Synthesis and Development of a Mucin Glycopeptide Microarray System for Evaluation of Protein-Interactions

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 Diplom Chemiker

CHRISTIAN PETT

aus Mainz.

Dortmund 2016

Die vorliegende Arbeit entstand im Zeitraum von Oktober 2009 bis September 2015 unter Anleitung von Dr. Ulrika Westerlind an der Fakultät Chemie und Chemische Biologie der Technischen Universität Dortmund und dem Leibniz-Institut für analytische Wissenschaften ISAS -e.V-, Dortmund.

Dekanin: Prof. Dr. Insa Melle

Erster Gutachter: Prof. Dr. Herbert Waldmann

Zweiter Gutachter: Prof. Dr. Christian Hedberg

Teile dieser Arbeit wurden bereits in folgenden Beiträgen veröffentlicht:

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.

A convergent Strategy for the Synthesis of Type-1 Elongated Mucin Cores 1-3 and the Corresponding Glycopeptides

C. Pett, U. Westerlind, *Chem Eur. J.* **2014**, 20, 7287-7299 (inkl. *Back Cover* der Ausgabe).

Weitere Beteiligung an folgenden Beiträgen:

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. Brinkmalm, 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. Westerlind, J. Nilsson, *Chem. Eur. J.* **2015**, 22, 1114-1124.

FÜR MARTHA

TABLE OF CONTENT

ABBREVIATIONS

aa	amino acid
AAA	amino acid analysis
Ac	acetyl
APC	antigen presenting cell
arom	aromatic
AuNP	gold nanoparticle
AW	acid washed
BCR	B-cell receptor
Boc	*tert*-butyloxycarbonyl
BSA	bovine serum albumin
Bu	butyl
Bn	benzyl
Bz	benzoyl
c	concentration
calc.	calculated
cat.	catalytic
CD	cluster of differentiation
^{C}Hex	cyclohexane
CID	collision induced dissociation
CMP	cytidine-5´-monophospho-
CFG	Consortium for Functional Glycomics
Cosmc	core 1 β3-Gal-T-specific molecular chaperone
COSY	correlated spectroscopy
CRD	carbohydrate recognition domain
CTL	cytotoxic T-lymphocyte
d	duplet or day
DBU	1,8-diazabicylo[5.4.0]undec-7-ene
DC	dendritic cell
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
DMTST	dimethyl(methylthio)sulfonium trifluoromethanesulfate
DNA	deoxyribonucleic acid
DPS	diphenyl sulfoxide
EDC	*N*-ethyl-*N'*-(dimethyaminopropyl)-carbodiimide
EGRF	epidermal growth factor receptor
ELISA	enzyme-linked immunosorbent assay
Em	emission
eq	equivalents
ESI	electrospray ionization
Et	ethyl
Ex	excitation
FA	formic acid
FAC	frontal affinity chromatography
Fmoc	*N*-(9H-fluoren-9-yl)-methoxycarbonyl
Fuc	L-fucose
Gal	D-galactose
GalNAc	*N*-acetyl-D-galactosamine
gg	gauche-gauche
Glc	D-glucose
GlcNAc	*N*-acetyl-D-glucosamine
grad	gradient
gt	gauche-trans
GTP	guanosine-5'-triphosphate
h	hour
HATU	*O*-(1H-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethylunronium hexafluorophosphate
HBTU	*O*-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethylunronium hexafluorophosphate
HCD	Higher energy C-trap dissociation
HER2 = ERBB2	human epidermal growth factor receptor 2
HMBC	heteronuclear multiple bond correlation
HMFG	human milk fat globule
HOAt	7-aza-1-hydroxybenzotriazole

HOBt	1-hydroxybenzotriazole
HPLC	high performance liquid chromatography
HR	high resolution
HSQC	heteronuclear single quantum coherence
Hz	hertz
IDCP	iodonium dicollidine perchlorate
Ig	immunoglobulin
IL	interleukin
*i*Pr	isopropyl
ITC	isothermal titration calorimetry
J	coupling constant
KLH	Keyhole Limpet Hemocyanin
Lac	lactose
LacNAc	*N*-acetyllactosamine
Le^x/Le^a	Lewis x/Lewis a
LG	leaving group
m	multiplet
M	molarity or mega
mAb	monoclonal antibody
MALDI	matrix assisted laser desorption ionization
Man	D-mannose
mbar	millibar
MD	molecular dynamics
Me	methyl
MGL	macrophage c-type lectin
MHC	major histocompatibility complex
min	minutes
miRNA	micro ribonucleic acid
MPL	monophosphoryl lipid A
mRNA	messenger ribonucleic acid
MS	molecular sieve or mass spectrometry
NBS	*N*-bromosuccinimide
NCE	normalized collision energy
Neu5Ac	*N*-acetylneuraminic acid
NHS	*N*-hydroxysuccinimide
NIS	*N*-iodosuccinimide
NMP	*N*-methylpyrrolidone
NMR	nuclear magnetic resonance
NSCLC	non-small cell lung cancer
OTf	trifluoromethanesulfonate (triflate)
p	*para*
Pam_3CSK_4	*N*-palmitoyl-s-[2,3-bis(palmitoyloxy)-(2RS)-propyl-[R]-cysteinyl-[S]-seryl-[S]-lysyl-[S]-lysyl-[S]-lysyl-[S]-lysyl-
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
PBS/PBST	phosphate buffered saline/PBS+Tween-20
PCR	polymerase chain reaction
PEG	polyethylene glycol
PG	protecting group
Ph	phenyl
pH	*potentia hydrogenii*
Phth	phthalimido
PMP	*para*-methoxy phenyl
pos	positive
PNGase	Peptide *N*-glycosidase
ppGalNAcT	polypeptide *N*-acetylgalactosamine transferase
ppm	parts per million
PS	polystyrene
PSGL-1	P-selectin glycoprotein ligand-1
p-TsOH	*para*-toluenesulfonic acid
PyBOP	(benzotriazol-1-yloxy)tris(pyrrolidino)phosphonium hexafluorophoshate
q	quartet
quart	quaternary
R	residue
Ras	rat sarcoma
R_f	retention factor
RNA	ribonucleic acid
RP	reverse phase
R_t	retention time
RTK	receptor tyrosine kinase
s	singlet
SEA	sea urchin sperm protein, enterokinase and agrin
sLe^x/sLe^a	sialyl-Lewis x/sialyl-Lewis a
SNP	single nucleotide polymorphism
SPE	solid phase extraction
SPPS	solid phase peptide synthesis

ST	sialyl-Thomsen-Friedenreich antigen
STD	saturation transfer difference
ST_N	sialyl-Thomsen Nouveau antigen
Su	succinimide
t	triplet
T/TF	Thomsen-Friedenreich antigen
TACA	tumor associated carbohydrate antigen
TBAF	*tert*-butylammonium fluoride
TBS	*tert*-butyldimethylsilyl
TBTU	*O*-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethylunronium tetrafluoroborate
TCR	T-cell receptor
TEG	triethylene glycol
tert	tertiary
TES	triethylsilyl
TFA	trifluoroacetic acid
TfOH	Trifluoromethanesulfonic acid
tg	trans-gauche
THF	tetrahydrofuran
TIPS	triisopropylsilane
TLC	thin layer chromatography
TLR	Toll-like receptor
TMS	tetramethylsilyl
T_N	Thomsen-Nouveau antigen
TOCSY	total correlation spectroscopy
TOF	time of flight
Troc	trichloroethoxycarbonyl
Trt	trityl
TT	tetanus toxoid
UV	ultraviolet
VEGF	vascular endothelial growth factor
VNTR	variable number of tandem repeats
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 INTRODUCTION

1.1 The mucin glycoprotein family

Mucins belong to a family of highly glycosylated proteins expressed by epithelial cells.[1] Up to now, 21 members of this family have been identified. The carbohydrate-content of the mucins is usually higher than 50 wt% of the protein. These glycoproteins are a major part of the innate immune system. Secreted mucins form the mucus layer that protects the epithelial tissues of the gastrointestinal and respiratory tract and the ductal surfaces of breast, pancreas and kidney tissue against invading pathogens and chemical or physical stress factors.[2] While most mucins are restricted to specific tissues, the MUC1 glycoprotein is ubiquitously distributed.[3] MUC1 was first identified by murine monoclonal antibodies raised against human milk fat globule (HMFG).[4] Mucins can be subdivided into two major classes, the secreted or gel-forming mucins and the membrane-bound mucins. The ubiquitous MUC1 belongs to the class of the membrane tethered mucins. It is embedded into the membrane lipid bilayer by a single trans-membrane domain. The N-terminal ectodomain has a rod-like structure and stretches out 200-500 nm into the extracellular space, significantly surpassing the thickness of the glycocalyx (ca. 10 nm).[5] The short transmembrane domain connects to the 72 amino acids cytoplasmatic domain, which is involved in cell signaling processes.[6,7]

The MUC5B glycoprotein is belongs to the class of the secreted mucins and is mainly expressed in the respiratory system. Together with MUC5AC, it is the major airway mucin. Most secreted mucins have cysteine-rich regions, enabling formation of disulfide bridges and thus protein networks.[8] Almost all mucins have an extracellular domain rich in proline, serine and threonine and arranged in domains called VNTR-regions (variable number of tandem repeats). In MUC1, a VNTR of 20 amino acids with the sequence PAHGVTSAPDTRPSAPGSTAP is repeated 20-125 times, in MUC5B a VNTR region consisting of 29 amino acids with the sequence ATGSTATPSSTPGTTHTPPVLTTTRTTPT is repeated in total 11 times.[9,10] The overall number of VNTRs in MUC1 depends on genetic polymorphism and is different in each individual. A MUC1 tandem repeat contains five glycosylation sites (3x threonine, 2x serine) and the high proline content disturbs the formation of secondary structures, giving MUC1 a stretched, rod-like appearance.[11,12] MUC1 is expressed as a single precursor protein with a proteolytic cleavage site within a SEA (sea urchin sperm protein, enterokinase and agrin) domain between the ecto- and transmembrane

domain. This site is auto-catalytically cleaved during posttranslational processing and the extracellular domain is then non-covalently associated with the transmembrane domain.[13] All *O*-glycosylated proteins are posttranslational modified during the passage through the golgi complex. Here, the carbohydrate decoration is added by the sequential addition of monosaccharides which requires an ensemble of various different glycosyltransferases. After the exocytotic transport to the cell membrane, MUC1 is internalized and transported back to the trans-golgi with subsequent re-localization to the cell membrane. During this cyclic process, the glycosylation pattern is further matured.[14]

Mucin-like *O*-glycosylation is commonly found on threonine and serine amino acid residues. Recently, tyrosine was identified as a mucin-type glycosylation site, however, only a few examples of glycosylated proteins with this new posttranslational modification were found until now.[15,16] Mucin-type glycosylation is characterized by eight core structures, namely core 1 to core 8.[17] All cores structures have in common, that the first added carbohydrate is *N*-acetylgalactosamine in α-configuration, which is attached by a family of polypeptide *N*-acetylgalactosaminyltransferases (ppGalNAcT). This structure alone is known as the T_N-antigen (Thomsen-Nouveau antigen). Further elongation by the action of a β1,3-galactose transferase (C1GalT, T-synthase), results in the basic core 1 motif, also named T-antigen (or TF for Thomsen-Friedenreich antigen). Instead of a galactose, a glucosamine can be attached by the core 3 *N*-acetylglucosaminyltransferase (C3GnT) resulting in the core 3 glycan. Core 3 structures are specifically found on mucins of the gastrointestinal tract.[18] Elongation on the 6-position of the first *N*-acetylgalactosamine of core 1 or core 3 by different members of the core 2 *N*-acetylglucosaminyltransferase family (C2GnT) results in the core 2 and core 4 glycans, respectively. Core 1 to core 4 are the most abundant core motifs, while core 5 to 8 are very rare and only found in a few mucins. The genes encoding for the glycosyltransferases involved in core 7 and core 8 biosynthesis have not been identified yet (*figure 1.1*).

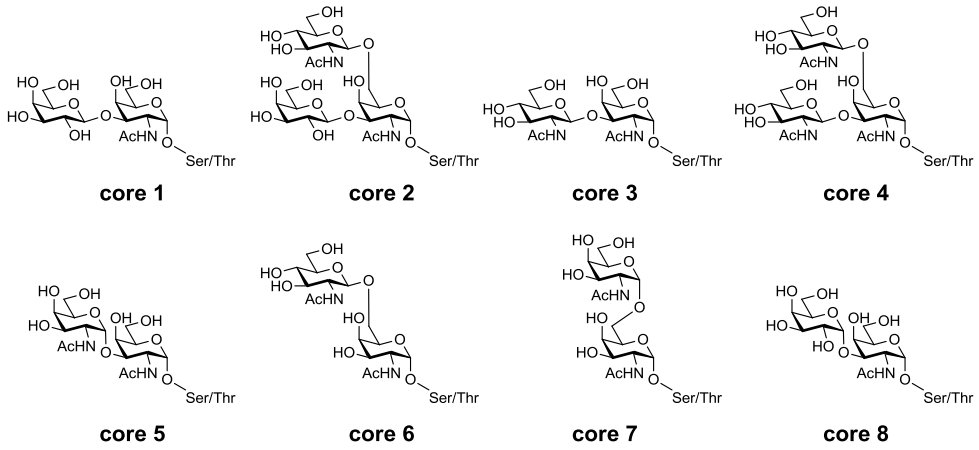


***Figure 1.1*:** Structures of the common *O*-glycan core motifs.

The *O*-glycan core structures can be elongated by the alternating action of a β1,3-GlcNAc-transferase (β1,3GlcNAcT) and by either a β1,3- or β1,4-Gal-transferase (β1,3GalT, β1,4GalT) (*figure 1.2*).

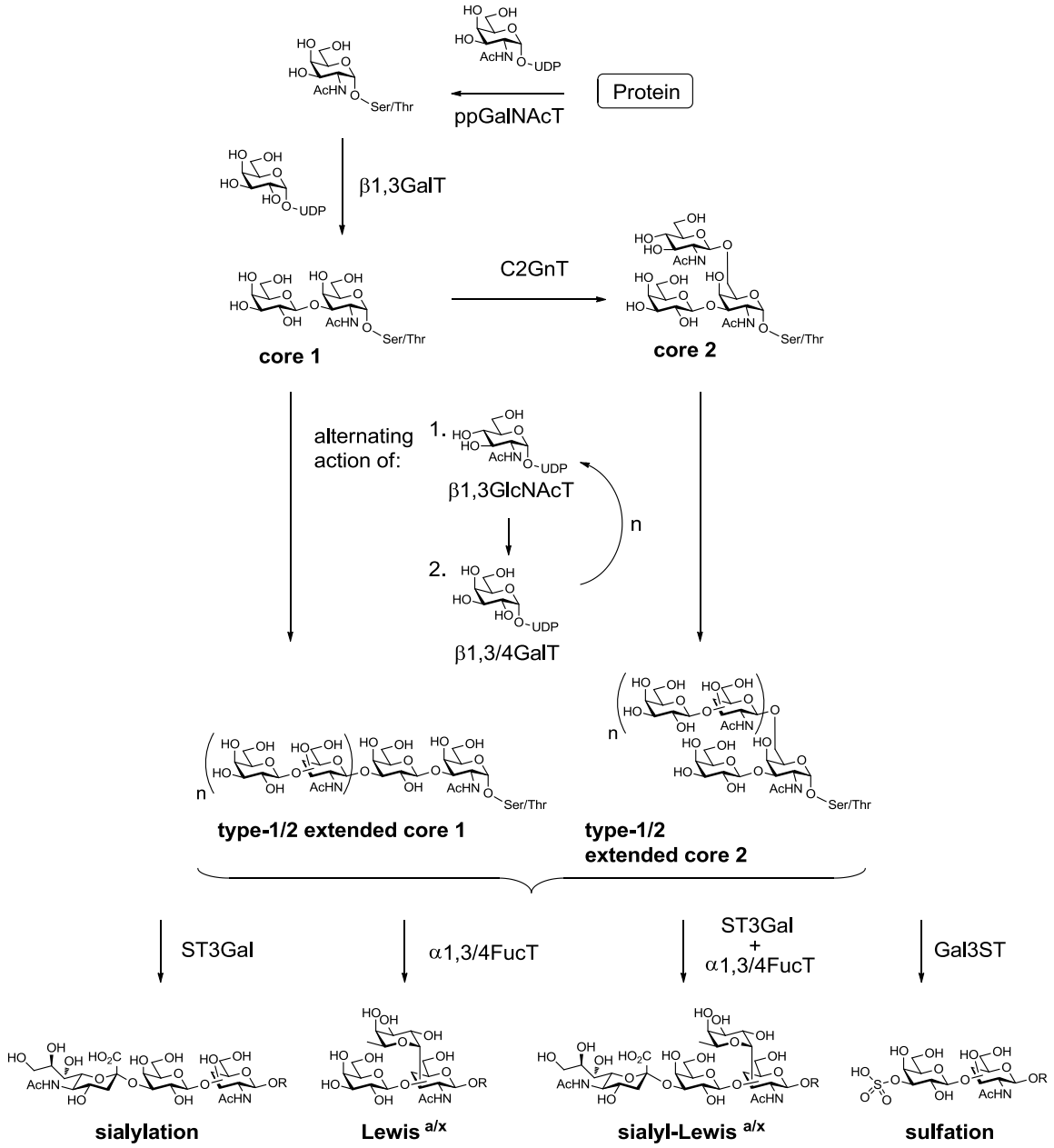


Figure 1.2: Biosynthesis of mucin type *O*-glycosylation.

Like this, the core structures are elongated either by a Galβ1,3-GlcNAcβ (type-1 *N*-acetyllactosamine, LacNAc) or by a Galβ1,4-GlcNAcβ (type-2 LacNAc) disaccharide unit (*figure 1.2*).[17] Chain termination may occur by the performance of a family of six α2,3-sialyltransferases (ST3Gal-I to -VI), which attach a sialic acid (e.g. *N*-acetylneuraminic acid, Neu5Ac) to the terminal galactose of the carbohydrate chains. Sialylation may also occur by a α2,6-linkage on a terminal Gal or the initial αGalNAc of the T_N- or T-antigen. Terminal α2,6-sialylation on galactose is mainly performed on *N*-glycans by a family of α2,6-galactose sialyltransferases (ST6Gal-I and -II), while *O*-glycans are mainly sialylated on the initial αGalNAc by a family of α2,6-*N*-acetylgalactosamine sialyltransferases (ST6GalNAc-I to -VI).[17,19] Many glycosyltransferases do not recognize sialylated glycans, and therefore further

chain growth is prevented. An exception is the α1,3/4-fucosyltransferases (FUT3) which adds a fucose monosaccharide, either to the 3- or the 4-position of the GlcNAc of the corresponding type-1 and type-2 LacNAc disaccharide units, even when sialylated. This results in the formation of the Lewis^x (Le^x) and the Lewis^a (Le^a) antigens or the corresponding sialyl-Lewis^x (sLe^x) and sialyl-Lewis^a (sLe^a) glycans. Sialylated Le antigens are important motifs, involved in leukocyte recruitment and emigration during inflammations.[20] The carbohydrates may also be terminated by sulfation (*figure 1.2*). The number of carbohydrates per *O*-glycan chain on mucins on healthy cells, usually varies between 2 and 20 monosaccharides.[21]

1.2 Abnormal glycosylation and the role of MUC1 in cancer

In epithelial carcinomas, expression of MUC1 undergoes a number of fundamental changes. The expression of the MUC1 glycoprotein is dramatically increased and the epithelial cell polarity is lost.[22] Usually, MUC1 and other membrane-bound mucins are expressed only on the apical site of the cells, which is exposed to the ductal environment. However, in carcinoma cells, MUC1 is found also on the lateral and basal sites of the cells. In cases of cellular stress, e.g. in acute or chronic inflammatory insults, the cell polarity may further be lost. Cells can perform a repair program to regain polarity as soon as the stress response is terminated.[23] In cancer however, the stress response is irreversible. During these processes, the cytoplasmatic domain of the dislocated MUC1 can now interact with receptor tyrosine kinases (RTKs) that are located on the basal cell membranes. For instance, MUC1 interacts with the RTK ERBB2 (HER2) and its activation is associated with the loss of *tight junctions*[24] between the cells.[25] Also, MUC1 binds β-catenin, which is usually bridging between actin filaments and the cell-cell-adherence molecule E-cadherin. As a result, *adherence junctions* between the cells are lost.[26] Overexpression of the rod-shaped MUC1 extracellular domain also masks the cell surface, resulting in destabilization of integrin-mediated adhesion of the cell with the extracellular matrix.[27,28] These factors promote a separation of tumor cells from normal cell clusters and trigger metastasis.

Additionally to overexpression of the MUC1 glycoprotein, the glycan structures attached to the mucin polypeptide tandem repeats are altered due to changes in the expression levels of various glycosyltransferases. These aberrantly structured glycans are referred to as TACAs (tumor associated carbohydrate antigens). In mammary cancer, the levels of C2GnT1 are downregulated or even completely obliterated.[29] The branching to core 2 glycan motifs is therefore inhibited and core 1 structure formation is favored instead. Concomitantly, an

overexpression of sialyltransferases results in premature sialylation of the truncated carbohydrates.[30,31,32,33] After sialylation further glycosylation is prevented. As a result, tumor cell glycosylation consists of short and often sialylated glycans. T_N- and T-antigens and the corresponding sialylated ST_N- and ST-antigens are the most prominent glycans on cancer-associated MUC1 (*figure 1.3*).

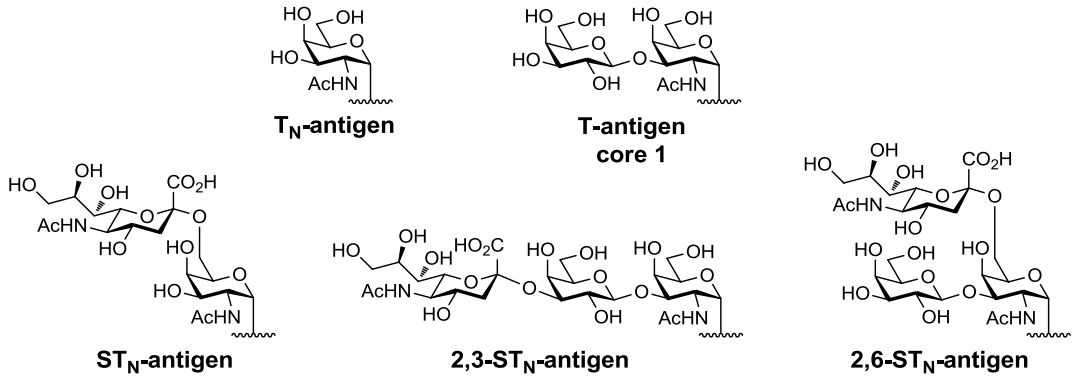


***Figure 1.3*:** Structures of tumor-associated carbohydrate antigens (TACAs) on mucins.

TACAs were found on several human and murine breast cancer cell lines, on tissue samples of breast cancer patients and as well as on shed or secreted MUC1 in the serum of advanced breast cancer patients.[30,34] Interestingly, expression of sLe^x epitopes, which are a ligands for E-selectins, also occurs in breast cancer.[35] For normal *O*-glycosylation in the gastric, colorectal, pancreatic and hepatic tissue, the expression of core 3 β1,3-*N*-acetylglucosaminyltransferase (C3GnT) is pivotal. Disturbed expression of this enzyme consequently results in higher levels of T_N-antigen instead, which is correlated to cancer growth and susceptibility for metastasis.[36,37,38]

The reasons for the aberrant levels of glycosyltransferase expression are not fully understood. The increased levels of T_N- and ST_N-antigen are related to a genetic mutation in the gene *Cosmc* (core 1 β3-Gal-T-specific molecular chaperone).[39,40,41] This gene encodes for the chaperone Cosmc, required for the correct folding of the C1GnT (β1,3GalT, T-synthase). A lack of this chaperone leads to C1GnT aggregation and subsequent degradation in the proteasome. Consequently, appearance of T_N-antigens increases because of faulty core 1 glycan elongation.[40] Increased levels of ST3Gal-I in cancer cells could be related to hypoxia, which also influences expression levels of various other glycosyltransferases.[42] In colon cancers, increased T-antigen and sLe^x and sLe^a glycosylation is not only related to glycosyltransferase expression rates. Rather higher transcription of the UDP-Gal transporter gene leads to increased amounts of galactose-donor in Golgi network.[43]

As a consequence, aberrant glycosylation results in a higher content of short and truncated carbohydrates. The tumor associated T_N- and T-antigen have been identified in about 90% of all carcinomas.[44,45] Because of the smaller glycans, the protein backbone is not effectively covered any more. Peptide epitopes are exposed, which in normal mucins would be masked by extensive glycosylation. These distinct differences of healthy and malignant cells, permits the humoral immune reactions to unique protein epitopes in cancer.

The PDTR peptide domain within the MUC1 tandem repeat sequence has been identified as an immune-relevant peptide epitope, which is recognized by many MUC1 antibodies.[46,47,48,49] A secondary structure in this peptide sequence was found and identified as either a type I or II β-turn.[50,51,52] Later NMR studies confirmed a type I turn,[53,54,55] but also consecutive γ-turns leading to s-shaped structures were proposed.[56,57] The reported variations might be referred to different experimental conditions and the interplay between the different analyzed peptide backbones and attached glycoforms, resulting in conformational variations. However, the secondary turn structure in the PDTR motif is usually regarded as the reason for its pronounced immune reactivity as it forms a “knob-like” structure that projects out of the straight protein. The GSTA domain of the tandem repeat was later identified as a novel cancer specific epitope, by the generation of MUC1 mAbs.[58] In further experiments, a MUC1 glycopeptide microarray system was presented for the screening of serum antibodies. Former breast cancer patients who were classified as disease-free were treated with a MUC1 triple tandem repeat sequence, perglycosylated with T_N-antigen (25 x T_N) and conjugated to KLH (keyhole limpet hemocyanin) as carrier protein. Antisera were extracted before and after administration. The sera of 18 out of 20 patients showed recognition of glycopeptides with double T_N-antigen glycosylation in the GSTA domain after the injection.[59,60] Further tests with sera from large cohorts of breast-, ovary- and prostate cancer patients, revealed auto-antibodies directed against the GSTA epitope.[61,62] Of the reported monoclonal antibodies against MUC1 tandem repeats, only a few were directed against the VTSA peptide domain, however, this region seems to be less immunogenic.[63,64] According to the site-specificities of the relevant ppGalNAcTs, involved in MUC1 glycosylation, the VTSA domain is likely to be glycosylated on the endogenous protein. However, in between the VT/SA motif is a cleavage site (/) for the protease cathepsin L, which is involved in antigen processing in the endosomes in mammals. Glycosylation in VTSA, whether on threonine or serine, obstructs the action of the protease. The resulting peptide fragments may not be in the optimal size range for MHC II (major histocompatibility complex II) protein presentation to the T-cells. This might be one reason for the low immunogenicity of this motif.[65,66] In many cases not only the peptide sequence is part of the epitope. A carbohydrate hapten may be included in the antigenic determinant. Several MUC1 antibodies that have been raised, show

better reactivity, when the peptide/protein contains glycosylation and many are exclusively reactive on glycosylated MUC1.[46]

1.3 Galectins and the role of galectin-3 in cancer

Galectins are a family of lectins with conserved C-terminal carbohydrate recognition domains (CDRs) for recognition of β-galactosides. The 15 found members of the family of mammal galectins are divided into three main classes: 1) The prototype galectins (galectin-1, -2, -5, -7, -10, -11, -13, -14, -15) consisting of two dimerized CDR chains, 2) the tandem repeat type galectins (galectin-4, -6, -8, -9, -12) consisting of a single chain with a CRD at each end and 3) chimera type galectin-3 with one CRD, connected to a tandem repeat of a collagen-like sequence, which enables oligomerization to pentamers. Although all members of the galectin family have a general affinity for β-galactosides, each galectin has a fine specificity regarding the type of galactosyl-oligosaccharide.[67] In contrast to other common mammalian lectins, such as C-type lectins, selectins and siglecs, galectins are not membrane-bound. They appear on the intracellular side in the cytosol, in the nucleus and are also secreted into the extracellular space. Accordingly, galectins are involved in a plethora of different physiological cell functions, such as involvement in cell-cell junctions, adhesion to extracellular matrix, signaling pathways, cell cycle progression, apoptosis, regulation of RNA splicing, modulation of signal transduction and cell migration.[68]

Galectin-3 is one of the most studied galectins and shows increased expression in various carcinomas. Further, galectin-3 plays several roles in tumorgenesis and metastasis.[69] For example, cytosolic galectin-3 binds selectively to the oncogene K-Ras but not to the isoforms H- and N-Ras.[70] The galectin-3 interaction with K-Ras increases the amount of activated GTP-K-Ras, leading to enhanced activation of downstream effectors, preventing apoptosis or promoting cell proliferation.[71] Further, upon an exogenous apoptotic stimuli, cytosolic galectin-3 is translocated to the mitochondrial membrane where it prevents release of cytochrome C, thereby preventing apoptosis.[72] The mechanism might be similar to the anti-apoptotic properties of Bcl-2 (B-cell lymphoma 2), which shows significant sequence homology to galectin-3. Tumor promotion may further be executed by galectin-3 involvement in the Wnt/β-catenin pathway. Galectin-3 prevents phosphorylation of β-catenin and thereby degradation in the proteasome. As a result β-catenin translocates into the nucleus, accumulates and activates gene expression of the cell cycle progressor cyclin D or the proto-oncogene c-Myc.[73,74] These examples show the various tumor-promoting effects of intracellular galectin-3. Galectin-3 was co-immunoprecipitated with several proteins involved

in the mentioned pathways, such as K-Ras[70] or β-catenin[75], indicating a recognition mode independent of glycosylation. Many more examples for intracellular, carbohydrate-independent interactions are reported.[76] In general, intracellular galectin-3 mediates cancer promoting effects by preventing apoptosis and providing tumor cell survival.

