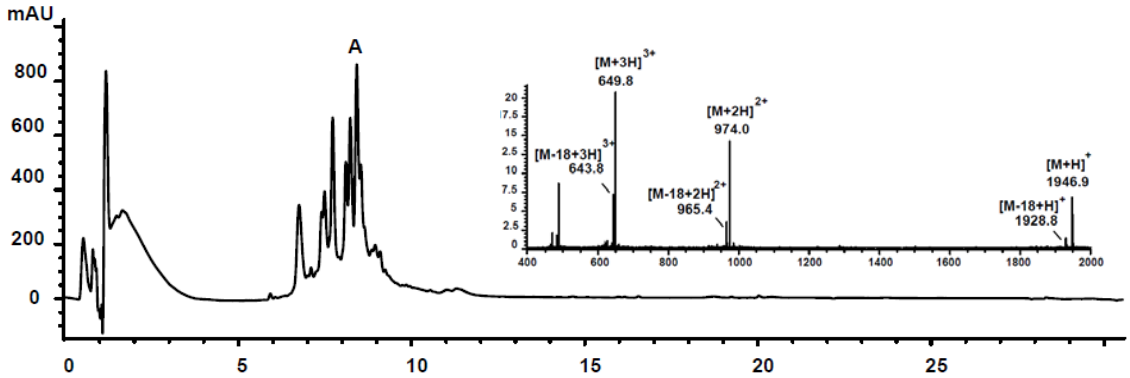
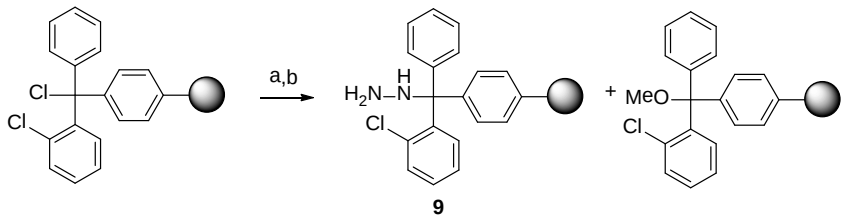


Figure 8 – LC-MS analysis of the crude fragment (2-18) synthesised with Wang resin. Chromatogram recorded at 214 nm after cleavage (RP-HPLC with C18 column, gradient: 5 to 70% of ACN in water + 0.1% of formic acid in 30 minutes). Mass spectra of peak A. Mass calculated = 1946.6 Da, Mass observed: $[M+H]^+$ = 1946.9 Da, $[M+2H]^{2+}$ = 974.0 Da, $[M+3H]^{3+}$ = 649.8 Da. $[M-18+H]^+$ = 1928.8 Da, $[M-18+2H]^{2+}$ = 965.4 Da, $[M-18+3H]^{3+}$ = 643.8 Da.

To overcome the formation of several truncated sequences and difficulties observed in the first synthesis of the peptide fragment **1**, special attention was given to the coupling of the first amino acid, the C-terminal aspartic acid, the selection of the resin and linker, and the selection of the protecting groups for specific amino acids in the sequence.

The 2-chloro-trityl (2-Cl-Trt or Cl-TCP) ProTide resin was selected as solid support for the next synthesis of this fragment. The ProTide resin contains a cover of polyethylene (PEG) chains and a polystyrene core. This resin is universally used for short peptides (< 30 amino acids) and presents a good combination of two important characteristics. The trityl linker is bulky and avoids by sterically hindrance side reactions involving intramolecular cyclization of the C-terminal amino acids, the formation of diketopiperazines, and aggregation between growing chains. In addition, the PEG coat increases the swelling properties of the resin in DCM and DMF exposing better the growing polypeptide chains to the reactants during couplings and deprotections.

The functionalization of the new resin involved the reaction of 2-Cl-Trt chloride resin with a large excess (77 equiv.) of hydrazine monohydrate in DMF (scheme 15).[144] The reaction was repeated twice to ensure a high level of functionalization and overcome the slow reaction due to the steric hindrance by the triphenyl-methyl group. There is no colorimetric test, such as the Kaiser test for amines, or other rapid methods to verify the complete functionalization of the resin.[145] Thus, possible unreacted sites were capped using a 5% solution of methanol in DMF to prevent the formation of possible side products in the next steps.

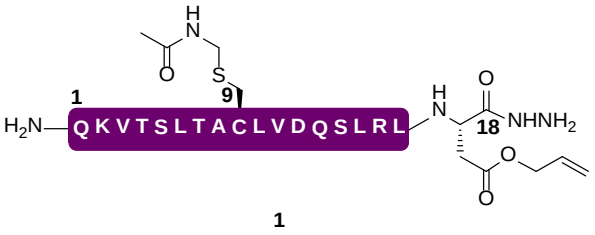


Scheme 15 - Functionalization of 2-Cl-(Cl)trityl resin with hydrazine monohydrate **9**. Conditions: a) hydrazine monohydrate (77 equiv.), DMF, r.t., 2h; b) MeOH 5% (v/v) in DMF, r.t., 10 min.

Before proceeding with the coupling of the aspartic acid, the side chain protection of the aspartate was considered. The intramolecular cyclization of C-terminal thioester of aspartic acid, glutamine, and asparagine has been documented before. Thus, aspartic acid containing an allyl protecting group at the side chain was coupled as a first amino acid.[143,146] The allyl group is stable to acids and bases and avoids the intramolecular rearrangement and formation of the succinic ring of peptides having this amino acid at the C-*terminus*.[143] In addition, this transient protecting group can be selectively removed with a palladium catalyst in the presence of appropriate scavengers.[115] Based on these considerations, the first amino acid (Fmoc-Asp(OAll)-OH) was coupled manually using HBTU as an activator and DIPEA as a base for 2 hours. The coupling was carried out using four-molar excess of the amino acid and the activator and a six-molar excess of the base in DMF as a solvent. To determine the efficiency of the modification of the resin and the coupling of the first amino acid to the hydrazide, the removal of the Fmoc group with the piperidine 20% in DMF from a small portion of the loaded resin was quantified by UV measurement. The quantification showed a new substitution of the resin of 0.26 mmol/g, corresponding to a yield of 62% conversion over two steps, based on the initial loading of the resin.

A second important consideration for the elongation of the peptide involved the selection of the protecting group of cysteine. In this peptide, there is an internal cysteine (Cys9) residue that is not required for the ligations. However, this residue is prone to give undesired side reactions and the formation of disulfide bridges. Thus, the thiol group at the side chain of the Cys9 was protected with an acetoamidomethyl group (Acm) instead of the standard trityl group.[56] The Acm group was selected as an orthogonal protecting group to avoid the intramolecular trans-thioesterification reaction between the thiol of Cys9 and the thioester of the Asp18. This reaction has been reported to occur during the conversion of the peptide hydrazide into peptide thioester and during the ligation of peptide having internal cysteines.[147] After all these considerations, the new target

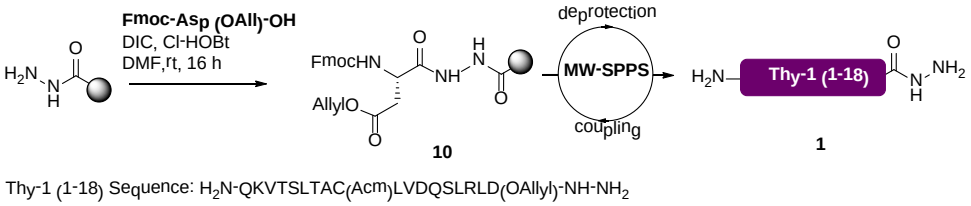
peptide fragment (1-18) was designed having an allyl protection of the β-carboxylate of Asp18 and a C-*terminus* hydrazide and Acm group protecting group on Cys9 (figure 9).



1

Figure 9 - Designed and synthesised sequence of fragment **1**.

The Fmoc group of the loaded resin **10** was removed, and the peptide was elongated in the MW-assisted peptide synthesizer using the corresponding amino acid/DIC/Oxyma (1:1:1) mixture in DMF for couplings (scheme 16). Each coupling cycle was performed at 75°C for 5 minutes and the Fmoc was removed with piperidine 20% in DMF at 75°C for 3 minutes. The analysis of peptide fragment **1** obtained under these conditions required particular attention to discard the presence of unwanted reactions on the side chain of the aspartate (aspartimide formation), cysteine (epimerization) and arginine (δ-lactam ring formation), which have been previously described for MW-SPPS and peptide synthesis at high temperatures.



Thy-1 (1-18) Sequence: H_2N-QKVTSLTAC(Acm)LVDQSLRLD(OAllyl)-NH-NH_2

Scheme 16– Elongation of Thy-1 fragment **1** (1-18). Fmoc removal was performed at 75°C for 3 min with piperidine 20% in DMF. Couplings were performed in 0.1 mmol scale using five-molar excess of Fmoc-Asp(OAllyl)-OH, DIC and Oxyma (1:1:1), at 75°C for 5 min.

After the elongation, a small-scale cleavage of the peptide from the resin was achieved using TFA/thioanisole/EDT/anisole in the ratio (90/5/3/2) as a cleavage cocktail. The cleaved peptide was analysed by analytical LC-MS using a RP-HPLC C18 column and a gradient of ACN in water in the presence of 0.1% of formic acid (figure 10).

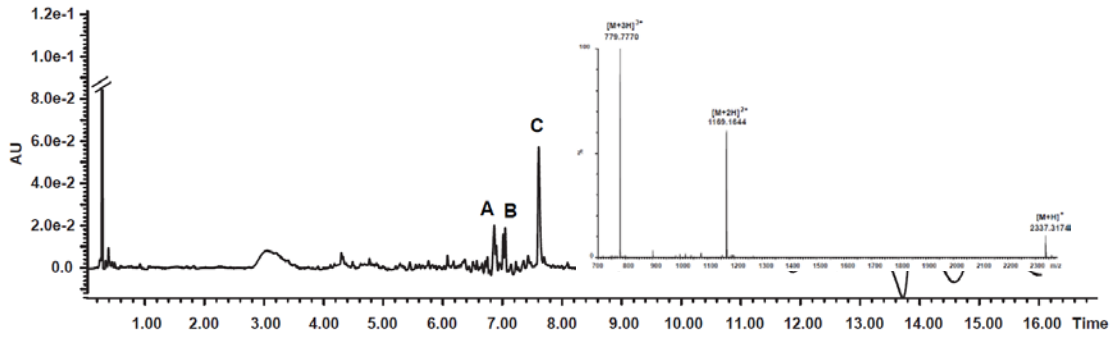
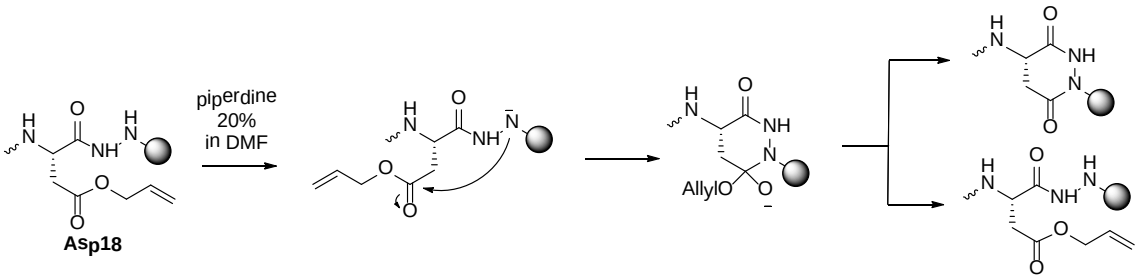


Figure 10 - LC-MS analysis of the crude fragment **1.** UV-Chromatogram recorded at 214 nm after cleavage (RP-HPLC with C18 column, gradient: 5 to 70% of ACN in water + 0.1% of formic acid in 18 minutes). A: Product -18 Da (Aspartimide formation at Asp 12); B: Product -58 Da (loss of allyl group and subsequent cyclization), C: Product **1**. High resolution mass spectra of 1 (peak C). Mass calculated = 2336.1518 Da, Mass observed: $[M+H]^+$ = 2337.3174 Da, $[M+2H]^{2+}$ = 1169.1644 Da, $[M+3H]^{3+}$ = 779.7770 Da.

The analysis by LC-MS and the profile of the chromatogram showed the presence of three major peaks (A, B, and C in figure 10). The first peak (A) corresponded to a product having a mass of the desired peptide minus 18 u.m.a. This difference corresponded to the loss of a molecule of water that was attributed to the undesired formation of an aspartimide on the second aspartic residue (Asp12). The second peak (B) showed a product having a mass corresponding to the loss of 58 a.m.u. that was attributed to the removal of the allyl group from Asp18 and the concomitant cyclization product at the C-*terminus*. The third peak (C), the major peak, corresponded to the desired peptide segment **1**. No further optimization was done to minimize the side reactions, the desired product was easily isolated from the mixture in good yields (21%).

It was no possible to determine in which the step of the process the formation of the cyclization of the C-*terminus* occurred. However, considering that the suggested product has a similar structure with an aspartimide, it is possible that one nitrogen of the hydrazide was deprotonated by piperidine during the removal of the Fmoc (scheme 17). Following, the deprotonated nitrogen attacked the carbonyl group of the β-aspartate and induced the cyclization and the elimination of the allyl ether.



Scheme 17 – Proposed mechanism for the formation of undesired cyclic by-product at the C-*terminus* under basic conditions.

Considering the obtained results with this strategy and the successful synthesis of the peptide without deletion sequences, the peptide-resin was treated with 20% piperidine in DMF to remove the N-terminal Fmoc and with a TFA/thioanisole/EDT/anisole (ratio 90/5/3/2) mixture to release the peptide from the resin under similar conditions as in the small cleavage. The peptide hydrazide product was purified by preparative RP-HPLC on a C18 hydrosphere column using the same gradient employed in the analytical conditions (5-70 % MeCN) to obtain (25 mg, 21% yield) of the purified peptide fragment **1**.

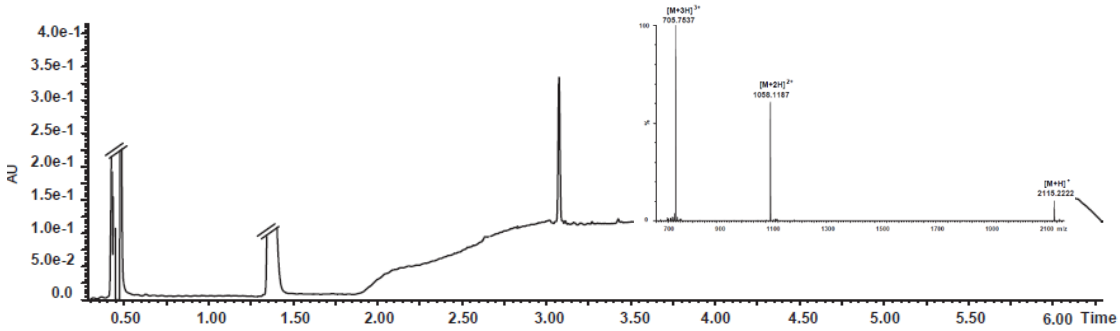


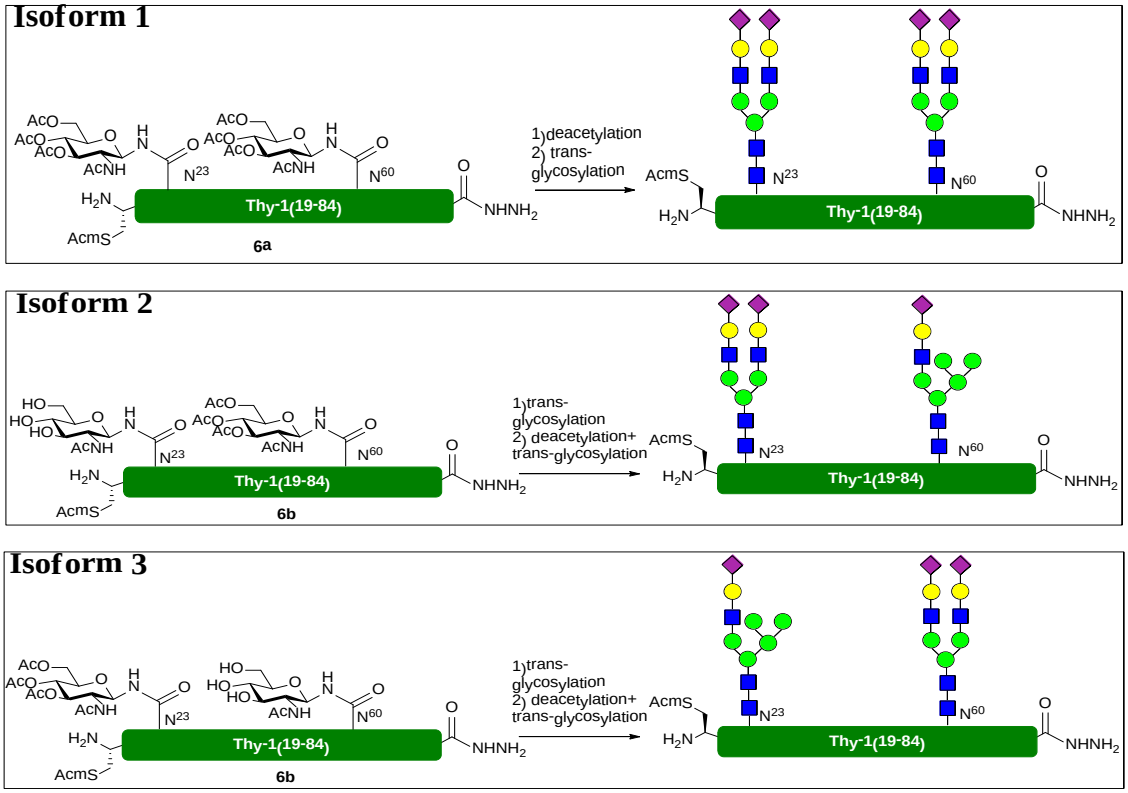
Figure 11 – LC-MS analysis of purified peptide fragment **1**. Chromatogram UV-profile was recorded at 214 nm, gradient: 5 to 70 of ACN in 6 minutes and high-resolution mass spectra of the product. $[M+H]^+_{cald}$ = 2115.1518. M_{obsd} = 2115.2222 $[M+H]^+$, 1058.1187 $[M+2H]^{2+}$, 705.7537 $[M+3H]^{3+}$.

2.4 Design and synthesis of glycoforms of fragment (19-84)

Three protocols were considered for the synthesis of isoforms of the glycopeptide fragment (19–84) and the differentiation of the glycosylation sites (scheme 18).

i. Coupling an *N*-acetyl-O-*per*-acetylated glucosamine on both Asn23 and Asn60
ii. Coupling an *N*-acetyl glucosamine on the Asn23 and an *N*-acetyl-O-*per*-acetylated glucosamine on the Asn60
iii. Coupling an *N*-acetyl-O-*per*-acetylated glucosamine on the Asn23 and an *N*-acetyl glucosamine on the Asn60.

In the first protocol, no differentiation was considered between the glycosylation sites and for the trans-glycosylation step. Thus, both glycosylated asparagine residues should contain the same monosaccharide unit. The second and third isoforms planned to distinct anchors for the transfer of oligosaccharides by the enzymes. These last two protocols required the incorporation of two and three pre-synthesised building blocks during the elongation of the peptide fragments. The procedure showing the synthesis of the designed isoforms and the required building blocks are explained in detail in the following sections.



Scheme 18 – Designed isoforms of the Thy-1 glycopeptide fragment (19-84).

The synthesis of the glycopeptide (19-84) required two important considerations. The first concerns the most convenient synthetic strategies to obtain this long peptide in good yield. The second aspect was the selection of the monosaccharides and their application for the differentiation for each glycosylation site.

The first strategy for the synthesis of the fragment (19-84) involved the stepwise assembly of the whole sequence using a combination of manual couplings and elongation in the MW-assisted synthesizer. Based on the observed side reactions during the SPPS and the conditions implemented for fragment **1**, the synthesis of the glycopeptide **6** (Thy-1 19-84) required similar considerations to avoid the side reactions during the elongation and the incorporation of pre-synthesised building

blocks on specific sites of the sequence. To obtain the three isoforms of glycopeptide **6**, three molecules were required and synthesised (table 6). Fmoc-asparagine glycosylated with O-*per*-acetylated *N*-acetyl- glucosamine (**2)**; the Fmoc-Thr-(OtBu)-Ser($\Psi^{Me,Me}$Pro)-OH (**3)**, a dipeptide having a pseudoproline, an aspartic acid protected by a photolabile phenacyl group on the β-carboxylic acid (**4)**, and the *N*-acetyl glucosamine (**5)**.

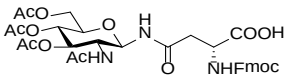
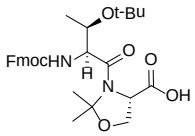
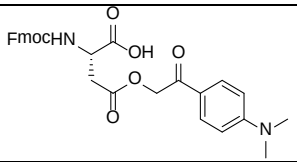
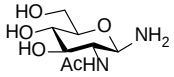
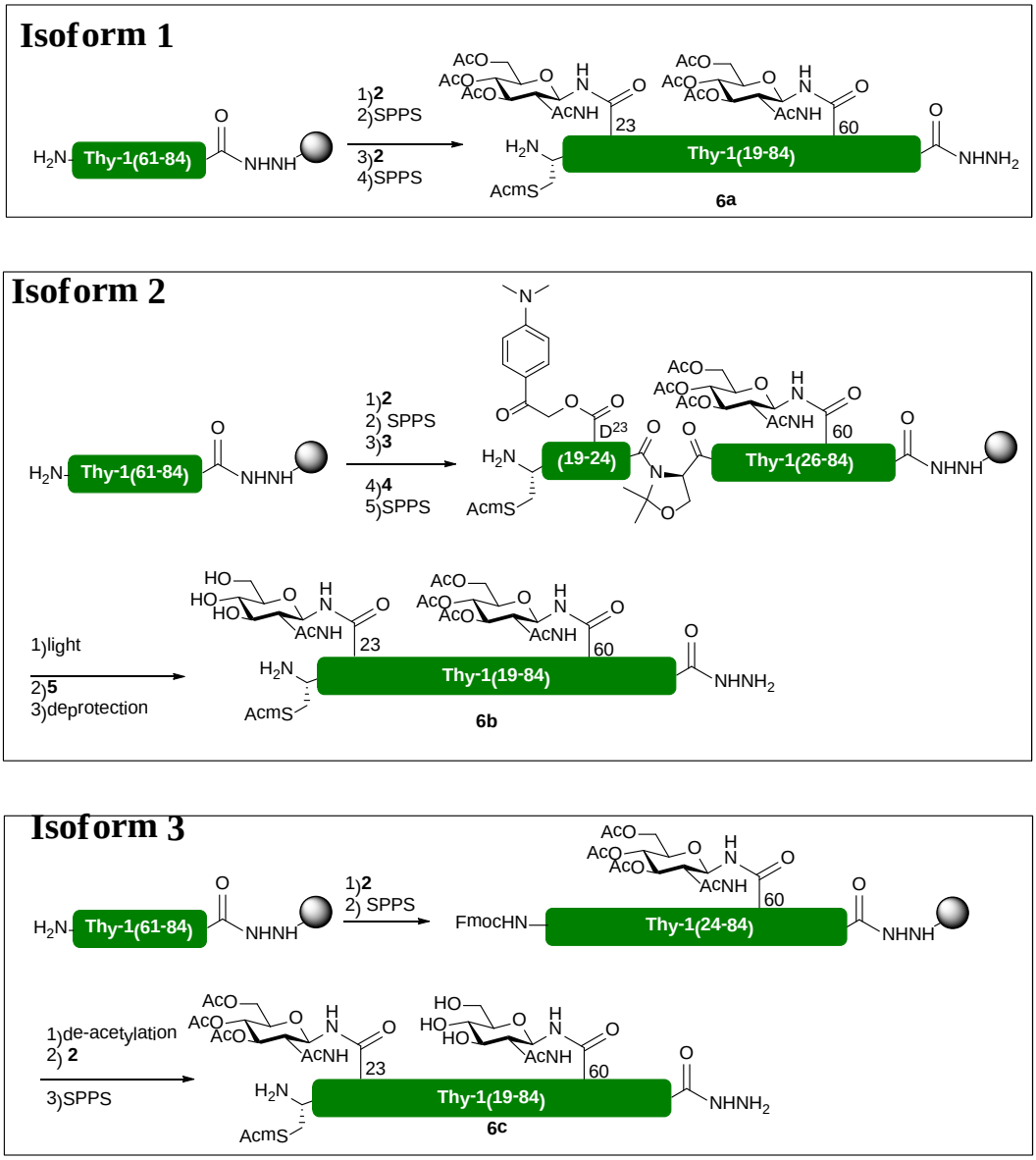
Compound	Structure	Name
2		N-acetyl-O-per-acetylated glucosyl N-Fmoc-asparagine
3		N-Fmoc-Thr(OtBu)-Ser($\Psi^{Me,Me}$ Pro)-OH
4		N-Fmoc-Asp(OMap)-OH
5		N-Acetylglucosamine amine

Table 6 - Building blocks for the synthesis of glycoforms **6** of Thy-1(19-84).

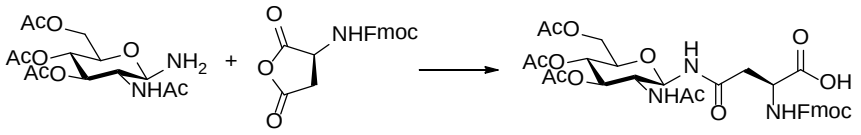
The glycosylated asparagine **2** and the C-terminal sequence of Thy-1 (61-84) were common building blocks required for synthesizing the three isoforms of the fragment **6** (scheme 19). The synthesis of isoform **6b** required the coupling of the pseudoproline dipeptide **3,** the protected aspartic acid **4**, and on-resin coupling of glucosamine **5** to a peptide having the β-carboxylate of aspartic acid unprotected. To synthesize the isoform **6c**, having a protected and a non-protected glycosylated asparagine, the glycosyl amino acid was **2** first incorporated and then deacetylated before the coupling of the next glycosylated residue during the elongation of the glycopeptide.



Scheme 19 –Synthetic routes for the synthesis of the isoforms **6a-c** of the Thy-1 fragment (19-84).

2.4.1 Synthesis of the glycosylated building blocks 2 and 5

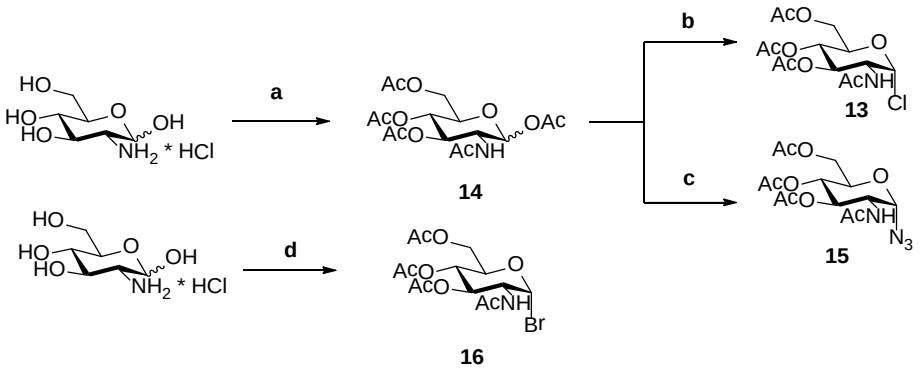
The required glycosylated asparagine **2** was synthesised and incorporated manually during the elongation of the peptide sequence at the corresponding glycosylation site. The synthesis was envisioned following the method reported by Ibatullin and Selivanov using a reaction between an N-acetylglucosamine amine and aspartic acid anhydride in equimolar concentration (scheme 20).[148]



11 **12** **2**

Scheme 20 - Reaction scheme for the synthesis of the glycosylated asparagine **2**.

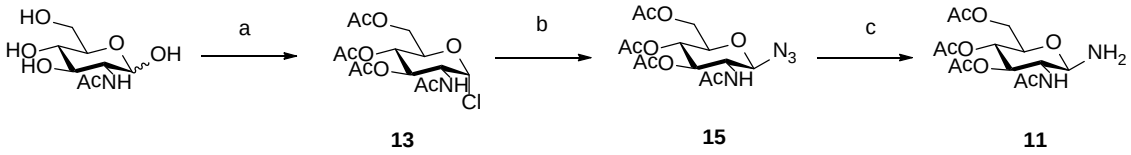
Different methods were used to synthesise the glycosylamine **11** starting from D-glucosamine. The first step involved the acetylation of glucosamine hydrochloride using acetic anhydride and pyridine (scheme 21). Following the method reported by Judeh *et al.*, the acetylated sugar **14** was reacted with 2-chloro-1,3-dimethylimidazolinium chloride (DMC) to generate the corresponding chloride **13**.[149] MS monitoring of the reaction showed the formation of several products that did not correspond to the desired halide. A second method involving tin tetrachloride as a Lewis acid activator and $TMSN_3$ was evaluated for the generation of the α-glycosyl azide **15**.[150] However, contrary to the literature reported and similar to the procedure with DMC, the formation of the product could not be detected by MS.



Scheme 21 - Synthesis of N-acetyl-O-per-acetylated glycosyl halide: Conditions: a) Ac_2O (6 equiv.), pyridine (0.3 equiv.), 16 h, r.t., quantitative; b) DMC (1equiv.), Et_3N (2 equiv.), 24 h, r.t.; c) $TMSN_3$ (2.5 equiv.), $SnCl_4$ (0.5 equiv.), DCM, r.t., 15 h; d) AcBr (4 equiv.), MeOH (4 equiv.), AcOH (1.5 equiv.), Ac_2O (1.6 equiv.) 15 h, r.t.

A third route to obtain the desired glycosyl amine involved the preparation corresponding to the O-per-acetylated glycosyl bromide **16** from glucosamine. Glucosamine hydrochloride was treated with acetyl bromide in acidic conditions. The formation of the bromide **16** product was observed by MS, but in low yield (9%) and with the formation of several by-products that hinder the purification of the product.[151]

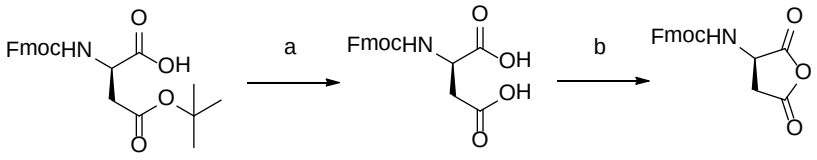
A different protocol was considered using commercially available *N*-acetyl-glucosamine, which was treated with an excess of acetyl chloride and hydrochloric acid for 16 hours.[152] The *N*-acetyl-O-*per*-acetylated-glycosyl chloride **13** was obtained in 55% yield. The purified product was converted into the correspondent compound **15** by reaction of the chloride at the anomeric position with an excess of sodium azide in DMF. A final reduction of the anomeric azide under hydrogen atmosphere in the presence of a palladium catalyst gave the desired 1-β-amino-O-*per*-acetylated glucosamine **11** (scheme 22).



13 **15** **11**

Scheme 22 - Synthesis of N-acetyl-O-per-acetylated glycosyl amine **11**: Conditions: a) AcCl (1 equiv.), HCl (0.1 equiv.), 16 h, r.t., 55%; b) NaN_3 (3.5 equiv.) in DMF, 3h, r.t., quantitative; c**)** H_2, Pd/C (0.1 equiv.), Et_3N (1 equiv.) in THF, 1h, r.t., quantitative.

For the synthesis of the next building block, the aspartic anhydride **12,** two steps were required starting from commercially Fmoc-protected amino acid (scheme 23).[148] First, Fmoc-Asp(OtBu)-OH was treated with an excess of trifluoroacetic acid (TFA) in dichloromethane to remove the *tert*-butyl protecting group from the carboxylate on the side chain of the amino acid. The deprotected dicarboxylic acid **17** was dissolved in acetic anhydride and heated to 80°C to induce the formation of the corresponding anhydride **12**. The formation of the anhydride was analysed by MS and NMR analysis.

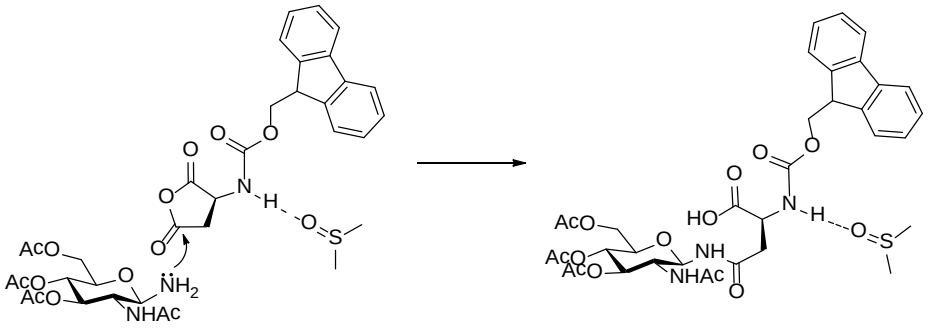


17 **12**

Scheme 23 - Synthesis of 2-N-Fmoc aspartic anhydride **12**: Conditions: a) TFA (20 equiv.) in DCM, 3 h, 0 °C to r.t., quantitative; b) Ac_2O (30 equiv.), 2 h, 80°C., 66%.

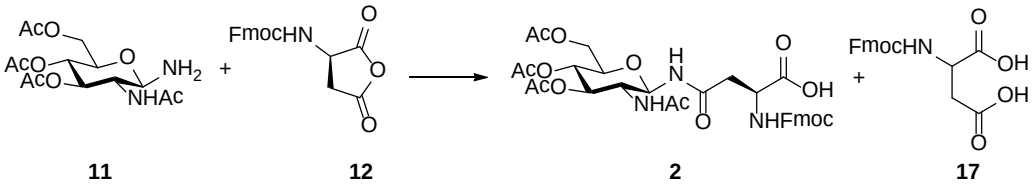
The two building blocks **11** and **12** were reacted in equimolar concentration in DMSO. The selected polar aprotic solvent was crucial to obtain the desired regioselectivity of the opening and the glycosylation at the β-carbonyl group, as indicated in previous studies.[153] The proposed mechanism for the selectivity is shown in scheme 24. The regioselectivity is attributed to the

hydrogen bond between the proton of the Fmoc-amine and the oxygen of the α-carbonyl group of the DMSO. This leads the nucleophile, the anomeric amino group on the N-acetylglucosamine amine, to attack preferentially at the less sterically hindered carbonyl group of the aspartic anhydride. It is important to highlight that the observed regioselectivity in the reaction with building block **11** required a large N-α-protecting group such as Fmoc.



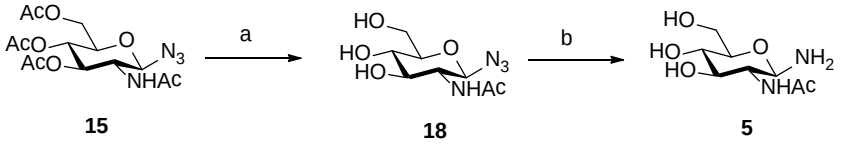
Scheme 24 - Proposed mechanism of the N-Fmoc aspartic anhydride opening in DMSO as solvent.

The reaction proceeded with the opening of the *N*-Fmoc aspartic anhydride at the β-carbonyl group of the ring. After 5 hours, an excess of the hydrolysed **17** was observed by MS analysis; the reaction was then stopped and the workup followed. The formed products were separated using flash chromatography to obtain the desired glycosylated amino acid **2** in 43% yield (scheme 25).



Scheme 25 - Synthesis of glycosylated asparagine **2**. Conditions: **11** (1 equiv.), **12** (1 equiv.), DMSO, 5 h, r.t., 43%.

The N-acetyl glucosamine **5** was obtained from the previously synthesised compound **15**. First, the glucosamine **15** was deacetylated using Zemplen conditions to obtain the triol **18**. Then, the azide group was reduced with hydrogen in the presence of palladium as a catalyst (scheme 26).

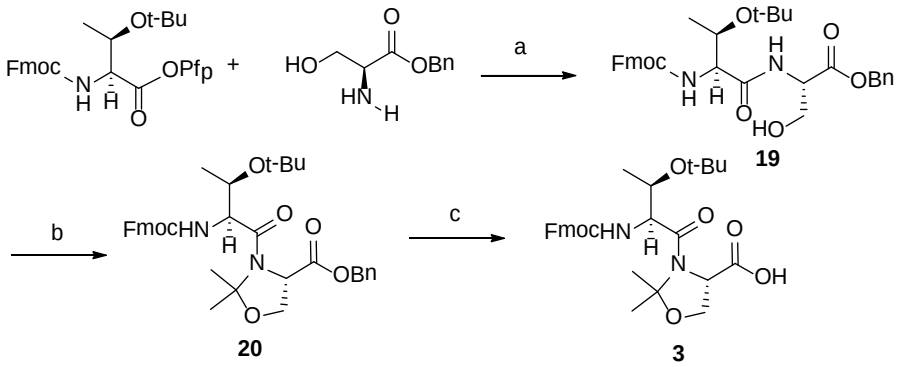


Scheme 26 - Synthesis of building block **5**: Conditions: a) NaOMe (freshly prepared from NaH in methanol), 1 h, r.t., quantitative; b) H_2, Pd/C (0.1 equiv) in methanol, 1 h, r.t, 98%.

2.4.2 Synthesis of the pseudoproline dipeptide 3

Pseudoprolines dipeptides were introduced in peptide synthesis to disrupt the formation of β-sheets and aggregation of the growing peptide chains during the SPPS. Similar to the natural function of proline, pseudoprolines disrupt secondary structures by inducing a "knick point" on the peptide chain, breaking the interactions within and between the chains inducing the aggregation and β-sheet formation.[154] Moreover, the presence of pseudoproline in the proximity of the aspartic acid can prevent the undesired intramolecular cyclization affecting this amino acid and the concomitant aspartimide formation.[155]

The dipeptide Thr-Ser present in the sequence of the fragment (19-84) was introduced as Fmoc-Thr(tBu)-Ser($\Psi^{Me,Me}$ Pro)-OH for the synthesis of the isoform glycopeptide **6b**. The synthesis of the dipeptide (Fmoc-Thr(tBu)-Ser($\Psi^{Me,Me}$ Pro)-OH was completed by adapting the protocol reported by Wöhr *et al.* (scheme 27).[154]



Scheme 27 Synthesis of dipeptide **3**. Conditions: a) Aq. sodium carbonate (10% w/v), H_2O, 2 h, r.t, 98%; b) 2,2-dimethoxypropane, pyridyl toluene-4-sulfonate, toluene, 110 °C, 5 h, 61%; c) H_2, Pd/C, methanol, 2 h, r.t, quantitative.

The dipeptide Fmoc-Thr(OtBu)-Ser(OH)-OBn **19** was formed in excellent yield by coupling the commercially available pentafluorophenyl ester of the protected threonine to benzyl ester of serine under basic conditions. Following, the free hydroxyl group of serine was reacted with 2,2-dimethoxypropane in the presence of pyridyl toluene-4 sulfonate in catalytic amounts to induce the formation of N,O-cyclic acetal with the nitrogen of the amide bond giving the fully protected pseudoproline dipeptide **20** in 61% yield. Finally, the benzyl group of **20** was removed by palladium catalysed hydrogenolysis giving the desired building block **3**.

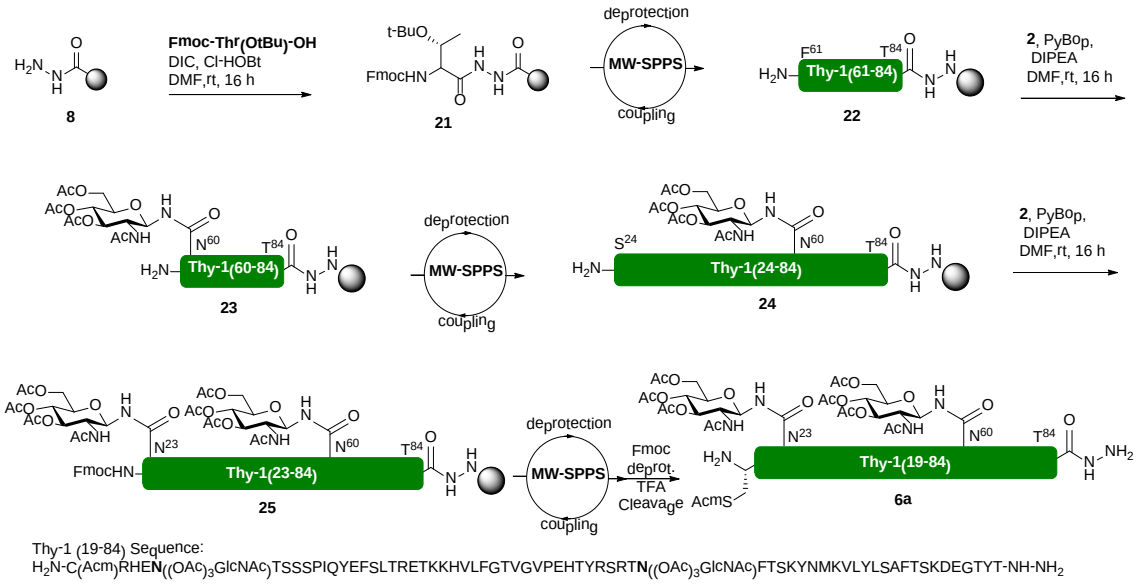
2.4.3 Fmoc-Asp(OMap)-OH 4

Protecting groups that are removable by light (photo-cleavable groups, PPGs) offer good alternatives to conventional temporary groups. These molecules are convenient for masking functional groups requiring stability in acidic and basic conditions and are removed orthogonally using an appropriate light source.[156] This orthogonality was studied to synthesise the glycopeptide **6b** that required differentiation of a glycosylation site from the other within the same fragment. For this purpose, aspartic acid having a photo-cleavable group at the β-carboxylate was considered for the glycosylation position (Asn23). Irradiation with light may allow the removal of the photolabile orthogonal protecting group and the generation of the corresponding free carboxylic acid. A Lansbury aspartylation, coupling the unmasked carboxylic acid with the glycosyl amine **5** on resin, will introduce the glycosylation and regenerate the asparagine residue.[157]

The 4-(dimethyl) amino-phenacyl group (Map) was selected as the most convenient group for this protection. This photo-removable protecting group is stable to TFA and piperidine and is removed by irradiation with light over 300 nm. The Map group has been previously synthesised and employed to protect the side chain of cysteine.[68] The synthesis of Fmoc-Asp(OMap)-OH building block was completed previously in our group following the report by Otaka *et al*., and it was used in the synthesis of the glycopeptide **6b** as described below.[158]

2.4.4 Synthesis of the glycopeptide fragment 6a (isoform 1)

The glycopeptide **6a**, having two *N*-acetyl-O-*per*-acetylated glucosamines on Asn23 and Asn60, was the first glycopeptide fragment to be synthesized. The incorporation of the same glycosylated building block will allow the attaching of the same oligosaccharide during the enzymatic trans-glycosylation. The synthesis of the isoform **6a** having an $(AcO)_3GlcNAc$-) unit in both glycosylation sites was designed involving three stage: manual modification of the resin, coupling of the non-glycosylated amino acids by MW-assisted SPPS, and a manual coupling of the pre-synthesized glycosylated asparagine **2** at the specific positions of the peptide (scheme 28).



Scheme 88 - Synthetic strategy to obtain the glycopeptide fragment **6a**.

The hydrazide-functionalized Wang hydroxymethylated resin **8** was prepared as described before and used to synthesize the glycopeptide **6a**. The Fmoc-Thr(OtBu)-OH was activated using DIC and Cl-HOBt and coupled manually to the resin. To ensure binding to all the reactive sites on the solid support, four molar excess of all reagents were used in the ratio 1:1:1. The efficiency of the coupling was confirmed by the determination of the released Fmoc protecting groups after a small-scale deprotection on few beads of **21** in a solution of 20% piperidine in DMF. A comparison of the theoretical loading of threonine as bound amino acid on the solid support with the observed loading (0.44 mmol/g) revealed that 84% of the reactive sites were successfully coupled. The remaining unreactive sites were capped using a solution of Ac_2O/DIPEA/DMF in ratio 1:1:8.

Due to the length and the complexity of the peptide, the deprotection and coupling cycles were carefully analysed and optimized for each amino acid. The pre-synthesised glycosylated asparagine was incorporated manually into the sequence to reduce the excess of required material and have a better control of the reaction, due to limited amount of building block **2**. Several attempts were necessary to optimize and get close to full conversion in deprotections and couplings for each amino acid and to minimize the number and amount of side reactions, such as undesired cyclization reactions.

The optimization started with the elongation of **21** to give the C-terminal peptide segment **22**. A first synthesis of this part resulted in several truncated sequences and undesired products (figure 12). Identification of the obtained peptides was performed using LC-MS analysis and it was used

to determine the critical amino acids in the sequence in pursuit of the optimal conditions for the synthesis of the fragment.

Seven major products (peaks A-G) were detected by LC-MS. The main peak C corresponded to the desired peptide; peaks A, B, and G corresponded to deletion sequences lacking phenylalanine, tyrosine, and glutamic acid. Peak D and F showed an incomplete removal of *tert*-butyl and Fmoc groups. Special attention was given to peak E corresponding to a base-catalysed aspartimide formation on the Asp79, a side reaction that occurs during the treatment with piperidine during the elongation of the peptide. Thus, the first parameter to be evaluated was the base for the Fmoc removal during the elongation of the peptide chain.

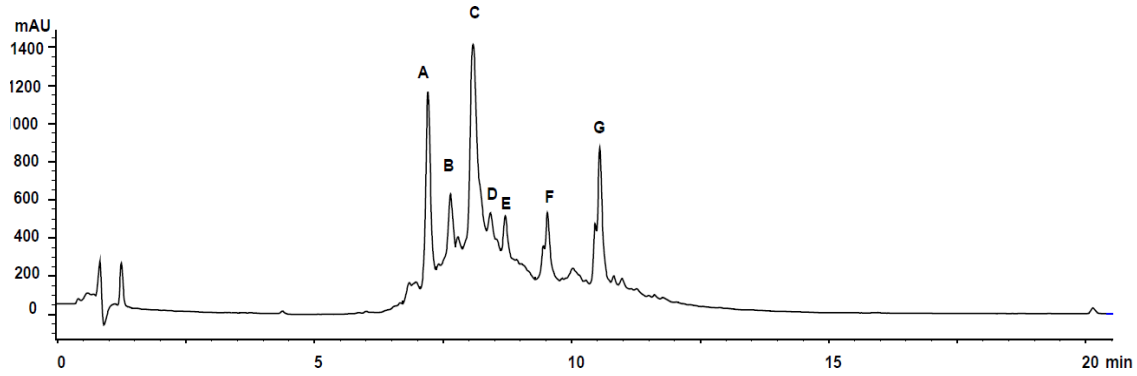


Figure 12 – LC analysis of the Thy-1 fragment 61-84 synthesised using MW-SPPS standard conditions. Chromatogram recorded at 214 nm of the crude fragment **22** after cleavage (RP-HPLC with C18 column, gradient: 5 to 70% of ACN in water + 0.1% of formic acid in 30 minutes). A: Product -147 Da (- phenylalanine); B: Product -163 Da (-tyrosine); C: Product **22**; D: Product + 57 Da (+ tert-butyl); E: Product – 18 Da (aspartimide formation); F: Product + 223 Da (+ Fmoc); G: Product – 129 Da (- glutamic acid).

Alternative reagents were considered for the Fmoc-removal and a new synthesis of the peptide was completed using a freshly prepared 10% (w/v) solution of piperazine in NMP/EtOH (90:10) for the deprotection in the MW-synthesizer.[159] The LC-MS analysis of the new synthesis after the test cleavage with TFA solution showed the formation of aspartimide (figure 13). The change of the base from piperidine to piperazine did not lead to a significant improvement in the synthesis. Considering that no progresses were observed for the synthesis of the peptide and the time-consuming dissolution of the piperazine in NMP/EtOH mixture, this deprotection cocktail was discarded for the following optimizations of the synthesis.

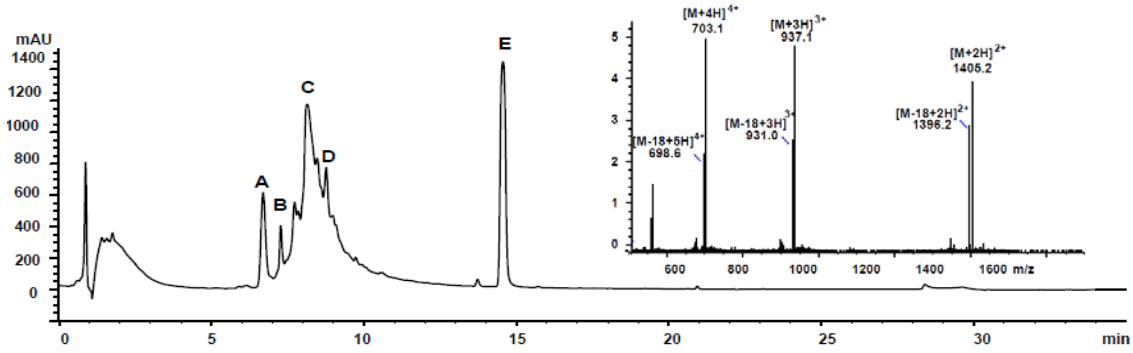
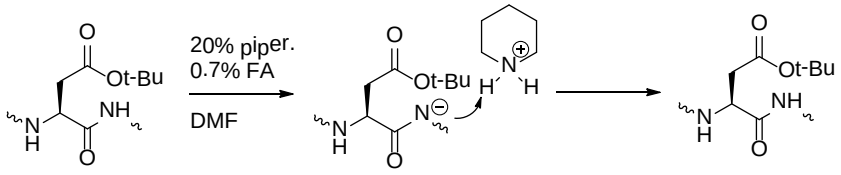


Figure 13 – LC-MS analysis the Thy-1 fragment 61-84 synthesised using 10% piperazine in NMP/EtOH as deprotecion solution. Chromatogram recorded at 214 nm of the crude fragment **22** after cleavage (RP-HPLC with C18 column, gradient: 5 to 70% of ACN in water + 0.1% of formic acid in 30 minutes). A: Undetected mass; B: Product – 18 Da (aspartimide formation); C: Product and Product -18 Da (aspartimide formation); D: Product + 57 Da (+ tert-butyl); E: Collection of the released protecting groups of side chains. Mass spectra of product (peak C). Mass cldc = 2808.6 Da. Mass obs, $[M+2H]^{2+}$ = 1405.2 Da, $[M+3H]^{3+}$ = 937.1 Da, $[M+4H]^{4+}$ = 703.1 Da, $[M+5H]^{5+}$ = 562.8 Da. $[M-18+2H]^{2+}$ = 1396.2 Da, $[M-18+3H]^{3+}$ = 931.0 Da, $[M-18+4H]^{4+}$ = 698.6 Da, $[M-18+5H]^{5+}$ = 559.1 Da

A second optimization synthesis was completed using the addition of 0.7% of formic acid in the solution of 20% piperidine in DMF (v/v). The base-promoted formation of the cyclic imide on the aspartic acid is a sequence-dependent reaction and could be reduced with the use of different organic and inorganic acids. Mier *et al.* proposed the use of formic acid as an effective suppressor of the formation of aspartimide. The formic acid works as an effective suppressor of the side reaction due to its ability to protonate the nitrogen of the piperidine and generate a less basic piperidinium ion as shown in the proposed mechanism (scheme 29).[160] Unfortunately, the addition of this inorganic acid did not eliminate completely the formation of the aspartimide in the peptide but minimized it remarkably (figure 14).

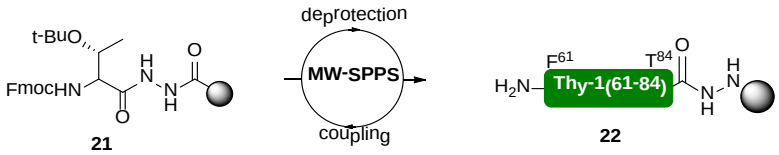


Scheme 29 - Proposed mechanism of the action of the formic acid in the deprotection solution.

Due to the high microwave absorption of the piperidine-formic acid mixture, the removal of the Fmoc group performed with the employ of the new deprotection solution in the synthesizer required lower temperatures (change from 90°C of the standard protocols to 75°C), longer reaction time (from 2 to 3 minutes), and lower MW energy (from 155 W to 50 W). Similarly, the synthesis of fragment **22** was improved by modifying the coupling conditions and using 75°C for 5 minutes.

Some amino acids required different cycles as observed from the previous synthesis. These modifications include the following changes:

- A preconditioning of the reaction vessel, which was cold at the beginning of the synthesis and warmed up as the process proceeds.
- The Fmoc removal of the Thr84 at the C-*terminus* was executed twice to assure fully amine deprotection.
- Double Fmoc removal also for Gly79 to facilitate the efficient coupling of the following amino acid, Glu80. The incorporation of the glutamic acid was performed twice to avoid deletion of the amino acid in the peptide as observed before (figure 12, peak **G)**.
- Another important modification was the incorporation of Phe75 and Phe61 using double coupling to ensure complete insertion of this sterically hindered amino acid. The same protocols were employed also for the coupling of Val69.



Scheme 30 – Elongation in the MW-assisted synthesizer of fragment **22** using optimized conditions for coupling and deprotection steps.

After deprotection of Phe61, the purity of the synthesised peptide fragment **22** was verified by small-scale cleavage. The LC-MS analysis showed the persistence of protecting groups on the side chains of amino acids. Thus, the cleavage cocktail was changed, and the TFA/phenol/water/thioanisole/EDT (82.5/5/5/5/2.5) mixture containing more scavengers was employed. LC-MS analysis of the products obtained by cleavage using this new solution showed that 75% of the crude material corresponded to the desired peptide and only a small amount portion corresponded to truncated sequences (figure 14).

