isas

technische universität dortmund

Synthesis of Glycopeptides for Exploration of Mucin-Type-Threonine and HexNAc-Tyrosine Modifications

DISSERTATION

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

der Fakultät Chemie und Chemische Biologie

der Technischen Universität Dortmund

Vorgelegt von

M. Sc. Chem. Biol.

Manuel Schorlemer

aus Anröchte.

Dortmund 2018

Die vorliegende Arbeit entstand im Zeitraum von April 2011 bis März 2018 unter Anleitung von Prof. Dr. Ulrika Westerlind an der Fakultät Chemie und Chemische Biologie der Technischen Universität Dortmund und dem Leibniz Institut für analytische Wissenschaften ISAS e.V. in Dortmund.

Erstgutachter:	Prof. Dr. Herbert Waldmann
Zweitgutachterin:	Prof. Dr. Ulrika Westerlind

Parts of this work have already been published in:

A Unified Strategy for the Synthesis of Mucin Cores 1-4 and the Assembled Multivalent Glycopeptides

C. Pett, M. Schorlemer, U. Westerlind, *Chem. Eur. J.* **2013**, 19, 17001-17010

Assignment of Saccharide identities through Analysis of Oxonium Ion Fragmentation Profiles in LC-MS/MS of Glycopeptides

A. Halim, U. Westerlind, C. Pett, M. Schorlemer, U. Rüetschi, G. Brinkmann, C. Sihlbom, J. lengquist, G. Larson, J. Nilsson, *J. Proteome Res.* **2014**, 13, 6024-6032

Distinctive MS/MS Fragmentation Pathway of Glycopeptide Generated Oxonium Ions Provide Evidence of the Glycan Structure

J. Yu, M. Schorlemer, A. Gomez Toledo, C. Pett, C. Sihlbom, G. Larson, U. Westerelind, J. Nilsson, *Chemistry* **2015**, 22, 1114-1124

Microarray Analysis of Antibodies Induced with Synthetic Antitumor Vaccines: Specificity against Diverse Mucin Core Structures

C. Pett, H. Cai, J. Liu, B. Palitzsch, M. Schorlemer, S Hartmann, N. Stergiou, M. Lu, H. Kunz, E. Schmitt, *Chem. Eur. J.* **2017**, 23, 3875-3884

Submitted papers:

Effective assignment of α2,3/α2,6 sialic acid isomers in LC-MS/MS based glycoproteomics

C. Pett, W. Nasir, C. Sihlbom, B.-M. Olsson, M. Schorlemer, R. Zahedi, G. Larson, J. Nilsson, U. Westerlind, *under review.*

Table of Contents

Abbreviations

AA	amino acid
Aβ	amyloid-β
Ac	acetyl
All	allyl
arom	aromatic
AspAT	aspartate amino transferase
Atp5a	ATP-synthase subunit α
Atp5b	ATP-synthase subunit β
Boc	*tert*-butyloxycarbonyl
BSA	bovine serum albumin
Bu	butyl
Bn	benzyl
c	concentration
calc.	calculated
cat.	catalytic
CAZy	carbohydrate active enzymes
CID	collision induced dissociation
CFG	Consortium for Functional Glycomics
COPD	chronic obstructive pulmonary disease
COSY	correlated spectroscopy
Cox4i1	cytochrome-c oxidase subunit 4 isoform 1
d	duplet, day
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCC	*N*,*N*'-dicyclohexylcarbodiimide
DCM	dichloromethane
dd	duplet of duplet
DHB	2,5-dihydroxy benzoic acid
DIC	*N*,*N*'-diisopropylcarbodiimide
DIPEA	diisopropylethylamine
DMAP	4-(dimethylamino(pyridine)
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
ELISA	enzyme-linked immunosorbent assay
eq.	equivalents
ESI	electron spray ionization
ETD	electron transfer dissociation
FA	formic acid
Fmoc	*N*-(9H-fluoren-9-yl)-methoxycarbonyl
Fuc	*L*-fucose
FliD	flagella hook associated protein 2
Gal	*D*-galactose
GalNAc	*N*-acetyl-*D*-galactosamine
gg	gauche-gauche
Glc	D-glucose
GlcNAc	*N*-acetyl-*D*-glucosamine
GSL II	griffonia simplicifolia lectin II
gt	gauche-trans
GTP	Guanosine-5'-triphosphate
h	hours
HATU	*O*-(1H-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HBTU	*O*-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HMBC	heteronuclear multiple bond correlation
HOAt	7-aza-1-hydroxybenzotriazole
HOBt	1-hydroxybenzotriazole
hOGA	human O-GlcNAc hydrolase
hOGT	human O-GlcNAc transferase
HPLC	high performance liquid chromatography
HR	high resolution
HSQC	heteronuclear single quantum coherence
Hz	hertz
Ig	immunoglobulin
*i*Pr	isopropyl
J	coupling constant
KLH	Keyhole Limpet Hemocyanin
LacNAc	*N*-acetyllactosamine
LC	liquid chromatography
Le^x/Le^a	Lewis x/Lewis a
LG	leaving group
LWAC	lectin weak affinity chromatography
m	multiplet
M	molarity
MALDI	matrix assisted laser desorption ionization
mbar	millibar
MD	Molecular dynamics
Me	methyl
MeCN	acetonitrile
MGL	macrophage c-type lectin
MHC	major histocompatibility complex
min	minutes
MS	mass spectrometry
NCE	normalized collision energy

Neu5Ac	*N*-acetylneuraminic acid
NHS	*N*-hydroxysuccinimide
NIS	*N*-iodosuccinimide
NMR	nuclear magnetic resonance
NucB2	nucleobindin-2
OGA	O-GlcNAc hydrolase
OGT	O-GlcNAc transferase
p	para
PBS	phosphate buffered saline
PBST	PBS + Tween-20
PcrV	type III secretion system protein
PEG	polyethylene glycol
PG	protecting group
Ph	phenyl
pH	potentia hydrogenii
PMP	para-methoxy phenyl
pos	positive
ppGalNAcT	polypeptide *N*-acetylgalactosamine transferase
ppm	parts per million
p-TsOH	*para*-toluenesulfonic acid
Pyr	pyridine
q	quartet
Rbl2	retinoblastoma-like protein 2
R_f	retention factor
s	singulet
SLe^x/SLe^a	sialyl-Lewis x/sialyl Lewis a
SPPS	solid phase peptide synthesis
ST	sialyl-Thomsen-Friedenreich
STD	saturation transfer difference
ST_N	sialyl-T-antigen nouvelle
Su	succinimide
t	triplet
T	Thomsen-Friedenreich antigen
TBAF	*tert*-butylammonium fluoride
TBS	*tert*-butylammonium fluoride
*t*Bu	*tert*-butyl
TEG	triethylene glycol
tert	tertiary
TFA	trifluoroacetic acid
TfOH	trifluoromethanesulfonic acid
tg	trans-gauche
THF	tetrahydrofuran
TIPS	triisopropylsilance
TLC	thin layer chromatography
TMS	tetramethylsilyl
T_N	T-antigen nouvelle
TOCSY	Total correlation spectroscopy
Troc	trichloroethoxycarbonyl
Trt	trityl
UDP-GlcNAc	Uridine diphosphate *N*-acetylglucosamine
UV	ultraviolet
VNTR	variable number of tandem repeats
VVA	vicia villosa agglutinin
WGA	wheat germ agglutinin
α	specific optical rotation
δ	chemical shift
λ	wavelength

Amino acid codes

Ala	A	Alanine
Arg	R	Arginine
Asn	N	Asparagine
Asp	D	Aspartate
Cys	C	Cysteine
Gln	Q	Glutamine
Glu	E	Glutamate
Gly	G	Glycine
His	H	Histidine
Ile	I	Isoleucine
Leu	L	Leucine
Lys	K	Lysine
Met	M	Methionine
Phe	F	Phenylalanine
Pro	P	Proline
Ser	S	Serine
Thr	T	Threonine
Trp	W	Tryptophan
Tyr	Y	Tyrosine
Val	V	Valine

1 Abstract

Protein *O*-glycosylation is a structurally and biologically versatile modification. The mucin-type *O*-glycosylation on extracellular mucin proteins is for instance involved in cell-cell and cell-matrix interactions, and play key roles in airway inflammations of COPD, asthma and cystic fibrosis patients by serving as binding receptors for virulence factors of airway disease-related pathogens (bacteria and viruses). Insights into molecular details in the binding specificities of virulence factors like adhesion proteins (lectins) or sialidases would support the development of glycomimetic antiadhesive drugs or enzyme inhibitors to suppress infections. Moreover, adenocarcinoma cells (from for instance breast, colon, lung and ovary cancer) have been found to exhibit altered mucin glycosylation and expression levels, which has been related to tumor progression and metastasis. MUC1-based vaccine constructs have been identified as promising for induction of a tumor-directed immune response, which in consequence might eradicate the tumor burden. However, without a structurally defined MUC1 glycopeptide library, a quality control and a precise epitope mapping of the induced antibodies is not possible. This work focuses on the synthesis of a well-defined peptide library (including extended core 2 and core 4 glycosylation) for application in microarray experiments to map lectin/glycopeptide and antibody/glycopeptide binding specificities. The gained information will aid to tackle the above-described challenges in airway disease and cancer progression. Further, this work contributed in the development of a new oxonium ion-based mass-spectrometric methodology for improved structural characterization of glycopeptides.

O-glycosylation is commonly found on serine (Ser) or threonine (Thr), but recent identifications suggest that similar modifications also occur on tyrosine (Tyr). These modifications are initiated by addition of a HexNAc (GalNAc or GlcNAc) monosaccharide residue to the respective amino acid. A major challenge for identification of the recently discovered HexNAc-Tyr modifications was to distinguish the structural similar isomers αGalNAc-, αGlcNAc- and βGlcNAc-Tyr. Assignments of the glycan structures were only possible by general knowledge of biosynthetic pathways or being based on lectin glycan-binding preferences when applied in lectin-affinity enrichment of protein-digested peptide fragments. However, lectin binding-preferences of HexNAc-Tyr have not been explored and lectins are not always capable to discriminate between different glycan isomers. In previous glycoproteomic studies complex Tyr glycosylation was assumed as mucin type αGalNAc-Tyr modification, based on knowledge of lectin specificity to Ser/Thr glycopeptides and the fact that mucin-type glycopeptides are generally extended with additional glycans. The αGlcNAc-Tyr was found

on host GTPases, a modification caused by intracellular infections and was assigned by enzymatic in vitro studies and NMR structural characterization. Further, in an extensive glycoproteomic study several HexNAc-Tyr modifications were found on mitochondrial proteins and some of the identified sites are also known phosphorylation sites. Therefore, the identified HexNAc-Tyr modifications were likely to be βGlcNAc-Tyr (although to date this modification is still not verified). In order to enable identification of HexNAc-Tyr as a possible modification by glycoproteomic analysis, lectins commonly used for glycopeptide affinity enrichment were evaluated towards binding recognition of HexNAc-Tyr glycopeptides compared to Ser/Thr analogs. To this extent, a versatile library of HexNAc model glycopeptides was synthesized, providing the basis for the lectin binding studies. The contribution of structural elements towards recognition by lectins commonly used for glycopeptide enrichment (Vicia villosa agglutinin - VVA, Wheat germ agglutinin - WGA and Griffonia simplicifolia lectin II - GSL II) was tested by using glycopeptide microarrays. The binding recognition influence of glycan (αGalNAc/αGlcNAc/βGlcNAc), acceptor amino acid (Tyr/Ser/Thr) and mono vs divalent glycan presentation was evaluated. The WGA and GSL II lectins showed a clear preference for Tyr glycopeptides over Ser/Thr. In contrast no clear Tyr preference was found for VVA. The WGA lectin showed strong binding to both GlcNAc- and GalNAc-Tyr glycopeptides and could not be considered for structural assignment of HexNAc-Tyr glycopeptides. In contrast VVA was highly specific for αGalNAc-containing glycopeptides and GSL II recognized both αGlcNAc- and βGlcNAc modified peptides. In analysis of WGA the impact of divalent glycan presentation was evaluated and found to be dependent on the peptide spacing between the glycans. A distance of 2-5 amino acids resulted in a generally enhanced recognition. Taken together, these results will enable a more reliable application of lectins for HexNAc-Tyr glycopeptide enrichment. Further, mass spectrometric oxonium ion fragmentation was elucidated for applicability towards HexNAc-Tyr glycopeptide identification. The methodology was found to be fully applicable for discrimination of GalNAc vs GlcNAc epimers, but not for assignment of the anomeric configuration. The knowledge about lectin affinities in combination with the oxonium-based glycan identification will enable large-scale glycoproteomic studies taking account of Tyr *O*-glycosylation, which eventually might help to clarify, whether βGlcNAc-Tyr is actually an existing modification. However, the ultimate proof of existence of the modification would be the observation of Tyr-specificity by the βGlcNAc biosynthetic enzymes O-GlcNAc transferase (OGT) and O-GlcNAcase (OGA). Accordingly, selected peptides and glycopeptides were used as substrate candidates for OGT and OGA. Thereby, OGA showed a clear preference for βGlcNAc-Tyr substrates over Ser/Thr analogs. The OGT substrate studies were inconclusive and needs further evaluation.

2 Zusammenfassung

Die Protein *O*-Glykosylierung ist eine biologisch sowie strukturell vielseitige Modifikation. Die *O*-Glykosylierung an extrazellulären Mucinproteinen ist z.B. involviert in Zell-Zell- sowie Zell-Matrix-Interaktionen und spielt eine Schlüsselrolle bei Atemwegsinfektionen von COPD-, Mukoviszidose- und Asthmapatienten. Dabei dienen die *O*-Glykane als Rezeptoren für Virulenzfaktoren von Bakterien und Viren. Detailliertere Einblicke in die Bindungsspezifitäten von Virulenzfaktoren wie Lektine oder Sialidasen würde die Entwicklung von glykomimetischen Antiadhäsionsmedikamenten und Enzyminhibitoren zur Infektionsbekämpfung unterstützen. Zudem wurde gezeigt, dass Adenokarzinomzellen veränderte Mucinglykosylierung und -expression aufweisen, was wiederum mit Tumorwachstum und Metastasierung in Verbindung gebracht wird. MUC1-basierte Vakzinkonstrukte gelten als vielversprechend, um eine tumorgerichtete Immunantwort zu induzieren, jedoch ist eine Qualitätskontrolle sowie präzise Aufklärung der induzierten Antikörper ohne eine Bibliothek bestehend aus MUC1-Gklykopeptiden definierter Struktur unmöglich. Diese Arbeit konzentriert sich auf die Synthese einer wohldefinierten Peptidbibliothek (mit verlängerter Core 2- und Core 4-Glykosylierung) zur Anwendung in Microarrayexperimenten um schließlich Lektin/Glykopeptid- und Antikörper/Glykopeptidspezifitäten aufzuklären und somit bei der Lösung der o.g. Herausforderungen bzgl. Atemwegserkrankungs- und Krebsverlauf zu helfen. Zudem trug diese Arbeit zur Entwicklung einer neuen massenspektrometrischen Methode zur verbesserten Charakterisierung von Glykopeptidstrukturen bei.

Jüngste Identifikationen zeigten, dass *O*-Glykosylierungen auch auf Tyr stattfinden können. Eine Herausforderung bei den kürzlich entdeckten HexNAc-Tyrosin-Modifikationen war die Unterscheidung zwischen den Isomeren αGalNAc-, αGlcNAc- und βGlcNAc-Tyr. Eine Zuordnung der Glykanstrukturen war nur möglich durch das generelle Wissen um Biosynthesewege oder beruhend auf bekannten Lektin/Glykan-Bindungspräferenzen zur Lektinaffinitätsanreicherung. Jedoch sind die Lektinbindungspräferenzen ggü. HexNAc-Tyr bislang unerforscht. Zudem können Lektine nicht immer zwischen verschiedenen Glykanisomeren unterscheiden. In bisherigen Studien wurde komplexe Tyr-Glykosylierung als mucinartige αGalNAc-Tyr-Modifikation betrachtet, basierend auf bekannten Lektinspezifitäten ggü. Ser/Thr-Glykopeptiden und der Tatsache, dass mucinartige Glykopeptide generell um weitere Glykane verlängert vorliegen. αGlcNAc-Tyr wurde an Host-GTPasen gefunden. Eine Glykanzuordnung fand mittels enzymatischen In-Vitro-Studien sowie NMR-basierten Strukturbestimmungen statt. Weitere HexNAc-Tyr-Modifikationen wurden im Rahmen einer

umfangreichen Glykoproteom-Studie an mitochondriellen Proteinen gefunden, wobei einige der identifizierten Glykosylierungsstellen gleichzeitig bekannte Phosphorylierungsstellen darstellen. Daher handelt es sich hier wahrscheinlich um βGlcNAc-Tyr (obwohl diese Modifikation bis dato nicht verifiziert werden konnte). Um eine Identifikation von HexNAc-Tyr als mögliche Modifikation durch Glykoproteomanalysen zu ermöglichen, wurden Lektine, welche üblicherweise zur Glykopeptidaffinitätsanreicherung verwendet werden, hinsichtlich ihrer Bindungserkennung von HexNAc-Tyr und im Vergleich zu dessen Ser/Thr-Analoga, untersucht. Zu diesem Zweck wurde eine vielseitige Bibliothek bestehend aus HexNAc-Modellglykopeptiden synthetisiert, welche die Grundlage für Lektinbindungsstudien bildet. Der Beitrag von Glykopeptidstrukturelementen zur Erkennung durch üblicherweise zur Glykopeptidanreicherung verwendete Lektine (Vicia Villosa Agglutinin - VVA, Wheat Germ Agglutinin – WGA und Griffonia Simplicifolia Lectin II – GSL II) wurde anhand von Glykopeptidmicroarrays getestet. Dabei wurde der Einfluss von Glykan (αGalNAc/αGlcNAc/βGlcNAc), Akzeptoraminosäure (Tyr/Ser/Thr) sowie die Art der Glykanpräsentation (mono- vs divalent) untersucht. WGA und GSL II, nicht aber VVA, zeigten eine deutliche Präferenz für Tyr-Glykopeptide ggü. Ser/Thr. WGA zeigte eine starke Bindung sowohl an GlcNAc-Tyr- als auch an GalNAc-Tyr-Glykopeptide und kommt somit für eine Strukturzuordnung von HexNAc-Tyr nicht in Betracht. Im Gegensatz dazu wurde für VVA eine hohe Spezifität für ausschließlich αGalNAc-Glykopeptide gefunden. GSL II erkannte sowohl αGlcNAc- als auch βGlcNAc-modifizierte Peptide. Für WGA wurde ein längenabhängiger Einfluss von divalenter Glykanpräsentation beobachtet, wobei eine Distanz von 2-5 Aminosäuren in erhöhter Glykopeptiderkennung resultierte. Zusammengenommen ermöglichen diese Ergebnisse eine verlässlichere Verwendung von Lektinen zum Zwecke der HexNAc-Tyr-Glykopeptidanreicherung. Des Weiteren wurde die o.g. massenspektrometrische Oxoniumfragmentierung hinsichtlich der Anwendbarkeit zur HexNAc-Tyr-Glykopeptididentifikation untersucht. Es ergab sich volle Anwendbarkeit zur Unterscheidung von GalNAc vs GlcNAc, jedoch nicht zur Zuordnung der jeweiligen anomerischen Konfiguration. Das hier generierte Wissen über Lektinaffinitäten in Kombination mit der oxoniumionenbasierten Glykanidentifikation wird großangelegte Glykoproteomstudien ermöglichen, in welchen die Tyr-*O*-Glykosylierung mit in Betracht gezogen wird. Dies könnte schließlich zur Klärung beitragen, ob βGlcNAc-Tyr eine tatsächlich existierende Modifikation ist. Der ultimative Beweis für dessen Existenz wäre jedoch die Beobachtung einer Tyr-Spezifität ausgehend von den βGlcNAc-Biosyntheseenzymen O-GlcNAc-Transferase (OGT) und O-GlcNAcase (OGA). Entsprechende, hier durchgeführte Inkubationen von Substratkandidaten mit OGA ergaben eine deutliche OGA-Präferenz ggü. βGlcNAc-Tyr-Substraten im Vergleich zu Ser/Thr-Analoga. Entsprechende Inkubationen mit OGT verliefen bislang ergebnislos und benötigen weitere Untersuchungen.

3 Introduction

3.1 *O*-glycosylation in eukaryotes

Among post-translational modifications (PTMs), protein glycosylation is one of the most abundant and versatile of its kind. Depending on the linkage connection protein glycosylation is divided into two main classes, the *N*-glycans, which are attached to asparagine (Asn) residues through an amide bond and the *O*-glycans that are linked to serine (Ser), threonine (Thr) or tyrosine (Tyr) amino acids forming a *O*-glycosidic bond (an acetal). *N*-glycans are divided into three major classes, the high-mannose, hybrid and complex type, which all contain a common pentasaccharide ($Man_3GlcNAc_2$) core structure. In *N*-glycan biosynthesis a presynthesized triantennary tetradecasaccharide ($Glc_3Man_9GlcNAc_2$) is transferred from a dolichol pyrophosphate glycolipid to a newly synthesized protein that often contain the Asn-Xaa-Ser/Thr motif[1]. Enzymatic processing in ER and Golgi are then followed to form more mature *N*-glycans.

The *O*-glycans are a more diverse group of modifications and among them the mucin-type *O*-glycosylation, attachment of α-N-acetylgalactosamine (α-GalNAc) to Ser/Thr/Tyr, and *O*-GlcNAcylation, attachment of the structural epimer β-N-acetylglucosamine (β-GlcNAc) to Ser/Thr, are most common.

Mucin-type *O*-glycosylation is the dominant glycan species found on mucin glycoproteins although the appearance is not limited to this group of proteins[2,3]. The biosynthesis is started by α-*N*-acetylgalactosamine (GalNAc) transfer to either Ser or Thr (the pathway for Tyr glycosylation is not explored), resulting in α-GalNAc-*O*-Ser/Thr, also known as *T-antigen nouvelle* or T_N-antigen[4]. The initial GalNAc transfer takes place under consumption of a uridine-diphosphate-*N*-acetylgalactosamine (UDP-GalNAc) molecule and is mediated by one of over twenty different tissue specific polypeptide *N*-acetylgalactosaminyl transferases (ppGalNAcTs), which are located in the luminal compartment of the rough ER or Golgi[45-11]. Some of the ppGalNAcT isoenzymes contain ricin-like lectin domains[12] that recognize *O*-GalNAc residues, which determines the site selectivity and glycan density on mucin polypeptide tandem repeats[13].

After GalNAc transfer, the glycoproteins are translocated via Golgi stacks to various destinations. During this transport, the glycans are further elongated inside the Golgi stacks to linear or branched structures by a set of thirty different glycosyl transferases[14-16]. These structures are divided into eight classes according to their core motifs[17] of which only core 1-4 are frequently found in nature while

core 5-8 are rather uncommon (***Figure 1***). Elongation of the T_N-antigen by the core 1 β-1,3-galactosyl transferase (core1 β3GalT) forms the core 1 motif, also called *Thomsen-Friedenreich-* or *T-antigen*. Transfer of a β-GlcNAc to the 6-OH of GalNAc at the T-antigen acceptor by the core 2 β-1,6-*N*-acetylglucosaminyl transferase I-III (core 2 β6GlcNAcT I-III) or the I-branching enzyme (IGnT) generates the core 2 structure. Alternatively, the T_N-antigen can be extended at C-3 by the core 3 β-1,3-*N*-acetylglucosaminyl transferase (core 3 β3GlcNAcT) forming the core 3 structure, which in turn can be elongated at GalNAc C-6 by the core 4 β-1,6-*N*-acetylglucosaminyl transferase (core 4 GlcNAcT) or core 2 β6GlcNAcT II or IGnT to form the core 4 structure.[18-19]

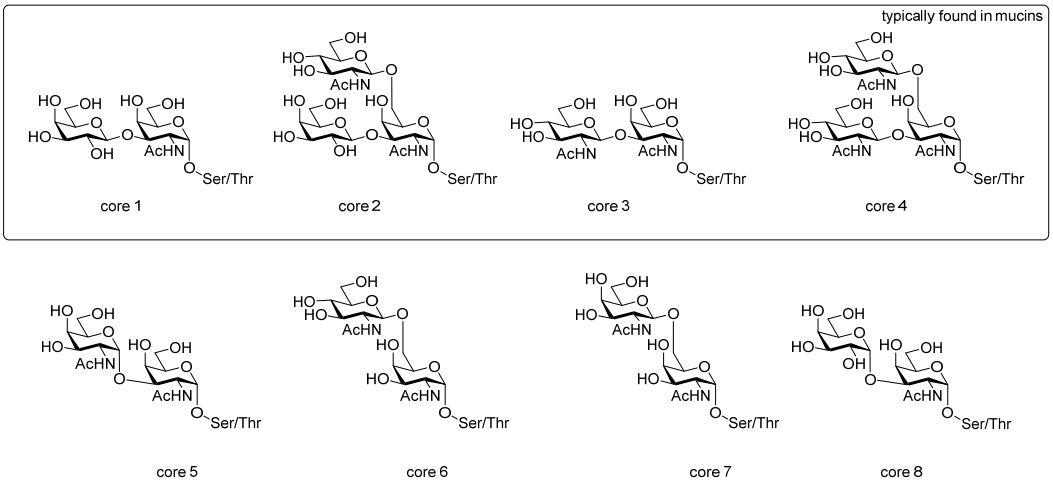


Figure 1: Overview of the *O*-glycan core structures.

The obtained mucin-type *O*-glycan core structures can be elongated by alternating Gal or GlcNAc residues in a linear or branched manner, e.g. elongation by one or several Galβ1,3-GlcNAcβ (LacNAc type-1) or Galβ1,4-GlcNAcβ (LacNAc type-2) repeats. This occurs more often at the GlcNAcβ1-6 residue of core 2 or core 4 structures and is catalyzed by a family of isoenzymes (β3GalT[20-21], β4GalT[22-25], iGnT[26] and IGnT[27-28]). Finally, the glycan chains can be sialylated, fucosylated, sulfated, methylated or acetylated[29], forming a considerable variety of glycan structures including histoblood group antigens such as the A, B or H or Lewis antigens ($Le^{a/b}$/$SLe^{a/b}$ or $Le^{x/y}$/$SLe^{x/y}$, ***Figure 2***)[30]. Among the terminal modifications, sialylation and sulfation provide a negative charge to glycoproteins and fucosylation increase the hydrophobicity. On mucins these terminal modifications contribute dramatically to physical and biological properties[19]. For instance, the mucosal glycobiome varies within different tissues depending on the different functional requirements of the mucosal surfaces and is altered in airway diseases and cancer as described in more detail in chapter 3.2[31, 18, 32].

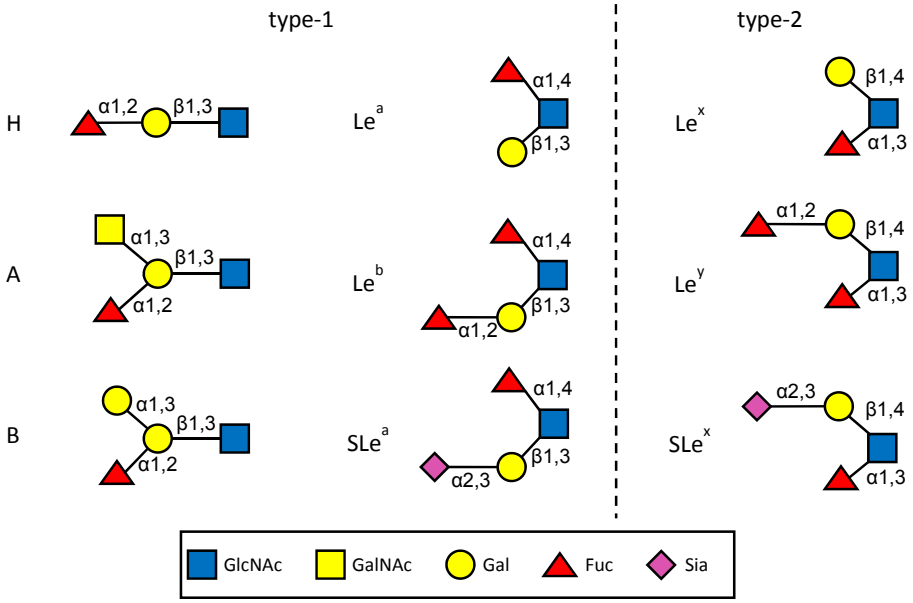


Figure 2: Selection of terminal antigenic determinants of mucin-type O-glycans. Shown are the histo blood group antigens H, A and B as well as the Lewis type antigens. Le^a and Le^b are mainly found in surface epithelium while Le^x and Le^y are commonly expressed in mucous cells.[18, 30]

The above described mucin-type *O*-glycosylation is usually found on proteins at the extracellular side of the cell plasma membrane being involved in important cell-cell and cell-matrix interactions. In contrast, another common group of *O*-glycosylation, *O*-GlcNAcylation, the transfer of a β-GlcNAc monosaccharide to Ser or Thr residues of a protein, is an intracellular modification that usually takes place in the nucleus[33-36], cytoplasm[37] or mitochondria[38-39]. This is mediated by one single human enzyme, the *O*-linked *N*-acetyl-glucosaminyl transferase (OGT)[37] under consumption of a UDP-GlcNAc molecule[40]. In turn, β-GlcNAc hydrolysis is mediated by the OGT counterpart *O*-GlcNAcase (OGA)[41]. Both enzymes are ubiquitously expressed and distributed in most mammalian tissues. As a consequence, *O*-GlcNAc is one of the most abundant PTMs within the nucleocytoplasmic compartments of mammalian cells. However, to date only in one case mammalian *O*-β-GlcNAcylation has been observed at the extracellular side of the cell plasma membrane[42]. However, on protozoa extracellular α-GlcNAcylation seems to be a common modification[43]. β-*O*-GlcNAcylation is a nutrient-responsive modification and is involved in the regulation of various processes[44] like signal transduction[45], cell cycle[46-47], gene expression[48-49] and proteasomal degradation[50-51].

Unlike α-GalNAc, the β-GlcNAc modification is usually not elongated or further modified and thus exists as monomeric glycan. A specific characteristic of β-*O*-GlcNAcylation is to show some

resemblance to protein *O*-phosphorylation. Both, β-*O*-GlcNAcylation and *O*-phosphorylation take place on Ser and Thr (and probably Tyr) and often show rapid cycling, commencing at a higher rate than protein turnover[52-53]. In addition, the enzymes responsible for β-GlcNAc cycling are regulated by *O*-phosphorylation and β-*O*-GlcNAcylation[37] like kinases and phosphatases[54]. Further, cross-talk of *O*-phosphorylation and β-*O*-GlcNAcylation is evident, as mammalian cells with attenuated *O*-GlcNAcylation show altered *O*-phosphorylation levels[55], in some cases *O*-GlcNAcylation could even be shown to be directly reciprocal to phosphorylation[56-57]. However, while phosphorylation could be linked to certain primary sequences, no such consensus motif has been identified for *O*-GlcNAc, although many *O*-GlcNAc sites contain nearby proline or valine residues[58-59]. Even though *O*-GlcNAcylation is distributed throughout the whole nucleocytoplasmic part of the cell, particularly high *O*-GlcNAc levels are found for example on nuclear pore proteins[33, 60] and the nuclear transcriptional regulatory machinery[36]. These regulatory proteins, are known to contribute to oncogenic phenotypes, which in turn suggests an involvement of *O*-GlcNAcylation in cancer. A prominent example is the *c-myc* oncogene, which is known to be *O*-GlcNAc modified in its transactivation domain and represents the mutational hotspot in many Burkitt's lymphomas patients[57, 61]. *O*-GlcNAcylation gains even more clinical relevance as it is hypothized[62], that a decreased glucose metabolism in aging neurons leads to reduced *O*-GlcNAcylation of central cytosolic or nuclear proteins. This might then result in abnormal phosphorylation of many proteins, which eventually results in neurodegenerative diseases. Finally, high cellular glucosamine levels (hyperglycemia) seem to promote insulin resistance (diabetes type 2)[63-64]. This might be explained with the fact, that glucosamine is a precursor in the UDP-GlcNAc formation, which might lead to hyper-*O*-GlcNAcylation of transcription factors or signaling proteins, which are involved in the insulin response[65]. In a few recent studies it was further shown that many of the metabolic enzymes involved in glycolysis and oxidative phosphorylation are highly *O*-GlcNAc modified, which probably is connected to functional regulation of these enzymes[66-67].

3.2 Mucin glycopeptides

In mammals, the epithelial surfaces of for instance the respiratory, reproductive and gastrointestinal tracts are coated by a mucus layer, an aqueous gel comprising ions, proteins, glycoproteins and lipids. This gel serves for lubrication and in addition, is a part of the innate immune system, that protects the host against pathogens and toxins[68]. The major constituent of the mucus layer are the secreted mucin glycoproteins for instance MUC2, MUC5AC, MUC5B and MUC6[69] that are covering the epithelial tissues. The secreted mucins constitute a mass amount of 10-50% (wt/wt) of the complete aqueous networks, which are held together by numerous disulfide bonds[70]. Below the mucus layer, the

membrane mucins are covering the epithelial cells, for instance MUC1, MUC4 and MUC16. Due to their extended (rod-like) conformation, these mucins are extended further from the membrane surface than most cell surface receptors, which enables a disruption of cell and pathogen adhesion[71-72]. Further they play important roles in the organization of cells and their cell surface-proteins (cell polarization)[73] and are involved in a number of signaling events related to for instance cellular motility, growth[74-75], adhesion[76] and proliferation[77] as well as T cell activation[78].[79]

Mucins exhibit a high and variable number of proline-, serine- and threonine-rich tandem repeats (VNTRs), which serve as scaffolds for dense *O*-glycosylation leading to 'bottle-brush' structures with a carbohydrate mass fraction of 50-90 % with respect to the whole mucin glycoproteins. With the different degrees of glycosylation and the variable number (resulting in variable glycoprotein length), the mucin VNTRs influence the biophysical properties of the mucus[80-82]. The mucins are classified according to the MUC protein backbones which are encoded by over twenty different genes and every mucin gene product shows a characteristic and unique tandem repeat sequence. The mucin genes are abbreviated with MUC, followed by a number, which reflects the order that the proteins were cloned for the first time. According to these Human Genome Mapping conventions, the MUC1 gene, which codes for a membrane tethered mucin glycoprotein, was the first mucin gene to be cloned. The corresponding glycoprotein exhibits a large extracellular *N*-terminal domain, comprising a signal sequence, which is responsible for mucin peptide insertion to the ER and translocation to the cytoplasm membrane. Further, the *N*-terminal domain contains 25 to 125 tandem repeats[83] with the sequence PAHGVTSAPDTRPAPGSTAP and a C-terminal region consisting of a hydrophobic membrane-spanning domain as well as a short phosphorylated cytoplasmic tail that is involved in signaling events. The membrane bound MUC1 glycoprotein is commonly expressed at the apical cell surface of normal epithelial tissues[84]. Besides the function of mucins in protection and lubrication, antibody staining experiments with the complete mucus gel have suggested that MUC1 is involved in the tethering of the mucus gel to the mucosal surface[85-86]. Furthermore, MUC1 play many important roles in cell adhesion lectin interactions of pathogens[72, 87], immune cells and tumor cell metastasis[19, 32, 84]. Additionally, the *O*-glycan structure, site specificity and density and the number of MUC1-VNTRs varies within different tissues and disease stages, which further influence the properties and functions of MUC1[88-90].

Besides MUC1 further mucins have received ongoing attention by researchers due to their involvement in chronic diseases: For instance increased expression of MUC4 has been linked to pancreatic cancer[91]. MUC5AC has been identified as marker for colorectal[92] and pancreatic cancer[93],

but altered MUC5AC levels have also been related to gastric cancer[94-95]. Decreased MUC5AC and MUC5B levels have been found in cystic fibrosis airways[96], while hypersecretion of these mucins have been found in the sputum of COPD patients[97]. Finally, MUC5AC and MUC6 serve as receptors for *Heliobacter pylori*, which enables colonization of the human gastric epithelium and thus promotes gastric inflammation[98].

3.2.1 Mucin glycoproteins in airway diseases

Mucins play an important role in airway diseases like asthma, chronic obstructive pulmonary disease (COPD), or cystic fibrosis (CF). If challenged with environmental toxins, allergens or infectious pathogens, a healthy respiratory tract reacts with the activation of lung immune response mediators of which some induce mucin hypersecretion by activation of a secretory cascade[19]. These hypersecreted mucins could entrap particles[99] or bind receptors of inflammatory cells via MUC domains or glycans[100]. Typically, these hypersecretions decrease to baseline levels after a couple of days, which might be connected to anti-inflammatory mediators[101]. However, patients with chronic airway diseases show a permanent overproduction of the airway mucins MUC5AC and MUC5B[97] caused by *goblet cell hyperplasia* and upon external insults, which mainly are triggered by bacteria and virus infections (in 75% of the cases), acute attacks (exacerbations) take place which result in shortage of breath and pulmonary obstruction. Every exacerbation leads to an advancement in disease stage and an increase in frequency of the attacks, which eventually lead to death. Even if the general symptoms in chronic pulmonary diseases are similar, the origin of these diseases vary. The fundamental cause of asthma is not known, but a combination of genetic predisposition and environmental exposure to inhaled substances that trigger allergy are considered major factors. One theory is that an imbalance in the T helper type 2 (Th2) cellular cytokine response is involved in the development of an asthma phenotype[102-103]. Cystic fibrosis is caused by mutations in the CFTR gene[104], which leads to a malfunction or failure of the chloride channel protein, resulting in a chloride ion imbalance between extra- and intracellular fluid. In consequence, the osmotic pressure leads to a water drain from the extracellular fluid, which increases its viscosity and eventually clogs the lung capillaries. Development of COPD is caused by repetitive insults of the airways triggered by extensive exposure to tobacco smoke[105] or in low income countries by pollution by biomass fuels from cooking and heating[106]. In addition to hypersecretion, characteristic changes in *O*-glycosylation can be attributed to pulmonary infections and inflammations. For example, increased fucosylation, sialylation and sulfation has been found in secreted mucins from CF patients[107-108], while the membrane bound mucins of these patients have been found to exhibit decreased sialylation[109-111]. Increased expression

of SLe^x epitopes, GlcNAc-6-sulphate and Gal-3-sulfate has been found in sputum of CF and COPD patients[112], which is in accordance with an observed shift from α2,6-sialylation to α2,3-sialylation on MUC5B from CF patients[113]. Glycans on mucins often serve as ligands for invading bacteria or viruses and the altered glycosylation pattern might promote infections.

Pseudomonas aeruginosa, the most prominent pathogen responsible for morbidity and mortality in CF, for instance often recognizes SLe^x and 6-sulfo-SLe^x [114]. To date, a few adhesins are known which serve as virulence factors for *P. aeruginosa* by mediating the bacterial adhesion to glycan structures. For example the type-4 pilus was found to recognize a GalNAcβ(1-4)Gal disaccharide motif as found internally in fucosylasialo-GM1 and asialo-GM1 as well as terminally in asialo-GM2[115-116]. LecB, which is secreted by *P. aeruginosa*, binds to fucosylated oligosaccharides and in particular to Le^b epitopes[117]. The flagellar cap protein, FliD, has been found to mediate *P. aeruginosa* adhesion, SLe^x and Le^x epitopes seem to be involved in these interactions[118]. The protein PcrV is a structural component of the type III secretion system of *P. aeruginosa* and serves as a virulence factor. This was demonstrated by neutralization experiments[119] which led to significantly reduced airway inflammation. Glycan structures are believed to be involved in PcrV adhesion, but to date the specific ligands are still unknown.

Adherence of *Haemophilus influenzae*, the most common pathogen involved in COPD, proceeds via fimbriae. Here, its subunit HifA could be identified to contain the adhesive domain for mucin binding but also in this case the underlying glycan ligand motifs are not identified yet.

For *Streptococcus pneumoniae* the adhesin PsrP (pneumococcal serine-rich repeat protein) was found to bind keratin10 in lung cells[120-121]. This particular binding was shown to be independent of a lectin-carbohydrate interaction, however *S. pneumoniae* has been found to interact with mucin glycoproteins and it is reasonable to expect that binding to glycans take place by other adhesin proteins.

Given the crucial role of adhesins in in the process of bacterial infections, a deeper understanding of the corresponding lectin specificities could pave the way for the development of antiadhesive drugs as part of an efficient pathogen-specific antivirulence therapy.

The above described bacterial adhesion via cell surface glycoproteins and glycolipids is further supported by sialidases (also known as neuraminidases), which are expressed by the major airway disease-related pathogens *P. aeruginosa*, *H. influenzae* and *S. pneumoniae*[122-126]. These enzymes catalyze the hydrolysis of terminal sialic acid residues, which can be metabolized and serve as nutrition for the invading pathogen or alternatively, be implemented to the bacterial glycocalyx to mimic the

host sialylation pattern and therefore deceive the immune system. Further, the host sialic acid removal reveals underlying glycan structures which sometimes can be recognized by corresponding bacterial cell surface receptors (adhesins)[127]. The bacterial sialidases are also involved in biofilm formation[122, 125]. In this case the desialylation results in physical changes of charge and hydrophobicity of the environment. Being involved in numerous virulence mechanisms, sialidases represent important and promising therapeutic targets. Knowledge about the substrate specificities of different sialidases would be of high value for the development of inhibitors, however past studies on sialidases were based on bacterial genetic analyses and specific sialidase substrate specificities remain to be elucidated.[127]

3.2.2 Mucins in cancer

In a plethora of epithelial cancers aberrant mucin *O*-glycosylation and altered mucin expression levels have been observed, for instance in breast-, colon-, gastric and lung cancer[70]. By influence on cell-cell- and cell-matrix-interactions as well as cell-recognition, trafficking and signaling, changes in *O*-glycosylation have an impact on tumor progression and metastasis[128]. A common characteristic of *O*-glycosylation in epithelial cancer tissue is that the glycans are truncated and short saccharides are formed with increased terminal sialylation compared to glycans in corresponding non-cancerous tissues[129-130]. On the mucin MUC1 which is highly overexpressed on epithelial cancer cells, the aberrant glycosylation additionally lead to an exposure of the immunodominant peptide backbone and specific glycopeptide epitopes are formed. This in turn enables induction and antibody recognition of unique glycopeptide structures, as exploited in the detection of aberrant presentation of MUC1 in breast carcinomas by SM-3 and DF-3P[84, 131-132]. Common glycans on cancer mucins are T_N-, Sialyl-T_N-, T-, Sialyl-T- and Le^x-antigens[133], whose dominance has several reasons. For example mutations in *Cosmc*, a chaperone responsible for the correct folding of the core-1-galactosyltransferase (C1GalT)[134], leads to the expression of the T_N-antigen in tissues with high core 1 and core 2 expression[135]. Alternatively, if α-*N*-acetylgalactosaminide α-2,6-sialyltransferase 1 (ST6GalNAc-I) is expressed, mutation of *Cosmc* results in formation of Sialyl-T_N instead. Another reason for high T_N-, ST_N-, T- and ST-antigen expression is the stimulation of the proto-oncogene *Src*, which regulates trafficking between *Golgi* and *Endoplasmatic Reticulum*. While mucin-type *O*-glycosylation is mainly initiated in the *Golgi*, activation of Src results in the relocation of ppGalNAc-Ts to the ER[136]. The resulting increased T_N-antigen density on the affected mucins blocks further glycosylation by C1GalT and C2GlcNT, which would facilitate the generation of core 1 and core 2 otherwise. However, the most common reason for aberrant mucin *O*-

glycosylation lies within altered expression levels of the different glycosylatransferases. These changes are versatile and dependent on the different cancerous tissues.[137]

On the glycopeptide level an overexpression of Mucins has been found in tumors, especially in adenocarcinomas. This correlates with a high risk of metastasis and as a consequence to lowered chances of survival. The mucin overexpression has been found to influence tumor cells in numerous approaches, for instance by cell growth, transformation, invasion, differentiation and adhesion. Here, MUC1 is a very prominent example, being massively overexpressed in cancer and not restricted to protein localization at apical surfaces[137]. Instead it is additionally found in other sub-cellular compartments and on the whole surface of tumor cells[84].

3.2.2.1 The role of galectin-3 in cancer

MUC1 overexpression plays a key role in the galectin mediated tumor survival and metastasis. Galectins are a family of lectins, which share affinity for β-galactosides[138]. Galectins are not membrane-bound and are localized in the nucleus, the cytosol as well as secreted in the extracellular space. They are involved in a number of cellular functions, for instance in cell adhesion to the extracellular matrix, signaling, cell cycle progression or apoptosis[139]. One of the most-studied galectins is galectin-3, which shows altered expression states in various carcinomas compared to normal cells[140]. The high amount of galectin-3 secreted from tumor cells has shown to be immunosuppressive towards T- and B-lymphocytes by induction of apoptosis of the affected immune cells upon galectin-3 exposure (***Figure 3***, mechanism **A**)[141-145]. Here, a high significance of lymphocyte *O*-glycan ligands have been found for the galectin-3 recognition[146]. In contrast tumor cells are protected from galectin-3 induced apoptosis by altered glycosylation. After sialidase treatment of colon carcinoma cells and subsequent contact with galectin-3, apoptosis was induced on those cells[147]. In these cells, the galectin-3 receptors were found to be heavily α2,6-sialylated. Consequently, it was proposed, that overexpression of galectin-3 is advantageous for the secreting tumor cells in several ways. It enables immune evasion from B- and T-cells via induction of apoptosis on those cells, the tumor itself is protected from this mechanism (***Figure 3***, mechanism **B**). In addition, the interplay of extracellular galectin-3 with overexpressed, altered and membrane-bound MUC1 supports the metastasis of primary tumors. Cancer associated MUC1 has been found to be a natural target for galectin-3[148] via the tumor-associated T-antigen[149]. It is proposed, that overexpressed MUC1 on circulating tumor cells is captured by galectin-3, leading to clustering of MUC1. This might reveal shorter cell adhesion molecules on the cell surface, which are then able to interact with corresponding binding partners on endothelial cells. This facilitates settling of the circulating tumor cell to the endothelium and in consequence mediates

metastasis (***Figure 3***, mechanism **C**)[149]. Alternatively, circulating cells can interact with each other, forming circulating tumor cell clusters. This cell adhesion protects the aggregated tumor cells from anoikis, an induced apoptosis by the loss of cell-cell or cell-matrix interactions[150-151]. The tumor cell clusters can – like the single tumor cells – settle to the surface of endothelial cells (***Figure 3***, mechanism **D**)[152].[153]

The mechanisms of metastasis promotion are also shared by other galectins[154]. Furthermore, the galectin-3 expression is enhanced by the cytosolic part of MUC1 via stabilization of the galectin-3 mRNA transcript. Accordingly, the tumor characteristic MUC1 overexpression directly correlates with the galectin-3 expression. In turn, galectin-3 binds to MUC1 N-glycosylation, bridging MUC1 to ERB1, which integrates MUC1 to EGFR signaling[155]. Summarized, galectins play an important role in the tumor progression of many carcinomas and are therefore an interesting subject for detailed studies of lectin/carbohydrate interactions.

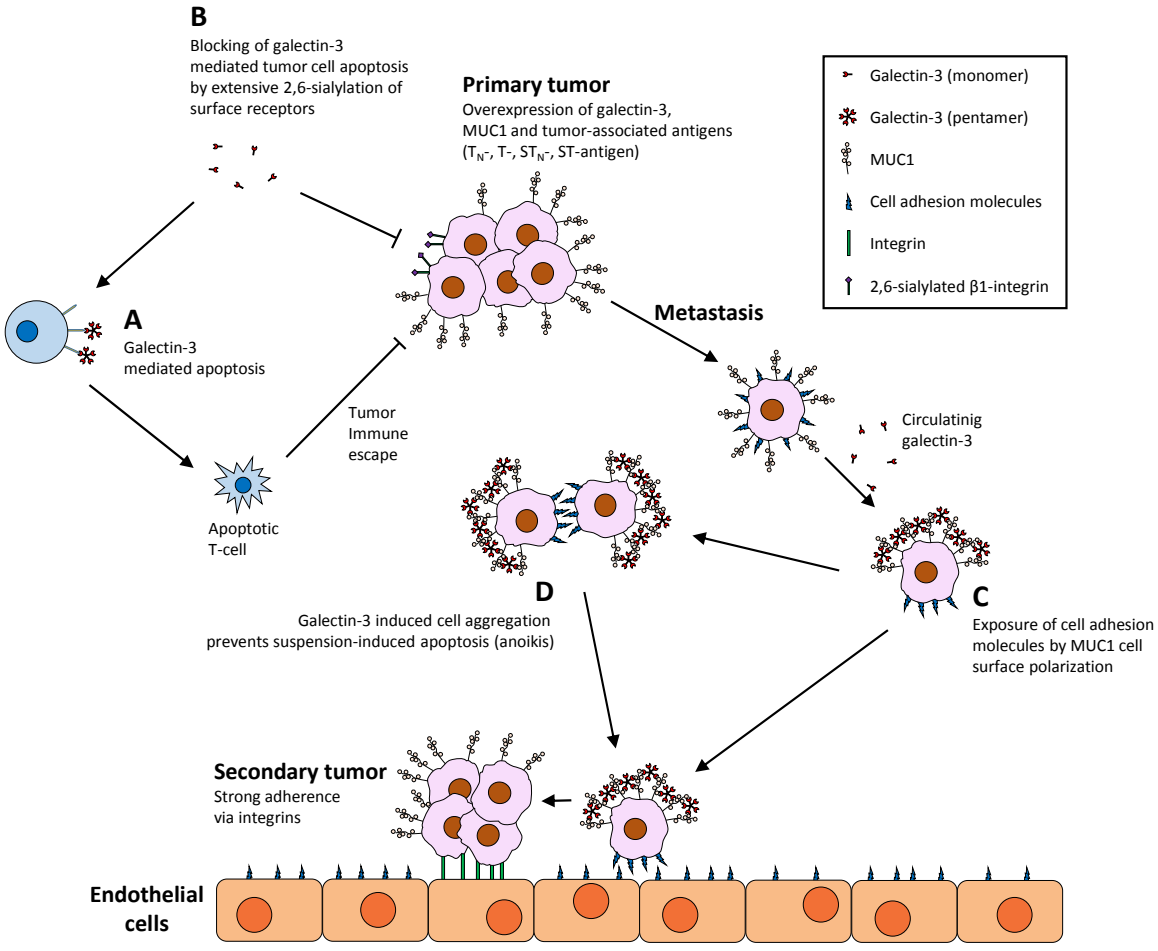


Figure 3: Proposed effects of galectin-3 on tumor immune escape and metastasis.

3.2.2.2 Mucin-directed anti-tumor vaccines

In addition to the galectin-3 mediated immune escape, tumor cells with aberrant MUC1 expression are protected by further mechanisms. The human body produces anti-MUC1 autoantibodies, which can be detected in the blood serum of cancer patients. However, the natural immune-response seems not to be strong enough to successfully defeat the entire tumor infestation[156-157]. The reason lies within the natural tolerance of the active immune system towards endogenous molecules, which in consequence also shows tolerance towards the structurally similar cancer-characteristic molecules[158]. To overcome this inherent low immunogenicity of mucin antigens, synthetic antitumor vaccines are promising constructs for active immunization of cancer patients. The concomitant generation of tumor-specific antibodies could be of particular benefit in regard of effective control of metastatic tissue. Here, MUC1 glycopeptides serve as valuable lead structures on the way towards new vaccines[159].

The goal in vaccine design is to induce a strong, specific and long-lived humoral immune response. Accordingly, the generation of long-lived IgG antibodies with high affinity is necessary, as the induction of short-lived IgM antibodies with low affinity does not fulfil these requirements. The generated antibody class depends on the cellular mechanism, which is induced by the presented vaccine antigen. If an exogenic antigen is bound by a B-cell receptor (BCR), the B-lymphocyte can proliferate into a plasma cell which starts to secrete antigen-directed IgM antibodies (***Figure 4***, **A**). This T-cell independent immune response is typically induced if the antigen consists of sole carbohydrates[160]. In order to provoke the aimed IgG antibody production, the T-cell dependent immune response has to be induced. This response is triggered, if an antigen presenting cell (macrophage, dendritic cell or B-cell) presents an unknown, exogenic antigen via major histocompatibility complex II (MHC II) to a T-cell with an appropriate T-cell receptor (TCR). As a result, the T-cell differentiates into a T-helper cell (T_h) and secretes cytokines. These cytokines are bound by cytokine receptors at a B-cell, inducing its differentiation into an IgG-antibody-producing plasma cell. Alternatively, B-cells and T-cells can differentiate to memory cells, enabling a long term immune recognition of the corresponding antigen structure (***Figure 4***, **B**).

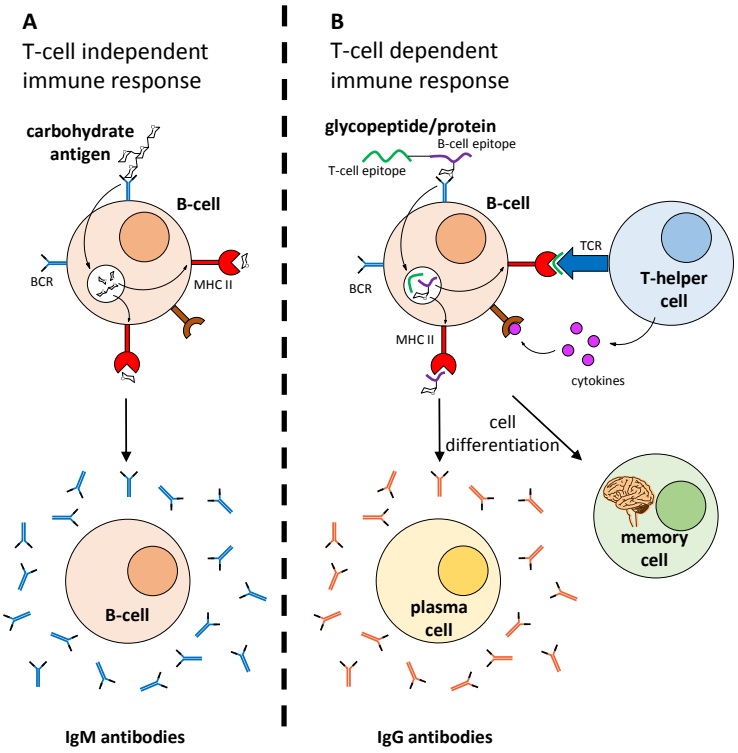


Figure 4: A: T-cell independent response induced by carbohydrates elicits production of IgM antibodies. **B**: T-cell dependent response induced by costimulation elicits IgG antibodies.

A common strategy in vaccine design is the coupling of B-cell epitopes with immunostimulants, such as the carrier proteins bovine serum albumin (BSA), keyhole limpet hemocyanin (KLH) or tetanus toxoid (TTox), which present T-cell epitopes in order to induce a T-cell dependent immune response. Thereby, the vaccines are endocytotically internalized by the antigen-presenting cell (dendritic cell, macrophage or B-cell), digested and the generated antigen fragments are presented on the cell surface via the MHC II complex. If a purely carbohydrate-based B-cell epitope is implemented, induction in mice models only lead to moderate immune responses mainly of the IgM-type, suggesting an insufficient T-cell dependent activation[161-162]. Strong immune responses have been induced by conjugation of MUC1 glycopeptides to the carrier proteins BSA[163], KLH[164-165] and TTox[166] (***Figure 5***, **A**). Here, multivalent B-cell epitope presentation on the carrier protein is possible, which increase the antigen density for possible uptake by antigen presenting cells (macrophages, dendritic cells, B-cells). An inherent risk of the carrier protein approach is a possible override of the B-cell epitope specific immune response by a stronger response to the carrier protein itself[167]. Alternatively, short T-cell epitope peptides can be conjugated to the B-cell epitope as for instance the tetanus toxoid-derived P30 peptide, which further do not require an extra adjuvants during immunization[168-169]. Further peptide-based examples are ovalbumin (OVA)[170] or the rationally designed Pan DR (PADRE) T-cell epitope[171] (***Figure 5***, **B**). Immunizations in mice with PADRE conjugates elicited strong IgG response[172].

A further alternative to generate a high immunogenicity for glycopeptide haptens is the conjugation of mitogens such as the Toll-like receptor2 ligand Pam_3CSK_4 (***Figure 5***, **C**). As a mitogen it induces the secretion of cytokines and chemokines[173]. Two-component conjugates of this lipopeptide with a MUC1 B-cell epitope, immunization in Balb/c-mice only elicited antibodies of the IgM type, but these showed tumor specificity[174]. The described components are conjugated via immunologically inert spacers, such as oligoethyleneglycol, resulting in two- or three-component vaccines (***Figure 5***, **D**).[153, 175]

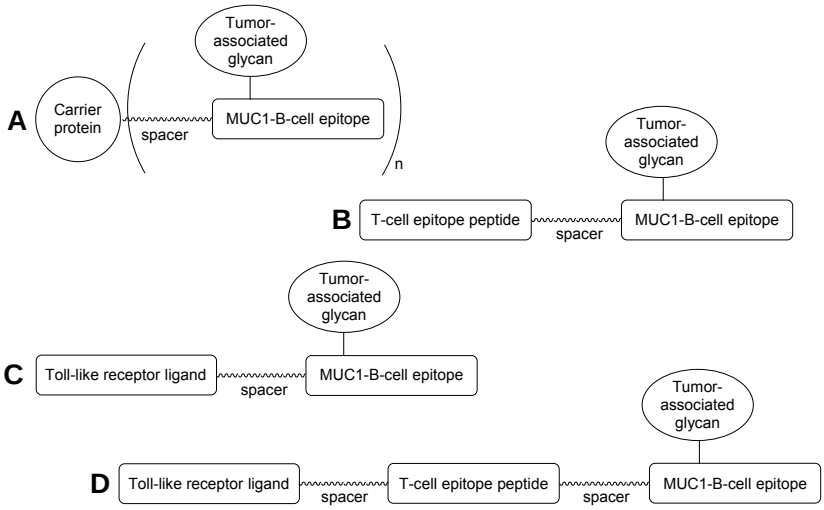


Figure 5: Design of synthetic vaccine conjugates: **A**: Two-component peptide-peptide vaccine. **B**: Two component peptide-protein vaccine. **C**: Two component peptide-TLR-lipopeptide vaccine. **D**: Three-component vaccine conjugate with B-cell epitope, T-cell epitope and TLR ligand.

Since tumor-associated MUC1 from biological samples is hetreogenious and may contain both tumor-associated truncated carbohydrate antigen structures as well as extended glycan structures, which are also presented on healthy cells, utilization of these structures in vaccines may result in an inefficient immune responses or may risk to build-up an immunological memory also directed to healthy epitopes and thereby induce auto-immune response[176]. Here, synthetic and structurally well-defined glycopeptide antigen conjugates represent a more reliable and safe alternative.

3.3 HexNAc *O*-glycosylation on tyrosine

O-glycosylation in mammalian cells is usually initiated with an *N*-acetyl hexosamine (HexNAc) residue linked to Thr and Ser amino acids. However, within the last decade several examples of HexNAc *O*-glycosylation on Tyr have been identified. The earliest discovery was made by *Jonas Nilsson* in 2011, by MS structural analysis of immunopurified amyloid-β isolated from human cerebrospinal fluid (CSF) samples via collision induced dissociation (CID) experiments[177]. Thereby Tyr10 of amyloid-β was

identified to be *O*-glycosylated with $(Neu5Ac)_{1-2}Hex(Neu5Ac)HexNAc$ structures. This site is a part of the neurotoxic amyloid precursor protein (APP) segment[178], which is proteolytically cleaved from the transmembrane APP mother protein to become a major component of the neuritic plaques[179]. These plaques are formed in the brain of *Alzheimer's Disease* (AD) patients[180]. It is suggested, that the discovered Tyr *O*-glycosylation modulate the APP proteolysis associated with the neurodegenerative disease[177].

By analysis of cell-lysates of colo-205 simple cells *Levery* et al in 2011 reported the identification of another *O*-HexNAc modification on Tyr, a tryptic peptide fragment from nucleobindin-2 (NucB2), LEYHQVIQQMEQK, after analysis by HPLC-MS ETD (**E**lectron **T**ransfer **D**issociation) fragmentation[181]. NucB2 is a Ca^{2+}-binding protein located in the golgi[182] as well as extracellular and membrane-bound[183] and is predominantly found in spleen, testicles and normal stomach. It represents the precursor for Nesfatin-1, a secreted nutrient-responsive anorexigenic peptide[184-185]. Furthermore, it has been found to be upregulated in breast carcinomas, having an amplifying effect on the cell-proliferation activity as well as the migration and invasion properties of tumor cells[186]. For this reason NucB2 is considered as a prognostic factor for breast cancer. With these identifications at hand, in 2013 *Levery* and *Clausen* were considering Tyr glycosylation in data analysis of further glycoproteomic work. In studies of lectin-enriched glycopeptides from colo-205 and capan-1 cell lysates, MS-ETD analysis resulted in identification of four new proteins modified with HexNAc-Tyr: Nucleobindin-1 (NucB1), CD44 antigen, extracellular matrix protein 1 (ECM1) and proline-rich acidic protein 1 (PRAP1)[187]. According to the utilized enrichment protocol *Levery* and co-workers assumed all identified Tyr glycosylations (including the one on NucB2) to be GalNAc, but acknowledge the possibility that some of them might also be GlcNAc due to potential cross-reactivity of the utilized *Vicia villosa agglutinin* (VVA).

In a large glycoproteomic study from 2013 on the murine synaptosome, *Medzihradszky* and co-workers identified that several mitochondrial glycopeptides were modified by HexNAc-Tyr. This included Tyr glycosylation of the proteins ATP synthase subunit beta (Atp5b), aspartate aminotransferase (mAspAT) and voltage-dependent anion-selective channel protein (Vdac1)[188]. Since all identifications were obtained by ETD mass-fragmentation experiments (only fragments the peptide backbone, not the glycans), differentiation between the structural epimers GalNAc and GlcNAc was not possible leaving the exact identification of the HexNAc unit for further investigation. In earlier collaborative work with Jonas Nilsson we showed that GalNAc and GlcNAc epimers linked on Thr or Ser was possible to be discriminated by HCD MS fragmentation[189-190].

Also in 2013 *Klaus Aktories* reported Tyr *O*-glycosylation on eukaryotic Rho GTPases in the course of bacterial infection by *Photorhabdus asymbiotica*[191]. By tandem-MS analysis *N*-acetyl hexosamine was

found to modify amino acid residues on Tyr34 of RhoA and Tyr32 of the RhoA equivalents Rac1 and Cdc42. NMR structural analysis confirmed that this modification was a αGlcNAc. The modification is introduced by the bacterial toxin PaTox on small GTPases in the GTP-bound state and is located in the highly conserved switch I region. This region rearranges between the inactive GDP- and the active GTP-bound form of the GTPases, which are controlled by the GTPase cycle (***Figure 6***). Glycosylation by PaTox inhibits the interaction of the small GTPases with their effectors, leading to blocking of downstream effector signaling. This results in disassembly of the actin cytoskeleton, inhibition of phagocytosis and finally to cell-death. Similarly, the toxin antifeeding prophage 18 (Afp18) from *Yersinia ruckeri* was reported to modify Tyr34 of RhoA with αGlcNAc, but not Rac1 or Cdc42 in zebrafish.[192] Both, PaTox and Afp18 show a preference for the active GTP-bound state of the GTPases. X-ray analysis revealed, that glycosylation of Tyr34 results in an opened protein conformation of the Switch-I region, which is different from the GDP- and GTP-bound conformations. The observed open and deactivated conformation is not compatible with effector or regulator interactions, which explains the impaired signaling and the toxicity that this modification causes.[191-194]

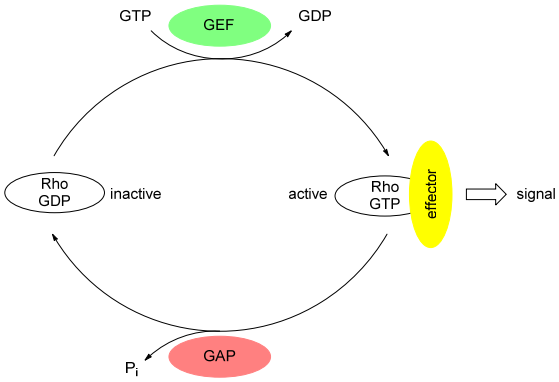


Figure 6: The GTPase activity is controlled by the GTPase cycle. Activation by exchange of GDP for GTP is promoted by G-nucleotide exchange factors (GEF). Deactivation by GTP hydrolysis is mediated by GTPase-activating proteins (GAP).

The large number of HexNAc-modifications discovered in recent years triggered our interest to generate tools that detect and selectively distinguish between the different structural epimers and between modifications on Thr/Ser vs Tyr. It also triggered a curiosity to explore the function of HexNAc-Tyr glycosylation.

4 Motivation

4.1 Project 1: Mucin glycosylation in cancer and airway diseases

Mucins and the mucin carbohydrate ligands that are densely presented on peptide tandem repeats and have been found to play important roles in the initiation and progression of bacteria and virus infections. A high frequency of infections is particularly a problem for patients suffering from chronic airway diseases such as COPD and Cystic fibrosis. A major aim of this work was to develop an efficient synthesis strategy to build-up a library of complex mucin *O*-glycopeptides. This would enable studies of the role of mucins and mucin carbohydrate ligands in virus and bacteria adhesion related to airway diseases. Carbohydrate binding proteins, lectins, are typically presented on the flagella or pili of bacteria or on the fiber knob of virus, which are involved in cell-surface adhesion via the glycans. Learning more about these carbohydrate-lectin interactions would enable future design of lectin inhibitors to block bacteria or virus cell-surface adhesion. Further, recognition of mucin glycan binding patterns of specific bacteria or virus strains are of interest for diagnostic applications.

In addition to play a role in infections, the mucin *O*-glycosylation and mucin expression is dramatically changed in epithelial cancers. The altered expression and glycosylation result in presentation of unique mucin glycopeptide epitopes on tumor cells, which makes them potential targets for immunotherapy. Within our research group extensive work have been made to develop vaccines to target altered structures found on the mucin MUC1. The synthetic glycopeptide library prepared in this work aimed to support elucidation of binding epitopes of antibodies generated from these vaccines. We were also interested to study mucin glycopeptide interactions with human lectins involved in tumor-progression such as macrophage galactose c-type lectin (MGL)[195] and galectin-3.

In recent years the development of the microarray technology has enabled efficient studies of carbohydrate binding interactions and with low consumption of valuable synthetic material. Glycopeptide microarrays was therefore the major method of choice in order to efficiently map protein interactions with the synthetic mucin glycopeptide library.

Due to the heterogenic nature of mucin glycoproteins, isolation of proteins with well-defined glycosylation from biological samples is problematic. Synthesis of mucin peptide tandem repeats with specific glycosylation pattern, site specificity and density was therefore preferable to address this

limitation. A synthetic strategy for efficient preparation of mucin glycopeptides and complex mucin-type glycosylated amino acids was aimed to be established within the frame of this work. The glycosylated amino acids served as building blocks for Fmoc-solid-phase peptide synthesis to generate a chemically well-defined glycopeptide library. Further enzymatic diversification or direct applications in microarray analysis was then possible. In an additional application, the well-defined glycopeptide structures were also interesting model compounds to define linkage connections and to discriminate between structural epimers in MS-structural analysis. This work is incredibly valuable for improving glycoproteomic analysis of intact glycopeptides.

The synthesis strategy was envisioned to constitute of a few common building blocks to be combined for build up into complex mucin-type glycan core and elongated core target structures. The synthesis includes well-matched glycosyl donors and acceptors, which are assembled stepwise into different individual connectivities. In the frame of this work, core 2 type-1 and -2 as well as core 4 type-1 and -2 target structures were aimed to be synthesized. These structures are build-up from the common T_N-antigen, the T-antigen and the core 3 (type-1 and type-2) acceptors, which can be elongated by introduction of LacNAc type-1 and type-2 disaccharide donors. The lactosamine disaccharides represent the alternating addition of *N*-acetylglucosamine and galactose units that build up and extend natural mucin-type *O*-glycan core structures. The glycosylated amino acid building block and glycopeptide generation in this work complements the work of *Dr. Christian Pett*, which covered the synthesis of elongated core 1, 2 and 3 glycosylated amino acids. The resulting glycosylated amino acid building blocks will be incorporated into synthetic peptides to generate a glycopeptide library with defined structures in high quantity. From each synthesized glycosylated Fmoc- amino acid a set of mono-, di- and trivalent glycosylated peptides will be generated, covering all possible combinations that include Thr *O*-glycosylation on MUC1. As peptide backbone, the VNTR region of MUC1 was chosen, since this region represents the peptide tandem repeats that are densely O-glycosylated and involved in diverse protein binding events. Furthermore, they are overexpressed and aberrantly *O*-glycosylated on tumor tissue. After successful glycopeptide synthesis, the structural diversity of the library can be further increased by chemo-enzymatic terminal glycosylation e.g. sialylation, fucosylation or poly-LacNAc. Finally depending on experimental requirements, selected glycopeptide constructs of the resulting library are printed on amine reactive *N*-hydroxy succinimide coated microarray slides. Here, the utilized peptide N-terminal triethylene glycol spacer fits the recommendation from earlier microarray studies that postulated an optimal spacer length of at least six carbons for ligand presentation.

The completed glycopeptide library was aimed to be used in microarray studies of glycopeptide-protein binding experiments to identify possible ligands and investigate the influence of structure, size

and distance of the presented glycans on the carbohydrate-protein-binding intensity. Binding preferences of two lectins from *P. aeruginosa,* FliD and PcrV will be evaluated and incubated on the mucin glycopeptide microarray. The array will include glycopeptides containing different sialylated and/or fucosylated core glycans to reflect the natural glycan diversity in respiratory mucins. For instance blood group antigens such as Sia-Le^x, Le^x, Sia-Le^a and Le^a have been found to be important for *P. aeruginosa* recognition and additionally mucins seem to be involved in these binding events. A further candidate for mucin glycopeptide binding studies are the fiber knob proteins from adenoviruses. Different strains of these viruses commonly cause respiratory infections, but are also involved in infections of the eyes as well as intestinal infections. In initial studies we were interested to explore glycopeptide interactions with AD37, from adenovirus type 37, a virus responsible for epidemic keratoconjunctivitis in eyes. Earlier work on the CFG (consortium for functional glycomics) glycan arrays revealed that the brain ganglioside GD1a is a particularly strong AD37 ligand[196-198]. However, additional mucin glycan binding motifs can be assumed, since GD1a is not highly abundant in the eyes, but highly abundant on mucin-rich targets that may mimic the GD1a structure. Besides focusing on specific lectin interactions, also general binding preferences of intact pathogens will be investigated by subjecting fluorescence-labelled bacteria and viruses to the array. Candidate examples are *P. aeruginosa*, *H. influenza* and *S. pneumoniae*, which are major players in chronic airway disease infections. This will enable to cover interactions while presenting the lectins in the natural protein complexes of the bacteria and the identification of the most important binding epitopes that are involved in the adhesion process. Another class of bacterial virulence factors are secreted enzymes such as sialidases. In future microarray studies the sialylated mucin glycopeptides are aimed to be elucidated as possible sialidase substrates.

As mentioned above the synthetic MUC1 glycopeptide library will also be used in microarray studies to support elucidation of binding epitopes of antibodies generated from anti-tumor vaccines as well as to study cancer relevant lectins such a as Galectin-3 and MGL.

4.2 Project 2: Chemical tools to explore HexNAc-Tyr *O*-glycosylation

In recent years a new group of protein modifications have been found: HexNAc Tyr *O*-glycosylation. The first example from 2011 described glycosylation with $(Neu5Ac)_{1-2}Hex(Neu5Ac)HexNAc$-Tyr structures on amyloid-β peptides which were isolated from CSF of Alzheimer's Disease patients. Based on the carbohydrate structure this PTM was considered to be mucin type Sia-Core 1 Tyr *O*-glycosylation since analog extension on Thr or Ser residues usually only take place on GalNAc linked peptides but not on GlcNAc. In a later glycoproteomic study on colo-205 SimpleCells, additional HexNAc-Tyr glycopeptides were identified from Nucleobindin-1 (NucB1), Nucleobindin-2 (NucB2), CD44 antigen, extracellular matrix protein 1 (Ecm1) and proline-rich acidic protein 1 (Prap1). The authors assigned these monosaccharides as GalNAc according to the GalNAc-preference of the VVA lectin used for enrichment. In another large glycoproteomic study on the murine synaptosome three mitochondrial proteins were found modified with HexNAc-Tyr. Here, differentiation and precise identification of the different possible HexNAc isomers was not possible even if localization in mitochondria, nearby phosphorylation and *O*-GlcNAcylation of Ser/Thr sites within the identified proteins would suggest that the found HexNAc-Tyr represents βGlcNAc. Further, a toxin from *Photorhabdus asymbiotica* has been found to modify host Rho GTPases with αGlcNAc-Tyr glycosylation. In this case enzymatic toxin modification of a recombinant protein and analysis with protein NMR was used to define the GlcNAc stereochemistry.

Taken together several examples of HexNAc-Tyr *O*-glycosylation were identified, but assignment of HexNAc as αGalNAc or as α- or βGlcNAc was in most cases not possible and βGlcNAc-Tyr is not completely proven to be a new PTM. The β-*O*-GlcNAc modification (β-*O*-GlcNAcylation) is further hard to detect because of its relatively high turnover and lower stability compared with α-*O*-GalNAc. Lectins used for weak affinity enrichment of β-*O*-GlcNAc peptides, mainly wheat germ agglutinin (WGA), are also not selective for βGlcNAc and partly cross-react with αGalNAc peptides. On Ser/Thr residues the biological functions of mucin type (GalNAc) *O*-glycosylation is dramatically different from β-*O*-GlcNAcylation and tools that can detect and discriminate these modifications on Tyr, would be desirable to understand the biology behind these new modifications.

The aim of this work was to develop chemistry based tools to enable analytical and functional studies of HexNAc-Tyr *O*-glycosylation. Further, existing molecular tools were aimed to be evaluated regarding their applicability to detect these new modifications. Enzyme substrate specificity studies using synthetic peptides aimed to be done to understand if GlcNAc-Tyr (or unglycosylated Tyr) are substrates for OGA and OGT enzymes and thereby to obtain insights into the biosynthetic pathways of HexNAc-Tyr *O*-glycosylation. To enable all these studies an extensive library of diverse HexNAc-Tyr

glycopeptides was prepared. A graphical overview of the complete project aims/workflow is given in ***Figure 7***.

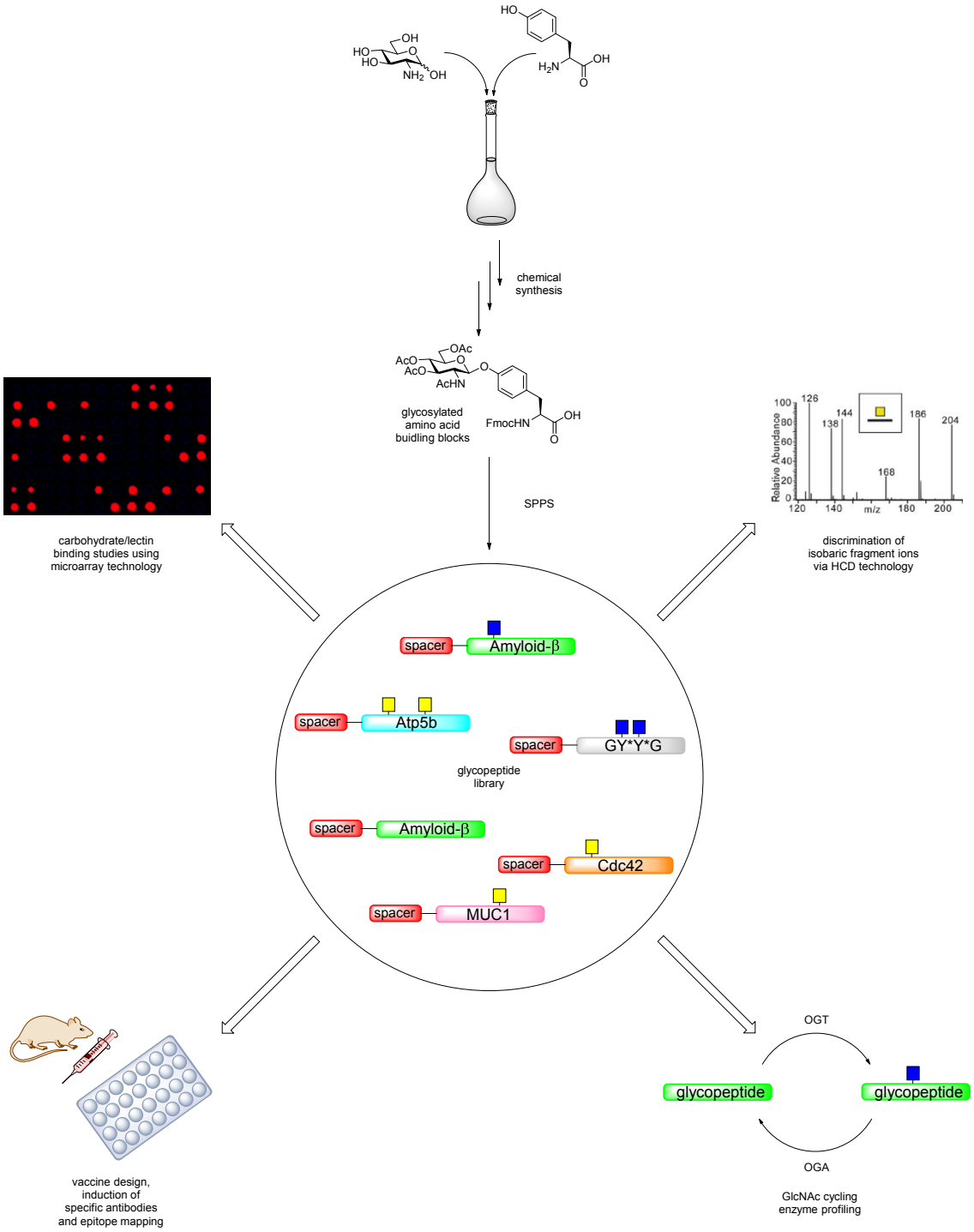


Figure 7: Schematic illustration of the project setup. A library of well-defined glycopeptide constructs including different glycan isomeres serves as basis for the development and evaluation of chemical biological tools for glycopeptide enrichment and identification.

The synthetic peptide library that was aimed to be prepared consisted of biologically relevant tryptic glycopeptides, which in glycoproteomic studies have been identified to contain HexNAc-Tyr *O*-glycosylation. Tryptic glycopeptides or analog glycopeptides exhibiting potential Ser, Thr and/or Tyr glycosylation sites in close vicinity to previously identified HexNAc-Tyr peptides were also interesting to prepare. Further, model peptides with natural Ser and Thr mucin type O-glycosylation such as the MUC1 VNTR were considered relevant to prepare as HexNAc Tyr analogs. All peptide sequences were aimed to be synthesized as mono- and divalent glycosylated structures in different glycan isoforms consisting of αGalNAc, βGlcNAc or αGlcNAc on Ser/Thr/Tyr residues.

To prove that the βGlcNAc-Tyr modification absolutely exists, a high-throughput methodology is needed that is able to discriminate between both diastereomers: αGalNAc and βGlcNAc. In a recent collaborative work with Dr. Jonas Nilsson (Gothenburg University), it was found that GalNAc- and GlcNAc-Ser/Thr structural epimers could be discriminated by MS-structural analysis based on high energy collision-induced dissociation (HCD) fragmentation. Discrimination of glycan isomers usually constitutes a huge challenge for mass spectrometry analysis, however characterization of low-mass oxonium ion fragment profiles made this possible. In the frame of this work, it was interesting to find out if this GalNAc- and GlcNAc discrimination also applies for HexNAc-Tyr glycopeptides. In that case this strategy may become a powerful glycoproteomic tool for detection and identification of such modifications. Synthetic HexNAc-Tyr model peptides were therefore aimed to be applied in HCD experiments with fragmentation at different collision energies.

In previous glycoproteomic studies, which identified examples of HexNAc-Tyr glycosylation, lectin affinity chromatography on tryptic digests was applied for standard glycopeptide enrichment prior to mass spectrometric analysis. Based on the utilized lectins the authors conclude that the identified HexNAc monomers are of mucin-type (αGalNAc). While this assumption might be verified by HCD oxonium ion analysis (if applicable on HexNAc-Tyr *O*-glycosylation), to date no information about the influence of Tyr towards the lectin-carbohydrate interaction is available. Additionally, the lectin WGA, which often is used for enrichment of O-GlcNAcylated Ser/Thr peptides, seems to have cross-reactivities to other glycans including GalNAc. Lack of specificity has also been reported for other commonly used lectins. However, informations about the lectin binding affinities are essential for comprehensive glycoproteomic studies. To elucidate the impact of Tyr glycosylation in the selectivity of lectin recognition, synthetic glycopeptides were aimed to be applied in microarray binding studies by incubation with plant-lectins which are commonly used for glycan detection or glycopeptide affinity enrichment. Based on the results, the lectins were aimed to be evaluated towards GalNAc- and GlcNAc-recognition. Further, differences in recognition between Tyr and Ser/Thr glycopeptides were interesting to be compared. Finally, the incorporation of mono- and diglycosylated glycopeptides to

the array will enable monitoring of effects from bidentate or multivalent lectin binding interactions. The impact of peptide sequence and nearby environment of the different glycosylation sites are also interesting to study.

As tools for specific detection (on western-blots or tissues) or enrichment of HexNAc-Tyr glycosylation it was also considered interesting to generate specific antibodies directed against GalNAc- or GlcNAc-Tyr modifications. Thereby vaccine constructs needed to be prepared and used in immunizations. A subsequent epitope mapping of the enriched serum antibodies via glycopeptide microarray screening would be needed to elucidate if desired antibody specificity had been obtained. Of particular interest would be the possible discrimination between HexNAc-Tyr vs HexNAc-Ser/Thr glycopeptides.

To gain insights into the possible existence of βGlcNAc-Tyr modifications it was interesting to elucidate if the enzymes which are known to modify Ser/Thr peptides also showed possible activity and substrate specificity towards Tyr peptides. Enzyme profiling with *O*-GlcNAc transferase (OGT) and *O*-GlcNAcase was therefore aimed to be made by incubations with synthesized peptides or glycopeptides respectively and simultaneous monitoring of the conversion via HPLC-MS.

5 Synthetic strategies

5.1 Carbohydrate synthesis

One of the key step in carbohydrate synthesis is the formation of glycosidic bonds. Carbohydrate chemistry generally comprises the generation of *O*-, *N*-, *S*- and *C*-glycosides, here only the synthesis of *O*-glycosides will be discussed, which is applied in this work. Generally, an *O*-glycoside can be considered as cyclic acetal with a substituent connected to the anomeric center via glycosidic bond. If the substituent is a non-carbohydrate it is called an aglycone while the carbohydrate part is called a glycone. The *O*-glycoside as shown in ***Figure 8*** can be obtained by coupling of an electrophilic glycosyl donor with a nucleophilic acceptor.

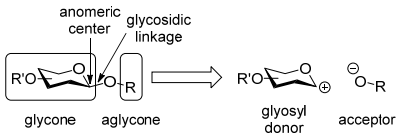


Figure 8: Retrosynthetic cut of a glycosidic linkage.

The most simple methodology to form an *O*-glycosidic bond is the *Fischer glycosylation*[199] (***Figure 9***). Here the hemiacetal of a free sugar is protonated to initiate the generation of a resonance-stabilized oxocarbenium intermediate. The sp^2-hybridized half-chair intermediate cation allows nucleophilic addition of an alcohol from each side of the planes resulting in formation of two stereoisomeric glycosides, an anomeric mixture. Alternatively, the reversed reaction with water leads to recovery of the starting material. To form the desired coupling product the equilibrium has to be shifted towards the product by employment of excessive amounts of acceptor alcohol.[200-201]

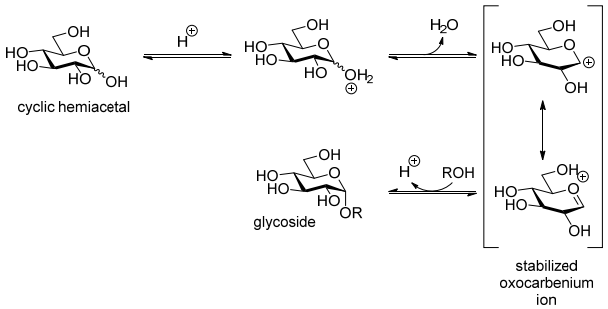


Figure 9: Principle of *Fischer glycosylation* exemplified with glucose.

The Fischer glycosylation works well with simple acceptors, for example the reaction with methanol, but the methodology is not applied in synthesis of more complex glycosides using more sophisticated

carbohydrate- or natural product-acceptors. The corresponding acceptor molecules are valuable and not available in sufficient quantities as needed for the use in large excess. Instead a number of other methodologies have been developed, which contain glycan donors with a better anomeric leaving group that can be activated by a suitable promotor. Because of the diversity of carbohydrates and the resulting amount of possible glycosides to be made, the number of glycosyl donors used in carbohydrate synthesis is manifold and continuously increasing. A selection of common glycosyl donors used in carbohydrate synthesis is shown in ***Figure 10***.

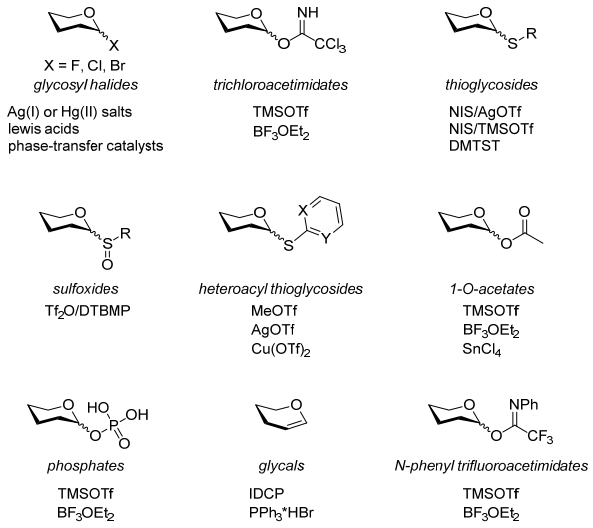


Figure 10: Selection of glycosyl donors and corresponding promotors as used in *O*-glycoside synthesis.[202]

One of the most important and frequently used glycosyl donors are glycosyl halides, which are commonly used in the course of *Koenigs-Knorr* reactions[203]. Trichloroacetimidates are used in the *Schmidt*-method[204-205]. Further, thioglycosides are suitable donors, which are particular stable until activation by a promotor (NIS/TfOH or DMTST) and are therefore also considered as a temporary protection group. All these methodologies follow the same fundamental principle of glycoside generation as described above by the *Fischer* glycosylation, the mechanisms of the leaving group detachment are dependent on the used leaving group/promotor system.

In addition to the donor activation there are two more major challenges to address in glycosylation reactions: the regiochemistry and the stereochemistry of the newly formed glycosidic bond. The reactivity difference of the various acceptor hydroxyl groups can be used to control the regiochemistry to some extent, but a suitable protection group strategy to block competing nucleophiles with exception of the targeted acceptor hydroxyl group are usually required. Furthermore, protection groups offer the opportunity to modulate the nucleophilicity of acceptors as well as the activation reactivity of glycosyl donors. The configuration of the generated glycosidic linkage can be controlled

by the reaction conditions (temperature, solvent, promotor) as well as the constitution of the glycosyl donor and its protection group pattern.

5.2 The anomeric effect

The stereochemical outcome of glycosylation reactions of D-pyranoses or D-pyranosides is influenced by the *anomeric effect* as first described by *R. Lemieux* in 1958[206-207]. In general equatorial substituents are energetically favored in carbohydrates and other molecules with a chair conformation, due to the strong steric repulsion of axial substituents among each other (1,3-diaxial interactions). The anomeric substituents of carbohydrates do not always follow this rule. Instead, the normally sterically unfavored α-isomer is pronounced in aqueous solutions of D-pyranoses. This effect increases with rising electronegativity of the anomeric substituent and decreases with rising dielectric constant of the solvent. Two main explanations are given for the *anomeric effect*. One focusses on dipole-dipole interactions from the lone electron pairs of the endocyclic oxygen atom and the anomeric substituent. Depending on the orientation of the anomeric substituent, the dipoles are orientated parallel or antiparallel to each other leading either to unfavored addition (equatorial orientation) or favored neutralization (axial orientation) of the dipoles (***Figure 11a***). The second and most accepted explanation is based on favored orbital-orbital interactions for the axial anomer. In this arrangement, the axial non-binding lone-pair orbital of the endocyclic oxygen overlaps with the antibonding σ*-orbital of the C-X bond (***Figure 11b***). In contrast, the resulting $n_O \rightarrow \sigma^*$ delocalization of the non-bonding electrons is not possible if the anomeric substituent is in equatorial orientation. In this case the lone-pair electrons of O5 are delocalized into the antibonding C-H orbital instead, which results in a higher energy level and thereby the equatorial orientation is unfavored. In the anomeric effect, the α-anomer is thermodynamically favored while the β-anomer is kinetically favored.[18, 200]

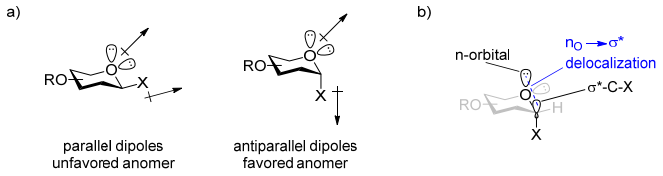


Figure 11: Explanations for the *anomeric effect* a) configuration dependent dipole-dipole interactions, b) delocalization of non-bonding electrons if anomeric substituent is axial.

In contrast to the above discussed *anomeric effect*, a β-directing *reverse anomeric effect* can be observed if the anomeric substituent is electropositive compared to the anomeric carbon[208]. In the *reverse anomeric effect* the dipole momentum of the C-X bond is reversed, which leads to dipole-

dipole neutralization when the substituent is in equatorial position. This effect is enhanced by the general rule that equatorial substituents are energetically favored over the axial.

5.3 Protecting groups in carbohydrate chemistry

The large number of functional groups, e.g. hydroxyl groups with similar reactivity of different mono- and oligosaccharides is a challenge for performing regioselective glycosylations. Some differences in reactivity are observed for the different hydroxyls in distinct sugars, but only a suitable protection group strategy can achieve complete regioselectivity in coupling reactions and meet the requirements of a complex carbohydrate synthesis. In general, the anomeric hydroxyl represents the most reactive group in a pyranose, followed by primary, equatorial and axial hydroxyls in that order[209]. Further, equatorial hydroxyl groups with a nearby axial hydroxyl are more reactive than a equatorial hydroxyls without axial neighbor. This trend can be used for regioselective introduction of protection groups. In principle all reactive groups (except the one involved in the reaction) have to be protected, sometimes two or three hydroxyls of a glycan acceptor remain unprotected. This can be done if their reactivity difference is large enough to maintain a regioselective glycosylation. In addition to regioselective considerations, protection groups on the hydrophilic carbohydrates also increase the solubility in organic solvents.

In general, carbohydrate synthesis involves more permanent protecting groups that are removed in the end of the synthesis route and temporary protection groups blocking specific reactive groups that for instance are removed just before or after a coupling reaction. Typical permanent protection groups are benzyl or allyl ether groups, which can be cleaved under hydrogenolytic conditions and ester groups (acetyl, benzoyl groups) that often are removed with base[210]. Ideally, all permanent protection groups in a target glycoside are cleavable (and stable) under the same conditions, which makes it possible to remove all groups in a single reaction step at the end of the synthesis. The temporary protection groups have to be selected to be orthogonal to the permanent groups and to be selectively cleavable under conditions, in which the permanent groups are stable. Temporary protection groups can be ethers (such as methoxybenzyl, benzyl, trityl, silyl and allyl ethers), esters (like acetyl, chloroacetyl, benzoyl groups) or Acetals which simultaneously protect vicinal hydroxyl groups. Amines are typically protected as carbamates (e.g. Troc or Fmoc groups), as phthalimido groups or masked as azides during synthesis. By proper arrangement of permanent and temporary protection groups, defined linear and branched glycans can be synthesized (***Figure 12***).

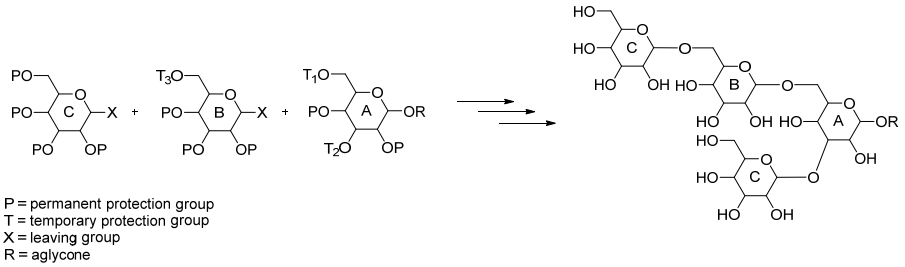


Figure 12: Oligosaccharide synthesis under utilization of a strategy including permanent and temporary protection groups.[18]

Besides the benefit in regard to regioselectivity protection groups have an impact on reactivity of donor molecules as well as a stereochemical impact. Alkyl groups generally enhance or at least do not reduce the reactivity of glycan donor or acceptor molecules, in contrast ester protection groups reduce reactivity. The different reactivities by diverse protecting groups are used in the *armed/disarmed concept*[211-212]. Here, armed refers to activated molecules, while disarmed refers to deactivated molecules. Ester groups, contribute with electron withdrawing effects that destabilizes the formation of an oxocarbenium ion of glycosyl donors, which is resulting in a lower activity (***Figure 13***). This effect is particularly strong if the electron withdrawing group is located at C-2 of the donor.

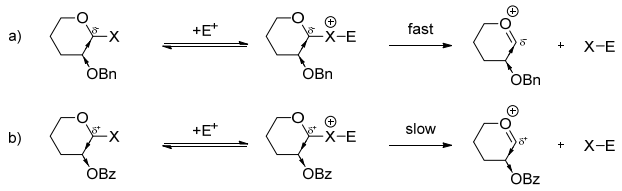


Figure 13: The electronic effects of substituents on oxocarbenium generation. a) The arming effect of alkyl ethers by induction of a negative polarization on C-1 leads to fast dissociation of the anomeric leaving group. b) The disarming effect of esters by induction of a positive polarization on C-1 leads to slow dissociation of the anomeric leaving group.[212]

In addition the substituent configuration of a carbohydrate ring has an impact on the reactivity of donor molecules. In general, the reactivity is increased by axial substituents over equatorial substituents. Furthermore fused ring systems reduce the reactivity by disturbing the planar sp^2-geometry of oxocarbenium ion intermediates.[213-214]

In addition, many protection groups have a stereochemical influence on glycosylations by neighboring group participation (also known as *anchimeric assistance*). Normally the stereochemical outcome of a glycosylation is dominated by the anomeric effect (***Figure 14a***), but the anchimeric assistance can overwrite this influence. If an ester, amide or carbamate is located at C-2, the lone pair electrons of its carbonyl oxygen stabilize a formed oxocarbenium ion after release of the anomeric leaving group. The formed dioxolenium ion intermediate (or oxazolinium ion if neighboring group is amide or carbamate) prevents a nucleophilic attack from the α-face and is thus promoting a 1,2-*trans* selective glycosylation reaction. Besides the stereochemical influence, neighboring group participation may

lead to formation of orthoester by-products, which further are intermediates in these reactions. (***Figure 14b***).[18]

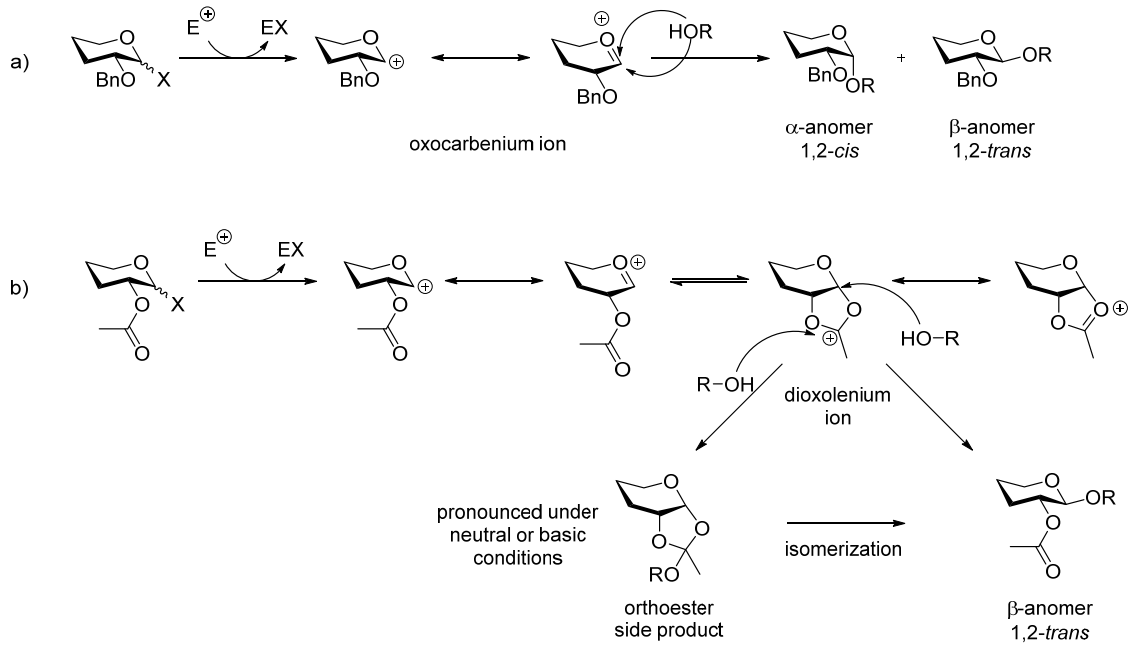


Figure 14: Influence of protection group at C-2. a) A nonparticipating group like ethers or azides has no effect on the stereochemical outcome of a glycosylation reaction and anomeric effect have major influence on stereoselectivity. b) A participation group leads to formation of the 1,2-trans product with potential formation of an orthoester by-product.

5.4 Solvent effects in carbohydrate synthesis

The solvents used in glycosylations are often of mild polar nature in order to stabilize the cation intermediate during the course of reaction[201]. The most frequently used glycosylation solvents are dichloromethane, diethyl ether, acetonitrile, toluene, 1,2-dichloroethane and nitromethane. The choice of the solvent influences – in addition to the *anomeric effect* (chapter 5.2) - the resulting anomeric configuration in a glycosylation reaction. Reactions with glucopyranosides in ether solvents like 1,4-dioxane, tetrahydrofuran or diethyl ether promote formation of 1,2-*cis* glycosides, reactions in acetonitrile support formation of 1,2-*trans* products[215]. The solvents can direct the outcome of glycosylations to the desired anomer if there is no participating neighboring group at C-2 (see chapter 5.3), which otherwise would dominate the stereoselectivity to 1,2-*trans* products. The most prominent example is the use of acetonitrile to achieve a 1,2-*trans* glycosylation (*nitrile effect*) or diethyl ether to achieve 1,2-cis glycosylation (*ether effect*). According to the *solvent coordination hypothesis*[216], these effects are based on coordination preferences of the solvent molecules on the anomeric carbon of the oxocarbenium ion intermediate. When acetonitrile is used, the mechanism proceeds via an α-nitrilium ion intermediate, which is generated in an S_N1 fashion via the

oxocarbenium ion[217]. The glycosidic linkage is then formed by nucleophilic substitution (S_N2) from the only possible β-face of the nitrilium ion by an alcohol leading to the β-glycosidic linkage (***Figure 15***). The covalent α-nitrilium ion was first isolated by Pougny and Sinay[218] and later confirmed by Ratcliffe and Fraser Reid[219]. It represents the kinetic anomer because of the reverse anomeric effect of the positively charged nitrilium residue (chapter 5.2). Ether solvents on the other hand coordinate preferably from the β-face, leading to a nucleophilic attack from the α-face under formation of the 1,2-*cis* glycosidic linkage. The ether effect is supported by the α-directing anomeric effect.

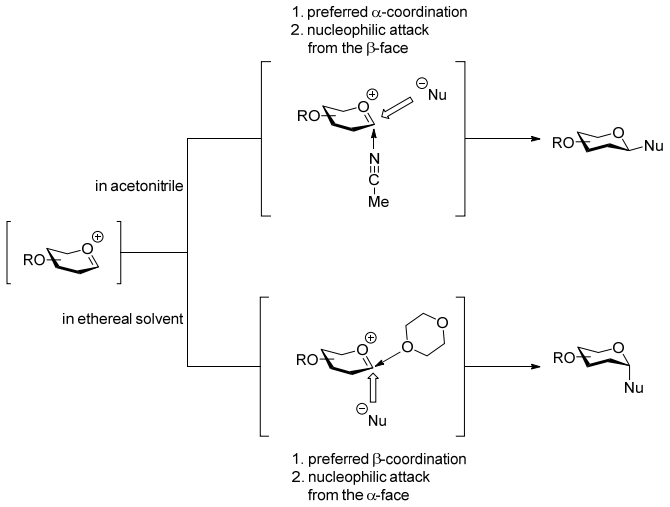


Figure 15: solvent coordination hypothesis.

An alternative explanation for the stereochemical effects of solvents is given by the *conformer and counterion distribution hypothesis* from *Satoh* and *Hünenberger*[220], which performed quantum mechanical calculations of the oxocarbenium-solvent interactions in vacuum and solvent together with molecular dynamics simulations. According to their results, the oxocarbenium ion adopts different conformations depending on the surrounding solvent. If the cation is surrounded by acetonitrile it adopts a $B_{2,5}$ boat conformation with the counterion (leaving group) residing in α-position while in toluene or ether solvents a $^4H_3$ half-chair conformation is suggested with the counterion in close proximity to the β-side. The nucleophile is then guided by the conformation of the cation and the relative position of the counterion to the cation.

5.5 Glycopeptide synthesis

Glycosylation of proteins results in an enormous structural diversity and the isolation of structurally well-defined glycopeptides from biological samples is a limited option. Instead chemical synthesis gives the possibility to prepare well-defined glycopeptides in desirable quantities. Four strategies are

applied in glycopeptide synthesis: a) coupling of the carbohydrate to the full-length peptide, b) stepwise glycopeptide assembly on solid-phase utilizing pre-synthesized fully glycosylated and normal Fmoc-amino acid building blocks and c) chemoenzymatic elongation of simple glycopeptides (***Figure 16***). Synthesis in solution enables the preparation of small or medium-sized glycopeptides.[221]

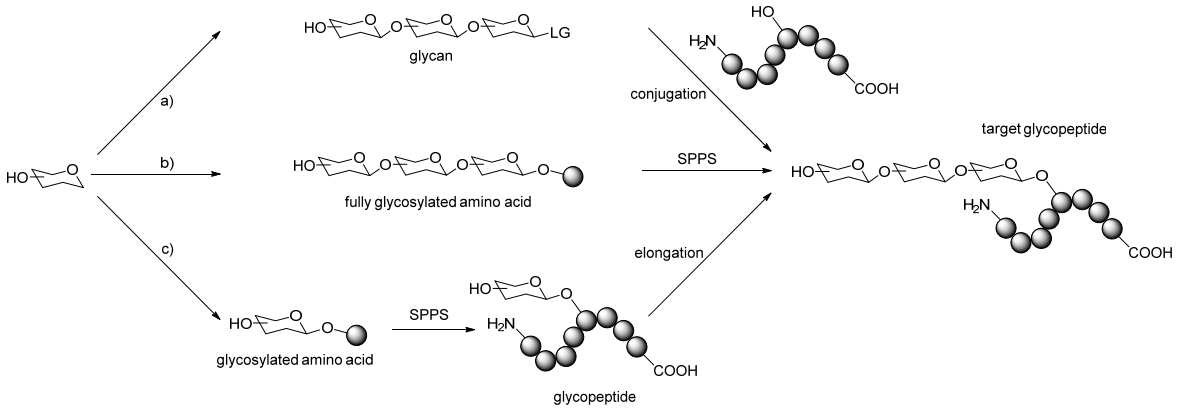


Figure 16: *O*-Glycopeptide synthesis strategies a) conjugation of the full glycan to the peptide, b) step by step synthesis on Fmoc-solid phase utilizing pre-synthesized fully glycosylated amino acids, c) on-resin glycosylation of resin-bound glycopeptides.

In this work the glycopeptides were synthesized by incorporation of pre-glycosylated amino acid building blocks via Fmoc-solid-phase peptide synthesis (Fmoc-SPPS, ***Figure 16***, route b)). In general, Fmoc-SPPS commences – in contrast to protein biosynthesis – from the C- to the N-terminus, typically starting with the C-terminal amino acid being preloaded via a cleavable linker to the solid-support (***Figure 17***). The peptide is then assembled in a stepwise fashion, coupling one amino acid per cycle to the N-terminus of the growing peptide. To prevent polymerization of the single amino acids instead of coupling to the growing peptide on the solid-support, the α-amino group of each utilized amino acid has to be temporarily protected during the coupling step, while the amino acid side-chains need to be protected during the complete peptide assembly. After complete sequence assembly, the peptide is released from the solid-support with simultaneous removal of amino acid side-chain protecting groups.

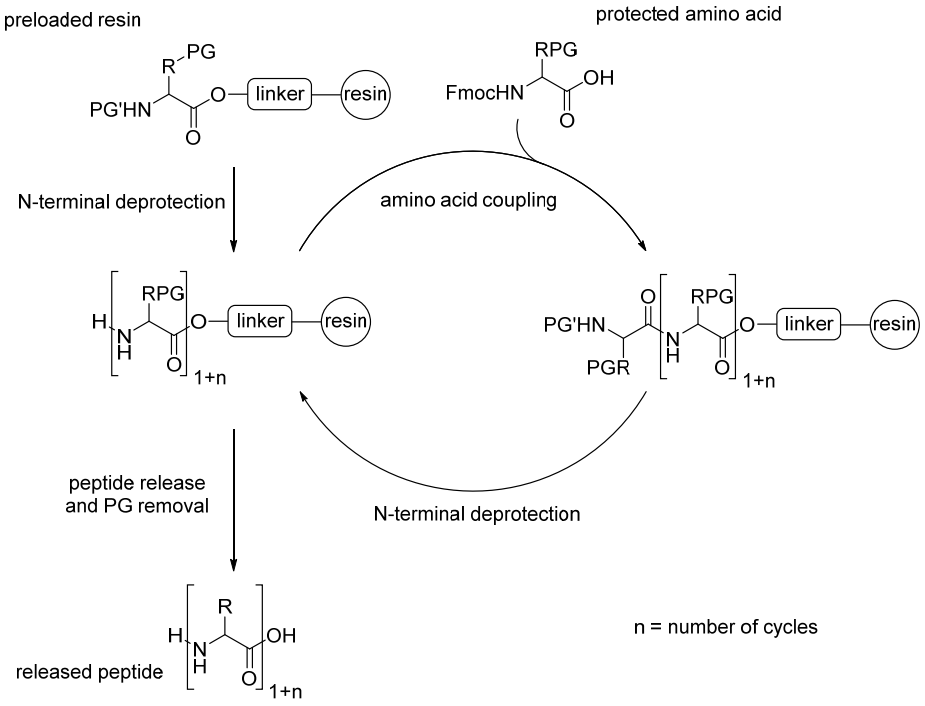


Figure 17: Principle of solid phase peptide synthesis.

When SPPS was introduced by *Merrifield* in 1963[222], an acid (with TFA) labile *tert*-butyloxycarbonyl (Boc) was used as a temporary amine protecting group during peptide assembly and aqueous hydrogen fluoride for the release of the peptide from the solid-support. This strategy is one suitable approach for synthesis of non-modified peptides, but is not applicable for glycopeptide synthesis due to the acid lability of *O*-glycosidic bonds. Even if the stability of the glycans can be improved by global acetylation, most glycopeptides would still suffer from the treatment with the very acidic hydrogen fluoride (HF). It has been found that the Fmoc-SPPS strategy is the most suitable method to address these problems[223]. Hereby piperidine in DMF is used, which facilitates removal of the temporary Fmoc-amine-protecting group without being basic enough to induce β-elimination of the carbohydrate from Ser and Thr or inducing racemization of the amino acids.[224] Following the principle of orthogonal protection, the fully assembled peptide is then released from the solid-support under acidic conditions using trifluoroacetic acid (TFA) with simultaneous global amino acid side-chain deprotection. The TFA treatment is mild compared to the HF conditions used in the Boc strategy and is therefore not affecting the peracetylated carbohydrates. Nevertheless, side-reactions might occur during TFA resin cleavage due to uncontrolled addition of carbocations, which are generated from the release of the side-chain protection groups. In this work addition of triisopropylsilane (TIPS) was used as a scavenger in the cleavage cocktail to prevent side-reactions.

As solid-support preloaded Fmoc-AA-TentaGel® R TRT resins (*Rapp Polymere GmbH*, Tübingen) were utilized in this work (***Figure 18***). The TentaGel beads consist of a low cross-linked polystyrene matrix with attached polyethylene glycol (PEG) chains. The PEG spacer dominates the physico-chemical properties of the co-polymer and provides with hydrophobic as well as hydrophilic properties, allowing the resin to swell in organic polar solvents like DMF, which is commonly used in peptide synthesis. At the end of the PEG chain a trityl linker is attached, which is stable during the conditions used for peptide assembly, but cleavable under the acidic conditions used for resin release of the synthesized peptide. The trityl group prevent diketopiperazin side-reactions and facilitates resin release under mild conditions. The cleavable trityl linker connects the resin with the preloaded amino acid, which represents the C-terminal and thus the first amino acid involved in the peptide synthesis. The amount of amino acids cross-linked to the polystyrene backbone defines the loading capacity of the resin.

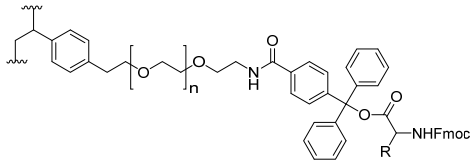


Figure 18: Preloaded Fmoc-AA-TentaGel® R TRT resin.

As coupling reagents for promoting the amide formation in each iterative step involving non-glycosylated amino acids a system of 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) and 1-hydroxybenzotriazole (HOBt) has been used. These reagents provide high coupling yields while rendering the reaction resistant to racemization[225]. The amino acids were usually applied in 8-fold excess with respect to the molar resin capacity and the coupling reagents were used in slightly less than equimolar quantities relative to the amino acids to prevent capping of the *N*-terminus of the growing peptide. Diisopropylethylamine (DIPEA) was additionally used as base. In total the relative stoichiometric amounts of the coupling reagents were 1 : 0.95 : 0.95 : 2 (amino acid : HBTU : HOBt : DIPEA).

For the more bulky and precious glycosylated amino acids the coupling system was exchanged for the more reactive 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate (HATU) with 1-hydroxy-7-azabenzotriazole (HOAT) coupling reagents. During these couplings only 1.5-2.0 fold excess of the valuable glycosylated amino acids were used, dissolved in reduced amounts of DMF and kept over longer reaction times. The amino acid/coupling reagent ratio was kept accordingly to the above mentioned HBTU/HOBt system. All used coupling reagents are depicted in ***Figure 19*****a**.

Both, HBTU and HATU exist in two isoforms: The uronium salt (*O*-form) and the less reactive, but more abundant aminium salt (*N*-form)[226]. The amide formation with the HATU *N*-form is shown in ***Figure 19b***. The mechanism with participation of HBTU is analogue. In the first step the deprotonated C-terminal carboxylic acid of amino acid A is activated via an *O*-Acyl(tetramethyl)isouronium salt intermediate. This is followed by a transesterification with acyl transfer to the exocyclic oxygen of the triazol anion. By hydrogen bonding to the endocyclic nitrogen atom in the pyridine ring amino acid B is positioned in close proximity to the activated amino acid A, leading to a seven-membered ring transition state under formation of the desired amide bond[227]. This neighboring effect might be responsible for the difference in reactivity of HATU towards HBTU, which is missing the endocyclic hydrogen acceptor.

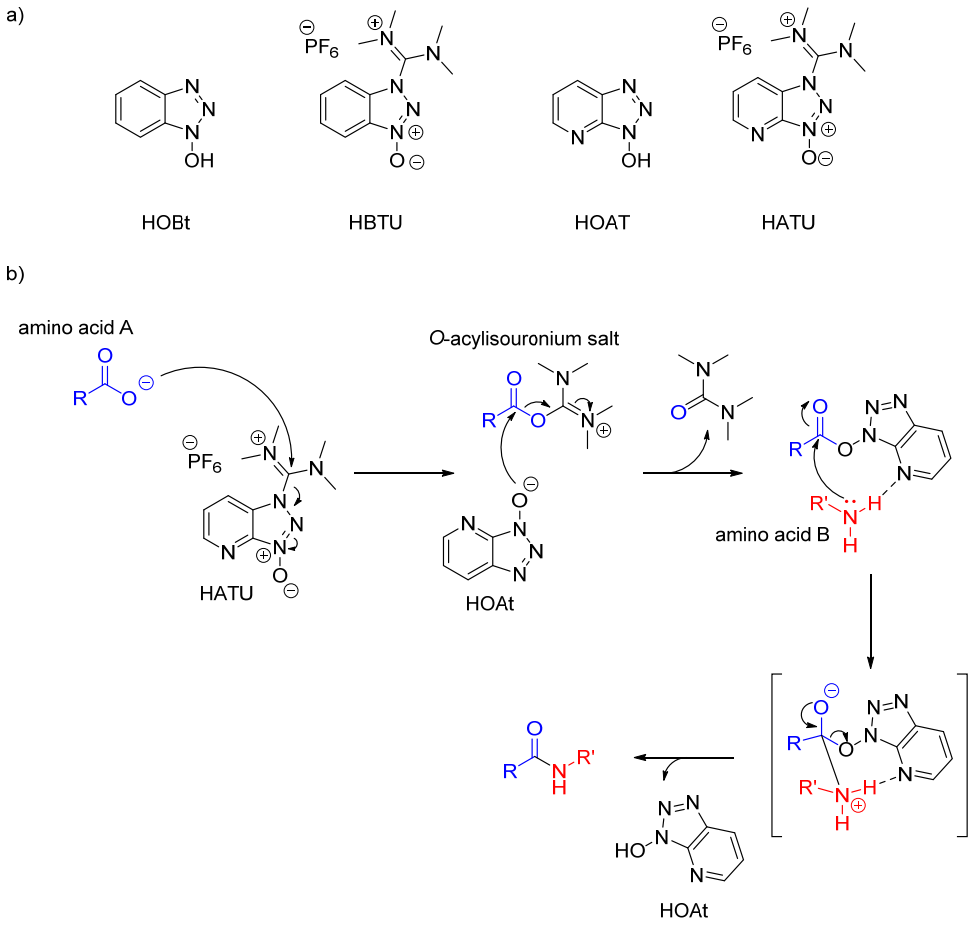


Figure 19: a) Coupling reagents used for SPPS. b) Mechanism of HATU mediated amide formation.

After completion of the target sequence and release from the resin with the cleavage cocktail (TFA : TIPS : H_2O – 15 : 1 : 1, v/v/v), unglycosylated peptides were directly purified by reverse-phase HPLC. The prepared glycopeptides were desalted via C18-Cartridge instead followed by a global glycan deacetylation step. This was carried out using a mild sodium methoxide/methanol system according

to Zemplén[228], keeping the *p*H at 9-10 for the large core 4 glycosylated peptides and sialylated peptides or with a sodium hydroxide/water/methanol system, keeping the *p*H at 10-11 for the core 2 glycosylated peptides and peptides containing monoglycosylated Tyr. During this step, the correct *p*H window is critical for the successful deacetylation without harming the glycopeptide.

At a *p*H lower than 8.5 no or slow conversion takes place, whereas a too high *p*H can cause deprotonation of the α-carbon within the glycosylated amino acid, leading to an equilibrium of D- and L-amino acid via an enolate intermediate (***Figure 20***). Additionally, irreversible β-elimination of the glycan from the peptide backbone with epimerization of the α-carbon within the glycosylated amino acid can occur.[229] In this case, the glycoside undergoes an irreversible E1cB mechanism to form the α,β-unsaturated alkene from glycosyl Thr or Ser. By contrast, glycosyl Tyr is unaffected by this reaction since it is lacking a corresponding leaving group. However, it has been shown that these base-catalyzed side reactions progress very slow compared to the glycan deacetylation process. This racemization protection of the α-carbon is attributable to competitive deprotonation of the adjacent amide group forming a protective aza-enolate[230]. For this reason, N^{α}-unprotected amino acids are more prone to undergo racemization and elimination.[231]

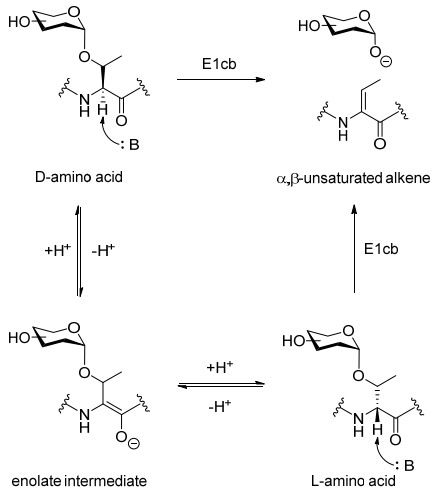


Figure 20: Potential side reactions of the base catalyzed global deacetylation shown exemplified on threonine.

A summarized overview of the Fmoc-SPPS glycopeptide synthesis individual steps and the corresponding reagents is shown in ***Figure 21***.

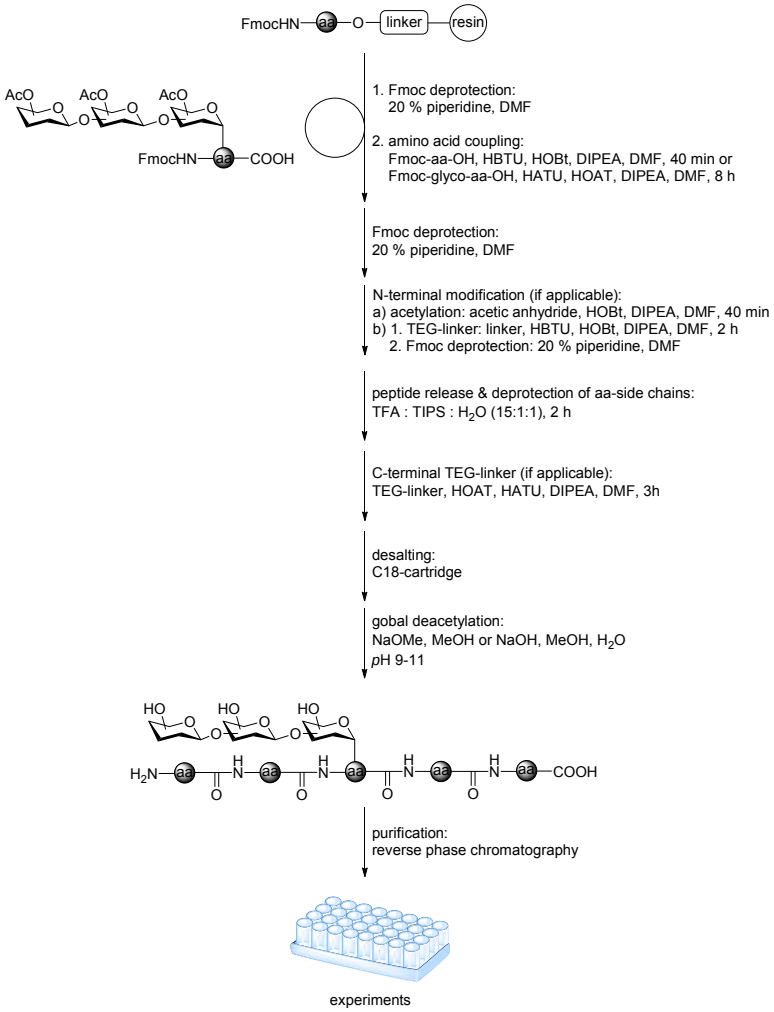


Figure 21: Steps and reagents of glycopeptide synthesis.

5.6 Enzymatic carbohydrate and glycopeptide synthesis

An alternative to the full-synthetic approach on the way towards regio- and stereochemically defined complex carbohydrate conjugates is the utilization of enzyme catalysis. Given that a suitable enzyme is at hand, carbohydrates can be conjugated via glycosyltransferases[232] to acceptor molecules or removed via glycosidases[233]. In general, glycosyltransferases can be sub-divided into non-Leloir-type and Leloir-type glycosyltransferases. The non-Leloir-type transferases use glycosylphosphates and oligosaccharides as carbohydrate donors and are rather uncommonly used in glycoconjugate synthesis. In contrast, Leloir-type transferases use monoglycosyl nucleotides as carbohydrate donors and many glycosyltransferases of this type have been isolated from biological sources and applied in synthesis. The acceptor can be a carbohydrate chain, a peptide/protein or a lipid. According to the

anomeric configuration of the sugar nucleotide and the product conjugate, Leloir-type glycosyltransferases are further subdivided into retaining and inverting transferases (***Figure 22***).

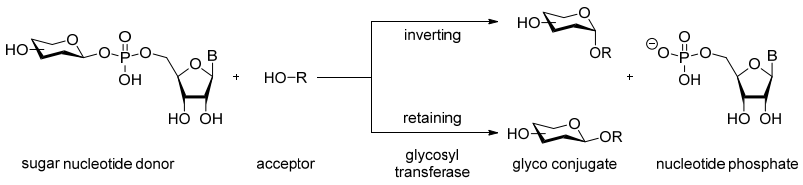


Figure 22: Reactions of Leloir-type inverting and retaining glycosyltransferases.

Due to the high acceptor selectivity of the glycosyltransferases and the fact that many transferases work at similar conditions it can be possible to generate complex carbohydrate structures in a one-pot approach by application of a set of different glycosyltransferases together with corresponding glycan donor conjugates. A prominent enzyme catalyzed *O*-glycosylation example is the synthesis of the SLe^x antigen and its analogs from LacNAc-based acceptors by tandem application of α3SiaT and α3FucT[234] or its synthesis from a GlcNAc-based acceptors by application of β4GalT, α3SiaT and α3FucT[235].

In addition to glycosyltransferases, glycosylation via glycosidases has been exploited for carbohydrate synthesis. The cellular role of glycosidases is the hydrolysis of glycosidic bonds. However, it is possible to reverse the equilibrium towards the glycoside-bond formation instead of glycan hydrolysis and thereby use the glycosidases for the formation of glycosidic linkages. The glycosidases are classified according to amino acid sequence similarities and to date over 140 family number entries are in the CAZy database (**C**arbohydrate **A**ctive en**Zy**mes). The glycosidases commonly used for carbohydrate synthesis are exo- and endoglycosidases. The difference lies within the transferred glycan subunit to the acceptor. While exoglycosidases only transfer the terminal monosaccharide, endoglycosidases transfer the complete oligosaccharide. Thereby the linkage formation can proceed in two different ways, depending on the reaction conditions and the structure of the glycosyl donor substrate (***Figure 23***). In the equilibrium process, which is called *reverse hydrolysis*, a free saccharide is linked to an acceptor with loss of one water molecule under thermodynamic control. Since the natural equilibrium strongly favors the hydrolysis over the linkage formation, it has to be shifted towards the latter. This can be achieved for instance by reducing the amount of water in the reaction solution by a high acceptor concentration[236], by salt addition[237] or by product removal from the system[238]. According to its limitation, this strategy is used mainly for the conjugation of sugars to primary alcohols. In contrast to the equilibrium process, the kinetically controlled transglycosylation uses sugar donors, which need to be activated[239] by a suitable leaving group. In addition, also this process requires reduction of the water activity. Since this process is not based on a chemical equilibrium, no product inhibition occurs,

allowing for higher product concentrations and in consequence higher yields[240] are obtained. Therefore this process is the most commonly used one.[241]

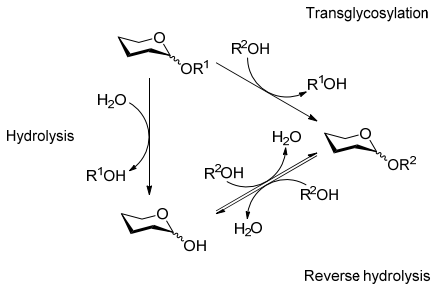


Figure 23: Glycosidase-catalyzed reactions.

In the context of this work, application of glycosyltransferases and glycosidases enable further diversification of carbohydrate core structures on readily synthesized glycopeptides in small quantities. Glycopeptides from this work have been modified with poly-*N*-acetyllactosamine (poly-LacNAc), *N*-acetylneuraminic acid (Neu5Ac), *N*-glycolylneuraminic acid (Neu5Gc), deaminated neuraminic acid (KDN) and fucose at different linkages by Dr. *Yu Jin* and Dr. *Christian Pett*.

6 Bioanalytical Methods

6.1 Glycan and glycopeptide microarrays for studies of binding interactions

Carbohydrate-protein interactions play a key role in many cellular processes like cell-recognition, adhesion or cell-signaling, which is of particular interest in the context of tumor progression and infections. The enormous structural diversity of carbohydrates in combination with the effort necessary for synthesis of defined glycans and glycopeptides represents a challenge in attempts to study these interactions. Thus, a sensitive high-throughput screening methodology is needed, which allows a efficient readout of a multitude of binding experiments under consumption of minor amounts of glycan or glycoconjugate analytes. In the past this challenge was addressed by the attachment of glycoconjugates to different kinds of carriers to create analyte arrays. In an early approach in 1985 neoglycolipids were attached to polyvinylchloride and silica plates[242]. Later in 1996 glycans were synthesized directly on resin beads[243] and in the same year another group immobilized neoglycolipids as mini- and maxiclusters on microplates[244]. The breakthrough was accomplished by several working groups in 2002 by using high-precision robotic arraying to order analytes on solid surfaces in a dense fashion[245-250]. This technology has become the dominant methodology for glycan-protein binding studies as it bears an additional advantage: The multivalent presentation of glycans. This feature mimics the arrangement of glycans as displayed in natural environments, for example on cell-surfaces and is thus allowing the formation of multivalent glycan-protein complexes. This allows for the recognition of glycan-protein interactions, which would be very weak if tested in solution instead, since many lectin- and antibody-recognitions depend on multivalent glycan binding[251]. The libraries for glycan microarray platforms can be generated using chemical or chemo-enzymatic synthetic approaches or by using glycans isolated from biological samples. Because of the very complex nature of a complete glycan microarray fabrication, the initiative *Consortium for Functional Glycomics* (CFG) was founded, which provides the research community with ready glycan microarray platforms. The current version (v5.3) contains 600 glycans on the mammalian-type platform and 300 targets on the microbial platform. Besides direct glycan immobilization, carbohydrates can also be immobilized as glycoconjugates like glycolipids or glycopeptides, as done in this work.[252]

6.1.1 Immobilization methods for glycopeptide microarrays

Many different techniques exist for the immobilization of analytes to a solid surface. In general they are divided into non-covalent and covalent immobilization methods of which each can be subdivided into site-nonspecific and site-specific methodologies. The term 'site-specific' means, that the probe is attached via a defined attachment site in a defined orientation to the slide surface. Accordingly, 'site-nonspecific' means that the probe orientation and linkage occurs randomly throughout the pool of immobilized analytes.

An example for non-covalent site-nonspecific immobilization is the spotting of unmodified glycans on glass slides modified with a nitrocellulose surface[246]. However, in this methodology the adsorbing strength is depending on the physical properties of the spotted analyte, which must fit to the properties of the slide surface coating. This principle is not suitable for glycopeptide libraries with their versatile properties as analyzed in this work. This strategy would result in partial analyte desorption during washing steps. In order to achieve a site-specific non-covalent immobilization, the analyte can be modified with a binding tag like as for example biotin and then be immobilized on a streptavidin-coated slide[253-256]. Here the strong biotin/streptavidin interaction ensures a comparably strong adhesion to the surface. Also this methodology was not applied in the course of this work since the glycopeptide conjugation with biotin would represent an additional step in the preparation of the analyte constructs. In contrast, this extra step is not necessary if the glycopeptides contain a suitable functional group and are covalently immobilized directly to the slide surface.

In covalent approaches the linkage to the slide surface is formed by condensation of intrinsic reactive groups of the probe with corresponding acceptors on the slide surface. Depending on the applied chemistry also the covalent approaches can be subdivided into site-nonspecific and site-specific methodologies. Examples for commonly applied strategies, which lead to site-nonspecific probe immobilization, proceed via coupling of terminal amines (i.e. N-terminal amine or spacer amine) to *N*-hydroxysuccinimide ester (NHS), epoxide or aldehyde coated surfaces[257-260]. Alternatively, the probes can be functionalized with a thiol group to be coupled to either maleimide groups, alkenes or to pyridyl disulfides on the slide surface[245, 261-262]. However, these approaches lack site-specificity if the probes exhibit a lysine or cysteine[263]. Further, the binding epitope of a probe might be disturbed, if an affected lysine or cysteine is part of it. To circumvent these problems and achieve complete site-specificity, several methodologies based on dedicated biorthogonal reaction sites apart from the active site have been developed. Besides Diels-Alder reactions[264-265], Staudinger ligation[266], native chemical ligation[267] also click chemistry based[268] approaches are applied[269-271]. An overview of the most common strategies for probe immobilization on microarray slides is shown in ***Figure 24***.

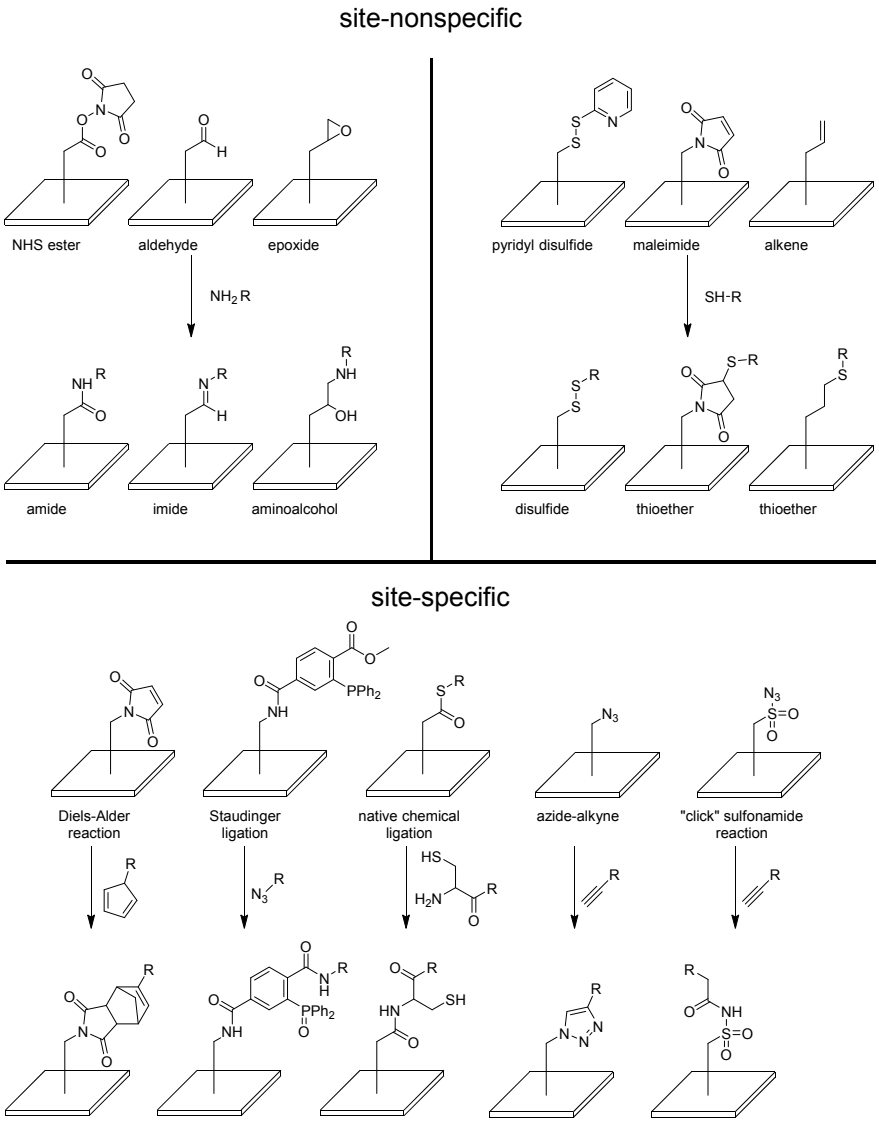


Figure 24: Examples for site-nonspecific and site-specific covalent immobilization strategies for microarray application.

In this work *Nexterion® slide H* microarray slides (Schott GmbH, Mainz, Germany) have been used, which generally immobilized the glycopeptides through an *N*-terminal triethylenglycol linker. The slides were coated with a hydrogel, which consists of a cross-linked hydrophilic polymer (***Figure 25***). This polymer network is functionalized with *N*-hydroxysuccinimide esters via flexible hydrophilic spacers. The whole hydrogel structure of the coating ensures a high accessibility of the covalently attached probes by the screened ligands in solution. This maintains the chemical conformation and specificity of probe and ligand since a direct contact of the biomolecules with the solid surface is prevented.