Extracellular galectin-3 is able to bind galactosides on glycoproteins presented on the cell surface. For instance, mucin glycoproteins on different epithelial carcinoma cells are primary binding partners for galectin-3.[77,78] Furthermore, co-immunoprecipitation from colon cancer cell lines showed that cancer associated MUC1 is a natural target for galectin-3 and binding seems to be mediated by glycosylation[77] with the tumor-associated T-antigen.[79] MUC1 proteins with T-antigen presentation on cancer cells are proposed to permit adhesion to endothelial cells in the presence of galectin-3.[80,81,82] Binding was increased when the cells were pretreated with sialidase and depletion of binding occurred by *O*-glycanase treatment (T-antigen specific endoglycosidase from *Streptococcus pneumoniae*). MUC1-transfected melanoma cells with T-antigen glycosylation exhibited stronger endothelial cell binding when treated with galectin-3, while no effect was visible on the MUC1 negative revertants.[83] It was therefore proposed that galectin-3 captures overexpressed MUC1 on the surface of circulating tumor cells and through clustering of MUC1, several shorter adhesion molecules, such as E-cadherin, CD44 and E-selectin ligands are revealed which are then exposed on the cell surface and are able to interact with their corresponding binding partners on the endothelial side (*figure 1.4*, mechanism **A**).[79] By this mechanism, galectin-3 supports settling of metastasizing tumor cells. Also, adhesion of the circulating tumor cells to each other is increased through mucin/galectin-3 interaction, leading to aggregation of tumor cells (*figure 1.4*, mechanism **B**).[84] The cell adhesion prevents apoptosis of the tumor cells, induced by the loss of cell-cell or cell-matrix interactions (anoikis).[85,86] By recognizing tumor related glycosylation on overexpressed MUC1 and following mediated adhesion to epithelial tissue, galectin-3 promotes installation of secondary tumors. Furthermore, extracellular galectin-3 has been shown to be immunosuppressive. T- and B-leukocytes, which were exposed to galectin-3, became apoptotic (*figure 1.4*, mechanism **C**).[87,88,89,90,91] As molecular receptors on the leukocytes CD7, CD29 (β1-integrin), CD45 (protein tyrosine phosphatase) and RAGE were identified.[87,88,92,93,94] All interactions are mediated by protein *N*- and *O*-glycosylation with varying contributions. For example, CD45 contains *N*- and *O*-glycans, but the splice isoform CD45RO, which lacks the heavily *O*-glycosylated protein domain, shows almost no galectin-3 binding, highlighting the significance of *O*-glycosylation for galectin-3 induced apoptosis.[94] Overexpressed galectin-3 in circulation and membrane-associated to the tumor cells is therefore suspected to be beneficial for the tumor cells and helps to escape immune surveillance. Apart from apoptosis of immune cells, the human colon adenocarcinoma cell line SW48 was also shown to enter apoptosis after removal of *N*-acetylneuraminic acid by

sialidase treatment and contact with galectin-3.[93] CD7 and CD29 were identified as galectin-3 receptors. CD29 was found to be α2,6-oversialylated on this carcinoma cell line, thus hindering galectin-3 interaction. It was therefore proposed, that overexpression of galectin-3 and oversialylation is advantageous for cancer progression in several ways. The cancer cells benefit from the intracellular anti-apoptotic effects and are at the same time protected from apoptotic effects of extracellular galectin-3 by extensive α2,6-sialylation of the galectin-3 receptor CD29 (*figure 1.4*, mechanism **D**).

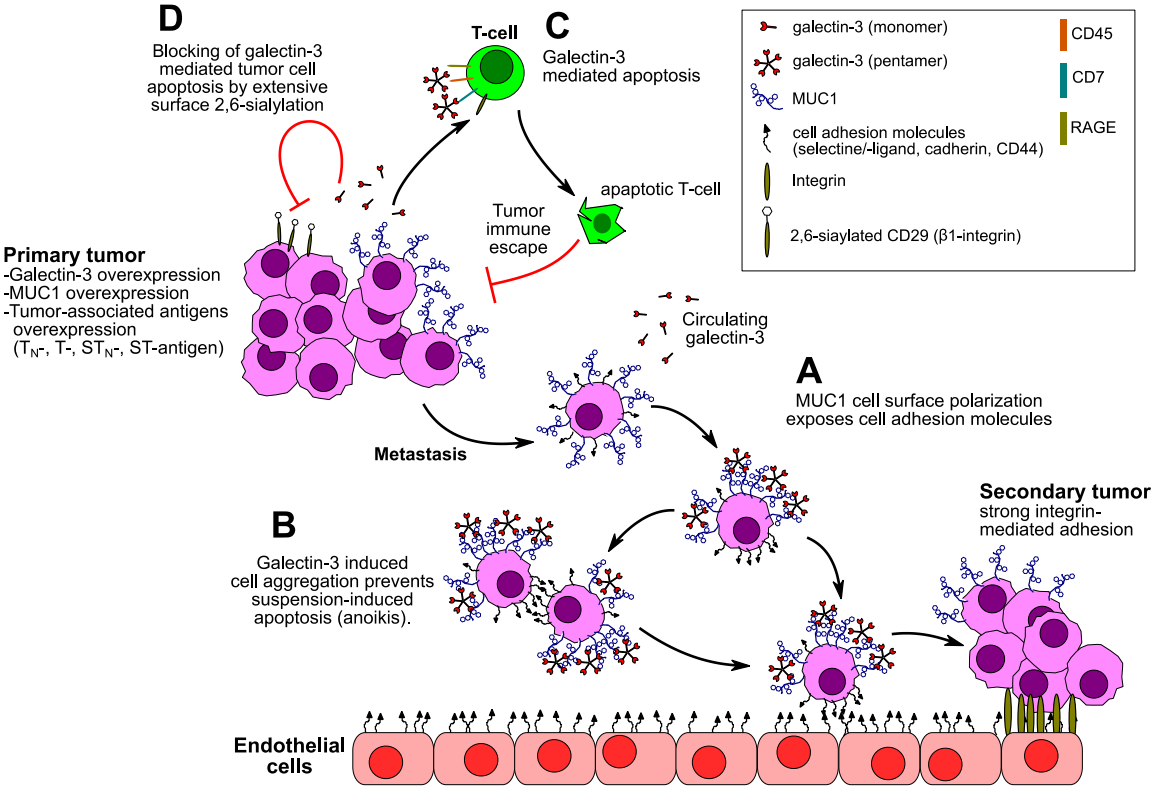


***Figure 1.4*:** Proposed effects of extracellular galectin-3 on tumor metastasis and immune evasion. **A:** Clustering of MUC1 by galectin-3 exposes cell adhesion molecules. **B:** Tumor **c**ell aggregation through exposed cell adhesion molecules prevents anoikis. **C**: Galectin-3 mediated apoptosis of T-cells results in tumor immune escape. **D:** Extensive α2,6-sialylation blocks galectin-3 mediated apoptosis of tumor cells.

Other members of the galectin family were found to share the same metastasis promoting features.[95] MUC1 and galectin-3 act in a tumor-promoting interplay. The *N*-glycosylated (Asn-36) intracellular, C-terminal part of MUC1 further signals an increase of galectin-3 mRNA translation by stabilizing the transcript through decrease of the inherent, galectin-3 suppressing miRNA-322. Tumor-related MUC1 overexpression is therefore correlated with increase in galectin-3 expression. In return, galectin-3 binds the MUC1 *N*-glycosylation at Asn-36 and bridges to ERBB1, integrating MUC1 into EGRF signaling.[96]

Galectin-3 may be involved in tumor progression in several ways. In general, intracellular galectin-3 acts anti-apoptotic and therefore promotes cancer development *via* carbohydrate-independent protein-protein interactions. Extracellular galectin-3 promotes metastasis and regulates cell apoptosis *via* carbohydrate-dependent mechanisms. Through direct contact with cell surface glycans, for example the tumor-associated T-antigen, it is probably involved in installment of daughter tumors, which makes it an interesting subject for further carbohydrate-lectin binding studies.

1.4 MUC1 carbohydrate vaccines

The fact that MUC1 is overexpressed and aberrantly glycosylated made it a prime target in immunotherapy and immunodiagnostics of carcinomas, including breast, ovarian, lung, colon and pancreatic cancers. Anti-MUC1 auto-antibodies are found in serum of cancer patients and their identification is in general associated with a better prognosis.[11,61,97] Treatment of cancer relies on the removal of the tumor burden by chemotherapy, radiation or surgery. Hence, immunotherapeutic methods are a valuable asset to established tumor therapy. Passive immunization with monoclonal antibodies is a successful and relatively safe approach in immunotherapy. Several cancer target reactive monoclonal antibody drugs have already entered the market and second generation approaches, such as engineered antibody fragments or antibody-drug-conjugates (ADCs), are in the focus of contemporary pharmaceutical research and clinical studies, with first products already launched.[98] However, the protection by passive immunization with monoclonal antibodies is not long-lasting, most available therapeutic antibodies on the market and in late stage clinical trials are chimeric or humanized antibody drugs,[99] which can still provoke an unwanted immune response[100] and full human sequence antibody production by *in vitro* affinity maturation through biological display systems or *in vivo* expression in transgenic animals is complex. Active immunization with an anti-tumor vaccine on the other hand, could induce inherent antibody production against tumor related tissue. Especially against metastasizing tumors, an augmented immune surveillance would be helpful. Although auto-antibodies are found in the serum of patients, the induced immune reaction by the inherent tumor-associated antigens is obviously not strong enough to eradicate the tumor burden. The reason is, that the tumor-associated antigens are often seen as self-antigens and are only poorly immunogenic, although being abnormal.[101]

An organism has a natural tolerance for antigens regarded as “own”. All lymphocytes are generated in the bone morrow. To avoid autoimmunity, lymphocytes are selected for

tolerance against endogenous tissues. Cells that recognize self-antigens are selected for cell lysis. For B-cells, this education happens directly in the bone morrow and for the T-cells in the thymus. Dendritic cells, macrophages and B-cells are antigen-presenting cells (APCs) and can detect and internalize exogenic pathogens. After digestions, parts of the pathogens are linked to major histocompatibility complex II (MHC II) proteins, which are presented at the surface of the APCs to $CD4^+$ T-cells.[102] If a presented antigen is of endogenous nature, the T-cell receptor (TCR) will only bind weakly to the MHC-antigen-complex, according to maturation in the thymus. Important costimulatory signals cannot emerge and the T-cell remains inactive. If the APC is a B-cell, it can proliferate into a plasma cell and start to secret antibodies in a T-cell independent reaction. These antibodies are of the IgM-type and are short-lived and of low affinity (*figure 1.5*, **A**). In general, sole carbohydrates are T-cell-independent antigens.[103] Tumor related MUC1 IgM antibodies are found in cancer patient sera but immune response is too weak to successfully generate immunity. On the other hand, if an APC presents an unknown antigen to the T-cell, then the MHC-II-antigen-complex is bound with high affinity. Costimulation takes place and the T-cell proliferates into a T-helper (T_h) cell and secretes cytokines which dock to cytokine receptors at the B-cell to induce proliferation into an antibody-producing plasma cell (*figure 1.5*, **B**). The antibodies generated by the T-cell-dependent pathway are long-lived, high affinity IgG antibodies. Furthermore, B-cells may also proliferate into memory B-cells. This class change from IgM to IgG antibodies characterizes an effective humoral immune response.

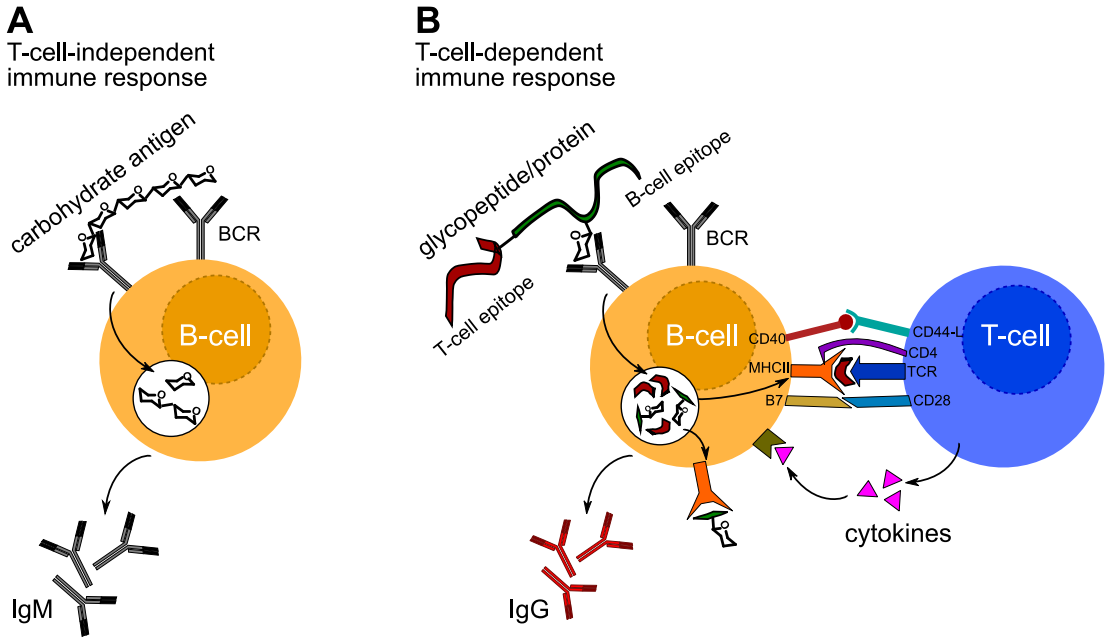


***Figure 1.5*: A:** T-cell independent response by carbohydrate antigens elicits IgM antibody production. **B:** Costimulatory signals enables proliferation of B-cells into plasma cells and IgG antibody production.

At any time, the immune effector cells patrol for unusual structures and start an immune reaction if necessary. However, tumor cells also develop mechanisms to escape the permanent immune surveillance. For instance, the ability to express MHC I molecules may be gradually lost during increasing stages of tumorgenesis.[104,105] Some solid tumors also start to overexpress cytokines with immunosuppressive functions, such as IL-10, IL-4 or vascular endothelial growth factor (VEGF).[106] It has been found, that MUC1 itself is an immunosuppressive factor. Shed and secreted MUC1 in tumor tissue arrests T-lymphocytes in their cell cycle and proliferation of naïve T-cells is blocked.[107,108]

A suitable vaccine must therefore be able to overcome natural tolerance and effectively augment the immune response. Commonly, a chemically synthesized antigenic B-cell epitope is conjugated to an immunostimulant. A common strategy is the conjugation to a carrier protein. Carbohydrate antigens may be directly linked to a carrier, such as keyhole limpet hemocyanin (KLH), bovine serum albumin (BSA) or tetanus toxoid. Immunological studies in mice revealed, that the induced immune reactions by such a conjugate elicits modest responses, often significantly of the IgM-type, indicating insufficient antibody class change.[109,110,111,112,113] Variants of using unglycosylated peptide epitopes, in form of multi tandem repeat peptides or as carrier protein conjugates have been tested. Most of the formulations used with unglycosylated peptides suffer from low humoral response or lack of class switching to IgG isoforms.[114,115,116,117,118] The reasons might be related to missing conformational effects of the carbohydrates on the peptide backbone. *Tecemotide* (*L-BLP25*, *Stimuvax®, Merck KGaA*) is a liposomal formulation of an unglycosylated MUC1 tandem repeat peptide (25mer), conjugated to monophosphoryl lipid A (MPL). The compound went into clinical trials with application in NSCLC (non-small cell lung cancer). After not meeting the primary end point in an initial phase III clinical trial,[119] the formulation was withdrawn in 2014, shortly after the beginning of a second phase III trial and is discontinued.[120]

Vaccine designs based on glycopeptides seem to be more promising.[103] In comparison, glycosylated peptide vaccine candidates usually elicit stronger immune responses than the unglycosylated counterpart.[121] An early study already showed that serum antibodies of breast cancer patients react stronger with glycosylated MUC1 tandem repeats than with the unglycosylated tandem repeats.[47] Commonly, MUC1 glycopeptides are also conjugated to immunostimulants to enhance immune response, to overcome tolerance and ensure antibody class switching. Several glycopeptide vaccine conjugate designs have been reported in the literature. A MUC1 glycopeptide may be attached to a carrier protein, for instance bovine serum albumin (BSA),[122,123] keyhole limpet hemocyanin (KLH)[124,125] or tetanus toxoid (TT) (*figure 1.6*, **A**).[126,127,128] Carrier proteins fortify immune responses by presenting T-cell-epitopes for T-cell-dependent immune reactions and multivalent

presentation of the B-cell epitopes enhances the tendency for receptor clustering on the antigen presenting cells (APC), such as dendritic cells (DCs), B-cells or macrophages. Exogenous proteins also provoke a response against carrier epitopes, which sustains the risk of overriding the immune reaction against the selected B-cell-epitope.[129,130,115] Alternatively, short peptide sequences may be used as T-cell epitopes (*figure 1.6*, **B**). Examples are peptides coming from ovalbumin,[131] the polio virus[132] or rationally designed peptides like the PADRE-sequence.[133] The P30-sequence, a 30mer peptide derived from tetanus toxoid, has recently been reported to give robust antibody titers. Also, this peptide may function as a built-in adjuvant, making the use of additional adjuvants in immunotherapy obsolete.[134,135] Other common immunostimulants are Toll-like-receptor (TLR) ligands, like the lipopeptide Pam_3CSK_4.[136] As a mitogen, the lipopeptide is not immunogenic but stimulates TLR-1/2,[137] which initiates induction of costimulatory proteins in B- and T-cells. Conjugation of these components, with spatial separation via a non-immunogenic spacer, results in two- and three component vaccines.

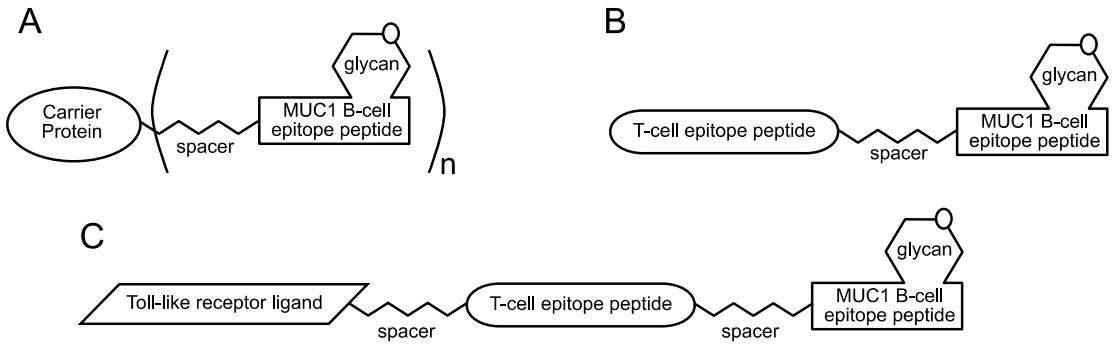


***Figure 1.6*:** Design examples for potential glycopeptide vaccine conjugates: **A:** Two-component peptide-protein vaccine. **B:** Two-component peptide-peptide vaccine. **C:** Three-component peptide vaccine with TLR-lipopeptide.

In addition to a humoral antibody response, it would be beneficial if the vaccine could induce cytotoxic T-cell responses. In that way, malignant tumor cells would also be opposed to cytotoxic T-lymphocytes (CTLs), which could directly lyse the cells or induce apoptosis. Exogenous substrates, such as the potential vaccine, are endocytotically internalized, processed by digestion and the resulting (glyco-)peptides loaded onto MHC II molecules. The loaded antigens are then exclusively presented to $CD4^+$-T-cells and differentiation into T-helper cells proceeds. CTL activation, on the other hand, relies on antigen presentation by MHC I to $CD8^+$ T-cells. Nevertheless, CTLs with specificity against the MHC-unrestricted epitope localized to the DTR domain have been isolated from tumor patients.[138,139] Since then, MHC I restricted peptide sequences within or outside the MUC1 tandem repeat were also reported and specific CTL-responses could be induced in mouse models.[140,141,142,143,144] Moreover, it was documented that glycosylation is able to stabilize the binding to MHC I.[142,144] It is not clear though, how the tandem repeat sequences get access to the MHC I pathway.

This would involve mechanisms of cross-presentation. The antigen could for example escape into the cytosol and be processed by the proteasome,[145,146] or be digested by cathepsins to yield 8-11mer peptides which would fit into the MHC I binding groove.[147] It was also shown, that the C-type lectin MGL (macrophage galactose c-type lectin) can mediate internalization of glycopeptides into APCs upon which the glycopeptides are directed into MHC II as well as into MHC I compartments.[148,149] The uncertainties regarding efficient MHC I presentation makes addressing cytotoxic immune responses by vaccines difficult and so far no example of a CTL-response, being able to successfully neutralize human cancer cells is reported (for example in transgenic mouse models).[143]

1.5 Mucins in airway diseases

The glycans on mucin glycoproteins are covering the epithelial cell-surface and determine physical and biological properties of the mucus. The glycans are effecting the conformation of the protein and the hydrophilic glycans are preventing dehydration, thus influencing the mucus flow properties. The different carbohydrate structures also function as ligands to different pathogens, which during normal flow properties to a large extend are cleared by the mucus.[150] This process is an important part of the innate immune system. In airway diseases, the binding of pathogens to glycan ligands results in inflammatory-infective responses in the respiratory tract contributing to the mucin overproduction and causing an aberrant flow of the mucus, which further is an important factor in the morbidity and mortality of COPD, CF and asthma patients. Changes in terminal mucin glycosylation has been found in these patient groups, e.g. altered levels of sialylation, sulfation and fucosylation.[151] These alterations of the *O*-glycosylation influence both biophysical and biological (binding to bacteria, viruses and cells) properties of specific mucins. The changes in glycosylation and mucin production in COPD, CF and Asthma are probably one of the major factors involved in the non-optimal transport properties of the mucus and rather than protecting the host, the gel instead provides an environment where the pathogens thrive and contribute even more to inflammation and disease.[152]

2 Motivation

Oligosaccharides attached to proteins and lipids form the glycocalyx, which covers the surface of most animal and some bacterial cells. These glycoproteins and lipids are the first interaction partners dealing with extracellular stimuli. The specific recognition by receptors, carbohydrate binding proteins or antibodies is the first step in a process leading to various kinds of biological outcomes. The development of microarray technology and the application in the field of glyco-sciences during the last decade, allows evaluation of carbohydrate-protein interactions in a high-throughput manner and with smaller amounts of precious, complex carbohydrates. The bottleneck in many cases is the availability of defined carbohydrate structures which are not easily isolated from biological material due to their heterogeneity regarding site-specific glycosylation and stoichiometric carbohydrate occupancy. The aim of this work was to establish a unified strategy for the synthesis of complex mucin type *O*-glycans and the incorporation into peptides in order set up a comprehensive glycopeptide library. The glycopeptides serve as substrates for the build-up of a glycopeptide microarray platform. The combination of peptide and glycan motifs, allows the evaluation of specificities of carbohydrate and glycoprotein binding proteins, such as lectins or antibodies, in the most natural way of glycan presentation.

As peptide scaffolds, tandem repeat sequences from MUC1 and MUC5B are chosen for solid phase peptide synthesis. MUC1 is overexpressed and aberrantly glycosylated on cancer cells, representing a potential candidate for vaccine development. Here, a designed MUC1 glycopeptide array will serve for epitope mapping of serum antibodies generated from administration of potential synthetic vaccine candidates to mice. Further, it is possible to elucidate MUC1 glycopeptide interactions with human lectins involved in cancer progression, such as the interaction of galectin-3 with MUC1. Also plant lectins, commonly used in affinity-based enrichment methods of glycans, glycopeptides and -proteins, will be screened for their binding specificity. The build-up of a library of MUC5B glycopeptides, in addition to the MUC1 peptides, can potentially serve as model system for multivalent lectin association of airway-invading pathogens. Knowledge about virus and microbe airway pathogen interactions are important in order to generate new point-of-care diagnostic tools or can function as starting points for development of new anti-adhesive drugs. An array of mucin glycopeptides may be useful to map pathogen strain fine specificity to different glycan ligands.

The synthesis strategy is envisioned to constitute from a few basic building blocks that could be assembled to diverse glycosylated amino acid building blocks. The synthesis involves construction of two glycosylated threonine amino acid acceptors, which allow chemical elongation into the common mucin-type *O*-glycan core structures. A second central motif is the assembly of type-1 and type-2 LacNAc (Galβ-1,3/4GlcNAcβ) disaccharides, for *O*-glycan core synthesis. The glycan elongation with the LacNAc disaccharides is mimicking the alternating biosynthetic action of GlcNAc- and Gal-transferases. From the stock of the common basic building blocks, a route was planned to synthesize T-antigen, extended core 1, 2 and 3 glycosylated amino acids (*figure 2.1*).

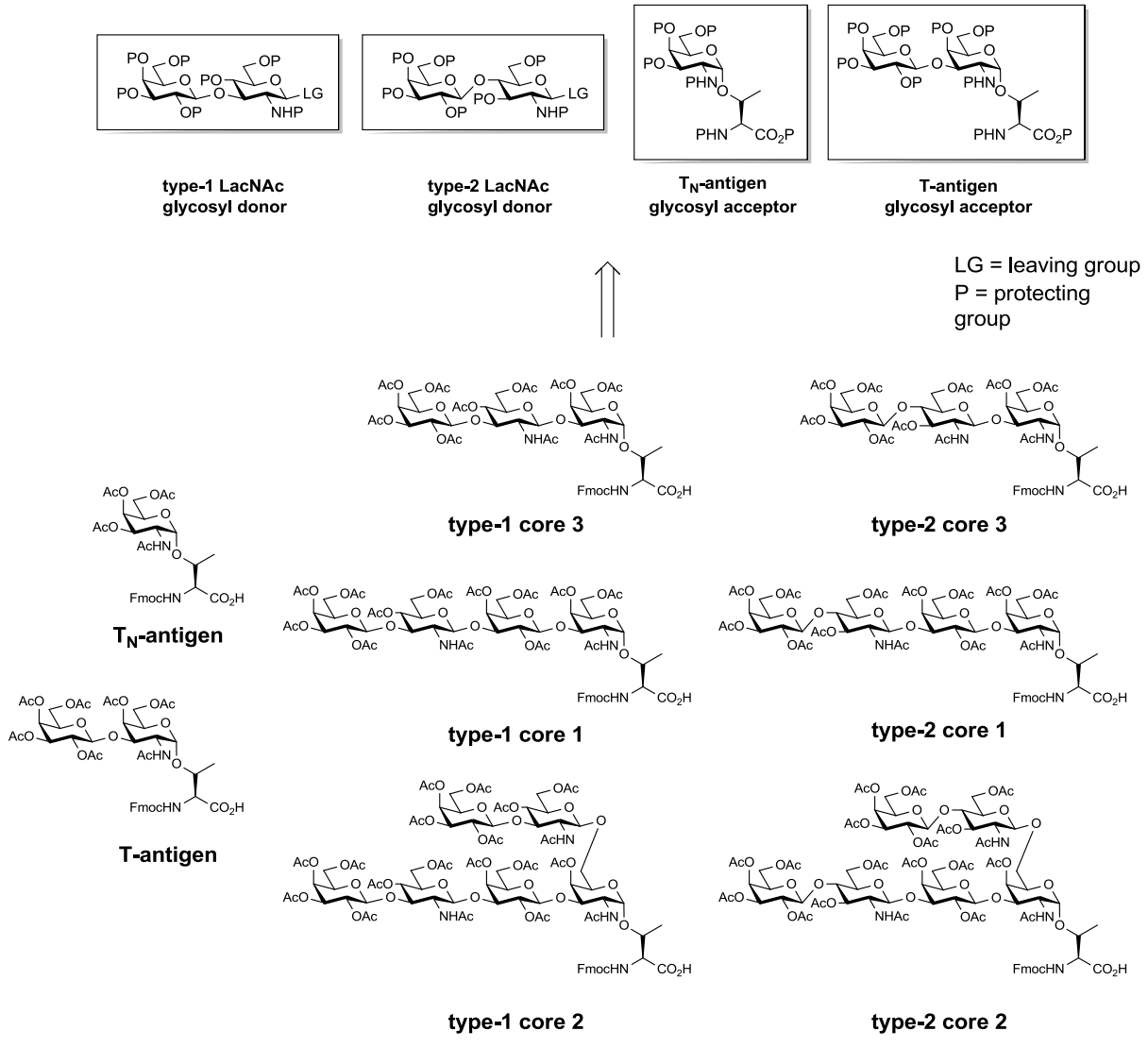


***Figure 2.1*:** Basic pool of glycosyl donors and acceptors for convergent assembly of *O*-glycan mucin core amino acids.

The threonine glycosyl amino acids were then aimed to be introduced into MUC1 and MUC5B peptide sequences of the corresponding tandem repeats. The focus lies on the synthesis of mono-, di- and triglycosylated glycosyl peptides which are homogeneously decorated and arranged in an identical pattern for each glycosyl amino acid (*figure* 2.1).

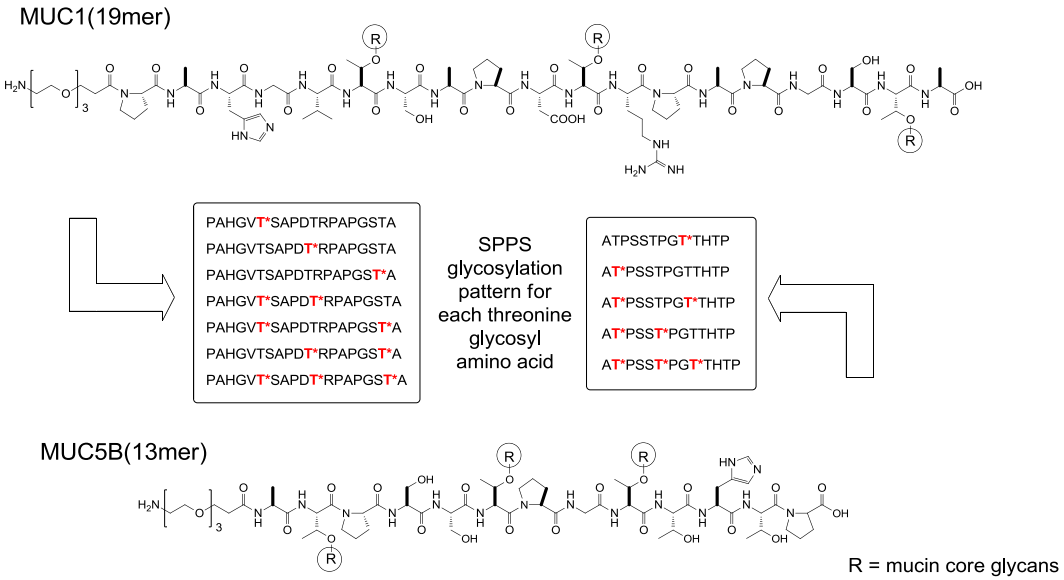


***Figure 2.2*:** Construction of mucin glycopeptide libraries with multivalent glycosylation patterns by solid phase synthesis using the O-glycosylated threonine amino acids.

Glycopeptide diversity can be further increased by applying glycosyl transferases such as sialyltransferases for termination of the glycan structures. Optimal would be a time and resources saving enzymatic approach on immobilized glycopeptides such as a glycopeptide microarray. Three sialyltransferases are planned to be used for the synthesis of sialylated glycopeptides. It is necessary to first evaluate the activity and specificity of the transferases towards the mucin glycoproteins in solution before applying them on a microarray surface and to further generate modified glycopeptide standards for later evaluation of the on-slide enzyme performance. With the sialic acid modified peptide standards in hand and knowledge about the substrate specificity regarding the glycan core structures on the mucin peptide backbones, on-slide enzymatic modification can be approached.