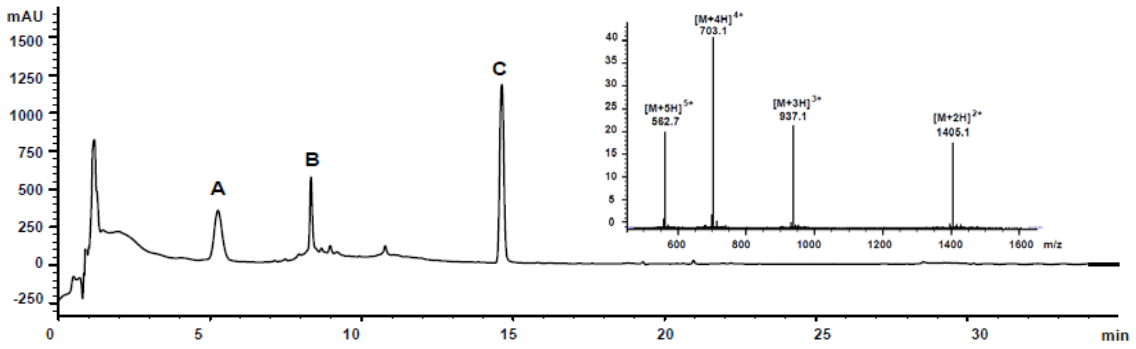
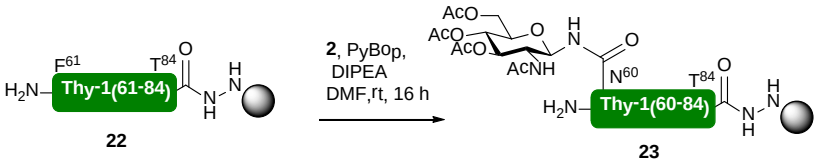


Figure 14 - LC-MS analysis the Thy-1 fragment synthesised 61-84 using piperidine 20% in DMF + 0.7 % FA. Chromatogram recorded at 214 nm after cleavage (RP-HPLC with C18 column, gradient: 5 to 70% of ACN in water + 0.1% of formic acid in 30 minutes). A: Undetected mass; B: Product **22**; C: Collection of the released protecting groups of side chains. Mass spectra of the product **22** (peak B). Mass cldc = 2808.6 Da, Mass obs $[M+2H]^{2+}$ = 1405.2 Da, $[M+3H]^{3+}$ = 937.1 Da, $[M+4H]^{4+}$ = 703.1 Da, $[M+5H]^{5+}$ = 562.7 Da

Having the C-terminal part of the sequence completed and in good quality, the peptide-resin was removed from the synthesizer for the manual incorporation of the pre-synthesised glycosylated asparagine **2**. The glycosyl amino acid was coupled manually to the free amino group of terminal Phe61 of the fragment **22** using PyBOP and DIPEA as a base for activation in DMF.



Scheme 31 – Incorporation of the glycosylated asparagine in DMF.

The efficiency of the reaction was based on careful solvent selection to dissolve building block **2**. The use of these conditions and activation of the glycosylated amino acid did not deliver a good coupling of **2** to obtain **23**, as determined from LC-MS analysis (figure 15). Therefore, a second coupling of the glycosylated asparagine was performed using PyBOP for the activation in the presence of Triton X100 in the activation mixture to increase the solubility of building block **2**. However, the addition of the detergent did not improve the outcome of the coupling compared to the use of DMF alone.

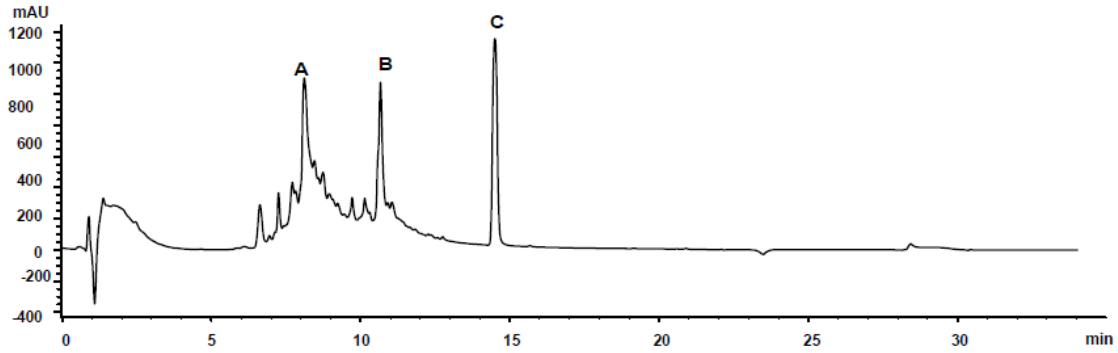
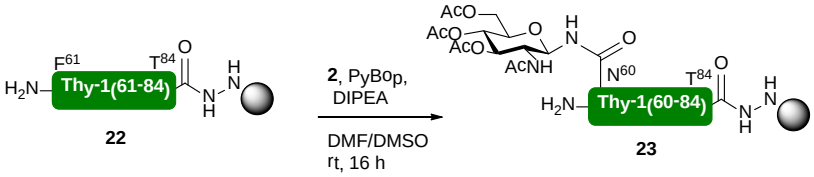


Figure 15– LC analysis the Thy-1 fragment (60-84) **23** after cleavage recorded at 214 nm (RP-HPLC with C18 column, gradient: 5 to 70% of ACN in water + 0.1% of formic acid in 30 minutes). A: Unreacted fragment **22**; B: Product **23**; C: Collection of the released protecting groups of side chains. The yield was calculated as 71:39 of **22** : **23**.

The solubility of the glycosylated asparagine was increased by the addition of DMSO as a co-solvent in the activation mixture. Building block **2** was activated with PyBOP and DIPEA using a mixture of DMF /DMSO in the ratio (200:1) as the solvent. The resin/amino acid slurry was mixed for 16 hours.



Scheme 32 – Incorporation of the glycosylated asparagine in DMF/DMSO (200:1).

To verify the efficiency of the coupling, a small portion of the peptide/resin having **23** was treated with 20% piperidine and 0.7 % of formic acid in DMF to remove the Fmoc-group and the peptides were cleaved from the resin with a TFA mixture. LC-MS analysis of the cleaved peptides revealed a significant improvement of the coupling and a positive effect resulting from the better solubilisation of building block **2** in the DMF/DMSO mixture (figure 16).

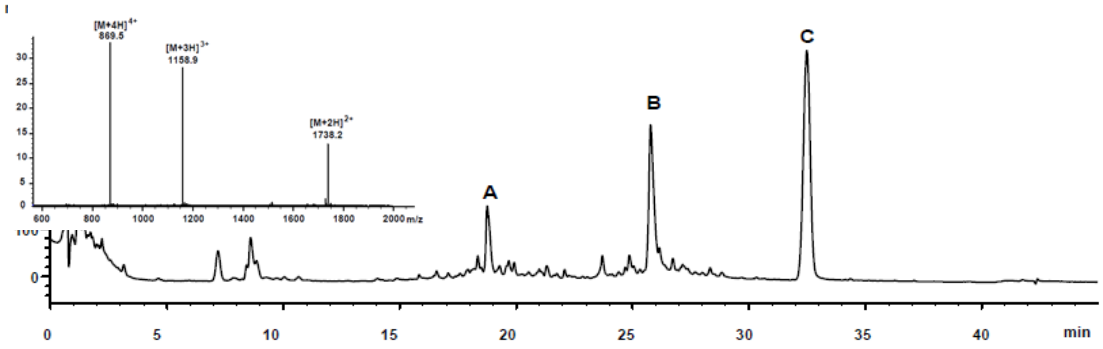
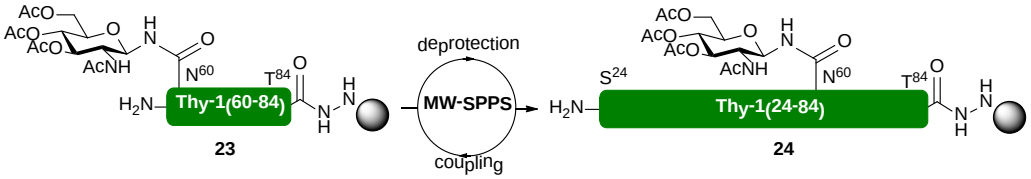


Figure 16 – LC-MS analysis the Thy-1 fragment **23** obtained using DMF/DMSO in the ratio 200:1 for the solubilisation of **2**. Chromatogram recorded at 214 nm after cleavage (RP-HPLC with C18 column, gradient: 10 to 40% of ACN in water + 0.1% of formic acid in 30 minutes). A: Unreacted fragment **22**; B: Product **23**; C: Collection of the released protecting groups of side chains. Mass spectra of the product **23** (peak B). Mass cldc = 3472.6 Da, Mass obs $[M+2H]^{2+}$ = 1738.2 Da, $[M+3H]^{3+}$ = 1158.9 Da, $[M+4H]^{4+}$ = 869.5 Da. The ratio of the products was calculated as 30:70 for **22**:**23** respectively.

The remaining amino groups of peptide resin **22** were capped with a solution of Ac_2O/DIPEA/DMF in a 1:1:8 ratio. The glycopeptide **23** was elongated in the MW-assisted synthesizer and the deprotection and coupling reactions were optimized. Based on the observed increase in the resin swelling with the increase of the peptide chain length, the volume of solvent was corrected to guarantee the interaction between the active groups on the resin and the reagents and to avoid a burning of the solid support under microwave irradiation. To keep the concentration of the coupling reagents constant, it was decided to split the resin-peptide **23** into two parts and elongate only one portion of the peptide-resin using the same volumes as in the first part.

The N-*terminus* (glycosylated Asn60) was deprotected twice at 75°C for 3 minutes and the following amino acid (Thr59) was coupled twice at 75°C for 5 minutes. The double coupling with threonine was required to obtain complete conversion and overcome the steric hindrance giving by the glycosylation on the side chain (the *N*-acetyl-O-*per*-acetylated glucosamine). To avoid the formation of a δ-lactam ring, a double coupling using a special cycle, first coupling at r.t. for 25 minutes, and the second coupling at 75°C for 2 minutes was used to introduce Arg58, Arg56, and Arg27 into the sequence.[124] Histidine and cysteine can racemize at the high temperatures.[125] Consequently, His53 and His42 were incorporated using double coupling at 50°C for 10 minutes each, after double amine deprotection of the previous amino acid (Thr54 and Val43 respectively). The Glu52, Glu38, Glu32, and Ile29 were also coupled twice at 75°C for 5 minutes to provide their binding efficiently and the amino acids that preceded them (His53, Thr39, Phe33 and Gln30, respectively from C-to-N *terminus*) were deprotected twice to guarantee liberation of the amino groups from the Fmoc protection.



Scheme 33– Elongation in the MW-assisted synthesised of glycopeptide **24**.

After the final deprotection of Ser26, a portion of the resulting glycopeptide **24** was cleaved and analysed by LC-MS (figure 17). Diverse truncated sequences, the capped fragment and product were observed, no side reactions were detected for the critical amino acids (Arg, Cys, His and Asp) but Ser and Lys were observed missing in different peaks. This material was used to explore the synthesis of the fully glycopeptide fragment of Thy1 with two glycosylations.

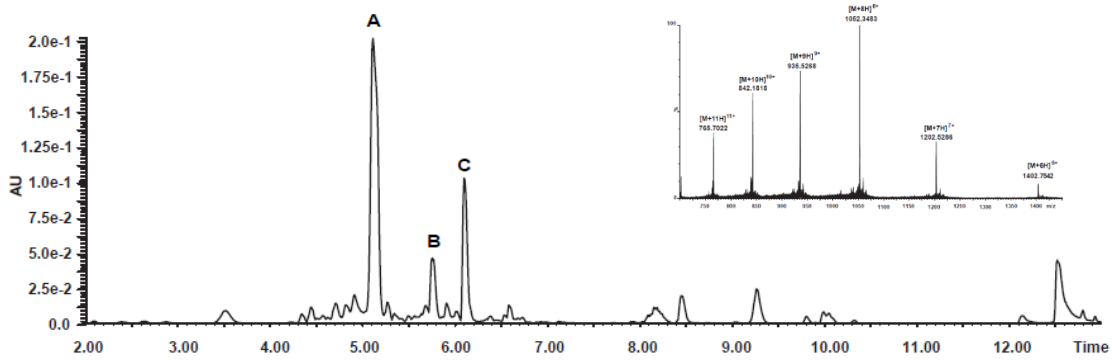
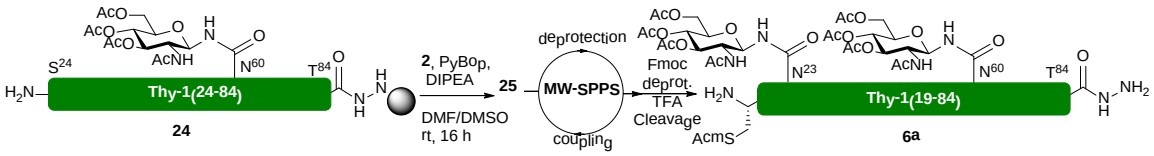


Figure 17 – LC-MS of Thy 1(24-84) **24**. Chromatogram recorded at 214 nm after cleavage (RP-UPLC with C18 column, gradient: 10 to 50% of ACN in water + 0.1% of formic acid in 14 minutes). A: Undetected mass (presumably, collection of the released protecting groups of side chains); B: Product **24**; C: Capped peptide fragment **22**. Mass spectra of the product **24** (peak B). Mass cldc = 7370.6242 Da, Mass obs: $[M+4H]^{4+}$ = 1843.7178 Da, $[M+5H]^{5+}$ = 1475.0087 Da, $[M+6H]^{6+}$ = 1229.5326 Da, $[M+7H]^{7+}$ = 1053.9023 Da, $[M+8H]^{8+}$ = 922.3040 Da, $[M+9H]^{9+}$ = 819.9453 Da.

The resulting peptide-resin **24** was elongated. First the glycosylated asparagine **2** was coupled manually to the fragment **24** using the conditions described for the coupling of the first glycosylated asparagine. (Fmoc-Asn((OAc)$_3$GlcNAc)-OH was activated with PyBOP and DIPEA and coupled for 12 hours at r.t.. After the coupling, the resin was capped with a solution of Ac_2O/DIPEA/DMF. To complete the peptide, the obtained glycopeptide **25** was elongated in the synthesiser using double deprotection and coupling cycles for the last four amino acids. His21 and (Acm)-protected Cys19 were coupled at 50°C; Arg20 was coupled at r.t. and 75 °C, and Glu 22 was coupled 75°C. After the final removal of the Fmoc-protecting group at the N-*terminus*, a portion of the glycopeptide product was cleaved from the resin using the cleavage reagent K (TFA/Phenol/water/thioanisole/EDT 82.5/5/5/5/2.5) at r.t. for 2.5 hours, and analysed by LC-MS.



Scheme 34 – Incorporation of glycosylated asparagine and elongation of the fragment **6a**.

The desired product **26** and the capped fragment **22** were observed. However, the resolution of the chromatogram was significantly reduced due to the presence of a mixture of compounds showing masses corresponding to the desired products and the loss of one or two acetyl groups from the glucosamine units on Asn23 or Asn60 (figure 18).

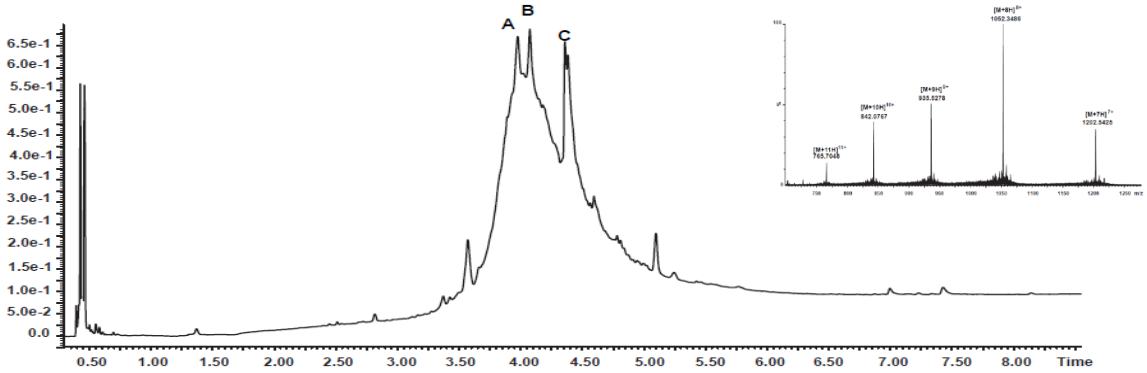


Figure 18 - LC-MS of crude fragment **6a** Thy 1(19-84) after TFA cleavage. Chromatogram recorded at 214 nm (RP-UPLC with C4 column, gradient: 5 to 70% of ACN in water + 0.1% of formic acid in 14 minutes). A: Product – 42 Da (-Acetyl) and -84 Da (-2 acetyls); B: Product **6a**; C: Capped peptide fragment 4. Mass spectra of the product **6a** (peak B). Mass cldc = 8410.9899 Da. Mass obs: $[M+7H]^{7+}$ = 1202.5425 Da, $[M+8H]^{8+}$ = 1052.3486 Da, $[M+9H]^{9+}$ = 935.5278 Da, $[M+10H]^{10+}$ = 842.0767 Da, $[M+11H]^{11+}$ = 765.7048 Da.

The glycopeptide **26** was purified by semi-preparative RP-HPLC on a C4 column using an extended gradient to improve the separation of the peaks and facilitate the isolation of compounds. The collected fractions were lyophilized to give 9.2 mg of the glycopeptide **6a** together with product minus 42 a.m.u., corresponding to a 1% yield based on the resin loading after the coupling of first amino acid and 22% based of the amount of crude material.

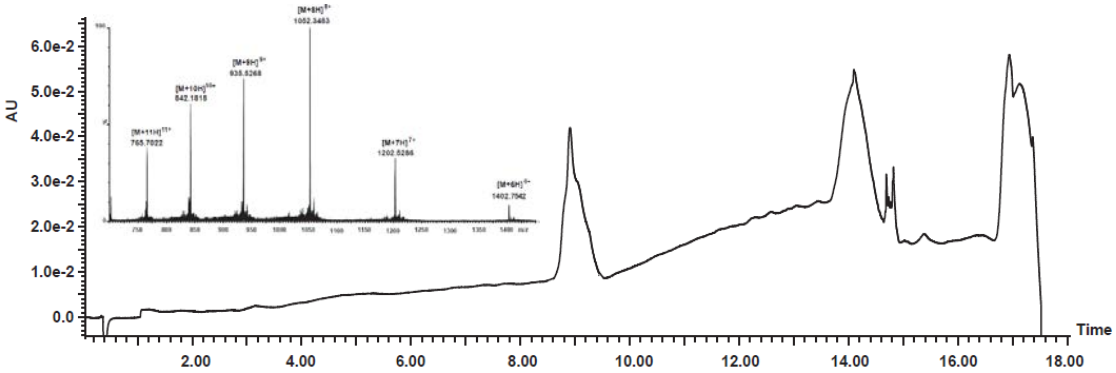


Figure 59 - LC-MS analysis of the purified glycopeptide **6a.** Chromatogram recorded at 214 nm (gradient: 10 to 50 of ACN in water in 18 minutes) and high resolution ESI-MS mass spectra of the product. Mass cldc = 8410.05 Da. Mass obs: $[M+7H]^{7+}$=1202.5425 Da, $[M+8H]^{8+}$=1052.3486 Da, $[M+9H]^{9+}$=935.5278 Da, $[M+10H]^{10+}$=842.0767 Da, $[M+11H]^{11+}$=765.7048 Da. Peaks at 14.00 and 17.00 min are artifacts of the column as confirmed by the blank run before the sample.

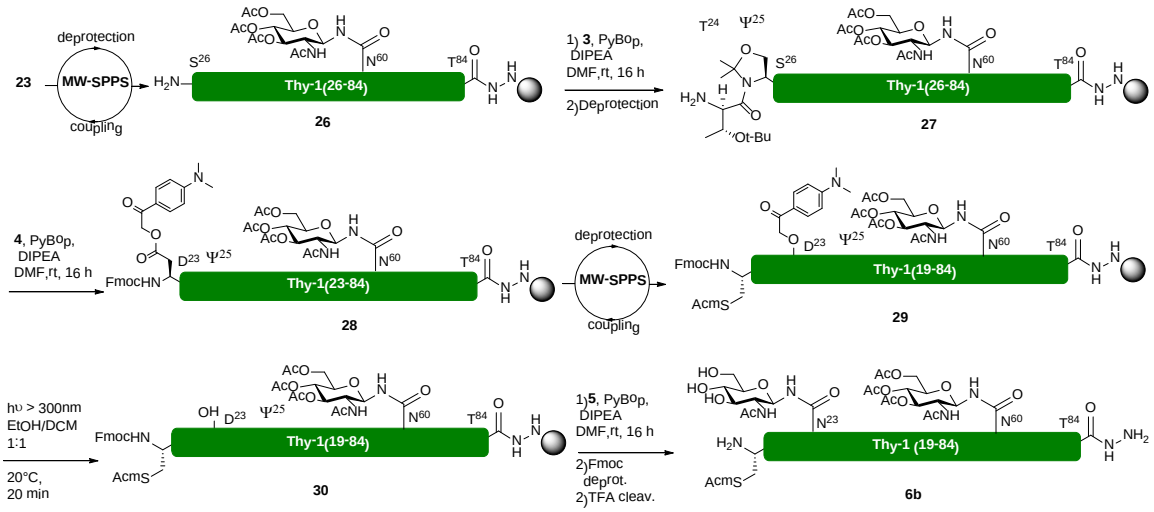
The glycoform **6a** was successfully obtained from the threonine anchored to the functionalized Wang resin **21**. The condition for the elongation of the peptide in the MW-assisted synthesiser, the selection of the deprotection solution and the manual coupling of the glycosylated asparagine **2** were investigated and optimized. Double coupling at 75° C for 5 minutes was required for glutamic acid, isoleucine, phenylalanine, and valine. The increase of the peptide length was considered, and appropriate amounts of solvents were employed. The temperature was adapted to different steps to assure the efficiency of reactions or to avoid the burning of the solid support. The addition of formic acid resulted effective in the minimization of the aspartimide formation affecting the aspartic acid during the removal of the Fmoc group. The reaction temperature for the deprotection was adjusted from 90°C to 75°C and MW power was reduced to avoid the rapid overheating of the reaction mixture in the synthesiser. The addition of DMSO facilitates the solubilisation of the glycosylated asparagine **2** in DMF and improves the coupling of the building block to peptide chains **22** and **24**.

2.4.5 Synthesis of the glycopeptide fragment 6b (isoform 2)

The installation of two different N-type glycans into the Thy-1 fragment (19-84) was a challenging goal of this project. The initial concept involved the installation of protected glucosamine on Asn60 and the glycosylation with a glycosyl amine at position 23 to have differentiated sites for the enzymatic elongation with oligosaccharides. The direct incorporation of an unprotected monosaccharide has disadvantages during the elongation of the peptide compared to the protected glucosamine, being the acylation of the free hydroxyl groups the main problem during the coupling reactions. Thus, a strategy to obtain the glycopeptide **6b** required an orthogonal protecting group

at one of the glycosylation positions that can be removed for the glycosylation at the end of the peptide assembly.

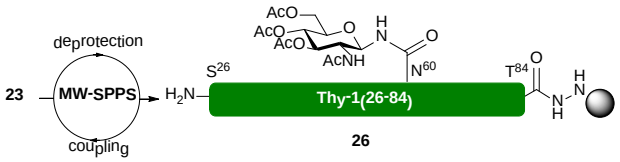
Starting from the glycopeptide-resin **23**, the peptide chain was elongated to obtain the fragment **26** followed the same conditions described to obtain **24**. Particular attention was given to the incorporation of two building blocks, the pseudoproline dipeptide **3** that was incorporated at the positions Thr24-Ser25, and the coupling of the Fmoc-Asp(OMap)-OH **4** replacing the asparagine at the glycosylated position (Asn23). Due to the limited availability, the pre-synthesised building blocks **3** and **4** were introduced manually and the efficiency of the couplings was verified before further elongation of the peptide.



Thy-1 (19-84) Sequence: H_2N-CRHE**N**($(OH)_3$-GlcNAc)TSSSPIQYEFSLTRETKKHVLFGTVGVPEHTYRSRT**N**($(OAc)_3$-GlcNAc)FTSKYNMKVLYLSAFTSKDEGTYT-NH-NH_2

Scheme 35 - Synthetic strategy to obtain the glycopeptide fragment **6b**.

Based on the analysis of the previous elongation of the fragment **23** by MW-SPPS and the presence of truncated sequences at amino acids having bulky side chains and protecting groups, the amino acids valine (at positions 50, 48, and 43), and phenylalanine (at positions 51, and 45). A further improvement in the quality of the peptide synthesis was achieved setting the double coupling of amino acids that are vicinal in the primary sequence, like in the case of the Lys40-Lys39, and Ser26-Ser25.



Scheme 36 – Elongation of fragment **26** with the optimized conditions.

The peptide was elongated until the Fmoc-removal from Ser26 and a portion of glycopeptide **26** was cleaved and analysed by LC-MS (figure 20). The UV profile showed the main peak corresponding to the desired product (peak B) and a successful elongation of the peptide and the optimization of the coupling and deprotection cycles.

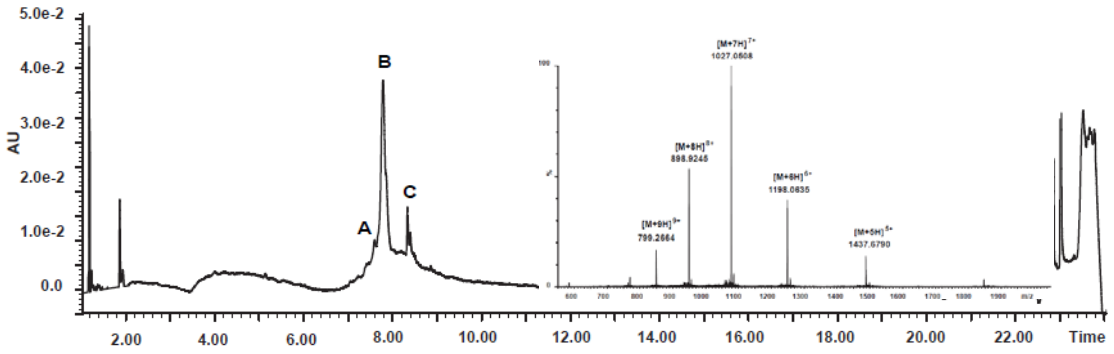
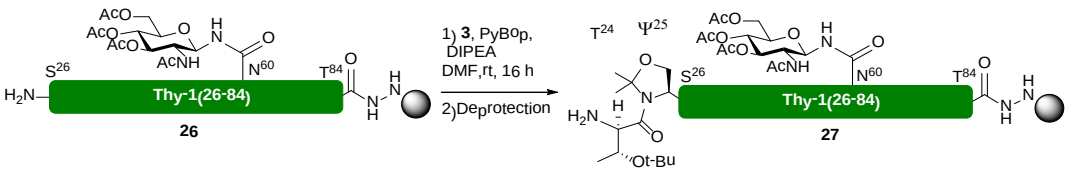


Figure 20 - LC-MS of Thy 1(26-84) **26.** Chromatogram recorded at 214 nm after cleavage (RP-UPLC with C18 column, gradient: 10 to 40% of ACN in water + 0.1% of formic acid in 23 minutes). A: Product – 42 Da (-Acetyl) and -84 Da (-2 acetyls); B: Product **26**; C: Capped peptide fragment **22**. Mass spectra of the product **26** (peak B). Mass cldc = 7182.5444 Da, Mass obs: $[M+5H]^{5+}$ = 1437.6790 Da, $[M+6H]^{6+}$ = 1198.0635 Da, $[M+7H]^{7+}$ = 1027.0508 Da, $[M+8H]^{8+}$ = 898.9245 Da, $[M+9H]^{9+}$ = 799.2664 Da.

The dipeptide Fmoc-Thr(tBu)-Ser($\Psi^{Me,Me}$ Pro)-OH **3** was coupled to the resin-bound fragment **26** using PyBOP and DIPEA during 16 hours. The efficiency of the coupling was analysed by LC-MS (figure 21). Three main peaks were identified in the resultant chromatogram: the desired glycopeptide fragment **27** at *Rt* = 8.03 min (peak B), the peak corresponding to the collection of all removed protecting groups and the capped peptide **22** at *Rt* = 8.40 min (peak C), and two peaks resulting from the loss of two different acetyl groups from glucosamine on Asn60 (peak A).



Scheme 37 – Incorporation of pseudoproline dipeptide in the sequence.

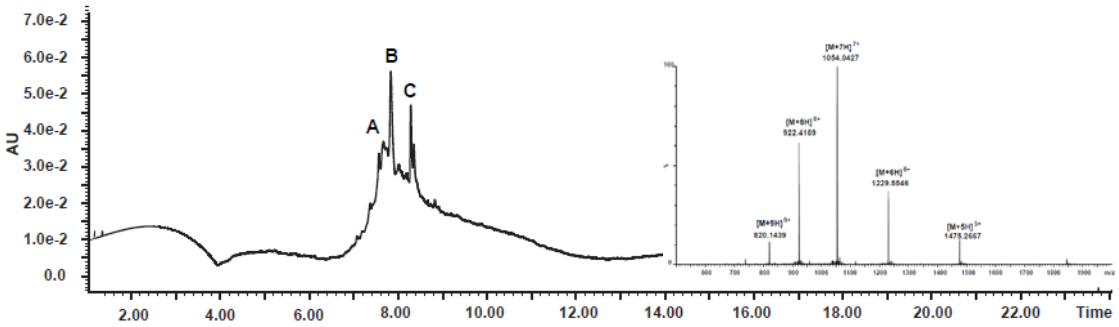
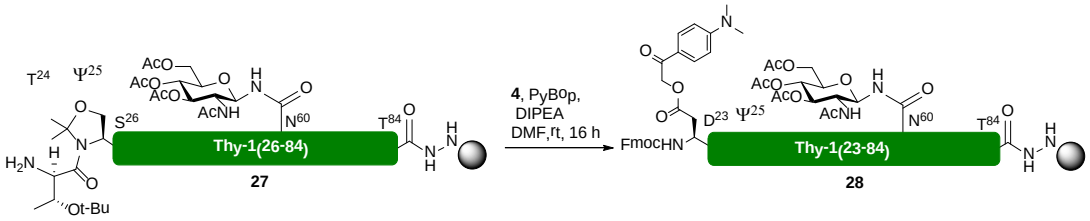


Figure 21– LC-MS analysis of fragment **27** after cleavage with TFA. UV-chromatogram obtained at 214 nm (RP-UPLC C18 column, gradient: 10 to 40% of ACN in water + 0.1% of formic acid in 23 minutes). A: Product – 42 Da (-Acetyl) and -84 Da (-2 acetyls); B: Product **27**; C: capped peptide fragment **22**. Mass spectra of the product **27** (peak B). Mass cldc = 7370.6352 Da, Mass obs: $[M+5H]^{5+}$ = 1475.2667 Da, $[M+6H]^{6+}$ = 1229.5546 Da, $[M+7H]^{7+}$ = 1054.0427 Da, $[M+8H]^{8+}$ = 922.4169 Da, $[M+9H]^{9+}$ = 820.1439 Da.

The Fmoc- group on **27** was removed and the Fmoc-Asp(OMap)-OH **4** was coupled to the free amino group using PyBOP and DIPEA for the activation. The reaction was carried out manually at r.t. and in absence of light to prevent the removal of the photo-cleavable group from the side chain of aspartic acid.



Scheme 38 – Coupling of Map-protected aspartic acid.

After 16 hours, a portion of the resin was treated with the Reagent K and the cleaved peptide was analysed by LC-MS (figure 22). Two main peaks corresponding to desired glycopeptide **28** and to the capped peptide fragment **22** were observed together with the products resulting from the removal of up to two acetyl groups. Spite of severe extractions with ether, some by-products from the protecting groups were also observed. The analysis and separation of the products by LC were difficult due to the similar polarity of the product and the capped fragment **22** that co-eluted at the same retention time.

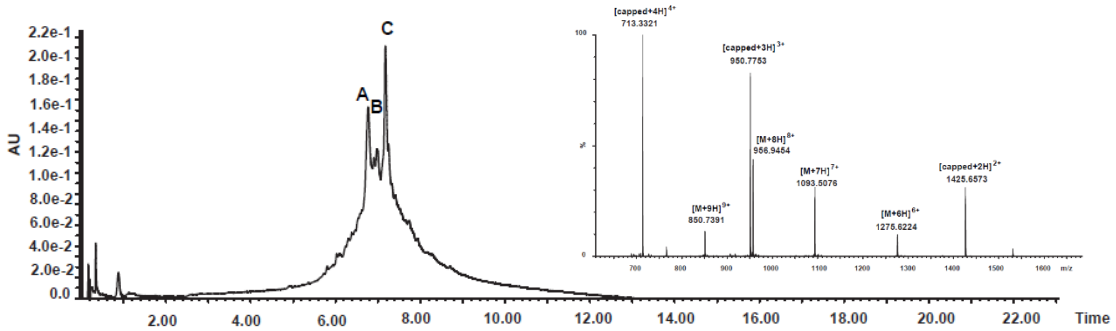
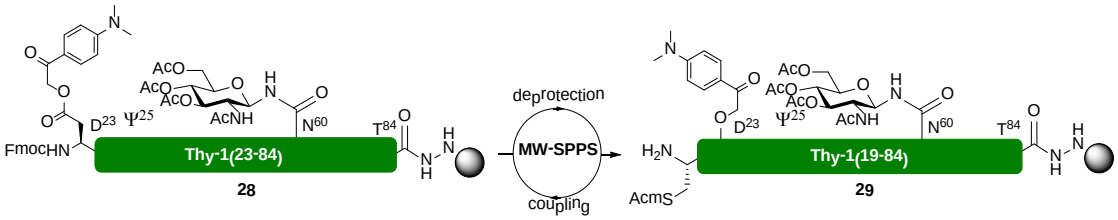


Figure 22 - LC-MS analysis of fragment **28** after cleavage with TFA. Chromatogram recorded at 214 nm (RP-UPLC with C18 column, gradient: 10 to 40% of ACN in water + 0.1% of formic acid in 23 minutes). A: Product +57 Da (+tert-butyl); B: Product – 42 Da (-acetyl) and -84 Da (- 2 acetyls); C: Product **28** + Capped fragment **22**. Mass spectra of the product **28** (peak C). Mass cldc = 7646.7420 Da, Mass obs: $[M+6H]^{6+}$ = 1275.6224 Da, $[M+7H]^{7+}$ = 1093.5076 Da, $[M+8H]^{8+}$ = 956.9454, $[M+9H]^{9+}$ = 850.7391 Da. $[M_{capped}+2H]^{2+}$ = 1425.6573 Da, $[M_{capped}+3H]^{3+}$ = 950.7753 Da, $[M_{capped}+4H]^{4+}$ = 713.3321 Da

The fragment **28** was elongated in the MW-assisted synthesizer until complete the sequence Thy-1 (19-84) **29** using the optimized deprotection and coupling cycles described before.



Scheme 39 – Elongation of fragment **29** in the MW-assisted synthesiser.

The Fmoc group of **29** was removed, a portion of the resin was treated with TFA and the resulting crude product was analysed by LC-MS (figure 23).

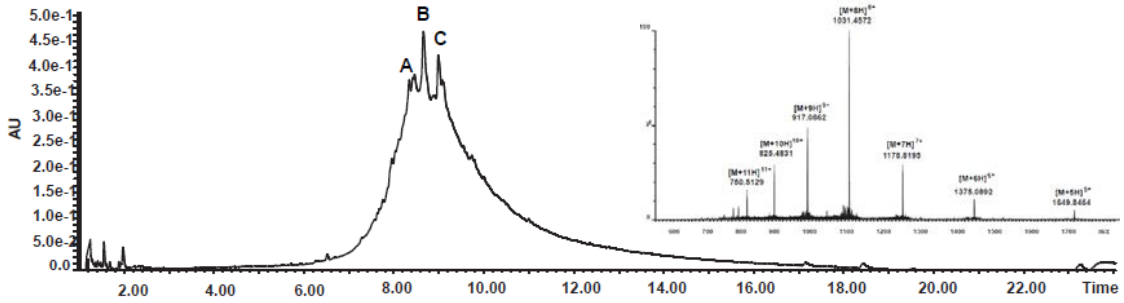
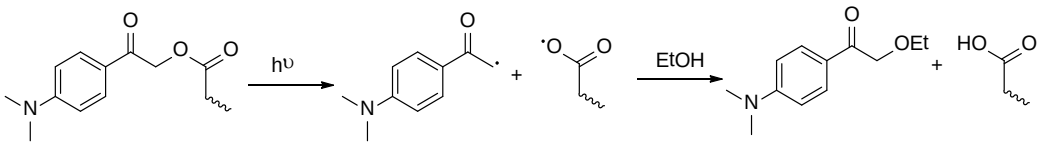


Figure 6 - LC-MS analysis of fragment **29** after cleavage with TFA. Chromatogram recorded at 214 nm (RP-UPLC with C18 column, gradient: 10 to 40% of ACN in water + 0.1% of formic acid in 23 minutes). A: Product – 42 Da (-acetyl) and -84 Da (- 2 acetyls); B: Product **29**; C: Capped fragment **22**. Mass spectra of the product **29** (peak B). Mass cldc = 8243.9794 Da, Mass obs: $[M+5H]^{5+}$ = 1649.8464 Da, $[M+6H]^{6+}$ = 1375.0892 Da, $[M+7H]^{7+}$ = 1178.8195 Da, $[M+8H]^{8+}$ = 1031.4572 Da, $[M+9H]^{9+}$ = 917.0862 Da, $[M+10H]^{10+}$ = 825.4831 Da, $[M+11H]^{11+}$ = 750.5129 Da

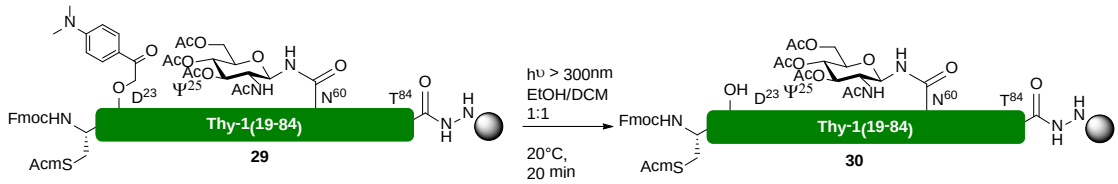
The chromatogram showed a broad signal containing a peak of the desired product **29** observed at *Rt* = 7.93 min; two peaks corresponding to fragment lacking one or two acetyl groups *Rt* = 6.40-6.53 min, and a peak from the capped peptide fragment **22** at *Rt* = 8.43 min. Different gradients and columns were tested to increase the separation of the peaks without a significantly improvement of the chromatographic profile. Thus, it was decided to continue the process and try to improve the separation of the products after the following steps.

The removal of the photo-labile group (Map) from the side chain of the Asp23 was performed using light irradiation at 320 nm of the swollen resin in a mixture of ethanol and dichloromethane (1:1). The proposed mechanism for the photochemical reaction involves the homolytic fission of the carbon-oxygen bond between the phenacyl group and the carboxylate of the aspartic acid to give two stable radicals (scheme 40).



Scheme 40 – Proposed mechanism for the photo-removal of the Map group from the side chain of the aspartic acid.

The selection of the appropriate solvent played an important role in the efficiency of the reaction. The resulting products of the Map-group removal in three different solvents were analysed by LC-MS after cleavage of the correspondent products with TFA: dichloromethane (Figure 24a), a mixture of methanol and dichloromethane 1:1 (Figure 24b), and the ethanol-dichloromethane solution 1:1 (Figure 24c).



Scheme 41 – Photo-deprotection of fragment **29**.

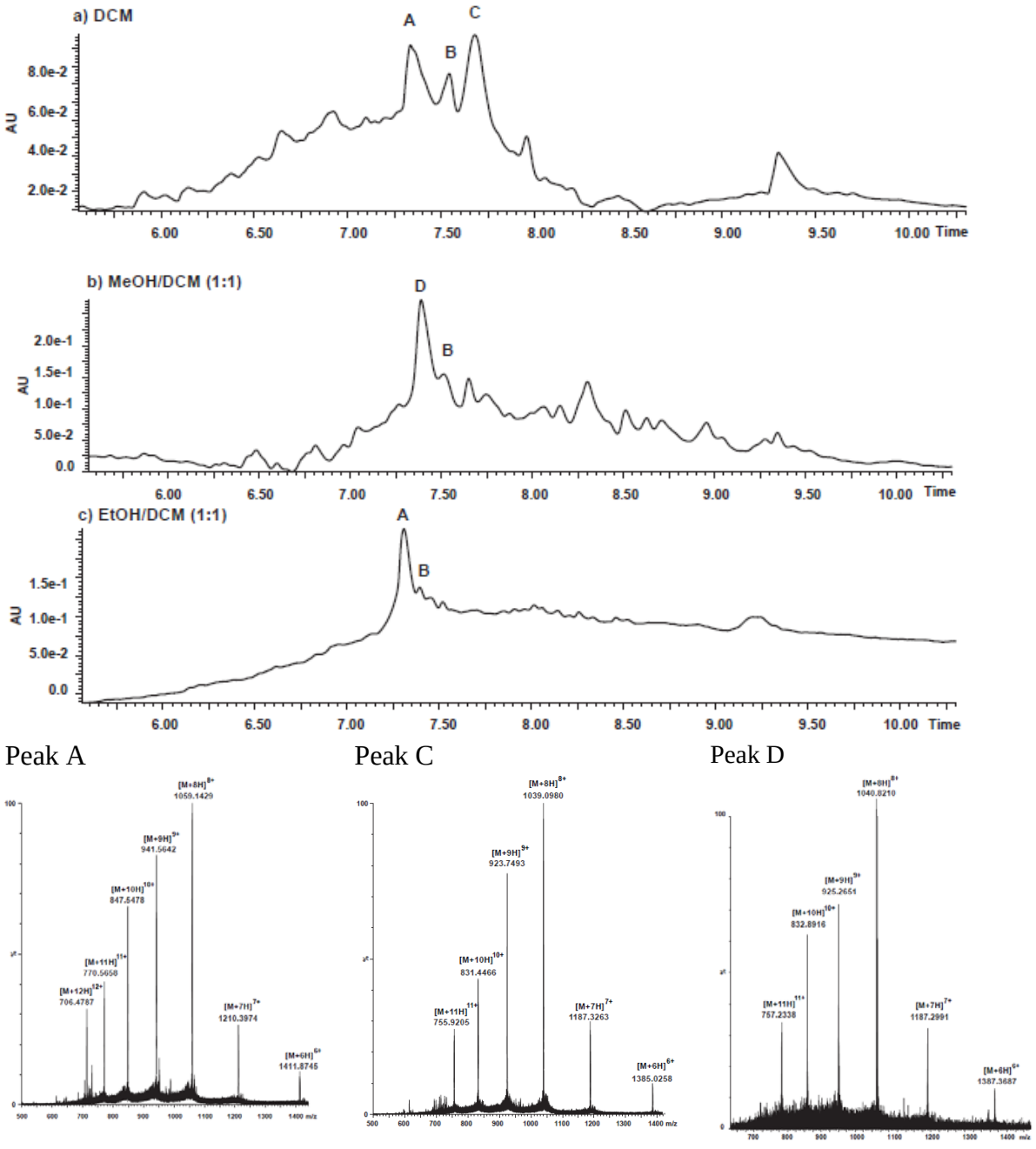
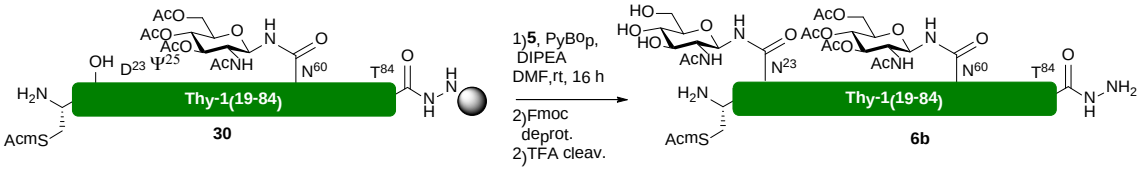


Figure 7 – LC-MS analysis of the removal of the Map-photolabile protecting group from **29**. UV-chromatograms recorded at 214 nm after photo-deprotection in different solvent. MS-spectra for peaks A, C, and D. Separation by RP-UPLC on a C18 column, gradient: 10 to 40% of ACN in water + 0.1% of formic acid in 23 minutes.a) in DCM Mass obs (peak A): $[M+6H]^{6+}$ = 1385.0258 Da, $[M+7H]^{7+}$ = 1187.3263 Da, $[M+8H]^{8+}$ = 1039.0980 Da, $[M+9H]^{9+}$ = 923.7493 Da, $[M+10H]^{10+}$ = 831.4466 Da, $[M+11H]^{11+}$ = 755.9205 Da. Mass obs (peak C): $[M+6H]^{6+}$ = 1411.8745 Da, $[M+7H]^{7+}$ = 1210.3974 Da, $[M+8H]^{8+}$ = 1059.1429 Da, $[M+9H]^{9+}$ = 941.5642 Da, $[M+10H]^{10+}$ = 847.5478 Da, $[M+11H]^{11+}$ = 770.5658 Da, $[M+12H]^{12+}$ = 706.4787 Da. b) in MeOH/DCM (1:1) Mass obs (peak D): $[M+6H]^{6+}$ = 1387.3687 Da, $[M+7H]^{7+}$ = 1187.2991 Da, $[M+8H]^{8+}$ = 1040.8210 Da, $[M+9H]^{9+}$ = 925.2651 Da, $[M+10H]^{10+}$ = 923.8916 Da, $[M+11H]^{11+}$ = 757.2338 Da.; c) EtOH/DCM (1:1). Peak B corresponded to the capped fragment **22.**

The deprotection of aspartate in DCM resulted in the incomplete removal of the protecting group with peaks A (desired peptide) and C (Map-protected peptide) as the main products. By the photoreaction in a mixture of methanol and dichloromethane, a methyl ester was obtained on the side chain of the aspartic acid (figure 24b, peak D). Better results were obtained by performing the reaction in the ethanol-dichloromethane mixture giving a major product corresponding to the selective removal of the photo-labile group. The peak A corresponding to the expected glycopeptide **30** appeared broad due to the co-elution of by-products derived from the loss of the acetyl groups from the protected glucosamine on the residue 60.

Finally, the 1-amino-*N*-acetyl glucosamine **5** was linked via Lansbury aspartylation to the peptide. The resin-bound glycopeptide **30** having unmasked carboxylic acid on the Asp23 was activated by the addition of PyBOP and DIPEA as a base and the glycosyl amine **5** was added. The reaction proceeded for 16 hours and the formation of the product was analysed by LC-MS. The MS analysis of the peaks A, B, and C of the chromatogram showed that the glycosylation occurred, although other products were also present (figure 25).



Scheme 42 – Lansbury aspartylation of fragment **30**.

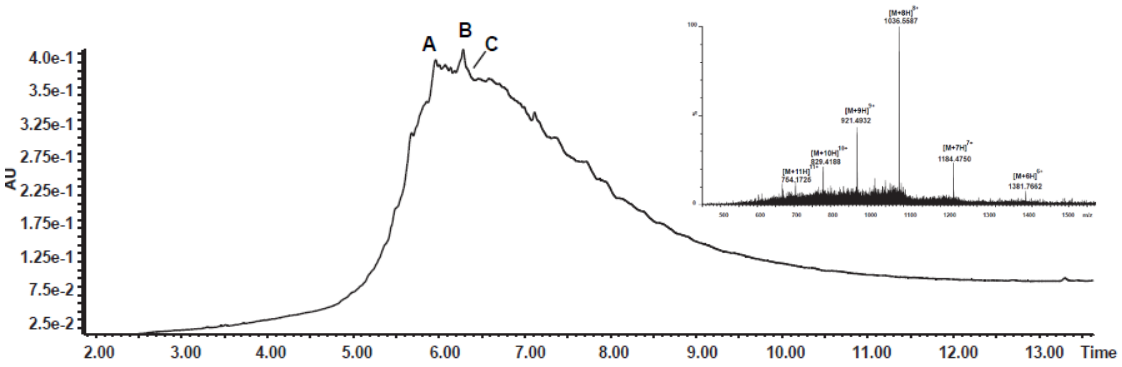


Figure 25– LC-MS of the crude glycoform of Thy-1 (19-84) **6b**. a) Chromatogram recorded at 214 nm of the crude fragment **6b** after cleavage (RP-UPLC with C4 column, gradient: 10 to 40% of ACN in water + 0.1% of formic acid in 23 minutes). A: Product – 42 Da (-acetyl) and -84 Da (- 2 acetyls); B: Product **6b**; C: Capped fragment 4. b) Mass spectra of the product **6b** (peak B). Mass cldc = 8243.01 Da, Mass obs: $[M+6H]^{6+}$=1381.7662 Da, $[M+7H]^{7+}$=1184.4750 Da, $[M+8H]^{8+}$=1036.5587 Da, $[M+9H]^{9+}$=921.4932 Da, $[M+10H]^{10+}$=829.4188 Da, $[M+11H]^{11+}$=754.1725 Da.

Probably due to their similar polarity, the glycoform **6b** and the by-products co-eluted making challenging the clear detection and the isolation of the desired product. The most difficult fragments to remove corresponded to the products losing one or more acetyl groups on the glucosamine at Asn60. The glycopeptide **6b** was obtained in 2% yield (9.2 mg) based on the loading of the first amino acid, 12% compared based on the crude peptide, and with acceptable grade purity. This fragment was used for the ligation reactions.

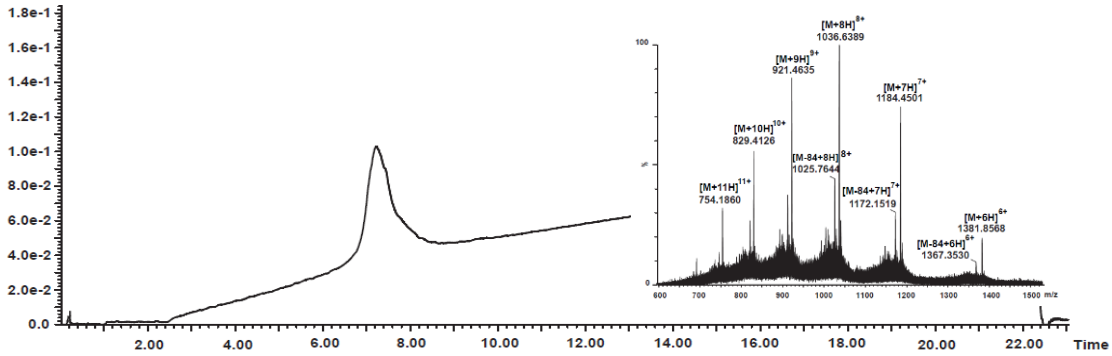
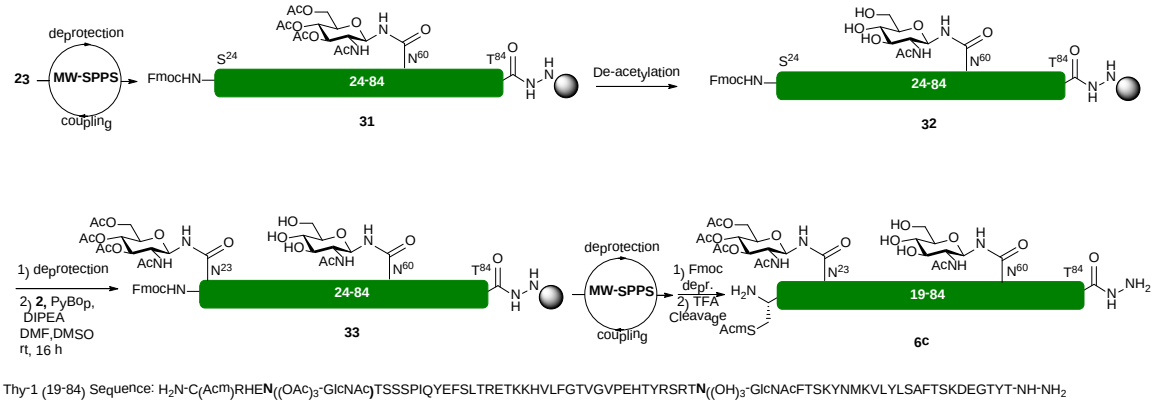


Figure 26– LC-MS analysis of the purified peptide fragment **6b**. UV-chromatogram recorded at 214 nm (gradient: 10 to 40 of ACN in water + 0.1% of formic acid in 23 minutes) and high-resolution ESI- mass spectra of the product. $[M+H]^{+}_{cald}$ = 8285.0125. M_{obsd} : $[M+6H]^{6+}$ = 1381.8568, $[M+7H]^{7+}$ = 1184.4501, $[M+8H]^{8+}$ = 1036.6389, $[M+9H]^{9+}$ = 921.4635, $[M+10H]^{10+}$ = 829.4126, $[M+11H]^{11+}$ = 754.1860; $[M-84+6H]^{6+}$ = 1367.3530, $[M-84+7H]^{7+}$ = 1172.1519, $[M-84+8H]^{8+}$ = 1025.7644, $[M-84+9H]^{9+}$ = 911.7928, $[M-84+10H]^{10+}$ = 820.6080, $[M-84+11H]^{11+}$ = 745.9812.