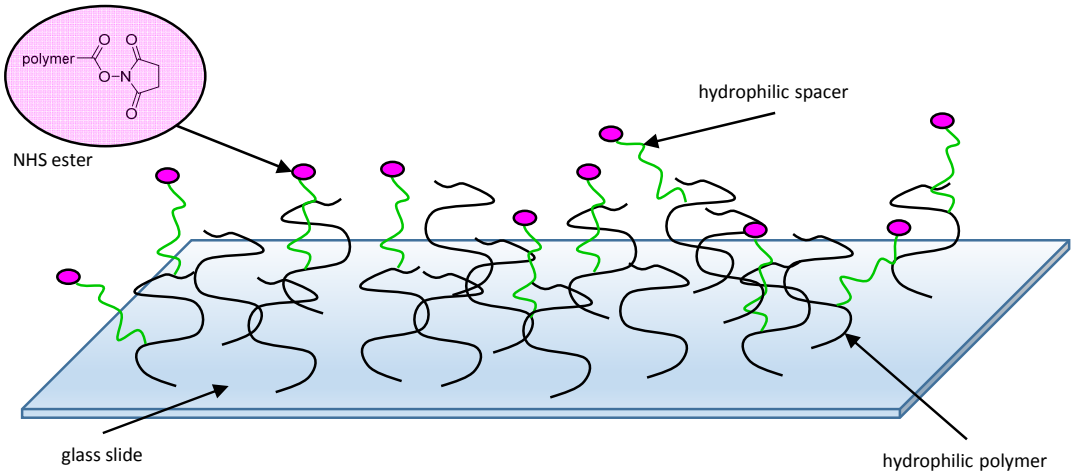


Figure 25: Structure of Nexterion® slide H coating.

6.1.2 General principle of glycopeptide microarray analysis

For each experiment a selection of glycopeptides was attached to NHS-functionalized microarray slides by piezo-driven non-contact spotting using an *iTwo-400* high-capacity micro-dispensing instrument (M2 Automation, Berlin, Germany). Thereby the different glycopeptide spots were arranged in grids, of which each replicate contained a repetitive sequence of glycopeptide spots. The spotted slides were incubated at high humidity (70 – 100 %) to allow maximum reaction between the amino functionalized probes and the NHS ester functionalized slides. Unreacted NHS esters were blocked by subsequent incubation with ethanolamine. By using a silicone structure containing specific frames with wells in various formats (8-, 16- and 48 wells), multiple incubations with different conditions on one slide were enabled. After slides were blocked, incubations were typically made with lectins, antibodies or bacteria diluted in PBS Tween (PBST) buffer, a low amount of Tween in the incubation buffer should prevent unspecific binding. The slides were kept on a shaker in a sealed environment at high humidity to prevent contamination and evaporation during incubation. PBS buffer was typically used for washing between the blocking or incubation steps. For detection, bound lectins, antibodies or bacteria were either directly labeled with a fluorophore or a second incubation step was included with a fluorescence labeled secondary antibody or with labeled streptavidin. The labeled binders could then be detected using a fluorescence scanner or fluorescence microscope with subsequent data processing and analysis using image quant and excel software.

6.2 Mass spectrometry-based analysis of the glycoproteome

The complexity of post-translational glycosylation makes the identification and characterization of glycans and glycosylation sites from biological samples a challenging field. In general, there are two disciplines focusing on the characterization of carbohydrate modifications: Proteomics and glycomics. The glycoproteome is the sum of all glycoproteins in one cell or tissue type, whereas the glycome is defined by the sum of all protein-bound glycans in a cell or tissue type. Accordingly, glycoproteomic studies focus on glycosylated proteins and glycosylation sites, whereas glycomics focus on glycan compositions[272]. In glycoproteomic studies, glycoproteins are typically enriched from cell-lysates in combination with digestion with for example trypsin. Common enrichment methods are for instance the covalent capture of glycoproteins/-peptides with hydrazide[273-274] or boronic acid[275] modified beads, chromatography-based approaches like for example *size-exclusion chromatography* (SEC)[276] or *hydrophilic-interaction chromatography* (HILIC)[277] or lectin-based methodologies like *lectin weak affinity chromatography* (LWAC)[278] or the isolation of lectin-glycopeptide complexes via ultrafiltration[279]. Lectin-based enrichment methods rely on the specificity and affinity of the utilized lectins and thus a detailed knowledge on the lectin binding properties is needed. Investigations on this topic for commonly used lectins are covered in this work (chapter 7.2.6). The enriched glycopeptides are then fragmented via HPLC-MS/MS methodologies in order to identify glycosylation sites. Here, the most common ESI-MS/MS fragmentation technologies are *collision induced dissociation* (CID), *high-energy collisional dissociation* (HCD) and *electron transfer dissociation* (ETD). According to the choice of applied technology, different fragment ion patterns can be induced. A nomenclature to peptide fragment ion patters was proposed by *Roepstorff* and *Fohlmann*[280] (***Figure 26*** A), while carbohydrate fragment ions are classified according to *Domon* and *Costello* nomenclature[281] (***Figure 26*** B).

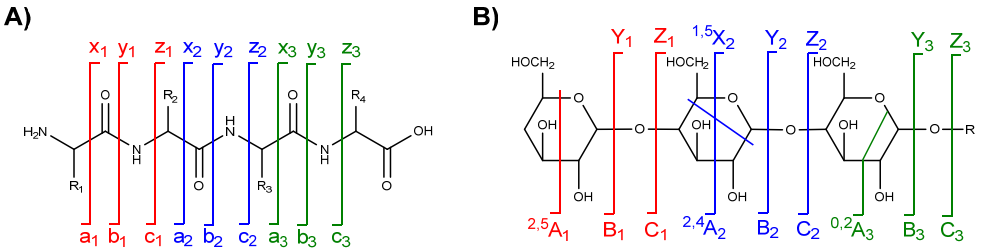


Figure 26: Fragment ion nomenclature for A) peptides and B) glycans.

6.2.1 Common glycoproteomic methodologies

CID experiments on ion-trap-quipped instruments often result in high fragmentation of the glycans and the formation of corresponding B and Y ions, whereas the peptide part remains intact[282-283]. The

evaluation of all different generated fragment ions allows the deduction of the underlying glycan sequence. However, since the different Hex as well as HexNAc isomers exhibit identical masses, a precise assignment of the corresponding monosaccharide subunits cannot be obtained. In an altered multistage procedure, the CID methodology can also be used to identify glycopeptide amino acid sequences[284]. This can be done by subjecting the MS^2 generated Y_1 fragment ions (peptide + HexNAc) as well as the Y_0 ions (peptide only) to CID-MS^3 fragmentation at an increased energy. However, the often occurring absence of glycopeptide Y_1-ion signals hampers the assignment of glycosylation sites. While this methodology is a viable option to obtain peptide sequence data, it is time consuming since many redundant spectra originating from glycan fragmentation are obtained[285].

The HCD method is a variation of the CID method with very high collision energies on quadrupole[286] and C-trap equipped Orbitrap instruments. This results in the fragmentation of the peptide backbone to mainly b- and y-ions, whereas glycans are fragmented to oxonium ions[285] in positive mode. Here, the type of fragmentation can influenced by adjustment of the collision energy. The b- and y-ions (peptide fragmentation) can be interpreted by software supported data analysis in order to deduce the original peptide sequence and – if glycopeptide fragment ions are present – the glycosylation site.

ETD is related closely to the CID technology and enables the peptide backbone fragmentation into c- and z-ions[287-289] with the exception of proline, which indeed is cleaved at the Cα-N-bond, but due to the N-C-ring structure, the peptide remains intact[289]. In contrast to CID and HCD, with ETD fragmentation the glycans remain intact. This allows identification of glycosylation sites, even if multiple Ser/Thr/Tyr residues are present in one glycopeptide. While this methodology is considered to deliver reliable results regarding peptide sequence, glycosylation site and glycan size, a drawback is the need for highly charged analyte ions.

While it is possible to resolve peptide sequences and pinpoint glycosylation sites with these methodologies, the structural information about the glycans is very limited. Neither is a discrimination of the different Hex and HexNAc subunit isomers possible, nor is reliable information about the glycosidic linkages in glycans always available. Without complementary analyses methods, the identification of for example GalNAc and GlcNAc can only be done based on knowledge of biosynthetic pathways and the subcellular localization of the mother glycoprotein. However, such assignments are unreliable, as the *O*-GlcNAc modification has also been found on extracellular proteins[290-291]. Here the complementing data is provided by the research field of glycomics. In typical glycomic approaches glycans are released from the peptide prior to LC-MS/MS analysis[292]. This is either accomplished by the use of enzymes (for example PNGase F for *N*-glycans) or chemically (typically for *O*-glycans). Experiments on isolated glycans simplify the glycan structure analysis by deletion of peptide fragment

ions from the spectra, which reduces the amount of recorded ion peaks. Further, the glycan release allows chemical modifications like permethylation of the carbohydrates[293-294], which in turn enable the elucidation of subunit linkages and the discrimination of Hex and HexNAc isomers. Alternatively, NMR spectroscopy on full glycoconjugates is used to resolve glycan structures. While this methodology enables precise and reliable results, it requires decent amounts of analyte and is very time-consuming. Thus this method is not suitable alone for studies, which aim for a broad characterization of the glycoproteome. Instead strategies should preferably be combined for desirable understanding of diseases and complex glycobiology; analysis of released glycans, analysis of glycosylation sites of deglycosylated peptides and analysis of intact glycopeptides giving information about both glycosylation sites and glycan structures. Within this thesis synthetic glycopeptides have been used as model compounds to enable analysis of isobaric glycans by HCD or CID fragmentation of intact glycopeptides and thereby discrimination of for instance GalNAc and GlcNAc epimers are now possible[189-190].

7 Results & Discussion

7.1 Project 1: Mucin glycosylation in cancer and airway diseases

7.1.1 Synthesis of mucin-type glycosylated amino acid building blocks

Mucin-type O-glycosylation represents a complex and diverse class of modifications, but to some extent the glycans share structural similarities. Basically, all mucin-type O-glycans are connected to a serine or threonine via a αGalNAc residue[295] and in the most common core structures core 1-4 this GalNAc is extended in the 3- and 6-position with βGal and βGlcNAc subunits. Further alternating elongation of the core structures with GlcNAc and 1,3- or 1,4-linked Gal leads to the generation of intramolecular N-acetyl lactosamine (LacNAc) disaccharide subunits, which are distinguished based on their connectivity: type-1 for Galβ1,3-GlcNAc and type-2 for Galβ1,4-GlcNAc[18]. To account for these structural properties, the target core structures were dissected into a convergent synthesis approach containing four common building blocks, the acceptor T_N-antigen **12**, and the donors galactosylbromide **5**, LacNAc type-1 **24** and LacNAc type-2 **29** disaccharide (**Figure 27**). This strategy enabled assembly of the LacNAc elongated core 2 and core 4 target glycosylated amino acids by stepwise extension of the T_N-antigen acceptor **12** with the different glycosyl donors as partly described in our recent published work[296].

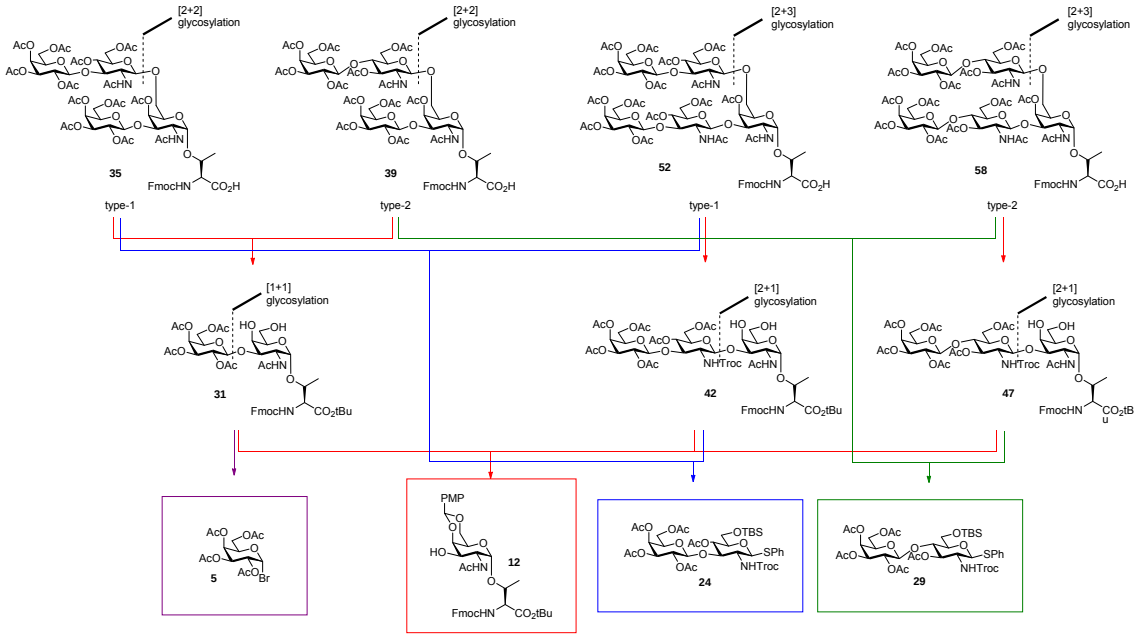


Figure 27: Retrosynthetic analysis of the core 2 and 4 glycosylated amino acid building blocks.

To generate building blocks suitable for Fmoc-SPPS global acetyl protection was selected for preparation of the final glycosylated amino acids and the majority of acetyl groups were introduced from early on in the synthesis route. Acyl esters stabilize glycosidic linkages due to their electron withdrawing nature, which protects the glycans from hydrolysis during the acidic conditions used during peptide cleavage from the solid-support. The acetyl groups can be removed later under mild basic conditions by using sodium methoxide in methanol according to *Zemplén*[228] or for larger glycan structures by using sodium hydroxide in methanol.

To be suitable for Fmoc-SPPS the glycan amino acid building blocks were protected with an Fmoc group at the N-terminus and during glycan synthesis the C-terminus was protected with a *tert*-butyl ester. The latter modification showed orthogonal stability towards the global acetyl protecting groups. Being labile at low *p*H, the *tert*-butyl ester could be hydrolyzed selectively to regenerate the free carboxylic acid in the final step prior to application of the Fmoc-glycan amino acid in Fmoc-SPPS.

During synthesis of the glycan building blocks, acetyl esters, silyl ethers and acetals have been applied for hydroxyl group protection and the amino groups were protected as carbamates or masked as azides. Silyl ethers are known to increase the reactivity of glycosyl donors due to their electron pushing effect[297-298] and can be removed under mild acidic conditions. Acetals allow selective protection of vicinal *cis*-diols in the presence of additional hydroxyl groups and are also hydrolysable under mild acidic conditions. N-Troc-carbamates and acetamides have a β-directing effect due to neighboring

group participation. Moreover, N-Troc groups are known to increase the reactivity of glycosyl donors when compared to acetamides due to less stable oxazolinium intermediate formation[299-300]. For this reason, the amines were temporary protected as carbamates, if the donor was planned to be applied in a β-selective glycosylation reaction. If the donor was designed to participate in an α-selective glycosylation, the amines were masked as azides which are non-participating protecting groups. However, in some cases usage of protection groups was avoided and reactivity differences of free hydroxyl groups was exploited instead. In general primary hydroxyl groups are more reactive then secondary[209] and equatorial hydroxyls are more reactive than axial groups[213]. This strategy as shown in the detailed synthesis description below minimized the number of required reaction steps and saved time and yield.

The common T_N-antigen acceptor **12** was generated by attachment of a galactopyranosyl bromide to the fully protected amino acid acceptor molecule using classical *Koenigs-Knorr* conditions as introduced by *Paulsen* and *Hölck*[301] and modified by *Kunz*[302-303]. To achieve regioselective elongation at the GalNAc 3-position, the T_N-antigen was protected with a 4,6-*O*-benzylidene acetal, making the hydroxyl in 3-position its only remaining nucleophilic reaction site. The benzylidene acetal was chosen since it can be introduced regioselectively at the desired 4- and 6-position and hydrolyzed under orthogonal conditions to the utilized permanent protection groups. The T_N-antigen acceptor represented the common precursor for all synthesized mucin type building blocks.

The T_N-antigen was extended with galactosylbromide **5** to generate the T-antigen (Galβ1,3-GalNAcα-Thr) as the common intermediate in the core 2 type-1 and type-2 synthesis route (chapters 7.1.7 and 7.1.8). This glycosylation was performed according to Helferich[304-306] using mercury(II) cyanide as promotor in a nitromethane/dichloromethane system. After removal of the 4,6-*O*-acetal from the T-antigen to generate the free diol **31** further extension of the GalNAc 6-position was possible with either LacNAc type-1 **24** or LacNAc type-2 **29** resulting in the type-1 or type-2 elongated core 2 amino acids **35** and **39**. Here the higher reactivity of the primary 6-OH group over the secondary 4-OH group in the acceptor was exploited to drive the regioselectivity without prior protection of the 4-OH group. The LacNAc disaccharides **24** and **29** have been introduced as thioglycoside donors using a NIS/TfOH promotor system[307-309]. The introduction as thioglycoside donor was chosen because the thiophenyl group is stable during the trichloroacetimidate coupling conditions, which were applied during the preceding LacNAc disaccharide synthesis (***Figure 35***, p. 61)[310].

In the other synthesis routes the T_N-antigen **12** were extended at GalNAc 3-OH with LacNAc type-1 **24** and type-2 **29** respectively with subsequent acetal removal to generate the intermediates **42** and **47.**

Analog to the core 2 synthesis, regioselective extension on the GalNAc 6-position with LacNAc leads to the target core 4 structures **52** and **58**.

7.1.2 Synthesis of the T_N-antigen acceptor

In the convergent synthesis of the mucin type *O*-glycosylated amino acids, the T_N-antigen (GalNAcα-*O*-Thr) represents the common glycan acceptor threonine amino acid. The strategy applied in this work was introduced by *Hans Paulsen* and *Jens –Peter Hölck* in 1982[301]. The building block is generated in four steps, using a silver carbonate and silver perchlorate system as promotor for the glycosylation of the threonine acceptor amino acid with a bromide donor using *Koenigs-Knorr* conditions[203].

The galactosyl acceptor is prepared in two steps (***Figure 28***), starting with the N-terminal Fmoc-protection of L-threonine using Fmoc-OSu and sodium bicarbonate in acetone/water[311].

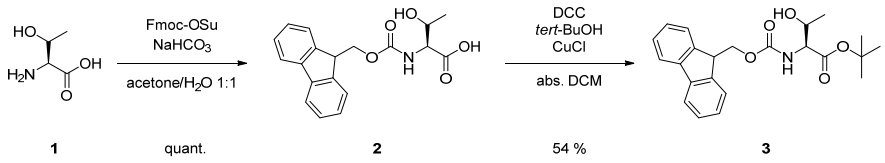


Figure 28: Synthesis of the threonine glycosyl acceptor **3**.

The base labile Fmoc serves as a permanent protecting group throughout the glycosylated amino acid building block synthesis, while during the following SPPS it serves as temporary protecting group. In the second step the C-terminal carboxylic acid was converted to a *tert*-butyl ester using DCC/CuCl as promotor system[312]. Here, the Copper(I) is coordinated by a nitrogen, leading to an enhanced reactivity of the carbodiimide carbon towards the *tert*-butanol. Nucleophilic attack of the alcohol leads to an *O*-*tert*-butyl-*N,N'*-dicyclohexylurea intermediate (***Figure 29***).[313-314] After butyl transfer, the *tert*-butyl ester serves as a permanent protecting group during the glycosylated amino acid synthesis and has to be removed before SPPS in order to enable C-terminal coupling. The threonine acceptor **3** was synthesized in two steps with a total yield of 54 %.

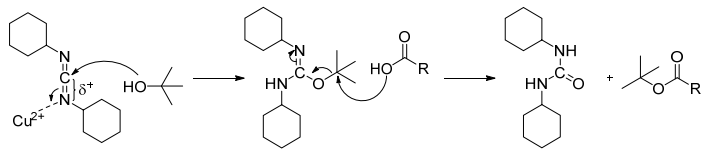


Figure 29: Mechanism of the tert-butyl esterification with CuCl/DCC.

The galactosyl donor **8** was prepared in four steps (***Figure 30***) starting with peracetylation of D-galactose using acetic anhydride and catalytic amounts of perchloric acid.

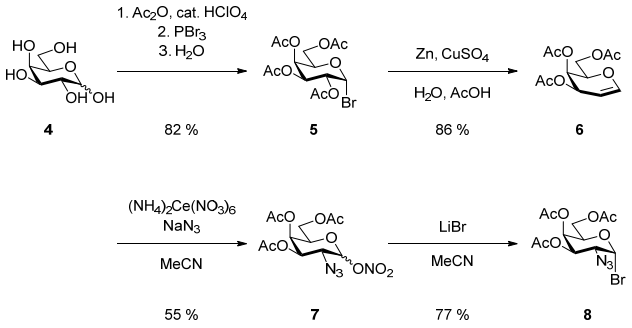


Figure 30: Synthesis of the galactopyranosyl donor **8**.

Subsequent addition of PBr_3 and water leads to the formation of α-galactosylbromide **5** in a one-pot reaction. After reductive elimination with copper(II) sulfate and activated zinc dust in H_2O/AcOH, 3,4,6-tri-*O*-acetyl galactal **6** was obtained under formation of ZnBrOAc.[315-317] For azidonitration according to *Lemieux*[318], the galactal was then treated with cerium(IV) ammonium nitrate and sodium azide in acetonitrile to generate 3,4,6-O-acetyl-2-azido-2-deoxy galactosyl nitrate **7**. This reaction proceeds via a radical mechanism[200], where the azide ion is oxidized first, giving one electron to Cer(IV) and thus reducing it to Cer(III). The generated azide radical then attacks the galactal double bond from the β-side. After following single electron transfer steps a stabilized 2-azido oxocarbenium ion is formed. This is then attacked by a nitrate ion to give the azido nitrate product in an anomeric mixture with α:β = 1:1 (***Figure 31***).

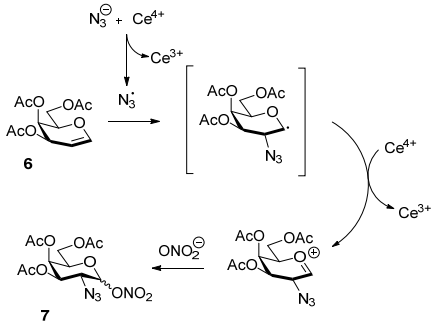


Figure 31: Mechanism of the azido nitration reaction.

Finally the anomeric nitrate was substituted using lithium bromide in acetonitrile to generate the α-anomeric 3,4,6-tri-*O*-acetyl-2-azido-2-deoxy-D-galactosyl bromide **8**.

Having the galactosyl donor **8** and the Thr acceptor **3** ready, the T_N-antigen was formed applying *Koenigs-Knorr* conditions[203]. Using a system of catalytic amount of silver perchlorate together with an excess of silver carbonate, the reaction proceeds in an S_N1-manner over an oxocarbenium ion

intermediate, leading to the formation of the thermodynamic α-glycoside[319] (***Figure 32***), which was obtained in a yield of 76 %. The excess of silver carbonate serves the regeneration of consumed silver perchlorate during the reaction. The stereoselective outcome of this reaction is determined by the anomeric effect, which is enabled by the non-participating 2-azido group of the galactosyl donor **8**.

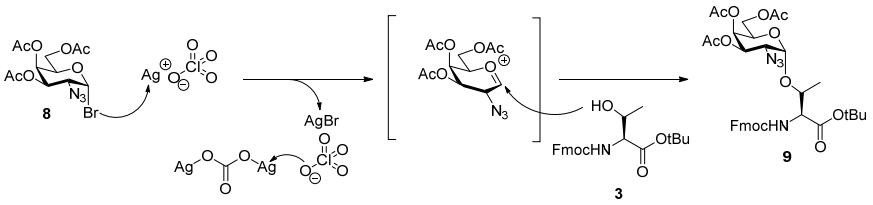


Figure 32: Mechanism of Koenigs-Knorr Glycosylation to generate the T_N-antigen.

After the glycosylation, further protection group manipulations have to be done to prepare the molecule for further selective glycosylation at the GalNAc 3-position (***Figure 33***)[304].

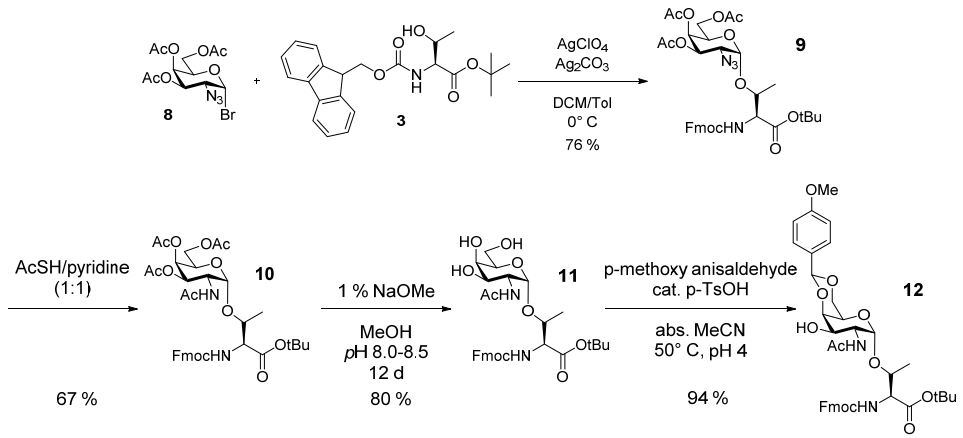


Figure 33: Synthesis of the T_N-antigen acceptor building block.

The azide **9** was reduced to the corresponding acetamide **10** using thioacetic acid in pyridine[320]. The mechanism starts with the formation of a bond between the terminal azide nitrogen and the sulfur, followed by the formation of a bond between the ring-connected nitrogen and the carbonyl carbon of the thioacetic acid. This leads to a thiatriazoline intermediate, which undergoes a retro-[3+2] cycloaddition to form the amide product (***Figure 34***).[321]

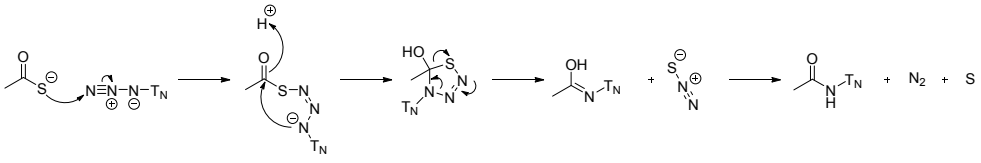


Figure 34: Mechanism of the azide reduction and amide formation.

The peracetylated GalNAc-Thr building block was then deacetylated selectively on the 3-,4- and 6-position by transesterification using sodium methoxide in methanol, as described by *Zemplén*[228] to generate compound **11**. During this reaction the *p*H was maintained between 8.0 and 8.5 enabling mild conditions and thus avoiding deprotection of the base labile Fmoc-group as well as racemization or β-elimination of the glycan. Finally the 4- and the 6-OH group were simultaneously protected as the *cis*-fused *para*-methoxybenzylidene acetal **12** using *para*-methoxybenzaldehyde dimethyl acetal in acetonitrile and catalytic amounts of *para*-toluenesulfonic acid[322]. While acetal formations have a general preference for *cis*-diols in a carbohydrate ring, benzylidene acetals have the additional property to be most stable at six-membered rings with the phenyl substituent being positioned in the thermodynamically favored equatorial position[200]. This, along with the highest reactivity of the primary 6-OH, leads to selective formation of the 4-6-*O*-acetal. This modification allows regioselective glycosylation at the 3-OH group. The protecting group manipulations proceeded with a total yield of 50 % over three steps. The total synthesis of the T_N-antigen glycosylated amino acid had an overall yield of 11 % starting from the galactose starting material **4**.

7.1.3 The LacNAc disaccharide

The LacNAc disaccharides consist of a galactose connected via a β-glycosidic bond to *N*-acetylglucosamine. Depending on the connectivity of the glycosidic bond the glycoside is distinguished as a type-1 (Galβ1,3GlcNAcβ) or type-2 (Galβ1,4GlcNAcβ) disaccharide. The synthesis of type-1 and type-2 disaccharide donors was done according to previous work from the group[296, 323]. In the synthetic approach, both regioisomer building blocks shared the same trichloroacetimidate glycosyl donor **15**, which was used for glycosylation of a 4,6-*O*-benzylidene acetal protected thioglycoside **21** (***Figure 35***, glycosylation A) to obtain the type-1 disaccharide[324] or the 6-*O*-TBS protected thioglycoside **25 (*Figure 35***, glycosylation B) to obtain the type-2 disaccharide. In the latter case, the regioselectivity of the reaction is defined by the difference in reactivity of the 3-OH and the 4-OH groups within the glucosamine acceptor[300, 325]. The participating acetyl group at the 2-position of the galactosyl donor promotes nucleophilic attack on the oxocarbenium ion from the β-side, circumventing the anomeric effect. In the formation of the LacNAc type-1 disaccharide, the 3-OH of the Troc-protected glucosamine donor were of low reactivity and for the reaction to proceed, azide protection was required instead.

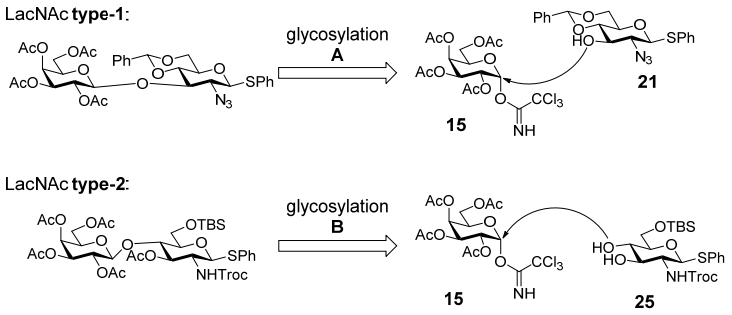


Figure 35: Strategy for synthesis of the type-1 and type 2 disaccharides.

7.1.4 Synthesis of the LacNAc type-1 disaccharide donor

The LacNAc type-1 disaccharide donor **24** was assembled by β-1,3-*O*-glycosylation of acceptor **21** with donor **15** ***(Figure 35***).

The synthesis of the donor **15** (***Figure 36***) started with the peracetylation of D-galactose **4**, using acetic anhydride and catalytic amounts of perchloric acid to obtain compound **13** as an anomeric mixture with α/β = 5:1.

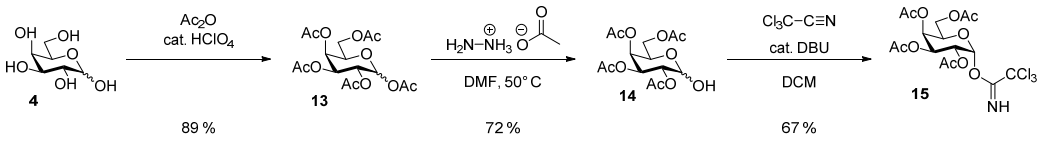


Figure 36: Synthesis of the galactosyl trichloroacetimidate donor **15**.

To enable trichloroacetimidate functionalization, a free anomeric hydroxyl group is needed. Thus the corresponding acetyl group was removed selectively using hydrazine acetate in dimethylformamide[326] to obtain compound **14**. Finally the trichloroacetimidyl residue was introduced according to *Schmidt*[205, 327] using trichloroacetonitrile in dichloromethane. The stereochemical outcome of this reaction was controlled by the choice of the base. Formation of α-acetimidates is favored in the presence of a strong base (e.g. sodium hydride, DBU), the presence of a weak base (e.g. potassium carbonate) leads to predominant formation of β-acetimidates. The reason for this lies in the different reactivities of the α- and β-oxyanion intermediates, being present in an equilibrium (***Figure 37***). Treatment with a weak base leads to fast reaction of the more reactive β-oxyanion to the corresponding β-acetimidate (kinetic product), pulling the oxyanion equilibrium to the β-anomer. In contrary, treatment with a strong base leads to backreaction of the β-acetimidate to the corresponding β-oxyanion. Irreversible reaction of the α-oxyanion to the corresponding α-acetimidate

pulls the oxyanion equilibrium to the α-anomer, resulting finally in exclusive formation of the α-acetimidate (thermodynamic product).[204]

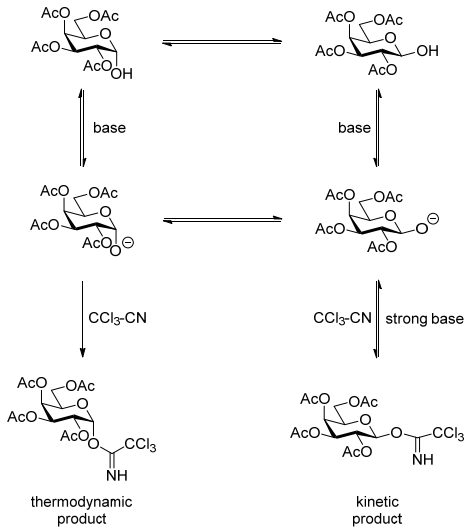


Figure 37: Formation of the galacopyranosyl trichloroaceimtimidate.[204]

During this synthesis the addition of catalytic amounts of the non-nucleophilic, strong base DBU resulted in formation of the α-acetimidate. The galactosyl trichloroacetimidate donor **15** was synthesized in three steps with a total yield of 43 %.

The thioglycoside acceptor **21** was synthesized with an azide in 2-position, which provides the glycosyl acceptor with a good reactivity while being little sterically demanding. To this extent, first imidazole-1-sulfonyl azide **16** has to be prepared. This diazotransfer reagent was described for the first time in 2007 by *Goddard-Borger* and *Stick* for the generation of an azide on D-glucosamine[328]. The compound is known to be more shelf-stable than the classically used triflic azide which is prone to explosions[329]. The diazotransfer donor was prepared in a one-pot reaction by generating the chlorosulfonyl azide in situ using equimolar amounts of sodium azide and sulfonyl chloride in acetonitrile. Then addition of imidazole gave the imidazole-1-sulfonyl azide, which was converted to the corresponding more stable hydrochloride **16** using methanolic hydrochloric acid (***Figure 38***). The qualification as a diazo donor is based on the electron-withdrawing effect of its imidazole-1-sulfonate (imidazylate) group.

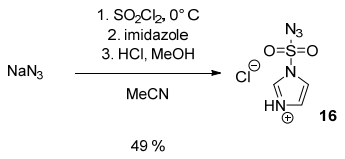


Figure 38: One-pot synthesis of imidazole-1-sulfonyl azide **16**.

With the diazo donor **16** at hand, the D-glucosamine hydrochloride **17** was converted to the azide **18** using catalytic amounts of copper(II) sulfate and sodium carbonate (***Figure 39***)[328-329].

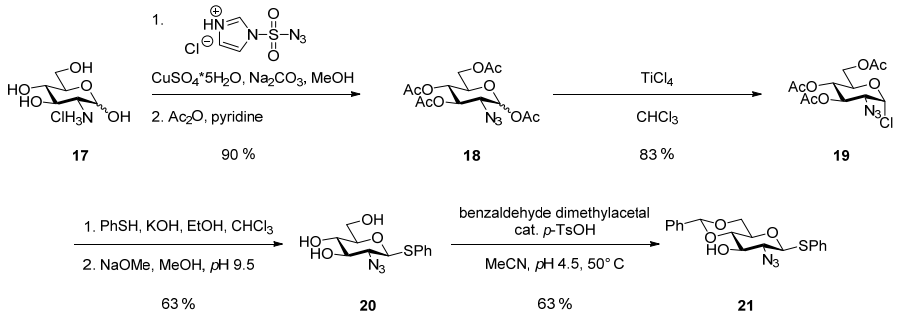


Figure 39: Synthesis of the LacNAc type-1 acceptor **21**.

Here the copper(II) coordinates to the proximal nitrogen of the donor azide and the amine, forming a mixed tetrazene (***Figure 40***)[330-331]. The electron withdrawing imidazylate causes a highly electron deficient terminal azide nitrogen, which is then attacked by the amine, being in close proximity due to the copper(II) coordination. This probably leads to a reverse [3+2] dipolar cycloaddition resulting in the formation of the target azide and the imidazylate amine. Subsequent acetylation affords compound **18** as an anomeric mixture with α/β = 5:1.

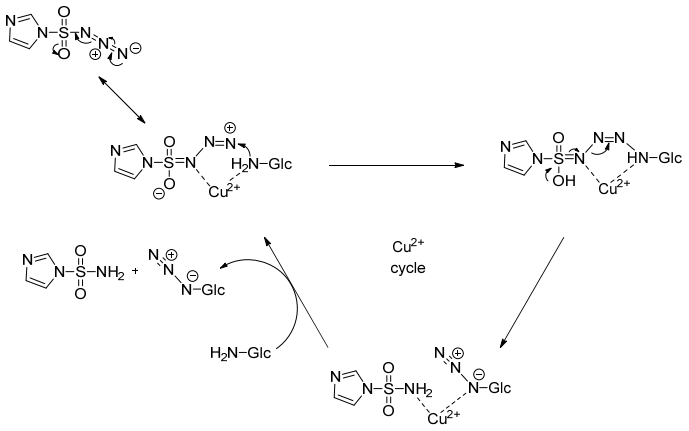


Figure 40: Mechanism of the copper(II) catalyzed diazotransfer reaction with imidazole-1-sulfonyl azide **16**.

Compound **18** could not directly be converted to a desired β-thioglycoside donor and because of the non-participating azide, direct introduction of a thiophenyl residue would result in an anomeric mixture. The thermodynamically more stable α-thioglycoside anomer is less reactive compared to the β-anomer, which would be a problem in the later glycosylation reaction. To address this problem, the α-glucopyranosyl chloride was generated from **18** by employing titanium(IV) chloride[332-333] in chloroform under thermodynamic conditions to substitute the anomeric acetyl group with an α-chloride. In the next step this chloride was substituted with a thiophenyl residue under inversion in an

S_N2-manner using thiophenol in alkaline conditions. After removal of the acetyl groups according to *Zemplén*[228] the thioglycoside **20** was afforded[334]. To ensure selective glycosylation on the 3-position of the glucopyranosyl acceptor, the equally or more reactive 4- and the 6-positions was protected. To this extent the acid labile benzylidene acetal was generated, forming a six-membered ring with the aryl residue in the thermodynamically favored equatorial position. This was done by treatment of thioglycoside **20** with benzaldehyde dimethylacetal and catalytic amounts of *para*-toluene sulfonic acid resulting in the LacNAc type-1 glucopyranosyl acceptor **21**[335]. In total the thioglycosyl acceptor **21** was synthesized from glucosamine in four steps with a total yield of 30 %.

The galactosyl trichloroacetimidate donor **15** and the glucosyl acceptor **21** were coupled in analogy to a similar procedure from *Toepfer* and *Schmidt*[324] applying catalytic amounts of trimethylsilyl trifluoromethanesulfonate in diethyl ether, which afforded the LacNAc disaccharide **22** (Galβ1,3GlcNAcβSPh) in quantitative yield (***Figure 41***)[323]. The participating character of the 2-*O*-acetyl group of the trichloroacetimidate donor ensures formation of the β-anomer solely. The LacNAc amine has to be masked as an azide during this glycosylation reaction. In earlier work from the group the analog reaction with an *N*-2,2,2,-trichloroethoxycarbonyl chloride (Troc) acceptor directly was additionally explored, which results in breakdown instead of disaccharide formation[153].

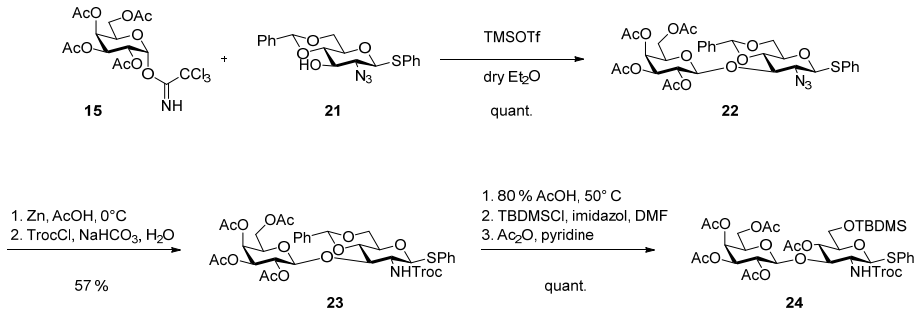


Figure 41: Synthesis of the LacNAc type-1 disaccharide.

To enable a participating effect in further glycosylation reactions of the complete LacNAc type-1 donor, the azide group was reduced to an amine which was subsequently protected with an *N*-Troc group. This protection group has the additional advantage of resulting in more reactive glycosyl donors when compared to other common *N*-protecting groups, like for example the phtalimido group or acetamides. Additionally, the Troc group can be removed under mild conditions, which is required after coupling of the disaccharide donor to the glycosylated amino acid. The reduction of the azide was carried out using zinc dust in acetic acid, leading to the free amine. These conditions were mild enough to promote the reduction of the azide, while still keeping the acid labile benzylidene acetal intact[153, 300]. After subsequent amine protection with Troc chloride, the carbamate **23** was obtained (***Figure 41***).

To have a LacNAc type-1 disaccharide donor building block suitable for the extension of the T_N-building block to generate the core 3 structure, the benzylidene acetal needed to be removed. The GalNAc residue from the T_N-antigen amino acid contained an acid labile 4,6-*O*-*para*-methoxybezylidene acetal, which protected the 6-OH from elongation during the core 3 trisaccharide synthesis. In the GalNAc 6-position, the acetal later needed to be removed prior to that final LacNAc elongation forming the core 4 pentasaccharide. To prevent simultaneous deprotection of the T_N-antigen acetal and the LacNAc acetal during that deprotection step, the GlcNAc acetal was removed already prior to conjugation of the completed LacNAc subunit with the T_N-antigen building block. In addition to these considerations, acetal protecting groups have been found to show partially torsional and partially electronic disarming effects on pyranosyl glycosides. *Fraser-Reid* et al. found, that acetalated pyranose glycosides showed lowered reactivity towards oxocarbenium generation compared to non-acetalated ones[214]. This is the result of a lock-in effect by the trans-fused acetal, forcing the saccharide to the $^4C_1$ conformation und thus restricting its flexibility. This makes it difficult to reach a half-chair conformation, as present in an oxocarbenium ion intermediate. This effect is even surpassed by the electronic effect[212]: Normally the hydroxymethyl group of a hexopyranoside occurs in three stagged conformations, the gg, gt and tg variant (***Figure 42***). The gg and tg conformations have a spatial arrangement of the 6-OH, that can be compared with the axial and equatorial 4-OH group in the hexopyranoside. According to the charge dipole hypothesis, in the tg conformer, the C6-O6 dipole is directed with the negative terminus away from the electron-deficient center in the oxocarbenium ion transition state and is therefore the least reactive conformer. In contrast the gg and gt conformers show a corresponding dipole, that is directed perpendicular to the charge of the potential oxocarbenium ion and is therefore favored and more reactive. Typically the tg conformer is the least abundant in hexopyranosides, but the 4,6-*O*-acetal locks the molecule in the tg conformation, which thus leads to a disarming effect for the glycosyl donor.[336]

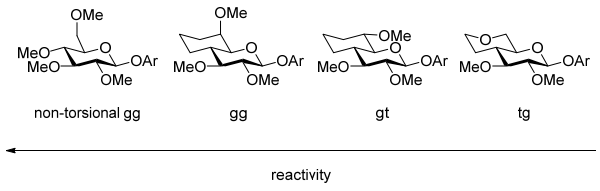


Figure 42: Torsional and electronic effects of 4,6-*O*-acetals[336].

Removal of the 4,6-*O*-benzylidene acetal was carried out using 80 % acetic acid at 50° C. To prevent polymerization of the resulting di-hydroxyl thioglycoside donor during the following glycosylation step, the free 4-OH group and the very reactive 6-OH group have to be reprotected. The 6-position was selectively protected as an *O*-*tert*-butyltrimethylsilyl ether (OTBDMS), which enhances the

reactivity of the thioglycoside donor due to the very low electron-withdrawing effect of silyl groups[298]. The regioselectivity of this protection reaction originates from the higher reactivity of primary hydroxyl groups compared to secondary ones[209]. This favors the development of a positive charge during thioglycoside activation, which is therefore armed[211]. The 4-position was finally acetylated to obtain the ready LacNAc type-1 disaccharide donor **24**. Starting from glucosamine hydrochloride **17**, the thioglycoside **24** was synthesized in seven steps with an overall yield of 17 %.

7.1.5 Synthesis of the LacNAc type-2 disaccharide donor

The LacNAc type-2 disaccharide donor (Galβ1,4GlcNAcβSPh) was assembled by glycosylation of glucopyranosyl thioglycoside acceptor **27** with the above described galactopyranosyl trichloroacetimidate donor **15** (***Figure 35*****,** glycosylation B). For the synthesis of acceptor **27**, the GlcN-Troc protected thioglycoside **26** was generated as described by *Yan* et al.[337], protecting the 2-position first with TrocCl. After subsequent peracetylation, the thermodynamic α-anomer was obtained as major product (***Figure 43***). This reaction is not influenced by neighboring group participation, because the acetylation is not proceeding via formation of an oxocarbenium ion. Nevertheless, the participating effect of the Troc group is driving the stereochemistry during the later LacNAc disaccharide elongation. Additionally, this modification is easily introduced and induces a considerably higher reactivity on glycosyl donors compared to the classical participating phtalimido group[338] thus providing the complete LacNAc type-2 disaccharide donor **29** with a high coupling efficiency during extension of the glycosylated amino acid. Further, *N*-phtalimido protection is not compatible with synthesis of glycosylated amino acids, as phtalimido-removal requires harsh basic conditions, which might result in amino acid epimerization or loss of the glycan via β-elimination. In contrast, the N-Troc protecting group can be removed under mild acidic/reductive conditions. The introduction of the thiophenyl leaving group was done by activation of the anomeric acetyl in compound **25** and its substitution with thiophenol.

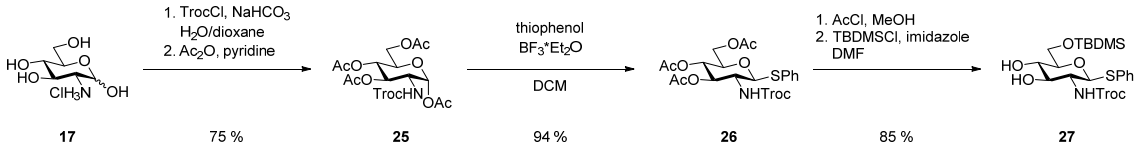


Figure 43: Synthesis of the LacNAc type-2 acceptor **27**.

In addition to the leaving group functionality, the thiophenyl residue serves as an orthogonal protective group, as the thioacetal is stable[310] at the conditions of the following protection group manipulations. Acid catalyzed deacetylation, which is compatible with the Troc-protection of

thioglycoside **26,** proceeded by in situ generated hydrochloric acid in methanol. This enabled a free hydroxyl group in the 4-position as needed for the extension to form the type-2 disaccharide. To achieve regioselective glycosylation at the 4-position, the more reactive primary alcohol[209] at the 6-position was blocked with equimolar amounts of TBDMS chloride. In contrast, the also free 3-hydroxyl was not needed to be blocked since it has been found to be less reactive than the 4-position in GlcNAc[300, 325]. Starting from glucosamine hydrochloride **17**, the glucopyranosyl thioglycoside acceptor **27** was synthesized in three steps with a total yield of 60 %.

The formation of the LacNAc type-2 disaccharide was accomplished by glycosylation of the glucopyranosyl thioglycoside acceptor **27** with the galactopyranosyl trichloroacetimidate **15** (***Figure 44***). Trimethylsilyl trifluoroemethanesulfonate has been used as a promotor for this reaction in an earlier work from the group[153], within this work slightly better results were obtained using trifluoromethanesulfonic acid (74 % yield for trimethylsilyl trifluoroemethanesulfonate, 78 % yield for trifluoromethanesulfonic acid). The stereochemical outcome of this reaction is influenced by the neighboring participation effect of the 2-*O*-acetyl group within the disaccharide donor, promoting 1,2-*trans* glycosylation. Supported by kinetic control at -50° C, the pure β-anomer **28** was obtained with good regioselectivity under formation of the 1,4-*O*-disaccharide exclusively. To prevent side reactions of the disaccharide donor in the following coupling steps, the remaining free GlcNAc 3-hydroxyl group was acetylated. Starting from acceptor **27**, the LacNAc type-2 disaccharide donor **29** was synthesized in two steps with a yield of 78 %.

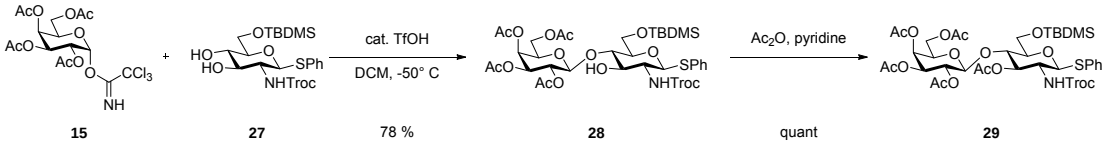


Figure 44: Synthesis of the LacNAc type-2 disaccharide donor **29**.

7.1.6 Synthesis of the T-antigen glycosyl acceptor

The T-antigen or *Thomsen-Friedenreich* antigen amino acid represents the common precursor for synthesis of both type-1 and type-2 elongated core 2 glycosylated amino acids. The T-antigen can be prepared by elongation of the T_N-antigen amino acid with a terminal galactose at the 3-position of the GalNAc subunit (Galβ1,3GalNAcαThr). Within the course of the T-antigen synthesis[304], the T_N-antigen acceptor **12** was coupled with peracetylated galactosyl bromide **8**, which also represents an intermediate of the T_N-antigen synthesis route. The glycosylation reaction was performed according to the *Helferich* conditions in a dichloromethane/nitromethane solvent system using mercury(II) cyanide as promotor (***Figure 45***).[304-306] This method is an advancement of the classical *Koenigs-Knorr*

method[203] for *O*-glycosylation with glycosyl bromides. Due to neighboring group participation of the 2-*O*-acetyl group in the galactopyranosyl bromide, the reaction proceeds with good stereoselectivity, leading to exclusive formation of the β-1,3-*O*-glycosidic bond. After acid promoted hydrolysis of the *para*-methoxybenzylidene acetal, the 4-OH and 6-OH are free on the GalNAc residue and the more reactive primary 6-OH group is ready for further extension. Starting from the T_N-antigen **12**, the T-antigen acceptor **31** was synthesized with a yield of 63 %.

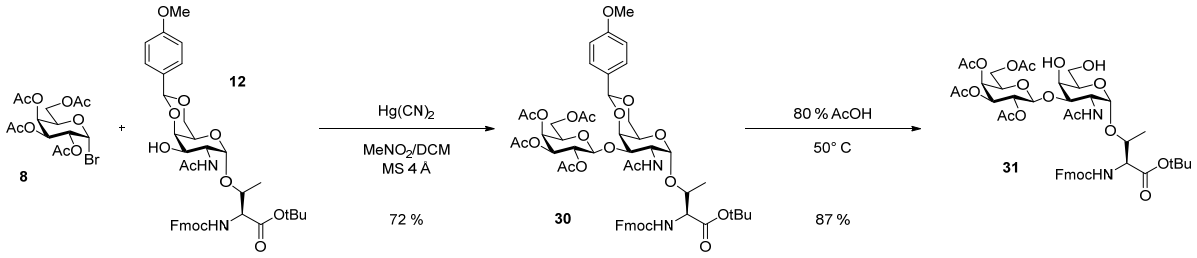


Figure 45: Synthesis of the T-antigen glycosyl acceptor **31**.

7.1.7 Synthesis of the core 2 type-1 glycosylated amino acid

The extended type-1 core 2 glycosylated amino acid is assembled in a [2+2] building block glycosylation of the T-antigen acceptor **31** with the LacNAc type-1 disaccharide donor **24**. The reaction was performed in dichloromethane using an *N*-iodosuccinimide/trifluoromethanesulfonic acid promoter system (NIS/TfOH, ***Figure 46***), which was reported independently by *Fraser-Reid*[307-308] and *Boom*[309] in 1990.

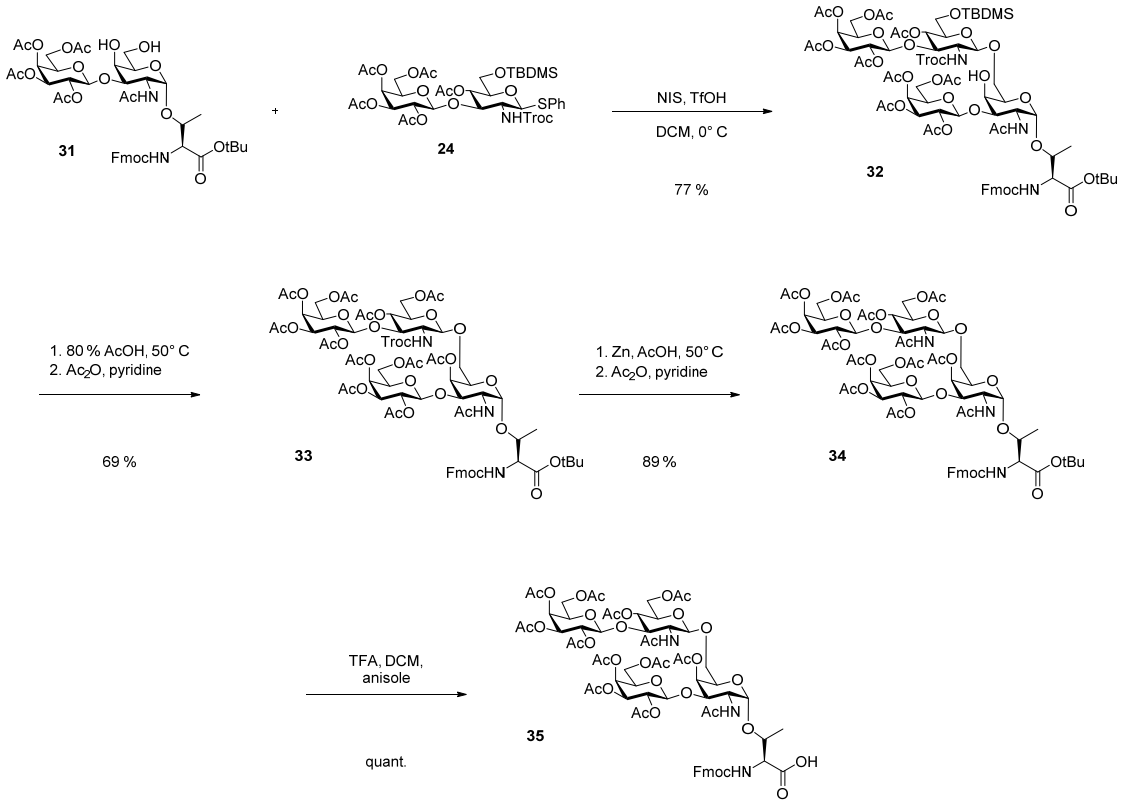


Figure 46: Synthesis of the extended type-1 core 2 glycosylated amino acid building block **35**.

The reaction is based on the *in situ* generation of thiophilic iodonium ions by acidic activation of *N*-iodosuccinimide, which leads to the formation of a sulfonium ion intermediate. This sulfonium ion represents a good leaving group, enabling the formation of the corresponding oxocarbenium ion intermediate, which is finally converted to the acetal by nucleophilic attack of the glycosyl acceptor (***Figure 47***).

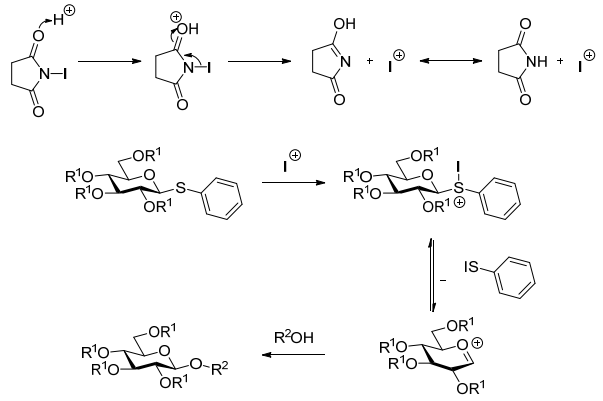


Figure 47: Thioglycoside activation with a NIS/TfOH promotor system.[309]

The regioselectivity of this glycosylation step is achieved by making use of the higher reactivity of primary hydroxyl groups over secondary ones and equatorial over axial ones[209], leading to a big

difference in the reactivities of the axial secondary 4- and the primary 6-position. Due to the participating character of the LacNAc 2-*N*-Troc group, the reaction proceeds with full stereoselectivity under formation of the β-1,6-*O*-glycosidic bond, as verified by gHMBC-NMR experiments (see appendix).

For later application in solid phase peptide synthesis, global acetyl group protection needed to be installed after coupling of the type-1 LacNAc disaccharide. A number of protection group manipulations were made. Acidic hydrolysis of the *O*-TBDMS group[339-340] and subsequent acetylation with acetic anhydride in pyridine afforded compound **33**. The Troc protecting group was then removed by reduction with zinc dust under acidic conditions[341] followed by generation of the corresponding acetamide **34**. Finally the *tert*-butyl ester was removed using trifluoroacetic acid in dichloromethane[342] with the addition of anisole as carbocation scavenger[343]. Starting from the T-antigen glycosylated amino acid **31** the extended type-1 core 2 glycosylated amino acid building block was synthesized in four steps with a total yield of 47 %.

7.1.8 Synthesis of the core 2 type-2 glycosylated amino acid

The extended type-2 core 2 glycosylated amino acid was assembled in a [2+2] building block glycosylation analogous to the above described synthesis for the type-1 variant (chapter 7.1.7). In this case the T-antigen acceptor **31** was glycosylated with the LacNAc type-2 donor **29** to form a β-1,6-*O*-glycosidic bond (***Figure 48***). The regio- and stereoselective considerations for this reaction correspond to the ones made for the type-1 core 2 assembly. However, the coupling proceeded much slower for the type-2 variant (8 – 10 h, further addition of donor and NIS) compared to the type-1 variant (1 h, no additional NIS necessary). The LacNAc type-1 and type-2 disaccharide donors only differ in the connectivity of the two pyranose subunits and it can be assumed that the difference in reactivity is associated with stereoelectronic effects related to the linkage connectivities or positions of the acetyl protecting group within the glucosamine residue. Substitution with *O*-glycosidic bonds have a negative impact on the reactivity of a donor[344], but this is also true for electron withdrawing acetyl groups which deactivate the release of the leaving group at C-1 as found for groups at C-2[211-212]. Later a disarming effect has also been found for C-3, 4 and 6 and it was rationalized, that the electron-withdrawing effect of acyl esters at C-3, -4 and -6 hamper the oxocarbenium stabilization by the O-5 lone pair electrons[345]. Introduction of TBDMS groups instead of acetyl groups would be an option to further activate the donor building blocks[346].

After LacNAc type-2 disaccharide coupling, a few protecting group manipulation steps followed. Acidic removal of a TBDMS protection group[339-340] and subsequent acetylation afforded compound **37**

followed by reduction of the Troc group[341] and acetylation gave the acetamide **38**. By acidic cleavage of the *tert*-butyl ester[342] with the addition of anisole[343], the extended type-2 core 2 glycosylated amino acid **39** was obtained. Starting from T-antigen **31** the synthesis of this building block proceeded in four steps with a total yield of 43 %.

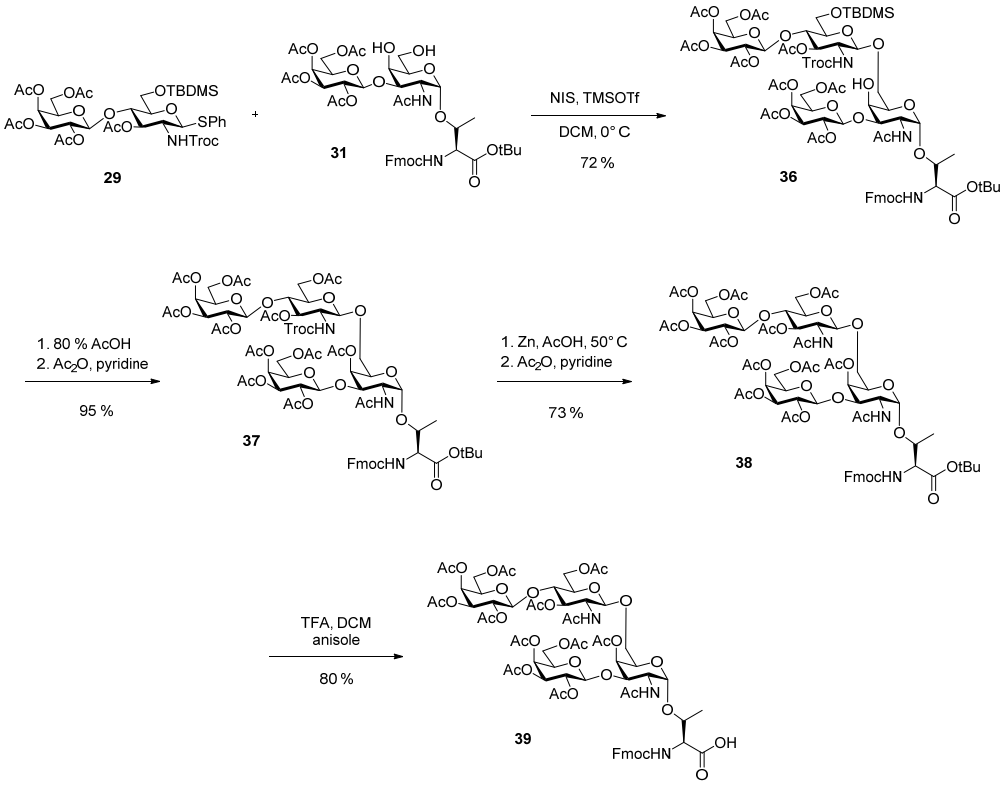


Figure 48: Synthesis of the extended type-2 core 2 glycosylated amino acid building block **39**.

7.1.9 Synthesis of the core 3 type-1 acceptor

The type-1 core 3 glycosylated amino acid acceptor is a precursor of the type-1 core 4 building block and is assembled by β-1,3-*O*-glycosylation of the T_N-antigen acceptor **12** with a LacNAc type-1 disaccharide donor **24**. Although the LacNAc donor has an *N*-Troc group in 2-position which has a β-directing effect, the reaction was found to proceed with low stereoselectivity[153, 323]. To address this problem, the T_N-antigen acceptor was modified and the acetamide was replaced by a less electron withdrawing azide group, which was expected to enhance acceptor reactivity. Additionally, the azide group is smaller and would contribute with less sterical hindrance. Unfortunately, the protecting group exchange resulted in almost identical R_f values for both, acceptor and donor in all tested solvent systems, making it difficult to monitor the reaction progress. Therefore this approach was not followed further and the reaction was done in accordance to the previous procedure developed in the group.

The thioglycoside was activated using an NIS/TfOH promotor system under kinetic conditions on ice leading to formation of the type-1 core 3 glycosylated amino acid **40** with a yield of 85 % and an α/β ratio of 2:3 (51 % β-anomer, ***Figure 49***).

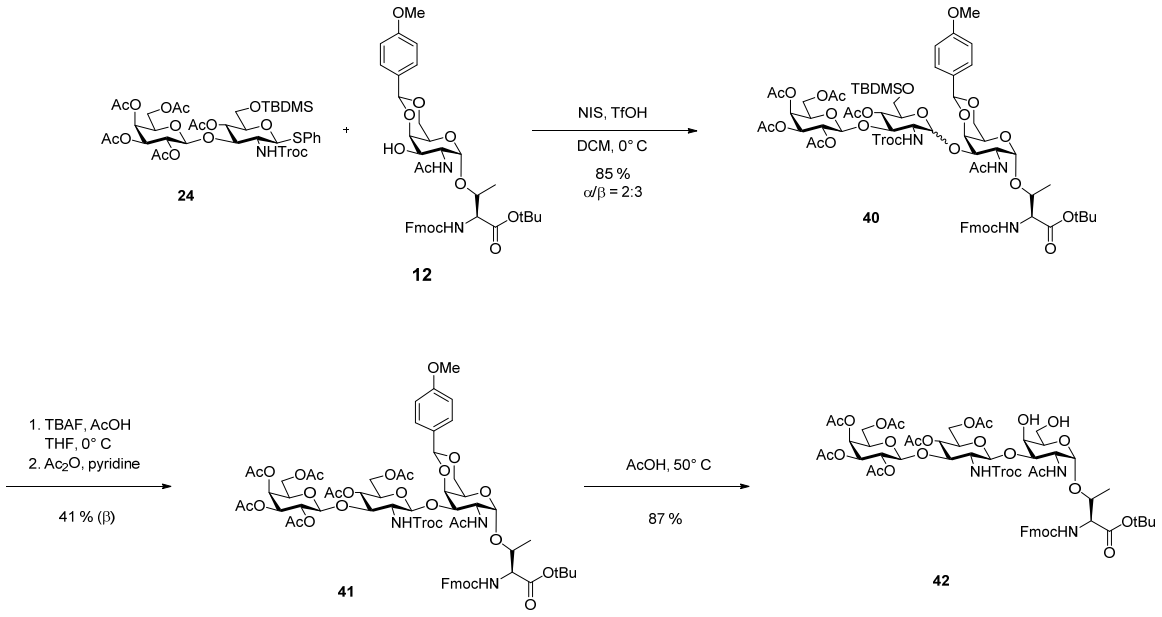


Figure 49: Synthesis of the type-1 core 3 acceptor.

In the following step the TBDMS was selectively removed and subsequently acetylated. For the removal of the silyl ether *tetra*-butylammonium fluoride (TBAF) and acetic acid in tetrahydrofuran was used[347]. Here, the acetic acid serves to moderate the strong basicity of the TBAF[348] to prevent side reactions with the base labile Fmoc amino acid protecting group. Finally, acidic removal of the *para*-methoxybenzylidene acetal generated the type-1 core 3 acceptor **42**. Starting from T_N-antigen **12** the trisaccharide acceptor **42** was synthesized in three steps with a total yield of 30 %.

7.1.10 Synthesis of the core 3 type-2 acceptor

The synthesis of the type-2 core 3 amino acid was performed analogous to the synthesis of the type-1 elongated core structure using a NIS/TfOH promotor system in dichloromethane[296]. Unlike in the type-1 case, the glycosylation proceeded with full stereoselectivity, leading to exclusive formation of the β-1,3-*O*-glycosidic bond. However, in addition to the desired trisaccharide **43**, formation of the byproduct **44** was observed (***Figure 51***). Mass spectrometry showed a mass difference of +108 u with respect to compound **43**, corresponding to an addition of a phenyl-sulfenyl group to the N-Troc amine. The phenyl-sulfenyl residue could be removed during the N-Troc reductive treatment with zinc to

generate the desired product in later steps of the synthesis. This further supported the conclusion, that the phenyl-sulfenyl leaving group was added to the N-Troc nitrogen during the glycosylation reaction, a plausible mechanism for the phenyl-sulfenyl addition is shown in ***Figure 50***. It suggests, that the endocyclic nitrogen in the oxazoline intermediate acts as a nucleophile on the phenyl-sulfenyl iodide intermediate in the course of the glycosylation reaction[296].

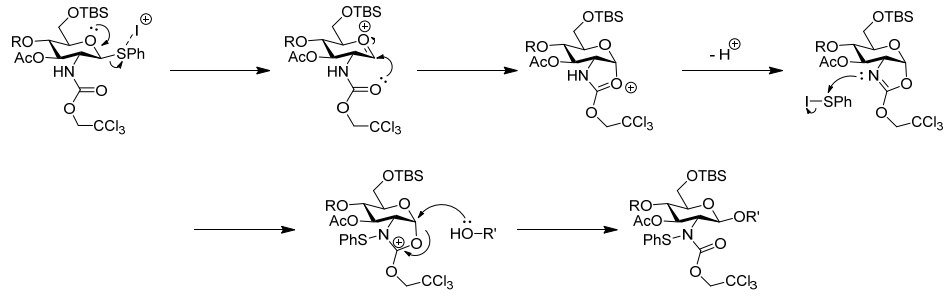


Figure 50: Possible mechanism of the phenyl sulfenyl adduct generation.[296]

Subsequent to the glycosylation reaction, a few protecting group manipulation steps were done to enable elongation on the GalNAc-6-position and formation of LacNAc core 4 structures. Initial attempts to selectively remove the highly acid labile *para*-methoxybenzylidene acetal in the presence of the silyl ether by using 80 % acetic acid on ice[349] were not efficient. Instead, the silyl ether was replaced by an acetyl group, using TBAF/AcOH[347-348] and subsequent acetylation wit Ac_2O/Pyr, before removing the acetal. Finally, acidic removal of the 4,6-*para*-methoxybenzylidenacetal afforded the diols **47** and **48** with a total yield of 32 % and 20 % respectively over three steps each starting from T_N-antigen **12**.

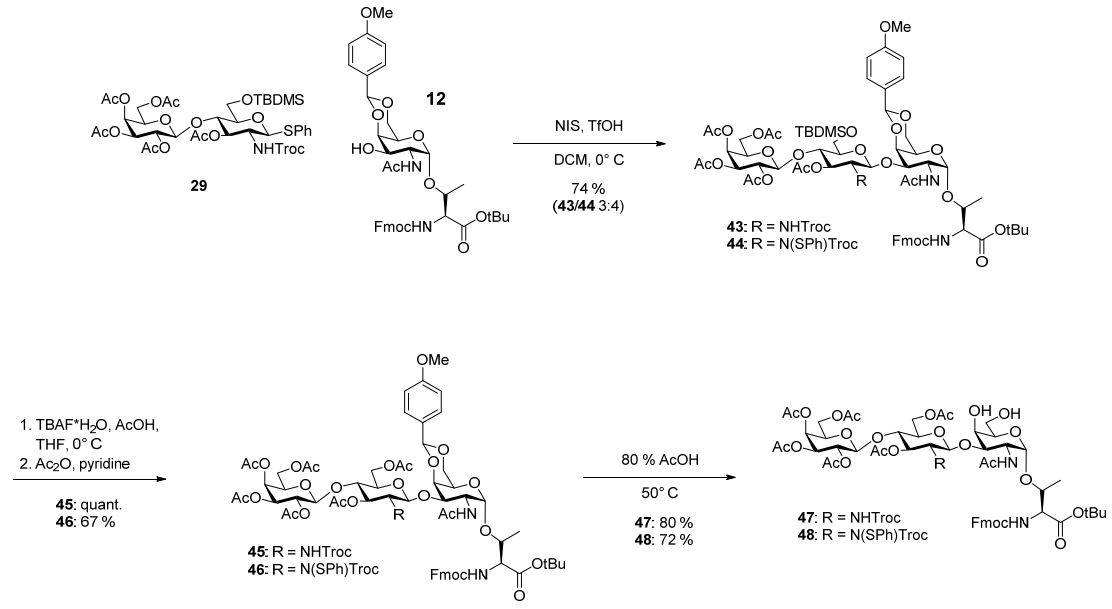


Figure 51: Synthesis of the type-2 core 3 acceptor.

7.1.11 Synthesis of the core 4 type-1 glycosylated amino acid

The type-1 core 4 glycosylated amino acid building block **49** was assembled in a [3+2] glycosylation of the type-1 core 3 glycosylated amino acid **42** with the LacNAc type-1 disaccharide donor **29** (***Figure 52***). The 1,6-*O*-glycosylation was carried out using conditions[307-309] analog to the above described [2+2] glycosylation to afford the type-1 core 2 glycosylated amino acid (***Figure 46***, p. 69). Despite the considerably higher reactivity of the primary hydroxyl group at the GalNAc-6-position over the secondary and axial 4 position, the coupling was not proceeding with full regioselectivity and formation of about 30% of the LacNAc-β-1,4-*O*-GalNAc adduct **49 A** was additionally observed. The desired type-1 core 4 glycosylated amino acid building block **49** was obtained in a yield of 59 % (1,6-product).

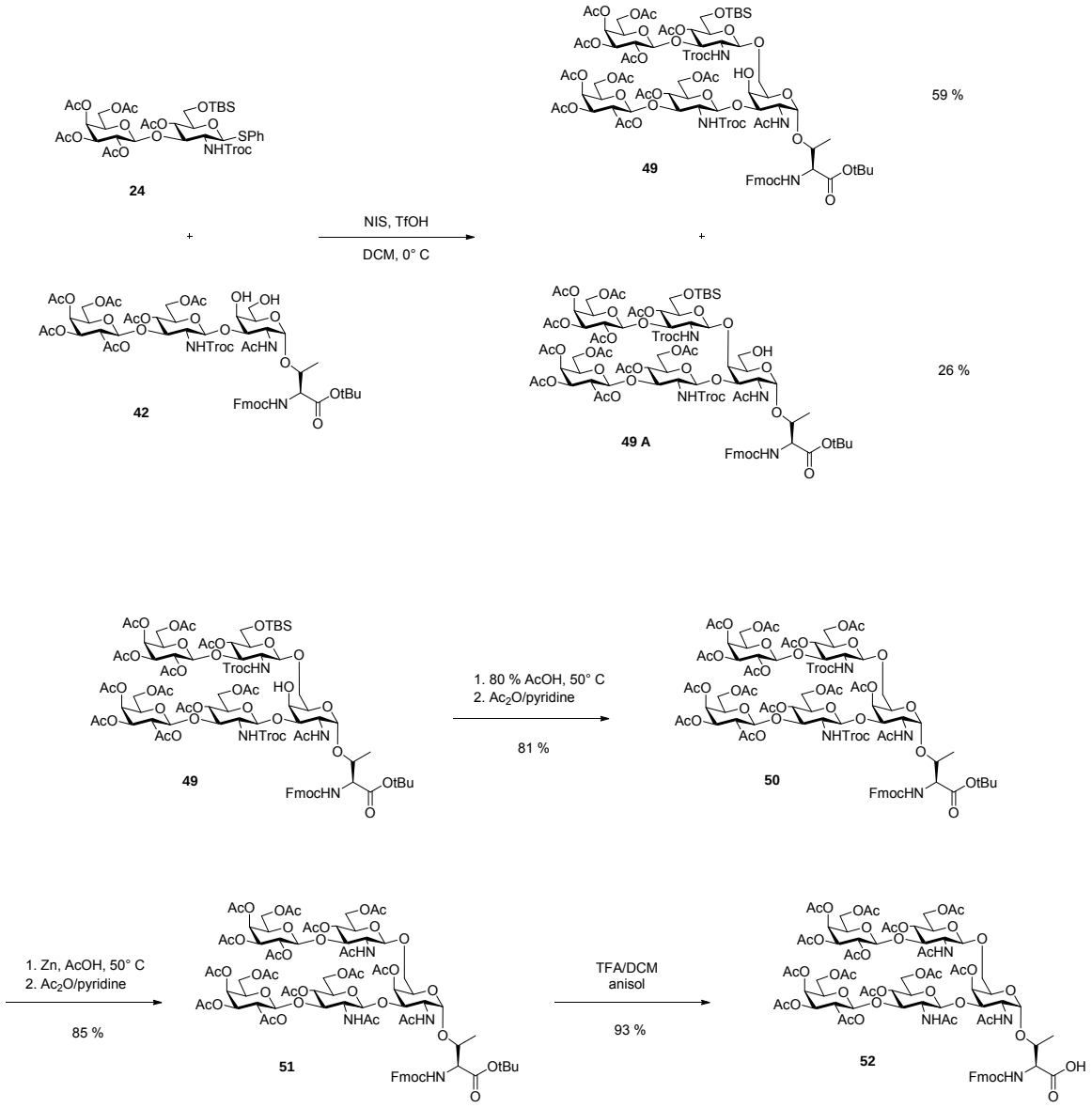


Figure 52: Synthesis of the type-1 core 4 glycosylated amino acid **52**.

To enable application of the building block in Fmoc-SPPS, the protecting groups had to be changed to global acetylation. The silyl ether was removed under acidic conditions, followed by subsequent acetylation to obtain compound **50**. Removal of the *N*-Troc groups was done by reductive elimination using zinc dust/acetic acid[341] with subsequent formation of the corresponding acetamide **51**. Finally, removal of the amino acid *tert*-butyl ester with trifluoroacetic acid in dichloromethane[342] under addition of anisole[343] afforded the desired type-1 core 4 amino acid **52**. Starting from compound **42** the type-1 core 4 glycosylated amino acid **49** was synthesized in four steps from the type-1 core 3 acceptor **42** in a yield of 38 %.

7.1.12 Synthesis of the core 4 type-2 glycosylated amino acid

The type-2 core 4 glycosylated amino acid building block was assembled by a [3+2] glycosylation of a LacNAc type-2 disaccharide donor **29** with either the type-2 core 3 acceptor **47** or its corresponding phenyl sulfenyl adduct **48** respectively (***Figure 53***). The glycosylation conditions were chosen analog to the formation of the core 2 type-1 glycosylated amino acid (chapter 7.1.7) and proceeded with full stereo- and regioselectivity under formation of the LacNAc-β-1,6-*O*-GalNAc products **53** and **54**. The stereochemical outcome of the reaction was controlled by the neighboring group participation of the *N*-Troc group in 2-position of the disaccharide donor **29** and the desired regioselectivity was obtained by making use of the difference in reactivity of the primary hydroxyl group in GalNAc-6-OH position over the GalNAc-4-OH.

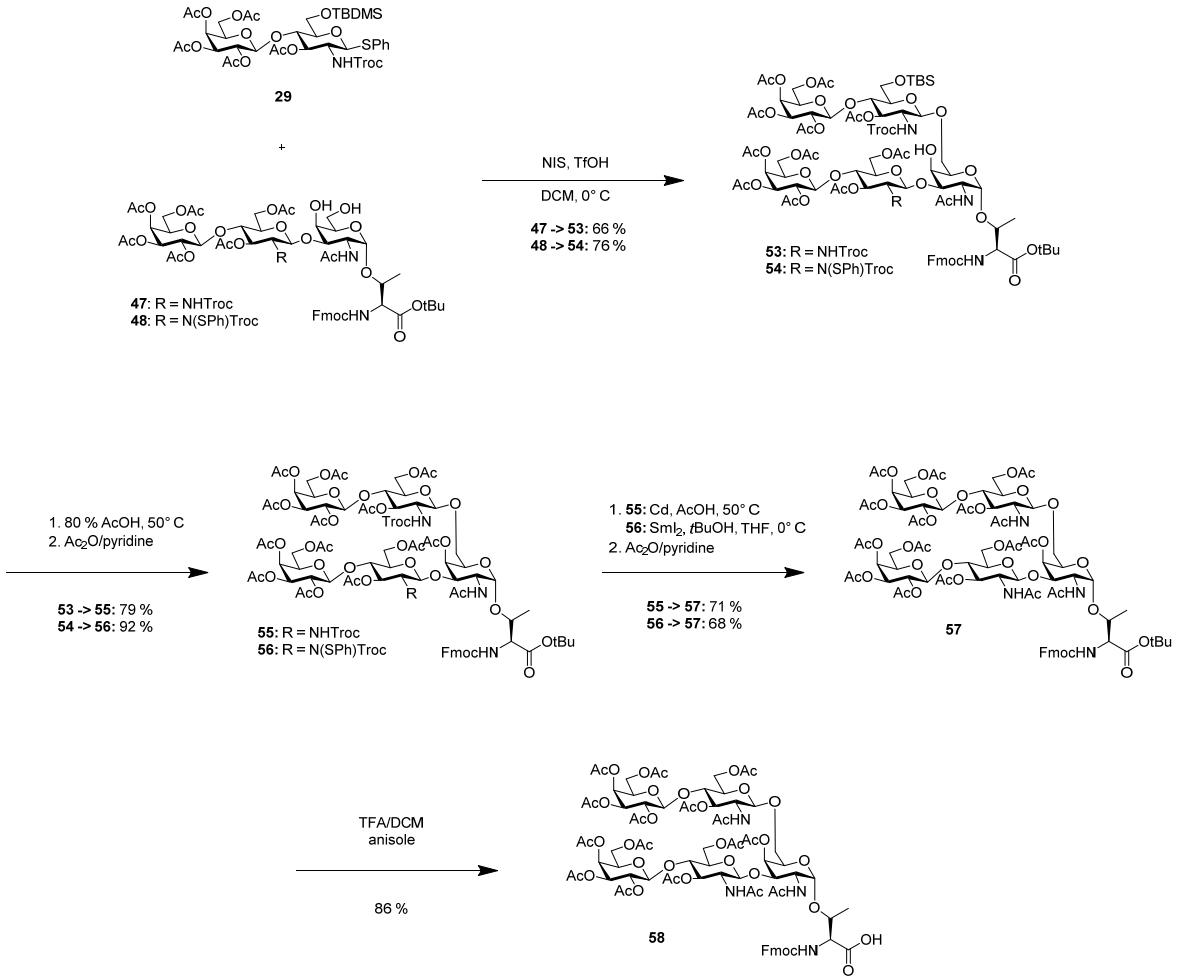


Figure 53: Synthesis of the type-2 core 4 glycosylated amino acid **58**.

The following protection group manipulations were performed on both compound **53** and **54**. Acidic removal of the silyl ether and subsequent acetylation gave compound **55** and **56**. Reduction of *N*-Troc

and N(SPh)Troc followed by acetylation was made to obtain the corresponding *N*-acetamides. Initial attempts using standard reductive zinc dust/acetic acid conditions[341] proceeded very slow and with low yields (4 days, 36 % for compound **55** and 6 days, 29 % for compound **56**). To improve the Troc deprotection, the zinc dust was replaced by cadmium dust, which has been reported to be more efficient compared to the Zn/AcOH conditions[350]. This approach improved the yield of the Troc reduction from compound **55** (NHTroc analog) from 36 % to 71 %. However, the reaction times were still long and undesirable for Troc removal of compound **56** (N(SPh)Troc analog). Samarium(II) iodide has been reported as reagent for removal of 2-chloroethyl carbamates on amines by *Philip Magnus*[351]. With $E_{0\ aq}Sm^{2+}/Sm^{3+}$ = -1.55 V, it exhibits one of the highest known reduction potentials for species known to be soluble in organic solvents[352-353] and surpasses the above-mentioned zinc ($E_{0\ aq}Zn/Zn^{2+}$ = -0.76 V) and cadmium ($E_{0\ aq}Cd/Cd^{2+}$ = -0.40 V) by far. By applying *in situ* generated samarium(II) iodide on the N(SPh)Troc compound **56**, the reduction efficiently proceeded over three hours with a yield of 68 % of the acetamide product. In the final step the *tert*-butyl ester was hydrolyzed using trifluoroacetic acid[342] with anisole[343] in dichloromethane to give the type-2 core 4 glycosylated amino acid in a total yield of 32 % (from the NHTroc coupling product) and 41 % (from the N(SPh)Troc coupling product) over four steps.

7.1.13 Synthesis of Fmoc-GalNAc-D_3-threonine

To support the MS structural work using model glycopeptides, a Fmoc-GalNAc-D_3-Thr was prepared and incorporated in Fmoc-SPPS. The Fmoc-GalNAc-D_3-Thr **60** could be synthesized starting from an intermediate of the T_N-antigen synthesis, GalN_3-Fmoc amino acid **9** (see chapter 7.1.2). After reduction of the azide using zinc dust in acetic acid[341], the corresponding deuterated acetamide **59** was generated by acetylation with *N*-acetoxy-D_3-succinimide. Removal of the *tert*-butyl ester was achieved by treatment with a trifluoroacetic acid/anisole/dichloromethane system. Starting from compound **9**, the deuterated T_N-antigen **60** was synthesized in two steps with a total yield of 32 % (***Figure 54***).[189]

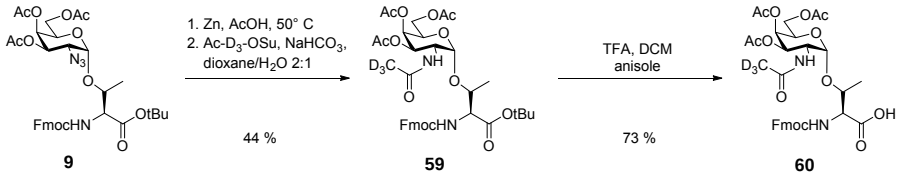


Figure 54: Synthesis of the Fmoc-GalNAc-D_3-Thr SPPS building block.

7.1.14 Generation of a MUC1 glycopeptide library

The synthesized mucin type glycosylated amino acids were used as Fmoc-SPPS building blocks to prepare a glycopeptide library based on the MUC1 tandem repeat sequence with the sequence PAHGVTSAPDTRPAPGSTA. A library was prepared with focus on the three Thr *O*-glycosylation sites of the MUC1 tandem repeat, glycosylated in different density and at different sites with the elongated core amino acids. All MUC1 glycopeptides were modified with an *N*-terminal triethyleneglycol (TEG) spacer, allowing immobilization of the peptides on a microarray surface for applications in protein binding studies (***Figure 55*** **a**). For optimization of the peptides synthesis, an unglycosylated peptide and the three possible monovalent T_N-antigen peptides were initially prepared (C-terminally unprotected T_N-antigen building block provided by Dr. Christian Pett). By application of the four different type-1/type-2 elongated core 2 and core 4 amino acids, mono-, di- and trivalent MUC1 glycopeptides were synthesized and combined with the previously synthesized library from Dr. Pett for further enzymatic modifications or directly applied in microarray binding studies[354] or MS structural analysis[189-190]. In order to acquire mechanistic insight to HexNAc oxonium ion fragmentation in the course of HCD experiments two more glycopeptides were synthesized containing the deuterium labelled T_N-antigen GalNAc-D_3. (***Figure 55*** **b**).

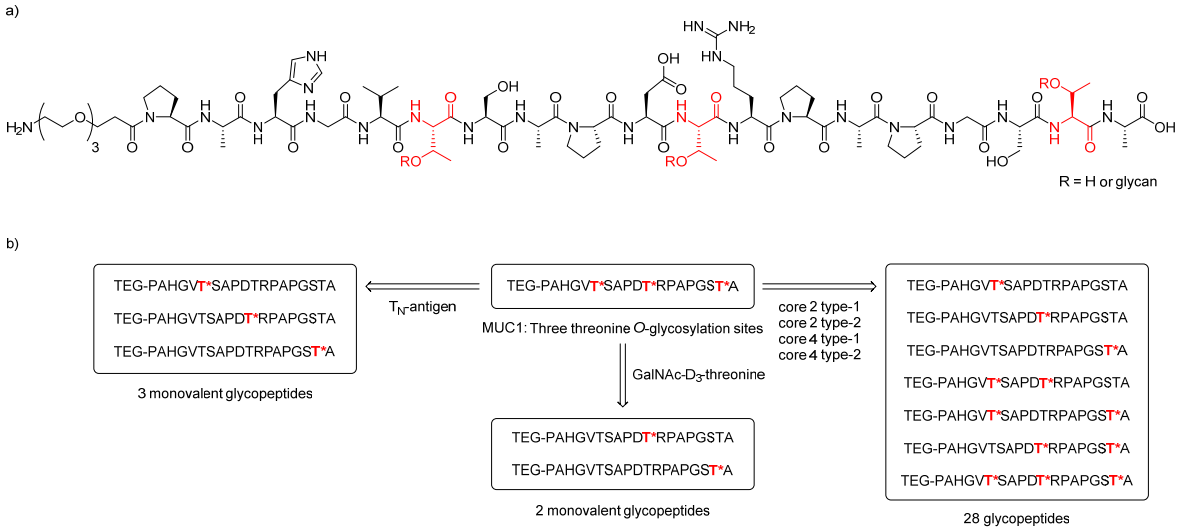


Figure 55: a) MUC1 glycopeptide construct with an *N*-terminal triethylene glycol spacer. b) Schematic overview of the MUC1 glycopeptide library that was prepared in this work.

7.1.14.1 Synthesis of mucin glycopeptides

All glycopeptides were synthesized on an Fmoc-alanine-TentaGel® trityl resin with a loading capacity of 0.15 to 0.17 mmol/g in a scale of 10 to 13 µmol per batch. To assure proper swelling of the resin, it was taken up in dichloromethane with subsequent shaking for 30 min. Dimethylformamide (DMF)

washing (5x 1 min) was performed prior to Fmoc deprotection of the preloaded amino acid in 20 % piperidine/DMF (2x 3 min, 1x 9 min). During coupling, the standard Fmoc-amino acids were applied in eight fold excess and activated using an HBTU/HOBt/DIPEA system. HBTU and HOBt were applied in slightly substochiometric amounts compared to the donor amino acid and DIPEA was used in two equivalents excess compared to the amino acid. Standard reaction time was 40 minutes with subsequent Fmoc removal (20 % piperidine, 2x 3 min, 1x 9 min). To compensate for possible sterical hindrance of the resin as well as from the glycosylated amino acids, a double coupling step was introduced after each glycosylated amino acid as well as for the first coupling to the resin-bound amino acid. The synthesized glycosylated amino acids are more valuable compounds and available only in limited quantities. For that reason these building blocks were used in only 1.5 – 2 fold excess with respect to the resin loading. Additionally, compared to normal Fmoc-amino acids, a more efficient coupling method was applied by using HATU/HOAT/DIPEA activation in a reduced volume of DMF and an increased reaction time (8h). After full peptide assembly, an *N*-terminal triethylene glycol (TEG) spacer was introduced with a 3-fold excess using the standard HBTU/HOBt/DIPEA coupling system and a reaction time of two hours. The on-resin peptide synthesis was completed by Fmoc removal from the *N*-terminal spacer using piperidine. Cleavage of the peptide from resin and simultaneous amino acid side chain deprotection was then done under acidic conditions using a TFA/TIPS/water system (15:1:1 v/v/v). The released glycopeptides were then desalted using a C18-cartridge followed by deacetylation to remove the glycan protection under *Zemplén* conditions, sodium methoxide/methanol, for the shorter saccharides and for the elongated core 4 glycans or using a sodium hydroxide/water/methanol system for deacetylation of elongated core 2 structures. The pH was kept between 10.0 and 11.0 for all reactions and the deacetylation process was monitored via HPLC. As an example, monitoring of the deacetylation of glycopeptide **P14** is shown in **Figure 56**.

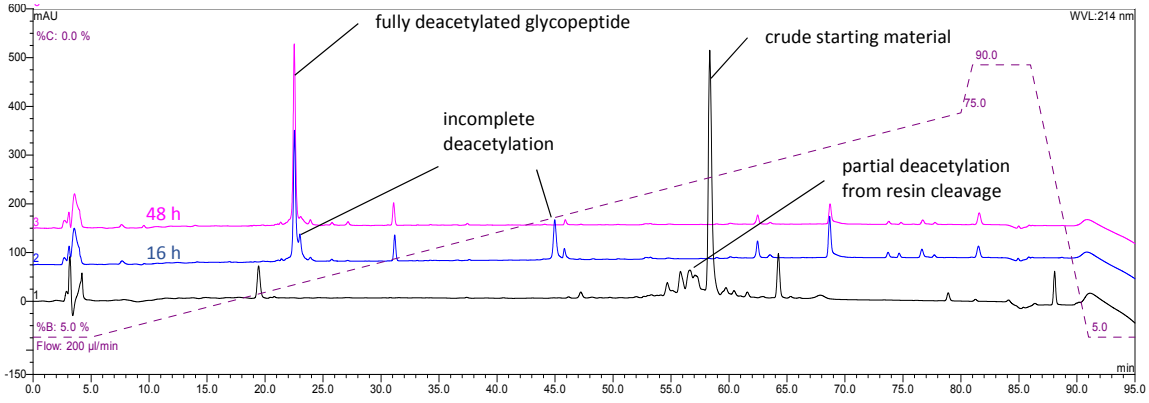


Figure 56: Glycopeptide deacetylation monitoring with reverse phase HPLC using the glycopeptide **P14** as an example.

Minor partial deacetylation is sometimes observed already before starting the deacetylation, which is caused by the acidic cleavage from the resin. After 16 hours the starting material is consumed completely, but the deacetylation process is still incomplete as indicated by peaks between the starting material and the product peak. These peaks disappeared after 48 hours in combination with a further growth of the product peak which indicates full conversion. After completion of the deacetylation the glycopeptide solutions were neutralized with acetic acid and subjected to purification by reversed phase HPLC. An overview over the synthesis conditions and yields for the different glycopeptides in the library is given in ***Table 1***.

Table 1: Synthesis conditions and yields for preparation of the mucin glycopeptide library.

glycan	ID	sequence	eqiv.	deacetylation		yield (%)
				pH	time (h)	
---	**P1**	TEG-PAHGVTSAPDTRPAPGSTA	---	---	---	40
T_N-antigen	**P2**	TEG-PAHGV**T***SAPDTRPAPGSTA	1.5	10.0	16	27
	P3	TEG-PAHGVTSAPD**T***RPAPGSTA	1.5	10.0	16	22
	P4	TEG-PAHGVTSAPDTRPAPGS**T***A	1.5	10.0	16	14
GalNAc-d_3	**P5**	TEG-PAHGVTSAPD**T***RPAPGSTA	1.5	10.0	16	51
	P6	TEG-PAHGVTSAPDTRPAPGS**T***A	1.5	10.0	16	52
core 2 type-1	**P7**	TEG-PAHGV**T***SAPDTRPAPGSTA	1.5	10.0 - 10.5	19	78
	P8	TEG-PAHGVTSAPD**T***RPAPGSTA	1.5	10.5	16	25
	P9	TEG-PAHGVTSAPDTRPAPGS**T***A	1.5	10.5	16	30
	P10	TEG-PAHGV**T***SAPD**T***RPAPGSTA	1.5	10.0 - 10.5	19	22
	P11	TEG-PAHGV**T***SAPDTRPAPGS**T***A	1.5	10.0 - 11.0	24	28
	P12	TEG-PAHGVTSAPD**T***RPAPGS **T***A	1.5	10.0 - 11.0	24	30
	P13	TEG-PAHGV**T***SAPD**T***RPAPGS**T***A	1.5	10.0 - 11.0	24	17
core 2 type-2	**P14**	TEG-PAHGV**T***SAPDTRPAPGSTA	1.5	10.5	48	38
	P15	TEG-PAHGVTSAPD**T***RPAPGSTA	1.5	10.5	48	39
	P16	TEG-PAHGVTSAPDTRPAPGS**T***A	1.5	10.5	144	35
	P17	TEG-PAHGV**T***SAPD**T***RPAPGSTA	1.5	11.0	48	28
	P18	TEG-PAHGV**T***SAPDTRPAPGS**T***A	1.5	11.0	48	32
	P19	TEG-PAHGVTSAPD**T***RPAPGS **T***A	1.5	11.0	48	15
	P20	TEG-PAHGV**T***SAPD**T***RPAPGS**T***A	1.5	11.0	48	14
core 4 type-1	**P21**	TEG-PAHGV**T***SAPDTRPAPGSTA	1.5	11.0	48	42
	P22	TEG-PAHGVTSAPD**T***RPAPGSTA	1.5	10.5	16	17
	P23	TEG-PAHGVTSAPDTRPAPGS**T***A	1.5	10.5	22	18
	P24	TEG-PAHGV**T***SAPD**T***RPAPGSTA	2.0	10.5	48	25
	P25	TEG-PAHGV**T***SAPDTRPAPGS**T***A	2.0	10.5	48	30
	P26	TEG-PAHGVTSAPD**T***RPAPGS**T***A	2.0	10.0	48	26
	P27	TEG-PAHGV**T***SAPD**T***RPAPGS**T***A	2.0	10.0	48	19
core 4 type-2	**P28**	TEG-PAHGV**T***SAPDTRPAPGSTA	1.5	11.0	48	51
	P29	TEG-PAHGVTSAPD**T***RPAPGSTA	1.5	10.5	16	8
	P30	TEG-PAHGVTSAPDTRPAPGS**T***A	1.5	10.5	16	12
	P31	TEG-PAHGV**T***SAPD**T***RPAPGSTA	2.0	11.0	72	28
	P32	TEG-PAHGV**T***SAPDTRPAPGS**T***A	2.0	11.0	72	29
	P33	TEG-PAHGVTSAPD**T***RPAPGS**T***A	2.0	11.0	48	30
	P34	TEG-PAHGV**T***SAPD**T***RPAPGS**T***A	2.0	11.0	48	15

7.1.15 Ligand specificity of adenovirus lectins

Adenoviruses represent a class of pathogens, which are considered as risk factors in the course of COPD[355]. Recently, the adenovirus lectins AD37 and AD52 have been identified to interact with glycan structures on host cell surfaces, but the lectin fine-specificities are still unclear.

The fiber knob protein AD37 from adenovirus type 37, which is also responsible for epidemic keratoconjuctivitis in eyes has been found to bind sialylated cell-surface molecules of unknown structure[198]. Strong binding could be shown for a branched hexasaccharide with two terminal sialic acids as present for example in the brain ganglioside GD1a[197]. However, while GD1a is highly abundant in brain tissue, the ocular mucus is rather characterized by a high abundance of sialylated mucins. Therefore it is unlikely, that GD1a is a native ligand of AD37 in the course of eye infection and the question of AD37 binding specificity remains unanswered. As the structure of GD1a (***Figure 57***) resembles the ones from common mucin core structures, it was considered to be likely, that sialylated mucin core structures might serve as AD37 ligands. With the help of structurally defined sialoglycopeptide microarrays from our research group, this theory should be verified.

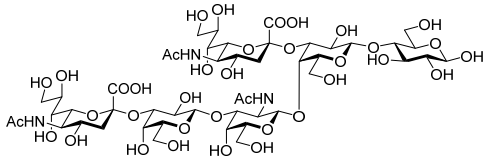


Figure 57: Structure of the GD1a glycan.

Similarly human adenovirus type 52 (AD52), which is also causing gastroenteritis[356], binds to its host cells via sialic acid terminated *O*-glycoproteins. Here, the virus is equipped with a short and a long fiber protein[357], of which the short one is responsible for sialic acid recognition. This sialic acid recognition has been found to play the major role over the CAR (coxsackievirus and adenovirus receptor protein) recognition of the long fiber. Earlier microarray screenings showed recognition of α2,3-sialylated probes and no recognition of α2,6-sialylated ones[358]. Further, a specificity towards LacNAc type-2 over LacNAc type-1 was reported by the authors. However, the ligand-positive probes of that studies were mainly glycolipids, while the corresponding glycan sequences are common to glycoproteins. Therefore, also this lectin was added to the sialoglycopeptide microarray incubation experiment. This enables a validation of the reported results with more natural ligand probes.

The lectins were incubated on glycan microarray **MA3** containing sialylated MUC1 glycopeptides, which was spotted by Dr. *Christian Pett* under utilization of glycopeptides from Prof. Dr. *Ulrika Westerlind*, Dr. *Cai Hui*, himself and glycopeptides from this work. The full list of glycopeptides is shown in ***Table 7***, p. 252. The lectins were expressed and provided with a penta-His tag by Prof. Dr.

Niklas Arnberg (Department of Clinical Microbiology, University of Umeå, Sweden). Lectin staining was then done using *Penta-His™ Alexa Fluor488® Conjugate* (*Quiagen*, Venlo, Netherlands).

However, no fluorescence signal could be detected. This was not an expected result, as the microarray slide contained several α2,3-sialylated glycans, which should be recognized by the reported α2,3-sialoglycan affinity of AD 52. In addition and as mentioned above, some glycans show high resemblance to the reported AD37 ligand GD1a. This in combination with the generally high presented sialoglycan diversity, the absence of binding detection can be considered highly unlikely. As possible causes of error the lectins, the antibodies and the antibody concentration were identified.

In order to isolate the cause of error a new microarray experiment was designed, based on the same immobilized MUC1 sialoglycopeptide library. In order to examine the functionality of the anti-penta-His antibody, the His-tagged AD37 lectin was switched for recombinant human His-tagged galectin-3 expressed from *E. coli* in two replicates. As galectin-3 is generally known to be β-galactoside-specific[359] and the lectin has been successfully used to bind parts of the present sialoglycopeptide library[153], at least partial binding to the whole sialoglycopeptide library galectin-3 can be expected. However, after staining with anti-penta-His antibody no signals could be detected in the fluorescence readout. Instead, staining with an anti-galectin-3 antibody resulted in staining on almost all presented glycopeptide spots. In order to test, whether the antibody concentration was too low in the initial adenovirus lectin/anti-penta-His antibody experiments, analog incubation experiments with higher antibody concentrations have been done, but also did not lead to signal detection. These results suggest that the utilized anti-penta-His antibody is likely not very sensitive for readout of AD37 and AD52 lectin binding. However, later experiments by Dr. Yu Jin, in which the His-tagged AD37 and AD52 lectins were stained with the anti-penta-His antibody *NTA-Atto 488* were also not successful in the fluorescence readout. The GD1a oligosaccharide was further included as positive control for AD37 fiber knob binding, which implies with the fluorescence readout or the lectins themselves in these experiments.

7.1.16 Binding specificities of human galectin-3

Integrated in a large glycopeptide microarray, some of the synthesized glycopeptides from this work were used in the group for specificity mapping of recombinant human galectin-3[153]. Extracellular galectin-3 is assumed to have an apoptotic function via interaction with carbohydrates of cell surface protein receptors[360]. As several kinds of tumor cells overexpress galectin-3, it is believed that the tumor benefits from the apoptotic effect of extracellular galectin-3 by suppression of for example T-lymphocytes[147]. The tumor cells themselves are protected by aberrant cell surface α2,6-sialylation,

blocking the interaction of galectin-3 with cell surface receptors and thus blocking induction of apoptosis. Further, galectin-3 has been found to interact with T-antigen structures on the aberrantly overexpressed MUC1 glycopeptides on cancer cell surfaces, which is thought to be involved in tumor metastasis[149]. Accordingly, the glycopeptide microarray was designed to contain typical tumor-associated carbohydrate motifs consisting of differently elongated *O*-glycan core structures with different sialic acid termination patterns in mono- and multivalent presentation on different glycosylation sites. As a result of incubations with galectin-3 it was found, that the lectin exhibits weak affinity towards the tumor-associated T- and T_N-antigen structures compared to common LacNAc elongated mucin-type O-glycans. No recognition difference was observed between LacNAc type-1 and type-2. Further, terminal α2,3-sialylation lead to increased recognition by the lectin, which is interesting as in earlier studies this modification has been found to have no impact or even reduce the affinity of galectin-3 towards the ligand. However, the results from our work can be regarded with high confidence, since all ligand candidates exhibit defined structures due to the synthetic approach. In contrast, terminal α2,6-sialylated glycopeptides showed decreased galectin-3 binding, which is in accordance with results from earlier studies. Glycan structures with additional internal α2,6-sialylation on the GalNAc residue of T-antigen or T_N substructure completely blocked binding to galectin-3. These results reflect the ability of galectin-3 to recognize internal galactose subunits and underline the impact of α2,6-sialylation on tumor cells in cell survival and tumor progression. An overview of the different glycan modifications on the galectin-3 binding affinity is given in ***Figure 58***.[153]

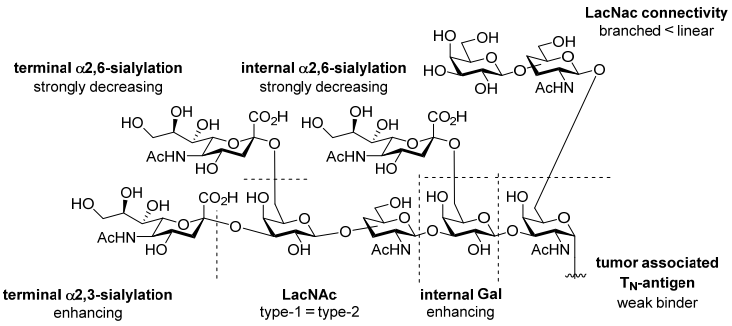


Figure 58: Impact of glycan modification on galectin-3 binding.

7.1.17 Evaluation of antitumor vaccines towards antibody specificity

Typically, adenocarcinoma cells show extensive overexpression of MUC1 glycopeptides, which in turn exhibit tumor-characteristic, premature sialylated short saccharides and an exposed tandem repeat peptide backbone[70]. The sialyl-T_N- and T_N-antigen structures as well as the sialyl-T- and T-antigen structures are dominant over the more complex branched and elongated core 2,3 and 4 structures.

The exposure of these characteristic glycopeptide antigen structures represents a valuable target for potential anti-tumor vaccines. For application in immunotherapy, vaccines have to elicit a strong and target-specific humoral immune response, overcome the natural immune tolerance towards mucin-type glycosylation and lead to immunological memory. Several successful methodologies have been developed to overcome the immune tolerance in order to elicit a strong immune response (see chapter 3.2.2.2), but evaluation of the raised antibodies towards quality and specificity is still challenging due to missing glycopeptide probes for screening experiments. Generation of an immunological memory with antibody specificity towards tumor cells without cross-reactivity to normal healthy cells is critical to avoid the risk of development of an auto immune disease upon vaccination. In order to perform a quality and specificity evaluation on polyclonal antibodies raised from MUC1-vaccine constructs, glycopeptides from this work have been incorporated into a large glycopeptide microarray library and used in binding studies[354]. Three vaccine candidates were selected for antibody induction, based on the T_N-antigen modified MUC1 20mer tandem repeat HGVTSAPDT(T_N)RPAPGS(T_N)T(T_N)APPA coupled to either a P30 T-cell epitope peptide **1**, a P30 T-cell epitope/Pam_3CSK_4 mitogen **2** and c) tetanus toxoid as carrier protein **3** (***Figure 59***).

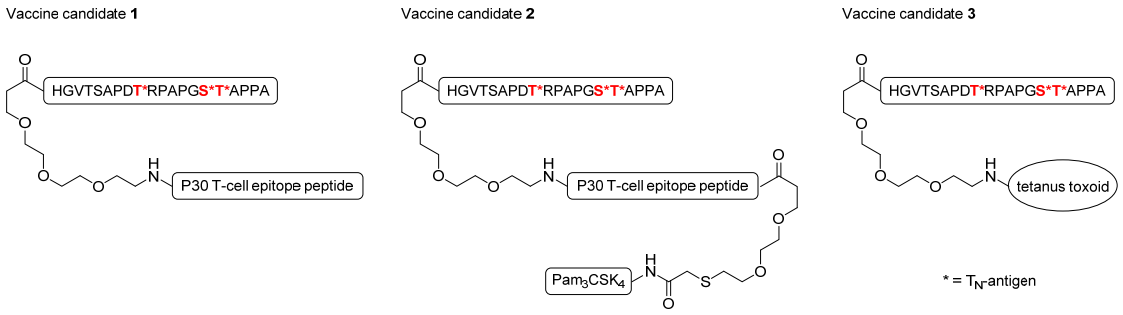


Figure 59: Vaccine candidates for antibody evaluation experiments.

The three vaccine candidates were administered in mice. Determination of the obtained antibody titers using *enzyme-linked immunosorbent assay* (ELISA) revealed a very strong immune response after induction with the T.Tox vaccine construct compared to the di- and tripartite vaccine constructs **1** and **2**. The fine-specificity of the induced antibodies was determined by incubation of the obtained sera on glycopeptide microarray slides. The microarray library consisted of over 100 different MUC1 glycopeptides, which were modified mono- and multivalently by β-GlcNAc, T_N-, ST- and T-antigen as well as the more complex mucin type core structures (LacNAc-elongated core 1, core 2 and core 3). The antisera, which were induced by the P30-based vaccine conjugates **1** and **2** showed a mixed affinity towards the immunodominant glycosylated epitopes PDT*R and GS*T*A, which were also glycosylated in the B-cell epitope of the vaccines. In general, various glycoforms of the PDT*R domain were recognized, but linear glycan structures (core 1, core 3) were dominant over branched ones (core

2) and a complete drop in recognition was often observed upon multivalent core 2 presentation on the MUC1 peptide tandem repeat. (***Figure 60***). Recognition of the GSTA domain was restricted to divalent T_N-antigen presentation. This strong selectivity might be the result of the strong rigidity of the doubly glycosylated domain, which is presented identically in the B-cell epitope. In contrast to the immune response induced from the fully synthetic vaccine conjugates 1 and 2, the antisera induced by the tetanus toxoid-based vaccine 3 showed a broader epitope recognition. The reason for this result might be, that the tetanus toxoid carrier protein exhibits a diverse set of T-cell epitopes, which leads to a strong activation of many different T-helper cells and ultimately to a strong polyclonal immune response. Core 2 glycosylated peptides were recognized weaker, analog to the behavior of the vaccine 1 and 2 antisera but less pronounced. Overall it can be concluded that tumor selective immune responses were generated with antibodies mainly directed to glycosylated GSTA and PDTR epitopes of different nature. Even if cross-reactivity to core 2 structures to some extend was observed from antibodies directed to the PDTR epitope, a clear preference of linear core structures over branched was found upon multivalent presentation.

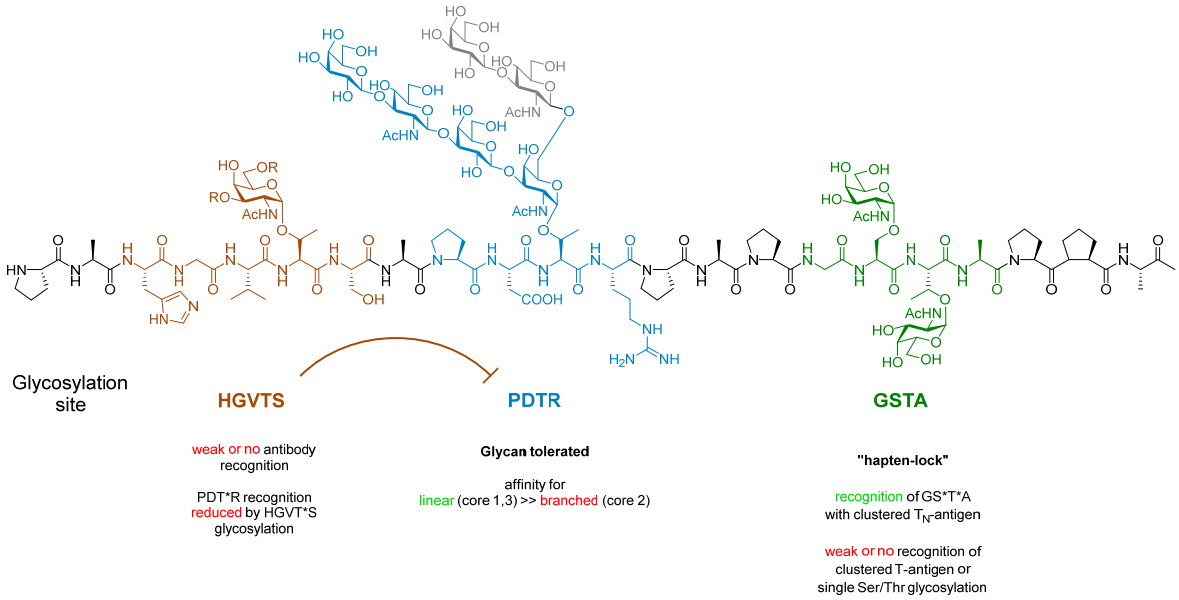


Figure 60: Contributions to antibody recognition by glycosylation-site-specific binding epitopes.

7.1.18 Oxonium ion-based HexNAc identification

Low mass diagnostic ions, oxonium ion patterns, are generated in the course of HCD fragmentation of oligosaccharides or glycopeptides and are highly valuable for data analysis[361]. The assignment of saccharide regio- and stereoisomers according to their characteristic oxonium ion profiles was studied (see chapter 7.1.18) within a collaboration of our group (*Synthetic Biomolecules*, ISAS, Dortmund,

Germany) and Dr. *Jonas Nilsson* (*Department of Clinical Chemistry and Transfusion Medicine*, *Institute of Biomedicine*, Gothenburg, Sweden)[190]. Different synthetic HexNAc containing mucin glycopeptides which were partly prepared in this work were employed in the studies. Analysis of decomposition pathways of the $[HexNAc]^+$ precursor oxonium ion was supported by preparation and fragmentation of the deuterium labelled GalNAc-D_3 glycopeptides **P5** and **P6** from this work[189]. Fragmentation of GalNAc oxonium ion m/z 204 resulted in the detection of m/z 186, 168, 144, 138 and 126 (***Figure 61***).

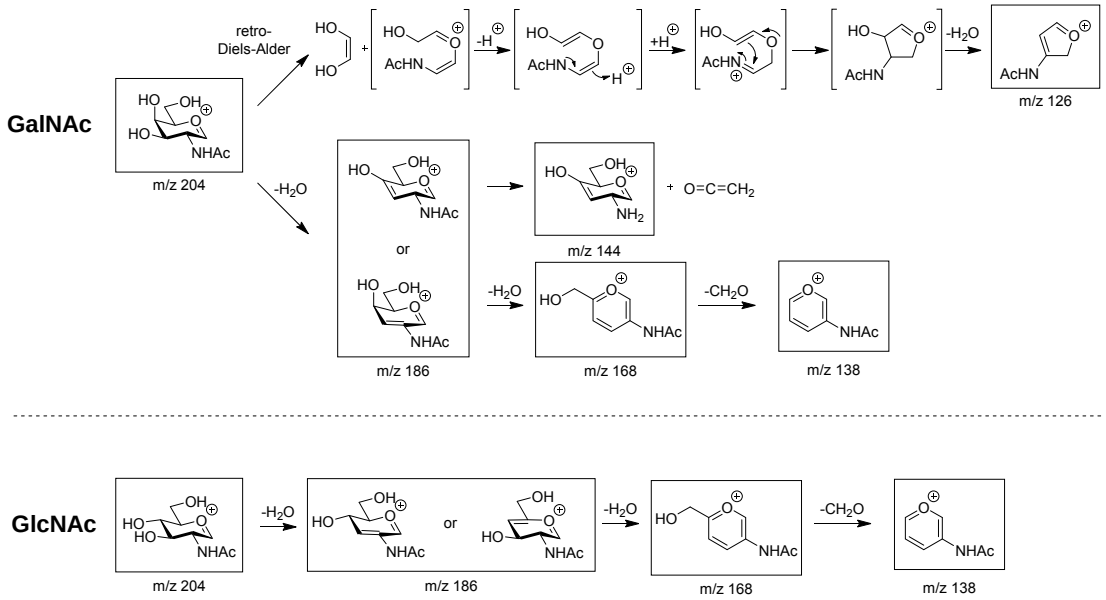


Figure 61: Decomposition pathways of GalNAc and GlcNAc oxonium ions by HCD fragmentation.

It was hypothized, that the $[GalNAc]^+$ ion m/z 204 fragments into isomeric structures with m/z 186 under the elimination of one H_2O. The isomer in which the generated π-bond is located between C3 and C4 undergoes further deacetylation under formation of the free amine with m/z 144. According to HCD experiments on the deuterated glycopeptide probes the fragmentation of m/z 189 (analog to undeuterated 186) generated an ion signal at m/z 145 (analog to undeuterated 144), which supported the proposed mechanism. Further loss of a H_2O from m/z 144 might be responsible for formation of the m/z 126 ion. However, CID-MS^4 on m/z 144 did not generate m/z 126 ions and also CID-MS4 on /z 189 only afforded minor m/z 127 signals (***Figure 62*** **A**).

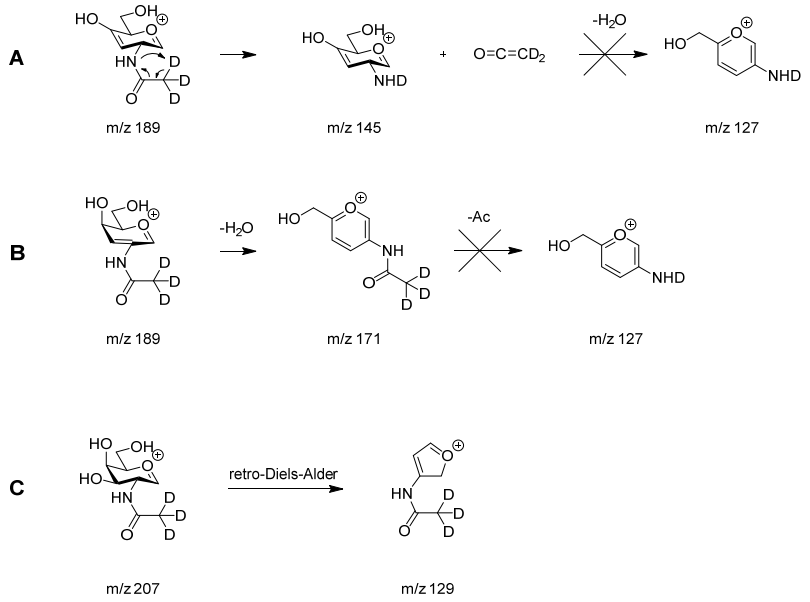


Figure 62: Results on HCD-MS/MS fragmentation on deuterium labelled **P5** and **P6**: **A**) Deacetylation of deuterium labelled m/z 189 under retention of on methyl deuterium. Further fragmentation to generate m/z 127 was not detected. **B**) Also no decomposition of m/z 171 to m/z 127 suggests, that m/z 126 does not originate from m/z 186 decomposition. **C**) Fragmentation of m/z 207 afforded detection of m/z 129, suggesting generation of m/z 126 via retro-Diels-Alder decomposition.

The other m/z 186 isomer in which the new π-bond is conjugated with the oxonium π-bond aromatizes to m/z 168 and further decomposes to m/z 138. Alternatively, deacetylation of m/z 168 might lead to formation of an alternative 126, however fragmentation of deuterium labeled glycopeptides only showed minor formation of m/z 127 (***Figure 62*** **B**). Instead and as expected the major formation of m/z 126 was obtained by carbohydrate ring cleavage of m/z 204 via a retro-Diels-Alder reaction. Since this reaction does not include deacetylation, CID-MS^3 on m/z 207 (deuterium-labelled analog of 204) resulted in the formation of m/z 129 (analog of 126, ***Figure 62*** **C**). It was concluded, that the HCD-induced fragmentation of the $[GalNAc]^+$ ion at m/z 204 resulted in decomposition along two pathways: The stepwise aromatization, leading to m/z 186, 168, 144 and 138 and on the other hand the retro-Diels-Alder ring cleavage under generation of m/z 126.

In the $[GlcNAc]^+$ fragmentation pathway the retro-Diels-Alder ring-cleavage to generate m/z 126 and amine deacetylation to generate m/z 144 is not pronounced and instead loss of water and aromatization was observed. As a consequence the GlcNAc diagnostic ions showed high abundance of m/z 168 and m/z 138. In order to discriminate GalNAc and GlcNAc oxonium ion patterns, we introduced the GlcNAc/GalNAc intensity ratio:

$$\frac{GlcNAc}{GalNAc} = \frac{m/z\ 138 + m/z\ 168}{m/z\ 126 + m/z\ 144}$$

with 0.2 ≤ GlcNAc/GalNAc ≤ 1 for GalNAc and 2 ≤ GlcNAc/GalNAc ≤ 50 for GlcNAc. With the help of this ratio, the isomer discrimination could be implemented into data analysis software, which would allow automated large scale analysis.

We also investigated the fragmentation patterns of core 3 modified glycopeptides[190]. If subjected to HCD, the oxonium ion profiles of type-1 core 3 (Gal-β-1,3-GlcNAc-β-1,3-GalNAc-*O*-) substituted glycopeptides were similar to those obtained GlcNAcβ1-*O*-substituted glycopeptides, suggesting that the terminal Gal in the GlcNAc-3 position did not affect the resulting oxonium pattern. In contrast, fragmentation of type-2 core 3 (Gal-β-1,3-GlcNAc-β-1,3-GalNAc-*O*-) substituted glycopeptides resulted in dramatically lower m/z = 204 values. Thus it is possible to discriminate the type-2 LacNAc glycopeptide from the type-1 LacNAc and GlcNAc-β-1-O substituted glycopeptides. Alternative CID-MS^3 fragmentation on the $[Gal\text{-}\beta\text{-}1,3\text{-}GlcNAc]^+$ and the $[Gal\text{-}\beta\text{-}1,4\text{-}GlcNAc]^+$ ions originating from CID-MS^2 precursor ions resulted in the same oxonium ion pattern as obtained from the HCD fragmentation experiments on the full glycopeptides. Thus it can be concluded, that the internal GalNAc-α-1-O-structure in core 3 glycans does not affect the resulting oxonium ion patterns. Further the results suggest, that the fragmentation behavior depends on the glycan structure but not whether HCD or CID is applied.

Summarized, the application of this methodology offers the possibility to resolve glycan structures and to identify individual terminal HexNAc subunits as well as the discrimination of type-1 and type-2 LacNAc subunits without the need of wet-chemical modifications of the glycan or glycoprotein. In addition, it bears the HCD-inherent feature to discriminate glycan and glycopeptide fragments from pure peptide fragments on the basis of oxonium fragment ion generation. This can be used for product-triggered approaches, where subsequent peptide fragmentation is only done, if characteristic oxonium ions are detected in the HCD MS/MS^{362}. This presorting avoids the analysis of non-glycosylated peptides and reduces the required computational effort in the software-aided fragment interpretation and generally enhances the reliability of the obtained results.

7.2 Project 2: Chemical tools to explore HexNAc-Tyr *O*-glycosylation

7.2.1 Overview of HexNAc-Tyr building block synthesis

In order to generate chemistry based tools to enable analytical and functional studies of HexNAc-Tyr *O*-glycosylation, a glycopeptide library modified with GalNAc- and GlcNAc-Tyr, Thr and Ser amino acids were synthesized by Fmoc-SPPS. Glycosylated Fmoc-Tyr amino acids were prepared by utilization of a common Tyr amino acid acceptor building block and coupled with $GlcN_3$-, GlcNTroc-trichloroacetimidate or $GalN_3Br$ glycosyl donors (***Figure 63***). The α-GalNAc-Tyr building block **64** was prepared in analogy to the previously described T_N-antigen Ser/Thr synthesis[203, 301, 363]. Trichloroacetimidate donors[204-205, 327] were used to generate the α- or β-GlcNAc-Tyr building blocks **69** and **74** and the configuration of the obtained glycosidic linkages was controlled by either installing a participating *N*-Troc at C-2 to promote β-glycosylation or by installing a non-participating azide to promote α-glycosylation. All glycans were protected globally with acetyl esters to achieve compatibility with Fmoc-SPPS. The HexNAc-Ser and –Thr Fmoc amino acids were prepared in previous work (provided by Dr. Yu, Dr. Pett and Prof. Dr. Westerlind).

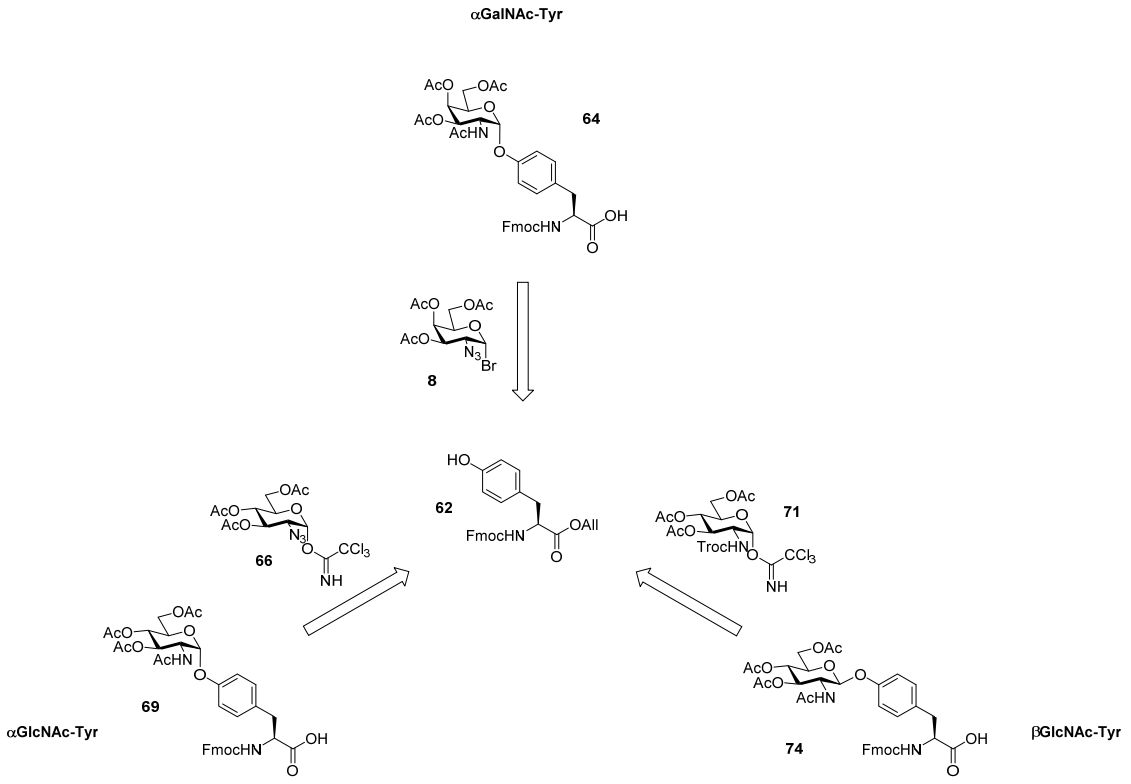


Figure 63: Retrosynthetical view on the glycosylated tyrosine synthesis plan.

The *L*-Tyr acceptor building block was prepared by Fmoc-protection of the N-terminus and allyl ester protection at the C-terminal carboxylic acid (***Figure 64***)[364-365].

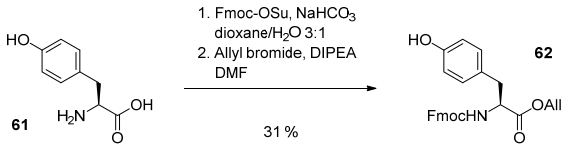


Figure 64: Synthesis of the tyrosine acceptor.

7.2.2 Synthesis of the α-GalNAc-Tyr building block

The synthesis of the α-GalNAc-Tyr Fmoc amino acid **64** were recently described[366] by applying the *Schmidt* method[205, 327]using a peracetylated galactopyranosyl trichloroacetimidate donor and a tyrosine acceptor. In this work the glycosylated amino acid was synthesized applying *Koenigs-Knorr* conditions[203] according to the previously described T_N-antigen Ser/Thr synthesis (compare chapter 7.1.2). This conveniently enabled the use of the acetobromogalactose donor **8** that additionally was applied as building block in the synthesis of mucin core glycosylated amino acids. Mechanistic and stereochemical considerations are equivalent to those discussed in chapter 7.1.2. After coupling of the glycoside donor, the allyl protecting group of the Tyr amino acid was removed by deallylation via the *Tsuji-Trost* reaction using catalytic amounts of tetrakis(triphenylphosphine)palladium(0) with phenylsilane in tetrahydrofuran[367-368] (***Figure 65***).

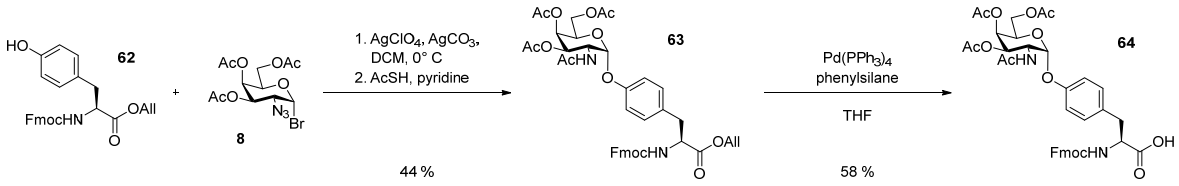


Figure 65: Synthesis of the α-GalNAc-Tyr building block.

The catalytic cycle of the deallylation is depicted in ***Figure 66*** and begins with complexation of the olefin by palladium(0), followed by ionization of the allylic leaving group. This leads to the formation of the free carboxylate anion and a π-allylpalladium(II) complex intermediate. Here, phenyltrihydrosilane serves as a scavenger by donating one hydride for nucleophilic addition to the π-palladium complex to regenerate the corresponding olefin. By decomplexation, the transition metal catalyst is regenerated and available for the next cycle. While the hydride is consumed for catalyst regeneration, the silyl residue of the scavenger adds to the carboxylate. The so formed silyl derivative is then hydrolyzed under formation of the desired carboxylic acid and a di(phenylsilyl) alkoxide.[367, 369]

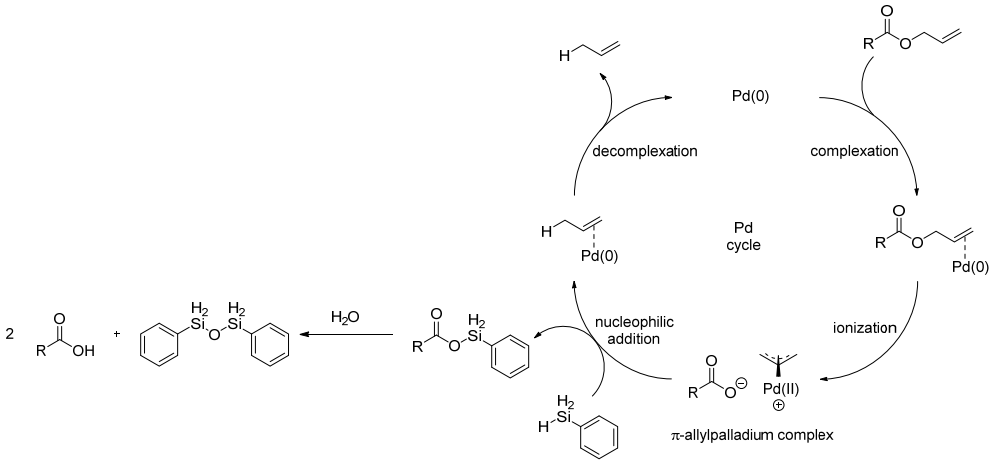


Figure 66: Catalytic cycle of the Tsuji-Trost deallylation.

Starting from Fmoc-tyrosine-*O*-allyl ester **62**, the α-GalNAc-Tyr building block was synthesized in two steps with a total yield of 26 %.

7.2.3 Synthesis of the α-GlcNAc-Tyr building block

The synthesis of the α-*O*-GlcNAc-Tyr building block **69** was partly carried out by Dennis Rosinski (***Figure 67***). As glycosyl donor the glycosyl trichloroacetimidate **66** was selected with the amine in 2-position masked as a non-participating azide to facilitate formation of the α-O-glycosidic linkage to the amino acid acceptor **62**. The donor was synthesized in three steps starting from an intermediate of the LacNAc type-1 synthesis, compound **18**. To enable introduction of the trichloroacetimidate leaving group, the anomeric acetate was removed by selective deacetylation with hydrazine acetate to give compound **65**. The trichloroacetimidate donor **66** was formed by using trichloroacetonitrile in alkaline conditions and was followed by coupling to the protected Tyr acceptor **62** applying *Schmidt* conditions. The glycan azide group was then converted to an NHAc group by reductive amidation using a thioacetic acid/pyridine system[320] to afford **68**. Finally, the free carboxylic acid was generated by deallylation according to the synthesis procedure described for α-GalNAc-Tyr in chapter 7.2.2. Starting from Fmoc-tyrosine-O-allyl ester **62** and monosaccharide **18** the synthesis of the α-GlcNAc-Tyr building block **69** proceeded with a total yield of 11% over five steps.

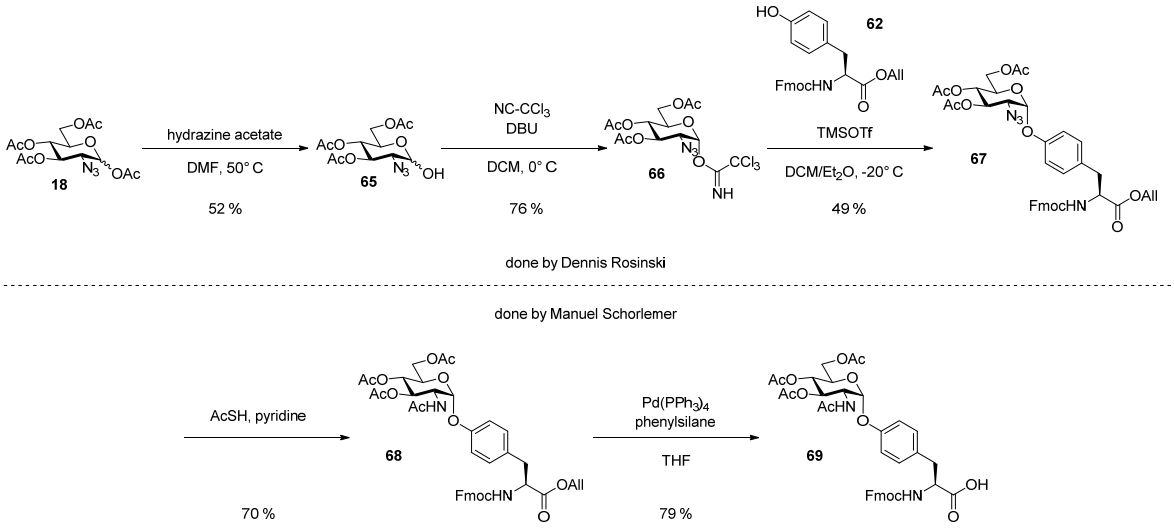


Figure 67: Synthesis of the α-GlcNAc tyrosine building block.

7.2.4 Synthesis of the β-GlcNAc-Tyr building block

The synthesis of the β-*O*-GlcNAc-Tyr amino acid building block was carried out by coupling according to the Schmidt method using the trichloroacetimidate donor **71** and the protected Tyr acceptor **62** as depicted in ***Figure 68***. The starting point for the donor synthesis was compound **25**, which constitutes a GlcNAc monosaccharide building block from the LacNAc type-2 disaccharide synthesis, equipped with an *N*-Troc group in the 2-position and global acetyl protection of the hydroxyl groups. Regioselective deacetylation with hydrazine acetate in THF afforded the free anomeric hydroxyl **70**, which was then converted to the corresponding α-trichloroacetimidate **71** using trichloroacetonitrile in dichloromethane under addition of the α-directing base DBU. The glycosylation step itself was carried out under addition of catalytic amounts of trimethylsilyl trifluoromethanesulfonate (TMSOTf) as promotor. Since the stereoselectivity of the reaction was guided by the donor carbamate, the reaction could be performed at room temperature without formation of the thermodynamic α-glycoside. The generated coupling product **72** was then converted to acetamide **73** by applying a zinc/acetic acid system with subsequent amine acylation using standard acetic anhydride/pyridine conditions. The final C-terminal *allyl ester* removal was done analog to the other HexNAc-Tyr building block synthesis in a *Tsuji-Trost* reaction. The complete βGlcNAc-Tyr building block **74** was prepared in five steps with a total yield of 15 % starting from the common Tyr acceptor **62** and the monosaccharide **25**.