For immobilization of the synthetic glycopeptides, amine reactive, *N*-hydroxysuccinimide functionalized, microarray slides were selected. All glycopeptides are going to be conjugated with a triethylene glycol spacer amino acid with a terminal amine to the *N*-terminal end. In previous studies it has been observed that ligand presentation to surface is dependent on the size of the spacer and a spacer length of at least 6-carbons is recommended. In both mucin peptide sequences the N-terminal linker contains the only free amine and will connect the glycopeptides site-specific to the slides. The established glycopeptide array will then be

utilized in carbohydrate/glycopeptide binding studies. The glycopeptide array is aimed for epitope mapping of serum antibodies from mice, which have been treated with potential synthetic vaccine candidates. Points of interest are the peptide epitopes and the dependence on glycosylation, influences of different glycans according to their structure or size and influences of multivalent glycosylation. An important issue regarding the safe application of potential anti-tumor vaccines, is the specificity of the induced immune reaction. Generated antibodies should selectively target tumor-associated epitopes and omit natural structures, in order to prevent autoimmune reactions. The planned glycopeptide microarray contains short tumor-associated and larger non-tumor-associated glycans in order to evaluate the fine-specificity of the induced antibodies. Further, the array system will be used with carbohydrate binding lectins, such as common plant lectins or human galectins with significance in tumor progression, for instance galectin-3. The carbohydrate binding properties are evaluated when connected to an authentic peptide as a carrier. This work is further supposed to lay the fundament for studies on host-pathogen interactions.

The chemically defined, synthetic glycopeptides synthesized in this work, will also be used in mass spectrometry HCD fragmentation studies. Access to defined carbohydrate or glycopeptide specimens is restricted due to carbohydrate heterogeneity. A biological sample of a glycoprotein has many different glycoforms at the same glycosylation site and variances in occupancy of the glycosylation sites. In typical bottom-up proteomic approaches, the digestion of a set of proteins gives rise to several peptides for mass spectrometric analysis. Protein glycosylation is important for the biological functionality of a protein and the diverse processes in which it is involved. However, glycopeptides obtained in bottom-up proteomics propose a special challenge in mass spectrometric analysis. Differentiation of isobaric fragments of glycan-derived ions and elucidation of the linkage connectivity of the carbohydrates is difficult by means of mass spectrometry. Often, when analysis of protein glycosylation is restricted due to technologic or methodic limitations, carbohydrate structure determination is based indirectly on the knowledge of specificities and expression of glycosyl transferases, glycosidases or carbohydrate binding lectins. However, the carbohydrate specificities of these proteins may not be fully explored and structural influences of the underlying glycan or protein in the substrate alter the activity of the recognizing proteins and enzymes. Absolute and direct methods of analysis of protein glycosylation are needed. The chemically defined, synthetic glycopeptides synthesized in this work, can serve as glycopeptide standards in HCD fragmentation studies. Typical carbohydrate fragmentation products of *O*-glycopeptides will be studied. The relative intensities of the extracted ions will be compared and examined for regularities which will aid in determination of carbohydrate structures from glycoprotein samples. The focus lies on structural glycan determination on the level of the glycopeptide which will facilitate typical glycoproteomic workflows.

3 SYNTHETIC STRATEGIES

3.1 Glycopeptide synthesis strategies

Obtaining chemically homogenous glycopeptides or glycoproteins from biological sources remains still a difficult task, due to the microheterogenic nature of carbohydrate biosynthesis.[153,154] Chemical or chemoenzymatic synthesis is a reliable method to make precise glycoforms available. For the synthesis of glycopeptides, three general approaches are commonly utilized (*figure 3.1*):

A) Synthesis of protected peptides and carbohydrates for convergent fragment condensation
B) Synthesis of simple glycosyl amino acids for solid phase peptide synthesis (SPPS) and chemical or chemoenzymatic elaboration
C) Synthesis of fully glycosylated amino acids for SPPS

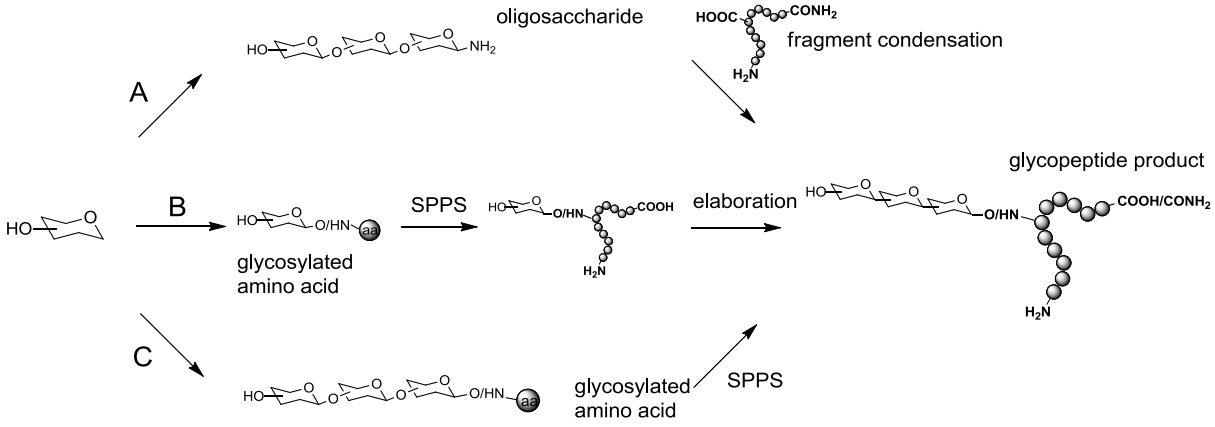


***Figure 3.1*:** Different strategies in glycopeptide synthesis. **A:** Carbohydrate and peptide fragment condensation. **B:** Glycopeptide elaboration on peptide level. **C:** Introduction of full extended glycosyl amino acids in SPPS.

Conveniently protected peptides and carbohydrates can be synthesized in a convergent fashion and then condensed to form glycopeptides (*figure 3.1*, **A**). Convergent assembly often results in low yields since the large structures suffer from sterical hindrance or from poor solubility of the peptides under common glycosylation conditions. This approach is usually applied for the syntheses of *N*-glycopeptides. The common procedure is to convert the unprotected saccharide by *Kochetkov* β-amination into a glycosylamine followed by

Lansbury aspartylation of these glycosylamines,[155,156] for example glycosylation of the *N*-glycan heptasaccharide $Man_5GlcNAc_2NH_2$ with aspartic acid containing peptides.[157] The approach of using simple glycosyl amino acids in SPPS and further elaboration on glycopeptide level is convenient for the synthesis of *N*-glycopeptides (*figure 3.1*, **B**). Endoglycosidases, such as Endo M (*Mucor hiemalis*) and Endo A (*Arthrobacter protophormiae*) can be used to extend the initial, asparagine-linked GlcNAcβ of the *N*-glycan chitobiose core *en bloc* with a presynthesized complex oligosaccharide.[158] Comparable enzymes for *O*-glycan synthesis are not known. *O*-glycans need to be assembled by a stepwise enzymatic elongation protocol using various different glycosyl transferases. An example for the stepwise enzymatic synthesis of a sLe^x containing PSGL-1 *O*-glycopeptide fragment was reported by *C.H. Wong*.[159] The most commonly applied method for generation of glycopeptides, especially of *O*-glycopeptides, is the use of Fmoc-protected, glycosyl amino acid building blocks in Fmoc-SPPS (*figure 3.1*, **C**).

In this research work, complex *O*-glycosylated amino acids were applied, since stepwise chemical and chemoenzymatical manipulations on peptide level are difficult when homogenous products, especially multivalent glycopeptides, should be formed. Thus, a retrosynthetic strategy for the generation of the various glycan core amino acid building blocks regarding glycosylation methods and protecting group chemistry needs to be worked out, also with regards to the demands of the following Fmoc-SPPS (see *chapter 3.2.3*). This approach can still be extended by enzymatic termination of the complex glycan moiety by e.g. sialylation, fucosylation or sulfation, to increase diversity and generate specific glycan motifs.

3.1.1 Glycopeptide synthesis on solid phase

Although glycosylation is a highly abundant posttranslational modification, isolation and characterization is difficult due to the microheterogenic nature of glycoproteins.[154,160] Often when glycopeptides are applied in biological experiments, the glycans need to be attached on specific glycosylation sites on the peptide backbone as a homogenous glycoforms. SPPS can provide defined glycopeptides in preparative amounts. When SPPS was introduced in 1963 by *Merrifield*, a simple tetrapeptide model was synthesized, to demonstrate the applicability of the method.[161] In SPPS the peptide sequence is build-up stepwise, amino acid by amino acid, on a polymer resin as insoluble support. Unreacted amino acids and coupling reagents are washed away after each step by simple filtration of the resin. One year after the method was introduced, it demonstrated a huge potential, by the synthesis of the biologically active nonapeptide bradykinin.[162] *Merrifields* method relied on N-terminal *tert*-butyl

oxycarbonyl (Boc) protection and aqueous hydrogen fluoride for release of the peptides from the resin. After Fmoc was introduced as a base labile N-terminal protecting group,[163] a milder variant of SPPS based on Fmoc protection was developed.[164,165] The Fmoc-SPPS is also suitable for the synthesis of glycopeptides, since the conditions applied, neither harm the acid sensitive acetal groups of the glycosidic bonds nor the base labile connection of *O*-glycans to the peptide backbone.

Fmoc-SPPS utilizes an orthogonal protecting group strategy. The base labile Fmoc is cleaved after each elongation step and the acid labile amino acid side-chain protecting groups are cleaved along with the release of the peptide from the resin. The peptide is attached by the C-terminal end to the resin and the chain grows towards the N-terminal end (opposed to natural protein biosynthesis). The general process consists of 1) attachment of a N^{α}-protected amino acid and 2) removal of the N^{α}-protection. These steps are repeated in cycles until the desired chain length is reached and the peptide is 3) deprotected on the N-terminal side and 4) released from the resin (*figure 3.2*).

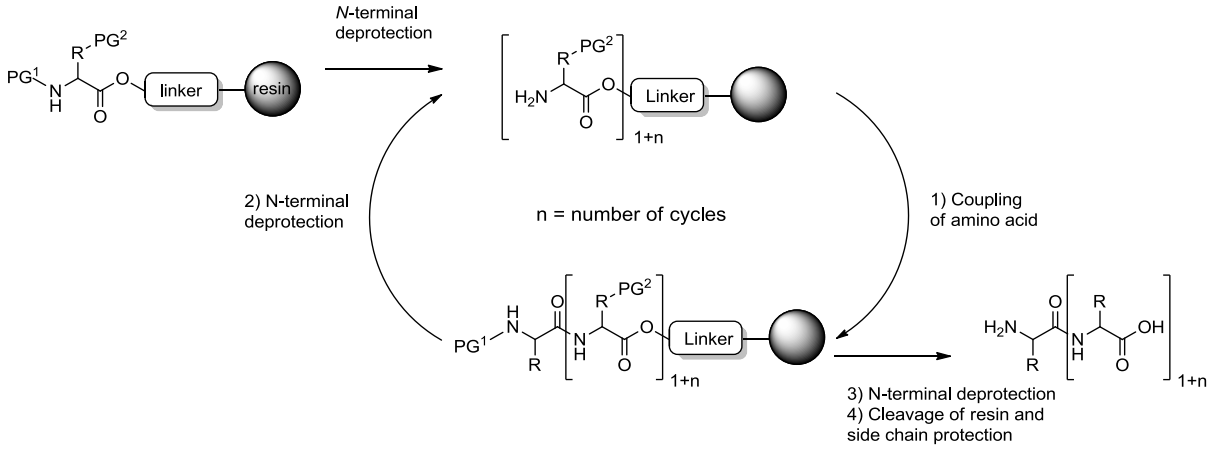


***Figure 3.2*:** Principle of solid phase peptide synthesis (SPPS). R = amino acid side-chain. PG^1 = N-terminal protecting group. PG^2 = Side-chain protecting group.

The peptide chain is synthesized on a polymer resin in the form of beads. The beads usually consist of polystyrene (PS), cross-linked with 1% divinylbenzene. *TentaGel*®-resins (*Rapp Polymere GmbH,* Tübingen) are copolymers of polyethylene glycol (PEG) attached to the PS matrix. The PEG chains contribute with hydrophilic properties and render the polymer more polar, which increases the ability to swell in polar solvents. Thereby, the reaction centers on the resin become more accessible. A linker moiety can be attached to the end of the PEG-chains. These linkers are stable under the conditions applied in peptide synthesis and serve as cleavable handles for the release of the peptide at the end of the synthesis. The linker chemistry determines the cleavage conditions and the nature of the C-terminal functional group of the peptide. Also, commercial resins can be obtained preloaded with the first C-

terminal Fmoc-protected amino acid. In this work, Fmoc-amino acid preloaded *TentaGel*® *R* Trt resins were used (*figure 3.3*).

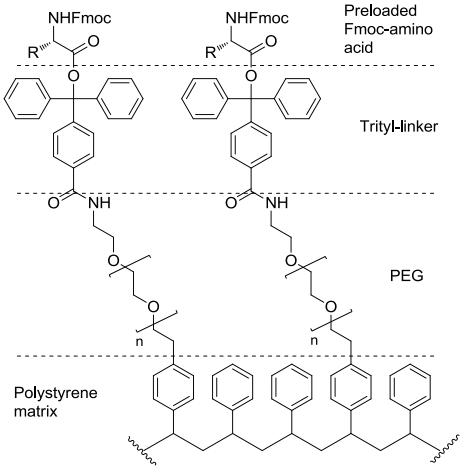


***Figure 3.3*:** Representation of *Tentagel*® trityl resin.

The peptide ester is cleaved from the trityl-linker under acidic conditions with neat trifluoroacetic acid (TFA), usually with addition of cation scavengers such as triisopropylsilane (TIPS), thiophenol or water. Albeit lower amounts of acid could be used for the trityl ester cleavage (1% TFA), higher amounts of TFA also remove all acid labile protecting groups from the side-chains of the amino acids. Among the amino acids used for the MUC1 and MUC5B sequences, arginine, histidine, threonine, serine and aspartate carried acid labile protecting groups (*figure 3.4*).

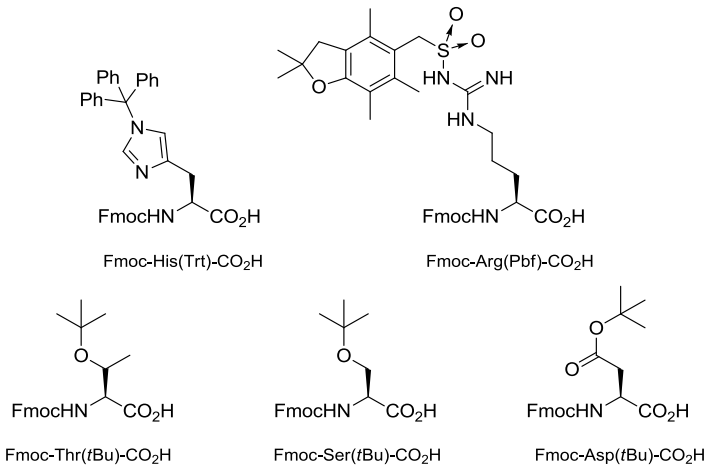


***Figure 3.4*:** Acid labile protecting groups on amino acids used for MUC1 and MUC5B glycopeptide syntheses.

The coupling of the Fmoc-amino acids is in general mediated by *in situ* activation of the C-terminus as an active anhydride/ester intermediate. Typical coupling reagents are carbodiimids like *N*,*N*'-dicyclohexylcarbodiimide (DCC),[166] *N*,*N*'-diisopropylcarbodiimde (DIC) or *N*-ethyl-*N*'-(dimethyaminopropyl)-carbodiimide-hydrochloride (EDC·HCl), phosphonium

salts like benzotriazol-1-yloxytris(pyrrolidino)phosphonium hexafluorophoshat (PyBOP)[167] or aminium/uronium salts like *O*-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethylunronium hexafluorophosphate (HBTU),[168] *O*-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethylunronium tetrafluoroborate (TBTU)[169] or *O*-(1H-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethylunronium hexafluorophosphate (HATU). Commonly, the hydroxybenzotriazoles HOBt[170] or HOAt[171] are used as additives along with the corresponding aminium/uronium salt, to further increase reaction speed and reduce the tendency for racemization (*figure 3.5*).[172]

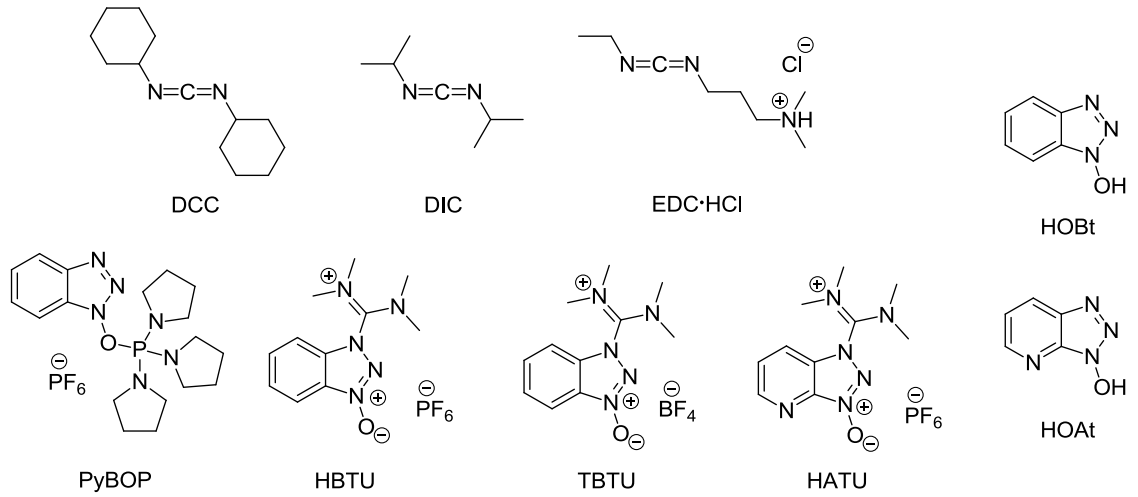


***Figure 3.5*:** Common used carbodiimide and benzotriazole peptide coupling reagents.

In terms of reactivity and lower tendency for racemization on C^{α} of the amino acids, the phosphonium and aminium salts are more effective than the carbodiimids. DCC forms dicyclohexylurea with low solubility. In Boc-SPPS the urea is dissolved in TFA during Boc-removal, but during Fmoc-SPPS the urea precipitates and cannot be removed by washing. Aminium salts derived from the corresponding benzotriazoles 1-hydroxybenzotriazole and 7-aza-1-hydroxybenzotriazole are very commonly used soluble coupling reagents. When the aminium salts were introduced, they were first described as uronium salts. Later, it was shown that HBTU and HATU actually crystallize in the aminium (guanidinium *N*-oxide) form, when synthesized by the originally reported procedure.[173,174] However, they are still often referred to as uronium salts. Both forms can be synthesized selectively but in solution and under basic conditions the *O*-form (uronium) isomerizes into the *N*-form (aminium). Both forms are able to activate carboxylic acids, albeit the *O*-isomer is reported to be more reactive than the thermodynamically more stable *N*-isomer (*figure 3.6*).

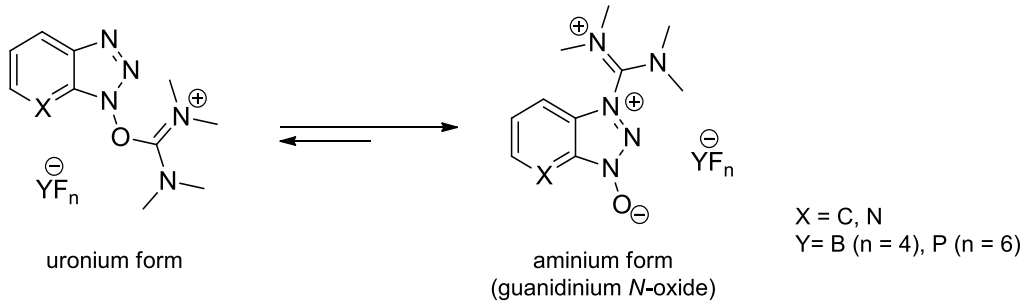


***Figure 3.6*:** Uronium and aminium forms of benzotriazole coupling reagents.

The HATU/HOAt-system is reported to be more effective in terms of reactivity and lower epimerization tendencies, compared to HBTU/HOBt-systems and is preferred in couplings with sterically demanding amino acids. The increased reactivity of the 7-aza-1-hydroxybenzotriazoles could is based on a neighboring effect of the pyridine nitrogen during formation of the active ester intermediate (*figure 3.7*).[171] The carboxylic acid component is preactivated with HATU and HOAt and *N*-ethyldiisopropylamine (DIPEA, *Hünigs Base*). The deprotonated acid attacks the electrophilic carbon in the guanidyl group of the guanidinium *N*-oxide to form an *O*-acylisouronium salt. This further reacts with the HOAt to give an active ester intermediate. The pyridine nitrogen of this intermediate presumably helps in coordination of the approaching amine by hydrogen bonding. Finally the amine can attack the ester carbonyl group and replace the benzotriazole to form a peptide bond (*figure 3.7*).

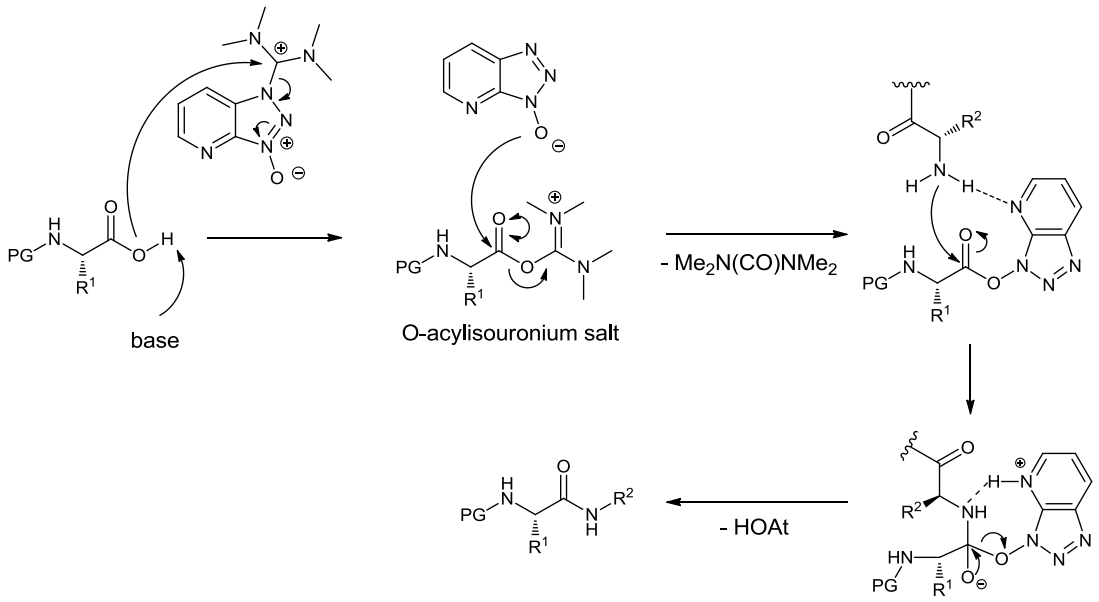


***Figure 3.7*:** Reaction mechanism of HATU mediated peptide coupling.

In this work HBTU/HOBt was used in formations of peptide bonds between unglycosylated amino acids. The more reactive HATU/HOAt system was utilized for couplings with the glycosylated amino acids. If not otherwise stated, commercial Fmoc amino acids were used

in excess of 8 equivalents relative to the molar resin binding capacity. The ratio of the reagents donor amino acid/HBTU/HOBt/DIPEA was 1 : 0.95 : 0.95 : 2. HBTU and HATU were applied in slightly substoichiometric amounts compared to the donor amino acid. The aminium salts are known to guanylate free amines and even hydroxyl groups. Capping of the N-terminus may occur, if too much of the coupling reagent is applied (*figure 3.8*).

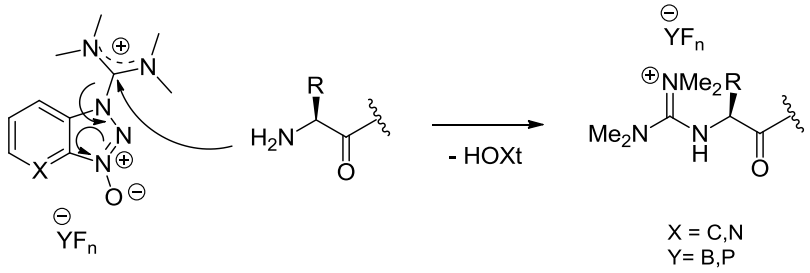


***Figure 3.8*:** Guanidylation of *N*-terminal amino groups by excess of used aminium/uronium reagent.

All peptide syntheses were conducted by applying the Fmoc-based SPPS protocol. The Fmoc group is base-labile and can be cleaved off with 20%-50% piperidine, morpholine or a mixture of DBU (1,8-diazabicycloundecene) and piperidine.[175,176] The piperidine serves also as a scavenger for the cleaved dibenzofulvene, which could otherwise block the liberated N-terminus. The base-catalyzed deprotection proceeds *via* an E1cb mechanism. The H9 proton of the fluorene ring is acidic due to the cyclopentadienyl-type anion as intermediate (*figure 3.9*).

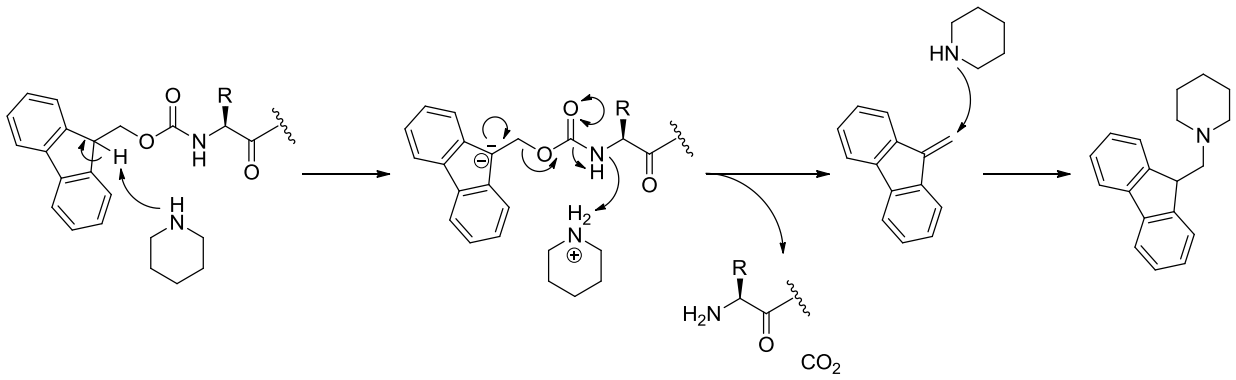


Figure 3.9: Piperidine mediated Fmoc-deprotection and reaction with dibenzofulvene.

3.1.2 Side reactions in solid phase glycopeptide synthesis

Side-reactions in SPPS can lower yields or even prevent product formation at all. A very prominent side-reaction is the diketopiperazine formation. Although also present in Boc-SPPS, diketopiperazine formation is a prominent problem in basic Fmoc-deprotection.[177] On

the level of a synthesized dipeptide, the liberated amine may attack the ester bond between the linker and the peptide leading to the cyclic piperazinedione (*figure 3.10*). Diketopiperazine formation is favored on sterical less demanding alkoxybenzyl-linkers (e.g. *Wang* resins). Bulky linkers, like the used trityl-linker, can suppress this side-reaction. C-terminal glycine and proline amino acids are prone to form diketopiperazines. In these cases, it may be helpful to elongate with a preformed dipeptide onto the first attached amino acid.

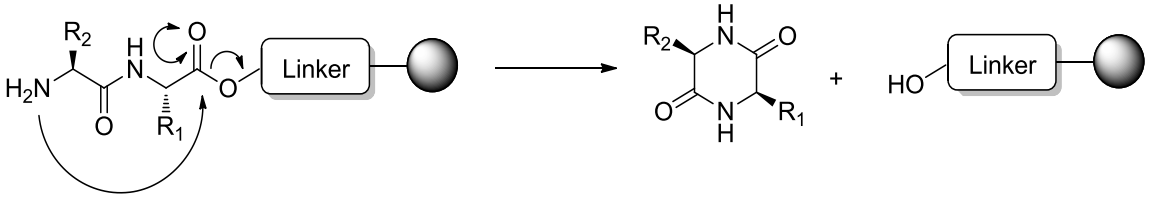


***Figure 3.10*:** Diketopiperazine secession from dipeptide loaded solid phase resin.

Another prominent side-reaction is the formation of aspartimides. The nitrogen on the α-carboxylic acid may react with the ester on the side chain of a protected aspartate ester to form a cyclic succinimide (*figure 3.12*). This intermediate can be hydrolyzed later in solution to form α-and β-aspartidyl peptides or it may react with piperidine during the Fmoc-deprotection procedure to yield α- and β-piperidyl peptides. Aspartimide formation may happen during chain elongation as well as under the acidic conditions applied during peptide release from the resin. It is sequence dependent and related to the amino acid coupled before the aspartate in the chain.[178] For Fmoc-SPPS it is reported that a preceding Thr(*t*Bu) and Thr residues can promote aspartimide formation.[179]

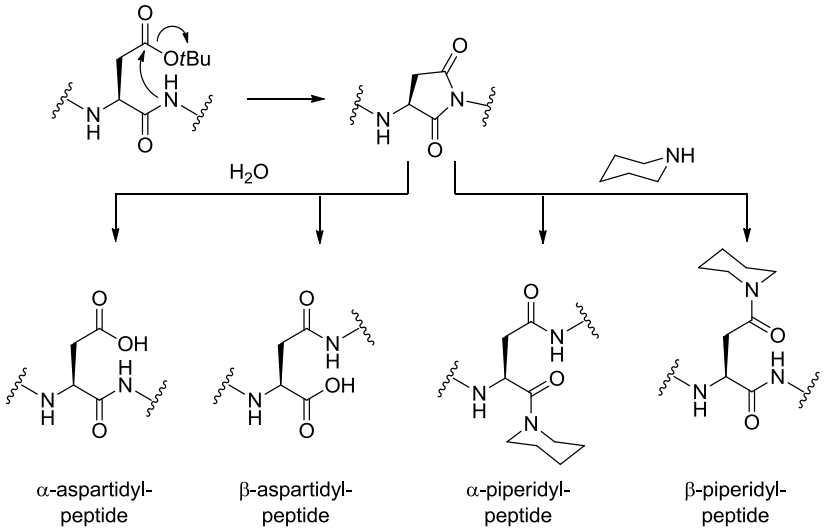


***Figure 3.11*:** Aspartimide formation and reaction with water during peptide cleavage and with piperidine during Fmoc-deprotection.