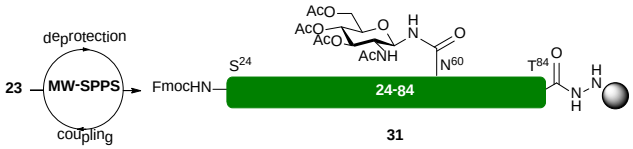
2.4.6 Synthesis of the glycopeptide fragment 6c (isoform 3)

The two strategies described above presented one main problem in common: the loss of one or more acetyl groups from glucosamine. The removal occurred during the elongation of the glycopeptide and the repeated treatment with piperidine resulting in lower yield and difficult purification of the product. An alternative strategy to generate the isoform (19-84) having a protected glucosamine on one glycosylation site and a deprotected glucosamine on the other asparagine is depicted in scheme 43. It involved the coupling of O-*per*-acetylated glucosamine and **23** to generate fragment **31** and the following elongation and deacetylation of glucosamine to give glycopeptide **32**. It was considered that any possible acylation during the final elongation will be removed during the treatment with piperidine. A second glycosylated asparagine was coupled to the N-*terminus* **33** and the glycopeptide was elongated carefully to obtain the glycopeptide. A final cleavage from the resin gave the desired peptide fragment **6c**.



Scheme 43- Synthetic strategy to obtain the glycopeptide fragment **6c**.

The glycosylated fragment **23** was synthesised as described in section 2.4.4 and it was used as starting material for this strategy. The elongation of **23** was performed until Ser24 using the optimized deprotection and coupling cycles. LC-MS analysis after TFA cleavage of the resulting glycopeptide **31** revealed the loss of an acetyl group from the glucosamine as expected and observed for the previously synthesised glycopeptide (figure 27).



Scheme 44 – Elongation of fragment **31** in the MW-assisted synthesiser.

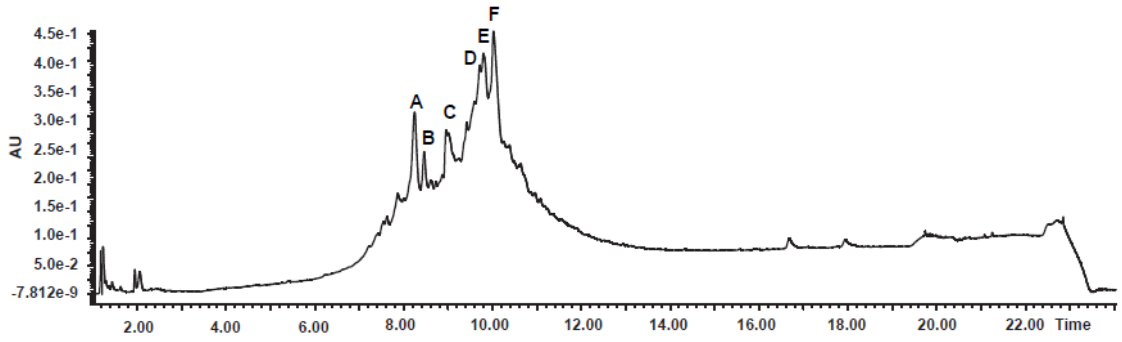
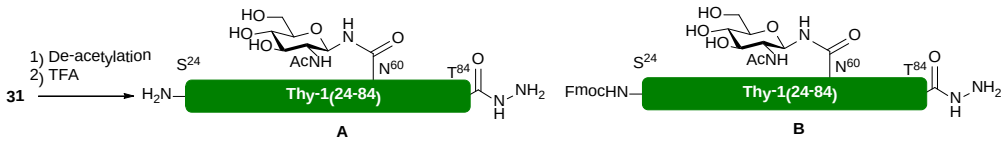


Figure 27 – LC-MS of the crude of Thy-1 (24-84) **31** recorded at 214 nm after cleavage (RP-UPLC with C18 column, gradient: 10 to 40% of ACN in water + 0.1% of formic acid in 23 minutes). A: Product – 42 Da (-acetyl) and -223 Da (- Fmoc); B**:** Unidentified mass; C: Capped peptide fragment **22**; D**:** Product -84 Da (-2 Acetyls); E: Product – 42 Da (Acetyl); F: Product **31**.

The level of removal and the site of the sugar that lose the acetyl group during the elongation was not investigated here. However, this disadvantage turned in favor of the planned strategy requiring the deacetylation of glucosamine in the glycopeptide **31**. The complete removal of the acetyl

groups on the solid support was investigated by a small scale by screening of various solvents and reaction time (table 7) and the following monitoring by LC-MS analysis (figure 28).



Scheme 45 - Deacetylation of **31** followed by acidic cleavage.

Entry	Conditions	Time (hours)	Conversion[a]
1	5% hydrazine in DCM/MeOH (1:1)	2	50%
2	5% hydrazine in THF/MeOH (1:1)	2	~ 57%
3	6mM MeONa in DMF/MeOH (17:3)	6	Undetected
4	5% hydrazine in THF/MeOH (1:1)	4	>98%
5	10% hydrazine in THF/MeOH (1:1)	2	~90%

Table 7 - Screening of different mixtures and reaction times for the O-deacetylation on resin of glycopeptide **31**. [a)] The conversion was calculated by integration of the areas from the chromatograms at 214 nm.

Partial de-acetylation was observed after treatment with 5% hydrazine in DCM/methanol (Entry 1). Similar results were obtained by changing the solvent to a THF and methanol mixture (Entry 2) and no de-acetylation was observed with 6 mM of sodium methoxide in DMF and methanol were used (Entry 3), showing that a proper swelling of the resin was necessary for the reaction progress. [161]

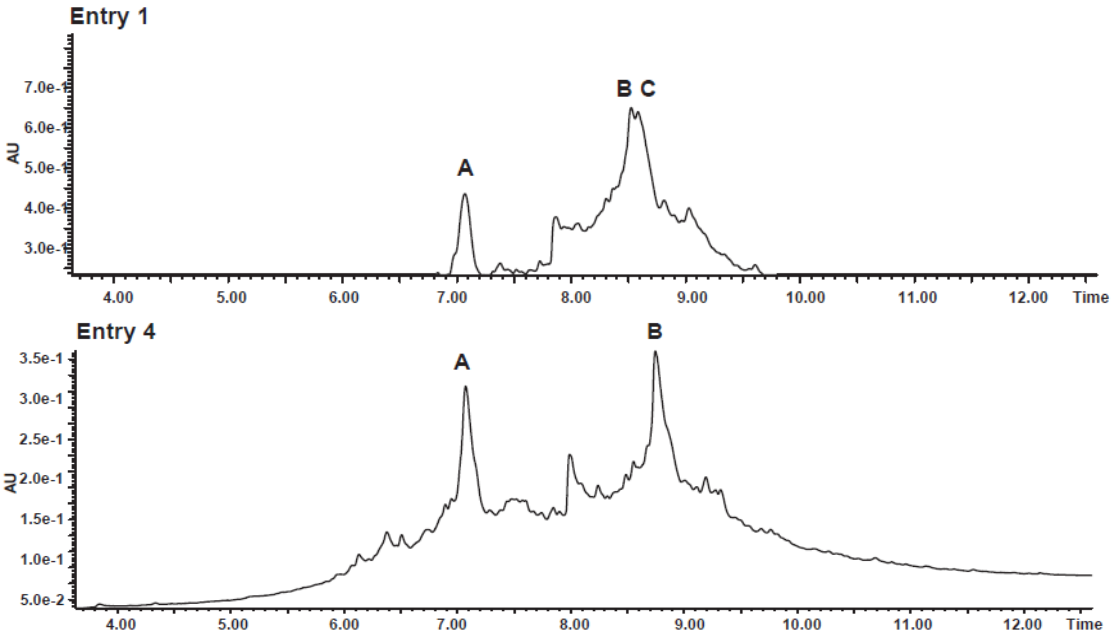
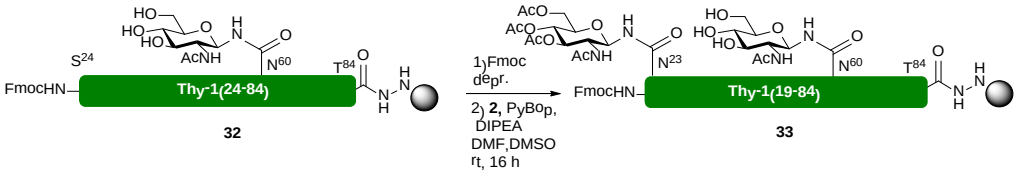


Figure 28– Comparison of the conditions for the deacetylation of the fragment **31** evaluated using RP-UPLC and ESI-QTof analysis. A: Product -222 Da (- Fmoc), B: Product **32**, C: Fragment **31** (No deacetylation), D: Unidentified mass, E: Fragment **32** -222 + 42 Da (- Fmoc, + Acetyl, respectively).

Since deacetylation in THF and methanol showed the best promising results, the de-acetylation was evaluated using 5% (Entry 4) and 10% of hydrazine (Entry 5) in this solvent. The last two conditions delivered similar results, with a better chromatographic profile for the 5% hydrazine solution. It is noteworthy that despite the low swelling of the resin in THF/methanol, the de-acetylation of glycopeptide **31** was completed after 4 hours using the appropriate amount of solvent.

The Fmoc group of deacetylated fragment **32** was removed and the second glycosylated asparagine **2** was coupled under the same reaction conditions used for Asn60. The LC-MS analysis of the cleaved peptides confirmed the completion of the reaction (figure 29).



Scheme 46 – Incorporation of glycosylated asparagine on **32** after Fmoc deprotection.

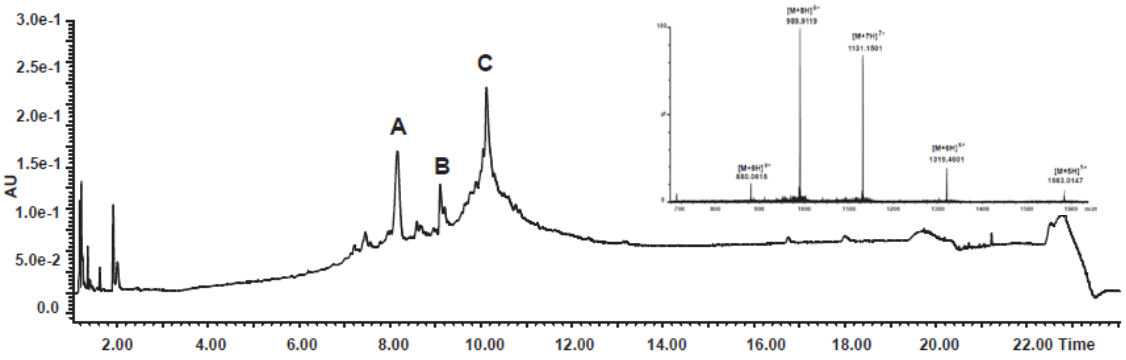
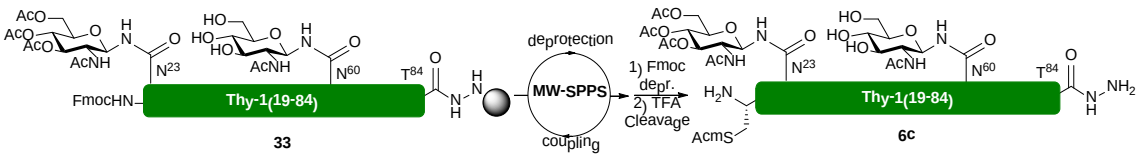


Figure 29 – LC-MS of the Thy 1 (23-84) **33**. Chromatogram recorded at 214 nm of the crude fragment **33** after cleavage (RP-UPLC with C18 column, gradient: 10 to 50% of ACN in water + 0.1% of formic acid in 23 minutes). A: Unidentified mass; B**:** Capped peptide fragment **22**; C**:** Product **33**. Mass spectra of the product (peak C). Mass cldc = 7909.8276 Da, Mass obs: $[M+5H]^{5+}$ = 1583.0147 Da, $[M+6H]^{6+}$ = 1319.4501 Da, $[M+7H]^{7+}$ = 1131.1501 Da, $[M+8H]^{8+}$ = 989.9119 Da, $[M+9H]^{9+}$ = 880.0615 Da.

The glycopeptide fragment **33** was elongated in the MW-assisted synthesizer selecting carefully the deprotection and coupling cycles for each amino acid as before. LC-MS analysis after the elongation showed the main peak corresponding to the desired product **6c**, which was resolved from the other by-products (figure 30). The glycopeptide **6c** was purified easily using semi-preparative

RP-HPLC to afford 6 mg of the pure product (1% yield base on resin loading, 11% base on the crude material). The strategy for the synthesis of fragment **6c** resulted in the best route to obtain the fragment of Thy-1 and a differentiation between the two glucosamines units.



Scheme 47 – Elongation of the isoform **6c**.

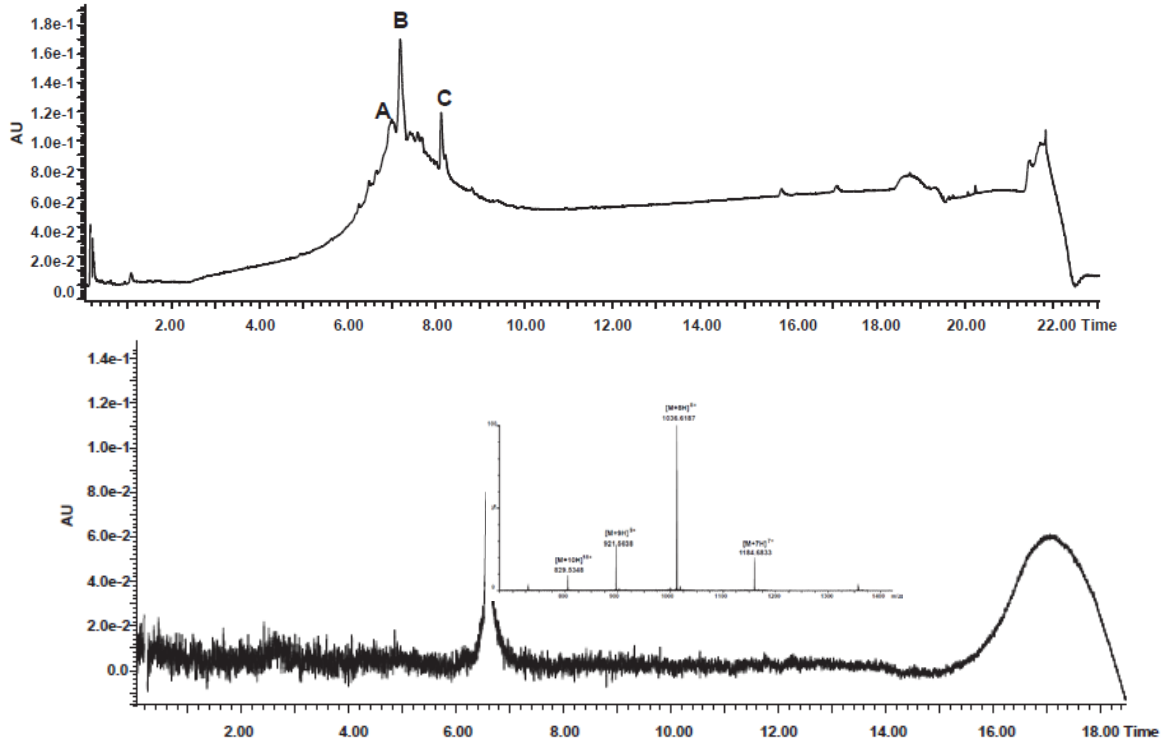
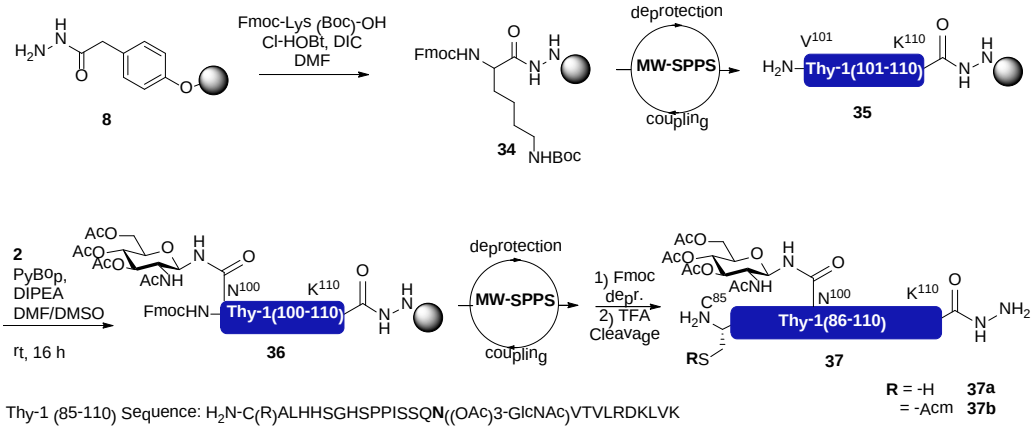


Figure 30 – LC-MS of the crude and the pure glycoform **6c**. Chromatogram recorded at 214 nm of the crude fragment **6c** after cleavage (RP-UPLC with C18 column, gradient: 10 to 50% of ACN in water + 0.1% of formic acid in 23 minutes). A: Product -42 Da (- acetyl); B**:** Product **6c**; C: Capped peptide fragment **22**. Chromatogram of pure peptide fragment **32** recorded at 214 nm (gradient: 10 to 40 of ACN in water in 20 minutes).Mass spectra of the product **6c** (peak B). Mass cldc = 8284.0127 Da, Mass obs: $[M+7H]^{5+}$ = 1184.6833 Da, $[M+8H]^{8+}$ = 1036.6137 Da, $[M+9H]^{9+}$ = 921.5638 Da, $[M+10H]^{10+}$ = 829.5348 Da.

2.5 Synthesis of the glycopeptide fragment (85-110)

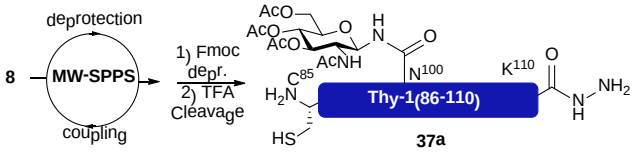
The third glycopeptide fragment required for the assembly of the Thy-1 glycoprotein contains 26 amino acids and a glycosylation on the Asn100. Similar to the other fragments, this glycopeptide was synthesised as peptide-hydrazine to allow the conversion into an appropriate thioester and

allow the ligation to GPI molecule by native chemical ligation. The general protocol for the synthesis of this glycopeptide **37** is presented in scheme 48. The identification of truncated sequences and side products was used to optimize the conditions for the peptide assembly. Two variants of the glycopeptide **37** were obtained, one having a terminal free cysteine (**37a**) and the second one having an Acm-protected cysteine (**37b**). These two glycopeptides were synthesised and used in the assembly of the Thy-1 protein.



Scheme 48 - Synthetic route for generating the two variants of glycopeptide **37**.

The synthesis started with the coupling of Fmoc-Lys(Boc)-OH to the hydrazide-functionalized Wang resin **8**. The amino acid was coupled to 93% of the active sites of the resin, showing a high efficiency of the reaction in this resin. The unreacted sites were capped with a solution of Ac_2O/DIPEA/DMF (1:1:8). The elongation of the **35** was carried out by MW-SPPS using the optimized coupling and deprotection conditions used before. After the removal of the Fmoc-group from Val101, the resin was treated for 16 hours with an excess of glycosylated asparagine **2** pre-activated with PyBOP and DIPEA in DMF. The peptide-resin was treated with Ac_2O/DIPEA/DMF to cap any unreacted groups and the sequence was elongated further in the synthesiser. The following amino acids were coupled twice (serine at position 97 and 98, proline 94 and 95); histidine at positions 92, 89, and 88 were coupled twice at 50 °C for 10 min after the double deprotection of the respective earlier amino acid (Ser93, Ser90, and His88). LC-MS analysis of cleaved peptide after the elongation showed that approximately 30% of the crude corresponded to the undesired cyclization of the aspartic acid to the correspondent succinimide (figure 31). However, the quality of the peptide was enough to continue the process.



Scheme 49 – Fragment **37a**.

The Fmoc of the peptide-resin was removed completely, the cysteine (Fmoc-Cys(Trt)-OH) was coupled manually using DIC and Cl-HOBt for activation. Finally, the Fmoc of Cys85 was removed and the peptide was cleaved from the resin using the reagent K. The product was purified by RP-HPLC on C18 column obtaining 10.35 mg of the fragment **37a** (4% based on initial resin loading, 21% on the crude obtained).

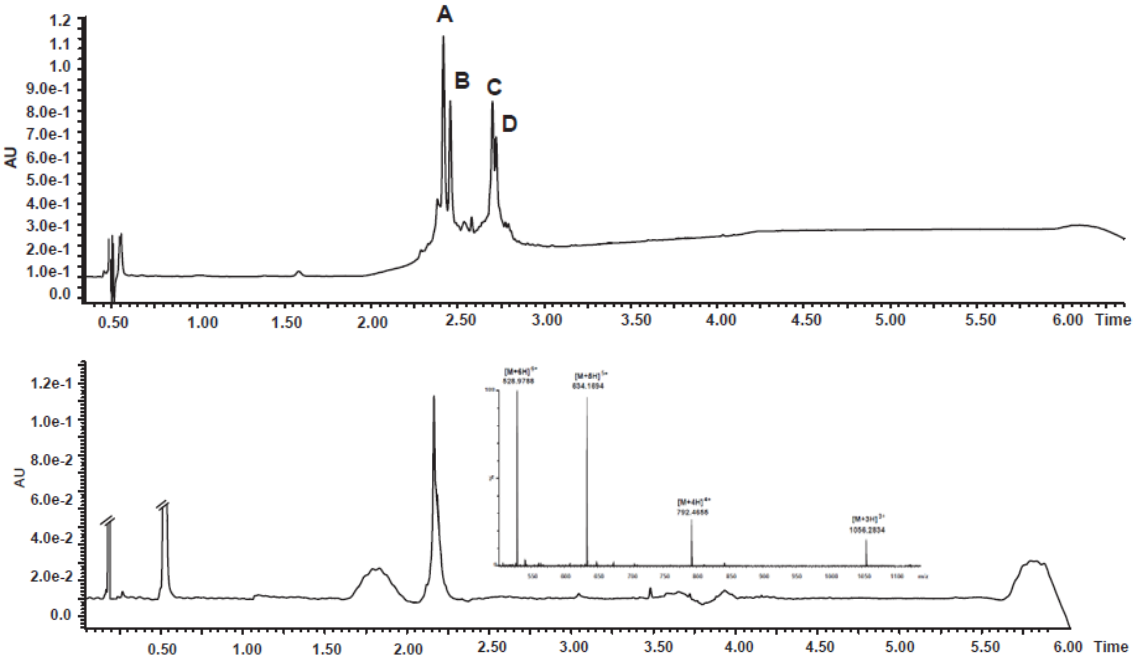
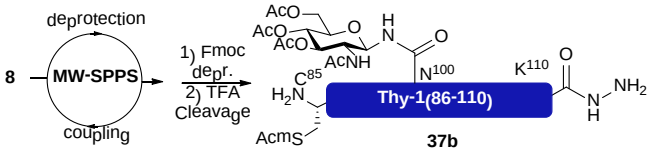


Figure 31- LC-MS of the crude and the pure glycopeptide Thy-1(85-110) **37a**. Chromatograms recorded at 214 nm of the crude fragment after cleavage (RP-UPLC with C18 column, gradient: 5 to 70% of ACN in water + 0.1% of formic acid in 6 minutes). A: Product **37a**; B: Product– 18 Da (aspartimide formation); C: Product + 222 Da (+Fmoc); D: Product +222 – 18 Da (Fmoc, aspartimide formation). High resolution mass spectra of the product **37a** (peak C). Mass cldc = 3165.6353 Da, Mass obs: $[M+3H]^{3+}$ = 1056.2834 Da, $[M+4H]^{4+}$ = 792.4655 Da, $[M+5H]^{5+}$ = 634.1694 Da, $[M+6H]^{6+}$ = 528.9788 Da.

The glycopeptide **37b** was obtained by elongation of **36** using the same coupling cycles used for **37a**. However, to reduce the aspartimide formation, the deprotection cycles were performed with 20% piperidine in DMF containing 0.7% of FA. The LC-MS analysis of the crude glycopeptide **37b** confirmed the efficient suppression of the aspartimide formation after the use of the acid in the deprotection cocktail.



Scheme 50 – Fragment **37b**.

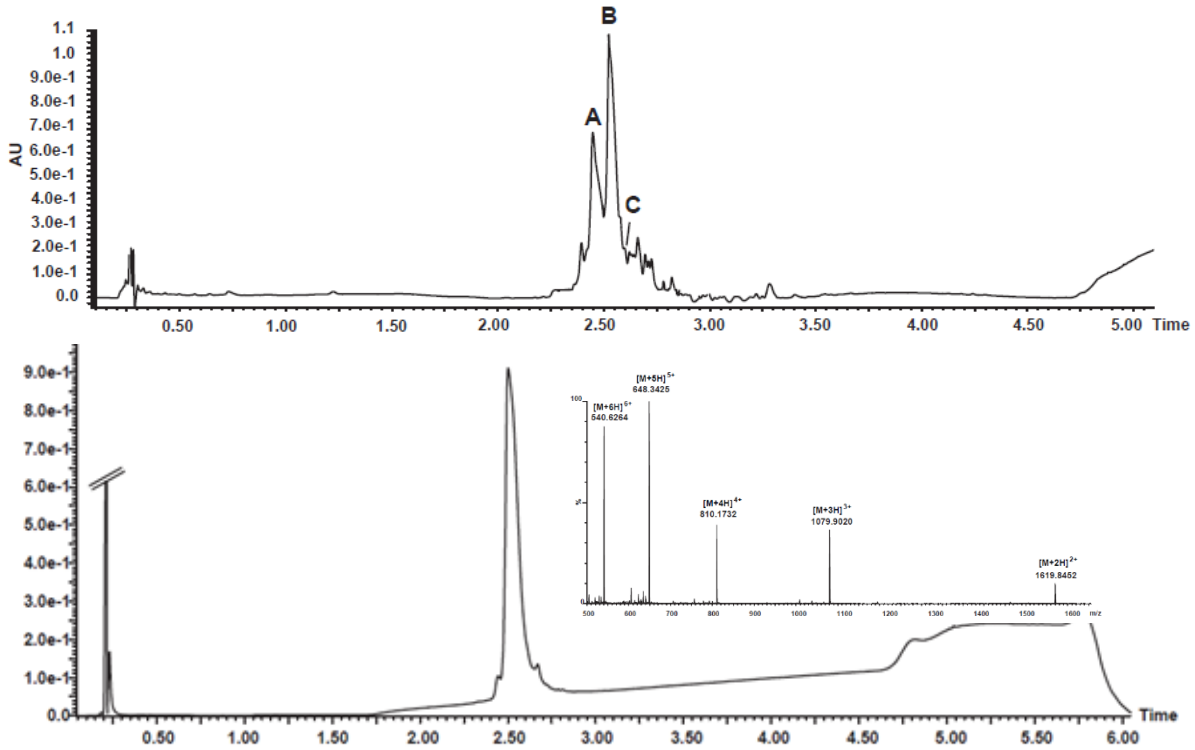


Figure 32 - LC-MS of the crude and the pure glycopeptide Thy-1(85-110) **37b**. **a)** Chromatogram recorded at 214 nm after cleavage (RP-UPLC with C18 column, gradient: 10 to 50% of ACN in water + 0.1% of formic acid in 6 minutes). A: Product – 128 Da (- lysine); B**:** Product **37b**; C: Product – 18 Da (aspartimide formation). b) … Mass spectra of the product 37b (peak B). Mass cldc = 3236.6764 Da, Mass obs: $[M+2H]^{2+}$ = 1619.8452 Da, $[M+3H]^{3+}$ = 1079.9020 Da, $[M+4H]^{4+}$ = 810.1732 Da, $[M+5H]^{5+}$ = 648.3425 Da, $[M+6H]^{6+}$ = 540.6264 Da.

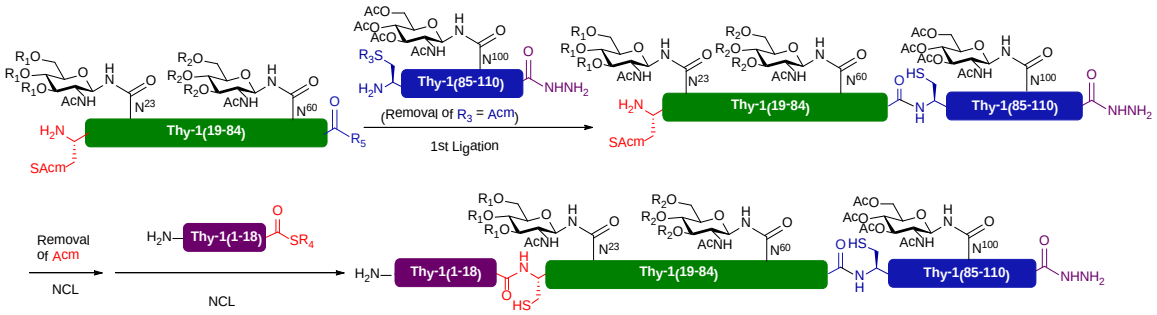
The glycopeptide was cleaved from the resin and purified by RP-HPLC to give the desired glycopeptide. The yield of the process improved to 6.7% based on the resin loading, 27% compared to the crude obtained giving 17.5 mg of the product. The formation of the truncated peptide detected by LC-MS was not observed in the elongation to obtain **37a**, the problem was attributed to a wrong delivery of some amino acid solutions in the synthesizer and not to problems associated with the optimized cycles.

2.6 Assembly of the Thy-1 Glycoprotein

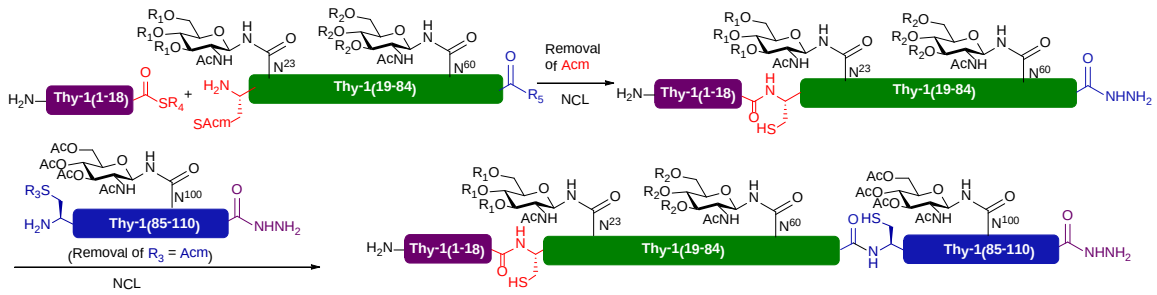
Native chemical ligation has been extensively employed in the synthesis of proteins starting from unprotected polypeptide chains. To assure the correct assembly of Thy-1, the fragments involved in the ligation required specific functional groups at the N- and C-*termini*. The selection of the route for the assembly of the protein was directed by several considerations: the availability of the synthesised fragments, the solubility of the starting materials, and the generated protein in the employed media.

Two general routes were considered and applied for the synthesis of the glycoprotein Thy-1 (scheme 51). The first route involved the ligation of the fragments (19-84) and (85-110) and the following ligation of the obtained fragment (19-120) with the fragment (1-18). The second route considered the generation of the fragment (1-84) by ligation of the fragment (1-18) and (19-84) and its subsequent ligation with the fragment (85-110). To find the most convenient strategy for the assembly of the protein, both approaches were evaluated. The hydrazide at the peptides C-*terminus* was converted into the corresponding thioesters for the fragments (1-18) and (19-84) before the ligation with the peptide having a free thiol at the N-*terminus* cysteine.

Strategy 1: C- to N-terminal assembly



Strategy 2: N- to C-terminal assembly



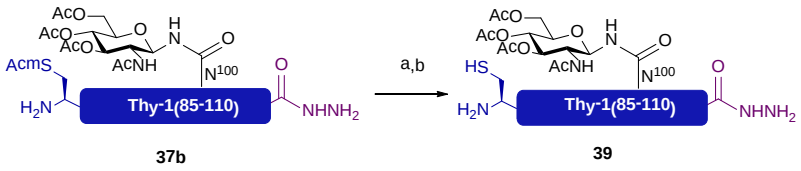
Scheme 51 - Strategies for the assembly of the Thy-1 by sequential native chemical ligation. $\mathbf{R_1}$, $\mathbf{R_2}$ = -Ac, -H; $\mathbf{R_3}$ = -Acm, -H; $\mathbf{R_4}$ = $-CH_2CH_2C(O)OCH_3$, $-4(CH_2C(O)OH)Ph$; $\mathbf{R_5}$ = $-NHNH_2$, $-4(CH_2C(O)OH)SPh$.

The type of glucosamine (protected or unprotected) defined the diverse isoforms of the glycoprotein. The selection of the reacting functionalities at the C- and N-*termini* of each fragment, hydrazide or thioester at the C-*terminus*, and protected or deprotected cysteine residues at the N-*terminus*, addressed the synthesis towards one strategy or the other in the assembly of the Thy-1.

2.6.1 Strategy 1: C- to N-terminal Assembly

Ligation of fragment 6a with 37b

The first attempt to assembly the Thy-1 glycoproteins involved the ligations of the glycopeptide **6a** having the *per*-acetylated glucosamines on the Asn23 and Asn60 and the Acm-protected glycopeptide **37b** containing the *per*-acetylated glucosamine on the Asn100. Two reactions were required to generate the thioester and free thiol for the native chemical ligation: the conversion of the hydrazide precursor **6a** into the corresponding thioester **38** and the removal of the acetoamidomethyl group in **37b** to obtain the fragment **39**.



Scheme 52 – Removal of the Acm- protecting group from **37b**. Conditions: a) **37b** (2.3 *10^{-4}mmol), $Hg(OAc)_2$ (10 equiv) in AcOH 10%, N_2, r.t., 90 min; b) 2-mercaptoethanol (20 equiv.), r.t., 1 h.

To obtain the peptide with a free cysteine at the N-terminus, the Acm-group was removed from peptide **37b** by treatment with mercury (II) acetate and following reduction of the Hg-S-Peptide intermediate with 2-mercaptoethanol. [58] The reaction was carried out in an inert atmosphere and reaction progress was monitored by LC-MS (figure 33). The reaction was characterized by the retention of mercury attached to a small portion of the peptide (peak A) spite of the treatment with β-mercaptoethanol.

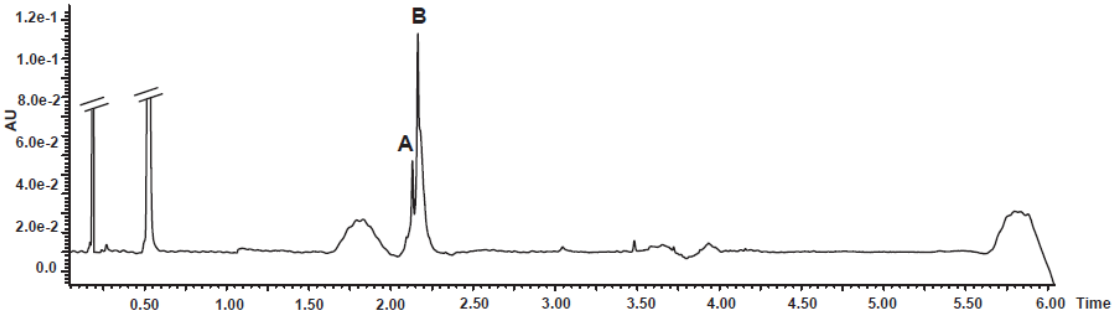
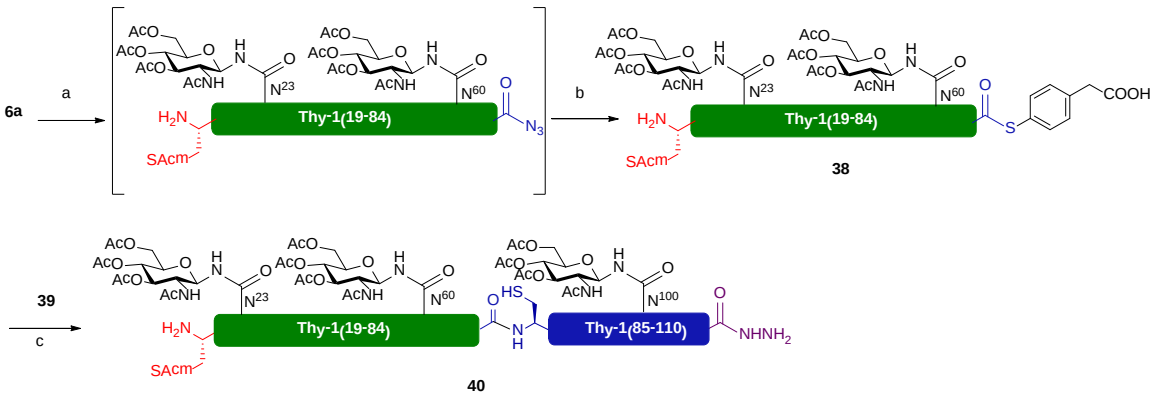


Figure 33 – LC analysis of the Acm deprotected glypeptide **39**. The chromatogram was recorded at 214 nm (RP-UPLC with C18 column, gradient: 10 to 50% of ACN in water + 0.1% of formic acid in 6 minutes). A: Product +200 Da (+ Hg); B: Product. Mass cldc (**39**, peak B) = 3165.6353 Da, Mass obs: $[M+3H]^{3+}$ = 1056.2902 Da, $[M+4H]^{4+}$ = 792.4655 Da, $[M+5H]^{5+}$ = 634.1694 Da, $[M+6H]^{6+}$ = 528.8142 Da. cldc (**Hg-S-39**, peak A) = 3363.4721 Da, Mass obs: $[M+3H]^{3+}$ = 1121.1688 Da, $[M+4H]^{4+}$ = 841.6359 Da.

The thioester **38** was obtained by converting the C–terminal hydrazide into the acyl azide intermediate at pH 3.0 – 3.2 by treatment with sodium nitrite. The reaction was performed at -10°C to avoid side reactions such as the Curtius rearrangement that leads to the formation of the isocyanates.[162] Following the formation of the azide, the addition of a large excess of the aryl thiol 4-mercaptophenylacetic acid (MPAA) and the subsequent adjustment of the pH from 3.0 to 6.8 allowed the thiolysis *in situ* and the formation of the corresponding thioester **38** (scheme 53).



Scheme 53 - Formation of the peptide thioester **38** and native chemical ligation for the generation of glycoprotein **40**. Conditions: a) dissolution of **6a** (2 mM) in Gdn*HCl 6M, Na_2HPO_4 0.2 M (pH 3.0 – 3.2), followed by reaction with $NaNO_2$ (10 equiv) at -10°C for 20 min; b) addition of MPAA (100 equiv), adjustment of temperature to 25 °C and pH to 6.7 – 6.9, 5 min c) Gdn*HCl 6M, Na_2HPO_4 0.1 M, MPAA (50 mM), TCEP (5 mM), pH 6.7 – 6.9, r.t., 24 h.

To ligate the two glycopeptides, the fragment **39** was dissolved in phosphate buffer at pH 6.8 containing MPAA (50 mM) and TCEP (5 mM) and added to the glycopeptide thioester **38**. The progress of the ligation was followed by LC-MS analysis after 0, 4, and 24 hours of reaction. After

1 day, the limiting glycopeptide **38** was completely consumed (figure 34). A considerable amount of the thioester **38** underwent hydrolysis, and co-eluted with the desired glycoprotein product **40** as shown in the LC-MS analysis. The formation of disulphide bonds between the thiol of cysteine and the MPAA reduced the availability of the cysteine thiol for the trans-thioesterification and the following rearrangement to the amide product **40**. The lack of the nucleophilic thiol on **39** reduced the level of conversion and thereby the consumption of the thioester **38** that hydrolysed with the time in the aqueous media.

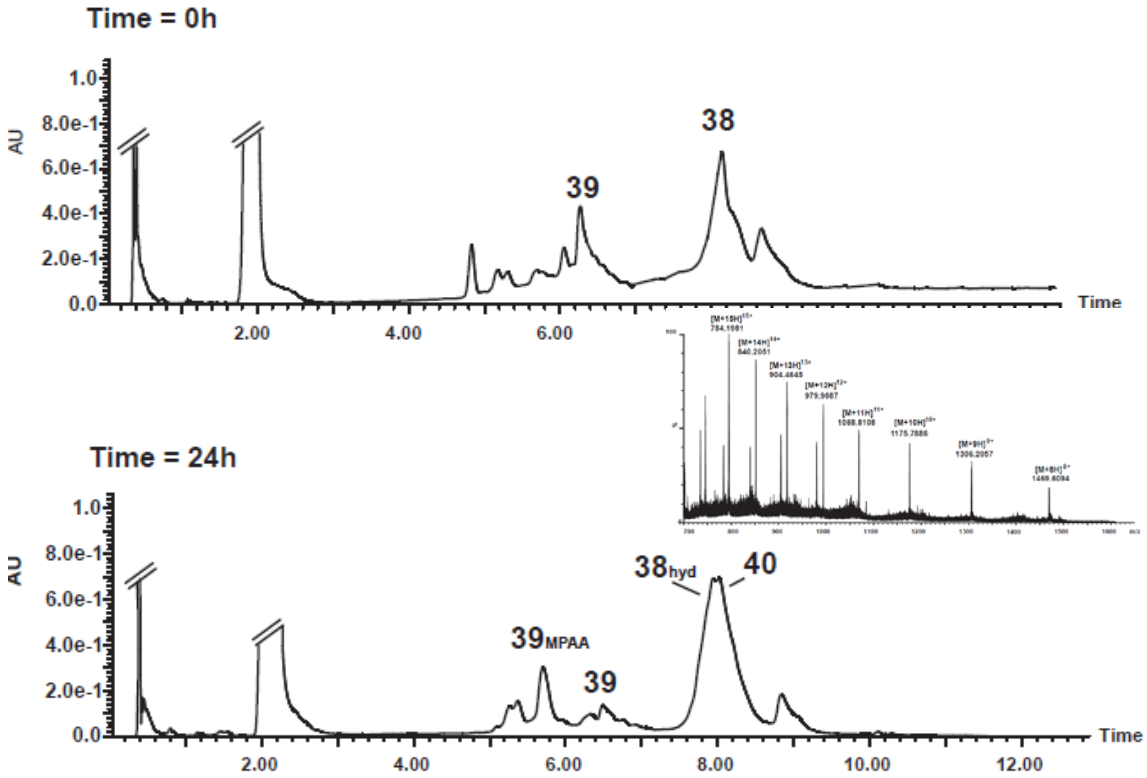


Figure 34 - Comparison of UV profile at time 0 and time 24 h. for ligation of fragments **38** and **39**. High resolution mass of the ligation product **40**. Mass cldc: 11543.6514 Da, Mass obs: $[M+8H]^{8+}$ = 1469.6094 Da, $[M+9H]^{9+}$ = 1306.2057 Da, $[M+10H]^{10+}$ = 1175.7886 Da, $[M+11H]^{11+}$ = 1068.8108 Da, $[M+12H]^{12+}$ = 979.9887 Da, $[M+13H]^{13+}$ = 9044645 Da, $[M+14H]^{14+}$ = 840.2051 Da, $[M+15H]^{15+}$ = 784.1981 Da. Mass of by-product $\mathbf{38_{hyd}}$ cldc: 8396.0325 Da. $[M_{hyd}+8H]^{8+}$ = 1050.6508 Da, $[M_{hyd}+9H]^{9+}$ = 934.1388 Da, $[M_{hyd}+10H]^{10+}$ = 840.7174 Da, $[M_{hyd}+11H]^{11+}$ = 746.4709 Da

Different eluents were tested to solubilise and purify by gel filtration the tri-glycosylated fragment **40**. LC analysis of the solubilization of the ligation product **40** in water showed that the protein has only limited solubility in this solvent and the product was present in both, the water-soluble and insoluble fractions. The insoluble fraction was separated, and its solubility was tested with different concentrations of acetonitrile (up to 40%). The LC analysis showed that product was still only partially soluble in this solvent mixture. The complete solubilisation of the glycoprotein **40**

was achieved in a buffer of 3 M guanidium hydrochloride (Gdn*HCl) in 0.2 M Na_2HPO_4, used as the eluent for the purification of the glycoprotein. The ligation mixture was purified by size exclusion chromatography (SEC) on a Superdex Peptide 10/300 GL column (GE Healthcare GmbH, Europe). However, glycoprotein **40** could not be obtained purely because of the co-elution of the hydrolysed fragment **38$_{hyd}$.** Noteworthy, the LC-MS analysis of the collected fractions after the SEC column showed that the high solubility of the small fragments **39** and **39$_{MPAA}$** in water could be used to remove these fragments from the reaction mixture before the purification (figure 35).

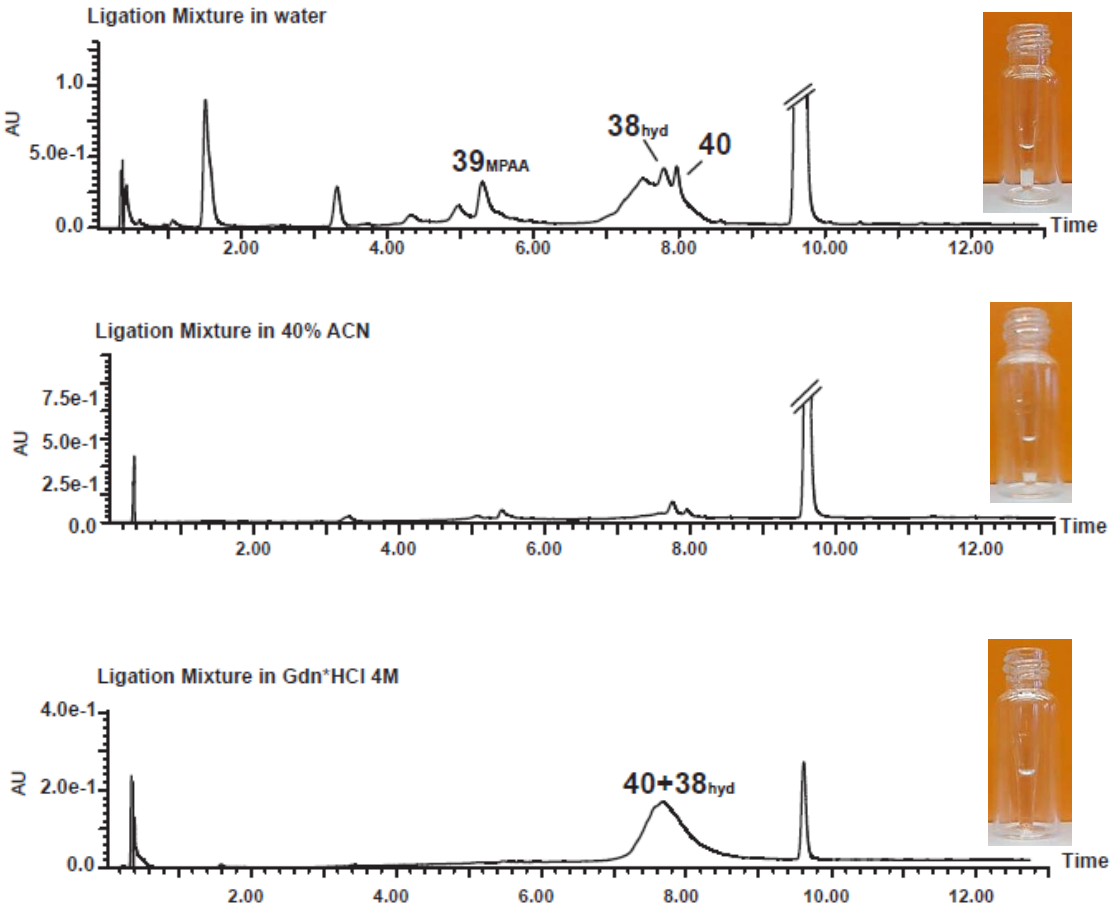


Figure 35 - Comparison of UV profile of the supernatant (left) and solubility test of small aliquots of ligation mixture (right) in water, 40% ACN in water, and Gdn*HCl 3 M.

A different strategy to purify the glycoprotein based on increasing the polarity of the glycoprotein **40** by removing the acetyl protecting groups from glucosamine and the removal of the Acm-protecting group on the cysteine. However, these reactions were not successful in the high concentration of salt required for the solubilisation of the protected glycoprotein. New strategies

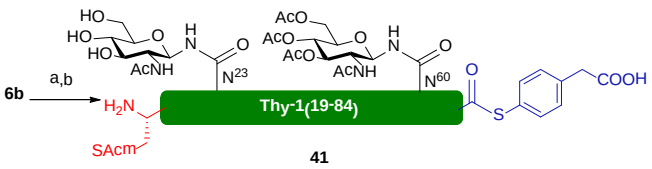
were then explored to generate more polar fragments that facilitate the purification of the ligation products.

Ligation of fragment 6b with fragment 37a

Solubility plays an important role in the synthesis and modification of peptides and proteins. The amino acids sequence can give an idea of the polar or nonpolar character of the compound and the presence of modification having polar hydrophilic groups such as hydroxyl or carboxylic groups helps the solubilisation of proteins in aqueous media (buffer) and organic solvent (acetonitrile, DMF, DMSO). The use of chaotropic solvents such as guanidine hydrochloride or urea can break hydrophobic interactions and hydrogen bonds that are the main cause of the aggregation of polypeptide chains. However, generally high concentrations of these salts are necessary.

The glycopeptide **6b** was the second designed and synthesized isoform of the fragment (19-110). The presence of free hydroxyl groups on the glycosylated Asn60 may offer a second property for the synthesis of Thy-1 by increasing the hydrophilicity of the glycoprotein intermediates. To evaluate the efficiency of this fragment in the synthesis of the glycoproteins, **6b** was converted into a thioester and ligated with the glycopeptide **37a**, which did not require the removal of Acm.

The fragment **6b** was converted into the thioester **41** by treatment with sodium nitrite and following reaction with thiols, as already described for **38**. Similar to fragment **6b,** the purification of the thioester **41** turned to be challenging and delivered a mixture of the glycopeptides having different acetylation on Asn60.



Scheme 54 – Thiolysis of fragment **6a**. Conditions: a) dissolution of **6b** (2.5mM) in Gdn*HCl 6M, Na_2HPO_4 0.2 M (pH 3.0 – 3.2), followed by reaction with $NaNO_2$ (10 equiv) at -10°C for 20 minutes; b) addition of MPAA (100 equiv), adjustment of temperature to 25 °C and pH to 6.7 – 6.9, 5 min

The LC-MS analysis showed the broad peak A containing the desired product **42** and the thioester of the glycopeptide lacking one or two acetyl groups. The figure 36 shows the high-resolution mass of the product eluting between 7.90 – 7.95 minutes.

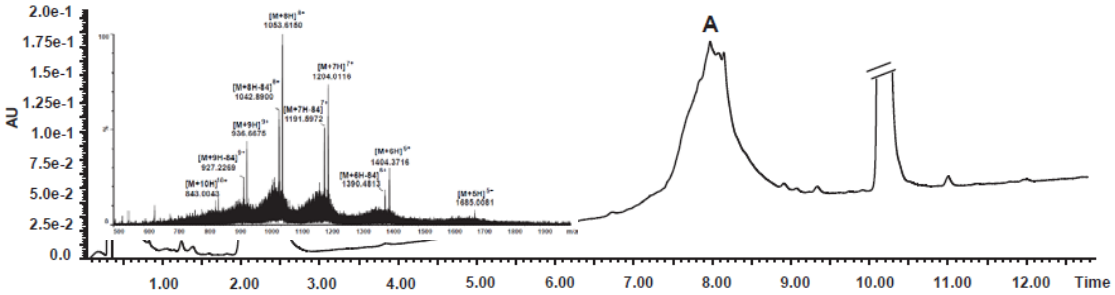
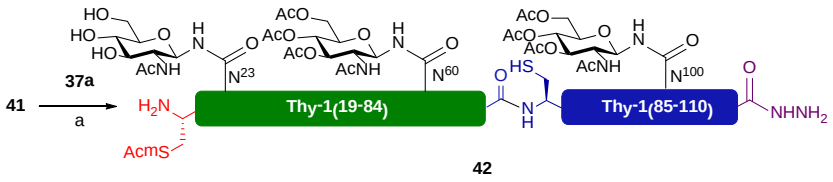


Figure 36 – LC-MS of the thioester glycopeptide **41**. Chromatogram recorded at 214 nm of mixture containing thioester **41** (RP-UPLC with C4 column, gradient: 10 to 50% of ACN in water + 0.1% of formic acid in 14 minutes). High resolution mass of the band A. Mass cldc: 8420.0103 Da, Mass obs: $[M+5H]^{5+}$ = 1685.0081 Da, $[M+6H]^{6+}$ = 1404.3716 Da, $[M+7H]^{7+}$ = 1204.0116 Da, $[M+8H]^{8+}$ = 1053.6150 Da, $[M+9H]^{9+}$ = 936.6675 Da, $[M+10H]^{10+}$ = 843.0043 Da. $[M+6H-2Ac]^{6+}$ = 1390.4813 Da, $[M+7H-2Ac]^{7+}$ = 1191.5972 Da, $[M+8H-2Ac]^{8+}$ = 1042.8900 Da, $[M+9H-2Ac]^{9+}$ = 927.2269 Da.

The ligation of **41** with the fragment **37a** was performed regardless the lower purity of the thioester and it was used to obtain a different isoform of the Thy-1 fragment (19-110). The increased polarity of the resulting glycoprotein **42** should contribute and improve the solubility of this protein in water and the following purification of the product, which was the major problem in the ligation of the glycopeptides having all glucosamines peracetylated.