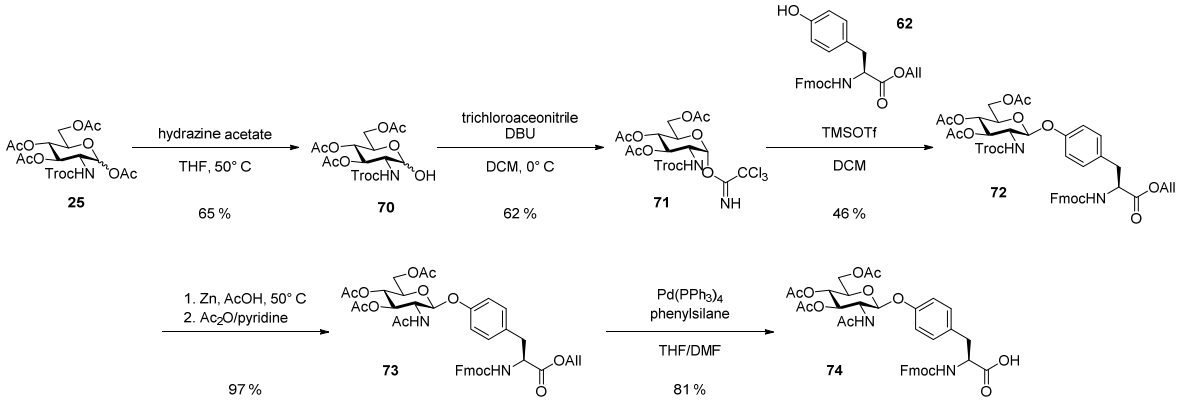


Figure 68: Synthesis of the β-GlcNAc-Tyr Fmoc-SPPS building block **74**.

7.2.5 Generation of the HexNAc glycopeptide library

In the frame of this project a glycopeptide library was generated, which consists of synthetic peptides such as MUC1 model peptides and glycopeptide sequences identified to contain HexNAc-Tyr/Ser/Thr *O*-glycosylation sites in eukaryotic cells. In addition to synthetic glycopeptides based on tryptic fragments identified in previous glycoproteomic work, additional HexNAc analogs were made to compare HexNAc glycosylation (αGalNAc vs βGlcNAc vs αGlcNAc) at the same site or to compare impact of glycan presentation by glycosylation at different amino acids (Tyr vs Ser vs Thr). Possible influence by bidentate or multivalent glycan presentation was also explored by preparation of di-valent glycopeptides. Unglycosylated sequences and mixed mitochondrial or small GTPase segments of glyco/phospho-peptides were also synthesized. These peptides enabled to study potential influences of vicinal phosphorylation to peptide conformation and stability as well as to elucidate substrate or binding specificity of *O*-GlcNAc enzymes involved in biosynthetic pathways or lectins commonly used for detection or enrichment of HexNAc glycopeptides. The general considerations and the synthesis strategy corresponds to the one used for build-up of the mucin type glycopeptide library described in chapter 5.5 and 7.1.14.1 with the following additions:

The sterically less demanding HexNAc amino acid building blocks were, unlike the mucin type amino acids, all coupled with donor excess of 1.5 compared to the resin loading without exception. For preparation of vaccine constructs to generate GalNAc-/GlcNAc-Tyr specific antibodies, *N*-terminal acetyl capping was performed on-resin with a preactivated solution of 500 mM acetic anhydride, 15 mM HOBt and 125 mM diisopropylethylamine in dimethylformamide for 40 minutes. Then release from the resin and a preceding desalting step via C18-cartridge (1 g) was performed. A C-terminal Boc-protected spacer **75** (**Figure 69 A**) was then coupled in solution with a HATU/HOAt/DIPEA system in 150 µl dimethylformamide within 8 h (peptide/spacer/HATU/HOAt/DIPEA 1:1.5:1.25:1.25:2.5, molar

ratios). The spacer was kindly provided by Dr. Yu Jin. Afterwards the reaction solutions were diluted with water and lyophilized. Acidic Boc removal was done by employing TFA/water/TIPS (15:1:1 v/v/v) for 2 h followed by a desalting step and glycan deprotection with sodium methoxide in methanol.

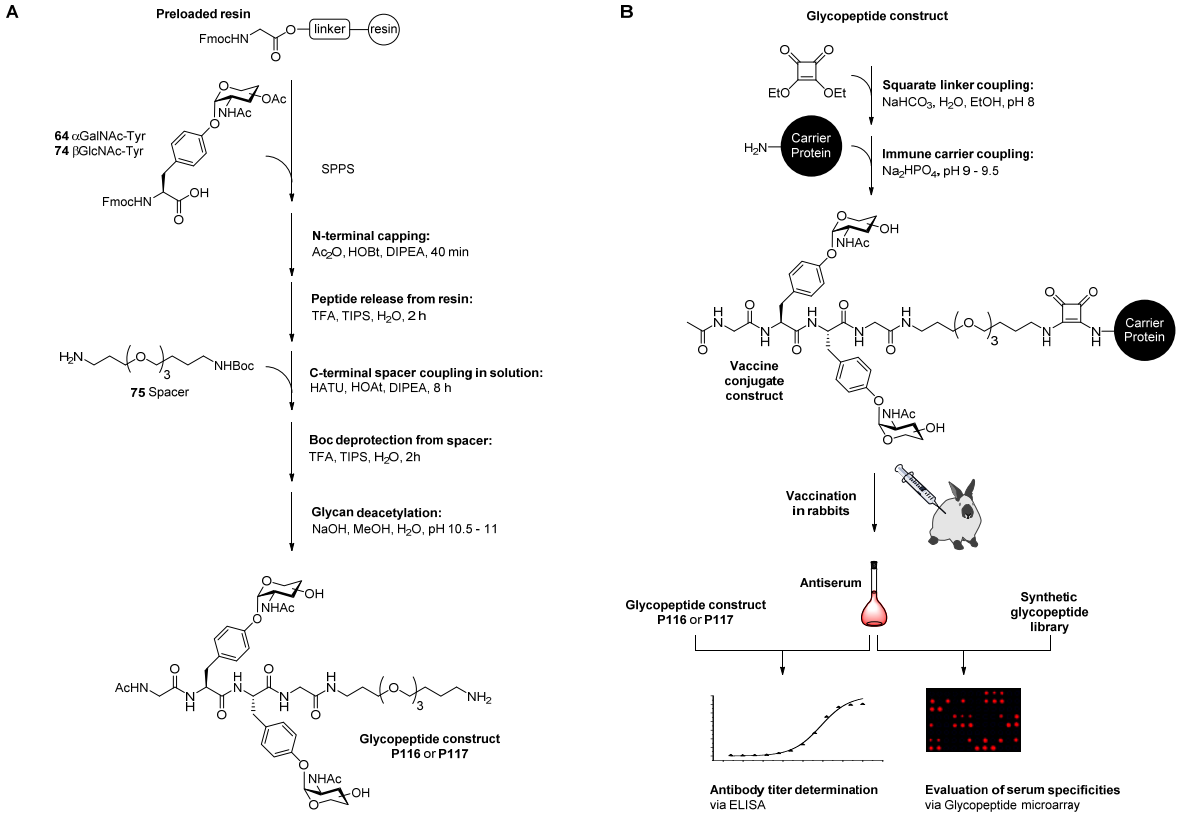


Figure 69: A) Synthesis of rational glycopeptides **P116** and **P117**, which did not follow the standard glycopeptide synthesis sequence from this work. **B**) Aimed generation of vaccine conjugate constructs from **P116** and **P117** for immunization experiments and antibody evaluation.

The glycopeptide constructs, which were synthesized according to this protocol, were aimed to be further modified to generate full vaccine conjugate constructs for generation of HexNAc-Tyr-specific antibodies as detection and enrichment tools (***Figure 69*** **B**). However, the synthesis of the GY*Y*G amino acid sequence proceeded with low yields (4 - 8 %, compare ***Table 2***), therefore the glycopeptides were not further worked up to full vaccine conjugates. Synthesis of the alternative amino acid sequence GY*Y*A in the group proceeded with considerably higher yields and was therefore utilized for vaccine design instead. In total the library includes 80 peptides, glycopeptides and phosphopeptides. A comprehensive overview is given in ***Table 2***.

Table 2: HexNAc-*O*-glycopeptide library.

protein	segment	PTM	ID	sequence	deacetylation		yield (%)
					*p*H	time (h)	
Amyloid-β (*homo sapiens*)	672-687	---	**P35**	TEG-DAEFRHDSGYEVHHQK	---	---	87
		αGalNAc	**P36**	TEG-DAEFRHDSGY*EVHHQK	11.0	16	53
		αGlcNAc	**P37**	TEG-DAEFRHDSGY*EVHHQK	10.5	16	63
		βGlcNAc	**P38**	TEG-DAEFRHDSGY*EVHHQK	11.0	16	49
		phospho	**P39**	TEG-DAEFRHDSGY*EVHHQK	---	---	60
Asp amino-transferase (*homo sapiens*)	91-107	---	**P40**	TEG-NLDKEYLPIGGLAEFCK	---	---	49
		αGalNAc	**P41**	TEG-NLDKEY*LPIGGLAEFCK	10.5	23	32
		αGlcNAc	**P42**	TEG-NLDKEY*LPIGGLAEFCK	10.5	16	41
		βGlcNAc	**P43**	TEG-NLDKEY*LPIGGLAEFCK	10.5	23	18
		phospho	**P44**	TEG-NLDKEY*LPIGGLAEFCK	---	---	41
	127-139	βGlcNAc	**P45**	TEG-FVTVQTIS*GTGALR	---	---	10
Asp amino-transferase (*rattus norvegicus*)	326-337	---	**P46**	TEG-IAATILTSPDLR	---	---	46
		βGlcNAc	**P47**	TEG-IAATILTS*PDLR	10.0	48	42
			P48	TEG-IAAT*ILTSPDLR	10.0	48	55
			P49	TEG-IAATILT*SPDLR	10.0	48	51
			P50	TEG-IAAT*ILT*SPDLR	10.0	48	36
			P51	TEG-IAAT*ILTS*PDLR	10.0	48	24
			P52	TEG-IAATILT*S*PDLR	10.0	48	31
ATP synthase-α (*homo sapiens/ mus musculus*)	335-347	---	**P53**	TEG-EAYPGDVFYLHSR	---	---	65
		βGlcNAc	**P54**	TEG-EAYPGDVFYLHS*R	10.0	48	13
			P55	TEG-EAY*PGDVFYLHSR	10.0	48	20
			P56	TEG-EAYPGDVFY*LHSR	10.0	48	8
			P57	TEG-EAY*PGDVFY*LHSR	10.0	48	22
			P58	TEG-EAYPGDVFY*LHS*R	10.0	48	17

protein	segment	PTM	ID	sequence	deacetylation		yield (%)
					*p*H	time (h)	
ATP synthase-β (*homo sapiens*)	311-324	---	**P59**	TEG-FTQAGSEVSALLGR	---	---	27
		βGlcNAc	**P60**	TEG-FTQAGS*EVSALLGR	10.0	48	22
			P61	TEG-FTQAGSEVS*ALLGR	10.0	48	64
			P62	TEG-FT*QAGSEVSALLGR	10.0	48	45
			P63	TEG-FT*QAGS*EVSALLGR	10.0	48	35
			P64	TEG-FTQAGS*EVS*ALLGR	10.0	48	36
	407-422	---	**P65**	TEG-IMDPNIVGSEHYDVAR	---	---	42
		βGlcNAc	**P66**	TEG-IMDPNIVGS*EHYDVAR	10.5	48	13
			P67	TEG-IMDPNIVGSEHY*DVAR	11.0	48	33
			P68	TEG-IMDPNIVGS*EHY*DVAR	10.5	48	20
		β GlcNAc*/ phospho#	**P69**	TEG-IMDPNIVGS*EHY#DVAR	10.0	16	59
		Phospho#/ βGlcNAc*	**P70**	TEG-IMDPNIVGS#EHY*DVAR	10.0	16	21
ATP synthase-β (*mus musculus*)	226-239	---	**P71**	TEG-AHGGYSVFAGVGER	---	---	6
		βGlcNAc	**P72**	TEG-AHGGYS*VFAGVGER	10.5	48	63
			P73	TEG-AHGGY*SVFAGVGER	11.0	48	63
			P74	TEG-AHGGY*S*VFAGVGER	10.0	48	12
	407-422	---	**P75**	TEG-IMDPNIVGNEHYDVAR	---	---	56
		αGalNAc	**P76**	TEG-IMDPNIVGNEHY*DVAR	9.5	16	31
		αGlcNAc	**P77**	TEG-IMDPNIVGNEHY*DVAR	10.5	16	33
		βGlcNAc	**P78**	TEG-IMDPNIVGNEHY*DVAR	9.5	16	33
		phospho	**P79**	TEG-IMDPNIVGNEHY*DVAR	---	---	21
cytochrome-c oxidase subunit 4 isoform 1 (*rattus norvegicus*)	30-42	---	**P80**	TEG-SEDYALPSTVDRR	---	---	36
		βGlcNAc	**P81**	TEG-S*EDTALPSTVDRR	10.0	48	54
			P82	TEG-TEDYALPS*TVDRR	10.0	48	58
			P83	TEG-TEDY*ALPSTVDRR	10.0	48	43
			P84	TEG-SEDYALPSY*VDRR	10.0	48	52
			P85	TEG-S*EDY*ALPSTVDRR	10.0	48	46
			P86	TEG-SEDY*ALPS*TVDRR	10.0	48	46
			P87	TEG-SEDYALPS*Y*VDRR	10.0	48	31

protein	segment	PTM	ID	sequence	deacetylation		yield (%)
					*p*H	time (h)	
CDC42 (*homo sapiens*)	27-38	---	**P88**	TEG-KFPSEYVPTVFD	---	---	51
		αGalNAc	**P89**	TEG-KFPSE**Y***VPTVFD	10.5	16	40
		αGlcNAc	**P90**	TEG-KFPSE**Y***VPTVFD	11.0	16	17
		βGlcNAc	**P91**	TEG-KFPSE**Y***VPTVFD	11.0	16	34
		phospho	**P92**	TEG-KFPSE**Y***VPTVFD	---	---	44
MUC1	VNTR	αGalNAc	**P93**	TEG-PAHGV**Y***SAPDTRPAPGSTA	10.5	16	50
			P94	TEG-PAHGVTSAPD**Y***RPAPGSTA	10.5	16	50
			P95	TEG-PAHGVTSAPDTRPAPGS**Y***A	10.5	16	54
		α-GlcNAc	**P96**	TEG-PAHGV**Y***SAPDTRPAPGSTA	11.0	16	47
			P97	TEG-PAHGVTSAPD**Y***RPAPGSTA	11.0	16	46
			P98	TEG-PAHGVTSAPDTRPAPGS**Y***A	11.0	16	42
		β-GlcNAc	**P99**	TEG-PAHGVTSAPD**S***RPAPGSTA	10.5	16	24
			P100	TEG-PAHGVTSAPDTRPAPGS**T***A	9.5	16	56
			P101	TEG-PAHGV**Y***SAPDTRPAPGSTA	10.5	16	52
			P102	TEG-PAHGVTSAPD**Y***RPAPGSTA	10.5	16	58
			P103	TEG-PAHGVTSAPDTRPAPGS**Y***A	10.5	16	52
Nucleobindin-2 (*homo sapiens*)	387-399	---	**P104**	TEG-LEYHQVIQQMEQK	---	---	59
		α-GalNAc	**P105**	TEG-LE**Y***HQVIQQMEQK	10.5	16	8
		β-GlcNAc	**P106**	TEG-LE**Y***HQVIQQMEQK	10.5	16	25
		phospho	**P107**	TEG-LE**Y***HQVIQQMEQK	---	---	45
Proto-oncogene tyrosine-protein kinase receptor Ret (*homo sapiens*)	680-692	---	**P108**	Ac-AQAFPVSYSSSGA	---	---	50
RAC1 (*homo sapiens*)	27-38	---	**P109**	TEG-AFPGEYIPTVFD	---	---	85
		α-GalNAc	**P110**	TEG-AFPGE**Y***IPTVFD	10.5	16	35
		α-GlcNAc	**P111**	TEG-AFPGE**Y***IPTVFD	11.0	16	17
		β-GlcNAc	**P112**	TEG-AFPGE**Y***IPTVFD	11.0	16	28
		phospho	**P113**	TEG-AFPGE**Y***IPTVFD	---	---	11
rational constructs	---	---	**P114**	Ac-GYYG-TEG	---	---	43
			P115	Ac-KKVPVSRA	---	---	92
		α-GalNAc	**P116**	Ac-G**Y***Y***G-TEG	9.0	48	4
		β-GlcNAc	**P117**	Ac-G**Y***Y***G-TEG	9.0	48	8
Retinoblastoma-like protein 2 (*homo sapiens*)	411-422	---	**P118**	Ac-KENPAVTPVSTA	---	---	74

protein	segment	PTM	ID	sequence	deacetylation		yield (%)
					*p*H	time (h)	
RHOA (*homo sapiens*)	29-40	---	**P119**	TEG-QFPEVYVPTVFE	---	---	54
		α-GalNAc	**P120**	TEG-QFPEV**Y***VPTVFE	10.5	16	22
		α-GlcNAc	**P121**	TEG-QFPEV**Y***VPTVFE	11.0	16	15
		β-GlcNAc	**P122**	TEG-QFPEV**Y***VPTVFE	11.0	16	7
		phospho	**P123**	TEG-QFPEV**Y***VPTVFE	---	---	6

7.2.5.1 MUC1 model peptides

The MUC1 VNTR model peptides were synthesized as HexNAc-Ser/Thr variants as well as corresponding Tyr analogs following the sequence TEG-PAHGVX*SAPDX*RPAPGX*TA. The resulting sequence selection enables comparative *O*-glycosylation studies while circumventing sequence dependent influence to the experimental results. While the monoglycosylated variants of the common αGalNAc-Thr glycopeptides are already covered in the mucin glycopeptide library, here additional αGlcNAc-Tyr as well as βGlcNAc-Tyr, -Thr and –Ser glycopeptide variants were synthesized. These analogs embody MUC1-glycopeptides with direct substitution of the αGalNAc-Thr with the different glycosylated amino acids

7.2.5.2 Amyloid-β

For SPPS the tryptic DAEFRHDSGYEVHHQK sequence has been chosen, which represents the *N*-terminal hexadecameric segment of the neurotoxic amyloid-β-APP peptide (***Figure 70***). CID experiments on immunopurified amyloid-β identified Tyr682 to be *O*-glycosylated with $(Neu5Ac)_{1\text{-}2}Hex(Neu5Ac)HexNAc$ structures[177]. Although a HexHexNAc structure hints to be of mucin type, the applied CID methodology does not provide waterproof conformational information about the carbohydrate subunits. Therefore three glycoforms of amyloid-β 672-687 containing either α-GalNAc-, α-GlcNAc- or β-GlcNAc-Tyr were generated. Additionally, one peptide analog containing phosphotyrosine was made.

661 694
IKTEEISEVKMDAEFRHDSGY*EVHHQKLVFFAED
695 729
VGSKKGAIIGLMVGGVVIATVIVITLVMLKKQWTS

Figure 70: The synthesized segment (red) of amyloid-β represents the N-terminal segment of the AD related amyloid-β-APP peptide (underlined).

7.2.5.3 Rho-GTPases

Upon intracellular infections the switch I region of RhoA, Rac1 and Cdc42 GTPases have been found to be modified with αGlcNAc at the Tyr32(34) site. To explore these modifications in more detail, undecameric glycopeptides were prepared containing the GlcNAc and GalNAc glycoforms of the switch I Tyr site. Further, a phosphotyrosine variant was included for each sequence homolog as Tyr32(34) in Rho GTPases, which also is a known phosphorylation site[370]. The synthesized segment was selected to position the Tyr residue in the center of the glycopeptides. In this way the length of the peptide was suitable for Fmoc-SPPS while still maintaining the amino acid environment around the glycosylation site. The prepared peptide sequences are shown in ***Figure 71***.

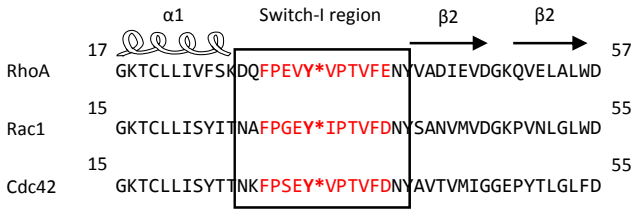


Figure 71: The synthesized sequences (red) of Rho A, Rac1 and Cdc42 are segments of the highly conserved Switch-I region.

7.2.5.4 Mitochondrial Proteins

In 2013 glycoproteomic LC/MS studies on murine synaptosomes from mice revealed Tyr *O*-glycosylation sites on several independent mitochondrial proteins. In ATP synthase subunit beta (Atp5b) from mice Tyr418 has been identified (IMDPNIVGNEHY*DVAR), in aspartate aminotransferase (mAspAT) Tyr96 was found (NLDKEY*LPIGGLAEFCK) and in voltage-dependent anion-selective channel protein (Vdac1) Tyr186 (VTQSNFAVGY*K) was identified as glycosylation site[188]. As the thereby utilized mass-fragmentation based methodology did not allow to distinguish GlcNAc and GalNAc subunits, the authors were not able to determine which structural HexNAc isomer was responsible for these modifications. However, the location in mitochondria and knowledge about nearby Ser/Thr HexNAc modifications and identified phosphorylation sites, would suspect that a new PTM was identified: βGlcNAc-Tyr. In earlier work we showed that HCD-MS fragmentation techniques allowed structural identification of HexNAc epimers such as GalNAc and GlcNAc attached on Thr/Ser modified glycopeptides and we were interested to explore if it was possible to distinguish these modification on Tyr residues as well. In that case the methodology would be available for future identification of possible βGlcNAc-Tyr modifications. For these reasons all three mitochondrial tryptic glycopeptide sequences with an identified HexNAc-Tyr residue were synthesized as GalNAc and GlcNAc glycoforms and added to the library. However, the synthesized Vdac1 glycopeptides were not included to the

library due to solubility problems upon purification. In addition to the different HexNAc-Tyr glycopeptides, phosphopeptide variants of the Atp5b(407-422) and mAspAT(91-107) sequences were synthesized and added to the library.

To evaluate differences of Ser, Thr and Tyr *O*-GlcNAcylation (βGlcNAc modification) in the context of lectin specificity as well as enzyme substrate specificity of *O*-GlcNAcase and GlcNAc transferase (OGT), which are involved in biosynthetic pathways, more peptide sequence examples containing known Ser and Thr *O*-GlcNAcylation were needed. Moreover, the above mentioned reported Tyr glycopeptides were not containing potential *O*-glycosylation in close vicinity, which would enable to study possible cooperative effects of multivalent *O*-GlcNAc presentation. To enable such studies of for instance a possible impact of bidentate and multivalent binding of lectins commonly used in affinity enrichment of tryptic glycopeptides or to explore cooperative effects during enzymatic transfer or removal of GlcNAc, different glycopeptide sequences were needed which simultaneously exhibit Ser, Thr and Tyr in close vicinity. Several mitochondrial tryptic glycoprotein fragments were chosen from a pool of reported[66] mitochondrial tryptic glycoprotein segments, which exhibit known *O*-GlcNAc sites and fulfil the needed criteria. Human Atp5a(335-347) EAYPGDVFYLHSR exhibits two additional Tyr amino acids in different distances from the known *O*-glycosylation site Ser346. Atp5b(226-239) from mouse AHGGYSVFAGVGER exhibits a known *O*-GlcNAc site at Ser231 with a neighboring Tyr downstream at -1. The segment Atp5b(407-422) from mice (IMDPNIVGNEHYDVAR) was of particular interest, which carries a known Tyr *O*-GlcNAc site on Tyr418 (see above), while the human Atp5b analog IMDPNIVGSEHYDVAR exhibits a identified *O*-GlcNAc site at the nearby Ser415 residue (Glu415 in mice). Mono and divalent HexNAc glycopeptides based on this sequence were therefore added to the library as well. To enhance the sequence versatility of Tyr absent glycopeptides, analogs of the Ser and Thr rich human Atp5b(311-324) segment FTQAGSEVSALLGR with an identified *O*-GlcNAc site on Ser316 were also prepared. Glycopeptides based on mAspAT(326-337) IAATILTSPDLR with the known *O*-GlcNAc site Thr329 and the segment SEDYALPSTVDRR from cytochrome-c oxidase subunit 4 isoform 1 (Cox4i1) with a known *O*-GlcNAc site on Ser37 were further included. All these sequences exhibit several Ser, Tyr and Thr, which allows for the generation of multiple mono and divalent isoforms.

7.2.5.5 Nucleobindin-2

Levery and *Clausen* reported the identification of an *O*-HexNAc modification on tyrosine in in the tryptic nucleobindin-2 (NucB2) fragment LEYHQVIQQMEQK as indicated by nLC-ETD experiments[181]. Due to isolation of the glycopeptides using the VVA-lectin, they assume the modification to be GalNAc.

Glycopeptide sequences were synthesized according to the fragment found by *Levery* and *Clausen* (NucB2 387-399) as GalNAc and GlcNAc variants as well as a corresponding phosphopeptide.

7.2.5.6 HexNAc-Tyr antigen glycopeptide sequences

In addition to biological relevant glycoprotein segments, several rational glycopeptide constructs were included to the library. The Ac-GY*Y*G-TEG sequence was designed as an efficient B-cell epitope for immunization experiments to generate GalNAc- and GlcNAc-Tyr specific antibodies via conjugation to an immune stimulant (***Figure 72***). The concept was to have a short glycopeptide with high HexNAc-Tyr density while keeping the immunogenicity of the peptide backbone low. To this extent, two neighboring Tyr glycosylation sites were framed by glycine amino acid spacer residues and the *N*-terminal amine was masked by acetyl capping to mimic a natural peptide backbone environment. It has been found, that synthetic glycopeptides, which are linked via oligoethylene glycol spacer to a T-cell epitope induce highly specific antibody sera without additional immune response to the spacer[170, 363, 371]. Therefore, a triethylene glycol spacer amino acid was used for conjugation to the C-terminal end of the tetrapeptide. The opposed spatial orientation of the glycosylated tyrosine side chains should minimize the risk to only generate antibodies with divalent selectivity. The synthesis of these tetrameric glycopeptides constructs turned out to progress with low yield (4-8 %), which seems to be connected with the utilized glycin rink amide resin (not fresh). Later attempts with a C-terminal Fmoc-Ala-rink amide resin instead, which were done by Dr. Yu Jin, progressed with better yields.

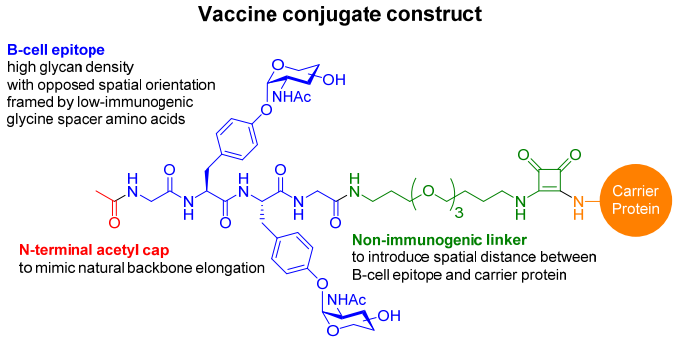


Figure 72: Components of aimed full vaccine construct.

7.2.5.7 Known OGT-substrates

Known substrates for human OGT have been included as positive control for enzymatic peptide glycosylation experiments. Among several identified[372] OGT substrates, a peptide derived from retinoblastoma-like protein 2 (Rbl2 411-422) Ac-KENPAVTPV**S**TA and a peptide derived from proto-

oncogene tyrosine protein kinase receptor Ret (Ret 680-692) Ac-AQAFPVSY**S**SSGA were selected for peptide synthesis. Both substrates include a Ser *O*-glycosylation site which is addressed by OGT. A rational glycopeptide Ac-KKVPVSRA-OH was also implemented to the library. This sequence has been reported[372] to be a very good substrate for the human OGT enzyme. Since the sequences were for use in enzymatic glycosylation experiments only, no glycosylated variants were synthesized.

7.2.6 Evaluation of lectins for HexNAc detection and enrichment

In glycoproteomic workflows, the lectins *Wheat Germ Agglutinin* (WGA), *Vicia Villosa Lectin* (VVA) and *Griffonia Simplicifolia Lectin II* (GSL II) are commonly utilized for enrichment or detection of HexNAc containing glycopeptides and glycoproteins. The lectin WGA is considered to be specific for terminal GlcNAc residues on oligosaccharides or from O-GlcNAcylated protein or peptide backbones and to further recognize 1,4-linked chitin oligomers[373-374]. Interactions of WGA with Neu5Ac and Neu5Gc sialic acid residues and terminal GalNAc have also been observed[375]. In contrast, VVA is known to specifically recognize α- and β-configured terminal GalNAc, in particular on GalNAc-Ser/Thr glycopeptides[376-377]. The lectin GSL II has been reported to bind terminal α/β-GlcNAc exclusively[378]. All this data is based on the recognition of either terminal monosaccharide subunits within larger oligosaccharides or of monosaccharides, which are directly linked to either Ser or Thr. Ultimately, there is no data on binding preferences of these lectins towards the recently discovered HexNAc-Tyr modifications. In order to evaluate these lectins binding specificities, a library of different αGalNAc-, αGlcNAc- and βGlcNAc-Tyr modified glycopeptide fragments were prepared and immobilized on microarray slides (slides with glycopeptide library **MA1**, details on p. 247). To enable a direct and sequence-independent comparison between HexNAc-Tyr and the corresponding HexNAc-Ser/Thr recognition, MUC1 sequence homologs were included to the microarray library, which only differ in the glycosylated amino acid. As an example **P93** TEG-PAHGVY(GalNAc)SAPDTRPAPGSTA can be directly compared with **P2** TEG-PAGHVT(GalNAc)SAPDTRPAPGSTA and the corresponding α- and β-configured GlcNAc analogs. Further, an unglycosylated variant of each peptide sequence was introduced as negative control. In total, on one slide nine incubation experiments were performed in nine wells with five replicates of the 39 glycopeptide library per well and incubation. In order to optimize the lectin concentration with respect to signal to noise ratio and possible saturation effects, each lectin was applied at three different concentrations: 10, 50 and 100 µg/ml. A schematic illustration is shown in ***Figure 73***. All lectins were obtained as biotinylated conjugates from *Vector Laboratories* (Burlingame, California). After lectin incubation, staining was done using *ZyMax™ Streptavidin-Cy™5* (*Life Technologies Corporation*, Frederick, Maryland).

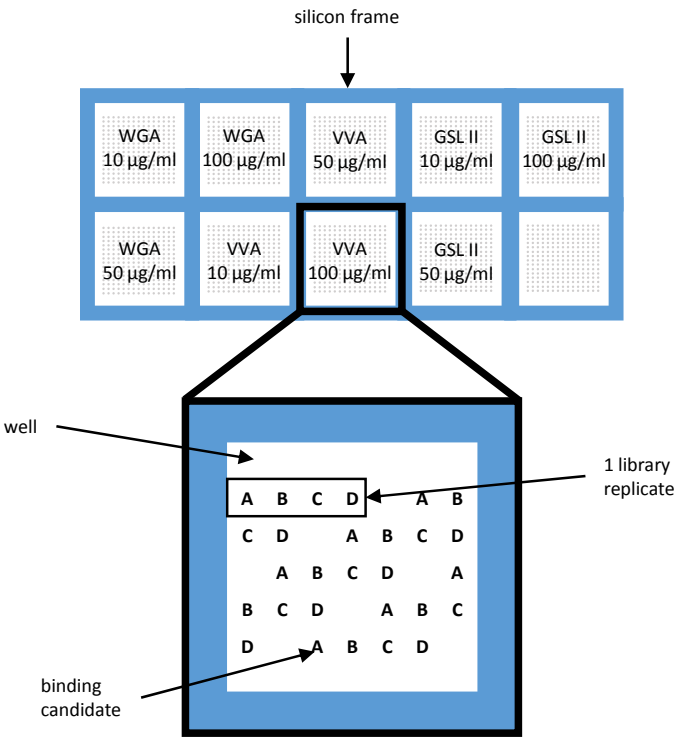


Figure 73: Schematic representation of the on-slide incubations.

In general, the experimental results from all three tested lectin concentrations do not differ significantly. In general, the fluorescence intensities of the 10 µg/ml batches were slightly lower if compared with results from the 50 and 100 µg/ml incubations. Nevertheless, at these rather high lectin concentrations, the recognition binding patterns did not differ with respect to the different lectin concentrations. As expected, none of the unglycosylated peptides were recognized by any of the tested lectins. Interestingly, none of the included MUC1 αGalNAc- and βGlcNAc-Thr glycopeptides were recognized by WGA (***Figure 74***). In contrast, all αGalNAc-Tyr as well as both, α- and βGlcNAc-Tyr glycopeptides were recognized very well by WGA, but with a generally higher affinity for GlcNAc than GalNAc. Here, even the MUC1-model peptides containing GalNAc-Tyr (**P93**, **P94** and **P95**) were recognized, although their Thr equivalents (**P2**, **P3** and **P4**) were not.

In contrast to WGA, the other two tested lectins VVA and GSL II showed clear selectivity each. In agreement with its postulated selectivity, VVA specifically recognized GalNAc on Thr and Tyr modified glycopeptides (***Figure 75***). Like WGA also VVA shows a preference towards GalNAc-Tyr over GalNAc-Thr as indicated by generally higher fluorescence intensities for **P93**, **P94** and **P95** if compared to their analogs **P2**, **P3** and **P4**. The purely GlcNAc selective GSL II also follows the trend of a generally high HexNAc-Tyr affinity (***Figure 76***). Almost all GlcNAc-Tyr peptides were recognized by GSL II, the tested βGlcNAc-Thr glycopeptide analogs **P100** and **P124** were not bound by the lectin. From these experiments it can be concluded, that only VVA and GSL II showed clear selectivity to one specific

HexNAc epimer on Tyr. In contrast WGA showed admittedly higher affinity towards βGlcNAc-Tyr, but since also αGalNAc-Tyr serves as receptor for this lectin, it cannot be considered to be HexNAc-selective. Interestingly, the data suggest a strong amplifying effect of Tyr towards the binding affinity of all tested GlcNAc specific lectins compared to Thr. The preference of both WGA and GSL II for Tyr modified glycopeptides might be explained by the occurance of a hydrophobic pocket that supports π-π stacking interactions. Additionaly less entropic penalty is probably caused upon binding to glycopeptides with glycans presentated on a less flexible Tyr residue.

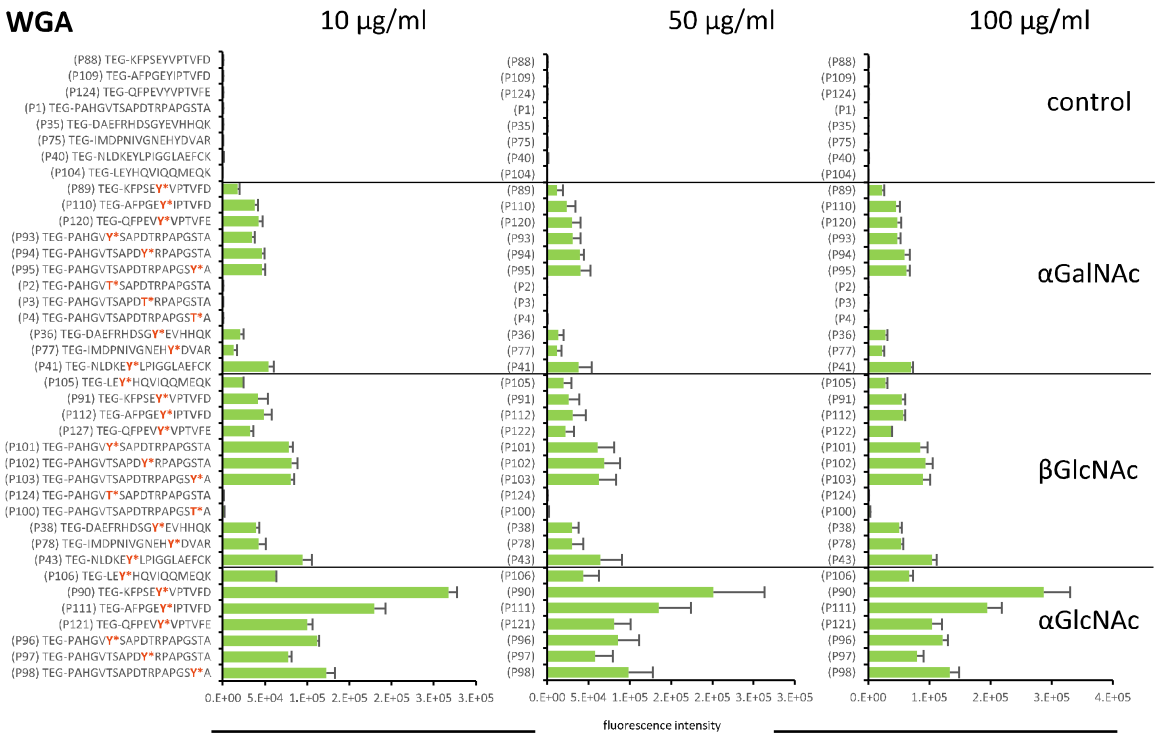


Figure 74: Fluorescence intensities obtained from WGA incubations on slide **MA1** using different lectin concentrations.

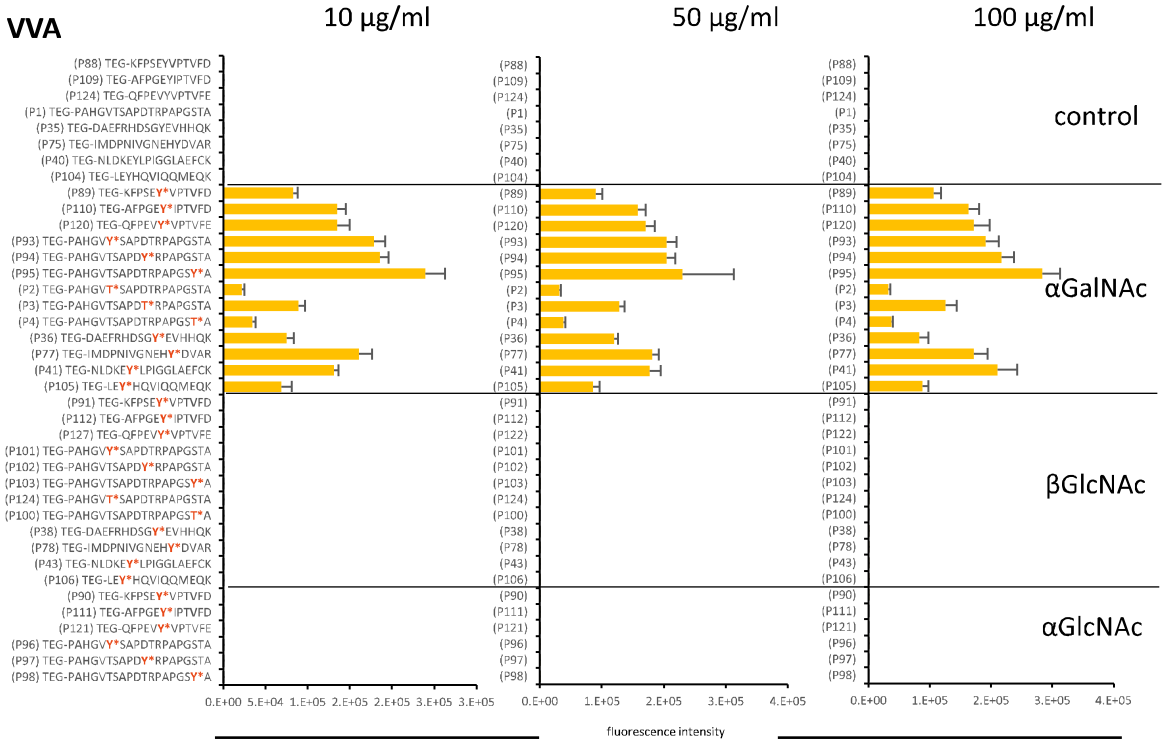


Figure 75: Fluorescence intensities obtained from VVA incubations on slide **MA1** using different lectin concentrations.

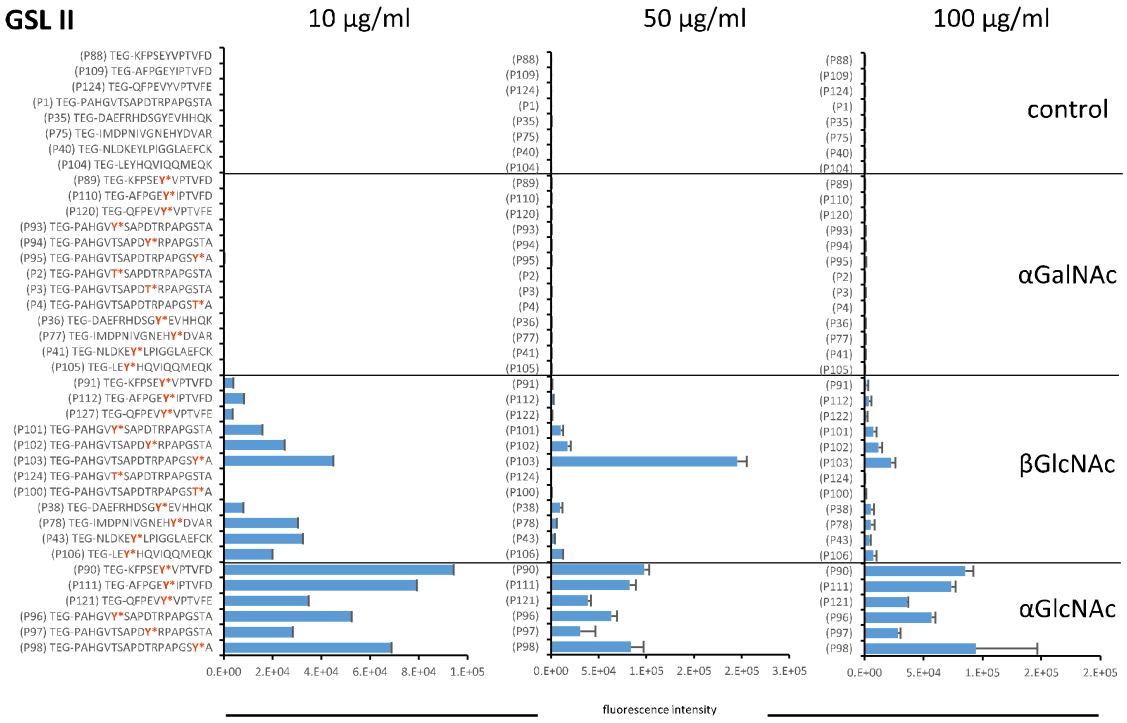


Figure 76: Fluorescence intensities obtained from GSL II incubations on slide **MA1** using different lectin concentrations.

With these results at hand a second generation of microarray incubations was performed on microarray slide **MA2** (for details see p. 249). Here, the aim was to expand the data base from GlcNAc-

Thr/Tyr recognition towards GlcNAc-Ser recognition to enable a direct comparison between all three amino acid acceptors. Since monovalent binding events of carbohydrates and proteins are typically weak, this is commonly compensated by the existence of multiple carbohydrate binding sites per lectin. This enables the simultaneous binding of several glycans per lectin, which results in multivalent ligand/lectin complexes with higher binding affinities[379]. To investigate the influence of Tyr on common multivalent presentation, glycopeptide fragment candidates were included, which exhibit *O*-glycosylation sites in various proximities. These peptide sequences were included in their monoglycosylated as well as diglycosylated forms. These divalent peptides were selected to maximize the versatility regarding distance and combination of the various amino acid acceptors. Finally, the heterodivalent phospho-/glycopeptides **P69** and **P70** were included to obtain clues, whether nearby phosphorylation has an impact on the lectin recognition and its corresponding binding affinity. This is of interest, as many GlcNAc sites have also been found to be phosphorylation sites and many GlcNAc sites are either in close proximity to further GlcNAc sites or genuine phosphorylation sites. In the preceding microarray experiments a lectin concentration of 10 µg/ml was proven to be sufficient and no noteworthy differences to higher concentrations were observed. The GlcNAc microarray library **MA2** was therefore incubated with 10 µg/ml lectin per experiment solely. Besides WGA and GSL II, which both show selectivity towards GlcNAc also the GalNAc selective VVA was incubated with the new library to screen for potential cross reactivity towards multivalently presented GlcNAc. However, no fluorescence could be observed after staining the VVA-incubated library grids, suggesting that no cross reactivity is induced by this kind of glycan presentation. Therefore, these VVA experiments are not in included in the overview in ***Figure 77***.

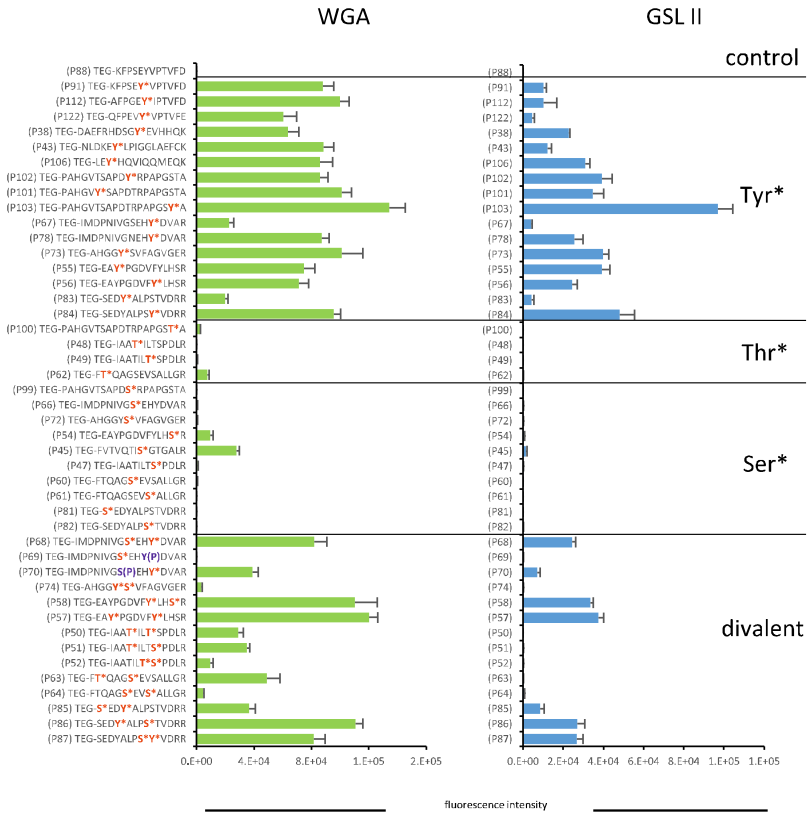


Figure 77: Results from WGA and GSL II incubations against the βGlcNAc microarray library on slide **MA2**. The lectins were incubated with 10 µg/ml each.

In general, with the GlcNAc microarray the results from the preceding incubation experiments with the general HexNAc array **MA1** could be reproduced. The binding affinity of both GlcNAc recognizing lectins was strong towards GlcNAc-Tyr exhibiting glycopeptides and the affinity towards their Thr analogs was considerably lower or even completely abolished. A similar picture was obtained for the corresponding Ser analogs. While some GlcNAc-Ser glycopeptides were very weak ligands for the lectins, most of the candidates were not even recognized at all. A more complex picture was obtained for the divalent presenting glycopeptides, which might be based on bidentate binding or multivalent effects, which may depend on specific lectin properties.

In general, the effects of multivalent glycan presentation can be classified as chelate effects, statistical rebinding and cross-linking effects[380]. The chelate effect is based on the simultaneous occupation of adjacent binding sites by one multivalent ligand. In contrast, the statistical rebinding is based on the occupation of only a single binding site by one ligand at a time. Here, the local epitope concentration allows for ligand rebinding before complete dissociation of this very ligand. Finally, lectin crosslinking occurs, if multivalent ligands are bound by multiple lectins at the same time. This can lead to the formation of lectin lattices[381].

Crystallographic studies in WGA[382] have revealed, that the lectin is a homodimer, while each dimer contains four unique carbohydrate binding sites. Thus, one complete lectin contains eight independent binding sites. All of these binding sites have been found to be simultaneously functional, which allows for bidentate or multivalent chelate binding. However, this recognition mode relies strongly on the spatial presentation of the carbohydrates and the flexibility of the linker (i.e. the peptide backbone)[380]. In addition, the high number of carbohydrate binding sites per homodimer allows for strong lectin cross-linking.

The results from both, WGA and GSL II incubations indicate clear effects of homodivalent glycan presentation on glycan recognition. Thereby, the length of the peptide backbone between the two glycans seems to be an important factor. A very short distance between the glycan residues presented on the peptide backbone led in the case of Atp5b 226-239 **P74** (TEG-AHGGY*S*VFAGVGER) to clearly decreased lectin binding compared to their monovalent analogs, as indicated by the fluorescence intensities. In contrast the divalent GlcNAc-Tyr glycopeptide Cox4i1 30-42 **P87** (TEG-SEDYALPS*Y*VDRR) and the monovalent counterpart were both strongly recognized by WGA and GSL II. Interestingly, AspAT 326-337 **P52** (TEG-IAATILT*S*PDLR) showed increased recognition compared to its monovalent analogs, although - just like in **P74** and **P87** (Y*S* and S*Y*) – the glycosylated amino acids are in direct neighboring position (***Figure 78***). Although the glycan presentation on these glycopeptides might look similar T*S* vs Y*S* vs S*Y*, the glycan orientation in the lectin binding pocket might be different in these cases depending on the order of the glycosylated Thr-, Ser-, and Tyr residues. Additionally the conformation at the glycosylation site and impact of steric and conformational differences depending on other nearby amino acids might be an issue for the lectin interaction. Further, differences in rotational degrees of freedom of Thr vs Tyr are expected. While Thr (and Ser) show almost full rotation at Cβ, the Tyr seems to be restricted to two favored conformations on its methylene Cβ. This might in some cases prevent the glycopeptide ligand to adopt a suitable conformation, which would fit for the simultaneous binding of both glycans by two different lectin binding sites. Glycopeptide candidates with longer amino acid residue distances between the glycan ligands showed throughout amplified or equal recognition compared to their respective monovalent analogs.

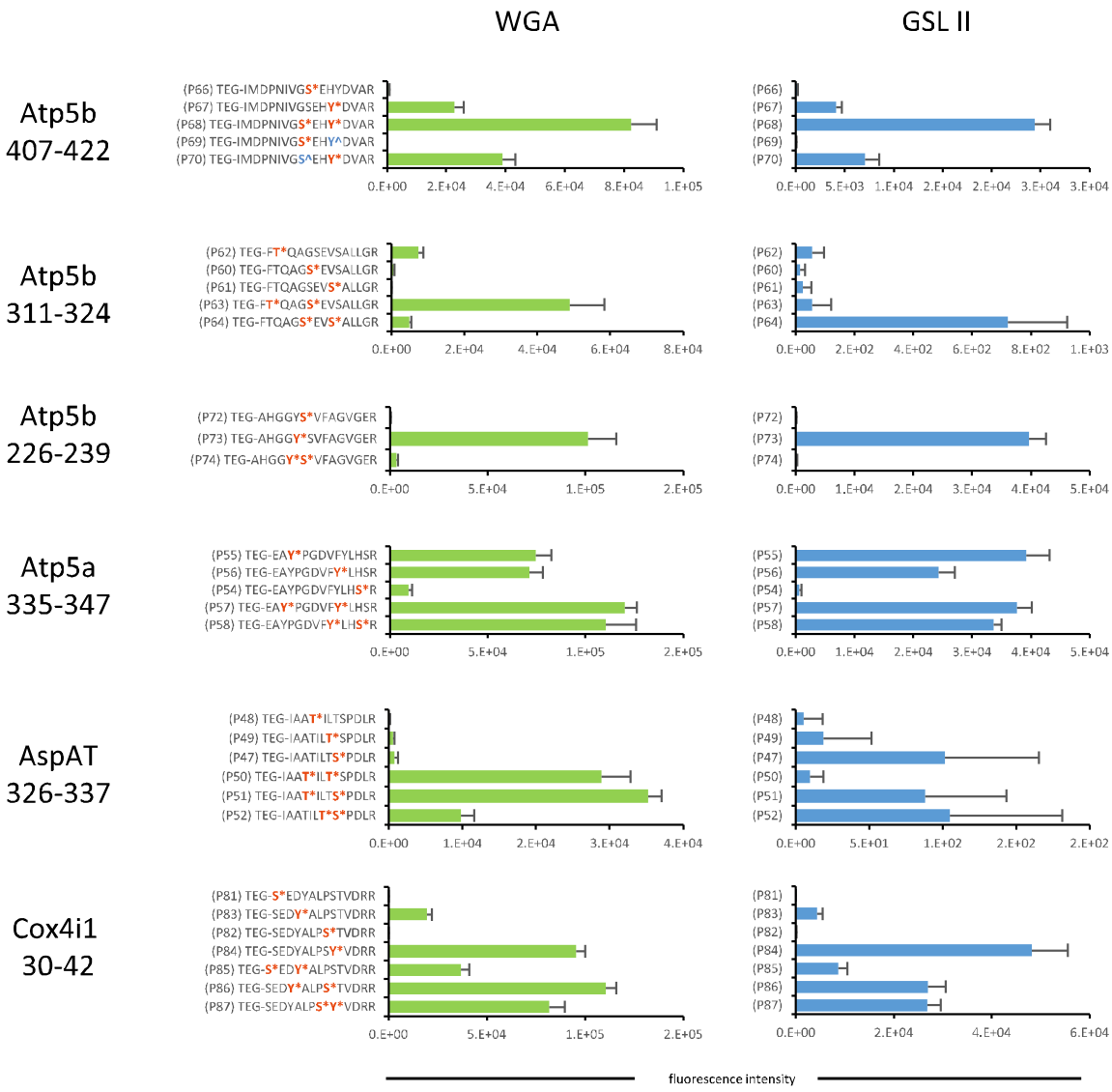


Figure 78: Influence of divalent glycan presentation towards WGA and GSL II recognition. * = βGlcNAc, ^ = phospho.

While these results show a clear recognition differences between Tyr and Ser/Thr glycosylation on the one hand and cooperative effects with divalent presentation on the other hand, these measurments were mainly made at one lectin concentration and did not allow determination of binding constants. Lectin binding was expected to be saturated at this concentration and also weak glycopeptide binders were expected to show maximum occupation in this case. Therefore, WGA follow-up experiments were performed on the basis of the present results. These experiments were done together with MSc. *Sandra Behren*. Incubations were made with lectin dilution series ranging from 3.1 ng/mL to 200 µg/mL and apparent K_D values were calculated from the obtained fluorescence data as described by *Gildersleeve* et al.[379]. To this extent, the data was plotted as a function of lectin concentration in logarithmic scale. Afterwards, nonlinear curve-fitting according to the equation

$$F_{obs} = \frac{F_{max}}{\frac{K_D}{[L]} + 1}$$

was done using the software *Origin 2017*. Here, F_{obs} is the mean fluorescence intensity from six replicate experiments at a lectin concentration *[L]*. F_{max} is the maximum fluorescence intensity and K_D is the apparent dissociation constant for binding of lectin and to a specific glycopeptide. The obtained results are shown in ***Figure 79***.

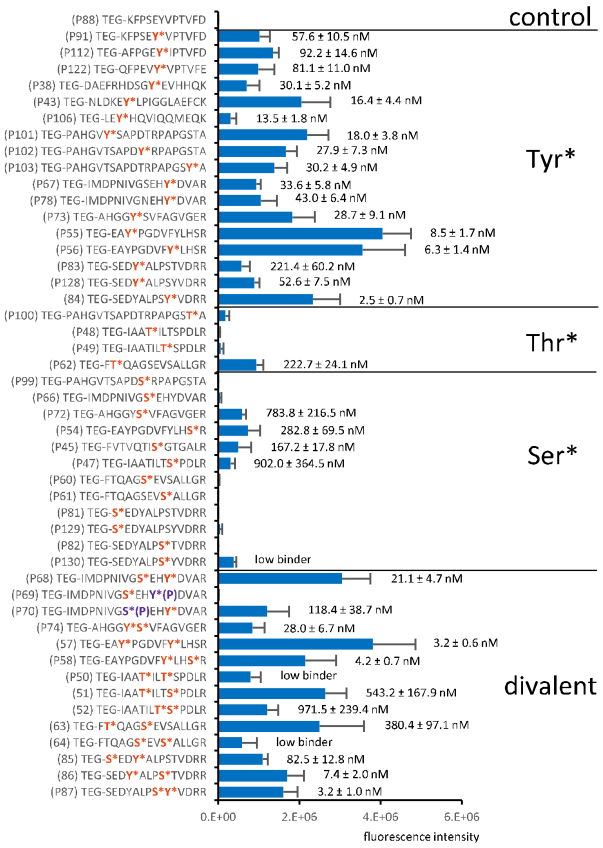


Figure 79: Fluorescence intensities and apparent dissociation constants as obtained from microarray binding experiments with WGA. Here, the fluorescence intensities are representatively shown for 100 µg/ml WGA experiments. K_D values over 1000 nM were marked as *low binder*. * = βGlcNAc, ^ = phospho.

The K_D values draw a similar picture as obtained in the preceding experiments, which were based purely on fluorescence intensities at a given lectin concentration. While monovalent GlcNAc-Tyr peptides were bound by WGA with low K_D values from 2.5-100 nM (one exception at 221 nM), the values from HexNAc-Ser and HexNAc-Thr experiments were found in the range between 100 and 1000 nM, if binding was observed at all. One example illustrating the influence of the backbone amino acid presenting the glycan structure is the difference observed in binding to a few MUC1 glycopeptide analogs, for instance replacement of a natural Thr glycosylation site with a Tyr (**P100** vs **P103**) resulted in a shift from no WGA binding to binding with 30 nM affinity. A similar observation was found by

replacing a Ser with a Tyr glycosylation site (**P99** vs **P102**), which resulted in a change from no binding to a 28 nM binder. Binding affinity was also found to be dependent on the peptide sequence and other amino acid residues nearby the glycosylation site, but the data clearly show that the amino acid presenting the glycan are most important upon WGA binding of monovalent peptides. A slightly more complex picture was obtained for glycopeptides with divalent glycan presentation. In agreement with earlier results, divalent glycan presentation with two to five spacing amino acids between the glycosylation sites generally led to positive cooperative effects (K_D values of down to 3.2 nM) compared to binding of their corresponding monovalent glycopeptide variants. However, directly neighboring glycan presentation did not lead to decreased K_D values. Instead, the K_D values were not affected at all. This indicates, that the peptide space need to exhibit a minimum length to bridge the distance between two glycan binding sites. This in turn supports the assumption of possible bidentate binding, in which one divalent glycopeptide is bound by two glycan binding sites of one very same lectin. These binding sites exhibit a defined distance and geometry to each other, which results in the necessity of peptide linkers of defined length in order to allow bidentate binding. Multivalent binding bridging two WGA molecules would also be possible and also require an optimal spacing for glycan ligand presentation.

In order to obtain more detailed insights into the role of peptide backbone distances for presentation of glycan ligands in the interactions with WGA during bidentate or multivalent binding, binding events of selected divalent glycopeptides are further studied by molecular modeling based on existing WGA crystal structures in a collaboration with Prof. Dr. *Robert J. Woods*, Georgia, USA. NMR conformational studies are also made in collaboration with Dr. Filipa Marcelo, Lisboa, Portugal

It can be concluded, that the common “GlcNAc” recognizing lectins WGA and GSL II showed a much higher binding affinity towards Tyr *O*-glycosylation compared to classical Ser and Thr *O*-glycosylation. However, discrimination between different HexNAc-Tyr epimers was not observed for WGA while GSL II showed selectivity for α- and βGlcNAc-Tyr glycopeptides. In contrast the lectin VVA was highly specific to αGalNAc containing glycopeptides, but showed no discrimination between glycan presentation on Tyr vs Ser or Thr residues. The lectins WGA and GSL II were also elucidated for binding to divalent glycopeptides with different peptide backbone spacing. The obtained data in most cases suggested a positive effect of divalent glycan presentation towards glycopeptide/lectin binding regardless of presentation on Ser/Thr or Tyr. This cooperative effect seems to be dependent on the peptide backbone distance between the two glycans on the divalent glycopeptide ligand. The results for WGA and GSL II did not show significant differences in this aspect.

These results highlight the influence of underlying peptide backbone and amino acid residues for glycan presentation in lectin binding events, which may be of great importance in diverse biological events. Further, the results enable a more precise utilization and interpretation of common lectin based enrichment strategies as used in glycomic and glycoproteomic workflows. It clearly shows, that the lectin preference towards Tyr glycosylation should always be taken into account, when using these technologies. Due to differences in lectin binding affinity to HexNAc residues presented on different peptide backbones, the pool of enriched glycopeptides may further not completely represent the natural HexNAc glycopeptide pool, which currently is not taken into consideration in glycoproteomic workflows. This impact should be the subject of further studies, where pools of selected glycopeptides are subjected directly to lectin weak affinity chromatography (LWAC). With this approach, the influence of directly competing glycopeptide candidates could be evaluated. Consequently, it would be possible to quantify the impact of the here presented phenomena on the enrichment results themselves.

7.2.7 Oxonium profiles of HexNAc-Tyr *O*-glycans

In our earlier work described above (chapter 7.1.18, p. 85)[189-190] we showed that oxonium ion profiles from HCD-MS fragmentation techniques allowed structural identification of HexNAc epimers such as GalNAc and GlcNAc attached on Thr/Ser modified glycopeptides and we were interested to explore if it was possible to distinguish these modification on Tyr residues as well. In that case the methodology would be available for future identification of a possible new PTM βGlcNAc-Tyr. Glycoproteomic studies of mitochondrial tryptic glycopeptides from ATP synthase subunit beta (Atp5b), aspartate aminotransferase (mAspAT) and in voltage-dependent anion-selective channel protein (Vdac1) Tyr186 (VTQSNFAVGY*K) were identified with HexNAc-Tyr glycosylation sites, which likely are βGlcNAc-Tyr modifications[188]. Unfortunately the methodology was at that time point not available to distinguish GalNAc and GlcNAc structural epimers by mass spectrometry, which motivated our current work on HCD-MS fragmentation. For HexNAc-Tyr *O*-glycans the oxonium ion fragmentation behavior was expected to be similar to the previous studies of HexNAc fragmentation on Ser/Thr glycopeptides with the difference that the HexNAc residues on Tyr was expected to be more labile based on that the aromatic residue of Tyr has a negative mesomeric effect on the glycosidic linkage. As a result, the cleavage of the glycosidic bond on Tyr should require less collision energy than glycosidic bonds on Ser and Thr.

In a first test the presence of all expected oxonium ions was examined for HexNAc-Tyr. The peptide sequence MUC1 VNTR was chosen, since the MUC1 αGalNAc-Thr variant **P2**

(TEG-PAHGVT*SAPDTRPAPGSTA) has been successfully used for fragmentation experiments in preceding studies[190]. The Thr variant **P2** was included as a control experiment and as HexNAc-Tyr glycopeptides, the αGalNAc modified **P94** and the βGlcNAc modified **P102** (both TEG-PAHGVTSAPDY*RPAPGSTA) were included in the fragmentation experiments. In the initial work the glycopeptides were subjected in a concentration of 400 fmol/15µl to a *Velos Orbitrap* mass spectrometer online-coupled to a nano RSLC HPLC system (both *Thermo Scientific*) for HCD fragmentation at an NCE level of 30 % in a buffer C/D system (C = 0.1 % FA in H_2O, D = 0.1 % FA in 84/16 MeCN/H_2O).

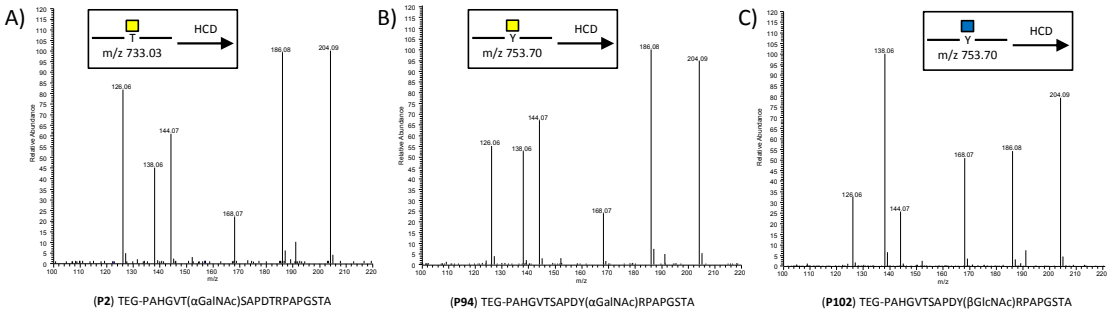


Figure 80: Average oxonium ion profiles obtained from test HCD fragmentation experiments to compare Thr and Tyr HexNAc-*O*-glycosylation. **A)** αGalNAc-Thr, **B)** αGalNAc-Tyr and **C)** βGlcNAc-Tyr.

All three tested glycopeptides showed expected MS^2-fragmentation oxonium ions of m/z 126, 138, 144, 168, 186 and 204 (***Figure 80***). Also, the obtained oxonium ion profiles suggested similar fragmentation characteristics for HexNAc-Tyr compared to HexNAc-Thr. Accordingly, the obtained GlcNAc/GalNAc ratios match to the postulated[190] rules for HexNAc-Ser/Thr (compare chapter 7.1.18). While the GlcNAc/GalNAc values for **P2** (αGalNAc-Thr) and **P94** (αGalNAc-Tyr) were obtained as 0.5 and 0.6 respectively, **P102** (βGlcNAc-Tyr) showed a value of 2.5.

In order to monitor the NCE-dependency of the fragmentation, oxonium ion profiles were recorded within a stepped NCE gradient between 10 and 50 % with intervals of 5 % in a follow-up experiment. Thereby, the glycopeptide sample pool was extended by the MUC1 αGlcNAc-Tyr isomer **P96** (TEG-PAHGVY*SAPDTRAPGSYA). This enables a view on the influence of the anomeric HexNAc configuration on the fragmentation results with Tyr as aglycone. The experiments were performed on an *Orbitrap Fusion™* mass spectrometer, online-coupled with a nano LC system (both *Thermo Scientific*) in a buffer C/D system. The glycopeptides were measured in separate experiments with a concentration of 400 fmol/15 µl each. The oxonium ion profile for each glycopeptide was extracted manually from the fragmentation experiment with the highest mother ion abundance at a given

retention time. Thereby, the oxonium ion abundances for each NCE value were normalized to the oxonium species with the highest abundance.

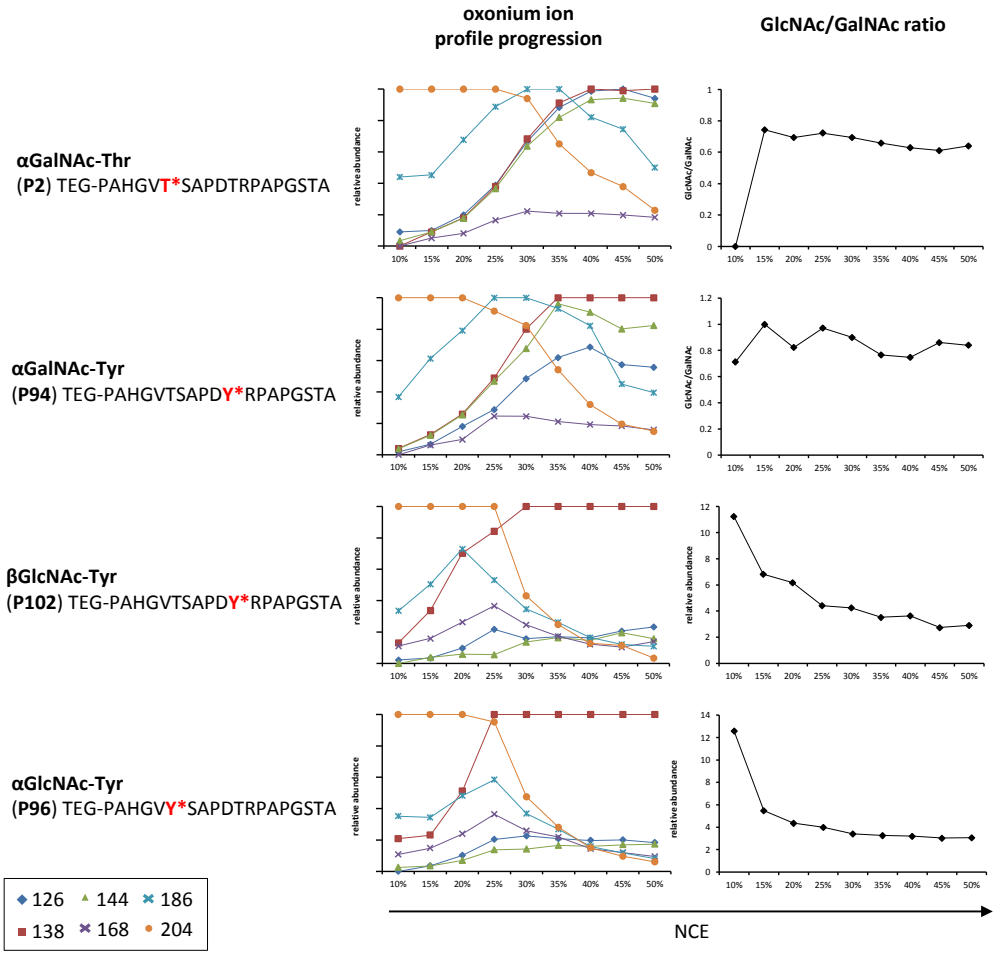


Figure 81: HCD-induced oxonium ion profiles and GlcNAc/GalNAc ratios at different NCE values.

The obtained fragmentation data was similar to earlier results on HexNAc-Thr/Ser glycopeptides. The oxonium profile progression generated from **P2** (αGalNAc-Thr) and **P94** (αGalNAc-Tyr) show a high similarity over the tested NCE interval (***Figure 81***). Also the profile progressions obtained from the *O*-GlcNAc peptides **P96** and **P102** show a high similarity between each other. As expected, the profile progressions obtained from the GalNAc peptides generally differ from those obtained from the GlcNAc peptides. This is in accordance with the postulated fragmentation pathways, which should not change without external influences. Thus it can be concluded, that the Tyr side chain does not influence the generated oxonium pattern. Accordingly, the obtained GlcNAc/GalNAc values are also in the isomer-typical magnitudes. The GlcNAc/GalNAc curves for **P2** and **P94** (GalNAc) reach a plateau located between 0.8 and 1, the curves of **P96** and **P102** reach a plateau between 3.0 and 4.0. The curve progressions clearly show a minimum NCE, which should not be undercut in order to obtain stable

GlcNAc/GalNAc values. In the present case this value is at approximately 30 % NCE. A significant difference in the ion profile progressions of the α- vs β-GlcNAc anomers was not observed.

The above described results indicate a portability of the results for Ser/Thr-*O*-glycosylation to Tyr-*O*-glycosylation, the influence of the peptide sequence remained unclear. In general, HCD is known and used for simultaneous glycan and peptide backbone fragmentation[285]. Thereby, lower collision energies typically lead to predominant glycan fragmentation, while higher collision energies can also fragment the peptide backbones[383-384]. Thus, it remained to be tested, whether peptide fragmentation interferes with the oxonium ion generation. Also here, a potential influence on HexNAc-Tyr fragmentation can be expected to be lower compared to HexNAc-Ser/Thr fragmentation, since the glycosidic bond in HexNAc-Tyr is less stable than in HexNAc-Ser/Thr. Therefore, additional tryptic glycopeptide sequences were subjected to HCD fragmentation. The experiments were again performed on the *Orbitrap Fusion™ Lumos™* mass spectrometer, online-coupled with a nano HPLC system (both *Thermo Scientific*) as described above. The data extraction was performed manually with Xcalibur. Therefore, the charge optima for maximal glycopeptide ionization were determined for each peptide sequence and the ion chromatogram was filtered for the corresponding m/z value. The oxonium ion abundances were extracted and normalized from the fragmentation experiments at the retention times with the three highest precursor ion abundances (***Figure 82***). Afterwards, an average spectrum including standard deviation was generated from three normalized spectra at a specific collision energy.

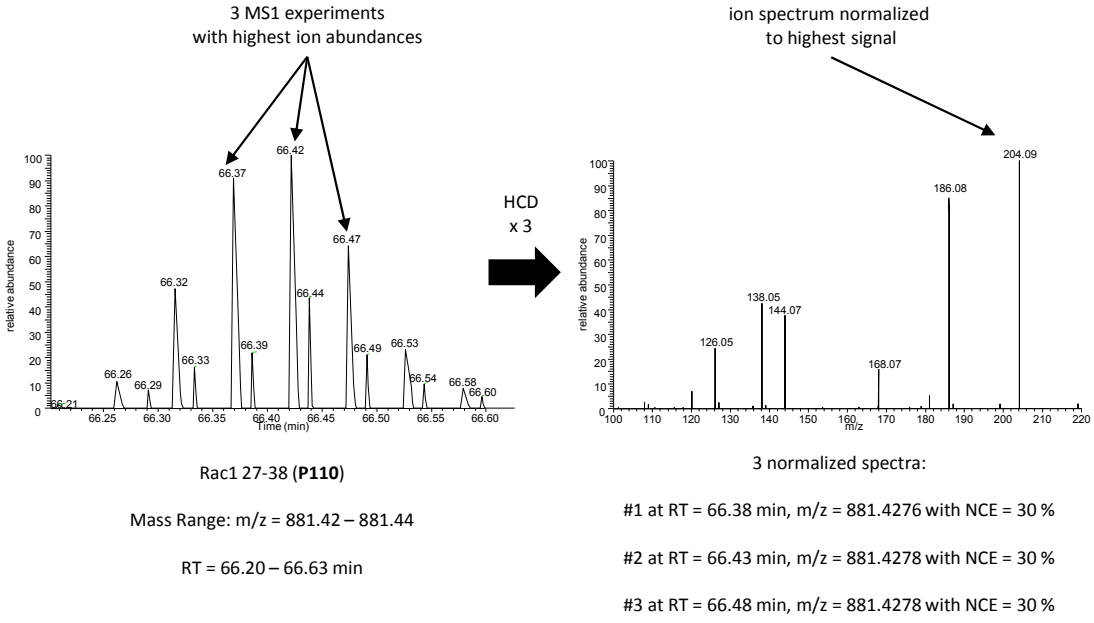


Figure 82: Data extraction from HCD fragmentation experiments on glycopeptide pools exemplified on the Rac1 αGalNAc variant **P110**.

The obtained oxonium ion profile progressions of all fragmented glycopeptides are characteristic for either GlcNAc or GalNAc and do not differ significantly with different peptide sequences (***Figure 83***). However, the ionization of the peptide backbone can influence the quality of the obtained data. While the signal intensities obtained from most glycopeptide precursors were in the magnitude of 10^5 to 10^6, the signal intensities of all amyloid-β glycoforms were in the magnitude of 10^4. Low intensities of precursor ions from the amyloid-β glycopeptides resulted in a very low amount of oxonium ions with problems to generate reliable profile progressions in this case.

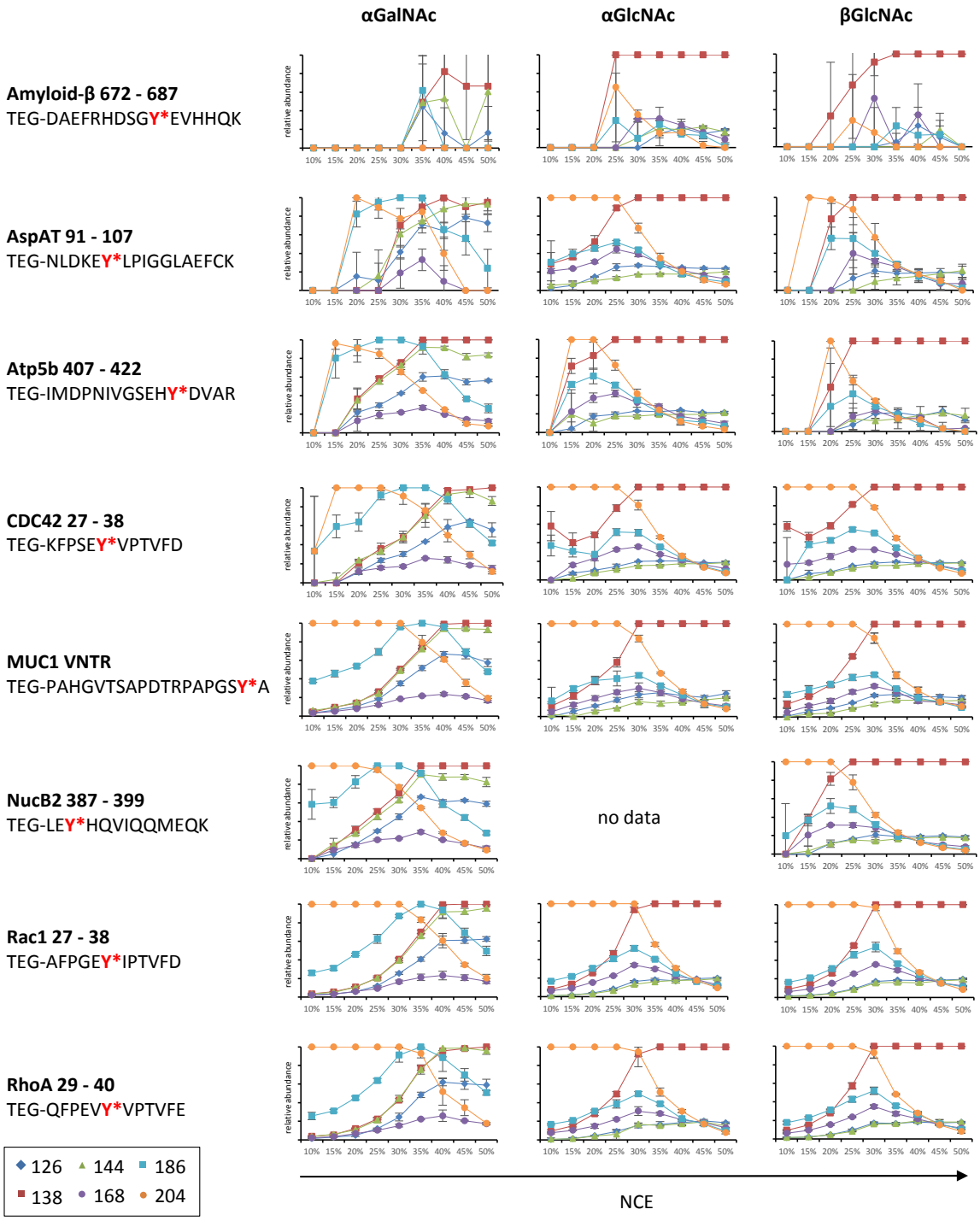


Figure 83: HCD-induced oxonium profile progressions for various glycopeptide sequences in different glycoforms.

For each glycopeptide fragmentation experiment the GlcNAc/GalcNAc ratio was calculated and the results were incorporated into HexNAc-epimer specific graphs as shown in ***Figure 84***. The individual curves of a given HexNAc-epimer glycopeptide pool showed huge variations at low collision energies, but the curves stabilized at increasing collision energy at HexNAc-isomer specific values. Stable

GlcNAc/GalNAc ratios were observed at NCE of 30 and 50 % with values of around 0.8 for GalNAc and 3-4 for GlcNAc. Like the general oxonium profile progression patterns (***Figure 83***), also the GlcNAc/GalNAc ratio was not influenced by the anomeric configuration of the fragmented glycan monomer. The GlcNAc/GalNAc curves of both GlcNAc anomers showed the same characteristic trend.

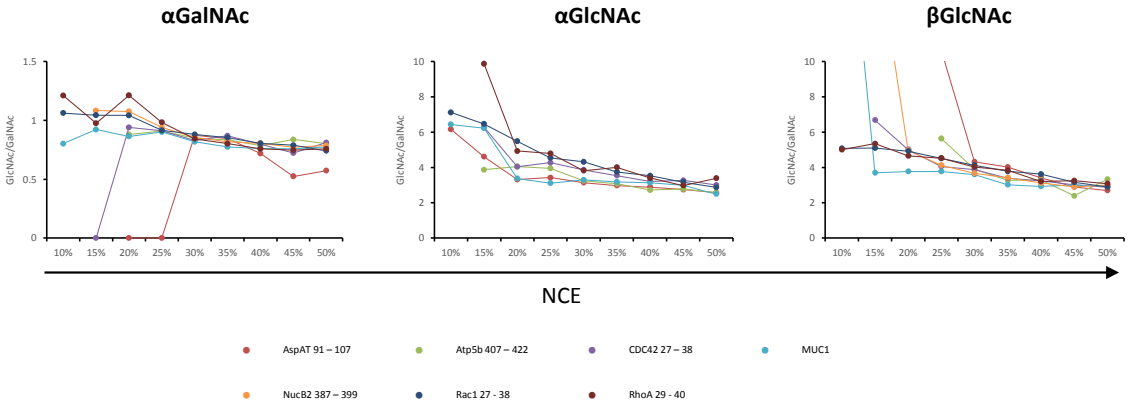


Figure 84: Overview of the calculated GlcNAc/GalNAc ratios. Amyloid-β is not included, since the high errors in the oxonium profile progressions do not allow the calculation of reasonable GlcNAc/GalNAc values.

In summary, the HCD-induced glycan fragmentation characteristics were explored using structurally defined synthetic HexNAc-Tyr glycopeptides. To this extent, MUC1 glycopeptide variants with HexNAc modifications on Thr and Tyr respectively were subjected to HCD-MS/MS and the generated oxonium ion signal profiles were recorded at 5 % NCE increments ranging from 10 % to 50 %. Here, no significant differences could be spotted between the obtained data for HexNAc-Thr and -Tyr. In order to proof the peptide sequence independence of this methodology, fragmentation data was further generated by elucidation of various synthetic Tyr glycopeptides presented as three different HexNAc glycoforms each: -αGalNAc, -αGlcNAc and -βGlcNAc. The obtained oxonium ion profiles showed that the methodology was applicable for MS-analysis of most glycopeptides without critical influence by the peptide backbone sequence. These profiles were assigned by the readout of HexNAc characteristic oxonium signal intensities and their evaluation with the help of the GlcNAc/GalNAc ratio and for optimal results a minimum of 30 % NCE was required. The obtained results will enable the integration of HexNAc-Tyr as a potential modification into large-scale glycoproteomic studies via oxonium ion profile evaluation. This is of particular interest, since to date there is only little information about the anomeric configuration and subcellular distribution of the different possible HexNAc-Tyr isomers.

7.2.8 β-GlcNAc-Tyr enzyme profiling

Identification of enzymes responsible for addition or removal of -βGlcNAc on Tyr on different peptide backbones would be the ultimate evidence to prove the existence of this new modification. Further, information regarding the βGlcNAc-Tyr cycling enzyme specificity and activity would be of great value. *O*-GlcNAcylation (modification with βGlcNAc) on Ser and Thr is known to often exhibit a fast turnover rate and it remains to be clarified whether this rule also applies to O-GlcNAcylation on Tyr. Further it is not clear[40, 372], if there are distinct consensus motifs or other peptide sequence properties, which contribute beneficial to the dynamic addition and removal of O-GlcNAc to tyrosine. To address these questions, the known *O*-GlcNAc cycling enzymes OGT[37] and its counterpart OGA[41] were incubated with potential peptide or glycopeptide substrates and the conversion was monitored via HPLC-MS.

7.2.8.1 O-GlcNAcase

The hydrolytical removal of *O*-GlcNAc from Ser and Thr is known to be mediated by only one single enzyme: The ubiquitously expressed *O*-GlcNAc hydrolase, commonly known as *O*-GlcNAcase (OGA) and a member of the CAZy-family GH84[385]. Two isoforms of human OGA have been identified: The long fOGA (full-length O-GlcNAcase), which is located in the cytoplasm and vOGA, a spliced variant of fOGA in the nucleus[386]. Among others, mechanistic studies on the bacterial homolog BtGH84[387] revealed, that this hydrolysis proceeds in a substrate-assisted manner under retention of the anomeric configuration.

Due to the lack of other βGlcNAc-Thr/Ser specific hydrolases and the lack of a defined recognition motif of OGA it can be considered as likely, that OGA is also responsible for βGlcNAc removal from Tyr glycopeptides. We were interested to verify this general assumption and a positive result would raise several additional questions. Taking the structural difference of Tyr and Ser/Thr into account, differences in substrate affinities and catalytic efficiencies of the enzyme towards these different glycosylated amino acids would be expected. Further it is unclear, if the selectivity for βGlcNAc anomer on Ser/Thr also applies for the same modification on Tyr. To verify this assumption different synthetic *O*-GlcNAcylated peptide substrates were prepared and incubated with OGA, while monitoring the glycopeptide conversion over different time points.

All reactions were performed at 37° C in 50 µl of a 20 mM HEPES buffer (*p*H 7.4)[388] with a glycopeptide concentration of 1 mM each. In initial experiments the *O*-GlcNAcase from *Streptococcus pyogenes* (*Prozomix* Ltd., specific activity: 5.56 U/mg at 37° C; *p*H 7.6; pNP-N-acetyl-β-D-glucosaminide (1 mM)) was utilized in an effective concentration of 25 mU/50 µl (0.5 U/ml) and after initial test reactions with

1, 2.5 and 5 mU per 50 µl. Samples of 2 µl were taken prior to enzyme addition as well as 1, 4, 8 and 24 h after enzyme addition each and subjected to HPLC-MS. Here, the integrals of the respective glycopeptide and product peptide peaks allowed quantitative monitoring of the reaction progression. An example of a corresponding HPLC-MS chromatogram series is shown in ***Figure 85a*** with the conversion of the βGlcNAc-Tyr MUC1 peptide **P102** with the sequence TEG-PAHGVTSAPDY(βGlcNAc)RPAPGSTA. Here the shrinking of the substrate peak indicated the ongoing consumption of the glycopeptide, while the growing product peak indicated product formation over time. In the sample taken after an incubation time of 24 h the substrate was completely consumed as indicated by the sole presence of the product peak at 19.4 min. These observations were supported by the recorded mass signals at m/z = 753.56 for the substrate peak and m/z = 686.21 for the product peak which correspond to the triple charged substrate and product ions respectively. Furthermore, the slight retention time shift of about + 1 min is characteristic for the loss of one HexNAc unit. After quantification of the conversion via peak integration, it was calculated as the quotient from product peak area over the combined substrate and product area. The resulting graph for the conversion of glycopeptide **P102** is shown in ***Figure 85b***.

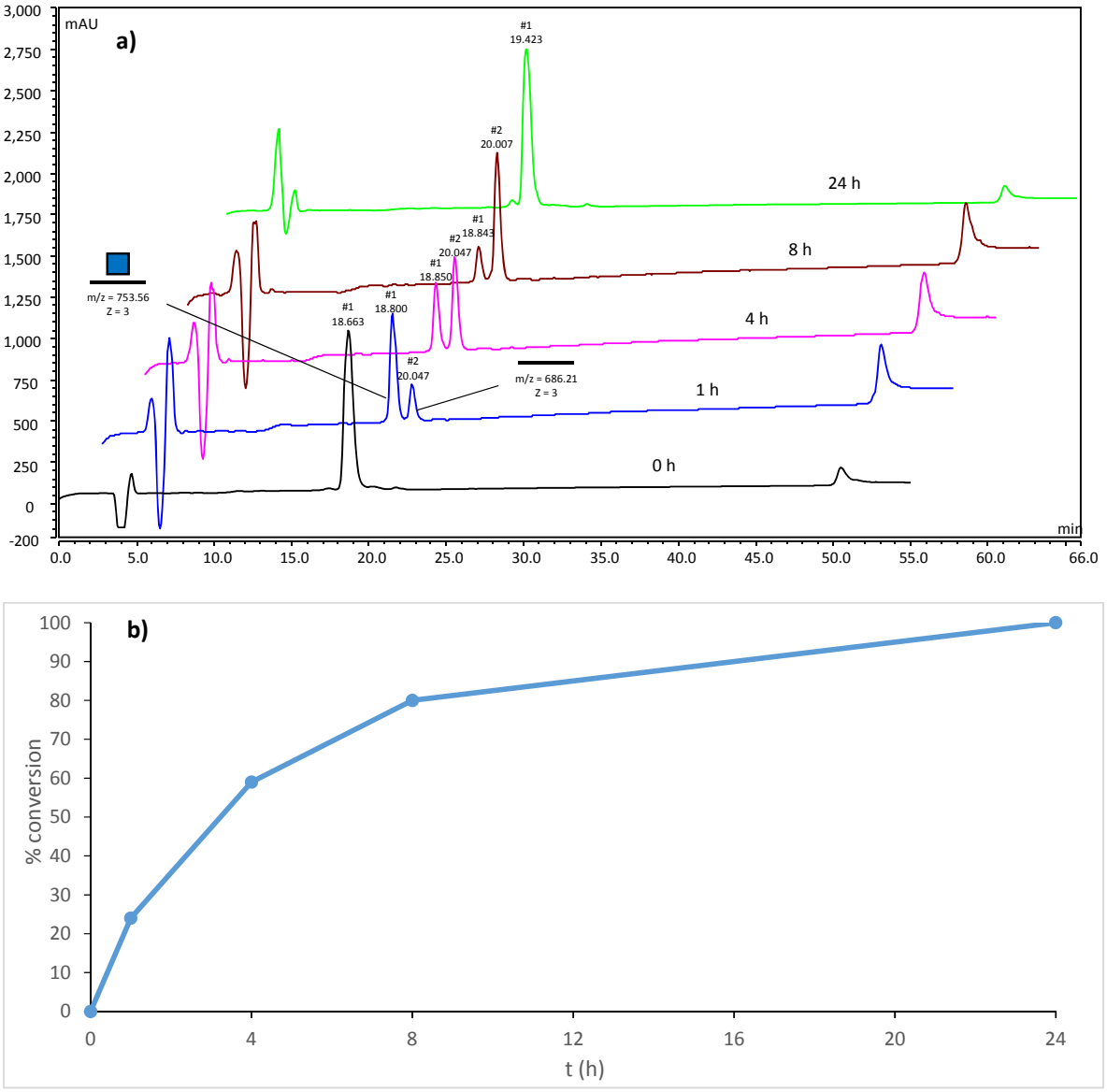


Figure 85: Evaluation of glycopeptide substrates for OGA exemplified on the conversion of the βGlcNAc-Tyr MUC1 peptide **P102**. a) Monitoring of substrate conversion via HPLC-MS at different time points throughout the reaction. b) Visualized substrate conversion over time.

A first series of experiments was designed, to verify the general conversion of HexNAc-Tyr by OGA. To this extent a set of monoglycosylated MUC1 VNTR analogs was subjected to individual enzyme incubation reactions. The substrate candidates were selected to cover βGlcNAc-Tyr, -Ser and -Thr analogs with substitutions of natural Thr glycosylation sites. The application of sequence analogs enabled the deduction of substrate preferences based on specific local differences without influences of the remaining peptide sequence. For Ser/Thr O-glycosylation, OGA is known to be selective for the βGlcNAc anomer solely, to know if this selectivity also applies to Tyr, αGlcNAc- and αGalNAc-Tyr

analogs were elucidated in addition to βGlcNAc. An overview of the experimental results is shown in ***Table 3***.

Table 3: MUC1 substrate candidates for OGA.

Substrate candidate		Conversion after 24 h		
fragment	sequence	βGlcNAc	αGlcNAc	αGalNAc
MUC1 VNTR	TEG-PAHGVTS*APDSRPAPGSTA	0 %		
	TEG-PAHGVTSAPDS*RPAPGSTA	0 %		
	TEG-PAHGVTSAPDSRPAPGS*TA	100 %		
MUC1 VNTR	TEG-PAHGVT*SAPDTRPAPGSTA	0 %		
	TEG-PAHGVTSAPDT*RPAPGSTA	0 %		
	TEG-PAHGVTSAPDTRPAPGST*A	0 %		
MUC1 VNTR	TEG-PAHGVY*SAPDTRPAPGSTA	97 %	0 %	
	TEG-PAHGVTSAPDY*RPAPGSTA	100 %	0 %	0 %
	TEG-PAHGVTSAPDTRPAPGSY*A	100 %	0 %	0 %

These results indicated, that the β-selectivity of OGA also applies for GlycNAc-hydrolysis on Tyr and no conversion was observed for the tested αGlcNAc and αGalNAc modified peptides. Furthermore and unexpectedly, also the majority of tested MUC1 βGlcNAc-Ser and -Thr analogs showed no conversion with the exception of βGlcNAc-Ser modification in the GSTA site. Obviously glycosylation with Ser or Thr on MUC1 VNTR glycosylation sites were not optimal for OGA βGlcNAc hydrolysis even if βGlcNAc-Ser and -Thr usually are natural substrates. Interestingly, βGlcNAc-Tyr OGA mediated hydrolysis was highly efficient on the MUC1 VNTR analogs. Comparison of the substrate conversions over time suggests, that the C-terminal PGSXA-site is a better substrate, than the GVXSAP and PDXRP sites and both, the PGSY*A and the PGS*TA glycopeptides were deglycosylated at a higher rate, than the remaining two Tyr analogs (***Figure 86***).

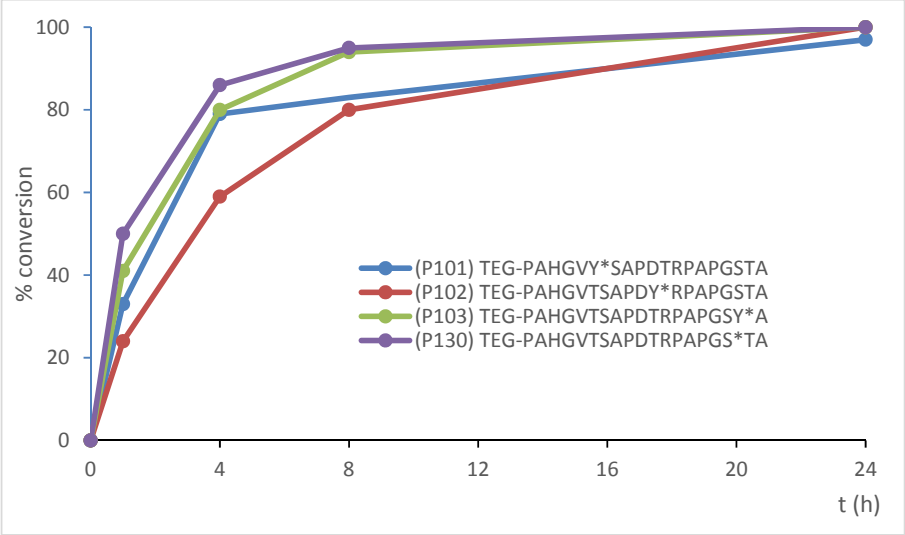


Figure 86: Substrate conversions of the recognized MUC1 VNTR variants.

The OGA experiments with MUC1 VNTR model glycopeptides were followed-up by additional substrate studies of synthetic glycopeptide fragments which contained previously identified βGlcNAc sites or analogs thereof. To examine the influence of divalent glycan presentation, additional diglycosylated peptide analogs were included to the substrate analysis pool. The OGA reaction conditions described above for the MUC1 VNTR peptides were used and samples were taken after 1, 4 and 24 h respectively. In summary the results, including the MUC1 peptide analogs, showed that from 31 monoglycosylated substrate candidates 19 were hydrolyzed by OGA (**Figure 87**). All of the 12 tested βGlcNAc-Tyr candidates showed complete or partly conversion after 24 h, while only 7 out of 12 tested βGlcNAc-Ser candidates were deglycosylated. In contrast none of the 6 tested βGlcNAc-Thr analogs showed conversion.

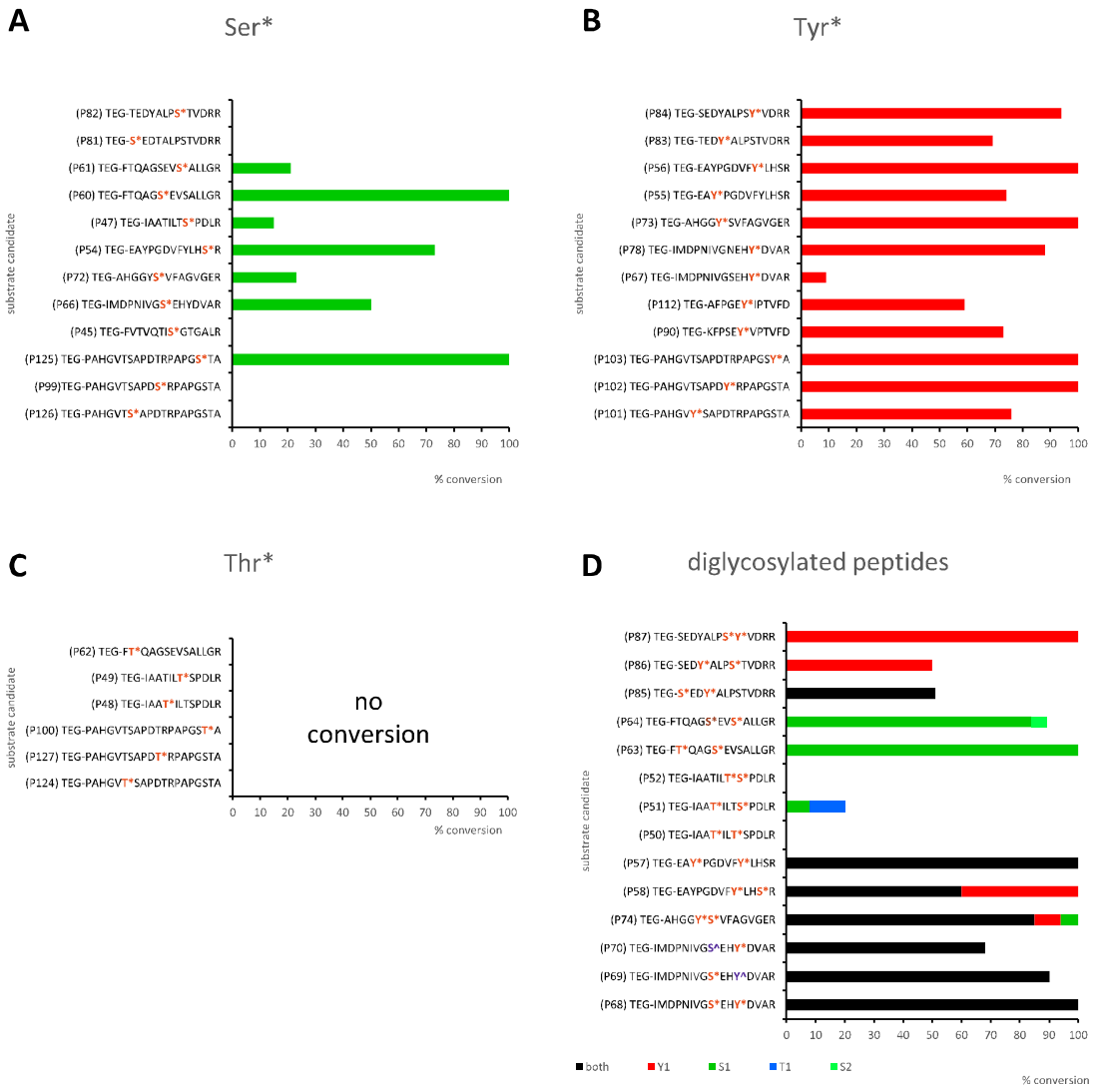


Figure 87: Conversions of OGA substrate candidates after 24 h. The substrate candidates are arranged according to the amino acid acceptor: **A)** serine **B)** tyrosine and **C)** threonine. **D)** divalent glycopeptides and mixed glyco-/phosphopeptides with various amino acid acceptors. * = β-GlcNAc, ^ = phospho.

Focusing on the Ser analogs, we observe a clear dependence of glycosylation site and peptide backbone for OGA activity, additional molecular modeling is required to better understand these results. Earlier X-ray crystallography studies on OGA/p53 complexes revealed a peptide backbone recognition by hydrophobic interactions between enzyme and substrate[44, 389]. The substrate binding cleft of the enzyme comprises for the most part conserved hydrophobic residues like tyrosine, phenylalanine or valine, which leads to π-stacking interactions with the substrate. According to X-ray crystallography on enzyme/substrate complexes, different substrates adopt similar conformations within the glycosylation site area ranging from -2 to +1 with a maximum Cα-shift of 1.2 Å[389].

In comparison, the enzyme showed a very different behavior towards Tyr substrate analogs and in general, incubation with these glycopeptides lead to throughout higher conversion results. All tested glycopeptides were hydrolyzed by the enzyme and 11 out of 12 glycopeptides showed a conversion of over 50 % after 24 h. The preference of GlcNAc-Tyr over GlcNAc-Ser substrates are illustrated by comparison of the MUC1 VNTR analogs **P99**, **P102** and **P127** (***Table 4***).

Table 4: Influence of glycosylated amino acids on glycopeptide recognition by OGA.

	substrate candidate	conversion after 24 h
P99	TEG-PAHGVTSAPDS*RPAPGSTA	0 %
P102	TEG-PAHGVTSAPDY*RPAPGSTA	100 %
P127	TEG-PAHGVTSAPDT*RPAPGSTA	0 %

For βGlcNAc-Tyr no clear influence of the peptide backbone amino acids were observed. A reason for this might be found in the different sterical geometry of the Tyr residue if compared to Ser or Thr in combination with the enzyme geometry. O-GlcNAcase is a homodimer of two sister monomers containing one catalytic domain and one stalk domain each. The catalytic domain consists of a $(\beta/\alpha)_8$-barrel (eight-stranded parallel β-sheet core surrounded by eight α-helices), which forms a deep pocket for glycan binding (***Figure 88*** **a**)[44]. Here, a tight binding of the GlcNAc residue is facilitated by an extensive number of hydrogen bonds within the catalytic core. Further, each catalytic domain is covered by the stalk domain of its sister monomer, creating a substrate binding cleft (***Figure 88*** **b**). This binding cleft comprises for the main part hydrophobic residues like tyrosine, phenylalanine, valine, methionine and tryptophan. It has been shown, that – among general hydrophobic interactions - π-stacking between these residues and the glycopeptide contributes to substrate binding. Here, the glycosylated Tyr might be directly involved in π-stacking interactions, which thus provides hydrophobic interaction with the substrate binding cleft without the need of additional adjacent hydrophobic amino acids. In addition, the spatial distance between Cα and the side chain hydroxyl might play a role. This distance ranges between 2.1 and 3.1 Å in Ser and Thr and is tripled in Tyr with 6.4 – 7.4 Å (Chemdraw calculation). This results in that the saccharide is positioned at a larger distance from the peptide backbone than in the common Ser/Thr case, which probably is critical to facilitate access of the GlcNAc-residue into the deep substrate binding/catalytical cleft without disturbance of adjacent amino acids.

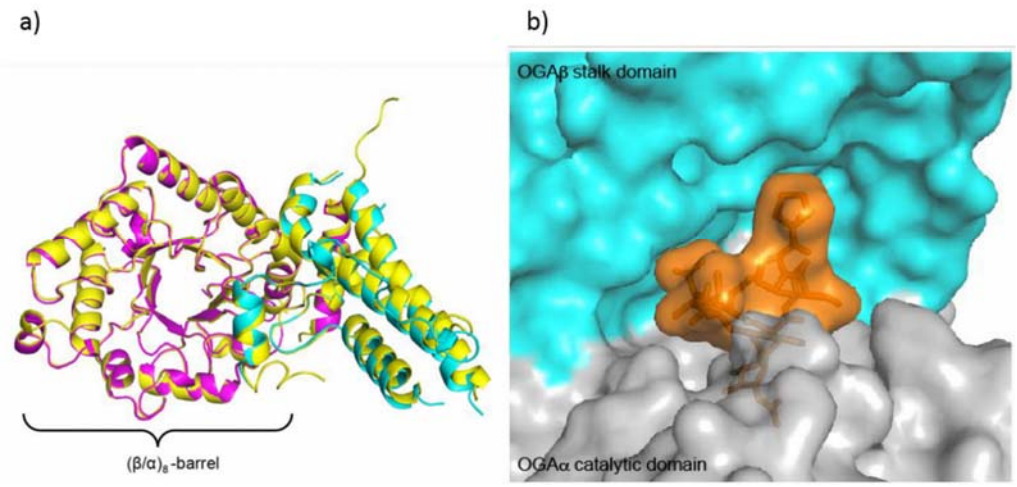


Figure 88: a) Sister monomer of OGA with catalytic domain (purple) and stalk domain (blue). b) X-ray structure of the OGA binding cleft in complex with a p53 glycopeptide.[44]

Divalent βGlcNAc presentation at the peptide backbone resulted in increased glycan hydrolysis. A good example for this effect is the conversion of the Tyr and Ser analogs of Atp5b 226-239 (mouse) **P73** and **P72** and the corresponding diglycosylated peptide **P74 (*Figure 89*a)**. The monoglycosylated Tyr peptide **P73** (TEG-AHGGY*SVFAGVGER) showed full conversion after 24 h and the serine analog **P72** (TEG-AHGGYS*VFAGVGER) only 23 %. In comparison **P74** (TEG-AHGGY*S*VFAGVGER) showed an enhanced hydrolysis of serine, which lead to 85 % of doubly deglycosylated peptide.

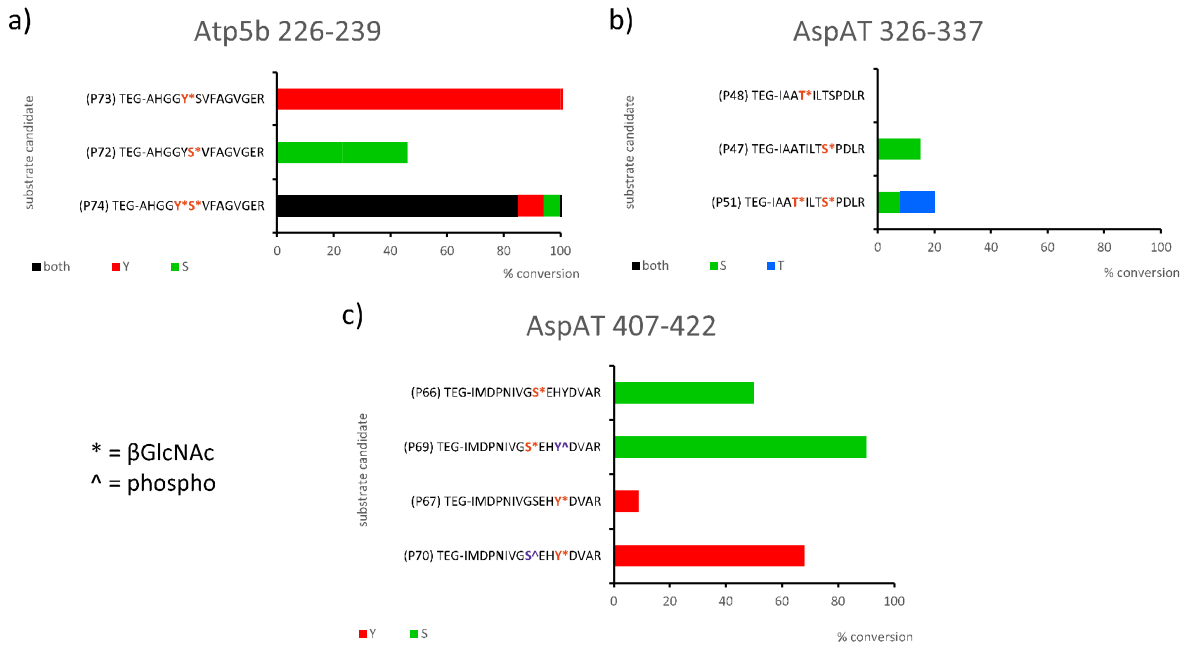


Figure 89: Influence of modifications in proximity to the glycosylation site. Phosphorylated amino acids are marked as X^. a) Amplification of serine deglycosylation by neighboring tyrosine glycosylation. b) Threonine deglycosylation enabled by nearby serine glycosylation. c) Amplifying effect of nearby phosphorylation.

The amplifying effect of divalent substrates even lead to the only case of GlcNAc-Thr hydrolysis among all tested substrate candidates as visualized in the comparison of Ser and Thr analogs of AspAT 326-337 (human/rat) in ***Figure 89*** **b**. The Thr analog **P48** (TEG-IAAT*ILTSPDLR) showed no conversion, the Ser analog **P47** (TEG-IAATILTS*PDLR) showed 15 % conversion after 24 h and the diglycosylated analog **P51** (TEG-IAAT*ILTS*PDLR) showed deglycosylation of either Ser or Thr. Interestingly, the divalent substrate showed a 20 % conversion, 8 % account for hydrolysis on Ser and 12 % on Thr.

In order to uncover a potential influence of nearby phosphorylation, mixed glyco/phospho-peptides were included to the list of OGA substrate candidates. The peptide scaffold Atp5b 407-422 (human) was chosen, as it contains a site for competing βGlcNAcylation and phosphorylation on Ser415[66, 390] as well as a Tyr in close proximity. Here, nearby phosphorylation resulted in a dramatic increase of substrate conversion by OGA if compared to the non-phosphorylated peptide variants (***Figure 89*** **c**). While the conversion of the Ser analog **P66** was amplified from 50 % to 90 % if the Tyr in +3 position was phosphorylated, the Tyr analog **P67** was amplified by a factor of 7.5 from 9 % to 68 % with a phosphoSer at -3.

Summarized, a clear discrimination of βGlcNAc-Tyr from its epimers αGlcNAc- and αGalNAc-Tyr by OGA was observed. Further, a strong correlation between glycosylated amino acid and conversion rate is visible with the trend Tyr > Ser > Thr. Thereby the recognition of βGlcNAc-Ser (and probably also the structurally similar -Thr) by OGA is strongly dependent on the polarity of the downstream

amino acids at -1 and -2 position with respect to the glycosylation site. In contrast, hydrolysis of βGlcNAc-Tyr commences with almost total disregard to this rule. A general amplification of the conversion was observed for substrate candidates with glycosylation sites in close proximity, enabling even deglycosylation on amino acids, where no conversion was observed with the monoglycosylated variant. This amplification effect was also observed, if a phosphorylated residue is present near the glycosylation site, which represents a further example of regulatory crosstalk between phosphorylation and βGlcNAcylation. In order to complete the data obtained from the already carried out OGA profiling experiments, mixed glycopeptide candidates need to be synthesized. Additionally, computer assisted docking studies might provide a more precise understanding of the enzyme/substrate interactions and the influence of Tyr and proximal modifications, a collaboration with Prof. Rob Woods (complex carbohydrate research center, Athens, Georgia, USA) is existing on this topic. The influence of Tyr glycosylation on the conformation of glycopeptides via STD-NMR is under investigation in cooperation with Dr. *Filipa Marcelo* (Universidade Nova de Lisboa Quinta da Torre, Portugal). Knowledge about the conformational impact will help to understand the recognition of Tyr glycosylation by OGA as well as other lectins and enzymes.

7.2.8.2 Human O-GlcNAcase

With the results from the OGA profiling experiments at hand, selected glycopeptide substrate candidates were incubated with its human homolog hOGA. Aim of these experiments was to find out, if the trends and correlations for OGA substrate recognition also apply for hOGA. The enzyme was expressed and kindly provided by Prof. Dr. *Jennifer Kohler* (Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas). The specific activity of the enzyme was given as 0.17 mU/mg at room temperature and *p*H 6.5 against 1 mM pNP-N-acetyl-β-D-glucosaminide in a 50 mM sodium cacodylate buffer including 5 mg/ml BSA and 4 mM GalNAc. In previous enzyme specificity studies by the Kohler group, different enzyme concentrations were applied for GlcNAc-hydrolysis with hOGA[391]. Due to limited amounts of enzyme and a rather dilute stock solution as starting point, the minimal required hOGA enzyme concentration needed to be further optimized using the MUC1 VNTR substrate candidate **P103** (TEG-PAHGVTSAPDTRPAPGSY*A) as model compound. This glycopeptide was selected since it has shown full conversion after 24 h using bacterial OGA (compare chapter 7.2.8.1). In addition to a batch with 50 nM hOGA, a second batch with higher enzyme concentration, 250 nM, was prepared. The reactions were, analog to the OGA experiments, set up in a total volume of 50 µl each with a glycopeptide concentration of 1 mM in sodium cacodylate buffer.

The substrate conversions were monitored according to the procedure used for the OGA experiments. After 24 h the batch with 250 nM (247 µU) hOGA showed 46 % conversion. In contrast, no conversion could be observed in the 50 nM (49 µU) batch. The fact that the same glycopeptide showed full conversion if incubated with 25 mU OGA instead, suggests that lower amounts of enzyme than the here applied 247 µU hOGA should not be used for reliable results.

However, since the conversion of **P103** at 250 nM hOGA was below the conversion observed with bacterial OGA, the enzyme concentration was raised to 388 nM in follow-up experiments. As substrate candidates the following Tyr glycopeptides were included: MUC1 VNTR **P103** (TEG-PAHGVTSAPDTRPAPGSY*A) and **P125** (TEG-PAHGVTSAPDTRPAPGS*TA) as well as Atp5b 407-422 (mouse) **P78** (TEG-IMDPNIVGNEHY*DVAR). **P103** represented a substrate candidate with a polar amino acid at -1, it also allowed direct comparison with the preceding hOGA incubations. In contrast, **P78** exhibited a polar amino acid (glutamic acid) at -2 with respect to the glycosylation site. **P125** represented an example with corresponding Ser βGlcNAcylation without polar amino acids on either -1 or -2. This substrate candidate was selected as it showed full conversion in earlier incubations with bacterial OGA. The observed conversions after 24 h are shown in ***Figure 90***.

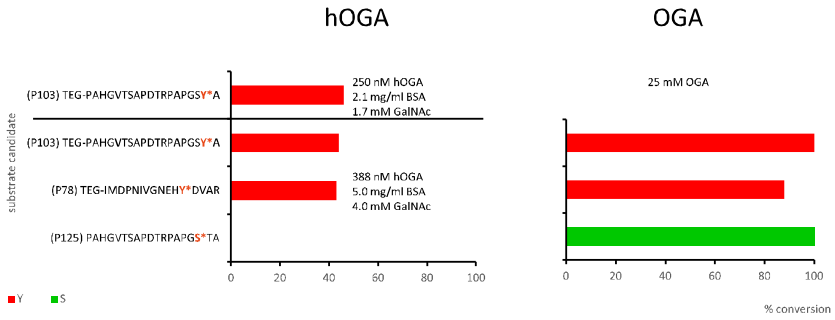


Figure 90: Overview about the conversions of all tested substrate candidates for hOGA.

After 24h incubation time **P103** showed 46 % conversion with 250 nM hOGA, 2.1 mg/ml BSA and 1.7 mM GalNAc, the same glycopeptide showed 44 % conversion at 388 nM hOGA, 5.0 mg/ml BSA and 4.0 mM GalNAc. **P78** showed conversion in the same magnitude with 43 %, while it showed almost full conversion by using bacterial OGA, just like **P103**. In contrast, the MUC1 VNTR Ser analog **P125** showed no conversion, other than the corresponding bacterial OGA experiment with 100 % conversion. From the results it can be concluded, that both tested Tyr analogs showed conversion with hOGA in contrast to the Ser analog, which corresponds well with the general trend for bacterial OGA substrates, which by far show better conversion if carrying a βGlcNAc-Tyr residue. Also just like with bacterial OGA, the βGlcNAc-Tyr substrates were not inhibited by polar amino acids at -1 or -2 with respect to the glycosylation site. Interestingly, the Ser analog showed no conversion, although it was expected based on the studies on the bacterial OGA homolog. However, much lower enzyme

concentrations were used by the hOGA incubations, which may influence the differences in hydrolysis of the Ser analog and further experiments are needed to draw conclusions on these very preliminary results.

7.2.8.3 Human O-GlcNAc transferase

While OGA catalyzes the hydrolytic removal of βGlcNAc from Ser, Thr and Tyr, *O*-GlcNAc transferase (OGT) represents its counterpart by facilitating the glycosylation of these amino acids under consumption of a uridine diphosphate N-acetylglucosamine (UDP-GlcNAc). Although a number of glycoproteomic studies[392-394] were performed trying to identify OGT substrates, no clear consensus motif has been identified. In 2015 *Pathak* and *van Aalten* measured the hOGT activity against a library of 720 peptides derived from the human proteome of which only 70 were glycosylated by the enzyme[372]. They identified a sequence for hOGT recognition: [TS][PT][VT]S/T*[RLV][ASY], with an amino acid probability cutoff of 0.5. From additional X-ray structural analyses of enzyme/substrate complexes they concluded, that the peptide substrates are bound in a common binding mode, constituting of a mixture of van der Waals interactions in the -3 to -1 subsites and hydrogen bonding with the substrate backbone in the +1 and +2 subsites. For the Ser and Thr in the 0 position, they postulate a high rigidity with hydrogen bonds between backbone amide and amino acid residue respectively to the donor substrate UDP-GlcNAc. In general they observed a higher affinity towards serine glycosylation compared to threonine glycosylation. Of particular benefit is a proline in -2 and -3, as it stabilizes the extended conformation of the peptide substrate, while being unpolar and thus participating in van der Waals interactions.

The performed experiments were only taking Ser and Thr residues into account, in this work we were interested to explore if OGT was capable to perform glycosylation on Tyr residues as well. To this extent unglycosylated peptide substrate candidates were incubated with hOGT and in the presence of UDP-GlcNAc donor. The selected candidates represent fragments of proteins, which have been identified to be naturally glycosylated on Tyr or contain peptide backbones with previously identified nearby GlcNAc-Ser/Thr sites. The recombinant enzyme was kindly provided by Prof. Dr. *Jennifer Kohler* (Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas). In the first set of experiments Atp5b 407-422 from mouse **P75** (TEG-IMDPNIVGNEHYDVAR), its human homolog **P65** (TEG-IMDPNIVGSEHYDVAR) as well as AspAT 91-107 from human and mouse **P40** (TEG-NLDKEYLPIGGLAEFCK) were incubated. All three sequences exhibit one Tyr, **P65** contained one additional Ser, which might serve as competing glycosylation site. The reaction conditions were set up according to reported hOGT incubations[391] in a Tris-HCl buffer including 500 µM UDP-GlcNAc at *p*H

8.0. Each reaction was done in a total volume of 50 µl with 0.1 mM peptide substrate and 1 µM enzyme at 37° C. After 24 h a sample of 20 µl was subjected to HPLC-ESI-MS to evaluate the reaction outcome. Surprisingly, neither the Atp5b peptides **P75** and **P65** nor the AspAT sequence **P40** showed any conversion. The obtained results may have two possible explanations: Either the enzyme is not responsible for Tyr-glycosylation in general or GlcNAcylation on tyrosine is site dependent and influenced by nearby amino acid residues. In the tested peptide sequences all Tyr residues exhibit polar amino acids in the -1 to -3 region, while according to *Pathak* and *van Aalten*[372] hOGT/substrate recognition on Ser/Thr peptides includes hydrophobic interactions of the catalytic core with amino acids in the -1 to -3 region. In contrast, the nearby Ser415 site to Tyr418 in **P65** (Atp5b 407-422) was also not modified in spite of that it showed favored amino acid environment (…IVGsEHY…) and that Ser415 has been reported[66] as GlcNAc site. Further, to date OGT is the only known glycosyl transferase responsible for *O*-GlcNAcylation and the peptide would be expected to be recognized by the enzyme. However, many unexplored parameters may influence the OGT activity as for instance enzyme/substrate conc., pH, impact of amino acids further apart from the glycosylation, conformation or glycosylation at other sites.

As no activity could be observed within the tested peptides, not even for those which were expected to be recognized, a further set of glycosylation experiments was set up to test the functionality of the enzyme and the reaction conditions. Here, peptide sequences were applied, which were found previously to be glycosylated by hOGT very efficiently and represent an optimized constructs for hOGT recognition: Human Rbl2 411-422 **P123** (Ac-KENPAVTPVSTA) and **P120** (KKVPVSRA)[372]. Both peptides showed full conversion after 24 h, thus proving the functionality of the expressed enzyme.

After showing, that the enzyme was active to these sequences, a new set of peptide/enzyme incubations was done, using sequences of known βGlcNAc peptides. Due to the low amount of applied enzyme and the possibility of slow conversions, the reaction time was increased from 24 h (preceding experiments) to 72 h. Samples of 15 µl were taken after 4 h, 24 h and 72 h each. Since some peptides included several potential glycosylation sites (Ser, Thr, Tyr), a product mixture might be expected for recognized peptide candidates. This, in combination with the low peptide amount per batch (5 nmol total per batch, 1.5 nmol per sample injection), resulted in very small chromatogram peak areas and the utilized low resolution ESI-MS instrument (*MSQ Plus*, Thermo Scientific) was not sensitive enough to obtain mass data for such small amounts, the method was modified to adopt to these challenges. Instead of the linear LC-ESI-MS setup used in the preceding experiments, the resulting chromatogram peaks were fractionated with subsequent analysis by MALDI-TOF instead. In addition, this allowed a change of the HPLC running solvent from FA (optimized for ionization in ESI-MS) to the known ion suppressor TFA. As a result the peak widths of the corresponding peptides were reduced, resulting in

better peak separation. All experiments were done using 0.1 mM peptide, 200 µM UDP-GlcNAc, 1 µM hOGT and *p*H 8.0 in a total volume of 50 µl. According to the collected mass data, only the reference peptides **P115** and **P118** were glycosylated by hOGA. Instead, the analyzed fractions of all other samples contained either m/z values of the unmodified substrate candidates or of peptide fragments, which corresponds to a high protease activity during the incubations, probably due to endopeptidase impurities of the hOGT that was used. In future experiments to explore hOGT activities protease inhibitors will be included during incubations.

8 Conclusion

Protein *O*-glycosylation takes place on serine (Ser), threonine (Thr) or tyrosine (Tyr) and are commonly initiated by glycosylation with *N*-acetylgalactosamine (α-GalNAc) or *N*-acetylglucosamine (β-GlcNAc) residues. The two modifications differ in structure, complexity, localization as well as biological functions. The mucin type *O*-glycosylation (α-GalNAc linked) presenting short or more complex linear or branched glycans is the dominant modification found on mucin glycoproteins and is common on other proteins that are localized in the extracellular environment either membrane bound or secreted. The Mucin type *O*-glycans at the cell-surface are typically involved in cell-cell and cell-matrix interactions and are subdivided to eight known core structures, of which core 1-4 are the most common. The core structures can be enzymatically elongated with alternating β-GlcNAc and β-Gal residues and thereby form extended N-acetyllactosamine (LacNAc) structures. Terminal sialylation, fucosylation or sulfation on LacNAc or on shorter core structures further enhances the mucin-type glycan diversity. In contrast, the β-GlcNAc modification is commonly found as a monosaccharide modification and is localized in an intracellular environment. *O*-GlcNAcylation is introduced by the single enzyme *O*-GlcNAc transferase (OGT) and removed by its counterpart O-GlcNAcase (OGA), which both are ubiquitously expressed and distributed in an intracellular environment throughout most mammalian tissues. β-GlcNAc is a nutrient responsive modification and is involved in cellular processes including signaling, gene transcription and proteasomal degradation. *O*-GlcNAcylation seems to be connected to phosphorylation, as both are intracellular modifications found at nearby or competing for the same sites on Ser, Thr and Tyr residues. An interplay of the two modifications is evident and it is believed that O-GlcNAcylation sometimes inhibit phosphorylation by covalently block the phosphorylation sites. Even the enzymes responsible for these modifications, OGA, OGT as well as many kinases and phosphatases, are regulated by phosphorylation or β-GlcNAc-modifications. This interplay is of particular relevance, as it is hypothized, that a decreased glucose metabolism in aging neurons leads to reduced GlcNAcylation of central proteins. In consequence, this may result in abnormal phosphorylation of the affected proteins and ultimately to progression of neurodegenerative diseases. In contrast, high glucosamine levels (hyperglycemia) and concomitant hyper-GlcNAcylation of transcription factors or signaling proteins are suspected to be responsible for the acquired insulin resistance in diabetes type-II.

8.1 Project 1: Mucin glycosylation in cancer and airway diseases

Epithelial surfaces of for instance the respiratory tract, the gastrointestinal or genital tracts are coated with a protective barrier consisting of secreted or membrane-bound mucin glycoproteins. Mucins exhibit an *N*-terminal *variable number of tandem repeats* (VNTRs), which are heavily *O*-glycosylated, a membrane spanning domain and a C-terminal cytoplasmic signaling domain.

The MUC1 glycoprotein has been found to play an important role in the adhesion of pathogens, immune cells as well as tumor cells (metastasis). Chronic airway diseases like COPD, cystic fibrosis or asthma are characterized by changes in mucin *O*-glycosylation including changes in sialylation, fucosylation and sulfation and bacterial and virus infections induce a permanent mucin hypersecretion, triggering exacerbations. The most prominent pathogens involved in airway disease-related infections are *Pseudomonas aeruginosa* (cystic fibrosis), *Haemophilus influenzae* (COPD) *Streptococcus pneumoniae* (asthma) and adenoviruses (COPD). Infections are promoted by cell-surface receptors (adhesins), which exhibit selectivity for mucin glycan structures. Further, bacterial infections are supported by bacterial sialidases, which reveal adhesion binding epitopes within mucin glycans and provide the pathogens with sialic acid for nutrition or as mimicry structures. The adhesin and sialidase specificities are of interest for diagnostic application and knowledge of glycan binding specificity would further enable development of glycomimetic antiadhesive drugs.

In epithelial carcinomas, mucins (in particular MUC1) are heavily overexpressed and often exhibit tumor-specific glycopeptide epitopes consisting of highly sialylated short saccharides. Due to the involvement of mucins in cell-cell, cell-matrix interactions, cell-recognition and signaling events, tumor-specific changes in *O*-glycosylation contribute to tumor proliferation, growth and metastasis. The formation of unique short saccharides and the exposure of the normally covered peptide backbone on tumor cells make these tandem repeat structures on mucins promising as targets for cancer immunotherapy. Another protein involved in tumor progression is the β-galactose-specific lectin galectin-3. The interplay of galectin-3 with tumor-specific mucin glycans is believed to be important in metastasis-mechanisms. In this context, a profound knowledge of the fine-specificity of galectin-3 towards tumor-associated glycans presented on the natural mucin tandem repeat peptide backbone would enable better understanding of these mechanisms. Further, better knowledge of galectin-3 specificity may support the development of new glycomimetic inhibitors.

This work aimed to explore the specificities of lectins, which play key-roles in tumor progression and pulmonary disease-related inflammations. To this extent, the microarray technology was used as it enables a multitude of parallel binding experiments with minimal consumption of precious analyte material. To date, the isolation of mucin-type glycopeptides from biological material still poses a

problem due to their heterogenic nature. To circumvent this limitation, a synthetic well-defined MUC1 glycopeptide library was built-up for microarray lectin/carbohydrate binding studies. Further, the glycopeptide library was used to evaluate specificity of antibodies generated from synthetic MUC1 glycopeptide based anti-tumor vaccines. For an efficient and encompassing screening, the underlying library has to contain a large variety of glycans presented in different density and distances on the mucin tandem repeat peptide backbones. An efficient strategy for the synthesis of mucin glycopeptides was developed, which included the preparation of glycosylated and appropriately protected Fmoc-amino acid building blocks and application in Fmoc-solid-phase peptide synthesis (Fmoc-SPPS). The glycosylated amino acids were assembled from a small group of building blocks, which represent common structural motifs within the target structures (***Figure 91***).

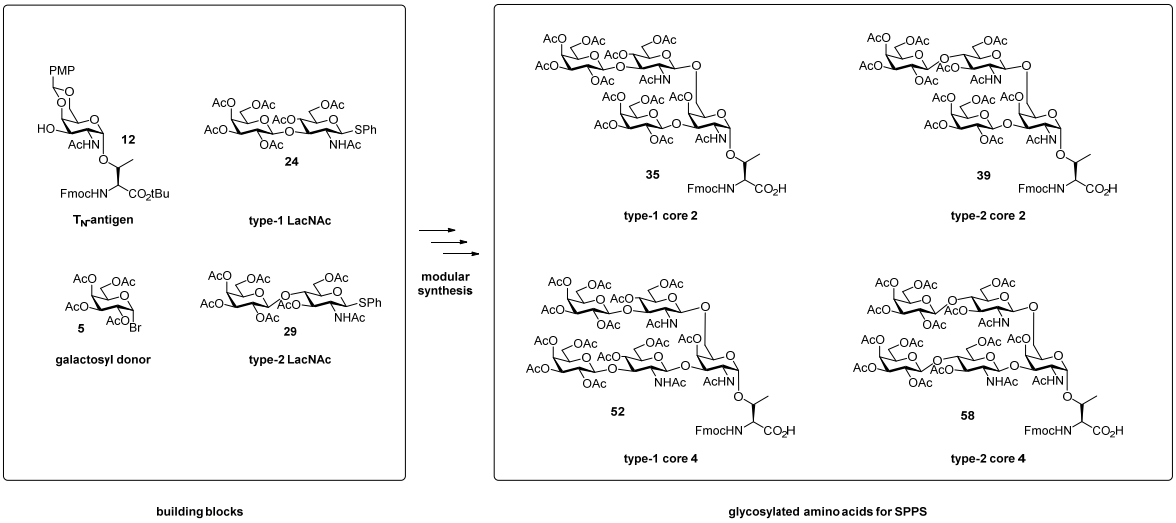


Figure 91: Efficient synthesis of complex mucin-type core and elongated core amino acids using common acceptor and donor building blocks.

The strategy involved the synthesis of the GalNAc-Thr intermediate (T_N-antigen) **12** as a common acceptor. Elongation of the T_N-antigen acceptor with the galactosyl bromide donor **5** and selective protecting group manipulations gave the core 1 disaccharide (T-antigen) acceptor **31**. The T- and T_N-acceptor building blocks were further elongated with type-1 (Galβ1,3GlcNAc) and type-2 (Galβ1,4GlcNAc) *N*-acetyllactosamine thioglycoside donors **24** and **29**. These LacNAc disaccharides were assembled using the trichloroacetimidate method and should elongate the mucin core *O*-glycans with naturally occurring alternating β-galactose and β-*N*-acetylglucosamine saccharides (LacNAc units). After complete assembly of the core and elongated core structures the carbohydrate protection groups were unified to Fmoc-SPPS-compatible global acetylation followed by C-terminal carboxylic acid deprotection.

After stereoselective glycosylation of the T_N-antigen acceptor **12** with galactosyl donor **5** and deprotection of the GalNAc 6- and 4-position of the generated T-antigen disaccharide **31**, the 6-position was elongated with either type-1 LacNAc disaccharide **24** to generate the type-1 core 2 glycosylated amino acid **35** or with type-2 LacNAc disaccharide **29** to generate the type-2 core 2 glycosylated amino acid **39** (**Figure 92**). Both glycosylations proceeded with full stereoselectivity, but coupling of the type-2 LacNAc disaccharide proceeded 10 times slower compared to coupling of the type-1 LacNAc disaccharide to the same glycosyl acceptor, suggesting a considerably higher reactivity of the type-1 donor **24**. The type-1 and type-2 LacNAc only differ in the connectivity of the sugar monomers at the GlcNAc subunit. Therefore it was concluded, that the deactivation of type-2 LacNAc is associated with stereoelectronic effects related to these glycosidic linkage position or the positions of the acetyl protecting group within the glucosamine residue.

In the frame of core 4 synthesis, linear LacNAc elongation of T_N-antigen **12** afforded core 3 trisaccharides **42, 47** and **48**. Despite the β-directing *N*-Troc group, the glycosylation with the type-1 LacNAc donor proceed with an undesired stereoselectivity resulting in a mixture of the α- and β-anomer products (α/β 2:3). Attempts to enhance the T_N-antigen acceptor reactivity was not successful and instead, the core 3 β-anomer **42** was isolated and further branched at the GalNAc-6-position to generate the branched core 4 glycosylated amino acid **52**.

In contrast, elongation of the T_N-antigen with the type-2 LacNAc disaccharide proceeded with full stereoselectivity. However, formation of a side product with addition an N-thiophenyl group on the desired trisaccharide coupling product **48** was further observed. Both, the acetamide **47** and the *N*-thiophenyl product **48** were employed for elongation at the GalNAc-6-position to form the desired type-2 core 4 pent saccharide product with full regio- and stereoselectivity. Both, the NHTroc and the *N*-thiophenyl analog were converted to the identical acetamide **58** during the deprotection of the N-Troc-group. Standard zinc reduction conditions were not optimal in this case and instead samarium diiodide was found to be a suitable reducing agent to efficiently remove the thiophenyl residue and simultaneously reduce the carbamate to an amine followed by acetylation to give the desired acetamide **58**.