In *O*-glycopeptide synthesis, a common side-reaction is related to the stability of the glycosidic bond connecting the αGalNAc to the threonine amino acid in the synthesized peptides. The *O*-glycosylated amino acids are introduced to the peptides during syntheses with *O*-acetyl ester protecting groups on the carbohydrate hydroxyl-groups. Treatment with a base is needed to liberate the glycans after cleavage from the resin. Strongly basic conditions may result in abstraction of the C^{α}-hydrogen by the base and initiate β-elimination of the glycan by an E1cB mechanism (*figure 3.12*).[180]

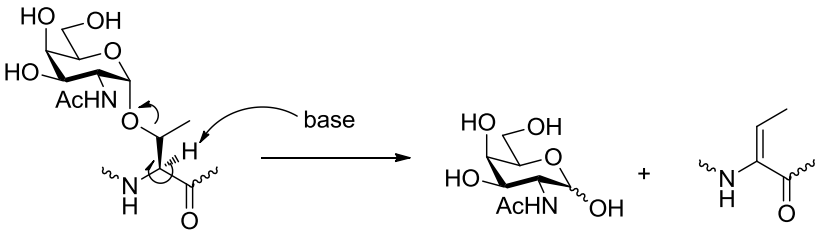


***Figure 3.12*:** β-Elimination of glycans at high pH.

Related to β-elimination, elevated pH during deacetylation may result in epimerization of the stereogenic centers on C^{α} of the amino acids (*figure 3.13*). However, epimerization and β-elimination are reported to be surprisingly low during basic acetyl removal, probably due to a protection by the deprotonation of the nitrogen in the peptide amide bond.[180]

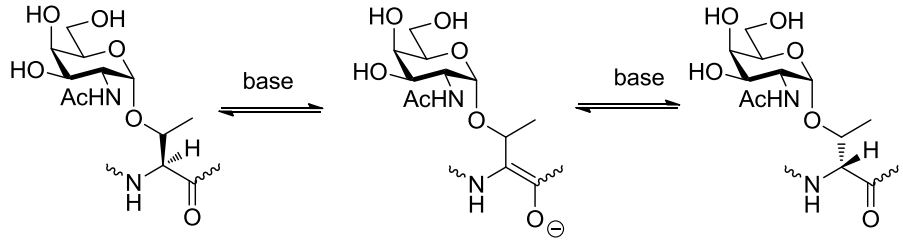


***Figure 3.13*:** Base-induced epimerization of glycosylated peptides.

3.2 Strategies in carbohydrate synthesis

3.2.1 Glycosylation methods

An important issue in chemical glycosylations, is the choice of appropriate anomeric leaving groups and suitable promoters or catalysts for glycosylation reactions. Many different glycosylation procedures are reported and it is not always obvious which is the most suitable for formation of a certain glycosidic bond. The first reported glycosylation reactions either used basic activation of the alcohol acceptor[181] or acidic activation of the glycosyl donor.[182] *W. Koenigs* und *E. Knorr* shortly afterwards reported reactions of C1 halogenated carbohydrates with alcohols and silver salts as promoters. The combination of a leaving group at C1 with a metal promoter provided irreversible glycosylations, which could be applied to various acceptors (*figure 3.14*).

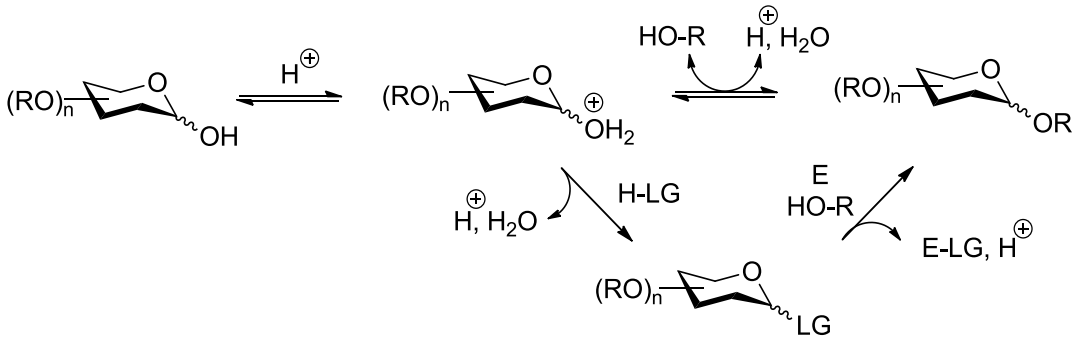


***Figure 3.14*:** General acid catalyzed glycosylation with and without leaving group. LG = leaving group. E = electrophilic promoter/catalyst.

Glycosyl halides are still very useful donors in glycosylation reactions, however they require stoichiometric amounts of a promoter that is consumed during the reaction (hence, not a catalyst). For a long time the *Koenigs-Knorr* glycosylation with glycosyl halides was the only reliable glycosylation method. In the last decades, further glycosylation strategies were established (*figure 3.15*).

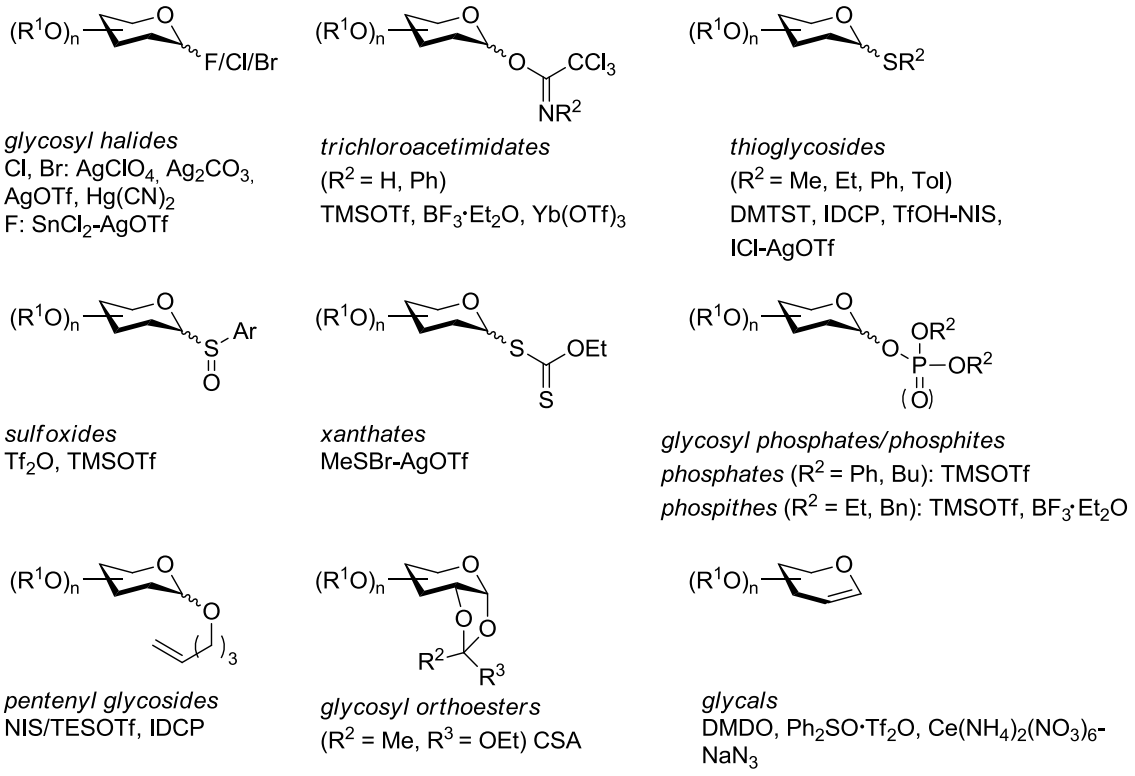


Figure 3.15: **C**ommon leaving groups and corresponding promoters/catalysts for chemical glycosylations.

Among the most commonly used glycosylation methods, besides the glycosyl halides, are glycosylations with thioglycosides and trichloroacetimidates. Because of their high stability, the thioalkyl/aryl moiety in thioglycosides can be introduced early in the synthesis of the buildings blocks and also function as anomeric protecting group. The thioglycosides are activated with thiophilic reagents commonly containing or generating iodonium- (e.g. iodonium dicollidine perchlorate (IDCP)[183], NIS/TfOH[184,185]) or sulfonium-ions (e.g. dimethyldithiosulfonium triflate (DMTST)[186]). Furthermore, glycosyl thioglycosides are usually stable under glycosylation conditions required for other glycosylation methods and can therefore orthogonally act as glycosyl acceptors at other positions of the carbohydrate.[187] Glycosyl trichloroacetimidates (reactions are often referred to as "*Schmidt-glycosylations*") are usually highly reactive and easy to handle.[188,189] They can be activated with catalytic amounts of a lewis acid without any further promoter (hence a catalyst), e.g. trimethylsilyl-triflate (TMSOTf), triflic acid (TfOH) or boron triflouride (BF_3).

3.2.2 Protecting group chemistry

A prerequisite for the successful synthesis of many different variants of complex glycans from basic building blocks, is a unified protecting group chemistry. Protecting groups allow temporary blocking of functional groups that would otherwise equally react and lead to unwanted results. Especially in carbohydrate synthesis, protecting groups have a significant influence on reactivity and stereoselectivity.[190] *B. Fraser-Reid* rationalized the relation of electronic effects by protecting groups and the reactivity of glycosyl donors.[191,192] Electron-withdrawing acyl protecting groups decrease the reactivity of the donors by decreasing the nucleophilicity of the leaving group in comparison to donors protected with electron-donating ether groups. Decreased electron density at the anomeric center further destabilizes the intermediate oxonium ion. Hence, protecting groups were divided into “disarming” (esters, amides) and “arming” (ethers) groups. However, an ester group in 2-position of the carbohydrate ring can stabilize the oxonium ion after leaving group separation together with the lone pair electrons of the ring oxygen. Hence, participating groups in 2-position in combination with global ether protecting make very reactive glycosyl donors (*figure 3.16*).[193,194]

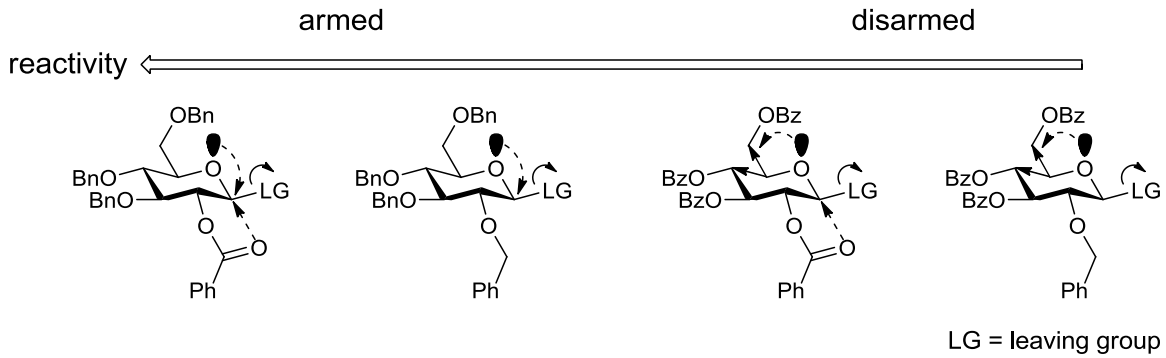


***Figure 3.16*:** The general reactivity of glycosyl donors according to the armed/disarmed concept.

Configurational changes on the carbohydrate have further influence on the donor reactivity and hence result in stereo-electronic effects. For example, axial configured substituents on the carbohydrate ring increase reactivity compared to equatorial oriented substituents and fused ring systems may disturb the planar sp^2-geometry of the oxonium ion intermediate after leaving group separation.[195,196]

Protecting groups at the 2-position have major influence on whether the glycosylation will give an α-(1,2-*cis*) or β-(1,2-*trans*) configured product.[190] Esters, amides and carbamates in 2-position behave as participating neighboring groups and will direct glycosylation reactions towards the β-product. Ethers and azides are non-participating neighboring groups and α- as well as β-glycosides can be formed (*figure 3.17*).

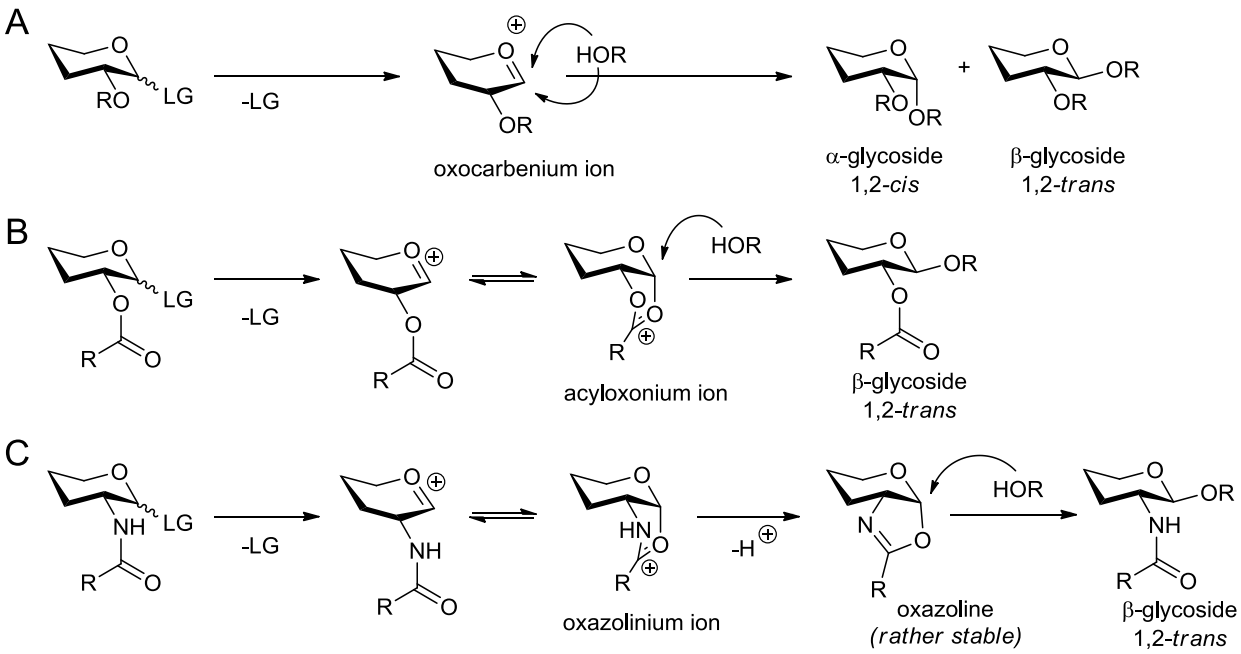


***Figure 3.17*:** Influence of neighboring groups. **A**: Ether groups adjacent to the C1 position do not participate in the glycosylation. **B:** β-directing influence of neighboring ester groups. **C:** β-directing influence of neighboring amide groups *via* an oxazoline intermediate.

Due to the directing influence of participating neighboring groups, stereoselective formation of 1,2-*trans*-glycosides is rather predictable, although unexpected exceptions may be possible. Synthesis of 1,2-*cis*-glycosides requires the absence of a participating group and it is challenging to perform the glycosylation with full stereoselectivity. Although the α-anomer is favored due the anomeric effect, diastereoselective mixtures are commonly obtained.[197] The strategies for preparing 1,2-*cis*-selective reactions are permanently extended since the need for comprehensive methods is still high. A widely-used method is the intramolecular aglycon delivery (IAD) introduced by *O. Hindsgaul* for the preparation of β-mannosides (*figure 3.18*, **A**).[198] The glycosyl donor and acceptor are first linked by a protecting group at the 2-position of the donor (an isopropenyl ether in the original publication) and upon activation of the leaving group an intramolecular S_N2-transfer of the acceptor takes place to form the 1,2-*cis*-linkage. Several modifications of IAD using different tethers are reported. A recent example by *Y. Ito* describes the oxidative coupling of donor and acceptor with a 2-naphtylmethyl ether group as tether.[199] The group of *G.J. Boons* presented an approach that makes use of a chiral auxiliary as protecting group in 2-position (*figure 3.18*, **B**).[200,201] A (S)-phenylthiomethylbenzyl ether participates by forming a six-membered ring, fused to the carbohydrate ring. Due to the equatorial orientation of the two phenyl groups, a *trans*-decalin-like intermediate favors an attack from the *cis*-face of the acceptor. A strategy for *trans*-selective glycosylations without an acyl group in 2-position was introduced by *A.V.*

Demchenko (*figure 3.18*, **C**). A 2-*O*-picolinyl ether group at the 2-position occupies the *cis*-site and forms stereoselectively 1,2-*trans*-glycosides. Recently, the same group described participation of picolinyl- and picoloyl substituents, also when located on one of the other positions of the carbohydrate ring. In this case, the participating group does not block the anomeric center from one site, but the pyridine ring presumably participates by hydrogen bonding with the hydrogen of the nucleophilic alcohol, facilitating deprotonation and putting the reaction partners in close proximity. The acceptor is *syn*-oriented with regards to the picolinyl/picoloyl-group. Depending on which position of the carbohydrate ring is substituted by the participating ether, either *cis*- or *trans*-products can be formed stereoselectively. This method for remote stereoselective control was termed H-bond mediated aglycone delivery (HAD).[202,203] Several more strategies for stereoselective glycosylations are reported and the field is permanently extended.[204]

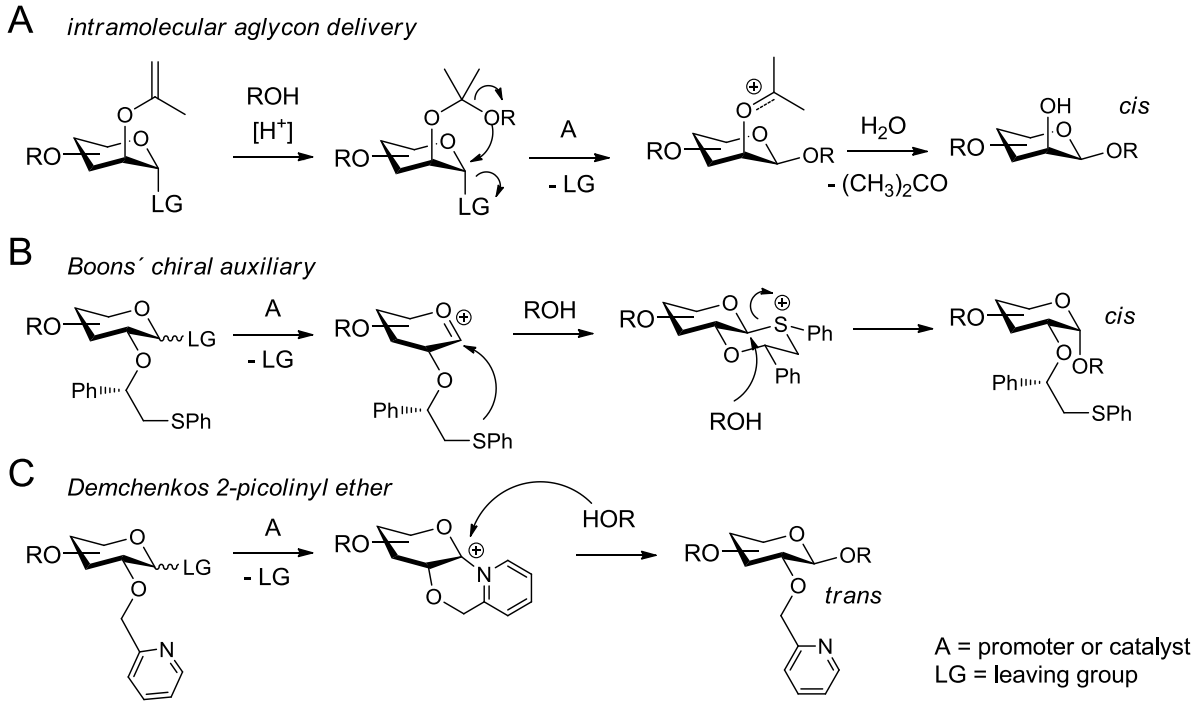


***Figure 3.18*:** Selected examples for stereoselective glycosylation methods.

4 RESULTS AND DISCUSSION

4.1 Part 1 – Synthesis of core 1, 2 and 3 glycosylated amino acids

4.1.1 Retrosynthetic considerations for glycosyl amino acid building block synthesis

Based on the inherent structural canon of mucin *O*-glycan biosynthesis (*chapter 1.1*), a retrosynthesis route for the complex core 1, 2 and 3 amino acid building blocks was envisioned. Biosynthetic backbone elongation happens by the combined alternating action of a β1,3-GlcNAc-transferase, followed by either a β1,3-Gal-transferase or a β1,4-Gal-transferase. The alternating transfer of either a 1,3- or 1,4-linked βGal and a βGlcNAc results in elongation by type-1 (Galβ1,3-GlcNAcβ) or type-2 (Galβ1,4-GlcNAcβ) *N*-acetyllactosamine disaccharides.[17,205,206] Thus, a synthetic route was imagined that includes chemically preformed type-1 and type-2 disaccharide units, with proper protecting groups for further glycosylation reactions and if possible already for later Fmoc-SPPS. The type-1 disaccharide **27** and type-2 disaccharide **31** for later elongation of the core structures should be formed in trichloroacetamide couplings of donor **18** and acceptors **24** (glycosylation **I**) and acceptor **30** (glycosylation **II**) (*figure 4.1*).

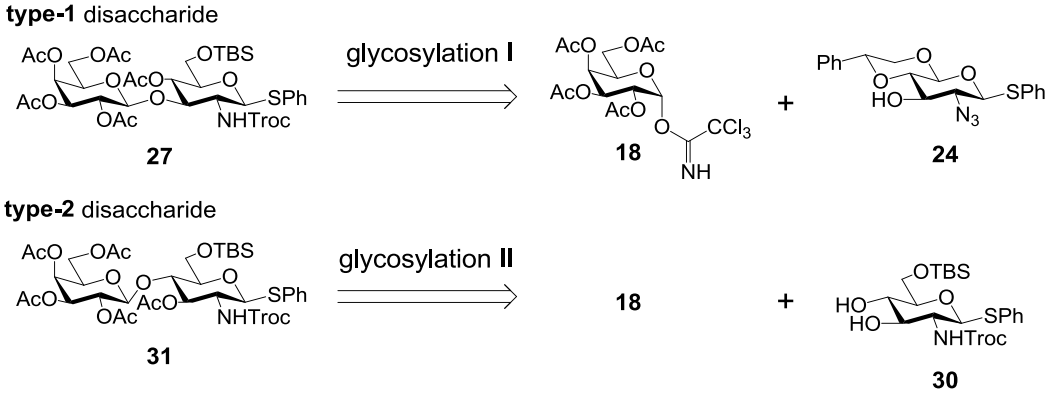


***Figure 4.1*:** Retrosynthesis of the type-1 and type-2 LacNAc disaccharide building blocks.

The disaccharides were chosen to be thioglycosyl donors, since the thiophenyl group in acceptors **24** and **30** is orthogonal to several glycosylation conditions and can be installed prior to trichloroacetamide couplings.[187]

In mucin glycoproteins, the *O*-glycans are linked to threonine and serine amino acid by an initial αGalNAc. In this synthetic approach the glycans are build-up on an appropriately protected threonine amino acid. Thus, the threonine amino group is protected by a Fmoc-group, according to the Fmoc-SPPS-protocol, while the carboxyl group is permanently blocked by a *tert*-butyl-ester during glycan assembly and cleaved off just before SPPS. The attachment of the first αGalNAc to the protected threonine amino acid to form the T_N-antigen **12** is to be carried out using a galactosyl bromide with a classical $AgClO_4/Ag_2CO_3$ promoter system, according to reported procedures.[207,208] Glycosylation of donors **27** and **31** with acceptor **12** will be carried out to give extended core 3 structures **35** and **39** (*figure 4.2*, **III**). Similarly, the extended core 1 tetrasaccharides amino acids **49** and **54** should be synthesized with the preceding glycosylation of the monosaccharide thiophenyl glycoside donor **42** (*figure 4.2*, **IV**+**V**). Finally the extended core 2 hexasaccharide amino acids **58** and **62** are branched from the synthesis routes of the extended core 1 structures by a further thioglycoside coupling with donors **27** and **31** in 6-position of the primary αGalNAc (*figure 4.2*, **IV**+**V**+**VI**).

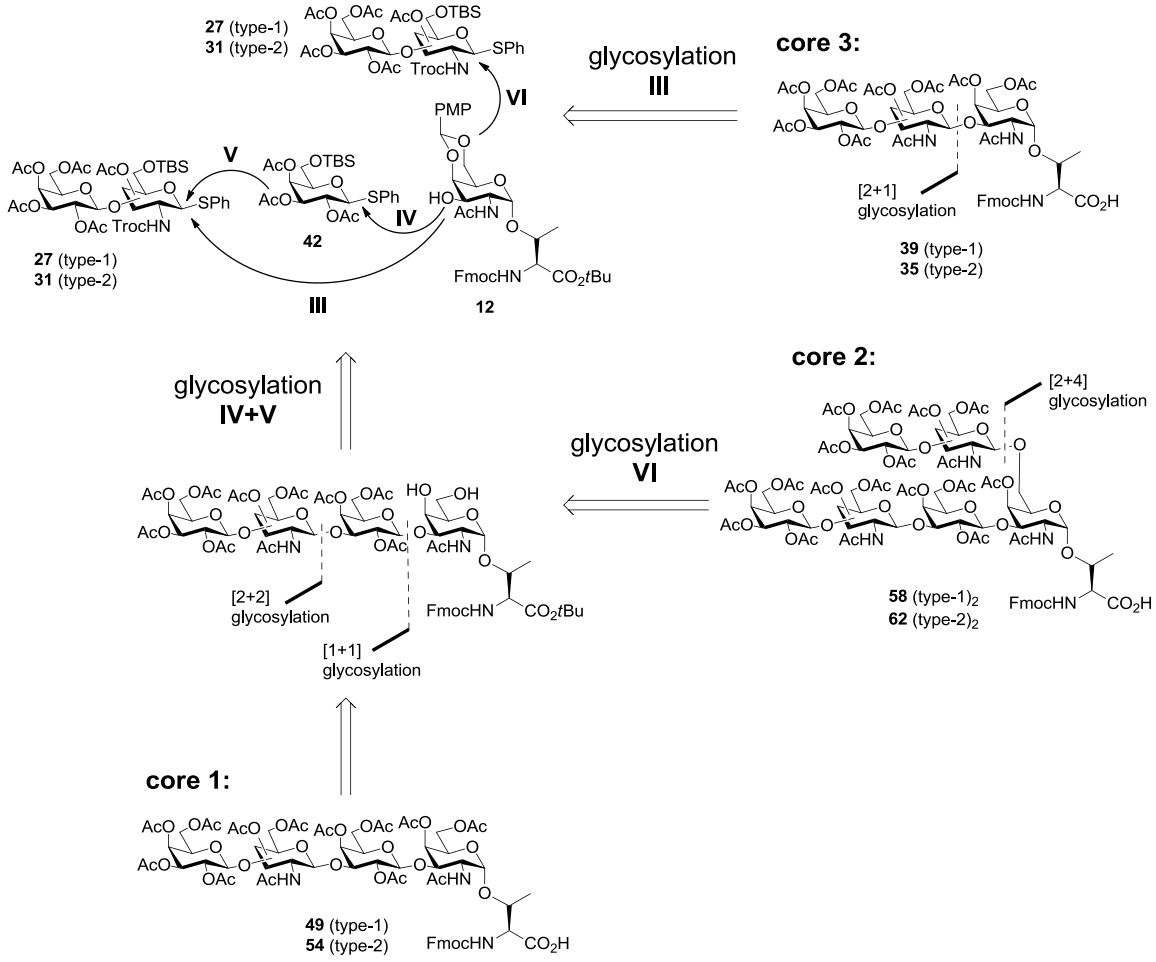


***Figure 4.2*:** Retrosynthesis of core 3 (glycosylation **III**), core 1 (glycosylations **IV**+**V**) and core 2 (glycosylations **IV**+**V**+**VI**) glycosylated threonine amino acids.

The final glycosyl amino acids for SPPS will be globally protected with *O*-acetyl-esters on the carbohydrate hydroxyl groups. Acetyl groups are known to stabilize glycosidic bonds due to electron withdrawing effects, especially during acidic trifluoroacetic acid treatment required for release of the synthesized glycopeptides from the resins.[209,210] After glycopeptide synthesis the acetyl groups can be cleaved by mild treatment with sodium methoxide in methanol (*Zemplén* conditions) or sodium hydroxide in methanol.[211] The temporary protecting group strategy relies on the use of silyl-ethers, acetyls and acetals for the carbohydrate hydroxyl-groups, while 2-*O*-deoxy-acetamide groups are protected by trichloroethoxycarbonyl groups (Troc) or are masked as azides.[212] The initial 2-amino-groups of the D-sugars are not directly converted to acetamide groups. During further glycosylations, the neighboring acetamide group would participate and stabilize the intermediate oxocarbenium ion by rapidly rearranging to an 1,2-oxazoline (*figure 3.17*, **B**). The 1,2-oxazoline intermediates are known to be rather stable and are poor glycosyl donors. In carbamates, in contrast to amides, the likeliness to form rather stable oxazolines is reduced because of the additional electron donating alkoxy group.[213] Therefore, the glycosyl donor reactivity, especially in the case of Troc-protection, is known to be increased during glycosylations.[214] Azides are employed during glycosylations either when a non-participating neighboring group is required for α-anomeric stereoselectivity, or when due to steric hindrance a small and linear group is beneficial. To keep the number of different protecting groups low, regioselective glycosylation preferences are used within the reactions. According to the slogan “The best protecting group is no protecting group”,[212] this represents an elegant way to minimize yield repressing and time consuming steps. For example, the superior reactivity of the equatorial 3-OH of galactose can often be used in glycosylations, without any protection on positions 2 and 4.[215,216,217] In contrast, equatorial 4-OH in GlcNAc is preferentially glycosylated over the equatorial 3-OH.[218,219] Further the advanced reactivity of primary alcohols over secondary alcohols can be used to glycosylate the carbohydrate in 6-position, without effecting the unprotected 4-position.[220]

The synthesis is highly convergent, meaning that from a few common buildings blocks in stock, a high diversity of various glycosyl amino acids can be obtained. The retrosynthetic strategy is finally based on four basic building blocks, namely the type-1 and -2 LacNAc disaccharides **27** and **31** as glycosyl donors and the T_N- and T-antigen as acceptors **12** and **44** (*figure 4.3*, **A**). Apart from the here presented core 1, 2 and 3 glycans, different synthesis variations thereof are further possible. Based on the developed synthesis route, core 2 tetrasaccharide and core 4 pentasaccharide were also synthesized outside the frame of this thesis during the course of this project (*figure 4.3*, **B**).