Scheme 55- Assembly of fragment **42** (19-110), N^{23}(GlcNAc) $N^{60,100}$($(OAc)_3$GlcNAc). Conditions: a) Gdn*HCl 6M, Na_2HPO_4 0.1 M, MPAA (50 mM), TCEP (5 mM), pH 6.7 – 6.9, r.t., 24-72 h.

The reaction of the fragments **41** and **37a** was carried out using the conditions used to get **40** and was monitored by LC-MS at different reaction time. Unfortunately, the ligation did not proceed and no-formation of the desired product **42** was detected after 60 hours (figure 37).

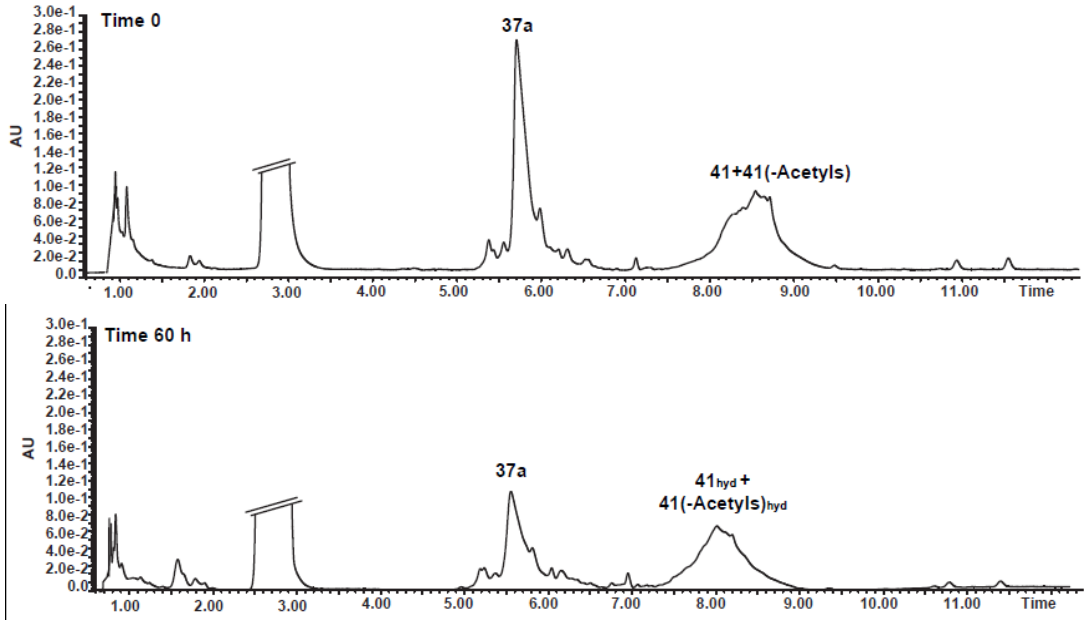
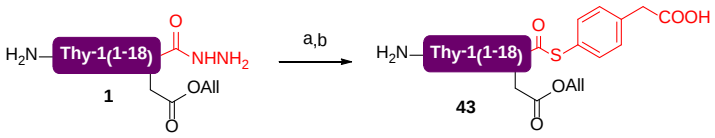


Figure 37 - Comparison of UV profile at time 0 and time 60 h. of ligation of fragments **37a** and **41.**

The main problem in this ligation was attributed to the lack of treatment with mercury acetate to remove the Acm from fragment **37** and the participation of a metal-thiol intermediate produced during the trans-thioesterification reaction. Brick *et al*. reported a mechanism of the removal of the Acm using palladium chloride and the subsequent one-pot ligation of the free-thiol fragment with a thioester peptide.[59] The authors assumed that the high affinity of the mercury for the sulphur leads to the formation of a complex S-Hg-S. However, no evidence has been found in the literature about the mechanism of the reaction.

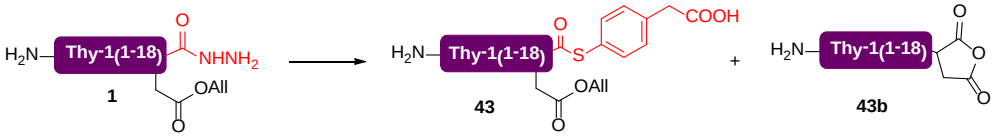
2.6.2 Strategy 2: N- to C-terminal ligations

A second strategy was considered for the assembly of Thy-1. It required different steps: i) removal of the Acm-group form fragment **6c (19-84)** and conversion of the hydrazide **1** into the correspondent thioester (**43**); ii) ligation of the free-thiol containing glycopeptide **43** with **44** to yield the glycoprotein **45**; iii) conversion of the glycoprotein **45** into the thioester **46** and removal of Acm-group from the fragment **37b**; iv) Ligation of the fragment **39** and **46** to Thy-1 (1-110) **47** (scheme 56).



Scheme 56 -Formation of the peptide thioester **43** Conditions: a) dissolution of **1** (3 mM) in Gdn*HCl 6M, Na_2HPO_4 0.2 M (pH 3.0 – 3.2), followed by reaction with $NaNO_2$ (10 equiv) at -10°C for 20 min; b) addition of MPAA (100 equiv), adjustment of temperature to 25 °C and pH to 6.7 – 6.9, 5 min.

The fragment **1** was first converted into the corresponding thioester using sodium nitrite and MPAA as described before and the reactions were monitored by LC-MS. Contrary to the expected results, the formation of succinic anhydride at the C-terminus was detected after the thiolysis (figure 38). The formation of this product was attributed to the loss of the allyl protecting group and the nucleophilic attack of the side chain of the aspartic acid to the high reactive aryl thioester (scheme 57).



Scheme 57 - Products after the conversion of the peptide hydrazide **1** into the thioester with excess of MPAA.

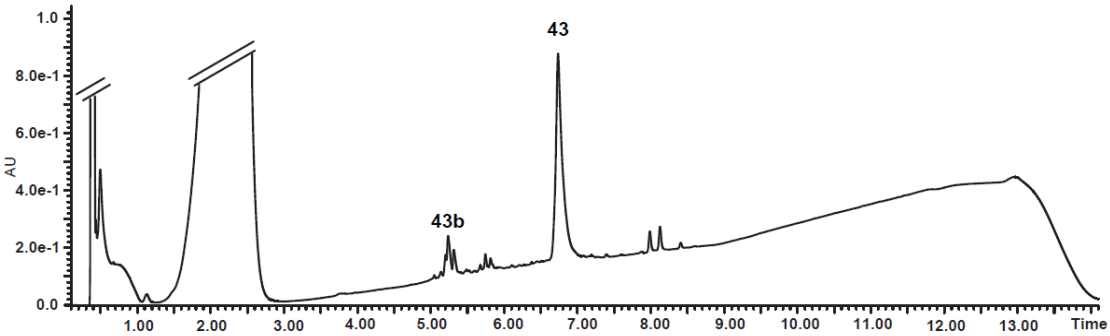
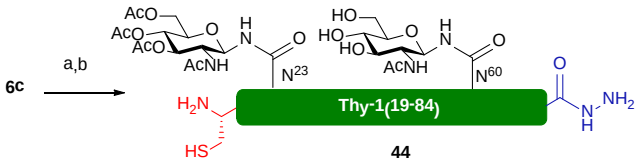


Figure 38 –LC-MS of the thioester **43**. Chromatogram recorded at 214 nm (RP-UPLC with C4 column, gradient: 10 to 50% of ACN in water + 0.1% of formic acid in 14 minutes). High resolution mass of **43**. Mass cldc: 2250.1404 Da, Mass obs: $[M+2H]^{2+}$ = 1126.0963 Da, $[M+3H]^{3+}$ = 751.0659 Da.

The pure glycopeptide **6c** having one protected glucosamine on the Asn23 and one deprotected glucosamine on the Asn60 was treated with mercury (II) acetate to remove the Acm protecting group from the N-terminal cysteine.



Scheme 58– Removal of Acm from fragment **6c**. Conditions: a) $Hg(OAc)_2$ (10 equiv) in AcOH 10%, N_2, r.t., 90 min; b) 2-mercaptoethanol (20 equiv.), r.t., 1h.

The reduction of the Hg-S complex with the excess of 2-mercaptoethanol was monitored by LC-MS (figure 39). As mentioned above, the complex S-Hg-S was generated and observed as intermediate during the reaction. This complex showed to be a reactive substrate in the ligation with the thioester and improve the efficiency of the ligation steps.

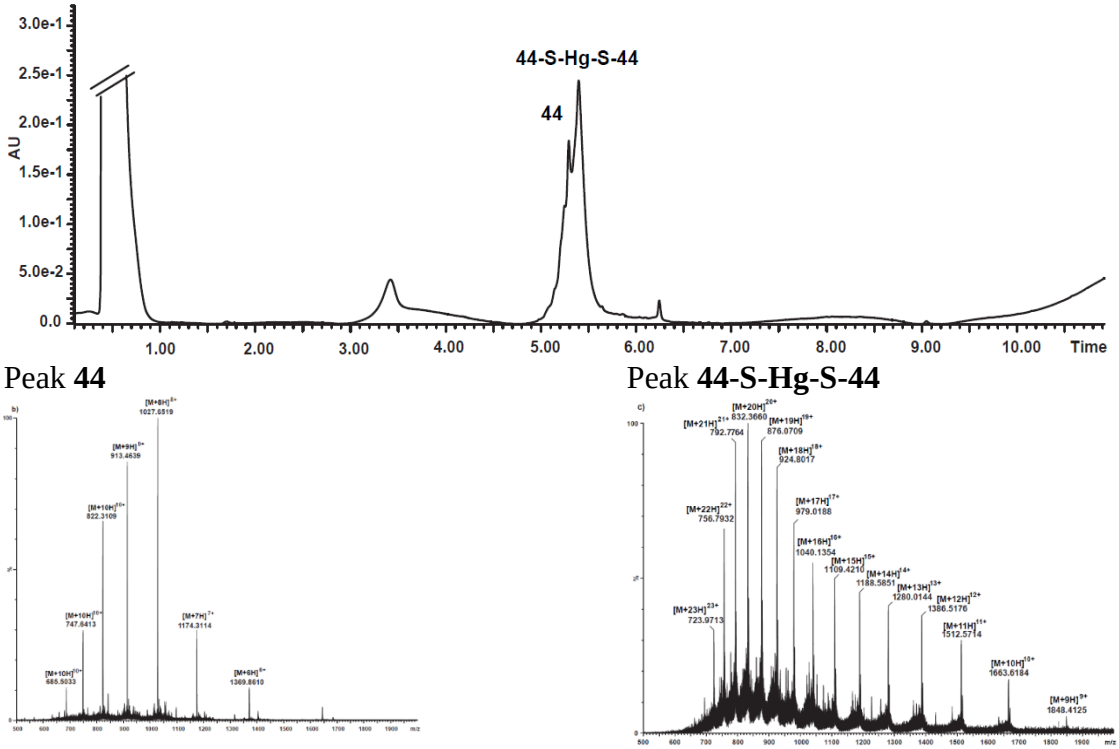
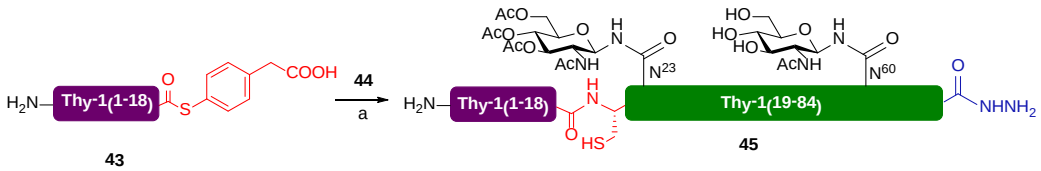


Figure 39 – LC-MS of the mixture **44** and ***44-S-Hg-S-44***. Chromatogram recorded at 214 nm (RP-UPLC with C4 column, gradient: 10 to 50% of ACN in water + 0.1% of formic acid in 14 minutes). High resolution mass of **44**. Mass cldc: 8212.9985 Da, Mass obs: $[M+6H]^{6+}$ = 1369.8610 Da, $[M+7H]^{7+}$ = 1174.3114 Da, $[M+8H]^{8+}$ = 1027.6519 Da, $[M+9H]^{9+}$ = 913.4639 Da, $[M+10H]^{10+}$ = 822.3109 Da, $[M+11H]^{11+}$ = 747.6413 Da, $[M+12H]^{12+}$ = 685.5033 Da. High resolution mass spectra of the **44-S-Hg-S-44**. Mass cldc: 16625.9376 Da, Mass obs: $[M+9H]^{9+}$ = 1848.4125 Da, $[M+10H]^{10+}$ = 1663.6184 Da, $[M+11H]^{11+}$ = 1512.5714 Da, $[M+12H]^{12+}$ = 1386.5176 Da, $[M+13H]^{13+}$ = 1280.0144 Da, $[M+14H]^{14+}$ = 1188.5851 Da, $[M+15H]^{15+}$ = 1109.4210 Da, $[M+16H]^{16+}$ = 1040.1354 Da, $[M+17H]^{17+}$ = 979.0188 Da, $[M+18H]^{18+}$ = 924.8017 Da, $[M+19H]^{19+}$ = 876.0709 Da, $[M+20H]^{20+}$ = 832.3660 Da, $[M+21H]^{21+}$ = 792.7764 Da, $[M+22H]^{22+}$ = 756.7932 Da, $[M+23H]^{23+}$ = 723.9713 Da.

The thioester **43** and the glycopeptide **44** were mixed in a buffer at pH 6.7-6.9 containing the MPAA and the TCEP as additives. The reaction mixture turned turbid after 30 minutes giving a hint of the hydrophobic character of the product **45** and the contribution of the highly hydrophobic thioester **43** that contains many non-polar amino acids (Ala, Leu, Val). The addition of DMF and DMSO into the ligation mixture did not improve the solubility of the mixture. Thus, the ligation was kept under these conditions and stirred for 48 hours under monitoring by LC-MS (figure 40).



Scheme 59- Ligation of fragments **43** and **44**. Conditions: a) Gdn*HCl 6M, Na_2HPO_4 0.1 M, MPAA (50 mM), TCEP (5 mM), pH 6.7 – 6.9, r.t., 24-48 h.

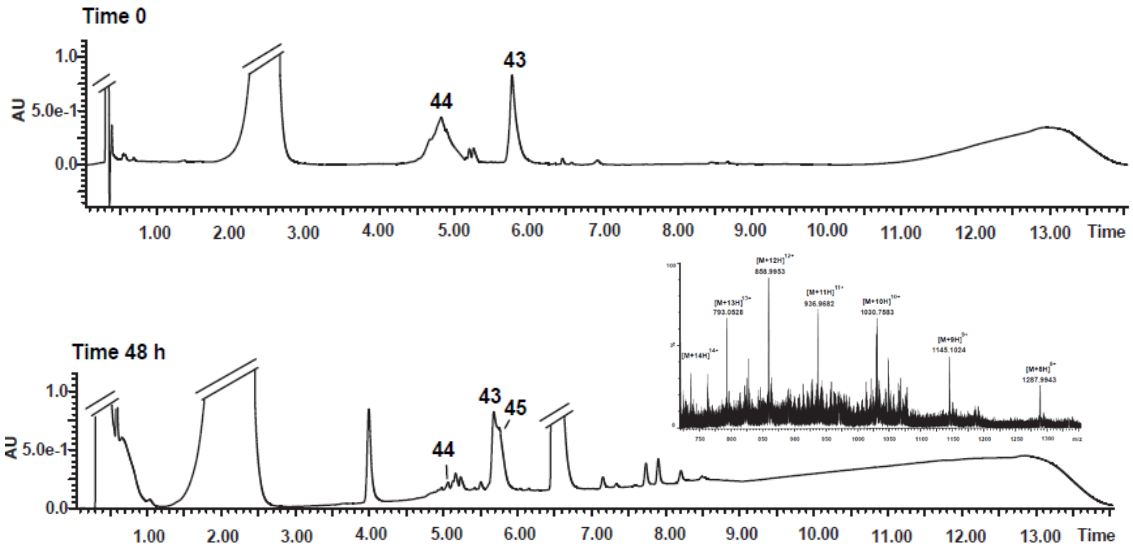


Figure 40 - Comparison of UV profile at time 0 and time 48 h. of ligation between fragments **43** and **44**; High resolution mass of **45**. Mass cldc: 10295.1007 Da, Mass obs: $[M+8H]^{8+}$ = 1287.9943 Da, $[M+9H]^{9+}$ = 1145.1024 Da, $[M+10H]^{10+}$ = 1030.7583 Da, $[M+11H]^{11+}$ = 936.9682 Da, $[M+12H]^{12+}$ = 858.9953 Da, $[M+13H]^{13+}$ = 793.0528 Da, $[M+14H]^{14+}$ = 736.4044 Da.

The desired glycoprotein **45** was separated from the unreacted peptide **43** and the side products by gel filtration and concentrated by ultracentrifugation in the buffer at pH 6.7-6.9. The pH of the solution containing the glycoprotein **45** was adjusted to pH to 3.0-3.2 with 1N HCl and the C-terminal hydrazide was converted into the thioester **46** following the described treatment with sodium nitrite and the MPAA. The thiolysis reaction presented some problems: the mixture turned turbid after the addition of the thiol. Moreover, the obtained thioester **46** was difficult to detect at the LC-MS due to its low solubility in the buffer (figure 41).

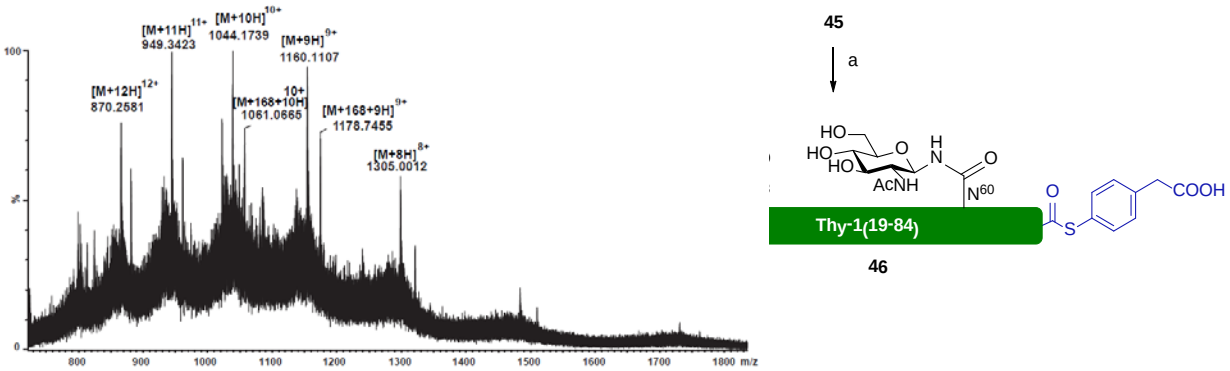
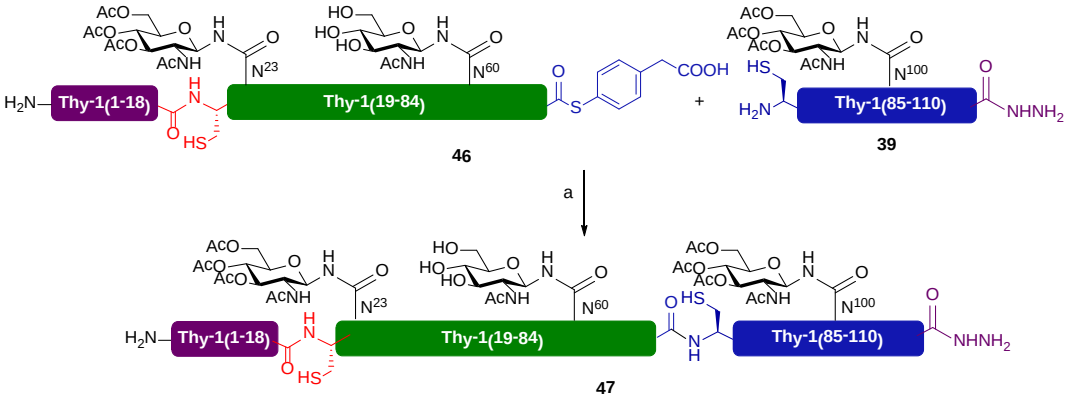


Figure 41 - High resolution mass of **46**. Mass cldc: 10431.0955 Da, Mass obs: $[M+8H]^{8+}$ = 1305.0012 Da, $[M+9H]^{9+}$ = 1160.1107 Da, $[M+10H]^{10+}$ = 1044.1739 Da, $[M+11H]^{11+}$ = 949.3423 Da, $[M+12H]^{12+}$ = 870.2581 Da. $[M+168+9H]^{9+}$ = 1178.7455 Da, $[M+168+10H]^{10+}$ = 1061.0665 Da, $[M+168+11H]^{11+}$ = 964.7102 Da, $[M+12H]^{12+}$ = 884.3952 Da.

The fragment **39** was dissolved in the ligation buffer at pH 6.7-6.9 containing MPAA (0.01 M) and TCEP (5 mM) and added to the solution with the thioester **46** (scheme 60). The ligation of the two fragments did not proceed as monitored by the LC-MS. After 48 hours, the reaction was warmed up to 35°C to improve the ligation of the fragments but the increased temperature did not have an effect reaction. The free thiol on the peptide **39** oxidized and formed a disulfide with the MPAA.



Scheme 60 – Ligation of fragments **46** and **39**. Conditions: a) Gdn*HCl 6M, Na_2HPO_4 0.1 M, MPAA (50 mM), TCEP (5 mM), pH 6.7 – 6.9, r.t., 24-48 h

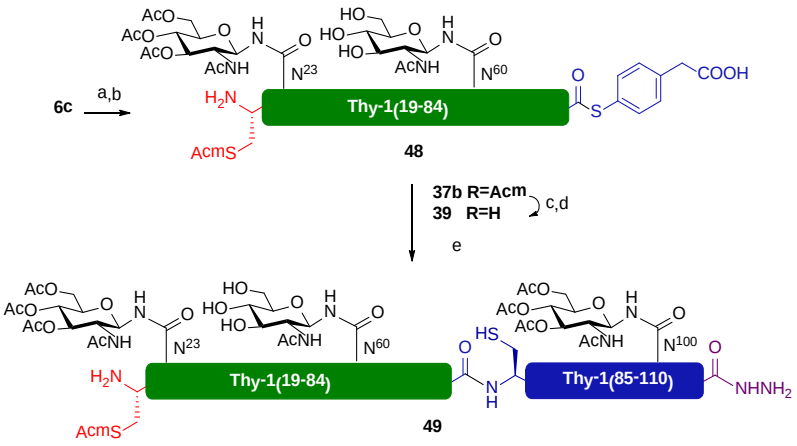
The lack of reaction or the difficult detection of any formed product was attributed to the poor solubility of the two fragments (**43** and **46**) and of product **45** in the ligation buffer. Additionally, the hydrophobic nature of fragment **1** and the aryl thiol used to obtain the thioester fragment **43**

decreased the solubility of the fragment in both aqueous and organic solvents. Thus, more soluble thiols may be necessary to enhance the applicability of this fragment.

2.6.3 Strategy 3: Thy-1 by ligation of fragment 6c, 37b and 1.

The polarity of the peptides, the reagents for the deprotection and the reaction conditions (solvent, pH, salt, etc) affected the ligation and the monitoring of the reactions with the peptide fragments. All these parameters were considered to test of a new strategy for the assembly of the glycoprotein using the available fragments.

The first step of this strategy was the ligation of the thioester **48** with fragment **39** under the same conditions described before. This ligation was carried out in the ligation buffer at pH 6.7-6.9 and monitored by LC-MS. After 16 h, the starting material (fragment **49**) was completely consumed indicating the completion of the reaction (figure 42).



Scheme 61 – Ligation of fragments **6c** and **37b** - Conditions: a) **6c**, Gdn*HCl 6M/Na_2HPO_4 0.2 M (pH 3.0), $NaNO_2$, -10°C, 20 min; b) MPAA, 25 °C, 6.7, 5 min; c) $Hg(OAc)_2$, AcOH 10%, N_2, r.t., 90 min; d) 2-mercaptoethanol, r.t., 1 h; e) Gdn*HCl 6M, Na_2HPO_4 0.1 M, MPAA (50 mM), TCEP (7 mM), pH 6.7, r.t., 48 h

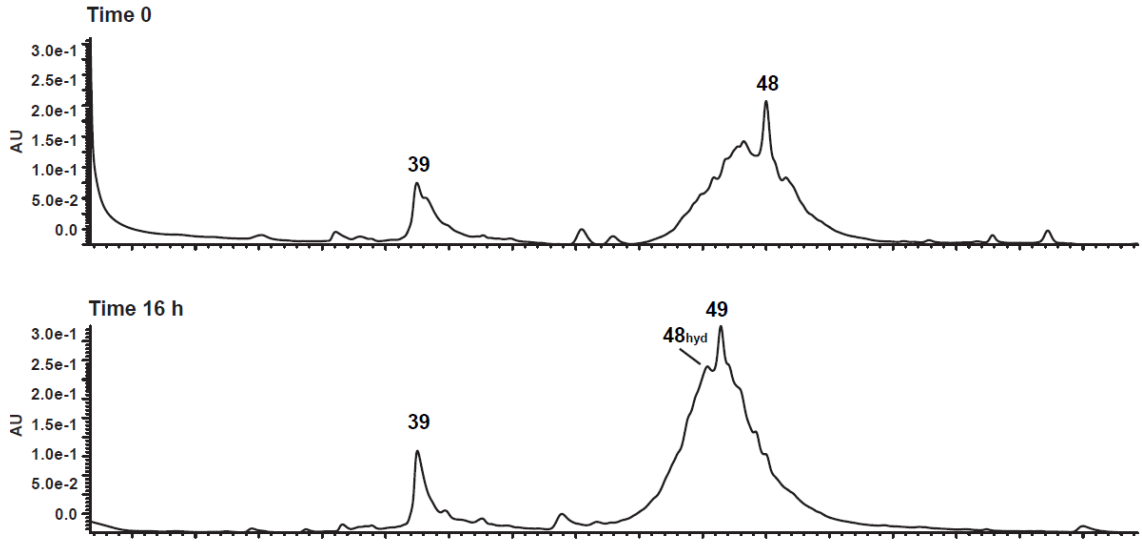
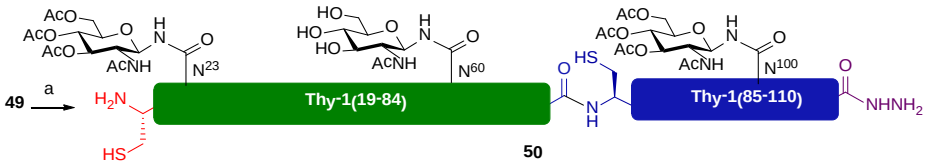


Figure 42 - Comparison of UV profile at time 0 and time 48 h. of ligation between fragments **39** and **48**

The produced glycoprotein **49** was purified by gel filtration and concentrated by ultracentrifugation. The Acm group was removed following the protocol from Brick *et al.*[59] The glycoprotein **49** was dissolved in a phosphate buffer containing 4 M guanidinium chloride and was treated with 15 equiv. of palladium chloride at 37°C in the same buffer for 2 hours and a large excess of DTT was added to stop the reaction.



Scheme 62 - Removal of Acm from **49**. Conditions: a) $PdCl_2$, Gdn*HCl 4 M, Na_2HPO_4 0.2 M, DTT (0.1M), in water.

LC-MS showed complete reaction and the formation of the glycoprotein **50** as the main product (figure 43). The reaction mixture was centrifuged to obtain the N-terminal deprotected glycoprotein **50**.

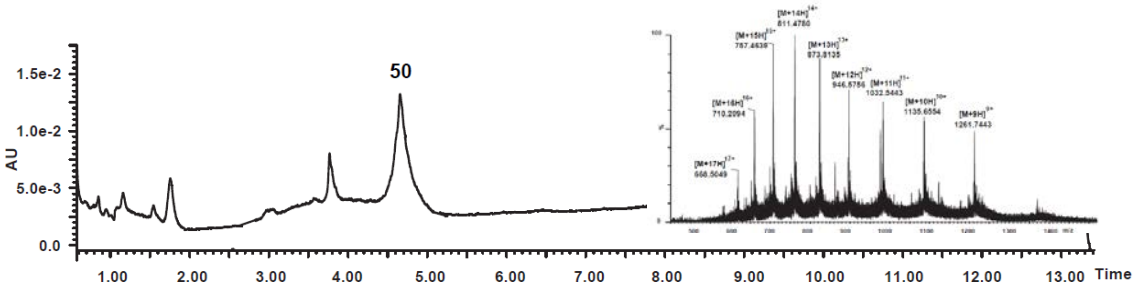
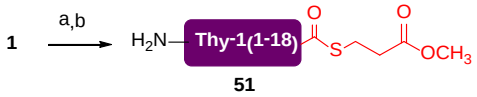


Figure 43 – LC-MS analysis of the glycoprotein **50**. Chromatogram recorded at 214 nm (RP-UPLC with C4 column, gradient: 10 to 50% of ACN in water + 0.1% of formic acid in 14 minutes). High resolution mass of **50**. Mass cldc: 11346.5724 Da, Mass obs: $[M+9H]^{9+}$ = 1261.7443 Da, $[M+10H]^{10+}$ = 1135.6554 Da, $[M+11H]^{11+}$ = 1032.5443 Da, $[M+12H]^{12+}$ = 946.5756 Da, $[M+13H]^{13+}$ = 873.8135 Da, $[M+14H]^{14+}$ = 811.4780 Da, $[M+15H]^{15+}$ = 757.4639 Da, $[M+16H]^{16+}$ = 710.2094 Da, $[M+17H]^{17+}$ = 668.5049 Da.

A new thioester of the fragment **1** was obtained under the described protocol by treating the peptide hydrazide with sodium nitrite and the addition of methyl-3-mercaptopropionate MMP as a thiol (figure 44). The obtained thioester **51** was immediately purified by RP-HPLC using a gradient of acetonitrile and water with 0.1% of formic acid from 5 to 70% in 35 min (77%, 4.1 mg).



Scheme 63 - Thiolysis of fragment **1**. Conditiond: a) Gdn*HCl 6M, Na_2HPO_4 0.2 M (pH 3.0 – 3.2), $NaNO_2$, -10°C 20 minutes; b) MMP, 25 °C, pH to 6.7 – 6.9, 5 min.

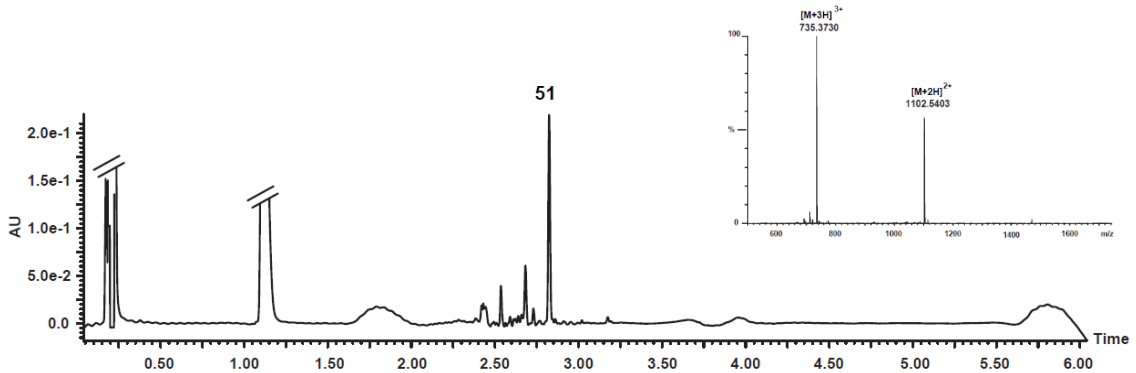
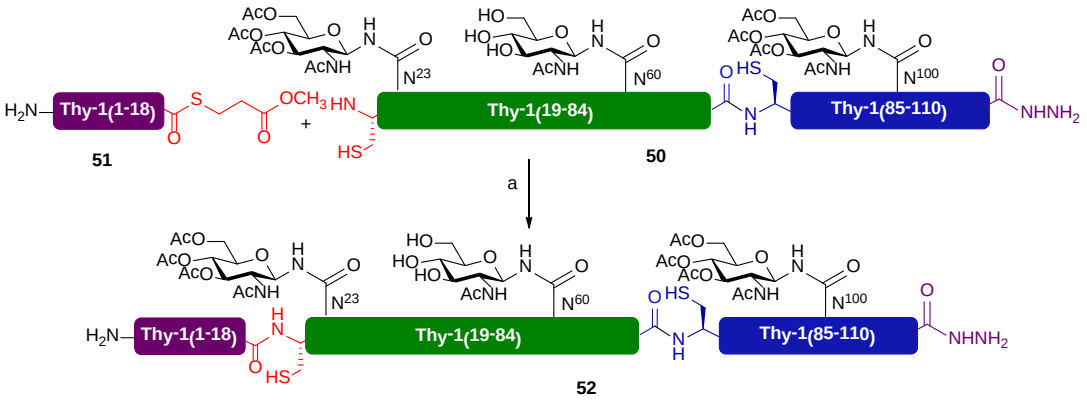


Figure 44 – LC-MS of the thioester **51**. Chromatogram recorded at 214 nm (RP-UPLC with C18 column, gradient: 5 to 70% of ACN in water + 0.1% of formic acid in 6 minutes). High resolution mass of **51**. Mass cldc: 2202.1404 Da, Mass obs: $[M+2H]^{2+}$ = 1102.5403 Da, $[M+3H]^{3+}$ = 735.3730 Da.

The thioester **51** and thiol **50** were ligated in buffer phosphate at pH 6.7-6.9 and the progress was monitored by LC-MS. After 4 h, the target glycoprotein **52** was detected (figure 45).



Scheme 64 – Ligation of thioester **51** and thiol **50**. Conditions: a) Gdn*HCl 6M, Na_2HPO_4 0.1 M, MPAA (50 mM), TCEP (7 mM), pH 6.7 – 6.9, r.t., 24-48 h

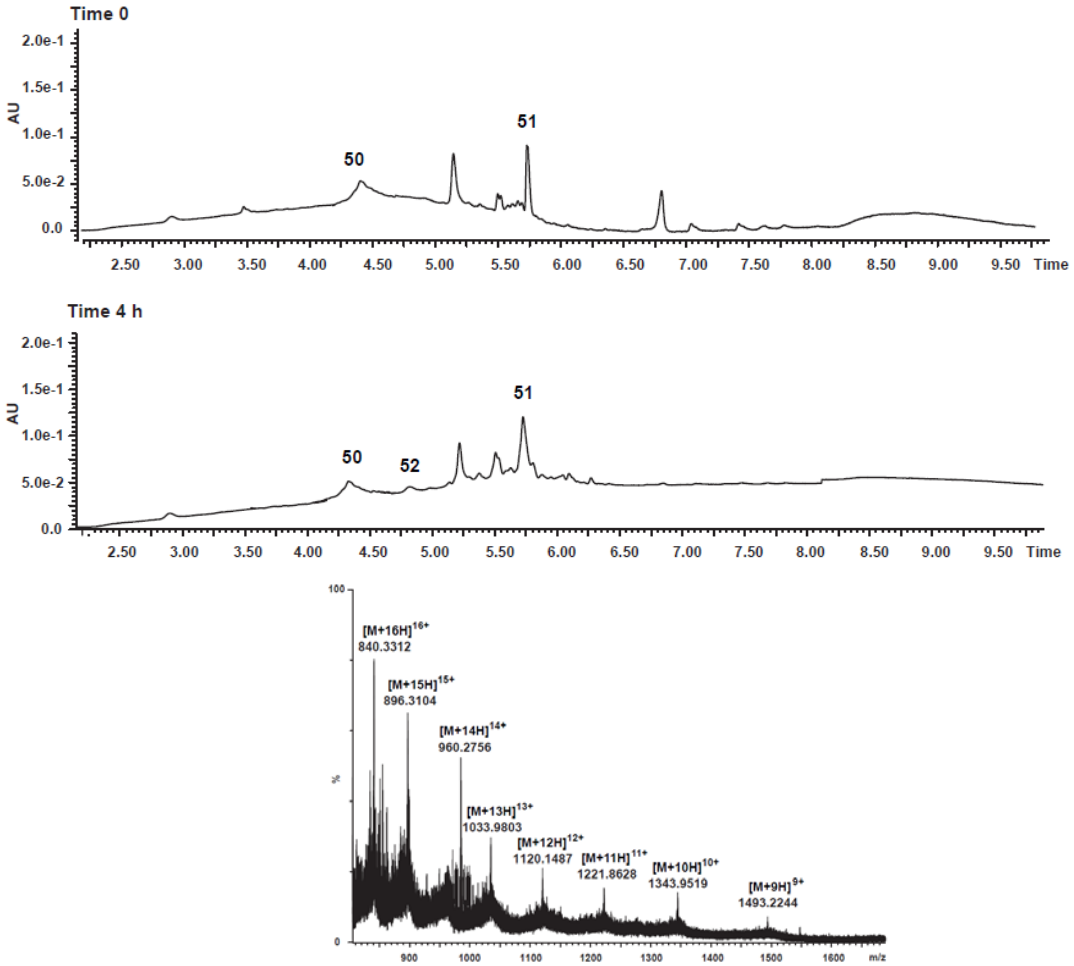


Figure 45 – LC-MS of the NCL of fragments **50** and **51**. Comparison of UV profile at time 0 and time 48 h. recorded at 214 nm. High resolution mass of the glycoprotein **52**. Mass cldc: 13428.70 Da, Mass obs: $[M+9H]^{9+}$=1493.2244 Da, $[M+10H]^{10+}$= 1343.9519 Da, $[M+11H]^{11+}$= 1221.8628 Da, $[M+12H]^{12+}$= 1120.1487 Da, $[M+13H]^{13+}$= 1033.9803 Da, $[M+14H]^{14+}$= 960.2756 Da, $[M+15H]^{15+}$= 896.3104 Da, $[M+16H]^{16+}$= 840.3312 Da.

Small aliquots of the ligation mixture were analysed by LC-MS to monitor the progress of the ligation. The detection resulted difficult due to the formation of a pale-yellow precipitate during the reaction that required special conditions (ligation buffer at pH 6.7-6.9, Gdn*HCl, Na_2HPO4 0.1 M, MPAA, TCEP) for its solubilization. The suspension was centrifuged and the precipitate and supernatant were analysed. LC-MS confirmed that the precipitate corresponded to the desired glycoprotein **52**. To determine the best conditions to solubilize the product, a portion of the ligation mixture was concentrated by ultracentrifugation. Different solvents were tested for the solubilization of the concentrated glycoprotein: 10% DMSO in water, 0.2 M Na_2HPO_4 containing 5% DMSO and 6 M Gdn*HCl, 50% acetonitrile in water, and acetonitrile. Unfortunately, none of these solvents could solubilize the product for further purifications and reactions with the glycoprotein.

The polarity of the peptide and glycopeptide fragments played a crucial role in the assembly of the Thy-1 glycoprotein. An increase in the polarity of the glycopeptide fragments assured the solubilisation of the reactants and facilitated the proceeding of the ligations and purification steps. The limitations to introduce hydrophilic functionalities into the protein fragments without modifying the amino acid side chains or the sequence of the protein left as the best alternative the use of modifications on the glycans. The glycopeptides for the assembly of Thy-1 were first synthesized having protected glycans that increased the hydrophobicity of the fragments. Thus, the removal of this protection and the free hydroxyl groups on a glucosamine residue was used to increase the polarity and the solubility of the glycopeptides. The polarity effect was also observed by comparing the ligation products **49** (obtained from ligation of **37b** with **6c**) and product **45** (obtained from the reaction of fragments **1** and **6b**). The glycoprotein fragment **49** (19-110) containing nine amino acids more than the fragment (1-84) showed higher polarity than **45** and better solubilization of the product, which was an important factor to allow further steps in the assembly of the glycoprotein. Although better results in the ligations were obtained having a deprotected glucosamine on N23 of **45**, this change was not enough to guarantee the complete dissolution of the glycoprotein and hampered the subsequent reaction of thioester **46** with fragment **37b**.

A second important phenomenon was observed during the ligation of isoforms **6a-c** with fragment **37a-b**. The treatment of fragment **37** with mercury (II) acetate and the following reduction with mercaptoethanol were not only necessary to remove the Acm-group, they also affected the efficiency of the reaction with the correspondent thioesters of fragments **6**. Although no further data were collected to confirm this hypothesis, the better behaviour in ligations of fragments **39**,

obtained by removal Acm-group of **37b**, compared with fragment **37a** suggests the formation of reactive species during the treatment with mercury that is beneficial for the trans-thioesterification reaction. Further studies and the application of other methods to remove Acm, such as the use of palladium and other metal complexes, may contribute to understand the involvement of sulphur complexes in the kinetic and outcome of the ligations and would open new directions in the chemical synthesis of glycoproteins.

The reactivity of the thioesters was an additional factor affecting the efficiency of the ligations. Highly reactive thioesters from aromatic thiols are less soluble and prompted to hydrolyse. Thus, the aliphatic thiol MMP was used for the thiolysis of the hydrazide precursor **1**. Despite the hydrophobic nature of the fragment **1,** as observed during the N- to C- strategy, a proper balance of reactivity/solubility properties of the generated peptide thioester **51** not only favored the isolation of the product, it also contributed to the successful assembly of the glycoprotein Thy-1 (1-110) having three glycosylations.

3 Conclusion and outlook

This project aimed at the total chemical synthesis of the N-linked glycoprotein Thy-1 or CD90. To investigate the role of this heavily glycosylated protein, it was chosen to synthesise *de novo* the mature form of the Thy-1.

Different strategies were designed for the assembly of the primary sequence of the glycoprotein Thy-1, composed of 110 amino acids and containing three glycosylation sites on the asparagine residues at positions 23, 60, and 100. The cysteine residues at position 19 and 85 were selected to build up the Thy-1 glycoprotein using sequential native chemical ligation to form the amide bond at these two junction points. Three fragments were designed for the assembly of the glycoprotein: i) the peptide fragment (1-18); ii) a glycopeptide fragment (19-84) having two glycosylated asparagine residues; iii) a glycopeptide fragment (85-110) containing a glycosylation site at the asparagine. Different processes were investigated and optimized for the solid-phase synthesis of the three peptide and glycopeptide fragments including the type of resin, coupling of the first amino acid, and elongation in the microwave-assisted peptide synthesizer. The synthesis of three isoforms of glycopeptide 19-84 having identical or differentiated glycosylation was optimized using two approaches: the incorporation of a glycosylated asparagine building block and the selective removal of a photo-cleavable group at aspartic acid at a specific position in the sequence followed by Lansbury aspartylation. Although the remarkable progresses obtained, the generation of the three isoforms of the glycopeptide 19-64 presented several problems, being the loss of one or more acetyl groups protecting hydroxyls on the glucosamine critical for the purification of the products and yield of the synthesis process.

Two strategies were designed and tested for the assembly of the Thy-1 glycoprotein. The C- to N-ligation approach involved the production of Thy-1 19-84 peptide fragment thioester from a peptide hydrazide and the ligation with the fragment 85-110. Two factors affected the low efficacy in forming the desired glycoprotein, the reactivity, and solubility of the reactants and the corresponding products. The solubility was affected by the natural properties of the amino acid in the protein sequence and the low polarity of the glycopeptide intermediates and the thiols used to form the thioester. The removal of the Acm group from the N-terminal cysteines affected the reactivity of the fragment in the ligation. However, major effects and limitations in ligation came from the polarity of the peptide and glycopeptide fragments and their solubility in the ligation buffers. To investigate the effect of the ligation sequence to overcome the solubility limitation, the N- to C-strategy was explored and used to obtain the Thy-1 (1-84) glycoprotein The limited

solubility of the thioester of this glycoprotein confirmed the need of introducing changes into the process to enhance the polarity of the fragments. The removal of the acetyl protecting groups from one glycan in the (19-84) fragment was enough to enhance the solubility of the products and the successful synthesis of a glycoform of Thy-1 (1-120) having three N-glycosylations, two of them having a protected N-acetylglucosamine, that may grant differentiation in one glycosylation site during enzymatic trans-glycosylations.

The synthetic strategy described here can deliver Thy-1 as glycoprotein and deliver material to investigate the next steps in the assembly of this complex molecule. Following steps include the folding of the protein and the anchor of oligosaccharides to the glucosamine residues to generate isoforms of the glycoprotein Thy-1. Mutant ENGases have been reported in the state-of-the-art and could be used to decorate the Thy-1 with different oligosaccharides. The reactivity of different enzymes towards the Thy-1 or single glycopeptide fragments as substrates should be explored for the anchoring of hybrid, complex, and oligo-mannose glycans. As an alternative, synthetic large oligosaccharides have been generated and could be anchored to the glucosamine residues on the target protein. The addition of glycans could enhance the solubility of the Thy-1 glycoprotein or simple glycopeptide fragments and improve the purification steps of the desired products. After the assembly of the full glycoprotein, the next steps in the process may include the formation of a glycoprotein thioester and its ligation with a GPI molecule. This additional reaction, together with the investigation of the conditions for the folding of the glycoprotein, could give the opportunity to understand the structure-activity relationship of a Thy-1 and it could give a complete protocol that could be extended to the synthesis of many other large glycoproteins.

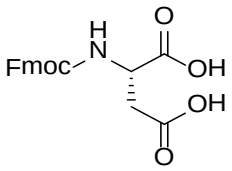
4 Experimental Part

4.1 General methods

Reagents were purchased from different suppliers in analytical grade or higher. N,N-dimethylformamide, dichloromethane, acetonitrile, formic acid, piperidine and trifluoroacetic acid were peptide synthesis grade. All Fmoc-L-amino acids, HOBt, PyBOP, HBTU, Cl-HOBt, and DIC were purchased from Iris Biotech GmbH (Germany). NovaPEG Wang resin (0.63 mmol/g) was purchased from Novabiochem Merck (Germany); Cl-TCP(Cl) ProTide resin (0.48 mmol/g) was purchased from CEM GmbH (Germany). Thin layer chromatography (TLC) was performed on coated silica gel glass plates (Merck, 60 F_{254} 0.25 mm). The compounds were detected by UV light at 254 nm or by heating after stained the plates with different solutions: basic potassium permanganate, cerium sulfate-ammonium molybdate (CAM), or a 3-methoxyphenol-sulfuric acid solution (Sugar stain). Flash column chromatography was carried out using air forced flow of the indicated eluent on silica gel high purity grade columns (60 Å, 230-400 mesh particle size, Sigma Aldrich). Solvents were removed under reduced pressure using rotary evaporator and high vacuum (<1 mbar). Freeze drying of the aqueous solutions was performed using Alpha 2-4 LD lyophilizer (Christ, Osterode am Harz, Germany). ^{1}H and ^{13}C NMR were recorded on a Varian 400 (400 MHz) and a Bruker Ascend 400 (400 MHz) spectrometers using $CDCl_3$ (7.26 ppm ^{1}H, 77.1 ppm ^{13}C) and DMSO-D_6 (2.5 ppm ^{1}H, 39.5 ppm ^{13}C) as solvents. The coupling constants (*J*) are reported in Hertz (Hz). Patterns of ^{1}H NMR signals are indicated as *s* for singlet, *d* for doublet, *t* for triplet, *q* for quadruplet, *m* for multiplet, *br* for broad singlet, *dd* for doublet of doublet, *dt* for doublet of triplet. Analytical LC-MS analysis was performed on a Agilent 1100 system using the column YMC hydrosphere C18 (50 x 3.0 mmI.D. S-3µm 12mm) and on a Waters Acquity H-class UPLC coupled to a Xevo G2-XS Q-TOF mass spectrometer using the column Acquity UPLC BEH C18 (1.7 µm 2.1 x 50 mm). Preparative RP-HPLC was performed by Knauer Semi-preparative HPLC system using a Synergi 4µm Fusion RP 80 Å (250 x 21.2 mm) column. The solvent system was: eluent **A**: 0.1 % of formic acid in water, eluent **B**: 0.1 % of formic acid in acetonitrile. Centrifugation was carried using a VWR MicroStar (Kinetic Energy 26 Joules Galaxy Mini Centrifuge) and VWR Micro Star 17R. UV absorbance measurements were recorded on the Shimadzu 1900i UV-Vis spectrophotometer using 1 cm quartz cuvettes. The propylene reactor vessels were shaken on the VWR Heidolph Vibramax 100 system.

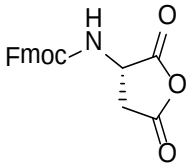
4.2 Synthesis of building blocks

N-Fmoc-L-Aspartic acid (**17**)



A suspension of N-Fmoc-L-aspartic acid beta-tert-butyl ester (Fmoc-Asp-(OtBu)-OH (3 g, 7.29 mmol) in DCM (10 ml) was cooled to 0°C and TFA (11.16 ml, 145.82 mmol) was added dropwise. Then, the mixture was warmed to r.t. and stirred for 3 h. Toluene (100 ml) was added to the reaction and the solvent mixture was removed under vacuum. The aspartic acid **17** was obtained as a white powder without purification (2.58 g, 7.26 mmol, quantitative). ^{1}H NMR (400 MHz, DMSO-D_6) δ(ppm): 7.97 (d, *J* = 7.5 Hz, 2H, Ar-H,Fmoc), 7.85 (d, *J* = 7.5 Hz, 2H, Ar-H, Fmoc), 7.57 (dd, *J* = 7.0 Hz, 7,0 Hz, 2H, Ar-H, Fmoc), 7.49 (dd *J* = 7.5 Hz, 2H, Ar-H, Fmoc), 7.39 (d, *J* = 7.37, 1H, N-H), 4.51 (d, J = 3.89 Hz, 1H CH_{2a}, Fmoc), 4.49 (d, J = 3.19 Hz, 1H, CH_{2b}, Fmoc), 4.42 (t, J = 7.1 Hz, 1H, CH, Fmoc), 3.10 – 2.92 (m, 2H, CH_2, Asp). ^{13}C NMR (101 MHz, DMSO-D_6) δ (ppm): 172.81 ($\underline{C}_q$, $\underline{C}O_2H_\beta$), 171.77 ($\underline{C}_q$, $\underline{C}O_2H_\alpha$), 155.93 ($\underline{C}$=O, Fmoc), [143.84, 140.78, 127.72, 127.17, 125.39 and 120.20 (C-Ar, Fmoc)], 65.78 (CH_2 Fmoc), 50.54 ($C_\alpha H$), 46.64 (CH Fmoc), 36.03(CH_2 Asp). MS (*m/z*): $[M+H]^+_{cald}$ = 356.11. $[M+H]^+_{obsd}$ = 356.1.

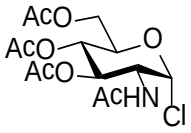
N-Fmoc-L-aspartic anhydride (**12**)



N-Fmoc-L-Aspartic acid **17** (2.58 g, 7.26 mmol) was dissolved in acetic anhydride (20.59 ml, 217.82 mmol) and the mixture was refluxed for 2 h at 95 °C. The mixture reaction was cooled to r.t. and the formed precipitate was filtrated, washed with ethyl ether and dried to give the N-Fmoc-aspartic anhydride **12** as a white powder (1.61 g, 4.77 mmol, 66%). ^{1}H NMR (400 MHz, DMSO-D_6): δ (ppm) 8.18 (d, *J* = 7.7 Hz, 1H), 7.87 (d, *J* = 7.6 Hz, 2H), 7.64 (d, *J* = 7.5 Hz, 2H), 7.39 (t, *J* = 7.5 Hz, 2H, Ar-H, Fmoc), 7.31 (t, *J* = 7.0 Hz, 2H, Ar-H, Fmoc), 4.62 (dq, J_1 = 10.0 Hz, J_2 = 6.2 Hz, 1H, H_α Asp), 4.39 (dd, $J_{2a,2b}$ = 10.55 Hz, 1H, CH_{2a} Fmoc), 4.35 (dd, $J_{2a,2b}$ = 10.49 Hz, 1H,

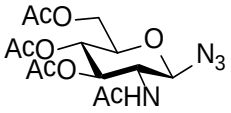
CH_{2b} Fmoc), 4.22 (t, *J* = 6.5 Hz, 1H, CH Fmoc), 3.21 (dd, $J_{1,2}$ = 18.5, 1H, $CH_{2(1)}$ Asp), 2.83 (dd, *J* = 18.5, 1H, $CH_{2(2)}$ Asp). ^{13}C NMR (101 MHz, DMSO-D_6): δ(ppm) 172.65 ($\underline{C}_q$, $\underline{C}O_2H_\beta$), 170.38 ($\underline{C}_q$, $\underline{C}O_2H_\alpha$), 156.32 ($\underline{C}$=O, Fmoc), [144.09, 144.05, 141.25, 128.17, 127.60, 127.56, 125.57, 125.49, 120.68, 120.65 (C-Ar, Fmoc)], 66.53 (CH_2 Fmoc), 50.82 ($C_\alpha H$), 47.00 (CH Fmoc), 35.17 (CH_2 Asp). MS (*m/z*): $[M+Na]^+_{cald}$ = 360.08. $[M+Na]^+_{obsd}$ = 360.2.