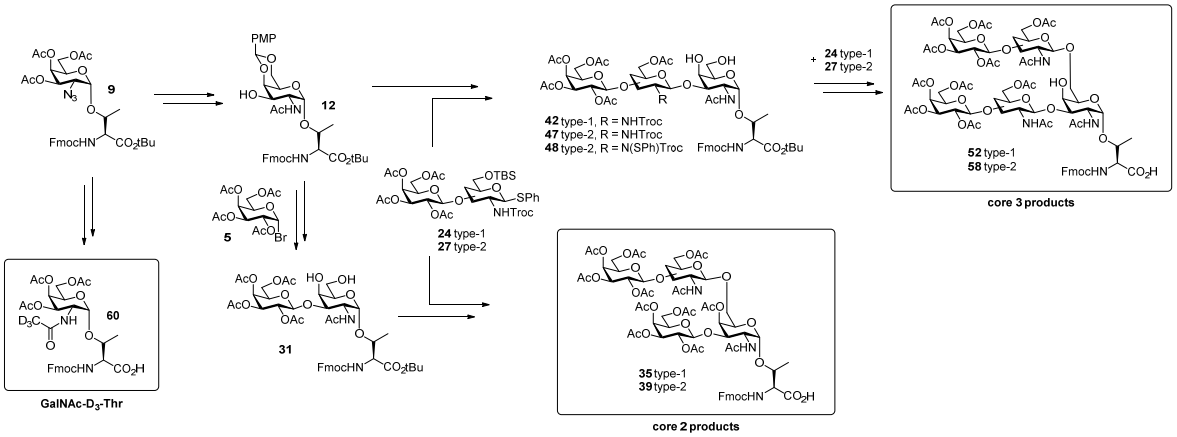


Figure 92: Synthesis of the mucin-type glycosylated amino acids from common building blocks.

In addition to the core 2 and core 4 glycosylated amino acids a deuterium labelled GalNAc-D_3-*O*-Thr glycosylated amino acid **60** was synthesized for application in mass spectrometric studies to resolve oxonium ion fractionation pathways upon high collision-induced dissociation (HCD) experiments. The azide intermediate **9** from the T_N-antigen synthesis was in this case converted to the corresponding deuterated acetamide by azide reduction and acetylation with D_3-labeled *N*-acetoxy succinimide.

The ready glycosylated amino acids were applied in Fmoc-protocol SPPS on *Tentagel*-trityl-resins. The nonadecameric glycopeptides were based on the MUC1 tandem-repeat sequence and were synthesized systematically as mono-, di- and trivalent analogs using the type-1 and type-2 elongated core 2 and core 4 glycan amino acids (***Figure 93***). MUC1 glycopeptides were also prepared carrying the deuterated α-GalNAc-D_3-Thr and normal α-GalNAc-Thr (T_N-antigen) modifications. In order to enable immobilization on microarray slides, all glycopeptides were equipped with an *N*-terminal triethylene glycol spacer (TEG). After on-resin-assembly and cleavage from resin, the glycan acetyl protecting groups were removed under *Zemplén* conditions for the T_N-antigen and core 4 peptides or by using sodium hydroxide at a pH of 10.5-11.0 for complete deacetylation of core 2 peptides.

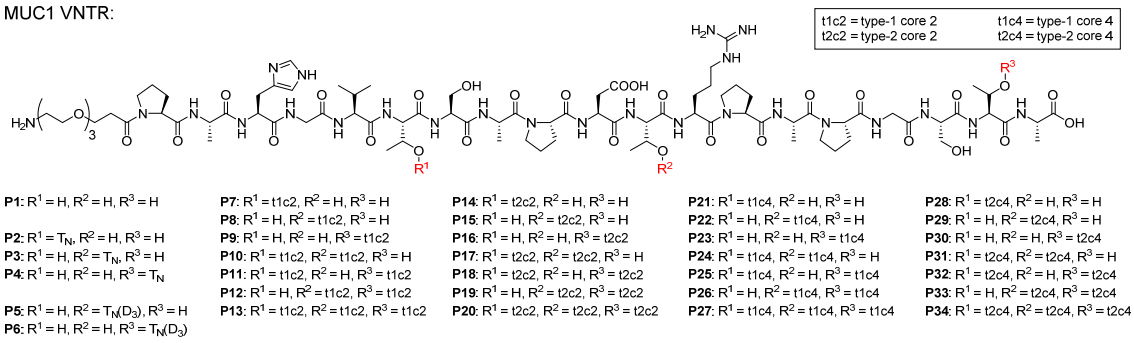


Figure 93: Synthesized MUC1 glycopeptides.

Parts of the prepared glycopeptides were incorporated into a large glycopeptide library and the peptides were further modified by enzymatic sialylation, fucosylation or polyLacNAc extension. The glycopeptides were immobilized on microarray slides for applications in lectin binding studies for instance of the lectins from COPD-related adenoviruses: Fiber-knob protein AD37 from adenovirus type 37 and the short fiber-knob protein AD52 from adenovirus type 52. Both lectins have been found to bind α2,3-sialylated *O*-glycan structures, but elucidation of the role of mucin-type O-glycosylation in these interactions are not explored. However, by the use of the AD37 ligand GD1a as positive control, we could conclude that the detection with a secondary penta-His antibody or by Ni-NTA-labeled fluorophores were not efficient to detect lectin binding and optimization of the detection conditions are still needed.

Parts of the glycopeptide library were also used in studies of human galectin-3 binding to elucidate recognition to tumor-associated glycans and elongated core structures. It was concluded that the tumor-associated T_N and T-antigen structures were recognized although weaker than the LacNAc-elongated core structures. In the natural environment the T_N and T-antigen structures are presented in high density on mucins and the observed binding interactions with Galectin-3 are likely of biological relevance. Increased binding was also observed for multivalent presented glycopeptides compared to monovalent glycosylated structures. From the microarray experiments a strong recognition was found to core glycopeptides elongated with β1,3- and β1,4-linked LacNAc-units. Further, linear α2,3-sialic acid termination enhanced binding, but branched α2,6-sialic acid termination reduced or if no internal Gal residue contained free 4- and 6-OH groups the α2,6-sialylation completely blocked galectin-3 binding.

Microarray binding studies were further performed to evaluate specificities of antibodies induced with MUC1-based anti-tumor vaccine conjugates. Elucidation of the specificity of antibodies raised from different vaccine constructs are of importance to make sure that no cross-reactivity is generated to glycan/glycopeptide epitopes found in healthy cells to avoid generation of auto-immune responses from the obtained immunological memory. In the three vaccines elucidated, the T_N-antigen MUC1 tandem repeat glycopeptides were connected to different immune stimulants; 1) a P30 T-cell epitope (peptide sequence from the tetanus toxoid protein), 2) the P30 T-cell epitope and a Pam_3CSK_4 lipopeptide adjuvant and 3) to the tetanus toxoid immune carrier protein. Antibody titers obtained after induction with the different MUC1 vaccine constructs were determined with ELISA. Further, the antisera were incubated on microarrays, which contained a library of over 100 different MUC1 glycopeptides presented with different glycans and glycan density. Evaluation of the ELISA and microarray results revealed a strong influence of the utilized vaccine conjugate components on the immune response and the generated antibody fine-specificity. The T.Tox-immune carrier vaccine

conjugate elicited very strong polyclonal immune responses with a broad epitope recognition, the two P30 T-cell epitope-based vaccines elicited a slightly weaker, but more specific immune responses. The antibodies induced from the P30 based vaccines showed binding with very high specificity towards the clustered T_N-antigen glycosylation in the GS*T*A region or recognized epitopes containing glycosylation in the PDT*R region. In the latter case extended O-glycan core structures were often tolerated, but recognition of core 2 glycan structures, expected to be found on healthy tissues, were lost upon multivalent glycan presentation on the peptide backbones. Our results confidently showed that tumor specific antibodies are efficiently induced from vaccines containing the P30 T-cell epitope as immune stimulant. These results will be valuable in the design of future anti-tumor vaccines.

In a collaboration with Dr. Jonas Nilsson (*Department of Clinical Chemistry and Transfusion Medicine*, *Institute of Biomedicine*, Gothenburg, Sweden) the deuterium-labelled T_N-antigen peptides **P5** and **P6** were used in order to resolve HexNAc oxonium ion decomposition pathways by HCD fragmentation (**Figure 94**). It was found that GalNAc oxonium ions can undergo a retro-Diels-Alder decomposition leading to the diagnostic ion m/z 126, which is not observed when GlcNAc oxonium ions are fragmented. Also the deacetylation of acetamide m/z 186 to ion m/z 144 is characteristic for GalNAc and was not found at GlcNAc fragmentation. In contrast, successive elimination of two H_2O to m/z 168 and further decomposition to m/z 138 was shared by both HexNAc precursors. Fragmentation of the deuterium-labelled acetamide revealed further, that the generation of an alternative m/z 126 by deacetylation of m/z 168 can be practically disregarded.

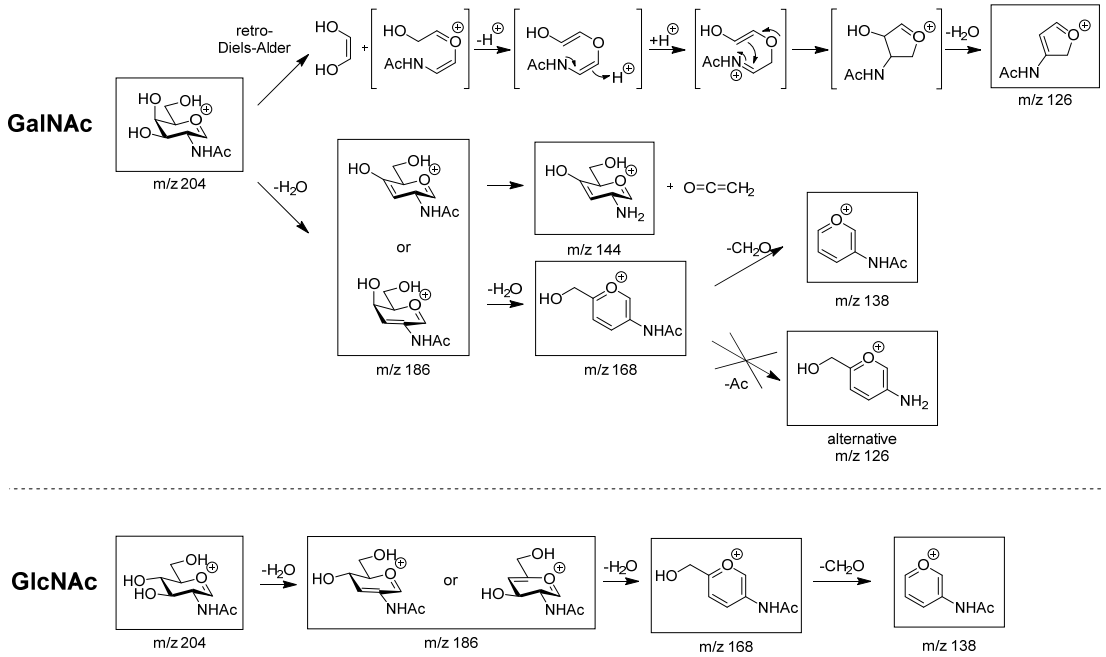


Figure 94: *HCD-induced oxonium ion decomposition pathways. Deacetylation of m/z 168 to generate an alternative m/z 126 can be practically disregarded.*

The characteristic decomposition pathways could be exploited in order to discriminate the different HexNAc isomers via the GlcNAc/GalNAc ratio:

$$\frac{GlcNAc}{GalNAc} = \frac{m/z\ 138 + m/z\ 168}{m/z\ 126 + m/z\ 144}$$

Ratio were fond of 0.2 ≤ GlcNAc/GalNAc ≤ 1 for GalNAc and between 2 ≤ GlcNAc/GalNAc ≤ 50 for GlcNAc.

Also type-1 and type-2 core 3 structures were discriminated with the use of oxonium ion fragmentation. Thereby the resulted fragmentation patterns were dependent on the type-1 and type-2-LacNAc disaccharides and not influenced by the internal GalNAc-1-*O* substructure. This methodology represents an efficient MS-based strategy that enables an oxonium ion-triggered and selective characterization of glycopeptides from complex samples.

In conclusion, a synthetic MUC1 glycopeptide library was generated, optimized for microarray-based evaluation of glycopeptide/lectin or glycopeptide/antibody interactions. The system will enable insights into specificities of known bacterial and virus virulence factors, which are related to infection and inflammation in the course of airway diseases. The gained knowledge may help in the development of potent inhibitors of antiadhesive drugs. Further, glycopeptides from this library were used to elucidate the fine-specificities of tumor-related galectin-3 and antibodies induced from different tumor-directed vaccine constructs were also evaluated. Isotope-labelled glycopeptides

prepared in this work enabled elucidation of HCD-induced oxonium ion fragmentation, providing improved methodology for mass-spectrometry based glycan identification in the course of glycoproteomic workflows.

8.2 Project 2: *O*-glycosylation on tyrosine

In 2011 HexNAc-Tyr modifications were discovered for the first time on amyloid-β peptides isolated from CSF of Alzheimer Disease patients. Due to the internal HexHexNAc structure, it was assumed, that the discovered glycans were of mucin type mono-, di- and trisialylated core 1 *O*-glycans (Galβ1,3GalNAc). In another glycoproteomic study a number of HexNAc-Tyr modifications were identified using colo205 SimpleCells: Nucleobindin-1 (NucB1), Nucleobindin-2 (NucB2), CD44 antigen, extracellular matrix protein 1 (Ecm1) and proline-rich acidic protein 1 (Prap1). In this work the glycopeptides were enriched using lectin weak affinity enrichment with the lectin *Vicia villosa agglutinin* (VVA), which were expected to be selective in recognition of terminal GalNAc residues. Further the SimpleCells contained enzyme and chaperon mutations to block GalNAc elongation of the modified proteins, the identified HexNAc-Tyr glycopeptides in this study were therefore assumed to be GalNAc-Tyr modifications. A later glycoproteomic study on the murine synaptosome identified three HexNAc-Tyr modifications on mitochondrial proteins: ATP synthase subunit β (Atp5b), aspartate aminotransferase (mAspAT) and voltage-dependent anion-selective channel protein (Vdac1). The utilized *Wheat germ agglutinin* (WGA) glycopeptide enrichment and ETD-mass spectrometric methodology were not able to identify if the detected HexNAc-Tyr glycopeptide modifications were GalNAc or GlcNAc residues. The localization in mitochondria and later identification of nearby Ser/Thr *O*-GlcNAcylation sites and the fact that the identified Tyr glycosylation sites are known phosphorylation sites would suggest that the found HexNAc-Tyr modifications are consisting of β-GlcNAcylation. In another work, PaTox, a toxin from *Photorhabdus asymbiotica* and the toxin antifeeding prophage 18 (Afp18) from *Yersinia ruckeri* were found to influence activity of small GTPases (Cdc42, Rac1, RhoA) by HexNAc-Tyr glycosylation in the course of infection. Based on enzyme studies with PaTox and subsequent NMR analysis the glycan on HexNAc-Tyr was in this case assigned as α-GlcNAc-Tyr.

Most of these identifications have in common, that the glycan assignment as α-GalNAc, α-GlcNAc or β-GlcNAc-Tyr modifications were done indirectly according to the utilized enrichment method or general knowledge about cellular glycan distribution. Also to date the β-GlcNAc-Tyr modification has not yet been proven to be a new PTM, although the conclusions we made regarding nearby modifications and localization of the identified mitochondrial HexNAc-Tyr glycopeptides would suggest that β-GlcNAc-Tyr modification exist. β-GlcNAc identification is more challenging due to its relatively high turnover compared to α-GalNAc glycosylation. In addition, some lectins which are used for β-GlcNAc enrichment show cross reactivity towards α-GalNAc, which in consequence might lead to false assignment. The biological roles of α-GalNAc vs β-GlcNAc O-glycosylation on Ser/Thr residues differs dramatically, thus a precise assignment is desirable in order to understand the functions of the recently discovered HexNAc-Tyr modifications.

In order to enable reliable functional and analytical studies of HexNAc-Tyr O-glycosylation existing molecular tools such as lectin and antibodies known to recognize GalNAc or GlcNAc residues on Ser/Thr glycopeptides were evaluated towards their applicability and reliability to detect these modifications on Tyr residues. In earlier work mentioned above we found that oxonium ion profiles from HCD MS-fragmentation could be used to distinguish GalNAc and GlcNAc epimers on Thr/Ser glycopeptides and it was therefore interesting to elucidate if this as expected also applies for Tyr modified peptides. Further, it was aimed to elucidate if the enzymes O-GlcNAc transferase (OGT) and O-GlcNAc hydrolase (OGA) were responsible for the biosynthesis of GlcNAc-Tyr modifications in accordance to the modifications on Ser/Thr. To enable these studies a glycopeptide library was prepared, containing fully synthetic peptides. The synthetic strategy followed an analog principle of the above described MUC1 glycopeptide synthesis (project 1), using pre-synthesized glycosylated amino acids in Fmoc-SPPS. To generate the different HexNAc-O-Tyr isomers, a common Fmoc-protected *L*-Tyr acceptor **62** was coupled with the corresponding GalNAc or GlcNAc glycosyl donors. Analog to the T_N-antigen synthesis (described in project 1), the α-GalNAc-Tyr amino acid **64** was prepared by using the acetobromogalactose donor **8** for in the coupling to the Tyr acceptor **62**. α- and β-GlcNAc were introduced as trichloroacetimidates, which were equipped with the 2-azide donor **66** for generation of α-GlcNAc-Tyr anomer **69** and with the 2-trichloroethyl carbamate donor **71** for generation of β-GlcNAc-Tyr anomer **74** (***Figure 95***). All glycosylations proceeded with full stereoselectivity. For compatibility with Fmoc-SPPS, the N-protecting groups were converted to NHAc groups and protection at the C-terminal carboxylic acid was removed.

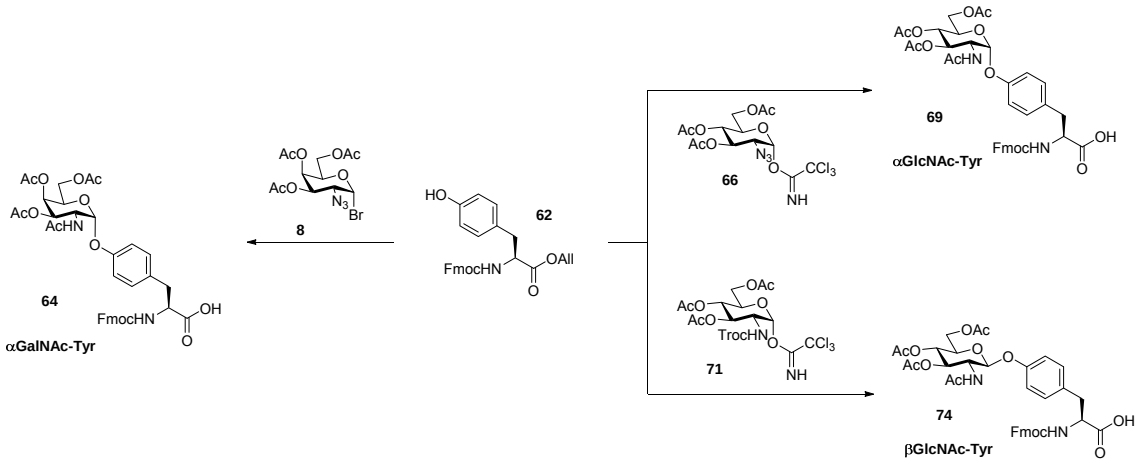


Figure 95: Synthetic overview for the HexNAc-*O*-Tyr amino acids.

Incorporation of the glycosylated amino acids via Fmoc-SPPS and synthesis of the full glycopeptide constructs were done according to the strategy described for project 1. In total a library of 80 peptides were synthesized. A full list is given on p. 95, ***Table 2*** and contained glycopeptides such as MUC1 model peptides and tryptic segments of biologically relevant glycopeptides with known Ser/Thr/Tyr *O*-glycosylation sites. To enable comparison of HexNAc glycosylations at the same site (α-GalNAc vs α-GlcNAc vs β-GlcNAc) or presentation of identical HexNAc isomers on different glycosylation sites, additional analogs were prepared, mainly based on tryptic glycopeptide fragments previously identified in glycoproteomic studies. Further, preparation of glycopeptides with divalent glycan presentation enabled studies of bidentate or multivalent glycan interactions. Unglycosylated peptides were applied to study the enzyme specificity of OGT. Finally, phosphorylated and mixed phospho/glyco analogs of mitochondrial HexNAc-O-Tyr peptides enabled to examine the influence of vicinal phosphorylation in studies of enzymes (OGA and OGT) responsible for β-GlcNAc biosynthesis and for elucidation of lectins.

In standard glycoproteomic and glycomic workflows, standard lectins are used for glycopeptide enrichment, as for instance: *Wheat Germ Agglutinin* (WGA), which has been reported to bind terminal glucosamine residues, *Vicia Villosa Lectin* (VVA) which is known to recognize α- and β-configured terminal GalNAc on Ser/Thr and *Griffonia Simplicifolia Lectin II* (GSL II), a lectin found to bind terminal α/β-GlcNAc exclusively. The binding specificity of these lectins towards terminal glycans and Ser/Thr-*O*-HexNAc glycans is extensively studied, while specificities towards the newly discovered HexNAc-Tyr modification are still unexplored. Further, lectin binding recognition dependence on glycan presentation at the peptide backbone or influences by peptide multivalent glycosylation is rather unexplored even if the above mentioned lectins, VVA and WGA, are commonly applied for weak-affinity enrichment in glycoproteomic workflows.

In order to enable an encompassing evaluation and comparison of HexNAc-Tyr and HexNAc-Ser/Thr recognition, glycopeptide/lectin binding assays were performed based on microarray format MA1. The array included various peptide sequences presenting α-GalNAc, α-GlcNAc or β-GlcNAc on Thr or Tyr as mentioned above.

Evaluation of the binding data suggested a strong preference of WGA and GSL II for HexNAc-Tyr glycopeptides over HexNAc-Thr glycopeptides, in contrast VVA recognized glycopeptides equally well independent on glycosylation on Tyr or Thr (***Figure 96***). Both WGA and GSL II contain a hydrophobic binding pocket, which supports π-π-stacking which may to some extend explain the observed binding preferences to HexNAc-Tyr glycopeptides. Further, binding to the less flexible Tyr side chain might result in less entropic penalty compared to Thr glycosylated peptides. Binding patterns of VVA (α-GalNAc) and GSL II (α/β-GlcNAc) followed those known from literature regardless of the acceptor amino

acid. WGA is often considered as a GlcNAc specific lectin, but earlier binding studies have shown that terminal GalNAc-residues are also well recognized, this was further supported in our work. Binding of WGA to GlcNAc-Tyr glycopeptides was generally stronger than binding to GalNAc-Tyr peptides, but in turn GalNAc-Tyr peptides were generally much better recognized than GlcNAc-Thr modified peptides.

In a follow-up assay the GlcNAc-binding lectins WGA and GSL II were incubated with an extended GlcNAc library MA2. In this library the selection of presented glycosylated amino acids was extended by HexNAc-Ser, which was found to be recognized in a similar manner as the above described HexNAc-Thr glycopeptides. Peptide sequences with multiple potential O-GlcNAc sites (Ser/Thr/Tyr) were further included and selected sequences were added as mono- and divalent analogs. The obtained binding data suggested a strong influence of the peptide backbone spacing for recognition of divalent glycopeptides and in most examples a clear cooperative effect could be observed. For directly neighboring glycan presentation an enhancing or diminishing effect was observed, depending on the glycopeptide sequence. This might be a result of different glycan orientations in the lectin binding pocket paired with different rotational freedom of Tyr vs Thr and Ser. In contrast, glycopeptides with longer distances between the glycosylated amino acids showed a generally enhanced recognition compared to their monovalent analogs.

In order to determine apparent K_D-values for each lectin/glycopeptide combination, the library has been incubated with a series of WGA concentrations. The results support the above described findings, which further indicates the necessity of a minimum peptide spacer length between the presented glycans for optimal multivalent or bidentate binding.

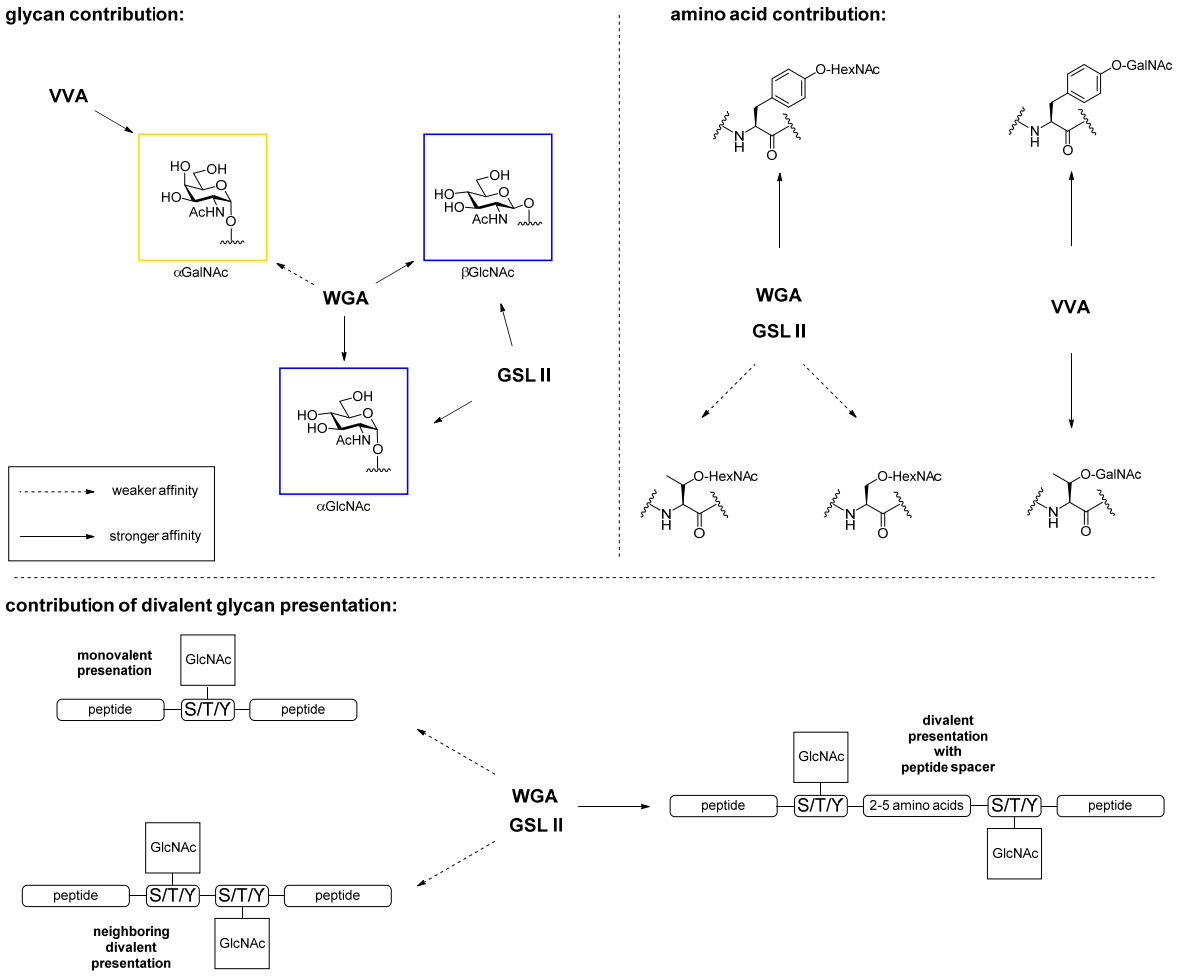


Figure 96: Ligand structure contributions to lectin affinity.

In project 1 we have shown, that HCD-induced oxonium ion fragmentation enables the discrimination of HexNAc structural epimers on Ser and Thr. If this methodology is also applicable on HexNAc-Tyr structures, it could enable the ultimate identification of β-GlcNAc-Tyr as a new PTM, which is expected to be present on mitochondrial glycoproteins like AspAT, Atp5b and Vdac1. Different glycoforms of known HexNAc-Tyr glycopeptides as well as MUC1 model peptides were applied in HCD-MS fragmentation and the results showed generation of the same oxonium ion pattern as observed for HexNAc-Ser/Thr glycopeptides: m/z 204, 186, 168, 144, 138 and 126 with characteristic ion abundance profiles for GalNAc vs GlcNAc. As expected no influence of the acceptor amino acid (Thr vs Tyr) could be identified, suggesting full applicability of the methodology on HexNAc-Tyr glycosylated peptide fragments. Fragmentation experiments over a stepped NCE (**N**ormalized **C**ollision **E**nergy) gradient generated NCE-dependent GlcNAc/GalNAc ratios (***Figure 97***), which reached HexNAc-characteristic values once an NCE threshold of 30 % was reached or exceeded (≤ 1 for GalNAc and 3-4 for GlcNAc).

No significant differences in ion profiles were observed when comparing α- vs β-HexNAc anomer decomposition and the methodology is thereby limited to the discrimination of GlcNAc vs GalNAc without further information on the anomeric configuration of the glycan. However, this work enables the integration of HexNAc-Tyr as potential modification into large-scale glycoproteomic studies via mass spectrometric analysis and computational discrimination of GalNAc-Tyr vs GlcNAc-Tyr.

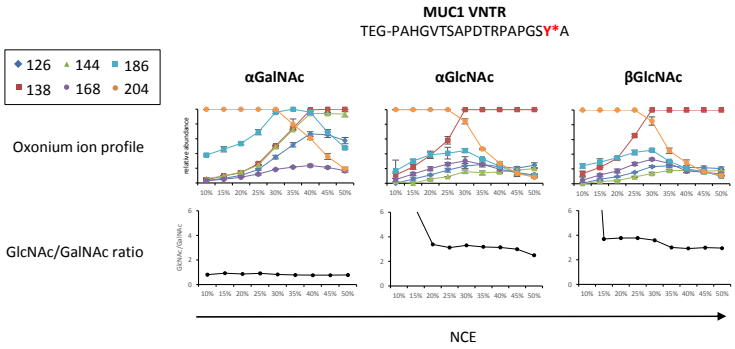


Figure 97: HexNAc-characteristic oxonium ion progressions exemplified on MUC1 glycopeptides. Calculation of the corresponding GlcNAc/GalNAc ratio allows computational discrimination of GlcNAc vs GalNAc.

The GlcNAc modification is introduced and removed by a set of two enzymes: O-GlcNAc transferase (OGT) and O-GlcNAc hydrolase (OGA). Evidence, that these enzymes also facilitate β-GlcNAc-Tyr biosynthesis would be the ultimate proof of this to date unverified modification.

The ubiquitously expressed OGA has been found to show broad substrate specificity towards GlcNAc-Ser/Thr peptides, lacking a defined consensus motif. Therefore it was considered likely that this enzyme also is responsible for removal of β-GlcNAc from Tyr residues. In order to verify this assumption and to obtain information on the catalytic efficiency and fine-specificity of OGA towards GlcNAc-Tyr, a selection of synthetic glycopeptides was applied in OGA experiments. The tested substrates included sequences with previously identified β-GlcNAc sites, corresponding Ser, Thr and Tyr analogs as well as divalent analogs and mixed glyco-/phosphopeptides. In addition, possible OGA discrimination of structural HexNAc epimers was elucidated by including model peptide constructs with αGalNAc, α-GlcNAc and β-GlcNAc residues. All glycopeptides were incubated with OGA from *Streptococcus pyogenes* and substrate conversion was followed with time via HPLC-MS.

In general, β-GlcNAc-Tyr was found to be a very good substrate for OGA and α-GalNAc and α-GlcNAc-Tyr peptides were not hydrolyzed. The data further suggested a strong correlation between glycosylated amino acid and conversion: Recognition of β-GlcNAc-Ser/Thr is dependent on the polarity of the downstream amino acids (-1 and -2 with respect to the glycosylation site). In contrast, deglycosylation of Tyr commences completely independent of the surrounding backbone sequence. This effect might be explained by the constitution of the deep OGA substrate binding cleft, which

comprises a catalytic $(\beta/\alpha)_8$-barrel with glycan recognition and a hydrophobic stalk domain, which interacts via aromatic amino acid side chains of the substrate backbone. Probably, π-stacking between substrate Tyr and the stalk domain facilitates a strong Tyr substrate recognition, which is not possible for Ser and Thr substrates. Further, the spatial distance between Tyr-Cα and Tyr-OH is much larger than the corresponding distance in Ser and Thr making GlcNAc-Tyr glycopeptides less dependent on the peptide backbone (***Figure 98***). Interestingly, enhanced conversion was observed for divalent glycopeptides compared to their monovalent analogs suggesting that cooperative effects are involved for substrate recognition and catalysis. Also nearby phosphorylation at the peptide backbone was found to increase β-GlcNAc hydrolysis.

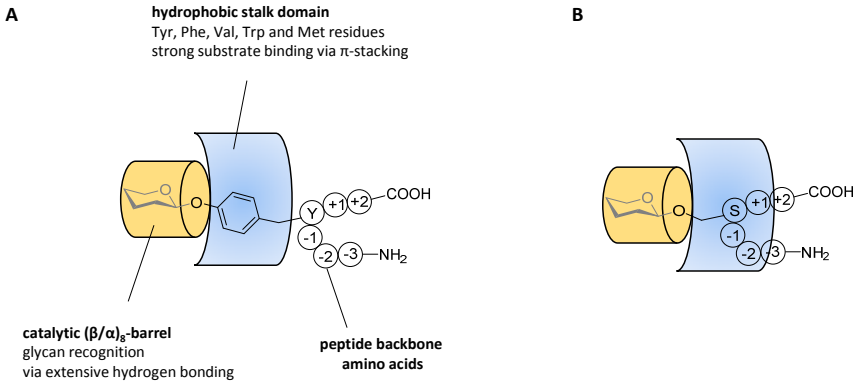


Figure 98: Schematic illustration of proposed OGA substrate recognition. **A)** Strong GlcNAc-Tyr recognition via π-stacking of Tyr and aromatic stalk domain amino acids. Due to the spatial dimensions of Tyr only minor peptide backbone interactions between stalk domain and substrate peptide backbone take place. **B)** The smaller spatial dimensions of Ser compared to Tyr enable interactions between stalk domain and substrate peptide backbone. As a result the OGA specificity towards Ser glycopeptides is dependent on the amino acid environment surrounding the substrate glycosylation site. Considerations for Thr are equivalent to Ser.

Follow-up glycopeptide incubations with human OGA (hOGA) supported the observed trends from bacterial OGA incubations. As substrate candidates, two β-GlcNAc-Tyr glycopeptides were selected, one with a polar amino acid at -1 and one with a polar amino acid at -2. Further one β-GlcNAc-Ser peptide without polar amino acid at -1 or -2 was included. Despite the presence of polar amino acids, both β-GlcNAc-Tyr candidates showed conversion (***Figure 99***). In contrast, the β-GlcNAc-Ser candidate showed no conversion, although no polar amino acids were present at -1 or -2. This highlights the better substrate specificity towards HexNAc-Tyr glycopeptides independent of peptide backbone residues.

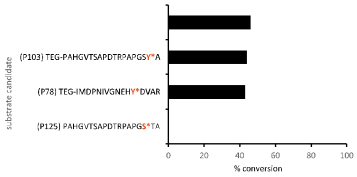


Figure 99: Observed substrate conversions from glycopeptide/hOGA incubations.

Enzymatic glycosylation of known GlcNAc sites (Tyr and Ser) by incubation with hOGT were not successful. Instead, m/z values of unglycosylated peptides and several shorter unglycosylated peptide fragments were detected, which suggested that the enzyme prep. might contain possible protease activity. In future experiments protease inhibitors will be included to the reaction mix. However, incubations of the recombinant enzyme with known and very reactive model substrates led to full conversion after 24 h, proving the functionality of the enzyme itself.

In conclusion, a versatile glycopeptide library was generated comprising peptide sequences with known HexNAc sites on Tyr, Ser and Thr and analogs thereof. Lectins commonly used for glycopeptide enrichment in glycoproteomic workflows have been evaluated towards their specificity to HexNAc-Tyr modifications, which was compared with glycosylation on Thr/Ser residues. As a result, HexNAc-Tyr glycopeptides were identified as superior ligands compared to glycosylation on Ser/Thr for WGA and GSL II lectins. The GalNAc-specific lectin VVA recognized glycopeptides equally well independent on presentation on Tyr vs Ser/Thr. Further, an oxonium ion-based mass-spectrometric methodology for HexNAc-isomer identification was verified to be applicable also for glycosylation on Tyr. Finally, the enzymes OGA and OGT responsible for β-GlcNAc biosynthesis were evaluated towards their substrate specificity towards Tyr glycosylation. β-GlcNAc-Tyr was found to be an excellent substrate for OGA, however Tyr glycosylation by OGT could not be verified. This work has resulted in new insights regarding binding recognition to HexNac-Tyr glycopeptides, which may be of importance for future work to understand biological functions of these modifications. The conclusions made regarding lectin specificity to HexNAc-Tyr glycopeptides is also of relevance for optimal affinity enrichment and detection of these modifications. The current work further enabled the use of a new methodology to distinguish HexNAc-Tyr epimers in glycoproteomic workflows, which might lead to new identifications of this recently discovered modification.

9 Experimentals

9.1 Building block synthesis

9.1.1 General

Solvents:

All solvents were purchased as quality grade *pro analysi* (p.a.). Acetonitrile, dichloromethane and nitromethane were dried by refluxing over calcium hydride and subsequent distillation under argon. Dry diethyl ether and toluene were purchased anhydrous with a purity of ≥ 99.5 %, dry methanol, dimethylformamide and tetrahydrofuran were purchased anhydrous with a purity of ≥ 99.8 %.

Thin layer chromatography (TLC):

For TLC silica gel (*Kieselgel 60 F_{254}*) coated aluminium plates from *Merck KGaA* (Darmstadt) were used. Spots were detected by UV light (254 nm) on the untreated TLC plate and/or by treatment of the TLC plate with 20 % sulfuric acid in water and subsequent burning of the carbohydrates in a hot airstream.

Column chromatography:

Silica column chromatography was performed using *Kieselgel 60* (0.04 – 0.063 nm) from *Carl Roth GmbH* (Karlsruhe).

Optical rotation:

The specific rotation ($[\alpha]_D^{20}$) was measured on a *Polaritronic HH8* from *Schmidt + Haensch* GmbH (Berlin) at the D-line of sodium (λ = 598.5 nm). The solvents and analyte concentrations are stated for each compound separately.

Mass spectrometry:

High resolution ESI-MS spectra were recorded with an *LTQ-FT ICR ultra* mass spectrometer and an *LTQ Orbitrap XL* mass spectrometer (both Thermo Scientific).

NMR-spectroscopy:

NMR spectra were measured at the following instruments at 295 K:

- *Varian Mercury 400* (*Agilent*): 400 MHz 1H-NMR, 100.6 MHz 13C-NMR, gCOSY, TOCSY, gHSQC, gHMBC.

- *Advance DRX 500* (*Bruker*): 500 MHz ^{1}H-NMR, 125.8 MHz ^{13}C-NMR, gCOSY, TOCSY, gHSQC and gHMBC.
- *Ascend 600* (*Bruker*)*:* 600 MHz ^{1}H-NMR, 150.9 MHz ^{13}C-NMR, gCOSY, TOCSY, gHSQC and gHMBC.
- *Avance III HD 700* (*Bruker*): 700 MHz ^{1}H-NMR, 176.0 MHz ^{13}C-NMR, gCOSY, TOCSY, gHSQC and gHMBC.

Elucidation of the ^{1}H and ^{13}C spectra was performed with the help of corresponding gCOSY, TOCSY, gHSQC and gHMBC experiments. The reported chemical shifts were calibrated to the residual solvent signals[395]. For assignment of the different carbohydrate ring systems, the reported signals are marked with apostrophes (‘) as shown in ***Figure 100***.

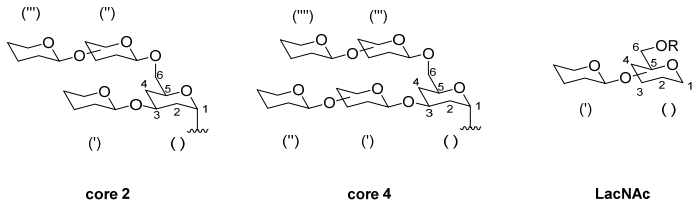
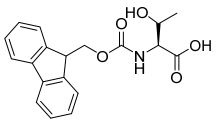


Figure 100: Assignment of carbohydrate ring systems and atom positions within the rings.

9.1.2 Synthesis of mucin-type glycosylated amino acid building blocks

9.1.2.1 Synthesis of the T_N-antigen amino acid acceptor

_N_-9-Fluorenylmethyloxycarbonyl-_L_-threonine (2)[311]

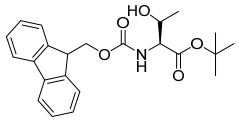


A solution of *L*-threonine **1** (26.6 g, 222.8 mmol, 1.0 eq.) and sodium bicarbonate (18.8 g, 222.9 mmol, 1.0 eq.) in acetone/water (1:1, 1.2 L) was stirred at and cooled on ice. Then Fmoc-*O*Su (75.0 g, 222.3 mmol, 1.0 eq.) was added in portions and stirring was continued for three days, allowing the solution to warm up to room temperature. Afterwards the *p*H was adjusted to 2 using concentrated hydrochloric acid and the acetone was removed in a rotary evaporator. The remaining aqueous solution was extracted three times with dichloromethane (300 mL each) before washing the combined organic layers twice with 1 N hydrochloric acid solution and subsequent drying over magnesium sulfate. The solvent was removed in vacuo by codistillation three time with toluene and three time with dichloromethane. Fmoc-threonine **2** was obtained as an amorphous solid (75.1 g, 220.0 mmol, quant.) and was used without further purification for the follow up reaction.

R_f = 0.51 (EtOAc/acetic acid 10:1). *ESI-MS*(*pos)*: m/z: 341.85 ($[M+H]^+$, calc.: 342.13).

^{1}H-NMR (400 MHz, MeOD), *δ (ppm)*: 7.79 (d, *J* = 7.9 Hz, 2H, H-4-; H-5-Fmoc), 7.42 – 7.36 (m, 2H, H-3-; H-6-Fmoc), 7.34 – 7.28 (m, 2H, H-2-; H-7-Fmoc), 4.41 – 4.36 (m, 2H, CH_2-Fmoc), 4.34 – 4.27 (m, 1H, T^{β}), 4.27 – 4.21 (m, 1H, H-9-Fmoc), 4.18 (d, *J* = 3.0 Hz, 1H, T^{α}), 1.21 (d, *J* = 6.5 Hz, 3H, T^{γ}).

_N_-9-Fluorenylmethyloxycarbonyl-_L_-threonine-_tert_-butylester (3)[312]

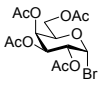


A mixture of dicyclohexylcarbodiimide (39.9 g, 193.4 mmol, 3.3 eq.), *tert*-butanol (10.9 g, 146.5 mmol, 2.5 eq.) and copper(I) chloride (406 mg, 4.1 mmol, 0.07 eq.) was stirred for three days under argon and exclusion of light. After addition of dry dichloromethane (15 mL), Fmoc-threonine **2** (20.0 g, 58.59 mmol, 1.0 eq., dissolved in 75 mL DCM) were added dropwise within 1 h. The solution was stirred until no further conversion could be observed (2 h) on TLC. The urea was filtrated off with subsequent washing with dichloromethane and the filtrate was concentrated, while more urea was falling out. The residue was washed three times with sat. sodium bicarbonate solution and dried over sodium sulfate. After removing the solvent in vacuo, the residue was dissolved in ethyl acetate and stored at -30° C overnight. Newly crystallized urea was filtered off and the residue was purified by column chromatography on silica (CH/EtOAc 2:1) to give of Fmoc-threonine *tert*-butylester **3** (12.5 g, 31.4 mmol, 54 %).

R_f = 0.33 (CH/EtOAc 2:1). *ESI-MS*(*pos)*: m/z: 420.10 ($[M+Na]^+$, calc.: 420.18).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.76 (d, *J* = 7.5 Hz, 2H, H-4-; H-5-Fmoc), 7.65 – 7.54 (m, 2H, H-1-; H-8-Fmoc), 7.43 – 7.37 (m, 2H, H-3-; H-6-Fmoc), 7.34 – 7.28 (m, 2H, H-5-; H-5-Fmoc), 5.54 (d, *J* = 9.0 Hz, 1H, NH), 4.42 (d, *J* = 7.2 Hz, 2H, CH_2-Fmoc), 4.32 – 4.18 (m, 3H, T^{α}; T^{β}; H-9-Fmoc), 1.49 (s, 9H, O*t*Bu), 1.25 (d, *J* = 7.0 Hz, 3H, T^{γ}).

2,3,4,6-Tetra-_O_-acetyl-α-_D_-galactopyranosyl bromide (5)



While stirring with a KPG-stirrer, $HClO_4$ (2.8 mL, 50.0 mmol, 0.09 eq.) were added to acetic anhydride (482 mL, 5.1 mol, 9.2 eq.) while cooling on ice. Then *D*-galactose **4** (100.0 g, 555.1 mmol, 1.0 eq.) were

added carefully, while the temperature did not exceed 40° C. After 1 h of stirring PBr_3 (104 mL, 721.6 mmol, 2.0 eq.) was added dropwise slowly at 0° C. Stirring was continued for 3 h, while the solution was allowed to reach room temperature. Subsequently the solution was poured into DCM (400 mL) and stirred with ice water (800 mL). The organic phase was stirred twice with sat. bicarbonate solution and once with brine before drying over sodium sulfate. After removal of the solvent in vacuo the residue was crystallized from diethyl ether/cyclohexane (14:10, 140 mL) to give of compound **5** (187.3 g, 455.0 mmol, 82 %).

R_f = 0.65 (Tol/EtOAc 3:2). *ESI-MS*(*pos)*: m/z: 330.93 ([Gal oxonium]$^+$, calc.: 331.10), 427.86 ([M+NH_4]$^+$, calc.: 428.06).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm):* 6.67 (d, *J* = 4.0 Hz, 1H, H-1), 5.54 – 5.45 (m, 1H, H-4), 5.37 (dd, *J* = 10.6, 3.3 Hz, 1H, H-3), 5.02 (dd, *J* = 10.6, 4.0 Hz, 1H, H-2), 4.49 – 4.41 (m, 1H, H-5), 4.20 – 4.03 (m, 2H, H-$6_{a,b}$), 2.16 – 1.92 (m, 12H, 4 Ac).

^{13}C-NMR: δ = 170.3, 170.1, 169.9, 169.8 (4 C=O(Ac)), 88.2 (C-1), 71.2 (C-5), 68.1, 67.9 (C-2; C-3), 67.1 (C-4), 60.9 (C-6), 20.8, 20.7, 20.6, 20.6 (4 Me(Ac)).

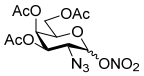
3,4,6-Tri-*O*-acetyl-D-galactal (6)[315-316]

A solution of water (435 mL) and acetic acid (326 mL) in a three-necked flask equipped with a mechanical stirrer was cooled to – 10° C. After addition of zinc dust (119.6 g, 1.83 mol, 6.3 eq.) and copper(II) sulfate pentahydrate (12.5 g, 49.9 mmol, 0.2 mmol), compound **5** (119.6 g, 290.8 mmol, 1.0 eq.) was added dropwise at -10° C. Additional zinc dust (58.9 g, 901.5 mmol, 3.1 eq.) and copper(II) sulfate pentahydrate (7.3g, 29.1 mmol, 0.1 eq.) was added and the mixture was stirred for 4.5 h. After filtering off the zinc over celite, the celite was washed with cold water/acetic acid (1:1) and dichloromethane. The organic layer was washed twice with sat. bicarbonate solution and dried over sodium sulfate. After removal of the solvent the crude was purified by column chromatography on silica (CH/EtOAc 3:1) to give compound **6** (68.3 g, 250.8 mmol, 86 %).

R_f = 0.40 (Tol/EtOAc 3:1). *ESI-MS*(*pos)*: m/z: 152.85 ([2-(acetomethyl)pyrylium]$^+$, calc.: 153.06), 289.79 ([M+NH_4]$^+$, calc.: 290.12).

[1]H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 6.45 (dd, *J* = 6.3, 1.7 Hz, 1H, H-1), 5.57 – 5.51 (m, 1H, H-3), 5.44 – 5.39 (m, 1H, H-4), 4.72 (ddd, *J* = 6.3, 2.7, 1.4 Hz, 1H, H-2), 4.36 – 4.15 (m, 3H, H-5; H-$6_{a,b}$), 2.14 – 1.99 (m, 9H, 3 Ac).

3,4,6-Tri-*O*-acetyl-2-azide-2-deoxy-α/β-D-galactopyranosyl nitrate (7)[318]

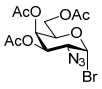


Dried ceric ammonium nitrate (242.4 g, 442.2 mmol, 3.0 eq.) were grinded and filled to a 2-L Schlenk flask under argon. After adding dry acetonitrile (500 mL), grinded sodium azide (23.0 g, 353.8 mmol, 2.4 eq.) were added and the resulting solution was cooled to -10° C. Then compound **6** (40.1 g, 147.4 mmol, 1.0 eq.) was dissolved in dry acetonitrile (250 mL) and added to the former solution. The temperature was kept between -20° C and -10° C while stirring was continued until no more conversion could be observed on TLC. The Solution was diluted with diethyl ether (700 mL) and poured into stirred ice water (600 mL). The organic layer was separated and washed with additional water (500 mL) before drying over sodium sulfate. The organic layer was removed in vacuo without heating and the residue was purified by column chromatography on silica (Tol/EtOAc 20:1) to give compound **7** (30.4 g, 80.8 mmol, 55 %).

R_f = 0.59 (Tol/EtOAc 4:1).

[1]H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 6.33 (d, *J* = 4.2 Hz, 1H, H-1^{α}), 5.57 (d, *J* = 8.8 Hz, 1H, H-1^{β}), 5.49 (d, *J* = 3.3 Hz, 1H, H-4^{α}), 5.38 (d, *J* = 3.3 Hz, 1H, H-4^{β}), 5.24 (dd, *J* = 11.3, 3.2 Hz, 1H, H-3^{α}), 4.96 (dd, *J* = 10.6, 3.3 Hz, 1H, H-3^{β}), 4.40 – 4.31 (m, 1H, H-5^{α}), 4.22 – 3.94 (m, 6H, H-2^{α}; H-5^{β}; H-$6_{a,b}{}^{\alpha}$; H-$6_{a,b}{}^{\beta}$), 3.82 (dd, *J* = 10.6, 8.8 Hz, 1H, H-2^{β}), 2.24 – 1.97 (m, 9H, 3 Ac).

3,4,6,-Tri-*O*-acetyl-2-azido-2-deoxy-α-D-galactopyranosyl bromide (8)[318]



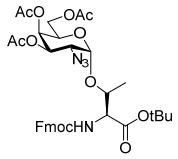
While stirring, lithium bromide (24.1 g, 277,5 mmol, 7.2 eq.) was added slowly to dry dichloromethane (174 mL). After adding compound **7** (14.4 g**,** 38.5 mmol, 1 eq.), the reaction was stirred overnight. The solution was diluted with dichloromethane, washed three times with brine and dried over sodium

sulfate. The crude was purified over a short silica column (CH/EtOAc 2:1) to give compound **8** (11.7 g, 29.8 mmol, 77 %).

R_f = 0.70 (Tol/EtOAc 7.3).

^{1}H-NMR (400 MHz,$CDCl_3$), *δ (ppm)*: 6.46 (d, *J* = 3.8 Hz, 1H, H-1), 5.48 (dd, *J* = 3.2, 1.4 Hz, 1H, H-4), 5.32 (dd, *J* = 10.7, 3.2 Hz, 1H, H-3), 4.48 – 4.43 (m, 1H, H-5), 4.19 – 4.05 (m, 2H, H-$6_{a,b}$), 3.97 (dd, *J* = 10.7, 3.8 Hz, 1H, H-1), 2.16 – 1.95 (m, 9H, 3 Ac).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-(3,4,6-O-acetyl-2-deoxy-2-azido-α-D-galactopyranosyl)-_L_-threonine-_tert_-butylester (9)[301-302, 314, 396]



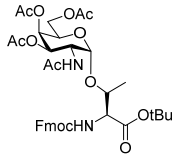
In a 2-L Schlenk flask, a suspension of Fmoc-threonine *tert*-butylester **3** (36.4 g, 91.6 mmol, 1.0 eq.) and molecular sieve (36.3 g, 4 Å) in dry dichloromethane/toluene (1:1, 760 mL) were stirred for 1 h at room temperature. Meanwhile silver(I) perchlorate (6.2 g, 27.5 mmol, 0.3 eq.) was dried by codistillation with dry toluene three times, while never removing the solvent completely to avoid explosive decomposition. Afterwards the slurry was cooled on ice and silver(I) carbonate (37.9 g, 137.4 mmol, 1.5 eq.) was added. Next a solution of the previously dried silver(I) perchlorate in toluene (20 mL) was added. Finally a solution of galactosyl bromide **8** (36.1 g, 91.6 mmol, 1.0 eq.) in dichloromethane/toluene (1:1, 170 mL) was added dropwise and stirring was continued overnight under exclusion of light, allowing the system to reach room temperature. Afterwards the slurry was diluted with dichloromethane (600 mL) and filtrated over a POR2 filter with celite. The filtrate was washed twice with sat. sodium bicarbonate solution and twice with brine before drying over sodium sulfate. The solvent was removed in vacuo at 35° C and the residue was purified by column chromatography on silica (CH/EtOAc 7:3) to give compound **9** (49.2 g, 69.2 mmol, 76 %).

R_f = 0.32 (CH/EtOAc 7:3). *ESI-MS(pos): m/z:* 710.72 ($[M+H]^+$, calc.: 711.28), 727.73 ($[M+NH_4]^+$, calc.: 728.31).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.76 (d, *J* = 7.5 Hz, 2H, H-4-; H-5-Fmoc), 7.63 (d, *J* = 7.6 Hz, 2H, H-1-; H-8-Fmoc), 7.43 – 7.36 (m, 2H, H-3-; H-6-Fmoc), 7.35 – 7.28 (m, 2H, H-2-; H-7-Fmoc), 5.65 (d, *J* = 9.5 Hz, 1H, N-H-Fmoc), 5.47 (d, *J* = 3.1 Hz, 1H, H-4), 5.34 (dd, *J* = 11.2, 3.3 Hz, 1H, H-3), 5.11 (d, *J* = 3.7

Hz, 1H, H-1), 4.50 – 4.20 (m, 5H, H-9-Fmoc; CH_2-Fmoc; T^{α}; H-5), 4.11 (t, *J* = 7.0 Hz, 3H, H-$6_{a,b}$; T^{β}), 3.64 (dd, *J* = 11.2, 3.7 Hz, 1H, H-2), 2.17 – 2.01 (m, 9H, 3 Ac), 1.51 (s, 9H, O*t*Bu), 1.35 (d, *J* = 6.4 Hz, 3H, T^{γ}).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-(2-acetamido-3,4,6-_O_-acetyl-2-deoxy-α-D-galactopyranosyl)-_L_-threonine-_tert_-butylester (10)[321]

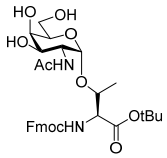


A solution of compound **9** (19.2 g, 27.0 mmol) in thioacetic acid/pyridine (1:1, 100 mL) was stirred at room temperature for 24 h. Afterwards the solvent was removed in vacuo by codistillation with toluene. The residue was taken up in cyclohexane/ethyl acetate (1:1, 200 mL) and precipitated sulfur was filtered off. The filtrate was purified by colum chromatography on silica (CH/EtOAc 2:1 → 1:1) to give compound **10** (13.2 g, 18.2 mmol, 67 %).

R_f = 0.54 (CH/EtOAc 1:2).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.78 (d, *J* = 7.5 Hz, 2H, H-4-; H-5-Fmoc), 7.64 (d, *J* = 7.5 Hz, 2H, H-1-; H-8-Fmoc), 7.44 – 7.38 (m, 2H, H-3-; H-6-Fmoc), 7.37 – 7.30 (m, 2H, H-2-; H-7-Fmoc), 5.98 (d, *J* = 9.8 Hz, 1H, N-H-Fmoc), 5.52 (d, *J* = 9.4 Hz, 1H, N-H), 5.41 – 5.36 (m, 1H, H-4), 5.09 (dd, *J* = 11.3, 3.1 Hz, 1H, H-2), 4.89 (d, *J* = 3.6 Hz, 1H, H-1), 4.67 – 4.56 (m, 1H, H-3), 4.52 – 4.40 (m, 1H, T^{α}), 4.32 – 4.02 (m, 7H, H-9-Fmoc; CH_2-Fmoc; H-5; H-$6_{a,b}$; T^{β}), 7.80 – 7.76 (m, 12H, 4 Ac), 1.46 (s, 9H, O*t*Bu), 1.32 (d, *J* = 6.4 Hz, 3H, T^{γ}).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-(2-acetamido-2-deoxy-α-D-galactopyranosyl)-_L_-threonine-_tert_-butylester (11)[302, 311]



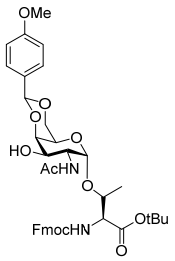
A solution of compound **10** (13.2 g, 18.2 mmol) in dry methanol (28 mL) was adjusted to *p*H 8 by adding a solution of sodium methoxide in methanol (1 %, w:w) and stirred for twelve days at room temperature, while maintaining a stable pH. After TLC indicated full conversion, the *p*H was

neutralized with acetic acid and the solvent was removed in vacuo. The residue was purified by column chromatography on silica (EtOAc/MeOH 25:1 → 15:1) to give compound **11** (8.8 g, 14.6 mmol, 80 %).

R_f = 0.24 (EtOAc/MeOH 15:1).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.76 (d, *J* = 7.6 Hz, 2H, H-4-; H-5-Fmoc), 7.60 (d, *J* = 7.4 Hz, 2H, H-1-; H-8-Fmoc), 7.43 – 7.35 (m, 2H, H-3-; H-6-Fmoc), 7.35 – 7.27 (m, 2H, H-2-; H-7-Fmoc), 6.93 (d, *J* = 7.8 Hz, 1H, N-H-Fmoc), 5.70 (d, *J* = 9.5 Hz, 1H, N-H), 4.88 (d, *J* = 3.9 Hz, 1H, H-1), 4.51 – 4.38 (m, 1H, H-3), 4.33 – 4.20 (m, 3H, H-2; T^{α}; H-9-Fmoc), 4.20 – 4.07 (m, 2H, H-4; T^{β}), 4.04 (d, *J* = 3.1 Hz, 1H, H-4), 3.96 – 3.76 (m, 5H, H-5; H-$6_{a,b}$; CH_2-Fmoc), 3.21 (s, 3H, hydroxyl-H), 2.11 (s, 3H, 1 Ac), 1.45 (s, 9H, O*t*Bu), 1.31 – 1.26 (m, 3H, T^{γ}).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-(2-acetamido-2-deoxy-4,6-_O-para_-methoxybenzylidene-α-D-galactopyranosyl)-_L_-threonine-_tert_-butylester (12)[153, 303]



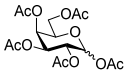
The *p*H of a solution of compound **11** (4.3 g, 7.1 mmol, 1.0 eq.) and *para*-anisaldehyde dimethylacetal (1.6 mL, 9.3 mmol, 1.3 eq.) in dry acetonitrile (42.8 mL) was adjusted to 4.5 using *para*-toluenesulfonic acid. After stirring for 3.5 h at room temperature, the *p*H was neutralized with triethylamine and the solvent was removed in vacuo by codistillation with toluene. The residue was purified by column chromatography on silica (Tol/EtOAc 1:1 → 1:2) to give compound **12** (4.8 g, 6.7 mmol, 94 %).

R_f = 0.20 (Tol/EtOAc 1:1).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.78 (d, *J* = 7.5 Hz, 2H, H-4-; H-5-Fmoc), 7.62 (d, *J* = 7.5 Hz, 2H, H-1-; H-8-Fmoc), 7.47 – 7.38 (m, 4H, H-2-; H-6-PMP; H-2-; H-7-Fmoc), 7.38 – 7.29 (m, 2H, H-3-; H-6-Fmoc), 6.89 (d, *J* = 8.6 Hz, 2H, H-3-; H-5-PMP), 6.46 (d, *J* = 8.4 Hz, 1H, N-H-Fmoc), 5.54 (s, 1H, CH(PMP)), 5.39 (d, *J* = 9.6 Hz, 1H, N-H), 4.96 (d, *J* = 3.5 Hz, 1H, H-1), 7.78 (d, *J* = 7.5 Hz, 2H, CH_2-Fmoc), 4.30 – 4.12 (m, 2H, H-$6_{a,b}$), 4.11 – 4.02 (m, 1H, H-3), 3.81 (s, 3H, OMe(PMP)), 3.72 – 3.70 (m, 1H, H-4), 2.09 (s, 3H, 1 Ac), 1.48 (s, 9H, O*t*Bu), 1.30 (d, *J* = 6.5 Hz, 3H, T^{γ}).

9.1.2.2 Synthesis of the type-1 N-acetyllactosamine glycosyl donor

1,2,3,4,6-Penta-*O*-acetyl-α/β-D-galactopyranoside (13)[153]

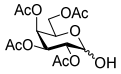


Acetic anhydride (793 mL, 7.8 mol, 20 eq.) and $HClO_4$ (3.67 mL, 64.7 mmol, 0.17 eq.) were cooled to 0° C. Afterwards D-galactose **4** (70.0 g, 388.5 mmol, 1.0 eq.) was added carefully and the solution was stirred until full conversion. The solution was washed with sat. $NaHCO_3$, concentrated in vacuo and codistilled with toluene. The crude was crystallized from diethyl ether (350 mL) to give compound **13**(136.0 g, 348.0 mmol, 89 %).

R_f = 0.60 (Tol/EtOAc 1:1).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ = 6.46 – 6.23 (m, 1H, H-1α), 5.68 (d, *J* = 8.3 Hz, 1H, H-1β), 5.48 (d, *J* = 1.3 Hz, 1H, H-4α), 5.40 (dd, *J* = 3.4, 1.1 Hz, 1H, H-4β), 5.33 – 5.27 (m, 3H, H-2α; H-2β; H-3α), 5.06 (dd, *J* = 10.4, 3.5 Hz, 1H, H-3β), 4.32 (t, *J* = 6.7 Hz, 1H, H-5α), 4.15 – 4.01 (m, 3H, H-5β, H-6αa,b; H-6βa,b), 2.19 – 1.87 (m, 15H, Ac-H).

2,3,4,6-Tetra-*O*-acetyl-α/β-D-galactopyranose (14)[153]



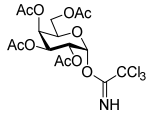
A solution of **13** (136.0 g, 348.4 mmol, 1.0 eq.) in dimethylformamide (550 mL) was heated to 50° C. After addition of hydrazine acetate (37.7 g, 418.1 mmol, 1.2 eq.) the solution was stirred until full conversion (3 h). The solvent was removed in vacuo and the residue was dissolved in ethyl acetate (1 L). After washing with sat. sodium bicarbonate solution, the aqueous layer was extracted three times with ethyl acetate. After concentrating the combined organic layers, compound **14** was obtained as an anomeric mixture (87.6 g 251.5 mmol, 72 %, α/β 3:1). The compound was subjected to follow up reactions without further purification.

R_f = 0.50 (Tol/EA 1:1).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ = 7.25 – 7.22 (m, toluene), 7.18 – 7.12 (m, toluene), 5.51 (t, *J* = 3.6 Hz, 1H, H-1α), 5.47 (dd, *J* = 3.4, 1.3 Hz, 1H, H-4α), 5.42 (d, *J* = 3.3 Hz, 1H, H-3β), 5.40 – 5.37 (m, 1H, H-3α), 5.15 (ddd, *J* = 10.8, 3.6, 1.0 Hz, 1H, H-2α), 5.09 – 5.04 (m, 2H, H-2β; H-4β), 4.69 (dd, *J* = 8.9, 7.8 Hz, 1H, H-1β), 4.46 (tdd, *J* = 6.5, 1.5, 0.6 Hz, 1H, H-5α), 4.15 (d, *J* = 6.5 Hz, 2H, H-6βa,b), 4.09 (dd, *J* = 6.6, 4.0

Hz, 2H, H-6αa,b), 3.95 (ddd, *J* = 6.8, 6.2, 1.2 Hz, 1H, H-5β), 3.78 (d, *J* = 8.9 Hz, 1H, β-OH), 3.37 (d, *J* = 3.6, 1.2 Hz, 1H, α-OH), 2.34 (s, toluene), 2.18 – 1.96 (m, 24H, Ac-H).

2,3,4,6-Tetra-*O*-acetyl-α-D-galactopyranosyl trichloroacetimidate (15)[397]

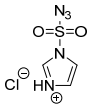


A solution of compound **14** (64.7 g, 185.8 mmol, 1.0 eq.) and trichloroacetonitrile (37.3 mL, 371.5 mmol, 2 eq.) in dry dichlormethane (650 mL) was cooled to 0° C. After addition of 1,8-Diazabicyclo(5.4.0)undec-7-ene (DBU) (5.56 mL, 37.2 mmol, 0.2 eq.) the solution was stirred for 2 h at 0° C. The solution was purified over a short silica column (pure DCM). The eluate was recrystallized from cyclohexane/ethyl acetate (7:1) at 40° C to give compound **15** (61.5 g, 124.8 mmol, 67 %).

R_f = 0.60 (CH/EtOAc 1:1). *ESI-MS*(*pos)*: m/z: 330.88 ([Gal oxonium]$^+$, calc.: 331.10).

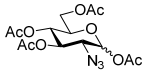
^{1}H-NMR (400 MHz, $CDCl_3$): δ = 8.66 (s, 1H, N-H), 6.60 (d, *J* = 3.5 Hz, 1H, 1-H), 5.56 (d, *J* = 3.0 Hz, 1H, H-4), 5.51 – 5.32 (m, 2H, H-2; H-3), 4.44 (t, *J* = 6.7 Hz, 1H, H-5), 4.21 – 4.01 (m, 2H, H-$6_{a,b}$), 2.16 (s, 3H, Ac-H), 2.04 – 2.00 (m, 9H, Ac-H).

Imidazole-1-sulfonyl azide hydrochloride (16)[328]



A suspension of sodium azide (37.7 g, 577 mmol, 1.0 eq.) in acetonitrile (750 mL) was cooled on ice before adding sulfuryl chloride (47 mL, 577 mmol, 1.0 eq.) dropwise. The solution was stirred for 18 h at room temperature and afterwards the solution was cooled to 0° C again followed by addition of imidazole (74.7 g, 1096 mmol, 1.9 eq.) in portions. After stirring for 4 h at room temperature, the solution was diluted with ethyl acetate (1.2 L) and washed twice with water (400 mL each) and once with sat. bicarbonate solution. The organic layer was dried over sodium sulfate and cooled on ice. Meanwhile a hydrochloric acid solution in methanol was prepared by adding acetyl chloride (70 mL) dropwise to dry methanol (240 mL) on ice. The latter solution was then added dropwise to the ice cooled organic layer and the resulting solution was stored overnight at -20° C. Precipitated solid was filtered off and washed with ethyl acetate to give imidazole-1-sulfonyl azide **16** (58.9 g, 281 mmol, 49 %).

1,3,4,6-Tetra-*O*-acetyl-2-azido-2-deoxy-α/β-D-glucopyranoside (18)[328]

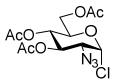


To a suspension of potassium carbonate (66.3 g, 480.0 mmol, 2.00 eq.) and copper(II) sulfate pentahydrate (380 mg, 2.4 mmol, 0.01 eq.) in MeOH (635 mL), first glucosamine hydrochloride **17** (48.4 g, 240.0 mmol, 1.00 eq.) and then imidazole-1-sulfonyl azide **16** (54.3 g, 259.2 mmol, 1.08 eq.) were added, followed by mechanical stirring for 4 h at room temperature. Afterwards the solvent was evaporated in vacuo following by codistillation with toluene three times. The residue was taken up in pyridine (480 mL) and acetic anhydride was added. After 12 h of mechanical stirring at room temperature, the solvent volume was halved by codistillation with toluene and water (500 mL) was added. The aqueous phase was extracted four times with ethyl acetate, the organic phase was dried over sodium sulfate and the solvent was removed in vacuo. After purification of the residue via column chromatography over silica (CH/EtOAc 2:1), compound **18** was obtained as a yellow liquid (72.7 g, 195.0 mmol, 81 %, α/β 1:2).

R_f = 0.59 (Tol/EtOAc 2:1).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 6.28 (d, *J* = 3.7 Hz, 1H, H-1α), 5.54 (d, *J* = 8.6 Hz, 1H, H-1β), 5.47 – 5.39 (m, 1H, H-3α), 5.14 – 4.97 (m, 3H, H-4α; H-3β; H-4β), 4.31 – 4.23 (m, 1H, H-6_a), 4.15 – 4.00 (m, 2H, H-6_b, H-5β), 3.83 – 3.75 (m, 1H, H-5β), 3.67 – 3.62 (m, 2H, H-2α; H-2β), 2.20 – 1.98 (m, 12H, 4 Ac).

3,4,6-Tri-*O*-acetyl-2-azido-2-deoxy-α-D-glucopyranosyl chloride (19)[318, 332-333]



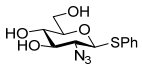
A solution of compound **18** (87.1 g, 233.3 mmol, 1.0 eq.) in chloroform (650 mL) was cooled on ice before adding titanium(IV) chloride (51.2 mL, 466.6 mmol, 2.0 eq.) dropwise while stirring. Meanwhile precipitation of solid titanium(IV) acetate was observed. The ice bath was removed and the solution was stirred additional 30 min at room temperature before stirring for 45 min at reflux. Thereby the solid dissolved at 63.5° C. The solution was diluted with dichlormethane, washed once with water, twice with sat. bicarbonate solution and once with brine and finally dried over sodium sulfate. The solvent was removed in vacuo and the residue was purified via column chromatography on silica (CH/EtOAc 2:1) to give compound **19** (67.5 g, 193.0 mmol, 83 %).

R_f = 0.58 (Tol/EtOAc 3:1). *ESI-MS*(*pos)*: m/z: 623.20 ($[M+Na]^+$, calc.: 623.26).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 6.11 (d, *J* = 3.8 Hz, 1H, H-1), 5.55 – 5.46 (m, 1H, H-3), 5.13 – 5.05 (m, 1H, H-4), 4.36 – 4.27 (m, 2H, H-6_a, H-5), 4.16 – 4.05 (m, 2H, H-6_b), 3.85 (dd, *J* = 10.3, 3.8 Hz, 1H, H-2), 2.11 – 2.02 (m, 9H, 3 Ac).

^{13}C-NMR: δ = 170.5, 169.8, 169.7 (3 C=O(Ac)), 91.7 (C-1), 70.8, 70.8 (C-3; C-5), 67.7 (C-4), 62.4 (C-2), 61.2 (C-6), 20.7, 20.7, 20.7 (3 Me(Ac)).

Phenyl 2-azido-2-deoxy-1-thio-β-D-glucopyranoside (20)[334]



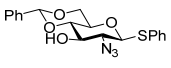
After a solution of thiophenol (23.4 mL, 212.3 mmol, 1.1 eq.) and potassium hydroxide (11.9 g 212.3 mmol, 1.1 eq.) in ethanol (390 mL) was prepared, a solution of compound **19** (67.5 g, 193.0 mmol, 1.0 eq.) in chloroform (390 mL) was added to the former dropwise. After stirring for 3 h at room temperature, the solution was diluted with dichlormethane, extracted twice with sat. sodium bicarbonate and the organic layer was dried over sodium sulfate. The solvent was removed in vacuo and the residue was dissolved in a suspension of sodium methoxide in methanol (390 mL, 0.225 M, *p*H = 9.5). The reaction mixture was stirred for 2 h at room temperature before neutralizing with Dowex 50WX8 ion exchange resin. The solvent was removed in vacuo and the residue was purified via column chromatography over silica (CH/EtOAc 1:3) to give compound **20** (36.0 g, 121.1 mmol, 63 %).

R_f = 0.56 (EtOAc/MeOH 15:1).

^{1}H-NMR (400 MHz, DMSO), *δ (ppm)*: 7.51 (d, *J* = 7.4 Hz, 2H, H-2-; H-6-SPh), 7.41 – 7.26 (m, 3H, H-3-; H-4-; H-5-SPh), 5.66 (d, *J* = 6.0 Hz, 1H, OH-3), 5.20 (d, *J* = 5.4 Hz, 1H, OH-4), 4.70 (d, *J* = 10.3 Hz, 1H, H-1), 4.67 – 4.58 (m, 1H, OH-6), 3.75 – 3.64 (m, 1H, H-6_a), 3.52 – 3.42 (m, 1H, H-6_b), 3.42 – 3.32 (m, 1H, H-3), 3.32 – 3.22 (m, 1H, H-5), 3.23 – 3.09 (m, 2H, H-2; H-4).

^{13}C-NMR: δ = 132.9, 131.0, 129.0 127.3 (Aryl-C), 84.3 (C-1), 81.0 (C-5), 76.6 (C-3), 69.6 (C-4), 65.4 (C-2), 60.8 (C-6).

Phenyl 2-azido-2-deoxy-3,6-*O*-benzylidene-1-thio-β-D-glucopyranoside (21)[335]



A solution of compound **20** (19.3 g, 64.8 mmol, 1.0 eq.) and benzaldehyde dimethylacetal (14.6 mL, 97.2 mmol, 1.5 eq.) in acetonitrile (190 mL) was heated to 50° C and adjusted to *p*H 2 with *para*-

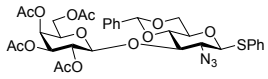
toluene sulfonic acid (aim was *p*H 4.5). After stirring for 2 h, TLC showed full conversion. The *p*H was neutralized with trimethylamine and the solution was washed twice with sat. sodium bicarbonate. The solvent volume was reduced in vacuo and the product was crystallized from EtOAc/CH. Remaining product in the mother liquid was recrystallized again to give compound **21** (15.6 g, 40.6 mmol, 63 %).

R_f = 0.80 (Tol/EtOAc 2:1).

^{1}H-NMR (400 MHz, DMSO), *δ (ppm)*: 7.53 – 7.47 (m, 2H, H-2-; H-6-SPh), 7.47 – 7.31 (m, 3H, H-3-; H-4-; H-5-SPh), 6.04 (d, *J* = 6.0 Hz, 1H, O-H), 5.63 (s, 1H, C-H-Bn), 4.96 (d, *J* = 10.3 Hz, 1H, H-1β), 4.24 (dd, *J* = 9.9, 4.7 Hz, 1H, H-6_a), 3.78 – 3.67 (m, 1H, H-6_b; H-3), 3.66 – 3.56 (m, 1H, H-5), 3.52 – 3.42 (m, 1, H-4), 3.39 – 3.26 (m, 1H).

^{13}C-NMR: δ = 137.5, 131.9, 131.4, 131.4, 129.2, 129.2, 128.9, 128.0, 128.0, 127.7, 126.3, 126.3 (Aryl-C), 100.7 (C-H-Bn), 84.7 (C-1), 80.1 (C-4), 72.8 (C-3), 69.6 (C-5), 67.5 (C-6), 65.8 (C-2).

Phenyl 3-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-2-azido-4,6-*O*-benzylidene-2-deoxy-1-thio-β-D-glucopyranoside (22)[324]



After dissolving glycosyl acceptor **21** (15.6 g, 40.5 mmol, 1.0 eq.) and glycosyl donor **15** (19.9 g, 40.5 mmol, 1.0 eq.) in dry diethyl ether (350 mL) under argon, a solution of TMSOTf (730 µL, 4.0 mmol, 0.1 eq.) in diethyl ether (8 mL) was added dropwise and the reaction solution was stirred for 30 min at room temperature. The *p*H was neutralized with trimethylamine before washing twice with sat. sodium bicarbonate solution and once with brine. The solution was dried over sodium sulfate and the solvent was removed in vacuo. The residue was purified by column chromatography over silica (CH/EtOAc 2:1) to give compound **22** as a solid (28.9 g, 40.5 mmol, quant.).

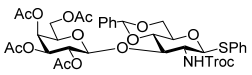
R_f = 0.42 (CH/EtOAc 2:1).). *HR-ESI-MS (pos), m/z*: 733.2387 ($[M+NH_4]^+$, calc.: 733.2391).

^{1}H-NMR (400 MHz, DMSO), *δ (ppm)*: 7.54 – 7.31 (m, 10H, H-2-; H-3-; H-4-; H-5-; H-6-SPh; H-2-; H-3-; H-4-; H-5-; H-6-Bn), 5.67 (s, 1H, C-H-Bn), 5.27 – 5.19 (m, 2H, H-3'; H-4'), 5.11 – 5.02 (m, 2H, H-1; H-1'), 5.01 – 4.91 (m, 1H, H-2'), 4.29 – 4.17 (m, 2H, H-5'; H-6_a), 4.05 – 3.96 (m, 1H, H-6'$_a$), 3.94 – 3.86 (m, 1H, H-5'), 3.84 – 3.73 (m, 3H, H-4; H-6_b; H-6'$_b$), 3.72 – 3.67 (m, 1H, H-5), 3.66 – 3.57 (m, 1H, H-2), 2.14 – 1.81 (m, 12H, 4 Ac).

^{13}C-NMR: δ = 169.9, 169.7, 169.5, 169.1 (4 C=O(Ac)), 162.9 (C-1(Bn)), 137.4 (C-1(SPh)), 132.0, 131.1, 131.1, 129.3, 129.3, 128.8, 127.9, 127.9, 126.1, 126.1 (10 Aryl-C), 100.1 (CH-Bn), 98.7 (C-1'), 86.0 (C-

1), 79.5 (C-3), 76.9 (C-4), 70.3 (C-3'), 69.8 (C-5'), 69.4 (C-5), 68.9 (C-2'), 67.5 (C-6), 67.2 (C-4'), 64.6 (C-2), 61.2 (C-6'), 26.3, 20.4, 20.3, 20.3 (4 Me(Ac)).

Phenyl 3-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-4,6-*O*-benzylidene-2-deoxy-1-thio-2-*N*-(2,2,2-trichloroethoxycarbonyl)-β-D-glucopyranoside (23)[323]



A solution of compound **22** (29.7 g, 41.5 mmol, 1.0 eq.) in dioxane/acetic acid (10:1, 340 mL) was cooled on ice before adding zinc dust (16.3 g, 249.2 mmol, 6.0 eq.). The suspension was stirred overnight, while allowing to temperate to room temperature. Afterwards additional zinc dust (5.0 g) was added and stirring was continued until TLC showed full conversion (R_f = 0 with Tol/EtOAc 3:1). The zinc was filtered off over celite on filter paper, which was washed with ethyl acetate (500 mL) afterwards. The filtrate was diluted with additional ethyl acetate (500 mL) and washed twice with sat. sodium bicarbonate solution and once with brine. After drying over sodium sulfate the solvent was removed in vacuo. The residue was dissolved in dioxane (300 mL) and sodium bicarbonate (8.0 g, 95.5 mmol, 2.3 eq.) in water (150 mL) was added. The resulting suspension was cooled on ice and TrocCl (6.9 mL, 49.9 mmol, 1.2 eq.) in dioxane (62 mL) was added dropwise. The ice bath was removed and stirring was continued for 3 h. Afterwards the dioxane was removed in a rotary evaporator and the remaining aqueous solution was extracted with three times with ethyl acetate (170 mL each). The combined organic layers were washed with water and brine before drying over sodium sulfate. The solvent was removed in vacuo and the residue was purified by column chromatography over silica (DCM/EtOAc 25:1 → 10:1) to give compound **23** as a solid (20.5 g, 23.7 mmol, 57 %).

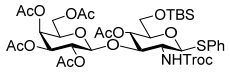
R_f = 0.26 (CH/EtOAc 2:1). *HR-ESI-MS (pos), m/z*: 864.1261 ($[M+H]^+$, calc.: 864.1257), 881.1525 ($[M+NH_4]^+$, calc.: 881.1522).

^{1}H-NMR (400 MHz, DMSO), *δ (ppm)*: 7.97 (d, *J* = 9.6 Hz, 1H, N-H), 7.57 – 7.21 (m, 10H, 10 Aryl-H), 5.68 (s, 1H, C-H-Bn), 5.22 (d, *J* = 3.6 Hz, 1H, H-4'),5.14 (dd, *J* = 10.2, 3.6 Hz, 1H, H-3'),5.04 (d, *J* = 12.3 Hz, 1H, CH_{2a}-Troc), 4.99 – 4.88 (m, 2H, H-1; H-2'),7.97 (d, *J* = 9.6 Hz, 1H, H-1'),4.65 (d, *J* = 12.3 Hz, 1H, CH_{2b}-Troc), 4.23 (dd, *J* = 10.1, 4.9 Hz, 1H, H-6_a), 4.11 – 4.05 (m, 1H, H-6_a), 4.03 – 3.93 (m, 3H, H-5'; H-$6'_a$; H-5'),3.85 – 3.74 (m, 2H, H-6_b; H-$6'_b$), 3.72 – 3.65 (m, 1H, H-4), 3.65 – 3.56 (m, 1H, H-2), 3.56 – 3.48 (m, 1H, H-5), 2.14 – 1.83 (m, 12H, 4 Me(Ac)).

^{13}C-NMR: δ = 169.8, 169.7, 169.4, 169.3 (4 C=O(Ac)), 154.2 (C=O(Troc)), 137.5 (C-1-SPh), 133.8 (C-1-Bn), 130.0, 130.0, 129.1, 129.1, 128.6, 127.8, 127.8, 127.0, 126.0, 126.0 (10 Aryl-C), 100.0 (C-1'), 99.7

(C-H-Bn), 96.0 (CCl_3-Troc), 87.0 (C-1), 79.0 (C-3), 78.3 (C-4), 73.6 (CH_2-Troc), 70.3 (C-3‘), 69.9 (C-5), 69.5 (C-5‘), 48.7 (C-2‘), 67.5 (C-6), 67.1 (C-4‘), 60.8 (C-6‘), 55.7 (C-2), 20.4, 20.3, 20.3, 20.3 (4 Me(Ac)).

Phenyl 3-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-4-*O*-acetyl-6-*O*-*tert*-butyldimethylsilyl-2-deoxy-1-thio-2-*N*-(2,2,2-trichloroethoxycarbonyl)-β-D-glucopyranoside (24)[323]



A solution of compound **23** (9.1 g, 10.5 mmol, 1.0 eq.) in acetic acid/water (4:1, 90 mL) was stirred for 10 h at 50° C. Afterwards the solvent was removed by codistillation with toluene and the residue was taken up in dry dimethylformamide (80 mL) under argon. Afterwards first imidazole (1.8 g, 26.3 mmol, 2.5 eq.) and then TBDMS chloride (1.8 g, 11.6 mmol, 1.1 eq.) was added. Stirring was continued at room temperature overnight before removing the solvent in vacuo at 40° C. The residue was dissolved in ethyl acetate, washed twice with water, twice with sat. sodium bicarbonate and twice with brine before drying the organic layer over sodium sulfate. The solvent was removed in vacuo and the residue was taken up in acetic anhydride/pyridine (1:2, 80 mL). After stirring overnight, the solvent was removed by codistillation with toluene and the residue was purified by column chromatography on silica (CH/EtOAc 2:1) to give compound **24** as a solid (10.0 g, 10.5 mmol, quant.).

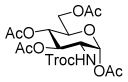
R_f = 0.54 (Tol/EtOAc 3:1). *HR-ESI-MS (pos), m/z*: 932.924 ($[M+H]^+$, calc.: 932.1914).

^{1}H-NMR (400 MHz, DMSO), *δ (ppm)*: 7.94 (d, *J* = 9.6 Hz, 1H, N-H), 7.43 – 7.39 (m, 2H, H-2-; H-6-SPh), 7.34 – 7.23 (m, 3H, H-3-; H-4-; H-5-SPh), 5.23 (d, *J* = 3.7 Hz, 1H, H-4’), 5.04 (dd, *J* = 9.8, 3.6 Hz, 1H, H-3’), 5.00 (d, *J* = 12.3 Hz, 1H, CH_{2a}-Troc), 4.85 – 4.75 (m, 3H, H-2’, H-1; H-1’), 4.75 – 4.64 (m, 2H, H-4; CH_{2b}-Troc), 4.15 – 4.05 (m, 2H, H-5’; H-6’$_a$), 4.02 – 3.87 (m, 2H, H-6’$_b$; H-3), 3.70 – 3.45 (m, 4H, H-6$_{a,b}$; H-5; H-2), 2.14 – 1.85 (m, 12H, 4 Me(Ac)), 0.85 (s, 9H, Si*t*Bu), 0.06 – -0.06 (m, 6H, 2 SiMe).

^{13}C-NMR (125.8 MHz, DMSO), : δ = 169.8, 169.7, 169.3, 169.1, 168.8 (5 C=O(Ac)), 154.2 (C=O(Troc)), 134.0 (C-1-SPh), 129.9, 129.9, 128.9, 128.9, 126.9 (C-2-; C-3-; C-4-; C-5-; C-6-SPh), 100.0 (C-1’), 95.9 (CCl_3-Troc), 86.2 (C-1), 78.8 (C-3), 77.8 (C-5), 73.5 (CH_2-Troc), 70.3 (C-3’), 69.6 (C-5’), 68.8 (C-2’), 68.5 (C-4), 67.1 (C-4’), 62.1 (C-6), 60.9 (C-6’), 55.8 (C-2), 20.5, 20.4, 20.3 20.3, 20.2 (5 Me(Ac)), -5.5, -5.5 (2 SiMe).

9.1.2.3 Synthesis of the type-2 N-acetyllactosamine glycosyl donor

1,3,4,6-Tetra-*O*-acetyl-2-deoxy-2-*N*-(2,2,2-trichloroethoxycarbonyl)-α/β-D-glucopyranoside (25)[338]

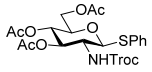


To a solution of D-glucosamine hydrochloride (30.0 g, 139.1 mmol, 1.0 eq.) in water/dioxane (2:1, 300 mL), sodium bicarbonate (32.4 g, 385.1 mmol, 2.8 eq.) was added slowly. Then the solution was cooled on ice before adding Troc chloride (30.0 mL, 217.7 mmol, 1.6 eq.) dropwise. Stirring was continued overnight while the system was allowed to warm up to temperate to room temperature. Afterwards the mixture was filtrated over a POR3-filter with subsequent washing with water/dioxane (2:1, 300 mL). The filtrate was cleared from dioxane in vacuo and the remaining aqueous solution was extracted five times with ethyl acetate (100 mL, each). The organic layers were combined with the filter cake and the solvent was removed in vacuo. The residue was then dissolved in acetic anhydride/pyridine (1:2, 300 mL) and stirred overnight at room temperature. Afterwards the solvent was removed by codistillation with toluene and the residue was purified by column chromatography on silica (CH/EtOAc 1:1) to give compound **25** (54.2 g, 103.7 mmol, 75 %).

R_f = 0.54 (CH/EtOAc 1:1).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 6.25 (d, *J* = 3.6 Hz, 1H, H-1), 5.33 – 5.24 (m, 1H, H-3), 5.24 – 5.16 (m, 1H, H-4), 4.86 – 4.60 (m, 2H, CH_2-Troc), 4.32 – 4.25 (m, 1H, H-6_a), 4.25 – 4.17 (m, 1H, H-2), 4.11 – 3.99 (m, 3H, H-5; H-6_b), 2.25 – 1.98 (m, 12H, 4 Ac).

Phenyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-*N*-(2,2,2-trichloroethoxycarbonyl)-β-D-glucopyranoside (26)[337]

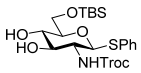


A solution of compound **25** (41.6 g, 79.6 mmol, 1.0 eq.) and thiophenol (10.6 mL, 103.5 mmol, 1.3 eq.) in dichloromethane (225 mL) was stirred for 45 min at 0° C. Then boron trifluoride diethyl etherate (14.7 mL, 103.5 mmol, 1.3 eq.) was added dropwise and stirring was continued for 4 h. Afterwards the reaction was stopped by adding trimethylamine (26 mL) before recrystallization from methanol. Remaining product in the mother liquid was purified by column chromatography on silica (CH/EtOAc 2:1) and combined with the crystallized product. In total 42.7 g of compound **26** were obtained (74.5 mmol, 94 %).

R_f = 0.65 (CH/EtOAc 1:1).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.53 – 7.42 (m, 2H, H-2-; H-6-SPh), 7.31 – 7.26 (m, 3H, H-3-; H-4-; H-5-SPh), 5.51 (d, *J* = 9.0 Hz, 1H, N-H), 5.31 – 5.23 (m, 1H, H-3), 5.05 – 4.95 (m, 1H, H-4), 4.86 (d, *J* = 10.4 Hz, 1H, H-1), 4.82 – 4.66 (m, 2H, CH_2-Troc), 4.27 – 4.01 (m, 2H, H-$6_{a,b}$), 3.77 – 3.66 (m, 2H, H-2; H-5), 2.09 – 1.94 (m, 9H, 3 Ac).

Phenyl 6-*O*-tert-butyldimethylsilyl-2-deoxy-2-*N*-(2,2,2-trichloroethoxycarbonyl)-β-D-glucopyranoside (27)[296]

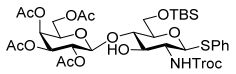


First dry methanol (440 mL) was cooled on ice, acetyl chloride (44.4 mL) was added dropwise and stirring was continued for 15 min. Then a solution of compound **26** (42.7 g, 74.5 mmol, 1.0 eq.) in dichloromethane (110 mL) was added dropwise. The solution was stirred overnight at room temperature before removing the solvent in vacuo. The residue was dissolved in ethyl acetate (800 mL) and washed once with water, twice with sat. sodium bicarbonate solution and once with brine (300 mL each). Finally the solution was dried over sodium sulfate and the solvent was removed in vacuo. The residue was then dissolved in dry dimethylformamide (300 mL) before adding TBDMSCl (11.0 g, 73.13 mmol, 1.0 eq.) and imidazole (11.3 g, 166.2 mmol, 2.5 eq.). The solution was stirred for 3 h at room temperature und argon and then the solvent was removed in vacuo at 50° C. The residue was diluted with ethyl acetate (500 mL) and washed twice with water, twice with sat. sodium bicarbonate solution and once with brine before drying over sodium sulfate. After evaporation of the solvent, the residue was purified by column chromatography on silica (DCM/MeOH 10:1) to give compound **27** (35.3 g, 62.9 mmol, 85 %).

R_f = 0.81 (DCM/MeOH 10:1). *ESI-MS*(*pos)*: m/z: 449.87 ([oxonium]$^+$, calc.: 450.07), 563.60 ([M+H]$^+$, calc.: 564.08).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.52 – 7.47 (m, 2H, H-2-; H-6-SPh), 7.32 – 7.26 (m, 2H, H-3-; H-4-; H-5-SPh), 5.60 (d, *J* = 8.2 Hz, 1H, N-H), 4.84 – 4.67 (m, 2H, H-1; CH_2-Troc), 3.93 – 3.87 (m, 2H, H-$6_{a,b}$), 3.79 – 3.64 (m, 1H, H-4), 3.58 – 3.51 (m, 1H, H-3), 3.44 – 3.38 (m, 2H, H-2; H-5), 0.91 (s, 9H, Si*t*Bu), 0.16 – 0.03 (m, 6H, 2 SiMe).

Phenyl 4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-6-*O*-tert-butyldimethylsilyl-2-deoxy-2-*N*-(2,2,2-trichloroethoxycarbonyl)-β-D-glucopyranoside (28)

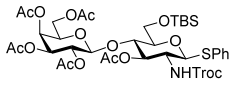


A slurry of galactopyranosyl trichloroacetimidate **15** (10.0 g, 20.3 mmol, 1.2 eq.), glucopyranosyl thioglycoside **27** (9.5 g, 16.9 mmol, 1.0 eq.) and molecular sieve (9.7 g, 4 Å) in dichloromethane (240 mL) was stirred for 1 h and thereby cooled to -50° C. Then trifluoromethanesulfonic acid (920 µL, 10.5 mmol, 0.6 eq.) was added dropwise and stirring was continued until TLC indicated full consumption of the donor (1 h) while keeping the temperature between -50° C and -35° C. Afterwards the *p*H was neutralized with *N*,*N*-diisopropylethylamine, the molecular sieve was filtered off and the solution was washed twice with sodium bicarbonate and once with brine. After drying over sodium sulfate the solvent was removed in vacuo and the residue was purified by column chromatography on silica (Tol/EtOAc 3:1 → 1:1) to give compound **28** (11.7 g, 13.2 mmol, 78 %).

R_f = 0.68 (Tol/EtOAc 2:1). *ESI-MS*(*pos)*: m/z: 891.44 ($[M+H]^+$, calc.: 890.18).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.52 – 7.46 (m, 2H, H-2-; H-6-SPh), 7.30 – 7.25 (m, 3H, H-3-; H-4-; H-5-SPh), 5.38 (d, *J* = 3.5 Hz, 1H, H-4'), 5.25 – 5.14 (m, 2H, H-2'; N-H), 4.97 (dd, *J* = 10.5, 3.4 Hz, 1H, H-3'), 4.86 (d, *J* = 10.4 Hz, 1H, H-1), 4.80 – 4.66 (m, 2H, CH_2-Troc), 4.61 (d, *J* = 8.0 Hz, 1H, H-1'), 4.13 (d, *J* = 6.5 Hz, 2H, H-6'$_{a,b}$), 4.01 – 3.94 (m, 1H, H-5), 3.87 – 3.79 (m, 2H, H-6$_a$; H-3), 3.71 (dd, *J* = 11.5, 3.7 Hz, 1H, H-6$_b$), 3.62 (dd, *J* = 9.7, 8.4 Hz, 1H, H-4), 3.44 – 3.32 (m, 2H, H-2; H-5), 2.21 – 1.92 (m, 12H, 4 Ac), 0.92 (s, 9H, Si*t*Bu), 0.17 – 0.02 (m, 6H, 2 SiMe).

Phenyl 3-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-6-*O*-tert-butyldimethylsilyl-2-deoxy-2-*N*-(2,2,2-trichloroethoxycarbonyl)-β-D-glucopyranoside (29)[296]



A solution of compound **28** (10.6 g, 11.9 mmol) in acetic anhydride/pyridine (1:2, 100 mL) was stirred overnight at room temperature. Afterwards the solvent was removed by codistillation with toluene and the residue was purified by column chromatography on silica (Tol/EtOAc 4:1) to give compound **29** (11.0 g, 11.8 mmol, quant).

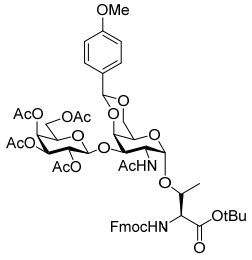
R_f = 0.60 (Tol/EtOAc 3:1). *ESI-MS*(*pos)*: m/z: 491.76 ([GlcNTroc oxonium]$^+$, calc.: 492.08), 821.74 ([LacNAc oxonium]$^+$, calc.: 822.17), 330.90 ([Gal oxonium]$^+$, calc.: 330.10), 933.14 ($[M+H]^+$, calc.: 933.19), 950.62 ($[M+Na]^+$, calc.: 950.22).

^{1}H-NMR (500 MHz, $CDCl_3$), *δ (ppm)*: 7.53 – 7.47 (m, 2H, H-2-; H-6-SPh), 7.33 – 7.21 (m, 3H, H-3-; H-4-; H-5-SPh), 5.34 (d, $J_{H4',H3'}$ = 2.8 Hz, 1H, H-4'), 5.26 (d, $J_{NH,H2}$ = 9.5 Hz, 1H, N-H), 5.12 – 5.03 (m, 2H, H-2'; H-3), 4.91 (dd, $J_{H3',H2'}$ = 10.4, $J_{H3',H4'}$ = 3.4 Hz, 1H, H-3'), 4.75 (q, *J* = 12.0 Hz, 2H, CH_2-Troc), 4.73 – 4.65 (m, 2H, H-1; H-1'), 4.14 – 4.07 (m, 2H, H-6'$_{a,b}$), 3.97 – 3.85 (m, 2H, H-4; H-6$_a$), 3.84 – 3.78 (m, 2H, H-6$_b$; H-5'), 3.70 (q, *J* = 9.9 Hz, 1H, H-2), 3.37 – 3.31 (m, 1H, H-5), 2.19 – 1.94 (m, 15H, 5 Ac), 0.93 (s, 9H, Si*t*Bu), 0.16 – 0.08 (m, 6H, 2 SiMe).

^{13}C-NMR (125.8 MHz, DMSO): δ = 170.8, 170.5, 170.3, 170.3, 169.0 (C=O(Ac)), 154.3 (C=O(Troc)), 132.9, 132.6 (C-1-; C-2-; C-6-SPh), 129.0, 128.4 (C-3-; C-4-; C-5-SPh), 100.4 (C-1'), 95.6 (CCl_3(Troc)), 87.1 (C-1), 79.7 (C-5), 100.4 (C-1), 79.7 (C-5), 74.6 (CH_2(Troc)), 74.0 (C-4), 73.7 (C-3), 71.2 (C-3'), 70.7 (C-5'), 69.3 (C-2'), 67.0 (C-4'), 61.2 (C-6), 55.2 (C-2), 26.0 (Me(*t*Bu)), 21.0, 20.9, 20.8, 20.8, 20.7 (Me(Ac)), 18.4 (C_q(*t*Bu)), -4.9, -5.1 (Me(*t*Bu)).

9.1.2.4 Synthesis of the T-antigen amino acid acceptor

_N_-9-Fluorenylmethyloxycarbonyl-_O_-(2-acetamido-3-_O_-[2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl]-2-deoxy-4,6-_O_-_para_-methoxybenzylidene-α-D-galactopyranosyl)-_L_-threonine-_tert_-butylester (30)[304-306]



A slurry of molecular sieve (4.0 g, 4 Å) and of T_N-antigen acceptor **12** (1.0 g, 1.4 mmol, 1.0 eq.) in dry nitromethane/dichloromethane (2:1, 45 mL) was stirred for 25 min under argon at room temperature. Meanwhile dry mercury(II) cyanide (1.1 g, 4.2 mmol, 3.0 eq.) was codistilled five times with dry toluene. Afterwards the mercury(II) cyanide was added to the slurry and the resulting mixture was stirred additional 20 min. Next, galactosyl bromide **8** (1.7 g, 4.2 eq.) was dissolved in dry dichloromethane (5 mL) and added dropwise to the reaction mixture. While stirring was continued, additional mercury(II) cyanide (351 mg) were added after one and three hours each, before the reaction was stirred for more 18 h. Finally, the slurry was diluted with dichloromethane (200 mL) and filtrated over celite with subsequent washing with dichloromethane. The filtrate was washed twice with sat. sodium bicarbonate solution (300 mL each) and twice with sodium iodide solution (10 , w:w, 150 mL each). After drying over sodium sulfate the solvent was removed in vacuo and the residue was

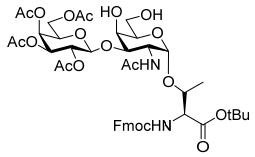
purified by column chromatography on silica (Tol/EtOAc 2:1 → 1:1) to give compound **30** (1.06 g, 1.0 mmol, 72 %).

R_f = 0.64 (Tol/EtOAc 1:3). *ESI-MS*(*pos)*: m/z: 1048.97 ($[M+H]^+$, calc.: 1049.41).

^{1}H-NMR (600 MHz, DMSO), *δ (ppm)*: 7.93 – 7.87 (m, 2H, H-4-; H-5-Fmoc), 7.77 – 7.72 (m, 2H, H-1-; H-8-Fmoc), 7.50 (d, $J_{NH,H2}$ = 10.0 Hz, 1H, N-H), 7.47 – 7.39 (m, 3H, H-3-; H-6-; N-H-Fmoc), 7.36 – 7.28 (m, 4H, H-2-; H-6-PMP; H-2-; H-7-Fmoc), 6.96 – 6.92 (m, 2H, H-3-; H-5-PMP), 5.47 (s, 1H, CH(PMP)), 5.29 (d, $J_{H4',H3'}$ = 3.4 Hz, 1H, H-4'), 5.04 (dd, $J_{H3',H2'}$ = 10.4, $J_{H3',H4'}$ = 3.6 Hz, 1H, H-3'), 4.92 (dd, $J_{H2',H3'}$ = 10.4, $J_{H2',H3'}$ = 8.0 Hz, 1H, H-2'), 4.76 (d, $J_{H1',H2'}$ = 8.1 Hz, 1H, H-1'), 4.71 (d, $J_{H1,H2}$ = 3.8 Hz, 1H, H-1), 4.55 – 4.46 (m, 2H, CH_2-Fmoc), 4.33 – 4.30 (m, 2H, H-9-Fmoc; H-3), 4.29 – 4.24 (m, 1H, T^{β}), 4.24 – 4.19 (m, 1H, H-2), 4.19 – 4.10 (m, 2H, H-5'; H-6'$_a$), 4.09 (dd, $J_{T\alpha,NH\text{-}Fmoc}$ = 9.7, $J_{T\alpha,T\beta}$ = 2.0 Hz, 1H, T^{α}), 4.07 – 3.97 (m, 3H, H-6'b; H-6$_{a,b}$), 3.78 (dd, $J_{H5,H6a}$ = 11.3, $J_{H5,H6b}$ = 3.4 Hz, 1H, H-5), 3.76 (s, 1H, OMe(PMP)), 3.68 (s, 1H, H-4), 2.13 – 1.78 (m, 15H, 5 Ac), 1.36 (s, 9H, O*t*Bu), 1.14 (d, $J_{T\gamma,T\beta}$ = 6.4 Hz, 3H, T^{γ}).

^{13}C-NMR (150.9 MHz, DMSO): δ = 170.0, 169.9, 169.5, 169.2, 168.8, 168.9 (5 C=O(Ac); 1 C=O(COO*t*Bu)), 159.6 (C_{para}(PMP), 156.8 (C=O(Fmoc), 143.7, 143.7 (C-1$_a$-; C-8$_a$-Fmoc), 140.8, 140.8 (C-4$_a$-; C-5$_a$-Fmoc), 130.8 (C_{arom}(PMP)), 128.9, 128.2, 127.7, 127.7, 127.5, 127.1, 127.0 (C-2-; C-3-; C-6-; C-7-Fmoc; 3x C_{arom}(PMP)), 125.2, 125.1 (C-1-; C-8-Fmoc), 120.2, 120.2 (C-4-; C-5-Fmoc), 113.4 (2x C_{arom}(PMP)), 101.4 (C-1'), 99.9 (C-H(PMP)), 99.1 (C-1), 81.4 (C_q(*t*Bu)), 75.3 (C-5), 75.0 (C-3), 73.5 (T^{β}), 70.4 (C-3'), 69.9 (C-5'), 68.5 (C-6), 68.3 (C-2'), 67.3 (C-4'), 65.6 (CH_2-Fmoc), 62.8 (C-4), 61.4 (C-6'), 59.3 (T^{α}), 55.1 (OMe(PMP)), 46.9, 46.8 (C-H(PMP); C-2), 27.6 (Me(O*t*Bu)), 21.1, 20.6, 20.5, 20.4, 20.3 (Me(Ac)), 19.3 (T^{γ}).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-(2-acetamido-3-_O_-[2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl]-2-deoxy-α-D-galactopyranosyl)-_L_-threonine-_tert_-butylester (31)[304]



A solution of compound **30** (3.0 g, 2.8 mmol) in acetic acid/water (4:1, 30 mL) was stirred for 3.5 h at 50° C. Afterwards the solvent was removed in vacuo and the residue was purified by column chromatography on silica (Tol/EtOAc 1:3) to give compound **31** (2.3 g, 2.4 mmol, 87 %).

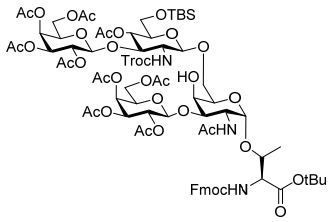
R_f = 0.12 (Tol/EtOAc 1:3). *ESI-MS*(*pos)*: m/z: 930.87 ($[M+H]^+$, calc.: 931.37).

^{1}H-NMR (600 MHz, DMSO), *δ (ppm)*: 7.93 – 7.86 (m, 2H, H-4-; H-5-Fmoc), 7.77 – 7.72 (m, 2H, H-1-; H-8-Fmoc), 7.51 – 7.45 (m, 4H, H-3-; H-6-Fmoc; N-H; N-H-Fmoc), 7.45 – 7.39 (m, 2H, H-1-; H-8-Fmoc), 5.28 (d, *J* = 3.7 Hz, 1H, H-2'), 5.03 (dd, *J* = 10.4, 3.5 Hz, 1H, H-3'), 5.00 – 4.95 (m, 1H, H-2'), 4.74 (d, *J* = 7.9 Hz, 1H, H-1'), 4.67 (t, *J* = 5.6 Hz, 1H, O-H-6), 4.59 (d, *J* = 4.1 Hz, 1H, H-1), 4.56 – 4.39 (m, 3H, CH_2-Fmoc; O-H-4), 4.32 (t, *J* = 6.8 Hz, 1H, H-9-Fmoc), 4.28 – 4.18 (m, 2H, H-2; T^{β}), 4.18 – 4.11 (m, 2H, H-5'; H-6'$_a$), 4.09 – 4.04 (m, 1H, T^{α}), 4.01 – 3.95 (m, 1H, H-6'$_b$), 3.92 – 3.88 (m, 1H, H-4), 3.65 (t, *J* = 6.2 Hz, 1H, H-5), 3.61 – 3.54 (m, 1H, H-3), 3.52 – 3.42 (m, 2H, H-6$_{a,b}$), 2.15 – 1.81 (m, 15H, 5 Ac), 1.35 (s, 9H, O*t*Bu), 1.15 (d, *J* = 6.4 Hz, 3H, T^{γ}).

^{13}C-NMR (150.9 MHz, DMSO): δ = 170.0, 169.9, 169.5, 169.1, 168.9, 168.7 (5 C=O(Ac), 1 C=O(COO*t*Bu)), 156.8 (C=O(Fmoc)), 143.7, 143.7 (C-1$_a$-; C-8$_a$-Fmoc), 140.8, 140.8 (C-4$_a$-; C-5$_a$-Fmoc), 127.7, 127.7 (C-3-; C-6-Fmoc), 127.1, 127.0 (C-2-; C-7-Fmoc), 125.3, 125.3 (C-1-; C-8-Fmoc), 120.2, 120.1 (C-4-; C-5-Fmoc), 101.3 (C-1'), 99.0 (C-1), 81.3 (C_q*t*Bu), 77.6 (C-3), 73.6 (T^{β}), 71.7 (C-5), 70.5 (C-3'), 69.9 (C-5'), 68.4 (C-2'), 67.5, 67.3 (C-4; C-4'), 65.5 (CH_2-Fmoc), 61.2 (C-6'), 60.5 (C-6), 59.4 (T^{α}), 47.0 (C-2), 46.8 (C-9-Fmoc), 27.6, 27.6, 27.6 (3 Me(O*t*Bu)), 22.9, 20.5, 20.4, 20.4, 20.3 (5 Me(Ac)), 19.1 (T^{γ}).

9.1.2.5 Synthesis of the extended core 2 type-1 amino acid

_N_-9-Fluorenylmethyloxycarbonyl-_O_-(2-acetamido-3-_O_-[2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl]-6-_O_-[3-_O_-{2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl}-4-_O_-acetyl-6-_O_-_tert_-butyldimethylsilyl-2-deoxy-2-_N_-{2,2,2-trichloroethoxycarbonyl}-β-D-glucopyranosyl]-2-deoxy-α-D-galactopyranosyl)-_L_-threonine-_tert_-butylester (32)



A slurry of T-antigen **31** (2.46 g, 2.64 mmol, 1.0 eq.), LacNAc type-1 disaccharide **24** (5.45 g, 3.70 mmol, 1.4 eq.) and molecular sieve (2.96 g, 4 Å) in dichloromethane (60 mL) was cooled on ice and stirred for 25 min. Afterwards first *N*-iodosuccinimide (713 mg, 3.17 mmol, 1.2 eq.) was added followed by dropwise addition of trifluoromethanesulfonic acid (116 µL, 1.32 mmol, 0.5 eq., dissolved in 5 mL diethyl ether). Stirring was continued for 1 h before the mixture was diluted with dichloromethane (60 mL) and filtrated over a POR3 filter. The filtrate was washed twice with 1 M sodium thiosulfate solution, twice with sat. sodium bicarbonate solution and once with brine. Finally the solution was

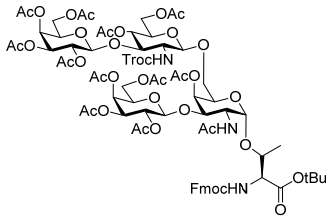
dried over sodium sulfate and the solvent was removed in vacuo. The residue was purified by column chromatography on silica (Tol/EtOAc 1:1 → 1:2) to give of compound **32** (3.58 g, 2.04 mmol, 77 %).

R_f = 0.42 (Tol/EtOAc 1:3). $[\alpha]_D^{20}$ = +21.02 (c = 0.98, $CHCl_3$). *HR-ESI-MS (pos), m/z*: 1754.5363 ($[M+H]^+$, calc.: 1754.5328), 1771.5625 ($[M+NH_3]^+$, calc.: 1771.5505).

^{1}H-NMR (600 MHz, DMSO), *δ (ppm)*: 7.92 – 7.89 (m, 2H, H-4-; H-5-Fmoc), 7.78 – 7.71 (m, 3H, H-1-; H-8-Fmoc; N-H-Troc), 7.53 – 7.46 (m, 2H, N-H-Fmoc; N-H), 7.45 – 7.38 (m, 2H, H-3-; H-6-Fmoc), 7.36 – 7.28 (m, 2H, H-2-; H-7-Fmoc), 5.27 (d, $J_{H4',H3'}$ = 3.3 Hz, 1H, H-4'), 5.22 (d, $J_{H4''',H3'''}$ = 3.7 Hz, 1H, H-4'''), 5.05 – 4.99 (m, 1H, H-3'), 4.99 – 4.94 (m, 1H, H-3'''), 4.84 (d, *J* = 12.2 Hz, 1H, CH_{2a}-Troc), 4.79 – 4.75 (m, 2H, H-2'; H-2'''), 7.92 – 7.89 (m, 4H, CH_{2b}-Troc; H-4''; H-1'; H-1'''), 4.53 (d, $J_{H1,H2}$ = 4.1 Hz, 1H, H-1), 4.51 – 4.44 (m, 2H, $CH_{2a,b}$-Troc), 4.39 (d, *JH1'',H2''* = 8.4 Hz, 1H; H-1''), 4.32 (t, $J_{H9Fmoc,CH2Fmoc}$ = 6.4 Hz, 1H, H-9-Fmoc), 4.23 – 4.17 (m, 1H, H-2), 4.17 – 4.05 (m, 5H, H-5'; H-5'''; H-6'$_a$; H-6'''$_a$; T^{β}), 4.04 – 3.93 (m, 3H, H-6'$_b$; H-6'''$_b$; T^{α}), 3.87 – 3.81 (m, 2H, H-3''; H-6$_a$), 3.80 – 3.71 (m, 2H, H-4; H-5), 3.67 (d, $J_{H6'',H5''}$ = 9.1 Hz, 1H, H-6''$_a$), 3.61 – 3.57 (m, 1H, H-6''$_b$), 3.56 – 3.49 (m, 2H, H-3; H-6$_b$), 3.48 – 3.43 (m, 1H, H-5''), 3.42 – 3.35 (m, 1H, H-2''), 2.15 – 1.79 (m, 30H, 10 Me(Ac)), 1.35 (s, 9H, *t*Bu(Thr)), 1.12 (d, $J_{T\gamma,T\beta}$ = 6.2 Hz, 3H, T^{γ}), 0.85 (s, 9H, Si*t*Bu), 0.08 – -0.04 (m, 6H, 2 SiMe).

^{13}C-NMR (150.9 MHz, DMSO): δ = 169.9, 169.9, 169.8, 169.6, 169.5, 169.4, 169.3, 169.0, 168.9, 168.9, 168.8 (10 C=O(Ac); 1 COOtBu), 156.8 (C=O(Fmoc)), 154.1 (C=O(Troc)), 143.7, 143.7 (C-1$_a$-; C-8$_a$-Fmoc), 140.8, 140.8 (C-4$_a$-; C-5$_a$-Fmoc), 127.7, 127.7 (C-3-; C-6-Fmoc), 127.0, 127.0 (C-2-; C-7-Fmoc), 125.2, 125.1 (C-1-; C-8-Fmoc), 120.2, 120.2 (C-4-; C-5-Fmoc), 101.4 (C-1'), 100.7 (C-1''), 100.0 (C-1'''), 96.0, 96.0 (C-1, CCl_3-Troc), 81.4 (C_q(O*t*Bu)), 78.0 (C-3''), 77.1 (C-3), 74.1 (T^{β}), 73.5, 73.5 (CH_2-Troc; H-5''), 70.5 (C-3'), 70.0 (C-4), 69.8, 69.6, 69.5 (C-5'; C-5'''; C-6), 68.9, 68.8, 68.8 (C-2'; C-2'''; C-4''), 68.4, 68.3 (C-3'''; C-5), 67.2 (C-4'''), 67.0 (C-4'), 65.6 (CH_2-Troc), 62.1 (C-6''), 60.9, 60.9 (C-6'; C-6'''), 59.5 (T^{α}), 56.6 (C-2''), 46.9 (C-2), 46.8 (C-9-Fmoc), 27.6 (O*t*Bu), 25.8 (Me(Si*t*Bu)), 21.1, 20.6, 20.5, 20.5, 20.5, 20.4, 20.4, 20.4, 20.3, 20.3 (10 Me(Ac)), 19.0 (T^{γ}), 18.0 (C_q(Si*t*Bu)), -5.4 (SiMe).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-(2-acetamido-4-_O_-acetyl-3-_O_-[2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl]-6-_O_-[3-_O_-{2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl}-4,6-di-_O_-acetyl-2-deoxy-2-_N_-{2,2,2-trichloroethoxycarbonyl}-β-D-glucopyranosyl]-2-deoxy-α-D-galactopyranosyl)-_L_-threonine-_tert_-butylester (33)



A solution of compound **32** (4.07 g, 2.32 mmol) in acetic acid/water (4:1, 40 mL) was heated to 50° C and stirred for 24 h under argon. Afterwards the solvent was removed by codistillation with toluene and the residue was dissolved in acetic anhydride/pyridine (1:2, 40 mL). The solution was stirred for 3 days at room temperature before removal of the solvent. The residue was purified by column chromatography on silica (Tol/EtOAc 1:1 → 0:1) to give compound **33** (2.61 g, 1.61 mmol, 69 %).

R_f = 0.70 (Tol/EtOAc 1:5). $[\alpha]_D^{20}$ = +25.15 (c = 0.99, $CHCl_3$). *HR-ESI-MS (pos), m/z*: 1015.3954 ([M-LacNTroc+Ac+H]$^+$, calc.: 1015.3918), 331.1027 (Gal oxonium]$^+$, calc.: 331.1029), 881.7175 ([M+H+K]$^+$, calc.: 881.7153), 1394.3890 ([M-Gal+H]$^+$, calc.: 1392.3754).*

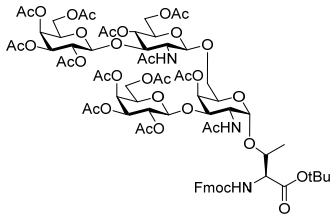
^{1}H-NMR (600 MHz, DMSO), *δ (ppm)*: 7.90 (d, $J_{H4Fmoc,H3Fmoc}$ = $J_{H5Fmoc,H6Fmoc}$ 7.4 Hz, 2H, C-4-; C-5-Fmoc), 7.79 (d, $J_{NH,H2''}$ = 9.6 Hz, 1H, N-H-Troc), 7.76 – 7.72 (m, 2H, H-1-; H-8-Fmoc), 7.56 (d, $J_{NH,H2}$ = 9.8 Hz, 1H, N-H), 7.48 – 7.39 (m, 3H, N-H-Fmoc; H-3-; H-6-Fmoc), 7.35 – 7.29 (m, 2H, H-2-; H-7-Fmoc), 5.28 – 5.20 (m, 3H, H-4; H-4'; H-4'''), 5.06 – 5.01 (m, 2H, H-3'; H-3'''), 4.85 – 4.75 (m, 3H, H-2'; H-2'''; CH_{2a}-Troc), 4.75 – 4.64 (m, 4H, H-1'; H-1'''; CH_{2b}-Troc; H-4''), 4.58 (d, $J_{H1,H2}$ = 4.1 Hz, 1H, H-1), 4.56 – 4.44 (m, 2H, CH_2-Fmoc), 4.40 (d, $J_{H1'',H2''}$ = 8.5 Hz, 1H, H-1''), 4.31 (t, $J_{H9Fmoc,CH2Fmoc}$ = 6.6 Hz, 1H, H-9-Fmoc), 4.18 – 4.03 (m, 8H, H-2; H-6'$_a$; H-6''$_a$; H-6'''$_a$; H-5; H-5'; H-5'''; T$^\beta$), 4.00 – 3.93 (m, 4H, H-6'$_b$; H-6''$_b$; H-6'''$_b$; T$^\alpha$), 3.87 – 3.77 (m, 3H, H-3; H-3''; H-6$_a$), 3.73 – 3.66 (m, 1H, H-5''), 3.44 – 3.35 (m, 2H, H-2''; H-6$_b$), 2.32 – 1.75 (m, 36H, 12 Me(Ac)), 1.34 (s, 9H, *t*Bu(Thr)), 1.14 (d, $J_{T\gamma,T\beta}$ = 6.2 Hz, 3H, T$^\gamma$).

^{13}C-NMR (150.9 MHz, DMSO): δ = 170.1, 169.9, 169.9, 169.8, 169.8, 169.7, 169.5, 169.4, 169.3, 169.2, 169.1, 168.9, 168.8 (12 C=O(Ac); 1 COO*t*Bu), 156.8 (C=O(Fmoc)), 154.1 (C=O(Troc)), 143.7, 143.7 (C-1$_a$-; C-8$_a$-Fmoc), 140.8, 140.8 (C-4$_a$-; C-5$_a$-Fmoc), 127.7, 127.7 (C-3-; C-6-Fmoc), 127.1, 127.0 (C-2-; C-7-Fmoc), 125.2, 125.1 (C-1-; C-8-Fmoc), 120.2, 120.1 (C-4-; C-5-Fmoc), 100.5 (C-1''), 100.4, 99.9 (C-1'; C-1'''), 98.7 (C-1), 95.8 (CCl_3-Troc), 81.5 (C_qO*t*Bu), 77.7 (C-3''), 74.2 (T$^\beta$), 73.7 (C-3), 73.6 (CH_2-Troc), 70.6 (C-5''), 70.3, 70.2, 70.2 (C-3'; C-3'''; C-4), 69.6, 69.5, 68.9, 68.9, 68.8, 68.8, 68.4 (C-2'; C-2'''; C-4''; C-5; C-5'; C-5'''; C-6), 67.2, 66.9 (C-4'; C-4'''), 65.5 (CH_2-Fmoc), 62.0 (C-6''), 60.8, 60.6 (C-6'; C-6'''), 59.5 (T$^\alpha$),

56.6 (C-2''), 47.6 (C-2), 46.8 (C-9-Fmoc), 27.6 (O*t*Bu), 21.1, 20.5, 20.5, 20.5, 20.5, 20.5, 20.5, 20.5, 20.4, 20.4, 20.3, 20.3 (12 Me(Ac)), 19.0 (T^{γ}).

*Under ESI-MS conditions, the LacNTroc-1,6-O-GalNAc glycosidic bond was frequently found to be fragmented in glycans.

_N_-9-Fluorenylmethyloxycarbonyl-_O_-(2-acetamido-4-_O_-acetyl-3-_O_-[2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl]-6-_O_-[2-acetamido-3-_O_-{2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl}-4,6-di-_O_-acetyl-2-deoxy-β-D-glucopyranosyl]-2-deoxy-α-D-galactopyranosyl)-_L_-threonine-_tert_-butylester (34)



While stirring a solution of compound **33** (2.63 g, 1.53 mmol, 1.0 eq.) in acetic acid (26 mL), zinc dust (2.00 g, 30.50 mmol, 20.0 eq.) was added. The slurry was heated to 50° C under argon and stirred for two days. Next, the zinc was filtered off and the solvent was removed in vacuo by codistillation with toluene. The residue was taken up in acetic anhydride/pyridine (1:2, 26 mL) and stirred 24 h at room temperature. After removal of the solvent the residue was purified by column chromatography on silica (Tol/EtOAc 1:4 → EtOAc → EtOAc/MeOH 20:1) to give compound **34** (2.17 g, 1.36 mmol, 89 %).

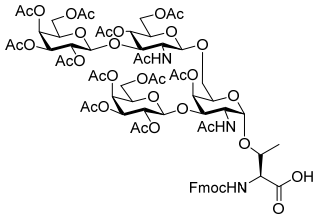
R_f = 0.58 (EtOAc/MeOH 20:1). $[\alpha]_D^{20}$ = +18.54 (c = 0.96, $CHCl_3$). *HR-ESI-MS (pos), m/z*: 795.7924 ($[M+2H]^{2+}$, calc.: 795.7920), 814.7713 $(M+H+K]^+$, calc.: 814.7700).

^{1}H-NMR (600 MHz, DMSO), *δ (ppm)*: 7.93 – 7.87 (m, 2H, H-4-; H-5-Fmoc), 7.82 – 7.77 (m, 1H, N-H''), 7.77 – 7.71 (m, 2H, H-1-; H-8-Fmoc), 7.61 (d, $J_{NH,H2}$ = 9.8 Hz, 1H, N-H), 7.45 – 7.40 (m, 3H, N-H-Fmoc; C-3-; C-6-Fmoc), 7.35 – 7.28 (m, 2H, C-2-; C-7-Fmoc), 5.26 – 5.19 (m, 3H, H-4; H-4'; H-4'''), 5.10 – 5.02 (m, 2H, H-3'; H-3'''), 4.85 – 4.68 (m, 4H, H-1'; H-1'''; H-2'; H-2'''), 4.67 – 4.61 (m, 1H, H-3''), 4.59 (d, $J_{H1,H2}$ = 4.1 Hz, 1H, H-1), 4.56 – 4.41 (m, 2H, CH_2-Fmoc), 4.40 – 4.34 (m, 1H, H-1''), 4.30 (t, $J_{H9Fmoc,CH2Fmoc}$ = 6.6 Hz, 1H, H-9-Fmoc), 4.20 – 3.87 (m, 12H, H-2; T^{α}; T^{β}; H-5; H-5'; H-5'''; H-6'$_{a,b}$; H-6''$_{a,b}$; H-6'''$_{a,b}$), 3.83 – 3.71 (m, 5H; H-2''; H-3; H-3''; H-5''; H-6$_a$), 3.35 – 3.28 (m, 1H, H-6$_b$), 2.13 – 1.79 (m, 39H, 13 Me(Ac)), 1.33 (s, 9H, O*t*Bu), 1.12 (d, $J_{T\gamma,T\beta}$ = 6.4 Hz, 3H, T^{γ}).

^{13}C-NMR (150.9 MHz, DMSO): δ = 170.3, 170.0, 169.9, 169.9, 169.8, 169.7, 169.5, 169.5, 169.4, 169.3, 169.2, 168.9, 168.9, 168.8 (13 C=O(Ac), 1 COO*t*Bu), 156.9 (C=O(Fmoc)), 143.7, 143.7 (C-1$_a$-; C-8$_a$-Fmoc), 140.8, 140.8 (C-4$_a$-; C-5$_a$-Fmoc), 127.7, 127.7 (C-3-; C-6-Fmoc), 127.1, 127.0 (C-2-; C-7-Fmoc), 125.3, 125.1 (C-1-; C-8-Fmoc), 120.2, 120.1 (C-4-; C-5-Fmoc), 101.0 (C-1''), 100.4, 100.0 (C-1'; C-1'''); 98.3 (C-1), 81.4 (C_qO*t*Bu), 77.8 (C-3''), 73.7 (C-3), 73.6 (T^{β}), 70.6, 70.5, 702., 70.1 (C-3'; C-3'''; C-4; C-5''), 69.5, 69.5, 69.1, 68.7, 68.7, 68.6, 68.4 (C-2'; C-2'''; C-4''; C-5; C-5'; C-5'''; C-6), 67.2, 66.9 (C-4'; C-5'''), 65.5

(CH_2-Fmoc), 62.0 (C-6''), 61.1, 60.7 (C-6'; C-6'''), 59.5 (T^{α}), 53.6 (C-2''), 47.7 (C-2), 46.8 (C-9-Fmoc), 27.6, (Me(O*t*Bu)), 20.8, 20.7, 20.5, 20.5, 20.5, 20.5, 20.5, 20.4, 20.4, 20.4, 20.4, 20.3, 20.3 (13 Me(Ac)), 19.2 (T^{γ}).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-(2-acetamido-4-_O_-acetyl-3-_O_-[2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl]-6-_O_-[2-acetamido-3-_O_-{2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl}-4,6-di-_O_-acetyl-2-deoxy-β-D-glucopyranosyl]-2-deoxy-α-D-galactopyranosyl)-_L_-threonine (35)



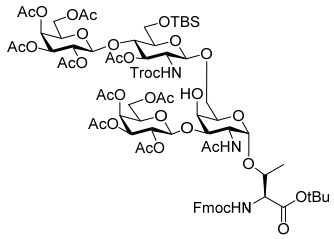
To a solution of compound **34** (2.09 g, 1.32 mmol) in dichloromethane (5 mL) anisole (2 mL) was added. After addition of trifluoroacetic acid (15 mL) the solution was stirred for 3 h$[\alpha]_D^{20}$ = +32.36 (c = 1.06, $CHCl_3$). *HR-ESI-MS (pos), m/z*: 767.7629 ($[M+2H]^{2+}$, calc.: 767.7607), 786.7374 $(M+H+K]^{+}$, calc.: 786.7387).

^{1}H-NMR (600 MHz, DMSO), *δ (ppm)*: 7.90 (d, $J_{H4Fmoc,H3Fmoc}$ = $J_{H5Fmoc,H6Fmoc}$ = 6.8 Hz, 2H, H-4-; H-5-Fmoc), 7.86 – 7.81 (m, 2H, N-H-Fmoc; N-H''), 7.73 (t, $J_{H1Fmoc,H2Fmoc}$ = $J_{H8Fmoc,H7Fmoc}$ = 7.5 Hz, 2H, H-1-; H-8-Fmoc), 7.45 – 7.39 (m, 3H, H-3-; H-6-Fmoc; N-H), 7.35 – 7.29 (m, 2H, H-2-; H-7-Fmoc), 5.26 – 5.19 (m, 3H, H-4; H-4'; H-4'''), 5.09 – 5.01 (m, 2H, H-3'; H-3'''), 4.86 – 4.67 (m, 5H, H-1; H-1'; H-1'''; H-2'; H-2'''), 4.67 – 4.60 (m, 1H, H-4''), 4.48 – 4.33 (m, 3H, CH_2-Fmoc; H-1''), 4.30 – 4.26 (m, 1H, H-9-Fmoc), 4.22 – 3.89 (m, 11H, T^{β}; H-2; H-5; H-5'; H-5'''; H-6'$_{a,b}$; H-6''$_{a,b}$; H-6'''$_{a,b}$), 3.82 (dd, $J_{H3,H2}$ = 11.1, $J_{H3,H4}$ = 3.3 Hz, 1H, H-3), 3.79 – 3.69 (m, 6H, T^{α}; H-2''; H-3; H-3''; H-5''; H-6_a), 3.39 – 3.29 (m, 1H, H-6_b), 2.11 – 1.80 (m, 39H, 13 Me(Ac)), 1.10 (d, $J_{T\gamma,T\beta}$ = 6.3 Hz, 1H, T^{γ}).

^{13}C-NMR (150.9 MHz, DMSO): δ = 170.1, 169.9, 169.9, 169.9, 169.9, 169.8, 169.6, 169.5, 169.5, 169.4, 169.3, 169.0, 169.0, 168.9 (13 C=O(Ac), 1 COOH), 143.8, 143.8 (C-1_a-; C-8_a-Fmoc), 140.8, 140.8 (C-4_a-; C-5_a-Fmoc), 127.7, 127,7 (C-3-; C-6-Fmoc), 127.1, 127.1 (C-2-; C-7-Fmoc), 125.3, 125.2 (C-1-; C-8-Fmoc), 120.2, 120.1 (C-4-; C-5-Fmoc), 101.0 (C-1''), 100.2, 100.0 (C-1'; C-1'''), 98.2 (C-1), 77.8 (C-5''), 74.1, 74.1 (T^{β}; C-3), 70.6, 70.6, 70.2 (C-3'; C-3''; C-3'''), 70.0 (C-4), 69.6, 69.6, 69.5, 68.8, 68.7, 68.6, 68.4 (C-2'; C-2'''; C-4''; C-5; C-5'; C-5'''; C-6), 67.2, 67.0 (C-4'; C-4'''), 65.5 (CH_2-Fmoc), 62.1, 62.0 (T^{α}; C-6''), 61.0, 60.7 (C-6'; C-6'''), 53.7 (C-2''), 47.9 (C-2), 46.8 (C-9-Fmoc), 25.5, 22.9, 20.7, 20.5, 20.5, 20.5, 20.5, 20.5, 20.4, 20.4, 20.4, 20.3, 20.3 (13 Me(Ac)), 18.5 (T^{γ}).

9.1.2.6 Synthesis of the extended core 2 type-2 amino acid

_N_-9-Fluorenylmethyloxycarbonyl-_O_-{2-acetamido-2-deoxy-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-6-_O_-[3-_O_-acetyl-2-deoxy-6-_O_-(_tert_-butyldimethylsilyl)-2-(2,2,2-trichloroethoxycarbonylamino)-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-α-_D_-galactopyranosyl}-_L_-threonine-_tert_-butylester (36)



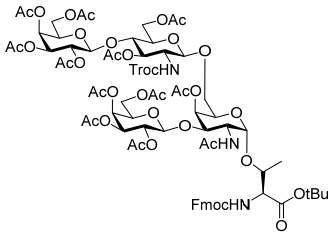
A suspension of T-antigen acceptor **31** (0.50 g, 0.54 mmol, 1.0 eq.), LacNAc type-2 donor **24** (0.67 g, 0.75 mmol, 1.4 eq.) and molecular sieve (0.59 g, 4 Å) in dry dichloromethane (12 mL) was stirred und argon at 0° C for 25 min. Afterwards *N*-iodosuccinimide (145 mg, 0.64 mmol, 1.2 eq.) and trimethylsilyl trifluoromethanesulfonate (TMSOTf, 23.5 µL, 0.27 mmol, 0.5 eq., suspended in 500 µL diethyl ether) were added to the reaction mixture and stirring was continued at 0° C for 4 h. Then additional donor **24** (96 mg, 0.10 mmol, 0.2 eq.) and *N*-iodosuccinimide (24 mg, 0.10 mmol, 0.2 eq.) were added followed by stirring at 0° C for another 6 h. The solution was diluted with dichloromethane (500 mL) and filtrated over celite. After washing twice with 0.5 M sodium thiosulfate solution, twice with sat. sodium bicarbonate solution and once with brine, the solution was dried over sodium sulfate. After removal of the solvent, the residue was purified by column chromatography on silica (Tol/EtOAc 2:3 → 1:2) to give compound **36** (1.14 g, 0.39 mmol, 72 %).

R_f = 0.4 (Tol/EtOAc 1:3). $[\alpha]_D^{20}$ = +15.3 (c = 1.00, $CHCl_3$). *HR-ESI-MS*(*pos)*: m/z: 1752.5333 ($[M+H]^+$, calc.: 1752.5363).

^{1}H-NMR (600 MHz, $CDCl_3$), *δ (ppm)*: 7.77 (d, $J_{H\text{-}4,H\text{-}5} = J_{H\text{-}5,H\text{-}6}$ = 7.6 Hz, 2H, H-4-; H-5-Fmoc), 7.61 (d, $J_{H\text{-}1,H\text{-}2} = J_{H\text{-}8,H\text{-}7}$ = 7.0 Hz, 2H, H-1-; H-8-Fmoc), 7.41 (t, $J_{H\text{-}3,H\text{-}2} = J_{H\text{-}6,H\text{-}7}$ = 7.1 Hz, 2H, H-3-; H-6-Fmoc), 7.32 (t, $J_{H\text{-}2,H\text{-}3} = J_{H\text{-}7,H\text{-}6}$ = 7.4 Hz, 2H, H-2-; H-7-Fmoc), 6.07 – 5.91 (m, 1H, N-H), 5.54 (d, $J_{NH'',H\text{-}2''}$ = 7.7 Hz, 1H, N-H''), 5.48 (d, $J_{NH\text{-}Fmoc,T\alpha}$ = 8.4 Hz, 1H, N-H-Fmoc), 5.40 – 5.31 (m, 2H, H-4'; H-4'''), 5.21 – 5.13 (m, 1H, H-2'), 5.12 – 5.02 (m, 2H, H-2'''; H-3''), 5.00 – 4.88 (m, 2H, H-3'; H-3'''), 4.82 (d, $J_{Ha\text{-}Troc,Hb\text{-}Troc}$ = 12.2 Hz, 1H, H_a-Troc), 4.76 (m, 1H, H-1), 4.66 (d, $J_{H\text{-}1''',H\text{-}2'''}$ = 7.8 Hz, 1H, H-1'''), 4.62 – 4.39 (m, 6H, H-2; Fmoc-CH_2; H_b-Troc; H-1'; H-1''), 4.29 – 4.21 (m, 1H, H-9-Fmoc), 4.20 – 4.02 (m, 6H, T^{α}; T^{β}; H-6'$_{a,b}$; H-6'''$_{a,b}$), 3.98 – 3.78 (m, 8H, H-6$_a$, H-6''$_{a,b}$; H-4; H-5'''; H-5''; H-5'; H-5), 3.78 – 3.69 (m, 1H, H-6$_b$), 3.70 – 3.58 (m, 2H, H-2''; H-3), 3.33 (d, $J_{H\text{-}4'',H\text{-}3''}$ = 8.7 Hz, 1H, H-4''), 2.18 – 1.88 (m, 30H, 10 CH_3(Ac)), 1.45 (s, 9H, O*t*Bu), 1.26 (bd, 3H, T^{γ}), 0.92 (s, 9H, Si*t*Bu), 0.10 (s, 6H, SiMe).