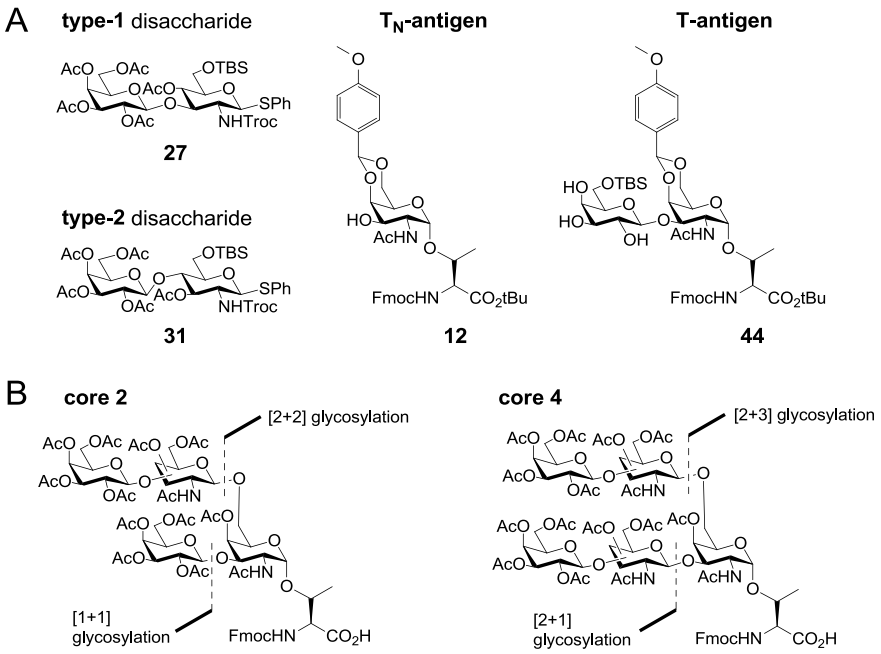


***Figure 4.3*: A:** Basic building blocks used for convergent core structure synthesis presented in this work. **B:** Core 2 and core 4 structures for further syntheses.

4.1.2 Examples for syntheses of glycosylated amino acids/glycopeptides in the field and comparison to the strategy of this work

Selected examples for core 1 structures came from early reports of *H. Paulsen* and *J.-P. Hölck*[221] and sialylated core 1 building blocks from *H. Kunz*,[222,223] applied in Fmoc-SPPS. In an approach to simplify the synthesis of core structures *K.J. Bock* and *H. Paulsen* reported a synthesis for the basic core 1-4 structures without further elongation or termination, ready for Fmoc-SPPS.[224] Elongated core 3 type-2 and core 1 type-2 amino acids are reported by *S.J. Danishefsky* as part of Le^y-antigens through glycal assembly and were incorporated into peptides.[225] There are some examples for core syntheses using a type-2 disaccharide for core extension by *Y. Nakahara* (core 1 and core 3 type-2),[226,227] *C.R. Bertozzi* (core 1 type-2)[216] and *H. Kunz* (core 2 type-2).[228] The protecting group and building block strategy of these examples were directed towards certain structures, specifically terminated (e.g. fucosylated) and therefore less universal for the design of various core structures, in contrast to the strategy presented in this thesis. The possibilities for glycan termination by means of chemical building blocks were not in the focus, but rather the convergent build-up of several scaffold core structures with the possibility of increasing the glycan diversity in a later step on glycopeptide level by a chemoenzymatic approach. Hence the protecting group chemistry should be kept as simple as possible for a high variability in core synthesis using only few

building blocks. Examples for chemoenzymatic modifications of mucins carrying a chemical presynthesized basic core 2 trisacharide are reported by *S.-I. Nishimura*.[229,230] The variable building block system presented in this research work, also includes type-1 elongation of core glycans, which has so far not been reported.

4.1.3 Synthesis of the T_N-antigen glycosyl acceptor

According to the mentioned convergent synthesis strategy (*chapter 4.1.1*), the synthesis of all mucin-type *O*-glycan core structures is based upon the generation of a central *N*-acetylgalactosamine glycosylated threonine amino acid (T_N-antigen), the fundamental motif for mucin-type glycosylation. The first GalNAc amino acid synthesis, employing the serine equivalent, has been described by *T.Osawa et al.*[231] *B. Ferrari* and *A. A. Paviat* later reported the synthesis of the threonine conjugate.[232] The methodology applied in this work is based on reported methodology developed by *H. Paulsen* and *J.-P. Hölck* using silver perchlorate/carbonate mixtures as glycosylation promoters.[221] The 3,4,6-tri-*O*-acetyl-2-azido-2-deoxy-galactosyl bromide donor **8** used in a *Keonigs-Knorr* glycosylation,[233] was generated from D-galactose in four steps. Initially, the D-galactose was peracetylated with acetic anhydride and catalytic amounts of perchloric acid. Subsequential addition of phosphorus tribromide and water results in the α-anomeric galactosylbromide **5**. The product was then reduced with zinc dust activated by copper(II) sulfate to give the 3,4,6-tri-*O*-acetyl-galactal **6**.[234,235] The galactal was further treated with cerium(IV) ammonium nitrate and sodium azide in dry acetonitrile. A radical mechanism for the azidonitration was proposed by *R.U. Lemieux*.[236] The product 3,4,6-tri-*O*-acetyl-2-azido-2-deoxy-galactosyl nitrate **7** was obtained as an anomeric mixture of α/β (1:1). Next, compound **7** was treated with lithium bromide in dry acetonitrile to give the desired α-anomer of the halide donor 3,4,6-tri-*O*-acetyl-2-azido-2-deoxy-galactosyl bromide **8**.[236] Donor **8** was synthesized in 4 steps in an overall yield of 32 % (*figure 4.4*).

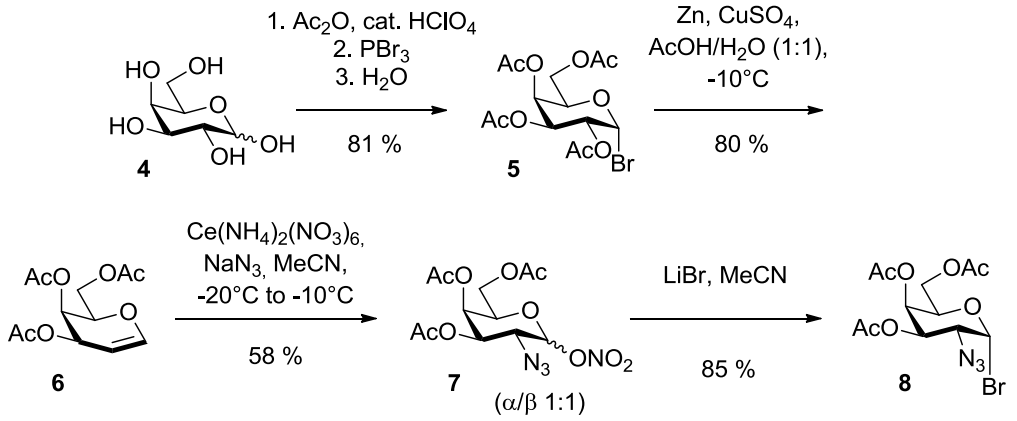


***Figure 4.4*:** Synthesis of 3,4,6-tri-*O*-acetyl-2-azido-2-deoxy-galactosyl bromide **8**.

For the synthesis of the amino acid glycosyl acceptor, L-threonine was protected with an *N*-(9-fluorenyl)-methoxycarbonyl-group (Fmoc) as a base labile protecting group during SPPS. Initially, L-threonine was reacted with *N*-(9-fluorenyl)-methoxycarbonyl-succinimidylcarbonate (Fmoc-OSu) under alkaline conditions.[237] The Fmoc amino acid **2** was then esterified with *tert*-butanol in a copper(I) catalyzed reaction with *N*,*N*'-dicylcohexylcarbodiimid (DCC) as coupling reagent to give glycosyl acceptor **3** (*figure 4.5*).[238,239,240]

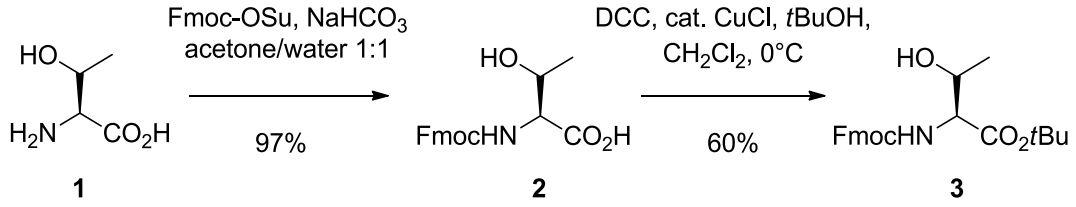


***Figure 4.5*:** Synthesis of threonine galactosyl acceptor **3**.

For the esterification, *tert*-butanol was initially activated with DCC catalyzed by copper(I). The *O-tert*-butyl-*N*,*N*'-isourea **I** was then treated with the amino acid to result in the ester product **III** and *N,N'*-dicyclohexyl urea **II** (*figure 4.6*).[238,239]

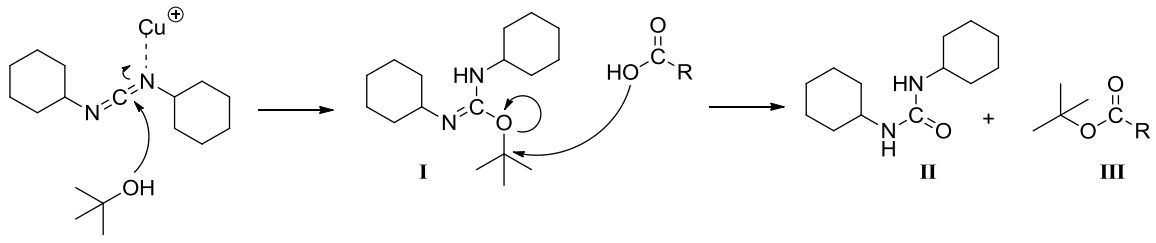


***Figure 4.6*:** Mechanism of DCC and copper(I) mediated esterification.

Finally, galactosylbromide **8** and amino acid acceptor **3** were glycosylated to form the acetyl protected T_N-antigen **9**. The *Koenigs-Knorr* reaction was performed with water-free silver perchlorate and silver carbonate as glycosylation promoters in an unpolar solvent mixture of

dry toluene/dichloromethane (1:1).[221,207,241] The azide group in 2-position of the galactosyl donor acts as a non-participating protecting group, promoting the α-anomer as the favored glycosylation product in a yield of 60% (*figure 4.7*).

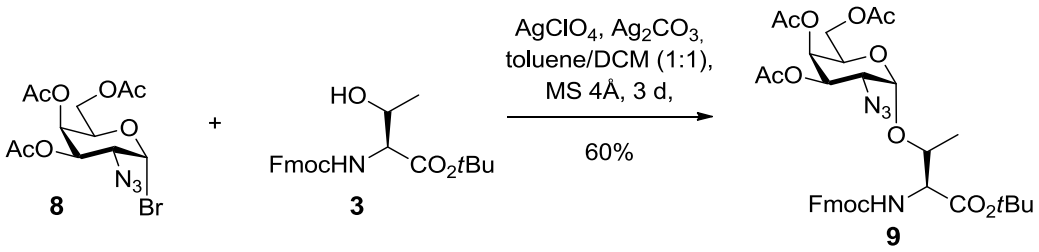


***Figure 4.7*: Synthesis of Fmoc-Thr(αAc$_3$GalN$_3$)-O*t*Bu **9**.**

Compound **9** was reduced to the acetamide using thioacetic acid (*figure 4.8*, **A**).[242] The reaction is base-promoted by pyridine and the azide is reduced to the acetamide *via* a thiatriazoline intermediate **I** with final loss of nitrogen and elemental sulfur (*figure 4.8*, **B**).[243]

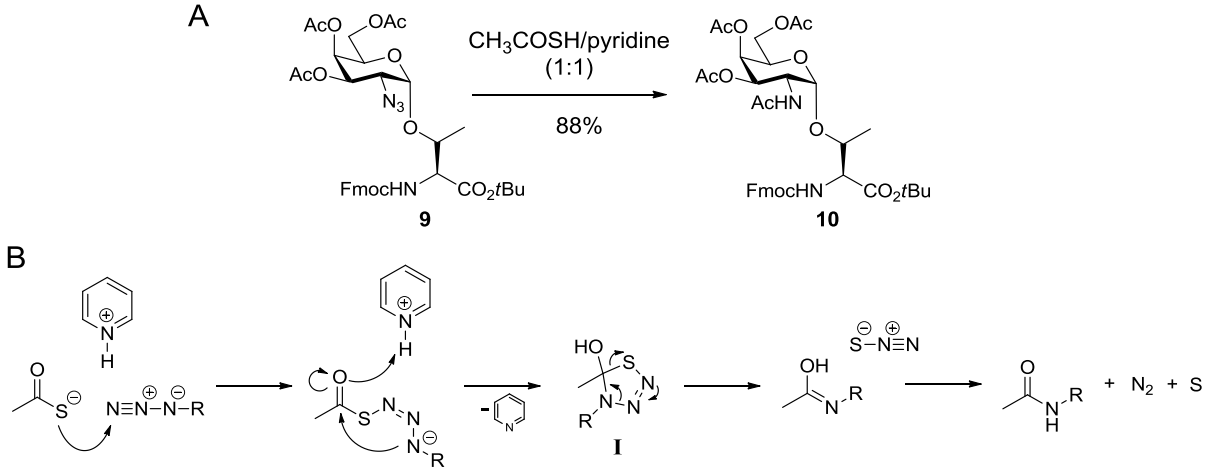


***Figure 4.8*: A:** Synthesis of **10** by reduction of **9**. **B:** Mechanism of the base promoted amidation.

The acetyl groups of compound **10** were removed by transesterification under *Zemplén* conditions with sodium methoxide in methanol.[211] These conditions are regarded as a mild deacetylation procedure and the pH is usually not raised above 8.5 in order to preserve the base labile Fmoc-group.[208] However, deacetylation was slow with reaction times of several hours to a few days and usually showed partial loss of Fmoc. It was found more convenient to transesterify at pH 10-10.5 in shorter times with direct subsequent re-protection of the α-amino group using Fmoc-OSu.[237] The amount of cleaved Fmoc was followed by thin layer chromatography (TLC) and an estimated amount of 0.5 eq of Fmoc-OSu was enough for complete re-protection. Further, β-elimination or epimerization by base-catalyzed removal of the H^{α} of threonine was not observed under these conditions.[180] Next, the 4-OH and 6-OH positions were protected with *para*-methoxybenzaldehyde dimethyl acetal in acetonitrile at

50°C and catalytic amounts of *para*-toluenesulfonic acid (*p*-TsOH) to give the corresponding acetal protected glycosyl acceptor **12**. The three modifications on the T_N-antigen from compound **9** gave **12** in 73% overall yield (*figure 4.9*).

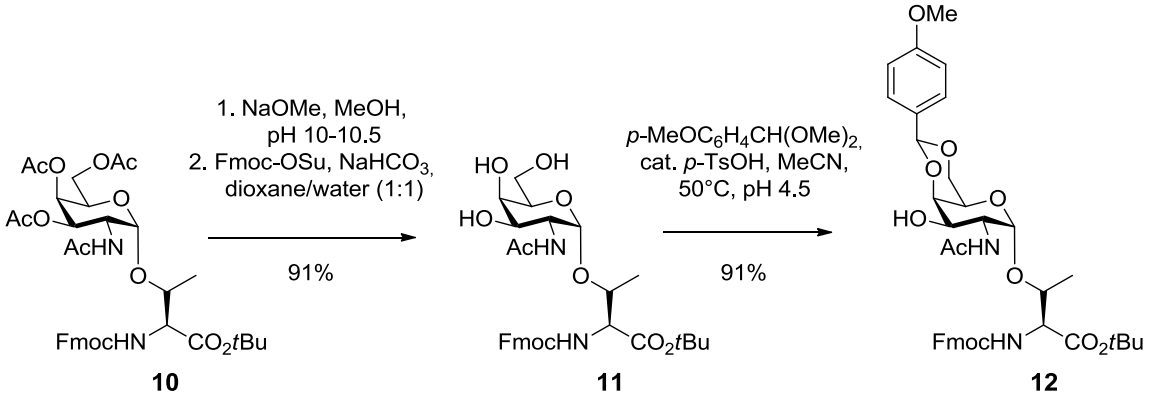


***Figure 4.9*:** Synthesis of T_N-antigen glycosyl acceptor **12**.

4.1.4 Synthesis of the T-antigen amino acid

The T- or T_F-antigen (Galβ1,3GalNAcαThr, Thomsen-Friedenreich antigen) was prepared from the suitable T_N-antigen **12** and the galactosyl bromide **5**, employing coupling conditions according to a *Koenigs-Knorr* variant developed by *Helferich* employing mercury(II) cyanide in dry nitromethane/dichloromethane (2:1).[244,245,220] As expected, only β-anomer of T-antigen **13** was obtained in 70% yield. The *para*-methoxybenzylidene group was hydrolyzed under acidic conditions and the liberated hydroxyl groups were acetylated with acetic anhydride to give compound **14**. Finally the *tert*-butyl ester was cleaved with trifluoroacetic acid and anisole as cation-scavenger, resulting in the peracetylated and Fmoc-protected T-antigen building block **15** ready for use in SPPS (*figure 4.10*).[220]

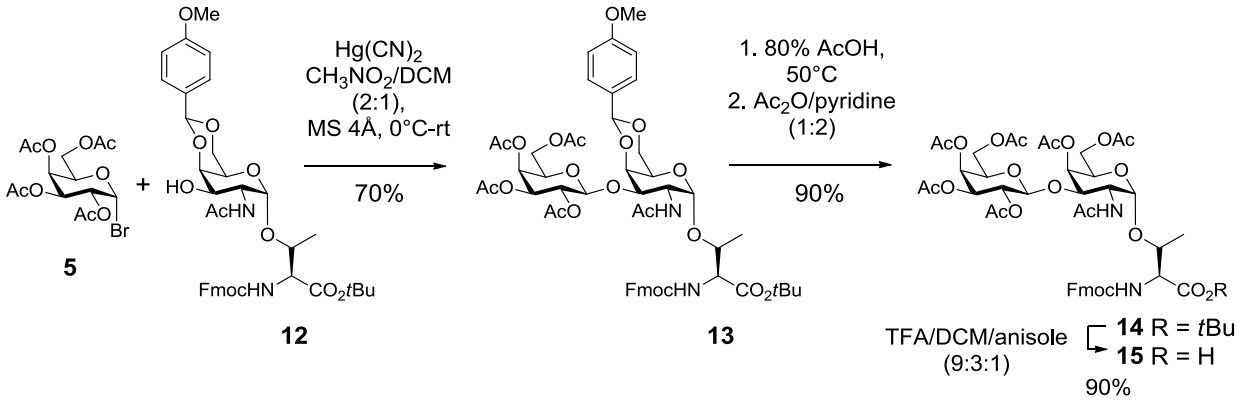


***Figure 4.10*:** Synthesis of T-antigen amino acid building block **15** for use in SPPS.

4.1.5 Synthesis of the type-1 *N*-acetyllactosamine glycosyl donor

The syntheses of the type-1 and type-2 disaccharide building blocks share the common trichloroacetimide galactosyl donor **18**. D-Galactose was peracetylated under the same conditions as mentioned for galactosyl bromide **8** generating peracetylated galactose **16**. The acetyl group on the anomeric position, forming an acetal, has an improved electrophilicity compared to the other acetyl groups and was selectively cleaved with hydrazine acetate in DMF. The resulting anomeric mixture of 2,3,4,6-tetra-*O*-acetyl-galactose **17** was then treated with trichloroacetonitrile and catalytic amounts of DBU (1,8-diazabicylo[5.4.0]undec-7-ene) to give the trichloroacetimidate donor **18**.[246] Donor **18** was synthesized in three steps in 66% yield (*figure 4.11*).

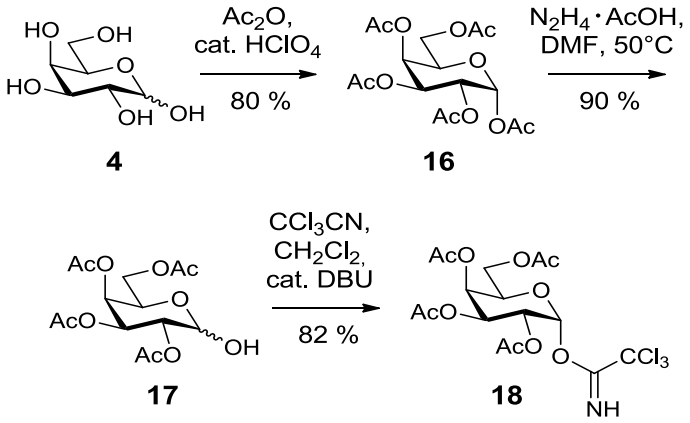


***Figure 4.11*:** Synthesis of galactosyl trichloroacetimidate donor **18**.

Galactosyl trichloroacetimidate **18** was obtained exclusively in α-configuration from the anomeric 2,3,4,6-tetra-*O*-acetyl-α/β-galactose **17**. In this base-catalyzed glycosylation, the intermediate anomeric oxyanion can either give α-anomer under thermodynamic control, stabilized by the anomeric effect, or β-anomeric product under kinetic control, due to the higher nucleophilicity of the β-oxyanion intermediate. Depending on the strength of the base that is used to deprotonate the anomeric OH-group, the kinetically formed β-anomeric product can react back, anomerize and finally form the α-anomer under thermodynamic reaction control. Bases like DBU or sodium hydride are strong enough to elicit retro-reaction, while potassium carbonate is commonly used to generate the β-anomer, since it is not able to deprotonate the β-imidate and initiate the back-reaction (*figure 4.12*).[247]

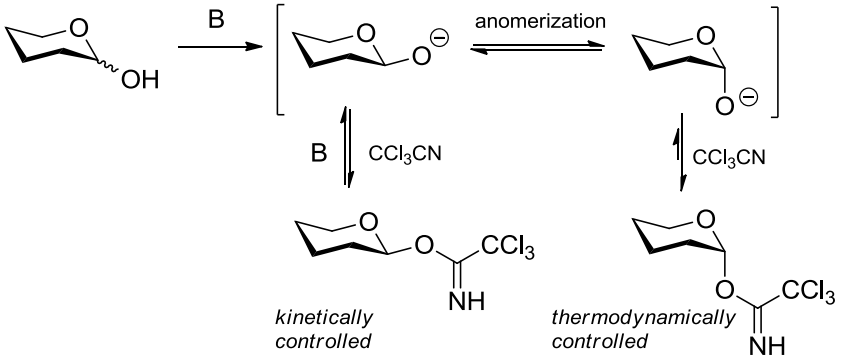


***Figure 4.12*:** Formation of α-imidates under kinetic control and formation β-imidates under thermodynamic control.

As acceptor for the synthesis of the type-1 disaccharide, the phenyl 2-azido-2-deoxy-4,6-benzylidene-1-thio-β-D-glucopyranoside **24** was synthesized. The 2-acetamido-group at the GlcNAc unit of the later disaccharide was initially masked as an azide during disaccharide formation. The azide group was introduced with the diazotransfer reagent imidazole-1-sulfonyl-azide hydrochloride **19**. Therefore, sodium azide was suspended in acetonitrile and sequentially treated with sulfurylchloride under ice-cooling followed by addition of imidazole. The imidazole-1-sulfonyl-azide was finally precipitated as the hydrochloride salt **19** with acetyl chloride in methanol (*figure 4.10*).[248] The hydrochloride **19** is regarded as less explosive and more shelf stable than other common diazotransfer reagents, e.g. trifluoromethanesulfonyl azide (TfN_3).[248] Anyhow, also compound **19** is impact sensitive and thermally labile, like other low-molecular weight azides, attached to electron withdrawing substituents. Concentration of mother liquids after filtration of the precipitated hydrochloride salt must be omitted, the hygroscopic salt must be stored dry to avoid water uptake (formation of hydrazoic acid) and the compound should always be added or used in solution for better heat dissipation.[249,250] The diazotransfer reagent **19** transferred a diazo group to the amine in 2-position of D-glucosamine in a copper(II) catalyzed reaction. Subsequentially, a peracetylation with acetic anhydride in pyridine was performed to give an anomeric mixture of 1,3,4,6-tetra-*O*-2-azido-2-deoxy-glucosylpyranoside **21** (*figure 4.13*).[248]

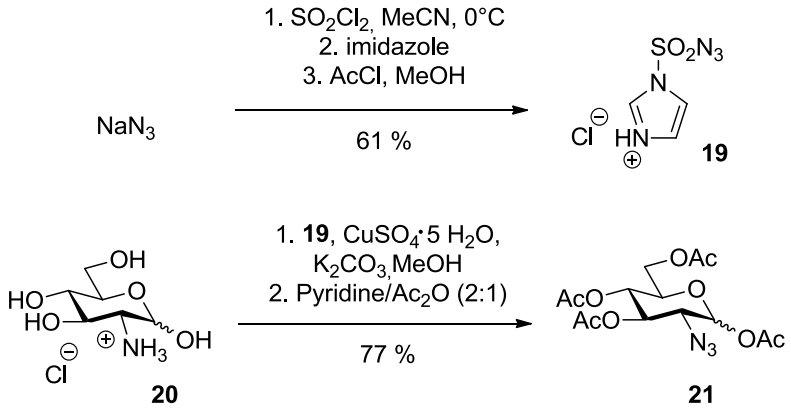


***Figure 4.13*:** Synthesis of 1,3,4,6-tetra-*O*-2-azido-2-deoxy-glucosylpyranoside **21** by diazotransfer with imidazole-1-sulfonyl-azide hydrochloride **19**.

Next, the thiophenyl group was installed on the anomeric position. Because of the non-participating azide on the 2-position, compound **21** was not transferred into the thiophenyl glucopyranoside directly, since the obtained thiophenyl donor would consist of a mixture of reactive β-anomer and less reactive, thermodynamically more stable α-anomer. Therefore, the anomeric acetyl group was first activated by the lewis acid titanium(IV) chloride with cleavage of the acetate and substitution by a chlorine atom, to result in the α-anomer of 3,4,6-tri-*O*-acetyl-2-azido-2-deoxy-glucopyranosyl chloride **22**.[251,252,236] Then, the chloride was substituted by a thiophenyl group in a *Walden inversion* to give the phenyl 2-azido-2-deoxy-1-thio-glucopyranoside **23**.[253] The 4- and 6-position were protected with an acid-labile benzylidene acetal to result in acceptor **24**.[254] The glycosyl acceptor **24** was synthesized in 4 steps with an overall yield of 25% (*figure 4.14*).

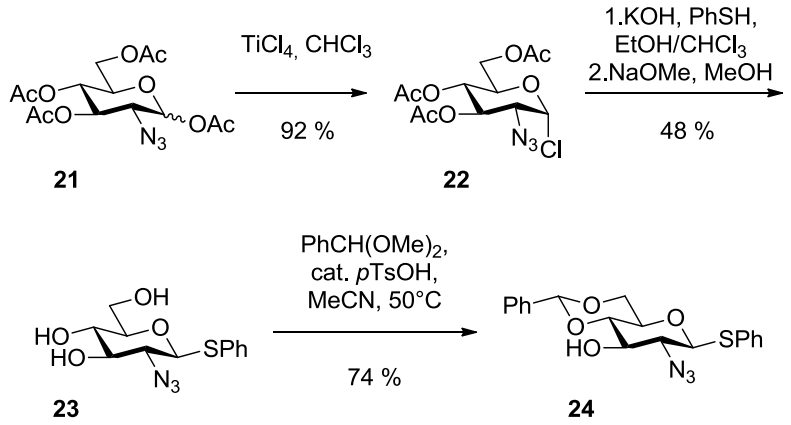


***Figure 4.14*:** Synthesis of glycosyl acceptor **24**.

The following glycosylation of acceptor **24** and trichloroacetimidate donor **18**, was performed according to a modified procedure originally described by *Toepfer* and *Schmidt*.[255] The glycosyl donor was activated by catalytic amounts of trimethylsilyl trifluoromethanesulfate

(TMSOTf) in diethyl ether at room temperature. The protected type-1 disaccharide **25** was obtained in good yields between 85-92% (*figure 4.15*).

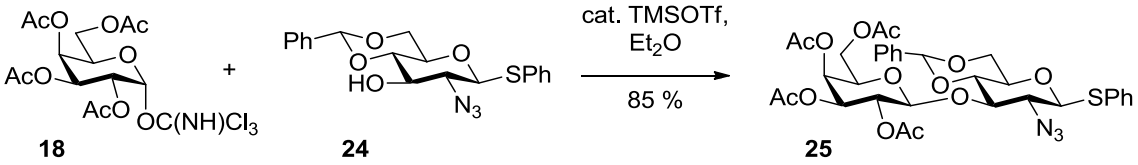
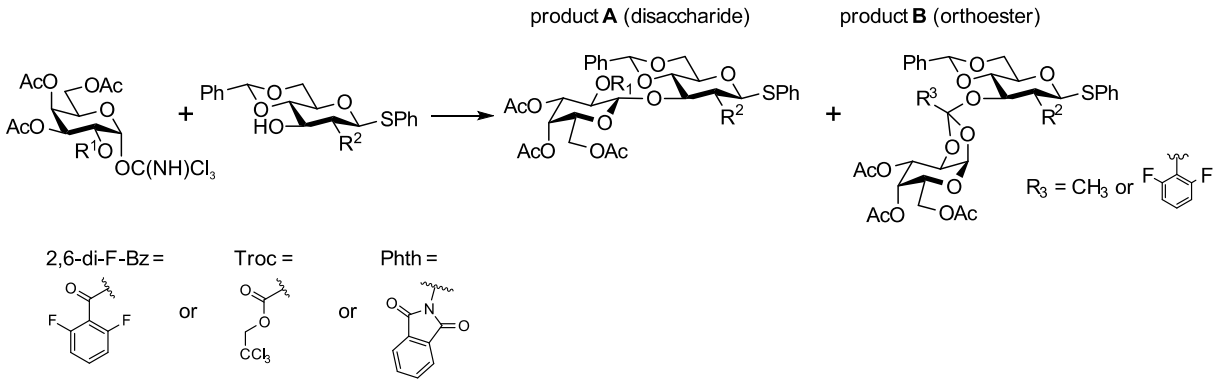


***Figure 4.15*:** Synthesis of type-1 disaccharide **25** by glycosylation of acceptor **24** with donor **18**.

Type-1 disaccharide **25**, which already contains a *trans*-configurated thiophenyl leaving group at the reducing end, will act as a donor in later glycosylation steps. To achieve desired β-stereoselectivity, the azide was transformed into a participating *N*-Troc protecting group. Implementation of the Troc directly on the monosaccharide stage, instead of masking the amine as an azide, would have been more convenient. Direct use of a *N*-Troc-containing acceptor with donor **18** under common conditions used for trichloroacetimidate couplings (e.g. DCM, T ≤ -40°C, molecular sieve) resulted in breakdown of the reactants (*table 4.1*, entry **a**). Exchange of *N*-Troc to a *N*-phthalimido group, aided glycosylation, but formed exclusively orthoester product (*table 4.1*, entry **b**). Common methods to repress orthoester formation, like the use of bulky and electron-withdrawing difluorobenzoyl esters in 2-position of the donor (*table 4.1*, entry **c**) or more acidic conditions created by acid washed molecular sieves (AW MS) (*table 4.1*, entry **d**), helped to reduce orthoester formation in favor for disaccharide product.[256,257] Reduction of the amount of molecular sieve eliminated orthoester formation in a test reaction (*table 4.1*, entry **e**), indicating beneficial effects from low or completely absent molecular sieve. However, keeping lower amount of AW MS again with a 2-*O*-acetyl group on the donor instead of the difluorobenzoyl ester, increased orthoester generation (*table 4.1*, entry **f**). If the 2-*O*-acetyl group remained on the donor and the phthalimido group was exchanged with an azide group on position-2 of the acceptor, complete disaccharide production was observed (although in low yield, *table 4.1*, entry **g**). The scale of the reaction also influenced the amount of orthoester formation. Small scale reactions (≤ 50 mg acceptor) gave the desired glycoside as a major product, while larger reactions (≥ 100 mg acceptor) and especially gram scale reactions resulted in higher amounts of orthoester (compare entry **g**/**h**). Great influence was observed by the solvent. Replacing DCM by diethyl ether, without MS, finally provided pure disaccharide glycosylation product, also in large scale glycosylations (*table 4.1*, entry **i**). A reaction was performed in order to control if complete disaccharide formation could be achieved in a diethyl ether system, without any MS by keeping the *N*-Troc on the acceptor, as initially planned (*table 4.1*, entry **j**), but similar to reaction conditions in entry **a**, no product formation was observed.