2-Acetamido-3,4,6-tri-O-acetyl-2-deoxy-α-D-glucopyranosyl chloride (**13**)



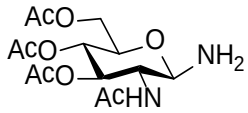
D-(+)-Glucosamine hydrochloride (10 g, 45.21 mmol) was suspended in acetyl chloride (50 ml, 45.21 mmol) and 37% aqueous hydrochloric acid (14 µl, 45.21 mmol) was added dropwise at 0°C. The mixture was warmed to r.t. and stirred for 16 h. The concentrated residue was dissolved in DCM (200 ml) and poured into ice-water solution. The organic phase was extracted, washed with saturated sodium bicarbonate (until pH 8.0) and with water (100 ml), dried over sodium sulfate and concentrated. The product was purified by silica gel flash column chromatography to give the chloride **13** as a withe solid (3.55 g, 9.71 mmol, 22%). R_f = 0.4 (DCM/MeOH 95:5). ^{1}H NMR (400 MHz, $CDCl_3$): δ(ppm) 6.18 (d, *J* = 3.7 Hz, 1H, H_1), 5.83 (d, *J* = 8.7 Hz, 1H, N-H), 5.32 (dd, *J* = 9.57, 9.55 Hz 1H, H_3), 5.21 (dd, *J* = 9.92, 9.92 Hz, 1H, H_4), 4.53 (ddd, *J* = 10.59, 8.79, 3.77 Hz, 1H, H_2), 4.30 – 4.24 (m, 2H, H_5 and H_6 or H_7), 4.15 – 4.11 (br d, *J* = 10.49 Hz, 1H, H_6 or H_7), 2.10 (s, 3H, Ac), 2.05 (d, *J* = 0.6 Hz, 6H, 2 Ac), 1.98 (s, 3H, Ac). ^{13}C NMR (101 MHz, $CDCl_3$): δ(ppm) 171.33 ($\underline{C}$=O, NH), [170.59, 170.32, 169.17 ($\underline{C}_q$, $\underline{C}$=O Ac)], 93.67 (C-Cl), 70.81 (C_3), 70.02 (C_5), 67.01(C_4), 61.12 (C_6), 53.31 (C_2), 22.97(CH_3 NH), [20.69, 20.55 (CH_3 3 OAc)]. MS (*m/z*): $[M+Na]^+_{cald}$ = 388.08. $[M+Na]^+_{obsd}$ = 388.0.

2-deoxy-2-N-acetamido-3,4,6-tri-O-acetyl-β-D-glucopyranosyl azide (**15**)



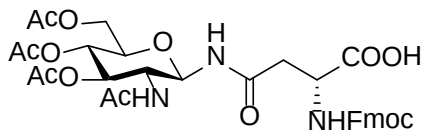
The glucopyranosyl chloride **13** (3.55 g, 9.71 mmol) was dissolved in DMF (10 ml) and sodium azide (1.58 g, 24.26 mmol) was added under stirring. After 4 h at r.t., the mixture was diluted with DCM (150 ml) and washed with water (6 X 70 ml), dried over sodium sulfate and concentrated. The glucosamine azide **15** was obtained without further purification as a white solid (3.02 g, 9.71 mmol, 84%). *Rf* = 0.4 (DCM/MeOH 95:5). ^{1}H NMR (400 MHz, $CDCl_3$): δ(ppm) 5.68 (d, *J* = 8.9 Hz, 1H, N-H), 5.24 (dd, *J* = 9.43, 9.42 Hz 1H, H_3), 5.10 (dd, , *J* = 9.87, 9.87 Hz 1H, H_4), 4.76 (dd, *J* = 9.28, 9.42 Hz, 1H, H_1), 4.27 (dd, *J* = 12.34, 4.81Hz 1H, H_6 or H_6), 4.16 (dd, *J* = 12.43, 2.19Hz 1H, H_6 or H_6), 3.91 (dt, *J* = 10.5, 9.1 Hz, 1H, H_2), 3.79 (ddd, *J* = 10.0, 4.8, 2.3 Hz, 1H, H_5), 2.10 (s, 3H, Ac), 2.04 (d, *J* = 3.0 Hz, 6H, 2 Ac), 1.98 (s, 3H, Ac). ^{13}C NMR (101 MHz, $CDCl_3$): δ(ppm) 171.07 (C=O, NH), [170.75, 170.45, 169.31 (C_q, C=O Ac)], 88.44 (C-N_3), 73.99 (C_5), 72.11 (C_3), 67.96 (C_4), 61.83 (C_6), 54.17 (C_2), 23.30 (CH_3 NH), [20.78, 20.68, 20.63 (CH_3 3 OAc)]. MS (*m/z*): $[M+H]^+_{cald}$ = 373.14. $[M+H]^+_{obsd}$ = 373.0.

2-deoxy-2-N-acetamido-3,4,6-tri-O-acetyl-β-D-glucopyranosyl amine (**11**)



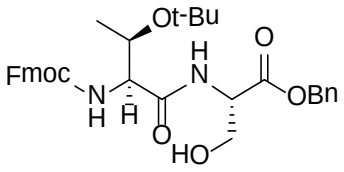
The glucopyranosyl azide **15** (0.9 g, 2.42 mmol) was dissolved in anhydrous THF (2 ml). Then, Pd/C 10% (25.7 mg, 0.24 mmol) and triethylamine (0.34 ml, 2.42 mmol) were added to the solution. The flask containing the mixture was purged with argon and and H_2 (3 bar). The suspension was stirred under a hydrogen atmosphere for 3 h. The flask was purged with argon and the reaction mixture was diluted with ethyl acetate (100 ml), filtered over Celite and concentrated. The crude amine was obtained as a white solid (0.82 g, 2.39 mmol, 98%). *Rf* = 0.1 (DCM/MeOH 95:5). ^{1}H NMR (400 MHz, $CDCl_3$): δ(ppm) 6.06 (d, *J* = 9.3 Hz, 1H, N-H), 5.06 – 4.96 (m, 2H, H_3 and H_4), 4.16 (dd, *J* = 12.2, 4.9 Hz, 1H, H_6 or H_7), 4.10 (d, *J* = 9.3 Hz, 1H, H_1), 4.05 (dd, *J* = 12.3, 2.4 Hz, 1H, H_6 or H_7), 4.00 – 3.91 (m, 1H, H_2), 3.61 (ddt, *J* = 7.3, 4.9, 2.4 Hz, 1H, H_5), 2.04 (s, 3H, Ac), 1.98 (d, *J* = 5.6 Hz, 6H, 2 Ac), 1.92 (s, 3H, Ac). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.42 (C=O, NH), [170.88, 170.81, 169.39 (C_q, C=O Ac)], 85.88 (C-NH_2), 73.39 (C_5), 72.64 (C_3), 68.60 (C_4), 62.43 (C_6), 54.74 (C_2), 23.29 (CH_3 NH), [20.81, 20.74, 20.65 (CH_3 3 OAc)]. MS (*m/z*): $[M+H]^+_{cald}$ = 347.15. $[M+H]^+_{obsd}$ = 347.1.

2-deoxy-2-N-acetamido-3,4,6-tri-O-acetyl-N-[N-Fmoc-L-aspart-4-oyl]-β-D-glucopyranosylamine (**2**)



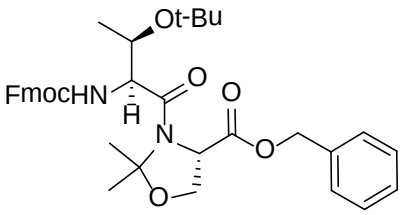
The crude glucopyranosyl amine **11** (0.82 g, 2.37 mmol) and the aspartic anhydride 12 (0.85 g, 2.37 mmol) were dissolved in anhydrous DMSO (2 ml) and stirred at r.t. during 5 h. The mixture was diluted with water (30 ml), and extracted with ethyl acetate (3 X 100 ml). The organic phase was collected, washed with a saturated solution of NaCl, dried over sodium sulfate and concentrate. The product **2** was purified by silica gel flash column chromatography using a gradient of MeOH in DCM and obtained as a white powder (695 mg, 1.02 mmol, 43%). *Rf* = 0.4 (DCM:MeOH 90:10). ^{1}H NMR (400 MHz, DMSO-D_6): δ(ppm) 8.62 (d, *J* = 8.9 Hz, 1H, NH-Glc), 7.90 (d, *J* = 9.47 Hz, 1H, NH-Ac) 7.87 (d, *J* = 7.3 Hz, 2H, Ar-H, Fmoc), 7.69 (d, *J* = 7.6 Hz, 2H, Ar-H, Fmoc), 7.48 (d, *J* = 8.1 Hz, 1H, NHα-Asp), 7.39 (t, *J* = 7.5 Hz, 2H, Ar-H, Fmoc), 7.30 (t, *J* = 7.5 Hz, 2H, Ar-H, Fmoc), 5.11 (dd, $J_{1,2}$ = 9.7 Hz, 1H, H_1-Glc), 5.07 (dd, $J_{3,4}$ = 9.7 Hz, 1H, H_3-Glc), 4.79 (t, *J* = 9.7 Hz, 1H, H_4-Glc), 4.33 (dd, *J* = 7.0, 6.5 Hz, 1H, $H_α$ Asn), 4.22 (dd, *J* = 2.7, 6.0 Hz, 1H, CH_{2b} Fmoc), 4.22 – 4.18 (m, 1H, H_9 Fmoc), 4.15 (dd, $J_{6,7}$ = 12.4 Hz, 1H, H_6-Glc), 3.94 (m, 1H, H_7-Glc), 3.85 (dd, *J* = 9.7, 9.7 Hz, 1H, H_2-Glc), 3.79 (m, 1H, H_5-Glc), 2.62 (dd, *J* = 16.3, 5.7 Hz, 1H, $H_{β1}$ Asn), 2.50 (m, 1H, $H_{β2}$ Asn), 1.97 (s, 3H, Ac), 1.94 (s, 3H, Ac), 1.88 (s, 3H, Ac), 1.69 (s, 3H, Ac-NH). ^{13}C NMR (101 MHz, DMSO- d_6): δ(ppm) 173.19 (COOH), 171.33 (C=O, NH Asn), 170.53 (C=O, NHAc), [170.35, 169.99, 169.80 (C=O, Ac)], 156.28 (C=O, Fmoc), [144.26 , 141.17, 128.12, 127.56, 125.74, 120.61 (C-Ar, Fmoc)], 78.50 (C_1), 73.81 (C_3), 72.71 (C_5), 68.78 (C_4), 66.14 (CH_2 Fmoc), 62.27 (C_6), 52.55 (C_2),50.56 (CH Asn), 47.04 (CH Fmoc), 23.07 (CH_3 NH), [21.02, 20.89, 20.86 (CH_3 3 OAc)]. MS (*m/z*): $[M+H]^+_{cald}$ = 684.24. $[M+H]^+_{obsd}$ = 684.2.

Fmoc-Thr(OtBu)-Ser-OBn (**19**)



N-α-Fmoc-L-Thr(O-tBu)-OPfp (1.92g, 3.32 mmol) was prepared following the procedure reported by Wöhr[154] and was suspended in acetone (15 ml). [154] H-L-Serine-OBn (1.95 g, 9.97 mmol) was dissolved in 4 ml of aqueous sodium carbonate solution 10% w/v (the pH of the resulting solution was 9.0) and added to the previous suspension. The mixture was stirred at r.t. for 3 h. The organic solvent was removed under vacuum and the product was extracted from the aqueous phase with ethyl acetate (3 X 100 ml). The organic phase was then washed with water and brine, dried over sodium sulfate and concentrated. The dipeptide **19** was purified over silica gel flash column chromatography using (a gradient of EtOAc in hexane and obtained as a white powder (1.87g, 3.25 mmol, 98%). *Rf* = 0.8 (Hexane:EtOAc 3:2). ^{1}H NMR (400 MHz, $CDCl_3$): δ(ppm) 8.10 (d, *J* = 7.1 Hz, 1H, NH Thr), 7.77 (dd, *J* = 7.7, 3.3 Hz, 2H, Ar-H, Fmoc), 7.61 (dd, *J* = 7.5, 3.0 Hz, 2H, Ar-H, Fmoc), 7.46 – 7.29 (m, 9H), 5.97 (m, *J* = 5.3 Hz, 1H, NH Fmoc), 5.24 (s, 2H, CH_2 Bn), 4.72 – 4.67 (m, 1H, CH_β Thr), 4.42 – 4.36 (m, 2H, CH_2 Fmoc), 4.24 (dt, *J* = 10.4, 6.4 Hz, 2H, CH fmoc and CHα Ser), 4.20 – 4.13 (m, 1H, CHα Thr), 4.05 (dd *J* = 11.3, 3.7 Hz, 1H, CH_{2a} Ser) 3.99 (dd, *J* = 44.4, 11.2, 3.6 Hz, 2H, CH_{2b} Ser), 1.30 (s, 9H, 3 CH_3 t-Bu), 1.05 (d, *J* = 6.5 Hz, 3H, CH_3 Thr). ^{13}C NMR (101 MHz, $CDCl_3$): δ(ppm) 170.07 ($\underline{C}_q$, $\underline{C}$=O Thr), 169.92 ($\underline{C}_q$, $\underline{C}$=O Ser), 155.75($\underline{C}_q$, $\underline{C}$=O Fmoc), [143.92, 143.68 (C-Ar, Fmoc)], [141.31, 128.68, 128.59, 128.31 (C-Ar, Bn)], [127.75, 127.09, 125.17, 120.03 (C-Ar, Fmoc)],75.62 ($\underline{C}_q$, $\underline{tBu}$ Thr), 67.54 (C_β Thr), 67.02 (CH_2 Fmoc), 66.68 (CH_2 Bn), 63.39 (C_α Thr), 58.61 (C_β Ser), 54.96 (C_α Ser), 47.16 (CH Fmoc), 28.13 (Me tBu Thr), 16.72 (Me Thr). MS (*m/z*): $[M+Na]^+_{cald}$ = 597.26. $[M+Na]^+_{obsd}$ = 597.2.

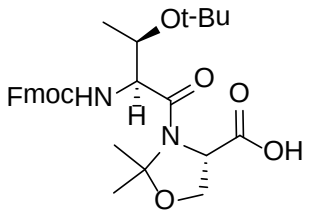
N-Fmoc-Thr(OtBu)-Ser($\Psi^{Me,Me}$ Pro)-OBn (**20**)



The dipeptide **19** (1.8 g, 3.13 mmol) was dissolved in dry toluene (100 ml) in a two-neck round flask equipped with a Dean-Stark apparatus. Pyridyl toluene-4-sulfonate (0.157 g, 0.626 mmol) and 2,2-dimethoxypropane (1.94 ml, 15.65 mmol) were added and the mixture was refluxed for 5 h. The reaction was cooled to r.t. and triethylamine (0.13 ml, 0.9 mmol) was added. The solvent was removed under vacuum and the residue was dissolved in ethyl acetate (100 ml). This solution was extracted with water (3 X 50 ml), dried over sodium sulfate and concentrated. The product was purified over silica gel using Hexane and EtOAc 1:1. The pseudoproline **20** was obtained as a white powder (1.18 g, 1.92 mmol, 61%). *Rf* = 0.7 (Hexane:EtOAc 1:1). ^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) 7.77 (d, *J* = 7.5 Hz, 2H, Ar-H, Fmoc), 7.60 (dd, *J* = 7.5, 1.5 Hz, 2H, Ar-H, Fmoc), 7.41 (dd, *J* = 7.5, 1.3 Hz, 2H, Ar-H, Fmoc), 7.34 – 7.27 (m, 7H), 5.87 (d, *J* = 7.1 Hz, 1H, NH Fmoc), 5.66 (d, *J* = 5.5 Hz, 1H, CHα Ψ^{Pro}), 5.19 (dd, *J* = 12.2, 8.5 Hz , 2H, 2H, CH_2 Bn), 4.35 (dq, *J* = 6.8, 3.2 Hz, 2H, CH_2 Fmoc), 4.29 (dd, *J* = 10.4, 3.4 Hz, 1H, CHα Thr), 4.19 (t, *J* = 7.2 Hz, 1H, CH Fmoc), 4.15 – 4.08 (m, 2H, CH_β Ψ^{Pro}), 3.95 – 3.87 (m, 1H, CH_β Thr), 1.68 (s, 3H, Me_1 Ψ^{Pro}), 1.58 (s, 3H, Me_2 Ψ^{Pro}), 1.21 (s, 9H, Me t-Bu Thr), 1.05 (d, *J* = 6.5 Hz, 3H, Me Thr).^{13}C NMR (101 MHz, $CDCl_3$): δ(ppm) 170.50 ($\underline{C}_q$, $\underline{C}$=O Ψ^{Pro}), 166.66 ($\underline{C}_q$, $\underline{C}$=O Thr), 155.18 ($\underline{C}_q$, $\underline{C}$=O Fmoc), [144.00 , 143.74 (C-Ar, Fmoc)], [134.94, 128.67, 128.50, 128.16 (C-Ar, Bn)], [127.73, 127.07, 125.27, 120.01 (C-Ar, Fmoc)], 96.84 ($C(CH_3)_2$ Ψ^{Pro}), 75.17 (C tBu Thr), 69.35 (C_β Thr), 67.91 (C_α Ψ^{Pro}), 66.88 (CH_2 Fmoc), 66.81 (CH_2 Bn), 60.44 (CH_2 Ψ^{Pro}), 59.67 (Cα Thr), 47.16 (CH Fmoc), 27.96 (Me tBu Thr), 25.34 (Me Ψ^{Pro}), 17.45 (Me Thr). MS (*m/z*): $[M+H]^+_{cald}$ = 615.31.

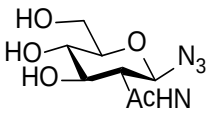
$[M+H]^+_{obsd}$ = 615.2.

N-Fmoc-Thr(OtBu)-Ser($\Psi^{Me,Me}$ Pro)-OH (**3**)



The protected pseudoproline dipeptide **20** (0.58 g, 0.94 mmol) and Pd/C 10% (0.02 g, 0.18 mmol) were suspended in MeOH anhydrous (10 ml). The flask containing the reaction mixture was purged with argon (3X) and the mixture was stirred at r.t under a hydrogen atmosphere. After 2 h, the solution was filtered through Celite and concentrated. The pseudoproline dipeptide **3** was obtained as a white solid (0.49 g, 0.93 mmol, quantitative) and used in SPPS without further purification. ^{1}H NMR (400 MHz, DMSO-D_6): δ (ppm) 7.87 (d, *J* = 7.5 Hz, 2H, Ar-H, Fmoc), 7.76 (t, *J* = 7.0 Hz, 2H, Ar-H, Fmoc), 7.39 (t, *J* = 7.5 Hz, 2H, Ar-H, Fmoc), 7.35 – 7.26 (m, 2H, Ar-H, Fmoc), 5.26 (d, *J* = 5.7 Hz, 1H, NH Fmoc), 4.26 – 4.13 (m, 5H, CH Fmoc, CH_2 Fmoc, CH_2 Ψ^{Pro}), 4.09 – 4.03 (m, 1H, CHα Thr), 3.99 (dd, *J* = 7.1 Hz, 1H, CHα Ψ^{Pro}), 3.77 – 3.70 (m, 1H, CH_β Thr), 1.56 (s, 3H, Me_1 Ψ^{Pro}), 1.44 (s, 3H, Me_2 Ψ^{Pro}), 1.13 (s, 9H, Me t-Bu Thr), 0.99 (d, *J* = 6.4 Hz, 3H, Me Thr). ^{13}C NMR (101 MHz, DMSO-d_6): δ (ppm) 172.43 (COOH), 166.67 (C_q, C=O Thr), 155.55 (C_q, C=O Fmoc), [144.41, 144.16, 128.14, 127.55, 126.03, 120.97 (C-Ar, Fmoc)], 95.91 ($C(CH_3)_2$ Ψ^{Pro}), 74.75 (C tBu Thr), 69.66 (C_β Thr), 66.94 (C_α Ψ^{Pro}), 66.39 (CH_2 Fmoc), 60.23 (CH_2 Ψ^{Pro}), 59.52 (Cα Thr), 47.10 (CH Fmoc), 28.09 (Me tBu Thr), 25.67 (Me Ψ^{Pro}), 18.46 (Me Thr). MS (*m/z*): $[M+H]^+_{cald}$ = 525.26. $[M+H]^+_{obsd}$ = 525.2.

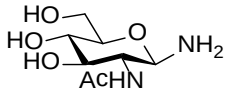
2-deoxy-2-N-acetamido-β-D-glucopyranosyl azide (**18**)



The O-*per*-acetylated glucopyranosyl azide **15** (0.23 g, 0.63 mmol) was dissolved in MeOH (4 ml) and fresh prepared sodium methoxide (sodium hydride was slowly dissolved in methanol, pH 9.0) was added (5 ml). The reaction was stirred for 4 h at r.t., the mixture was neutralized with amberlite IR120 resin [H^+] and filtrated. The solvent was removed under vacuum and the product **18** was obtained without further

purification as a white solid (0.15 g, 0.63 mmol, quantitative). ^{1}H NMR (400 MHz, MeOD): δ(ppm) 4.44 (d, *J* = 9.3 Hz, 1H, H_1), 3.81 (dd, *J* = 12.2, 1.8 Hz, 1H, H_6), 3.63 (dd, *J* = 8.2, 3.8 Hz, 1H, H_7), 3.61 – 3.59 (m, 1H, H_2), 3.42 – 3.36 (m, 1H, H_3), 3.29 (d, *J* = 5.7 Hz, 1H, H_5), 3.26 (d, *J* = 3.1 Hz, 1H, H_4), 1.92 (s, 3H, Ac). ^{13}C NMR (101 MHz, MeOD): δ(ppm) 172.54 ($\underline{C}_q$, $\underline{C}$=O), 88.49 (C_1), 78.90 (C_5), 74.37 (C_3), 70.23 (C_4), 61.19 (C_6), 55.27 (C_2), 21.62 (Ac). MS (*m/z*): $[M+Na]^+_{cald}$ = 269.09. $[M+Na]^+_{obsd}$ = 269.0.

2-deoxy-2-N-acetamido-β-D-glucopyranosyl amine (**5**)



The N-acetylglucopyranosyl azide **18** (0.15 g, 0.63 mmol) was dissolved in MeOH anhydrous (5 ml). Then, Pd/C (10% Pd, 0.07 g, 0.063 mmol) was added. The mixture was purged with argon and hydrogen. The suspension was stirred for 1 h under a hydrogen atmosphere. The catalyst was removed by filtration over Celite and the filtrate was concentrated to give the amine **5** as a white solid (0.13 g, 0.62 mmol, 98%). ^{1}H NMR (400 MHz, MeOD): δ(ppm) 3.95 (d, *J* = 9.2 Hz, 1H, H_1), 3.80 (dd, *J* = 11.8, 2.1 Hz, 1H, H_6), 3.59 (dd, *J* = 11.8, 5.6 Hz, 1H, H_7), 3.52 (t, *J* = 9.7 Hz, 1H, H_3), 3.36 (dd, *J* = 10.1, 8.0 Hz, 1H, H_2), 3.24 (m, 1H, H_4), 3.23 – 3.18 (m, 1H, H_5), 1.96 (s, 3H, Ac). ^{13}C NMR (101 MHz, MeOD): δ(ppm) 172.99 ($\underline{C}_q$, $\underline{C}$=O), 84.92 (C_1), 77.61 (C_5), 75.07 (C_3), 70.90 (C_4), 61.59 (C_6), 56.77 (C_2), 21.62 (Ac). MS (*m/z*): $[M+Na]^+_{cald}$ = 221.11. $[M+Na]^+_{obsd}$ = 221.1.

4.3 General method for the solid phase peptide synthesis (SPPS)

The synthesis of the peptide fragments was performed both manually and in the automated approach. The functionalization of the resins, the coupling of the first amino acid to the resin, and the coupling of the pseudo-prolines and valuable glycosylated building blocks were performed manually in a 5, 10 ml or 20 ml disposable polypropylene syringes equipped with 25 µm bottom Teflon filter (VWR, Germany).

The elongation of the peptide and glycopeptide fragments was performed by automated microwave-assisted solid phase peptide synthesis using a Liberty Blue^{TM} peptide synthesizer (CEM Corporation, Germany). The main operations (washes, addition of the deprotection solution, and microwave method) were calibrated before every synthesis. The automated peptide synthesis

was carried in a 30 ml Teflon reaction vessel using the CEM suggested conditions for 0.1 mmol scale (table 8). The temperature inside the reaction vessel was controlled by a fiber optic probe during the entire process. The solutions used for the peptide synthesis were prepared freshly as shown in the following table:

Solution	Concentration	Volume
Main solvent	DMF	4 ml each wash
Deprotection	20% piperidine in DMF[a]	4 ml
Amino acids (a.a.)	0.2 M in DMF	2.5 ml
Activator (Act.)	0.5 M DIC in DMF	1 ml
Activator base (Act. Base)	1.0 M Oxyma in DMF	0.5 ml

Table 8 - CEM peptide Synthesizer suggested conditions for 0.1 mmol scale; [a] 0.7% of FA (v/v of the total volume prepared) was added for the optimized conditions

4.3.1 Qualitative amine test (Kaiser Test)

The presence of free amino groups were detected with ninhydrin after manual couplings and Fmoc-deprotections following the method reported by Kaiser.[145] **Solution A**: 80% phenol in ethanol (w/v) and KCN in pyridine (2 ml of KCN 1M in 98 ml of pyridine) in proportion 1:4 (v/v). **Solution B**: 5% ninhydrine in ethanol (w/v). Two 1.5 ml Eppendorf tubes were used: one containing few mg of the resin labeled as Sample, and one empty designated as Reference. Two drops of **solution A** and one drop of **solution B** were added to both tubes. Both samples were heated at 100° C for 5 min. The presence of free amino groups were indicated by an intense blue color of the solution and the beads in the tube labelled as Sample.

4.3.2 Fmoc removal quantification and determination of the resin loading

Peptide-resin (between 0.3 mg to 1.2 mg) having the Fmoc-protected amino acid were added into a 2 ml Eppendorf tube. A freshly prepared 20% solution of piperidine in DMF v/v (1.8 ml) was added, the mixture was shaken at r.t. for 15 min, and centrifuged. The supernatant (1 ml) was poured in a clean 2 ml Eppendorf tube and was diluted with 1 ml of the deprotection solution. The absorbance at 290 nm was then measured in a 1 ml quartz cuvette using the 20% solution of piperidine in DMF as a reference. The loading of the resin was calculated according the following equation:

$$loading^{obs}\left(\frac{mmol}{g}\right) = \frac{Abs_{290\,nm}}{mg_{resin} \ * \ 1.75}$$

where 1.75 is referred to ε = 5253 M^{-1} cm^{-1}.

The observed value was compared with the theoretical loading which was obtained as follows:

$$loading^{\,theor}\left(\frac{mmol}{g}\right) = \frac{A * 1000}{1000 + A(B - C)}$$

where A is the initial substitution of the resin (mmol/g); B is the molecular weight of the loaded amino acid included all the protecting groups; C is the mass lost in Dalton after the coupling of the amino acid and it corresponds to 18 for hydroxymethyl resins and 36 for trityl chloride resins.

4.3.3 Automated deprotection/coupling cycles used for the peptide elongation

The elongation of the designed peptide and glycopeptide fragments were performed using the setup for 0.1 mmol scale (as describe before), using 5 molar equiv. of amino acids, activator, and activator base for each coupling cycle. The optimized deprotection and coupling cycles for each microwave method were as follow:

0.10 - Single Deprotection Single Coupling 75°C

Step	Operation	Parameters
1	Deprotection	75°C, 50W, 3min (Volume delivered = 4ml)
2	Wash	DMF, 4ml, 17 sec, draining time 10 sec
3	Wash	DMF, 4ml, 17 sec, draining time 10 sec
4	Wash	DMF, 4ml, 17 sec, draining time 10 sec
5	Wash	DMF, 4ml, 17 sec, draining time 10 sec
6	Coupling	75°C 30W 5 min, a.a (2.5 ml), Act. (1 ml), Act. base (0.5 ml)
Total time		**12 min**

0.10 - Double Deprotection Double Coupling 75°C

Step	Operation	Parameters
1	Deprotection	75°C, 50W, 3min (Volume delivered = 4ml)
2	Wash	DMF, 4ml, 17 sec, draining time 10 sec
3	Deprotection	75°C, 50W, 3min (Volume delivered = 4ml)
4	Wash	DMF, 4ml, 17 sec, draining time 10 sec
5	Wash	DMF, 4ml, 17 sec, draining time 10 sec
6	Wash	DMF, 4ml, 17 sec, draining time 10 sec
7	Coupling	75°C 30W 5 min, a.a (2.5 ml), Act. (1 ml), Act. base (0.5 ml)
8	Coupling	75°C 30W 5 min, a.a (2.5 ml), Act. (1 ml), Act. base (0.5 ml)
Total time		**19 min**

0.10 - Single Deprotection Double Coupling 75°C

Step	Operation	Parameters
1	Deprotection	75°C, 50W, 3min (Volume delivered = 4ml)
2	Wash	DMF, 4ml, 17 sec, draining time 10 sec
3	Wash	DMF, 4ml, 17 sec, draining time 10 sec
4	Wash	DMF, 4ml, 17 sec, draining time 10 sec
5	Coupling	75°C 30W 5 min, a.a (2.5 ml), Act. (1 ml), Act. base (0.5 ml)
6	Coupling	75°C 30W 5 min, a.a (2.5 ml), Act. (1 ml), Act. base (0.5 ml)
Total time		**22 min**

0.10 - Double Deprotection Double Coupling 50°C

Step	Operation	Parameters
1	Deprotection	75°C, 50W, 3min (Volume delivered = 4ml)
2	Wash	DMF, 4ml, 17 sec, draining time 10 sec
3	Deprotection	75°C, 50W, 3min (Volume delivered = 4ml)
4	Wash	DMF, 4ml, 17 sec, draining time 10 sec
5	Wash	DMF, 4ml, 17 sec, draining time 10 sec
5	Wash	DMF, 4ml, 17 sec, draining time 10 sec
7	Coupling	25°C 0W 2 min + 50°C 35W 8 min a.a. (2.5 ml), Act. (1 ml), Act. base (0.5 ml)
8	Coupling	25°C 0W 2 min + 50°C 35W 8 min a.a. (2.5 ml), Act. (1 ml), Act. base (0.5 ml)
Total time		**30 min**

0.10 - Double Arg Coupling

Step	Operation	Parameters
1	Deprotection	75°C, 50W, 3min (Volume delivered = 4ml)
2	Wash	DMF, 4ml, 17 sec, draining time 10 sec
3	Deprotection	75°C, 50W, 3min (Volume delivered = 4ml)
4	Wash	DMF, 4ml, 17 sec, draining time 10 sec
5	Wash	DMF, 4ml, 17 sec, draining time 10 sec
5	Wash	DMF, 4ml, 17 sec, draining time 10 sec
7	Coupling	25°C 0W 25 min + 75°C 35W 2 min a.a. (2.5 ml), Act. (1 ml), Act. base (0.5 ml)
8	Coupling	25°C 0W 25 min + 75°C 35W 2 min a.a. (2.5 ml), Act. (1 ml), Act. base (0.5 ml)
Total time		**66 min**

4.3.4 Small- and large-scale cleavage of the peptide and glycopeptide fragments from the resin

A small-scale cleavage was performed to verify the quality of the desired products and to establish the best conditions (cocktail solution and time) for the release of the peptide fragments from the resin and the removal of all temporary protecting groups on the side chains of the amino acids.

Two cleavage cocktails mixtures of TFA and scavengers were used for the cleavage of peptide and glycopeptide fragments:

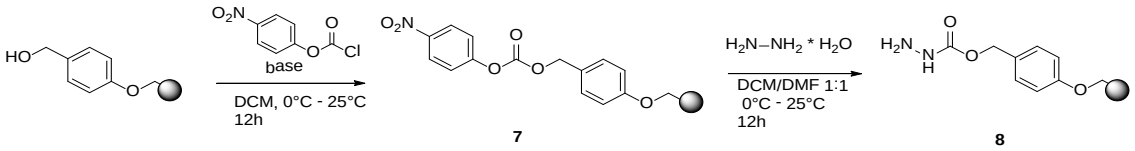
Solution	Reagents	Ratio
K	TFA/phenol/water/thioanisole/EDT	82.5/5/5/5/2.5
R	TFA/thioanisole/EDT/anisole	90/5/3/2

Few milligrams of the dried resin bearing the peptide were transferred into a 1.5 ml Eppendorf tube and 250 µl of the TFA cocktail cleavage solution were added. The slurry was shaken at r.t. for at least 2 h. Cold ethyl ether (1 ml) was added to precipitate the peptide, and the mixture was centrifugated. The supernatant containing the released protecting groups was removed. This operation was repeated at least other three times. The precipitate was dried under air and dissolved in 10% acetonitrile in water (some intermediates required different conditions). The solution was analyzed by RP-HPLC and LC-MS.

The large-scale cleavage of peptides was performed as described for the small-scale process using a 10 ml disposable polypropylene syringe equipped with a bottom Teflon frit and using 7 ml/200 mg of peptide-resin of the chosen cocktail solution. The filtrate was collected in 50 ml Falcon tube and the peptide was precipitated and triturated with 45 ml of cold diethyl ether. The centrifugation and the removal of the supernatant were carried out as described above and repeated for at least three times.

4.4 Functionalization of the resin to obtain C-terminal hydrazide peptides

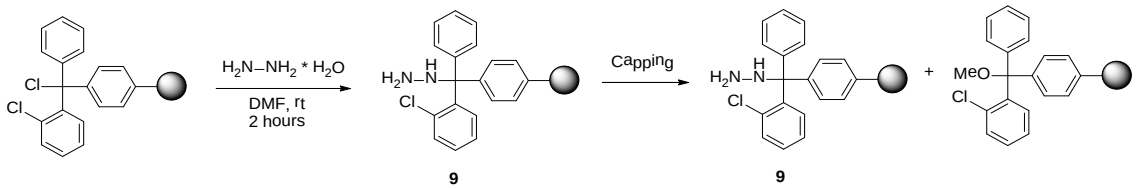
*Functionalization of NovaPEG Wang resin (**8**)*



Scheme 65- Functionalization of Wang resin with p-nitrophenyl chloroformate and base, and hydrazine monohydrate

NovaPEG Wang resin (635 mg, 0.39 mmol according to the supplier's loading of the resin) was swollen in 8 ml of anhydrous DCM for 1 h at r.t. Then, the solvent was drained and a cold solution of *para*-nitrophenyl chloroformate (390.6 mg, 1.9 mmol) in DCM (4 ml) was added to the resin. The reactor vessel was cooled to -20°C, and after 1 h a solution of N-methylmorpholine (212.2 µl, 1.9 mmol) in DCM (4 ml) was added to the resin. The mixture was warmed to r.t. and shaken overnight. The solvent was drained, and the resin was washed with DCM (5 X 4 ml), DMF (5 X 4 ml), MeOH (5 X 4 ml), and DCM (5 X 4 ml). A cold solution of hydrazine monohydrate (92.2 µl, 1.9 mmol) in DCM/DMF 1:1 (8 ml) was then added to the resin and the mixture was shaken overnight at r.t. The solvent was drained, and the resin was washed with DCM (5 X 4 ml), DMF (5 X 4 ml), MeOH (5 X 4 ml), and DCM (5 X 4 ml), dried under vacuum and stored at -20°C.

*Functionalization of 2-Cl-(Cl)trityl ProTide resin (**9**)*

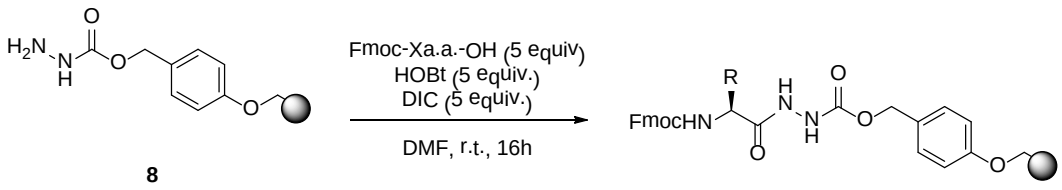


Scheme 66 - Functionalization of Cl-TCP(Cl) ProTide resin with hydrazine monohydrate

Cl-TCP(Cl) ProTide resin (200 mg, 0.096 mmol according to the supplier's loading of the resin) was swollen with a solution of DMF/DCM (1:1, 4 ml) for 30 min. The solvent was removed and a solution 5% v/v of hydrazine monohydrate in DMF (2 ml) was added to the resin. After 1.5 h, the resin was drained and washed with DMF (2 X 4 ml). A new solution of hydrazine was added and the mixture was shaked for additional 1.5 h. Finally the resin was washed with DMF (5 X 4

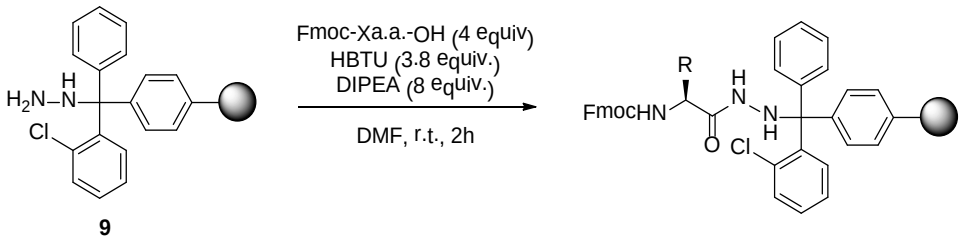
ml), DCM (5 X 4 ml), and DMF (5 X 4 ml). Then, a solution of 5% v/v of methanol in DMF (2 ml) was added to the resin and the mixture was shaken for 20 min. The resin was drained and washed with DMF (5 X 4 ml), DCM (5 X 4 ml), and DMF (5 X 4 ml). The functionalized resin was immediately used for the coupling of the first amino acid.

General method for the loading of the first amino acid on functionalized NovaPEG Wang resin and 2-Cl-(Cl)trityl ProTide resin



Scheme 67 –Loading of the Fmoc-protected amino acid on the functionalized Wang resin.

The hydrazine functionalized trityl resin **9** (0.16 g, 0.10 mmol) was swollen in DCM (4 ml) at r.t. for 30 min. The amino acid (0.5 mmol) was pre-activated with HOBt (67.5 mg, 0.5 mmol), and DIC (77.4 µl, 0.5 mmol) in DMF (4 ml) for 2 min. The amino acid solution was added to the resin and the slurry was shaken for 16 h. The resin was drained and washed with DMF (4 X 4 ml), DCM (4 X 4 ml), DMF (4 X 4 ml). A small amount of the dried resin was taken to verify the efficiency of the co upling and treated with the solution of piperidine 20% in DMF v/v following the protocol described in 4.3.2.



Scheme 68 - Loading of the Fmoc-protected amino acid on the functionalized trityl resin.

The hydrazine functionalized trityl resin **9** (0.20 g, 0.10 mmol) was swollen in DCM (4 ml) at r.t. for 30 min. The amino acid (0.40 mmol) was pre-activated with HBTU (0.13 g, 0.38 mmol) and DIPEA (0.13 ml, 0.8 mmol) in DMF (2 ml) for 2 min. The amino acid solution was added to the resin and the slurry was shaken for 1 h. The resin was drained and washed with DMF (4 X 4 ml), DCM (4 X 4 ml), DMF (4 X 4 ml). A small amount of the dried resin was taken to verify the

efficiency of the coupling and treated with the solution of piperidine 20% in DMF v/v following the protocol described in 4.3.2.

4.5 Synthesis of peptide and glycopeptide fragments

The following table lists the peptide and glycopeptide fragments synthesized in the scope of this work for the assembly of the mature form Thy-1 glycoprotein.

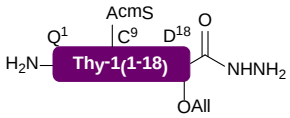
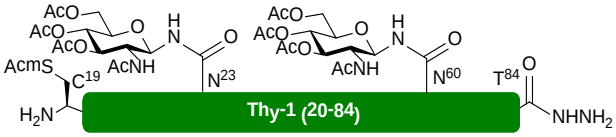
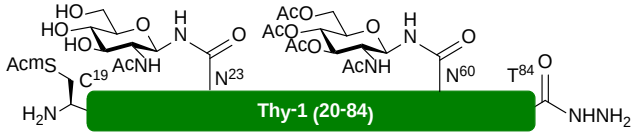
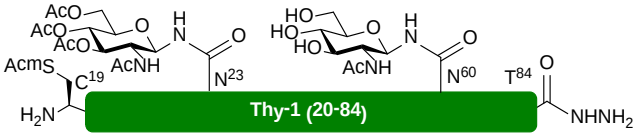
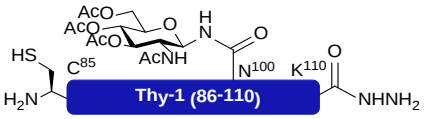
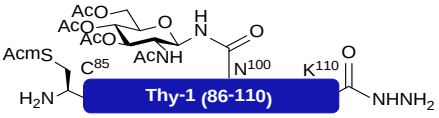
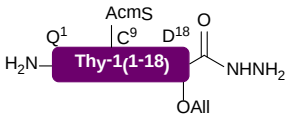
Fragment	Compound	Paragraph n.
	**1**	**4.5.1**
	**6a**	**4.5.2**
	**6b**	**4.5.3**
	**6c**	**4.5.4**
	**37a**	**3.5.5**
	**37b**	**3.5.5**

Table 8 – Required fragments for the assembly of the glycoprotein Thy-1 (1-120).

4.5.1 Synthesis of Thy-1 Peptide Fragment 1

H-Gln-Lys-Val-Thr-Ser-Leu-Thr-Ala-Cys(Acm)-Leu-Val-Asp-Gln-Ser-Leu-Arg-Leu-Asp(OAll)-NH-NH₂ (**1**)



The synthesis of this peptide fragment was accomplished using functionalized 2-Cl-trt-ProTide resin **9** (200 mg, 0.096 mmol) loaded with Fmoc-Asp(OAll)-OH as described before. The calculated loading was 0.24 mmol/g, 62% compared to the theoretical loading. The resin was placed in the reactor and the peptide was elongated by MW-SPPS on the synthesizer using 5 molar excess of the amino acid and DIC/Oxima activation in DMF. The amino acid and DIC were used at 0.2 M and 0.5 M, respectively, the Oxyma was 1 M concentration (Table 1). The Fmoc-group was removed using 20 % piperidine in DMF. The following table contain the deprotection/coupling cycles used for each amino acid according to the parameter described in **4.3.3**

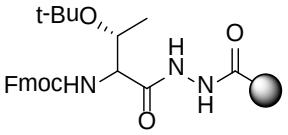
Position	A.A.	Cycle
17	Fmoc-Leu-OH	*0.10-Double Deprotection Double Coupling 75°C*
16	Fmoc-Arg(Pbf)-OH	*0.10 - Double Arg Coupling*
15	Fmoc-Leu-OH	*0.10-Single 75°C Coupling*
14	Fmoc-Ser(tBu)-OH	*0.10-Single 75°C Coupling*
13	Fmoc-Gln(Trt)-OH	*0.10-Single 75°C Coupling*
12	Fmoc-Asp(OtBu)-OH	*0.10-Single 75°C Coupling*
11	Fmoc-Val-OH	*0.10-Single 75°C Coupling*
10	Fmoc-Leu-OH	*0.10-Single 75°C Coupling*
9	Fmoc-Cys(Acm)-OH	*0.10- Double Deprotection Double Coupling 50°C*
8	Fmoc-Ala-OH·H_2O	*0.10-Single 75°C Coupling*
7	Fmoc-Thr(tBu)-OH	*0.10-Single 75°C Coupling*
6	Fmoc-Leu-OH	*0.10-Single 75°C Coupling*
5	Fmoc-Ser(tBu)-OH	*0.10-Single 75°C Coupling*
4	Fmoc-Thr(tBu)-OH	*0.10-Single 75°C Coupling*
3	Fmoc-Val-OH	*0.10-Double Deprotection Double Coupling 75°C*
2	Fmoc-Lys(Boc)-OH	*0.10-Single 75°C Coupling*
1	Fmoc-Gln(Trt)-OH	*0.10-Single 75°C Coupling (HS)*

After the removal of Fmoc-group at the *N*-terminal amino acid in the synthesizer, the peptide was released from the solid support using reagent R for 2 h, and precipitated and triturated with cold diethyl ether (3 X 45 ml). The crude peptide (119 mg) was separated by semi-preparative RP-HPLC on a Synergi 4µm Fusion RP 80 Å (250 x 21.2 mm) column using a gradient of 5 to 70% of ACN in water containing 0.1% of formic acid, in 25 min and at a flow rate of 8 ml/min. The separation of the compounds was followed using UV-detection at 214 and 280 nm.

Yield: 25 mg (20% compared to the calculated loading and the amount of resin used, 21% based on the peptide crude obtained). $C_{90}H_{159}N_{27}O_{29}S$. ESI-MS (*m/z*): $[M+H]^+_{cald}$ = 2115.16. M_{obsd} = 2115.2222 $[M+H]^+$, 1058.1187 $[M+2H]^{2+}$, 705.7537 $[M+3H]^{3+}$, 4228.4772 [2M].

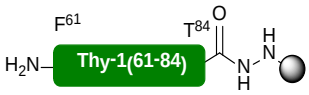
4.5.2 Synthesis of Thy-1 Peptide Fragment 6a

Fmoc-Thr(OtBu)-NH-NH-Resin (**21**)



The synthesis of this peptide fragment was accomplished using functionalized NovaPEG Wang reson **8** (163.6 mg, 0.1 mmol) loaded with Fmoc-Thr(Ot-Bu)-OH (199 mg, 0.5 mmol) as described before.. The loading calculated was 0.44 mmol/g (84% of coupled sites compared to the theoretical loading). The resin was capped with 5 ml solution of DMF/Ac_2O/DIPEA (4/0.5/0.5) for 10 min and then washed with DMF (4 X 4 ml), and DCM (4 X 4 ml).

H-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Tyr(OtBu)-Asn(Trt)-Met-Lys(Boc)-Val-Leu-Tyr(OtBu)-Leu-Ser(OtBu)-Ala-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Asp(OtBu)-Glu(OtBu)-Gly-Thr(OtBu)-Tyr(OtBu)-Thr(OtBu)-NH- NH-resin **(22)**

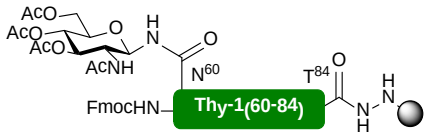


The loading Wang resin loaded the Fmoc-Thr(OtBu)-OH as a first amino acids was elongated with the following amino acids by automated MW-SPPS using the following deprotection/coupling cycles in the synthesizer as described in 4.3.3.

Position	A.A.	Cycle
83	Fmoc-Tyr(OtBu)-OH	*0.10-Double Deprotection Single Coupling 75°C*
82	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
81	Fmoc-Gly-OH	*0.10-Single 75°C Coupling*
80	Fmoc-Glu(OtBu)-OH	*0.10- Double Deprotection Double Coupling 75°C*
79	Fmoc-Asp(OtBu)-OH	*0.10-Single 75°C Coupling*
78	Fmoc-Lys(Boc)-OH	*0.10-Single 75°C Coupling*
77	Fmoc-Ser(tBu)-OH	*0.10-Single 75°C Coupling*
76	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
75	Fmoc-Phe-OH	*0.10- Double Deprotection Double Coupling 75°C*
74	Fmoc-Ala-OH · H_2O	*0.10-Single 75°C Coupling*
73	Fmoc-Ser(OtBu)-OH	*0.10-Single 75°C Coupling*
72	Fmoc-Leu-OH	*0.10-Single 75°C Coupling*
71	Fmoc-Tyr(OtBu)-OH	*0.10-Single 75°C Coupling*
70	Fmoc-Leu-OH	*0.10-Single 75°C Coupling*
69	Fmoc-Val-OH	*0.10- Double Deprotection Double Coupling 75°C*
68	Fmoc-Lys(Boc)-OH	*0.10-Single 75°C Coupling*
67	Fmoc-Met-OH	*0.10-Single 75°C Coupling*
66	Fmoc-Asn(Trt)-OH	*0.10-Single 75°C Coupling*
65	Fmoc-Tyr(OtBu)-OH	*0.10-Single 75°C Coupling*
64	Fmoc-Lys(Boc)-OH	*0.10-Single 75°C Coupling*
63	Fmoc-Ser(OtBu)-OH	*0.10-Single 75°C Coupling*
62	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
61	Fmoc-Phe-OH	*0.10- Double Deprotection Double Coupling 75°C*

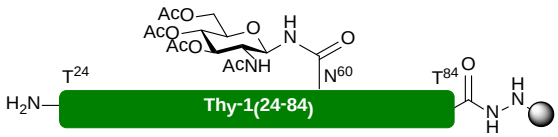
After the Fmoc removal on the residue Phe61, a small-scale cleavage was performed on using ~2 mg of the peptide-resin and 250 µl of the TFA cleavage cocktail K for 2 h at r.t. The resulting peptide was precipitated and washed 4 times with cold diethyl ether and dried. The obtained peptide was analyzed with LC-MS on a YMC hydrosphere C18 column (gradient from 5 to 70% of ACN in water + 0.1 % of formic acid in 30 minutes, 0.5 ml/min). $C_{128}H_{194}N_{30}O_{39}S$ calculated for the peptide *H-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH_2*. MS (*m/z*): $M+H]^+_{cald}$ = 2808.39. M_{obsd} = 1738.2 $[M+2H]^{2+}$, 937.1 $[M+3H]^{3+}$, 703.1 $[M+4H]^{4+}$, 562.7 $[M+5H]^{5+}$.

Fmoc-Asn((OAc)$_3$-GlcNAc)-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Tyr(OtBu)-Asn(Trt)-Met-Lys(Boc)-Val-Leu-Tyr(OtBu)-Leu-Ser(OtBu)-Ala-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Asp(OtBu)-Glu(OtBu)-Gly-Thr(OtBu)-Tyr(OtBu)-Thr(OtBu)-NH- NH-resin (**23**)



The glycosylated amino acid was coupled manually. A solution of **2** (138.4 mg, 0.2 mmol), PyBop (104 mg, 0.2 mmol), and DIPEA (70 µl, 0.4 mmol) in DMF (3 ml) and DMSO (6 µl) was prepared and added to the resin. The mixture was shaken at r.t. overnight. After the removal of the solvent, the resin was washed with DMF (4 X 4 ml), and DCM (4 X 4 ml), and a small amount was taken and submitted to the small-scale cleavage. The efficacy of the coupling was checked with the LC-MS system on a YMC hydrosphere C18 column (gradient from 10 to 40% of ACN in water + 0.1 % of formic acid in 30 min, 0.5 ml/min). $C_{161}H_{229}N_{33}O_{51}S$ calculated for the peptide *Fmoc-Asn((OAc)$_3$-GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH$_2$*. MS (*m/z*): $[M+H]^+_{cald}$ = 3473.61. M_{obsd} = 1405.1 $[M+2H]^{2+}$, 1158.9 $[M+3H]^{3+}$, 869.5 $[M+4H]^{4+}$.

H-Thr(OtBu)-Ser(OtBu)-Ser(OtBu)-Ser(OtBu)-Pro-Ile-Gln(Trt)-Tyr(OtBu)-Glu(OtBu)-Phe-Ser(OtBu)-Leu-Thr(OtBu)-Arg(Pbf)-Glu(OtBu)-Thr(OtBu)-Lys(Boc)-Lys(Boc)-His(Trt)-Val-Leu-Phe-Gly-Thr(OtBu)-Val-Gly-Val-Pro-Glu(OtBu)-His(Trt)-Thr(OtBu)-Tyr(OtBu)-Arg(Pbf)-Ser(OtBu)-Arg(Pbf)-Thr(OtBu)-Asn((OAc)$_3$-GlcNAc)-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Tyr(OtBu)-Asn(Trt)-Met-Lys(Boc)-Val-Leu-Tyr(OtBu)-Leu-Ser(OtBu)-Ala-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Asp(OtBu)-Glu(OtBu)-Gly-Thr(OtBu)-Tyr(OtBu)-Thr(OtBu)-NH-NH-resin (**24**)

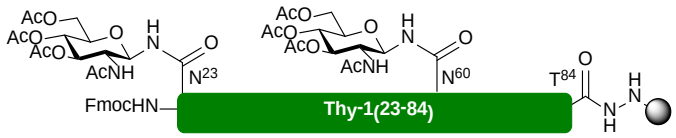


The glycopeptide **23** was elongated in the microwave-assisted synthesizer and the deprotection/coupling set wee the following:

Position	A.A.	Cycle
59	Fmoc-Thr(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*
58	Fmoc-Arg(Pbf)-OH	*0.10 - Double Arg Coupling*
57	Fmoc-Ser(OtBu)-OH	*0.10-Single 75°C Coupling*
56	Fmoc-Arg(Pbf)-OH	*0.10 - Double Arg Coupling*
55	Fmoc-Tyr(OtBu)-OH	*0.10-Single 75°C Coupling*
54	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
53	Fmoc-His(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*
52	Fmoc-Glu(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*
51	Fmoc-Pro-OH	*0.10-Single 75°C Coupling*
50	Fmoc-Val-OH	*0.10-Single 75°C Coupling*
49	Fmoc-Gly-OH	*0.10-Single 75°C Coupling*
48	Fmoc-Val-OH	*0.10-Single 75°C Coupling*
47	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
46	Fmoc-Gly-OH	*0.10-Single 75°C Coupling*
45	Fmoc-Phe-OH	*0.10-Single 75°C Coupling*
44	Fmoc-Leu-OH	*0.10-Single 75°C Coupling*
43	Fmoc-Val-OH	*0.10-Single 75°C Coupling*
42	Fmoc-His(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*
41	Fmoc-Lys(Boc)-OH	*0.10-Single 75°C Coupling*
40	Fmoc-Lys(Boc)-OH	*0.10-Single 75°C Coupling*
39	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
38	Fmoc-Glu(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*
37	Fmoc-Arg(Pbf)-OH	*0.10 - Double Arg Coupling*
36	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
35	Fmoc-Leu-OH	*0.10-Single 75°C Coupling*
34	Fmoc-Ser(OtBu)-OH	*0.10-Single 75°C Coupling*
33	Fmoc-Phe-OH	*0.10-Single 75°C Coupling*
32	Fmoc-Glu(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*
31	Fmoc-Tyr(OtBu)-OH	*0.10-Single 75°C Coupling*
30	Fmoc-Gln(Boc)-OH	*0.10-Single 75°C Coupling*
29	Fmoc-Ile-OH	*0.10-Double Deprotection Double Coupling 75°C*
28	Fmoc-Pro-OH	*0.10-Single 75°C Coupling*
27	Fmoc-Ser(OtBu)-OH	*0.10-Single 75°C Coupling*
26	Fmoc-Ser(OtBu)-OH	*0.10-Single 75°C Coupling*
25	Fmoc-Ser(OtBu)-OH	*0.10-Single 75°C Coupling*
24	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*

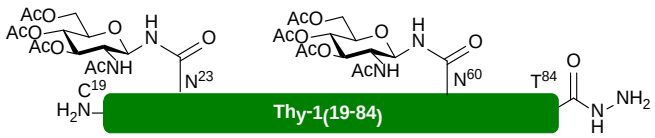
A small-scale cleavage followed the deprotection of the Fmoc-group to check the successful of the elongation. The freshly prepared TFA reagent K was added to a 1.5 ml Eppendorf tube containing few beads of **24**. After 2 h, the crude was analyzed by RP-UPLC ESI-QTof system using a Acquity UPLC BEH C18 column (gradient from 10 to 50% of ACN in water + 0.1% of formic acid in 14 minutes, 0.95 ml/min). $C_{330}H_{505}N_{85}O_{105}S$ calculated for the peptide *H-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OAc)$_3$-GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH$_2$*. ESI-MS (*m/z*): $[M+H]^+_{cald}$ = 7361.66. M_{obsd} = 1843.7178 $[M+4H]^{4+}$, 1475.0087 $[M+5H]^{5+}$, 1229.5326 $[M+6H]^{6+}$, 1053.9023 $[M+7H]^{7+}$, 922.3040 $[M+8H]^{8+}$, 819.9453 $[M+9H]^{9+}$.