^{13}C-NMR (150.9 MHz, $CDCl_3$): δ = 170.69, 170.48, 170.44, 170.40, 170.35, 170.33, 170.30, 170.27, 170.03, 169.73, 169.10 (10 C=O(Ac); 1 C=O(COO*t*Bu)), 156.54 (C=O(Fmoc)), 154.40 (C=O(Troc)), 143.82, 143.82 (C-1_a-, C-8_a-Fmoc), 141.46, 141.46 (C-4_a-, C-5_a-Fmoc), 127.97, 127.97 (C-3-; C-6-Fmoc), 127.26, 127.26 (C-2-Fmoc; C-7-Fmoc), 125.06, 125.06 (C-1-; C-8-Fmoc), 120.21, 120.21 (C-4-; C-5-Fmoc), 101.76 (C-1'), 101.09 (C-1''), 100.41 (C-1; C-1'''), 95.68 (CCl_3(Troc)), 83.46 (C_q(O*t*Bu)), 77.84 (C-3), 76.10 (T^{β}) 75.60 (C-4''), 74.54 (Troc-CH_2), 74.17 (C-5'''), 72.70 (C-3''), 71.22, 71.14 (C-3', C-3'''), 70.82, 70.76, 70.63 (C-5; C-5'; C-5''), 69.76 (C-4), 69.29 (C-2'''), 68.94 (C-6), 68.83 (C-2'), 67.06, 66.99, 66.91, 66.88 (C-4'; C-4''', CH_2(Fmoc)), 61.20, 61.12, 60.98 (C-6'; C-6''; C-6'''), 59.11 (T^{α}), 56.09 (C-2''), 47.74 (C-2), 47.37 (C-9-Fmoc), 29.81 (CH2-group (HSQC)→ impurity?), 28.22 (Me(O*t*Bu)), 26.01 (Me(Si*t*Bu)), 23.46, 20.94, 20.86, 20.86, 20.80, 20.78, 20.78, 20.78, 20.69, 20.69 (10 CH_3(Ac)), 18.83 (Tγ), 18.38 (C_q(Si*t*Bu)), (Spectrum only goes to 0 ppm).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-{2-acetamido-4-_O_-acetyl-2-deoxy-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-6-_O_-[3,6-di-_O_-acetyl-2-deoxy-2-(2,2,2-trichloroethoxycarbonylamino)-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-α-_D_-galactopyranosyl}-_L_-threonine-_tert_-butylester (37)



A solution of compound **36** (1.18 g, 670 µmol) in acetic acid/water (4:1, 12 mL) was stirred overnight at 50° C. Afterwards the reaction mixture was codistilled three times with toluene before dissolving the residue in acetic anhydride/pyridine (1:2, 12 mL) and stirring at room temperature overnight. The solution was codistilled three times with toluene and the residue was purified by column chromatography on silica (Tol/EtOAc 1:4) to give compound **37** (1.09 g, 634 µmol, 95 %).

R_f = 0.60 (Tol/EtOAc 1:4). $[\alpha]_D^{20}$ = +29.5 (c = 1.00, $CHCl_3$). *HR-ESI-MS(pos): m/z:* 1722.4661($[M+H]^+$, calc.: 1722.4710).

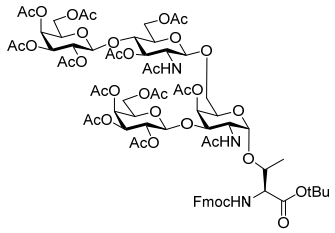
^{1}H-NMR (600 MHz, $CDCl_3$, gCOSY, TOCSY, gHSQC, gHMBC), *δ (ppm)*: 7.79(d, $J_{H\text{-}3,H\text{-}4} = J_{H\text{-}5,H\text{-}6}$ = 7.3 Hz, 2H, H-4-;H-5-Fmoc), 7.65 – 7.56 (m, 2H, H-1; H-8-Fmoc), 7.45 – 7.30 (m, 4H, H-2-; H-3-; H-6-; H-7-Fmoc), 5.93 (d, $J_{NH,H\text{-}2}$ = 9.5 Hz, 1H, N-H), 5.53 (d, $J_{NH\text{-}Fmoc,T\alpha}$ = 9.3 Hz, 1H, N-H-Fmoc), , 5.34 (m, 2H, H-4'; H-4'''), 5.29 (m, 2H, H-4, N-H''), 5.17 – 5.04 (m, 3H, H-3''; H-2'; H-2'''), 4.99 – 4.91 (m, 2H, H-3'; H-3'''), 4.86 –

4.78 (m, 2H, H_a-Troc; H-1); 4.62 – 4.46 (m, 5H, H-2; Fmoc-CH_{2a}; H_b-Troc; H-1''; H-1'/H-1'''*), 4.42 (m, 3H, H6''a; CH_{2b}-Fmoc; H-1'/H-1'''*), 4.27 (t, $J_{H\text{-}9\text{-}Fmoc,CH2\text{-}Fmoc}$ = 5.0 Hz, 1H, H-9-Fmoc), 4.23 – 4.19 (m, 1H, T^{α}), 4.19 – 4.02 (m, 7H, H-6'$_{a,b}$; H-6'''$_{a,b}$; H-6''$_b$; H-5; T^{β}), 3.90 – 3.76 (m, 4H, H-6_a; H-5'; H-5'''; H-3), 3.71 (m, 1H, H-4''), 3.65 – 3.56 (m, 2H, H-2''; H-5''), 3.48 (t, *J* = 9.6 Hz, 1H, H-6_b), 2.19 – 1.91 (m, 36H, 12 CH_3(Ac)), 1.47 (s, 9H, (O*tBu*), 1.33 – 1.24 (m, 3H, T^{γ}).

(*, signals could not be assigned to specific proton)

^{13}C-NMR (150.9 MHz, $CDCl_3$): δ = 170.59, 170.56, 170.54, 170.48, 170.37, 170.33, 170.27, 170.27, 170.22, 170.13, 170.75, 170.69, 170.24 (12 C=O(Ac); 1 C=O(COO*t*Bu)), 156.64 (C=O(Fmoc)), 154.26 (C=O(Troc)), 143.85, 143.80 (C-1_a-; C-8_a-Fmoc), 141.49, 141.46 (C-4_a-; C-5_a-Fmoc), 128.02, 128.02 (C-3-; C-6-Fmoc), 127.32, 127.26 (C-2-; C-7-Fmoc), 125.13, 125.03 (C-1-; C-8-Fmoc), 120.27, 120.25 (C-4-; C-5-Fmoc), 101.20 (C-1'/C-1'''), 101.,06 101.06 (C-1''; C-1'/C-1'''), 99.69 (C-1), 95.60 (CCl_3(Troc)), 83.40 (C_q(O*t*Bu)), 76.26, 76.26 (T^{β}; C-4''), 74.58 (CH_2-Troc), 73.38, 73.04 (C-3; C-5''), 72.19 (C-3''), 71.02, 70.87, 70.82, 70.69 (C-3'; C-3'''; C-5'; C-5'''), 69.86, 69.86 (C-4; C-5), 69.35 (C-6), 69.18 (C-2'; C-2'''), 67.03 (CH_2-Fmoc), 66.81, 66.71 (C-4'; C-4'''), 62.07 (C-6''), 60.98, 60.93 (C-6'; C-6'''), 59.20 (T^{α}), 56.13 (C-2''), 48.71 (C-2), 47.38 (C-9-Fmoc), 28.24 (3 Me(O*tBu*)), 23.52, 21.02, 20.92, 20.92, 20.88, 20.85, 20.85, 20.79, 20.79, 20.79, 20.73, 20.67 (12 CH_3(Ac)), 18.84 (T^{γ}).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-{2-acetamido-4-_O_-acetyl-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-6-_O_-[2-acetamido-3,6-di-_O_-acetyl-2-deoxy-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-α-_D_-galactopyranosyl}-_L_-threonine-_tert_-butylester (38)



Zinc dust (616 mg, 9.42 mmol, 20.0 eq.) was added to a solution of compound **38** (812 mg, 0.47 mmol 1.0 eq.) in acetic acid (8.1 mL) and the reaction was stirred at room temperature overnight. For completion of the reaction additional zinc dust (616 mg, 9.42 mmol, 20.0 eq.) was added and stirring was continued for 8 h. Then the slurry was filtrated over celite, washed with acetic acid and the filtrate was concentrated and codistilled with toluene three times. The residue was dissolved in acetic anhydride/pyridine (1:2, 8.1 mL) and stirred at room temperature overnight. After codistillation with toluene three times the residue was purified by column chromatography on silica (EtOAc/MeOH 25:1) to give compound **38** (546 mg 343 µmol, 73 %).

R_f = 0.60 (EtOAc/MeOH 20:1). $[\alpha]_D^{20}$ = +17.70 (c = 1.00, $CHCl_3$). *HR-ESI-MS(pos)*: *m/z:* 1590.5739 ($[M+H]^+$, calc.: 1590.5774).

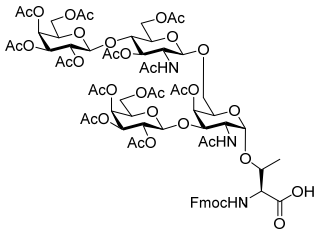
^{1}H-NMR (600 MHz, DMSO, gCOSY, TOCSY, gHSQC, gHMBC), *δ (ppm)*: 7.93 – 7.88 (m, 2H, H-4-; H-5-Fmoc), 7.87 – 7.78 (m, 1H, NH''), 7.77 – 7.71 (m, 2H, H-1-; H-8-Fmoc), 7.60 (d, $J_{NH,H-2}$ = 10.4 Hz, 1H, N-H), 7.47 – 7.38 (m, 3H, H-3-; H-6-; N-H-Fmoc), 7.35 – 7.28 (m, 2H, H-2-; H-7-Fmoc), 5.27 – 5.13 (m, 4H, H-4; H-4'; H-4'''; H-3'/H-3'''*), 5.08 – 4.99 (m, 1H, H-3'/H-3'''*), 4.93 – 4.88 (m, 1H, H-3''), 4.82 (dd, *J* = 9.2, 8.2 Hz, 2H, H-2'; H-2'''), 4.75 – 4.70 (m, 2H, H-1'; H-1'''), 4.58 (m, 1H, H-1), 4.55 – 4.43 (m, 3H, CH_2-Fmoc; H-1''), 4.33 – 4.22 (m, 2H, H-9-Fmoc; H-6''$_a$), 4.20 (t, $J_{H-5',H-6'}$ = 7.1 Hz, 1H, H-5'/H-5'''*), 4.17 – 3.86 (m, 10H, H-2; T^{α}; H-6'$_{a,b}$; H-6'''$_{a,b}$; H-6''$_b$; H-5; H-5'/H-5'''*, T^{β}), 3.80 (d, $J_{H-2,H-3}$= 10.6 Hz, 1H, H-3), 3.75 (d, $J_{H-6a,H-6b}$ = 10.6 Hz, 1H, H-6$_a$), 3.72 – 3.59 (m, 3H, H-2''; H-5''; H-4''), 3.29 (m, 1H, H-6$_b$), 2.14 – 1.68 (m, 39H, 13 Ac), 1.33 (s, 9H, O*t*Bu), 1.10 (d, $J_{T\gamma,T\beta}$ = 3.8 Hz, 3H, T^{γ}).

(*, signals could not be assigned to specific proton)

^{13}C-NMR (150.9 MHz, DMSO): δ = 170.26, 169.96, 169.92, 169.89, 169.86, 169.66, 169.54, 169.54, 169.39, 169.12, 169.12, 169.12, 168.96, 168.90 (13 C=O(Ac), 1 C=O(COO*t*Bu)), 156.87 (C=O(Fmoc)), 143.73, 143.69 (C-1$_a$-; C-8$_a$-Fmoc), 140.83, 140.80 (C-4$_a$-, C-5$_a$-Fmoc), 127.74, 127.69 (C-3-; C-6-Fmoc), 127.07, 127.03 (C-2-; C-7-Fmoc), 125.26, 125.12 (C-1-; C-8-Fmoc), 120.23, 120.17 (C-4-; C-5-Fmoc), 100.43, 100.39 (C-1''; C-1'/C-1'''), 99.92 (C-1'/C-1'''), 98.45 (C-1), 81.42 (C_q(O*t*Bu)), 76.53 (C-4''), 73.80, 73.71 (T^{β}; C-3), 73.32 (C-3''), 71.69 (C-5''), 70.23, 70.18, 70.00 (C-4; C-3'; C-3'''), 69.63, 69.50 (C-5'; C-5'''), 69.09 (C-6), 69.92 (C-2'/C-2'''), 68.57 (C-5), 68.38 (C-2'/C-2'''), 67.05, 66.93 (C-4'; C-4'''), 65.57 (Fmoc-CH_2), 62.23 (C-6''), 60.83, 60.67 (C-6'; C-6'''), 59.47 (T^{α}), 53.12 (C-2''), 47.66 (C-2), 46.79 (C-9-

Fmoc), 27.57 (3 Me(O*t*Bu)), 22.79, 22.69, 20.65, 20.57, 20.49, 20.49, 20.49, 20.40, 20.40, 20.40 , 20.34, 20.34, 20.34 (13 CH_3(Ac)), 19.00 (T^{γ}).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-{2-acetamido-4-_O_-acetyl-2-deoxy-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-6-_O_-[2-acetamido-3,6-di-_O_-acetyl-2-deoxy-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-α-_D_-galactopyranosyl}-_L_-threonine (39)



A solution of compound **38** (585 mg, 0.37 mmol, 1.0 eq.) in trifluoroacetic acid/dichloromethane (3:1, 6 mL) and anisole (585 µL, 5.38 mmol, 14.5 eq.) was stirred at room temperature for 1.5 h. After codistillation with toluene the residue was purified by column chromatography on silica (EtOAc/MeOH/AcOH/H_2O 60:3:3:2) to give of the extended core 2 type-2 glycosylated amino acid **39** (457 mg, 0.30 mmol, 80 %).

R_f = 0.40 (EtOAc/MeOH/AcOH/H_2O 30:3:3:2). $[\alpha]_D^{20}$ = +29.2 (c = 1.00, $CHCl_3$). *HR-ESI-MS(pos)*: *m/z:* 786.7338 ($[M+H+K]^{2+}$, calc.: 786.7393), 1534.5119 ($[M+H]^+$, calc.: 1534.5148), 1551.5404 ($[M+NH\text{-}4]^+$, calc.: 1551.5413).

^{1}H-NMR (600 MHz, $CDCl_3$), *δ (ppm)*: 7.90 (d, $J_{H\text{-}3,H\text{-}4} = J_{H\text{-}5,H\text{-}6}$ =7.6 Hz, 2H, H-4-; H-5-Fmoc), 7.83 (d, $J_{NH'',H\text{-}2''}$ = 9.2 Hz, 1H, NH''), 7.69 (d, $J_{H\text{-}1,H\text{-}2} = J_{H\text{-}7,H\text{-}8}$ = 7.5 Hz, 2H, H-1-; H-8-Fmoc), 7.45 – 7.30 (m, 4H, H-2-; H-3-; H-6-; H-7-Fmoc), 5.25 – 5.20 (m, 3H, H-4; H-4'; H-4'''), 5.15 (d, $J_{H\text{-}2,H\text{-}3}$ = 10.1 Hz, 1H, H-3'/H-3'''*), 5.04 (d, $J_{H\text{-}2,H\text{-}3}$ = 10.1 Hz, 1H, H-3'/H-3'''*), 4.93 – 4.87 (m, 1H, H-3''), 4.86 – 4.80 (m, 2H, H-2'; H-2'''), 4.80 – 4.69 (m, 3H, H-1; H-1'; H-1'''), 4.49 (d, $J_{H\text{-}1'',H\text{-}2''}$ = 8.3 Hz, 1H, H-1''), 4.37 – 4.22 (m, 4H, Fmoc-CH_2; H-9-Fmoc; H-6''$_a$), 4.20 (t, $J_{H\text{-}5,H\text{-}6ab}$ = 6.4 Hz, 1H, H-5'/H-5'''*), 4.18 – 4.12 (m, 1H, H-2), 4.12 – 3.89 (m, 8H, H-6'$_{a,b}$; H-6'''$_{a,b}$; H-6''$_b$; H-5; H-5'/H-5'''*; T^{β}), 3.79 (d, $J_{H\text{-}2,H\text{-}3}$ = 11.3 Hz, 1H, H-3), 3.72 – 3.63 (m, 4H, H-2''; T^{α}; H-6$_a$; H-4''), 3.63 – 3.59 (m, 1H, H-5''), 3.32 (m, 1H, H-6$_b$), 2.15 – 1.64 (m, 39H, Me(Ac)), 1.04 (d, $J_{T\beta,T\gamma}$ = 6.3 Hz, 3H).

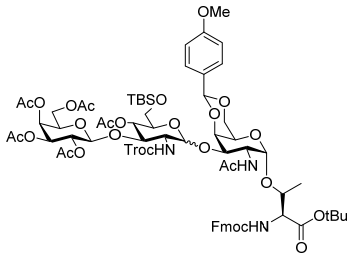
(*, signals could not be assigned to specific proton)

^{13}C-NMR (150.9 MHz, DMSO): δ = 171.72, 170.29, 170.17, 169.96, 169.93, 169.93, 169.89, 169.65, 169.54, 169.54, 169.40, 169.13, 169.05, 168.09 (13 C=O(Ac), 1 C=O(COOH)), 155.77 (C=O(Fmoc)),

143.95, 143.89 (C-1_a-; C-8_a-Fmoc), 140.75, 140.75 (C-4_a-; C-5_a-Fmoc), 127.65, 127.65 (C-3-; C-6-Fmoc), 127.12, 127.12 (C-2-; C-7-Fmoc), 125.17, 125.12 (C-1-; C-8-Fmoc), 120.17, 120.17 (C-4-; C-5-Fmoc), 100.46 (C-1''), 99.95, 99.77 (C-1'; C-1'''), 98.19 (C-1), 76.62 (C-4''), 75.77 (T^{β}), 74.97 (C-3), 73.38 (C-3''), 71.69 (C-5''), 70.32, 70.26 (C-3'; C-3'''), 69.63, 69.63 (C-5'; C-5'''), 69.41 (C-4), 68.92, 68.77, 68.68, 68.55 (C-6; C-2'; C-2'''; C-5), 67.06, 67.06 (C-4'; C-4'''), 65.44 (Fmoc-CH_2), 62.31 (C-6''), 60.83, 60.70 (C-6'; C-6'''), 58.94 (T^{α}), 53.14 (C-2''), 47.99 (C-2), 46.80 (C-9-Fmoc), 22.81, 22.65, 20.79, 20.67, 20.56, 20.51, 20.51, 20.47, 20.47, 20.42, 20.39, 20.36, 20.33 (13 Me(Ac)), 17.71 (T^{γ}).

9.1.2.7 Synthesis of the extended core 3 type-1 amino acid acceptor

_N_-9-Fluorenylmethyloxycarbonyl-_O_-[2-acetamido-2-deoxy-3-{2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-4-_O_-acetyl-6-_tert_-butyldimethyldimethylsilyl-β-D-glucopyranosyl}-4,6-_O_-_para_-methoxybenzylidenacetal-α-D-galactopyranosyl]-_L_-threonine-_tert_-butylester (40)[323]



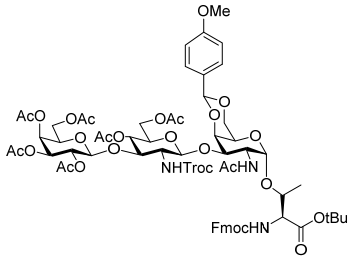
A solution of T_N-antigen acceptor **12** (1.78 g, 2.48 mmol, 1.0 eq.) and LacNAc type-1 acceptor **24** (3.00 g, 3.22 mmol, 1.3 eq.) in dry dichloromethane (50 mL) was stirred with molecular sieve (2.39 g, 4 Å) for 25 min at room temperature under argon. Afterwards the solution was cooled on ice and *N*-iodosuccinimide (724 mg, 3.22 mmol, 1.3 eq.) was added. Next, a solution of trifluoromethanesulfonic acid (65 µL, 743 µmol, 0.3 eq.) in dichloromethane (1 mL) was added dropwise. The resulting dark red solution was stirred for 1 h at 0° C before diluting the solution with dichloromethane (100 mL), filtering over celite and washing twice with 1 M sodium thiosulfate, sat. sodium bicarbonate and once with brine in that order. Finally the organic phase was dried over sodium sulfate and purified by column chromatography on silica (CH/EtOAc 2:1 → 1:1) to afford compound **42** (3.24 g, 85 %, α/β = 2:3).

R_f = 0.35 (CH/EtOAc 1:1).). *HR-ESI-MS (pos), m/z*: 790.7192 ($[M+H+K]^{2+}$, calc.: 790.7214), 730.0686 (LacNTroc oxonium – TBS + $Na]^+$, calc.: 730.0679), 331.1026 ([Gal oxonium]$^+$, calc.: 331.1029), 1542.4898 ($[M+H]^+$, calc.: 1542.4796).

^{1}H-NMR (600 MHz, DMSO), *δ (ppm)*: 7.93 – 7.86 (m, 4H, H-$4_{\alpha,\beta}$-; H-$5_{\alpha,\beta}$-Fmoc), 7.77 – 7.68 (m, 5H, H-$1_{\alpha,\beta}$-; H-$8_{\alpha,\beta}$-Fmoc; N-H_{β}-Troc), 7.57 (d, $J_{NH\alpha Fmoc,T\alpha}$ = 9.7 Hz, 1H, N-H_{α}-Fmoc), 7.51 (d, $J_{NH\beta Fmoc,T\alpha}$ = 9.6 Hz, 1H, N-H_{β}-Fmoc), 7.46 – 7.28 (m, 14H, N-H'$_{\alpha,\beta}$; H-$2_{\alpha,\beta}$ -; H-$7_{\alpha,\beta}$ -Fmoc; H-$3_{\alpha,\beta}$ -; H-$6_{\alpha,\beta}$ -Fmoc; H-$2_{\alpha,\beta}$ -; H-$6_{\alpha,\beta}$ -PMP), 7.09 (d, *J* = 9.3 Hz, 1H, N-H_{α}-Troc), 6.96 – 6.92 (m, 4H, H-$3_{\alpha,\beta}$ -; H-$5_{\alpha,\beta}$ -PMP), 5.70 (s, 1H, C-H_{α}-PMP), 5.45 (s, 1H, C-H_{β}-PMP), 5.25 (d, $J_{H4'\alpha,H3'\alpha}$ = 3.3 Hz, 1H, H-4'$_{\alpha}$), 5.23 (d, $J_{H4'\beta,H3'\beta}$ = 3.2 Hz, 1H, H-4'$_{\beta}$), 5.03 (d, $J_{H1'\alpha,H2'\alpha}$ = 3.6 Hz, 1H, H-1'$_{\alpha}$), 4.99 – 4.95 (m, 1H, H-3'$_{\beta}$), 4.91 – 4.82 (m, 3H, H-4'$_{\alpha}$, H-3''$_{\beta}$, CH_{2a}-Troc), 4.82 – 4.71 (m, 7H, H-2''$_{\alpha,\beta}$; CH_{2a}-Troc, CH_{2b}-Troc, H-$1_{\alpha,\beta}$; H-1'$_{\alpha}$), 4.64 – 4.55 (m, 4H, H-4'$_{\beta}$; CH_{2b}-Troc, H-1''$_{\beta}$; H-1'$_{\beta}$), 4.52 – 4.37 (m, 5H, $CH_{2\alpha,\beta}$-Fmoc; H-4_{α}), 4.34 – 4.22 (m, 6H, H-$9_{\alpha,\beta}$-Fmoc; H-2_{β}; $T^{\beta}_{\alpha,\beta}$; H-4_{β}), 4.21 – 4.15 (m, 1H, H-2_{α}), 4.14 – 3.89 (m, 11H, $T^{\alpha}_{\alpha,\beta}$; H-6'$_{a\alpha,\beta}$; H-$6_{a,b\alpha,\beta}$; H-3'$_{\beta}$; H-5''$_{\alpha,\beta}$), 3.84 – 3.73 (m, 9H, $OMe_{\alpha,\beta}$; H-$3_{\alpha,\beta}$; H-3'$_{\alpha}$), 3.70 – 3.53 (m, 6H, H-2'$_{\alpha}$; H-6'$b_{\alpha,\beta}$; H-$5_{\alpha,\beta}$; H-5'$_{\alpha}$), 3.52 – 3.44 (m, 1H, H-5'$_{\beta}$), 3.20 – 3.10 (m, 1H, H-2'$_{\beta}$), 2.54 – 1.79 (m, 36H, $Me(Ac)_{\alpha,\beta}$), 1.36 (s, 18H, O*t*$Bu_{\alpha,\beta}$), 1.19 – 1.12 (m, 6H, $T^{\gamma}_{\alpha,\beta}$), 0.91 – 0.77 (m, 18H, Si*t*$Bu_{\alpha,\beta}$), 0.09 – -0.07 (m, 12H, $SiMe_{\alpha,\beta}$).

^{13}C-NMR (150.9 MHz, DMSO): δ = 170.4, 169.9, 169.9, 169.9, 169.5, 169.4, 169.3, 169.2, 169.2, 169.0, 168.5 (C=O(Ac), C=O(COO*t*Bu)), 159.6, 159.5 (C-1-$PMP_{\alpha,\beta}$), 156.8, 1568.8 (C=O$(Fmoc)_{\alpha,\beta}$), 154.4, 153.7 (C=O$(Troc)_{\alpha,\beta}$), 143.9, 143.7, 143.7, 143.7 (C-1_a-; C-8_a-$Fmoc_{\alpha,\beta}$), 140.8, 140.8, 140.8 (C-4_a-; C-5_a-$Fmoc_{\alpha,\beta}$), 130.8, 130.7 (C-4-$PMP_{\alpha,\beta}$), 127.7, 127.7, 127.7, 127.6, 127.5, 127.1 (C-2-; C-7-$Fmoc_{\alpha,\beta}$; C-2-; C-6-$PMP_{\alpha,\beta}$), 125.2, 125.2, 125.2, 125.2, 124.9 (C-3-; C-6-$Fmoc_{\alpha,\beta}$; C-1-; C-8-$Fmoc_{\alpha,\beta}$), 120.2, 120.2, 120.1 (C-4-; C-5-$Fmoc_{\alpha,\beta}$), 113.4, 113.4 (C-3-; C-5-$PMP_{\alpha,\beta}$), 102.0 (C-1'$_{\beta}$), 100.1 (C-H-PMP_{β}), 100.0 (C-1''$_{\beta}$), 99.5 (C-H-PMP_{α}), 99.3, 99.1 (C-$1_{\alpha,\beta}$; C-1''$_{\alpha}$), 97.1 (C-1'$_{\alpha}$), 95.9, 95.8 (CCl_3-$Troc_{\alpha,\beta}$), 81.5, 81.4 (C_q(O*t*$Bu)_{\alpha,\beta}$), 76.8 (C-3'$_{\beta}$), 76.2 (C-3_{β}), 75.8 (C-4_{β}), 74.7 (C-3''$_{\beta}$), 74.1 (C-3_{α}), 73.9 (T^{β}_{β}), 73.8 (CH_2-$Troc_{\alpha,\beta}$), 73.6 (T^{β}_{α}), 73.4 (C-5'$_{\beta}$), 72.7 (C-4_{α}), 70.4 (C-3''$_{\alpha}$), 70.3 (C-3''$_{\beta}$), 70.0 (C-5'$_{\alpha}$), 69.8 (C-5''$_{\alpha}$), 69.7 (C-5''$_{\beta}$), 68.9 (C-4'$_{\alpha}$), 68.8 (C-2''$_{\alpha}$), 68.5, 68.4 (C-$6_{\alpha,\beta}$; C-2''$_{\beta}$), 67.9 (C-4'$_{\alpha}$), 67.2, 67.2 (C-4''$_{\alpha,\beta}$), 65.8, 65.6 (CH_2-$Fmoc_{\alpha,\beta}$), 62.8 (C-5_{α}), 62.7 (C-5_{β}), 62.2 (C-6'$_{\beta}$), 61.1, 61.0 (C-6''$_{\alpha,\beta}$), 60.9, (C-6'$_{\alpha}$), 59.4 ($T^{\alpha}_{\alpha,\beta}$), 57.3 (C-2'$_{\beta}$), 55.6 (C-2'$_{\alpha}$), 55.1, 55.1 ($OMe_{\alpha,\beta}$), 47.3, 47.1 (C-$2_{\alpha,\beta}$), 46.8, 46.8 (C-9-$Fmoc_{\alpha,\beta}$), 27.6, 25.8 (Me(O*t*$Bu)_{\alpha,\beta}$), 23.1, 23.0, 20.8, 20.6, 20.6, 20.4, 20.4, 20.3, 20.2(Me(Ac)), 19.2, 19.2 ($T^{\gamma}_{\alpha,\beta}$), 17.9, 17.9 ($C_q$(Si*t*$Bu)_{\alpha,\beta}$). -5.3, -5.4, -5.6, -5.7 ($SiMe_{\alpha,\beta}$).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-[2-acetamido-2-deoxy-3-{2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-4,6-di-_O_-acetyl-β-D-glucopyranosyl}-4,6-_O_-_para_-methoxybenzylidenacetal-α-D-galactopyranosyl]-_L_-threonine-_tert_-butylester (41)



A solution of tetrabutylammonium fluoride (6.81 g 21.6 mmol, 10.0 eq.) and acetic acid (3.7 mL, 63.0 mmol, 30.0 eq.) in tetrahydrofuran (33 mL) was added to a solution of compound **40** (3.24 g, 2.1 mmol, 1.0 eq., α/β = 2:3) in tetrahydrofuran (33 mL) at 0° C. Stirring was continued overnight, while temperature was allowed to raise to room temperature. Afterwards the solution was diluted with ethyl acetate (200 mL) and washed twice with sat. sodium bicarbonate and once with brine. The organic phase was dried over sodium sulfate and concentrated in vacuo. The residue was taken up in acetic anhydride/pyridine (1:2, 60 mL) and stirred overnight at room temperature. After full conversion, the solvent was removed by codistillation with toluene. The residue was purified by column chromatography on silica (Tol/EA 1:1 → 1:2) to give the α-anomer (1.13 g, 766 µmol, 36 %) and the β-anomer **41** (1.28 g, 869 µmol, 41 %) as colorless solids.

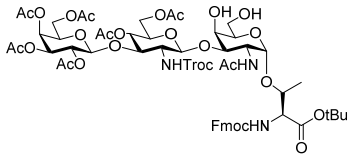
R_f(α) = 0.54, R_f(β) = 0.71 (Tol/EA 1:3).). $[\alpha]_D^{20}$ = +46.25 (c = 11.2, DMF). *HR-ESI-MS (pos), m/z*: 1468.4074 ($[M+H]^+$, calc. 1468.4067), 719.3173 ($[GalNAcThr + H]^+$, calc.: 719.3174).

^{1}H-NMR (600 MHz, DMSO), *δ (ppm)*: 7.95 – 7.86 (m, 2H, H-4-; H-5-Fmoc), 7.76 – 7.71 (m, 2H, H-1-; H-8-Fmoc), 7.69 (d, $J_{NHTroc,H2'}$ = 9.0 Hz, 1H, N-H-Troc), 7.48 (d, $J_{NHFmoc,T\alpha}$ = 9.5 Hz, 1H, N-H-Fmoc), 7.45 – 7.39 (m, 2H, H-3-; H-6-Fmoc), 7.37 – 7.27 (m, 5H, N-H; H-2-; H-7-Fmoc, H-2-; H-6-PMP), 6.94 (d, *J* = 8.7 Hz, 2H, H-3-; H-5-PMP), 5.46 (s, 1H, CH-PMP), 5.23 (d, $J_{H4'',H3''}$ = 3.2 Hz, 1H, H-4''), 4.89 (dd, $J_{H3'',H2''}$ = 10.4, $J_{H3'',H4''}$ = 3.7 Hz, 1H, H-3''), 4.83 – 4.74 (m, 2H, H-2''; CH_{2a}-Troc), 4.73 (d, $J_{H1,H2}$ = 3.8 Hz, 1H, H-1), 4.70 – 4.64 (m, 1H, H-4'), 4.62 – 4.55 (m, 2H, H-1'; H-1''; CH_{2b}-Troc), 4.49 (d, $J_{CH2Fmoc,H9Fmoc}$ = 6.6 Hz, 2H, CH_2-Fmoc), 4.34 – 4.27 (m, 2H, H-9-Fmoc; H-5), 4.26 – 4.21 (m, 1H, T^{β}), 4.20 – 3.94 (m, 10H, H-2; H-3; H-3'; H-$6_{a,b}$; H-$6'_{a,b}$; H-$6''_{a,b}$; T^{α}), 3.80 – 3.72 (m, 4H, OMe; H-5''), 3.71 – 3.62 (m, 2H, H-4; H-5'), 3.28 – 3.19 (m, 1H, H-2'), 2.13 – 1.83 (m, 21H, 7 Me(Ac)), 1.36 (s, 9H, O*t*Bu), 1.14 (d, $J_{T\gamma,T\beta}$ = 6.4 Hz, 3H, T^{γ}).

^{13}C-NMR (150.9 MHz, DMSO), *δ (ppm)*: 170.1, 169.9, 169.9, 169.9, 169.3, 169.2, 169.2, 169.0 (8 C=O(Ac); 1 COO*t*Bu), 159.5 (C-1-PMP), 156.7 (C=O(Fmoc)), 153.7 (C=O(Troc)), 143.7, 143.7 (C-1_a-; C-8_a-Fmoc), 140.8, 140.8 (C-4_a-; C-5_a-Fmoc), 130.8 (C-4-PMP), 127.5, 127.5 (C-3-; C-6-Fmoc), 127.3, 127.3 (C-2-; C-7-Fmoc), 127.0, 127.0 (C-2-; C-6-PMP), 125.2, 125.1 (C-1-; C-8-Fmoc), 120.2, 120.2 (C-4-; C-5-Fmoc), 101.6 (C-1'), 99.9 (C-1''), 99.8 (C-H-PMP), 99.1 (C-1), 95.7 (CCl_3-Troc), 81.4 (C_q(O*t*Bu)), 76.7 (C-3'), 75.3 (C-5), 74.3 (C-5''), 73.8 (CH_2-Troc), 73.6 (T^{β}), 70.5 (C-5'), 70.2 (C-3''), 69.7 (C-3), 68.7 (C-4'),

68.5 (C-6), 68.4 (C-2''), 67.2 (C-4''), 65.6 (CH_2-Fmoc), 62.8 (C-4), 61.8, 61.0 (C-6'; C-6''), 59.4 (T^{α}), 57.2 (C-2'), 51.9 (OMe), 47.0 (C-2), 46.8 (H-9-Fmoc), 27.6 (Me(O*t*Bu)), 22.9, 20.7, 20.6, 20.5, 20.5, 20.4, 20.4, 20.2 (8 Me(Ac)), 19.2 (T^{γ}).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-[2-acetamido-2-deoxy-3-{2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-4,6-di-_O_-acetyl-β-D-glucopyranosyl}-α-D-galactopyranosyl]-_L_-threonine-_tert_-butylester (42)[323]



A solution of compound **41** (1.28 g, 869 µmol) in acetic acid (12.8 mL) was stirred overnight at 50° C. Afterwards the solvent was removed by codistillation and the residue was purified by column chromatography on silica (CH/EA 1:5 → 0:1) to give the core 3 type-1 glycosylated amino acid **42** as a colorless solid (1.02 g, 752 µmol, 87 %).

R_f = 0.36 (EA). $[\alpha]_D^{20}$ = +32.38 (c = 7.0, DMF). *HR-ESI-MS (pos), m/z*: calcd for $[M+H]^+$: 1352.3579, found: 1352.3613.

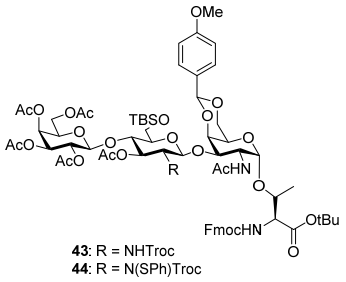
^{1}H-NMR (600 MHz, DMSO), *δ (ppm)*: 7.93 – 7.87 (m, 2H, H-4-; H-5-Fmoc), 7.76 – 7.72 (m, 2H, H-1-; H-8-Fmoc), 7.65 (d, $J_{NHTroc,H2'}$ = 9.3 Hz, 1H, N-H-Troc), 7.51 (d, $J_{NHFmoc,T\alpha}$ = 9.7 Hz, 1H, N-H-Fmoc), 7.45 – 7.39 (m, 2H, H-3-; H-6-Fmoc), 7.35 – 7.28 (m, 2H, H-2-; H-7-Fmoc), 5.24 (d, $J_{H4'',H3''}$ = 3.8 Hz, 1H, H-4''), 4.91 (dd, $J_{H3'',H2''}$ = 10.4 Hz, $J_{H3'',H4''}$ = 3.7 Hz, 1H, H-3''), 4.86 (d, $J_{a,b}$ = 12.4 Hz, 1H, CH_{2a}-Troc), 4.77 (dd, $J_{H2'',H3''}$ = 10.3, $J_{H2'',H1''}$ = 8.0 Hz, 1H, H-2''), 4.69 – 4.47 (m, 6H, H-1; H-1'; H-1''; H-4'; CH_{2b}-Troc, O-H-6), 4.44 (dd, $J_{a,b}$ = 10.9 Hz, $J_{CH2Fmoc,H9Fmoc}$ = 6.9 Hz, 1H, CH_2-Fmoc), 4.40 (d, $J_{OH4,H4}$ = 5.2 Hz, 1H, O-H-4) 4.31 (t, $J_{H9Fmoc,CH2Fmoc}$ = 6.8 Hz, 1H- H-9-Fmoc), 4.23 – 3.93 (m, 9H, H-2; H-3'; H-5; H-6a,b; H-6'a; H-6''a; T^{α}; T^{β}), 3.88 – 3.84 (m, 1H, H-4), 3.66 – 3.60 (m, 2H, H-5'; H-5''), 3.53 (dd, $J_{H3,H2}$ = 11.0 Hz, $J_{H3,H4}$ = 2.9 Hz, 1H, H-3), 3.50 – 3.43 (m, 2H, H-6'$_b$; H-6''$_b$), 3.40 – 3.35 (m, 1H, H-2'), 2.13 – 1.83 (m, 21H, 7 Me(Ac)), 1.35 (s, 9H, O*t*Bu), 1.19 – 1.13 (m, 3H, T^{γ}).

^{13}C-NMR (150.9 MHz, DMSO), *δ (ppm)*: 170.1, 169.9, 169.9, 169.4, 169.3, 169.3, 169.1, 168.9 (7 C=O(Ac), 1 COO*t*Bu), 156.8 (C=O(Fmoc)), 153.7 (C=O(Troc)), 143.7, 143.7 (C-1$_a$-; C-8$_a$-Fmoc), 140.8, 140.8 (C-4$_a$-; C-5$_a$-Fmoc), 127.7, 127.7 (C-3-; C-6-Fmoc), 127.1, 127.0 (C-2-; C-7-Fmoc), 125.3, 125.2 (C-1-; C-8-Fmoc), 120.2, 120.2 (C-4-; C-5-Fmoc), 100.0 (C-1'), 99.0 (C-1''), 95.8 (C-1), 81.3 (C_q(O*t*Bu)), 76.9, 76.9 (C-3; C-3'), 73.8 (CH_2-Troc), 73.6 (T^{β}), 71.5, 70.3, 70.2, 69.7 (C-3''; C-5; C-5'; C-5''), 68.9 (C-4'), 68.4

(C-2''), 67.8 (C-4), 67.2 (C-4''), 65.6 (CH_2-Fmoc), 62.1 (C-6), 61.1, 60.7 (C-6'; C-6''), 59.8 (T^{α}), 57.2 (C-2'), 47.1 (C-2), 46.8 (C-9-Fmoc), 27.6 (Me(O*t*Bu)), 22.9, 20.8, 20.5, 20.5, 20.4, 20.4, 20.3 (7 Me(Ac)), 19.0 (T^{γ}).

9.1.2.8 Synthesis of the extended core 3 type-2 amino acid acceptor

_N_-9-Fluorenylmethyloxycarbonyl-_O_-[2-acetamido-2-deoxy-3-{2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-3-_O_-acetyl-6-_tert_-butyldimethyldimethylsilyl-β-D-glucopyranosyl}-4,6-_O_-_para_-methoxybenzylidenacetal-α-D-galactopyranosyl]-_L_-threonine-_tert_-butylester (43, 44)



A slurry of T_N-antigen acceptor **12** (4.15 g, 5.77 mmol, 1.0 eq.), LacNAc type-2 donor **29** (7.54 g, 8.08 mmol, 1.4 eq.) and molecular sieve (5.85 g, 4 Å) in dichloromethane (120 mL) was stirred for 25 min on ice. Afterwards *N*-iodosuccinimide (3.90 g, 17.32 mmol, 3.0 eq.) was added followed by dropwise addition of trifluoromethanesulfonic acid (150 µL, 1.73 mmol, 0.3 eq.) in diethyl ether (8 mL). The mixture was stirred for 3 h, then an additional portion of *N*-iodosuccinimide (260 mg, 1.15 mmol, 0.2 eq.) was added and stirring was continued for an additional hour. Finally the system was diluted with dichloromethane (120 mL) and filtrated over celite. The filtrate was washed twice with 1 M sodium thiosulfate solution, twice with sat. sodium bicarbonate solution and once with brine before drying over sodium sulfate. After removal of the solvent in vacuo, the residue was purified by column chromatography on silica (Tol/EtOAc 2:1 → 1:1) to give the core 3 type-2 glycosylated amino acid **43** (2.69 g, 1.74 mmol, 30 %) and the corresponding phenyl sulfenyl adduct **44** (4.18 g, 2.53 mmol, 44 %).

R = NHTroc (**43**):

R_f(**43**) = 0.59, (Tol/EtOAc 1:1). $[\alpha]_D^{20}$ = +45.78**,** c = 1.05, $CHCl_3$). *HR-ESI-MS (pos), m/z*: 1540.4841 ($[M+H]^+$, calc. 1540.4831), 1557.5123 ($[M+NH_4]^+$, calc.: 1557.5097).

^{1}H-NMR (500 MHz, DMSO), *δ (ppm)*: 7.90 (d, $J_{H\text{-}3,H\text{-}4} = J_{H\text{-}5,H\text{-}6}$ = 6.9 Hz, 2H, H-4-; H-5-Fmoc), 7.76 – 7.69 (m, 2H, H-1-; H-8-Fmoc), 7.62 (d, $J_{H\text{-}2,NH}$ = 8.7 Hz,, 1H, N-H'), 7.45 – 7.39 (m, 2H, H-3-; H-6-Fmoc), 7.32

(m, 5H, H-2-; H-7-Fmoc; N-H-Fmoc; H-2-; H-6-PMP), 7.28 – 7.21 (m, 1H, N-H), 6.93 (d, $J_{H\text{-}2,H\text{-}3} = J_{H\text{-}6,H\text{-}7}$ = 8.7 Hz, 2H, H-3-; H-5-PMP), 5.42 (s, 1H, CH-PMP), 5.23 (d, $J_{H\text{-}3,H\text{-}4}$ = 3.4 Hz, 1H, H-4''), 5.17 (dd, $J_{H\text{-}2,H\text{-}3}$ = 10.3, $J_{H\text{-}3,H\text{-}4}$ = 3.5 Hz, 1H, H-3''), 5.05 (m, 1H, H-3'), 4.93 (d, $J_{Ha,Hb}$ = 12.4 Hz, 1H, Troc-H_a), 4.85 (dd, $J_{H\text{-}2,H\text{-}3}$ = 10.0, $J_{H\text{-}1,H\text{-}2}$ = 8.1 Hz, 1H, H-2''), 4.80 – 4.74 (m, 2H, H-1; H-1''), 4.71 (d, $J_{H\text{-}1,H\text{-}2}$ = 8.7 Hz, 1H, H-1'), 4.54 – 4.37 (m, 4H, H-6'$_a$; Fmoc-CH_2; Troc-H_b), 4.35 – 4.31 (m, 1H, H-4), 4.29 (t, $J_{H\text{-}9,CH2}$ = 6.1 Hz, 1H, H-9-Fmoc), 4.23 – 4.15 (m, 3H, H-2; H-5''; T^{β}), 4.12 – 4.08 (m, 1H, H-6'$_b$), 4.08 – 4.02 (m, 2H, T^{α}; H-6_a), 4.01 – 3.93 (m, 3H, H-6''$_{a,b}$; H-6_b), 3.81 – 3.70 (m, 5H, H-3; H-4'; OMe), 3.69 – 3.60 (m, 2H, H-5; H-5'), 3.40 – 3.26 (m, 1H, H-2'), 2.13 – 1.75 (m, 21H, CH_3(Ac)), 1.37 (s, 9H, Me(O*t*Bu)), 1.15 (d, $J_{T\beta,T\gamma}$ = 6.5 Hz, 3H, T^{γ}), 0.90 (Me(Si*t*Bu)), 0.12 (s, 6H, SiMe).

^{13}C-NMR (150.9 MHz, DMSO), *δ (ppm)*: 170.25, 169.90, 169.88, 169.51, 169.28, 169.17, 169.17, 169.06 (7 C=O(Ac); 1 C=O(COO*t*Bu)), 159.50 (C_{ipso}-PMP), 156.64, (C=O(Fmoc)), 153.95 (C=O(Troc)), 143.73, 143.73 (C-1_a-; C-8_a-Fmoc), 140.84, 140.81 (C-4_a-; C-5_a-Fmoc), 130.85 (C_{para}-PMP), 127.69, 127.67 (C-3-; C-6-Fmoc); 127.51, 127.49, 127.06, 127.06 (C-2-; C-7-Fmoc; C-2-; C-6-PMP), 125.11, 125.07 (C-1-; C-8-Fmoc), 120.23, 120.19 (C-4-; C-5-Fmoc), 113.36, 113.36 (C-3-; C-5-PMP), 101.26 (C-1'), 100.04, 99.88 (CH-PMP; C-1''), 98.98 (C-1), 96.06 (CCl_3(Troc)), 81.31 (C_q(OtBu)), 76.42 (C-4'), 75.48, 75.06 (C-4; C-3), 73.67 (T^{β}), 73.40 (Troc-CH_2), 73.21 (C-3'), 71.93 (C-5'); 70.26 (C-3''), 69.62 (C-5''), 69.02 (C-2''), 68.67 (C-6), 67.04 (C-4''), 65.50 (Fmoc-CH_2), 62.75 (C-5), 62.09 (C-6'), 60.77 (C-6''), 59.52 (T^{α}), 55.80 (C-2'), 55.05 (OMe), 47.05 (C-2), 46.79 (C-9-Fmoc), 27.58 (Me(O*t*Bu)), 25.8 (Me(Si*t*Bu)), 22.67, 20.80, 20.53, 20.46, 20.37, 20.36, 20.33 (Me(Ac)), 19.07 (T^{γ}), 17.9 (C_q(Si*t*Bu)).

R = N(SPh)Troc (**44**):

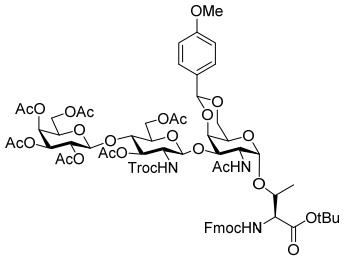
R_f(**44**) = 0.27 (Tol/EtOAc 1:1). $[\alpha]_D^{20}$ +51.73 (c = 0.52, $CHCl_3$). *HR-ESI-MS (pos), m/z:* 1648.4844 ($[M+H]^+$, calc. 1648.4865).

^{1}H-NMR (600 MHz, MeOD, T = 278 K), *δ (ppm)*: 7.85 – 7.77 (m, 4H, H-4^A-; H-5^A-Fmoc; N-H^A; N-H^B), 7.70 (d, $J_{H\text{-}1,H\text{-}2} = J_{H\text{-}7,H\text{-}8}$ = 6.8 Hz, 3H, H-1^A-; H-8^A-Fmoc, N-H^A-Fmoc), 7.63 – 7.59 (m, 2H, H-4^B-; H-5^B-Fmoc), 7.53 – 7.47 (m, 2H, H-1^B-; H-8^B-Fmoc), 7.45 – 7.37 (m, 4H, H-3^A-; H-6^A-Fmoc, H-2-; H-6-PMP), 7.37 – 7.30 (m, 5H, H-2^A-; H-7^A-Fmoc; H-2-; H-6-SPh; N-H^B-Fmoc), 7.27 – 7.18 (m, 4H, H-2^B-; H-7^B-Fmoc; H-3-; H-5-SPh), 7.14 – 7.08 (m, 1H, H-4-SPh), 6.94 – 6.86 (m, 2H, H-3-; H-5-PMP), 5.53 (s, 1H, CH-PMP), 5.44 – 5.27 (m, 3H, H-4''^{A}; H-4''^{B}; Troc-H_a^A), 5.26 – 5.20 (m, 1H, H-3''^{A}), 5.17 – 5.07 (m, 2H, H-3'B; Troc-H_b^A), 5.07 – 4.94 (m, H, Troc-H_a^B; H-3''^{A}; H-2''^{A}), 4.94 – 4.65 (m, 8H, H-2''^{B}; H-3''^{B}; H-1^A; H-1^B; H-1'^{A}; H-1'^{B}; H-1''^{A}; H-1''^{B}), 4.64 – 4.56 (m, 5H, H-2^A; H-2'^{A}; Fmoc-CH_2; Troc-H_b^B), 4.53 – 4.37 (m, 5H, H-2^B; H-2'^{B}; $T^{\beta A}$; H-4^A; H-4^B), 4.34 – 4.23 (m, 2H, H-9-Fmoc^A; H-9-Fmoc^B), 4.23 – 4.01 (m, 6H, $T^{\alpha A}$; H-6''$_{a,b}$; H-$6_{a,b}$; $T^{\beta B}$), 4.01 – 3.83 (m, 6H, H-6'$_{a,b}$; H-5''^{A}; H-5''^{B}; H-4'; H-3), 3.82 – 3.69 (m, 5H, CH_3(OMe); $T^{\alpha B}$; H-5), 3.50 –

3.43 (m, 1H, H-5‘), 2.18 – 1.87 (m, CH_3(Ac)), 1.51 – 1.38 (m, O*t*Bu), 1.36 – 1.19 (m, 6H, $T^{\gamma A}$; $T^{\gamma B}$), 1.03 – 0.87 (m, Si*t*Bu), 0.25 – 0.05 (m, 6H, 2 SiMe).

^{13}C-NMR (150.9 MHz, MeOD, T = 278 K): δ = 174.41, 172.17, 172.09, 172.06, 172.04, 172.00, 171.98, 171.95, 171.88, 171.54, 171.51, 171.19, 171.06, 170.77 (C=O(Ac), C=O(COO*t*Bu)), 161.65 (C_{ipso}-PMP), 159.52 (C=O(Fmoc)), 158.33 (C=O(Troc)), 156.77 (C=O(Troc)), 145.42, 145.31 (C-1_a-; C-8_a-Fmoc), 142.86, 142.84 (C-4_a-; C-5_a-Fmoc), 138.48 (C_{ipso}(SPh)), 131.94, 131.81 (C_{para}-PMP), 130.35, 130.35 (2 C_{arom}(SPh)), 129.39, 129.19, 129.02, 128.99, 128.97 (C-2-; C-7-PMP; C-3-; C-6-Fmoc, C_{arom}SPh), 128.37, 128.37 (C-2-; C-7-Fmoc), 126.60, 126.30 (2 C_{arom}(SPh), 126.18, 126.18 (C-1-; C-8-Fmoc), 121.22, 121.18 (C-4-; C-5-Fmoc), 114.51, 114.51 (C-3-; C-6-PMP), 104.05 (C-1’), 102.10 (CH-PMP), 101.53, 101.06, 101.00 (C-1; C-1‘‘; C_q(OTBS)), 83.69, (C_q(OtBu), 78.03, 77.74 (CH_2(Troc); C-3), 77.38, 77.30 (CH_2(Troc); C-4), 76.60, 76.33 (C-4‘; C-5‘), 75.65, 75.39, 75.00 (CH_2(Troc); T^{β}), 72.69 (C-3‘‘); 72.18, 72.03 (C-5‘‘^{A}; C-5’’^{B}), 70.75 (C-2’’), 70.38 (C-6), 68.82, 68.71 (C-4‘‘^{A}; C-4’’^{B}), 67.61 (Fmoc-CH_2), 64.96, 64.80 (C-5^A; C-5^B), 63.70 (C-2‘), 62.98, 62.65, 62.50, 62.28 (C-6‘^{A}; C-6’^{B}; C-6‘‘^{A}; C-6’’^{B}), 60.90, 60.87 ($T^{\alpha A}$; $T^{\alpha A}$), 49.13, 48.90, 48.63 (C-2^A; C-2^B; C-9-Fmoc), 28.55, 28.50 (O*t*Bu^A; O*t*Bu^B), 26.87, 26.83, 26.76, 26.76, 26.76, 26.62 (Si*t*Bu^A; Si*t*Bu^B), 21.83, 21.68, 21.60, 21.51, 21.42, 21.37, 21.26, 21.05, 21.04, 20.99, 20.95, 20.92, 20.89, 20.84, 20.68, 20.63, 20.60, 20.21, 19.62 (CH_3(Ac)), 19.36, 19.29 (C_q(Si*t*Bu)^A; C_q(Si*t*Bu)^B), 16.75 (T^{γ}).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-[2-acetamido-2-deoxy-3-{3,6-di-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl}-4,6-_O_-_para_-methoxybenzylidenacetal-α-D-galactopyranosyl]-_L_-threonine-_tert_-butylester (45)



A solution of compound **43** (3.39 g, 2.2 mmol, 1.0 eq.) in tetrahydrofuran (70.0 mL) was cooled on ice. Meanwhile, tetrabutylammonium fluoride trihydrate (6.94 g, 22.0 mmol, 10.0 eq.) and acetic acid (3.8 mL) were dissolved in tetrahydrofuran (35.0 mL) and added to the former solution. After maintaining the ice bath for 2 h, the reaction was stirred overnight, allowing the system to reach room temperature. Afterwards the solution was diluted with ethyl acetate (250 mL), washed twice with sat. bicarbonate solution and once with brine. The solvent was removed in vacuo, the residue was taken up in acetic anhydride/pyridine (1:2, 100 mL) and the resulting solution was stirred overnight at room temperature. Finally the solvent was removed in vacuo by codistillation with toluene and the residue was purified by column chromatography on silica (Tol/EtOAc 1:1 → 1:2) to give compound **45** (3.27 g, 2.2 mmol, quant).

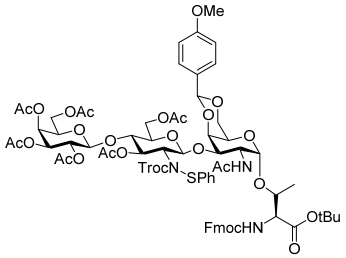
R_f = 0.44 (Tol/EtOAc 1:2). $[\alpha]_D^{20}$ = +46.82 (c = 0.22, $CHCl_3$). *HR-ESI-MS (pos), m/z*: 1490.3917 ($[M+Na]^+$, calc. 1490.3892).

^{1}H-NMR (500 MHz, DMSO), *δ (ppm)*: 7.90 (d, $J_{H\text{-}3,H\text{-}4}$ = $J_{H\text{-}5,H\text{-}6}$ = 6.9 Hz, 2H, H-4-; H-5-Fmoc), 7.76 – 7.69 (m, 2H, H-1-; H-8-Fmoc), 7.62 (d, $J_{H\text{-}2,NH}$ = 8.7 Hz,, 1H, N-H'), 7.45 – 7.39 (m, 2H, H-3-; H-6-Fmoc), 7.32 (m, 5H, H-2-; H-7-Fmoc; N-H-Fmoc; H-2-; H-6-PMP), 7.28 – 7.21 (m, 1H, N-H), 6.93 (d, $J_{H\text{-}2,H\text{-}3}$ = $J_{H\text{-}6,H\text{-}7}$ = 8.7 Hz, 2H, H-3-; H-5-PMP), 5.42 (s, 1H, CH-PMP), 5.23 (d, $J_{H\text{-}3,H\text{-}4}$ = 3.4 Hz, 1H, H-4''), 5.17 (dd, $J_{H\text{-}2,H\text{-}3}$ = 10.3, $J_{H\text{-}3,H\text{-}4}$ = 3.5 Hz, 1H, H-3''), 5.05 (m, 1H, H-3'), 4.93 (d, $J_{Ha,Hb}$ = 12.4 Hz, 1H, Troc-H_a), 4.85 (dd, $J_{H\text{-}2,H\text{-}3}$ = 10.0, $J_{H\text{-}1,H\text{-}2}$ = 8.1 Hz, 1H, H-2''), 4.80 – 4.74 (m, 2H, H-1; H-1''), 4.71 (d, $J_{H\text{-}1,H\text{-}2}$ = 8.7 Hz, 1H, H-1'), 4.54 – 4.37 (m, 4H, H-6'$_a$; Fmoc-CH_2; Troc-H_b), 4.35 – 4.31 (m, 1H, H-4), 4.29 (t, $J_{H\text{-}9,CH2}$ = 6.1 Hz, 1H, H-9-Fmoc), 4.23 – 4.15 (m, 3H, H-2; H-5''; T^{β}), 4.12 – 4.08 (m, 1H, H-6'$_b$), 4.08 – 4.02 (m, 2H, T^{α}; H-6$_a$), 4.01 – 3.93 (m, 3H, H-6''$_{a,b}$; H-6$_b$), 3.81 – 3.70 (m, 5H, H-3; H-4'; OMe), 3.69 – 3.60 (m, 2H, H-5; H-5'), 3.40 – 3.26 (m, 1H, H-2'), 2.13 – 1.75 (m, 21H, CH_3(Ac)), 1.37 (s, 9H, CH_3(O*t*Bu)), 1.15 (d, $J_{T\beta,T\gamma}$ = 6.5 Hz, 3H, T^{γ}).

^{13}C-NMR (150.9 MHz, DMSO), *δ (ppm)*: 170.25, 169.90, 169.88, 169.51, 169.28, 169.17, 169.17, 169.06 (7 C=O(Ac); 1 C=O(COO*t*Bu)), 159.50 (C_{ipso}-PMP), 156.64, (C=O(Fmoc)), 153.95 (C=O(Troc)), 143.73, 143.73 (C-1$_a$-; C-8$_a$-Fmoc), 140.84, 140.81 (C-4$_a$-; C-5$_a$-Fmoc), 130.85 (C_{para}-PMP), 127.69, 127.67 (C-3-; C-6-Fmoc); 127.51, 127.49, 127.06, 127.06 (C-2-; C-7-Fmoc; C-2-; C-6-PMP), 125.11, 125.07 (C-1-; C-8-Fmoc), 120.23, 120.19 (C-4-; C-5-Fmoc), 113.36, 113.36 (C-3-; C-5-PMP), 101.26 (C-1'), 100.04, 99.88

(CH-PMP; C-1''), 98.98 (C-1), 96.06 (CCl_3(Troc)), 81.31 (C_q(OtBu)), 76.42 (C-4'), 75.48, 75.06 (C-4; C-3), 73.67 (T^{β}), 73.40 (Troc-CH_2), 73.21 (C-3'), 71.93 (C-5'); 70.26 (C-3''), 69.62 (C-5''), 69.02 (C-2''), 68.67 (C-6), 67.04 (C-4''), 65.50 (Fmoc-CH_2), 62.75 (C-5), 62.09 (C-6'), 60.77 (C-6''), 59.52 (T^{α}), 55.80 (C-2'), 55.05 (OMe), 47.05 (C-2), 46.79 (C-9-Fmoc), 27.58 (Me(O*t*Bu)), 22.67, 20.80, 20.53, 20.46, 20.37, 20.36, 20.33 (Me(Ac)), 19.07 (T^{γ}).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-[2-acetamido-2-deoxy-3-{3,6-di-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-2-_N_-phenylsulfenyl-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl}-4,6-_O_-_para_-methoxybenzylidenacetal-α-D-galactopyranosyl]-_L_-threonine-_tert_-butylester (46)



While a solution of compound **44** (5.09 g, 3.08 mmol, 1.0 eq.) in tetrahydrofuran (100.0 mL) was stirred on ice, another solution of tetrabutylammonium fluoride trihydrate (9.73 g, 30.84 mmol, 10 eq.) and acetic acid (5.3 mL, 92.54 mmol, 30 eq.) in tetrahydrofuran (50 mL) was prepared and added dropwise to the former solution. While stirring, the ice bath was maintained for 2 h before allowing the system to reach room temperature overnight. Afterwards the solution was diluted with ethyl acetate (200 mL), washed twice with sat. sodium bicarbonate solution and once with brine. After removal of the solvent the residue was taken up in acetic anhydride/pyridine (1:2, 100 mL) and the system was stirred overnight at room temperature. Finally the solvent was removed by codistillation with toluene and the residue was purified by column chromatography on silica (Tol/EtOAc 2:1 → 1:1) to give compound **46** (3.28 g, 2.08 mmol, 67 %).

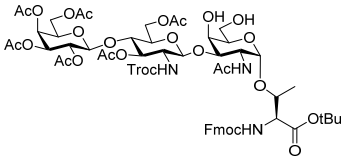
R_f = 0.38 (Tol/EtOAc 1:1). $[\alpha]_D^{20}$ = +41.40 (c = 0.88, $CHCl_3$). *HR-ESI-MS (pos), m/z*: 1576.4115 ($[M+H]^+$, calc.: 1576.4106), 1593.4397 ($[M+NH_4]^+$, calc. 1593.4371).

^{1}H-NMR (600 MHz, MeOD, T = 278 K), *δ (ppm)*: 7.86 – 7.79 (m, 3H, H-4^A-; H-5^A-Fmoc, N-H^A), 7.72 – 7.67 (m, 3H, H-1-; H-8-Fmoc, N-H^B), 7.63 – 7.59 (m, 1H, H-4^B-; H-5^B-Fmoc), 7.48 – 7.22 (m, 11H, H-2-; H-3-; H-6-; H-7-Fmoc; H-2-; H-6-PMP, H-2-; H-3-; H-4-; H-5-; H-6-SPh), 6.92 – 6.75 (m, 2H, H-3-; H-5-PMP), 5.76 – 5.50 (m, 1H, CH-PMP), 5.42 – 5.22 (m, 2H, H-4''^{A}; H-4''^{B}), 5.20 – 5.02 (m, 4H, H-3''; H-3'^{A}; H-3'^{B}; Troc-CH_{2a}), 5.02 – 4.88 (m, 2H, H-2''; H-1'), 4.84 (m, 1H, H-1), 4.80 – 4.47 (m, 8H, H-2^A; H-2'; H-

6'$_a$; Fmoc-CH_2; H-4; Troc-CH_{2b}; H-1''), 4.47 – 4.34 (m, 2H, H-2^B; T$^\beta$), 4.30 (t, $J_{H\text{-}9,CH2}$ = 6.0 Hz, 1H, H-9-Fmoc^A), 4.27 – 4.23 (m, 1H, H-9-Fmoc^B), 4.23 – 3.97 (m, 7H, T$^{\alpha A}$; H-6''$_{a,b}$; H-6'$_b$; H-6$_{a,b}$; H-5'), 3.88 – 3.46 (m, 8H, CH_3-PMP; T$^{\alpha B}$; H-5; H-5''; H-4'; H-3), 2.22 – 1.87 (m, CH_3(Ac)), 1.42 (s, 9H, Me(O*t*Bu)), 1.26 – 1.16 (m, 3H, T$^\gamma$).

^{13}C-NMR (150.9 MHz, MeOD, T = 278 K): δ = 173.41, 172.33, 172.14, 172.02, 171.50, 171.50, 171.09, 171.03 (C=O(Ac); C=O(COO*t*Bu)), 161.85, 161.69 (C_{ipso}-PMP^A; C_{ipso}-PMP^B), 159.49, 159.31 (C=O(Fmoc)^A; C=O(Fmoc)^B), 158.48 (C=O(Troc)), 145.45, 145.33 (C-1$_a$-; C-8$_a$-Fmoc), 142.90, 142.90 (C-4$_a$-; C-5$_a$-Fmoc), 132.17 (C_{para}-PMP, 130.42, 130.42, 129.10, 129.04, 129.04, 129.00, 128.38, 128.38 (C-2-; C-3-; C-6-; C-7-Fmoc; C-2-6-SPh; C-2-; C-6-PMP), 126.15, 126.15 (C-1-; C-8-Fmoc), 121.24, 121.19 (C-4-; C-5-Fmoc), 114.58, 114.58 (C-3-; C-5-PMP), 103.36 (C-1'), 101.77 (C-1''), 101.28 (C-1), 96.58 (CCl_3(Troc)), 78.53, 78.24 (C-3^A; C-3^B), 77.66, 77.38 (Troc-CH_2^A; Troc-CH_2^B), 77.15 (C-4'), 76.88 (C-4), 76.12 (T$^\beta$), 74.15 (C-5''), 72.59 (C-3''), 71.83, 71.55 (C-3'; C-5'), 70.73 (C-2''), 70.42 (C-6), 68.80 (C-4''), 67.52 (Fmoc-CH_2), 64.89 (C-5), 64.83 (C-2'), 63.62, 62.86 (C-6'; C-6''), 60.87 (T$^\alpha$), 56.12, 55.91 (CH_3-PMP^A; CH_3-PMP^B), 48.82, 48.69, 48.65 (C-2^A; C-2^B; C-9-Fmoc), 28.54 (Me(O*t*Bu)), 24.36, 24.13, 21.02, 21.93, 20.88, 20.67, 20.62 (Me(Ac)), 20.13 (T$^\gamma$).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-[2-acetamido-2-deoxy-3-{3,6-di-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl}-α-D-galactopyranosyl]-_L_-threonine-_tert_-butylester (47)



A solution of compound **45** (2.50 g, 1.7 mmol) in 80 % acetic acid (100 mL) was stirred for 2.5 h at 50° C. Afterwards the solvent was removed by codistillation with toluene and the residue was purified by column chromatography on silica (EtOAc) to give compound **47** (1.84 g, 1.36 mmol, 80 %).

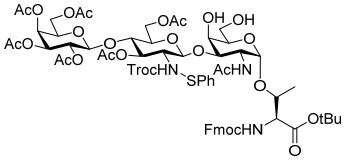
R_f = 0.26 (EtOAc/MeOH 20:1). $[\alpha]_D^{20}$ = +32.25 (c = 0.92, $CHCl_3$). *HR-ESI-MS (pos), m/z*: 694.6603 ($[M+K+H]^{2+}$, calc. 694.6646), 1350.3676 ($[M+H]^+$, calc. 1350.3654), 1367.3953 ($[M+NH_4]^+$, calc.: 1367.3919).

^{1}H-NMR (500 MHz, DMSO), *δ (ppm)*: 7.92 – 7.88 (m, 2H, H-4-; H-5-Fmoc), 7.77 – 7.70 (m, 2H, H-1-; H-8-Fmoc), 7.63 (d, $J_{NH',H2'}$ = 8.9 Hz, 1H, N-H'), 7.46 – 7.36 (m, 3H, H-3-; H-6-Fmoc; N-H-Fmoc), 7.36 – 7.28 (m, 2H, H-2-; H-7-Fmoc), 7.23 (d, $J_{NH,H2}$ = 9.4 Hz, 1H, N-H), 5.23 (d, $J_{H4'',H3''}$ = 3.5 Hz, 1H, H-4''), 5.16 (dd, $J_{H3'',H2''}$ = 10.3, $J_{H3'',H4''}$ = 3.6 Hz, 1H, H-3''), 5.06 – 4.99 (m, 1H, H-3'), 4.94 (d, $J_{CH2a,CH2b}$ = 12.4 Hz, 1H, CH_{2a}-

Troc), 4.84 (dd, $J_{H2'',H3''}$ = 10.2, $J_{H2'',H1''}$ = 8.0 Hz, 1H, H-2''), 4.77 (d, $J_{H1'',H2''}$ = 8.0 Hz, 1H, H-1''), 4.68 – 4.61 (m, 2H, H-1; H-1'), 4.50 – 4.44 (m, 3H, CH_2-Fmoc; CH_{2b}-Troc), 4.32 – 4.27 (m, 2H, H-9-Fmoc; H-6'$_a$), 4.27 – 4.12 (m, 3H, H-2; H-3; T^{β}), 4.09 (dd, *J* = 11.7, 7.5 Hz, 1H, H-6'$_b$), 4.06 – 3.98 (m, 3H, T^{α}; H-6$_a$; H-6''$_a$), 3.97 – 3.91 (m, 1H, H-5), 3.71 – 3.53 (m, 4H, H-4'; H-5'; H-5''; H-4), 3.52 – 3.40 (m, 3H, H-6$_b$; H-6''$_b$; H-2'), 2.13 – 1.77 (m, 21H, 7 Ac), 1.35 (s, 9H, O*t*Bu), 1.20 – 1.13 (m, 3H, T^{γ}).

^{13}C-NMR *(150.9 MHz, DMSO), δ (ppm):* 170.3, 169.9, 169.5, 169.3, 169.1, 169.0 (7 C=O(Ac); 1 C=O(COO*t*Bu)), 156.7 (C=O(Fmoc)), 154.0 (C=O(Troc)), 143.8, 143.7 (C-1$_a$-; C-8$_a$-Fmoc), 140.8, 140.8 (C-4$_a$- ; C-5$_a$-Fmoc), 127.7, 127.7 (C-3-; C-6-Fmoc), 127.1 (C-2-; C-7-Fmoc), 125.2, 125.1 (C-1-; C-8-Fmoc), 120.2, 120.2 (C-4-; C-5-Fmoc), 101.3 (C-1'), 100.1 (C-1''), 98.9 (C-1), 96.1 (CCl_3-Troc), 81.2 (C_q(O*t*Bu)), 78.0 (C-4), 77.0 (C-4'), 73.6 (T^{β}), 73.4 (CH_2-Troc), 73.1 (C-3'), 71.7 (C-5''), 71.5 (C-5'), 70.3 (C-3''), 69.7 (C-3), 69.0 (C-2''), 67.5 (C-5), 67.1 (C-4''), 65.5 (CH_2-Fmoc), 62.4 (C-6'), 60.8, 60.6 (C-6; C-6''), 59.5 (T^{α}), 55.8 (C-2'), 47.1 (C-2), 46.8 (C-9-Fmoc), 27.6 (Me(O*t*Bu)), 22.7, 20.5, 20.5, 20.4, 20.3 (7 Me(Ac)), 18.9 (T^{γ}).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-[2-acetamido-2-deoxy-3-{3,6-di-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-2-_N_-phenylsulfenyl-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl}-α-D-galactopyranosyl]-_L_-threonine-_tert_-butylester (48)



A solution of compound **46** (3.23 g, 2.05 mmol) in 80 % acetic acid (32 mL) was stirred for 2.5 h at 50° C. Afterwards the solvent was removed by codistillation with toluene and the residue was purified by column chromatography on silica (Tol/EtOAc 1:5 → EtOAc/MeOH 20:1) to give compound **48** (2.16 g, 1.48 mmol, 72 %).

R_f = 0.13 (Tol/EtOAc 1:3). $[\alpha]_D^{20}$ = +32.30 (c = 0.97, $CHCl_3$). *HR-ESI-MS (pos), m/z*: 1458.3694 ($[M+H]^+$, calc.: 1458.3687), 1475.3981 ($[M+NH_4]^+$, calc.: 1475.3953).

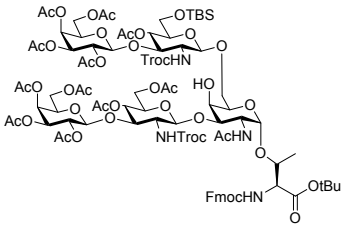
^{1}H-NMR (600 MHz, MeOD, T = 278 K), *δ (ppm)*: 7.86 – 7.75 (m, 3H, H-4^A-; H-5^A-Fmoc, N-H), 7.73 – 7.67 (m, 2H, H-1-; H-8-Fmoc), 7.63 – 7.56 (m, 2H, H-4^B-; H-5^B-Fmoc), 7.45 – 7.24 (m, 9H, H-2-; H-3-; H-6-; H-7-Fmoc; H-2-6-SPh), 5.40 – 5.25 (m, 2H, H-4''; Troc-H$_a$), 5.12 – 5.01 (m, 2H, H-3'; H-3''), 5.01 – 4.86 (m, 3H, H-2''^{A}; H-2''^{B}; H-1'), 4.77 (d, $J_{H\text{-}1,H\text{-}2}$ = 3.9 Hz, 1H, H-1^A), 4.74 (d, $J_{H\text{-}1,H\text{-}2}$ = 4.0 Hz, 1H, H-1^B), 4.69 – 4.54

(m, 7H, H-2'; H-6'$_a$; Fmoc-CH_2; Troc-H_b; H-1''^{A}; H-1''^{B}), 4.54 – 4.47 (m, 1H, H-2^A), 4.44 – 4.33 (m, 2H, H-2^B; T$^{\beta A}$), 4.33 – 4.27 (m, 1H, H-9-Fmoc^A), 4.27 – 4.13 (m, 4H, H-9-Fmoc^B; H-4^A; H-4^B; T$^{\beta B}$), 4.13 – 3.97 (m, 5H, T$^{\alpha A}$; H-6'$_b$; H-6''$_{a,b}$; H-5''), 3.87 – 3.78 (m, 2H, H-5; H-4'^{A}), 3.77 – 3.62 (m, 5H, H-6$_{a,b}$; H-5'; H-4'^{B}; H-3), 2.16 – 1.87 (m, CH_3(Ac)), 1.78 – 1.70 (m, CH_3(Ac)), 1.43 (s, 9H, O*t*Bu^A), 1.40 (s, 9H, O*t*Bu^B), 1.28 – 1.19 (m, 3H, T$^\gamma$).

^{13}C-NMR (150.9 MHz, MeOD, T = 278 K): δ = 173.56, 172.43, 172.33, 172.07, 170.01, 171.97, 171.50, 171.05, 171.01 (C=O(Ac); C=O(COO*t*Bu)), 159.50, 159.36 (C=O(Fmoc)^A; C=O(Fmoc)^B), 158.81, 158.44 (C=O(Troc)^A; C=O(Troc)^B), 145.43, 145.27 (C-1$_a$-; C-8$_a$-Fmoc), 142.85, 142.85 (C-4$_a$-; C-5$_a$-Fmoc), 138.50, 138.16 (C_{ipso}(SPh)^A; C_{ipso}(SPh)^B), 130.44, 130.44, 130.44, 128.99, 128.99, 128.99, 128.36, 126.36, 126.36 (C-2-; C-3-; C-6-; C-7-Fmoc; C-2-6-SPh), 126.16 (C-1-; C-8-Fmoc), 121.19, 121.16 (C-4-; C-5-Fmoc), 103.59 (C-1'), 102.09, 101.93 (C-1''^{A}; C-1''^{B}), 100.72, 100.60 (C-1^A; C-1^B), 96.72, 96.49 (CCl_3(Troc)^A; CCl_3(Troc)^B), 83.55 (C_q(O*t*Bu)), 80.82 (C-3), 78.63, 78.52 (C-4'^{A}; C-4'^{B}), 77.58, 77.41 (Troc-CH_2^A; Troc-CH_2^B), 75.98, 75.94 (T$^{\beta A}$; T$^{\beta B}$), 74.03 (C-5'), 72.80 (C-5), 72.56, (C-3''), 71.87, 71.76 (C-5''^{A}; C-5''^{B}), 71.35 (C-3'), 70.64 (C-2''), 70.17 (C-4), 68.58 (C-4''), 67.61, 67.41 (Fmoc-CH_2^A; Fmoc-CH_2^B), 63.50, 63.37 (C-6'^{A}; C-6'^{B}), 63.03 (C-2'), 62.87, 62.79, 62.71 (C-6^A; C-6^B; C-6''), 60.91, (T$^\alpha$), 48.86, 48.58, 48.56, 48.50 (C-2^A; C-2^B; C-9-Fmoc^A; C-9-Fmoc^B), 28.48, 28.48, 28.48 (CH_3(O*t*Bu)), 24.64, 24.35, 24.09, 21.83, 21.66, 21.04, 21.00, 20.88, 20.83, 20.78, 20.65, 20.60 (Me(Ac)), 19.99 (T$^\gamma$).

9.1.2.9 Synthesis of the extended core 4 type-1 amino acid

_N_-9-Fluorenylmethyloxycarbonyl-_O_-{2-acetamido-2-deoxy-3-_O_-[4,6-di-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-3-_O_-(2,3,4,6,-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-6-_O_-[4-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-6-_O_-_tert_-butyldimethylsilyl-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl}-β-_D_-glucopyranosyl]-α-_D_-galactopyranosyl)-_L_-threonine-_tert_-butylester (49)



A slurry of core 3 type-1 glycosyl acceptor **42** (1.14 g, 840 µmol, 1.0 eq.), LacNAc type-1 disaccharide donor **29** (0.98 g, 1.1 mmol, 1.3 eq.) and molecular sieve (1.06 g, 4 Å) in dry dichloromethane (27 mL) were stirred for 25 min at room temperature. Next, an ice bath was installed and *N*-iodosuccinimide (247 mg, 1.1 mmol, 1.3 eq.) was added. Finally, trifluoromethanesulfonic acid (22.4 µL, 250 µmol, 0.3 eq., dissolved in 500 µL diethyl ether) was added dropwise and the system was stirred for 2 h, while ice cooling was continued. Afterwards the slurry was diluted with additional dichloromethane (30 mL) and filtrated over celite. The filtrate was washed twice with 1 M sodium thiosulfate solution, twice with sat. sodium bicarbonate and once with brine before drying over sodium sulfate. The solvent was removed in vacuo and the residue was purified by column chromatography on silica (Tol/EtOAc 1:1 → 1:3) to give the core 4 type-1 glycosylated amino acid **49** (1.08 g, 495 µmol, 59 %).

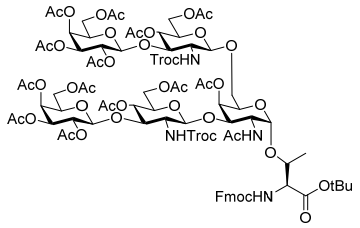
R_f = 0.48 (Tol/EtOAc 1:3). R_f = 0.36 (EA). $[\alpha]_D^{20}$ = +20.96 (c = 5.9, DMF). *HR-ESI-MS (pos), m/z*: 1352.3639 ([M-LacNAc'+H]$^+$, calc.: 1352.3619), 331.1024 ([GalNAc oxonium]$^+$, calc.: 331.1024), 730.0683 ([LacNAc' oxonium -TBS -H +Na]$^+$, calc.: 730.0679), 1394.3738 ([M-LacNAc'+Ac+H]$^+$, calc.: 1394.3724), 1107.2413 ([M+H+K]$^{2+}$, calc.: 1107.2436).

^{1}H-NMR (600 MHz, DMSO), *δ (ppm)*: 7.93 – 7.88 (m, 2H, H-4-; H-5-Fmoc), 7.78 – 7.71 (m, 3H, N-H'''; H-1-; H-8-Fmoc), 7.63 (d, $J_{NH',H2'}$ = 8.2 Hz, 1H, N-H'), 7.52 (d, $J_{NHFmoc,T\alpha}$ = 9.7 Hz, 1H, N-H-Fmoc), 7.45 – 7.39 (m, 2H, C-3-; C-6-Fmoc), 7.35 – 7.28 (m, 3H, N-H; C-2-; C-7-Fmoc), 5.24 (d, $J_{H4'',H3''}$ = 3.2 Hz, 1H, H-4''), 5.22 (d, $J_{H4'''',H3''''}$ = 3.5 Hz, 1H, H-4''''), 5.01 (dd, $J_{H3'''',H2''''}$ = 10.4, $J_{H3'''',H4''''}$3.5 Hz, 1H, H-3''''), 7.91 (dd, $J_{H3'',H2''}$ = 7.6, $J_{H3'',H4''}$ = 4.3 Hz, 1H), 4.88 – 4.81 (m, 1H, CH_{2a}-Troc), 4.79 – 4.75 (m, 2H, H-2''; H-2''''), 4.72 – 4.62 (m, 5H, C-1''; C-1''''; H-4'; H-4'''; CH_{2b}-Troc), 4.59 (d, $J_{OH4,H4}$= 5.1 Hz, 1H, O-H-4), 4.53 (d, $J_{H1,H2}$ = 4.0 Hz, 1H, H-1), 4.52 – 4.43 (m, 3H, H-1'; CH_2-Fmoc), 4.38 (d, $J_{H1''',H2'''}$ = 8.2 Hz, 1H, H-1'''), 4.31 (t, $J_{H9Fmoc,CH2Fmoc}$ = 6.6 Hz, 1H, H-9-Fmoc), 4.22 – 4.18 (m, 1H, H-6'$_a$), 4.17 – 4.04 (m, 6H, H-2; H-6''$_a$; H-

6’’’’$_a$; H-5’’; H-5’’’’; T$^\beta$), 4.04 – 3.92 (m, 5H, T$^\alpha$; H-6’’$_b$; H-6’’’’$_b$; H-6’$_b$; H-3’), 3.87 – 3.79 (m, 2H, H-3’’’; H-6$_a$), 3.79 – 3.72 (m, 2H, H-4; H-5), 3.69 – 3.64 (m, 1H, H-6’’’$_a$), 3.63 – 3.56 (m, 2H, H-6’’’$_b$; H-5’), 3.56 – 3.51 (m, 1H, H-6$_b$), 3.48 – 3.42 (m, 2H, H-3: H-5’’’), 3.42 – 3.36 (m, 2H, H-2’; H-2’’’), 2.32 – 1.84 (m, 33H, 11 Me(Ac)), 1.34 (s, 9H, O*t*Bu), 1.12 (d, $J_{T\gamma,T\beta}$ = 6.4 Hz, 3H, T$^\gamma$), 0.84 (s, 9H, Si*t*Bu), 0.02, 0.01 (s, 6H, SiMe).

^{13}C-NMR (150.9 MHz, DMSO), *δ (ppm)*: 169.9, 169.9, 169.4, 169.4, 169.3, 169.2, 169.2, 169.1, 168.9, 168.9 (C=O(Ac); C=O(COO*t*Bu)), 156.8 (C=O(Fmoc)), 154.1, 153.7 (C=O(Troc)), 143.7 (C-1$_a$-; C-8$_a$-Fmoc), 140.8, 140.8 (C-4$_a$-; C-5$_a$-Fmoc), 128.9 (C-3-; C-6-Fmoc), 128.2 (C-2-; C-7-Fmoc), 125.3 (C-1-; C-8-Fmoc), 120.2, 120.2 (C-4-; C-5-Fmoc), 102.1 (C-1’), 100.6 (C-1’’’), 100.0 (C-1’’; C-1’’’’), 99.1 (C-1), 96.0, 95.8 (CCl_3-Troc), 81.3 (C$_q$(O*t*Bu)), 78.0 (C-3’’’), 76.9 (C-3’), 76.4 (C-3), 74.0 (T$^\beta$), 73.8, 73.5, 73.5 (C-5’’’; CH_2-Troc), 70.4 (C-3’’’’; C-5’), 70.2 (C-3’’), 69.8, 69.7, 69.6 (C-5; C-5’’; C-5’’’’; C-6), 68.9, 68.8, 68.7, 68.6, 68.4 (C-4; C-4’; C-4’’’; C-2’’; C-2’’’’), 67.2, 67.2 (C-4’’; C-4’’’’), 65.6 (CH_2-Fmoc), 62.1 (C-6’’’), 61.8 (C-6’), 61.0, 60.9 (C-6’’; C-6’’’’), 59.6 (T$^\alpha$), 57.2 (C-2’), 56.6 (C-2’’’), 47.0 (C-2), 46.8 (C-9-Fmoc), 27.6 (Me(O*t*Bu)), 25.8 (Me(Si*t*Bu)), 22.9, 21.1, 21.0, 20.6, 20.5, 20.5, 20.4, 20.4, 20.3, 20.2 (Me(Ac)), 19.1 (T$^\gamma$), 18.0 (C$_q$(Si*t*Bu)), -5.4 (SiMe).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-{2-acetamido-4-_O_-acetyl-2-deoxy-3-_O_-[4,6-di-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-3-_O_-(2,3,4,6,-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-6-_O_-[4,6-di-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-α-_D_-galactopyranosyl}-_L_-threonine-_tert_-butylester (50)



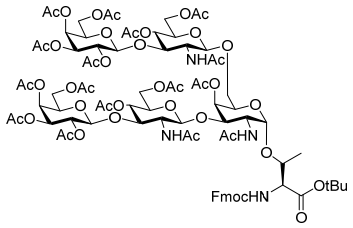
A solution of compound **49** (556 mg, 260 µmol) in 80 % acetic acid (16 mL) was stirred for 16 h at 50° C. Then, the solvent was removed by codistillation with toluene and the residue was dissolved in acetic anhydride/pyridine (1:2, 16 mL). After addition of 4-dimethylaminopryidine (10 mg), the resulting solution was stirred another 16 h at room temperature. Finally, the solvent was removed by codistillation with toluene and the residue was purified by column chromatography (Tol/EtOAc 1:1 → 0:1) to give compound **50** (449 mg, 209 µmol, 81 %).

R_f = 0.61 (Tol/EtOAc 1:3). R_f = 0.36 (EA). $[\alpha]_D^{20}$ = +219.68 (c = 3.1, DMF). *HR-ESI-MS (pos), m/z*: 685.2964 ([GalNAc-Thr+H]$^+$, calc.: 685.2967), 750.0963 ([LacNAc oxonium]$^+$, calc.: 750.0965), 1434.3865 ([M-LacNAc+Ac+H]$^+$, calc.: 1434.3859), 1392.3760 ([M-LacNAc+H]$^+$, calc.: 1392.3754).

^{1}H-NMR (600 MHz, DMSO), *δ (ppm)*: 7.93 – 7.88 (m, 2H, H-4-; H-5-Fmoc), 7.77 (d, $J_{NH''',H2'''}$ = 9.7 Hz, 1H, N-H'''), 7.75 – 7.71 (m, 2H, H-1-; H-8-Fmoc), 7.57 (d, $J_{NH',H2'}$ = 9.7 Hz, 1H, N-H'), 7.48 – 7.40 (m, 4H, N-H-Fmoc; N-H; H-3-; H-6-Fmoc), 7.36 – 7.29 (m, 2H, H-2-; H-7-Fmoc), 5.26 – 5.19 (m, 3H, H-4; H-4''; H-4''''), 5.03 (dd, $J_{H3'''',H2''''}$ = 10.4, $J_{H3'''',H4''''}$3.6 Hz, 1H, H-3''''), 4.92 (dd, $J_{H3'',H2''}$ = 10.4, $J_{H3'',H4''}$ = 3.6 Hz, 1H, H-3''), 4.88 – 4.74 (m, 3H, CH_{2a}-Troc; H-2''; H-2''''), 4.73 – 4.58 (m, 6H, H-1; H-1''; H-1''''; CH_{2b}-Troc; H-4'; H-4''''), 4.51 – 4.47 (m, 2H, CH_2-Fmoc), 4.46 (d, $J_{H1',H2'}$ = 8.4 Hz, 1H, H-1'), 4.40 (d, $J_{H1''',H2'''}$ = 8.5 Hz, 1H, H-1'''), 4.31 (t, $J_{H9Fmoc,CH2Fmoc}$ = 6.6 Hz, 1H, H-9-Fmoc), 4.17 – 4.01 (m, 9H, H-2; H-5; H-5''; H-5''''; H-6'$_a$; H-6''$_a$; H-6'''$_a$; H-6''''$_a$; T$^\beta$), 4.01 – 3.89 (m, 5H, H-6'$_b$; H-6''$_b$; H-6'''$_b$; H-6''''$_b$; T$^\alpha$), 3.87 – 3.81 (m, 1H, H-3'''), 3.78 – 3.67 (m, 3H, H-3; H-5'''; H-6$_a$), 3.58 – 3.53 (m, 1H, H-5'), 3.44 – 3.35 (m, 2H, H-2'''; H-6$_b$), 3.23 – 3.14 (m, 1H, H-2'), 2.14 – 1.80 (m, 42H, 14 Me(Ac)), 1.34 (s, 9H, O*t*Bu), 1.13 (d, $J_{T\gamma,T\beta}$ = 6.1 Hz, 3H, T$^\gamma$).

^{13}C-NMR (150.9 MHz, DMSO), *δ (ppm)*: 170.1, 169.9, 169.8, 169.6, 169.4, 169.3, 169.3, 169.3, 169.2, 169.2, 169.0, 169.0 (C=O(Ac); C=O(COO*t*Bu)), 156.8 (C=O(Fmoc)), 154.1, 153.8 (C=O(Troc)), 143.7, 1437 (C-1$_a$-; C-8$_a$-Fmoc), 140.8, 140.8 (C-4$_a$-; C-5$_a$-Fmoc), 127.7, 127.7 (C-3-; C-6-Fmoc), 127.1, 127.0 (C-2-; C-7-Fmoc), 125.2, 125.1 (C-1-; C-8-Fmoc), 120.2, 120.2 (C-4-; C-5-Fmoc), 101.3 (C-1'), 100.6 (C-1'''), 99.9, 99.9 (C-1''; C-1''''), 98.9 (C-1), 95.9, 95.8 (CCl_3-Troc), 81.4 (C_q(O*t*Bu)), 77.7 (C-3'''), 76.6 (C-3'), 74.3 (T$^\beta$), 73.8, 73.6 (CH_2-Troc), 72.9 (C-3), 70.6, 70.6 (C-4; C-5'''), 70.3, 70.2 (C-3''; C-3''''; C-5'), 69.7, 69.6 (C-5''; C-5''''), 69.1 (C-6), 68.9 (C-4'''), 68.8, 68.6, 68.5, 68.5 (C-2''; C-2''''; C-4'; C-5), 67.2, 67.1 (C-4''; C-4''''), 65.5 (CH_2-Fmoc), 62.0, 61.4, 61.0, 60.8 (C-6'; C-6''; C-6'''; C-6''''), 59.6 (T$^\alpha$), 57.3 (C-2'), 56.6 (C-2'''), 47.6 (C-2), 46.8 (C-9-Fmoc), 27.6 (Me(O*t*Bu)), 22.8, 20.7, 20.5, 20.5, 20.5, 20.4, 20.4, 20.3, 20.3 (Me(Ac)), 19.0 (T$^\gamma$).

***N*-9-Fluorenylmethyloxycarbonyl-*O*-{2-acetamido-4-*O*-acetyl-2-deoxy-3-*O*-[4,6-di-*O*-acetyl-2-deoxy-2-acetamido-3-*O*-(2,3,4,6,-tetra-*O*-acetyl-β-*D*-galactopyranosyl)-β-*D*-glucopyranosyl]-6-*O*-[4,6-di-*O*-acetyl-2-deoxy-2-acetamido-3-*O*-(2,3,4,6-tetra-*O*-acetyl-β-*D*-galactopyranosyl)-β-*D*-glucopyranosyl]-α-*D*-galactopyranosyl}-*L*-threonine-*tert*-butylester (51)**



To zinc dust (279 mg, 4.27 mmol, 20 eq.) and compound **50** (449 mg, 209 µmol, 1 eq.), acetic acid (4.5 mL) was added under argon. Then the system was heated to 50° C and stirred for 16 h. Afterwards the slurry was filtrated over celite and the filtrate was concentrated in vacuo. The residue was taken up in acetic anhydride/pyridine (1:2, 4.5 mL) and the resulting solution was stirred for 16 h at room temperature. After removal of the solvent by codistillation with toluene, the residue was purified by column chromatography on silica to give compound **51** (332 mg, 177 µmol, 85 %).

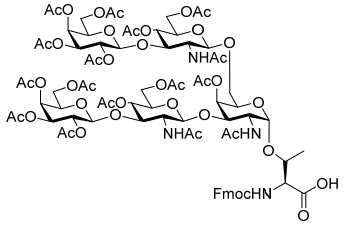
R_f = 0.26 (EtOAc/MeOH 15:1). R_f = 0.36 (EA). $[\alpha]_D^{20}$ = +22.92 (c = 3.2, DMF). *HR-ESI-MS (pos), m/z*: 958.3177 ($[M+H+K]^{2+}$, calc.: 958.3202), 939.8440 ($[M+H+Na]^{2+}$, calc.: 939.8349).

^{1}H-NMR (600 MHz, DMSO), *δ (ppm)*: 7.92 – 7.89 (m, 2H, H-4-; H-5-Fmoc), 7.78 (d, $J_{NH'''',H2''''}$ = 9.0 Hz, 1H, N-H'''), 7.74 (t, $J_{H1,8Fmoc,H2,7Fmoc}$ = 8.3 Hz, 2H, H-1-; H-8-Fmoc), 7.50 – 7.46 (m, 2H, N-H; N-H-Fmoc), 7.43 (td, *J* = 7.5, 3.2 Hz, 2H, H-2-; H-7-Fmoc), 7.35 – 7.30 (m, 2H, H-3-; H-6-Fmoc), 7.12 (d, $J_{NH',H2'}$ = 10.0 Hz, 1H, N-H'), 5.22 (d, $J_{H4'',H3''} = J_{H4'''',H3''''}$ = 3.7 Hz, 2H, H-4''; H-4''''), 5.20 (d, $J_{H4,H3}$ = 2.9 Hz, 1H, H-4), 5.10 – 5.05 (m, 2H, H-4''; H-4''''), 4.75 (dd, $J_{H2,H3}$ = 10.2, $J_{H2,H1}$ = 7.9 Hz, 2H, H-2''; H-2''''), 4.72 – 4.62 (m, 4H, H-1''; H-1''''; H-4'''; H-4'''), 4.62 – 4.59 (m, 2H; H-1; H-4'), 4.55 (d, $J_{H1',H2'}$ = 10.4 Hz, 1H, H-1'), 4.54 – 4.44 (m, 2H, $CH_{2a,b}$-Fmoc), 4.36 (d, $J_{H1''',H2'''}$ = 7.5 Hz, 1H, H-1'''), 4.30 (t, $J_{H9Fmoc,CH2Fmoc}$ = 6.6 Hz, 1H, H-9-Fmoc), 4.19 – 3.90 (m, 15H. T^{α}; T^{β}; H-3'; H-5; H-5''; H-5''''; H-6'$_{a,b}$; H-6''$_{a,b}$; H-6'''$_{a,b}$; H-6''''$_{a,b}$), 3.78 – 3.72 (m, 4H, H-2'''; H-3'''; H-5'''; H-6_a), 3.70 (dd, $J_{H3,H2}$ = 11.4, $J_{H3,H4}$3.2 Hz, 1H, H-3), 3.61 – 3.57 (m, 1H, H-5'), 3.42 – 3.28 (m, 2H, H-2'; H-6_b), 2.15 – 1.75 (m, 48H, 16 Me(Ac)), 1.34 (s, 9H, O*t*Bu), 1.11 (d, $J_{T\gamma,T\beta}$ = 6.3 Hz, 1H).

^{13}C-NMR (176.0 MHz, DMSO), *δ (ppm)*: 170.1, 170.0, 169.9, 169.9, 169.9, 169.7, 169.5, 169.4, 169.4, 169.3, 169.2, 169.1, 169.0, 168.8 (C=O(Ac), C=O(COO*t*Bu)), 156.8 (C=O(Fmoc)), 143.7, 143.7 (C-1_a-; C-8_a-Fmoc), 140.8, 140.8 (C-4_a-; C-5_a-Fmoc), 127.7, 127.7 (C-3-; C-6-Fmoc), 127.1,27.0 (C-2-; C-7-Fmoc), 125.2, 125.1 (C-1-; C-8-Fmoc), 120.2, 120.2 (C-4-; C-5-Fmoc), 101.0 (C-1'''), 100.7 (C-1'), 100.0, 99.9 (C-1''; C-1''''), 98.4 (C-1), 81.4 (C_q(O*t*Bu)), 77.9 (C-3'''), 76.9 (C-3'), 73.6 (T^{β}), 73.1 (C-3), 70.6, 70.5, 70.5,

70.3, 70.2 (C-4; C-3''; C-3''''; C-5'''; C-5'), 69.5, 69.4 (C-5''; C-5''''), 69.4 (C-6), 68.7, 68.6, 68.6 (C-5; C-4'; C-4'''; C-2''; C-2''''), 65.5 (CH_2-Fmoc), 62.0, 61.6, 61.1, 61.1 (C-6'; C-6''''), 59.5 (T^{α}), 54.9 (C-2'), 53.6 (C-2'''), 47.6 (C-2), 46.8 (C-9-Fmoc), 27.6 (Me(O*t*Bu)), 22.9, 22.8, 20.7, 20.5, 20.5, 20.5, 20.4, 20.4, 20.3 (Me(Ac)), 19.1 (T^{γ}).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-{2-acetamido-4-_O_-acetyl-2-deoxy-3-_O_-[4,6-di-_O_-acetyl-2-deoxy-2-acetamido-3-_O_-(2,3,4,6,-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-6-_O_-[4,6-di-_O_-acetyl-2-deoxy-2-acetamido-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-α-_D_-galactopyranosyl}-_L_-threonine (52)



After dissolving compound **51** (332 mg, 0.18 mmol, 1 eq.) in dichloromethane (830 µL), anisole (664 mL, 6.11 mmol, 34 eq.) was added. Next, trifluoroacetic acid (2.49 mL) was added and the resulting solution was stirred for 2.5 h at room temperature. Afterwards the solvent was removed in vacuo and the residue was purified by column chromatography on silica (EtOAc/MeOH/AcOH/H_2O) 60:3:3:2 → 30:3:3:2) to give of the core 4 type-1 glycosylated amino acid **52** (300 mg, 165 µmol, 93 %).

R_f = 0.21 (EtOAc/MeOH/AcOH/H_2O 60:3:3:2). R_f = 0.36 (EA). $[\alpha]_D^{20}$ = +46.77 (c = 3.1, DMF). *HR-ESI-MS (pos), m/z*: 930.2856 ($[M+H+K]^{2+}$, calc.: 930.2889), 911.3127 ($[M+2H]^{2+}$, calc.: 911.3110).

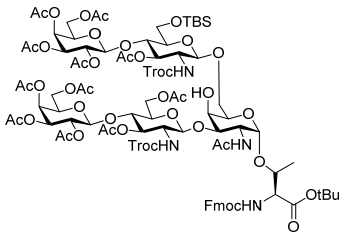
^{1}H-NMR (700 MHz, DMSO), *δ (ppm)*: 7.92 – 7.89 (m, 2H, H-4-; H-5-Fmoc), 7.82 – 7.77 (m, 2H, NH'; NH'''), 7.76 – 7.72 (m, 2H, H-1-; H-8-Fmoc), 7.45 – 7.41 (m, 3H, H-3-; H-6-Fmoc; N-H-Fmoc), 7.36 – 7.29 (m, 2H, H-2-; H-7-Fmoc), 5.24 – 5.21 (m, 2H, H-4''; H-4''''), 5.21 – 5.20 (m, 1H, H-4), 5.10 – 5.04 (m, 2H, H-3''; H-5''''), 4.75 (dd, $J_{H2,H3}$ = 10.2, $J_{H2,H1}$ = 7.9 Hz, 2H, H-2''; H-2''''), 4.71 – 4.62 (m, 5H, H-1''; H-1''''; H-1; H-4'''; H-4'), 4.62 – 4.57 (m, 1H, H-4'), 4.56 – 4.54 (m, 1H, H-1'), 4.54 – 4.50 (m, 1H, CH_{2a}-Fmoc), 4.47 – 4.41 (m, 1H, CH_{2b}-Fmoc), 4.39 – 4.34 (m, 1H, H-1'''), 4.33 – 4.28 (m, 1H, H-9-Fmoc), 4.23 – 4.19 (m, 1H, T^{β}), 4.17 – 3.94 (m, 14H, H-5''; H-5''''; H-$6_{a,b}$'; H-$6_{a,b}$''; H-$6_{a,b}$'''; H-$6_{a,b}$''''; H-5; H-2; T^{α}), 3.79 – 3.70 (m, 6H, H-2'; H-2'''; H-5'''; H-6_a; H-3; H-3'''), 3.64 – 3.58 (m, 1H, H-5'), 3.35 – 3.27 (m, 1H, H-6_b), 2.12 – 1.77 (m, 48H, 16 Me(Ac)), 1.10 (d, $J_{T\gamma,T\beta}$ = 6.3 Hz, 3H, T^{γ}).

^{13}C-NMR (176.0 MHz, DMSO), *δ (ppm)*: 170.1, 170.1, 169.9, 169.9, 169.9, 169.9, 169.9, 169.7, 169.5, 169.5, 169.5, 169.4, 169.4, 169.3, 169.2, 169.2, 168.9 (16 C=O(Ac); 1 COOH), 143.8, 143.8 (C-1_a-; C-8_a-

Fmoc), 140.8, 140.8 (C-4_a-; C-5_a-Fmoc), 127.7, 127.7 (C-3-; C-6-Fmoc), 127.1, 127.1 (C-2-; C-7-Fmoc), 125.3, 125.1 (C-1-; C-8-Fmoc), 120.2, 120.2 (C-4-; C-5-Fmoc), 100.9 (C-1'''), 100.7 (C-1'), 100.0, 99.9 (C-1''; C-1''''), 98.3 (C-1), 77.8 (C-3'''), 74.0 (T^{β}), 73.1 (C-3), 70.5, 70.5, 70.5 (C-3''; C-3''''; C-5'''), 70.4, 70.4 (C-4; C-5'), 69.5, 69.5 (C-5''; C-5''''), 69.1 (C-6), 68.7, 68.6, 68.6, 68.6, 68.6 (C-2''; C-2''''; C-4'; C-4'''; C-5), 67.2, 67.2 (C-4''; C-4''''), 65.5 (CH_2-Fmoc), 62.0, 61.7, 61.1, 61.0 (C-6'; C-6''; C-6'''; C-6''''), 58.5 (T^{α}), 53.7, 53.7 (C-2'; C-2'''), 46.8 (C-9-Fmoc), 22.9, 22.9, 20.7, 20.5, 20.5, 20.5, 20.5, 20.5, 20.5, 20.5, 20.4, 20.4, 20.4, 20.3, 20.3, 20.3 (16 CH_3(Ac)), 18.8 (T^{γ}).

9.1.2.10 Synthesis of the extended core 4 type-2 amino acid

_N_-9-Fluorenylmethyloxycarbonyl-_O_-{2-acetamido-2-deoxy-3-_O_-[3,6-di-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbony)-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-6-_O_-[3-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-6-_O_-_tert_-butyldimethylsilyl-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-α-_D_-galactopyranosyl}-_L_-threonine-_tert_-butylester (53)



A suspension of core 3 type-2 acceptor **47** (1.94 g, 1.43 mmol, 1.0 eq.), LacNAc type-2 donor **29** (1.79 g, 2.01 mmol, 1.4 eq.) and molecular sieve (1.86 g, 4 Å) in dry dichloromethane (45 mL) was stirred for 25 min at 0° C. Then *N*-Iodosuccinimide (387 mg, 1.72 mmol, 1.2 eq.) was added followed by a solution of TfOH (63.4 µL, 0.72 mmol, 0.5 eq.) in diethyl ether (1 mL). The slurry was stirred at 0° C for 5 h, diluted with dichloromethane (50 mL) and filtrated over celite. The filtrate was washed twice with sat. sodium bicarbonate, twice with 1 M sodium thiosulfate, once with Brine and dried over sodium sulfate. The solvent was removed in vacuo and the crude material was purified by silica column chromatography (Tol/EtOAc 1:2 → 1:5) to give compound **53** (2.05 g, 0.94 mmol, 66 %).

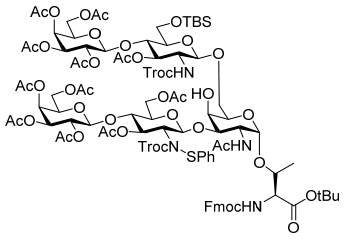
R_f = 0.60 (Tol/EtOAc 1:5). $[\alpha]_D^{20}$ = +15.96 (c = 0.80, $CHCl_3$). *HR-ESI-MS (pos), m/z*: 1106.2454 ($[M+K+H]^{2+}$, calc.: 1106.2457).

^{1}H-NMR (600 MHz, DMSO), *δ (ppm)*: 7.92 – 7.86 (m, 2H, H-4-; H-5-Fmoc), 7.78 (d, $J_{NH,H-2}$ = 9.2 Hz, 1H, NH'''), 7.75 – 7.70 (m, 2H, H-1-; H-8-Fmoc), 7.67 (d, $J_{NH,H-2}$ = 9.0 Hz, 1H, NH'), 7.45 (d, $J_{NH,T\alpha}$ = 9.6 Hz, 1H, N-H-Fmoc), 7.45 – 7.38 (m, 2H, H-3-; H-6-Fmoc), 7.32 (m, 2H, H-2-; H-7-Fmoc), 7.23 – 7.21 (m, 1H, N-

H), 5.25 (d, $J_{H\text{-}3,H\text{-}4}$ = 3.6 Hz, 1H, H-4’’’’), 5.22 (d, $J_{H\text{-}3,H\text{-}4}$ = 3.6 Hz, 1H, H-4’’), 5.16 (dd, $J_{H\text{-}2,H\text{-}3}$ = 10.3, $J_{H\text{-}3,H\text{-}4}$ = 3.5 Hz, 1H, H-3’’), 5.08 (dd, $J_{H\text{-}2,H\text{-}3}$ = 10.3, $J_{H\text{-}3,H\text{-}4}$ = 3.7 Hz, 1H, H-3’’’’), 5.02 (t, $J_{H\text{-}3,H\text{-}4}$ = 9.5 Hz, 1H, H-3’), 4.96 – 4.88 (m, 3H, H-3’’’; 2 CH_{2a}-Troc), 4.86 – 4.80 (m, 2H, H-2’’; H-2’’’’), 4.75 (d, $J_{H\text{-}1,H\text{-}2}$ = 8.0 Hz, 1H, H-1’’), 4.70 (d, $J_{H\text{-}1,H\text{-}2}$ = 8.0 Hz, 1H, H-1’’’’), 4.65 (m, 1H, H-1’), 4.59 (d, $J_{Ha,Hb}$ = 12.4 Hz, 1H, CH_{2b}-Troc), 4.56 (d, $J_{H\text{-}1,H\text{-}2}$ = 4.1 Hz, 1H, H-1), 4.52 – 4.43 (m, 4H, Fmoc-CH_2; CH_{2b}-Troc; H-1’’’), 4.40 (d, $J_{H\text{-}4,OH}$ = 4.8 Hz, 1H, O-H-4), 4.33 – 4.26 (m, 2H, H-9-Fmoc; H-6’$_a$), 4.21 (t, $J_{H\text{-}5,H\text{-}6}$ = 6.8 Hz, 1H, H-3), 4.18 – 4.12 (m, 1H, H-2), 4.12 – 3.94 (m, 8H, H-6’$_b$; H-6’’$_{a,b}$; H-6’’’’$_{a,b}$; T^{α}; H-5’’’’; T^{β}), 3.87 – 3.65 (m, 7H, H-6’’’a,b; H-4; H-4’’’; H-4’; H-5, H-6$_a$), 3.62 – 3.49 (m, 3H, H-6$_b$; H-5’’; H-5’), 3.48 – 3.42 (m, 1H, H-2’), 3.43 – 3.35 (m, 1H, H-2’’’), 3.36 – 3.29 (m, 1H, H-5’’’), 2.14 – 1.73 (m, 36H, Me(Ac)), 1.34 (s, 9H, O*t*Bu), 1.09 (d, $J_{T\beta,T\gamma}$ = 6.2 Hz, 3H, T^{γ}), 0.89 (s, 9H, Si*t*Bu), 0.08 (s, 6H, SiMe).

^{13}C-NMR (150.9 MHz, DMSO): δ = 170.00, 169.89, 169,89, 169.89, 169.87, 169.51, 169.46, 169.29, 169.26, 169.10, 169.10, 169.01, 169.01 (12 C=O(Ac); 1 C=O(COO*t*Bu)), 156.73 (C=O(Fmoc)), 154.13, 153.96 (2 C=O(Troc)), 143.74, 143.71 (C-1$_a$-; C-8$_a$-Fmoc), 140.83, 140.80 (C-4$_a$-; C-5$_a$-Fmoc), 127.68, 127.65 (C-3-; C-6-Fmoc), 127.04, 126.99 (C-2-; C-7-Fmoc), 125.14, 125.09 (C-1-; C-8-Fmoc), 120.22, 120.18 (C-4-; C-5-Fmoc), 101.30 (C-1’), 100.11 (C-1’’), 100.03 (C-1’’’), 99.64 (C-1’’’’), 99.11 (C-1), 96.17, 96.09 (2 CCl_3(Troc)), 81.23 (C_q(O*t*Bu)), 77.30 (C-3), 76.68 (C-4’), 74.81, 74.58, 74.58 (C-5; C-5’’’; C-4’’’), 74.24 (T^{β}), 73.39, 73.32 (2 Troc-CH_2), 73.01 (C-3’), 71.06 (C-5’), 70.30, 70.30 (C-3’’; C-3’’’’), 72.27, 72.27 (C-5’’’’; C-6), 69.94 (C-5’’), 69.00, 69.00 (C-2’’; C-2’’’’), 68.66 (C-4), 67.11, 67.05 (C-4’’; C-4’’’’), 65.49 (CH_2-Fmoc), 62.09 (C-6’), 61.10, 61.05, 60.83 (C-6’’; C-6’’’; C-6’’’’), 59.76 (T^{α}), 55.76 (C-2’), 55.62 (C-2’’’), 46.87, 46.79 (C-9-Fmoc; C-2), 27.62 (CH_3(O*t*Bu)), 25.77 (Me(Si*t*Bu)), 22.78, 20.60, 20.54, 20.47, 20.44, 20.44, 20.37, 20.37, 20.37, 20.33, 20.33, 20.33 (Me(Ac)), 18.81 (T^{γ}), 17.96 (C_q(Si*t*Bu)).

***N*-9-Fluorenylmethyloxycarbonyl-*O*-{2-acetamido-2-deoxy-3-*O*-[3,6-di-*O*-acetyl-2-deoxy-2-*N*-(2,2,2-trichloroethoxycarbonyl-*N*-phenylsulfidyl)-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-*D*-galactopyranosyl)-β-*D*-glucopyranosyl]-6-*O*-[3-*O*-acetyl-2-deoxy-2-*N*-(2,2,2-trichloroethoxycarbonyl)-6-*O*-*tert*-butyldimethylsilyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-*D*-galactopyranosyl)-β-*D*-glucopyranosyl]-α-*D*-galactopyranosyl}-*L*-threonine-*tert*-butylester (54)**



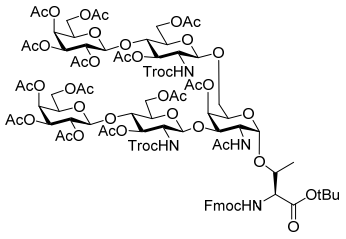
A suspension of core 3 type-2 acceptor **48** (2.06 g, 1.41 mmol, 1.0 eq.), LacNac type-2 donor **29** (1.76 g, 1.98 mmol, 1.4 eq.) and molecular sieve (1.91 g, 4 Å) in dry dichloromethane (50 mL) was stirred at 0° C for 25 min. Then *N*-iodosuccinimide (381 mg, 1.69 mmol, 1.2 eq.) was added to the slurry followed by a solution of trifluoromethanesulfonic acid (62.2 µL) in diethyl ether (1 mL). After stirring at 0° C for 9 h, the reaction mix was diluted with dichloromethane (50 mL) and filtrated over celite. The filtrate was washed twice with sat. sodium bicarbonate, once with 1 M sodium thiosulfate, once with Brine and dried over sodium sulfate. The solvent was removed in vacuo and the residue was purified by silica column chromatography (Tol/EtOAc 3:2 → 1:1 → 2:3) to give compound **54** (2.44 g, 1.07 mmol, 76 %).

R_f = 0.50 (Tol/EtOAc 1:3). $[\alpha]_D^{20}$ = +13.83 (c = 1.13, $CHCl_3$). *HR-ESI-MS (pos), m/z*: 1163.7549 ($[M+2Na]^{2+}$, calc.: 1163.7530).

^{1}H-NMR (600 MHz, MeOD, T = 278 K), *δ (ppm)*: 7.87 – 7.78 (m, 3H, H-4^A-; H-5^A-Fmoc, N-H^A), 7.75 (d, $J_{NH,H-2}$ = 10.0 Hz, 1H, N-H^B), 7.71 (d, $J_{H-1,H-2}$ = $J_{H-7,H-8}$ = 7.0 Hz, 2H, H-1-; H-8-Fmoc), 7.59 (d, $J_{H-3,H-4}$ = $J_{H-5,H-6}$ = 7.5 Hz, 2H, H-4^B-; H-5^B-Fmoc), 7.45 – 7.24 (m, 9H, H-2-; H-3-; H-6-; H-7-Fmoc, H-2-6-SPh), 5.40 – 5.25 (m, 3H, H-4''; H-4''''; CH_{2a}-Troc^A), 5.19 – 5.03 (m, 3H, H-3'''; H-3''; H-3'), 5.02 – 4.94 (m, 3H, H-2''''; H-3''''; CH_{2a}-Troc^B), 4.94 – 4.86 (m, 2H, H-2''; H-1'), 4.84 – 4.75 (m, 1H, H-1''''), 4.74 – 4.51 (m, 11H, H-6'$_a$; CH_2-Fmoc^A; CH_2-Fmoc^B; CH_{2b}-Troc^A; CH_{2b}-Troc^B; H-1; H-1''^{A}; H-1''^{B}; H-1'''), 4.51 – 4.43 (m, 1H, H-2^A), 4.40 – 4.33 (m, 1H, H-2^B), 4.33 – 4.20 (m, 3H, H-9-Fmoc^A; H-9-Fmoc^B; T$^\beta$), 4.20 – 4.03 (m, 9H, T$^\alpha$; H-6''''$_{a,b}$; H-6''$_{a,b}$; H-6'$_b$; H-4^A; H-4^B; H-5''), 4.03 – 3.94 (m, 3H, H-6'''$_a$; H-6$_a$; H-5''''), 3.94 – 3.87 (m, 2H, H-6'''$_b$; H-4'''), 3.87 – 3.78 (m, 1H, H-4'^{A}), 3.77 – 3.61 (m, 5H, H-6$_b$; H-5; H-5'; H-4'^{B}; H-3), 3.54 (dd, $J_{H-2,H-3}$ = 9.6 Hz, 1H, H-2'''), 3.41 (d, $J_{H-4,H-5}$ = 8.9 Hz, 1H, H-5'''), 2.19 – 1.87 (m, Me(Ac)), 1.80 – 1.70 (m, Me(Ac)), 1.44 (s, 9H, Me(O*t*Bu)), 1.25 – 1.06 (m, 6H, T$^{\gamma A}$; T$^{\gamma B}$), 0.97 (s, 9H, Me(Si*t*Bu)), 0.16 (s, 6H, SiMe).

^{13}C-NMR (150.9 MHz, MeOD, T = 278 K): δ = 173.55, 172.16, 172.08, 172.02, 171.99, 171.56, 171.52, 171.14, 171.03, 170.98 (C=O(Ac); C=O(COO*t*Bu)), 159.53, 159.48, 159.40, 159.34 (C=O(Fmoc)), 156.71, 156.71 (2 C=O(Troc)), 145.45, 145.30 (C-1$_a$-; C-8$_a$-Fmoc), 142.06, 142.06 (C-4$_a$-; C-5$_a$-Fmoc), 138.49, 138.12 (C_{ipso}(SPh)^A; C_{ipso}(SPh)^B), 130.40, 130.40, 129.02, 129.02, 128.87, 128.40, 128.40, 126.95, 126.95 (C-2-; C-3-; C-6-; C-7-Fmoc; C-2-6-SPh), 126.20, 126.20 (C-1-; C-8-Fmoc), 121.22, 121.19 (C-4-; C-5-Fmoc), 103.56 (C-1'), 102.36 (C-1'''), 102.08, 101.90 (C-1''^{A}; C-1''^{B}), 101.61 (C-1''''), 100.67 (C-1), 97.35, 97.35, (2 CCl_3(Troc)), 83.66 (C_q(O*t*Bu)), 80.55 (C-3), 78.20 (C-4'), 77.57, 77.45 (CH_2-Troc^{A1}; CH_2-Troc^{B1}), 76.62 (C-5'''), 76.45 (T$^\beta$), 76.17 (C-4'''), 75.50, 75,50 (CH_2-Troc^{A2}; CH_2-Troc^{B2}), 75.21 (C-5'), 74.75 (C-3'''), 74.01 (C-5), 72.64, 72.57, 72.45 (C-3''; C-3''''), 71.90 (C-5''^{B}), 71.77 (C-5''^{A}), 71.52(C6), 71.45 (C-5''''), 71.13 (C-3''), 71.05 (C-4^C), 70.68, 70.68 (C-2''; C-2''''), 70.43 (C-4^B), 70.37 (C-4^A), 68.72, 68.58 (C-4''; C-4''''), 67.71, 67.48 (CH_2-Fmoc^A; CH_2-Fmoc^B), 62.67, 62.67, 62.43, 62.35 (C-6'; C-6''; C-6'''; C-6''''), 61.12, 61.04 (T$^{\alpha A}$; T$^{\alpha B}$), 57.47 (C-2'''), 48.74, 48.50, 48.65, 48.45 (C-2^A; C-2^B; C-9-Fmoc^A; C-9-Fmoc^B), 28.61 (Me(O*t*Bu)), 26.69 (Me(Si*t*Bu)), 24.67, 24.34, 24.09, 21.85, 21.68, 21.43, 21.30, 20.95, 20.87, 20.81, 20.68, 20.62, 20.10, 20.01 (Me(Ac); C_q(Si*t*Bu)), 19.40, (T$^\gamma$).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-{2-acetamido-4-_O_-acetyl-2-deoxy-3-_O_-[3,6-di-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-6-_O_-[3,6-di-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-α-_D_-galactopyranosyl}-_L_-threonine-_tert_-butylester (55)



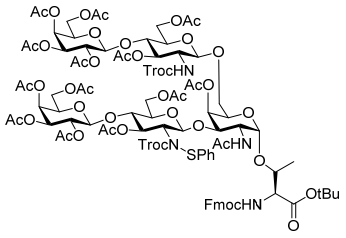
A solution of compound **53** (1.98 g, 909 µmol) in 80 % aqueous acetic acid (100 mL) was stirred at 50° C for 24 h. The solvent was removed and the residue was dissolved in Ac_2O/pyridine (1:2, 100 mL) and stirred overnight at room temperature. The reaction mix was concentrated and the residue was purified by silica column chromatography (Tol/EtOAc 1:2) to give compound **55** (1.55 g, 720 µmol, 79 %).

R_f = 0.70 (Tol:EtOAc 1:5). $[\alpha]_D^{20}$ = +30.31 (c = 0.87, $CHCl_3$) *HR-ESI-MS (pos), m/z*: 1072.2395 ($[M+2H]^+$, calc.: 1072.2380).

^{1}H-NMR (500 MHz, DMSO), *δ (ppm)*: 7.90 (d, $J_{H\text{-}3,H\text{-}4}$ = $J_{H\text{-}5,H\text{-}6}$ = 7.5 Hz, 2H, H-4-; H-5-Fmoc), 7.81 – 7.69 (m, 3H, H-1-; H-8-Fmoc; N-H'), 7.69 – 7.57 (m, 1H, N-H'''), 7.48 – 7.37 (m, 3H, H-3-; H-6-Fmoc; N-H-Fmoc), 7.37 – 7.28 (m, 3H, H-2-; H-7-Fmoc; N-H), 5.29 – 5.19 (m, 3H, H-4; H-4''; H-4''''), 5.16 (dd, $J_{H\text{-}2,H\text{-}3}$ = 10.3, $J_{H\text{-}3,H\text{-}4}$ = 3.6 Hz, 2H, H-3''; H-3''''), 5.06 – 4.98 (m, 1H, H-3''''), 4.98 – 4.87 (m, 3H, H-3'; 2 CH_{2a}-Troc), 4.83 (t, $J_{H\text{-}1,H\text{-}2}$ = $J_{H\text{-}2,H\text{-}3}$ = 9.0 Hz, 2H, H-2''; H-2''''), 4.76 – 4.68 (m, 2H, H-1''; H-1''''), 4.67 – 4.58 (m, 3H, CH_{2b}-Troc; H-1; H-1'''), 4.56 – 4.42 (m, 4H, CH_2-Fmoc; CH_{2b}-Troc; H-1'), 4.35 – 4.25 (m, 3H, H-9-Fmoc; H-6'$_a$; H-6''$_a$), 4.24 – 4.16 (m, 2H, H-5''; H-5''''), 4.14 – 3.89 (m, 9H, H-2; H-6'$_b$; H-6'''$_b$; H-6''$_{a,b}$; H-6''''$_{a,b}$; H-5; T^{β}), 3.84 – 3.63 (m, 4H, H-6$_a$; H-3; H-4'; H-4'''), 3.60 (t, $J_{H\text{-}5,H\text{-}6}$ = 7.1 Hz, 1H, H-5'), 3.55 – 3.47 (m, 1H, H-5'''), 3.43 (dd, *J* = 9.4 Hz, 1H, H-2'), 3.37 – 3.19 (m, 2H, H-6$_b$; H-2'''), 2.13 – 1.75 (m, 42H, Me(Ac)), 1.35 (s, 9H, O*t*Bu), 1.11 (d, $J_{T\beta,T\gamma}$ = 6.1 Hz, 3H, T^{γ}).

^{13}C-NMR (150.9 MHz, DMSO): δ = 170.31, 170.23, 169.91, 169.88, 169.88, 169.88, 169.88, 169.88, 169.52, 169.52, 169.52, 169.24, 169.21, 169.10, 169.10, 169.06, 169.05 (14 C=O(Ac), 1 C=O(COO*t*Bu)), 156.77 (C=O(Fmoc)), 154.14, 154.00 (2 C=O(Troc)), 143.73, 143.73 (C-1$_a$-; C-8$_a$-Fmoc), 140.84, 140.81 (C-4$_a$-; C-5$_a$-Fmoc), 127.67, 127.67 (C-3-; C-6-Fmoc), 127.06, 127.06 (C-2-; C-7-Fmoc), 125.11, 125.07 (C-1-; C-8-Fmoc), 120.23, 120.19 (C-4-; C-5-Fmoc), 100.51 (C-1'''), 100.16 (C-1'), 100.05, 99.91 (C-1''; C-1''''), 98.93 (C-1), 96.12, 96.08 (2 CCl_3(Troc)), 81.33 (C_q(O*t*Bu)), 76.42, 76.27 (C-4'; C-4'''), 74.59 (T^{β}), 73.76 (C-3), 73.40, 73.36 (2 Troc-CH_2), 73.15 (C-3'), 72.82 (C-3'''), 71.73 (C-5'), 71.48 (C-5'''), 70.55 (C-4), 70.25, 70.24 (C-3''; C-3''''), 69.67, 69.62 (C-5''; C-5''''), 69.42 (C-6), 68.96, 68.89, 68.68 (C-5; C-2''; C-2''''), 67.05, 67.05 (C-4''; C-4''''), 65.48 (CH_2-Fmoc), 62.24, 61.77 (C-6'; C-6'''), 60.87, 60.80 (C-6''; C-6''''), 59.71 (T^{α}), 56.01 (C-2'''), 55.64 (C-2'), 47.60 (C-2), 46.79 (C-9-Fmoc), 27.59 (CH_3(O*t*Bu)), 22.67, 20.61, 20.50, 20.46, 20.46, 20.41, 20.41, 20.41, 20.31, 20.31, 20.26, 20.26, 20.26, 20.26 (14 CH_3(Ac)), 18.76 (T^{γ}).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-{2-acetamido-4-_O_-acetyl-2-deoxy-3-_O_-[3,6-di-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl-_N_-phenylsulfidyl)-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-6-_O_-[3,6-di-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-α-_D_-galactopyranosyl}-_L_-threonine-_tert_-butylester (56)



A solution of compound **54** (2.28 g, 997 µmol) in 80 % aqueous acetic acid (100 mL) was stirred at 50° C for 24 h. Then the solvent was removed in vacuo and the residue was dissolved in Ac_2O/pyridine (100 mL, 1:2) and stirred overnight at room temperature. The reaction mix was concentrated and the residue was purified by silica column chromatography (Tol/EtOAc 1:1 → 1:2) to give compound **56** (2.06 g, 920 µmol, 92 %).

R_f = 0.85 (Tol/EtOAc 1:5). $[\alpha]_D$ = 28.30 (c = 0.98, $CHCl_3$). *HR-ESI-MS (pos), m/z*: 1126.7408 ($[M+2H]^{2+}$, calc.: 1126.7384), 1145.7152 ($[M+K+H]^{2+}$, calc.: 1145.7163).

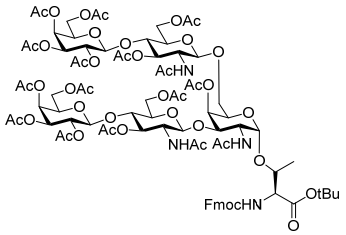
^{1}H-NMR (600 MHz, MeOD, T = 278 K), *δ (ppm)*: 7.90 – 7.75 (m, 3H, H-4^A-; H-5^A-Fmoc; N-H-Fmoc), 7.71 (d, $J_{H\text{-}1,H\text{-}2}$ = $J_{H\text{-}7,H\text{-}8}$ = 6.8 Hz, 3H, H-1-; H-8-Fmoc; N-H), 7.62 (d, $J_{H\text{-}3,H\text{-}4}$ = $J_{H\text{-}5,H\text{-}6}$ = 7.5 Hz, 3H, H-4^B-; H-5^B-Fmoc), 7.52 – 7.24 (m, 9H, H-2-; H-3-; H-6-; H-7-Fmoc; C-2-6-SPh), 5.45 – 5.25 (m, 3H, H-4''; H-4''''; H-3'^{A}), 5.21 – 5.03 (m, 5H, H-3''; H-3''''; H-4; H-3''', CH_{2a}-Troc^A), 5.03 – 4.97 (m, 2H, H-2''; H-2''''), 4.97 – 4.86 (m, 3H, H-2'; CH_{2a}-Troc^B, H-1'^{A}), 4.77 – 4.56 (m, 9H, H-6'$_a$; CH_2-Fmoc; CH_{2b}-Troc^A; CH_{2b}-Troc^B; H-1; H-1'^{B}; H-1''; H-1''''), 4.57 – 4.46 (m, 2H, H-6'''$_a$; H-1'''), 4.38 – 4.22 (m, 3H, H-2; H-9-Fmoc; T$^\beta$), 4.21 – 3.93 (m, 3H, T$^\alpha$; H-6'''$_b$; H-6'$_b$; H-6''$_{a,b}$; H-6''''$_{a,b}$; H-5; H-5''; H-5''''), 3.92 – 3.81 (m, 1H, H-6$_a$), 3.81 – 3.71 (m, 3H, H-3; H-4'; H-4'''), 3.71 – 3.64 (m, 1H, H-5'''), 3.63 – 3.57 (m, 1H, H-5'); 3.54 (dd, *J* = 9.5 Hz, 1H, H-2'''), 3.49 – 3.41 (m, 1H, H-6$_b$), 2.19 – 1.87 (m, Me(Ac)), 1.44 (s, 9H, Me(O*t*Bu)), 1.26 – 1.18 (m, 3H, T$^\gamma$).

^{13}C-NMR (150.9 MHz, MeOD, T = 278 K): δ = 173.60, 172.65, 172.57, 172.13, 172.08, 172.02, 171.95, 171.81, 171.56, 171.53, 171..,21, 171.03, 170.91 (C=O(Ac); C=O(COO*t*Bu)), 159.45 (C=O(Fmoc)), 156.72, 156.65 (C=O(Troc)^A; C=O(Troc)^B), 145.45, 145.30 (C-1$_a$-; C-8$_a$-Fmoc), 142.87, 142.87 (C-4$_a$-; C-5$_a$-Fmoc), 138.53 (C_{ipso}(SPh)), 130.47, 130.47, 130.47, 129.05, 129.05, 129.05, 128.43, 128.43, 128.43 (C-2-; C-3-; C-6-; C-7-Fmoc; C-2-6-SPh), 126.26, 126.18 (C-1-; C-8-Fmoc), 121.24, 121.20 (C-4-; C-5-Fmoc), 102.81 (C-2'); 102.19, 102.14, 101.94 (C-1''; C-1'''; C-1''''), 100.70 (C-1), 97.35 (CCl_3(Troc)),

83.73 (C_q(O*t*Bu)), 78.28, 78.21, 78.10 (C-3; C-4'; C-4'''), 77.35 (CH_2-Troc^A), 76.84 (T^β), 75.51 (CH_2-Troc^B), 74.50 (C-5'''), 74.04 (C-5'), 72.53, 72.53, 72.53 (C-3'; C-3''; C-3''''), 71.78, 71.78, (C-5''; C-5''''), 70.76 (C-6), 70.68, 70.68, 70.68, 70.56 (C-2'; C-2''; C-2''''; C-5), 68.63, 68.60 (C-4''; C-4''''), 67.65 (CH_2-Fmoc), 63.75 (C-6'''), 62.65, 62.40, 62.32 (C-6'; C-6''; C-6''''), 61.08, 61.00 ($T^{\alpha A}$; $T^{\alpha B}$), 57.55, 57.46 (C-2'''^{A}; C-2'''^{B}), 48.60, 48.60 (C-2; C-9-Fmoc), 28.64, 28.64, 28.64 (Me(O*t*Bu)), 24.28, 23.98, 21.41, 21.31, 21.08, 20.92, 20.82, 20.81, 20.68, 20.64 (Me(Ac)), 20.09 (T^γ).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-{2-acetamido-4-_O_-acetyl-2-deoxy-3-_O_-[2-acetamido-3,6-di-_O_-acetyl-2-deoxy-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-6-_O_-[2-acetamido-3,6-di-_O_-acetyl-2-deoxy-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-_D_-galactopyranosyl)-β-_D_-glucopyranosyl]-α-_D_-galactopyranosyl}-_L_-threonine-_tert_-butylester (57)

From **55**:



A suspension of compound **55** (1.37 g, 639 µmol, 1.0 eq.) and cadmium powder (1.44 g, 12.8 mmol, 20.0 eq.) in acetic acid (14 mL) was stirred at 50° C for five days. The reaction mix was concentrated and the residue dissolved in acetic anhydride/pyridine (14 mL, 1:2) and stirred at room temperature for 24h. Then the solvent was removed in vacuo and the residue was purified by silica column chromatography (EtOAc/MeOH 1:0 → 20:1 → 10:1) to give compound **57** (856 mg, 456 µmol, 71 %).

From **56**:

To a dry Schlenk tube, samarium (1.48 g, 9.85 mmol, 20.0 eq., 40 mesh) was added and the system was flushed with argon. Then dry tetrahydrofuran (44 mL) and methyl diiodide (397 µL, 4.93 mmol, 10.0 eq.) were added. The formation of samarium(II) iodide was observed after a few minutes recognized by conversion to a deep blue solution. Stirring was continued for 1 h and the reaction mix was then cooled to 0 °C. A solution of compound **56** (1.11 g, 493 µmol, 1.0 eq.) and *tert*-butanol (4.93 mmol, 10.0 eq.) indry tetrahydrofuran (1 mL) was added dropwise to the samarium(II) iodide reaction mix. After stirring for 3 h at room temperature, all the starting material was consumed and a solution of acetic acid (1.7 mL) in water (40 mL) was poured into the reaction mix. Next, the *p*H was adjusted

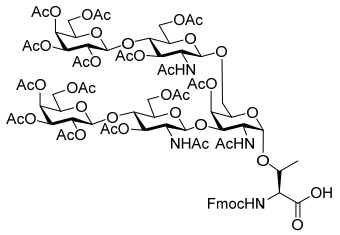
to 7 using solid sodium bicarbonate. The suspension was diluted with ethyl acetate (40 mL) and a small portion of celite was added. After 5 min of stirring the mixture was filtered over celite. The filtrate was transferred to a separating funnel and sodium chloride was added until saturation. After separation of the organic phase, the aqueous phase was washed with additional ethyl acetate. The organic phases were pooled, dried with sodium sulfate and concentrated in vacuo. The residue was dissolved in ethyl acetate and additional insoluble salts were removed followed by evaporation of the ethyl acetate again. Finally the residue was taken up in acetic anhydride/pyridine (1:2, 10 mL) and stirred at room temperature overnight. Then the solvent was removed in vacuo and the residue was purified by silica column chromatography (EA/MeOH 50:1 → 0:1) to give compound **57** (627 mg, 334 µmol, 68 %).

R_f = 0.3 (EtOAc/MeOH/AcOH/H_2O 30:3:3:2). $[\alpha]_D^{20}$ = +18.16 (c = 0.98, $CHCl_3$). *HR-ESI-MS (pos), m/z*: 939.3433 ($[M+2H]^{2+}$, calc.: 939.3429), 958.3170 ($[M+K+H]^{2+}$, calc.: 958.3208).