***Table 4.1*:** Comparison of various reaction conditions and protecting group patterns for the synthesis of a type-1 disaccharide.



entry	R_1	R_2	scale in g_{donor}	conditions	molecular sieve	yield in %	product ratio A:B
a	Ac	*N*-Troc	0.05	DCM, -40°C	100 mg MS 4Å	breakdown	n.a.
b	Ac	*N*-Phth	0.05	DCM, -40°C	100 mg MS 4Å	70	0 : 1
c	2,6-di-F-Bz	*N*-Phth	0.05	DCM, -40°C	100 mg MS 4Å	72	1 : 3
d	2,6-di-F-Bz	*N*-Phth	0.05	DCM, -40°C	100 mg AW MS 4Å	52	1 : 1
e	2,6-di-F-Bz	*N*-Phth	0.05	DCM, -40°C	30 mg AW MS 4Å	44	1 : 0
f	Ac	*N*-Phth	0.05	DCM, -40°C	30 mg AW MS 4Å	53	1 : 1.7
g	Ac	N_3	0.02	DCM, -40°C	25 mg AW MS 4Å	25	1 : 0
h	Ac	N_3	1	DCM, -40°C	1 g AW MS 4Å	92	1 : 3
i	Ac	N_3	10	Et_2O, RT	-	85	1 : 0
j	Ac	*N*-Troc	0.05	Et_2O, RT	-	breakdown	n.a.

The methodology reported by *Toepfer* and *Schmidt*, using an azide acceptor and diethyl ether as solvent was successful, also in gram scale reactions.[255] It is interesting, that especially a reaction in diethyl ether represses orthoester formation, since orthoester formation usually benefits from diethyl ether as a solvent.[258] Sometimes orthoesters can rearrange into the corresponding glycosidic bond by treatment with higher amounts of lewis acids like TMSOTf.[259] Here, these attempts lead to breakdown of the starting materials. The synthesis of type-1 LacNAc disaccharide, was dependent on subtle interactions of various reaction parameters, like protecting groups, solvents, batch scale or molecular sieve can influence the product formation in chemical glycosylations.

After the 1,3-glycosylation was carried out, the azide-group was first reduced and subsequently protected with a *N*-Troc group. Disaccharide **25** was treated with activated zinc powder (previously suspensed in 1 M HCl and washed with water, methanol and diethyl

ether) in a mixture of 1,4-dioxane/acetic acid (10:1). Under these conditions the acidity was high enough to promote reduction of the azide with the zinc, while the acid labile benzylidene acetal remained stable. Then TrocCl was used to protect the free amine to give disaccharide **26** in 72% yield (*figure 4.16*).

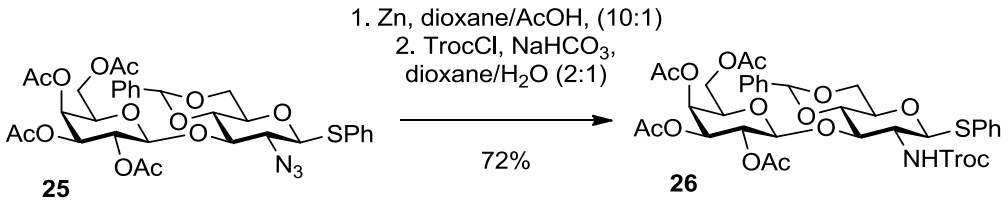


***Figure 4.16*:** Synthesis of compound **26**. Change from azide to *N*-Troc protection.

Considerations regarding the reactivity of the thiophenyl disaccharide **26** as glycosyl donor were made. The 4,6-*O*-benzylidene acetal group is well known to lower the reactivity of glycosyl donors. This effect was explained by a conformational lock of the pyranose in the $^4C_1$ chair configuration because of the rigidity introduced by the fused 1,3-dioxane ring. As a consequence, the formation of an intermediate oxocarbenium ion with planar sp^2-geometry at C1, after separation of the leaving group, is hampered during glycosylation, due to non-optimal orbital overlap of the π-orbital forming the double bond.[196] Additionally to this torsional effect, reactivity is also reduced by an electronic effect.[260] The C6-O6 bond (and the negative dipole it represents) can adopt three staggered conformations, the *gg*, *gt* and *tg* conformation. In *gg* and *gt*, the dipole is oriented more perpendicular to the positive oxocarbenium transition state, for better dipole-charge stabilization, while in *tg* the negative dipole is directed away from the positive center (*figure 4.17*).

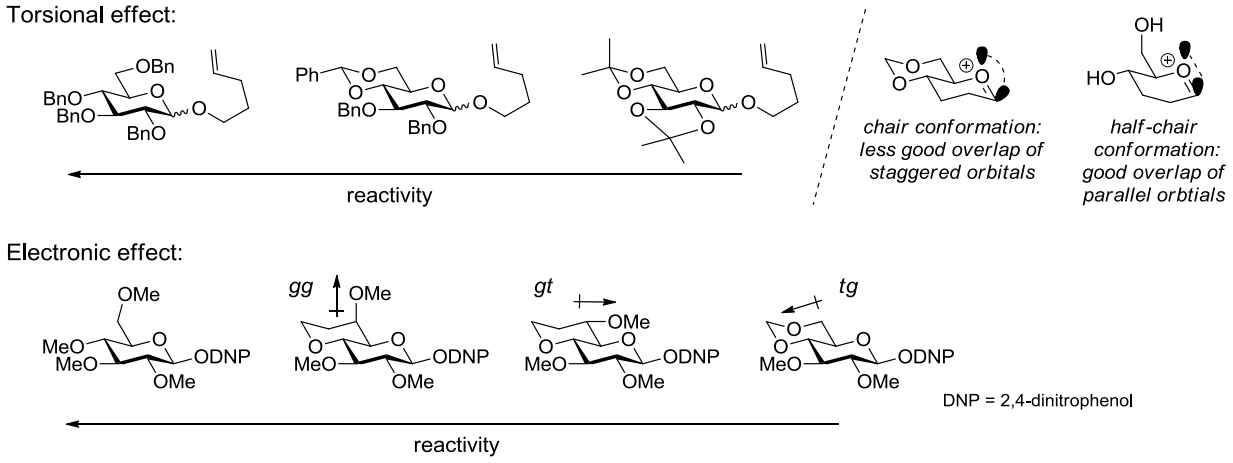


***Figure 4.17*:** Torsional[196] and electronic [260] disarming properties of 4,6-*O*-benzylidene acetal protecting groups.

Additionally to reactivity issues, retrosynthetic planning involved disaccharide extension of the T-antigen acceptor amino acid, forming a common core 1 tetrasaccharide amino acid, which contains an acid labile *para*-methoxy benzylidene acetal at the 6-position of the αGalNAc. For further extension of the core 1 tetrasaccharide, in order to form core 2 structures, a step involving selective GalNAc benzylidene acetal removal was required before the core 2 structure branching could be established (retrosynthesis, *figure 4.2*). Therefore, the 4,6-*O*-benzylidene acetal at disaccharide **26** was exchanged to a 4-*O*-acetyl and 6-*O*-*tert*-butyldimethylsilyl protecting (TBS) group.[261] The TBS-group has the advantage that it can be selectively cleaved by a fluoride source in the presence of other acid labile protecting groups, such as the benzylidene acetal on the αGalNAc of the glycosylated amino acids. The benzylidene acetal of compound **26** was cleaved in 80% acetic acid, followed by regioselective protection at 6-position with TBSCl and acetylation with acetic anhydride to give the final type-1 disaccharide donor **27** (*figure 4.18*).

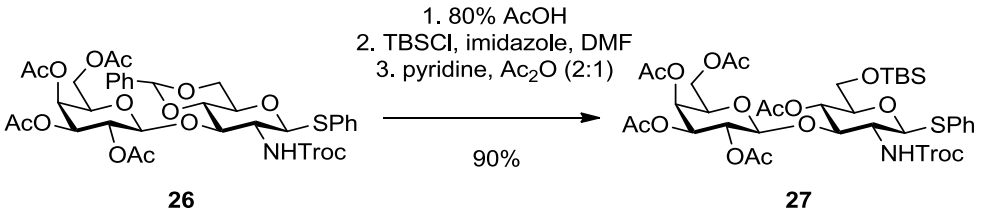


***Figure 4.18*:** Change from a 4,6-*O*-benzylidene acetal to *tert*-butyldimethylsilyl ether and acetyl protecting groups for the synthesis of the type-1 thiophenyl donor **27**.

4.1.6 Synthesis of the type-2 *N*-acetyllactosamine glycosyl donor

The type-2 disaccharide building block was formed from the described galactosylpyranosyl trichloroacetimidate donor **18** and glucosaminylpyranosyl thiophenol acceptor **30**. Similar to the type-1 building block synthesis, the acceptor **30** contained a C1-thiophenyl leaving group for later elongation towards the desired extended core amino acids. To obtain acceptor **30**, glucosamine hydrochloride was directly protected with a Troc group at position 2 and all further hydroxyl groups were acetylated to give compound **28** as an anomeric mixture (α/β = 8.3 : 1). Next, the anomeric acetyl leaving group of compound **28** was activated with boron trifluoride and substituted with a thiophenyl group. Since compound **28** had the anomeric acetyl group predominantly in the thermodynamically stable α-anomeric configuration, the reaction was stirred for 3 d with excess of three equivalents of thiophenol for full conversion into the thioglycoside **29**.[262] In order to liberate the acetylated 4-position for glycosylation with donor **18**, compound **29** was completely deacetylated with methanolic HCl solution, prepared from acetyl chloride in methanol. Standard basic alcoholysis for the removal of the acetyl

groups, converts the *N*-Troc group into the corresponding alkylcarbamate, which cannot be cleaved under reductive conditions commonly applied for *N*-Troc deprotection.[214] As an alternative, a guanidine nitrate/sodium methoxide system was reported to deacetylate under basic conditions, without harming the *N*-Troc group.[263] Subsequently the primary alcohol at the 6-position was protected with a TBS ether group to give glycosyl acceptor **30** in 3 steps and 60% yield (*figure 4.19*).

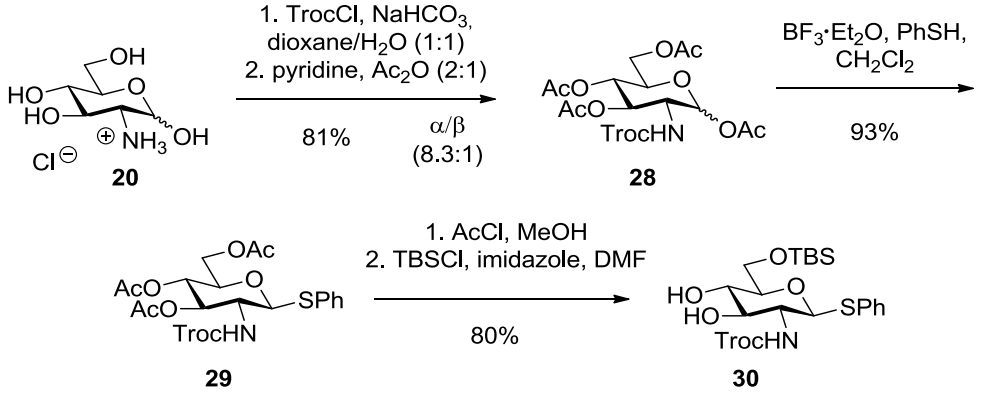


***Figure 4.19*:** Synthesis of type-2 glycosyl acceptor **30**.

Glycosyl acceptor **30** was regioselectively glycosylated in the more reactive[214,219] 4-position with trichloroacetimidate **18** and catalytic amounts of TMSOTf and subsequently acetylated in 3-position in 80% overall yield (*figure 4.20*). Problems with orthoester formation were not observed.

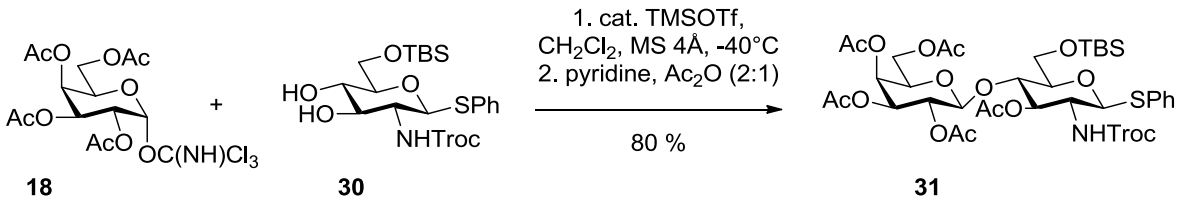


***Figure 4.20*:** Synthesis of type-2 thiophenyl donor **31**.

4.1.7 Synthesis of the galactose extended core 3 amino acids

As outlined in the retrosynthetic analysis (*figure 4.2*), the synthesis of the core 3 amino acids was conducted by the glycosylation of the T_N-antigen acceptor amino acid **12** with either the type-1 or type-2 disaccharide donor. For the synthesis of the type-1 core 3 building block, thioglycoside donor **31** was activated by the promoter-catalyst system *N*-iodosuccinimide (NIS) and trifluoromethanesulfonic acid (TfOH) (*figure 4.21*).[184]

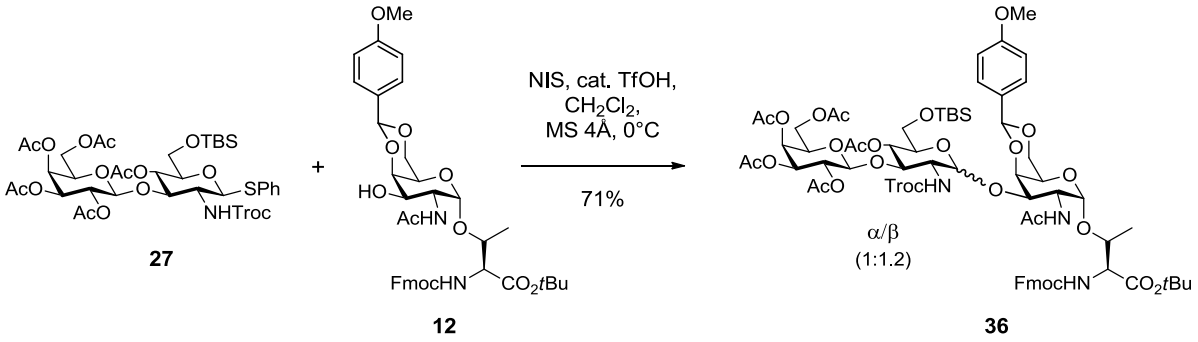


***Figure 4.21*:** Preparation of type-1 core 3 glycosylated amino acid **36**.

Although the *N*-Troc group is a participating neighboring group, directing glycosylations towards the β-anomer, a mixture of α/β-anomers (α/β = 1:1.2) of compound **36** was obtained in 71% yield. The thiophilic iodonium ion liberated from NIS at acidic conditions, was attacked by the phenyl sulfide to leave as phenylsulfenyl iodide. Since most of the product formed was the desired β-anomer, it can be assumed that the reaction proceeded by a S_N1-mechanism. The neighboring *N*-Troc group participated in the reaction, since more α-product instead of β-product would be expected due to the α-directing anomeric effect. However, *N*-Troc participation was likely somehow hampered, since some donor was also attacking from the *cis* side, instead from the *trans* side of the oxazolinium intermediate (*figure 4.22*).

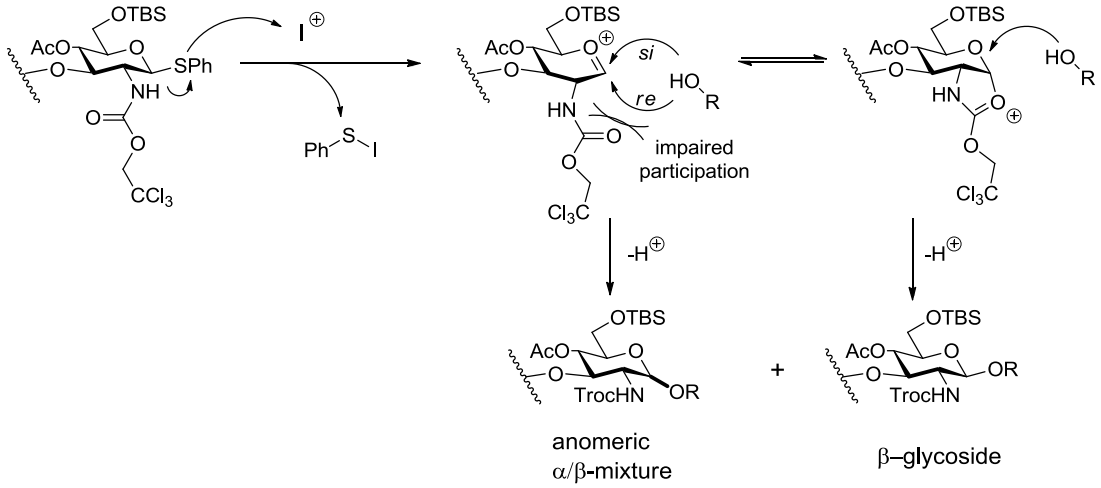


***Figure 4.22*:** Formation of an anomeric mixture of compound **36**.

Separation of the two diastereomeric forms of compound **36** was not possible by means of standard silica gel chromatography. The diastereomeric mixture was proceeded to the next step and treated with 80% acetic acid to remove the acid labile *para*-methoxybenzylidene acetal and the *tert*-butyldimethylsilyl ether to give compound **37** in 50% yield of the desired β-anomer and 29% of the α-anomer. The anomeric products had sufficient distinct retention

times to be separated by silica gel chromatography. Next, the *N*-Troc group was cleaved under reductive conditions with zinc in acetic acid and the released amine was acetylated along with the remaining free hydroxyl groups to give compound **38** in 85% yield.[264] In the final step, the *tert*-butyl group was removed in a mixture of TFA/DCM and anisole as a cation-scavenger. The type-1 core 3 building block was synthesized over 4 steps and in 26% yield starting from the T_N-antigen acceptor **12** (*figure 4.23*).

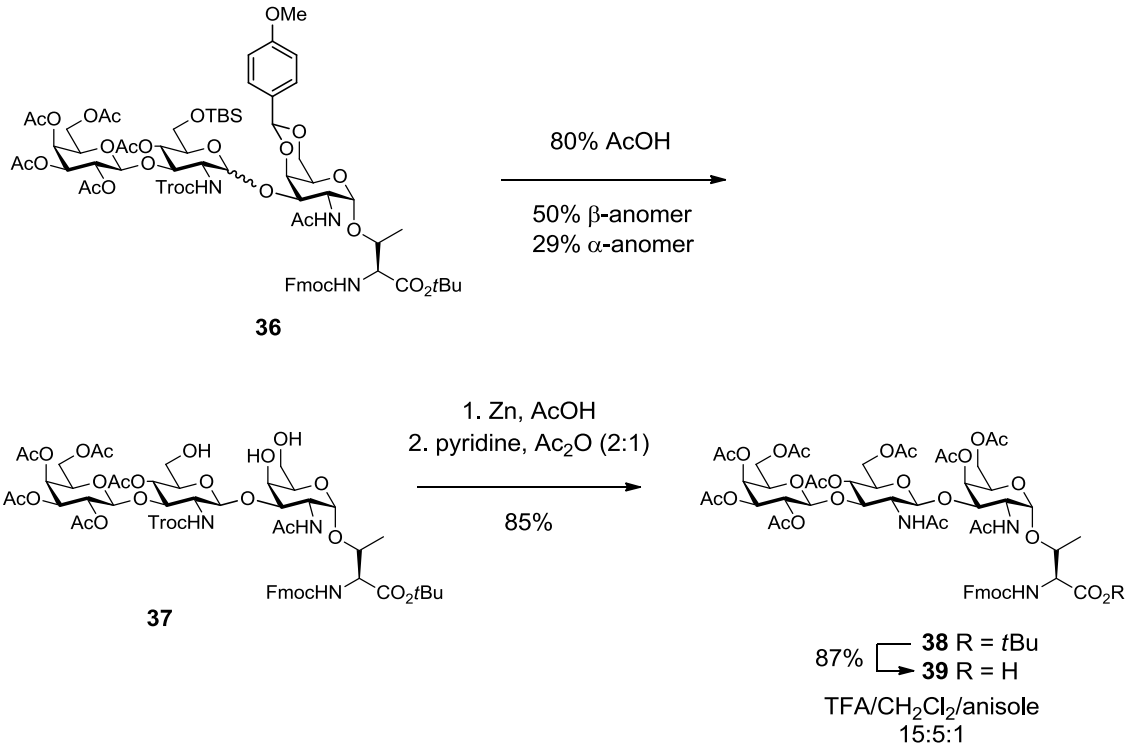


***Figure 4.23*:** Synthesis of type-1 core 3 amino acid building block **39** for SPPS.

The synthesis of the type-2 core 3 amino acid followed the same strategy as the corresponding type-1 extended amino acid. T_N-antigen glycosyl acceptor **12** was glycosylated with the type-2 LacNAc thioglycoside donor **31** with NIS/TfOH as promoter system. The glycosylation product **32** was generated with full β-stereoselectivity as expected and in contrast to type-1 core 3 glycosylation product **37**. However, a significant amount of the obtained product was an adduct consisting of the glycosylation product and the leaving phenyl sulfide group, as indicated by a mass increase of 108 u. NMR-spectroscopy showed proton-signal doubling and two-dimensional NMR-experiments showed two independent spin-systems. This separation indicated on one compounds with axial chirality due to restricted rotation around a single bond. Therefore it was assumed that the phenyl sulfide was added to the carbamate nitrogen of the *N*-Troc group. Different promoter systems did not improve reaction outcome. Dimethyl(methylthio)sulfonium trifluoromethanesulfate (DMTST)[186] did not activate the reaction at all, while *N*-bromosuccinimide (NBS)/TfOH[185] or

diphenyl sulfoxide (DPS)/Tf_2O[265] showed the same results as the NIS/TfOH system. Thus, NIS was used and the phenyl sulfide adduct was obtained as the main product in 70% yield (*figure 4.24*, **A**).

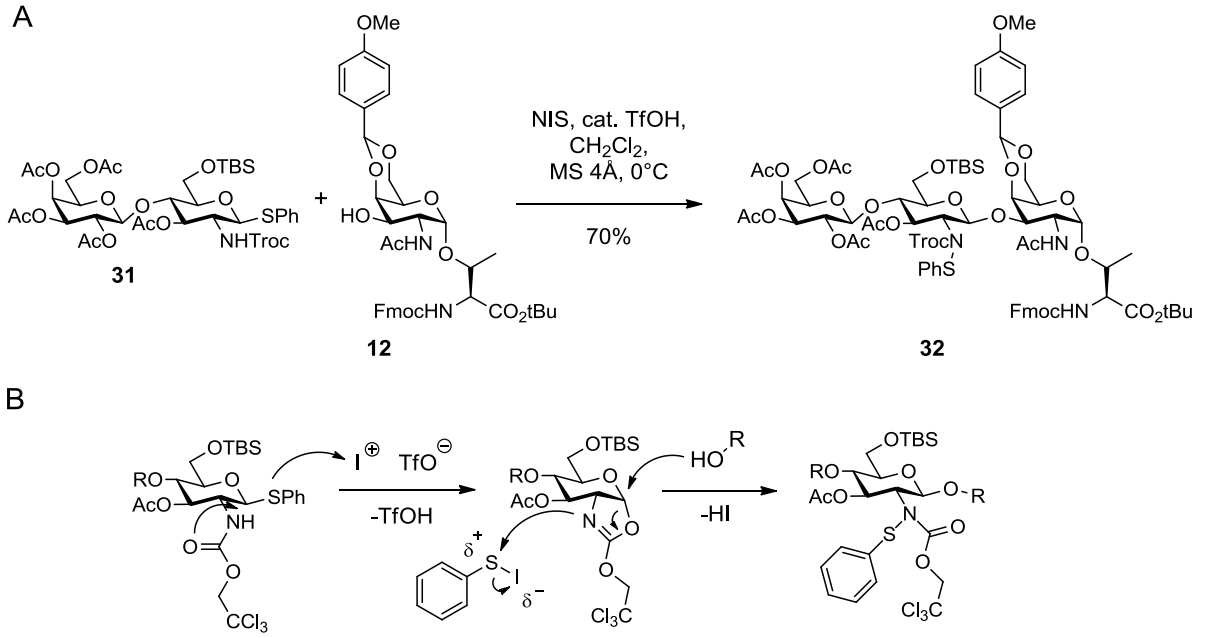


***Figure 4.24*: A:** Glycosylation of donor **31** and acceptor **12** for preparation of type-1 core 3 glycosylated amino acid **32**. **B:** Proposed formation of phenyl sulfide adduct **32**.

The glycosylation worked with full β-stereoselectivity, implying full neighboring group participation of the *N*-Troc group. The oxazoline intermediate may be attacked by the donor at C1 and probably along with oxazoline ring opening, the phenyl sulfide group added by substitution on sulfur (*figure 4.23*, **B**). According to *Pauling* electronegativity, sulfur is slightly positive polarized by the more electronegative iodine. Also, an orbital stabilization from the aromatic π-system to the transition state of a S_N2-reaction in benzylic position may promote the substitution. Since the formation of the type-1 core 3 analogue **36** showed unexpected stereoselectivity and later glycosylations with the type-1 and type-2 donors on extended core 1 and 2 glycans were not forming the phenyl sulfide byproducts, it was assumed that the sterical environment at the T_N-antigen acceptor had a strong influence on the reaction outcome. *T. Reipen* and *H. Kunz* have reported unexpected orthoester formation on an equally protected serine analogue of the T_N-antigen, when peracetylated galactosylpyranosides with halide or trichloroactimidate were evaluated as donors.[223]

Compound **32** followed protecting group transformations according to the corresponding type-1 core 3 analog **36**. The acid labile *para*-methoxybenzylidene acetal and the *tert*-butyldimethylsilyl ether groups were hydrolyzed in 80% acetic acid to give compound **33**. The *N*-Troc group was removed under reductive conditions with zinc in acetic acid with

simultaneous removal of the phenyl sulfide group. The remaining amine and the hydroxyl groups were further acetylated in pyridine and acetic anhydride (2:1). The *tert*-butyl group was removed in TFA/DCM with anisole as cation-scavenger. The type-2 core 3 amino acid building block was obtained in 4 steps and 41% yield overall, starting from the T_N-antigen glycosyl acceptor **12** (*figure 4.25*).

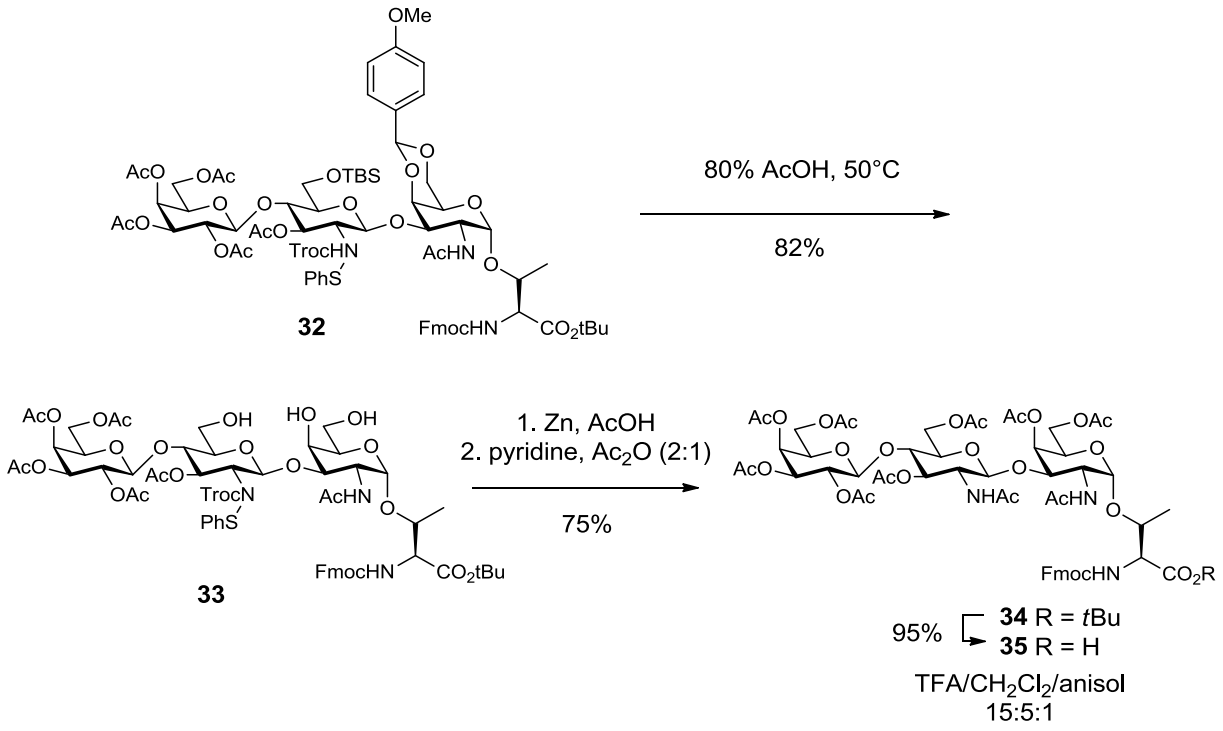


***Figure 4.25*:** Synthesis of type-1 core 3 amino acid building block **35** for SPPS.

4.1.8 Synthesis of the extended core 1 tetrasaccharide amino acids

For the synthesis of the linear core 1 tetrasaccharides and branched core 2 hexasaccharides a common core 1 acceptor (a T-antigen acceptor) building block was constructed (retrosynthesis, *figure 4.2*, glycosylation **IV**). In contrast to the previously synthesized T-antigen building block **13** (*figure 4.10*), a different monosaccharide galactosyl donor was applied. The galactose donor was protected temporarily in 6-position with a protecting group, stable under the basic conditions which are used for deacetylation. Like this, the reactive primary alcohol was blocked, which later allowed regioselective modifications on the remaining secondary alcohol groups. A TBS-group was selected, which later was cleaved together with the TBS-group of the elongating LacNAc disaccharide units by using a fluoride source, orthogonal to the *para*-methoxy benzylidene acetal of the initial αGalNAc unit. Peracetylated galactose **16** was first transformed into the β-thiophenyl-galacosylpyranoside

40 using boron trifluoride as a lewis acid promoter and thiophenol as glycosyl acceptor. Due to the α-anomeric configuration of **16,** excess of thiophenol (3 eq) was used and long reaction times were required for full conversion. The acetyl groups of the obtained thioglycoside **40** were then removed by transesterification with sodium methoxide, the 6-position was selectively protected with a TBS group and the remaining hydroxyl groups were again acetylated. The thiophenyl glycoside **42**[266] was prepared in 3 steps and 57% yield starting from 1,2,3,4,6-penta-*O*-acetyl-α-galactosylpyranoside **16** (*figure 4.26*).