Fmoc-Asn((OAc)$_3$-GlcNAc)-Thr(OtBu)-Ser(OtBu)-Ser(OtBu)-Ser(OtBu)-Pro-Ile-Gln(Trt)-Tyr(OtBu)-Glu(OtBu)-Phe-Ser(OtBu)-Leu-Thr(OtBu)-Arg(Pbf)-Glu(OtBu)-Thr(OtBu)-Lys(Boc)-Lys(Boc)-His(Trt)-Val-Leu-Phe-Gly-Thr(OtBu)-Val-Gly-Val-Pro-Glu(OtBu)-His(Trt)-Thr(OtBu)-Tyr(OtBu)-Arg(Pbf)-Ser(OtBu)-Arg(Pbf)-Thr(OtBu)-Asn((OAc)$_3$-GlcNAc)-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Tyr(OtBu)-Asn(Trt)-Met-Lys(Boc)-Val-Leu-Tyr(OtBu)-Leu-Ser(OtBu)-Ala-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Asp(OtBu)-Glu(OtBu)-Gly-Thr(OtBu)-Tyr(OtBu)-Thr(OtBu)-NH- NH-resin (**25**)



The glycosylated amino acid was coupled manually. A solution of **2** (138.4 mg, 0.2 mmol), PyBop (104 mg, 0.2 mmol), and DIPEA (70 µl, 0.4 mmol) in DMF (3 ml) and DMSO (6 µl) was prepared and added to the resin. The mixture was shaken at r.t. overnight. After the removal of the solvent, the resin was washed with DMF (4 X 4 ml), and DCM (4 X 4 ml).

H-Cys(Acm)-Arg-His-Glu-Asn((OAc)$_3$-GlcNAc)-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OAc)$_3$-GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH$_2$ **(6a)**



The glycopeptide **25** was elongated in the microwave-assisted synthesizer and the deprotection/coupling set wee the following:

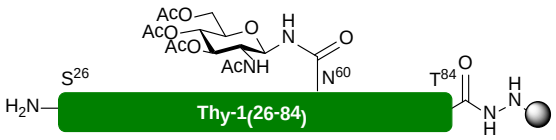
Position	A.A.	Cycle
22	Fmoc- Glu(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*
21	Fmoc-His(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*
20	Fmoc-Arg(Pbf)-OH	*0.10 - Double Arg Coupling*
19	Fmoc-Cys(Acm)-OH	*0.10-Double Deprotection Double Coupling 50°C*

A solution of 20% piperidine in DMF v/v with 0.7% of FA (4 ml) was added to the Fmoc-protected glycopeptide-resin **6a** and the slurry was shaken at r.t. for 20 min. The resin was then washed with DMF (4 X 4 ml), and DCM (4 X 4 ml) and dried over air. Reagent K (7ml) was used to cleaved the glycopeptide **6a** from the solid support and the crude (41 mg) was purified by semi-preparative RP-HPLC on Synergi 4µm Fusion RP 80 Å (250 x 21.2 mm) column (gradient 10 to 50% of ACN in water + 0.1% of formic acid in 35 minutes, flow rate 8 ml/min).

Yield: 9.2 mg (1% compared to the calculated loading and the amount of resin-glycopeptide **22** used, 22% compared to the crude obtained). $C_{371}H_{566}N_{98}O_{122}S_2$. ESI-MS (*m/z*): $[M+H]^+_{cald}$ = 8410.05. M_{obsd} = 1202.5425 $[M+7H]^{7+}$, 1052.3486 $[M+8H]^{8+}$, 935.5278 $[M+9H]^{9+}$, 842.0767 $[M+10H]^{10+}$, 765.7048 $[M+11H]^{11+}$.

4.5.3 Synthesis of Thy-1 Peptide Fragment 6b

*H-Ser(OtBu)-Ser(OtBu)-Pro-Ile-Gln(Trt)-Tyr(OtBu)-Glu(OtBu)-Phe-Ser(OtBu)-Leu-Thr(OtBu)-Arg(Pbf)-Glu(OtBu)-Thr(OtBu)-Lys(Boc)-Lys(Boc)-His(Trt)-Val-Leu-Phe-Gly-Thr(OtBu)-Val-Gly-Val-Pro-Glu(OtBu)-His(Trt)-Thr(OtBu)-Tyr(OtBu)-Arg(Pbf)-Ser(OtBu)-Arg(Pbf)-Thr(OtBu)-Asn((OAc)$_3$-GlcNAc)-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Tyr(OtBu)-Asn(Trt)-Met-Lys(Boc)-Val-Leu-Tyr(OtBu)-Leu-Ser(OtBu)-Ala-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Asp(OtBu)-Glu(OtBu)-Gly-Thr(OtBu)-Tyr(OtBu)-Thr(OtBu)-NH- NH-resin (**26**)*



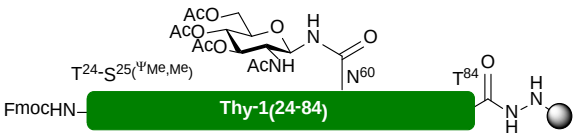
The glycopeptide-resin **23** (100 mg, originally prepared from 163.3 mg of hydrazine functionalized Wang resin, 0.1 mmol) was elongated into the synthesizer for the following amino acids:

Position	A.A.	Cycle
59	Fmoc-Thr(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*
58	Fmoc-Arg(Pbf)-OH	*0.10 - Double Arg Coupling*
57	Fmoc-Ser(OtBu)-OH	*0.10-Single 75°C Coupling*
56	Fmoc-Arg(Pbf)-OH	*0.10 - Double Arg Coupling*
55	Fmoc-Tyr(OtBu)-OH	*0.10-Single 75°C Coupling*
54	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
53	Fmoc-His(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*
52	Fmoc-Glu(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*
51	Fmoc-Pro-OH	*0.10-Single 75°C Coupling*
50	Fmoc-Val-OH	*0.10-Single Deprotection Double Coupling 75°C*
49	Fmoc-Gly-OH	*0.10-Single 75°C Coupling*
48	Fmoc-Val-OH	*0.10- Single Deprotection Double Coupling 75°C*
47	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
46	Fmoc-Gly-OH	*0.10-Single 75°C Coupling*
45	Fmoc-Phe-OH	*0.10- Single Deprotection Double Coupling 75°C*
44	Fmoc-Leu-OH	*0.10-Single 75°C Coupling*
43	Fmoc-Val-OH	*0.10- Single Deprotection Double Coupling 75°C*
42	Fmoc-His(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*
41	Fmoc-Lys(Boc)-OH	*0.10-Single 75°C Coupling*
40	Fmoc-Lys(Boc)-OH	*0.10- Single Deprotection Double Coupling 75°C*
39	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
38	Fmoc-Glu(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*

37	Fmoc-Arg(Pbf)-OH	*0.10 - Double Arg Coupling*
36	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
35	Fmoc-Leu-OH	*0.10-Single 75°C Coupling*
34	Fmoc-Ser(OtBu)-OH	*0.10-Single 75°C Coupling*
33	Fmoc-Phe-OH	*0.10- Single Deprotection Double Coupling 75°C*
32	Fmoc-Glu(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*
31	Fmoc-Tyr(OtBu)-OH	*0.10-Single 75°C Coupling*
30	Fmoc-Gln(Boc)-OH	*0.10-Single 75°C Coupling*
29	Fmoc-Ile-OH	*0.10-Double Deprotection Double Coupling 75°C*
28	Fmoc-Pro-OH	*0.10-Single 75°C Coupling*
27	Fmoc-Ser(OtBu)-OH	*0.10-Single 75°C Coupling*
26	Fmoc-Ser(OtBu)-OH	*0.10- Single Deprotection Double Coupling 75°C*

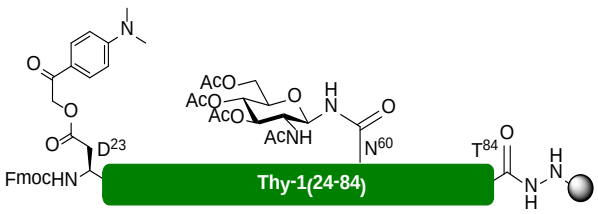
After the Fmoc removal, a cleavage was performed on 2 mg circa of resin using the 250 µl of the Reagent K for 2 h at r.t. The deprotected peptide was washed 4 times with cold diethyl ether and analyzed with RP-UPLC ESI-QTof system using a Acquity UPLC BEH C18 column (gradient from 10 to 50% of ACN in water + 0.1% of formic acid in 23 minutes, 0.95 ml/min). $C_{323}H_{493}N_{83}O_{101}S$ calculated for the peptide *H-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OAc)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH$_2$*. ESI MS (*m/z*): $[M+H]^+_{cald}$ = 7183.58. M_{obsd} = 1437.6790 $[M+5H]^{5+}$, 1198.0635 $[M+6H]^{6+}$, 1027.0508 $[M+7H]^{7+}$, 898.9245 $[M+8H]^{8+}$, 799.2664 $[M+9H]^{9+}$.

Fmoc-Thr(OtBu)-Ser($\Psi^{Me,Me}$Pro)-*Ser(OtBu)-Ser(OtBu)-Pro-Ile-Gln(Trt)-Tyr(OtBu)-Glu(OtBu)-Phe-Ser(OtBu)-Leu-Thr(OtBu)-Arg(Pbf)-Glu(OtBu)-Thr(OtBu)-Lys(Boc)-Lys(Boc)-His(Trt)-Val-Leu-Phe-Gly-Thr(OtBu)-Val-Gly-Val-Pro-Glu(OtBu)-His(Trt)-Thr(OtBu)-Tyr(OtBu)-Arg(Pbf)-Ser(OtBu)-Arg(Pbf)-Thr(OtBu)-Asn((OAc)$_3$-GlcNAc)-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Tyr(OtBu)-Asn(Trt)-Met-Lys(Boc)-Val-Leu-Tyr(OtBu)-Leu-Ser(OtBu)-Ala-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Asp(OtBu)-Glu(OtBu)-Gly-Thr(OtBu)-Tyr(OtBu)-Thr(OtBu)-NH-NH-resin (**27**)*



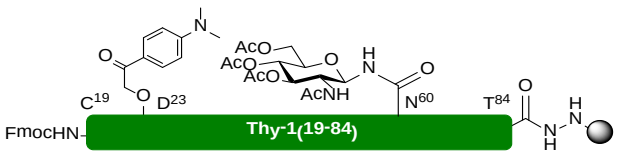
The pseudoproline was coupled manually. A solution of **3** (165.3 mg, 0.32 mmol), PyBop (163.9 mg, 0.32 mmol), and DIPEA (77 µl, 0.44 mmol) in DMF (3 ml) was prepared and added to the resin. The mixture was shaken at r.t. overnight. After the removal of the solvent, the resin was washed with DMF (4 X 4 ml), and DCM (4 X 4 ml), and a small amount was taken and submitted to the small-scale cleavage. The efficacy of the coupling was checked on Acquity UPLC BEH C18 column (gradient from 10 to 40% of ACN in water + 0.1% of formic acid in 23 minutes, 0.95 ml/min). $C_{330}H_{506}N_{85}O_{105}S$ calculated for the peptide *H-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OAc)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH$_2$*. MS (*m/z*): $[M+H]^+_{cald}$ = 7371.66. M_{obsd} = 1475.2667 $[M+5H]^{5+}$, 1229.5546 $[M+6H]^{6+}$, 1054.0427 $[M+7H]^{7+}$, 922.4169 $[M+8H]^{8+}$, 820.1439 $[M+9H]^{9+}$.

Fmoc-Asp(OMap)-Thr(OtBu)-Ser($\Psi^{Me,Me}$Pro)-*Ser(OtBu)-Ser(OtBu)-Pro-Ile-Gln(Trt)-Tyr(OtBu)-Glu(OtBu)-Phe-Ser(OtBu)-Leu-Thr(OtBu)-Arg(Pbf)-Glu(OtBu)-Thr(OtBu)-Lys(Boc)-Lys(Boc)-His(Trt)-Val-Leu-Phe-Gly-Thr(OtBu)-Val-Gly-Val-Pro-Glu(OtBu)-His(Trt)-Thr(OtBu)-Tyr(OtBu)-Arg(Pbf)-Ser(OtBu)-Arg(Pbf)-Thr(OtBu)-Asn((OAc)$_3$-GlcNAc)-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Tyr(OtBu)-Asn(Trt)-Met-Lys(Boc)-Val-Leu-Tyr(OtBu)-Leu-Ser(OtBu)-Ala-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Asp(OtBu)-Glu(OtBu)-Gly-Thr(OtBu)-Tyr(OtBu)-Thr(OtBu)-NH- NH-resin (**28**)*



The Fmoc-Asp(OMap)-OH was coupled manually. A solution of **4** (162.5 mg, 0.32 mmol), PyBop (163.9 mg, 0.32 mmol), and DIPEA (77 µl, 0.44 mmol) in DMF (3 ml) was prepared and added to the resin. The mixture was shaken at r.t. overnight. After the removal of the solvent, the resin was washed with DMF (4 X 4 ml), and DCM (4 X 4 ml), and a small amount was taken and submitted to the small-scale cleavage. The efficacy of the coupling was checked on Acquity UPLC BEH C18 column (gradient from 10 to 40% of ACN in water + 0.1% of formic acid in 23 minutes, 0.95 ml/min). $C_{344}H_{521}N_{87}O_{109}S$ calculated for the peptide *Fmoc-Asp(OMap)-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OAc)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH$_2$*. MS (*m/z*): $[M+H]^+_{cald}$ = 7647.77. M_{obsd} = 1275.5076 $[M+6H]^{6+}$, 1093.5076 $[M+7H]^{7+}$, 956.9454 $[M+8H]^{8+}$, 850.7391 $[M+9H]^{9+}$.

Fmoc-Cys(Acm)-Arg(Pbf)-His(Trt)-Glu(OtBu)-Asp(OMap)-Thr(OtBu)-Ser($\Psi^{Me,Me}$Pro)-*Ser(OtBu)-Ser(OtBu)-Pro-Ile-Gln(Trt)-Tyr(OtBu)-Glu(OtBu)-Phe-Ser(OtBu)-Leu-Thr(OtBu)-Arg(Pbf)-Glu(OtBu)-Thr(OtBu)-Lys(Boc)-Lys(Boc)-His(Trt)-Val-Leu-Phe-Gly-Thr(OtBu)-Val-Gly-Val-Pro-Glu(OtBu)-His(Trt)-Thr(OtBu)-Tyr(OtBu)-Arg(Pbf)-Ser(OtBu)-Arg(Pbf)-Thr(OtBu)-Asn((OAc)$_3$-GlcNAc)-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Tyr(OtBu)-Asn(Trt)-Met-Lys(Boc)-Val-Leu-Tyr(OtBu)-Leu-Ser(OtBu)-Ala-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Asp(OtBu)-Glu(OtBu)-Gly-Thr(OtBu)-Tyr(OtBu)-Thr(OtBu)-NH- NH-resin (**29**)*

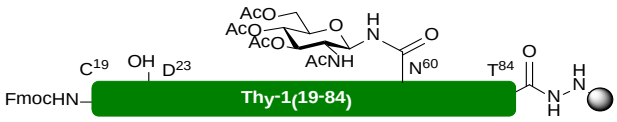


The glycopeptide-resin **28** was elongated in the microwave-assisted synthesizer and the deprotection/coupling set wee the following:

Position	A.A.	Cycle
22	Fmoc- Glu(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*
21	Fmoc-His(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*
20	Fmoc-Arg(Pbf)-OH	*0.10 - Double Arg Coupling*
19	Fmoc-Cys(Acm)-OH	*0.10-Double Deprotection Double Coupling 50°C*

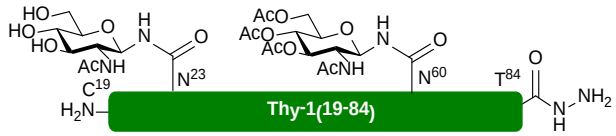
Few beads of the Fmoc-protected glycopeptide **29** were submitted to the small-scale cleavage. The efficacy of the coupling was checked on Acquity UPLC BEH C18 column (gradient from 10 to 40% of ACN in water + 0.1% of formic acid in 23 minutes, 0.95 ml/min). $C_{382}H_{569}N_{97}O_{118}S_2$ calculated for the peptide *Fmoc-Cys(Acm)-Arg-His-Glu-Asp(OMap)-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OAc)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH$_2$*. MS (*m/z*): $[M+H]^+_{cald}$ = 8467.09. M_{obsd} = 1649.8464 $[M+5H]^{5+}$, 1375.0892 $[M+6H]^{6+}$, 1178.8195 $[M+7H]^{7+}$, 1031.4572 $[M+8H]^{8+}$, 917.0862 $[M+9H]^{9+}$, 825.4831 $[M+10H]^{10+}$, 750.5129 $[M+11H]^{11+}$.

Fmoc-Cys(Acm)-Arg(Pbf)-His(Trt)-Glu(OtBu)-Asp(OH)-Thr(OtBu)-Ser($\Psi^{Me,Me}$Pro)-*Ser(OtBu)-Ser(OtBu)-Pro-Ile-Gln(Trt)-Tyr(OtBu)-Glu(OtBu)-Phe-Ser(OtBu)-Leu-Thr(OtBu)-Arg(Pbf)-Glu(OtBu)-Thr(OtBu)-Lys(Boc)-Lys(Boc)-His(Trt)-Val-Leu-Phe-Gly-Thr(OtBu)-Val-Gly-Val-Pro-Glu(OtBu)-His(Trt)-Thr(OtBu)-Tyr(OtBu)-Arg(Pbf)-Ser(OtBu)-Arg(Pbf)-Thr(OtBu)-Asn((OAc)$_3$-GlcNAc)-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Tyr(OtBu)-Asn(Trt)-Met-Lys(Boc)-Val-Leu-Tyr(OtBu)-Leu-Ser(OtBu)-Ala-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Asp(OtBu)-Glu(OtBu)-Gly-Thr(OtBu)-Tyr(OtBu)-Thr(OtBu)-NH- NH-resin (**30**)*



The Fmoc-protected glycopeptide-resin **29** was swollen with 4 ml of DCM for 30 min in a 10 ml polypropylene syringe equipped with a frit. After the removal of the solvent, the wet resin was transferred to a 20 ml polypropylene syringe using a solution of EtOH/DCM 1:1 (18 ml). The slurry was pushed through a 40 cm loop into a UV-Vis reactor core during a total time of 25 min. The resin was then irradiated at 360 nm for other 25 min. Additional 20 ml solution of EtOH/DCM 1:1 allowed the expulsion of the resin from the reactor and its collection in a 10 ml polypropylene syringe equipped with a frit. After the draining of the solvent, the resin was ultimately washed with DCM (4 X 4 ml). Few beads of the photo-deprotected glycopeptide **30** were treated with reagent K for 2.5 h at r.t. The cleaved glycopeptide was analyzed by RP-UPLC ESI-QTof system on Acquity UPLC BEH C18 column (gradient from 10 to 40% of ACN in water + 0.1% of formic acid in 14 minutes, 0.95 ml/min). $C_{372}H_{558}N_{96}O_{117}S_2$ calculated for the peptide *Fmoc-Cys(Acm)-Arg-His-Glu-Asp-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OAc)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH$_2$*. ESI-MS (*m/z*): $[M+H]^+_{cald}$ = 8306.01. M_{obsd} = 1385.0258$[M+6H]^{6+}$, 1187.3263 $[M+7H]^{7+}$, 1039.0980 $[M+8H]^{8+}$, 923.7493 $[M+9H]^{9+}$, 831.4466 $[M+10H]^{10+}$, 755.9205 $[M+11H]^{11+}$.

H-Cys(Acm)-Arg-His-Glu-Asn((OH)$_3$GlcNAc)-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OAc)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH$_2$ **(6b)**

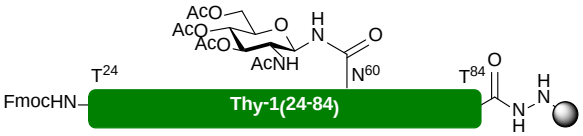


The Fmoc-protected glycopeptide-resin **30** was swollen with 4 ml of DCM for 30 min in a 10 ml polypropylene syringe equipped with a frit. After the removal of the solvent, a solution of **5** (28 mg, 0.13 mmol), PyBop (66 mg, 0.13 mmol), and DIPEA (44 µl, 0.25 mmol) in DMF was added to the reactor vessel and the mixture was gently shaken for 16 h at r.t. Finally the solvent was removed, the resin was washed with DMF (4 X 4 ml), and DCM (4 X 4 ml). A solution of 20% piperidine in DMF v/v with 0.7% of FA (4 ml) was added and the slurry was shaken at r.t. for 20 min. The resin was then washed with DMF (4 X 4 ml), and DCM (4 X 4 ml) and dried over air. Reagent K (7 ml) was used to cleaved the glycopeptide **6b** from the solid support and the crude (74.6 mg) was purified by semi-preparative RP-HPLC on Synergi 4µm Fusion RP 80 Å (250 x 21.2 mm) column (gradient 10 to 50% of ACN in water + 0.1% of formic acid in 35 minutes, flow rate 8 ml/min).

Yield: 9.2 mg (2% compared to the calculated loading and the amount of resin-glycopeptide **23** used, 12% compared to the crude obtained). $C_{365}H_{561}N_{98}O_{119}S_2$. ESI-MS (*m/z*): $[M+H]^+_{cald}$ = 8285.03. M_{obsd} = 1381.8568 $[M+6H]^{6+}$, 1184.4501 $[M+7H]^{7+}$, 1036.6389 $[M+8H]^{8+}$, 921.4635 $[M+9H]^{9+}$, 829.4126 $[M+10H]^{10+}$, 754.1860 $[M+11H]^{11+}$; 1367.3530 $[M-84+6H]^{6+}$, 1172.1519 $[M-84+7H]^{7+}$, 1025.7644 $[M-84+8H]^{8+}$, 911.7928 $[M-84+9H]^{9+}$, 820.6080 $[M-84+10H]^{10+}$, 745.9812 $[M-84+11H]^{11+}$.

4.5.4 Synthesis of Thy-1 Peptide Fragment 6c

Fmoc-Thr(OtBu)-Ser(OtBu)-Ser(OtBu)-Ser(OtBu)-Pro-Ile-Gln(Trt)-Tyr(OtBu)-Glu(OtBu)-Phe-Ser(OtBu)-Leu-Thr(OtBu)-Arg(Pbf)-Glu(OtBu)-Thr(OtBu)-Lys(Boc)-Lys(Boc)-His(Trt)-Val-Leu-Phe-Gly-Thr(OtBu)-Val-Gly-Val-Pro-Glu(OtBu)-His(Trt)-Thr(OtBu)-Tyr(OtBu)-Arg(Pbf)-Ser(OtBu)-Arg(Pbf)-Thr(OtBu)-Asn((OAc)$_3$-GlcNAc)-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Tyr(OtBu)-Asn(Trt)-Met-Lys(Boc)-Val-Leu-Tyr(OtBu)-Leu-Ser(OtBu)-Ala-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Asp(OtBu)-Glu(OtBu)-Gly-Thr(OtBu)-Tyr(OtBu)-Thr(OtBu)-NH-NH-resin ***(31)***



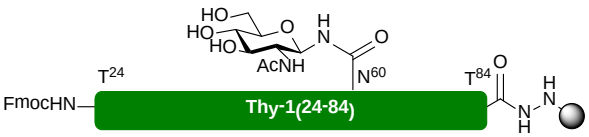
The glycopeptide-resin **23** (151.2 mg, originally prepared from 163.3 mg of hydrazine functionalized Wang resin, 0.1 mmol) was elongated into the synthesizer for the following amino acids:

Position	A.A.	Cycle
59	Fmoc-Thr(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*
58	Fmoc-Arg(Pbf)-OH	*0.10 - Double Arg Coupling*
57	Fmoc-Ser(OtBu)-OH	*0.10-Single 75°C Coupling*
56	Fmoc-Arg(Pbf)-OH	*0.10 - Double Arg Coupling*
55	Fmoc-Tyr(OtBu)-OH	*0.10-Single 75°C Coupling*
54	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
53	Fmoc-His(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*
52	Fmoc-Glu(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*
51	Fmoc-Pro-OH	*0.10-Single 75°C Coupling*
50	Fmoc-Val-OH	*0.10-Single Deprotection Double Coupling 75°C*
49	Fmoc-Gly-OH	*0.10-Single 75°C Coupling (HS)*
48	Fmoc-Val-OH	*0.10- Single Deprotection Double Coupling 75°C*
47	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
46	Fmoc-Gly-OH	*0.10-Single 75°C Coupling*
45	Fmoc-Phe-OH	*0.10- Single Deprotection Double Coupling 75°C*
44	Fmoc-Leu-OH	*0.10-Single 75°C Coupling*
43	Fmoc-Val-OH	*0.10- Single Deprotection Double Coupling 75°C*
42	Fmoc-His(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*
41	Fmoc-Lys(Boc)-OH	*0.10-Single 75°C Coupling*
40	Fmoc-Lys(Boc)-OH	*0.10- Single Deprotection Double Coupling 75°C*

39	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
38	Fmoc-Glu(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*
37	Fmoc-Arg(Pbf)-OH	*0.10 - Double Arg Coupling*
36	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
35	Fmoc-Leu-OH	*0.10-Single 75°C Coupling*
34	Fmoc-Ser(OtBu)-OH	*0.10-Single 75°C Coupling*
33	Fmoc-Phe-OH	*0.10- Single Deprotection Double Coupling 75°C*
32	Fmoc-Glu(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*
31	Fmoc-Tyr(OtBu)-OH	*0.10-Single 75°C Coupling*
30	Fmoc-Gln(Boc)-OH	*0.10-Single 75°C Coupling*
29	Fmoc-Ile-OH	*0.10-Double Deprotection Double Coupling 75°C*
28	Fmoc-Pro-OH	*0.10-Single 75°C Coupling*
27	Fmoc-Ser(OtBu)-OH	*0.10-Single 75°C Coupling*
26	Fmoc-Ser(OtBu)-OH	*0.10- Single Deprotection Double Coupling 75°C*
25	Fmoc-Ser(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*
24	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*

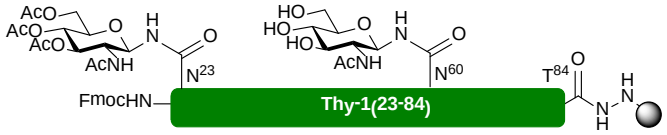
A small amount of resin submitted to cleavage with 250 µl of the Reagent K for 2 h at r.t. The glycopeptide **31** was analyzed with RP-UPLC ESI-QTof system using a Acquity UPLC BEH C18 column (gradient from 10 to 40% of ACN in water + 0.1% of formic acid in 23 minutes, 0.95 ml/min). $C_{345}H_{515}N_{85}O_{107}S$ calculated for the peptide *H-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OAc)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH$_2$*. MS (*m/z*): $[M+H]^+_{cald}$ = 7596.74. M_{obsd} = 1520.3555 $[M+5H]^{5+}$, 1266.8344 $[M+6H]^{6+}$, 1086.0321 $[M+7H]^{7+}$, 950.4299 $[M+8H]^{8+}$, 845.0830 $[M+9H]^{9+}$.

Fmoc-Thr(OtBu)-Ser(OtBu)-Ser(OtBu)-Ser(OtBu)-Pro-Ile-Gln(Trt)-Tyr(OtBu)-Glu(OtBu)-Phe-Ser(OtBu)-Leu-Thr(OtBu)-Arg(Pbf)-Glu(OtBu)-Thr(OtBu)-Lys(Boc)-Lys(Boc)-His(Trt)-Val-Leu-Phe-Gly-Thr(OtBu)-Val-Gly-Val-Pro-Glu(OtBu)-His(Trt)-Thr(OtBu)-Tyr(OtBu)-Arg(Pbf)-Ser(OtBu)-Arg(Pbf)-Thr(OtBu)-Asn((OH)$_3$-GlcNAc)-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Tyr(OtBu)-Asn(Trt)-Met-Lys(Boc)-Val-Leu-Tyr(OtBu)-Leu-Ser(OtBu)-Ala-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Asp(OtBu)-Glu(OtBu)-Gly-Thr(OtBu)-Tyr(OtBu)-Thr(OtBu)-NH-NH-resin ***(32)***



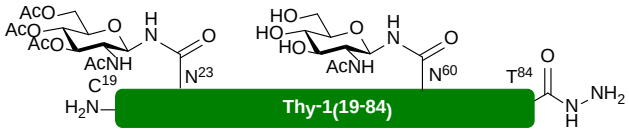
The glycopeptide-resin **31** was transferred into a 25 ml polypropylene disposable syringe equipped with a septum. A solution of 5% hydrazine monohydrate in MeOH/THF (20 ml) was added and the mixture was shaken at r.t. for 4h. The solvent was drained and the resin was washed with DMF (4 X 4 ml), and DCM (4 X 4 ml). The Fmoc-protecting group was then removed with 4 ml of piperidine 20% in DMF + 0.7% FA for 20 min at r.t. Few beads were submitted to cleavage and the resulting glycopeptide **32** was analysed by RP-UPLC ESI-QTof system using a Acquity UPLC BEH C18 column (gradient from 10 to 40% of ACN in water + 0.1% of formic acid in 23 minutes, 0.95 ml/min). $C_{324}H_{500}N_{85}O_{102}S$ calculated for the peptide *H-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OH)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH$_2$*. MS (*m/z*): $[M+H]^+_{cald}$ = 7245.63. M_{obsd} = 1449.9825 $[M+5H]^{5+}$, 1208.4348 $[M+6H]^{6+}$, 1036.1250 $[M+7H]^{7+}$, 906.6407 $[M+8H]^{8+}$, 806.1406 $[M+9H]^{9+}$.

*Fmoc-Asn((OAc)$_3$-GlcNAc)-Thr(OtBu)-*Ser*(OtBu)-Ser(OtBu)-Ser(OtBu)-Pro-Ile-Gln(Trt)-Tyr(OtBu)-Glu(OtBu)-Phe-Ser(OtBu)-Leu-Thr(OtBu)-Arg(Pbf)-Glu(OtBu)-Thr(OtBu)-Lys(Boc)-Lys(Boc)-His(Trt)-Val-Leu-Phe-Gly-Thr(OtBu)-Val-Gly-Val-Pro-Glu(OtBu)-His(Trt)-Thr(OtBu)-Tyr(OtBu)-Arg(Pbf)-Ser(OtBu)-Arg(Pbf)-Thr(OtBu)-Asn((OH)$_3$-GlcNAc)-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Tyr(OtBu)-Asn(Trt)-Met-Lys(Boc)-Val-Leu-Tyr(OtBu)-Leu-Ser(OtBu)-Ala-Phe-Thr(OtBu)-Ser(OtBu)-Lys(Boc)-Asp(OtBu)-Glu(OtBu)-Gly-Thr(OtBu)-Tyr(OtBu)-Thr(OtBu)-NH-NH-resin (**33**)*



The glycosylated amino acid was coupled manually. A solution of **2** (208.3 mg, 0.3 mmol), PyBop (156 mg, 0.3 mmol), and DIPEA (104.5 µl, 0.6 mmol) in DMF (3 ml) and DMSO (6 µl) was prepared and added to the resin. The mixture was shaken at r.t. overnight. After the removal of the solvent, the resin was washed with DMF (4 X 4 ml), and DCM (4 X 4 ml). Few dry beads were treated with reagent K at r.t. for 2.5 h and the cleaved glycopeptide **33** was analyzed on RP-UPLC ESI-QTof system using a Acquity UPLC BEH C18 column (gradient from 10 to 40% of ACN in water + 0.1% of formic acid in 23 minutes, 0.95 ml/min). $C_{342}H_{525}N_{88}O_{112}S$ calculated for the peptide *Fmoc-Asn((OAc)$_3$-GlcNAc)-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OH)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH$_2$*. MS (*m/z*): $[M+H]^+_{cald}$ = 7910.85. M_{obsd} = 1583.0147 $[M+5H]^{5+}$, 1319.4601 $[M+6H]^{6+}$, 1131.1501 $[M+7H]^{7+}$, 989.9119 $[M+8H]^{8+}$, 880.0615 $[M+9H]^{9+}$.

H-Cys(Acm)-Arg-His-Glu-Asn((OAc)$_3$GlcNAc)-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OH)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH$_2$ (**6c**)



The glycopeptide-resin **35** was elongated in the microwave-assisted synthesizer and the deprotection/coupling were set a follows:

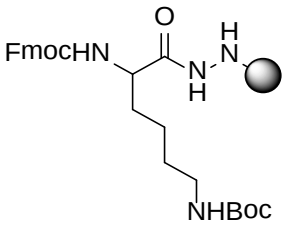
Position	A.A.	Cycle
22	Fmoc- Glu(OtBu)-OH	*0.10-Double Deprotection Double Coupling 75°C*
21	Fmoc-His(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*
20	Fmoc-Arg(Pbf)-OH	*0.10 - Double Arg Coupling*
19	Fmoc-Cys(Acm)-OH	*0.10-Double Deprotection Double Coupling 50°C*

After the removal of Fmoc-protecting group at N-*terminus*, the glycopeptide **6c** was cleaved from solid support with reagent K for 2.5 h at r.t. and the crude (52.3 mg) was purified by semi-preparative RP-HPLC on Synergi 4µm Fusion RP 80 Å (250 x 21.2 mm) column (gradient 10 to 50% of ACN in water + 0.1% of formic acid in 35 minutes, flow rate 8 ml/min).

Yield: 6 mg (1% compared to the calculated loading and the amount of resin-glycopeptide **23** used, 11% compared to the crude obtained). $C_{365}H_{561}N_{98}O_{119}S_2$. ESI-MS (*m/z*): $[M+H]^+_{cald}$ = 8285.03. M_{obsd} = 1184.6833 $[M+7H]^{7+}$, 1036.6187 $[M+8H]^{8+}$, 921.5638 $[M+9H]^{9+}$, 829.5348 $[M+10H]^{10+}$.

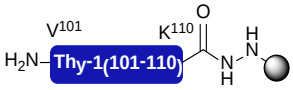
4.5.5 Synthesis of Thy-1 Peptide Fragments 37a-b

Fmoc-Thr(OtBu)-NH-NH-Resin (**34**)



The synthesis of this peptide fragment was accomplished using functionalized NovaPEG Wang resin **8** (123 mg, 0.08 mmol) loaded with Fmoc-Lys(Boc)-OH (110 mg, 0.23 mmol) as described before. The loading calculated was 0.49 mmol/g (96% of coupled sites compared to the theoretical loading). The resin was capped with 5 ml solution of DMF/Ac_2O/DIPEA (4/0.5/0.5) for 10 min and then washed with DMF (4 X 4 ml), and DCM (4 X 4 ml).

Fmoc-Val-Thr(OtBu)-Val-Leu-Arg(Pbf)-Asp(OtBu)-Lys(Boc)-Leu-Val-Lys(Boc)-NH-NH-resin (**35**)



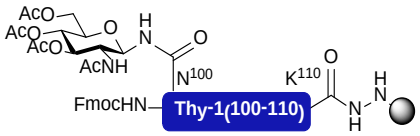
The peptide was elongated in the synthesizer for the first 9 amino acids using the following deprotection/coupling cycles:

Position	A.A.	Cycle
119	Fmoc-Val-OH	*0.10-Double Deprotection Double Coupling 75°C*
118	Fmoc-Leu-OH	*0.10-Single 75°C Coupling*
117	Fmoc-Lys(Boc)-OH	*0.10-Single 75°C Coupling*
116	Fmoc-Asp(OtBu)-OH	*0.10-Single 75°C Coupling*
115	Fmoc-Arg(Pbf)-OH	*0.10 - Double Arg Coupling*
114	Fmoc-Leu-OH	*0.10-Single 75°C Coupling*
113	Fmoc-Val-OH	*0.10-Single Deprotection Double Coupling 75°C*
112	Fmoc-Thr(OtBu)-OH	*0.10-Single 75°C Coupling*
111	Fmoc-Val-OH	*0.10-Single Deprotection Double Coupling 75°C*

A small amount of resin submitted to cleavage with 250 µl of the Reagent K for 2 h at r.t. The peptide **35** was analyzed with RP-UPLC ESI-QTof system using a Acquity UPLC BEH C18

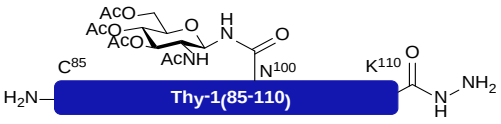
column (gradient from 5 to 70% of ACN in water + 0.1% of formic acid in 6 minutes, 0.95 ml/min). $C_{68}H_{112}N_{17}O_{15}$ calculated for the peptide *H-Val-Thr-Val-Leu-Arg-Asp-Lys-Leu-Val-Lys-NH-NH₂*. MS (*m/z*): $[M+H]^+_{cald}$ = 1406.85. M_{obsd} = 1406.93 $[M+H]^+$, 703.9730 $[M+2H]^{2+}$, 592.9284 $[M+3H]^{3+}$.

Fmoc-Asn((OAc)₃GlcNAc)-Val-Thr(OtBu)-Val-Leu-Arg(Pbf)-Asp(OtBu)-Lys(Boc)-Leu-Val-Lys(Boc)-NH-NH-resin (**36**)



The glycosylated amino acid was coupled manually. A solution of **2** (157 mg, 0.23 mmol), PyBop (119.7 mg, 0.23 mmol), and DIPEA (80 µl, 0.46 mmol) in DMF (3 ml) and DMSO (6 µl) was prepared and added to the resin. The mixture was shaken at r.t. overnight. After the removal of the solvent, the resin was washed with DMF (4 X 4 ml), and DCM (4 X 4 ml). Few dry beads were treated with reagent K at r.t. for 2.5 h and the cleaved glycopeptide **39** was analyzed on RP-UPLC ESI-QTof system using a Acquity UPLC BEH C18 column (gradient from 5 to 70% of ACN in water + 0.1% of formic acid in 6 minutes, 0.95 ml/min). $C_{86}H_{137}N_{20}O_{25}$ calculated for the peptide *Fmoc-Asn((OAc)₃GlcNAc)-Val-Thr-Val-Leu-Arg-Asp-Lys-Leu-Val-Lys-NH-NH₂*. MS (*m/z*): $[M+H]^+_{cald}$ = 1850.01. M_{obsd} = 850.1273 $[M+H]^+$, 925.5729 $[M+2H]^{2+}$, 617.3857 $[M+3H]^{3+}$.

H-Cys-Ala-Leu-His-His-Ser-Gly-His-Ser-Pro-Pro-Ile-Ser-Ser-Gln-Asn((OAc)₃GlcNAc)-Val-Thr-Val-Leu-Arg-Asp-Lys-Leu-Val-Lys-NH-NH₂ (**37a**)



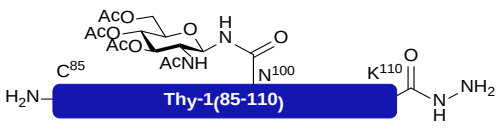
The glycopeptide-resin **36** was elongated into the synthesizer for the following amino acids and piperidine 20% in DMF (v/v) was used as deprotection cocktail:

Position	A.A.	Cycle
99	Fmoc-Gln(Boc)-OH	*0.10-Double Deprotection Double Coupling 75°C*
98	Fmoc-Ser(OtBu)-OH	*0.10-Single 75°C Coupling*
97	Fmoc-Ser(OtBu)-OH	*0.10-Single Deprotection Double Coupling 75°C*
96	Fmoc-Ile-OH	*0.10-Double Deprotection Double Coupling 50°C*
95	Fmoc-Pro-OH	*0.10-Single 75°C Coupling*
94	Fmoc-Pro-OH	*0.10-Single Deprotection Double Coupling 75°C*
93	Fmoc-Ser(OtBu)-OH	*0.10-Single Deprotection Double Coupling 75°C*
92	Fmoc-His(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*
91	Fmoc-Gly-OH	*0.10-Single 75°C Coupling*
90	Fmoc-Ser(OtBu)-OH	*0.10-Single Deprotection Double Coupling 75°C*
89	Fmoc-His(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*
88	Fmoc-His(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*
87	Fmoc-Leu-OH	*0.10-Single 75°C Coupling*
86	Fmoc-Ala-OH	*0.10-Single 75°C Coupling*
85	Fmoc-Cys(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*

The resin was treated for 20 min with 20% piperidine in DMF v/v (4 ml), washed with DMF (4 X 4 ml), and DCM (4 X 4 ml), and dried over air. Reagent K (7 ml) was added to the resin and after 2 h at r.t. the glycopeptide **37a** was triturated with ethyl ether (3 X 50 ml). The crude (70 mg) was purified by semi-preparative RP-HPLC on Synergi 4µm Fusion RP 80 Å (250 x 21.2 mm) column (gradient 10 to 50% of ACN in water + 0.1% of formic acid in 35 minutes, flow rate 8 ml/min).

Yield: 10.35 mg (4% compared to the calculated loading, 21% compared to the crude obtained). $C_{136}H_{224}N_{42}O_{43}S$. ESI-MS (*m/z*): $[M+H]^{+}_{cald}$ = 3165.64. M_{obsd} = 1056.1727 $[M+3H]^{3+}$, 792.3904 $[M+4H]^{4+}$, 634.1694 $[M+5H]^{5+}$, 528.6403 $[M+6H]^{6+}$.

H-Cys(Acm)-Ala-Leu-His-His-Ser-Gly-His-Ser-Pro-Pro-Ile-Ser-Ser-Gln-Asn((OAc)$_3$GlcNAc)-Val-Thr-Val-Leu-Arg-Asp-Lys-Leu-Val-Lys-NH-NH$_2$ (**37b**)



The glycopeptide-resin **36** was elongated into the synthesizer for the following amino acids and piperidine 20% in DMF (v/v) with 0.7% of FA was used as deprotection cocktail:

Position	A.A.	Cycle
99	Fmoc-Gln(Boc)-OH	*0.10-Double Deprotection Double Coupling 75°C*
98	Fmoc-Ser(OtBu)-OH	*0.10-Single 75°C Coupling*
97	Fmoc-Ser(OtBu)-OH	*0.10-Single Deprotection Double Coupling 75°C*
96	Fmoc-Ile-OH	*0.10-Double Deprotection Double Coupling 50°C*
95	Fmoc-Pro-OH	*0.10-Single 75°C Coupling*
94	Fmoc-Pro-OH	*0.10-Single Deprotection Double Coupling 75°C*
93	Fmoc-Ser(OtBu)-OH	*0.10-Single Deprotection Double Coupling 75°C*
92	Fmoc-His(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*
91	Fmoc-Gly-OH	*0.10-Single 75°C Coupling*
90	Fmoc-Ser(OtBu)-OH	*0.10-Single Deprotection Double Coupling 75°C*
89	Fmoc-His(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*
88	Fmoc-His(Trt)-OH	*0.10-Double Deprotection Double Coupling 50°C*
87	Fmoc-Leu-OH	*0.10-Single 75°C Coupling*
86	Fmoc-Ala-OH	*0.10-Single 75°C Coupling*
85	Fmoc-Cys(Acm)-OH	*0.10-Double Deprotection Double Coupling 50°C*

The resin was treated with 20% piperidine in DMF v/v + 0.7% FA (4 ml), washed with DMF (4 X 4 ml), and DCM (4 X 4 ml), and dried over air. Reagent K (7 ml) was added to the resin and after 2 h at r.t. the glycopeptide **37b** was triturated with ethyl ether (3 X 50 ml). The crude (64 mg) was purified by semi-preparative RP-HPLC on Synergi 4µm Fusion RP 80 Å (250 x 21.2 mm) column (gradient 10 to 50% of ACN in water + 0.1% of formic acid in 35 minutes, flow rate 8 ml/min). Yield: 17.5 mg (6.7% compared to the calculated loading, 27% compared to the crude obtained). $C_{139}H_{230}N_{43}O_{44}S$. ESI-MS (*m/z*): $[M+H]^{+}_{cald}$ = 3237.68. M_{obsd} = 1619.8452 $[M+2H]^{2+}$, 1079.9020 $[M+3H]^{3+}$, 810.1732 $[M+4H]^{4+}$, 648.3425 $[M+5H]^{5+}$, 540.6264 $[M+6H]^{6+}$.

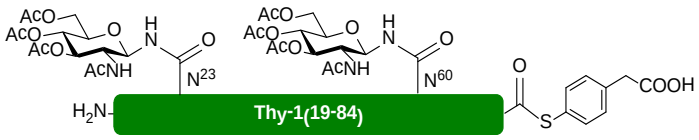
4.6 Ligation of peptide and glycopeptide fragments

The buffers were freshly prepared before every ligation reaction. The pH was checked by pH meter VWR phenomenal (VWR international GmbH) previously calibrated at pH 4.0, 7.0, and 10.0, and with pH paper. The reactions were carried out in 1.5 ml Eppendorf tubes and a small stirring magnet ensured the mixing of the reactants. The following solutions were prepared and employed:

Buffer	Volume	pH	A.A.	Conc.	Amount
A	10 ml	3.0	Gdn*HCl	6 M	5.73 g
			NaH_2PO_4*H_2O	0.2 M	0.356g
B	10 ml	7.0	Gdn*HCl	6 M	5.73 g
			NaH_2PO_4*H_2O	0.1 M	0.178 g
			MPAA	0.05 M	0.084 g
			TCEP	0.007 M	0.017g
$NaNO_2$	0.5 ml		$NaNO_2$	0.5 M	17 mg

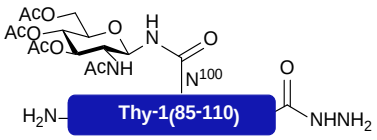
4.6.1 Synthesis of 44

H-Cys(Acm)-Arg-His-Glu-Asn((OAc)$_3$-GlcNAc)-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OAc)$_3$-GlcNAc)-Phe-Thr-Ser-Lys-Tyr-asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-S-para-(CH_2CO_2H)Ph (**38**)



Glycopeptide **6a** (4.0 mg, $4.7 \cdot 10^{-4}$ mmol) was dissolved in buffer A (317 µl, 1.5 mM) and the tube was cooled to -15°C. The $NaNO_2$ solution (9.6 µl) was added and the mixture was stirred for 20 min. Then, MPAA (8 mg, 0.048 mmol) was added, the vessel was warmed slowly to r.t. and the pH was adjusted to 6.7-6.9 with 5 M NaOH aqueous solution.

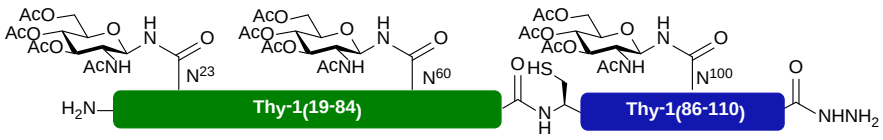
H-Cys-Ala-Leu-His-His-Ser-Gly-His-Ser-Pro-Pro-Ile-Ser-Ser-Gln-Asn((OAc)$_3$GlcNAc)-Val-Thr-Val-Leu-Arg-Asp-Lys-Leu-Val-Lys-NH-NH$_2$ (**39**)



Glycopeptide **37b**(10 mg, $3.0 \cdot 10^{-3}$ mmol) was dissolved in aqueous solution AcOH 10% v/v (0.5 ml); $Hg(OAc)_2$ (9 mg, $3.0 \cdot 10^{-2}$ mmol) was added and the mixture was stirred under N_2 atmosphere for 90 min. Next, β-mercaptoethanol (4 µl, $9.0 \cdot 10^{-2}$ mmol) was added and the reaction was mixed for additionally 60 min. The solution was centrifuged for 5 min at 3000 rpg and the supernatant was collected in a new 1.5 ml Eppendorf tube. The glycopeptide **39** was analyzed by LC-MS system and purified by semi-preparative RP-HPLC on Synergi 4µm Fusion RP 80 Å (250 x 21.2 mm) column (gradient 10 to 50% of ACN in water + 0.1% of formic acid in 35 minutes, flow rate 4 ml/min).

Yield: 3.5 mg (37%). $C_{136}H_{224}N_{42}O_{43}S$. ESI-MS (*m/z*): $[M+H]^+_{cald}$ = 3165.64. M_{obsd} = 1056.1860 $[M+3H]^{3+}$, 792.4010 $[M+4H]^{4+}$, 634.2943 $[M+5H]^{5+}$.

H-Cys(Acm)-Arg-His-Glu-Asn((OAc)$_3$GlcNAc)-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OAc)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-Cys-Ala-Leu-His-His-Ser-Gly-His-Ser-Pro-Pro-Ile-Ser-Ser-Gln-Asn((OAc)$_3$GlcNAc)-Val-Thr-Val-Leu-Arg-Asp-Lys-Leu-Val-Lys-NH-NH$_2$ (**40**)



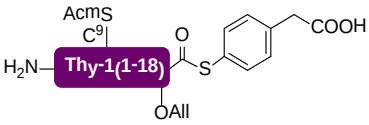
A solution of **39** (2.34 mg, $7.4 \cdot 10^{-4}$ mmol) in buffer B (100 µl) was added to the Eppendorf tube containing **38** and the mixture was stirred at r.t. for 24 h. The glycoprotein **40** was purified by FPLC using a Superdex Peptide 10/300 GL (GE Healthcare GmbH, Europe) column (eluent: Gdn*HCl 4 M in water pH 6.7, 0.5 ml/min flow rate).

Yield: Not detected. $C_{507}H_{786}N_{138}O_{165}S_3$. ESI-MS (*m/z*): $[M+H]^+_{cald}$ = 11543.65. M_{obsd} = 1468.9667 $[M+200+8H]^{8+}$, 1305.8658 $[M+200+9H]^{9+}$, 1175.3818 $[M+200+10H]^{10+}$, 1068.6237

$[M++200+11H]^{11+}$, 979.6558 $[M+200+12H]^{12+}$, 904.3669 $[M+200+13H]^{13+}$, 839.8452 $[M+200+14H]^{14+}$, 783.9282 $[M+200+15H]^{15+}$; $[M(\mathbf{38})_{hydrolized}+H]^{+}_{cald}$ = 8397.03. $[M(\mathbf{38})_{hydrolized}]_{obsd}$ = 1200.6006 $[M+7H]^{7+}$, 1050.6508 $[M+8H]^{8+}$, 934.0263 $[M+9H]^{9+}$, 840.7226 $[M+10H]^{10+}$, 764.3900 $[M+11H]^{11+}$.

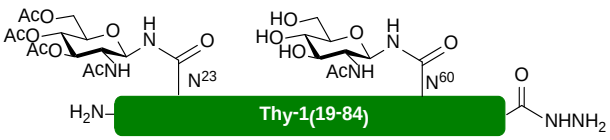
4.6.2 Synthesis of 51

H-Gln-Lys-Val-Thr-Ser-Leu-Thr-Ala-Cys(Acm)-Leu-Val-Asp-Gln-Ser-Leu-Arg-Leu-Asp(OAll)-S-para-(CH_2CO_2H)Ph (**43**)



Peptide **1** (2.0 mg, $9.4 \cdot 10^{-4}$ mmol) was dissolved in buffer A (315 µl, 3 mM) and the tube was cooled to -15°C. The $NaNO_2$ solution (18.8 µl, $9.4 \cdot 10^{-3}$) was added and the mixture was stirred for 20 min. Then, MPAA (15.8 mg, $9.4 \cdot 10^{-2}$ mmol) was added, the vessel was warmed slowly to r.t. and the pH was adjusted to 6.7-6.9 with 5 M NaOH aqueous solution.