^{1}H-NMR (600 MHz, MeOD, T = 278 K, gCOSY, TOCSY, gHSQC, gHMBC), *δ (ppm)*: 8.28 (d, $J_{N\text{-}H,H\text{-}2}$ = 9.3 Hz, 1H, N-H'), 8.05 (d, $J_{N\text{-}H,H\text{-}2}$ = 8.6 Hz, 1H, N-H'''), 7.84 (d, $J_{H\text{-}3,H\text{-}4}$ = $J_{H\text{-}5,H\text{-}6}$ = 7.6 Hz, 2H, H-4-; H-5-Fmoc), 7.73 – 7.68 (m, 2H, H-1-; H-8-Fmoc), 7.64 (d, $J_{N\text{-}H,H\text{-}2}$ = 10.0 Hz, 1H, NH), 7.48 (d, $J_{N\text{-}H,T\alpha}$ = 10.0 Hz, 1H, N-H-Fmoc), 7.45 – 7.39 (m, 2H, H-3-; H-6-Fmoc), 7.38 – 7.31 (m, 2H, H-2-; H-7-Fmoc), 5.39 – 5.25 (m, 4H, H-4''; H-4''''; H-4; H-3'''), 5.16 – 5.08 (m, 2H, H-3''; H-3''''), 5.08 – 4.94 (m, 3H, H-2''; H-2''''; H-3'), 4.82 (d, $J_{H\text{-}1,H\text{-}2}$ = 6.2 Hz, 1H, H-1'''), 4.74 – 4.66 (m, 3H, H-1; H-1''; H-1''''), 4.63 – 4.53 (m, 3H, H-6'''$_a$; Fmoc-CH_2), 4.52 – 4.46 (m 2H, H-1'; H-6'$_a$), 4.31 – 4.22 (m, 3H, H-2; H-9-Fmoc; T^{β}), 4.20 – 4.07 (m, 9H, T^{α}; H-6''$_{a,b}$; H-6''''$_{a,b}$; H-6'$_b$; H-5; H-5''; H-5''''), 4.05 (dd, $J_{H\text{-}6a,H\text{-}6b}$ = 12.2, $J_{H\text{-}5,H\text{-}6}$ = 4.4 Hz, 1H, H-6'''$_b$), 3.92 – 3.83 (m, 2H, H-2'; H-6$_a$), 3.83 – 3.71 (m, 3H, H-3; H-4'; H-4'''), 3.66 (t, $J_{H\text{-}5,H\text{-}6}$ = $J_{H\text{-}4,H\text{-}5}$ = 8.1 Hz, 1H, H-5'), 3.58 (bd, $J_{H\text{-}5,H\text{-}6}$ = 8.4 Hz, 1H, H-5'''), 3.42 – 3.34 (m, 2H, H-6$_b$; H-2'''), 2.21 – 1.81 (m, 48H, CH_3(Ac)), 1.42 (s, 9H, O*t*Bu), 1.20 (d, $J_{T\beta,T\gamma}$ = 6.3 Hz, 3H, T^{γ}).

^{13}C-NMR (150.9 MHz, MeOD, T = 278 K): δ = 173.55, 173.43, 173.34, 173.20, 172.78, 172.56, 172.09, 172.09, 172.09, 172.07, 172.07, 171.82, 171.56, 171.56, 171.25, 171.22, 171.11 (16 C=O(Ac); 1 C=O(COO*t*Bu)), 159.38 (C=O(Fmoc)), 145,45, 145.28 (C-1$_a$-; C-8$_a$-Fmoc), 142.82, 142.82 (C-4$_a$-; C-5$_a$-Fmoc), 129.04, 129.02 (C-3-; C-6-Fmoc), 128.41, 128.40 (C-2-; C-7-Fmoc), 126.25, 126.18 (C-1-; C-8-Fmoc), 121.23, 121.20 (C-4-; C-5-Fmoc), 102.50 (C-1'), 102.23, 102.17 (C-1''; C-1''''), 101.67 (C-1'''), 100.41 (C-1), 83.61 (C_q(O*t*Bu)), 78.00, 77.74 (C-4'; C-4'''), 76.30 (T^{β}), 74.99 (C-3), 74.87 (C-3'), 74.04 (C-5'), 73.80 (C-5'''), 73.54 (C-3'''), 72.49 (C-5'), 72.49 (C-5'''), 72.16 (C-4), 71.73, 71.71 (C-5''; C-5''''), 71.43 (C-6), 70.77, 70.69, 70.59 (C-2''; C-2''''; C-5), 68.60, 68.60 (C-4''; C-4''''), 67.68 (Fmoc-CH_2), 63.71 (C-6'), 62.91 (C-6'''), 62.36, 62.28 (C-6''; C-6''''), 61.01 (T^{α}), 56.71 (C-2'''), 55.26 (C-2'), 49.59, 48.82 (C-2; C-9-Fmoc), 48.49, 48.49, 48.49 (CH_3(O*t*Bu)), 23.60, 23.18, 23.18, 21.17, 21.17, 21.17, 21.09, 21.00, 20.97, 20.91, 20.81, 20.79, 20.69, 20.69, 20.69, 20.69 (16 CH_3(Ac)), 19.98 (T^{γ}).

***N*-9-Fluorenylmethyloxycarbonyl-*O*-{2-acetamido-4-*O*-acetyl-2-deoxy-3-*O*-[2-acetamido-3,6-di-*O*-acetyl-2-deoxy-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-*D*-galactopyranosyl)-β-*D*-glucopyranosyl]-6-*O*-[2-acetamido-3,6-di-*O*-acetyl-2-deoxy-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-*D*-galactopyranosyl)-β-*D*-glucopyranosyl]-α-*D*-galactopyranosyl}-*L*-threonine (58)**



To a solution of compound **57** (1.52 g, 808 µmol) in dichloromethane (4 mL), anisole (1.6 mL) and trifluoroacetic acid (12 mL) were added. After stirring for 2.5 h, the solvent was removed in vacuo and the crude residue was purified by silica column chromatography (EtOAc/MeOH/AcOH/H_2O 10:1:0:0 → 60:3:3:2 → 30:3:3:2) to give compound **58** (1.26 g, 692 µmol, 86 %).

R_f = 0.30 (EtOAc/MeOH/AcOH/H_2O 60:3:3:2). $[\alpha]_D^{20}$ = +49.59 (c = 0.98, $CHCl_3$). *HR-ESI-MS (pos), m/z*: 930.2850 ($[M+K+H]^{2+}$, calc.: 930.2895), 933.2941 ($[M+2Na]^{2+}$, calc.: 933.2935).

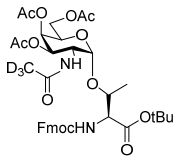
^{1}H-NMR (600 MHz, MeOD, T = 278 K), *δ (ppm)*: 8.25 (d, $J_{N\text{-}H,H\text{-}2}$ = 9.4 Hz, 1H,N-H'), 8.06 (d, $J_{N\text{-}H,H\text{-}2}$ = 8.8 Hz, 1H, N-H'''), 7.84 (d, $J_{H\text{-}3,H\text{-}4}$ = $J_{H\text{-}5,H\text{-}6}$ = 7.7 Hz, 2H, H-4-; H-5-Fmoc), 7.74 – 7.66 (m, 2H, H-1-; H-8-Fmoc), 7.52 (d, $J_{N\text{-}H,H\text{-}2}$ = 9.7 Hz, 1H, N-H), 7.45 – 7.31 (m, 5H, N-H-Fmoc; H-2-; H-3-; H-6-; H-7-Fmoc), 5.36 – 5.31 (m, 3H, H-4; H-4''; H-4''''), 5.27 – 5.19 (m, 1H, H-3'''), 5.14 – 4.98 (m, 5H, H-3', H-3''; H-3''''; H-2''; H-2''''), 4.82 (m, 1H, H-1), 4.78 (d, $J_{H\text{-}1,H\text{-}2}$ = 8.2 Hz, 1H, H-1'''), 4.74 – 4.65 (m, 2H, H-1''; H-1''''), 4.61 – 4.45 (m, 5H, H-6'$_a$; H-6'''$_a$; CH_2-Fmoc; H-1'), 4.34 – 4.20 (m, 3H, H-9-Fmoc; H-2; T^{β}), 4.20 – 4.02 (m, 10H, T^{α}; H-6'$_b$; H-6''$_{a,b}$; H-6'''$_b$; H-6''''$_{a,b}$; H-5; H-5''; H-5''''), 3.91 – 3.80 (m, 3H, H-2'; H-6$_a$; H-3), 3.79 – 3.73 (m, 2H, H-4'; H-4'''), 3.69 – 3.62 (m, 1H, H-5'), 3.61 – 3.55 (m, 1H, H-5'''), 3.49 – 3.43 (m, 1H, H-2'''), 3.41 – 3.36 (m, 1H, H-6$_b$), 2.17 – 1.83 (m, 48H, Me(Ac)), 1.20 (d, $J_{H\beta,H\gamma}$ = 6.3 Hz, 3H, H^{γ}).

^{13}C-NMR (150.9 MHz, MeOD, T = 278 K): δ = 173.78, 173.53, 173.37, 172.78, 172.57, 172.10, 172.10, 172.10, 172.08, 172.08, 172.02, 171.83, 171.56, 171.56, 171.56, 171.27, 171.23 (16 C=O(Ac); 1 C=O(COOH)), 159.18 (C=O(Fmoc)), 145.52, 145.32 (C-1$_a$-; C-8$_a$-Fmoc), 142.81, 142.81 (C-4$_a$-; C-5$_a$-Fmoc), 129.06, 129.01 (C-3-; C-6-Fmoc), 128.45, 128.41 (C-2-; C-7-Fmoc), 126.31, 126.25 (C-1-; C-8-Fmoc), 121.22, 121.20 (C-4-; C-5-Fmoc), 102.33 (C-1'), 102.24, 102.18 (C-1''; C-1''''), 101.90 (C-1'''), 100.51 (C-1), 78.09, 77.79 (C-4'; C-4'''), 77.40 (T^{β}), 75.26 (C-3), 74.90 (C-3'), 74.04 (C-5'), 73.85, 73.82 (C-3'''; C-5'''), 72.50, 72.50 (C-3''; C-3''''), 72.22 (C-4), 71.72, 71.72 (C-5''; C-5''''), 71.16 (C-6), 70.77, 70.69 (C-2''; C-2''''), 70.53 (C-5), 68.60, 68.60 (C-4''; C-4''''), 67.73 (CH_2-Fmoc), 63.77 (C-6'), 63.05 (C-

6’’’), 62.35, 62.28 (C-6’’; C-6’’’’), 60.60 (T^{α}), 56.45 (C-2’’’), 55.35 (C-2’), 50.02 (C-2), 48.77 (C-9-Fmoc), 23.46, 23.18, 23.10, 21.28, 21.28, 21.28, 21.09, 21.02, 20.96, 20.91, 20.81, 20.81, 20.79, 20.68, 20.68, 20.68 (16 CH_3(Ac)), 19.52 (T^{γ}).

9.1.2.11 Synthesis of the deuterium-labelled T_N-antigen amino acid

N-9-Fluorenylmethyloxycarbonyl-O-[2-(2,2,2-trideutero)acetamido-2-deoxy-3,4,6-O-acetyl-α-D-galactopyranosyl]-L-threonine-*tert*-butylester (59)[189]



A suspension of galactosyl-threonine **9** (500 mg, 703 µmol) and zinc dust (920 mg, 14.1 mmol) in acetic acid (5.0 mL) was stirred for 16 h at 50° C under argon. Afterwards, the formed salt was filtered off over celite and the remaining acetic acid was codistilled with toluene to dryness. Next, the residue (370 mg, 540 µmol, 1.0 eq.) was dissolved in dioxane (2.5 mL) before adding a solution of sodium bicarbonate (136 mg, 1.6 mmol, 3.0 eq.) in water (1.2 mL) was added. Finally, Ac-D_3-OSu (130 mg, 810 µmol, 1.5 eq.) was added and the solution was stirred at room temperature for further 16 h. After dilution with ethyl acetate (5 mL), the solution was washed twice with water and once with brine before drying over sodium sulfate. After removal of the solvent in vacuo, the residue was purified by column chromatography on silica (CH/EtOAc 1:1 → 1:2) to give the deuterated acetamide **59** (172 mg, 236 µmol, 44 %).

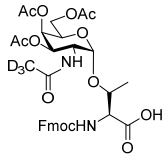
R_f = 0.38 (CH/EtOAc 1:3).

^{1}H-NMR (600 MHz, DMSO), δ (*ppm*): 7.90 (d, $J_{H4,H3} = J_{H5,H6}$ = 7.7 Hz, 2H, H-4-; H-5-Fmoc), 7.73 (m, 3H, NH; H1-; H-8-Fmoc), 7.63 (d, $J_{NH,T\alpha}$ = 9.7 Hz, 1H, NH-Fmoc), 7.42 (m, 2H, H-3-; H-6-Fmoc), 7.33 (m, 2H, H-2-; H-7-Fmoc), 5.32 (d, $J_{H4,H3}$ = 3.2 Hz, 1H, H-4), 5.02 (dd, $J_{H3,H2} = J_{H3,H4}$ = 11.6, 3.3 Hz, 1H, H-3), 4.78 (d, $J_{H1,H2}$ = 3.9 Hz, 1H, H-1), 4.51 – 4.37 (m, 2H,CH_2-Fmoc), 4.33 – 4.24 (m, 3H, H-9-Fmoc; H-2; T^{β}), 4.21 (t, $J_{H5,H6}$ = 6.3 Hz, 1H, H-5), 4.10 (d, $J_{T\alpha,NH}$ = 10.7 Hz, 1H, T^{α}), 4.06 – 4.00 (m, 2H, H-$6_{a,b}$), 2.11 (s, 3H, CH_3(Ac)), 2.00 (s, 3H, CH_3(Ac)), 1.91 (s, 3H, CH_3(Ac)), 1.36 (s, 9H, O*t*Bu), 1.19 (d, $J_{T\gamma,T\beta}$ = 6.4 Hz, 2H, T^{γ}).

^{13}C-NMR (150.9 MHz, DMSO): δ = 170.04, 169.97, 169.93, 169.26, 169.00 (3 C=O(Ac), 1 C=O(Ac-d3), 1 C=O(COOtBu)), 156.87 (C=O(Fmoc)), 143.71 (C1_a-; C-8_a-Fmoc), 140.81, 140.77 (C-4_a-; C-5a-Fmoc), 127.72, 127.68 (C-3-; C-6-Fmoc), 127.09, 127.09 (C-2-; C-7-Fmoc), 125.23, 125.14 (C-1-; C-8-Fmoc), 120.22, 120.18 (C-4-; C-5-Fmoc), 98.80 (C-1), 81.44 (C_q(*t*Bu)), 74.59 (T^{β}), 67.80 (C-3), 67.24 (C-4), 66.51

(C-5), 66.65 (Fmoc-CH_2), 62.03 (C-6), 59.15 (T^{α}), 46.75 (C-9-Fmoc), 46.23 (C-2), 27.60 (Me(*t*Bu)), 20.52, 20.50, 20.50, 20.47 (4 Me(Ac)), 18.91 (T^{γ}).

N-9-Fluorenylmethyloxycarbonyl-O-[2-(2,2,2-trideutero)acetamido-2-deoxy-3,4,6-O-acetyl-α-D-galactopyranosyl]-L-threonine (60)[189]



After dissolving compound **59** (170 mg, 233 µmol) in dichloromethane (425 µL), anisole (340 µL) and trifluoroacetic acid (1.28 mL) were added in that order. After stirring for 3.5 h at room temperature the solvent was removed by codistillation with toluene and the residue was purified by column chromatography on silica (EtOAc/MeOH/AcOH/H_2O 1:0:0:0 → 60:3:3:2) to give compound **60** (114 mg, 169 µmol, 73 %).

R_f = 0.56 (EtOAc/MeOH/AcOH/H_2O 60:3:3:2).

^{1}H-NMR (600 MHz, DMSO), δ (*ppm*): 7.90 (d, $J_{H4,H3} = J_{H5,H6}$ = 7.7 Hz, 2H, H-4-; H-5-Fmoc), 7.74 (m, 2H, H-1-; H-8-Fmoc), 7.64 (d, $J_{NH,H2}$ = 9.5 Hz, 1H, N-H), 7.56 (d, $J_{NH,T\alpha}$ = 9.7 Hz, 1H, N-H-Fmoc), 7.42 (m, 2H, H-3-; H-6-Fmoc), 7.33 (m, 2H, H-2-; H-7-Fmoc), 5.31 (d, $J_{H4,H3}$ = 4.4 Hz, 1H, H-4), 5.04 (dd, $J_{H3,H2}$ = 11.7, $J_{H3,H4}$ = 3.2 Hz, 1H, H-3), 4.81 (d, $J_{H1,H2}$ = 3.8 Hz, 1H, H-1), 4.50 – 4.37 (m, 2H, CH_2-Fmoc), 4.33 – 4.27 (m, 2H, H-9-Fmoc; T^{β}), 4.26 – 4.18 (m, 2H, H-2; H-5), 4.13 (d, $J_{T\alpha,NH}$ = 10.2 Hz, 1H, T^{α}), 4.05 – 3.98 (m, 2H, H-$6_{a,b}$), 2.11 (s, 3H, CH_3(Ac)), 1.99 (s, 3H, CH_3(Ac)), 1.91 (s, 3H, CH_3(Ac)), 1.17 (d, $J_{T\gamma,T\beta}$ = 6.4 Hz, 3H, T^{γ}).

^{13}C-NMR (100.6 MHz, DMSO): δ = 171.62, 170.01, 169.91, 169.85, 169.48 (3 C=O(Ac), 1 C=O(Ac-d3), 1 C=O(COOH), 156.84 (C=O(Fmoc)), 143.71 (C-1_a-; C-8_a-Fmoc), 140.77, 140.72 (C-4_a-; C-5_a-Fmoc), 127.65, 127.61 (C-3-; C-6-Fmoc), 127.04, 127.04 (C-2-; C-7-Fmoc), 125.20, 125.10 (C-1-; C-8-Fmoc), 120.16, 120.11 (C-4-; C-5-Fmoc), 98.72 (C-1), 74.88 (T^{β}), 67.62 (C-3), 67.21 (C-4), 66.42 (C-5), 65.57 (Fmoc-CH_2), 61.94 (C-6), 58.37 (T^{α}), 46.74 (C-9-Fmoc), 46.49 (C-2), 20.50, 20.44, 20.44, 20.44 (4 Me(Ac)), 18.52 (T^{γ}).

9.1.3 Synthesis of the HexNAc-Tyr amino acid building blocks

9.1.3.1 Synthesis of the tyrosine amino acid acceptor

N-9-Fluorenylmethyloxycarbonyl-L-tyrosine-allyl ester (62)[365]

To a suspension of L-Tyr (10.00 g, 55.19 mmol, 1.0 eq.) in dioxane/water (3:1, 150 mL), sodium bicarbonate (13.91 g, 165.57 mmol, 3.0 eq.) was added and stirred. Then an ice bath was installed and Fmoc-OSu (18.62 g, 55.19 mmol, 1.0 eq.) was added in portions. Afterwards stirring was continued for four days at room temperature and completion of the reaction was followed by TLC (R_f = 0.5, CH/EtOAc 1:2 + 1% AcOH). Next, more water (450 mL) was added and it was washed with diethyl ether. The aqueous phase was acidified to *p*H 3 using potassium hydrogen sulfate (approx. 3 eq.) and saturated with sodium chloride. After extraction with ethyl acetate, the organic phase was washed with brine and dried over sodium sulfate. After removal of the solvent 11.83 g of a solid crude was obtained. The crude dried in vacuo and dissolved in dry tetrahydrofuran (77 mL) before addition of diisopropylamine (7.58 mL, 58.65 mmol, 2 eq.). Then allyl bromide (14.13 mL, 58.65 mmol, 2 eq.) were added dropwise and the solution was stirred overnight at room temperature. A solution of ice water (250 mL) with citric acid (17 g) was prepared and the reaction mixture was poured into it. The resulting suspension was extracted with three times ethyl acetate (130 mL) and the organic phase was then washed with sat. sodium bicarbonate and brine. After drying over sodium sulfate and removal of the solvent, the crude was purified by column chromatography on silica (CH/EtOAc 4:1 → 1:1). Finally recrystallization from ethyl acetate/diethyl ether (1:1) and cyclohexane was giving Fmoc-tyrosine-allyl ester **62** as a colorless solid (7.67 g, 17.3 mmol, 31 %).

R_f = 0.41 (CH/EtOAc 2:1). $[\alpha]_D^{20}$ = +14.14 (c = 9.90, $CHCl_3$). *HR-ESI-MS (pos), m/z:* 444.1806 ($[M+H]^+$, calc.: 444.1805), 466.1620 ($[M+Na]^+$, calc.: 466.1625).

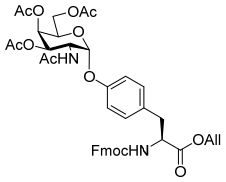
^{1}H-NMR (500 MHz, DMSO), *δ (ppm)*: 7.91 – 7.79 (m, 3H, NH-; H-4-; H-5-Fmoc), 7.68 – 7.62 (m, 2H, H-1-; H-8-Fmoc), 7.45 – 7.38 (m, 2H, H-3-; H-6-Fmoc), 7.35 – 7.28 (m, 2H, H-2-; H-7-Fmoc), 7.06 (d, $J_{Tyr\text{-}o,Tyr\text{-}m}$ = 8.3 Hz, 2H, Tyr-$o_{a,b}$), 6.68 (d, $J_{Tyr\text{-}m,Tyr\text{-}o}$ = 8.4 Hz, 2H, Tyr-$m_{a,b}$), 5.90 – 5.80 (m, 1H, H-2-OAll), 5.33 – 5.15 (m, 2H, H-3-$OAll_{a,b}$), 4.60 – 4.53 (m, 2H, H-1-$OAll_{a,b}$), 4.31 – 4.11 (m, 4H, CH_2-Fmoc; Tyr^{α}; H-9-Fmoc), 3.00 – 2.78 (m, 2H, Tyr^{β}).

^{13}C-NMR (125.8 MHz, DMSO): *δ* = 171.6 (C=O (COOAll)), 156.0, 156.0, 155.9 (C=O(Fmoc); Tyr-*y*; Tyr-*p*), 143.7, 143.7 (C-1_a-; C-8_a-Fmoc), 140.7, 140.7 (C-4_a-; C-5_a-Fmoc), 132.3 (C-2-OAll), 130.0, 130.0 (Tyr-

$o_{a,b}$), 127.6, 127.6 (C-3-; C-6-Fmoc), 127.0, 127.0 (C-2-; C-7-Fmoc), 125.2, 125.2 (C-1-; C-8-Fmoc), 120.0, 120.0 (C-4-; C-5-Fmoc), 117.6 (C-3-OAll), 115.0, 115.0 (Tyr-$m_{a,b}$), 65.6 (Fmoc-CH_2), 64.8 (C-1-OAll), 55.9 (Tyr$^{\alpha}$), 46.6 (C-9-Fmoc), 35.7 (Tyr$^{\beta}$).

9.1.3.2 Synthesis of the α-GalNAc-Tyr amino acid

_N_-9-Fluorenylmethyloxycarbonyl-_O_-[2-acetamido-2-deoxy-3,4,6-_O_-acetyl-α-D-galactopyranosyl]-_L_-tyrosine-allylester (63)



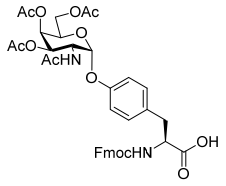
A solution of Fmoc-tyrosine-*O*-allyl ester **8** (1.56 g, 3.50 mmol, 1.0 eq.) in dry dichloromethane (30.0 mL) was added to a dry flask, containing molecular sieve (1.81 g, 4 Å). An ice bath was installed and the suspension was stirred for 1 h. Meanwhile silver perchlorate (236 mg, 1.05 mmol, 0.3 eq.) was codistilled with dry toluene (three times, each time close to dryness) and then dissolved in toluene (3.5 mL). After the stirring of the suspension was finished, first silver(II) carbonate (1.45 g, 5.24 mmol, 1.5 eq.) and then the silver perchlorate/toluene solution were added. Stirring was continued for 30 min before adding galactopyranosyl bromide **8** (2.07 g, 5.24 mmol, 1.5 eq.) in dichloromethane (13.0 mL) dropwise. Then the solution was stirred overnight, while the ice bath was allowed to melt. Next morning, additional bromide **8** (275 mg, 0.70 mmol, 0.2 eq.) was dissolved in dichloromethane (3.0 mL) and added dropwise. Stirring was continued overnight again. When the reaction was complete (R_f = 0.14, CH/EtOAc 4:1), the solution was filtered over celite with subsequent washing of the celite with dichloromethane. The filtrate was washed twice with sat. sodium bicarbonate and once with brine before drying it over sodium sulfate. After removal of the solvent, the residue was purified by column chromatography on silica (CH/EtOAc 3:1 → 2:1) to obtain a colorless solid (1.66 g, 2.20 mmol) which was directly subjected for acetylation. The compound was dissolved in pyridine (4.3 mL) before adding further thioacetic acid (4.3 mL) and the resulting solution was stirred for 72 h. Afterwards the solvent was removed by codistillation with toluene. The crude was then purified by column chromatography on silica (CH/EtOAc 2:1 → 1:2) to obtain glycoside **63** as a colorless solid (1.18 g, 1.53 mmol, 44 %).

R_f = 0.36 (CH/EtOAc 1:2). $[\alpha]_D^{20}$ = +58.79. (c = 10.50, $CHCl_3$). *HR-ESI-MS (pos), m/z:* 795.2719 ($[M+Na]^+$, calc.: 795.2736), 790.3172 ($[M+NH_4]^+$, calc.: 790.3181), 773.2903 ($[M+H]^+$, calc.: 773.2916.

[1]H-NMR (600 MHz, DMSO), *δ (ppm)*: 8.24 (d, $J_{NH,H2}$ = 8.2 Hz, 1H, N-H), 7.91 – 7.83 (m, 3H, N-H-; H-4-; H-5-Fmoc), 7.67 – 7.59 (m, 2H, H-1-; H-8-Fmoc), 7.45 – 7.37 (m, 2H, H-3-; H-6-Fmoc), 7.35 – 7.26 (m, 2H, H-2-; H-7-Fmoc), 7.22 (d, $J_{Tyr\text{-}o,Tyr\text{-}m}$ = 8.3 Hz, 2H, Tyr-$o_{a,b}$), 7.00 (d, $J_{Tyr\text{-}m,Tyr\text{-}o}$ = 8.5 Hz, 2H, Tyr-$m_{a,b}$), 5.91 – 5.77 (m, 1H, H-2-OAll), 5.49 (d, $J_{H1,H2}$ = 3.5 Hz, 1H, H-1), 5.38 (d, $J_{H4,H3}$ = 2.7 Hz, 1H, H-4), 5.30 – 5.15 (m, 3H, H-3, H-3-OAll), 4.59 – 4.53 (m, 2H, H-1-OAll), 4.42 – 4.34 (m, 1H, H-2), 4.29 – 4.20 (m, 4H, H-5; CH_2-Fmoc; Tyr^{α}), 4.21 – 4.12 (m, 1H, H-9-Fmoc), 4.07 – 3.89 (m, 2H, H-$6_{a,b}$), 3.07 – 2.82 (m, 2H, Tyr^{β}), 2.37 – 1.81 (m, 12H, Ac-H).

[13]C-NMR (150.9 MHz, DMSO), *δ (ppm)*: 171.5, 170.0, 169.9, 169.8, 169.8 (C=O(COOAll); 4 Ac), 156.0, 156.0, 155.1 (C=O(Fmoc); Tyr-*y*; Tyr-*p*), 143.7, 143.7 (C-1_a-; C-8_a-Fmoc), 140.7, 140.7 (C-4_a-; C-5_a-Fmoc), 132.3 (C-2-OAll), 130.2, 130.2 (Tyr-$o_{a,b}$), 127.6, 127.6 (C-3-; C-6-Fmoc), 127.0, 127.0 (C-2-; C-7-Fmoc), 125.2, 125.2 (C-1-; C-8-Fmoc), 120.1, 120.1 (C-4-; C-5-Fmoc), 117.8 (C-3-OAll), 117.3, 117.3 (Tyr-$m_{a,b}$), 96.5 (C-1), 67.4 (C-3), 67.0, 67.0 (C-4; C-5), 65.7 (CH_2-Fmoc), 64.9 (C-1-OAll), 61.6 (C-6), 55.7 (Tyr^{α}), 47.2 (C-2), 46.6 (C-9-Fmoc), 35.7 (Tyr^{β}), 22.4, 20.6, 20.4, 20.4 (4 Me(Ac)).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-[2-acetamido-2-deoxy-3,4,6-_O_-acetyl-α-D-galactopyranosyl]-_L_-tyrosine (64)[366]



To a solution of compound **63** (970 mg, 1.26 mmol, 1.00 eq.) in dry tetrahydrofuran (5 mL), phenylsilane (317 µL, 2.57 mmol, 2.05 eq.) was added. Then tetrakis(triphenylphosphine)palladium(0) (73 µg, 0.06 mmol, 0.05 eq.) was added and the reaction was stirred. After 30 min full conversion could be observed and the solvent was removed in vacuo. The dark brown crude was purified by column chromatography on silica (EtOAc/MeOH 20:1 → 1:1). For further purification, it was passed over a C18-Cartridge (H_2O/MeCN 100:0 → 0:100) to give the carboxylic acid **64** (530 mg, 723 µmol, 58 %).

R_f = 0.69 (EtOAc/MeOH/AcOH/H_2O 60:3:3:2). $[\alpha]_D^{20}$ = +96.84 (c = 0.95, DMF).

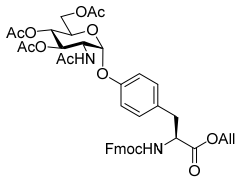
[1]H-NMR (500 MHz, DMSO) , *δ (ppm)*: 8.23 (d, $J_{NH,H2}$ = 8.1 Hz, 1H, N-H), 7.90 – 7.82 (m, 3H, NH-; H-4-; H-5-Fmoc), 7.67 – 7.59 (m, 2H, H-1-; H-8-Fmoc), 7.44 – 7.37 (m, 2H, H-3-; H-6-Fmoc), 7.36 – 7.23 (m, 2H, H-2-; H-7-Fmoc), 7.18 (d, $J_{Tyr\text{-}o,Tyr\text{-}m}$ = 8.3 Hz, 2H, Tyr-$o_{a,b}$), 6.96 (d, $J_{Tyr\text{-}m,Tyr\text{-}o}$ = 8.4 Hz, 2H, Tyr-$m_{a,b}$),

5.47 (d, $J_{H1,H2}$ = 3.3 Hz, 1H, H-1), 5.38 (d, $J_{H4,H3}$ = 3.4 Hz, 1H, H-4), 5.22 (dd, $J_{H3,H2}$ = 11.8, $J_{H3,H4}$ = 3.2 Hz, 1H, H-3), 4.43 – 4.35 (m, 1H, H-2), 4.30 – 4.21 (m, 2H, H-5; CH_{2a}-Fmoc), 4.20 – 4.11 (m, 2H, H-9-Fmoc; CH_{2b}-Fmoc), 4.10 – 4.02 (m, 1H, Tyr$^{\alpha}$), 4.00 – 3.89 (m, 2H, H-$6_{a,b}$), 3.13 – 3.04 (m, 1H, Tyr$^{\beta}{}_{a}$), 2.91 – 2.80 (m, 1H, Tyr$^{\beta}{}_{b}$), 2.55 – 1.75 (m, 12H, Me(Ac)).

^{13}C-NMR (150.9 MHz, DMSO), *δ (ppm)*: 170.0, 169.9, 169.8, 169.8 (C=O(Ac)), 155.5, 155.5, 154.7 (C=O(Fmoc); Tyr-*y*; Tyr-*p*), 143.9, 143.9 (C-1_a-; C-8_a-Fmoc), 140.7, 140.7 (C-4_a-; C-5_a-Fmoc), 130.4, 130.3 (Tyr-$o_{a,b}$), 127.6, 127.6 (C-3-; C-6-Fmoc), 127.0, 127.0 (C-2-; C-7-Fmoc), 125.3, 125.2 (C-1-; C-8-Fmoc), 120.0, 120.0 (C-4-; C-5-Fmoc), 117.0, 117.0 (Tyr-$m_{a,b}$), 96.6 (C-1), 67.4 (C-3), 67.0, 67.0 (C-4; C-5), 65.3 (CH_2-Fmoc), 61.6 (C-6), 56.7 (Tyr$^{\alpha}$), 47.2 (C-2), 46.7 (C-9-Fmoc), 39.8 (Tyr$^{\beta}$), 22.4, 20.6, 20.4, 20.4 (4 Me(Ac)).

9.1.3.3 Synthesis of the α-GlcNAc-Tyr amino acid

_N_-9-Fluorenylmethyloxycarbonyl-_O_-[2-acetamido-2-deoxy-3,4,6-_O_-acetyl-α-D-glucopyranosyl]-_L_-tyrosine-allylester (68)



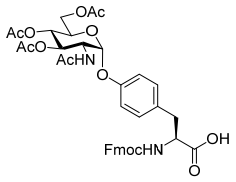
To a solution of compound **67** (1.72 g, 2.27 mmol) in pyridine 4.5 mL, thioacetic acid (4.5 mL) were added and the solution was stirred overnight at room temperature. Afterwards the solvent was removed and the residue was codistilled with toluene. The crude was purified by column chromatography on silica (CH/EtOAc 3:1 → 1:2) to give acetamide **68** as a colorless amorphous solid (1.25 g, 1.62 mmol, 70 %).

R_f = 0.16 (CH/EtOAc 1:1). $[\alpha]_D^{20}$ = +76.29 (c = 12.50, $CHCl_3$). *HR-ESI-MS (pos), m/z:* 795.2720 ($[M+Na]^+$, calc.: 795.2736).

^{1}H-NMR (500 MHz, $CDCl_3$) , *δ (ppm)*: 7.76 (d, $J_{H4,H3}$ = $J_{H5,H6}$ = 7.5 Hz, 2H, H-4-; H-5-Fmoc), 7.58 – 7.52 (m, 2H, H-1-; H-8-Fmoc), 7.42 – 7.36 (m, 2H, H-3-; H-6-Fmoc), 7.33 – 7.27 (m, 2H, H-2-; H-7-Fmoc), 7.05 – 6.95 (m, 4H, Tyr-$m_{a,b}$; Tyr-$o_{a,b}$), 5.90 – 5.83 (m, 2H, N-H, H-2-OAll), 5.50 (d, $J_{H1,H2}$ = 3.5 Hz, 1H, H-1), 5.41 (dd, *J* = 10.9, 9.4 Hz, 1H, H-4), 5.35 – 5.23 (m, 2H, H-3-OAll), 4.70 – 4.58 (m, 3H, Tyr$^{\alpha}$, H-1-OAll), 4.55 – 4.47 (m, 1H, H-2), 4.47 – 4.32 (m, 2H, $CH_{2a,b}$-Fmoc), 4.24 – 4.16 (m, 2H, H-6_a, H-9-Fmoc), 4.04 – 3.96 (m, 2H, H-6_b, H-5), 3.16 – 2.99 (m, 2H, Tyr$^{\beta}$), 2.10 – 1.90 (m, 12H, Me(Ac)).

^{13}C-NMR (125.8 MHz, $CDCl_3$), *δ (ppm)*: 171.6, 171.2, 170.6, 170.2, 169.3 (4 C=O(Ac); 1 C=O (COOAll)), 155.6 (C=O(Fmoc)), 155.1, 155.1 (Tyr-*y*; Tyr-*p*), 143.9, 143.8 (C-1_a-; C-8_a-Fmoc), 141.4, 141.4 (C-4_a-; C-5_a-Fmoc), 131.4 (C-2-OAll), 130.8, 130.8 (Tyr-$o_{a,b}$), 127.9, 127.9 (C-3-; C-6-Fmoc), 127.1, 127.1 (C-2-; C-7-Fmoc), 125.1, 125.0 (C-1-; C-8-Fmoc), 120.1, 120.1 (C-4-; C-5-Fmoc), 119.3 (C-3-OAll), 116.5, 116.5 (Tyr-$m_{a,b}$), 95.8 (C-1), 71.2 (C-4), 68.6 (C-5), 68.0 (C-3), 67.0 (CH_2-Fmoc), 66.2 (C-1-OAll), 61.7 (C-6), 54.9 (Tyr$^{\alpha}$), 52.0 (C-2), 47.3 (C-9-Fmoc), 37.5 (Tyr$^{\beta}$), 23.2 (Me(NHAc)), 20.8, 20.7, 20.7 (Me(OAc)).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-[2-acetamido-2-deoxy-3,4,6-_O_-acetyl-α-D-glucopyranosyl]-_L_-tyrosine (69)



To a solution of compound **69** (918 mg, 1.19 mmol, 1.00 eq.) in dry tetrahydrofuran (4.5 mL), phenylsilane (300 µL) was added. Next a solution tetrakis(triphenylphosphine)palladium(0) (69 mg, 0.06 mmol, 0.05 eq.) in another tetrahydrofuran (4.5 mL) was added to the reaction mixture and it was stirred at room temperature. After 30 min TLC indicated full conversion. The solvent was removed in vacuo and the residue was purified over a C18-cartridge (H_2O/MeCN 40:60 → 60:40). After removal of the solvent the residue was further purified by column chromatography on silica (DCM/MeOH 1:0 → 1:1) and subsequent C18-cartridge (MeCN in H_2O, 40% → 60%) to give the carboxylic acid **69** as a colorless solid (686 mg, 0.94 mmol, 79 %).

R_f = 0.85 (EtOAc/MeOH/AcOH/H_2O 60:3:3:2). $[\alpha]_D^{20}$ = +87.90 (c = 8.40, $CHCl_3$). *HR-ESI-MS (pos), m/z:* 733.2592 ($[M+H]^+$, calc.: 733.2603), 755.2408 ($[M+Na]^+$, calc.: 755.2423), 750.2859 ($[M+NH_4]^+$, calc.: 750.2868.

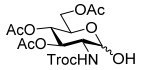
^{1}H-NMR (600 MHz, $CDCl_3$) , *δ (ppm)*: 7.76 (d, $J_{H4,H3} = J_{H5,H6}$ = 7.7 Hz, 2H, H-4-; H-5-Fmoc), 7.59 – 7.51 (m, 2H, H-1-; H-8-Fmoc), 7.41 – 7.36 (m, 2H, H-3-; H-6-Fmoc), 7.32 – 7.27 (m, 2H, H-2-; H-7-Fmoc), 7.07 – 6.89 (m, 4H, Tyr-$o_{a,b}$; Tyr-$m_{a,b}$), 6.17 (d, $J_{NH,H2}$ = 9.3 Hz, 1H, N-H), 5.50 (d, $J_{H1,H2}$ = 1.8 Hz, 1H, H-1), 5.46 – 5.37 (m, 1H, H-4), 5.25 – 5.18 (m, 1H, H-3), 4.66 – 4.62 (m, 1H, Tyr$^{\alpha}$), 4.55 – 4.45 (m, 2H, H-2; CH_{2a}-Fmoc), 4.39 – 4.33 (m, 1H, CH_{2b}-Fmoc), 4.22 – 4.15 (m, 2H, H-9-Fmoc; H-6_a), 4.05 – 3.92 (m, 2H, H-6_b; H-5), 3.17 – 3.02 (m, 2H, Tyr$^{\beta}$), 2.12 – 1.88 (m, 12H, Ac-H).

^{13}C-NMR (150.9 MHz), *δ (ppm)*: 174.0, 171.7, 171.3, 170.9, 169.5 (4 C=O(Ac); 1 C=O(COOH)), 155.8, 155.8 (Tyr-*y*; Tyr-*p*), 143.9, 143.8 (C-1_a-; C-8_a-Fmoc), 141.4, 141.4 (C-4_a-; C-5_a-Fmoc), 130.9, 130.9 (Tyr-

$m_{a,b}$), 127.9, 127.9 (C-3-; C-6-Fmoc), 127.1, 127.1 (C-2-; C-7-Fmoc), 125.2, 125.1 (C-1-; C-8-Fmoc), 120.1, 120.1 (C-4-; C-5-Fmoc), 116.4, 116.4 (Tyr-$o_{a,b}$), 95.5 (C-1), 71.1 (C-4), 68.5 (C-5), 68.0 (C-3), 67.0 (CH_2-Fmoc), 61.8 (C-6), 54.7 (Tyr^{α}), 52.1 (C-2), 47.3 (C-9-Fmoc), 37.1 (Tyr^{β}), 23.1, 20.9, 20.8, 20.7 (Me(Ac)).

9.1.3.4 Synthesis of the β-GlcNAc-Tyr amino acid

3,4,6-Tri-*O*-acetyl-2-deoxy-2-*N*-(2,2,2-trichloroethoxycarbonyl)-α/β-D-glucose (70)[338]

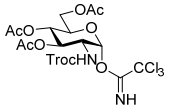


After dissolving acetate **25** (15.0 g, 28.7 mmol, 1.0 eq.) in dry tetrahydrofuran (75 mL), hydrazine acetate (3.4 g, 37.3 mmol, 1.3 eq.) was added. The reaction flask was equipped with a cooler and flushed with argon before heating the system to 50° C. After stirring for 4 h, TLC indicated full conversion. The solution was diluted with ethyl acetate (150 mL) and washed twice with sat. sodium bicarbonate, once with brine and dried over sodium sulfate. After removal of the solvent in vacuo the residue was purified by column chromatography on silica (CH/EtOAc 2:1 → 3:2) to give compound **70** (9.0 g, 18.7 mmol, 65 %).

R_f = 0.55 (CH/EtOAc 1:1).

^{1}H-NMR (700 MHz, DMSO) , *δ (ppm)*: 7.19 (d, 1H, N-H-Troc), 7.22 – 7.17 (m, 1H, O-H-1), 5.19 (dd, 1H, H-3), 5.07 (dd, 2H, H-1), 7.22 – 7.17 (m, 3H, H-4; $CH_{2a,b}$-Troc), 4.16 – 4.11 (m, 2H, H-5; H-6_a), 3.99 (d, 1H, H-6_b), 3.80 – 3.74 (m, 1H, H-2), 2.03 – 1.84 (m, 9H, 3 Me(Ac)).

3,4,6-Tri-*O*-acetyl-2-deoxy-2-*N*-(2,2,2-trichloroethoxycarbonyl)- α-D-glucopyranosyl trichloroacetimidate (71)[398]

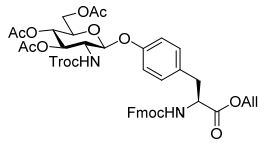


A solution of compound **90** (8.9 g, 18.5 mmol, 1.0 eq.) in dichloromethane (90.0 mL) was cooled on ice. Afterwards, first trichloroacetonitrile (2.4 mL, 24.1 mmol, 1.3 eq.) and then DBU (550 µl, 3.7 mmol, 0.2 eq.) were added in that order. After stirring for 75 min TLC indicated full conversion. The solution was filtrated over a silica pad and washed with dichloromethane to obtain trichloroacetimidate **71** (7.2 g, 11.5 mmol, 62 %).

R_f = 0.71 (CH/EtOAc 1:1).

^{1}H-NMR (700 MHz, DMSO) , *δ (ppm)*: 8.24 (d, $J_{NH,H2}$ = 7.3 Hz, 1H, N-H-Troc), 8.27 – 8.22 (m, 1H, H-1), 5.27 – 5.22 (m, 1H, H-3), 5.06 – 5.02 (m, 1H, H-4), 8.27 – 8.22 (m, 2H, $CH_{2a,b}$-Troc), 4.20 (dd, *J* = 12.4, 4.8 Hz, 1H, H-6_a), 4.15 – 4.09 (m, 2H, H-2; H-5), 4.05 (dd, *J* = 12.5, 2.4 Hz, 1H, H-6_b), 2.04 – 1.95 (m, 9H, 3 Me(Ac)).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-[3,4,6-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-β-D-glucopyranosyl]-_L_-tyrosine-allylester (72)



A solution of Fmoc-tyrosine allyl ester **62** (3.0 g, 6.76 mmol, 1.0 eq.), trichloroacetimidate **71** (5.5 g, 8.8 mmol, 1.3 eq.) and molecular sieve (4.2 g, 4 Å) in dichloromethane (85.0 mL) was stirred for 25 min at room temperature under argon. Afterwards an ice bath was installed and TMSOTf (245 µL, 2.0 mmol, 0.2 eq.) were added dropwise. Stirring was continued overnight, while the temperature was allowed to reach room temperature. Finally, the solution was washed twice with sat. sodium bicarbonate solution, once with brine and dried over sodium sulfate. The solvent was removed in vacuo and the residue was purified by column chromatography on silica (CH/EtOAc 3:1 → 2:1) to give of glycoside **72** (2.8 g, 3.1 mmol, 46 %).

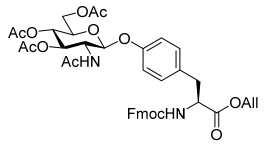
R_f = 0.33 (CH/EtOAc 2:1). $[\alpha]_D^{20}$ = +-4.64 (c = 8.70, $CHCl_3$). *HR-ESI-MS (pos), m/z:* 929.1634 ($[M+Na]^+$, calc.: 929.1643), 922.2107 ($[M+NH_4]^+$, calc.: 922.2073), 905.1840 ($[M+H]^+$, calc.: 905.1853.

^{1}H-NMR (600 MHz, $CDCl_3$), *δ (ppm)*: 7.83 – 7.77 (m, 2H, H-4-; H-5-Fmoc), 7.60 – 7.54 (m, 2H, H-1-; H-8-Fmoc), 7.48 – 7.38 (m, 2H, H-3-; H-6-Fmoc), 7.35 – 7.30 (m, 2H, H-2-; H-7-Fmoc), 7.08 – 6.86 (m, 4H, Tyr-$o_{a,b}$; Tyr-$m_{a,b}$), 5.95 – 5.81 (m, 1H, H-2-OAll), 5.41 – 5.24 (m, 3H, H-$3_{a,b}$-OAll; N-H-Fmoc), 5.24 – 5.15 (m, 1H, H-4), 5.10 (d, $J_{NH,H2}$ = 9.2 Hz, 1H, N-H), 5.07 – 5.02 (m, 1H, H-3), 4.79 – 4.67 (m, 4H, H-1; Tyr$^\alpha$; CH_2-Troc), 4.63 (d, $J_{H1OAll,H2OAll}$ = 6.0 Hz, 2H, H-1-OAll), 4.54 – 4.46 (m, 1H, CH_{2a}-Fmoc), 4.23 – 4.13 (m, 2H, H-6_a; CH_{2b}-Fmoc), 4.12 – 4.00 (m, 2H, H-9-Fmoc; H-6_b), 3.89 – 3.81 (m, 1H, H-2), 3.52 – 3.42 (m, 1H, H-5), 3.21 – 2.95 (m, 2H, Tyr$^\beta$), 2.07 – 2.02 (m, 9H, Me(Ac)).

^{13}C-NMR (150.9 MHz, $CDCl_3$), *δ (ppm)*: 171.5, 171.3, 170.7, 169.5 (3 C=O(Ac); 1 C=O(OAll), 156.3, 156.3 (Tyr-*y*; Tyr-*p*), 155.6 (C=O(Fmoc)), 154.2 (C=O(Troc)), 144.1, 143.6 (C-1_a-; C-8_a-Fmoc), 141.3, 141.3 (C-

4_a-; C-5_a-Fmoc), 131.4 (C-2-OAll), 130.6, 130.6 (Tyr-$o_{a,b}$), 128.0, 128.0 (C-3-; C-6-Fmoc), 127.3, 127.3 (C-2-; C-7-Fmoc), 125.6, 125.2 (C-1-; C-8-Fmoc), 120.3, 120.3 (C-4-; C-5-Fmoc), 117.8, 117.8 (Tyr-$m_{a,b}$), 99.8 (C-1), 95.6 (CCl_3-Troc), 74.6 (CH_2-Troc), 72.0, 71.8 (C-3-OAll; C-4), 68.5 (C-3), 67.1 (CH_2-Fmoc), 66.3 (C-1-OAll), 62.0 (C-6), 56.1 (C-2), 54.7 (Tyr$^{\alpha}$), 47.3 (C-9-Fmoc), 37.8 (Tyr$^{\beta}$), 20.8, 20.8, 20.7 (CH_3-Ac).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-[2-acetamido-3,4,6-_O_-acetyl-2-deoxy-β-D-glucopyranosyl]-_L_-tyrosine-allylester (73)



After dissolving compound **72** (2.01 g, 2.2 mmol, 1.0 eq.) in a solution of dry acetic acid/tetrahydrofuran (1:1, 40.0 mL), zinc dust (2.90 g, 44.4 mmol, 20.0 eq.) were added while stirring. A reflux condenser was installed, the system was flushed with argon and the slurry was stirred overnight at 50° C. Afterwards the zinc was filtered off over a POR3-filter and the filtrate was concentrated in vacuo. The residue was dissolved in acetic anhydride/pyridine (1:2, 20.0 mL), 4-dimethylaminopyridine (26.9 mg, 0.2 mmol, 0.1 eq.) were added and the solution was stirred overnight at room temperature. Next, the solvent was evaporated in vacuo, followed by codistillation of the residue with toluene. The residue was purified by column chromatography on silica (CH/EtOAc 1:2) to give acetamide **73** as a colorless solid (1.67 g, 2.2 mmol, 97 %).

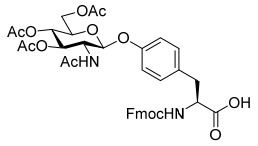
R_f = 0.45 (CH/EtOAc 1:4). $[\alpha]_D^{20}$ = +39.05 (c = 0.70, DMF). *HR-ESI-MS (pos), m/z:* 795.2716 ($[M+Na]^+$, calc.: 795.2736), 773.2902 ($[M+H]^+$, calc.: 773.2916), 790.3169 ($[M+NH_4]^+$, calc.: 790.3181.

^{1}H-NMR (600 MHz, DMSO) , *δ (ppm)*: 8.05 (d, $J_{NH,H2}$ = 6.8 Hz, 1H, N-H), 7.93 – 7.84 (m, 2H, H-4-; H-5-Fmoc), 7.68 – 7.62 (m, 2H, H-1-; H-8-Fmoc), 7.44 – 7.37 (m, 2H, H-3-; H-6-Fmoc), 7.35 – 7.27 (m, 2H, H-2-; H-7-Fmoc), 7.24 – 6.88 (m, 2H, Tyr-$m_{a,b}$; Tyr-$o_{a,b}$), 5.93 – 5.79 (m, 1H, H-2-OAll), 5.29 (d, $J_{H1,H2}$ = 12.8 Hz, 1H), 5.32 – 5.26 (m, 2H, H-1; H-3_a-OAll), 5.24 – 5.17 (m, 2H, H-4; H-3_b-OAll), 4.98 – 4.87 (m, 1H, H-3), 4.61 – 4.52 (m, 2H, H-1-OAll), 4.35 – 4.14 (m, 5H, CH_2-Fmoc; Tyr$^{\alpha}$; H-6_a; H-9-Fmoc), 4.13 – 3.94 (m, 3H, H-5; H-6_b; H-2), 3.06 – 2.81 (m, 2H, Tyr$^{\beta}$), 2.07 – 1.69 (m, 12H, 4 Me(Ac)).

^{13}C-NMR (150.9 MHz, DMSO), *δ (ppm)*: 171.5 (C=O(OAll)), 170.0, 170.0, 169.4, 169.3 (C=O(Ac)), 155.9 (C=O(Fmoc)), 155.5 (Tyr-*y*; Tyr-*p*), 143.7, 143.7 (C-1_a-Fmoc; C-8_a-Fmoc), 140.7, 140.7 (C-4_a-; C-5_a-Fmoc), 132.3 (C-2-OAll), 130.2, 130.2 (Tyr-$o_{a,b}$), 127.6, 127.6 (C-3-; C-6-Fmoc), 127.1, 127.1 (C-2-; C-7-Fmoc), 125.2, 125.2 (C-1-; C-8-Fmoc), 120.1, 120.1 (C-4-; C-5-Fmoc), 117.8 (C-3-OAll), 116.3 (Tyr-$m_{a,b}$),

97.9 (C-1), 72.4 (C-4), 70.8 (C-5), 68.4 (C-3), 65.7 (CH_2-Fmoc), 64.9 (C-1-OAll), 61.7 (C-6), 55.8 (Tyr^{α}), 53.2 (C-2), 46.6 (C-9-Fmoc), 35.6 (Tyr^{β}), 22.7, 20.5, 20.4, 20.4 (4 CH_3-Ac).

_N_-9-Fluorenylmethyloxycarbonyl-_O_-[2-acetamido-3,4,6-_O_-acetyl-2-deoxy-β-D-glucopyranosyl]-_L_-tyrosine (74)



A dry flask was flushed with argon and filled with dry tetrahydrofuran/dimethylformamide (1:1, 4.0 mL). Afterwards allyl ester **73** (527 mg, 682 µmol, 1.0 eq.) and phenylsilane (172 µL, 1.398 mmol, 2.05 eq.) were added. Finally tetrakis(triphenylphosphine)palladium(0) (39 mg, 34 µmol, 0.05 eq.) were dissolved in dry tetrahydrofuran (2.0 mL) and added dropwise. The solution was stirred for 1 h at room temperature, before the solvent was removed in vacuo. The residue was purified by column chromatography on silica (EtOAc/MeOH 10:1 → 1:1) and subsequent C18-cartridge (MeCN in H_2O, 40% → 60%) to give the carboxylic acid **74** as a colorless solid (404 mg, 551 µmol, 81 %).

R_f = 0.81 (EtOAc/MeOH/H_2O/AcOH 60:3:3:2). $[\alpha]_D^{20}$ = -3.94 (c = 1.10, DMF). *HR-ESI-MS (pos), m/z:* 755.2405 ($[M+Na]^+$, calc.: 755.2423), 733.2590 ($[M+H]^+$, calc.: 733.2603), 750.2856 ($[M+NH_4]^+$, calc.: 750.2868.

^{1}H-NMR (600 MHz, DMSO) , *δ (ppm)*: 8.05 (d, $J_{NH,H2}$ = 9.1 Hz, 1H, N-H), 7.88 (d, $J_{H4,H3}$ = $J_{H5,H4}$ = 7.5 Hz, 2H, H-4-; H-5-Fmoc), 7.73 (d, $J_{NH,CH2\text{-}Fmoc}$ = 8.4 Hz, 2H, CH_2-Fmoc), 7.68 – 7.64 (m, 2H, H-3-; H-6-Fmoc), 7.44 – 7.38 (m, 2H, H-2-; H-7-Fmoc), 7.22 (d, $J_{Yo,Ym}$ = 8.5 Hz, 2H, Tyr-$o_{a,b}$), 6.92 (d, $J_{Ym,Yo}$ = 8.6 Hz, 2H, Tyr-$m_{a,b}$), 5.28 (d, $J_{H1,H2}$ = 8.5 Hz, 1H, H-1), 5.20 (t, *J* = 9.9 Hz, 1H, H-3), 4.91 (t, *J* = 9.8 Hz, 1H, H-4), 4.25 – 4.16 (m, 4H, H-$6_{a,b}$; CH_{2a}-Fmoc; H-9-Fmoc), 4.14 – 4.09 (m, 1H, Tyr^{α}), 4.09 – 4.04 (m, 1H, H-5), 4.04 – 3.95 (m, 1H, H-2), 3.05 – 2.79 (m, 2H, Tyr^{β}), 2.02 – 1.92 (m, 12H, Me(Ac)).

^{13}C-NMR (150.9 MHz, DMSO), *δ (ppm)*: 173.4, 169.7, 169.5, 169.3 (4 C=O(Ac); 1 COOH), 156.0 (C=O(Fmoc)), 155.5 (Tyr-γ; Tyr-p), 143.7 (C-1_a-; C-8_a-Fmoc), 140.7 (C-4_a-; C-5_a-Fmoc), 130.2 (Tyr-$o_{a,b}$), 127.6 (C-3-; C-6-Fmoc), 127.1 (C-2-; C-7-Fmoc), 125.3, 125.3 (C-1-; C-8-Fmoc), 120.1 (C-4-; C-5-Fmoc), 116.2 (Tyr-$m_{a,b}$), 97.8 (C-1), 72.5 (C-3), 70.8 (C-5), 68.4 (C-4), 65.6 (C-6), 61.6 (CH_2-Fmoc), 55.7 (Tyr^{α}), 53.2 (C-2), 46.7 (C-9-Fmoc), 35.6 (Tyr^{β}), 20.5, 20.4, 20.4 (4 Me(Ac)).

9.2 Solid phase peptide synthesis

9.2.1 General protocol for solid phase (glyco-)peptide synthesis

General

Solid phase peptide synthesis of peptides, glycopeptides and phosphopeptides was done at a *Syro* I peptide synthesizer (*Multisynth* GmbH). All used resins were purchased from *Rapp Polymere* GmbH and the protected standard amino acids were purchased from *Merck* KGaA and Merck Suchhardt OHG. Benzyl protected phosphotyrosine was purchased from *Merck KGaA*. Solvents were purchased from *Biosolve Chimie SARL*. The coupling reagents HBTU and HATU (*Novabiochem®*) were purchased from *Merck* KGaA. HOBt monohydrate was purchased from *Sigma-Aldrich* GmbH and was recrystallized prior to utilization from absolute methanol and dried in vacuo. HOAT was purchased from *GL Biochem*, Shanghai. The N-terminal tri-ethylene glycol spacer was kindly provided by Dr. *Christian Pett* and the C-terminal spacer was kindly provided by Dr. *Yu Jin*.

Automated peptide synthesis

The peptide synthesis was done automatically using a peptide synthesizer according to a standard Fmoc protocol. This excludes the coupling of glycosylated amino acids, which was carried out manually. The N-terminally Fmoc-protected amino acids were prepared as 0.5 M solutions in DMF and loaded to corresponding reservoir bottles of the peptide synthesizer. HBTU and HOBt were prepared each as 0.45 M solutions in DMF, a 2 M DIPEA/NMP solution and a 20 vol% piperidine/DMF solution were prepared and stored separately in dedicated bottles. The preloaded *TentaGel*-Fmoc-aa-Trt resins (batch size depending on individual peptide) were filled into 2 ml synthesis reactors equipped with a terminal filter frit and swelled with dichloromethane (600 µL) each for 30 min. The solvent was switched to dimethylformamide by multiple washing steps with this solvent (5 x 500 µL, 15 s vortex, 45 s break) and the Fmoc protection group was removed by triple addition of 20 vol% piperidine/DMF (500 µL each, 2 x 3 min, 1 x 9 min, 15 s vortex, 45 s break). Afterwards, the resin was washed with DMF (5 x 500 µL, 15 s vortex, 45 s break) and the automated cycle of amino acid coupling, washing, Fmoc deprotection and washing started. For amino acid coupling 8 eqiv. Fmoc-aa-OH, 7.9 eqiv. HBTU and 8 equiv. HOBt were mixed in the synthesis reactor and shaken 40 min (15 s vortex, 2:45 min break). DMF washing, Fmoc cleavage and final DMF washing proceeded as described above.

Manual coupling of glycosylated amino acids

The glycosylated amino acids were dissolved in DMF (20 µL per µmol batch size) and activated with HOAT, HATU and DIPEA. Thereby 1.5 to 2.0 equiv. (in relation to resin loading) of the glycosylated amino acids were used (compare ***Table 1***, p. 80 and ***Table 2*** p. **Fehler! Textmarke nicht definiert.**). The activating agents, HOAT and HATU, were added in slightly substochiometric amounts with respect to the amino acid (1.45-1.95 equiv.) followed by addition of DIPEA (3-4 equiv.). Afterwards this solution was quickly added to the resin in the syringe reactor and shaken for 8 hours (15 s vortex, 2:45 min break) at room temperature. Finally, the resin was washed with DMF (5 x 500 µL, 1 min each, 15 s vortex, 45 s break). If the glycosylated amino acid did not represent the N-terminal amino acid, the peptide synthesis was continued according to the automated protocol. If the glycosylated amino acid represented the *N*-terminal amino acid, the Fmoc protection group was cleaved with 20 vol% piperidine/DMF (500 µL each, 2 x 3 min, 1 x 9 min, 15 s vortex, 45 s break).

Coupling of N-terminal TEG spacers

The coupling of *N*-terminal spacers was performed manually on-resin after complete synthesis of the peptide. A 3-fold excess (with respect to the resin batch size) of the Fmoc protected spacer was dissolved in dimethylformamide (300 µL) and activated with 3 equiv. HOBt, 3 equiv. HBTU and 6 equiv DIPEA. The solution was quickly added to the resin and the system was shaken for 40 min (15 s vortex, 2:45 break). Afterwards, the resin was washed with DMF (500 µL, 5 x 1 min, 15 s vortex, 45 s break) and Fmoc was cleaved off with 20 vol% piperidine/DMF (500 µL each, 2 x 3 min, 1 x 9 min, 15 s vortex, 45 s break).

Release of the peptides from the resin

After completion of the peptide synthesis, the resin was washed with DMF (500 µL, 5 x 1 min, 15 s vortex, 45 s break) and then DCM, *i*PrOH and diethyl ether (each 700 µL, 5 x 1 min, 15 s vortex, 45 s break) in that order. The resin was dried for 30 min in an airstream and then transferred to a 2 mL syringe filter frit. The cleavage was performed by addition of TFA/TIPS/H_2O (1 x 120 min, 2 x 10 min, 1 mL each) and shaking (15 s vortex, 2:45 min break). The filtrates were combined and the solvent was removed by codistillation with toluene *in vacuo*.

Removal of glycan acetyl groups

The glycopeptide was desalted via C18-cartridge (*Waters SepPak® Vac 6cc* (1g)). Therefore, the residue was washed with buffer A (5 x 2 mL) prior to elution of the glycopeptide with a gradient of 30 %, 50 % and 70 % buffer B in this order (10 mL each). The fractions, which contained buffer B were combined and the solvent was removed in vacuo. The residue was then applied to deacetylation. Glycopeptides containing core 4 glycans or sialic acid were dissolved in methanol (5 mL). Then the *p*H was adjusted to 9-10 using portions of 1 % sodium methanolate in methanol. Glycopeptides containing core 2 glycans or HexNAc-Tyr were dissolved in methanol (4.5 mL) and water (3.9 mL), which was then adjusted to *p*H 10-11 by stepwise addition of aqueous sodium hydroxide (200 mM). The system was stirred at room temperature and the deacetylation process was monitored via HPLC (16 – 144 h). Finally, the system was acidified using acetic acid (20 µL) and the solvent was removed *in vacuo*.

Purification of peptides and glycopeptides

The peptides and deacetylated glycopeptides were purified via preparative HPLC. Afterwards MUC1 peptides were freeze-dried from water and other peptides were freeze-dried from 20 vol% acetonitrile in water.

9.2.2 Synthesis of mucin-type glycosylated peptides

All mucin-type glycosylated peptides were synthesized as MUC1 VNTR sequences. As resin the preloaded *TentaGel R Fmoc-Ala-Trt* resin (loading capacity: 0.17 mmol/g, *Rapp Polymere*, Tübingen) was utilized. The HPLC eluents were composed of gradients of buffer A (water + 0.1 % TFA) and buffer B (84 % acetonitrile + 0.1 % TFA) or C (water + 0.1 % FA) and D (84 % acetonitrile + 0.1 % FA). The chromatogram was recorded at λ = 214 nm (peptide bonds).

MUC1 unglycosylated: TEG-PAHGVTSAPDTRPAPGSTA (P1)

Resin amount: 13 µmol; Analytical HPLC R_f = 21.58 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 17.51 min (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 655.3387 ($[M+3H]^{3+}$, calc.: 665.3388); *Yield*: 40 % (10.5 mg, 5.2 µmol).

MUC1 T_N-antigen: TEG-PAHGVT*SAPDTRPAPGSTA (P2)

Resin amount: 13 µmol; Analytical HPLC R_f = 22.88 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 15.30 min (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 733.0326 ($[M+3H]^{3+}$, calc.: 733.0139); *Yield*: 27 % (7.6 mg, 3.5 µmol).

MUC1 T_N-antigen: TEG-PAHGVTSAPDT*RPAPGSTA (P3)

Resin amount: 13 µmol; Analytical HPLC R_f = 22.25 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 15.87 min (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 733.0320 ($[M+3H]^{3+}$, calc.: 733.0139); *Yield*: 22 % (6.3 mg, 2.9 µmol).

MUC1 T_N-antigen: TEG-PAHGVTSAPDTRPAPGST*A (P4)

Resin amount: 13 µmol; Analytical HPLC R_f = 14.93 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 16.25 min (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 733.0319 ($[M+3H]^{3+}$, calc.: 733.0139); *Yield*: 14 % (4.0 mg, 1.8 µmol).

MUC1 GalNAc-D_3: PAHGVTSAPDT*RPAPGSTA (P5)

Resin amount: 13 µmol; Analytical HPLC R_f = 22.13 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 13.74 min (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 666.3335 ($[M+3H]^{3+}$, calc.: 666.3330); *Yield*: 51 % (13.2 mg, 6.6 µmol).

MUC1 GalNAc-D_3: PAHGVTSAPDTRPAPGST*A (P6)

Resin amount: 13 µmol; Analytical HPLC R_f = 22.49 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 14.02 min (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 665.9991 ($[M+3H]^{3+}$, calc.: 665.9985); *Yield*: 52 % (13.4 mg, 6.7 µmol).

MUC1 core 2 type-1: TEG-PAHGVT*SAPDTRPAPGSTA (P7)

Resin amount: 10 µmol; Analytical HPLC R_f = 22.49 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 18.70 min (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 908.7606 ($[M+3H]^{3+}$, calc.: 908.7603); *Yield*: 78 % (21.3 mg, 7.8 µmol).

MUC1 core 2 type-1: TEG-PAHGVTSAPDT*RPAPGST*A (P8)

Resin amount: 13 µmol; Analytical HPLC R_f = 22.20 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 14.59 min (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 908.7622 ($[M+3H]^{3+}$, calc.: 908.7603); *Yield*: 25 % (8.7 mg, 3.2 µmol).

MUC1 core 2 type-1: TEG-PAHGVTSAPDTRPAPGST*A (P9)

Resin amount: 13 µmol; Analytical HPLC R_f = 22.79 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 15.24 min (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 908.7618 ($[M+3H]^{3+}$, calc.: 908.7603); *Yield*: 30 % (10.5 mg, 3.9 µmol).

MUC1 core 2 type-1: TEG-PAHGVT*SAPDT*RPAPGSTA (P10)

Resin amount: 10 µmol; Analytical HPLC R_f = 16.39 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 17.23 min (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1152.1834 ($[M+3H]^{3+}$, calc.: 1152.1817), 864.3903 ($[M+4H]^{4+}$, calc.: 864.3881); *Yield*: 22 % (7.6 mg, 2.2 µmol).

MUC1 core 2 type-1: TEG-PAHGVT*SAPDTRPAPGST*A (P11)

Resin amount: 10 µmol; Analytical HPLC R_f = 17.44 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 22.25 min (A/B, (95/5) 5 min isocratic; (95/5) → (70/30) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1152.1834 ($[M+3H]^{3+}$, calc.: 1152.1817), 864.3899 ($[M+4H]^{4+}$, calc.: 864.3881); *Yield*: 28 % (9.6 mg, 2.8 µmol).

MUC1 core 2 type-1: TEG-PAHGVTSAPDT*RPAPGST*A (P12)

Resin amount: 10 µmol; Analytical HPLC R_f = 17.93 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 22.91 min (A/B, (95/5) 5 min isocratic; (95/5) →

(70/30) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1152.1838 ($[M+3H]^{3+}$, calc.: 1152.1817), 864.3901 ($[M+4H]^{4+}$, calc.: 864.3881); *Yield*: 30 % (10.3 mg, 3.0 µmol).

MUC1 core 2 type-1: TEG-PAHGVT*SAPDT*RPAPGST*A (P13)

Resin amount: 10 µmol; Analytical HPLC R_f = 16.50 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 20.10 min (A/B, (95/5) 5 min isocratic; (95/5) → (70/30) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1395.9409 ($[M+3H]^{3+}$, calc.: 1395.9376), 1047.2059 ($[M+4H]^{4+}$, calc.: 1047.2051); *Yield*: 17 % (7.0 mg, 1.7 µmol).

MUC1 core 2 type-2: TEG-PAHGVT*SAPDTRPAPGSTA (P14)

Resin amount: 13 µmol; Analytical HPLC R_f = 20.14 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 13.77 min (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 908.7631 ($[M+3H]^{3+}$, calc.: 908.7603); *Yield*: 38 % (13.4 mg, 4.9 µmol).

MUC1 core 2 type-2: TEG-PAHGVTSAPDT*RPAPGSTA (P15)

Resin amount: 13 µmol; Analytical HPLC R_f = 20.84 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 14.13 min (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 908.7622 ($[M+3H]^{3+}$, calc.: 908.7603); *Yield*: 39 % (13.8 mg, 5.1 µmol).

MUC1 core 2 type-2: TEG-PAHGVTSAPDTRPAPGST*A (P16)

Resin amount: 13 µmol; Analytical HPLC R_f = 21.74 min (A/B, (95:5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 20.32 min (A/B, (95:5) 5 min isocratic; (95/5) → (67/33) in 43 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 908.7625 ($[M+3H]^{3+}$, calc.: 908.7603); *Yield*: 35 % (12.2 mg, 4.5 µmol).

MUC1 core 2 type-2: TEG-PAHGVT*SAPDT*RPAPGSTA (P17)

Resin amount: 10 µmol; Analytical HPLC R_f = 15.06 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 24.87 min (A/B, (95/5) 5 min isocratic; (95/5) →

(75/25) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1152.1838 ($[M+3H]^{3+}$, calc.: 1152.1817), 864.3898 ($[M+4H]^{4+}$, calc.: 864.3881); *Yield*: 28 % (12.3 mg, 3.6 µmol).

MUC1 core 2 type-2: TEG-PAHGVT*SAPDTRPAPGST*A (P18)

Resin amount: 10 µmol; Analytical HPLC R_f = 16.29 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 26.43 min (A/B, (95/5) 5 min isocratic; (95/5) → (75/25) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1152.1836 ($[M+3H]^{3+}$, calc.: 1152.1817), 864.3901 ($[M+4H]^{4+}$, calc.: 864.3881); *Yield*: 32 % (11.1 mg, 3.2 µmol).

MUC1 core 2 type-2: TEG-PAHGVTSAPDT*RPAPGST*A (P19)

Resin amount: 10 µmol; Analytical HPLC R_f = 23.45 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 24.97 min (A/B, (95/5) 5 min isocratic; (95/5) → (75/25) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1152.1835 ($[M+3H]^{3+}$, calc.: 1152.1817), 864.3898 ($[M+4H]^{4+}$, calc.: 864.3881); *Yield*: 15 % (5.1 mg, 1.5 µmol).

MUC1 core 2 type-2: TEG-PAHGVT*SAPDT*RPAPGST*A (P20)

Resin amount: 10 µmol; Analytical HPLC R_f = 16.29 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 19.79 min (A/B, (95/5) 5 min isocratic; (95/5) → (75/25) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1047.2064 ($[M+4H]^{4+}$, calc.: 1047.2051), 1395.9412 ($[M+3H]^{3+}$, calc.: 1395.9412); *Yield*: 14 % (6.0 mg, 1.4 µmol).

MUC1 core 4 type-1: TEG-PAHGVT*SAPDTRPAPGSTA (P21)

Resin amount: 10 µmol; Analytical HPLC R_f = 15.57 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 27.53 min (A/B, (95/5) 5 min isocratic; (95/5) → (75/25) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 976.4544 ($[M+3H]^{3+}$, calc.: 976.4534); *Yield*: 42 % (12.4 mg, 4.2 µmol).

MUC1 core 4 type-1: TEG-PAHGVTSAPDT*RPAPGSTA (P22)

Resin amount: 13 µmol; Analytical HPLC R_f = 19.82 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 14.16 min (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 976.4546 ($[M+3H]^{3+}$, calc.: 976.4534), 732.5932 ($[M+4H]^{4+}$, calc.: 732.5919); *Yield*: 17 % (6.5 mg, 2.2 µmol).

MUC1 core 4 type-1: TEG-PAHGVTSAPDTRPAPGST*A (P23)

Resin amount: 13 µmol; Analytical HPLC R_f = 20.78 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 15.53 min (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 976.4557 ($[M+3H]^{3+}$, calc.: 976.4534), 732.5929 ($[M+4H]^{4+}$, calc.: 732.5919); *Yield*: 18 % (6.8 mg, 2.3 µmol).

MUC1 core 4 type-1: TEG-PAHGVT*SAPDT*RPAPGSTA (P24)

Resin amount: 10 µmol; Analytical HPLC R_f = 13.14 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 23.06 min (A/B, (95/5)5 min isocratic (95/5) → (75/25) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1287.9059 ($[M+3H]^{3+}$, calc.: 1287.9024), 966.1803 ($[M+4H]^{4+}$, calc.: 966.1786); *Yield*: 25 % (9.5 mg, 2.5 µmol).

MUC1 core 4 type-1: TEG-PAHGVT*SAPDTRPAPGST*A (P25)

Resin amount: 10 µmol; Analytical HPLC R_f = 13.74 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 15.84 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1287.9057 ($[M+3H]^{3+}$, calc.: 1287.9024), 966.1804 ($[M+4H]^{4+}$, calc.: 966.1786); *Yield*: 30 % (11.5 mg, 3.0 µmol).

MUC1 core 4 type-1: TEG-PAHGVTSAPDT*RPAPGST*A (P26)

Resin amount: 9 µmol; Analytical HPLC R_f = 14.96 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 24.70 min (A/B, (95/5) 5 min isocratic; (95/5) → (75/25) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1287.9061 ($[M+3H]^{3+}$, calc.: 1287.9024), 966.1803 ($[M+4H]^{4+}$, calc.: 966.1786); *Yield*: 26 % (8.7 mg, 2.3 µmol).

MUC1 core 4 type-1: TEG-PAHGVT*SAPDT*RPAPGST*A (P27)

Resin amount: 9 µmol; Analytical HPLC R_f = 13.05 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 20.94 min (A/B, (95/5) 5 min isocratic; (95/5) → (75/25) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1199.5175 ($[M+4H]^{4+}$, calc.: 1199.5146), 1595.0222 ($[M+3H]^{3+}$, calc.: 1595.0170), 959.8154 ($[M+5H]^{5+}$, calc.: 959.8131); *Yield*: 19 % (8.1 mg, 1.7 µmol).

MUC1 core 4 type-2: TEG-PAHGVT*SAPDTRPAPGSTA (P28)

Resin amount: 10 µmol; Analytical HPLC R_f = 14.78 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 28.13 min (A/B, (95/5) 5 min isocratic; (95/5) → (75/25) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 976.4541 ($[M+3H]^{3+}$, calc.: 976.4534), 732.5945 ($[M+4H]^{4+}$, calc.: 732.5919); *Yield*: 51 % (14.8 mg, 5.1 µmol).

MUC1 core 4 type-2: TEG-PAHGVTSAPDT*RPAPGSTA (P29)

Resin amount: 13 µmol; Analytical HPLC R_f = 20.03 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 14.47 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 976.4543 ($[M+3H]^{3+}$, calc.: 976.4534), 732.5923 ($[M+4H]^{4+}$, calc.: 732.5919); *Yield*: 8 % (3.3 mg, 1.1 µmol).

MUC1 core 4 type-2: TEG-PAHGVTSAPDTRPAPGST*A (P30)

Resin amount: 13 µmol; Analytical HPLC R_f = 21.77 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 19.86 min (A/B, (95/5) 5 min isocratic; (95/5) → (82/18) in 13 min; (88/18) → (72/28) in 100 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 732.5929 ($[M+4H]^{4+}$, calc.: 732.5919), 976.4559 ($[M+3H]^{3+}$, calc.: 976.4534); *Yield*: 12 % (4.4 mg, 1.5 µmol).

MUC1 core 4 type-2: TEG-PAHGVT*SAPDT*RPAPGSTA (P31)

Resin amount: 10 µmol; Analytical HPLC R_f = 15.65 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 24.97 min (A/B, (95/5) 5 min isocratic; (95/5) → (75/25) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1287.9057 ($[M+3H]^{3+}$, calc.: 1287.9024), 966.1801 ($[M+4H]^{4+}$, calc.: 966.1786); *Yield*: 28 % (10.7 mg, 2.8 µmol).

MUC1 core 4 type-2: TEG-PAHGVT*SAPDTRPAPGST*A (P32)

Resin amount: 10 µmol; Analytical HPLC R_f = 15.37 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 26.45 min (A/B, (95/5) 5 min isocratic; (95/5) → (75/25) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1287.9054 ($[M+3H]^{3+}$, calc.: 1287.9024), 966.1800 ($[M+4H]^{4+}$, calc.: 966.1786); *Yield*: 29 % (11.3 mg, 2.9 µmol).

MUC1 core 4 type-2: TEG-PAHGVTSAPDT*RPAPGST*A (P33)

Resin amount: 10 µmol; Analytical HPLC R_f = 23.15 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 27.56 min (A/B, (95/5) 5 min isocratic; (95/5) → (75/25) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 966.1801 ($[M+4H]^{4+}$, calc.: 966.1786), 1287.9059 ($[M+3H]^{3+}$, calc.: 1287.9024); *Yield*: 30 % (11.7 mg, 3.0 µmol).

MUC1 core 4 type-2: TEG-PAHGVT*SAPDT*RPAPGST*A (P34)

Resin amount: 10 µmol; Analytical HPLC R_f = 13.95 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 23.52 min (A/B, (95/5) 5 min isocratic; (95/5) → (75/25) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1199.5175 ($[M+4H]^{4+}$, calc.: 1199.5146), 1595.0222 ($[M+3H]^{3+}$, calc.: 1595.0170), 959.8154 ($[M+5H]^{5+}$, calc.: 959.8131); *Yield*: 15 % (7.3 mg, 1.5 µmol).

9.2.3 Synthesis of HexNAc glycopeptides

For (glyco-)peptide synthesis for the HexNAc library the following preloaded resins have been utilized. Alanine: *TentaGel R Fmoc-Ala-Trt* resin (loading capacity: 0.17 mmol/g); Arginine*: TentaGel R Fmoc-Arg(Pbf)-Trt* resin (loading capacity: 0.16 mmol/g); Aspartate: *TentaGel R Moc-Asp(tBu)-Trt* resin (loading capacity: 0.16 mmol/g); Glycine: *TentaGel R Fmoc-Gly-Trt* resin (loading capacitiy 0.18 mmol/g); Lysine: *Tenta Gel R Fmoc-Lys(Boc)-Trt* resin (loading capacity 0.17 mmol/g). All resins were purchased from *Rapp Polymere*, Tübingen. The HPLC eluents were composed of gradients of buffer A (water + 0.1 % TFA) and buffer B (84 % acetonitrile + 0.1 % TFA) or C (water + 0.1 % FA) and D (84 % acetonitrile + 0.1 % FA). The chromatogram was recorded at λ = 214 nm (peptide bonds).

Amyloid-β (human) 672 – 687 unglycosylated: TEG-DAEFRHDSGEYEVHHQK (P35)

Resin amount: 10 µmol; Analytical HPLC R_f = 15.32 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 20.69 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 730.9989 ($[M+3H]^{3+}$, calc.: 731.0009); *Yield*: 87 % (18.7 mg, 8.7 µmol).

Amyloid-β (human) 672 – 687 αGalNAc: TEG-DAEFRHDSGEY*EVHHQK (P36)

Resin amount: 13 µmol; Analytical HPLC R_f = 23.35 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 19.45 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 473.2213 ($[M+5H]^{5+}$, calc.: 473.2213), 591.2770 ($[M+4H]^{4+}$, calc.: 591.2748); *Yield*: 53 % (16.2 mg, 6.9 µmol).

Amyloid-β (human) 672 – 687 αGlcNAc: TEG-DAEFRHDSGEY*EVHHQK (P37)

Resin amount: 10 µmol; Analytical HPLC R_f = 15.10 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 21.63 min (A/B, (95/5) 5 min isocratic; (95/5) → (70/30) in 25 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 473.2212 ($[M+5H]^{5+}$, calc.: 473.2213), 591.2750 ($[M+4H]^{4+}$, calc.: 591.2748); *Yield*: 63 % (14.8 mg, 6.3 µmol).

Amyloid-β (human) 672 – 687 βGlcNAc: TEG-DAEFRHDSGEY*EVHHQK (P38)

Resin amount: 13 µmol; Analytical HPLC R_f = 23.16 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 16.51 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 473.2212 ($[M+5H]^{5+}$, calc.: 473.2213), 591.2765 ($[M+4H]^{4+}$, calc.: 591.2748); *Yield*: 49 % (15.2 mg, 6.4 µmol).

Amyloid-β (human) 672 – 687 phospho: TEG-DAEFRHDSGEY*EVHHQK (P39)

Resin amount: 10 µmol; Analytical HPLC R_f = 22.07 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 19.90 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 35 min at 20 mL/min); Preparative HPLC R_f = 16.51 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 560.4962 ($[M+4H]^{4+}$, calc.: 560.4966), 448.5988 ($[M+5H]^{5+}$, calc.: 5987), 746.9944 ($[M+3H]^{3+}$, calc.: 746.9930); *Yield*: 60 % (13.4 mg, 6.0 µmol).

AspAT (human) 91 - 107 unglycosylated: TEG-NLDKEYLPIGGLAEFCK (P40)

Resin amount: 10 µmol; Analytical HPLC R_f = 41.81 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 39.53 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 705.3687 ($[M+3H]^{3+}$, calc.: 705.3688); *Yield*: 49 % (10.3 mg, 4.9 µmol).

AspAT (human) 91 - 107 αGalNAc: TEG-NLDKEY*LPIGGLAEFCK (P41)

Resin amount: 13 µmol; Analytical HPLC R_f = 39.66 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 32.70 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 773.0636 ($[M+3H]^{3+}$, calc.: 773.0620); *Yield*: 32 % (9.4 mg, 4.1 µmol).

AspAT (human) 91 - 107 αGlcNAc: TEG-NLDKEY*LPIGGLAEFCK (P42)

Resin amount: 10 µmol; Analytical HPLC R_f = 33.42 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 37.64 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 773.0624 ($[M+3H]^{3+}$, calc.: 773.0620); *Yield*: 41 % (9.5 mg, 4.1 µmol).

AspAT (human) 91 - 107 βGlcNAc: TEG-NLDKEY*LPIGGLAEFCK (P43)

Resin amount: 13 µmol; Analytical HPLC R_f = 39.32 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 32.54 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 773.0638 ($[M+3H]^{3+}$, calc.: 773.0620); *Yield*: 18 % (5.3 mg, 2.3 µmol).

AspAT (human) 91 - 107 phospho: TEG-NLDKEY*LPIGGLAEFCK (P44)

Resin amount: 10 µmol; Analytical HPLC R_f = 38.86 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 38.71 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 732.0242 ($[M+3H]^{3+}$, calc.: 732.0243); *Yield*: 41 % (9.1 mg, 4.1 µmol).

AspAT (human) 127 - 139 βGlcNAc: TEG-FVTVQTIS*GTGALR (P45)

Resin amount: 10 µmol; Analytical HPLC R_f = 27.51 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 27.63 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 928.5042 ($[M+2H]^{2+}$, calc.: 928.5042), 619.3402 ($[M+3H]^{3+}$, calc.: 619.3386); *Yield*: 10 % (1.9 mg, 1.0 µmol).

AspAT (rat) 326 – 337 unglycosylated: TEG-IAATILTSPDLR (P46)

Resin amount: 10 µmol; Analytical HPLC R_f = 32.68 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 37.57 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 737.4301 ($[M+2H]^{2+}$, calc.: 737.4298); *Yield*: 46 % (6.8 mg, 4.6 µmol).

AspAT (rat) 326 – 337 βGlcNAc: TEG-IAATILTS*PDLR (P47)

Resin amount: 10 µmol; Analytical HPLC R_f = 29.15 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 34.33 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 838.9700 ($[M+2H]^{2+}$, calc.: 838.9695); *Yield*: 42 % (7.0 mg, 4.2 µmol).

AspAT (rat) 326 – 337 βGlcNAc: TEG-IAAT*ILTSPDLR (P48)

Resin amount: 10 µmol; Analytical HPLC R_f = 28.19 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 32.59 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 838.9702 ($[M+2H]^{2+}$, calc.: 838.9695); *Yield*: 55 % (9.3 mg, 5.5 µmol).

AspAT (rat) 326 – 337 βGlcNAc: TEG-IAATILT*SPDLR (P49)

Resin amount: 10 µmol; Analytical HPLC R_f = 28.89 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 33.43 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 838.9701 ($[M+2H]^{2+}$, calc.: 838.9695); *Yield*: 51 % (8.6 mg, 5.1 µmol).

AspAT (rat) 326 – 337 βGlcNAc: TEG-IAAT*ILT*SPDLR (P50)

Resin amount: 10 µmol; Analytical HPLC R_f = 26.63 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 30.15 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 940.5092 ($[M+2H]^{2+}$, calc.: 940.5092); *Yield*: 36 % (6.7 mg, 3.6 µmol).

AspAT (rat) 326 – 337 βGlcNAc: TEG-IAAT*ILTS*PDLR (P51)

Resin amount: 10 µmol; Analytical HPLC R_f = 26.88 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 30.39 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 940.5097 ($[M+2H]^{2+}$, calc.: 940.5092); *Yield*: 24 % (4.5 mg, 2.4 µmol).

AspAT (rat) 326 – 337 βGlcNAc: TEG-IAATILT*S*PDLR (P52)

Resin amount: 10 µmol; Analytical HPLC R_f = 26.62 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 31.39 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 940.5092 ($[M+2H]^{2+}$, calc.: 940.5092); *Yield*: 31 % (5.9 mg, 3.1 µmol).

Atp5a (human) 335 – 347 unglycosylated: TEG-EAYPGDVFYLHSR (P53)

Resin amount: 10 µmol; Analytical HPLC R_f = 27.73 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 32.57 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 878.9309 ($[M+2H]^{2+}$, calc.: 878.9307), 586.2894 ($[M+3H]^{3+}$, calc.: 586.2894); *Yield*: 65 % (11.4 mg, 6.5 µmol).

Atp5a (human) 335 – 347 βGlcNAc: TEG-EAYPGDVFYLHS*R (P54)

Resin amount: 10 µmol; Analytical HPLC R_f = 26.21 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 31.40 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 980.9723 ($[M+2H]^{2+}$, calc.: 980.9720), 654.3168 ($[M+3H]^{3+}$, calc.: 654.3171); *Yield*: 13 % (2.5 mg, 1.3 µmol).

Atp5a (human) 335 – 347 βGlcNAc: TEG-EAY*PGDVFYLHSR (P55)

Resin amount: 10 µmol; Analytical HPLC R_f = 24.71 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 29.63 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 980.9725 ($[M+2H]^{2+}$, calc.: 980.9720); *Yield*: 20 % (3.9 mg, 2.0 µmol).

Atp5a (human) 335 – 347 βGlcNAc: TEG-EAYPGDVFY*LHSR (P56)

Resin amount: 10 µmol; Analytical HPLC R_f = 25.50 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 30.54 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 980.9724 ($[M+2H]^{2+}$, calc.: 980.9720), 654.3169 ($[M+3H]^{3+}$, calc.: 654.3171); *Yield*: 8 % (1.6 mg, 0.8 µmol).

Atp5a (human) 335 – 347 βGlcNAc: TEG-EAY*PGDVFY*LHSR (P57)

Resin amount: 10 µmol; Analytical HPLC R_f = 23.79 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 28.21 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1082.5121 ($[M+2H]^{2+}$, calc.: 1082.5117), 722.0101 ($[M+3H]^{3+}$, calc.: 722.0102); *Yield*: 22 % (4.7 mg, 2.2 µmol).

Atp5a (human) 335 – 347 βGlcNAc: TEG-EAYPGDVFY*LHS*R (P58)

Resin amount: 10 µmol; Analytical HPLC R_f = 26.40 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 32.91 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 722.0098 ($[M+3H]^{3+}$, calc.: 722.0102); *Yield*: 17 % (2.0 mg, 1.7 µmol).

Atp5b (human) 311 – 324 unglycosylated: TEG-FTQAGSEVSALLGR (59)

Resin amount: 10 µmol; Analytical HPLC R_f = 34.74 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 38.12 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 819.9387 ($[M+2H]^{2+}$, calc.: 819.9385); *Yield*: 27 % (5.0 mg, 2.7 µmol).

Atp5b (human) 311 – 324 βGlcNAc: TEG-FTQAGS*EVSALLGR (60)

Resin amount: 10 µmol; Analytical HPLC R_f = 32.65 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 37.53 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 921.9777 ($[M+2H]^{2+}$, calc.: 921.9799); *Yield*: 22 % (4.0 mg, 2.2 µmol).

Atp5b (human) 311 – 324 βGlcNAc: TEG-FTQAGSEVS*ALLGR (61)

Resin amount: 10 µmol; Analytical HPLC R_f = 31.13 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 40.69 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 921.9773 ($[M+2H]^{2+}$, calc.: 921.9799); *Yield*: 64 % (11.7 mg, 6.4 µmol).

Atp5b (human) 311 – 324 βGlcNAc: TEG-FT*QAGSEVSALLGR (62)

Resin amount: 10 µmol; Analytical HPLC R_f = 33.12 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 36.91 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 921.9772 ($[M+2H]^{2+}$, calc.: 921.9799); *Yield*: 45 % (8.3 mg, 4.5 µmol).

Atp5b (human) 311 – 324 βGlcNAc: TEG-FT*QAGS*EVSALLGR (63)

Resin amount: 10 µmol; Analytical HPLC R_f = 32.20 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 35.70 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1023.5175 ($[M+2H]^{2+}$, calc.: 1023.5196); *Yield*: 35 % (7.2 mg, 3.5 µmol).

Atp5b (human) 311 – 324 βGlcNAc: TEG-FTQAGS*EVS*ALLGR (64)

Resin amount: 10 µmol; Analytical HPLC R_f = 30.15 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 33.55 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1023.5178 ($[M+2H]^{2+}$, calc.: 1023.5196); *Yield*: 36 % (7.3 mg, 3.6 µmol).

Atp5b (human) 407 – 422 unglycosylated: TEG-IMDPNIVGSEHYDVAR (65)

Resin amount: 10 µmol; Analytical HPLC R_f = 27.42 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 30.91 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1010.4976 ($[M+2H]^{2+}$, calc.: 1010.4979); *Yield*: 42 % (8.5 mg, 4.2 µmol).

Atp5b (human) 407 – 422 βGlcNAc: TEG-IMDPNIVGS*EHYDVAR (66)

Resin amount: 10 µmol; Analytical HPLC R_f = 25.60 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 40.10 min (A/B, (95/5) 5 min isocratic; (95/5) → (65/35) in 55 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 741.6937 ($[M+3H]^{3+}$, calc.: 741.6941); *Yield*: 13 % (3.0 mg, 1.3 µmol).

Atp5b (human) 407 – 422 βGlcNAc: TEG-IMDPNIVGSEHY*DVAR (67)

Resin amount: 10 µmol; Analytical HPLC R_f = 24.78 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 38.67 min (A/B, (95/5) 5 min isocratic; (95/5) → (65/35) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 741.6938 ($[M+3H]^{3+}$, calc.: 741.6941); *Yield*: 33 % (7.3 mg, 3.3 µmol).

Atp5b (human) 407 – 422 βGlcNAc: TEG-IMDPNIVGS*EHY*DVAR (68)

Resin amount: 10 µmol; Analytical HPLC R_f = 23.77 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 37.24 min (A/B, (95/5) 5 min isocratic; (95/5) → (65/35) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 809.3868 ($[M+3H]^{3+}$, calc.: 809.3873); *Yield*: 20 % (4.9 mg, 2.0 µmol).

Atp5b (human) 407 – 422 βGlcNAc (*), phospho ($^{\#}$): TEG-IMDPNIVGS*EHY$^{\#}$DVAR (69)

Resin amount: 10 µmol; Analytical HPLC R_f = 26.34 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 60 min at 200 µL/min); Preparative HPLC R_f = 40.75 and 41.11 min (peak splitting) (A/B, (95/5) 5 min isocratic; (95/5) → (80/20) in 5 min; (80/20) → (60/40) in 30 min; (60/40) → (90/10) in 1 min at 20

mL/min); *HR-ESI-MS (pos), m/z:* 1152.0210 ($[M+2H]^{2+}$, calc.: 1152.0208); *Yield*: 59 % (13.5 mg, 5.9 µmol).

Atp5b (human) 407 – 422 βGlcNAc (*), phospho (#): TEG-IMDPNIVGS#EHY*DVAR (70)

Resin amount: 10 µmol; Analytical HPLC R_f = 26.09 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 60 min at 200 µL/min); Preparative HPLC R_f = 29.38 min (A/B, (95/5) 5 min isocratic; (95/5) → (60/40) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1152.0195 ($[M+2H]^{2+}$, calc.: 1152.0208); *Yield*: 21 % (4.8 mg, 2.1 µmol).

Atp5b (mouse) 226 – 239 unglycosylated: TEG-AHGGYSVFAGVGER (71)

Resin amount: 10 µmol; Analytical HPLC R_f = 23.77 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 29.49 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 805.4021 ($[M+2H]^{2+}$, calc.: 805.4021); *Yield*: 6 % (1.0 mg, 0.6 µmol).

Atp5b (mouse) 226 – 239 βGlcNAc: TEG-AHGGYS*VFAGVGER (72)

Resin amount: 10 µmol; Analytical HPLC R_f = 21.43 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 43.51 min (A/B, (95/5) 5 min isocratic; (95/5) → (70/30) in 50 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 604.9635 ($[M+3H]^{3+}$, calc.: 604.9636); *Yield*: 63 % (11.4 mg, 6.3 µmol).

Atp5b (mouse) 226 – 239 βGlcNAc: TEG-AHGGY*SVFAGVGER (73)

Resin amount: 10 µmol; Analytical HPLC R_f = 20.96 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 42.35 min (A/B, (95/5) 5 min isocratic; (95/5) → (70/30) in 50 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 604.9635 ($[M+3H]^{3+}$, calc.: 604.9636); *Yield*: 63 % (11.5 mg, 6.3 µmol).

Atp5b (mouse) 226 – 239 βGlcNAc: TEG-AHGGY*S*VFAGVGER (74)

Resin amount: 10 µmol; Analytical HPLC R_f = 20.85 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 29.88 min (A/B, (95/5) 5 min isocratic; (95/5) → (65/35) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 672.9912 ($[M+3H]^{3+}$, calc.: 672.9912); *Yield*: 12 % (2.5 mg, 1.2 µmol).

Atp5b (mouse) 407 – 422 unglycosylated: TEG-IMDPNIVGNEHYDVAR (75)

Resin amount: 10 µmol; Analytical HPLC R_f = 27.35 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 17.59 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1023.5024 ($[M+2H]^{2+}$, calc.: 1023.5017)*; Yield*: 56 % (11.5 mg, 5.6 µmol).

Atp5b (mouse) 407 – 422 αGalNAc: TEG-IMDPNIVGNEHY*DVAR (76)

Resin amount: 13 µmol; Analytical HPLC R_f = 25.61 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 25.77 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 750.6996 ($[M+3H]^{3+}$, calc.: 750.6978); *Yield*: 31 % (8.9 mg, 4.0 µmol).

Atp5b (mouse) 407 – 422 αGlcNAc: TEG-IMDPNIVGNEHY*DVAR (77)

Resin amount: 10 µmol; Analytical HPLC R_f = 25.53 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 30.13 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 750.6974 ($[M+3H]^{3+}$, calc.: 750.6978); *Yield*: 33 % (7.5 mg, 3.3 µmol).

Atp5b (mouse) 407 – 422 αGlcNAc: TEG-IMDPNIVGNEHY*DVAR (78)

Resin amount: 13 µmol; Analytical HPLC R_f = 30.92 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 27.34 min (A/B, (95/5) 5 min isocratic; (95/5) → (82/18) in 13 min; (82/18) → (72/28) in 100 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 750.6998 ($[M+3H]^{3+}$, calc.: 750.6978); *Yield*: 33 % (9.6 mg, 4.3 µmol).

Atp5b (mouse) 407 – 422 phospho: TEG-IMDPNIVGNEHY*DVAR (79)

Resin amount: 10 µmol; Analytical HPLC R_f = 28.50 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 19.90 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 709.6603 ($[M+3H]^{3+}$, calc.: 709.6601); *Yield*: 21 % (4.4 mg, 2.1 µmol).

Cox4i1 (rat) 30 – 42 unglycosylated: TEG-SEDYALPSTVDRR (80)

Resin amount: 10 µmol; Analytical HPLC R_f = 21.91 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 26.25 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 856.4287 ($[M+2H]^{2+}$, calc.: 856.4285), 571.2880 ($[M+3H]^{3+}$, calc.: 571.2880); *Yield*: 36 % (6.1 mg, 3.6 µmol).

Cox4i1 (rat) 30 – 42 βGlcNAc: TEG-S*EDTALPSTVDRR (81)

Resin amount: 10 µmol; Analytical HPLC R_f = 18.21 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 22.47 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 926.9600 ($[M+2H]^{2+}$, calc.: 926.9604), 618.3087 ($[M+3H]^{3+}$, calc.: 618.3093); *Yield*: 54 % (10.0 mg, 5.4 µmol).

Cox4i1 (rat) 30 – 42 βGlcNAc: TEG-TEDYALPS*TVDRR (82)

Resin amount: 10 µmol; Analytical HPLC R_f = 20.99 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 25.24 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 957.9681 ($[M+2H]^{2+}$, calc.: 957.9682), 638.9808 ($[M+3H]^{3+}$, calc.: 638.9808); *Yield*: 58 % (11.1 mg, 5.8 µmol).

Cox4i1 (rat) 30 – 42 βGlcNAc: TEG-TEDY*ALPSTVDRR (83)

Resin amount: 10 µmol; Analytical HPLC R_f = 19.71 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 23.57 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 957.9676 ($[M+2H]^{2+}$, calc.: 957.9682), 638.9804 ($[M+3H]^{3+}$, calc.: 638.9808); *Yield*: 43 % (8.2 mg, 4.3 µmol).

Cox4i1 (rat) 30 – 42 βGlcNAc: TEG-SEDYALPSY*VDRR (84)

Resin amount: 10 µmol; Analytical HPLC R_f = 21.91 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 25.81 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 659.9865 ($[M+3H]^{3+}$, calc.: 659.9876), 989.4769 ($[M+2H]^{2+}$, calc.: 989.4777); *Yield*: 52 % (10.3 mg, 5.2 µmol).

Cox4i1 (rat) 30 – 42 βGlcNAc: TEG-S*EDY*ALPSTVDRR (85)

Resin amount: 10 µmol; Analytical HPLC R_f = 19.17 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 23.10 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1060.0095 ($[M+2H]^{2+}$, calc.: 1060.0095), 707.0083 ($[M+3H]^{3+}$, calc.: 707.0088); *Yield*: 46 % (9.8 mg, 4.6 µmol).

Cox4i1 (rat) 30 – 42 βGlcNAc: TEG-SEDY*ALPS*TVDRR (86)

Resin amount: 10 µmol; Analytical HPLC R_f = 19.42 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 23.05 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1060.0093 ($[M+2H]^{2+}$, calc.: 1060.0095), 707.0082 ($[M+3H]^{3+}$, calc.: 707.0088); *Yield*: 46 % (9.7 mg, 4.6 µmol).

Cox4i1 (rat) 30 – 42 βGlcNAc: TEG-SEDYALPS*Y*VDRR (87)

Resin amount: 10 µmol; Analytical HPLC R_f = 21.67 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 25.29 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 727.6801 ($[M+3H]^{3+}$, calc.: 727.6801), 1091.0171 ($[M+2H]^{2+}$, calc.: 1091.0174); *Yield*: 31 % (6.8 mg, 3.1 µmol).

CDC42 (human) 27 – 38 unglycosylated: TEG-KFPSEYVPTVFD (88)

Resin amount: 13 µmol; Analytical HPLC R_f = 37.53 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 26.34 min (A/B, (90/10 → (45/55) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 816.4152 ($[M+2H]^{2+}$, calc.: 816.4138); *Yield*: 51 % (10.8 mg, 6.6 µmol).

CDC42 (human) 27 – 38 αGalNAc: TEG-KFPSEY*VPTVFD (89)

Resin amount: 10 µmol; Analytical HPLC R_f = 31.88 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 31.43 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 918.4552 ($[M+2H]^{2+}$, calc.: 918.4552); *Yield*: 40 % (7.3 mg, 4.0 µmol).

CDC42 (human) 27 – 38 αGlcNAc: TEG-KFPSEY*VPTVFD (90)

Resin amount: 10 µmol; Analytical HPLC R_f = 29.82 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 31.27 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 918.4545 ($[M+2H]^{2+}$, calc.: 918.4552); *Yield*: 17 % (3.1 mg, 1.7 µmol).

CDC42 (human) 27 – 38 βGlcNAc: TEG-KFPSEY*VPTVFD (91)

Resin amount: 13 µmol; Analytical HPLC R_f = 31.30 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 23.78 min (A/B, (90/10) → (45/55) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 918.4569 ($[M+2H]^{2+}$, calc.: 918.4552); *Yield*: 34 % (8.1 mg, 4.4 µmol).

CDC42 (human) 27 – 38 phospho: TEG-KFPSEY*VPTVFD (92)

Resin amount: 13 µmol; Analytical HPLC R_f = 34.63 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 32.50 min (A/B, (95/5) 5 min isocratic, (95/5) → (45/55) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 856.3981 ($[M+2H]^{2+}$, calc.: 856.3970); *Yield*: 44 % (7.5 mg, 4.4 µmol).

MUC1 VNTR αGalNAc: TEG-PAHGVY*SAPDTRPAPGSTA (93)

Resin amount: 5 µmol; Analytical HPLC R_f = 23.48 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 20.36 min (A/B, (95/5) 5 min isocratic, (95/5) → (45/55) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 753.7031 ($[M+2H]^{2+}$, calc.: 753.7038); *Yield*: 50 % (5.7 mg, 2.5 µmol).

MUC1 VNTR αGalNAc: TEG-PAHGVYSAPDY*RPAPGSTA (94)

Resin amount: 5 µmol; Analytical HPLC R_f = 23.61 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 20.61 min (A/B, (95/5) 5 min isocratic, (95/5) → (45/55) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 753.7032 ($[M+2H]^{2+}$, calc.: 753.7038); *Yield*: 50 % (5.6 mg, 2.5 µmol).

MUC1 VNTR αGalNAc: TEG-PAHGVYSAPDYRPAPGSY*A (95)

Resin amount: 5 µmol; Analytical HPLC R_f = 23.95 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 20.97 min (A/B, (95/5) 5 min isocratic, (95/5) → (45/55) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 753.7039 ($[M+2H]^{2+}$, calc.: 753.7038); *Yield*: 54 % (2.6 mg, 2.7 µmol).

MUC1 VNTR αGlcNAc: TEG-PAHGVY*SAPDTRPAPGSTA (96)

Resin amount: 10 µmol; Analytical HPLC R_f = 16.79 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 20.46 min (A/B, (95/5) 5 min isocratic, (95/5) → (45/55) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 753.7037 ($[M+2H]^{2+}$, calc.: 753.7038); *Yield*: 47 % (10.6 mg, 4.7 µmol).

MUC1 VNTR αGlcNAc: TEG-PAHGVYSAPDY*RPAPGSTA (97)

Resin amount: 10 µmol; Analytical HPLC R_f = 18.39 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 20.88 min (A/B, (95/5) 5 min isocratic, (95/5) → (45/55) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 753.7035 ($[M+2H]^{2+}$, calc.: 753.7038); *Yield*: 46 % (10.5 mg, 4.6 µmol).

MUC1 VNTR αGlcNAc: TEG-PAHGVYSAPDYRPAPGSY*A (98)

Resin amount: 10 µmol; Analytical HPLC R_f = 17.55 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 20.82 min (A/B, (95/5) 5 min isocratic, (95/5) → (45/30) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 753.7038 ($[M+2H]^{2+}$, calc.: 753.7038); *Yield*: 42 % (9.6 mg, 4.2 µmol).

MUC1 VNTR βGlcNAc: TEG-PAHGVTSAPDS*RPAPGSYA (99)

Resin amount: 10 µmol; Analytical HPLC R_f = 25.67 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 22.19 min (A/B, (95/5) 5 min isocratic, (95/5) → (70/30) in 25 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 728.3600 ($[M+3H]^{3+}$, calc.: 728.3600); *Yield*: 24 % (5.2 mg, 2.4 µmol).

MUC1 VNTR βGlcNAc: TEG-PAHGVTSAPDTRPAPGST*A (100)

Resin amount: 10 µmol; Analytical HPLC R_f = 16.44 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 32.43 min (A/B, (95/5) 5 min isocratic, (95/5) → (75/25) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 733.0320 ($[M+3H]^{3+}$, calc.: 733.0319); *Yield*: 56 % (12.2 mg, 5.6 µmol).

MUC1 VNTR βGlcNAc: TEG-PAHGVY*SAPDTRPAPGSTA (101)

Resin amount: 5 µmol; Analytical HPLC R_f = 17.34 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 20.00 min (A/B, (95/5) 5 min isocratic, (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 753.7033 ($[M+2H]^{2+}$, calc.: 753.7038); *Yield*: 52 % (5.8 mg, 2.6 µmol).

MUC1 VNTR βGlcNAc: TEG-PAHGVTSAPDY*RPAPGSTA (102)

Resin amount: 5 µmol; Analytical HPLC R_f = 17.86 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 20.38 min (A/B, (95/5) 5 min isocratic, (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 753.7034 ($[M+3H]^{3+}$, calc.: 753.7038); *Yield*: 58 % (6.5 mg, 2.9 µmol).

MUC1 VNTR βGlcNAc: TEG-PAHGVTSAPDTRPAPGSY*A (103)

Resin amount: 5 µmol; Analytical HPLC R_f = 17.98 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 20.59 min (A/B, (95/5) 5 min isocratic, (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 753.7035 ($[M+3H]^{3+}$, calc.: 753.7038); *Yield*: 52 % (5.9 mg, 2.6 µmol).

NucB2 (human) 387 – 399 unglycosylated: TEG-LEYHQVIQQMEQK (104)

Resin amount: 10 µmol; Analytical HPLC R_f = 34.41 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 29.36 min (A/B, (95/5) 5 min isocratic, (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 626.6554 ($[M+3H]^{3+}$, calc.: 626.6551); *Yield*: 59 % (10.3 mg, 5.9 µmol).

NucB2 (human) 387 – 399 αGalNAc: TEG-LEY*HQVIQQMEQK (105)

Resin amount: 10 µmol; Analytical HPLC R_f = 31.34 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 26.99 min (A/B, (95/5) 5 min isocratic, (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 699.6798 ($[M(Met396\ oxidized)+3H]^{3+}$, calc.: 699.6798); *Yield*: 8 % (1.7 mg, 0.8 µmol).

NucB2 (human) 387 – 399 βGlcNAc: TEG-LEY*HQVIQQMEQK (106)

Resin amount: 10 µmol; Analytical HPLC R_f = 31.27 min (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 26.71 min (A/B, (95/5) 5 min isocratic, (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 699.6800 ($[M(Met396\ oxidized)+3H]^{3+}$, calc.: 699.6798); *Yield*: 25 % (5.1 mg, 2.5 µmol).

NucB2 (human) 387 – 399 phospho: TEG-LEY*HQVIQQMEQK (107)

Resin amount: 10 µmol; Analytical HPLC R_f = 30.81 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 27.99 min (A/B, (95/5) 5 min isocratic, (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 653.3105 ($[M+3H]^{3+}$, calc.: 653.3105); *Yield*: 45 % (8.9 mg, 4.5 µmol).

Ret (human) 680 – 692 unglycosylated: TEG-AQAFPVSYSSSGA (108)

Resin amount: 10 µmol; Analytical HPLC R_f = 33.00 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 29.41 min (A/B, (95/5) 5 min isocratic, (95/5) → (40/60) in 55 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1313.6015 ($[M+H]^{+}$, calc.: 1313.6008); *Yield*: 50 % (6.6 mg, 5.0 µmol).

RAC1 (human) 27 – 38 unglycosylated: TEG-AFPGEYIPTVFD (109)

Resin amount: 13 µmol; Analytical HPLC R_f = 41.69 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 33.53 min (A/B, (85/15) → (45/55) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 779.8883 ($[M+2H]^{2+}$, calc.: 779.8874); *Yield*: 85 % (17.4 mg, 11.1 µmol).

RAC1 (human) 27 – 38 αGalNAc: TEG-AFPGEY*IPTVFD (110)

Resin amount: 13 µmol; Analytical HPLC R_f = 38.42 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 35.29 min (A/B, (95/5) 5 min isocratic, (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 881.4272 ($[M+2H]^{2+}$, calc.: 881.4271); *Yield*: 35 % (6.2 mg, 3.5 µmol).

RAC1 (human) 27 – 38 αGlcNAc: TEG-AFPGEY*IPTVFD (111)

Resin amount: 10 µmol; Analytical HPLC R_f = 35.69 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 35.88 min (A/B, (95/5) 5 min isocratic, (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 881.4275 ($[M+2H]^{2+}$, calc.: 881.4271), 892.4205 ($[M+H+Na]^{2+}$, calc.: 892.4181); *Yield*: 17 % (3.0 mg, 1.7 µmol).

RAC1 (human) 27 – 38 βGlcNAc: TEG-AFPGEY*IPTVFD (112)

Resin amount: 13 µmol; Analytical HPLC R_f = 37.88 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 27.56 min (A/B, (90/10) → (45/55) in 45 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 881.9314 ($[M+2H]^{2+}$, calc.: 881.9288); *Yield*: 28 % (6.5 mg, 3.7 µmol).

RAC1 (human) 27 – 38 phospho: TEG-AFPGEY*IPTVFD (113)

Resin amount: 10 µmol; Analytical HPLC R_f = 37.87 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 39.18 min (A/B, (95/5) 5 min isocratic, (95/5) → (40/60) in 55 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 819.8701 ($[M+2H]^{2+}$, calc.: 819.8706); *Yield*: 11 % (1.8 mg, 1.1 µmol).

Rational construct unglycosylated: Ac-GYYG-TEG (114)

Resin amount: 13 µmol; Analytical HPLC R_f = 24.87 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 19.18 min (A/B, (95/5) → (55/45) in 35 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 703.3657 ($[M+H]^+$, calc.: 703.3667); *Yield*: 43 % (3.3 mg, 5.6 µmol).

Rational construct unglycosylated: Ac-KKVPVSRA (115)

Resin amount: 10 µmol; Analytical HPLC R_f = 21.20 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 17.13 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 926.5787 ($[M+H]^+$, calc.: 926.5782); *Yield*: 92 % (8.5 mg, 9.2 µmol).

Rational construct αGalNAc: Ac-GY*Y*G-TEG (116)

Resin amount: 13 µmol; Analytical HPLC R_f = 21.04 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 15.64 min (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 555.2658 ($[M+2H]^{2+}$, calc.: 555.2661); *Yield*: 4 % (0.6 mg, 0.5 µmol).

Rational construct βGlcNAc: Ac-GY*Y*G-TEG (117)

Resin amount: 13 µmol; Analytical HPLC R_f = 20.18 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 13.79 and 14.59 min (peak splitting) (A/B, (95/5) → (55/45) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 555.2660 ($[M+2H]^{2+}$, calc.: 555.2661); *Yield*: 8 % (1.2 mg, 1.0 µmol).

Rbl2 (human) 411 – 422 unglycosylated: Ac-KENPAVTPVSTA (118)

Resin amount: 10 µmol; Analytical HPLC R_f = 25.40 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 22.05 (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 1255.6531 ($[M+H]^+$, calc.: 1255.6531); *Yield*: 74 % (9.3 mg, 7.4 µmol).

RhoA (human) 29 – 40 unglycosylated: TEG-QFPEVYVPTVFE (119)

Resin amount: 13 µmol; Analytical HPLC R_f = 41.68 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 33.79 (A/B, (85/15) → (45/55) in 40 min at 20 mL/min); *HR-ESI-MS (pos), m/z:* 829.4226 ($[M+2H]^{2+}$, calc.: 829.4216); *Yield*: 54 % (11.7 mg, 7.0 µmol).

RhoA (human) 29 – 40 αGalNAc: TEG-QFPEVY*VPTVFE (120)

Resin amount: 10 µmol; Analytical HPLC R_f = 38.73 min (C/D, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 35.53 (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 35 min at 20 mL/min); *ESI-MS (pos), m/z:* 931.4628 ($[M+2H]^{2+}$, calc.: 931.4630); *Yield*: 22 % (4.2 mg, 2.2 µmol).

RhoA (human) 29 – 40 αGlcNAc: TEG-QFPEVY*VPTVFE (121)

Resin amount: 10 µmol; Analytical HPLC R_f = 35.95 min (C/D, (95/5) 5 min isocratic; (95/5) → (55/45) in 40 min at 200 µL/min); Preparative HPLC R_f = 35.29 (A/B, (95/5) 5 min isocratic; (95/5) → (55/45) in 35 min at 20 mL/min); *ESI-MS (pos), m/z:* 931.4631 ($[M+2H]^{2+}$, calc.: 931.4630); *Yield*: 15 % (2.7 mg, 1.5 µmol).

RhoA (human) 29 – 40 βGlcNAc: TEG-QFPEVY*VPTVFE (122)

Resin amount: 13 µmol; Analytical HPLC R_f = 39.50 min (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µL/min); Preparative HPLC R_f = 28.38 (A/B, (90/10) → (45/55) in 45 min at 20 mL/min); *ESI-MS (pos), m/z:* 931.4643 ($[M+2H]^{2+}$, calc.: 931.4630), 633.9608 ($[M+2H+K]^{3+}$, calc.: 633.9630; *Yield*: 7 % (1.7 mg, 0.9 µmol).

RhoA (human) 29 – 40 phospho: TEG-QFPEVY*VPTVFE (123)

Resin amount: 13 µmol; Analytical HPLC R_f = 38.31 min (A/A, (95:5) 5 min isocratic; (95/5) → (40/60) in 55 min at 200 µl/min); Preparative HPLC R_f = 39.23 (A/B, (95/5) 5 min isocratic; (95/5) → (40/60) in 55 min at 20 mL/min); *MS (pos), m/z:* 869.4054 ($[M+2H]^{2+}$, calc.: 869.5054); *Yield*: 6 % (1.1 mg, 0.6 µmol).

9.3 Microarray experiments

9.3.1 General procedure

As microarray glass slides NHS-coated *Nexterion H*® slides were utilized (*Schott GmbH*, Mainz). The slides were spotted in an *iTwo 400* spotter (*M2 Automation*) equipped with an additional air humidity control system. The fluorescence readout was performed with a *Typhoon Trio+* by *Amersham Biosciences* (*GE Healthcare*), a tube voltage of 470 – 480 V and a resolution of 10 µm. The excitation of *Cy5*-labelled conjugates was done with a red laser at 633 nm and a band-pass emission filter at 670 nm. The excitation of *Alexa Fluor 488*-labelled conjugates was done with a blue laser at 488 nm with a band-pass emission filter at 520 nm. With the help of *Amersham Typhoon Array* software the scans were processed performing a background subtraction *via spot edge average* background subtraction. The obtained data was processed in *Microsoft Excel*. The presented values in this work represent the mean values of all spot replicates with corresponding standard deviation.

9.3.2 Lectins and antibodies

Biotinylated plant lectins (purchased from *Vector Laboratories Inc*):

- Biotinylated Wheat germ agglutinin (WGA) from *Triticum aestivum*, 5 mg/mL.
- Biotinylated Vicia Villosa (VVL), 2 mg/mL.
- Biotinylated Griffonia (Bandeiraea) simplicifolia lectin II (GSL II), 2 mg/mL.

His-tagged lectins:

- Penta-His Fiber nob protein AD37, recombinant (*E. coli*) obtained from Prof. Dr. *Niklas Arnberg* (Department of Clinical Microbiology, University of Umeå, Sweden).
- Penta-His Fiber nob protein AD52, recombinant (*E. coli*) obtained from Prof. Dr. *Niklas Arnberg* (Sweden).
- Human His-tagged, recombinant (*E. coli*) galectin-3 (1-250aa) purchased from *ATGen Ltd*, South Korea.

Cy5-Streptavidin conjugate for detection of biotinylated lectins:

- Cy5-Streptavidin conjugate (*ZyMax*™ grade), 1 mg/mL, purchased from *Invitrogen* (*Thermo Scientific*).

Anti-penta-His antibody for detection of His-tagged lectins:

- *Penta-His Alexa Fluor® 488* mouse IgG1 conjugate, 200 µg/mL, purchased from *Quiagen* (Venlo, Netherlands).

Anti-galectin-3 antibody:

- *Alexa Fluor®* 488 anti-mouse/human Mac-2 (galectin-3), 500 µg/mL, *Biolegends Inc* (San Diego, USA).

9.3.3 General spotting conditions

Glycopeptide stock solutions were prepared in a concentration of 50 µM in spotting buffer (150 mM NaH_2PO_4/Na_2HPO_4, pH 8.5) and 20 µL of each solution was pipetted into a *Nunc®* 384-well plate. The well plate and the *Nexterion H®* slides were then inserted to the spotter. The droplets were generated in a piezo-driven method with a size of 100 ± 3 pL each. While during the spotting process the air humidity was kept at 50 – 60 %, the spotted microarray slides were stored at 90 – 99 % overnight in order to achieve full surface immobilization. Afterwards, the slides were stored at -20° C until experimental application.

9.3.4 Microarray slides MA1, MA2 and MA3

In the frame of this work, three microarray formats have been applied. **MA1** and **MA2** have been spotted with the glycopeptides from the HexNAc glycoeptpide library (chapter 7.2.5) and some additional glycopeptides from Dr. *Christian Pett* and Dr. *Cai Hui* (both *Leibniz-Institut für Analytische Wissenschaften – ISAS – e.V., Dortmund*), while **MA3** was spotted and provided by Dr. *Christian Pett*.

9.3.4.1 Format of microarray MA1

Microarray slide MA1 contains a library of immobilized αGalNAc and βGlcNAc glycopeptides (***Table 5***) and was used for plant lectin incubations with WGA, GSL and VVA (see chapter 7.2.6). The slide contains 16 identical arrays with 14x14 spots (196 spots total) each with a pitch of 350 µm in each direction (**Figure 101**). The wells were incubated with 50 µL plant lectin solution each. Detection of the lectins was done with Cy5-Streptavidin conjugate (dilution 1:1000).

Table 5: List of glycopeptides immobilized on microarray **MA1**.

position on microarray	peptide	protein segment	sequence	glycan	origin
1	**P91**	Cdc42 27 - 38	TEG-KFPSEY*VPTVFD	βGlcNAc	this work
2	**P112**	Rac1 27 - 38	TEG-AFPGEY*IPTVFD	βGlcNAc	this work
3	**P127**	RhoA 29 - 40	TEG-QFPEVY*VPTVFE	βGlcNAc	this work
4	**P88**	Cdc42 27 - 38	TEG-KFPSEYVPTVFD	---	this work
5	**P109**	Rac1 27 - 38	TEG-AFPGEYIPTVFD	---	this work
6	**P119**	RhoA 29 - 40	TEG-QFPEVYVPTVFE	---	this work
7	**P1**	MUC1 VNTR	TEG-HGVTSAPDTRPAPGSTA	---	this work
8	**P3**	MUC1 VNTR	TEG-PAHGVTSAPDT*RPAPGSTA	T_N	this work
9	**P2**	MUC1 VNTR	TEG-PAHGVT*SAPDTRPAPGSTA	T_N	this work
10	**P4**	MUC1 VNTR	TEG-PAHGVTSAPDTRPAPGST*A	T_N	this work
11	**P124**	MUC1 VNTR	TEG-PAHGVT*SAPDTRPAPGSTA	βGlcNAc	Dr. C. Pett
12	**P125**	MUC1 VNTR	TEG-PAHGVTSAPDTRPAPGS*TA	βGlcNAc	Dr. C. Pett
13	**P16**	MUC1 VNTR	TEG-PAHGVTSAPDTRPAPGST*A	core 2 type-2	this work
14	**P36**	Aβ 672 - 687	TEG-DAEFRHDSGY*EVHHQK	αGalNAc	this work
15	**P38**	Aβ 672 - 687	TEG-DAEFRHDSGY*EVHHQK	βGlcNAc	this work
16	**P76**	Atp5b 407 - 422	TEG-IMDPNIVGNEHY*DVAR	αGalNAc	this work
17	**P78**	Atp5b 407 - 422	TEG-IMDPNIVGNEHY*DVAR	βGlcNAc	this work
18	**P41**	AspAT 91 - 107	TEG-NLDKEY*LPIGGLAEFCK	αGalNAc	this work
19	**P43**	AspAT 91 - 107	TEG-NLDKEY*LPIGGLAEFCK	βGlcNAc	this work
20	**P105**	NucB2 387 - 399	TEG-LEY*HQVIQQMEQK	αGalNAc	this work
21	**P106**	NucB2 387 - 399	TEG-LEY*HQVIQQMEQK	βGlcNAc	this work
22	**P30**	MUC1 VNTR	TEG-PAHGVTSAPDTRPAPGST*A	core 4 type-2	this work
23	**P89**	Cdc42 27 - 38	TEG-KFPSEY*VPTVFD	αGalNAc	this work
24	**P110**	Rac1 27 - 38	TEG-AFPGEY*IPTVFD	αGalNAc	this work
25	**P120**	RhoA 29 - 40	TEG-QFPEVY*VPTVFE	αGalNAc	this work
26	**P35**	Aβ 672 - 687	TEG-DAEFRHDSGYEVHHQK	---	this work
27	**P75**	Atp5b 407 - 422	TEG-IMDPNIVGNEHYDVAR	---	this work
28	**P40**	AspAT 91 - 107	TEG-NLDKEYLPIGGLAEFCK	---	this work
29	**P104**	NucB2 387 - 399	TEG-LEYHQVIQQMEQK	---	this work
30	**P94**	MUC1 VNTR	TEG-PAHGVTSAPDY*RPAPGSTA	αGalNAc	this work
31	**P93**	MUC1 VNTR	TEG-PAHGVY*SAPDTRPAPGSTA	αGalNAc	this work
32	**P95**	MUC1 VNTR	TEG-PAHGVTSAPDTRPAPGSY*A	αGalNAc	this work
33	**P102**	MUC1 VNTR	TEG-PAHGVTSAPDY*RPAPGSTA	βGlcNAc	this work
34	**P101**	MUC1 VNTR	TEG-PAHGVY*SAPDTRPAPGSTA	βGlcNAc	this work
35	**P103**	MUC1 VNTR	TEG-PAHGVTSAPDTRPAPGSY*A	βGlcNAc	this work

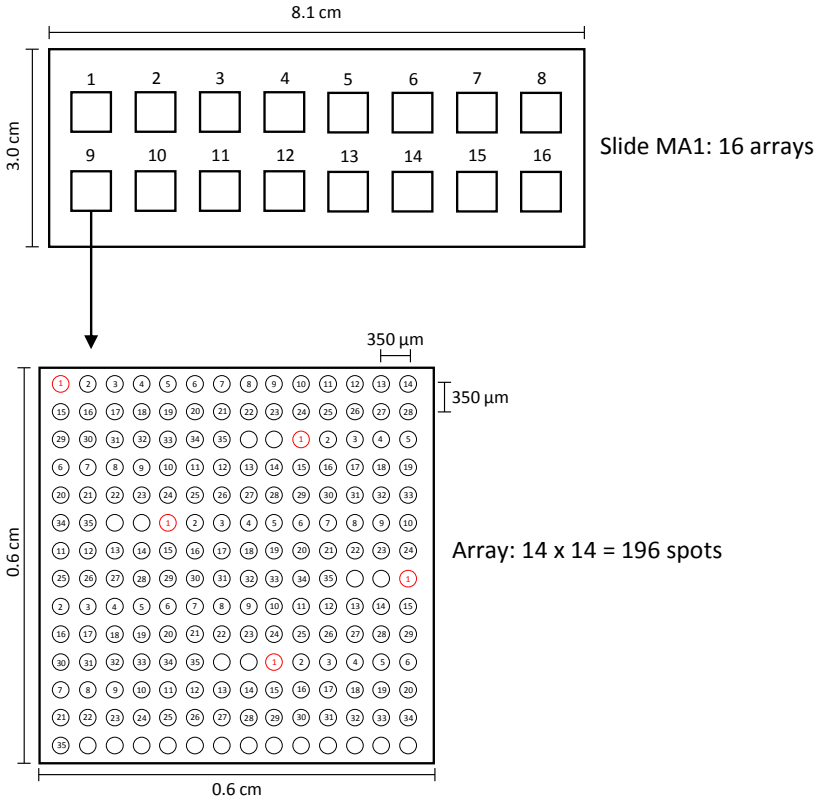


Figure 101: Format of microarray **MA1**.

9.3.4.2 Format of microarray MA2

Microarray slide MA2 contains an immobilized βGlcNAc glycopeptide library (***Table 6***) and was used for plant lectin incubations with WGA and GSL II (chapter 7.2.6). It contains 16 identical arrays with 15x15 substrate spots (225 spots total) with a pitch of 330 µm in each direction. The wells were incubated with 50 µL plant lectin solution each. Detection of the lectins was done with Cy5-Streptavidin conjugate (dilution 1:1000).

Table 6: List of glycopeptides immobilized on **MA2**.

position on microarray	peptide	protein segment	sequence	glycan	origin
1	P88	Cdc42 27 - 38	TEG-KFPSEYVPTVFD	---	this work
2	P91	Cdc42 27 - 38	TEG-KFPSEY*VPTVFD	βGlcNAc	this work
3	P112	Rac1 27 - 38	TEG-AFPGEY*IPTVFD	βGlcNAc	this work
4	P122	RhoA 29 - 40	TEG-QFPEVY*VPTVFE	βGlcNAc	this work
5	P38	Aβ 672 - 687	TEG-DAEFRHDSGY*EVHHQK	βGlcNAc	this work
6	P43	AspAT 91 - 107	TEG-NLDKEY*LPIGGLAEFCK	βGlcNAc	this work
7	P106	NucB2 387 - 399	TEG-LEY*HQVIQQMEQK	βGlcNAc	this work
8	P102	MUC1 VNTR	TEG-PAHGVTSAPDY*RPAPGSTA	βGlcNAc	this work
9	P101	MUC1 VNTR	TEG-PAHGVY*SAPDTRPAPGSTA	βGlcNAc	this work
10	P103	MUC1 VNTR	TEG-PAHGVTSAPDTRPAPGSY*A	βGlcNAc	this work
11	P100	MUC1 VNTR	TEG-PAHGVTSAPDTRPAPGST*A	βGlcNAc	this work
12	P99	MUC1 VNTR	TEG-PAHGVTSAPDS*RPAPGSTA	βGlcNAc	this work
13	P66	Atp5b 407 - 422	TEG-IMDPNIVGS*EHYDVAR	βGlcNAc	this work
14	P67	Atp5b 407 - 422	TEG-IMDPNIVGSEHY*DVAR	βGlcNAc	this work
15	P68	Atp5b 407 - 422	TEG-IMDPNIVGS*EHY*DVAR	βGlcNAc	this work
16	P69	Atp5b 407 - 422	TEG-IMDPNIVGS*EHY°DVAR	*βGlcNAc, °phospho	this work
17	P70	Atp5b 407 - 422	TEG-IMDPNIVGS°EHY*DVAR	*βGlcNAc, °phospho	this work
18	P78	Atp5b 407 - 422	TEG-IMDPNIVGNEHY*DVAR	βGlcNAc	this work
19	P73	Atp5b 226 - 239	TEG-AHGGY*SVFAGVGER	βGlcNAc	this work
20	P66	Atp5b 226 - 239	TEG-AHGGY*S*VFAGVGER	βGlcNAc	this work
21	P68	Atp5b 226 - 239	TEG-AHGGY*S*VFAGVGER	βGlcNAc	this work
22	Unrelated pepetides				
23					
24					
25					
26					
27	P55	Atp5a 335 - 347	TEG-EAY*PGDVFYLHSR	βGlcNAc	this work
28	P56	Atp5a 335 - 347	TEG-EAYPGDVFY*LHSR	βGlcNAc	this work
29	P54	Atp5a 335 - 347	TEG-EAYPGDVFYLHS*R	βGlcNAc	this work
30	P58	Atp5a 335 - 347	TEG-EAYPGDVFY*LHS*R	βGlcNAc	this work
31	P57	Atp5a 335 - 347	TEG-EAY*PGDVFY*LHSR	βGlcNAc	this work
32	P45	Atp5a 127 - 139	TEG-FVTVQTIS*GTGALR	βGlcNAc	this work
33	P48	AspAT 326 - 337	TEG-IAAT*ILTSPDLR	βGlcNAc	this work
34	P49	AspAT 326 - 337	TEG-IAATILT*SPDLR	βGlcNAc	this work
35	P47	AspAT 326 - 337	TEG-IAATILTS*PDLR	βGlcNAc	this work
36	P50	AspAT 326 - 337	TEG-IAAT*ILT*SPDLR	βGlcNAc	this work
37	P51	AspAT 326 - 337	TEG-IAAT*ILTS*PDLR	βGlcNAc	this work
38	P52	AspAT 326 - 337	TEG-IAATILT*S*PDLR	βGlcNAc	this work
39	P62	Atp5b 311 - 324	TEG-FT*QAGSEVSALLGR	βGlcNAc	this work
40	P60	Atp5b 311 - 324	TEG-FTQAGS*EVSALLGR	βGlcNAc	this work
41	P61	Atp5b 311 - 324	TEG-FTQAGSEVS*ALLGR	βGlcNAc	this work
42	P63	Atp5b 311 - 324	TEG-FT*QAGS*EVSALLGR	βGlcNAc	this work
43	P64	Atp5b 311 - 324	TEG-FTQAGS*EVS*ALLGR	βGlcNAc	this work
44	P81	Cox4i1 30 - 42	TEG-S*EDTALPSTVDRR	βGlcNAc	this work
45	P83	Cox4i1 30 - 42	TEG-TEDY*ALPSTVDRR	βGlcNAc	this work
46	P82	Cox4i1 30 - 42	TEG-TEDYALPS*TVDRR	βGlcNAc	this work
47	P84	Cox4i1 30 - 42	TEG-SEDYALPSY*VDRR	βGlcNAc	this work
48	P85	Cox4i1 30 - 42	TEG-S*EDY*ALPSTVDRR	βGlcNAc	this work
49	P86	Cox4i1 30 - 42	TEG-SEDY*ALPS*TVDRR	βGlcNAc	this work
50	P87	Cox4i1 30 - 42	TEG-SEDYALPS*Y*VDRR	βGlcNAc	this work

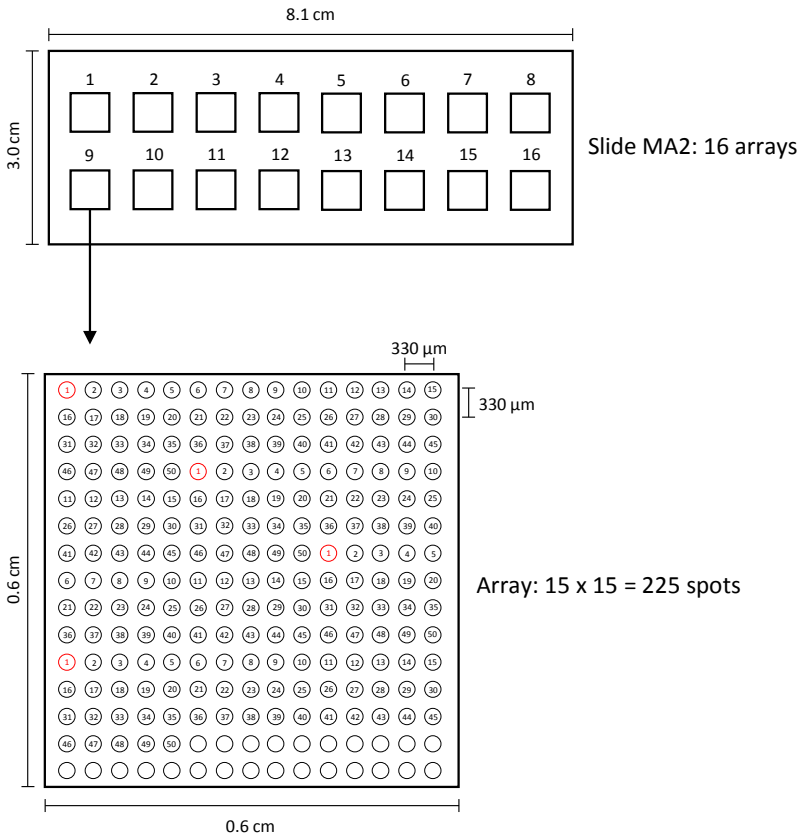


Figure 102: Format of microarray **MA2**.

9.3.4.3 Format of microarray MA3

Microarray slide MA3 was utilized for incubations with adenovirus nob proteins AD37, AD52 and galectin-3 (chapter 7.1.15). It contains eight identical arrays with a sialoglycopeptide library (***Figure 103***, ***Table 7***). Each array contains 5 replicates of this library in a 42x16 = 672 spots with a pitch of 340x343 µm. An illustration can be found in the dissertation[153] of Dr. *C. Pett*.

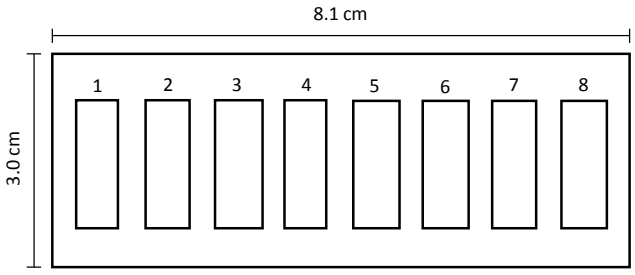


Figure 103: Format of microarray slide **MA3**.