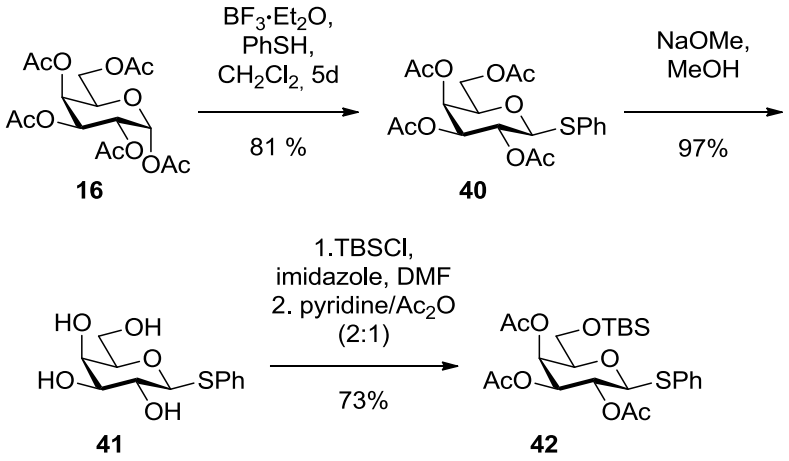


***Figure 4.26*:** Synthesis of glycosyl donor **42**.

The thiophenyl glycosyl donor **42** was glycosylated with the T_N-antigen acceptor **12**, promoted by a NIS/TfOH system. The glycosylation product **43** was obtained in 74% yield and with full β-stereoselectivity. No problems with stereoselectivity or orthoester formation were monitored in this reaction (compared with type-1 core 3 synthesis, *chapter 4.1.5*). The obtained compound **43** was deacetylated with sodium methoxide in methanol. The pH was set to 9.5 to efficiently remove the acetyl esters. A reaction to reinsert the partially removed Fmoc group followed. The core 1 glycosyl acceptor **44** was synthesized in 72% yield and displays unprotected hydroxyl groups 2, 3 and 4 on the galactose unit (*figure 4.27*).

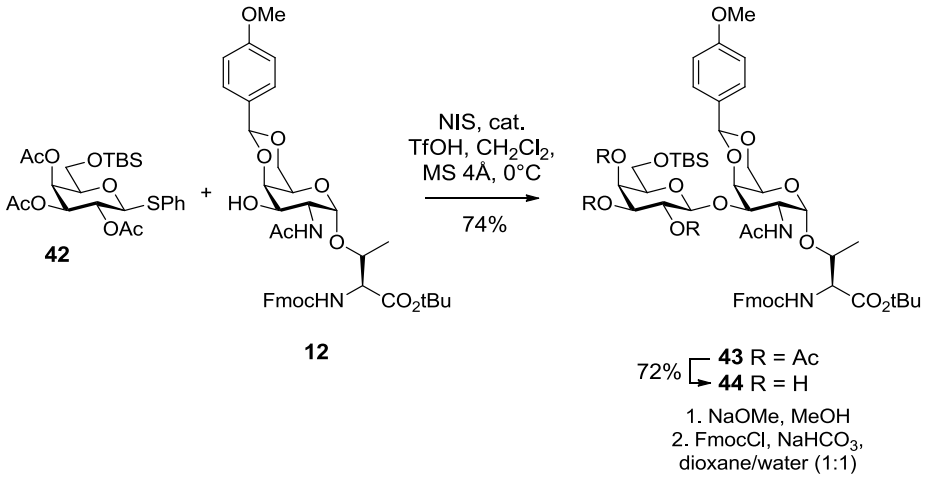


***Figure 4.27*:** Preparation of T-antigen **44** by glycosylation of donor **42** and acceptor **12**.

Next, T-antigen acceptor **44** was glycosylated with either the type-1 or type-2 LacNAc disaccharide in a [2+2]-block synthesis (the reaction pathways for both the disaccharides were analog from this stage, thus, presentation of both pathways will be presented simultaneously). The thioglycoside donors were activated by the approved conditions with NIS/TfOH in DCM. It is well known that equatorial hydroxyl groups tend to react preferentially in glycosylations, if a vicinal axial hydroxyl group is present. Therefore the hydroxyl group at the galactose position 3 of acceptor **44** was glycosylated regioselectively in the presence of the unprotected hydroxyl groups on position 2 and 4.[222] Regioselectivity was proven by HMBQ-NMR spectrometry, which showed spin coupling between H1(donor)-C3'(acceptor) and H3'(acceptor)-C1(donor). Tetrasaccharide core 1 structures were obtained in 64% (**45**) and 65% (**50**) yields, respectively (*figure 4.28*).

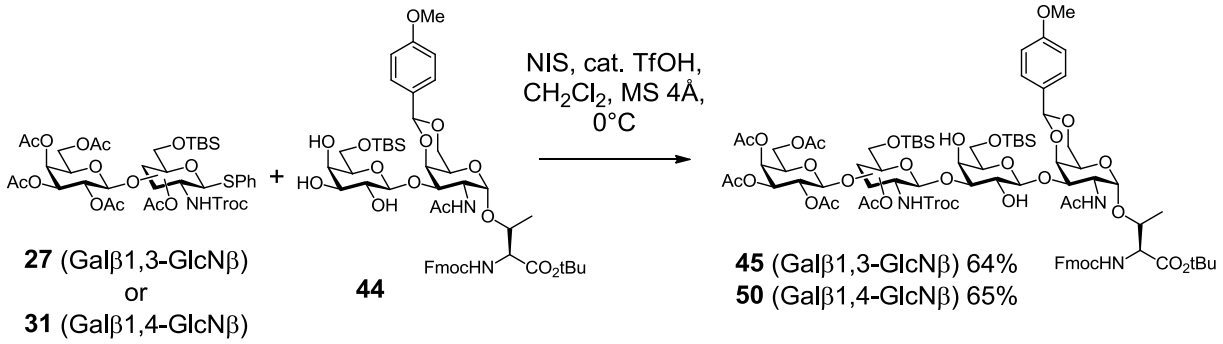


***Figure 4.28*:** Glycosylations of glycosyl donors **27** and **32** with acceptor **44**.

The two TBS protecting groups of glycosylation products **45** and **50** were cleaved with tetra-n-butylammonium fluoride (TBAF) trihydrate. The reasonable basicity of the fluoride anion in aprotic solvents demands addition of acetic acid to buffer the reaction and prevent unwanted cleavage of the Fmoc group.[267,268] Subsequent, acetylation of the hydroxyl groups resulted in

compounds **46** and **51**. Treatment with acetic anhydride in pyridine of the complex glycans required addition of catalytic amounts of 4-(dimethylamino)pyridine (DMAP, *Steglich Base*) for complete acetylation.[269] Compounds **46** and **51** were treated with 80% acetic acid to hydrolyze the acid labile acetal, revealing 4- and 6-OH of the prime αGalNAc. The resulting compounds **47** and **52** in this stage represented branching points for either further protecting group manipulations to complete type-1 and -2 elongated core 1 glycosyl amino acids for SPPS or for selective glycosylations on 6-OH to give branched core 2 hexasaccharides (*figure 4.29*).

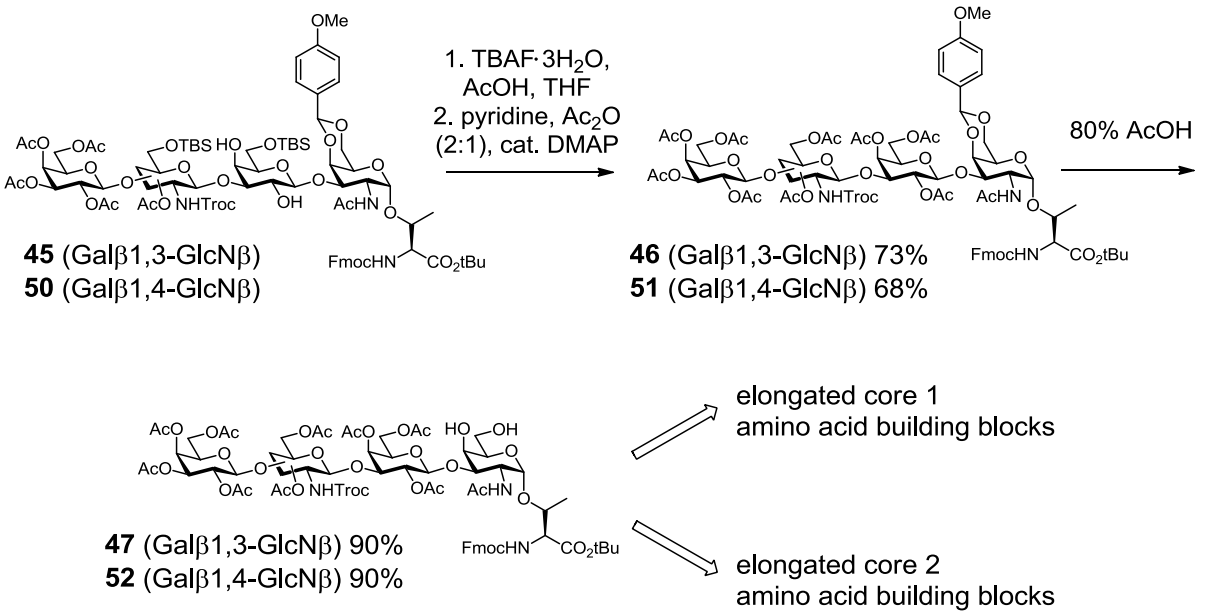


***Figure 4.29*:** Synthesis of the extended core 1 tetrasaccharides **47** and **52**, as branching compounds in the synthesis routes of extended core 1 and core 2 structures.

To continue the synthesis route of the linear elongated core 1 building blocks, the *N*-Troc groups on compounds **47** and **52** were removed by reduction using activated zinc in acetic acid and the obtained free amines and the unprotected hydroxyl groups were acetylated. In the final step, the *tert*-butyl esters on **48** and **53** were cleaved in TFA/DCM with anisole as a scavenger to give the type-1 and -2 elongated core 1 amino acids **49** and **54** ready for SPPS. Compounds **49** and **54** were both synthesized over 5 steps in 32% (type-1, **49**) and 29% (type-2, **54**) overall yield starting from the glycosylation with the common T-antigen acceptor **12** (*figure 4.30*).

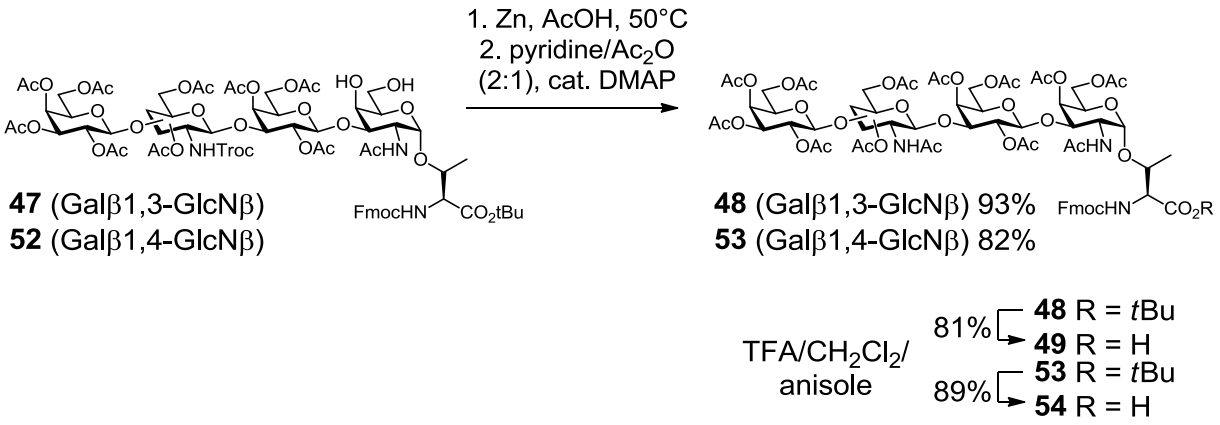


Figure 4.30: Syntheses of extended core 1 amino acid building blocks **49** and **54** for SPPS.

4.1.9 Synthesis of the extended core 2 hexasaccharide amino acids

The branched core 2 building blocks were constructed from the common tetrasaccharide building blocks **47** for type-1 and **52** for type-2 and were assembled in a [2+4] building block glycosylation (retrosynthesis, *figure 4.2*). In this work the two homogeneous variants core 2 (type-1)$_2$ and core 2 (type-2)$_2$ were synthesized. For further biological evaluation on microarrays, it was considered more interesting to see effects of glycan size and core 2 branching of homogenous glycans, for comparing the two different elongation variants. The type-1 tetrasaccharide acceptor **47** was reacted with type-1 disaccharide donor **27** and type-2 tetrasaccharide acceptor **52** with type-2 disaccharide donor **31** in a NIS/TfOH promoted glycosylation (*figure 4.31*).

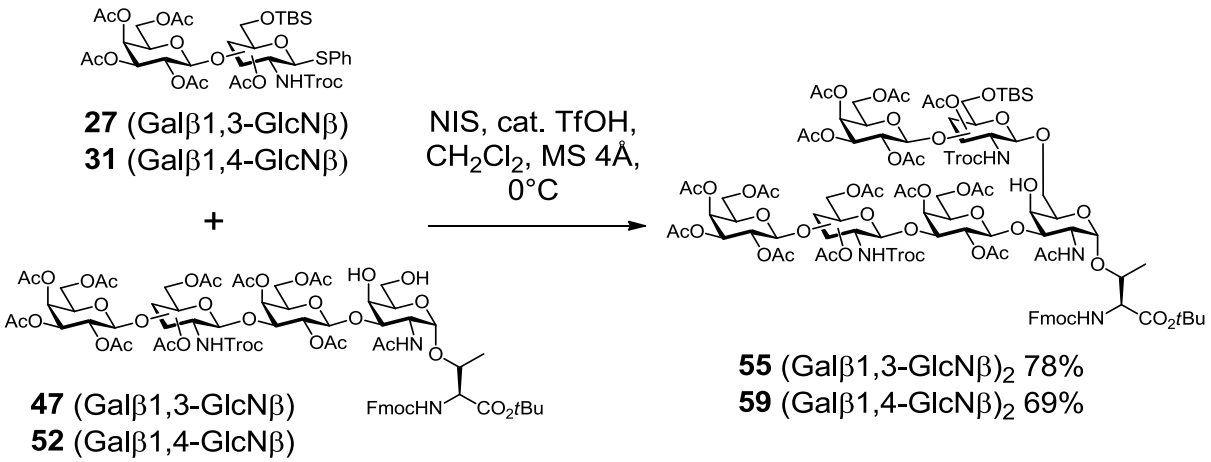


Figure 4.31: Glycosylation of donors **27** and acceptor **47** and glycosylation of donor **31** and acceptor **52**.

Both reactions resulted in the corresponding core 2 hexasaccharides **55** ((Galβ1,3-GlcNβ)$_2$) and **59** ((Galβ1,4-GlcNβ)$_2$), in 78% and 69% yield, respectively. In both cases, full regioselectivity on position 6 of the acceptor amino acid, verified by HMBC-NMR spectroscopy, and full β-stereoselectivity was obtained. Next, the TBS groups were hydrolyzed in 80% acetic acid and acetylation of the free hydroxyl groups followed, resulting in compounds **56** and **60**. The subsequent acetylation reactions after each protecting group manipulation step had the purpose to neutralize effects of acetyl migration, commonly observed with carbohydrates partly acetylated and partly unprotected.[270] Here, intra- or intermolecular acetyl migrations lead to ambiguities during reaction monitoring by thin layer chromatography and purification by silica chromatography due to differently acetylated species of the same scaffold. After removal of the TBS groups, the two *N*-Troc groups on each hexasaccharide were removed with zinc in acetic acid and the amines were acetylated to give compounds **57** and **61**, respectively. Finally, the *tert*-butyl ester at the carboxy-terminus of the amino acid was removed using TFA/DCM with anisole to give the final core 2 hexasaccharides **58** and **62**, ready for SPPS (*figure 4.32*).

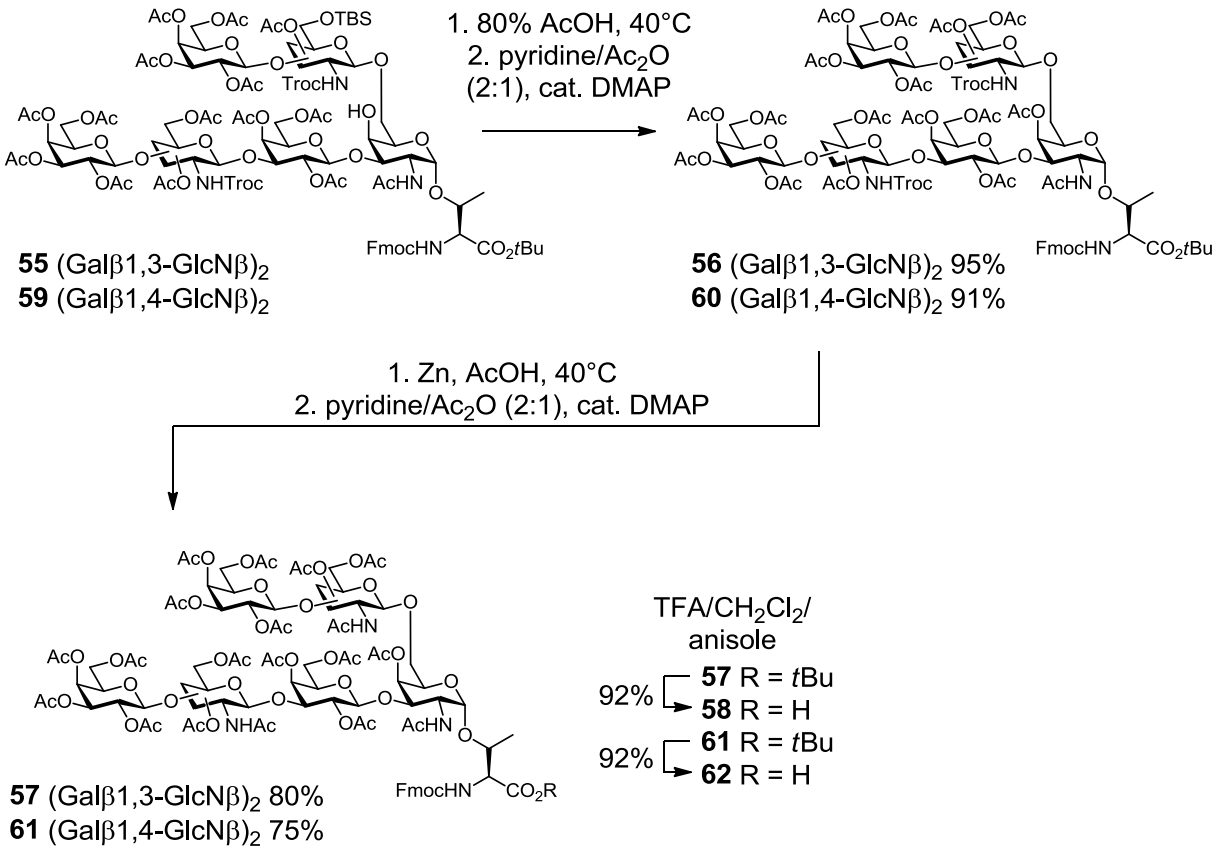


***Figure 4.32*:** Synthesis of extended core 2 amino acid building blocks **58** and **62** for SPPS.

4.2 Part 2 - Syntheses of mucin glycopeptide libraries

4.2.1 Synthesis plan for MUC1 and MUC5B glycopeptide library build-up

An extended library of mucin glycoproteins was synthesized with the synthetic T_N- and T-antigen and the extended core 1, 2 and 3 amino acids. These glycosylated amino acid building blocks were incorporated into the tandem repeat peptide sequences of MUC1 and MUC5B to generate several peptide isoforms/glycoforms for each of the mucins tandem repeat sequences. The different variants were chosen with respect to different glycosylation sites and multivalent occupation of these glycosylation sites. A MUC1 tandem repeat consists of 20 amino acids of the sequence PAHGVTSAPDTRPAPGSTAP. The sequence was first synthesized without glycosylation. The decision was made to synthesize a 19mer of the MUC1 peptide of the sequence PAHGVTSAPDTRPAPGSTA to avoid potential problems with diketopiperazine formation and increase the yields. The MUC1 19mer contains five potential glycosylation sites, of which three are composed of threonines and were occupied by the synthesized threonine amino acid building blocks. The MUC5B tandem repeat consists of 29 amino acids of the sequence ATGSTATPSSTPGTTHTPPVLTTTATTPT. For MUC5B the 13mer ATPSSTPGTTHTP from A^6 to P^{18} was synthesized. It contains seven potential glycosylation sites, of which five are located on threonine. *Figure 4.33* shows the chosen multivalent glycosylation pattern for the synthesis of the MUC1 and MUC5B peptides. This pattern was applied to the synthetic glycosylated amino acids. Like this, 8 (glycosylated amino acids) x 7 (peptide isoforms) = 56 MUC1(19mer) and 7 (glycosylated amino acids) x 5 (peptide isoforms) = 42 MUC5B(13mer) glycopeptides were synthesized (*figure 4.33*).

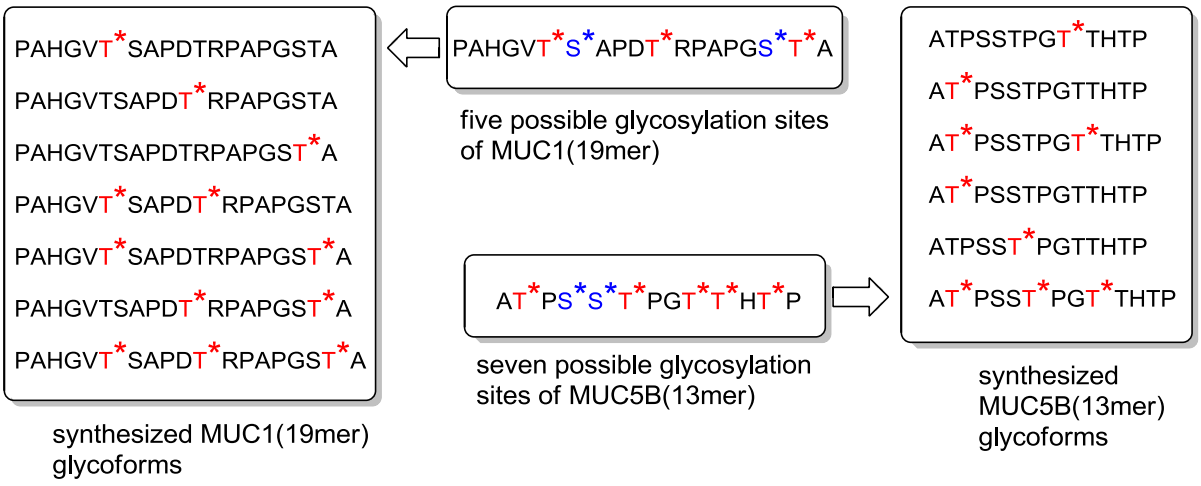


***Figure 4.33*:** Synthesized MUC1(19mer) and MUC5B(13mer) glycopeptides.

In the case of the MUC1 peptides, which were already used in the first microarray studies, also three MUC1 22mers and a few MUC1 sequences with migrating PDTR and GSTA domains were synthesized. Full lists of all synthesized MUC1 and MUC5B peptides are shown in *table 4.2* (MUC1, *chapter 4.2.3*) and in *table 4.3* (MUC5B, *chapter 4.2.4*).

4.2.2 Synthesis of a triethlyene glycol linker

Each synthesized mucin sequence, was extended with a triethylene glycol (TEG) linker amino acid at the N-terminal end. In the synthesis of immunogenic vaccine conjugates based on (glyco-) peptide antigens, spacers between the antigen and applied carriers or mitogens provide spatial separation of the different components and prohibit unwanted conformational influences.[271] Similar to this approach, the triethylene glycol attached to the N-terminus of the glycopeptides provides additional spacing to the surface of a microarray. In the field of carbohydrate microarrays it is common to use various spacers between the surface and the glycan.[272] Signal intensity variations depending on the length of spacers are reported and due to steric hindrance, too short linkers usually disturb the interactions between the glycan and the protein binding partners.[273] Oligoethylene glycol spacers are reported to have minimal nonspecific interaction with proteins on microarrays.[274] The utilized amine reactive microarray slides have a three-dimensional hydrophilic polymer coating with a hydrophilic tether connected to the reactive site. The triethylene glycol spacer is incorporated additionally to the surface tether of the commercial slide. The spacer displays an amine group and is therefore suited for the use with functionalized amine reactive microarray slides. Furthermore, with a glycol spacer attached, each glycopeptide can be used also for the synthesis of mucin glycopeptide vaccine conjugates.[271]

The synthesis of the triethylene glycol spacer amino acid followed a reported strategy.[275,276,240] The alkoxide of triethylene glycol was initially reacted in a *Michael*-addition with *tert*-butylacrylate to give the 12-hydroxy-4,7,10-trioxa-dodecanate-*tert*-butylester **63**.[275] The free hydroxyl group was transferred into a better leaving group by mesylation with methanesulfonyl chloride and then substituted with an azide group. The resulting 12-azido-4,7,10-trioxa-dodecanate-*tert*-butylester **64** was further reduced with raney-nickel in isopropanol to give the corresponding amine **65**.[276] The free amine was then protected using Fmoc-OSu and the orthogonally protected Fmoc amino acid **66** was obtained.[240] By deblocking the *tert*-butyl-carboxyl group in TFA, the adequate Fmoc amino acid **67** was received and ready for SPPS.[240] Compound **67** was synthesized in five steps and 30% yield overall (*figure 4.34*).

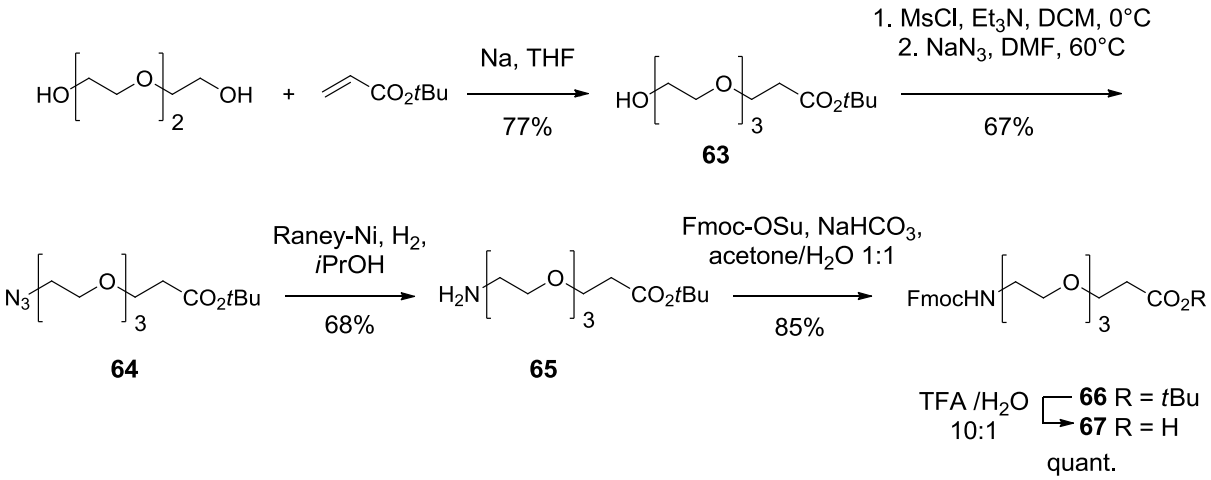


***Figure 4.34*:** Synthesis of triethylene glycol spacer amino acid **67**.

4.2.3 Synthesis of a MUC1 glycopeptide library

Since all MUC1 as well as all MUC5B glycopeptides follow the same SPPS-protocol, representative syntheses are shown in detail only for MUC1(19mer) **108,** MUC1(19mer) **109,** MUC1(19mer) **98** and MUC5B(13mer) **161**. Glycopeptide numbering, characteristic reaction parameters and yields for each synthesized glycopeptide are combined in *table 4.2* (MUC1) and *table 4.3* (MUC5B).