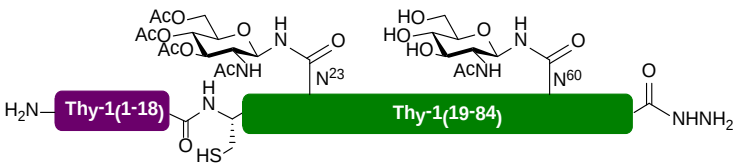
H-Cys-Arg-His-Glu-Asn((OAc)$_3$GlcNAc)-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OH)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH$_2$ (**44**)



Glycopeptide **6c** (1.92 mg, $2.3 \cdot 10^{-4}$ mmol) was dissolved in aqueous solution AcOH 10% v/v (0.2 ml); $Hg(OAc)_2$ (0.74 mg, $2.3 \cdot 10^{-3}$ mmol) was added and the mixture was stirred under N_2 atmosphere for 90 min. Next, β-mercaptoethnol (0.32 µl, $4.6 \cdot 10^{-3}$ mmol) was added and the reaction was mixed for additionally 60 min. The solution was centrifuged for 5 min at 3000 rpg and the supernatant was collected in a new 1.5 ml Eppendorf tube. The glycopeptide **44** was analyzed by LC-MS system and used without further purification.

Yield: 1.87 mg (quant.). $C_{362}H_{555}N_{97}O_{118}S_2$. ESI-MS (*m/z*): $[M+H]^+_{cald}$ = 8213.99. M_{obsd} = 1369.8610 $[M+6H]^{6+}$, 1174.3114 $[M+7H]^{7+}$, 1027.6519 $[M+8H]^{8+}$, 913.4639 $[M+9H]^{9+}$, 822.3109 $[M+10H]^{10+}$, 747.6413 $[M+11H]^{11+}$, 685.5033 $[M+12H]^{12+}$; $[M+H+200]^+_{cald}$ = 16627.98. M_{obsd} = 1848.4125 $[9M+9H+200]^+$, 1663.6184 $[10M+10H+200]^+$, 1512.5714 $[11M+11H+200]^+$, 1386.5176 $[12M+12H+200]^+$, 1280.0144 $[13M+13H+200]^+$, 1188.5851 $[14M+14H+200]^+$, 1109.4210 $[15M+15H+200]^+$, 1040.1354 $[16M+16H+200]^+$, 979.0188 $[17M+17H+200]^+$, 924.8017 $[18M+18H+200]^+$, 876.0709 $[19M+19H+200]^+$, 832.3660 $[20M+20H+200]^+$, 792.7764 $[21M+21H+200]^+$, 756.7932 $[22M+22H+200]^+$, 723.9713 $[23M+23H+200]^+$.

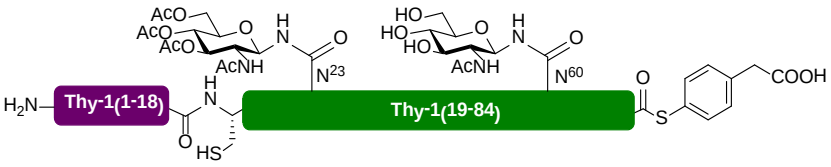
H-Gln-Lys-Val-Thr-Ser-Leu-Thr-Ala-Cys(Acm)-Leu-Val-Asp-Gln-Ser-Leu-Arg-Leu-Asp(OAll)-Cys-Arg-His-Glu-Asn((OAc)$_3$GlcNAc)-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OH)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-NH-NH$_2$ (**45**)



A solution of **44** (1.87 mg, $2.3 \cdot 10^{-4}$ mmol) in buffer B (77.9 µl) was added to the Eppendorf tube containing **43** and the mixture was stirred at r.t. for 48 h. The glycoprotein **45** was purified by FPLC using a Superdex Peptide 10/300 GL (GE Healthcare GmbH, Europe) column (eluent: Gdn*HCl 4 M in water pH 6.7, 0.5 ml/min flow rate) and concentrated with centricon 5kDa cutoff 0.5 ml Vial (Fisher Scientific GmbH, Germany).

Yield: Not detected (3 mg of product with salts). $C_{452}H_{711}N_{122}O_{147}S_3$. ESI-MS (*m/z*): $[M+H]^+_{cald}$ = 10297.12. M_{obsd} = 1287.9943 $[M+8H]^{8+}$, 1145.1024 $[M+9H]^{9+}$, 1030.7583 $[M+10H]^{10+}$, 936.9682 $[M+11H]^{11+}$, 858.9953 $[M+12H]^{12+}$, 793.0528 $[M+13H]^{13+}$, 736.4044 $[M+14H]^{14+}$.

H-Gln-Lys-Val-Thr-Ser-Leu-Thr-Ala-Cys(Acm)-Leu-Val-Asp-Gln-Ser-Leu-Arg-Leu-Asp(OAll)-Cys-Arg-His-Glu-Asn((OAc)$_3$GlcNAc)-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OH)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-S-para-(CH$_2$CO$_2$H)Ph (**46**)

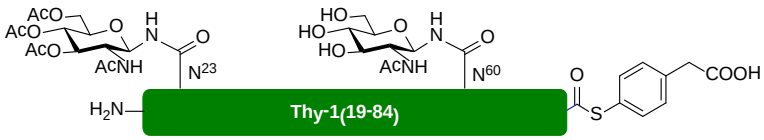


Few drops of a solution of HCl 1 M (3 µl) were added to the buffer pH 6.7 containing **45** (total volume 80 µl), and the pH was adjusted to 3.0. The mixture was cooled to -20° C and $NaNO_2$ (6 µl) was added After 20 min, the solution was warmed to r.t., MPAA (5 mg, $3.0 \cdot 10^{-2}$ mmol) was added and the pH was adjusted to 6.7- 6.9 with a 1M NaOH aqueous solution.

Yield: Not detected. $C_{460}H_{716}N_{120}O_{149}S_4$. ESI-MS (*m/z*): $[M+H]^+_{cald}$ = 10433.10. M_{obsd} = 1305.0012 $[M+8H]^{8+}$, 1160.1107 $[M+9H]^{9+}$, 1044.1739 $[M+10H]^{10+}$, 949.3423 $[M+11H]^{11+}$, 870.2581 $[M+12H]^{12+}$; 1178.7455 $[M+168+9H]^{9+}$, 1061.0665 $[M+168+10H]^{10+}$, 964.7103 $[M+168+11H]^{11+}$.

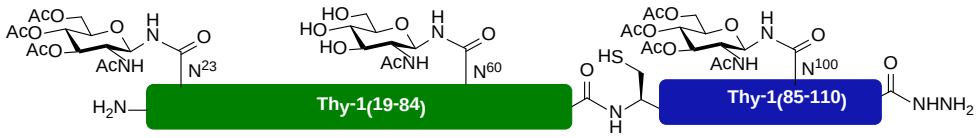
4.6.3 Synthesis of 52

H-Cys(Acm)-Arg-His-Glu-Asn((OAc)$_3$GlcNAc)-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OH)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-S-para-(CH$_2$CO$_2$H)Ph (**48**)



Glycopeptide **6c** (1.2 mg, $1.4 \cdot 10^{-4}$ mmol) was dissolved in buffer A (100 µl, 1.4 mM) and the tube was cooled to -15°C. The $NaNO_2$ solution (3 µl, $1.4 \cdot 10^{-3}$) was added and the mixture was stirred for 20 min. Then, MPAA (2.3 mg, $1.4 \cdot 10^{-2}$ mmol) was added, the vessel was warmed slowly to r.t. and the pH was adjusted to 6.7-6.9 with 5M NaOH aqueous solution.

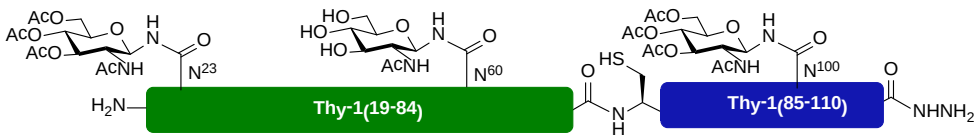
H-Cys(Acm)-Arg-His-Glu-Asn((OAc)$_3$GlcNAc)-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OH)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-Cys-Ala-Leu-His-His-Ser-Gly-His-Ser-Pro-Pro-Ile-Ser-Ser-Gln-Asn((OAc)$_3$GlcNAc)-Val-Thr-Val-Leu-Arg-Asp-Lys-Leu-Val-Lys-NH-NH$_2$ (**49**)



A solution of **39** (1.0 mg, 2.8 · 10^{-4} mmol) in buffer B (50 µl) was added to the Eppendorf tube containing **48** and the mixture was stirred at r.t. for 16 h. The glycoprotein **49** was purified by FPLC using a Superdex Peptide 10/300 GL (GE Healthcare GmbH, Europe) column (eluent: Gdn*HCl 4 M in water pH 6.7, 0.5 ml/min flow rate) and concentrated with centricon 5kDa MWCO 0.5 ml Vial (Fisher Scientific GmbH, Germany).

Yield: Not detected (2 mg of product with salts). $C_{501}H_{780}N_{138}O_{162}S_3$. ESI-MS (*m/z*): $[M+H]^+_{cald}$ = 11418.63. M_{obsd} = 1428.3601 $[M+8H]^{8+}$, 1269.6698 $[M+9H]^{9+}$, 1142.8961 $[M+10H]^{10+}$, 1039.0643 $[M+11H]^{11+}$, 952.5305 $[M+12H]^{12+}$, 879.3248 $[M+13H]^{13+}$, 816.6185 $[M+14H]^{14+}$, 762.2490 $[M+15H]^{15+}$, 714.6745 $[M+16H]^{16+}$.

H-Cys-Arg-His-Glu-Asn((OAc)$_3$GlcNAc)-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OH)$_3$GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-Cys-Ala-Leu-His-His-Ser-Gly-His-Ser-Pro-Pro-Ile-Ser-Ser-Gln-Asn((OAc)$_3$GlcNAc)-Val-Thr-Val-Leu-Arg-Asp-Lys-Leu-Val-Lys-NH-NH$_2$ (**50**)

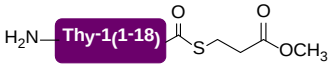


A solution of $PdCl_2$ (0.55 mg, 3.1 · 10^{-3} mmol) in 7 µl of buffer of 3 M Gdn*HCl and 0.1 M Na_2HPO_4 at pH 6.7 - 6.9 was incubated at 37°C for 30 min. Glycoprotein **49** (2.0 mg, 1.8 · 10^{-4} mmol) was dissolved in 5 µl solution of $PdCl_2$ and the mixture was stirred for 37°C for 2 h. A solution 0.1 M of DTT in water (7.7 mg in 0.5 ml of water) was prepared and added to the mixture containing **50**. The solution was centrifuged for 5 min at 3000 rpg and the supernatant was

collected in a new 1.5 ml Eppendorf tube. The glycopeptide **50** was analyzed by LC-MS system and used without further purification.

Yield: Not detected (2 mg of product with salts). $C_{498}H_{776}N_{137}O_{161}S_3$. ESI-MS (*m/z*): $[M+H]^+_{cald}$ = 11346.58. M_{obsd} = 1261.7443 $[M+9H]^{9+}$, 1135.6554 $[M+10H]^{10+}$, 1032.5443 $[M+11H]^{11+}$, 946.5756 $[M+12H]^{12+}$, 873.8135 $[M+13H]^{+13}$, 811.4780 $[M+14H]^{+14}$, 757.4639 $[M+15H]^{+15}$, 710.2094 $[M+16H]^{+16}$, 668.5049 $[M+17H]^{+17}$.

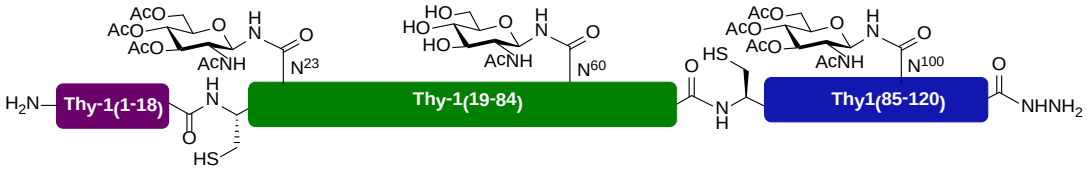
H-Gln-Lys-Val-Thr-Ser-Leu-Thr-Ala-Cys(Acm)-Leu-Val-Asp-Gln-Ser-Leu-Arg-Leu-Asp(OAll)-S-$(CH_2)_2C(O)OCH_3$ (**51**)



Peptide **1** (5.6 mg, $2.4 \cdot 10^{-3}$ mmol) was dissolved in buffer A (1.2 ml, 2 mM) and the tube was cooled to -15°C. The $NaNO_2$ solution (48 µl, $2.4 \cdot 10^{-2}$) was added and the mixture was stirred for 20 min. Then, MMP (26.6 µl, $2.4 \cdot 10^{-1}$ mmol) was added, the vessel was warmed slowly to r.t. and the pH was adjusted to 6.7-6.9 with 5 M NaOH aqueous solution. After 5 min, the solution was analyzed and purified by semi-preparative RP-HPLC on Synergi 4µm Fusion RP 80 Å (250 x 21.2 mm) column (gradient 5 to 70% of ACN in water + 0.1% of formic acid in 35 minutes, flow rate 4 ml/min).

Yield: 4.1 mg (77%). $C_{94}H_{163}N_{25}O_{31}S_2$. ESI-MS (*m/z*): $[M+H]^+_{cald}$ = 2203.15. M_{obsd} = 1102.54.3 $[M+2H]^{2+}$, 735.3730 $[M+3H]^{3+}$.

H-Gln-Lys-Val-Thr-Ser-Leu-Thr-Ala-Cys(Acm)-Leu-Val-Asp-Gln-Ser-Leu-Arg-Leu-Asp(OAll)-Cys-Arg-His-Glu-Asn((OAc)$_3$-GlcNAc)-Thr-Ser-Ser-Ser-Pro-Ile-Gln-Tyr-Glu-Phe-Ser-Leu-Thr-Arg-Glu-Thr-Lys-Lys-His-Val-Leu-Phe-Gly-Thr-Val-Gly-Val-Pro-Glu-His-Thr-Tyr-Arg-Ser-Arg-Thr-Asn((OH)$_3$-GlcNAc)-Phe-Thr-Ser-Lys-Tyr-Asn-Met-Lys-Val-Leu-Tyr-Leu-Ser-Ala-Phe-Thr-Ser-Lys-Asp-Glu-Gly-Thr-Tyr-Thr-Cys-Ala-Leu-His-His-Ser-Gly-His-Ser-Pro-Pro-Ile-Ser-Ser-Gln-Asn((Ac)$_3$GlcNac)-Val-Thr-Val-Leu-Arg-Asp-Lys-Leu-Val-Lys-NH-NH$_2$ (**52**)



Thioester **51**(0.52 mg, 2.3 · 10^{-4} mmol) was dissolved in 30 µl of buffer B and added to a solution of thiol **50** (2 mg, 1.8· 10^{-4} mmol) previously dissolved in 100 µl of buffer B. The solution was mixed at r.t. and monitored time by time by LC-MS.

Yield: Not detected. $C_{588}H_{931}N_{162}O_{190}S_4$. ESI-MS (*m/z*): $[M+H]^+_{cald}$ = 13429.71. M_{obsd} = 1493.2244 $[M+9H]^{9+}$, 1343.9519 $[M+10H]^{10+}$, 1221.8628 $[M+11H]^{11+}$, 1120.1487 $[M+12H]^{12+}$, 1033.9803 $[M+13H]^{+13}$, 960.2756 $[M+14H]^{+14}$, 896.3104 $[M+15H]^{+15}$, 840.3312 $[M+16H]^{+16}$.

References

[1] M. Karplus, J. Kuriyan, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6679–6685.

[2] R. Nygaard, J. Kim, F. Mancia, *Curr. Opin. Struct. Biol.* **2020**, *64*, 26–33.

[3] H. Zhang, R. Aebersold, *Methods Mol. Biol.* **2006**, *328*, 177–185.

[4] T. Tuller, A. Carmi, K. Vestsigian, S. Navon, Y. Dorfan, J. Zaborske, T. Pan, O. Dahan, I. Furman, Y. Pilpel, *Cell* **2010**, *141*, 344–354.

[5] A. P. Lothrop, M. P. Torres, S. M. Fuchs, *FEBS Lett.* **2013**, *587*, 1247–1257.

[6] C. T. Walsh, *Posttranslational Modification of Proteins : Expanding Nature's Inventory.*, Englewood: Roberts And Co. Publ., **2006**.

[7] L. R. Singh, T. Ali Dar, P. Ahmad, *Proteostasis Chaperone Surveill.* **2015**, 1–180.

[8] P. Gagneux, A. Varki, *Glycobiology* **1999**, *9*, 747–755.

[9] T. W. Rademacher, R. B. Parekh, R. A. Dwek, *Annu. Rev. Biochem.* **1988**, *57*, 785–838.

[10] R. G. Spiro, *Glycobiology* **2002**, *12*, 43R-56R.

[11] A. Varki, *Glycobiology* **2017**, *27*, 3–49.

[12] A. Helenius, M. Aebi, *Annu. Rev. Biochem.* **2004**, *73*, 1019–1049.

[13] C. Reily, T. J. Stewart, M. B. Renfrow, J. Novak, *Nat. Rev. Nephrol.* **2019**, *15*, 346–366.

[14] M. A. J. Ferguson, A. F. Williams, **1988**, *57*, 285–320.

[15] G. Müller, S. Wied, E. A. Dearey, G. Biemer-Daub, *Metabolism.* **2011**, *60*, 1021–1037.

[16] M. A. J. Ferguson, *J. Cell Sci.* **1999**, *112*, 2799–2809.

[17] R. A. Laine, *Glycobiology* **1994**, *4*, 759–767.

[18] F. W. Johann Lechner, *Annu. Rev. Biochem* **1989**, *58*, 173–94.

[19] A. Gottschalk, *Nature* **1969**, *222*, 452–454.

[20] F. Schwarz, M. Aebi, *Curr. Opin. Struct. Biol.* **2011**, *21*, 576–582.

[21] R. Kornfeld, S. Kornfeld, *Annu. Rev. Biochem.* **1985**, *54*, 631–664.

[22] G. Spik, G. Strecker, B. Fournet, S. Bouquelet, J. Montreuil, L. Dorland, H. Van Halbeek, J. F. G. Vliegenthart, *Eur. J. Biochem.* **1982**, *121*, 413–419.

[23] P. Stanley, N. Taniguchi, M. Aebi, in *Essentials Glycobiol. 3rd Ed.*, Cold Spring Harbor (NY), **2017**.

[24] A. J. Parodi, *Biochem. J.* **2000**, *348*, 1–13.

[25] A. Helenius, E. S. Trombetta, D. N. Hebert, J. F. Simons, *Trends Cell Biol.* **1997**, 193–200.

[26] S. A. Middleton, F. P. Barbone, D. L. Johnson, R. L. Thurmond, Y. You, F. J. McMahon, R. Jin, O. Livnah, J. Tullai, F. X. Farrell, M. A. Goldsmith, I. A. Wilson, L. K. Jolliffe, *J. Biol. Chem.* **1999**, *274*, 14163–14169.

[27] A. Sternberg, C. Naujokat, *Life Sci.* **2020**, *257*, 118056.

[28] P. Burda, M. Aebi, *Biochim. Biophys. Acta - Gen. Subj.* **1999**, *1426*, 239–257.

[29] B. Imperiali, K. W. Rickert, *Biophysics (Oxf).* **1995**, *92*, 97–101.

[30] S. Silberstein, R. Gilmore, *FASEB J.* **1996**, *10*, 849–858.

[31] S. C. Hubbard, R. J. Ivatt, *Ann. Rev. Biochem* **1981**, *50*, 555–583.

[32] K. Moremen, R. B. Trimble, A. Herscovics, *Glycobiology* **1994**, *4*, 113–125.

[33] J. D. Jamieson, E. Palade, *Cell* **1968**, 589–603.

[34] S. Yu, S. Ito, I. Wada, N. Hosokawa, *J. Biol. Chem.* **2018**, *293*, 10663–10674.

[35] A. Herscovics, *Biochim. Biophys. Acta - Gen. Subj.* **1999**, *1473*, 96–107.

[36] A. Herscovics, *Biochim. Biophys. Acta - Gen. Subj.* **1999**, *1426*, 275–285.

[37] A. Helenius, M. Aebi, *Science (80-.).* **2001**, *291*, 2364–2369.

[38] T. Kobori, S. Iwamoto, K. Takeyasu, T. Ohtani, *Pept. Sci.* **2007**, *88*, 308–324.

[39] T. Buskas, S. Ingale, G. J. Boons, *Glycobiology* **2006**, *16*.

[40] U. Carlo, K. Yasuhiro, *Curr. Opin. Chem. Biol.* **2018**, *46*, 130–137.

[41] L. X. Wang, *Carbohydr Res.* **2008**, *343*, 1509–1522.

[42] L. X. Wang, J. V. Lomino, *ACS Chem. Biol.* **2012**, *7*, 110–122.

[43] T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 249–289.

[44] R. F. Roller, A. Malik, M. A. Carillo, M. Garg, A. Rella, M. Raulf, B. Lepenies, P. H. Seeberger, D. Varón Silva, *Angew. Chemie* **2020**, *132*, 12133–12138.

[45] S. Weikert, D. Papac, J. Briggs, D. Cowfer, S. Tom, M. Gawlitzek, J. Lofgren, S. Mehta, V. Chisholm, N. Modi, S. Eppler, K. Carroll, S. Chamow, D. Peers, P. Berman, L. Krummen, *Nat. Biotechnol.* **1999**, *17*, 1116–1121.

[46] W. L. Xi, W. Huang, *Curr Opin Chem Biol.* **2009**, *13*, 592–600.

[47] R. Kowalczyk, M. A. Brimble, Y. Tomabechi, A. J. Fairbanks, M. Fletcher, D. L. Hay, *Org. Biomol. Chem.* **2014**, *12*, 8142–8151.

[48] A. J. Fairbanks, *Chem. Soc. Rev.* **2017**, *46*, 5128–5146.

[49] W. Huang, J. Giddens, S. Q. Fan, C. Toonstra, L. X. Wang, *J. Am. Chem. Soc.* **2012**.

[50] T. Li, X. Tong, Q. Yang, J. P. Giddens, L. X. Wang, *J. Biol. Chem.* **2016**, *291*, 16508–16518.

[51] S. Aimoto, *Tanpakushitsu Kakusan Koso.* **2007**, *52*, 1804–1805.

[52] M. Murakami, T. Kiuchi, M. Nishihara, K. Tezuka, R. Okamoto, M. Izumi, Y. Kajihara, *Sci. Adv.* **2016**, *2*, 1–13.

[53] M. R. Pratt, C. R. Bertozzi, *J. Am. Chem. Soc.* **2003**, *20*, 6149–6159.

[54] S. Mezzato, M. Schaffrath, C. Unverzagt, *Angew. Chemie - Int. Ed.* **2005**, *44*, 1650–1654.

[55] V. Agouridas, O. El Mahdi, V. Diemer, M. Cargoët, J. C. M. Monbaliu, O. Melnyk, *Chem. Rev.* **2019**, *119*, 7328–7443.

[56] D. F. Veber, J. D. Milkowski, S. L. Varga, R. G. Denkewalter, R. Hirschmann, *J. Am.*

Chem. Soc. **1972**, *94*, 5456–5461.

[57] D. Bang, S. B. H. Kent, *Angew. Chemie - Int. Ed.* **2004**, *43*, 2534–2538.

[58] Y. Y. Yang, S. Ficht, A. Brik, C. H. Wong, *J. Am. Chem. Soc.* **2007**, *24*, 7690–7701.

[59] S. K. Maity, M. Jbara, S. Laps, A. Brik, *Angew. Chemie* **2016**, *128*, 8240–8244.

[60] D. Bang, N. Chopra, S. B. H. Kent, *J. Am. Chem. Soc.* **2004**, *126*, 1377–1383.

[61] K. Sato, K. Kitakaze, T. Nakamura, N. Naruse, K. Aihara, A. Shigenaga, T. Inokuma, D. Tsuji, K. Itoh, A. Otaka, *Chem. Commun.* **2015**, *51*, 9946–9948.

[62] Z. Harpaz, P. Siman, K. S. Ajish Kumar, A. Brik, *ChemBioChem* **2010**, *11*, 1232–1235.

[63] M. Jbara, S. K. Maity, M. Seenaiah, A. Brik, *J. Am. Chem. Soc.* **2016**.

[64] M. Jbara, S. Laps, M. Morgan, G. Kamnesky, G. Mann, C. Wolberger, A. Brik, *Nat. Commun.* **2018**, *9*.

[65] F. Diezmann, H. Eberhard, O. Seitz, *Biopolymers* **2010**, *94*, 397–404.

[66] N. Ollivier, J. Vicogne, A. Vallin, H. Drobecq, R. Desmet, O. El Mahdi, B. Leclercq, G. Goormachtigh, V. Fafeur, O. Melnyk, *Angew. Chemie* **2012**, *124*, 213–217.

[67] S. Dong, S. Shang, Z. Tan, S. J. Danishefsky, *Isr. J. Chem.* **2011**, *51*, 968–976.

[68] S. Ueda, M. Fujita, H. Tamamura, N. Fujii, A. Otaka, *ChemBioChem* **2005**, *6*, 1983–1986.

[69] R. K. McGinty, J. Kim, C. Chatterjee, R. G. Roeder, T. W. Muir, *Nature* **2008**, *453*, 812–816.

[70] A. B. Smith, S. N. Savinov, U. V. Manjappara, I. M. Chaiken, *Org. Lett.* **2002**, *4*, 4041–4044.

[71] L. Raibaut, M. Cargoët, N. Ollivier, Y. M. Chang, H. Drobecq, E. Boll, R. Desmet, J. C. M. Monbaliu, O. Melnyk, *Chem. Sci.* **2016**, *7*, 2657–2665.

[72] E. Boll, J. P. Ebran, H. Drobecq, O. El-Mahdi, L. Raibaut, N. Ollivier, O. Melnyk, *Org. Lett.* **2015**, *17*, 130–133.

[73] L. E. Canne, S. J. Bark, S. B. H. Kent, *J. Am. Chem. Soc.* **1996**, *118*, 5891–5896.

[74] S. F. Loibl, Z. Harpaz, O. Seitz, *Angew. Chemie - Int. Ed.* **2015**, *54*, 15055–15059.

[75] Q. Wan, S. J. Danishefsky, *Angew. Chemie* **2007**, *119*, 9408–9412.

[76] R. E. Thompson, B. Chan, L. Radom, K. A. Jolliffe, R. J. Payne, *Angew. Chemie* **2013**, *125*, 9905–9909.

[77] N. J. Mitchell, L. R. Malins, X. Liu, R. E. Thompson, B. Chan, L. Radom, R. J. Payne, *J. Am. Chem. Soc.* **2015**, *137*, 14011–14014.

[78] X. Wang, A. S. Ashhurst, L. J. Dowman, E. E. Watson, H. Y. Li, A. J. Fairbanks, M. Larance, A. Kwan, R. J. Payne, *Org. Lett.* **2020**, *22*, 6863–6867.

[79] C. L. Lee, H. Liu, C. T. T. Wong, H. Y. Chow, X. Li, *J. Am. Chem. Soc.* **2016**, *138*, 10477–10484.

[80] A. Brik, S. Ficht, Y. Y. Yang, C. S. Bennett, C. H. Wong, *J. Am. Chem. Soc.* **2006**, *128*, 15026–15033.

[81] H. Chai, K. Le Mai Hoang, M. D. Vu, K. Pasunooti, C. F. Liu, X. W. Liu, *Angew. Chemie - Int. Ed.* **2016**, *55*, 10363–10367.

[82] D. J. Craik, D. P. Fairlie, S. Liras, D. Price, *Chem. Biol. Drug Des.* **2013**, *81*, 136–147.

[83] S. S. Usmani, G. Bedi, J. S. Samuel, S. Singh, S. Kalra, P. Kumar, A. A. Ahuja, M. Sharma, A. Gautam, G. P. S. Raghava, *PLoS One* **2017**.

[84] D. Chow, M. Nunalee, L. D. Woo, C. Ashutosh, A. J. Simnick, *Mater Sci Eng R Rep.* **2008**, *62*, 125–155.

[85] K. Fosgerau, T. Hoffmann, *Drug Discov. Today* **2015**, *20*, 122–128.

[86] E. Fischer, *Berichte der Dtsch. Chem. Gesellschaft* **1905**, *38*, 605–619.

[87] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.

[88] R. B. Merrifield, *Biochemistry* **1964**, *3*, 1385–1390.

[89] J. C. Sheehan, G. P. Hess, *J. Am. Chem. Soc.* **1955**, *77*, 1067–1068.

[90] W. König, R. Geiger, *Chem. Ber.* **1970**, *103*, 788–798.

[91] R. Subirós-Funosas, R. Prohens, R. Barbas, A. El-Faham, F. Albericio, *Chem. - A Eur. J.* **2009**, *15*, 9394–9403.

[92] J. C. Sheehan, G. L. Boshart, P. A. Cruickshank, *J. Org. Chem.* **1961**, *26*, 2525–2528.

[93] M. F. Chen, *Chem. Commun.* **1981**, 543–545.

[94] G. Sabatino, B. Mulinacci, M. C. Alcaro, M. Chelli, P. Rovero, A. M. Papini, *Lett. Pept. Sci.* **2002**, *9*, 119–123.

[95] L. A. Carpino, *J. Am. Chem. Soc.* **1993**, *115*, 4397–4398.

[96] R. F. Poulain, A. L. Tartar, B. P. Déprez, *Tetrahedron Lett.* **2001**, DOI 10.1016/S0040-4039(00)02293-0.

[97] V. Dourtoglou, B. Gross, V. Lambropoulou, C. Zioudrou, *Synth.* **1984**, *1984*, 572–574.

[98] F. Albericio, J. M. Bofill, A. El-Faham, S. A. Kates, *J. Org. Chem.* **1998**, *63*, 9678–9683.

[99] A. El-Faham, R. S. Funosas, R. Prohens, F. Albericio, *Chem. - A Eur. J.* **2009**, *15*, 9404–9416.

[100] B. Castro, J. R. Dormoy, G. Evin, C. Selve, *Tetrahedron Lett.* **1975**, *16*, 1219–1222.

[101] J. Coste, D. Le-Nguyen, B. Castro, *Tetrahedron Lett.* **1990**, *31*, 205–208.

[102] R. Subirós-Funosas, A. El-Faham, F. Albericio, *Org. Biomol. Chem.* **2010**, *8*, 3665–3673.

[103] D. M. M. Jaradat, *Amino Acids* **2018**, *50*, 39–68.

[104] S. B. H. Kent, *Chemical Synthesis of Peptides and Proteins*, **1988**.

[105] F. García-Martín, M. Quintanar-Audelo, Y. García-Ramos, L. J. Cruz, C. Gravel, R. Furic, S. Côté, J. Tulla-Puche, F. Albericio, *J. Comb. Chem.* **2006**, *8*, 213–220.

[106] A. R. Vaino, K. D. Janda, *J. Comb. Chem.* **2000**, *2*, 579–596.

[107] C. Blackburn, *Wiley & Sons,Inc* **1999**, *30*, 311–351.

[108] S. sun Wang, *J. Am. Chem. Soc.* **1973**, *95*, 1328–1333.

[109] H. Rink, *Tetrahedron Lett.* **1987**, *28*, 3787–3790.

[110] L. E. Canne, S. M. Walker, S. B. H. Kent, *Tetrahedron Lett.* **1995**, *36*, 1217–1220.

[111] W. F. DeGrado, E. T. Kaiser, *J. Org. Chem.* **1980**, *45*, 1295–1300.

[112] L. A. Carpino, G. Y. Han, *J. Am. Chem. Soc.* **1970**, *92*, 5748–5749.

[113] F. C. McKay, N. F. Albertson, *J. Am. Chem. Soc.* **1957**, *79*, 4686–4690.

[114] L. A. Carpino, *Acc. Chem. Res.* **1987**, *20*, 401–407.

[115] A. Isidro-Llobet, M. Álvarez, F. Albericio, *Chem. Rev.* **2009**, *109*, 2455–2504.

[116] Fernando Albericio Judit Tulla-Puche Steven A. Kates, in *Amin. Acids, Pept. Proteins Org. Chem. Build. Blocks, Catal. Coupling Chem. Vol. 3* (Ed.: Andrew B. Hughes), Copyright © 2011 Wiley-VCH Verlag GmbH & Co. KGaA, **2011**.

[117] R. C. de L. Milton, S. C. F. Milton, P. A. Adams, *J. Am. Chem. Soc.* **1990**, *112*, 6039–6046.

[118] M. Meli, G. Morra, G. Colombo, *Biophys. J.* **2008**, *94*, 4414–4426.

[119] M. Mergler, F. Dick, *J. Pept. Sci.* **2005**, *11*, 650–657.

[120] B. Due Larsen, A. Holm, *J. Pept. Res.* **1998**, *52*, 470.

[121] A. R. Katritzky, D. N. Haase, J. V. Johnsons, A. Chung, *J. Org. Chem.* **2009**, *74*, 2028–2032.

[122] Y. Hui-Ming, S. T. Chen, K. T. Wang, K. T. Wang, *J. Org. Chem.* **1992**, *57*, 4781–4784.

[123] M. Terêsa Machini Miranda Cleber W. Liria Cesar Remuzgo, in *Amin. Acids, Pept. Proteins Org. Chem. Build. Blocks, Catal. Coupling Chem. Vol. 3* (Ed.: Andrew B. Hughes), Copyright © 2011 Wiley-VCH Verlag GmbH & Co. KGaA, **2011**, pp. 549–569.

[124] M. H. S. Cezari, L. Juliano, *Pept. Res.* **1996**, *9*, 88–91.

[125] M. Kuczer, D. Konopińska, G. Rosiński, *J. Pept. Sci.* **2007**, 16–26.

[126] M. Paradís-Bas, J. Tulla-Puche, F. Albericio, *Chem. Soc. Rev.* **2016**, *45*, 631–654.

[127] T. A. Rege, M. A. Pallero, C. Gomez, H. E. Grenett, J. E. Murphy-Ullrich, J. S. Hagood, *Exp. Cell Res.* **2006**, *312*, 3752–3767.

[128] T. Spleen, *Nature* **1964**, *203*, 886–887.

[129] A. E. Reif, J. M. Allen, *J. Exp. Med.* **1964**, *120*, 413–433.

[130] S. M. M. Haeryfar, D. W. Hoskin, *J. Immunol.* **2004**, *173*, 3581–3588.

[131] T. Seki, N. Spurr, F. Obata, S. Goyert, P. Goodfellow, J. Silver, *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82*, 6657–6661.

[132] S. Pont, *Biochimie* **1987**, *69*, 315–320.

[133] C. Sauzay, K. Voutetakis, A. A. Chatziioannou, E. Chevet, T. Avril, *Front. Cell Dev. Biol.* **2019**, *7*, 1–11.

[134] S. W. Homans, M. A. J. Ferguson, R. A. Dwek, T. W. Rademacher, R. Anand, A. F. Williams, *Lett. to Nat.* **1988**, *333*, 269–272.

[135] D. G. Campbell, J. Gagnon, K. B. M. Reid, A. F. Williams, *Rat Brain Thy-I Glycoprotein The Amino Acid Sequence, Disulphide Bonds and an Unusual Hydrophobic Region*, **1981**.

[136] T. A. Rege, J. S. Hagood, *FASEB J.* **2006**, *20*, 1045–1054.

[137] J. D. Lin, X. W. Liu, *Chem. - An Asian J.* **2020**, *15*, 2548–2557.

[138] M. Izumi, Y. Makimura, S. Dedola, A. Seko, A. Kanamori, M. Sakono, Y. Ito, Y. Kajihara, *J. Am. Chem. Soc.* **2012**, *134*, 7238–7241.

[139] B. Aussedat, B. Fasching, E. Johnston, N. Sane, P. Nagorny, S. J. Danishefsky, *J. Am. Chem. Soc.* **2012**, *134*, 3532–3541.

[140] K. C. Nicolaou, H. J. Mitchell, *Angew. Chemie - Int. Ed.* **2001**, *40*, 1576–1624.

[141] J. M. C. M. J. Collins, in *Org. Synth. Second Ed.* (Ed.: André Loupy), Copyright © 2006 WILEY-VCH Verlag GmbH & Co. KGaA, **2006**.

[142] Yi Yang, *Side Reactions in Peptide Synthesis*, Academic Press, **2015**.

[143] G.-M. Fang, Y.-M. Li, F. Shen, Y.-C. Huang, J.-B. Li, Y. Lin, H.-K. Cui, L. Liu, *Angew. Chemie* **2011**, *123*, 7787–7791.

[144] J. S. Zheng, S. Tang, Y. K. Qi, Z. P. Wang, L. Liu, *Nat. Protoc.* **2013**, *8*, 2483–2495.

[145] E. Kaiser, R. L. Colescott, C. D. Bossinger, P. I. Cook, *Anal. Biochem.* **1970**, DOI 10.1016/0003-2697(70)90146-6.

[146] P. N. L. Campayo , B. Jimenez, T. Manzano, *Synth.* **1985**, *Num 2*, 197–200.

[147] Andreas Reif Sebastian Siebenhaar Andrea Tröster Marina Schmälzlein Carolin Lechner Phanindra Velisetty Karen Gottwald Claudia Pöhner Irene Boos Volker Schubert Prof. Stefan Rose-John Prof. Carlo Unverzagt, *Angew. Chemie - Int. Ed.* **2014**, *53*, 12125–12131.

[148] F. M. Ibatullin, S. I. Selivanov, *Tetrahedron Lett.* **2009**, *50*, 6351–6354.

[149] M. B. Tatina, D. T. Khong, Z. M. A. Judeh, *European J. Org. Chem.* **2018**, *2018*, 2208–2213.

[150] M. Tosin, P. V. Murphy, *Org. Lett.* **2002**, *4*, 3675–3678.

[151] M. Hunsen, D. A. Long, C. R. D'Ardenne, A. L. Smith, *Carbohydr. Res.* **2005**.

[152] S. K. A. IBATULLIN F.M., *Carbohydr. Lett.* **2000**, *3*, 427–429.

[153] X. Huang, X. Luo, Y. Roupioz, J. W. Keillor, *J. Org. Chem.* **1997**, *62*, 8821–8825.

[154] T. Wöhr, F. Wahl, A. Nefzi, B. Rohwedder, T. Sato, X. Sun, M. Mutter, *J. Am. Chem. Soc.* **1996**, *118*, 9218–9227.

[155] V. Cardona, I. Eberle, S. Barthélémy, J. Beythien, B. Doerner, P. Schneeberger, J. Keyte, P. D. White, *Int. J. Pept. Res. Ther.* **2008**, *14*, 285–292.

[156] P. Klán, T. Šolomek, C. G. Bochet, A. Blanc, R. Givens, M. Rubina, V. Popik, A. Kostikov, J. Wirz, *Chem. Rev.* **2013**, *113*, 119–191.

[157] S. T. Anisfeld, P. T. Lansbury, *J. Org. Chem.* **1990**, *55*, 5560–5562.

[158] D. Michel, Synthesis of Homogeneous Nglycosylated- and GPI-Anchored Peptides for Semi-Synthesis of the Prion Protein, **2017**.

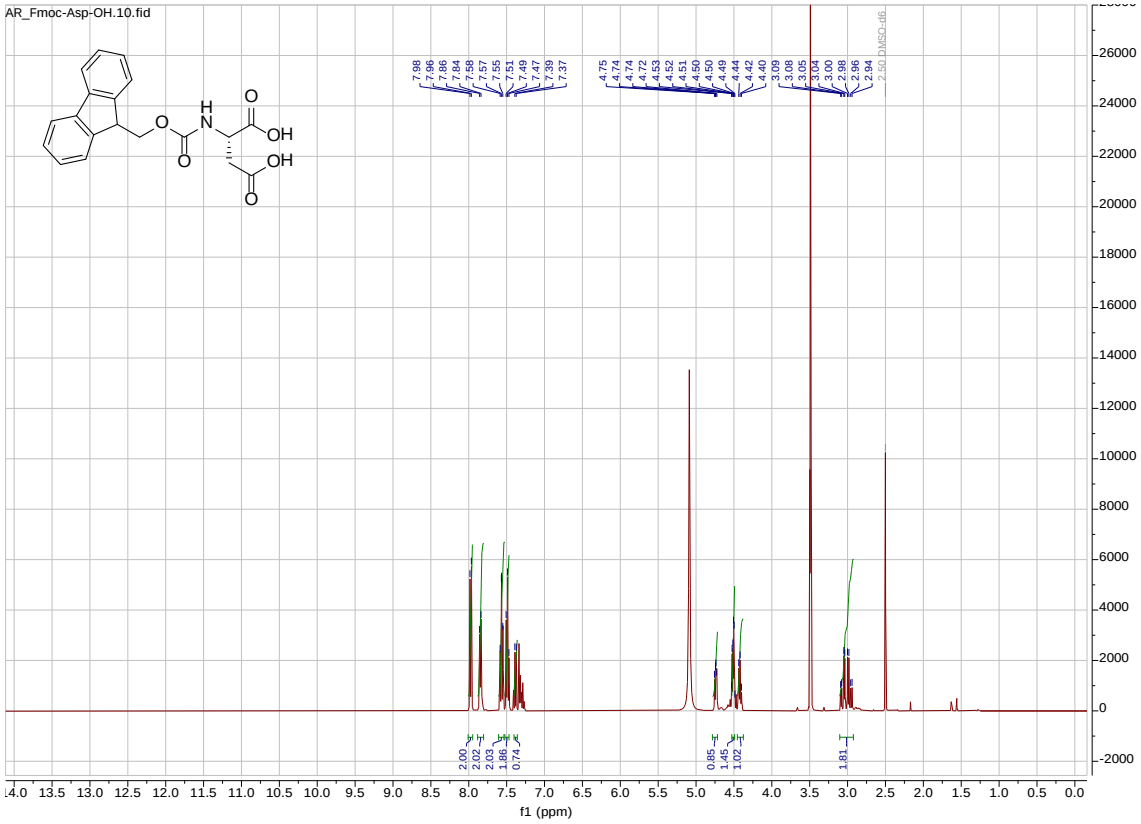
[159] J. D. Wade, M. N. Mathieu, M. Macris, G. W. Tregear, *Lett. Pept. Sci.* **2000**, *7*, 107–112.

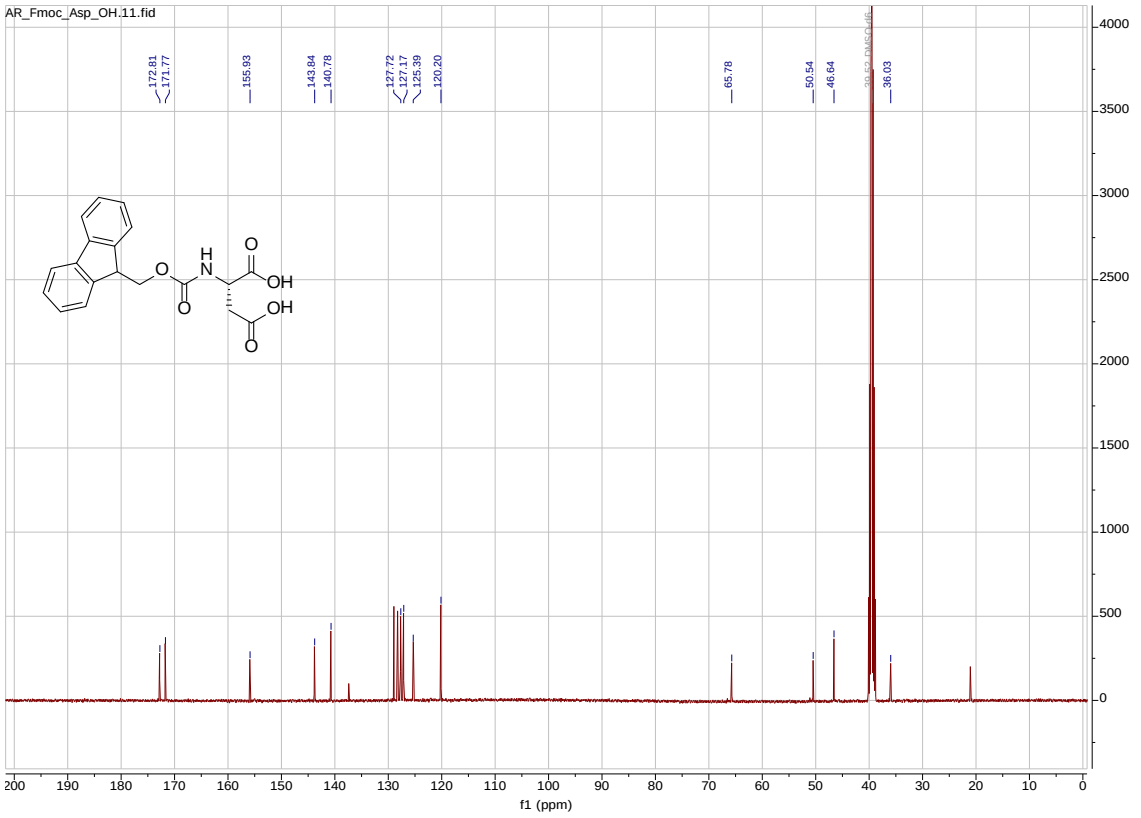
[160] T. Michels, R. Dölling, U. Haberkorn, W. Mier, *Org. Lett.* **2012**, *14*, 5218–5221.

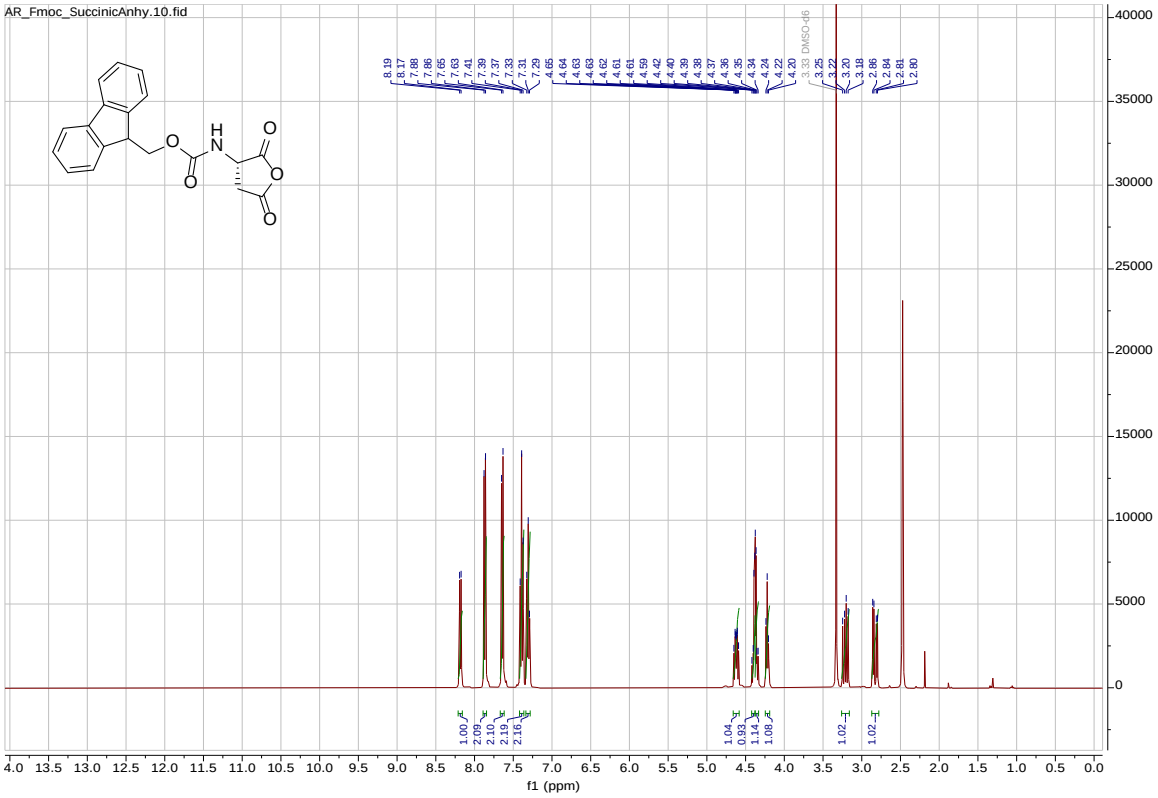
[161] M. Liu, G. Barany, D. Live, *Carbohydr. Res.* **2005**, *340*, 2111–2122.

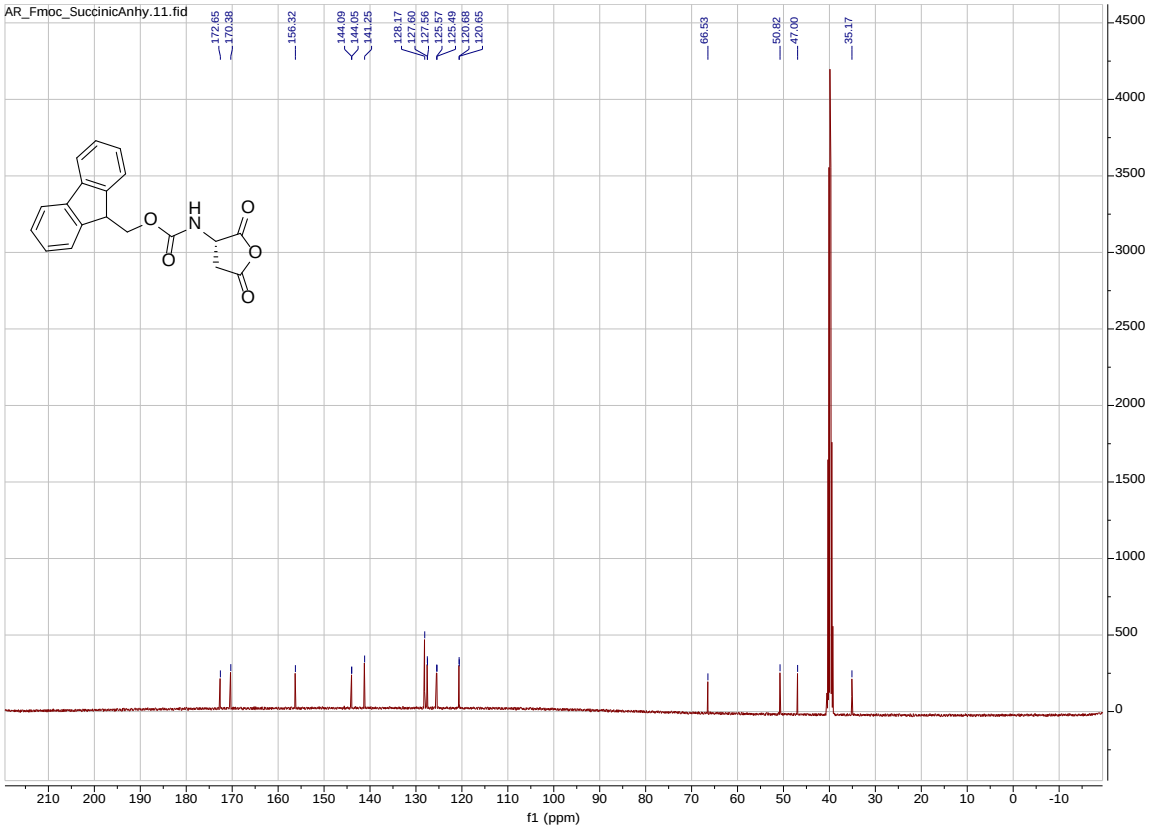
[162] A. A. Vinogradov, M. D. Simon, B. L. Pentelute, *Org. Lett.* **2016**, *18*, 1222–1225.