Table 7: List of glycopeptides immobilized on **MA3**. The slide was kindly provided by Dr. *Christian Pett*.

position on microarray	peptide	MUC1 sequence	glycan	origin
1	**P131**	TEG-HGVTSAPDTRPAPGSTAPPA	---	Dr. H Cai
2	**P132**	TEG-PAHGVT*SAPDTRPAPGSTAP	ST_N	Prof. Dr. U. Westerlind
3	**P133**	TEG-PAHGVTSAPDT*RPAPGSTAP	ST_N	Prof. Dr. U. Westerlind
4	**P134**	TEG-PAHGVTSAPDTRPAPGST*AP	ST_N	Prof. Dr. U. Westerlind
5	**P135**	TEG-PAHGVT*SAPDTRPAPST*AP	ST_N	Prof. Dr. U. Westerlind
6	**P136**	TEG-PAHGVT*SAPDTRPAP	ST_N	Prof. Dr. U. Westerlind
7	**P137**	TEG-GSTAPPAHGVT*SAP	ST_N	Prof. Dr. U. Westerlind
8	**P138**	TEG-PAHGVT*SAPDTRPAPGST°AP	*ST_N, °T_N	Prof. Dr. U. Westerlind
9	**P139**	TEG-PAHGVT*SAPDT°RPAPGST*AP	*ST_N, °T_N	Prof. Dr. U. Westerlind
10	**P140**	TEG-PAHFVT*SAPDTRPAPGST°APPAHGVT*SAPDTRPAPGST°AP	*ST_N, °T_N	Prof. Dr. U. Westerlind
11	**P141**	TEG-PAHGVTS*APDTRPAPGSTAP	T_N	Prof. Dr. U. Westerlind
12	**P142**	TEG-PAHGVTSAPDTRPAPGS*TAP	T_N	Prof. Dr. U. Westerlind
13	**P143**	TEG-APDTRPAPGST*AP	T_N	Prof. Dr. U. Westerlind
14	**P144**	TEG-HGVTSAPDT*PAPGS*T*APPA	T_N	Dr. H. Cai
15	**P145**	TEG-HGVTSAPDTRPAPGS*T*APPA	T_N	Dr. H. Cai
16	**P146**	TEG-HGVTSAPDTRPAPGST*APPA	T_N	Dr. H. Cai
17	**P147**	TEG-HGVTSAPDTRPAPGS*TAPPA	T_N	Dr. H. Cai
18	**P148**	TEG-HGVTSAPDT*RPAPGS*TAPPA	T_N	Dr. H. Cai
19	**P149**	TEG-HGVTSAPDT*RPAPGST*APPA	T_N	Dr. H. Cai
20	**P150**	TEG-HGVTSAPDT*RPAPGSTAPPA	T_N	Dr. H. Cai
21	**P151**	TEG-PAHGVT*SAPDTRPAPGSTA	T_N	Dr. C. Pett
22	**P152**	TEG-PAHGVTSAPDT*RPAPGSTA	T_N	Dr. C. Pett
23	**P153**	TEG-PAHGVTSAPDTRPAPGST*A	T_N	Dr. C. Pett
24	**P154**	TEG-PAHGVT*SAPDT*RPAPGSTA	T_N	Dr. C. Pett
25	**P155**	TEG-PAHGVT*SAPDTRPAPGST*A	T_N	Dr. C. Pett
26	**P156**	TEG-PAHGVTSAPDT*RPAPGST*A	T_N	Dr. C. Pett
27	**P157**	TEG-PAHGVT*SAPDT*RPAPGST*A	T_N	Dr. C. Pett
28	**P158**	TEG-PAHGVTS*APDTRPAPGSTAP	T	Prof. Dr. U. Westerlind
29	**P159**	TEG-PAHGVTSAPDTRPAPGST*AP	T	Prof. Dr. U. Westerlind
30	**P160**	TEG-PAHGVT*SAPDTRPAP	T	Prof. Dr. U. Westerlind
31	**P161**	TEG-PAHGVTS*APDTRPAP	T	Prof. Dr. U. Westerlind
32	**P162**	TEG-HGVTSAPDT*RPAPGS°T°APPA	*T, °T_N	Dr. H. Cai
33	**P163**	TEG-HGVTSAPDTRPAPGS°T°APPA	*T, °T_N	Dr. H. Cai
34	**P164**	TEG-PGSTAPPAHGVTSAPDT*RPA	T	Dr. C. Pett
35	**P165**	TEG-APDT*RPAPGSTAPPAHGVTSA	T	Dr. C. Pett
36	**P166**	TEG-APDT*RPA	T	Dr. C. Pett
37	**P167**	TEG-PAHGVT*SAPDTRPAPGSTA	T	Dr. C. Pett
38	**P168**	TEG-PAHGVTSAPDT*RPAPGSTA	T	Dr. C. Pett
39	**P169**	TEG-PAHGVTSAPDTRPAPGST*A	T	Dr. C. Pett
40	**P170**	TEG-PAHGVT*SAPDT*RPAPGSTA	T	Dr. C. Pett
41	**P171**	TEG-PAHGVT*SAPDTRPAPGST*A	T	Dr. C. Pett
42	**P172**	TEG-PAHGVTSAPDT*RPAPGST*A	T	Dr. C. Pett
43	**P173**	TEG-PAHGVT*SAPDT*RPAPGST*A	T	Dr. C. Pett

44	**P174**	TEG-PAHGVT*SAPDTRPAPGSTA	ST	Dr. C. Pett
45	**P175**	TEG-PAHGVTSAPDT*RPAPGSTA	ST	Dr. C. Pett
46	**P176**	TEG-PAHGVTSAPDTRPAPGST*A	ST	Dr. C. Pett
47	**P177**	TEG-PAHGVT*SAPDT*RPAPGSTA	ST	Dr. C. Pett
48	**P178**	TEG-PAHGVT*SAPDTRPAPGST*A	ST	Dr. C. Pett
49	**P179**	TEG-PAHGVTSAPDT*RPAPGST*A	ST	Dr. C. Pett
50	**P180**	TEG-PAHGVT*SAPDT*RPAPGST*A	ST	Dr. C. Pett
51	**P181**	TEG-PAHGVT*SAPDTRPAPGSTA	Core 3 type-1	Dr. C. Pett
52	**P182**	TEG-PAHGVTSAPDT*RPAPGSTA	Core 3 type-1	Dr. C. Pett
53	**P183**	TEG-PAHGVTSAPDTRPAPGST*A	Core 3 type-1	Dr. C. Pett
54	**P184**	TEG-PAHGVT*SAPDT*RPAPGSTA	Core 3 type-1	Dr. C. Pett
55	**P185**	TEG-PAHGVT*SAPDTRPAPGST*A	Core 3 type-1	Dr. C. Pett
56	**P186**	TEG-PAHGVTSAPDT*RPAPGST*A	Core 3 type-1	Dr. C. Pett
57	**P187**	TEG-PAHGVT*SAPDT*RPAPGST*A	Core 3 type-1	Dr. C. Pett
58	**P188**	TEG-PAHGVT*SAPDTRPAPGSTA	Core 3 type-2	Dr. C. Pett
59	**P189**	TEG-PAHGVTSAPDT*RPAPGSTA	Core 3 type-2	Dr. C. Pett
60	**P190**	TEG-PAHGVTSAPDTRPAPGST*A	Core 3 type-2	Dr. C. Pett
61	**P191**	TEG-PAHGVT*SAPDT*RPAPGSTA	Core 3 type-2	Dr. C. Pett
62	**P192**	TEG-PAHGVT*SAPDTRPAPGST*A	Core 3 type-2	Dr. C. Pett
63	**P193**	TEG-PAHGVTSAPDT*RPAPGST*A	Core 3 type-2	Dr. C. Pett
64	**P194**	TEG-PAHGVT*SAPDT*RPAPGST*A	Core 3 type-2	Dr. C. Pett
65	**P195**	TEG-PAHGVT*SAPDTRPAPGSTA	Core 3 2,3-sialyl type-2	Dr. C. Pett
66	**P196**	TEG-PAHGVTSAPDT*RPAPGSTA	Core 3 2,3-sialyl type-2	Dr. C. Pett
67	**P197**	TEG-PAHGVTSAPDTRPAPGST*A	Core 3 2,3-sialyl type-2	Dr. C. Pett
68	**P198**	TEG-PAHGVT*SAPDT*RPAPGST*A	Core 3 2,6-sialyl type-2	Dr. C. Pett
69	**P199**	TEG-PAHGVT*SAPDTRPAPGSTA	Core 3 2,6-sialyl type-2	Dr. C. Pett
70	**P200**	TEG-PAHGVTSAPDT*RPAPGSTA	Core 3 2,6-sialyl type-2	Dr. C. Pett
71	**P201**	TEG-PAHGVTSAPDTRPAPGST*A	Core 3 2,3-sialyl type-2	Dr. C. Pett
72	**P202**	TEG-PAHGVT*SA	Core 3 type-2	Dr. C. Pett
73	**P203**	TEG-PAHGVT*SAPDTRPAPGSTA	Core 1 type-1	Dr. C. Pett
74	**P204**	TEG-PAHGVTSAPDT*RPAPGSTA	Core 1 type-1	Dr. C. Pett
75	**P205**	TEG-PAHGVTSAPDTRPAPGST*A	Core 1 type-1	Dr. C. Pett
76	**P206**	TEG-PAHGVT*SAPDT*RPAPGSTA	Core 1 type-1	Dr. C. Pett
77	**P207**	TEG-PAHGVT*SAPDTRPAPGST*A	Core 1 type-1	Dr. C. Pett
78	**P208**	TEG-PAHGVTSAPDT*RPAPGST*A	Core 1 type-1	Dr. C. Pett
79	**P209**	TEG-PAHGVT*SAPDT*RPAPGST*A	Core 1 type-1	Dr. C. Pett
80	**P210**	TEG-PAHGVTSAPDTRPAPGST*A	Core 1 2,6-sialyl type-1	Dr. C. Pett
81	**P211**	TEG-PAHGVTSAPDTRPAPGST*A	Core 1 2,6-sialyl type-1	Dr. C. Pett
82	**P212**	TEG-PAHGVTSAPDT*RPAPGSTA	Core 1 2,3-sialyl type-1	Dr. C. Pett
83	**P213**	TEG-PAHGVTSAPDTRPAPGST*A	Core 1 2,3-sialyl type-1	Dr. C. Pett
84	**P214**	TEG-PAHGVT*SAPDTRPAPGSTA	Core 1 type-2	Dr. C. Pett
85	**P215**	TEG-PAHGVTSAPDT*RPAPGSTA	Core 1 type-2	Dr. C. Pett
86	**P216**	TEG-PAHGVTSAPDTRPAPGST*A	Core 1 type-2	Dr. C. Pett
87	**P217**	TEG-PAHGVT*SAPDT*RPAPGSTA	Core 1 type-2	Dr. C. Pett
88	**P218**	TEG-PAHGVT*SAPDTRPAPGST*A	Core 1 type-2	Dr. C. Pett
89	**P219**	TEG-PAHGVTSAPDT*RPAPGST*A	Core 1 type-2	Dr. C. Pett
90	**P220**	TEG-PAHGVT*SAPDT*RPAPGST*A	Core 1 type-2	Dr. C. Pett
91	**P221**	TEG-PDTRPAPGSTAPPAHGVT*SA	Core 1 type-2	Dr. C. Pett
92	**P222**	TEG-PDTRPAPGST*APPAHGVTSA	Core 1 type-2	Dr. C. Pett
93	**P223**	TEG-GST*APPAHGVTSAPDTRPA	Core 1 type-2	Dr. C. Pett
94	**P224**	TEG-PAHGVTSAPDT*RPAPGSTA	Core 1 2,6-sialyl type-2	Dr. C. Pett
95	**P225**	TEG-PAHGVTSAPDTRPAPGST*A	Core 1 2,6-sialyl type-2	Dr. C. Pett
96	**P226**	TEG-PAHGVTSAPDTRPAPGST*A	Core 1 2,6-sialyl type-2	Dr. C. Pett
97	**P227**	TEG-PAHGVT*SAPDTRPAPGSTA	Core 1 2,3-sialyl type-2	Dr. C. Pett
98	**P228**	TEG-PAHGVTSAPDT*RPAPGSTA	Core 1 2,3-sialyl type-2	Dr. C. Pett
99	**P229**	TEG-PAHGVTSAPDTRPAPGST*A	Core 1 2,3-sialyl type-2	Dr. C. Pett
100	**P230**	TEG-PAHGVT*SAPDT*RPAPGST*A	Core 1 2,3-sialyl type-2	Dr. C. Pett
101	**P231**	TEG-PAHGVT*SAPDTRPAPGSTA	Core 2 type-1 (hexa)	Dr. C. Pett
102	**P232**	TEG-PAHGVTSAPDT*RPAPGSTA	Core 2 type-1 (hexa)	Dr. C. Pett
103	**P233**	TEG-PAHGVTSAPDTRPAPGST*A	Core 2 type-1 (hexa)	Dr. C. Pett

104	**P234**	TEG-PAHGVT*SAPDT*RPAPGSTA	Core 2 type-1 (hexa)	Dr. C. Pett
105	**P235**	TEG-PAHGVT*SAPDTRPAPGST*A	Core 2 type-1 (hexa)	Dr. C. Pett
106	**P236**	TEG-PAHGVTSAPDT*RPAPGST*A	Core 2 type-1 (hexa)	Dr. C. Pett
107	**P237**	TEG-PAHGVT*SAPDT*RPAPGST*A	Core 2 type-1 (hexa)	Dr. C. Pett
108	**P238**	TEG-PAHGVTSAPDTRPAPGST*A	Core 2 2,6-sialyl type-1 (hexa)	Dr. C. Pett
109	**P239**	TEG-PAHGVT*SAPDTRPAPGSTA	Core 2 type-2 (hexa)	Dr. C. Pett
110	**P240**	TEG-PAHGVTSAPDT*RPAPGSTA	Core 2 type-2 (hexa)	Dr. C. Pett
111	**P241**	TEG-PAHGVTSAPDTRPAPGST*A	Core 2 type-2 (hexa)	Dr. C. Pett
112	**P242**	TEG-PAHGVT*SAPDT*RPAPGSTA	Core 2 type-2 (hexa)	Dr. C. Pett
113	**P243**	TEG-PAHGVT*SAPDTRPAPGST*A	Core 2 type-2 (hexa)	Dr. C. Pett
114	**P244**	TEG-PAHGVTSAPDT*RPAPGST*A	Core 2 type-2 (hexa)	Dr. C. Pett
115	**P245**	TEG-PAHGVT*SAPDT*RPAPGST*A	Core 2 type-2 (hexa)	Dr. C. Pett
116	**P246**	TEG-PAHGVTSAPDT*RPAPGSTA	Core 2 2,6-sialyl type-2 (hexa)	Dr. C. Pett
117	**P247**	TEG-PAHGVTSAPDTRPAPGST*A	Core 2 2,6-sialyl type-2 (hexa)	Dr. C. Pett
118	**P248**	TEG-PAHGVTSAPDTRPAPGST*A	Core 2 2,6-sialyl type-2 (hexa)	Dr. C. Pett
119	**P15**	TEG-PAHGVTSAPDT*RPAPGSTA	Core 2 type-2 (tetra)	this work
120	**P249**	TEG-PAHGVTSAPDT*RPAPGSTA	Core 2 2,3-sialyl(core 1) type-2 (tetra)	Dr. C. Pett
121	**P250**	TEG-PAHGVTSAPDT*RPAPGSTA	Core 2 2,3-sialyl(LacNAc) type-2 (tetra)	Dr. C. Pett
122	**P251**	TEG-PAHGVTSAPDT*RPAPGS#T#A	*Core 1 type-2 + $^{\#}2xT_N$	Dr. C. Pett
123	**P252**	TEG-PAHGVTSAPDT*RPAPGS#T#A	*Core 3 type-2 + $^{\#}2xT_N$	Dr. C. Pett
124	**P253**	TEG-PAHGVTSAPDT*RPAPGS#T#A	*Core 2 type-2 + $^{\#}2xT_N$	Dr. C. Pett

Incubation strategies with MA3:

The wells were incubated with 100 µL lectin and antibody solution each.

Experiment 1:

well	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Lectin	AD37 20 µg/mL	AD37 20 µg/mL	AD37 200 µg/mL	AD37 200 µg/mL	AD52 20 µg/mL	AD52 20 µg/mL	AD52 200 µg/mL	AD52 200 µg/mL
Antibody	Anti-Penta-His 200 ng/mL (dilution 1:1000)	Anti-Penta-His 40 ng/mL (dilution 1:5000)	Anti-Penta-His 200 ng/mL (dilution 1:1000)	Anti-Penta-His 40 ng/mL (dilution 1:5000)	Anti-Penta-His 200 ng/mL (dilution 1:1000)	Anti-Penta-His 40 ng/mL (dilution 1:5000)	Anti-Penta-His 200 ng/mL (dilution 1:1000)	Anti-Penta-His 40 ng/mL (dilution 1:5000)

Experiment 2:

well	**1**	**2**	**3**	**4**	**5**
Lectin	Galectin-3 10 µg/mL	Galectin-3 10 µg/mL	AD37 200 µg/mL	AD37 200 µg/mL	Galectin-3 10 µg/mL
Antibody	Anti-Penta-His 1 µg/mL (dilution 1:1000)	Anti-Penta-His 4 µg/mL (dilution 1:5000)	Anti-Penta-His 2 µg/mL (dilution 1:1000)	Anti-Penta-His 2 µg/mL (dilution 1:5000)	Anti-Galectin-3 1 µg/mL (dilution 1:500)

9.3.5 Protocol for microarray binding assays

Blocking:

- The slide was bathed in an aqueous solution with 25 mM ethanolamine and 100 mM sodium borate for one hour. Afterwards it was rinsed five times with water and dried via centrifugation.

Lectin incubation:

- The slide was mounted into a slide holder with a silicon superstructure, which forms the wells.
- The lectin was diluted with incubation buffer (PBS + 0.2 % Tween + 0.1 mM $CaCl_2$) and pipetted into the wells.
- The slide holder was transferred to a humidity chamber (70 % air humidity) and slowly shaken for one hour.

Washing:

- The lectin solutions were removed and each well was washed twice with washing buffer (PBS + 0.05 % Tween + 0.1 mM $CaCl_2$) and once with PBS buffer (PBS + 0.01 mM $CaCl_2$). Thereby each washing solution was kept in the wells for 15 min under slow shaking in the humidity chamber.

Incubation with secondary conjugates or antibodies:

- The conjugate or antibodies were diluted in incubation buffer and pipetted into the wells.
- Incubation was done for one hour while shaking slowly in the humidity chamber.

Washing and readout

- Washing was done as described in step “washing”.
- The slide was demounted from the slide holder.
- The slide was rinsed five times with water and dried via centrifugation.
- The slides were scanned with a fluorescence scanner.

9.4 HCD-based oxonium ion profiling

9.4.1 Profiling with single collision energy on MUC1 glycopeptides

Glycopeptide solutions of **P2**, **P94** and **P102** respectively were prepared with 400 fmol analyte in 15 µl of 0.1 % (w/v) formic acid each and subjected to LC-MS/MS using an *Orbitrap Velos Pro* mass spectrometer coupled in line to a nano RSLC HPLC system (both systems from *Thermo Scientific*). LC separation was performed using a linear gradient from 3 – 42 % buffer D (C/D system, compare chapter 9.2.2) in 70 min. Samples were measured in data-dependent acquisition mode by acquiring full MS scans at a resolution of 60,000, an m/z range of 300 – 2,000 , a target value of 1 x 10^6 and a maximum injection time of 100 ms. The seven most abundant ions were selected for HCD fragmentation with a resolution of 6,500, an m/z range of 100 – 1,060, a target value of 0.2 x 10^6, an isolation window of 2.0, a dynamic exclusion of 10 seconds and a maximum injection time of 100 ms at a normalized collision energy at 30 %.

9.4.2 Profiling with stepped collision energy on MUC1 glycopeptides

Glycopeptide solutions of **P2**, **P94**, **P96** and **P102** respectively were prepared with 400 fmol analyte in 15 µl of 0.1 % (w/v) formic acid each and subjected to LC-MS/MS using an *Orbitrap Fusion* mass spectrometer coupled in line to a nano RSLC HPLC system (both systems from *Thermo Scientific*). LC separation was performed using a linear gradient from 3 – 35 % buffer D (C/D system, compare chapter 9.2.2) in 30 min. Samples were measured in data-dependent acquisition mode by acquiring full MS scans at a resolution of 120,000, an m/z range of 300 – 1,500 and a target value of 200,000 with a maximum injection time of 50 ms. The ten most abundant ions were selected for HCD fragmentation with a resolution of 30,000, an m/z range of 100 – 1,500, a target value of 50,000, an isolation window of 1.6 and a maximum injection time of 40 ms. Each mother ion was subjected to nine consecutive HCD fragmentation scans with individual collision energies from 10 to 50 % in increments of 5 % respectively.

9.4.3 Profiling with stepped collision energy on tryptic glycopeptides

The glycopeptide analytes were grouped according to their HexNAc glycoforms and dissolved in solutions of 0.1 % (w/v) formic acid (Pool 1- 3). In order to remeasure glycopeptides with weak or unstable ion signals, higher concentrated solutions of these glycopeptides have been generated (Pool 4 and Aβ 1- 3). The resulting glycopeptide pools are shown in ***Table 8***. Each glycopeptide solution contained a total concentration of 400 fmol/15 µL with equal distribution over all included analytes.

Table 8: Composition of analyte pools used for HCD fragmentation.

Protein segment	sequence	Pool 1 (αGalNAc)	Pool 2 (αGlcNAc)	Pool 3 (βGlcNAc)	Pool 4 (αGlcNAc)	Aβ 1	Aβ 2	Aβ 3
Aβ 672-687	TEG-DAEFRHDSGY*EVHHQK	P36	P37	P38	---	P36	P37	P38
AspAT 91 - 107	TEG-NLDKEY*LPIGGLAEFCK	P41	P42	P43	P42	---	---	---
Atp5b 407 - 422	TEG-IMDPNIVGNEHY*DVAR	P76	P77	P78	P76	---	---	---
Cdc42 27 – 38	TEG-KFPSEY*VPTVFD	P89	P90	P91	---	---	---	---
MUC1 VNTR	TEG-PAHGVTSAPDTRPAPGSY*A	P95	P98	P103	---	---	---	---
NucB2 387 - 399	TEG-LEY*HQVIQQMEQK	P105	---	P106	---	---	---	---
Rac1 27 - 38	TEG-AFPGEY*IPTVFD	P110	P111	P112	---	---	---	---
RhoA 29 - 40	TEG-QFPEVY*VPTVFE	P120	P121	P122	---	---	---	---

The analyte solutions were subjected to LC-MS/MS using an *Orbitrap Fusion Lumos* mass spectrometer coupled in line to a nano RSLC HPLC system (both systems from *Thermo Scientific*). LC separation was performed using a linear gradient from 3 – 45 % buffer D (C/D system, compare chapter 9.2.2) in 90 min. Samples were measured in data-dependent acquisition mode by acquiring full MS scans at a resolution of 120,000, an m/z range of 300 – 1,500 and a target value of 200,000 with a maximum injection time of 50 ms. The ten most abundant ions were selected for HCD fragmentation with a resolution of 30,000, an m/z range of 100 – 1,500, a target value of 50,000, an isolation window of 1.6 and a maximum injection time of 40 ms. Each mother ion was subjected to nine consecutive HCD fragmentation scans with individual collision energies from 10 to 50 % in increments of 5 % respectively.

9.5 Enzyme incubation experiments

Used Enzymes:

- *O*-GlcNAcase from *Streptococcus pyogenes* (*Prozomix* Ltd., United Kingdom), 9.27 mg/mL in 3.2 M ammonium sulfate. Specific activity: 5.56 U/mg at 37° C; *p*H 7.6; pNP-N-acetyl-β-D-glucosaminide (1 mM).

- Recombinant O-GlcNAcase (human) expressed in *E. coli*, obtained from Prof. Dr. *Jennifer Kohler* (Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas), 431 nM solution in PBS, *p*H 7.4. Specific activity: 0.17 mU/mg at room temperature and *p*H 6.5; 1 mM pNP-N-acetyl-β-D-glucosaminide in a 50 mM sodium cacodylate buffer including 5 mg/ml BSA and 4 mM GalNAc

- Recombinant *O*-GlcNAc transferase (human) expressed in *E. coli*, obtained from Prof. Dr. *Jennifer Kohler* (Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas), 8.1 µM solution in PBS + 10 % glycerol, *p*H 7.0.

Analysis:

LC-MS analysis was performed with a *Phenomenex Luna* C18 column (150 x 2 mm, 3 µm particle size) inline coupled with a *MSQ Plus* mass spectrometer (*Thermo Scientific*) at a flow rate of 200 µL/min at λ = 214 nm (peptide bonds). MALDI-TOF analysis was performed using an *UltrafleXtreme* mass spectrometer (*Bruker*).

9.5.1 OGA-incubations

Each batch was prepared as 50 µL of a 1 mM glycopeptide solution containing 25 mU OGA. To this extent, 25 µL of a 2 mM glycopeptide solution in water were added to 22.6 µL 20 mM HEPES buffer (*p*H 7.4) in a 500 µL safe-lock eppendorf tube. The OGA stock solution was diluted 1:10 with HEPES buffer and 2.4 µL of the diluted solution was pipetted to the glycopeptide/buffer solution. The system was kept for 24 h in a thermocycler at 37° C while being shaken. For monitoring via HPLC-MS, 2 µL aliquots were taken and diluted with 68 µL of 20 % buffer D.

9.5.2 hOGA incubations

First generation enzyme concentration test:

A buffer of 50 mM sodium cacodylate, 5 mg/mL BSA and 4 mM GalNAc was prepared and the *p*H was adjusted to 6.5 using hydrochloric acid. Further, a glycopeptide stock solutions was prepared containing 2.5 mM **P103** in sodium cacodylate buffer. Each batch was prepared by mixing enzyme stock solution with glycopeptide stock solution in a 500 µL safe-lock eppendorf tube, while filling up the effective volume to 50 µl by addition of more cacodylate buffer. Afterwards the *p*H was adjusted to 6.5 using 3.7 % hydrochloric acid. The pipetted volumes are shown in ***Table 9***. The reaction mixture was shaken at room temperature for 24 h, 2 µL aliquots were taken and diluted with 68 µL 20 % buffer D at each measured time point for reaction monitoring via HPLC-MS.

Table 9: Compositions of the test batches with **P103** (TEG-PAHGVTSAPDTRPAPGSY*A) for the evaluation of different enzyme concentrations.

target enzyme concentration	effective activity (in 50 µl)	enzyme solution (431 nM)	glycopeptide solution (2.5 mM)	cacodylate buffer	3.7 % HCl (to adjust *p*H to 6.5)
50 nM	49 µU	5.8 µl	20.0 µl	24.2 µl	---
250 nM	247 µU	29.0 µl	20.0 µl	1.0 µl	0.1 µl

Second generation hOGT incubation test:

First, 10 mM glycopeptide stock solutions in 50 mM sodium cacodylate, 50 mg/mL BSA and 40 mM GalNAc were prepared (for glycopeptides **P78**, **P103** and **P125**). Afterwards, each batch was prepared by adding 5 µL of this glycopeptide stock solution to 45 µL of the 431 nM hOGT stock solution in a 500 µL safe-lock eppendorf tube. The composition of the stock solutions and the effective conditions are shown in ***Table 10***. Finally, the *p*H of the solutions was adjusted to 6.5 by adding 0.2 µL 3.7 % hydrochloric acid to each batch. The reaction mixtures were shaken at room temperature for 24 h, 2 µL aliquots were taken and diluted with 68 µL 20 % buffer D at each measured time point for reaction monitoring via HPLC-MS.

Table 10: Composition of the stock and reaction solutions in the 2nd generation of hOGA incubation experiments.

component	enzyme solution (45 µl)	glycopeptide solution (5 µl)	effective reaction solution (50 µl)
hOGA	431 nM	---	388 nM
substrate candidate	---	10 mM	1.0 mM
BSA	---	50 mg/ml	5 mg/ml
GalNAc	---	40 mM	4.0 mM

9.5.3 hOGT incubations

First hOGT test incubations with tyrosine peptides:

A 20 mM Tris-HCl buffer with 12.6 mM magnesium(II) chloride, 20 mM β-mercaptoethanol, 500 µM UDP GlcNAc was prepared and the *p*H was adjusted to 8.0 by use of an aqueous sodium hydroxide solution. For each batch, 42 µL of this buffer were mixed with 2.5 µL of a 2 mM aqueous peptide solution in a 500 µL safe-lock eppendorf tube. After addition of 6.2 µL hOGT stock solution (8.1 µM) the tube was transferred to a thermocycler and incubated at 37° C for 24 h while shaking. Afterwards, 20 µL were taken from each batch and analyzed via LC-MS.

hOGT incubations with known substrates **P115** and **P118**:

In a 500 µL safe-lock eppendorf tube 42 µL of a 20 mM Tris-HCl buffer with 12.6 mM magnesium(II) chloride, 20 mM β-mercaptoethanol and 200 µM UDP GlcNAc was mixed with 2.5 µL of a 2 mM aqueous peptide stock solution. Afterwards, 6.2 µL hOGT stock solution (8.1 µM) was added and the tube was transferred to a thermocycler. The system was incubated at 37° C for 24 h, while samples of 15 µL each were taken for LC-MS monitoring after 4 and 24 h respectively.

Second hOGT test incubations with tyrosine peptides:

For the hOGT incubations with **P40**, **P71** and **P75** the same protocol was applied as for the known substrates **P115** and **P118** (see above) with the following adjustments:

- The incubation time was extended to 72 h
- The reaction monitoring was done by HPLC fractionation and subsequent MALDI-TOF analyses of the different peak eluates. To this extent a MALDI matrix solution was prepared by dissolving 20 mg 2,5-dihydroxybenzoic acid (DHB) in 1 mL buffer A/methanol (1:1, v:v). Then 0.7 µL of each fraction was mixed with 0.7 µL of matrix solution on the MALDI plate.

10 References

1. Montreuil, J., *Pure and applied Chemistry* **1975,** *42* (3), 431.
2. Carraway, K. L.; Hull, S. R., *Bioessays* **1989,** *10* (4), 117.
3. Brockhausen, I.; Schachter, H.; Stanley, P., *O-GalNAc glycans*. 2009.
4. Moreau, R.; Dausset, J.; Bernard, J.; Moullec, J., *Bulletins et mémoires de la Société médicale des hôpitaux de Paris* **1957,** *73* (20-21), 569.
5. Schwientek, T.; Bennett, E. P.; Flores, C.; Thacker, J.; Hollmann, M.; Reis, C. A.; Behrens, J.; Mandel, U.; Keck, B.; Schäfer, M. A., *J Biol Chem* **2002,** *277* (25), 22623.
6. Heise, N.; Singh, D.; Van der Wel, H.; Sassi, S. O.; Johnson, J. M.; Feasley, C. L.; Koeller, C. M.; Previato, J. O.; Mendonça-Previato, L.; West, C. M., *Glycobiology* **2009,** *19* (8), 918.
7. Wojczyk, B. S.; Stwora-Wojczyk, M. M.; Hagen, F. K.; Striepen, B.; Hang, H. C.; Bertozzi, C. R.; Roos, D. S.; Spitalnik, S. L., *Molecular and biochemical parasitology* **2003,** *131* (2), 93.
8. Wandall, H. H.; Hassan, H.; Mirgorodskaya, E.; Kristensen, A. K.; Roepstorff, P.; Bennett, E. P.; Nielsen, P. A.; Hollingsworth, M. A.; Burchell, J.; Taylor-Papadimitriou, J., *J Biol Chem* **1997,** *272* (38), 23503.
9. Gerken, T. A.; Ten Hagen, K. G.; Jamison, O., *Glycobiology* **2008,** *18* (11), 861.
10. Clausen, H.; Bennett, E. P., *Glycobiology* **1996,** *6* (6), 635.
11. Pedersen, J. W.; Bennett, E. P.; Schjoldager, K. T. B. G.; Meldal, M.; Holmer, A. P.; Blixt, O.; Clo, E.; Levery, S. B.; Clausen, H.; Wandall, H. H., *J Biol Chem* **2011,** *286* (37), 32684.
12. Imberty, A.; Piller, V.; Piller, F.; Breton, C., *Protein engineering* **1997,** *10* (12), 1353.
13. Wandall, H. H.; Irazoqui, F.; Tarp, M. A.; Bennett, E. P.; Mandel, U.; Takeuchi, H.; Kato, K.; Irimura, T.; Suryanarayanan, G.; Hollingsworth, M. A., *Glycobiology* **2007,** *17* (4), 374.
14. Brockhausen, I., *Glycoprotein Methods and Protocols: The Mucins* **2000**, 273.
15. Brockhausen, I.; Schachter, H., *Glycosciences: status and perspectives* **1997**, 79.
16. Schachter, H.; Williams, D., Biosynthesis of mucus glycoproteins. In *Mucus in Health and Disease—II*, Springer: 1982; pp 3.
17. Brockhausen, I., *New comprehensive biochemistry* **1995,** *29*, 201.
18. Kamerling, J. P., *Comprehensive Glycoscience*. Elsevier: 2007.
19. Rose, M. C.; Voynow, J. A., *Physiological reviews* **2006,** *86* (1), 245.
20. Almeida, R.; Amado, M.; Carneiro, F.; Levery, S.; Holmes, E.; Nomoto, M.; Hollingsworth, M. A.; Hassan, H.; Schwientek, T.; Nielsen, P., *J Biol Chem* **1998,** (273), 12770.
21. Isshiki, S.; Togayachi, A.; Kudo, T.; Nishihara, S.; Watanabe, M.; Kubota, T.; Kitajima, M.; Shiraishi, N.; Sasaki, K.; Andoh, T., *J Biol Chem* **1999,** *274* (18), 12499.
22. Shaper, N.; Shaper, J.; Bertness, V.; Chang, H.; Kirsch, I.; Hollis, G., *Somatic cell and molecular genetics* **1986,** *12* (6), 633.
23. Almeida, R.; Amado, M.; David, L.; Levery, S. B.; Holmes, E. H.; Merkx, G.; van Kessel, A. G.; Rygaard, E.; Hassan, H.; Bennett, E., *J Biol Chem* **1997,** *272* (51), 31979.

24. Schwientek, T.; Almeida, R.; Levery, S. B.; Holmes, E. H.; Bennett, E.; Clausen, H., *J Biol Chem* **1998,** *273* (45), 29331.

25. Sato, T.; Furukawa, K.; Bakker, H.; Van den Eijnden, D. H.; Van Die, I., *Proceedings of the National Academy of Sciences* **1998,** *95* (2), 472.

26. Sasaki, K.; Kurata-Miura, K.; Ujita, M.; Angata, K.; Nakagawa, S.; Sekine, S.; Nishi, T.; Fukuda, M., *Proceedings of the National Academy of Sciences* **1997,** *94* (26), 14294.

27. Bierhuizen, M.; Mattei, M.-G. v.; Fukuda, M., *Genes & Development* **1993,** *7* (3), 468.

28. Inaba, N.; Hiruma, T.; Togayachi, A.; Iwasaki, H.; Wang, X.-H.; Furukawa, Y.; Sumi, R.; Kudo, T.; Fujimura, K.; Iwai, T., *Blood* **2003,** *101* (7), 2870.

29. Patsos, G.; Corfield, A., *O-Glycosylation: structural diversity and functions*. Wiley-VCH: Weinheim, Germany: 2009.

30. Jin, C.; Kenny, D. T.; Skoog, E. C.; Padra, M.; Adamczyk, B.; Vitizeva, V.; Thorell, A.; Venkatakrishnan, V.; Lindén, S. K.; Karlsson, N. G., *Mol Cell Proteomics* **2017,** *16* (5), 743.

31. Hart, G. W.; Akimoto, Y., *Essentials of Glycobiology*. 2009.

32. Gendler, S. J.; Spicer, A., *Annual review of physiology* **1995,** *57* (1), 607.

33. Hanover, J.; Cohen, C.; Willingham, M.; Park, M., *J Biol Chem* **1987,** *262* (20), 9887.

34. Love, D. C.; Ghosh, S.; Mondoux, M. A.; Fukushige, T.; Wang, P.; Wilson, M. A.; Iser, W. B.; Wolkow, C. A.; Krause, M. W.; Hanover, J. A., *Proceedings of the National Academy of Sciences* **2010,** *107* (16), 7413.

35. Hanover, J. A.; Krause, M. W.; Love, D. C., *Nature reviews Molecular cell biology* **2012,** *13* (5), 312.

36. Jackson, S. P.; Tjian, R., *Cell* **1988,** *55* (1), 125.

37. Kreppel, L. K.; Blomberg, M. A.; Hart, G. W., *J Biol Chem* **1997,** *272* (14), 9308.

38. Hanover, J. A.; Yu, S.; Lubas, W. B.; Shin, S.-H.; Ragano-Caracciola, M.; Kochran, J.; Love, D. C., *Archives of biochemistry and biophysics* **2003,** *409* (2), 287.

39. Love, D. C.; Kochran, J.; Cathey, R. L.; Shin, S.-H.; Hanover, J. A., *Journal of cell science* **2003,** *116* (4), 647.

40. Ma, J.; Hart, G. W., *Clinical proteomics* **2014,** *11* (1), 8.

41. Dong, D.; Hart, G. W., *J Biol Chem* **1994,** *269* (30), 19321.

42. Marino, F.; Bern, M.; Mommen, G. P.; Leney, A. C.; van Gaans-van den Brink, J. A.; Bonvin, A. M.; Becker, C.; van Els, C. c. A.; Heck, A. J., *J Am Chem Soc* **2015,** *137* (34), 10922.

43. Jung, E.; Gooley, A. A.; Packer, N. H.; Karuso, P.; Williams, K. L., *The FEBS Journal* **1998,** *253* (2), 517.

44. Li, B.; Li, H.; Lu, L.; Jiang, J., *Nat Struct Mol Biol* **2017,** *24* (4), 362.

45. Hardivillé, S.; Hart, G. W., *Cell metabolism* **2014,** *20* (2), 208.

46. Lazarus, M. B.; Jiang, J.; Kapuria, V.; Bhuiyan, T.; Janetzko, J.; Zandberg, W. F.; Vocadlo, D. J.; Herr, W.; Walker, S., *Science* **2013,** *342* (6163), 1235.

47. Drougat, L.; Olivier-Van Stichelen, S.; Mortuaire, M.; Foulquier, F.; Lacoste, A.-S.; Michalski, J.-C.; Lefebvre, T.; Vercoutter-Edouart, A.-S., *Biochimica et Biophysica Acta (BBA)-General Subjects* **2012,** *1820* (12), 1839.

48. Lewis, B. A.; Hanover, J. A., *J Biol Chem* **2014,** *289* (50), 34440.

49. Lewis, B. A., *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **2013,** *1829* (11), 1202.

50. Zachara, N. E.; Hart, G. W., *Trends in cell biology* **2004,** *14* (5), 218.

51. Guinez, C.; Mir, A.-M.; Dehennaut, V.; Cacan, R.; Harduin-Lepers, A.; Michalski, J.-C.; Lefebvre, T., *The FASEB Journal* **2008,** *22* (8), 2901.

52. Chou, C.-F.; Smith, A. J.; Omary, M. B., *J Biol Chem* **1992,** *267* (6), 3901.

53. Roquemore, E. P.; Chevrier, M. R.; Cotter, R. J.; Hart, G. W., *Biochemistry* **1996,** *35* (11), 3578.

54. Butkinaree, C.; Park, K.; Hart, G. W., *Biochimica et Biophysica Acta (BBA)-General Subjects* **2010,** *1800* (2), 96.

55. Wang, Z.; Gucek, M.; Hart, G. W., *Proceedings of the National Academy of Sciences* **2008,** *105* (37), 13793.

56. Kelly, W. G.; Dahmus, M. E.; Hart, G. W., *J Biol Chem* **1993,** *268* (14), 10416.

57. Chou, T.-Y.; Dang, C. V.; Hart, G. W., *Proceedings of the National Academy of Sciences* **1995,** *92* (10), 4417.

58. Hart, G. W., *Annual review of biochemistry* **1997,** *66* (1), 315.

59. Haltiwanger, R. S.; Busby, S.; Grove, K.; Li, S.; Mason, D.; Medina, L.; Moloney, D.; Philipsberg, G.; Scartozzi, R., *Biochemical and biophysical research communications* **1997,** *231* (2), 237.

60. Holt, G. D.; Snow, C. M.; Senior, A.; Haltiwanger, R. S.; Gerace, L.; Hart, G. W., *The Journal of cell biology* **1987,** *104* (5), 1157.

61. Chou, T.-Y.; Hart, G. W.; Dang, C. V., *J Biol Chem* **1995,** *270* (32), 18961.

62. Comer, F. I.; Hart, G. W., *J Biol Chem* **2000,** *275* (38), 29179.

63. Marshall, S.; Rumberger, J., *Diabetes Annual* **2000,** *13*, 97.

64. Hebert Jr, L. F.; Daniels, M. C.; Zhou, J.; Crook, E. D.; Turner, R. L.; Simmons, S. T.; Neidigh, J. L.; Zhu, J.-S.; Baron, A. D.; McClain, D. A., *Journal of Clinical Investigation* **1996,** *98* (4), 930.

65. Bond, M. R.; Hanover, J. A., *The Journal of cell biology* **2015,** *208* (7), 869.

66. Ma, J. F.; Liu, T.; Wei, A. C.; Banerjee, P.; O'Rourke, B.; Hart, G. W., *J Biol Chem* **2015,** *290* (49), 29141.

67. Yi, W.; Clark, P. M.; Mason, D. E.; Keenan, M. C.; Hill, C.; Goddard, W. A.; Peters, E. C.; Driggers, E. M.; Hsieh-Wilson, L. C., *Science* **2012,** *337* (6097), 975.

68. Corfield, A. P.; Shukla, A. K., *Genomic/Proteomic Technology* **2003,** *3*, 20.

69. Perez-Vilar, J.; Hill, R. L., *J Biol Chem* **1999,** *274* (45), 31751.

70. Corfield, A. P., *Biochimica et Biophysica Acta (BBA)-General Subjects* **2015,** *1850* (1), 236.

71. Wesseling, J.; Van Der Valk, S. W.; Vos, H. L.; Sonnenberg, A.; Hilkens, J., *The Journal of cell biology* **1995,** *129* (1), 255.

72. Bramwell, M.; Wiseman, G.; Shotton, D., *Journal of cell science* **1986,** *86* (1), 249.

73. Braga, V.; Pemberton, L.; Duhig, T.; Gendler, S. J., *Development* **1992,** *115* (2), 427.

74. Schroeder, J. A.; Thompson, M. C.; Gardner, M. M.; Gendler, S. J., *J Biol Chem* **2001,** *276* (16), 13057.

75. Li, Y.; Ren, J.; Yu, W.-h.; Li, Q.; Kuwahara, H.; Yin, L.; Carraway, K. L.; Kufe, D., *J Biol Chem* **2001,** *276* (38), 35239.

76. Rahn, J. J.; Chow, J. W.; Horne, G. J.; Mah, B. K.; Emerman, J. T.; Hoffman, P.; Hugh, J. C., *Clinical and Experimental Metastasis* **2005,** *22* (6), 475.

77. Funes, M.; Miller, J. K.; Lai, C.; Carraway, K. L.; Sweeney, C., *J Biol Chem* **2006,** *281* (28), 19310.

78. Mukherjee, P.; Tinder, T.; Basu, G.; Gendler, S. J., *Journal of leukocyte biology* **2005,** *77* (1), 90.

79. Hattrup, C. L.; Gendler, S. J., *Annu. Rev. Physiol.* **2008,** *70*, 431.

80. Cheng, P.; Boucher, R.; Yankaskas, J.; Boat, T., *Cellular and Molecular Basis of Cystic Fibrosis* **1988**, 233.

81. Rose, M.; Voter, W. A.; Brown, C.; Kaufman, B., *Biochem J* **1984,** *222* (2), 371.

82. Sheehan, J.; Oates, K.; Carlstedt, I., *Biochem J* **1986,** *239* (1), 147.

83. Gendler, S. J.; Lancaster, C. A.; Taylor-Papadimitriou, J.; Duhig, T.; Peat, N.; Burchell, J.; Pemberton, L.; Lalani, E. N.; Wilson, D., *J Biol Chem* **1990,** *265* (25), 15286.

84. Gendler, S. J., *Journal of mammary gland biology and neoplasia* **2001,** *6* (3), 339.

85. Parmley, R. R.; Gendler, S. J., *Journal of Clinical Investigation* **1998,** *102* (10), 1798.

86. Hinojosa-Kurtzberg, M.; Hansson, G.; Gendler, S., *Pediatr Pulmonol* **2000,** *30* (Suppl.), 220.

87. Fontenot, J. D.; Tjandra, N.; Bu, D.; Ho, C.; Montelaro, R. C.; Finn, O. J., *Cancer research* **1993,** *53* (22), 5386.

88. Dekker, J.; Rossen, J. W.; Büller, H. A.; Einerhand, A. W., *Trends in biochemical sciences* **2002,** *27* (3), 126.

89. Rose, M. C., *American Journal of Physiology-Lung Cellular and Molecular Physiology* **1992,** *263* (4), L413.

90. Baeckström, D.; Hansson, G. C.; Nilsson, O.; Johansson, C.; Gendler, S. J.; Lindholm, L., *J Biol Chem* **1991,** *266* (32), 21537.

91. Jonckheere, N.; Van Seuningen, I., *Biochimie* **2010,** *92* (1), 1.

92. Bara, J.; Chastre, E.; Mahiou, J.; Singh, R. L.; Forgue-Lafitte, M. E.; Hollande, E.; Godeau, F., *International journal of cancer* **1998,** *75* (5), 767.

93. Kim, G. E.; Bae, H. I.; Park, H. U.; Kuan, S. F.; Crawley, S. C.; Ho, J. J.; Kim, Y. S., *Gastroenterology* **2002,** *123* (4), 1052.

94. Baldus, S. E.; Mönig, S. P.; Arkenau, V.; Hanisch, F.-G.; Schneider, P. M.; Thiele, J.; Hölscher, A. H.; Dienes, H. P., *Annals of surgical oncology* **2002,** *9* (9), 887.

95. Machado, J. C.; Nogueira, A. M.; Carneiro, F.; Reis, C. A.; Sobrinho-Simões, M., *The Journal of pathology* **2000,** *190* (4), 437.

96. Henke, M. O.; Renner, A.; Huber, R. M.; Seeds, M. C.; Rubin, B. K., *American journal of respiratory cell and molecular biology* **2004,** *31* (1), 86.

97. Kirkham, S.; Kolsum, U.; Rousseau, K.; Singh, D.; Vestbo, J.; Thornton, D. J., *American journal of respiratory and critical care medicine* **2008,** *178* (10), 1033.

98. Van den Brink, G.; Tytgat, K.; Van der Hulst, R.; Van der Loos, C.; Einerhand, A.; Büller, H.; Dekker, J., *Gut* **2000,** *46* (5), 601.

99. Hollingsworth, M. A.; Swanson, B. J., *Nature Reviews Cancer* **2004,** *4* (1), 45.

100. Van de Bovenkamp, J. H.; Mahdavi, J.; Male, K. V.; Anita, M.; Büller, H. A.; Einerhand, A. W.; Borén, T.; Dekker, J., *Helicobacter* **2003,** *8* (5), 521.

101. Folkerts, G.; Busse, W. W.; Nijkamp, F. P.; Sorkness, R.; Gern, J. E., *American journal of respiratory and critical care medicine* **1998,** *157* (6), 1708.

102. Robinson, D. S.; Hamid, Q.; Ying, S.; Tsicopoulos, A.; Barkans, J.; Bentley, A. M.; Corrigan, C.; Durham, S. R.; Kay, A. B., *New England Journal of Medicine* **1992,** *326* (5), 298.

103. Holt, P. G.; Strickland, D. H., *Journal of Allergy and Clinical Immunology* **2010,** *125* (5), 963.

104. Riordan, J. R.; Rommens, J. M.; Kerem, B.-s.; Alon, N.; Rozmahel, R.; Grzelczak, Z.; Zielenski, J.; Lok, S.; Plavsic, N.; Chou, J.-L., *Science* **1989,** *245* (4922), 1066.

105. Laniado-Laborín, R., *International Journal of Environmental Research and Public Health* **2009,** *6* (1), 209.

106. Kennedy, S. M.; Chambers, R.; Du, W.; Dimich-Ward, H., *Proceedings of the American Thoracic Society* **2007,** *4* (8), 692.

107. Cheng, P.-W.; Boat, T. F.; Cranfill, K.; Yankaskas, J. R.; Boucher, R. C., *Journal of Clinical Investigation* **1989,** *84* (1), 68.

108. RousselS, P.; Lamblinh, G., *Ile* **1994,** *23*, 2.55.

109. Rhim, A. D.; Kothari, V. A.; Park, P. J.; Mulberg, A. E.; Glick, M. C.; Scanlin, T. F., *Glycoconjugate journal* **2000,** *17* (6), 385.

110. Jiang, X.; Hill, W. G.; Pilewski, J. M.; Weisz, O. A., *American Journal of Physiology-Lung Cellular and Molecular Physiology* **1997,** *273* (5), L913.

111. Leir, S.-H.; Parry, S.; Palmai-Pallag, T.; Evans, J.; Morris, H. R.; Dell, A.; Harris, A., *American journal of respiratory cell and molecular biology* **2005,** *32* (5), 453.

112. Davril, M.; Degroote, S.; Humbert, P.; Galabert, C.; Dumur, V.; Lafitte, J.-J.; Lamblin, G.; Roussel, P., *Glycobiology* **1999,** *9* (3), 311.

113. Shori, D.; Genter, T.; Hansen, J.; Koch, C.; Wyatt, H.; Kariyawasam, H.; Knight, R.; Hodson, M.; Kalogeridis, A.; Tsanakas, I., *Pflügers Archiv European Journal of Physiology* **2001,** *443*, S55.

114. Scharfman, A.; Delmotte, P.; Beau, J.; Lamblin, G.; Roussel, P.; Mazurier, J., *Glycoconjugate journal* **2000,** *17* (10), 735.

115. Krivan, H. C.; Roberts, D. D.; Ginsburg, V., *Proceedings of the National Academy of Sciences* **1988,** *85* (16), 6157.

116. Hahn, H. P., *Gene* **1997,** *192* (1), 99.

117. Topin, J.; Arnaud, J.; Sarkar, A.; Audfray, A.; Gillon, E.; Perez, S.; Jamet, H.; Varrot, A.; Imberty, A.; Thomas, A., *Plos One* **2013,** *8* (8), e71149.

118. Sampaio, S. C.; Luiz, W. B.; Vieira, M. A.; Ferreira, R. C.; Garcia, B. G.; Sinigaglia-Coimbra, R.; Sampaio, J. L.; Ferreira, L. C.; Gomes, T. A., *Infection and immunity* **2016,** *84* (4), 1112.

119. Milla, C. E.; Chmiel, J. F.; Accurso, F. J.; VanDevanter, D. R.; Konstan, M. W.; Yarranton, G.; Geller, D. E., *Pediatric pulmonology* **2014,** *49* (7), 650.

120. Shivshankar, P.; Sanchez, C.; Rose, L. F.; Orihuela, C. J., *Molecular microbiology* **2009,** *73* (4), 663.

121. Thamadilok, S.; Roche-Håkansson, H.; Håkansson, A. P.; Ruhl, S., *Molecular oral microbiology* **2016,** *31* (2), 175.

122. Soong, G.; Muir, A.; Gomez, M. I.; Waks, J.; Reddy, B.; Planet, P.; Singh, P. K.; Kanetko, Y.; Wolfgang, M. C.; Hsiao, Y.-S., *The Journal of clinical investigation* **2006,** *116* (8), 2297.

123. Xu, G.; Ryan, C.; Kiefel, M. J.; Wilson, J. C.; Taylor, G. L., *Journal of molecular biology* **2009,** *386* (3), 828.

124. Lillehoj, E. P.; Hyun, S. W.; Liu, A.; Guang, W.; Verceles, A. C.; Luzina, I. G.; Atamas, S. P.; Kim, K. C.; Goldblum, S. E., *J Biol Chem* **2015,** *290* (30), 18316.

125. Parker, D.; Soong, G.; Planet, P.; Brower, J.; Ratner, A. J.; Prince, A., *Infection and immunity* **2009,** *77* (9), 3722.

126. King, S., *Molecular oral microbiology* **2010,** *25* (1), 15.

127. Lewis, A. L.; Lewis, W. G., *Cellular microbiology* **2012,** *14* (8), 1174.

128. Brockhausen, I., *EMBO reports* **2006,** *7* (6), 599.

129. Varki, A., *Trends in molecular medicine* **2008,** *14* (8), 351.

130. Cazet, A.; Julien, S.; Bobowski, M.; Burchell, J.; Delannoy, P., *Breast cancer research* **2010,** *12* (3), 204.

131. Burchell, J.; Gendler, S.; Taylor-Papadimitriou, J.; Girling, A.; Lewis, A.; Millis, R.; Lamport, D., *Cancer research* **1987,** *47* (20), 5476.

132. Perey, L.; Hayes, D. F.; Kufe, D., *Cancer research* **1992,** *52* (22), 6365.

133. Remmers, N.; Anderson, J. M.; Linde, E. M.; DiMaio, D. J.; Lazenby, A. J.; Wandall, H. H.; Mandel, U.; Clausen, H.; Yu, F.; Hollingsworth, M. A., *Clinical Cancer Research* **2013,** *19* (8), 1981.

134. Wang, Y.; Ju, T.; Ding, X.; Xia, B.; Wang, W.; Xia, L.; He, M.; Cummings, R. D., *Proceedings of the National Academy of Sciences* **2010,** *107* (20), 9228.

135. Ju, T.; Cummings, R. D., *Nature* **2005,** *437* (7063), 1252.

136. Gill, D. J.; Clausen, H.; Bard, F., *Trends in cell biology* **2011,** *21* (3), 149.

137. Taylor-Papadimitriou, J.; Burchell, J., *Mucins and cancer*. 2013.

138. Hirabayashi, J.; Hashidate, T.; Arata, Y.; Nishi, N.; Nakamura, T.; Hirashima, M.; Urashima, T.; Oka, T.; Futai, M.; Muller, W. E., *Biochimica et Biophysica Acta (BBA)-General Subjects* **2002,** *1572* (2), 232.

139. Viguier, M.; Advedissian, T.; Delacour, D.; Poirier, F.; Deshayes, F., *Tissue barriers* **2014,** *2* (3), e29103.

140. Ahmed, H.; Guha, P.; Kaptan, E.; Bandyopadhyaya, G., *Trends in carbohydrate research* **2011,** *3* (2), 13.

141. Fukumori, T.; Takenaka, Y.; Yoshii, T.; Kim, H.-R. C.; Hogan, V.; Inohara, H.; Kagawa, S.; Raz, A., *Cancer research* **2003,** *63* (23), 8302.

142. Suzuki, Y.; Inoue, T.; Yoshimaru, T.; Ra, C., *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **2008,** *1783* (5), 924.

143. Guha, P.; Kaptan, E.; Bandyopadhyaya, G.; Kaczanowska, S.; Davila, E.; Thompson, K.; Martin, S. S.; Kalvakolanu, D. V.; Vasta, G. R.; Ahmed, H., *Proceedings of the National Academy of Sciences* **2013,** *110* (13), 5052.

144. Peng, W.; Wang, H. Y.; Miyahara, Y.; Peng, G.; Wang, R.-F., *Cancer research* **2008,** *68* (17), 7228.

145. Suzuki, O.; Abe, M., *Oncology reports* **2008,** *19* (3), 743.

146. Xue, J.; Gao, X.; Fu, C.; Cong, Z.; Jiang, H.; Wang, W.; Chen, T.; Wei, Q.; Qin, C., *Febs Lett* **2013,** *587* (24), 3986.

147. Zhuo, Y.; Chammas, R.; Bellis, S. L., *J Biol Chem* **2008,** *283* (32), 22177.

148. Bresalier, R. S.; Byrd, J. C.; Wang, L.; Raz, A., *Cancer research* **1996,** *56* (19), 4354.

149. Yu, L.-G.; Andrews, N.; Zhao, Q.; McKean, D.; Williams, J. F.; Connor, L. J.; Gerasimenko, O. V.; Hilkens, J.; Hirabayashi, J.; Kasai, K., *J Biol Chem* **2007,** *282* (1), 773.

150. Matarrese, P.; Fusco, O.; Tinari, N.; Natoli, C.; Liu, F. T.; Semeraro, M. L.; Malorni, W.; Iacobelli, S., *International journal of cancer* **2000,** *85* (4), 545.

151. Frisch, S. M.; Screaton, R. A., *Current opinion in cell biology* **2001,** *13* (5), 555.

152. Zhao, Q.; Barclay, M.; Hilkens, J.; Guo, X.; Barrow, H.; Rhodes, J. M.; Yu, L.-G., *Molecular cancer* **2010,** *9* (1), 154.

153. Pett, C. Synthesis and Development of a MUC1 Glycopeptide Microarray System for Evaluation of Protein-Glycopeptide and Protein-Carbohydrate Interactions. 2015.

154. Barrow, H.; Guo, X.; Wandall, H. H.; Pedersen, J. W.; Fu, B.; Zhao, Q.; Chen, C.; Rhodes, J. M.; Yu, L.-G., *Clinical cancer research* **2011,** *17* (22), 7035.

155. Ramasamy, S.; Duraisamy, S.; Barbashov, S.; Kawano, T.; Kharbanda, S.; Kufe, D., *Molecular cell* **2007,** *27* (6), 992.

156. Blixt, O.; Bueti, D.; Burford, B.; Allen, D.; Julien, S.; Hollingsworth, M.; Gammerman, A.; Fentiman, I.; Taylor-Papadimitriou, J.; Burchell, J. M., *Breast Cancer Research* **2011,** *13* (2), R25.

157. Graves, C.; Robertson, J.; Murray, A.; Price, M.; Chapman, C., *Chemical Biology & Drug Design* **2005,** *66* (6), 357.

158. Rabinovich, G. A.; Van Kooyk, Y.; Cobb, B. A., *Annals of the New York Academy of Sciences* **2012,** *1253* (1), 1.

159. Cheever, M. A.; Allison, J. P.; Ferris, A. S.; Finn, O. J.; Hastings, B. M.; Hecht, T. T.; Mellman, I.; Prindiville, S. A.; Viner, J. L.; Weiner, L. M., *Clinical Cancer Research* **2009,** *15* (17), 5323.

160. Carbone, F. R.; Gleeson, P. A., *Glycobiology* **1997,** *7* (6), 725.

161. Beuvery, E.; Van Rossum, F.; Nagel, J., *Infection and immunity* **1982,** *37* (1), 15.

162. Ragupathi, G.; Slovin, S. F.; Adluri, S.; Sames, D.; Kim, I. J.; Kim, H. M.; Spassova, M.; Bornmann, W. G.; Lloyd, K. O.; Scher, H. I., *Angewandte Chemie International Edition* **1999,** *38* (4), 563.

163. Kunz, H.; Birnbach, S., *Angewandte Chemie International Edition* **1986,** *25* (4), 360.

164. Ragupathi, G.; Koide, F.; Livingston, P. O.; Cho, Y. S.; Endo, A.; Wan, Q.; Spassova, M. K.; Keding, S. J.; Allen, J.; Ouerfelli, O., *J Am Chem Soc* **2006,** *128* (8), 2715.

165. Zhu, J.; Wan, Q.; Lee, D.; Yang, G.; Spassova, M. K.; Ouerfelli, O.; Ragupathi, G.; Damani, P.; Livingston, P. O.; Danishefsky, S. J., *J Am Chem Soc* **2009,** *131* (26), 9298.

166. Kaiser, A.; Gaidzik, N.; Westerlind, U.; Kowalczyk, D.; Hobel, A.; Schmitt, E.; Kunz, H., *Angewandte Chemie International Edition* **2009,** *48* (41), 7551.

167. Herzenberg, L. A.; Tokuhisa, T.; Herzenberg, L. A., *Nature* **1980,** *285* (5767), 664.

168. Cai, H.; Chen, M. S.; Sun, Z. Y.; Zhao, Y. F.; Kunz, H.; Li, Y. M., *Angewandte Chemie International Edition* **2013,** *52* (23), 6106.

169. Cai, H.; Sun, Z. Y.; Chen, M. S.; Zhao, Y. F.; Kunz, H.; Li, Y. M., *Angewandte Chemie International Edition* **2014,** *53* (6), 1699.

170. Dziadek, S.; Hobel, A.; Schmitt, E.; Kunz, H., *Angewandte Chemie International Edition* **2005,** *44* (46), 7630.

171. Alexander, J.; Sidney, J.; Southwood, S.; Ruppert, J.; Oseroff, C.; Maewal, A.; Snoke, K.; Serra, H. M.; Kubo, R. T.; Sette, A., *Immunity* **1994,** *1* (9), 751.

172. Cremer, G. A.; Bureaud, N.; Piller, V.; Kunz, H.; Piller, F.; Delmas, A. F., *ChemMedChem* **2006,** *1* (9), 965.

173. Spohn, R.; Buwitt-Beckmann, U.; Brock, R.; Jung, G.; Ulmer, A. J.; Wiesmüller, K.-H., *Vaccine* **2004,** *22* (19), 2494.

174. Kaiser, A.; Gaidzik, N.; Becker, T.; Menge, C.; Groh, K.; Cai, H.; Li, Y. M.; Gerlitzki, B.; Schmitt, E.; Kunz, H., *Angewandte Chemie International Edition* **2010,** *49* (21), 3688.

175. Gaidzik, N.; Westerlind, U.; Kunz, H., *Chem Soc Rev* **2013,** *42* (10), 4421.

176. Becker, T.; Dziadek, S.; Wittrock, S.; Kunz, H., *Current cancer drug targets* **2006,** *6* (6), 491.

177. Halim, A.; Brinkmalm, G.; Ruetschi, U.; Westman-Brinkmalm, A.; Portelius, E.; Zetterberg, H.; Blennow, K.; Larson, G.; Nilsson, J., *P Natl Acad Sci USA* **2011,** *108* (29), 11848.

178. Masters, C. L.; Simms, G.; Weinman, N. A.; Multhaup, G.; McDonald, B. L.; Beyreuther, K., *Proceedings of the National Academy of Sciences* **1985,** *82* (12), 4245.

179. Glenner, G. G.; Wong, C. W., *Biochemical and biophysical research communications* **1984,** *120* (3), 885.

180. Denman, R. B.; Rosenzcwaig, R.; Miller, D. L., *Biochemical and biophysical research communications* **1993,** *192* (1), 96.

181. Steentoft, C.; Vakhrushev, S. Y.; Vester-Christensen, M. B.; Schjoldager, K. T. B. G.; Kong, Y.; Bennett, E. P.; Mandel, U.; Wandall, H.; Levery, S. B.; Clausen, H., *Nat Methods* **2011,** *8* (11), 977.

182. Lin, P.; Le-Niculescu, H.; Hofmeister, R.; McCaffery, J. M.; Jin, M.; Hennemann, H.; McQuistan, T.; De Vries, L.; Farquhar, M. G., *The Journal of cell biology* **1998,** *141* (7), 1515.

183. Nesselhut, J.; Jurgan, U.; Onken, E.; Götz, H.; Barnikol, H. U.; Hirschfeld, G.; Barnikol-Watanabe, S.; Hilschmann, N., *Febs Lett* **2001,** *509* (3), 469.

184. Oh, S.; Shimizu, H.; Satoh, T.; Okada, S.; Adachi, S.; Inoue, K.; Eguchi, H.; Yamamoto, M.; Imaki, T.; Hashimoto, K., *Nature* **2006,** *443* (7112), 709.

185. Mohan, H.; Ramesh, N.; Mortazavi, S.; Le, A.; Iwakura, H.; Unniappan, S., *Plos One* **2014,** *9* (12), e115102.

186. Suzuki, S.; Takagi, K.; Miki, Y.; Onodera, Y.; Akahira, J. i.; Ebata, A.; Ishida, T.; Watanabe, M.; Sasano, H.; Suzuki, T., *Cancer science* **2012,** *103* (1), 136.

187. Vakhrushev, S. Y.; Steentoft, C.; Vester-Christensen, M. B.; Bennett, E. P.; Clausen, H.; Levery, S. B., *Mol Cell Proteomics* **2013,** *12* (4), 932.

188. Trinidad, J. C.; Schoepfer, R.; Burlingame, A. L.; Medzihradszky, K. F., *Mol Cell Proteomics* **2013,** *12* (12), 3474.

189. Yu, J.; Schorlemer, M.; Gomez Toledo, A.; Pett, C.; Sihlbom, C.; Larson, G.; Westerlind, U.; Nilsson, J., *Chemistry* **2016,** *22* (3), 1114.

190. Halim, A.; Westerlind, U.; Pett, C.; Schorlemer, M.; Ruetschi, U.; Brinkmalm, G.; Sihlbom, C.; Lengqvist, J.; Larson, G.; Nilsson, J., *Journal of proteome research* **2014,** *13* (12), 6024.

191. Jank, T.; Bogdanovic, X.; Wirth, C.; Haaf, E.; Spoerner, M.; Bohmer, K. E.; Steinemann, M.; Orth, J. H. C.; Kalbitzer, H. R.; Warscheid, B.; Hunte, C.; Aktories, K., *Nat Struct Mol Biol* **2013,** *20* (11), 1273.

192. Jank, T.; Eckerle, S.; Steinemann, M.; Trillhaase, C.; Schimpl, M.; Wiese, S.; Van Aalten, D. M.; Driever, W.; Aktories, K., *Nature communications* **2015,** *6*.

193. Aktories, K., *Nat Rev Microbiol* **2011,** *9* (7), 487.

194. Jank, T.; Trillhaase, C.; Brozda, N.; Steinemann, M.; Schwan, C.; Süss, R.; Aktories, K., *The FASEB Journal* **2015,** *29* (7), 2789.

195. Saeland, E.; van Vliet, S. J.; Bäckström, M.; van den Berg, V. C. M.; Geijtenbeek, T. B. H.; Meijer, G. A.; van Kooyk, Y., *Cancer Immunology, Immunotherapy* **2007,** *56* (8), 1225.

196. FORD, E.; NELSON, K. E.; WARREN, D., *Epidemiologic reviews* **1987,** *9* (1), 244.

197. Nilsson, E. C.; Storm, R. J.; Bauer, J.; Johansson, S. M. C.; Lookene, A.; Angstrom, J.; Hedenstrom, M.; Eriksson, T. L.; Frangsmyr, L.; Rinaldi, S.; Willison, H. J.; Domellof, F. P.; Stehle, T.; Arnberg, N., *Nat Med* **2011,** *17* (1), 105.

198. Arnberg, N., *Reviews in medical virology* **2009,** *19* (3), 165.

199. Fischer, E., *Berichte der deutschen chemischen Gesellschaft* **1895,** *28* (1), 1167.

200. Lindhorst, T. K., *Essentials of Carbohydrate Chemistry and Biochemistry*. 3rd ed.; 2007.

201. Kafle, A.; Liu, J.; Cui, L., *Canadian Journal of Chemistry* **2016,** *94* (11), 894.

202. Yang, Y.; Zhang, X.; Yu, B., *Nat Prod Rep* **2015,** *32* (9), 1331.

203. Koenigs, W.; Knorr, E., *Berichte der deutschen chemischen Gesellschaft* **1901,** *34* (1), 957.

204. Schmidt, R. R.; Michel, J., *Tetrahedron Lett* **1984,** *25* (8), 821.

205. Schmidt, R. R.; Michel, J., *Angewandte Chemie-International Edition in English* **1980,** *19* (9), 731.

206. Lemieux, R.; Chu, P., *Am Chem Soc* **1958,** *133*, 31N.

207. Lemieux, R.; de Mayo, P., *by P. De Mayo, Interscience, New York* **1964**.

208. Lemieux, R., *Pure and Applied Chemistry* **1971,** *25* (3), 527.

209. Haines, A. H., *Advances in Carbohydrate Chemistry and Biochemistry* **1976,** *33*, 11.

210. Yamaguchi, M.; Ishida, H.; Galustian, C.; Feizi, T.; Kiso, M., *Carbohyd Res* **2002,** *337* (21), 2111.

211. Mootoo, D. R.; Konradsson, P.; Udodong, U.; Fraser-Reid, B., *J Am Chem Soc* **1988,** *110* (16), 5583.

212. Fraser-Reid, B.; Wu, Z.; Udodong, U. E.; Ottosson, H., *The Journal of Organic Chemistry* **1990,** *55* (25), 6068.

213. McDonnell, C.; López, O.; Murphy, P.; Fernández Bolaños, J. G.; Hazell, R.; Bols, M., *J Am Chem Soc* **2004,** *126* (39), 12374.

214. Fraser-Reid, B.; Wu, Z.; Andrews, C. W.; Skowronski, E.; Bowen, J. P., *J Am Chem Soc* **1991,** *113* (4), 1434.

215. Demchenko, A. V., *Current organic chemistry* **2003,** *7* (1), 35.

216. Demchenko, A. V., *Synlett* **2003,** *2003* (09), 1225.

217. Marra, A.; Sinaÿ, P., *Carbohyd Res* **1990,** *195* (2), 303.

218. Braccini, I.; Derouet, C.; Esnault, J.; Herv, C.; Mallet, J.-M.; Michon, V.; Sinaÿ, P., *Carbohyd Res* **1993,** *246* (1), 23.

219. Ratcliffe, A. J.; Fraser-Reid, B., *Journal of the Chemical Society, Perkin Transactions 1* **1990,** (3), 747.

220. Satoh, H.; Hansen, H. S.; Manabe, S.; van Gunsteren, W. F.; Hünenberger, P. H., *Journal of chemical theory and computation* **2010,** *6* (6), 1783.

221. Seitz, O., *Chembiochem* **2000,** *1* (4), 214.

222. Merrifield, R., *J. Am. Chem. Soc* **1963,** *85*, 2149.

223. Vuljanic, T.; Bergquist, K.-E.; Clausen, H.; Roy, S.; Kihlberg, J., *Tetrahedron* **1996,** *52* (23), 7983.

224. Kihlberg, J.; Vuljanic, T., *Tetrahedron Lett* **1993,** *34* (38), 6135.

225. Carpino, L. A.; El-Faham, A.; Albericio, F., *Tetrahedron Lett* **1994,** *35* (15), 2279.

226. Carpino, L. A.; Imazumi, H.; El-Faham, A.; Ferrer, F. J.; Zhang, C.; Lee, Y.; Foxman, B. M.; Henklein, P.; Hanay, C.; Mügge, C., *Angewandte Chemie International Edition* **2002,** *41* (3), 441.

227. Carpino, L. A.; Imazumi, H.; Foxman, B. M.; Vela, M. J.; Henklein, P.; El-Faham, A.; Klose, J.; Bienert, M., *Org Lett* **2000,** *2* (15), 2253.

228. G. Zemplén, A. K., *Ber. Dtsch. Chem. Ges. A/B* **1923,** *56*, 1705.

229. Kunz, H., *Angewandte Chemie International Edition in English* **1987,** *26* (4), 294.

230. Seebach, D., *Aldrichimica Acta* **1992,** *25*, 59.

231. Sjölin, P.; Elofsson, M.; Kihlberg, J., *The Journal of organic chemistry* **1996,** *61* (2), 560.

232. Endo, T.; Koizumi, S., *Current opinion in structural biology* **2000,** *10* (5), 536.

233. Koizumi, S.; Endo, T.; Tabata, K.; Ozaki, A., *Nature Biotechnology* **1998,** *16* (9), 847.

234. de Vries, T.; van den Eijnden, D. H.; Schultz, J.; O'Neill, R., *Febs Lett* **1993,** *330* (3), 243.

235. Ichikawa, Y.; Lin, Y. C.; Dumas, D. P.; Shen, G. J.; Garcia-Junceda, E.; Williams, M. A.; Bayer, R.; Ketcham, C.; Walker, L. E., *J Am Chem Soc* **1992,** *114* (24), 9283.

236. Ajisaka, K.; Nishida, H.; Fujimoto, H., *Biotechnology letters* **1987,** *9* (4), 243.

237. Rajnochová, E.; Dvorakova, J.; Kren, V., *Biotechnology letters* **1997,** *19* (9), 869.

238. Ajisaka, K.; Nishida, H.; Fujimoto, H., *Biotechnology letters* **1987,** *9* (6), 387.

239. Rabate, M., *Bull. soc. chim. biol* **1935,** *17*, 572.

240. Hedbys, L.; Larsson, P.-O.; Mosbach, K.; Svensson, S., *Biochemical and biophysical research communications* **1984,** *123* (1), 8.

241. Kamerling, J. P.; Boons, G.-J., **2007**.

242. Tang, P.; Gool, H.; Hardy, M.; Lee, Y.; Felzi, T., *Biochemical and biophysical research communications* **1985,** *132* (2), 474.

243. Liang, R.; Yan, L.; Loebach, J.; Ge, M.; Uozumi, Y.; Sekanina, K.; Horan, N.; Gildersleeve, J.; Thompson, C.; Smith, A., *Science* **1996,** *274* (5292), 1520.

244. Roy, R., *Current opinion in structural biology* **1996,** *6* (5), 692.

245. Park, S.; Shin, I., *Angewandte Chemie International Edition* **2002,** *41* (17), 3180.

246. Wang, D.; Liu, S.; Trummer, B. J.; Deng, C.; Wang, A., *Nature biotechnology* **2002,** *20* (3), 275.

247. Houseman, B. T.; Mrksich, M., *Chem Biol* **2002,** *9* (4), 443.

248. Willats, W. G.; Rasmussen, S. E.; Kristensen, T.; Mikkelsen, J. D.; Knox, J. P., *Proteomics* **2002,** *2* (12), 1666.

249. Fazio, F.; Bryan, M. C.; Blixt, O.; Paulson, J. C.; Wong, C. H., *J Am Chem Soc* **2002,** *124* (48), 14397.

250. Fukui, S.; Feizi, T.; Galustian, C.; Lawson, A. M.; Chai, W., *Nat Biotechnol* **2002,** *20* (10), 1011.

251. Mammen, M.; Choi, S.-K.; Whitesides, G. M., *Angewandte Chemie International Edition* **1998,** *37* (20), 2754.

252. Park, S.; Gildersleeve, J. C.; Blixt, O.; Shin, I., *Chem Soc Rev* **2013,** *42* (10), 4310.

253. Galanina, O.; Mecklenburg, M.; Nifantiev, N.; Pazynina, G.; Bovin, N., *Lab on a Chip* **2003,** *3* (4), 260.

254. Guo, Y.; Feinberg, H.; Conroy, E.; Mitchell, D. A.; Alvarez, R.; Blixt, O.; Taylor, M. E.; Weis, W. I.; Drickamer, K., *Nat Struct Mol Biol* **2004,** *11* (7), 591.

255. Bochner, B. S.; Alvarez, R. A.; Mehta, P.; Bovin, N. V.; Blixt, O.; White, J. R.; Schnaar, R. L., *J Biol Chem* **2005,** *280* (6), 4307.

256. Godula, K.; Bertozzi, C. R., *J Am Chem Soc* **2010,** *132* (29), 9963.

257. Blixt, O.; Cló, E.; Nudelman, A. S.; Sørensen, K. K.; Clausen, T.; Wandall, H. H.; Livingston, P. O.; Clausen, H.; Jensen, K. J., *Journal of proteome research* **2010,** *9* (10), 5250.

258. Blixt, O.; Head, S.; Mondala, T.; Scanlan, C.; Huflejt, M. E.; Alvarez, R.; Bryan, M. C.; Fazio, F.; Calarese, D.; Stevens, J.; Razi, N.; Stevens, D. J.; Skehel, J. J.; van Die, I.; Burton, D. R.; Wilson, I. A.; Cummings, R.; Bovin, N.; Wong, C. H.; Paulson, J. C., *P Natl Acad Sci USA* **2004,** *101* (49), 17033.

259. Lee, M. r.; Shin, I., *Angewandte Chemie International Edition* **2005,** *44* (19), 2881.

260. Westerlind, U.; Schroder, H.; Hobel, A.; Gaidzik, N.; Kaiser, A.; Niemeyer, C. M.; Schmitt, E.; Waldmann, H.; Kunz, H., *Angew Chem Int Edit* **2009,** *48* (44), 8263.

261. Weinrich, D.; Kohn, M.; Jonkheijm, P.; Westerlind, U.; Dehmelt, L.; Engelkamp, H.; Christianen, P. C. M.; Kuhlmann, J.; Maan, J. C.; Nusse, D.; Schroder, H.; Wacker, R.; Voges, E.; Breinbauer, R.; Kunz, H.; Niemeyer, C. M.; Waldmann, H., *Chembiochem* **2010,** *11* (2), 235.

262. Smith, E. A.; Thomas, W. D.; Kiessling, L. L.; Corn, R. M., *J Am Chem Soc* **2003,** *125* (20), 6140.

263. Blixt, O.; Westerlind, U., *Curr Opin Chem Biol* **2014,** *18*, 62.

264. De Araújo, A. D.; Palomo, J. M.; Cramer, J.; Köhn, M.; Schröder, H.; Wacker, R.; Niemeyer, C.; Alexandrov, K.; Waldmann, H., *Angewandte Chemie International Edition* **2006,** *45* (2), 296.

265. Beckmann, H. S.; Niederwieser, A.; Wiessler, M.; Wittmann, V., *Chem-Eur J* **2012,** *18* (21), 6548.

266. Köhn, M.; Wacker, R.; Peters, C.; Schröder, H.; Soulère, L.; Breinbauer, R.; Niemeyer, C. M.; Waldmann, H., *Angewandte Chemie International Edition* **2003,** *42* (47), 5830.

267. Helms, B.; Van Baal, I.; Merkx, M.; Meijer, E., *Chembiochem* **2007,** *8* (15), 1790.

268. Govindaraju, T.; Jonkheijm, P.; Gogolin, L.; Schroeder, H.; Becker, C. F.; Niemeyer, C. M.; Waldmann, H., *Chem Commun* **2008,** (32), 3723.

269. Bryan, M. C.; Fazio, F.; Lee, H.-K.; Huang, C.-Y.; Chang, A.; Best, M. D.; Calarese, D. A.; Blixt, O.; Paulson, J. C.; Burton, D., *J Am Chem Soc* **2004,** *126* (28), 8640.

270. Yi, L.; Sun, H.; Wu, Y. W.; Triola, G.; Waldmann, H.; Goody, R. S., *Angewandte Chemie International Edition* **2010,** *49* (49), 9417.

271. Cló, E.; Blixt, O.; Jensen, K. J., *European Journal of Organic Chemistry* **2010,** *2010* (3), 540.

272. Hizal, D. B.; Wolozny, D.; Colao, J.; Jacobson, E.; Tian, Y.; Krag, S. S.; Betenbaugh, M. J.; Zhang, H., *Clinical proteomics* **2014,** *11* (1), 15.

273. Zhang, H.; Li, X.-j.; Martin, D. B.; Aebersold, R., *Nature biotechnology* **2003,** *21* (6), 660.

274. Zou, Z.; Ibisate, M.; Zhou, Y.; Aebersold, R.; Xia, Y.; Zhang, H., *Analytical chemistry* **2008,** *80* (4), 1228.

275. Zhou, W.; Yao, N.; Yao, G.; Deng, C.; Zhang, X.; Yang, P., *Chem Commun* **2008,** (43), 5577.

276. Alvarez-Manilla, G.; Atwood III, J.; Guo, Y.; Warren, N. L.; Orlando, R.; Pierce, M., *Journal of proteome research* **2006,** *5* (3), 701.

277. Alpert, A. J., *Journal of Chromatography A* **1990,** *499*, 177.

278. Vosseller, K.; Trinidad, J. C.; Chalkley, R. J.; Specht, C. G.; Thalhammer, A.; Lynn, A. J.; Snedecor, J. O.; Guan, S.; Medzihradszky, K. F.; Maltby, D. A., *Mol Cell Proteomics* **2006,** *5* (5), 923.

279. Yodoshi, M.; Oyama, T.; Masaki, K.; Kakehi, K.; Hayakawa, T.; Suzuki, S., *Analytical Sciences* **2011,** *27* (4), 395.

280. Roepstorff, P.; Fohlman, J., *Biomedical mass spectrometry* **1984,** *11* (11), 601.

281. Domon, B.; Costello, C. E., *Glycoconjugate journal* **1988,** *5* (4), 397.

282. Carr, S. A.; Huddleston, M. J.; Bean, M. F., *Protein science* **1993,** *2* (2), 183.

283. Conboy, J. J.; Henion, J. D., *Journal of the American Society for Mass Spectrometry* **1992,** *3* (8), 804.

284. Demelbauer, U. M.; Zehl, M.; Plematl, A.; Allmaier, G.; Rizzi, A., *Rapid communications in mass spectrometry* **2004,** *18* (14), 1575.

285. Nilsson, J., *Glycoconjugate journal* **2016**.

286. Chalkley, R. J.; Burlingame, A. L., *Mol Cell Proteomics* **2003,** *2* (3), 182.

287. Mirgorodskaya, E.; Hassan, H.; Clausen, H.; Roepstorff, P., *Analytical chemistry* **2001,** *73* (6), 1263.

288. Håkansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L., *Analytical chemistry* **2001,** *73* (18), 4530.

289. Mikesh, L. M.; Ueberheide, B.; Chi, A.; Coon, J. J.; Syka, J. E. P.; Shabanowitz, J.; Hunt, D. F., *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2006,** *1764* (12), 1811.

290. Sakaidani, Y.; Nomura, T.; Matsuura, A.; Ito, M.; Suzuki, E.; Murakami, K.; Nadano, D.; Matsuda, T.; Furukawa, K.; Okajima, T., *Nature communications* **2011,** *2*, 583.

291. Matsuura, A.; Ito, M.; Sakaidani, Y.; Kondo, T.; Murakami, K.; Furukawa, K.; Nadano, D.; Matsuda, T.; Okajima, T., *J Biol Chem* **2008,** *283* (51), 35486.

292. Ashline, D.; Singh, S.; Hanneman, A.; Reinhold, V., *Analytical chemistry* **2005,** *77* (19), 6250.

293. Kang, P.; Mechref, Y.; Kyselova, Z.; Goetz, J. A.; Novotny, M. V., *Analytical chemistry* **2007,** *79* (16), 6064.

294. HAKOMORI, S.-I., *The Journal of Biochemistry* **1964,** *55* (2), 205.

295. Carraway, K. L.; Hull, S. R., *Glycobiology* **1991,** *1* (2), 131.

296. Pett, C.; Schorlemer, M.; Westerlind, U., *Chem-Eur J* **2013,** *19* (50), 17001.

297. Pedersen, C. M.; Marinescu, L. G.; Bols, M., *Chem Commun* **2008,** (21), 2465.

298. Bols, M. P., C. M., *J Org Chem* **2017,** *13*, 93.

299. Boons, G.; Hale, K., Organic Synthesis with Carbohydrates. Sheffield Academic Press/Blackwell Science: 2000.

300. Ellervik, U.; Magnusson, G., *Carbohyd Res* **1996,** *280* (2), 251.

301. Paulsen, H.; Holck, J. P., *Carbohyd Res* **1982,** *109* (Nov), 89.

302. Liebe, B.; Kunz, H., *Angewandte Chemie-International Edition in English* **1997,** *36* (6), 618.

303. Dziadek, S.; Brocke, C.; Kunz, H., *Chemistry–A European Journal* **2004,** *10* (17), 4150.

304. Brocke, C.; Kunz, H., *Synthesis* **2004,** *2004* (04), 525.

305. Helferich, B.; Weis, K., *Chem Ber-Recl* **1956,** *89* (2), 314.

306. Helferich, B.; Wedemeyer, K. F., *Liebigs Ann Chem* **1949,** *563* (1-2), 139.

307. Konradsson, P.; Udodong, U. E.; Fraser-Reid, B., *Tetrahedron Lett* **1990,** *31* (30), 4313.

308. Konradsson, P.; Mootoo, D. R.; McDevitt, R. E.; Fraser-Reid, B., *Journal of the Chemical Society, Chemical Communications* **1990,** (3), 270.

309. Veeneman, G.; Van Leeuwen, S.; Van Boom, J., *Tetrahedron Lett* **1990,** *31* (9), 1331.

310. Fugedi, P.; Garegg, P. J.; Lonn, H.; Norberg, T., *Glycoconjugate journal* **1987,** *4* (2), 97.

311. Paquet, A., *Can J Chem* **1982,** *60* (8), 976.

312. Cai, H.; Huang, Z. H.; Shi, L.; Zou, P.; Zhao, Y. F.; Kunz, H.; Li, Y. M., *European Journal of Organic Chemistry* **2011,** *2011* (20-21), 3685.

313. Dabritz, E., *Angew Chem Int Edit* **1966,** *5* (5), 470.

314. Dziadek, S., *Johannes Gutenberg Universität Mainz* **2005**.

315. Broddefalk, J.; Nilsson, U.; Kihlberg, J., *J Carbohyd Chem* **1994,** *13* (1), 129.

316. Kozikowski, A. P.; Lee, J. M., *J Org Chem* **1990,** *55* (3), 863.

317. Fischer, E., *Chemische Berichte* **1914,** *47*, 196.

318. Lemieux, R. U.; Ratcliffe, R. M., *Canadian Journal of Chemistry* **1979,** *57* (10), 1244.

319. László Kürti, B. C., *Strategic Applications of Named Reactions in Organic Synthesis*. Elsevier: 2005.

320. Rosen, T.; Lico, I. M.; Chu, D. T. W., *J Org Chem* **1988,** *53* (7), 1580.

321. Kolakowski, R. V.; Shangguan, N.; Sauers, R. R.; Williams, L. J., *J Am Chem Soc* **2006,** *128* (17), 5695.

322. Florent, J. C.; Monneret, C., *Synthesis-Stuttgart* **1982,** (1), 29.

323. Pett, C.; Westerlind, U., *Chem-Eur J* **2014,** *20* (24), 7287.

324. Toepfer, A.; Schmidt, R. R., *J Carbohyd Chem* **1993,** *12* (7), 809.

325. Ellervik, U.; Magnusson, G., *J Org Chem* **1998,** *63* (25), 9314.

326. Excoffier, G.; Gagnaire, D.; Utille, J. P., *Carbohyd Res* **1975,** *39* (2), 368.

327. Schmidt, R. R., *Angewandte Chemie-International Edition in English* **1986,** *25* (3), 212.

328. Goddard-Borger, E. D.; Stick, R. V., *Org Lett* **2007,** *9* (19), 3797.

329. Cavender, C. J.; Shiner, V. J., *J Org Chem* **1972,** *37* (22), 3567.

330. Nyffeler, P. T.; Liang, C. H.; Koeller, K. M.; Wong, C. H., *J Am Chem Soc* **2002,** *124* (36), 10773.

331. Pandiakumar, A. K.; Sarma, S. P.; Samuelson, A. G., *Tetrahedron Lett* **2014,** *55* (18), 2917.

332. Paulsen, H.; Richter, A.; Sinnwell, V.; Stenzel, W., *Carbohyd Res* **1978,** *64* (Jul), 339.

333. De Silva, R. A.; Wang, Q. L.; Chidley, T.; Appulage, D. K.; Andreana, P. R., *J Am Chem Soc* **2009,** *131* (28), 9622.

334. Luning, B.; Norberg, T.; Tejbrant, J., *Glycoconjugate journal* **1989,** *6* (1), 5.

335. Vasan, M.; Wolfert, M. A.; Boons, G. J., *Org Biomol Chem* **2007,** *5* (13), 2087.

336. Jensen, H. H.; Nordstrøm, L. U.; Bols, M., *J Am Chem Soc* **2004,** *126* (30), 9205.

337. Yan, F.; Mehta, S.; Eichler, E.; Wakarchuk, W. W.; Gilbert, M.; Schur, M. J.; Whitfield, D. M., *The Journal of organic chemistry* **2003,** *68* (6), 2426.

338. Dullenkopf, W.; CastroPalomino, J. C.; Manzoni, L.; Schmidt, R. R., *Carbohyd Res* **1996,** *296*, 135.

339. Kerns, R. J.; Zha, C.; Benakli, K.; Liang, Y.-Z., *Tetrahedron Lett* **2003,** *44* (44), 8069.

340. Crouch, R. D., *Tetrahedron* **2004,** *60* (28), 5833.

341. Just, G.; Grozinger, K., *Synthesis* **1976,** *1976* (07), 457.

342. Bryan, D. B.; Hall, R. F.; Holden, K. G.; Huffman, W. F.; Gleason, J. G., *J Am Chem Soc* **1977,** *99* (7), 2353.

343. Martin, S. F.; Chen, K. X.; Eary, C. T., *Org Lett* **1999,** *1* (1), 79.

344. Zhang, Z. Y.; Ollmann, I. R.; Ye, X. S.; Wischnat, R.; Baasov, T.; Wong, C. H., *J Am Chem Soc* **1999,** *121* (4), 734.

345. Kamat, M. N.; Demchenko, A. V., *Org Lett* **2005,** *7* (15), 3215.

346. Pedersen, C. M.; Nordstrøm, L. U.; Bols, M., *J Am Chem Soc* **2007,** *129* (29), 9222.

347. Corey, E.; Venkateswarlu, A., *J Am Chem Soc* **1972,** *94* (17), 6190.

348. Smith, A. B.; Ott, G. R., *J Am Chem Soc* **1996,** *118* (51), 13095.

349. Smith, M.; Khorana, H. G.; Rammler, D. H.; Goldberg, I. H., *J Am Chem Soc* **1962,** *84* (3), 430.

350. Hancock, G.; Galpin, I. J.; Morgan, B. A., *Tetrahedron Lett* **1982,** *23* (2), 249.

351. Ananthanarayan, T. P.; Gallagher, T.; Magnus, P., *J Chem Soc Chem Comm* **1982,** (12), 709.

352. Morss, L. R., *Chem Rev* **1976,** *76* (6), 827.

353. Kagan, H.; Namy, J.; Girard, P., *Tetrahedron* **1981,** *37*, 175.

354. Pett, C.; Cai, H.; Liu, J.; Palitzsch, B.; Schorlemer, M.; Hartmann, S.; Stergiou, N.; Lu, M.; Kunz, H.; Schmitt, E., *Chem-Eur J* **2017,** *23* (16), 3875.

355. Hayashi, S., *Chest* **2002,** *121* (5, Supplement), 183S.

356. Jones, M. S.; Harrach, B.; Ganac, R. D.; Gozum, M. M.; dela Cruz, W. P.; Riedel, B.; Pan, C.; Delwart, E. L.; Schnurr, D. P., *Journal of virology* **2007,** *81* (11), 5978.

357. Roy, S.; Sandhu, A.; Medina, A.; Clawson, D. S.; Wilson, J. M., *Emerging infectious diseases* **2012,** *18* (7), 1081.

358. Lenman, A.; Liaci, A. M.; Liu, Y.; Årdahl, C.; Rajan, A.; Nilsson, E.; Bradford, W.; Kaeshammer, L.; Jones, M. S.; Frängsmyr, L.; Feizi, T.; Stehle, T.; Arnberg, N., *PLOS Pathogens* **2015,** *11* (2), e1004657.

359. Barondes, S. H.; Castronovo, V.; Cooper, D.; Cummings, R. D.; Drickamer, K.; Feizi, T.; Gitt, M. A.; Hirabayashi, J.; Hughes, C.; Kasai, K.-i., *Cell* **1994,** *76* (4), 597.

360. Di Lella, S.; Sundblad, V.; Cerliani, J. P.; Guardia, C. M.; Estrin, D. A.; Vasta, G. R.; Rabinovich, G. A., *Biochemistry* **2011,** *50* (37), 7842.

361. Huddleston, M. J.; Bean, M. F.; Carr, S. A., *Analytical chemistry* **1993,** *65* (7), 877.

362. Yu, Q.; Wang, B.; Chen, Z.; Urabe, G.; Glover, M. S.; Shi, X.; Guo, L. W.; Kent, K. C.; Li, L., *J Am Soc Mass Spectrom* **2017**.

363. Westerlind, U.; Hobel, A.; Gaidzik, N.; Schmitt, E.; Kunz, H., *Angew Chem Int Edit* **2008,** *47* (39), 7551.

364. Ficht, S.; Payne, R. J.; Guy, R. T.; Wong, C. H., *Chemistry–A European Journal* **2008,** *14* (12), 3620.

365. Wang, S.; Dupin, L.; Noël, M.; Carroux, C. J.; Renaud, L.; Géhin, T.; Meyer, A.; Souteyrand, E.; Vasseur, J. J.; Vergoten, G., *Chem-Eur J* **2016,** *22* (33), 11785.

366. Wang, P.; Nilsson, J.; Brinkmalm, G.; Larson, G.; Huang, X., *Chem Commun* **2014,** *50* (95), 15067.

367. Dessolin, M.; Guillerez, M.-G.; Thieriet, N.; Guibé, F.; Loffet, A., *Tetrahedron Lett* **1995,** *36* (32), 5741.

368. Smit, C.; Blümer, J.; Eerland, M. F.; Albers, M. F.; Müller, M. P.; Goody, R. S.; Itzen, A.; Hedberg, C., *Angewandte Chemie* **2011,** *123* (39), 9367.

369. Trost, B. M.; Zhang, T.; Sieber, J. D., *Chem Sci* **2010,** *1* (4), 427.

370. Uezu, A.; Okada, H.; Murakoshi, H.; del Vescovo, C. D.; Yasuda, R.; Diviani, D.; Soderling, S. H., *Proceedings of the National Academy of Sciences* **2012,** *109* (43), E2929.

371. Kaiser, A.; Gaidzik, N.; Westerlind, U.; Kowalczyk, D.; Hobel, A.; Schmitt, E.; Kunz, H., *Angew Chem Int Edit* **2009,** *48* (41), 7551.

372. Pathak, S.; Alonso, J.; Schimpl, M.; Rafie, K.; Blair, D. E.; Borodkin, V. S.; Schuttelkopf, A. W.; Albarbarawi, O.; van Aalten, D. M. F., *Nat Struct Mol Biol* **2015,** *22* (9), 744.

373. Nagata, Y.; Burger, M. M., *J Biol Chem* **1974,** *249* (10), 3116.

374. Wright, C. S., *Journal of molecular biology* **1984,** *178* (1), 91.

375. Bhavanandan, V. P.; Katlic, A. W., *J Biol Chem* **1979,** *254* (10), 4000.

376. Tollefsen, S.; Kornfeld, R., *J Biol Chem* **1983,** *258* (8), 5172.

377. Puri, K. D.; Gopalakrishnan, B.; Surolia, A., *Febs Lett* **1992,** *312* (2-3), 208.

378. Iyer, P. S.; Wilkinson, K. D.; Goldstein, I. J., *Archives of biochemistry and biophysics* **1976,** *177* (1), 330.

379. Oyelaran, O.; Li, Q.; Farnsworth, D.; Gildersleeve, J. C., *Journal of proteome research* **2009,** *8* (7), 3529.

380. Schwefel, D.; Maierhofer, C.; Beck, J. G.; Seeberger, S.; Diederichs, K.; Möller, H. M.; Welte, W.; Wittmann, V., *J Am Chem Soc* **2010,** *132* (25), 8704.

381. Brewer, C. F.; Miceli, M. C.; Baum, L. G., *Current opinion in structural biology* **2002,** *12* (5), 616.

382. Wright, C. S., *Journal of molecular biology* **1980,** *141* (3), 267.

383. Hinneburg, H.; Stavenhagen, K.; Schweiger-Hufnagel, U.; Pengelley, S.; Jabs, W.; Seeberger, P. H.; Silva, D. V.; Wuhrer, M.; Kolarich, D., *Journal of The American Society for Mass Spectrometry* **2016,** *27* (3), 507.

384. Cao, L.; Tolić, N.; Qu, Y.; Meng, D.; Zhao, R.; Zhang, Q.; Moore, R. J.; Zink, E. M.; Lipton, M. S.; Paša-Tolić, L., *Analytical biochemistry* **2014,** *452*, 96.

385. Henrissat, B.; Davies, G., *Current opinion in structural biology* **1997,** *7* (5), 637.

386. Li, J.; Huang, C.-l.; Zhang, L.-w.; Lin, L.; Li, Z.-h.; Zhang, F.-w.; Wang, P., *Biochemistry (Moscow)* **2010,** *75* (7), 938.

387. Dennis, R. J.; Taylor, E. J.; Macauley, M. S.; Stubbs, K. A.; Turkenburg, J. P.; Hart, S. J.; Black, G. N.; Vocadlo, D. J.; Davies, G. J., *Nat Struct Mol Biol* **2006,** *13* (4), 365.

388. Sheldon, W. L.; MacAuley, M. S.; Taylor, E. J.; Robinson, C. E.; Charnock, S. J.; Davies, G. J.; Vocadlo, D. J.; Black, G. W., *Biochem J* **2006,** *399*, 241.

389. Schimpl, M.; Borodkin, V. S.; Gray, L. J.; van Aalten, D. M. F., *Chem Biol* **2012,** *19* (2), 173.

390. Zhou, H.; Di Palma, S.; Preisinger, C.; Peng, M.; Polat, A. N.; Heck, A. J. R.; Mohammed, S., *Journal of proteome research* **2013,** *12* (1), 260.

391. Rodriguez, A. C.; Kohler, J. J., *Medchemcomm* **2014,** *5* (8), 1227.

392. Alfaro, J. F.; Gong, C.-X.; Monroe, M. E.; Aldrich, J. T.; Clauss, T. R.; Purvine, S. O.; Wang, Z.; Camp, D. G.; Shabanowitz, J.; Stanley, P., *Proceedings of the National Academy of Sciences* **2012,** *109* (19), 7280.

393. Trinidad, J. C.; Barkan, D. T.; Gulledge, B. F.; Thalhammer, A.; Sali, A.; Schoepfer, R.; Burlingame, A. L., *Mol Cell Proteomics* **2012,** *11* (8), 215.

394. Chalkley, R. J.; Thalhammer, A.; Schoepfer, R.; Burlingame, A., *Proceedings of the National Academy of Sciences* **2009,** *106* (22), 8894.

395. Gottlieb, H. E.; Kotlyar, V.; Nudelman, A., *The Journal of organic chemistry* **1997,** *62* (21), 7512.

396. Liebe, B.; Kunz, H., *Helvetica chimica acta* **1997,** *80* (5), 1473.

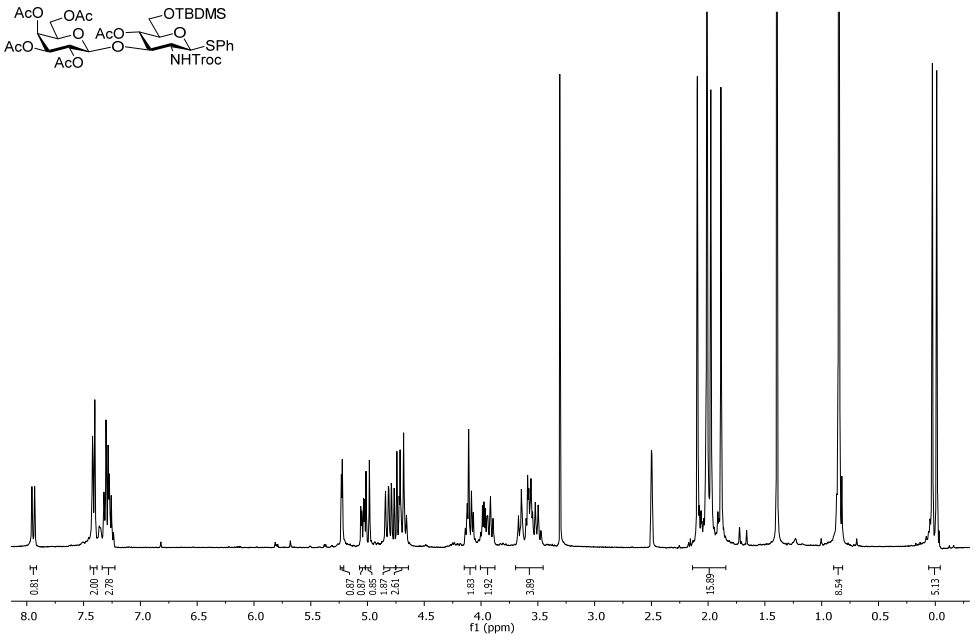
397. Schmidt, R. R.; Stumpp, M., *European Journal of Organic Chemistry* **1983,** *1983* (7), 1249.

398. Mitachi, K.; Mohan, P.; Siricilla, S.; Kurosu, M., *Chemistry–A European Journal* **2014,** *20* (16), 4554.

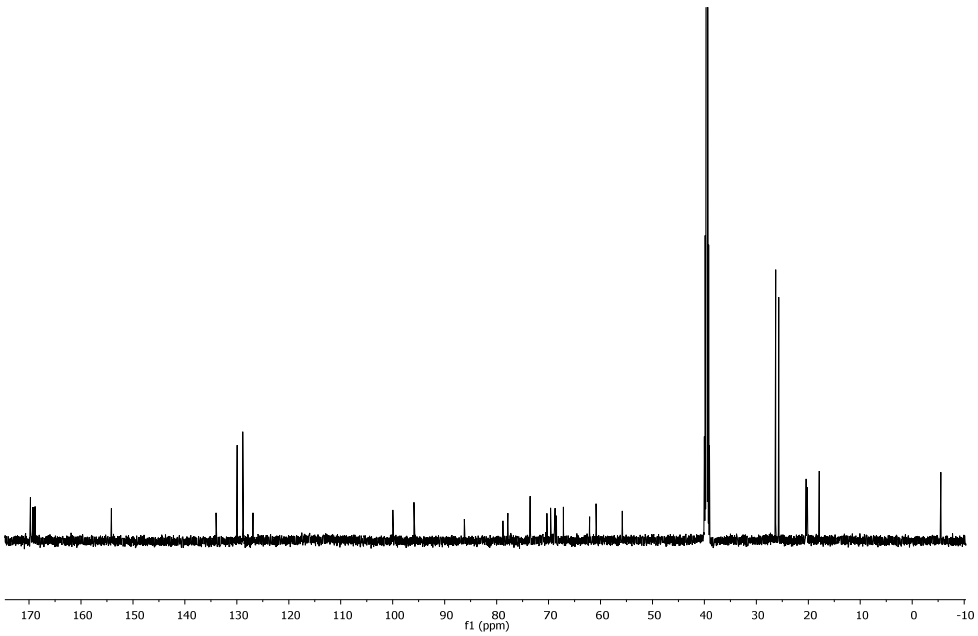
11 Appendix

11.1 NMR spectroscopic data

Compound **24**

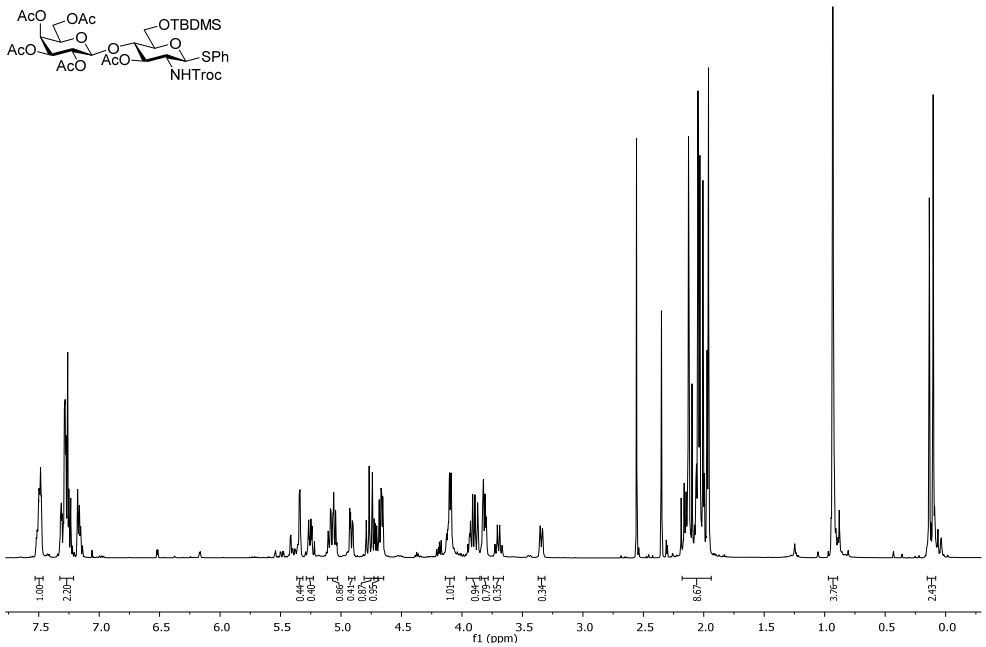


[1]H-NMR (400 MHz, DMSO)

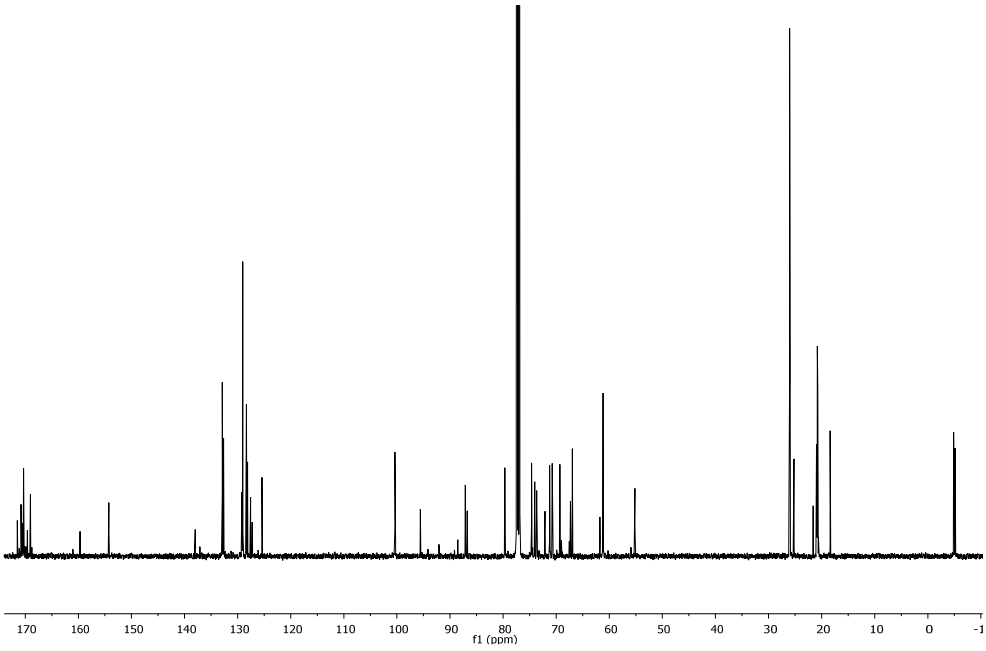


[13]C-NMR (125.8 MHz, DMSO

Compound **29**

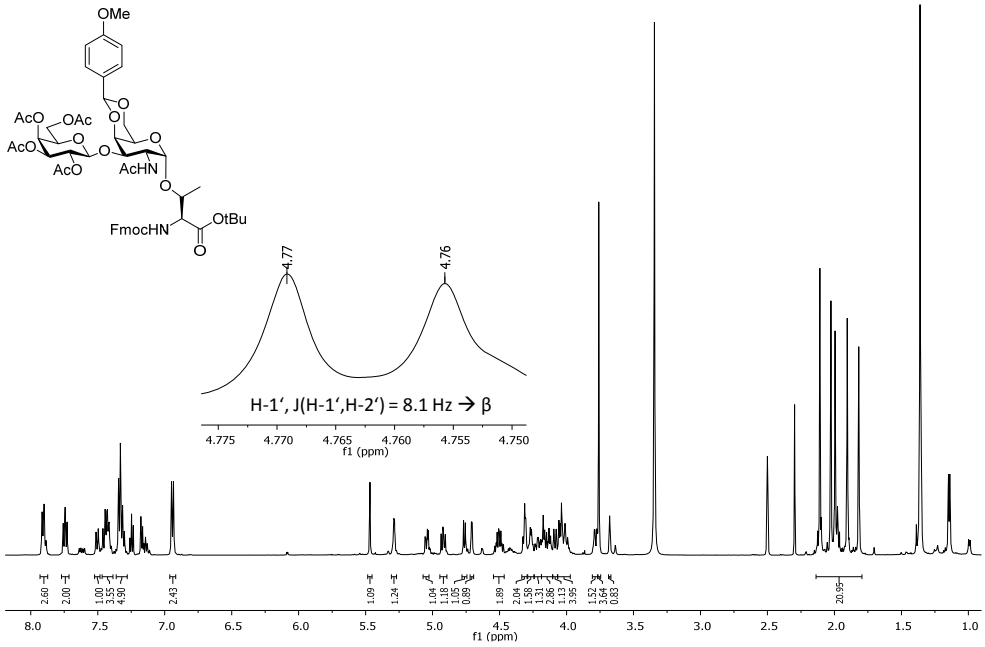


[1]H-NMR (400 MHz, $CDCl_3$)

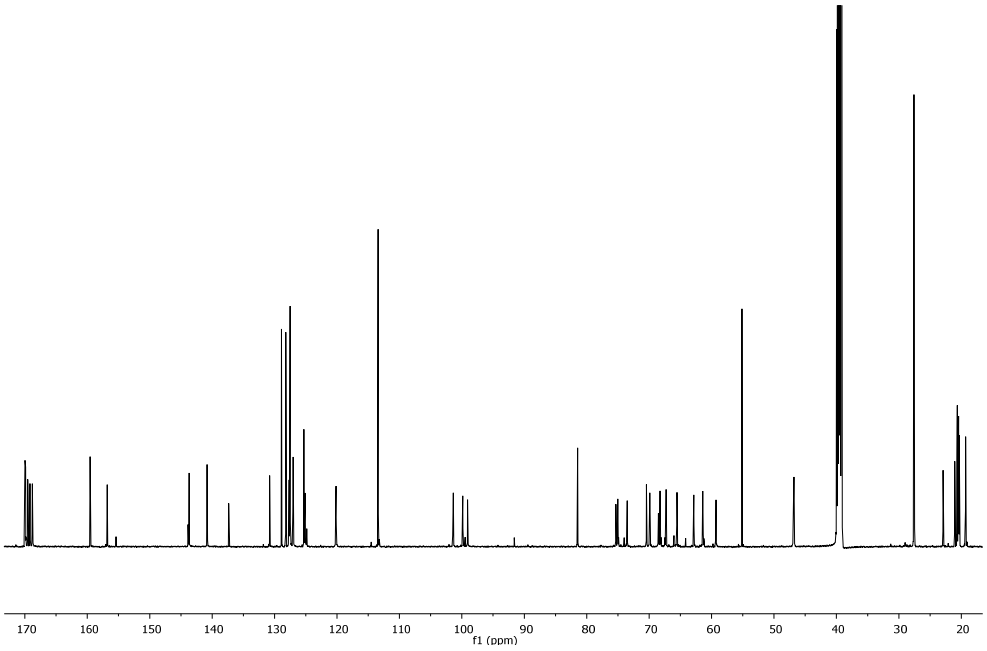


[13]C-NMR (125.8 MHz, $CDCl_3$)

Compound **30**

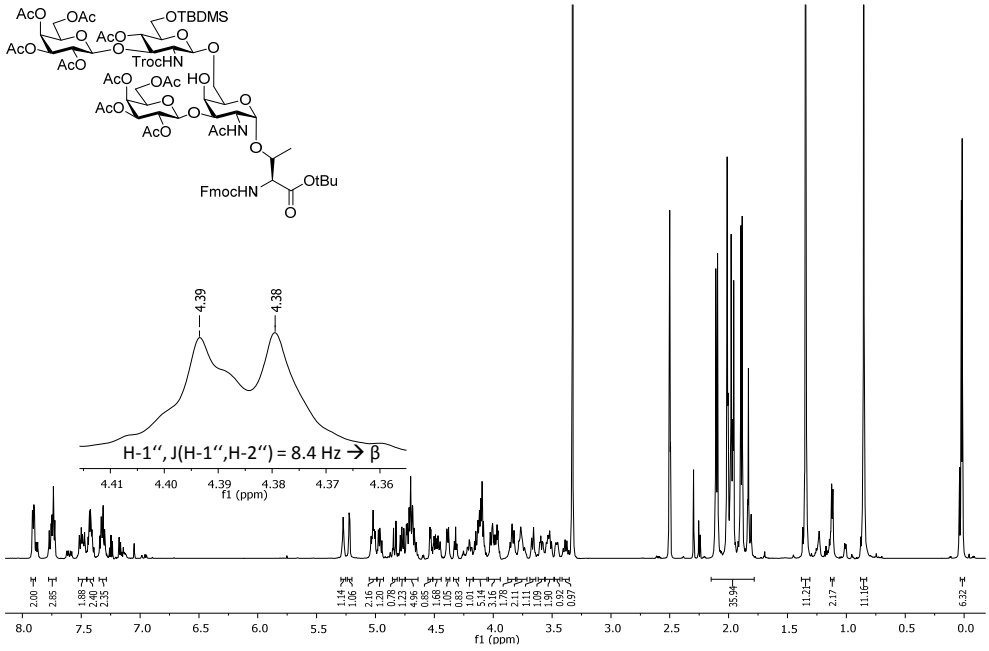


[1]H-NMR (600 MHz, DMSO)

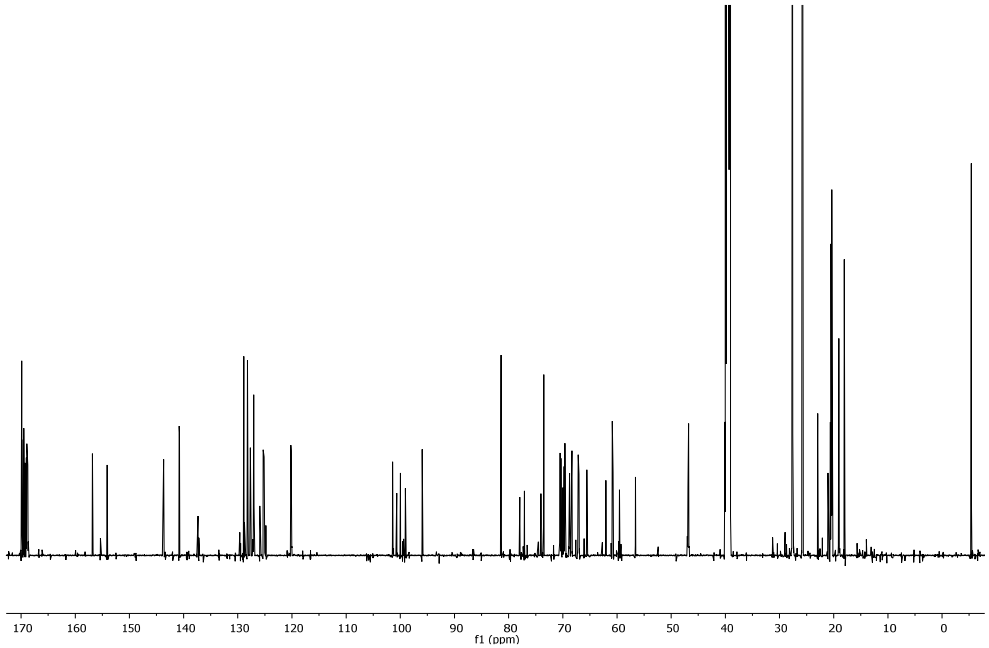


[13]C-NMR (150.9 MHz, DMSO)

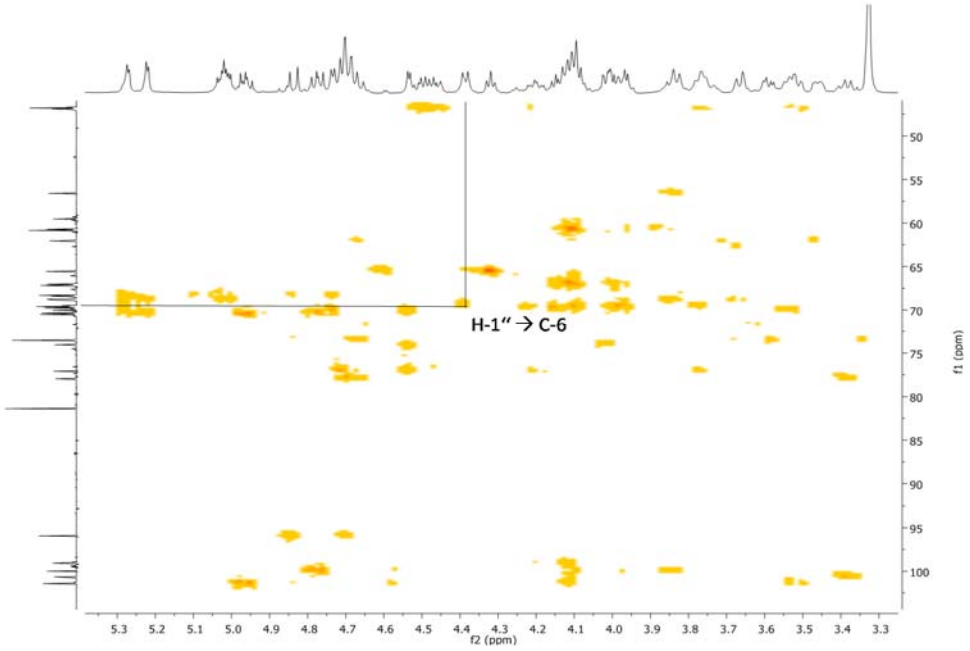
Compound **32**



[1]H-NMR (600 MHz, DMSO)

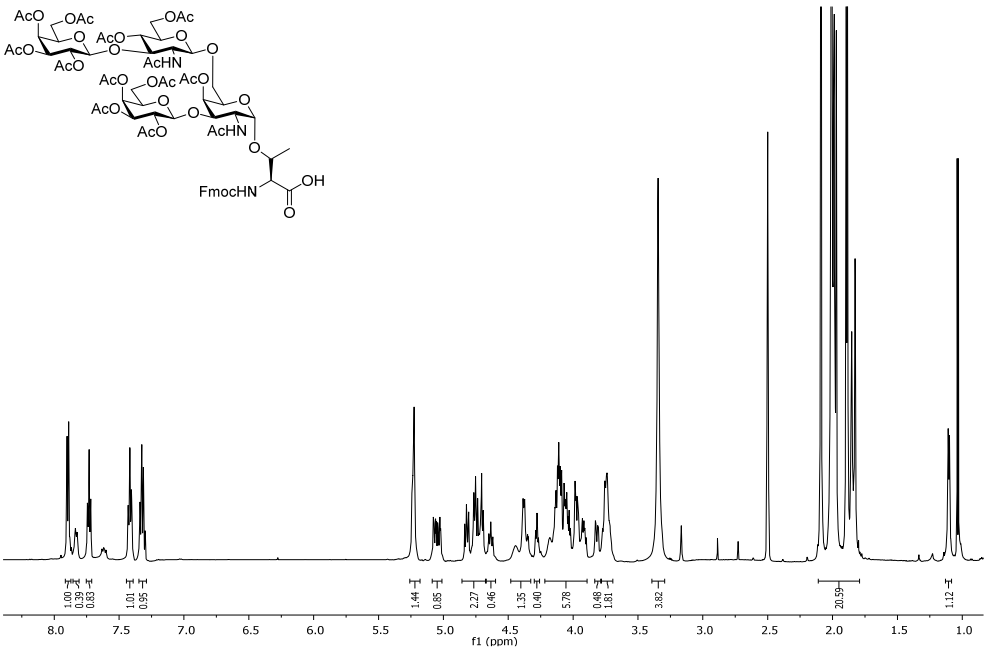


[13]C-NMR (150.9 MHz, DMSO)

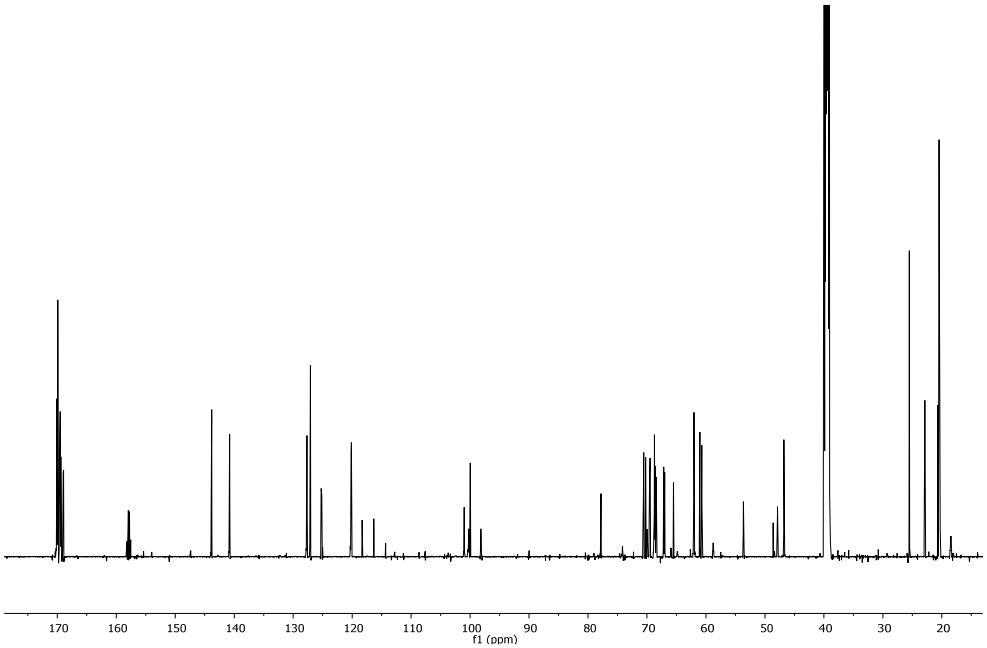


^{1}H-^{13}C-HMBC

Compound **35**

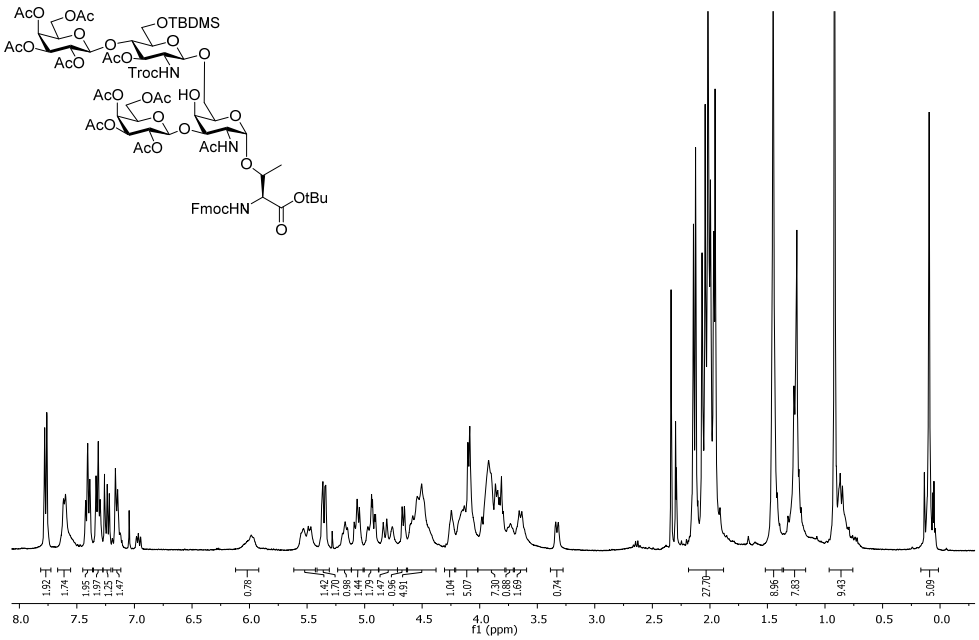


[1]H-NMR (600 MHz, DMSO)

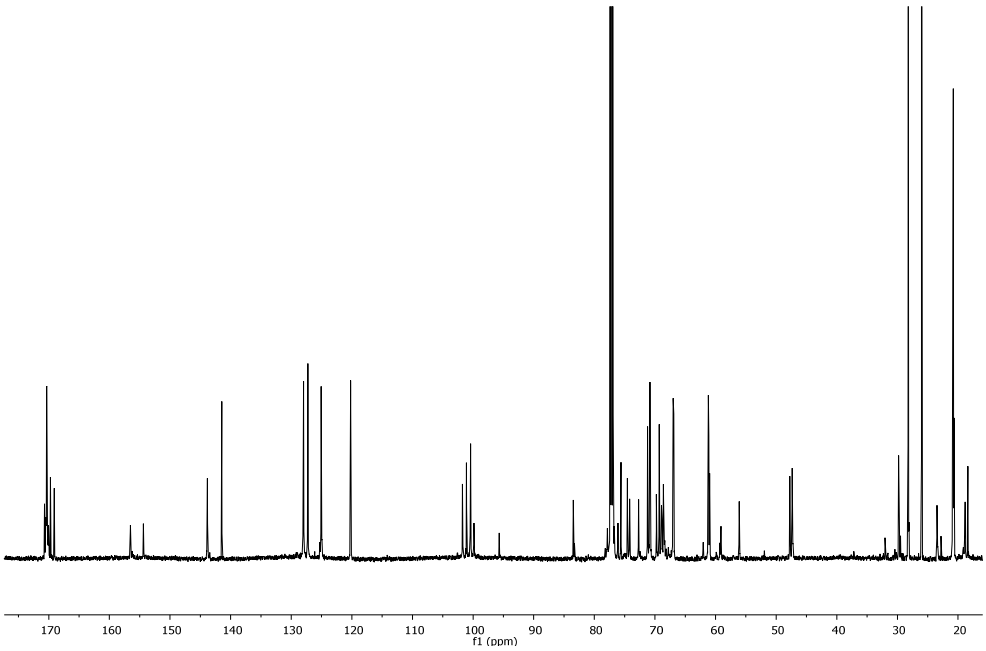


[13]C-NMR (150.9 MHz, DMSO)

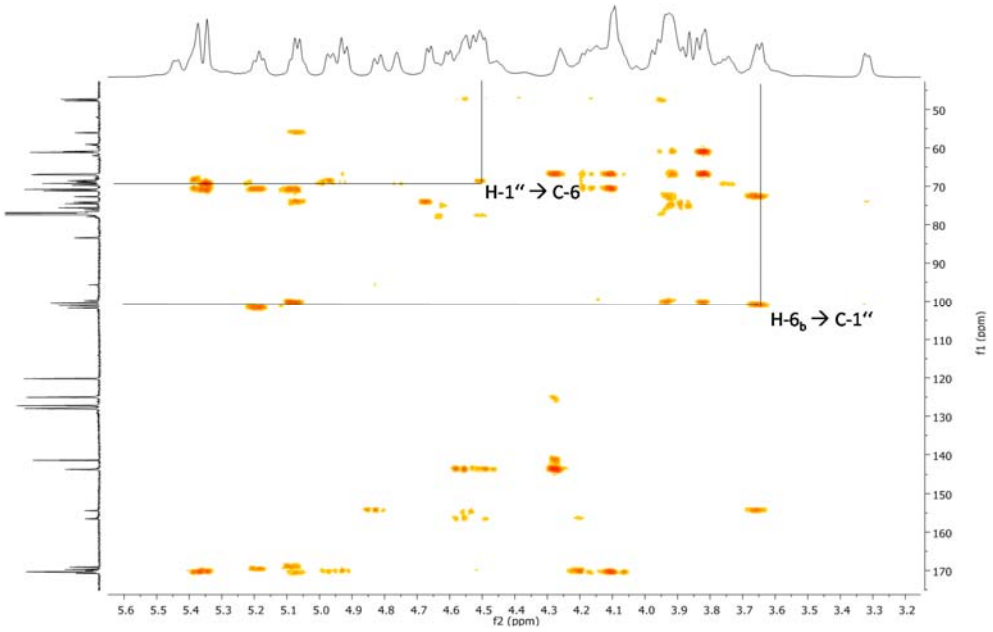
Compound **36**



[1]H-NMR (600 MHz, $CDCl_3$)

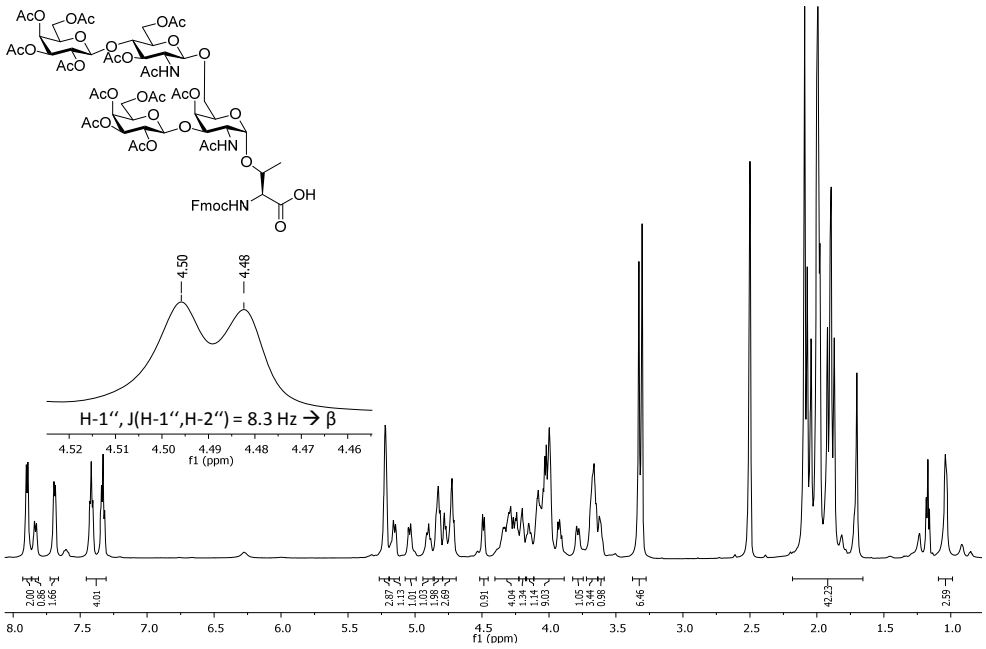


[13]C-NMR (150.9 MHz, $CDCl_3$)

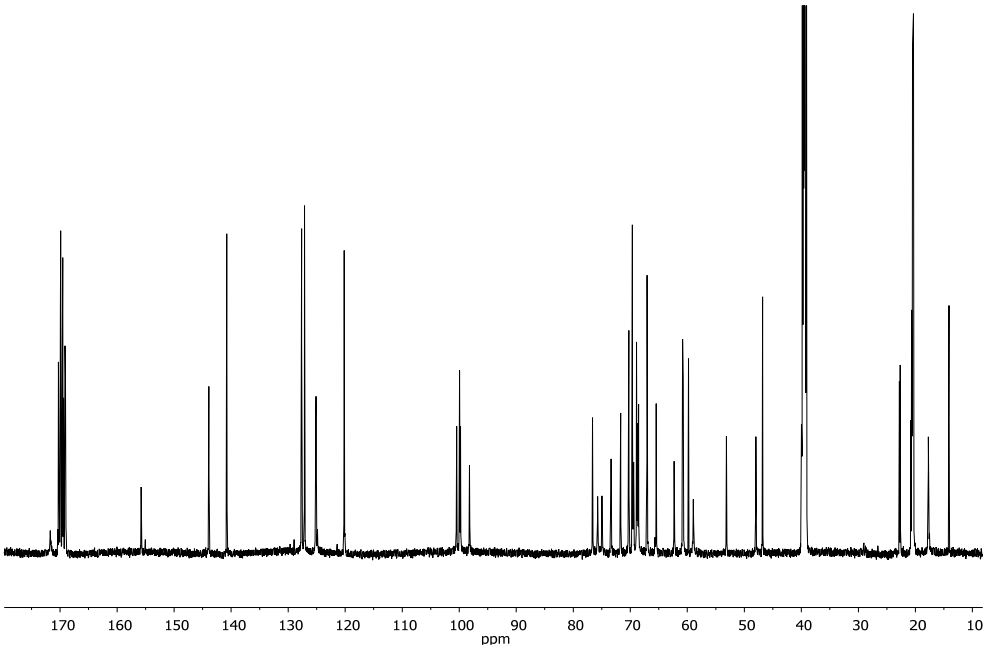


^{1}H-^{13}C-HMBC

Compound **39**

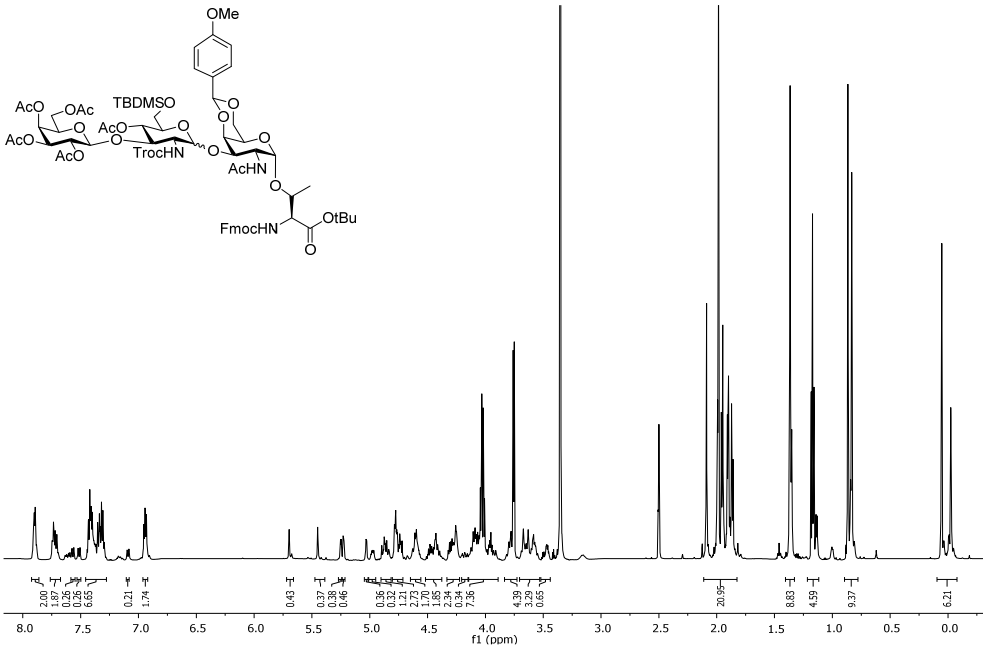


[1]H-NMR (600 MHz, $CDCl_3$)