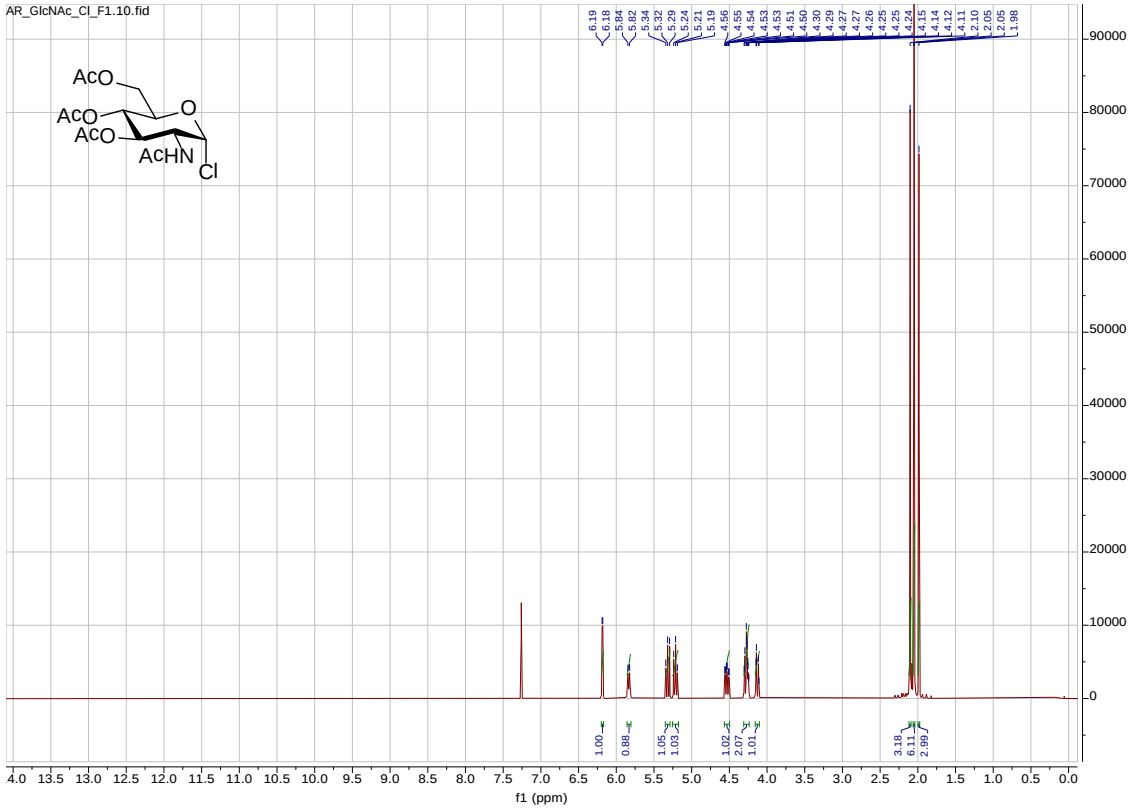
Appendix

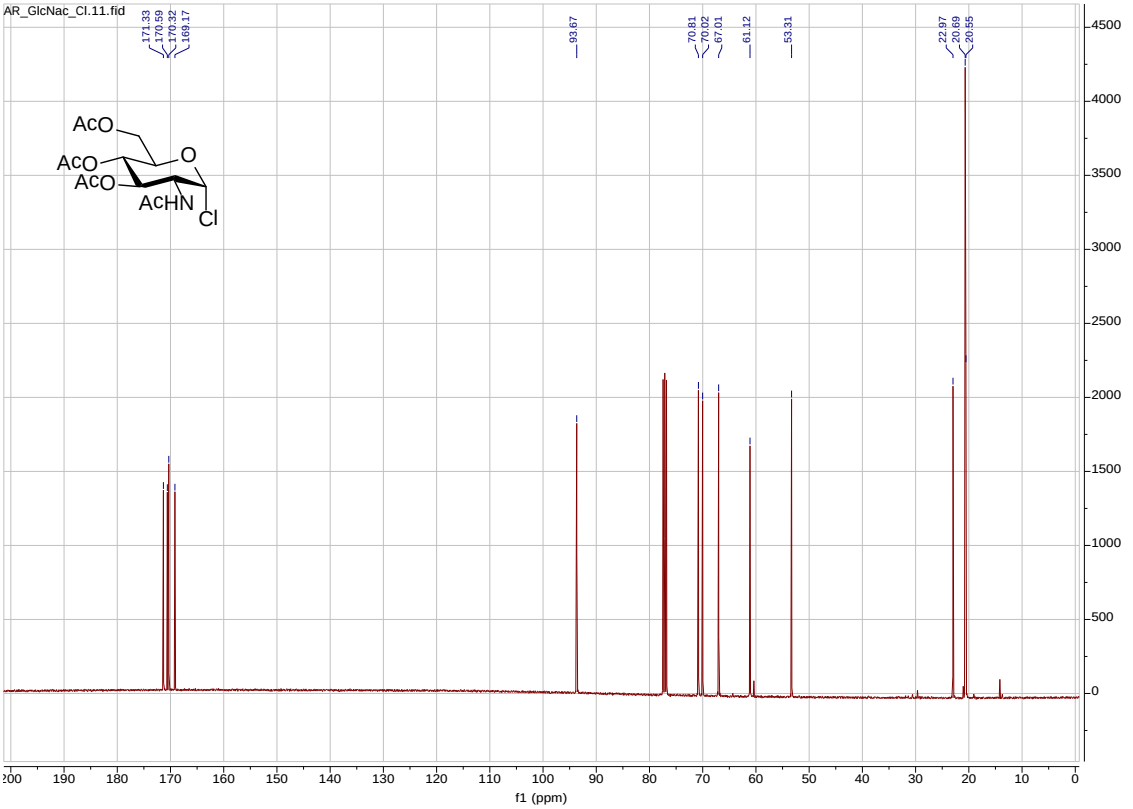

AR_Fmoc-Asp-OH.10.fid
7.98
7.96
7.86
7.84
7.58
7.57
7.55
7.51
7.49
7.47
7.39
7.37
4.75
4.74
4.74
4.72
4.53
4.52
4.51
4.50
4.50
4.49
4.44
4.42
4.40
3.09
3.08
3.05
3.04
3.00
2.98
2.96
2.94
2.50 DMSO-d6
2.00
2.02
2.03
1.86
0.74
0.85
1.45
1.02
1.81
f1 (ppm)

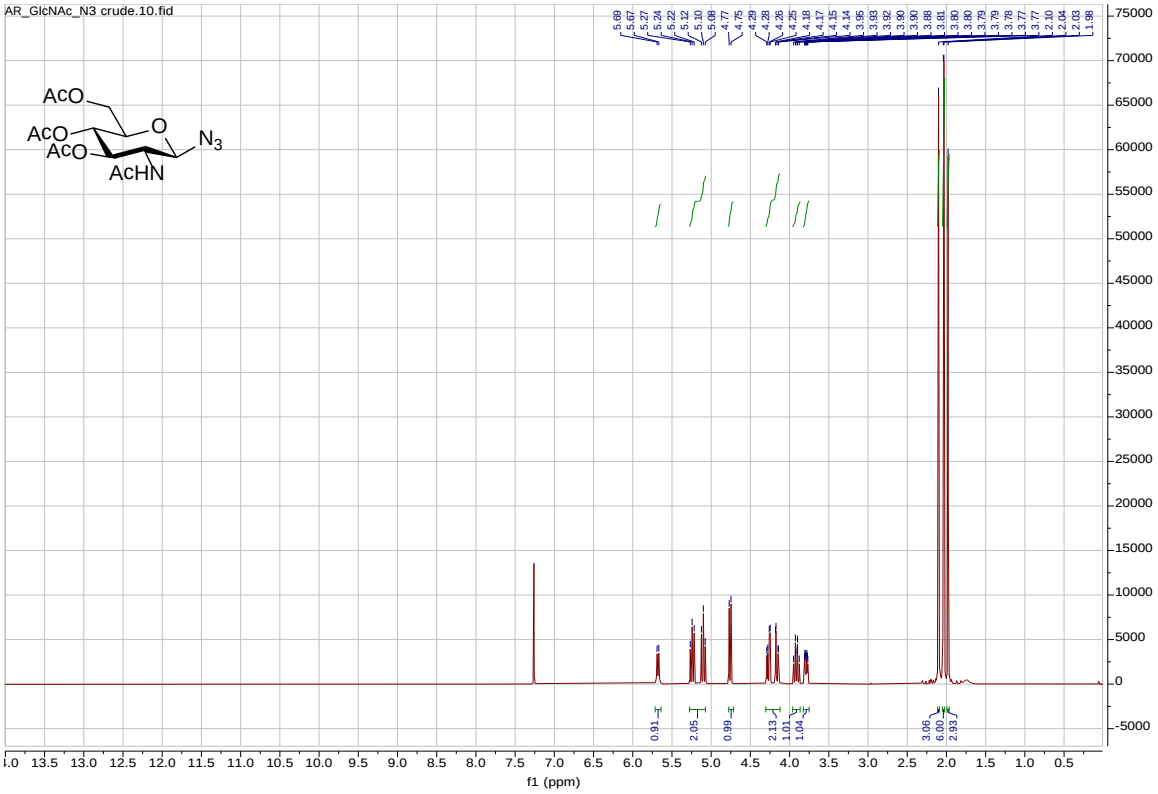

AR_Fmoc_Asp_OH.11.fid
172.81
171.77
155.93
143.84
140.78
127.72
127.17
125.39
120.20
65.78
50.54
46.64
39.52 DMSO-d6
36.03
f1 (ppm)

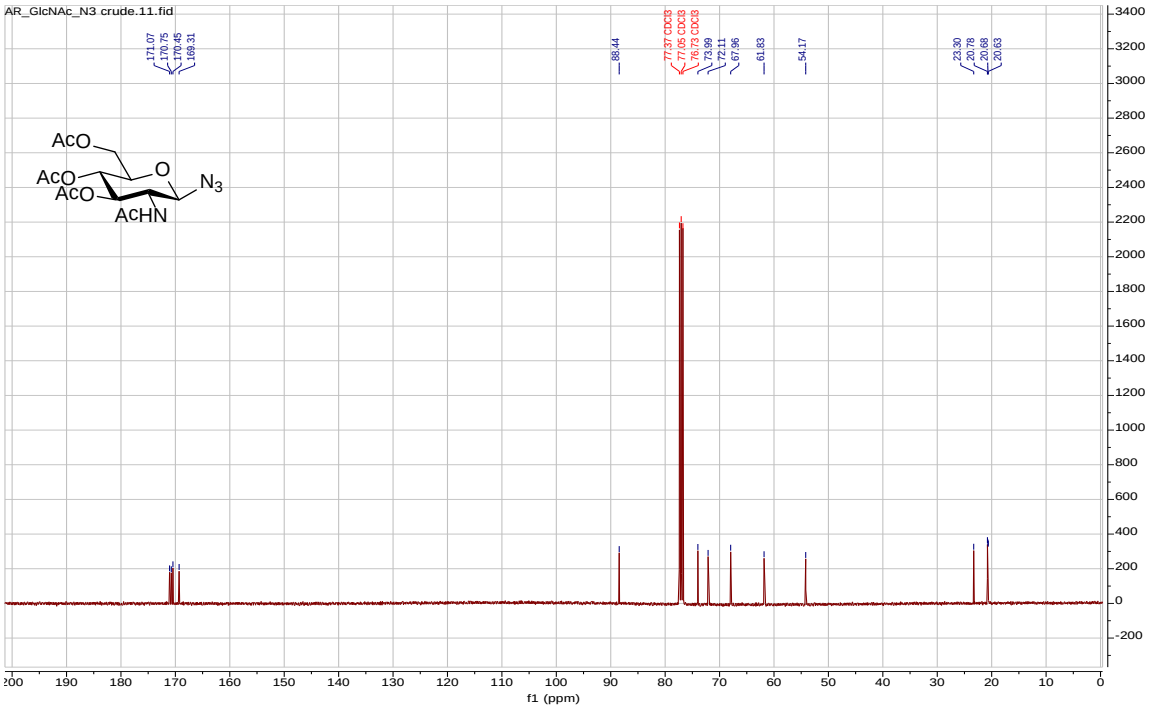

AR_Fmoc_SuccinicAnhy.10.fid
8.19
8.17
7.88
7.86
7.65
7.63
7.41
7.39
7.37
7.33
7.31
7.29
4.65
4.64
4.63
4.62
4.61
4.61
4.59
4.42
4.40
4.39
4.38
4.37
4.36
4.35
4.34
4.24
4.22
4.20
3.33 DMSO-d6
3.25
3.22
3.20
3.18
2.86
2.84
2.81
2.80
1.00
2.09
2.10
2.19
2.16
1.04
0.93
1.14
1.08
1.02
1.02
f1 (ppm)

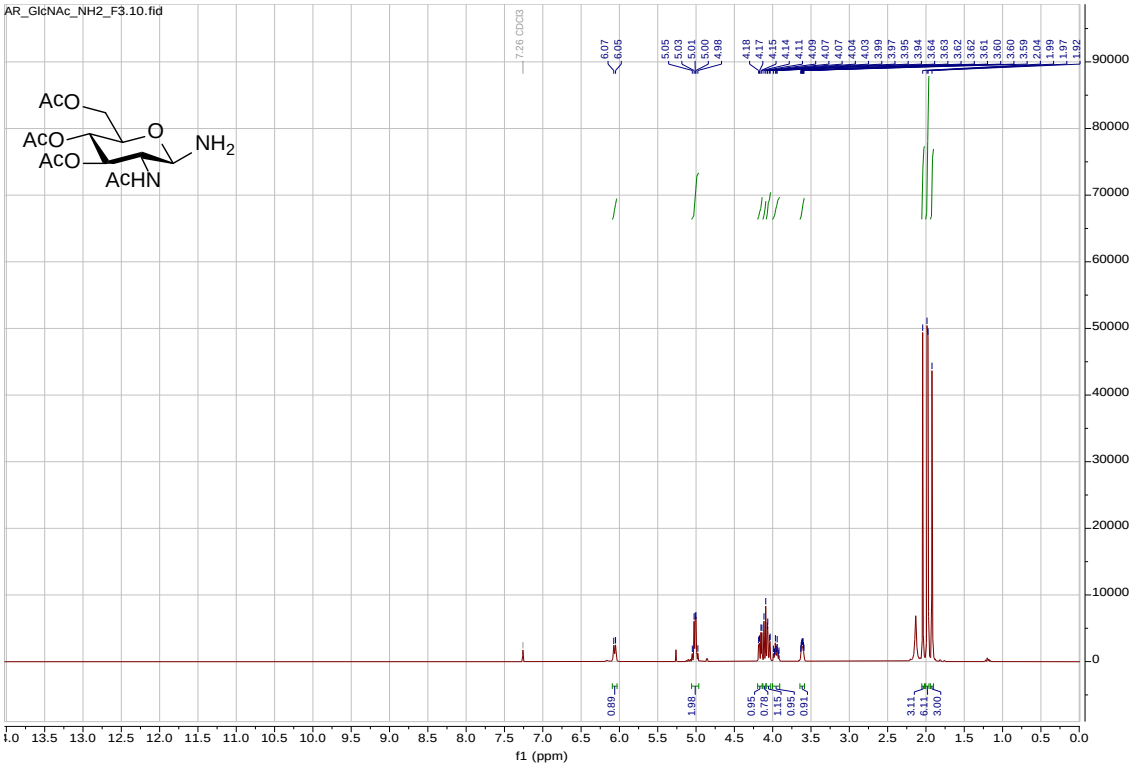

AR_Fmoc_SuccinicAnhy.11.fid
172.65
170.38
156.32
144.09
144.05
141.25
128.17
127.60
127.56
125.57
125.49
120.68
120.65
66.53
50.82
47.00
35.17
f1 (ppm)

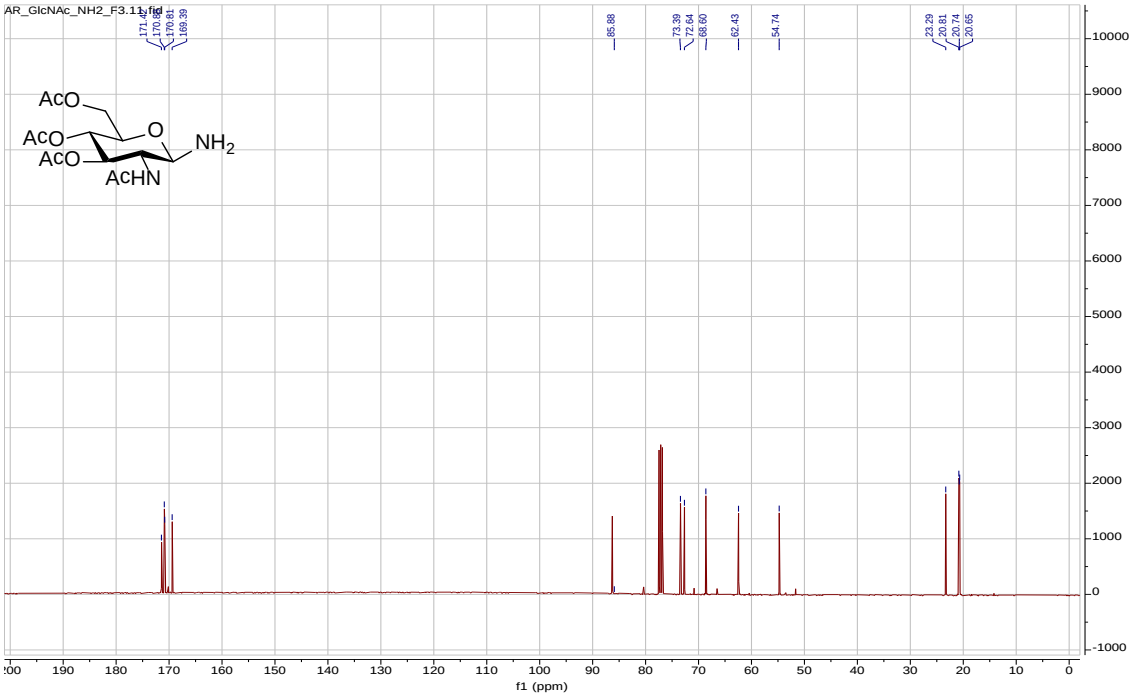

AR_GlcNAc_Cl_F1.10.fid
AcO
AcO
AcO
AcHN
O
Cl
f1 (ppm)

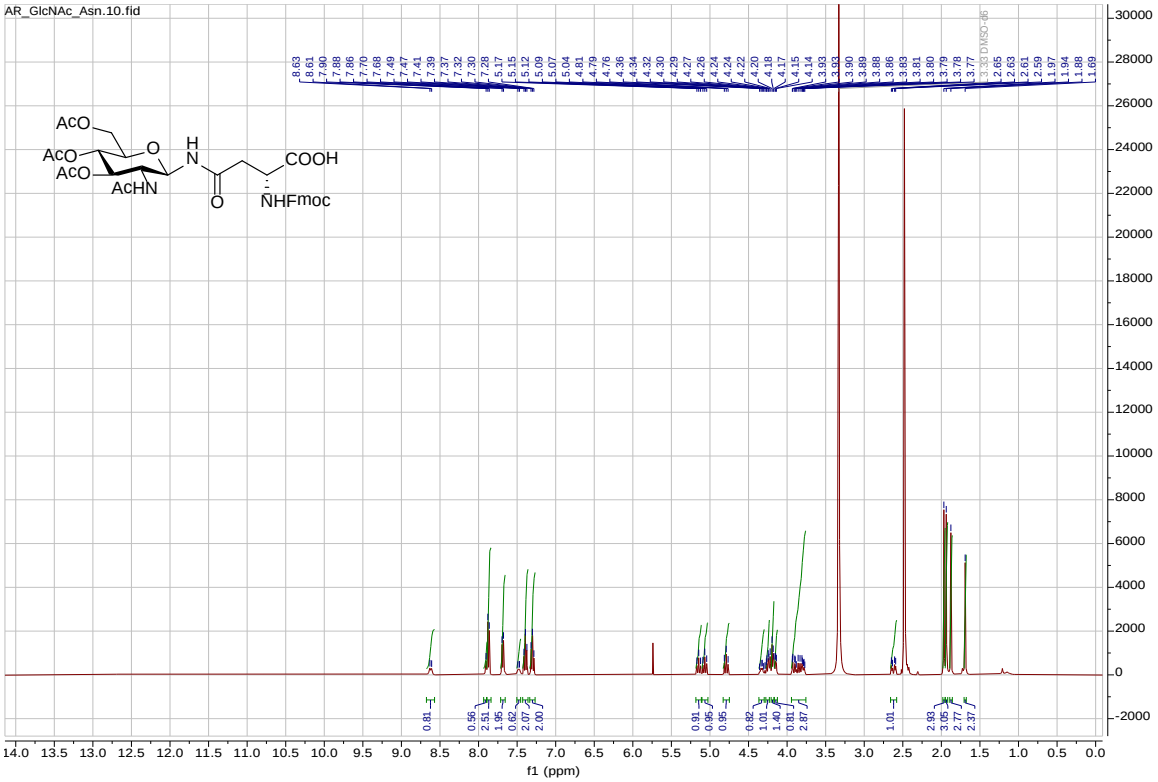

AR_GlcNac_Cl.11.fid
AcO
AcO
AcO
AcHN
O
Cl
f1 (ppm)

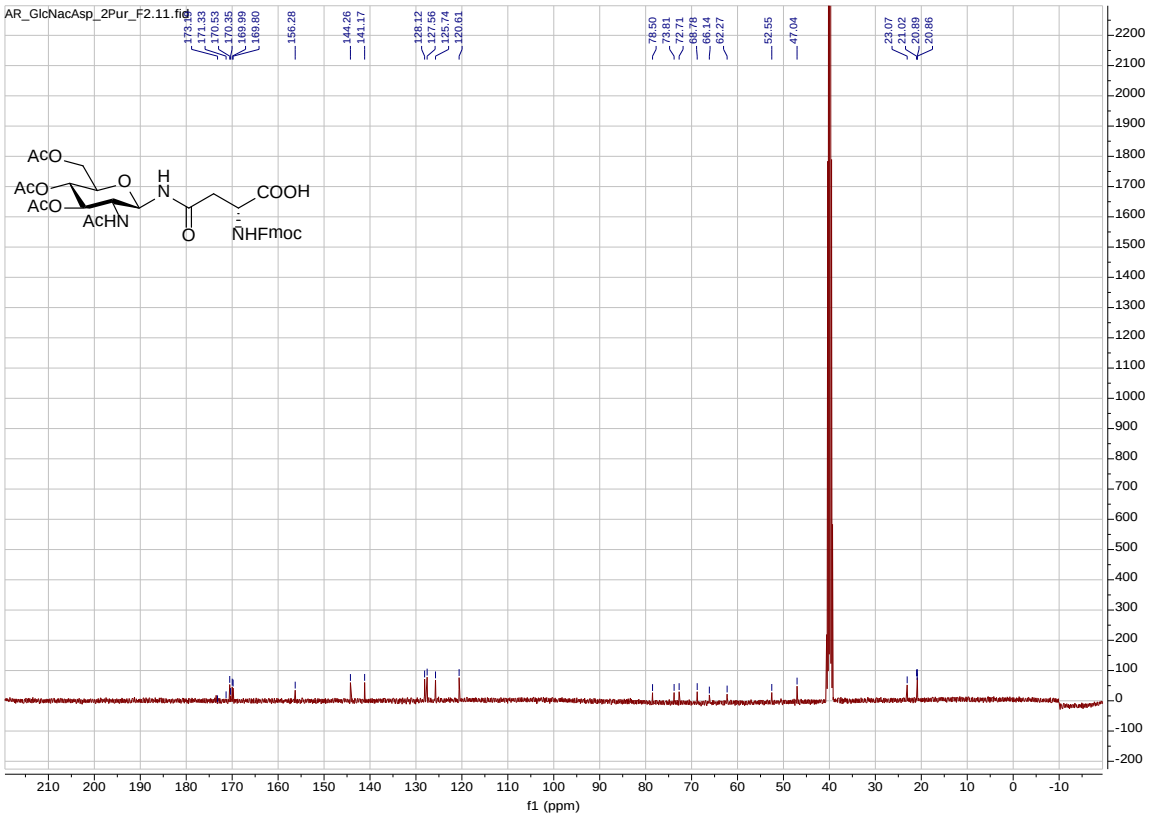

AR_GlcNAc_N3 crude.10.fid
AcO
AcO
AcO
O
N3
AcHN
f1 (ppm)

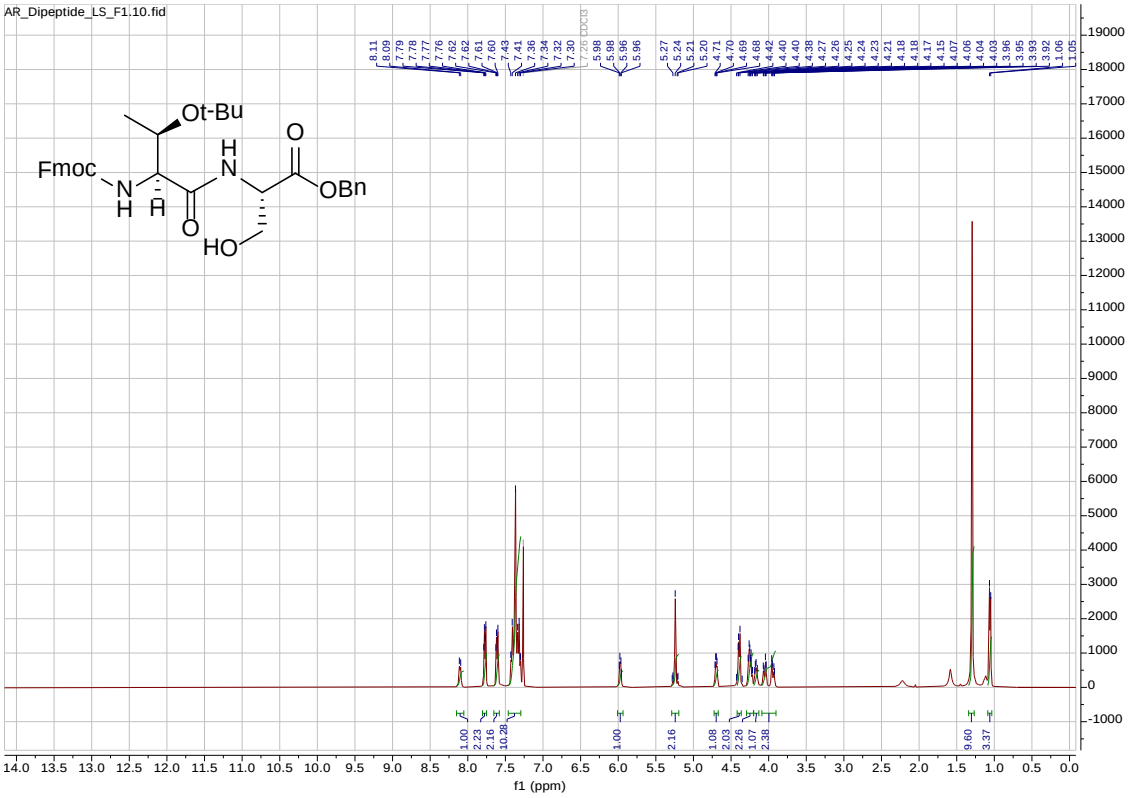

AR_GlcNAc_N3 crude.11.fid
AcO
AcO
AcO
O
N3
AcHN
f1 (ppm)

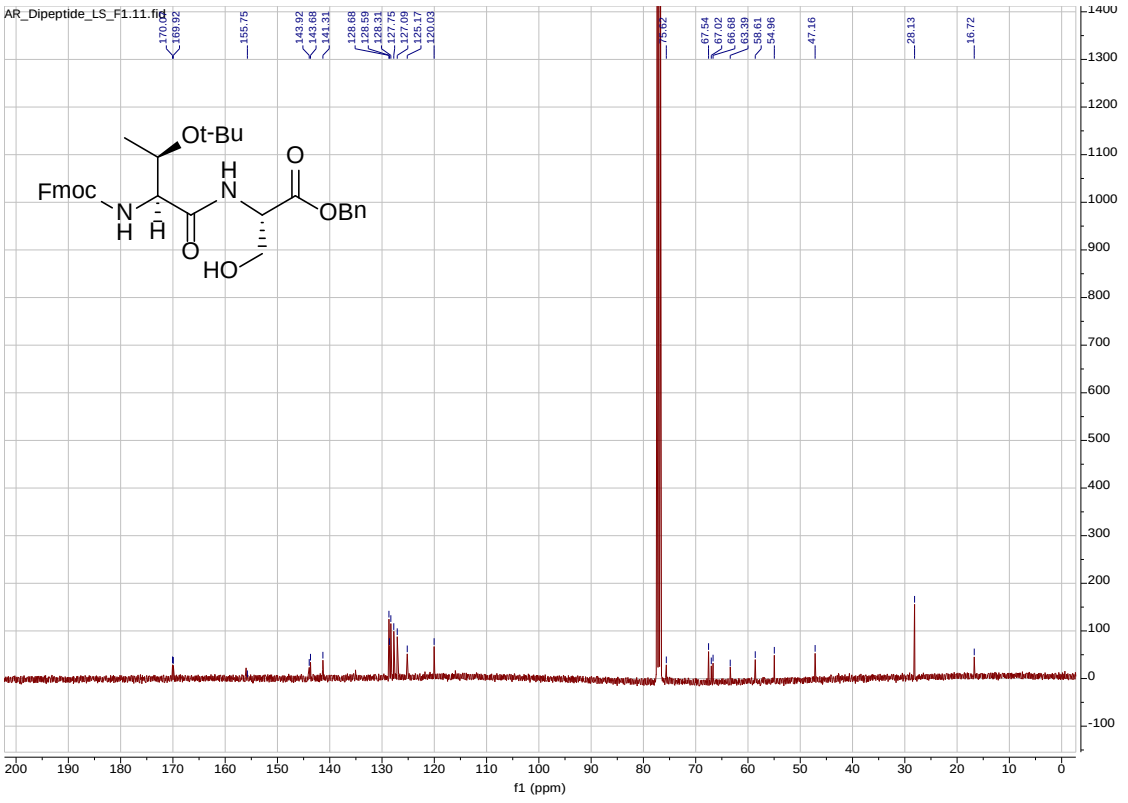

AR_GlcNAc_NH2_F3.10.fid
AcO
AcO
AcO
AcHN
O
NH2
f1 (ppm)

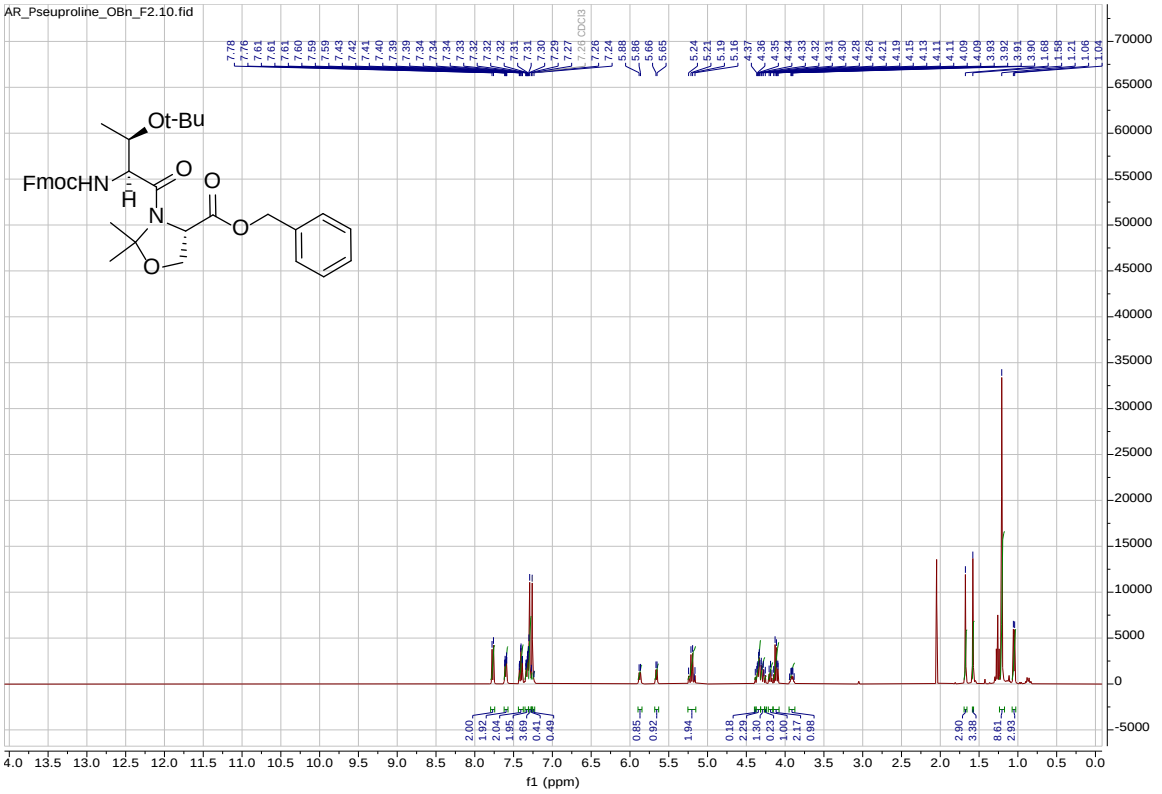

AR_GlcNAc_NH2_F3.11.fid
AcO
AcO
AcO
AcHN
O
NH2
f1 (ppm)

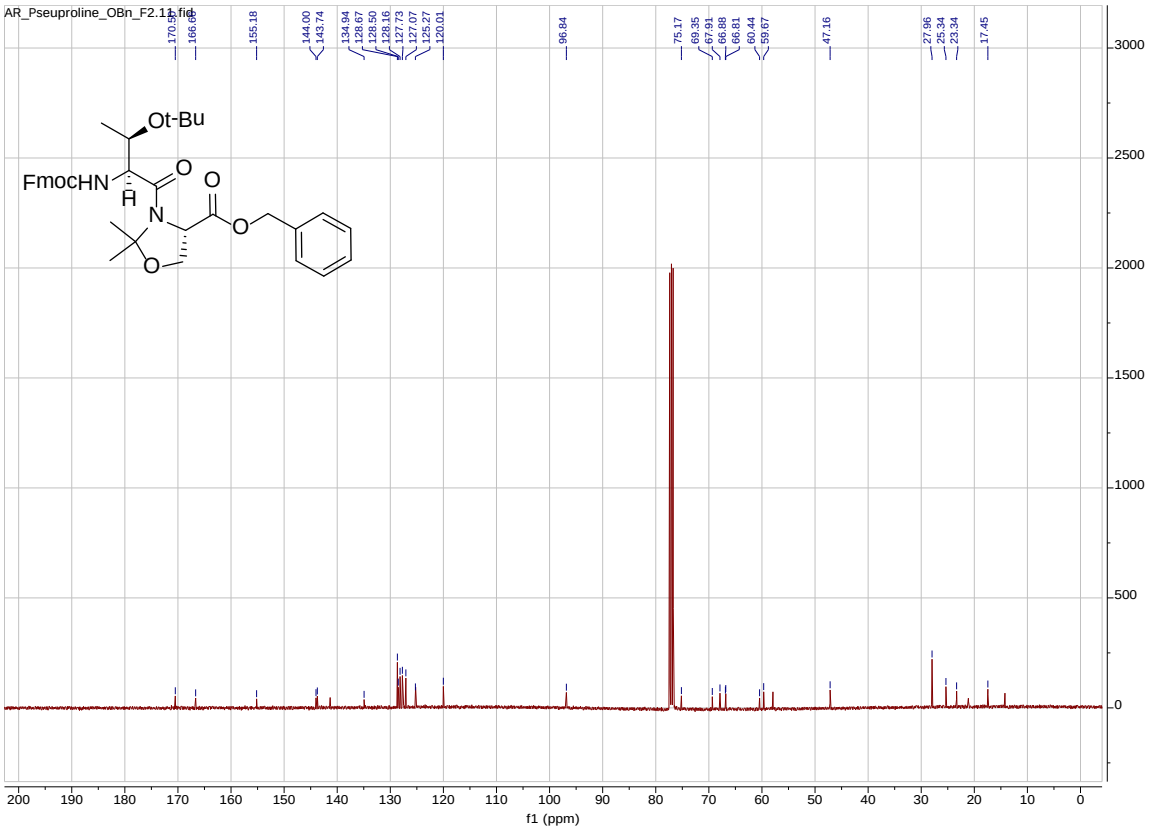

AR_GlcNAc_Asn.10.fid
AcO
AcO
AcO
AcHN
O
H
N
O
COOH
NHFmoc
f1 (ppm)

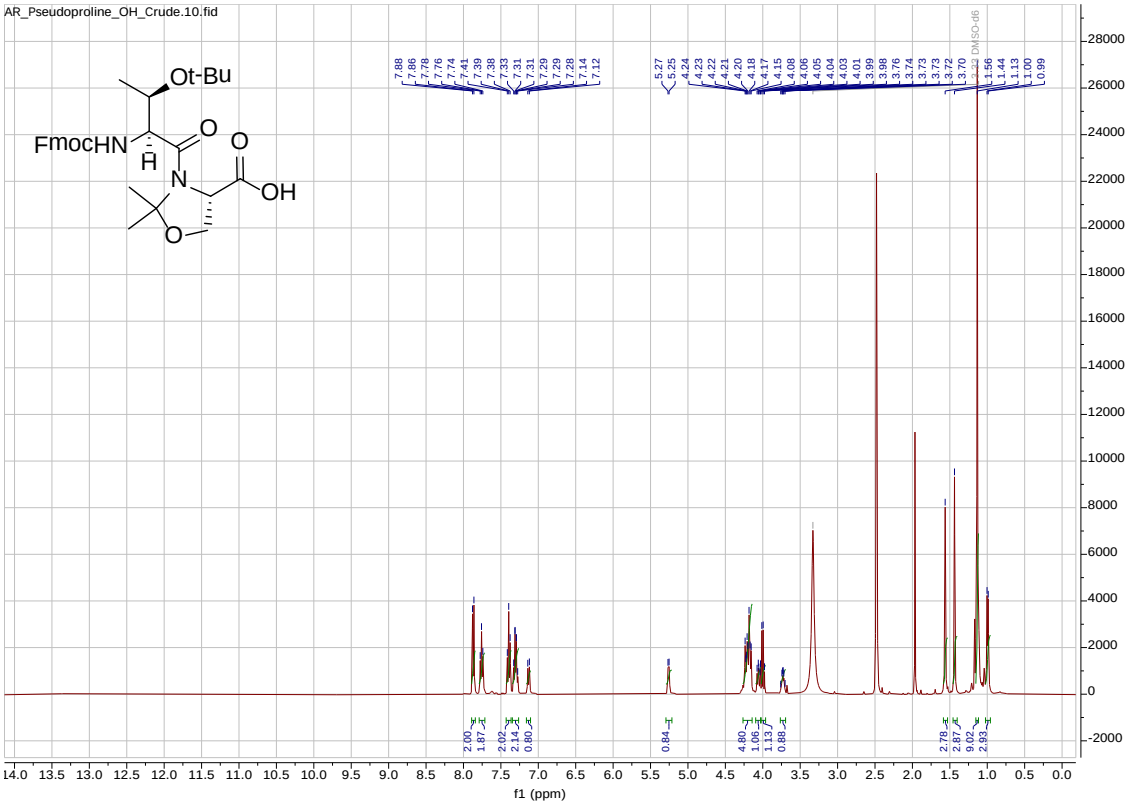

AR_GlcNacAsp_2Pur_F2.11.fid
AcO
AcO
AcO
AcHN
O
H
N
O
COOH
NHFmoc
f1 (ppm)

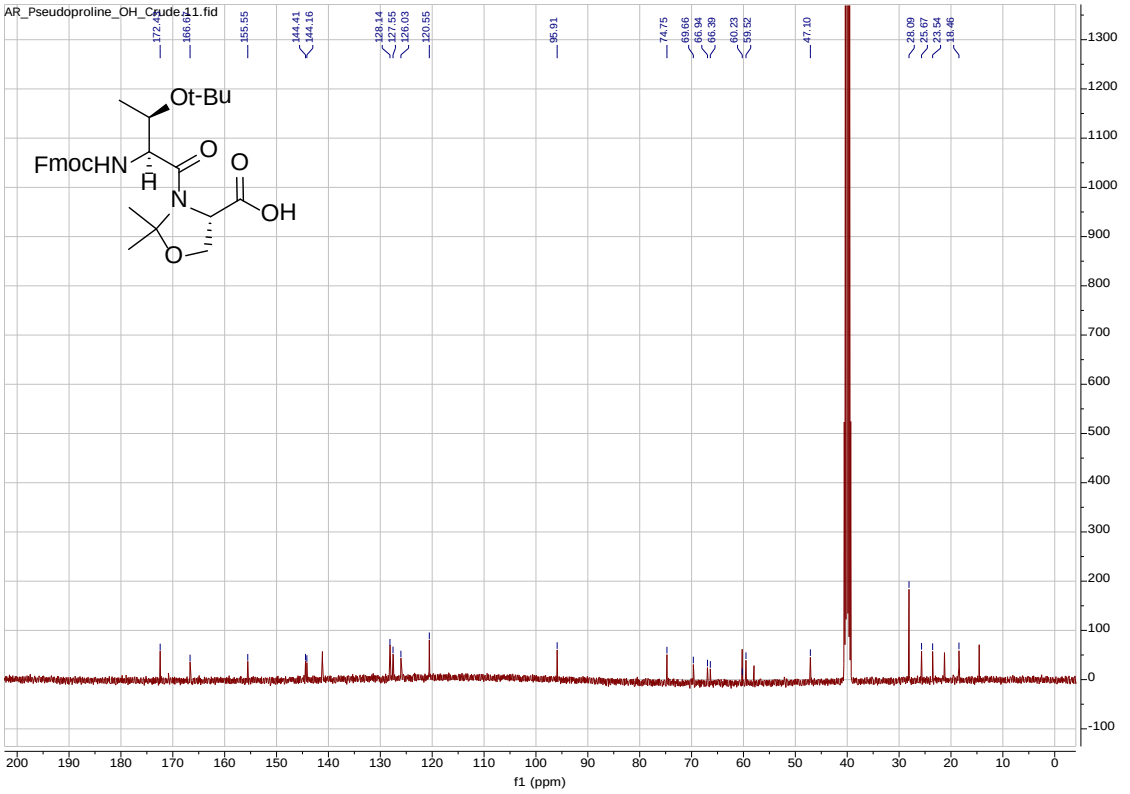

AR_Dipeptide_LS_F1.10.fid
Ot-Bu
Fmoc
N
H
H
O
H
N
O
OBn
HO
f1 (ppm)

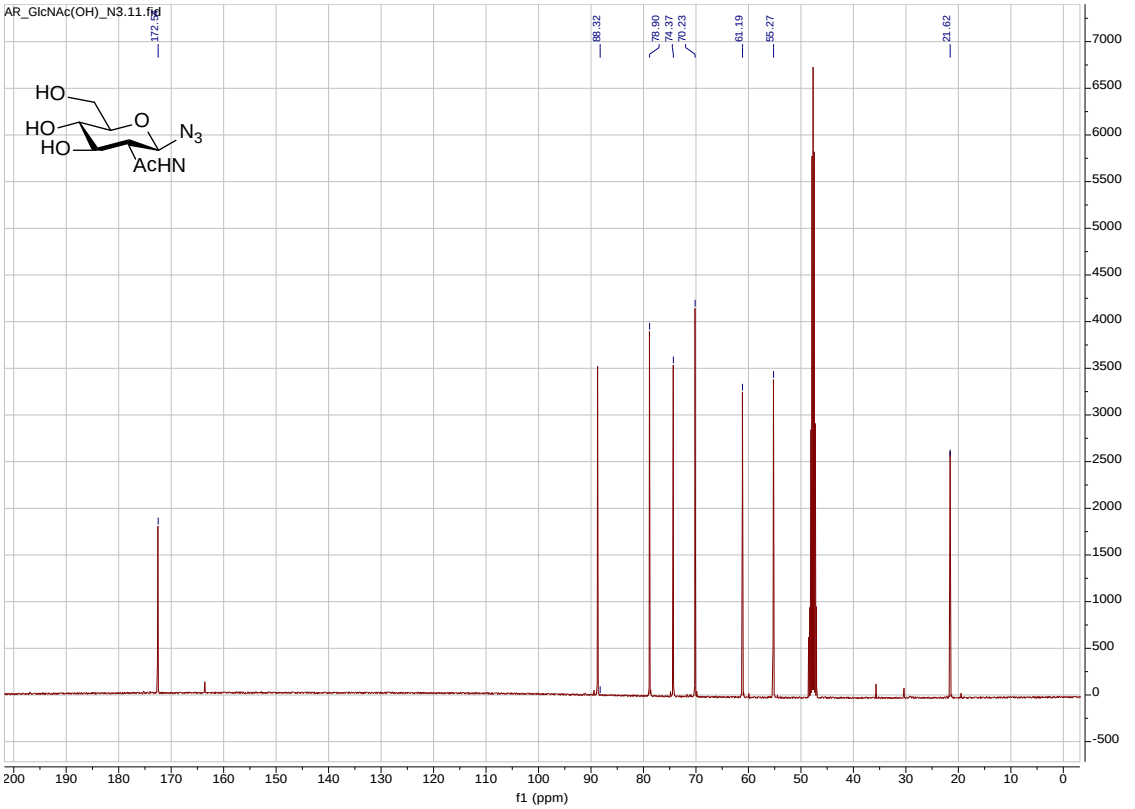

AR_Dipeptide_LS_F1.11.fid
Ot-Bu
Fmoc
N
H
H
O
H
N
O
OBn
HO
f1 (ppm)

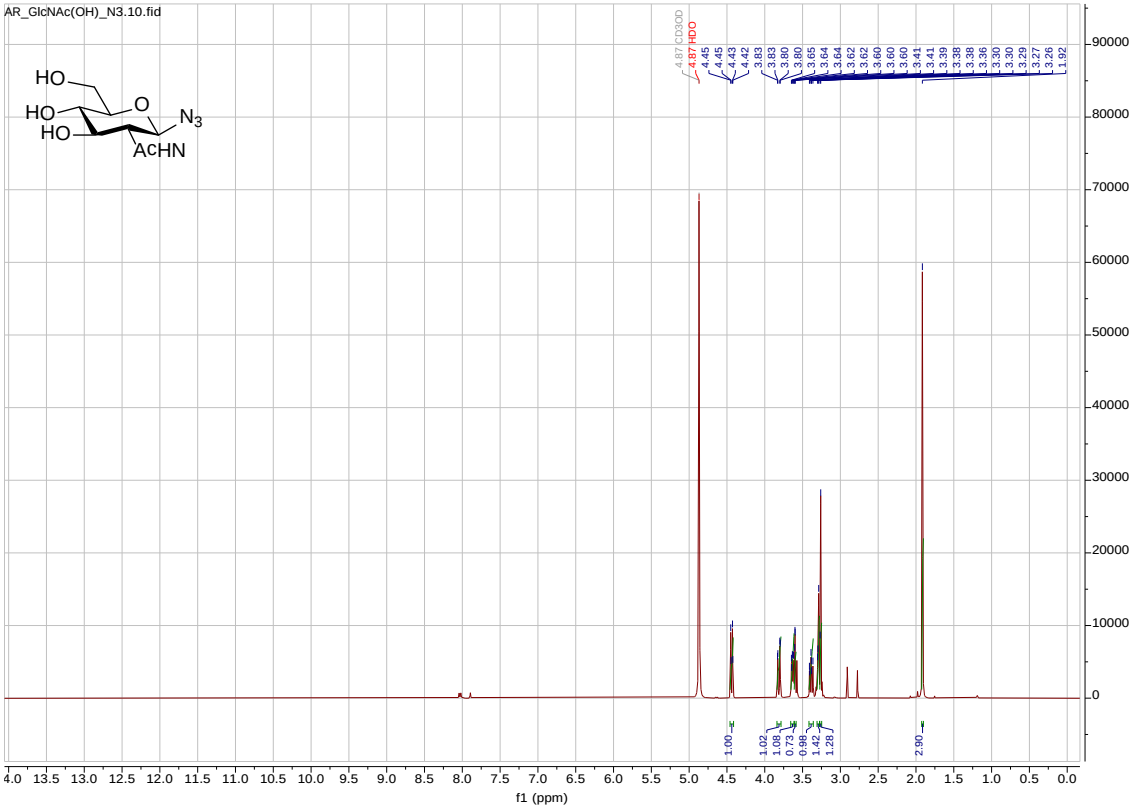

AR_Pseuproline_OBn_F2.10.fid
Ot-Bu
FmocHN
O
O
H
N
O
O
O
f1 (ppm)

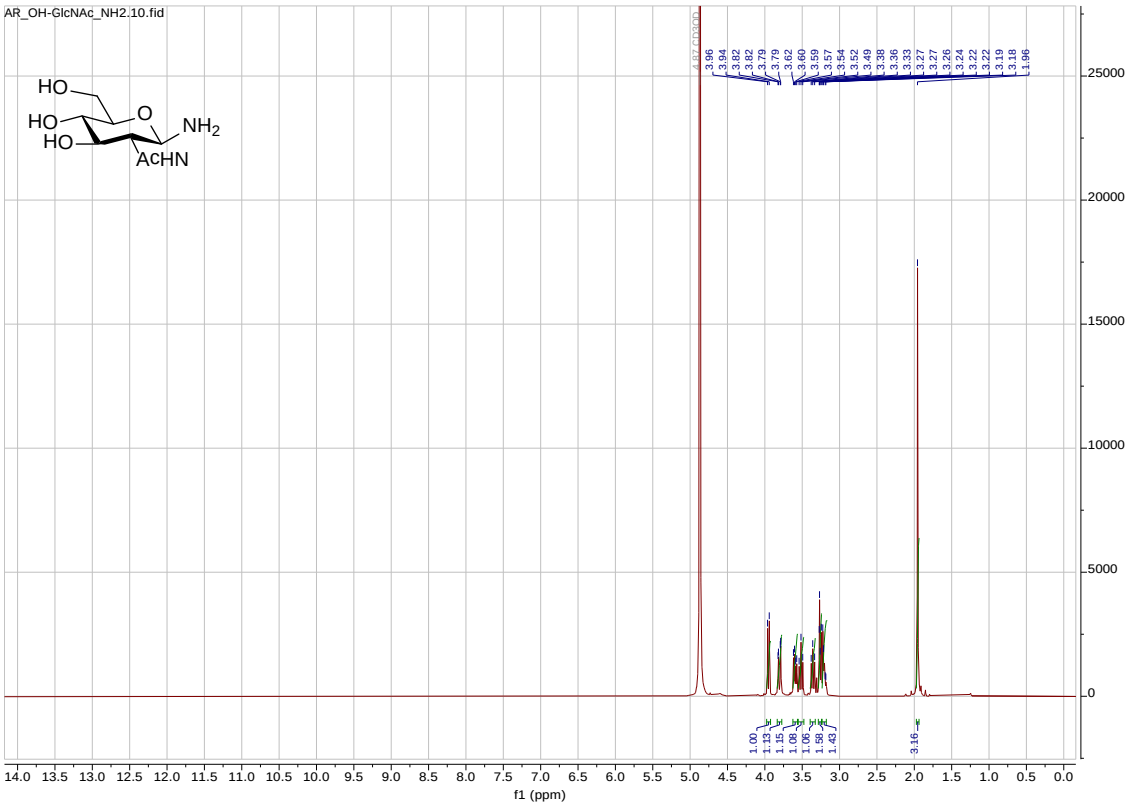

AR_Pseuproline_OBn_F2.11.fid
Ot-Bu
FmocHN
O
O
H
N
O
O
O
f1 (ppm)

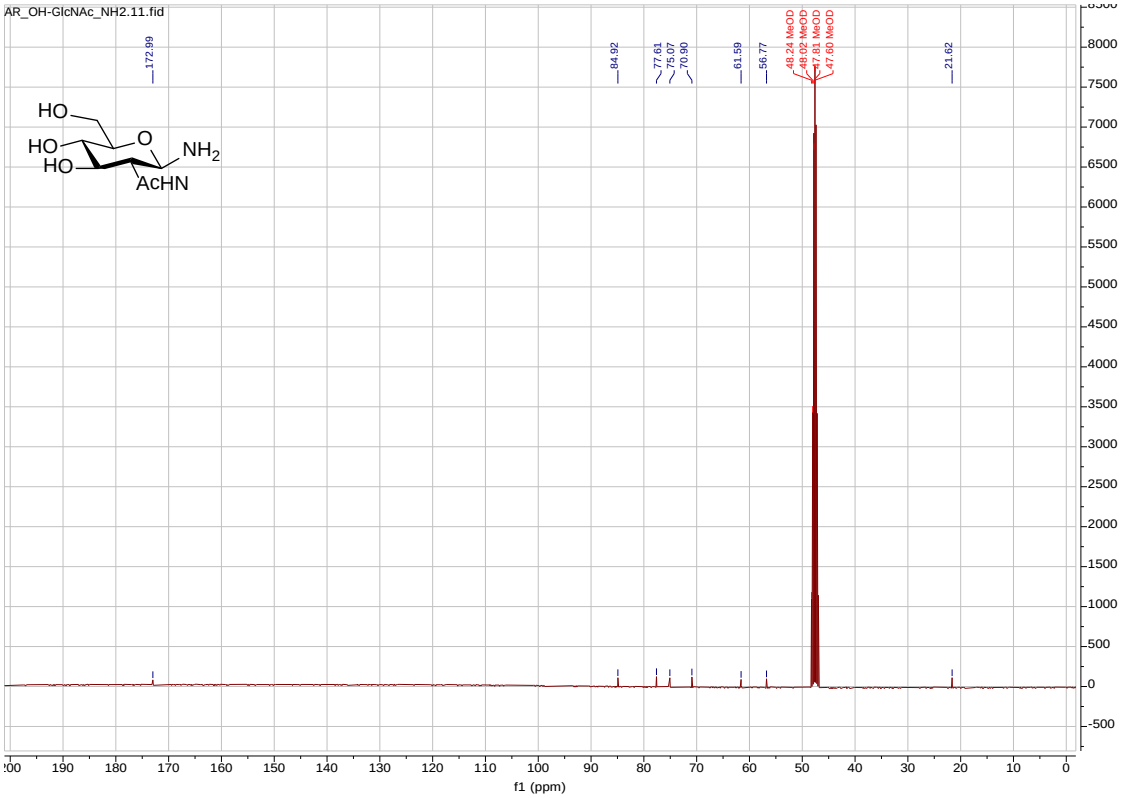

AR_Pseudoproline_OH_Crude.10.fid
Ot-Bu
FmocHN
H
N
O
O
OH
O
f1 (ppm)


AR_Pseudoproline_OH_Crude.11.fid
Ot-Bu
FmocHN
H
N
O
O
OH
O
f1 (ppm)


AR_GlcNAc(OH)_N3.11.fid
172.45
88.32
78.90
74.37
70.23
61.19
55.27
21.62
HO
HO
HO
O
N3
AcHN
f1 (ppm)


AR_GlcNAc(OH)_N3.10.fid
4.87 CD3OD
4.87 HDO
HO
HO
HO
O
N3
AcHN
1.00
1.02
1.08
0.73
0.98
1.42
1.28
2.90
f1 (ppm)


AR_OH-GlcNAc_NH2.10.fid
HO
HO
HO
O
NH2
AcHN
3.96
3.94
3.82
3.82
3.79
3.79
3.62
3.60
3.59
3.57
3.54
3.52
3.49
3.38
3.36
3.33
3.27
3.27
3.26
3.24
3.22
3.22
3.19
3.18
1.96
1.00
1.13
1.15
1.08
1.06
1.58
1.43
3.16
f1 (ppm)


AR_OH-GlcNAc_NH2.11.fid
172.99
84.92
77.61
75.07
70.90
61.59
56.77
48.24 MeOD
48.02 MeOD
47.81 MeOD
47.60 MeOD
21.62
HO
HO
HO
O
NH2
AcHN
f1 (ppm)