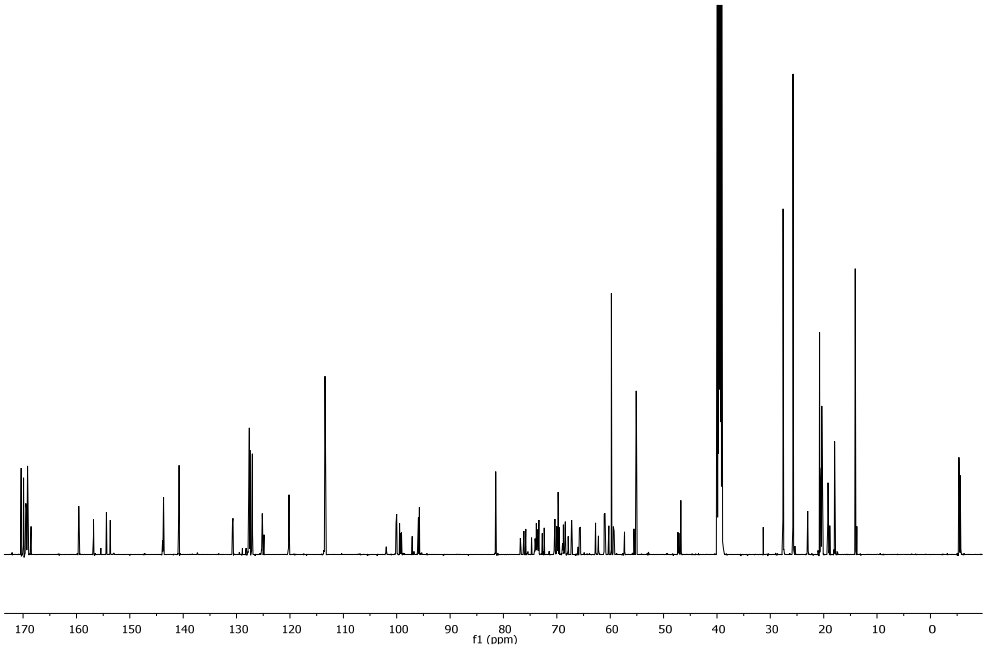


[13]C-NMR (150.9 MHz, DMSO)

Compound **40**

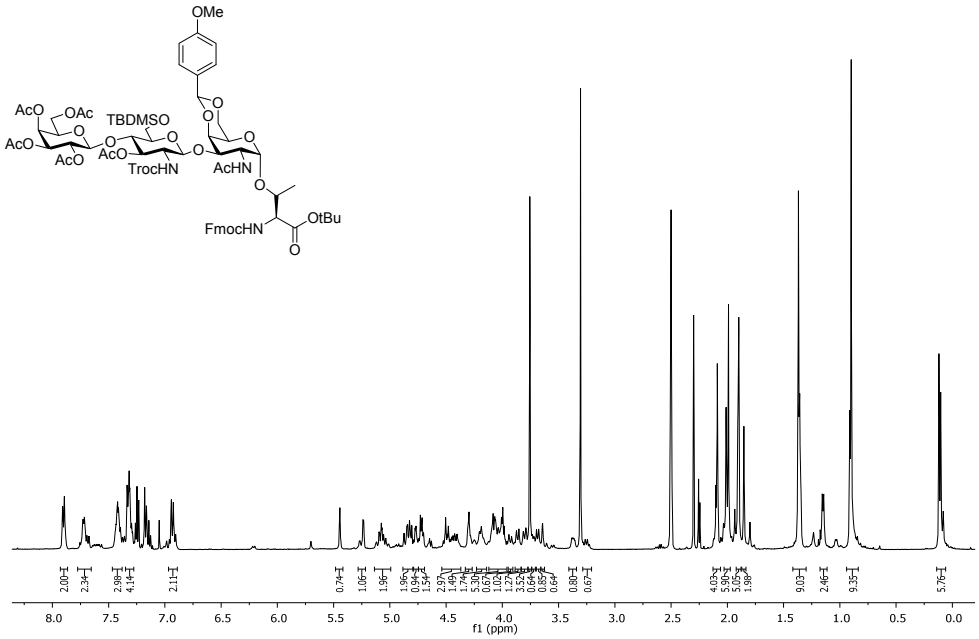


[1]H-NMR (600 MHz, DMSO)

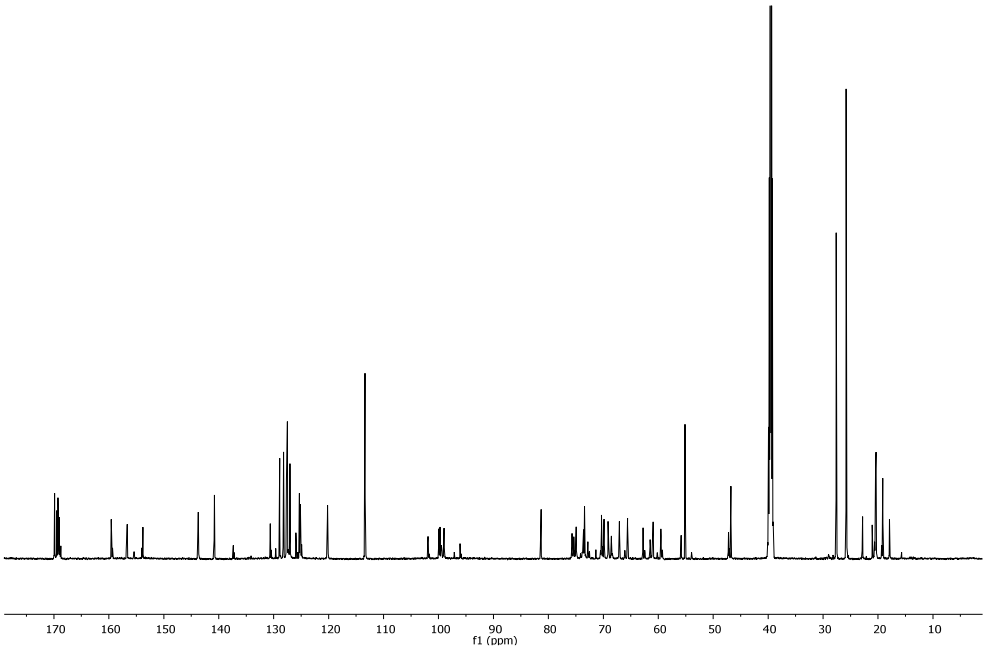


[13]C-NMR (150.9 MHz, DMSO)

Compound **43**

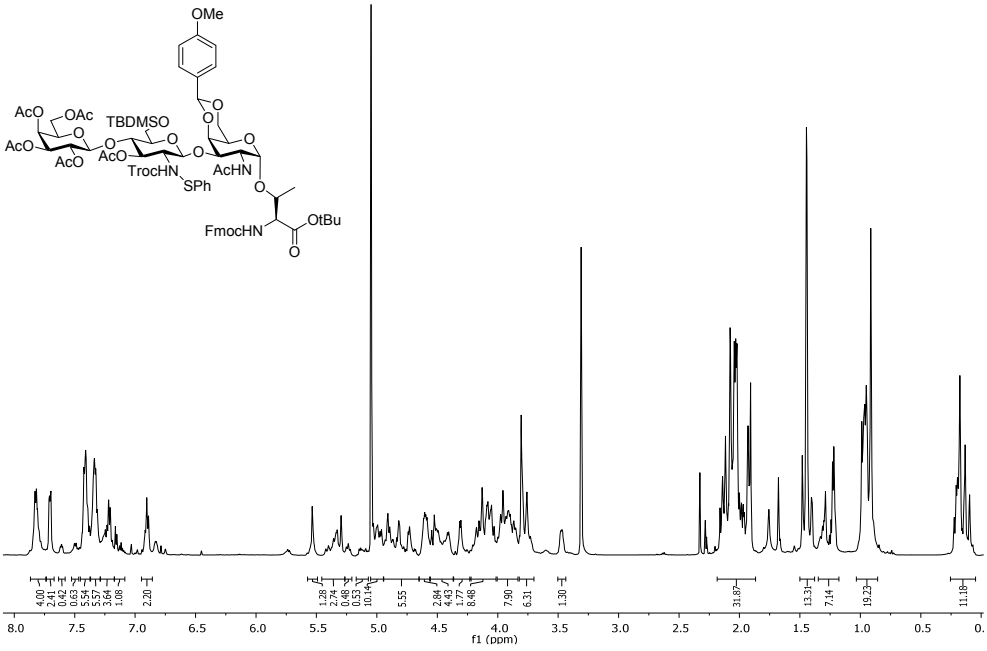


[1]H-NMR (500 MHz, DMSO)

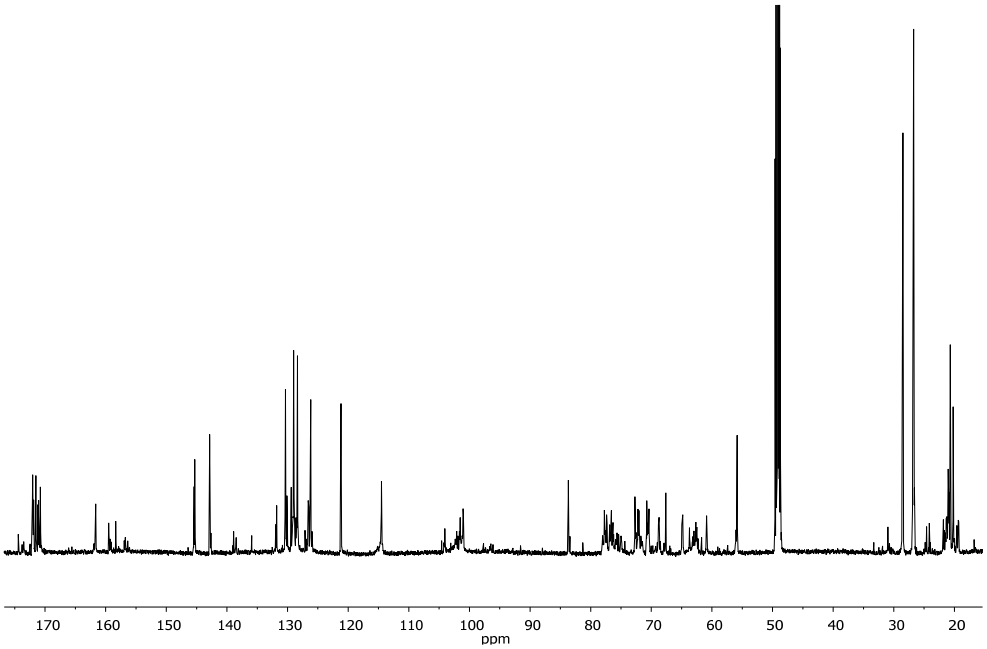


[13]C-NMR *(150.9 MHz, DMSO)*

Compound **44**

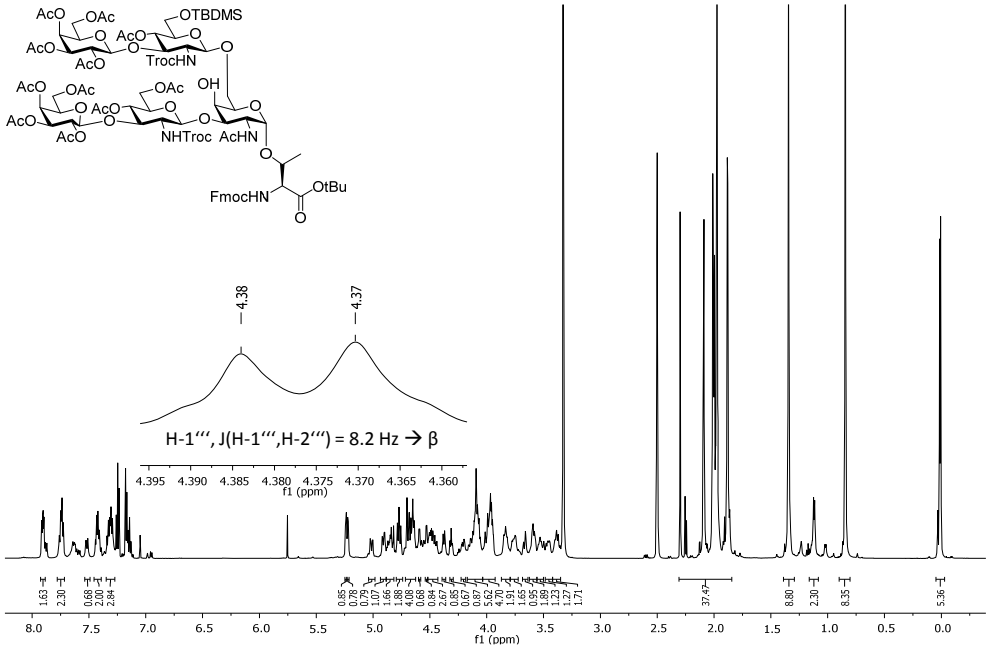


[1]H-NMR (600 MHz, MeOD, T = 278 K)

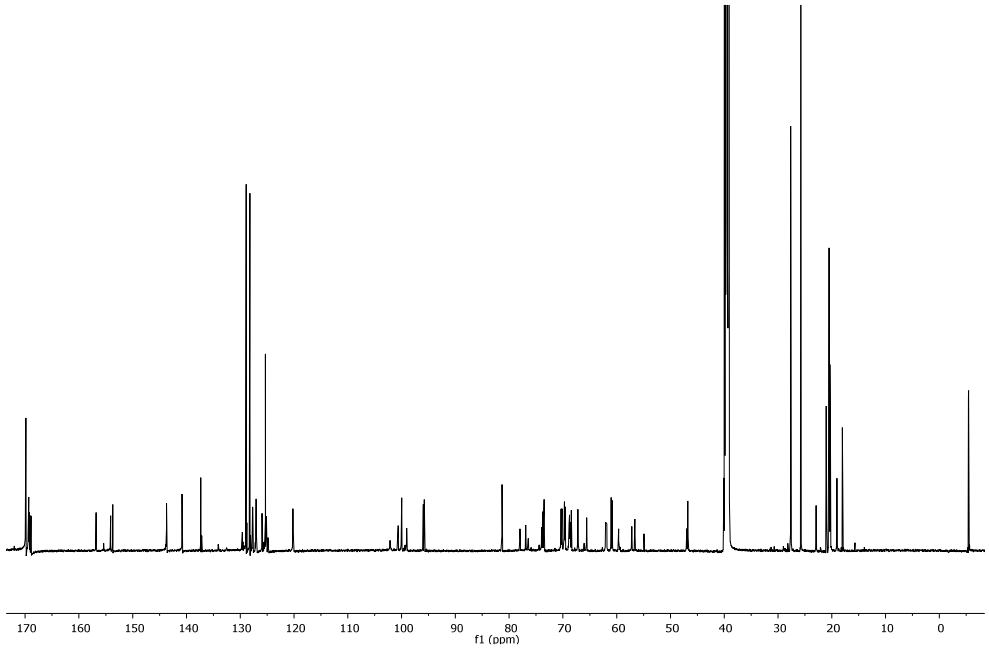


[13]C-NMR (150.9 MHz, MeOD, T = 278 K)

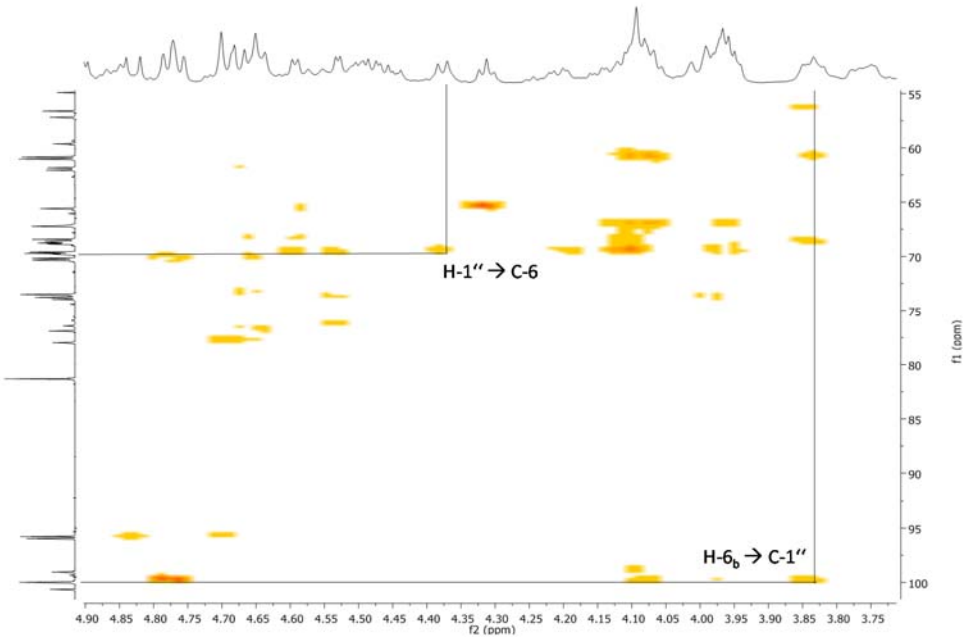
Compound **49**



[1]H-NMR (600 MHz, DMSO)

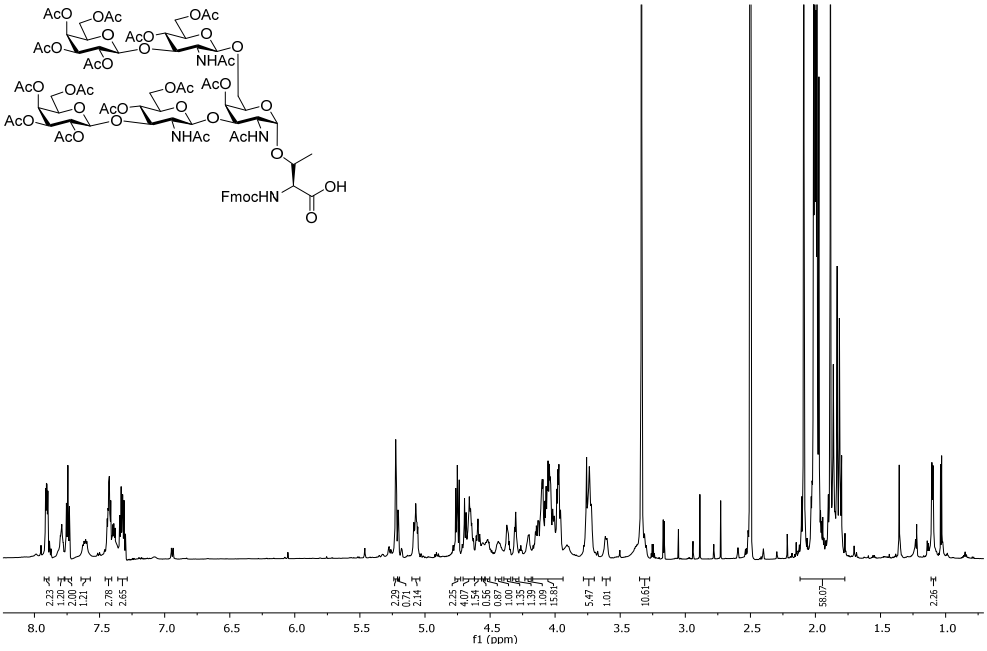


[13]C-NMR (150.9 MHz, DMSO)

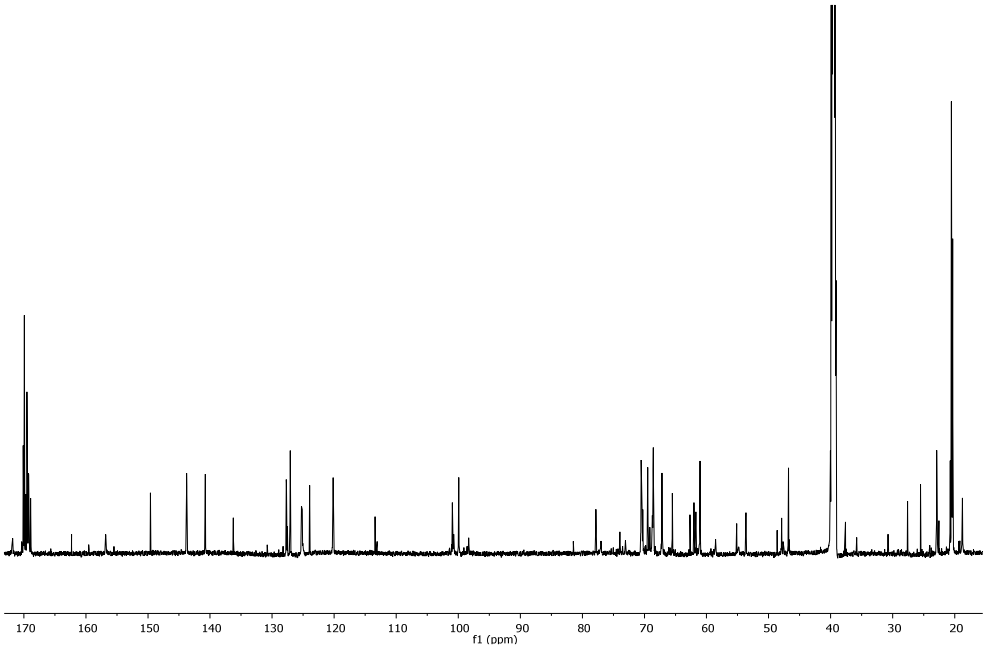


^{1}H-^{13}C-HMBC

Compound **52**

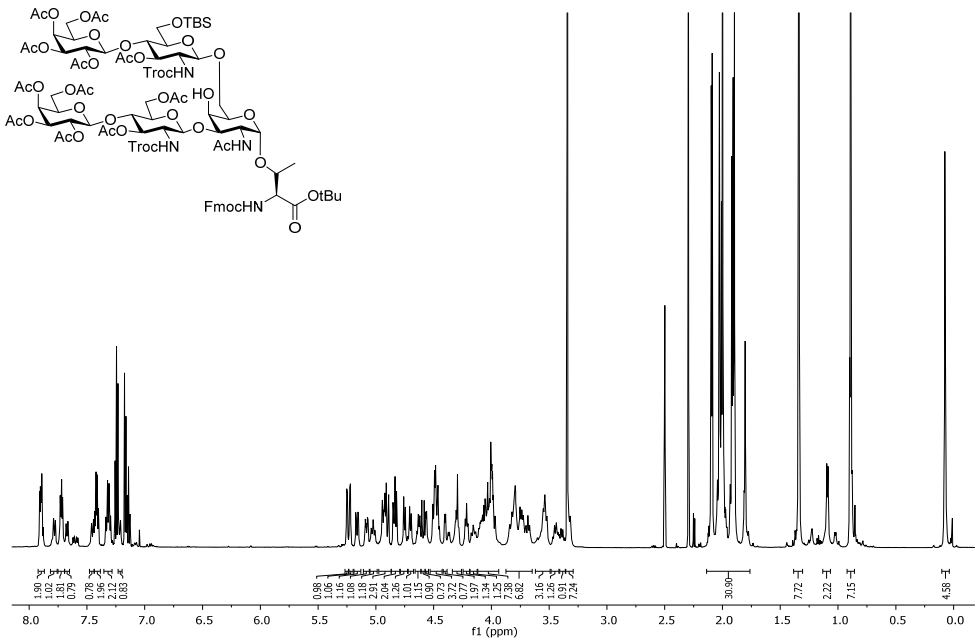


[1]H-NMR (700 MHz, DMSO)

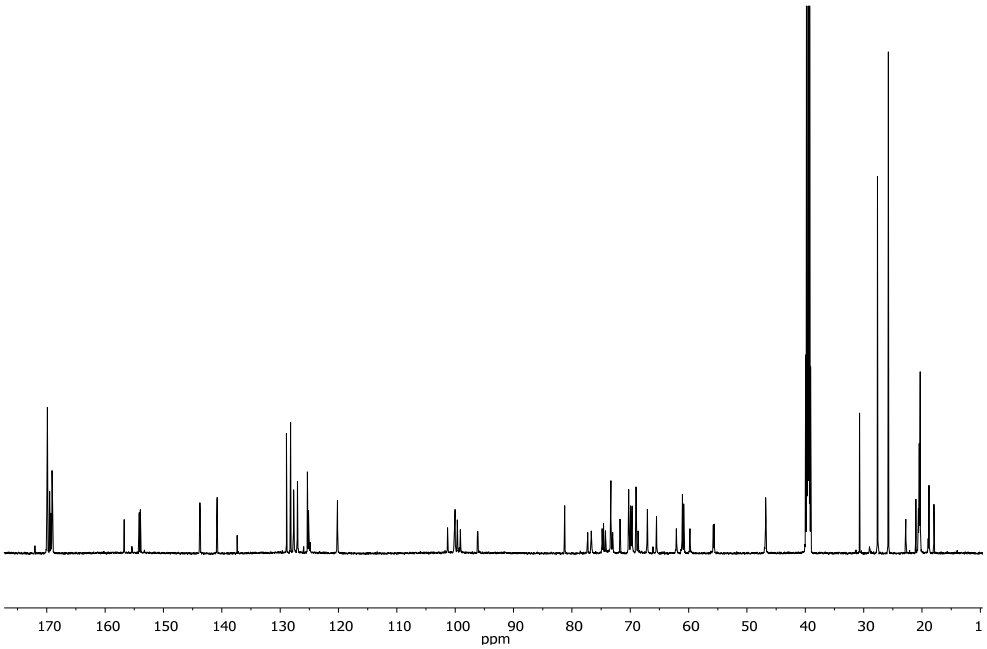


[13]C-NMR (176.0 MHz, DMSO)

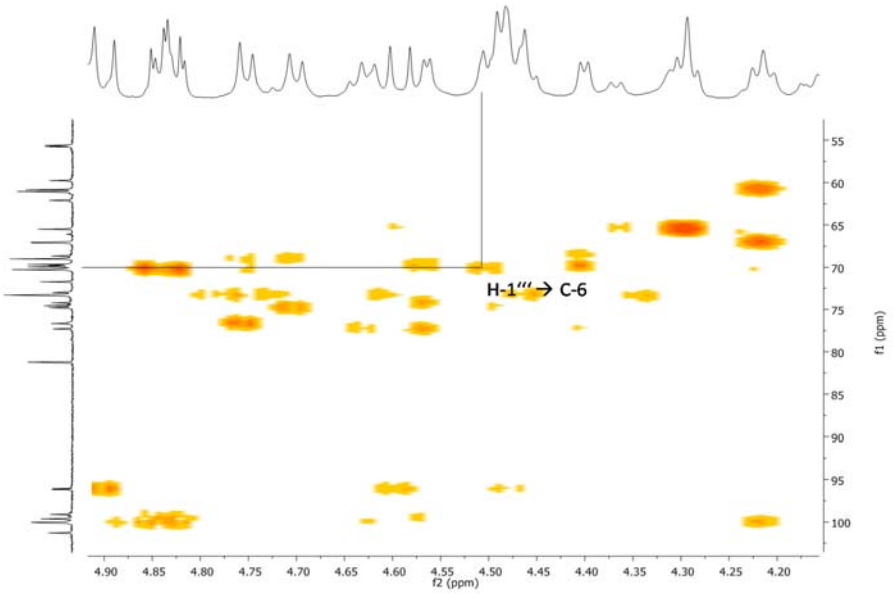
Compound **53**



[1]H-NMR (600 MHz, DMSO)

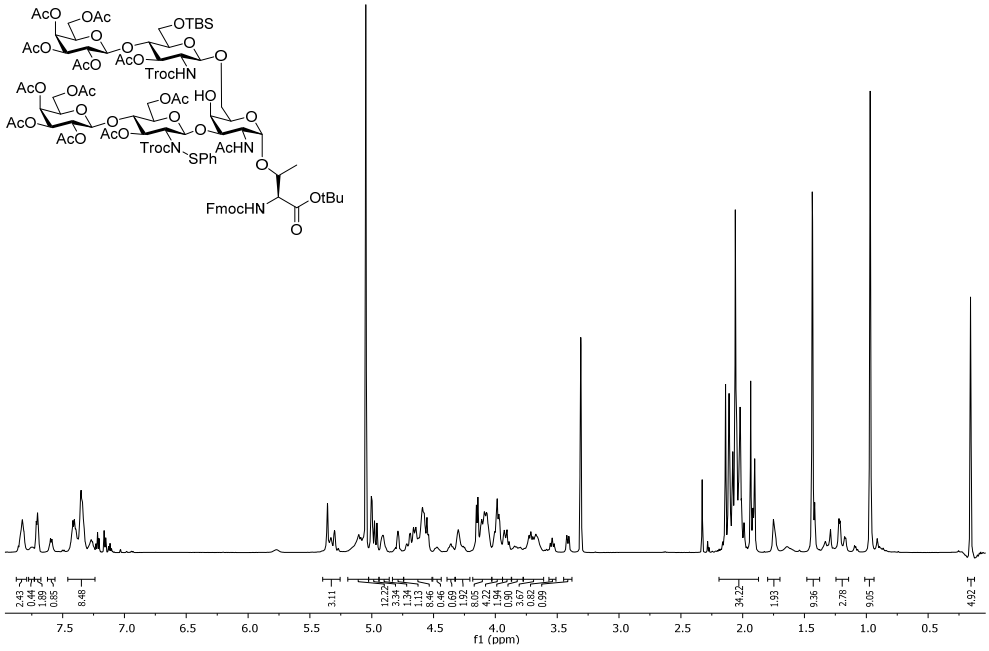


[13]C-NMR (150.9 MHz, DMSO)

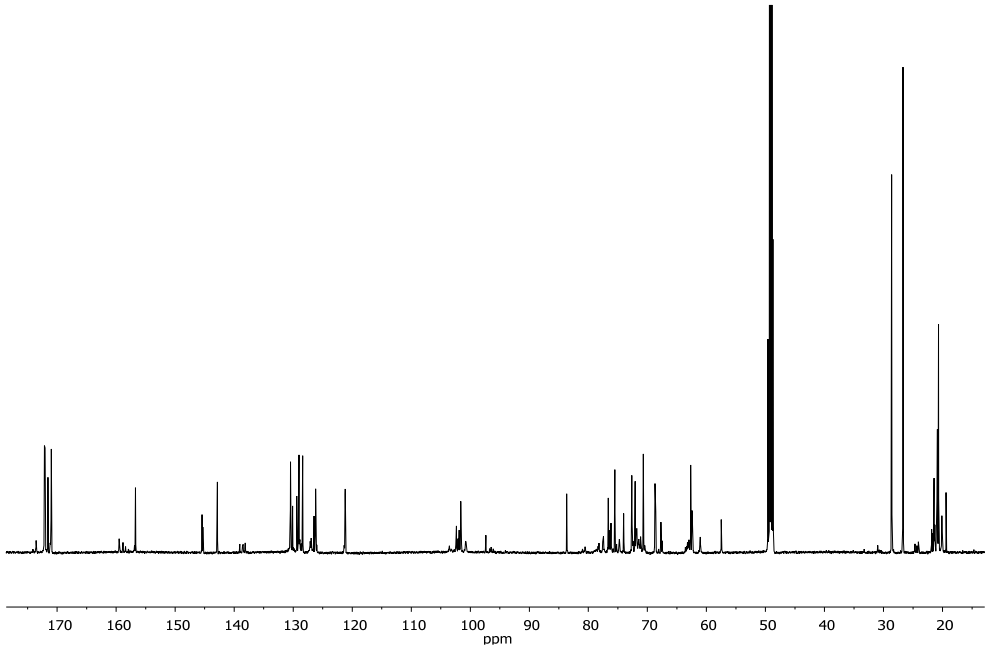

H-1‴→ C-6
f1 (ppm)
f2 (ppm)
4.90
4.85
4.80
4.75
4.70
4.65
4.60
4.55
4.50
4.45
4.40
4.35
4.30
4.25
4.20
55
60
65
70
75
80
85
90
95
100

^{1}H-^{13}C-HMBC

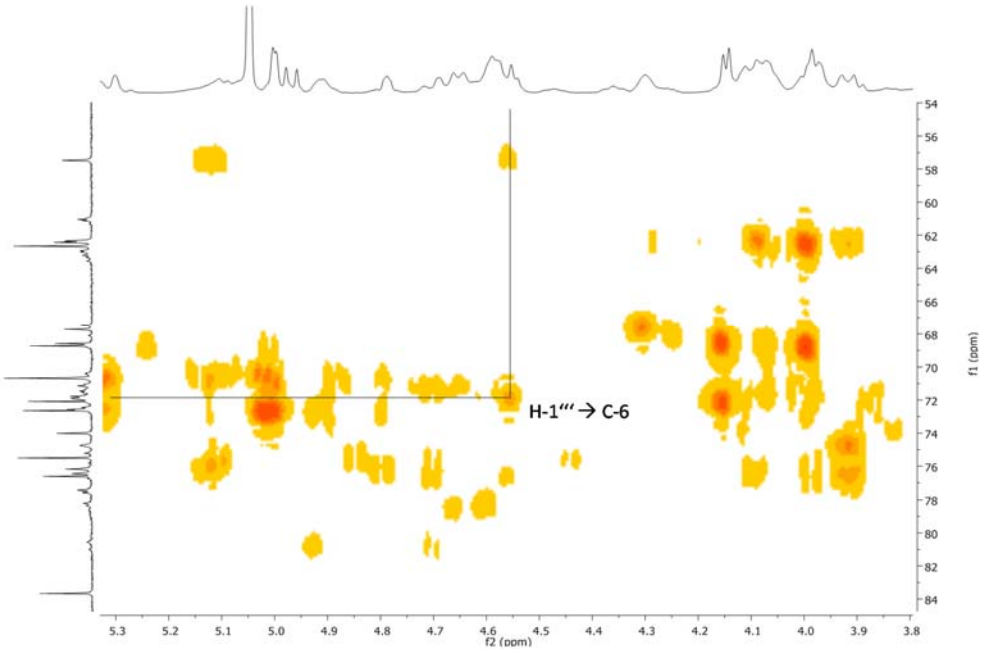
Compound **54**



[1]H-NMR (600 MHz, MeOD, T = 278 K)

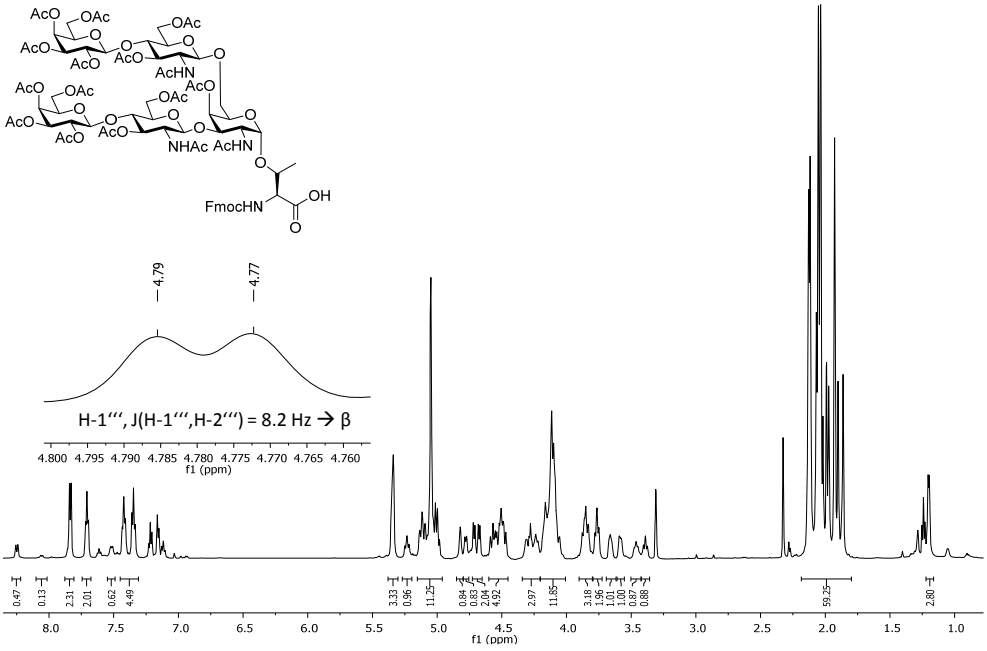


[13]C-NMR (150.9 MHz, MeOD, T = 278 K)

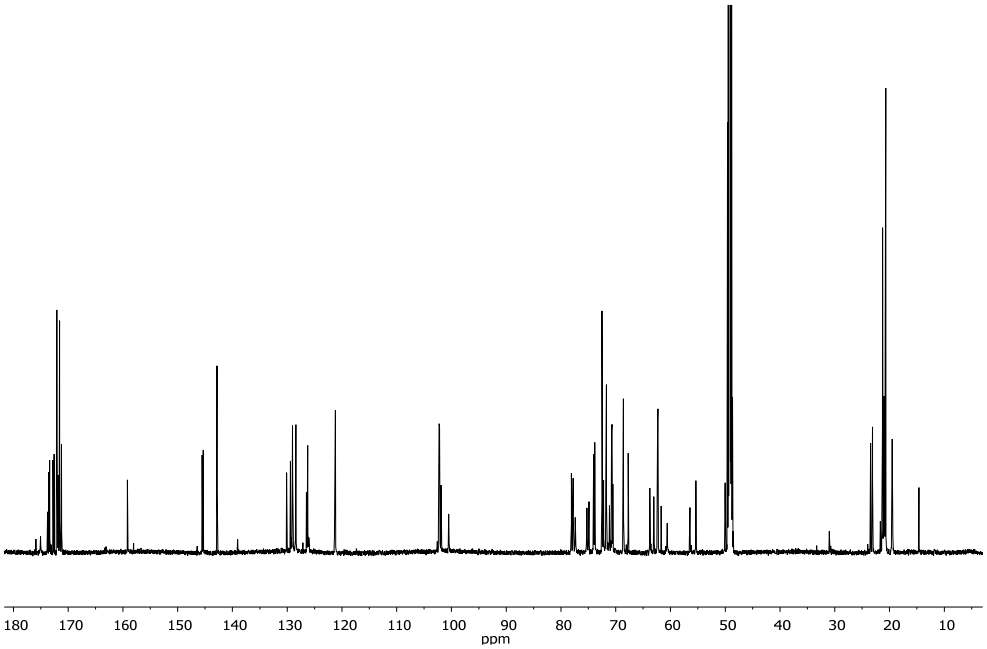


^{1}H-^{13}C-HMBC

Compound **58**

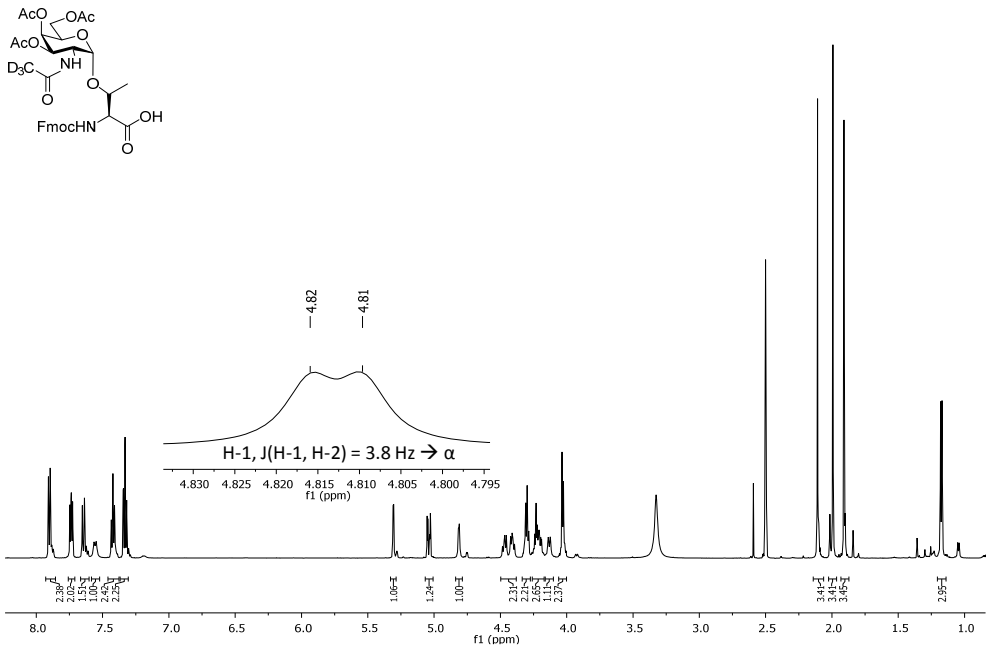


[1]H-NMR (600 MHz, MeOD, T = 278 K)

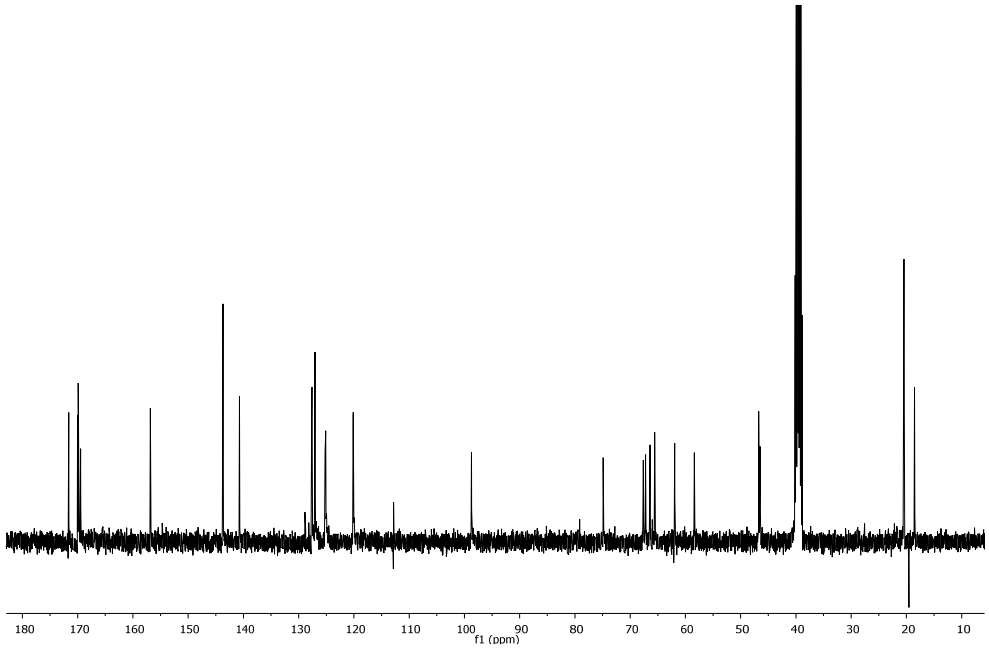


[13]C-NMR (150.9 MHz, MeOD, T = 278 K)

Compound **60**

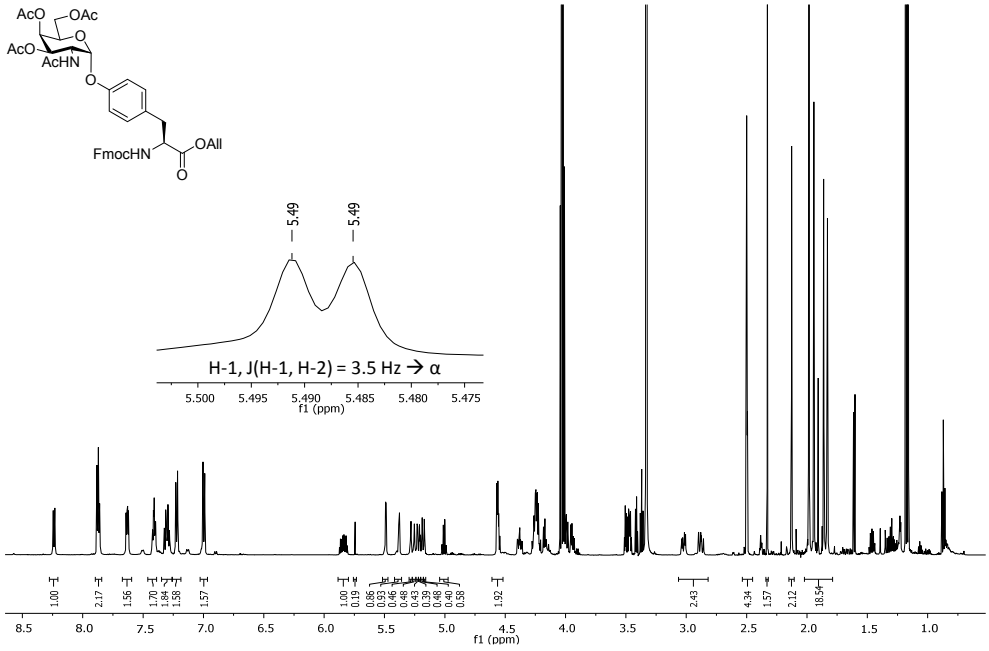


[1]H-NMR (600 MHz, DMSO)

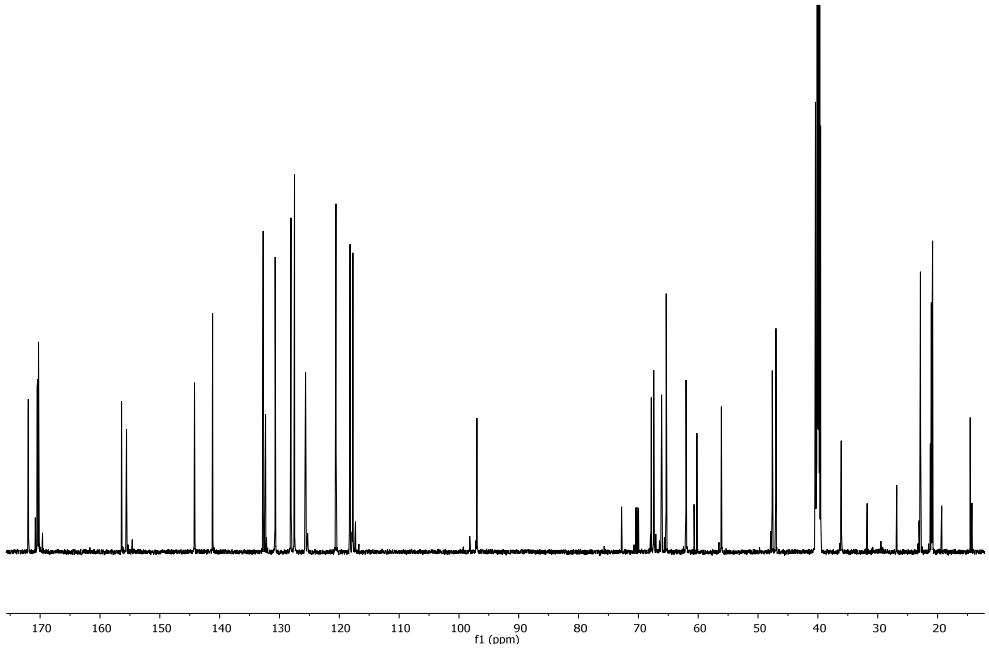


[13]C-NMR (100.6 MHz, DMSO)

Compound **63**

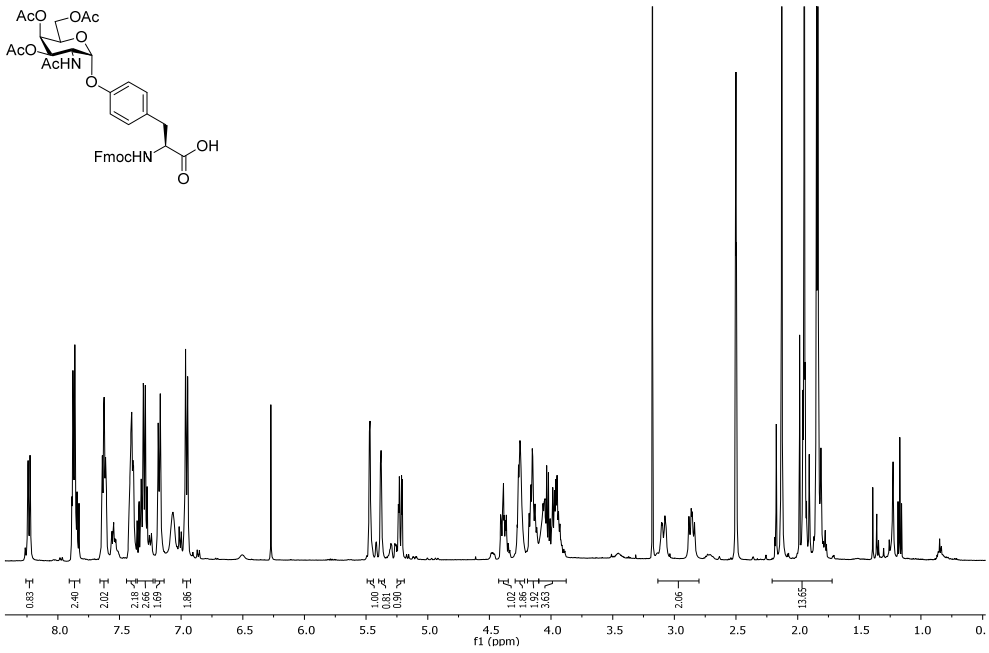


[1]H-NMR (600 MHz, DMSO)

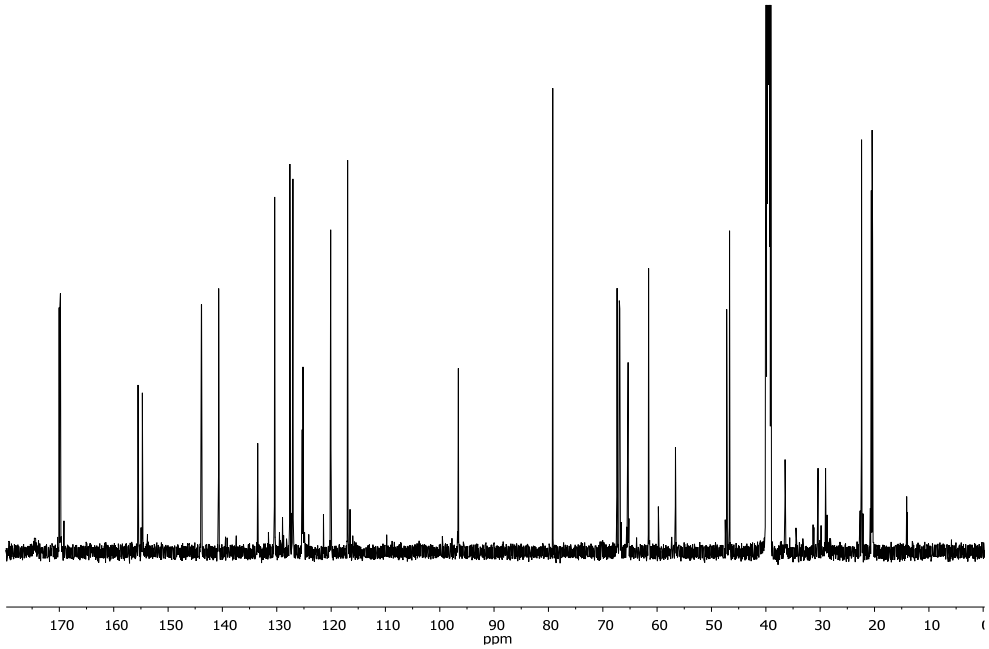


[13]C-NMR (150.9 MHz, DMSO)

Compound **64**

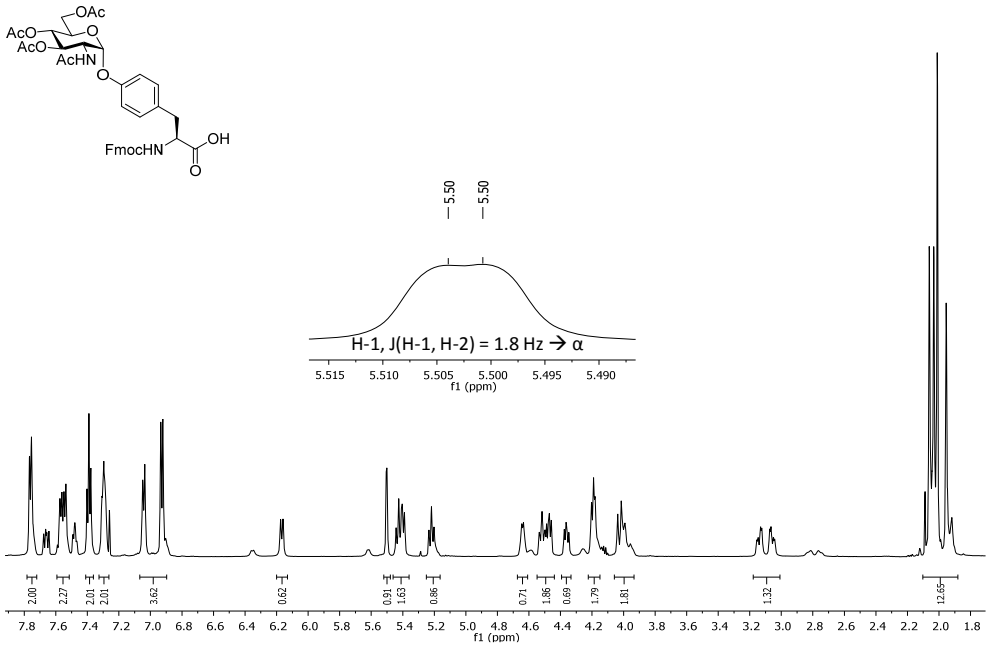


[1]H-NMR (500 MHz, DMSO)

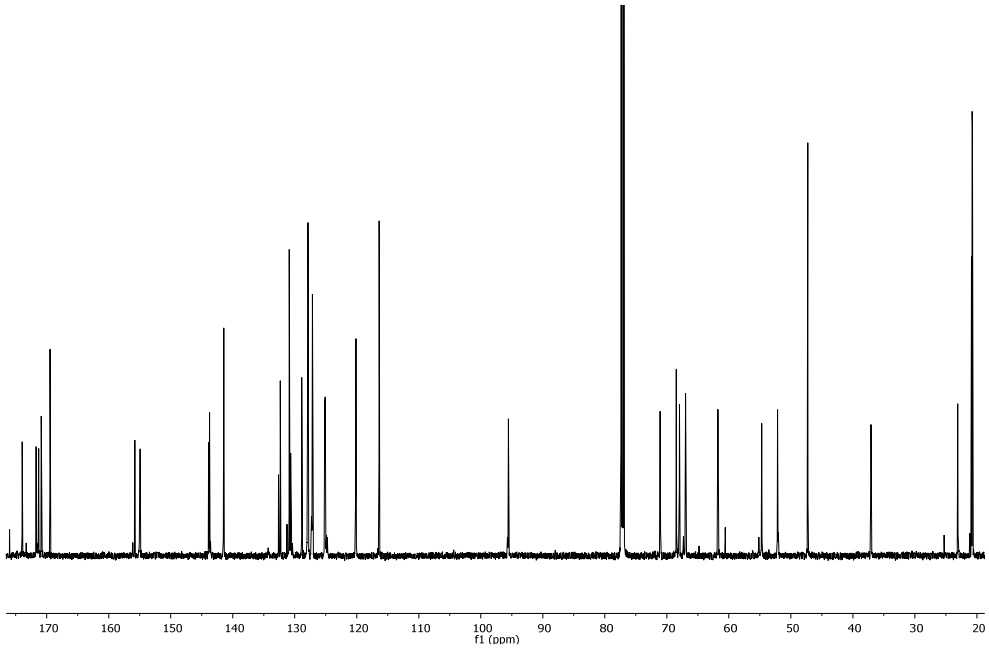


[13]C-NMR (150.9 MHz, DMSO)

Compound **69**

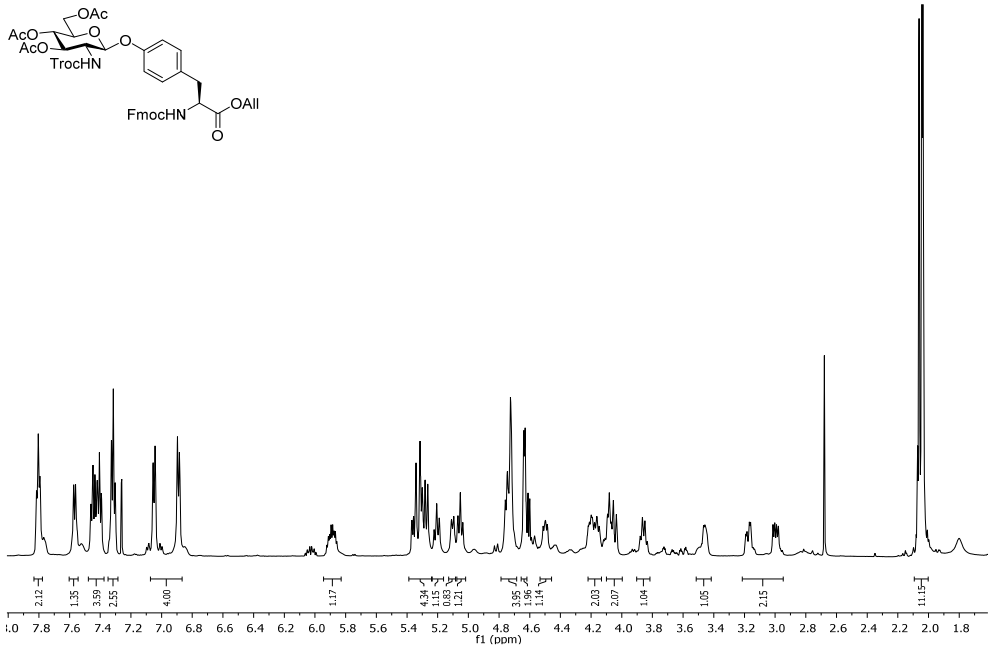


^{1}H-NMR (600 MHz, $CDCl_3$)

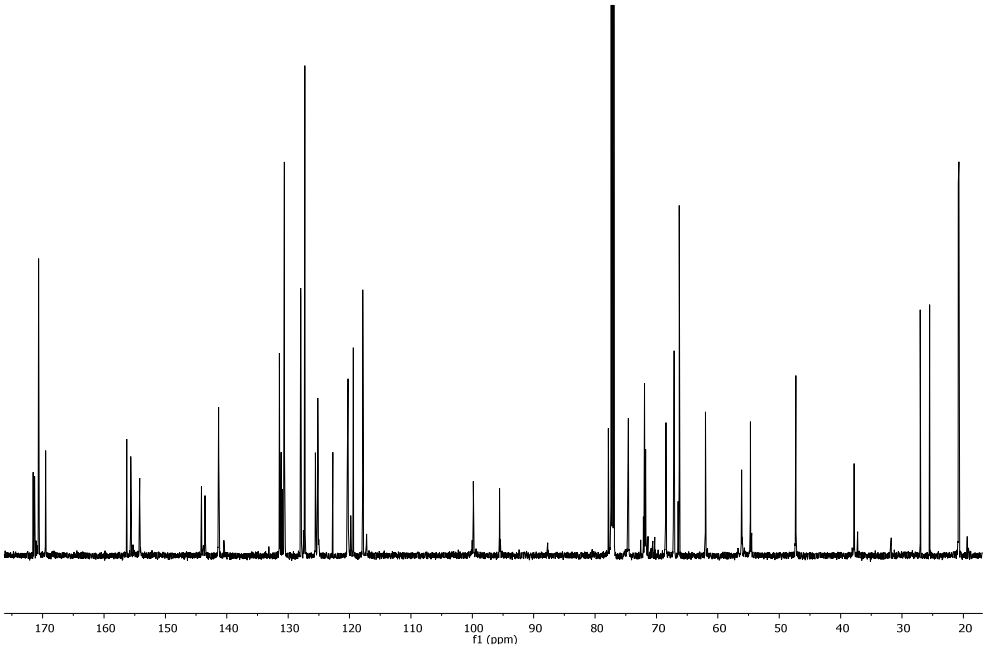


^{13}C-NMR (150.9 MHz)

Compound **72**

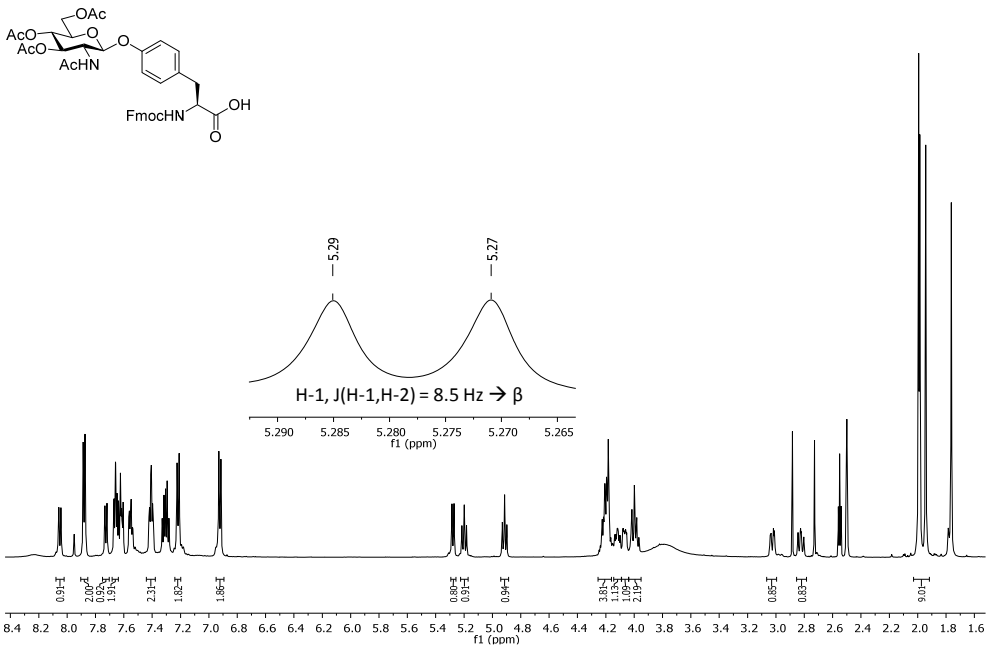


[1]H-NMR (600 MHz, $CDCl_3$)

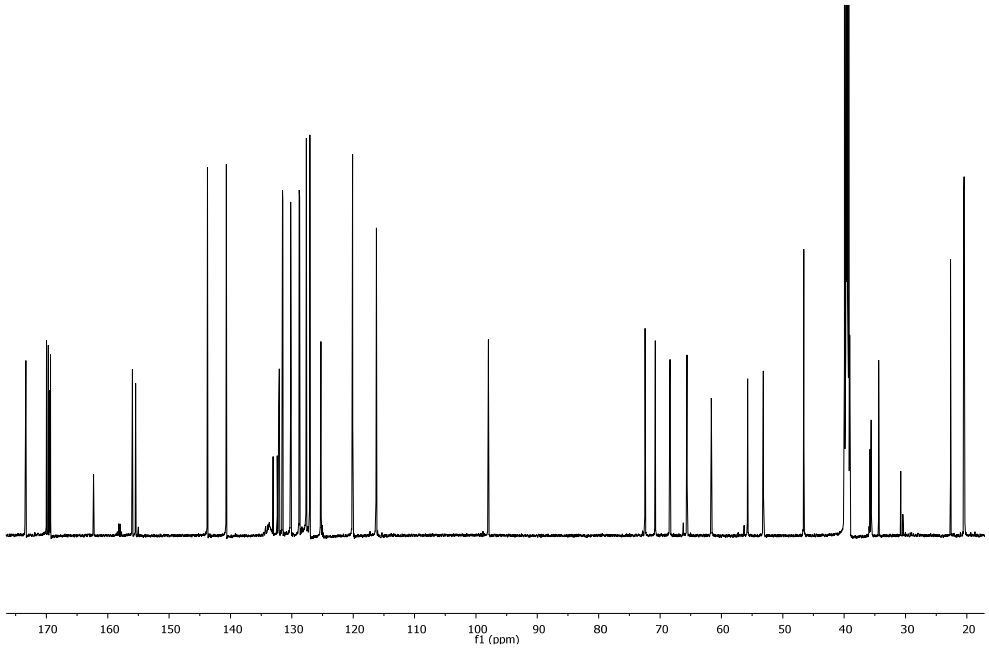


[13]C-NMR (150.9 MHz, $CDCl_3$)

Compound **74**



[1]H-NMR (600 MHz, DMSO)



[13]C-NMR (150.9 MHz, DMSO)

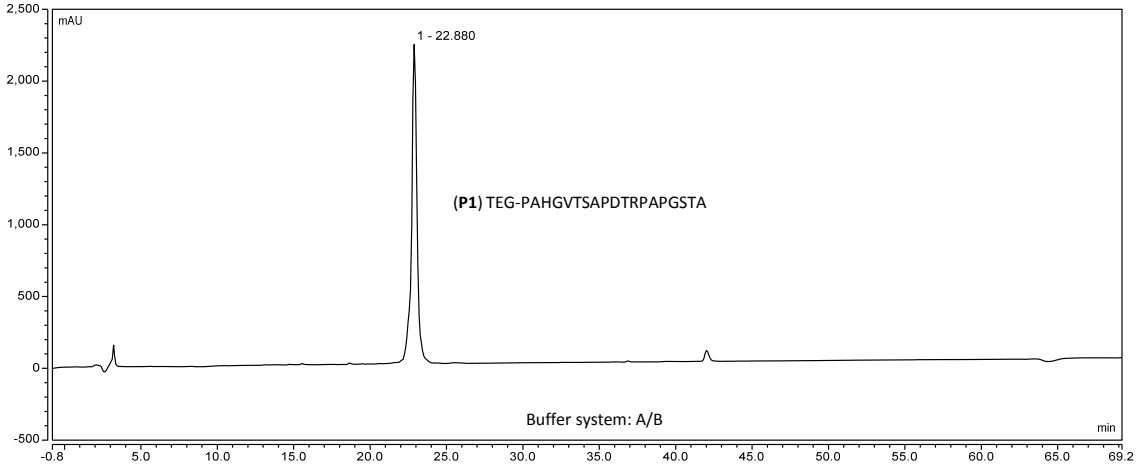
11.2 HPLC chromatograms

Buffer systems:

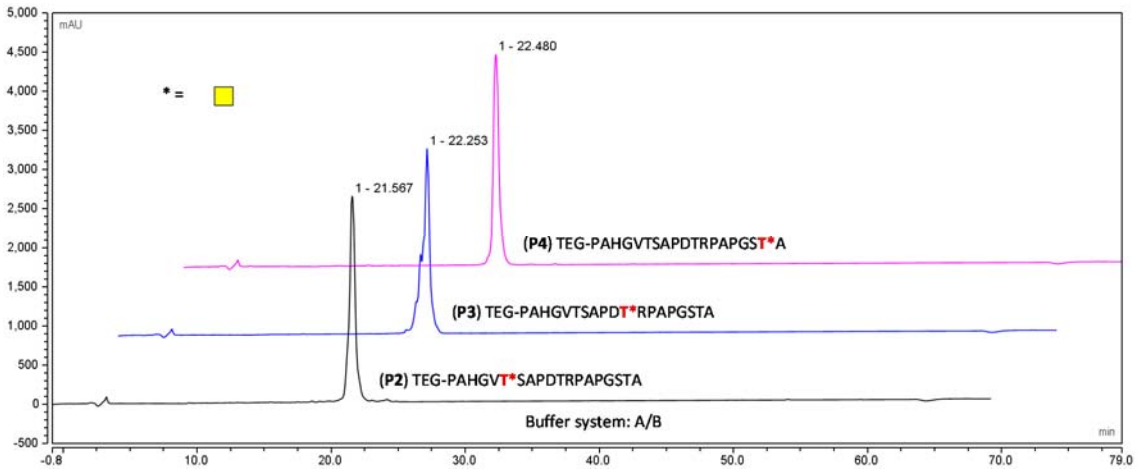
A: water + 0.1 % TFA, B: 84 % acetonitrile + 0.1 % TFA,

C: water + 0.1 % FA, D: 84 % acetonitrile + 0.1 % FA.

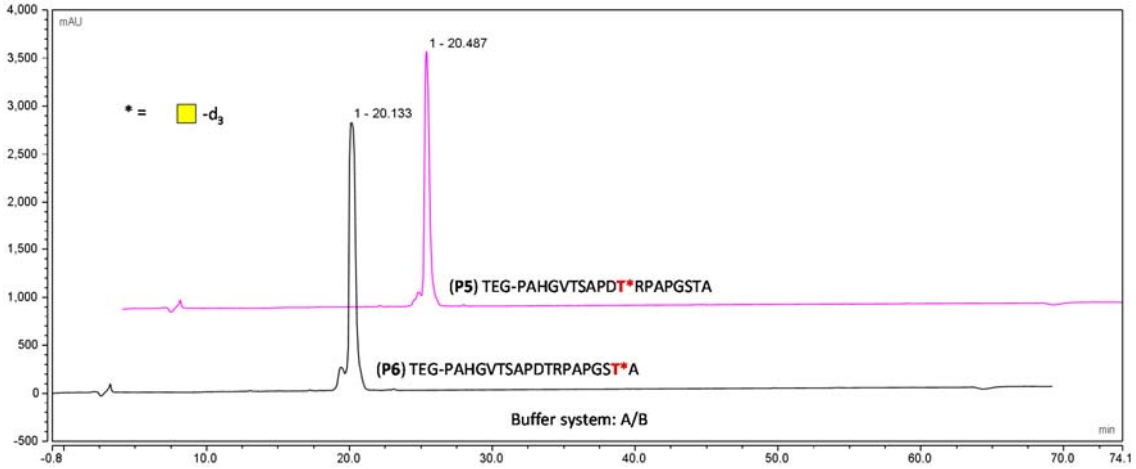
MUC1 (unglycosylated)



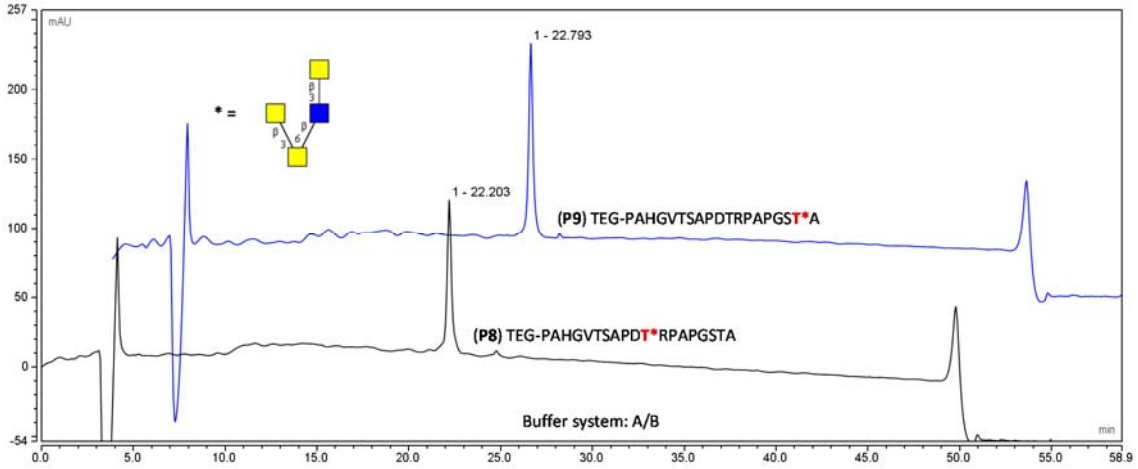
MUC1 (T_N)

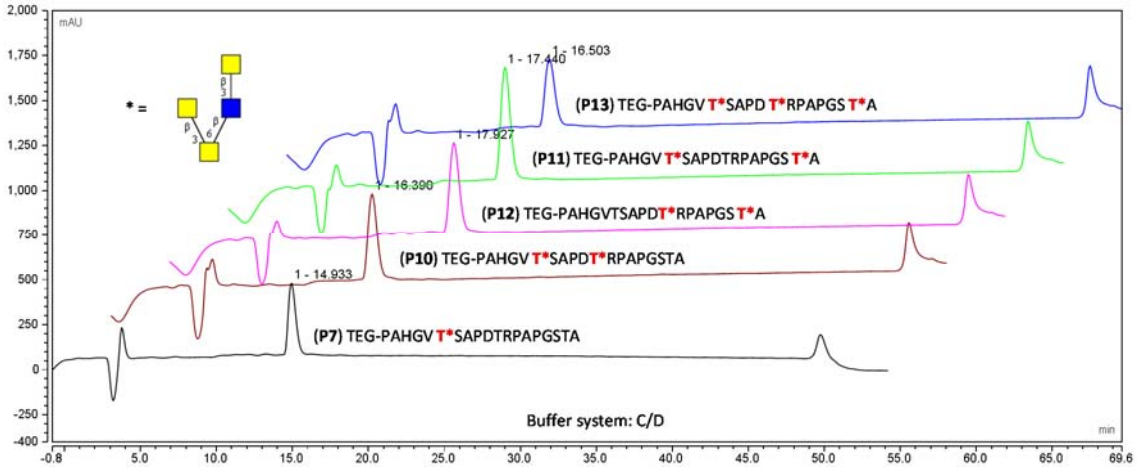


MUC1 (GalNAc-D_3)

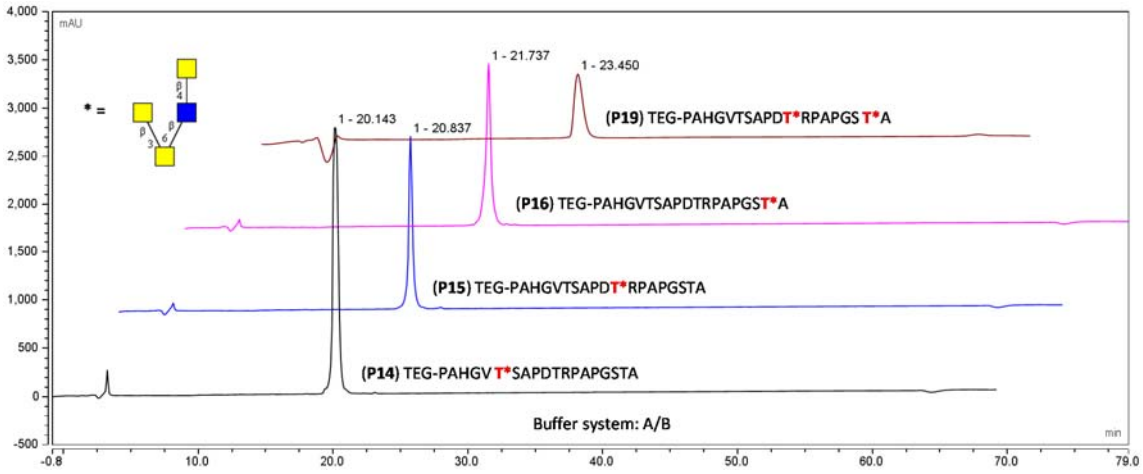


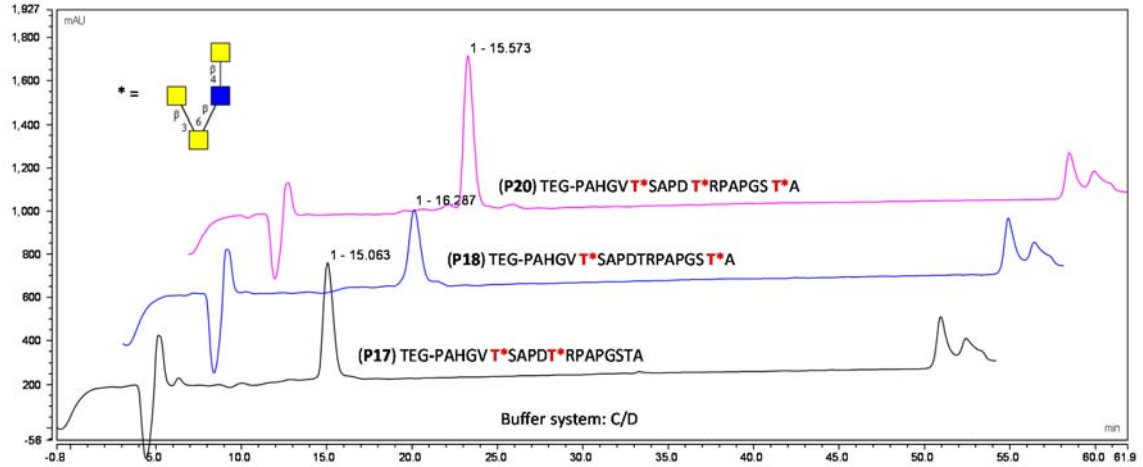
MUC1 (core 2 type-1)



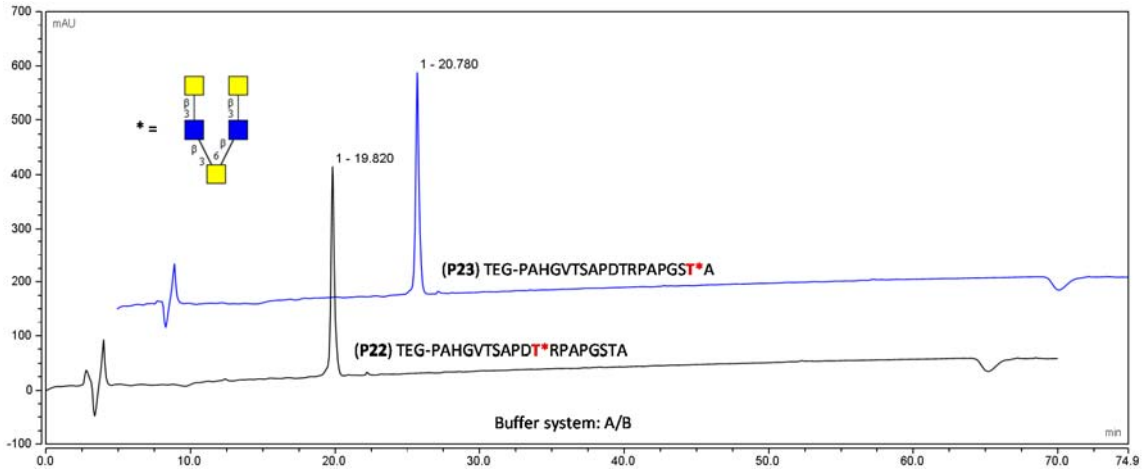


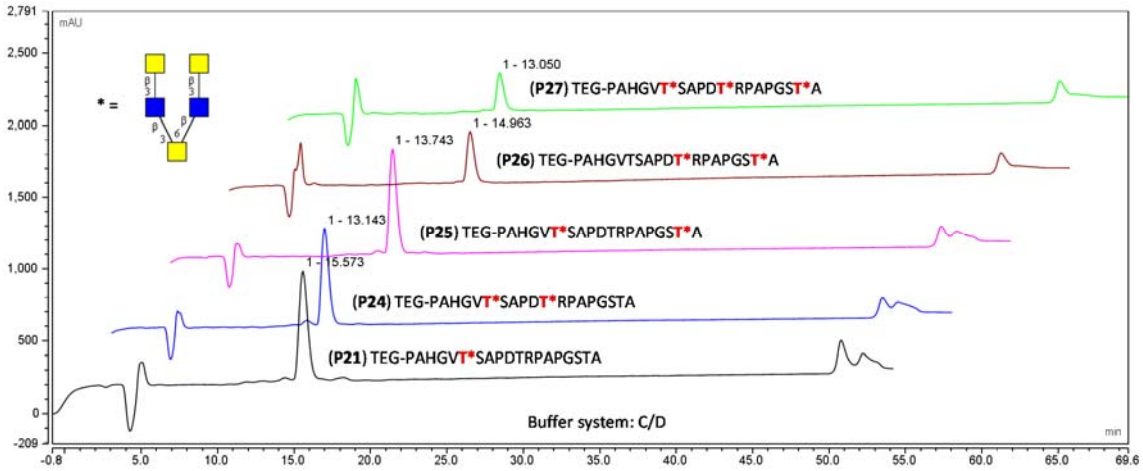
MUC1 (core 2 type-2)



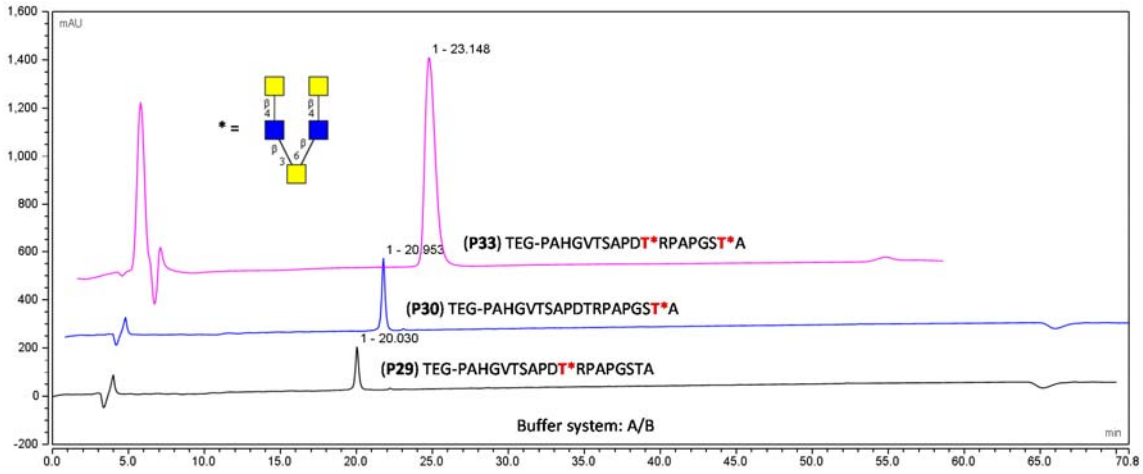


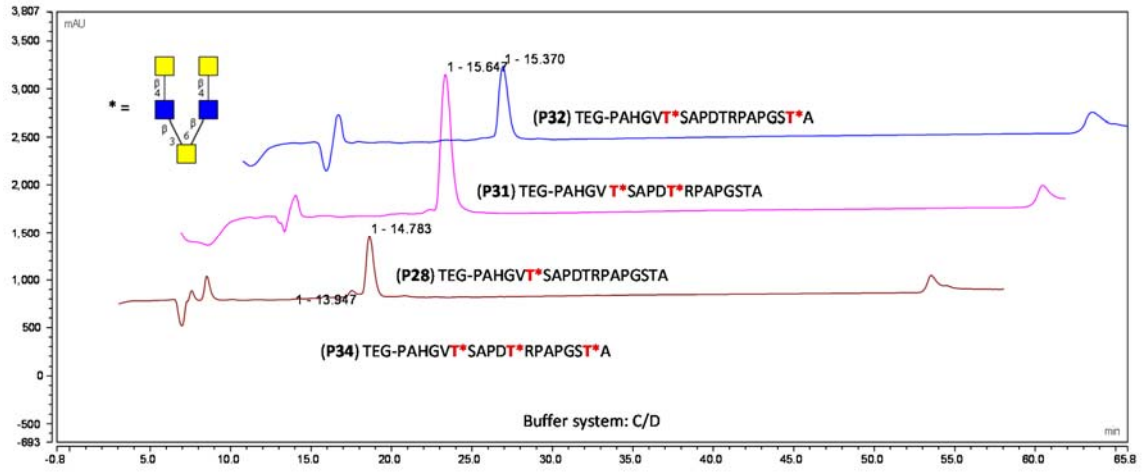
MUC1 (core 4 type-1)



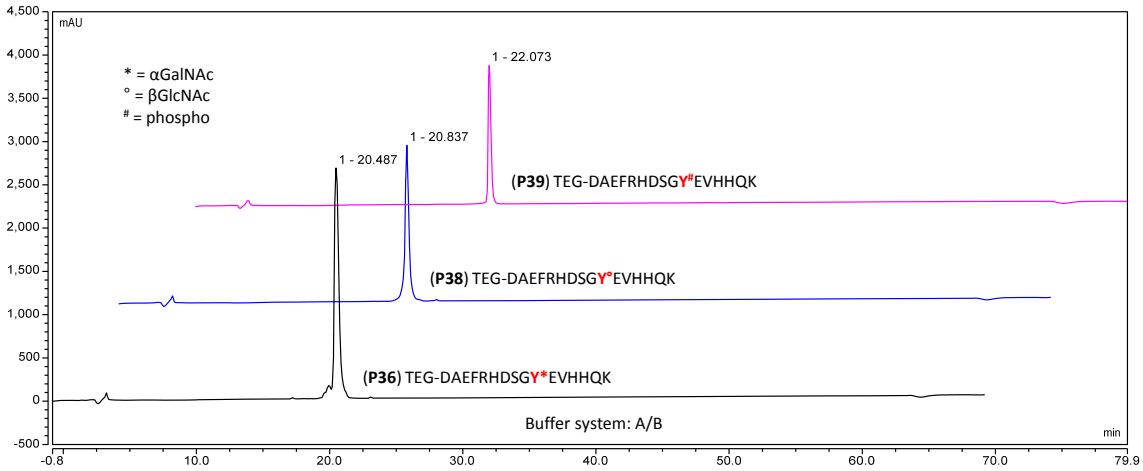


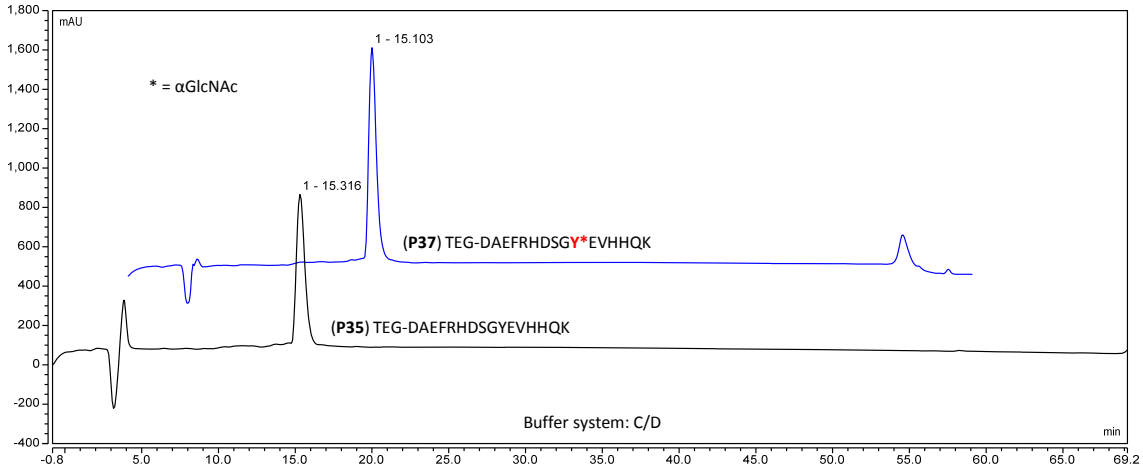
MUC1 (core 4 type-2)



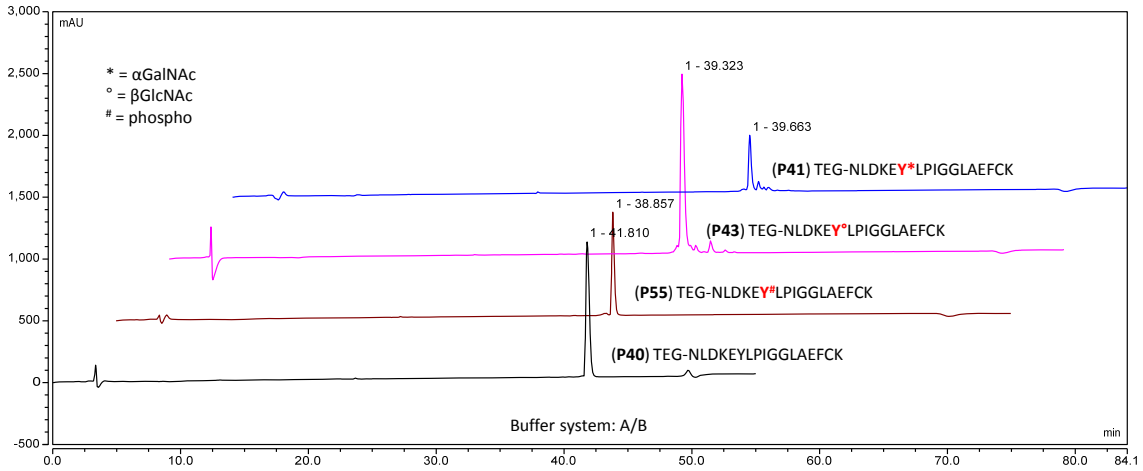


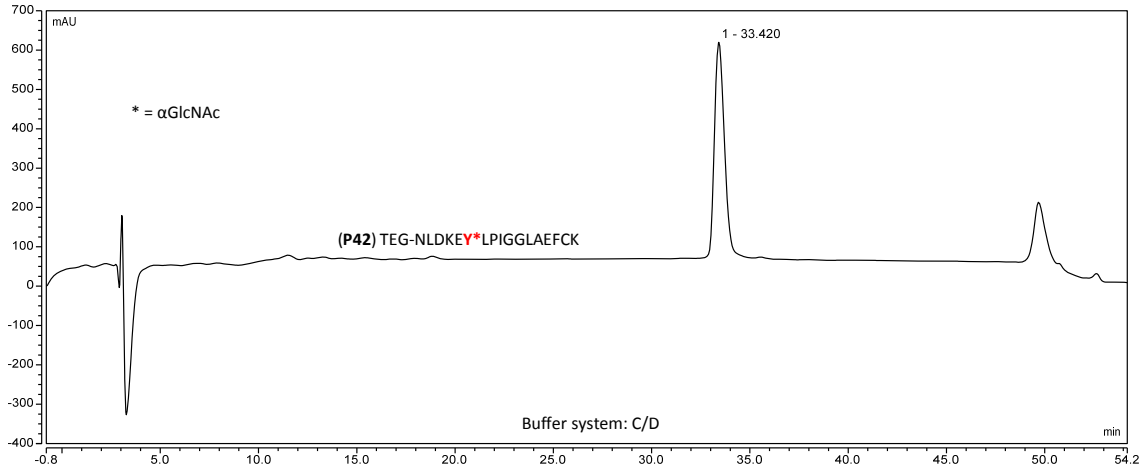
Amyloid-β 627-687



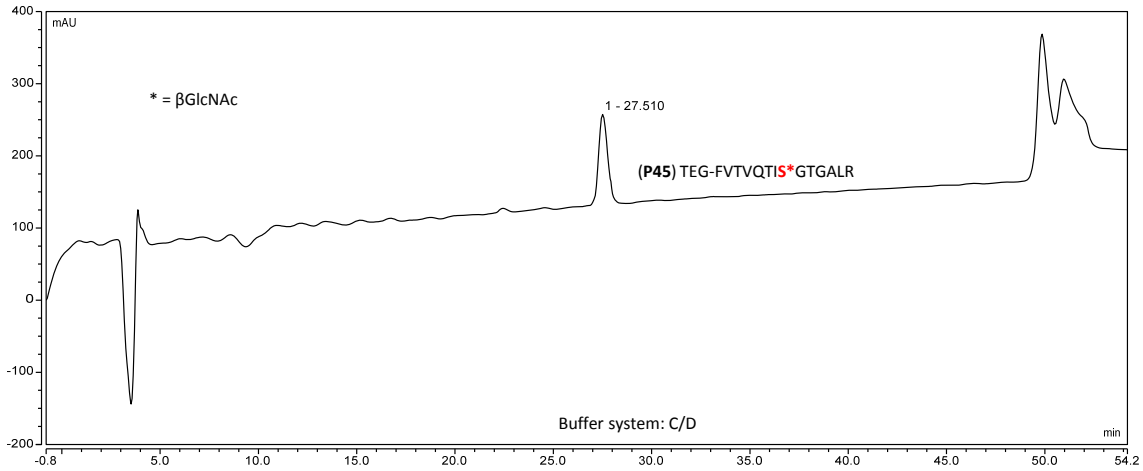


AspAT 91 – 107

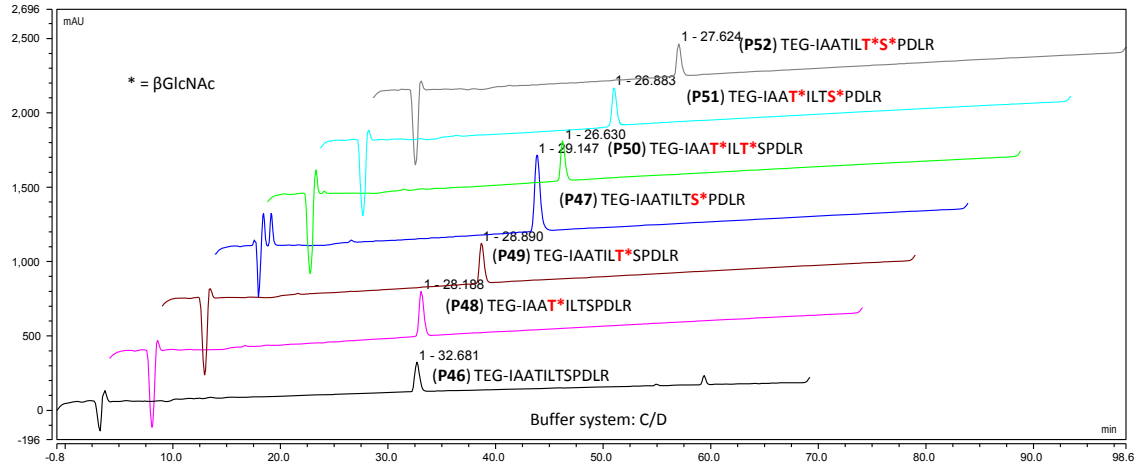

mAU
* = αGalNAc
° = βGlcNAc
= phospho
1 - 39.323
1 - 39.663
(P41) TEG-NLDKEY*LPIGGLAEFCK
1 - 38.857
1 - 41.810
(P43) TEG-NLDKEY°LPIGGLAEFCK
(P55) TEG-NLDKEY#LPIGGLAEFCK
(P40) TEG-NLDKEYLPIGGLAEFCK
Buffer system: A/B
min


mAU
* = αGlcNAc
1 - 33.420
(P42) TEG-NLDKEY*LPIGGLAEFCK
Buffer system: C/D
min

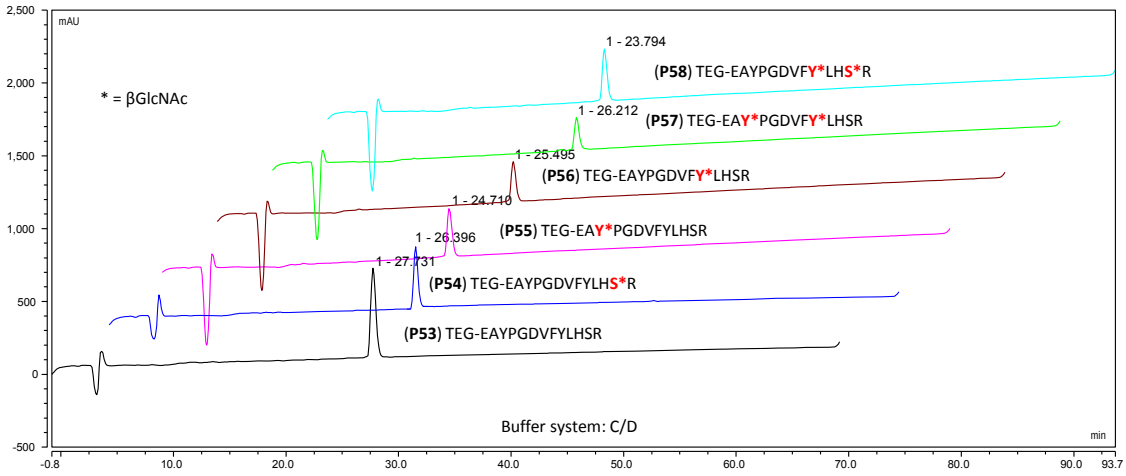
AspAT 127 – 139



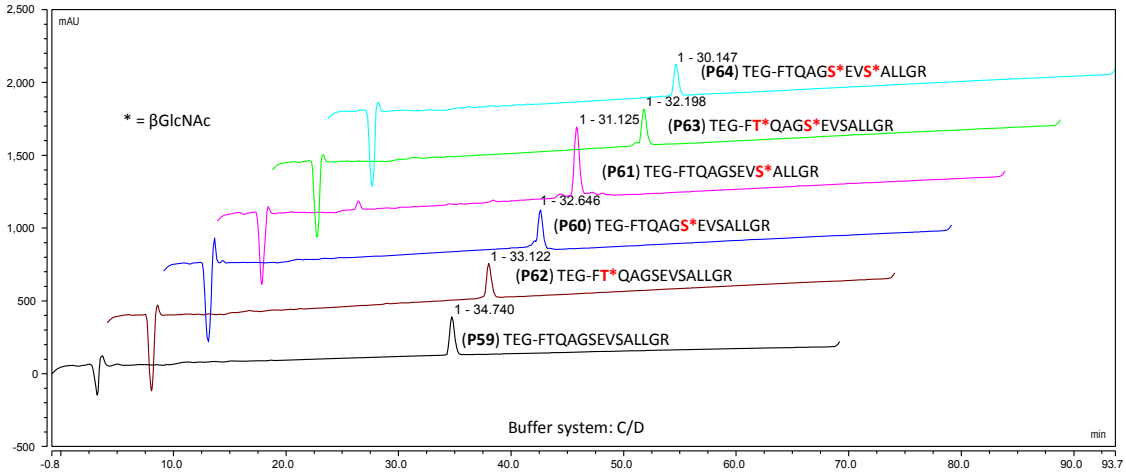
AspAT 326 – 337



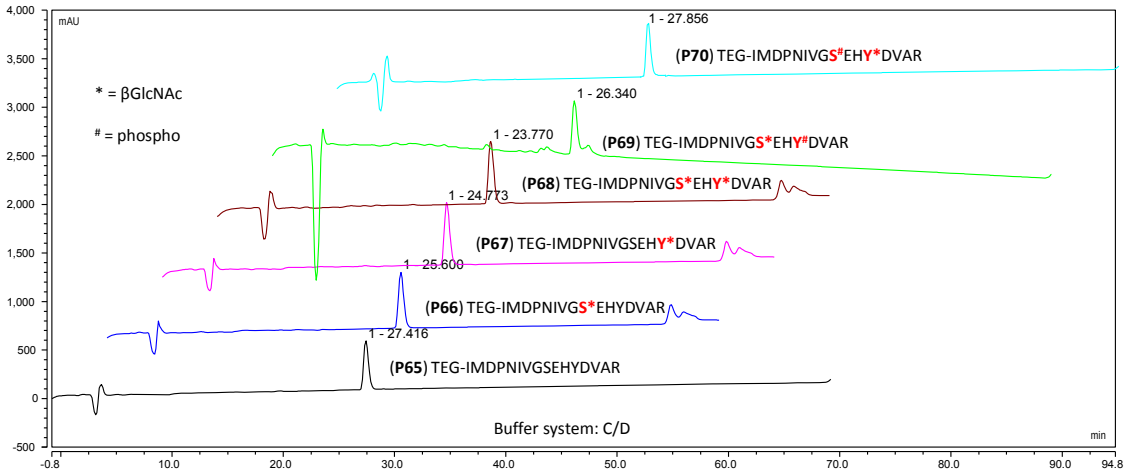
Atp5a 335 – 347



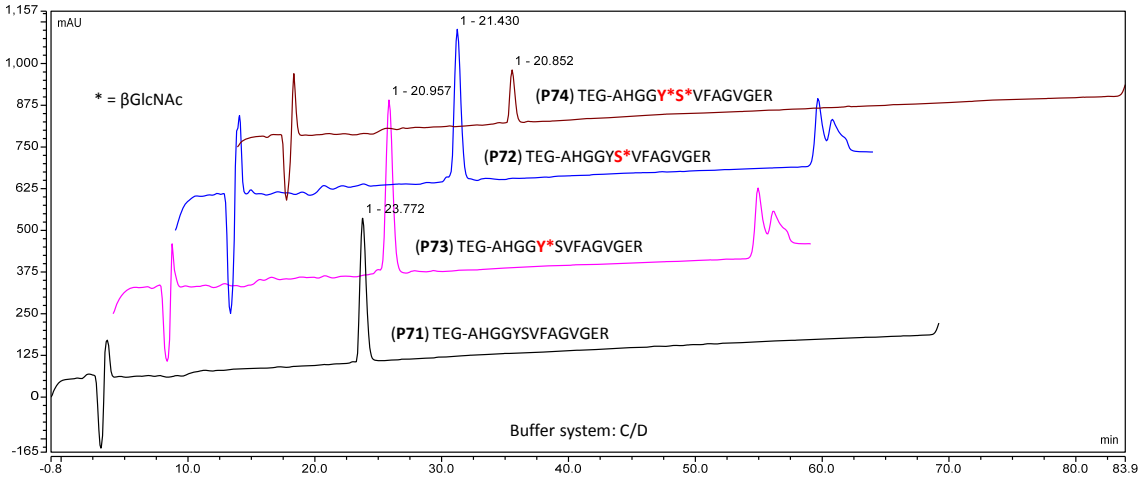
Atp5b 311 – 324



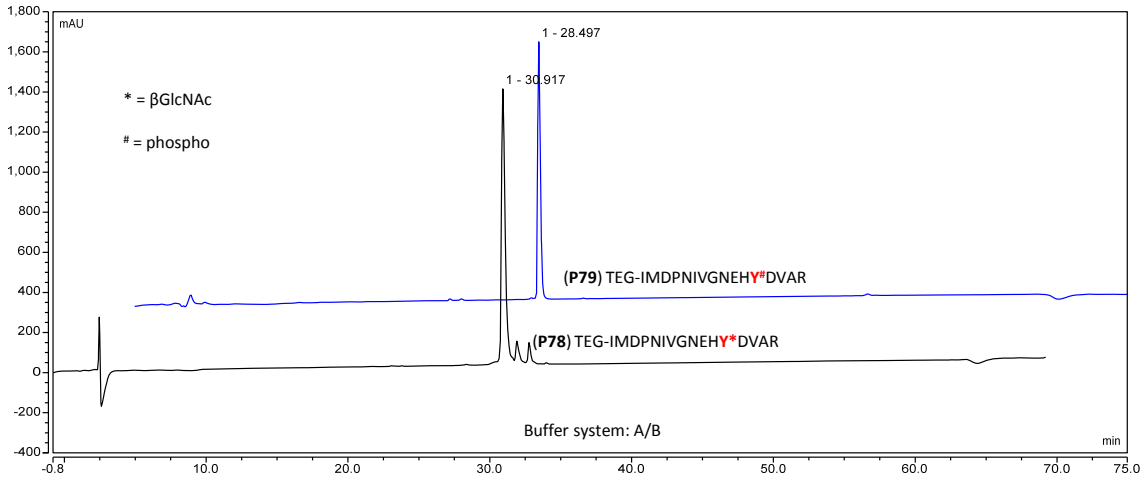
Atp5b 407 – 422 (human)

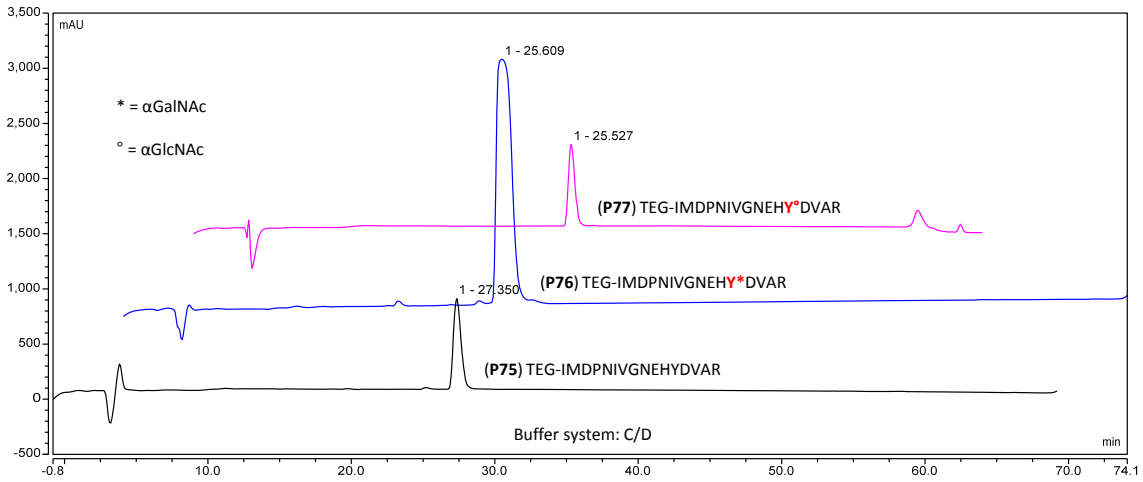


Atp5b 226 – 239

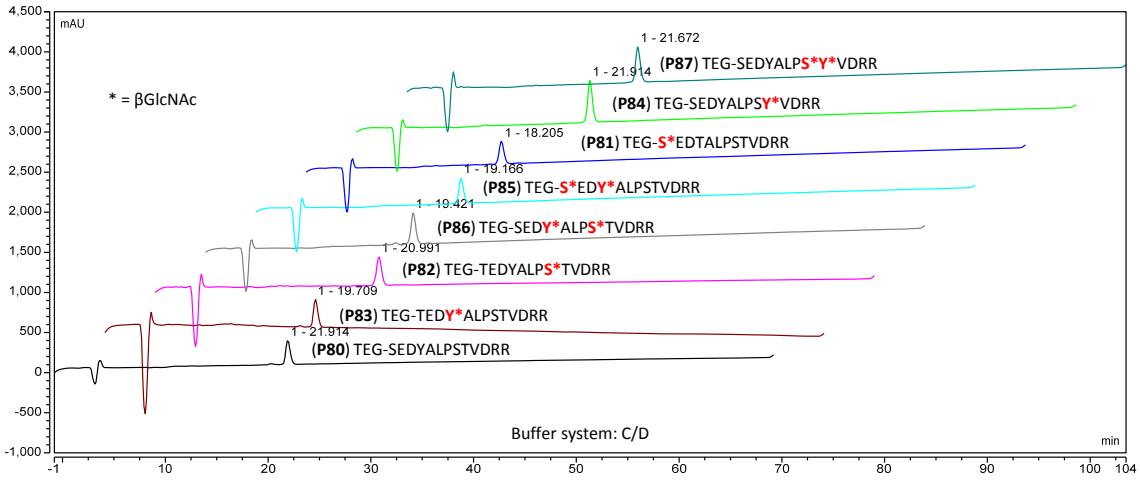


Atp5b 402 – 422 (mouse)

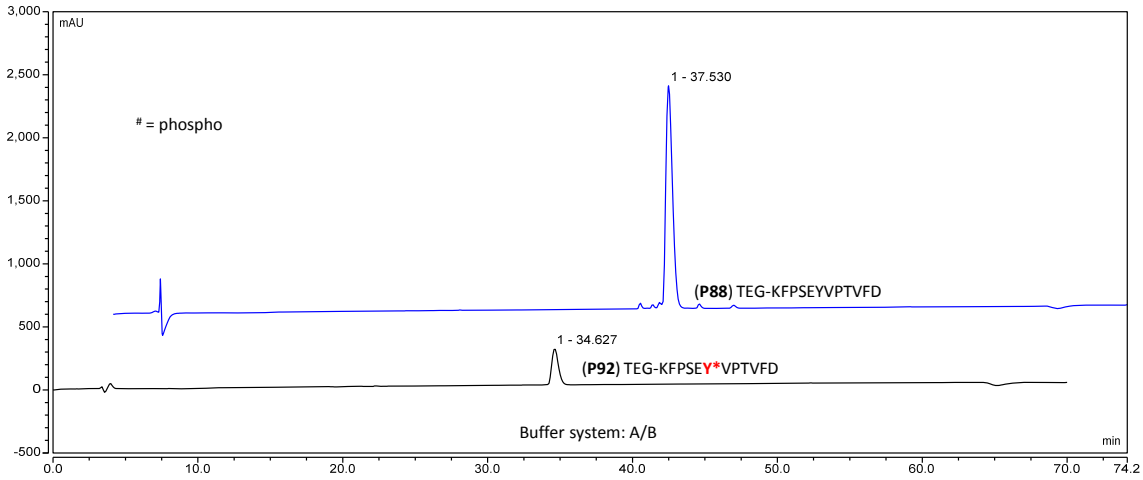


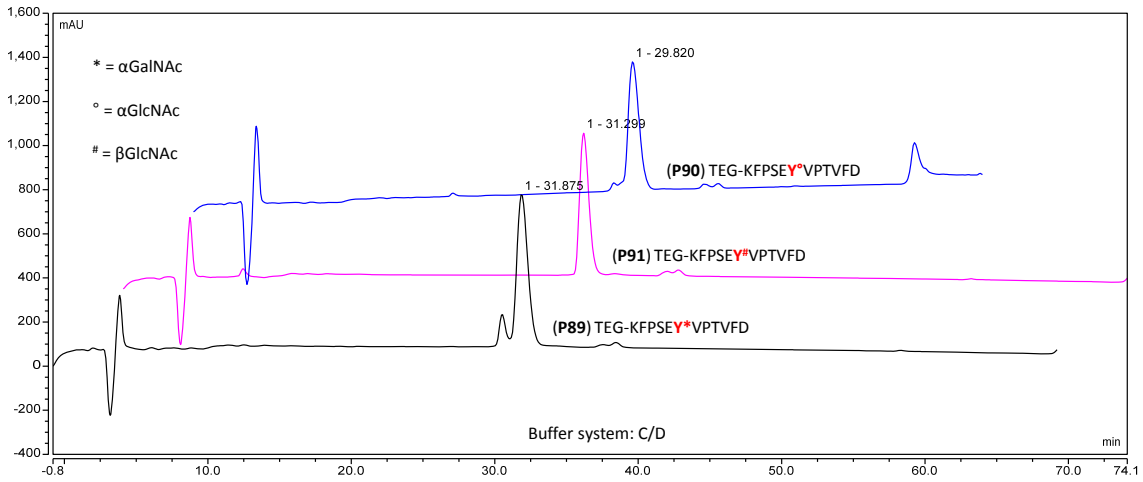


Cox4i1 30 – 42

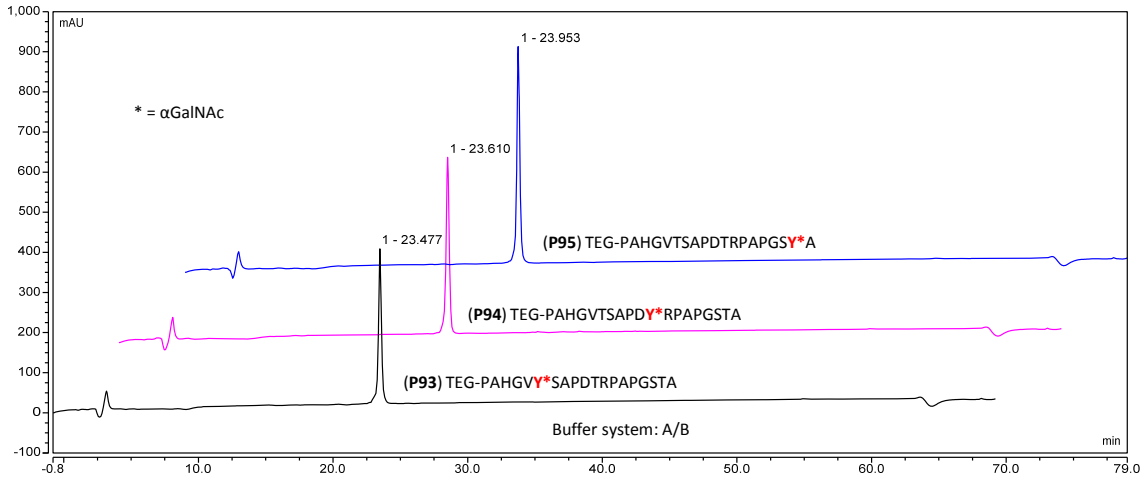

mAU
* = βGlcNAc
1 - 21.672
(P87) TEG-SEDYALPS*Y*VDRR
1 - 21.914
(P84) TEG-SEDYALPSY*VDRR
1 - 18.205
(P81) TEG-S*EDTALPSTVDRR
1 - 19.166
(P85) TEG-S*EDY*ALPSTVDRR
1 - 19.421
(P86) TEG-SEDY*ALPS*TVDRR
1 - 20.991
(P82) TEG-TEDYALPS*TVDRR
1 - 19.709
(P83) TEG-TEDY*ALPSTVDRR
1 - 21.914
(P80) TEG-SEDYALPSTVDRR
Buffer system: C/D
min

Cdc42 27 – 38

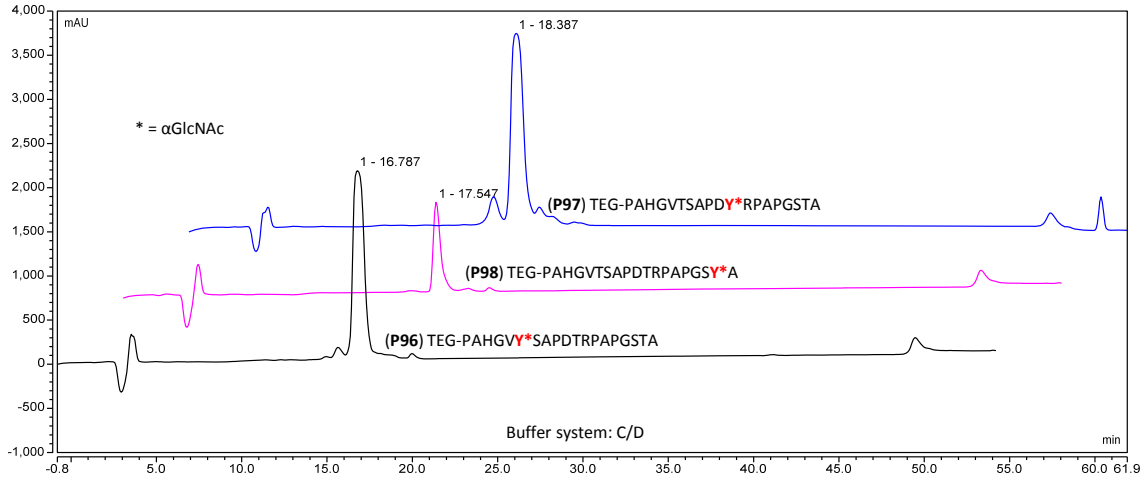

mAU
= phospho
1 - 37.530
(P88) TEG-KFPSEYVPTVFD
1 - 34.627
(P92) TEG-KFPSEY*VPTVFD
Buffer system: A/B
min


mAU
* = αGalNAc
° = αGlcNAc
= βGlcNAc
1 - 29.820
1 - 31.299
1 - 31.875
(P90) TEG-KFPSEY°VPTVFD
(P91) TEG-KFPSEY#VPTVFD
(P89) TEG-KFPSEY*VPTVFD
Buffer system: C/D
min

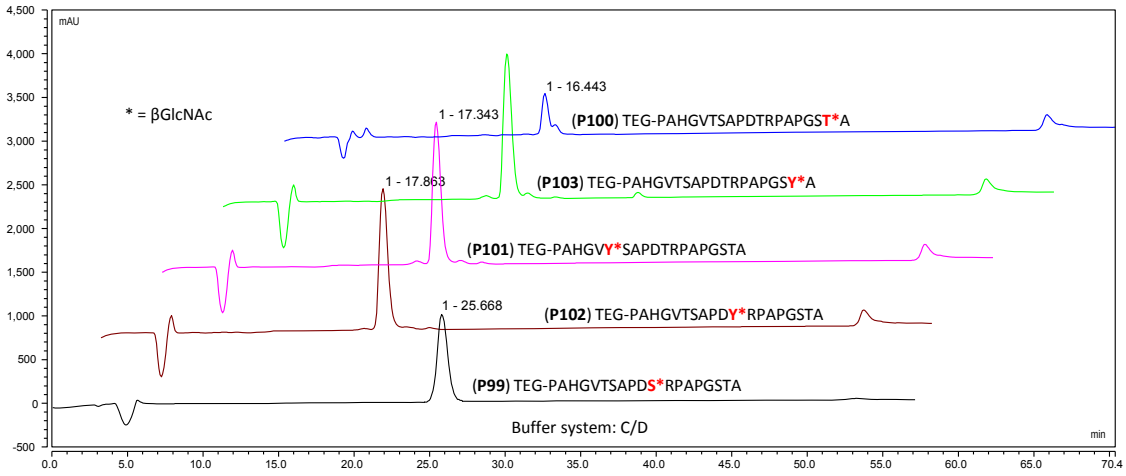
MUC1 αGalNAc-Tyr



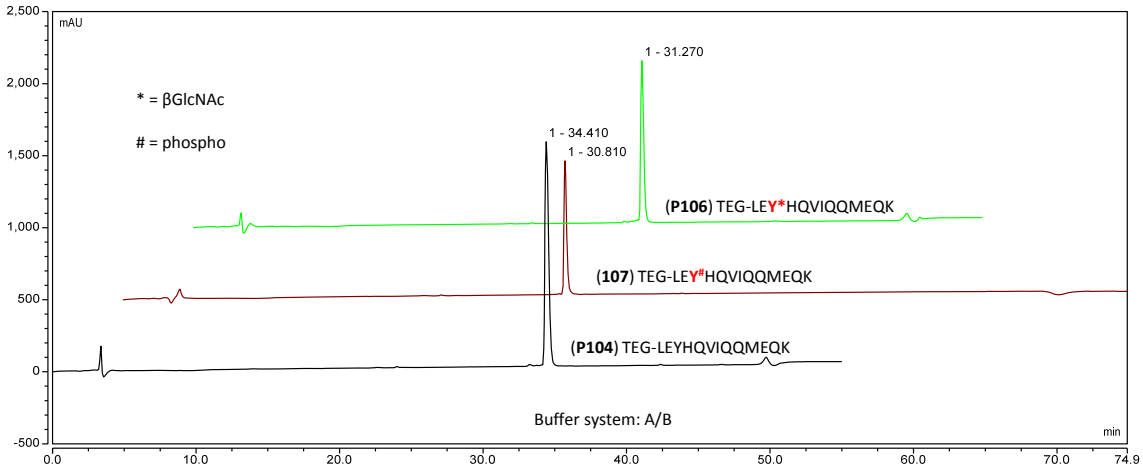
MUC1 αGlcNAc-Tyr

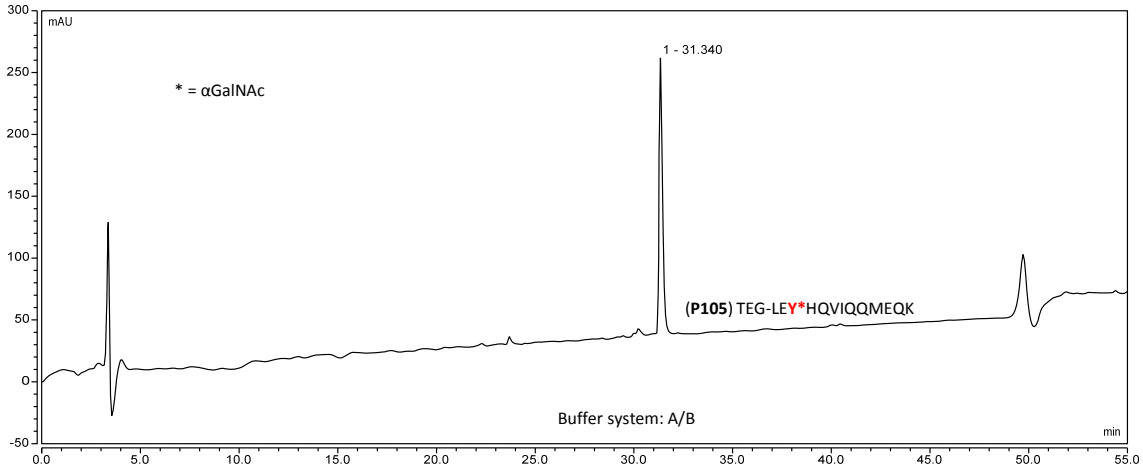


MUC1 βGlcNAc

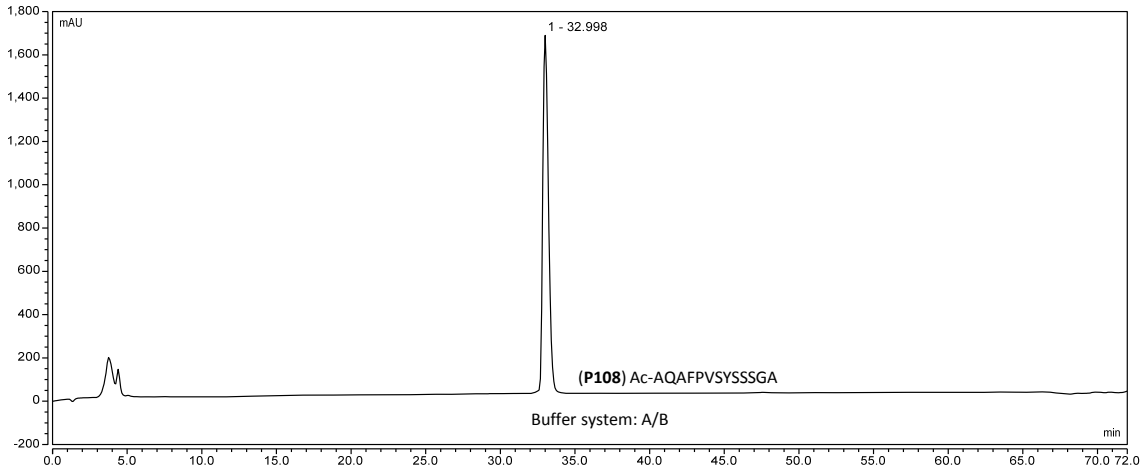


NucB2 387 – 399

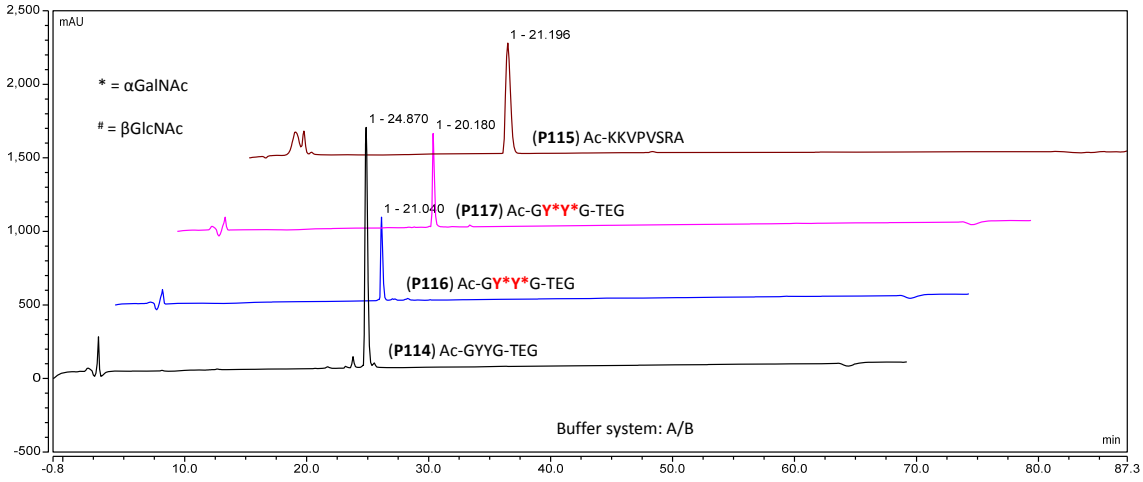




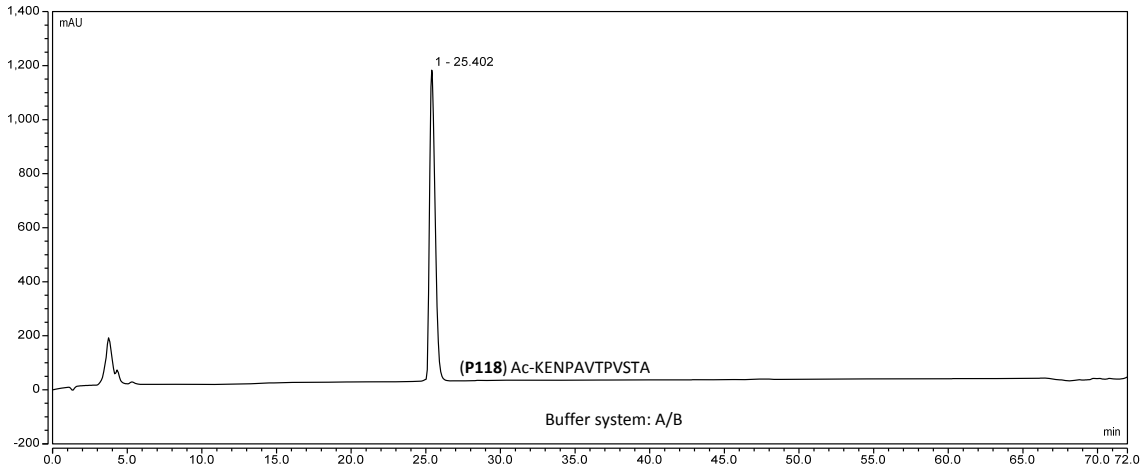
Ret 680 – 692



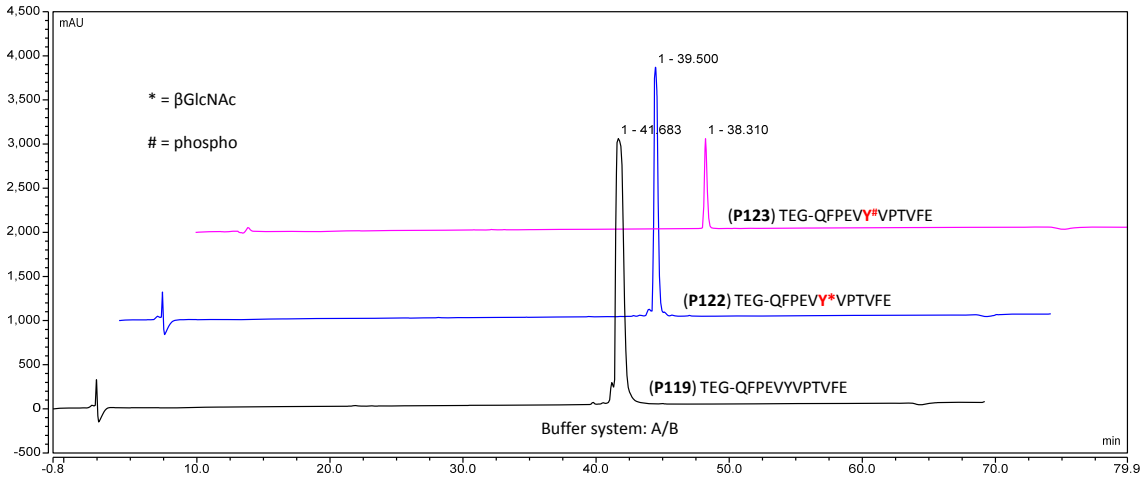
Rational constructs

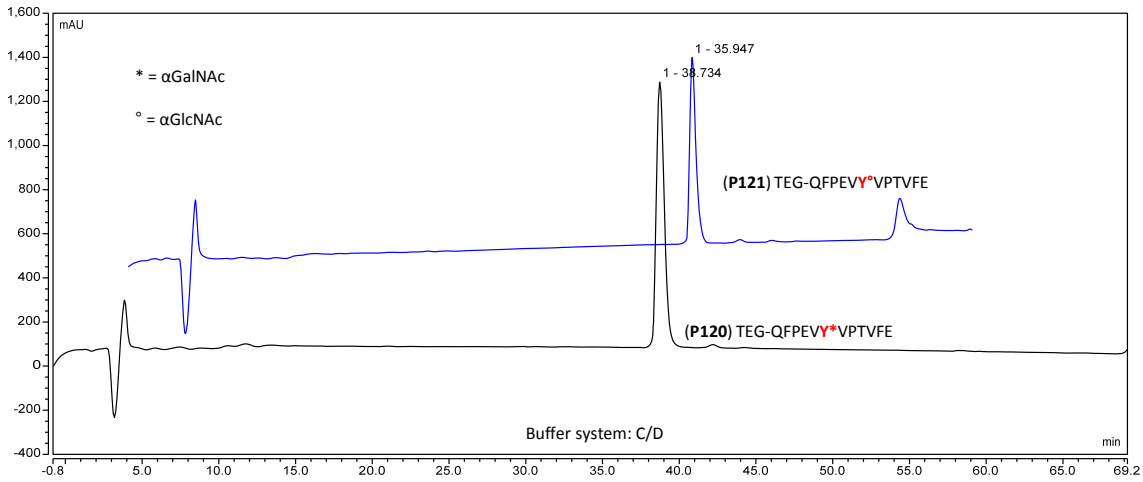


Rbl2 411 – 422


mAU
1 - 25.402
(P118) Ac-KENPAVTPVSTA
Buffer system: A/B
min
1,400
1,200
1,000
800
600
400
200
0
-200
0.0
5.0
10.0
15.0
20.0
25.0
30.0
35.0
40.0
45.0
50.0
55.0
60.0
65.0
70.0
72.0

RhoA 29 – 40


mAU
* = βGlcNAc
# = phospho
1 - 39.500
1 - 41.683
1 - 38.310
(P123) TEG-QFPEVY#VPTVFE
(P122) TEG-QFPEVY*VPTVFE
(P119) TEG-QFPEVYVPTVFE
Buffer system: A/B
min


mAU
* = αGalNAc
° = αGlcNAc
1 - 35.947
1 - 38.734
(P121) TEG-QFPEVY°VPTVFE
(P120) TEG-QFPEVY*VPTVFE
Buffer system: C/D
min

Acknowledgements

First I want my gratitude to Prof. Dr. Ulrika Westerlind for giving me the opportunity to work on these projects, for guidance and for critical reviewing of this manuscript. Also, I want to thank Prof. Dr. Herbert Waldmann for providing a productive working environment and for enabling the work on this thesis. Further, I want to thank Prof. Dr. Albert Sickmann for welcoming me at ISAS and providing an advanced and reliable infrastructure.

Also my thanks go to Dr. Jonas Nilsson for the cooperation in the oxonium ion fragmentation project, Dr. Filipa Marcelo for cooperation in NMR-studies to resolve the impact of glycosylation on peptide conformation, Prof. Dr. Jennifer Kohler for providing the recombinant hOGA and hOGT enzymes and Prof. Dr. Niklas Arnberg for providing the recombinant adenovirus lectins.

A lot of gratitude goes to the technical staff from ISAS: Ingo Feldmann, Dr. Konstantin Shuavev and Rolf Bandur for help and guidance regarding peptide synthesis and HPLC purification, Rita Fobbe for recording many HiRes-ESI spectra. Further my thanks go to the TU Dortmund NMR team for measuring countless products and intermediates.

A special thank you goes to Dr. Christian Pett for many constructive discussions, ideas and practical tips throughout all stages of work on these projects, for providing glycopeptide samples for completing glycopeptide libraries and for technical help with spotting of the microarray slides. I want to thank Oliver Pagel for technical help with the oxonium fractionations and Stefan Loroch for help with HPLC-MS characterization of glycosylated amino acid building blocks and glycopeptides.

I want to thank my fellow lab and office mates Vanessa Caixeta Pereira, Fiorella Solari Zubiaga, Dr. Yu Jin, Sandra Behren and Dr. Cai Hui for lots of discussions, ideas, incentives and a generally nice and warm atmosphere. Also I want to thank all the institute members, which have sweetened the time with events and discussions and thereby made it a nice and memorable one.

I want to thank my family for the support they have given me on the way to and also throughout this work.

Finally, I want to thank Anna for her unconditional support, smile and the countless little things, which to say would go beyond the constraints of this thesis.

Eidesstattliche Versicherung (Affidavit)

_____ _____

Name, Vorname
(Surname, first name)

Matrikel-Nr.
(Enrolment number)

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden, § 63 Abs. 5 Hochschulgesetz NRW.

Die Abgabe einer falschen Versicherung an Eides statt ist strafbar.

Wer vorsätzlich eine falsche Versicherung an Eides statt abgibt, kann mit einer Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft werden, § 156 StGB. Die fahrlässige Abgabe einer falschen Versicherung an Eides statt kann mit einer Freiheitsstrafe bis zu einem Jahr oder Geldstrafe bestraft werden, § 161 StGB.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offence can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offences of this type is the chancellor of the TU Dortmund University. In the case of multiple or other serious attempts at deception, the candidate can also be unenrolled, Section 63, paragraph 5 of the Universities Act of North Rhine-Westphalia.

The submission of a false affidavit is punishable.

Any person who intentionally submits a false affidavit can be punished with a prison sentence of up to three years or a fine, Section 156 of the Criminal Code. The negligent submission of a false affidavit can be punished with a prison sentence of up to one year or a fine, Section 161 of the Criminal Code.

I have taken note of the above official notification.

_____ _____

Ort, Datum
(Place, date)

Unterschrift
(Signature)

Titel der Dissertation:
(Title of the thesis):

_____

_____

_____

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel selbstständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht.
Die Arbeit hat in gegenwärtiger oder in einer anderen Fassung weder der TU Dortmund noch einer anderen Hochschule im Zusammenhang mit einer staatlichen oder akademischen Prüfung vorgelegen.

I hereby swear that I have completed the present dissertation independently and without inadmissible external support. I have not used any sources or tools other than those indicated and have identified literal and analogous quotations.

The thesis in its current version or another version has not been presented to the TU Dortmund University or another university in connection with a state or academic examination.*

*Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the PhD thesis is the official and legally binding version.

_____ _____

Ort, Datum
(Place, date)

Unterschrift
(Signature)