The MUC1 glycopeptides were generated on *TentaGel*® R Trt-Ala Fmoc resins with loading capacities of 0.17 mmol/g or 0.15 mmol/g. The comparatively low loading capacities support good availability of reaction centers and prohibit peptide aggregation. The synthesis of glycopeptide **108** was initiated by the cleavage of the amino acid Fmoc protecting group on the preloaded resin by addition of 20% piperidine in DMF (*figure 4.32*). Automated chain elongation of the first eight amino acids followed. In each elongation cycle, eight equivalents of the standard Fmoc protected amino acid were utilized in a peptide coupling mediated by HBTU/HOBt. Each cycle was finished by Fmoc deprotection with 20% piperidine in DMF. The glycosylated type-2 core 1 amino acid **54** (1.5 eq) was preactivated with HATU/HOAt and DIPEA in DMF in a separate vessel and then manually pipetted to the synthesis reactor. The reaction was periodically shaken by vortex for 6 h. The two preceding amino acids after the glycosylated amino acids were coupled in double couplings. After the completion of the MUC1(19mer) sequence by automated synthesis steps, the Fmoc protected spacer amino acid **67** (3 eq) was preactivated with HBTU/HOBt and DIPEA and then added manually to the reactor. The reaction was shaken for 2 h. Final Fmoc deprotection preceded detachment of the peptide from the resin using TFA/TIPS/H_2O (15:1:1) with simultaneous cleavage of all acid sensitive protecting groups on the side chains of the amino acids (*figure 4.35*). The crude peptide was lyophilized and then loaded onto a C18 cartridge (1 g) and desalted. The

glycopeptide was eluted successively with 30, 50, and 70% of acetonitrile in water. The eluted glycopeptide was again lyophilized before the following deacetylation step

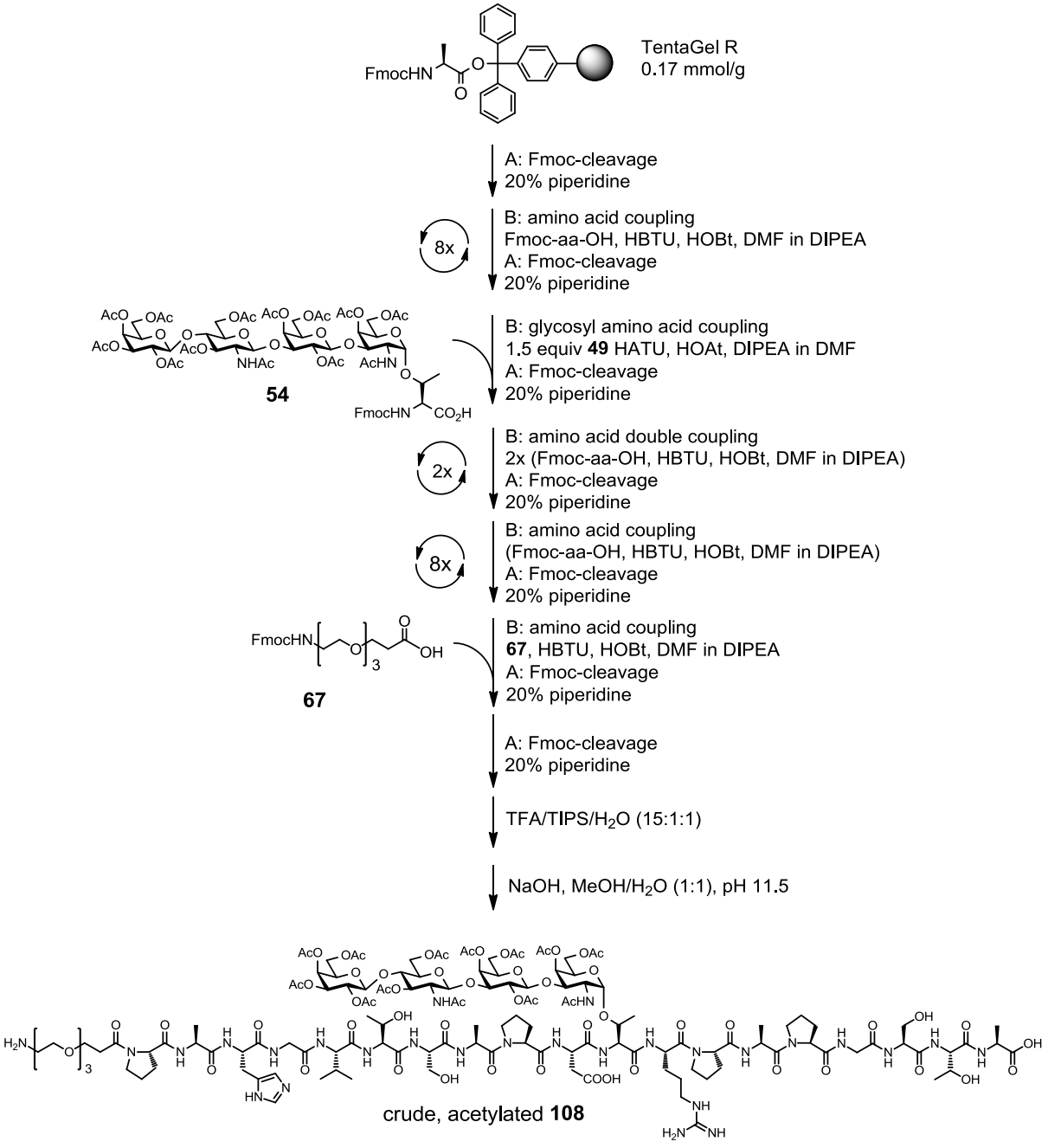


***Figure 4.35*:** SPPS of MUC1 glycopeptide **108**.

In order to hydrolyze the acetyl groups of the glycan moiety, the glycopeptides were dissolved in methanol/water (1:1) and 150 mM NaOH solution was added in small portions until a pH of 11.5 was reached. Several additions were necessary, due to buffering effects. After completion of the deacetylation the reaction was acidified with acetic acid and glycopeptide **108** was purified by preparative HPLC. After lyophilization, glycopeptide **108** was obtained in 53% yield overall, regarding the resin loading capacity (*figure 4.36*).

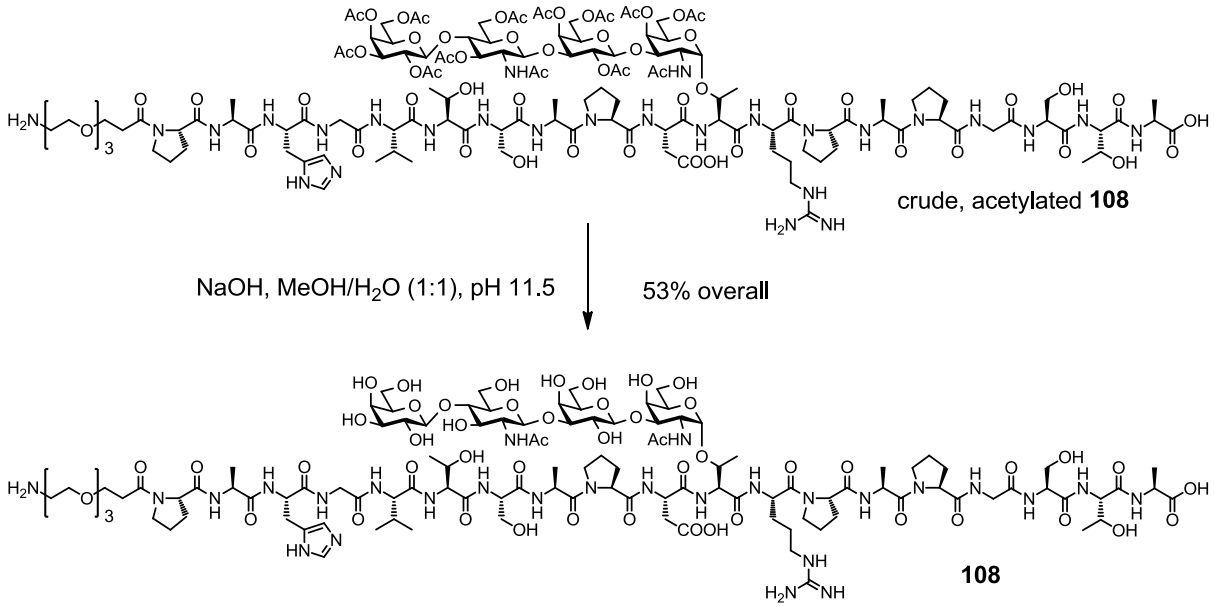


***Figure 4.36*:** Deacetylation and synthesis of MUC1(19mer) **108**.

Standard deacetylation conditions, like the mentioned *Zemplén* deacetylation at pH 8.5,[211] were not sufficient to the remove all acetyl groups. Chromatographic and mass spectrometric analysis showed that under these mild conditions one acetyl group was omitted. *S. Dziadek et al.* reported the synthesis of α2,3-sialyl-T-antigen glycopeptides, where the 4'-position of the galactose unit was sterically hindered and required harsher conditions for deacetylation.[215] Here, ESI-mass-spectrometry with collision induced dissociation (CID) fragmentation of monoacetylated compound **108** indeed revealed that a hydroxyl group of the core 1 galactose remained acetylated at standard *Zemplén* conditions.

All glycopeptides carrying an elongated tetrasaccharide core 1 motif, including the elongated hexasaccharide core 2 structures, share the hindered 4'-OAc of the core 1 galactose. All of these glycopeptides were treated at higher pH (11.5) for full deacetylation. Extended core 3 glycopeptides, devoid of a core 1 galactose, but also the not extended T-antigen glycopeptides, could be deacetylated under standard *Zemplén* conditions. Stringent chromatographic reaction control by HPLC was necessary to keep unavoidable β-elimination at pH 11.5 at acceptable levels. The sodium hydroxide solution was added carefully in small doses to the dissolved glycopeptide in the methanol/water mixture. Setting the pH to 11.5 often took a long time, since glycopeptides with large glycans showed strong buffering effects, especially in the case multiglycosylated glycopeptides. The deacetylation step, resulting in glycopeptide **108**, is shown in the HPLC chromatograms in *figure 4.37*.

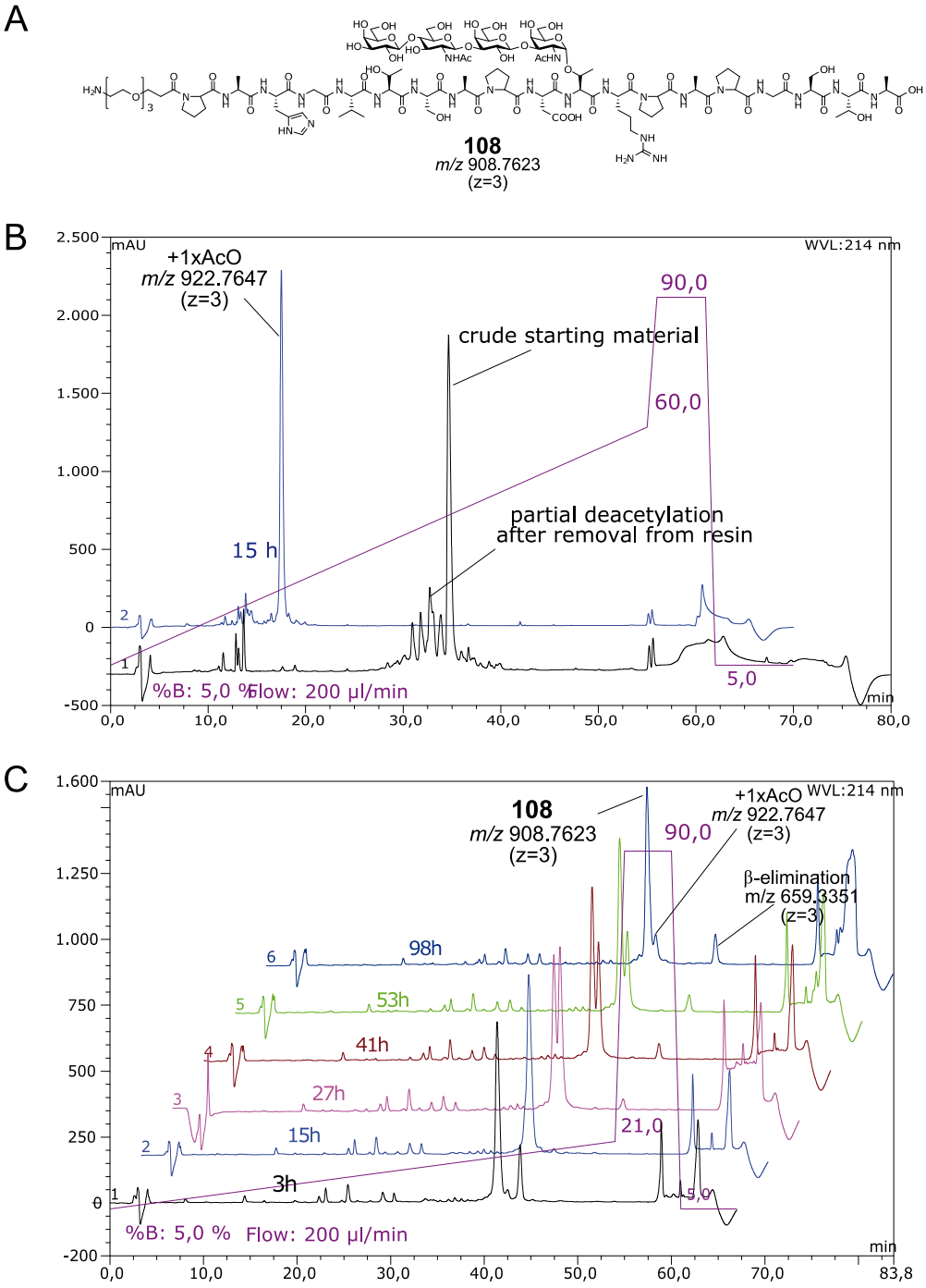


***Figure 4.37*:** Analytical HPLC control of deacetylation of glycopeptide **108**. **A**: Structure of deacetylated product **108**. **B:** After 15 h (grad.: water/84% ACN + 0.1% TFA (95:5) → (40:60), 55 min = 1% min^{-1}). Monoacetylated glycopeptide (blue curve) is left. All starting material (black curve) is completely consumed. **C:** Until 98 h (grad.: water/84% ACN + 0.1% TFA (95:5) → (79:21), 54 min, 0.3% min^{-1}). Deacetylation of monoacetylated glycopeptide (+1xAcO) and increasing β-elimination at pH 11.5.

Partial deacetylation commonly occurred during the release of the glycopeptides from the resin by the treatment with TFA (*figure 4.37*, **B**, black curve). After 15 h at pH 9.5, one acetyl group was left. Adjustment of the pH to 11.5 and HPLC monitoring with a slower gradient

(water/84% ACN (95:5)→(79:21), 54 min, 0.3% min^{-1}) enabled removal of the last acetyl group (*figure 4.37*, **C**). After 98 h only small amounts of monoacetylated product (+1x AcO) were left. Due to the increased pH, 14% of β-elimination was observed (by peak integration). Purification by preparative HPLC separated the main product peak from the leftover monoacetylated glycopeptide (*figure 4.38*). Analytical HPLC of the main fraction F1 showed a single product and with the expected molecular weight (*figure 4.38*, **B** and **C**). The separated fraction F2 was a mixture of fully deacetylated and monoacetylated **108** (*figure 4.38*, **D**). The last fraction F3 contained peptide with β-elimination (*figure 4.38*, **E**).

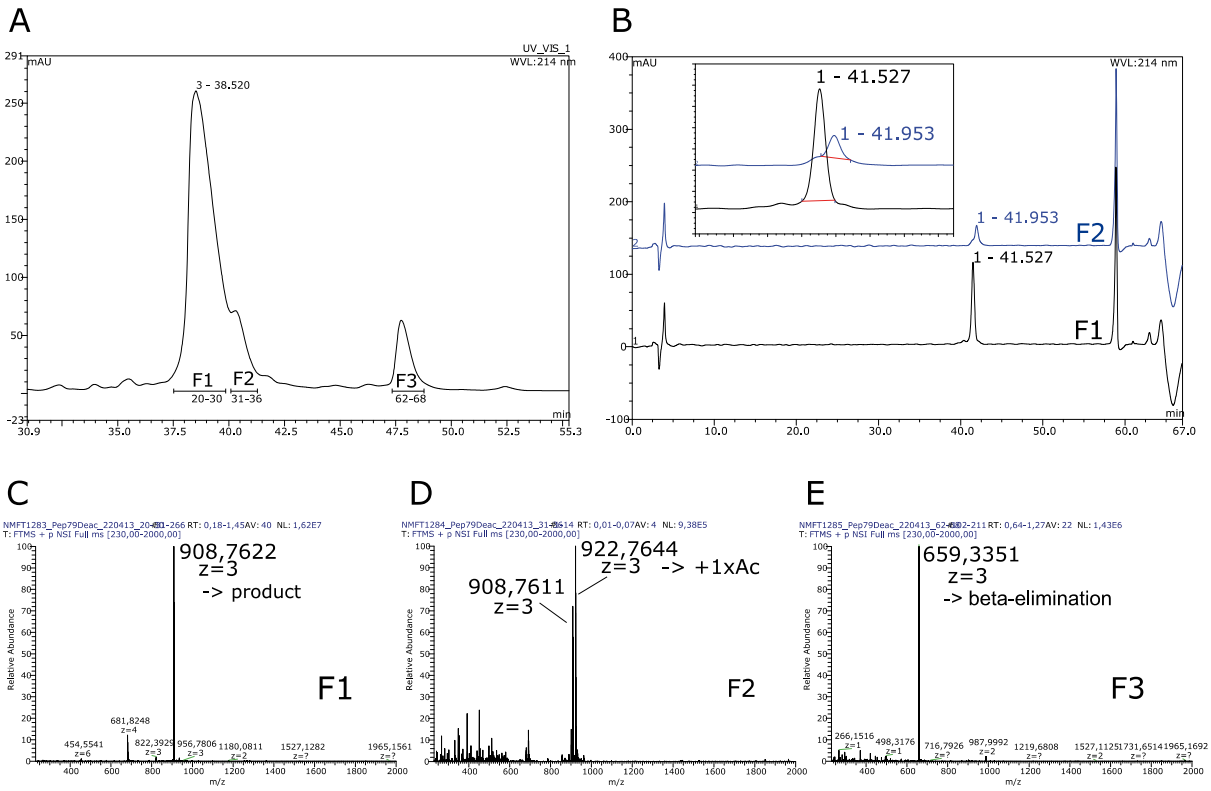


Figure 4.38: HPLC and HR-ESI-MS analysis after deacetylation of **108**. **A**: Preparative HPLC (grad.: water/84% acetonitrile + 0.1% TFA (95:5) → (79:21), 60 min = 0.26% min^{-1}) separation of fractions F1 = 20-30, F2 = 31-36 and F3 = 62-68. **B**: Analytical HPLC (grad.: water/84% acetonitrile + 0.1% TFA (95:5) → (21:79), 54 min = 0.27% min^{-1}) of main peak F1 and F2. **C:** HR-ESI-MS of F1, *m/z*: 908.7621 ($[M+3H]^{3+}$, calc. 908.7603); 681.8247 ($[M+4H]^{4+}$, calc. 681.8220) → complete deacetylation. **D:** HR-ESI-MS of F1, *m/z*: 922.7644 ($[M+3H]^{3+}$, calc. 922.7638) → monoacetylation; **E:** HR-ESI-MS of F1, *m/z*: 659.0008 ($[M+3H]^{3+}$, calc. 659.0007)→ β-elimination.

Glycopeptide **109** carried the type-2 core 1 glycan in the GST*A domain. Deacetylation of **109** was performed under the same conditions as **108**. The majority of acetyl groups were hydrolyzed rapidly (*figure 4.39*, **B**), while further hydrolysis of the remaining 4'-OAc on galactose was again slow (*figure 4.39*, **C**).

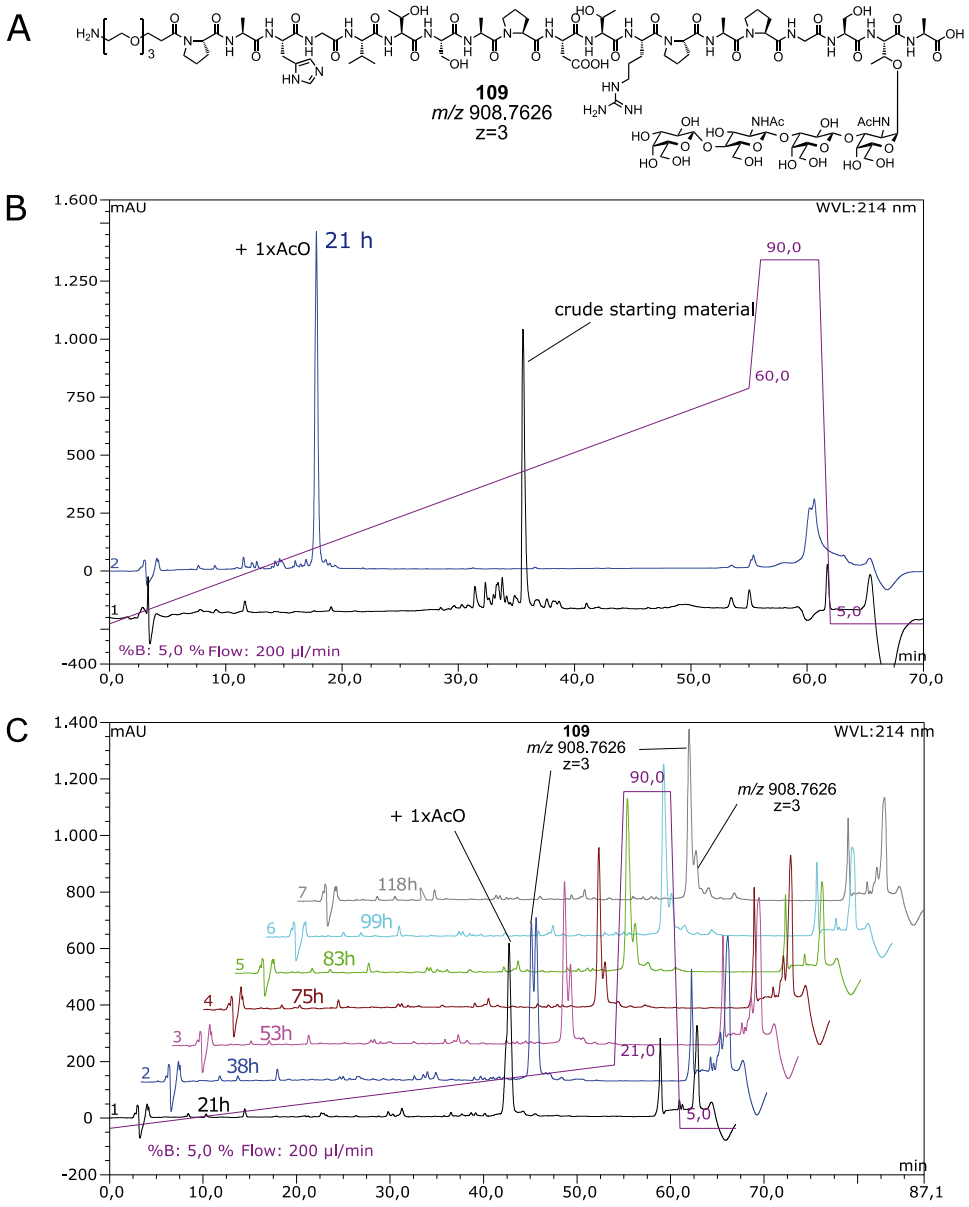


***Figure 4.39*:** Analytical HPLC control of deacetylation of glycopeptide **109**. **A**: Structure of deacetylated product **109**. **B:** After 21 h (grad.: water/84% acetonitrile + 0.1% TFA (95:5) → (40:60), 55 min = 1% min^{-1}): Monoacetylated glycopeptide (blue curve) is left. All starting material (black curve) is completely consumed. **C:** Until 118 h (grad.: water/84% acetonitrile + 0.1% TFA (95:5) → (21:79), 54 min = 0.3% min^{-1}): No further product formation from 75 h.

No further change was visible in the HPLC chromatogram after 75 h. ESI-MS analysis after separation of the main peak and the minor site product in preparative HPLC revealed that the minor product was not leftover monoactylated glycopeptide but a product with the same mass as the desired deacetylated glycopeptides (*figure 4.40*). It was assumed that the minor peak was epimerization product. Depending on the glycosylation site and extend of glycosylation this side reaction appeared for some peptides in varying amounts and the by-product was usually effectively removed by preparative HPLC.

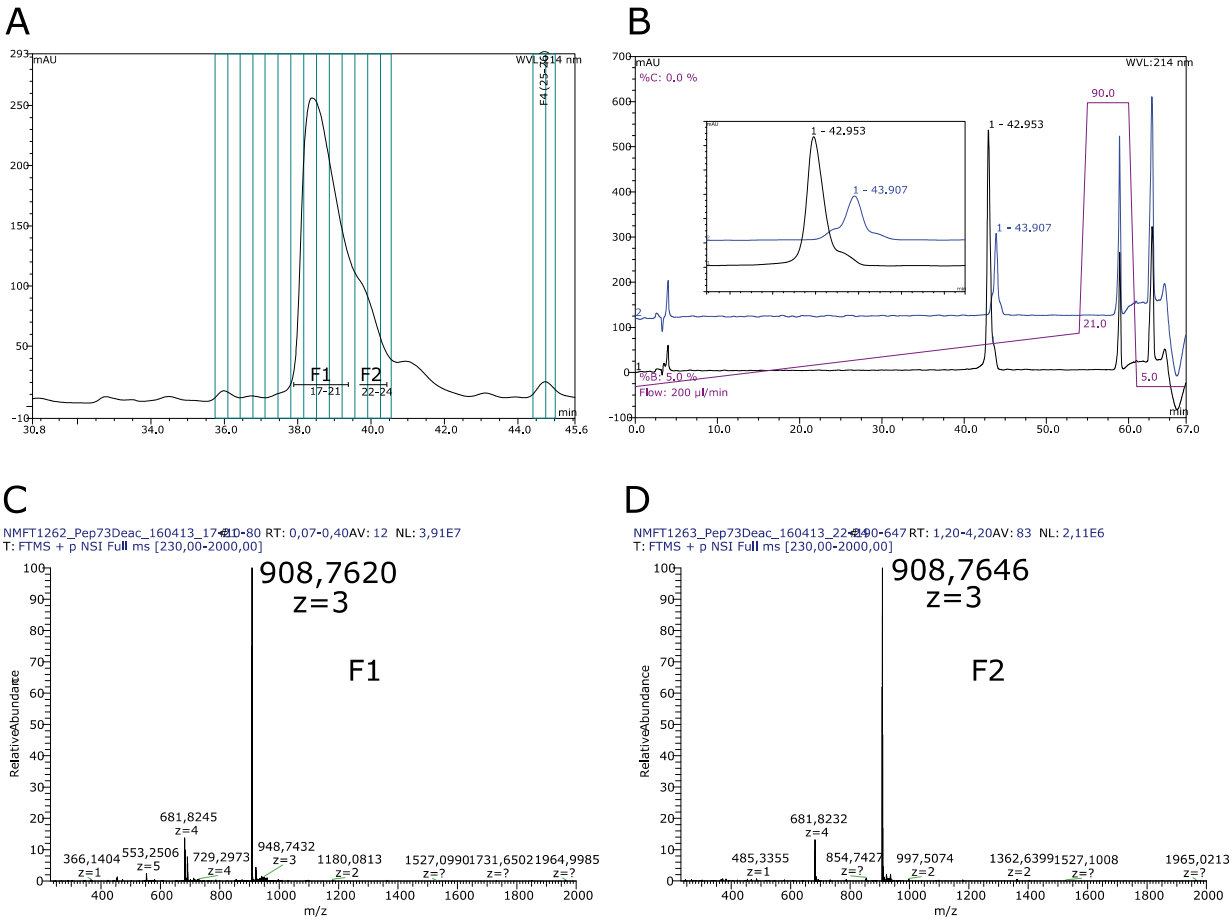


***Figure 4.40*:** HPLC and HR-ESI-MS analysis after deacetylation of **109**. **A:** Preparative HPLC (grad.: water/84% acetonitrile + 0.1% TFA (95:5) → (40:60), 55 min = 1% min^{-1}) separation of fractions F1 = 17-21, F2 = 22-24. **B:** Analytical HPLC (grad.: water/84% acetonitrile + 0.1% TFA (95:5) → (79:21), 54 min = 0.3% min^{-1}) of fractions F1 (black curve) and F2 (blue curve). **C:** HR-ESI-MS of fraction F1, *m/z*: 908.7620 ($[M+3H]^{3+}$, calc. 908.7603); 681.8245 ($[M+4H]^{4+}$, calc. 681.8220) → complete deacetylation. **D:** HR-ESI-MS of fraction F2, *m/z*: 908.7646 ($[M+3H]^{3+}$, calc. 908.7603) → assumed epimerization.

Glycopeptides without the internal galactose of the central core 1 motif, such as the T-antigen or core 3 glycopeptides, were deacetylated at pH 9.5 in a few hours. As an example, *figure 4.41* shows HPLC chromatograms of trivalent type-2 core 3 glycopeptide **98** (*figure 4.41*, **A**), after SPPS (*figure 4.41*, **B**) and at deacetylation conditions (*figure 4.41*, **C**). The deacetylation reaction was stopped after 20 h by addition of acetic acid.

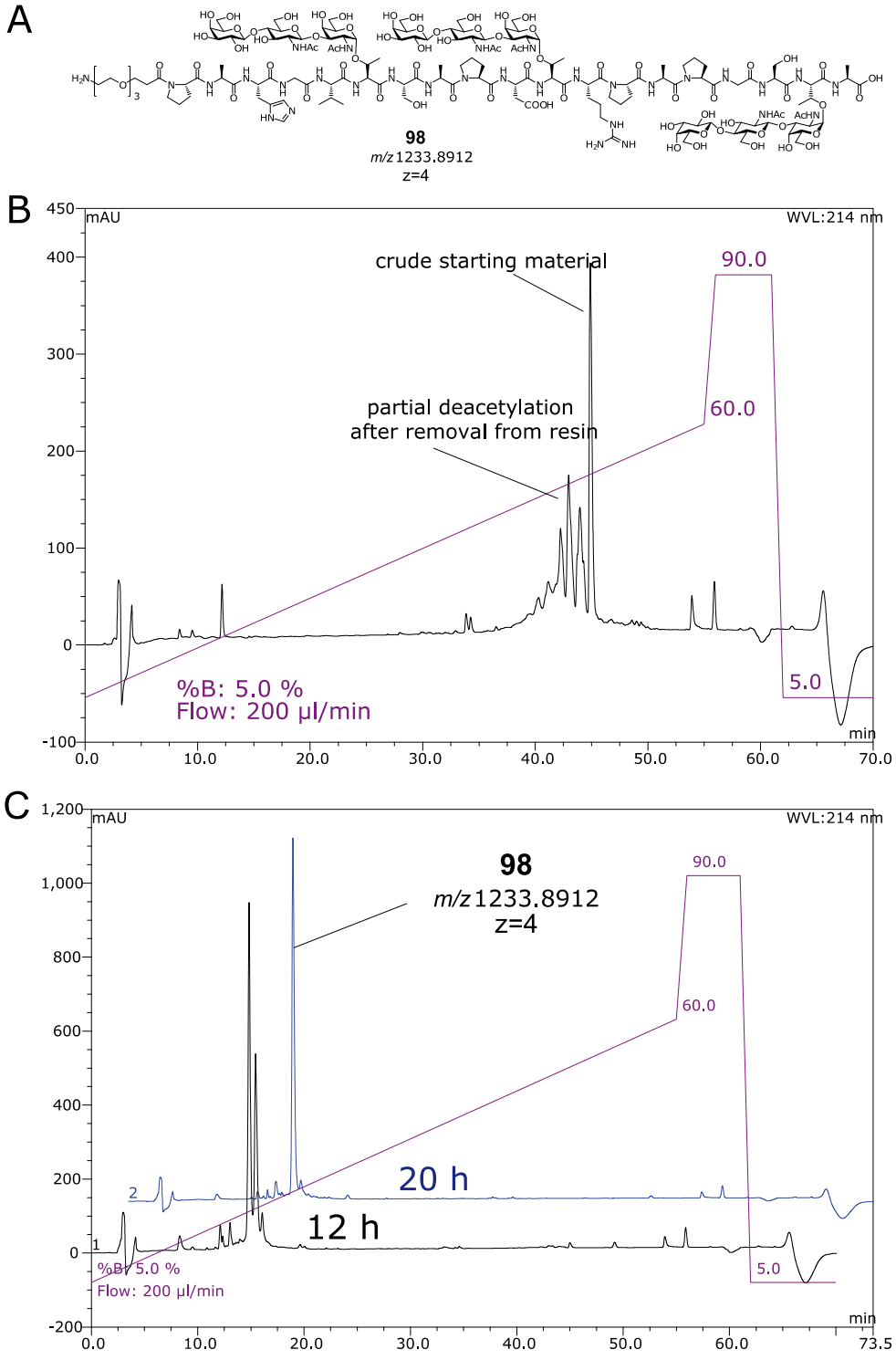


***Figure 4.41*: A:** Structure of triglycosylated type-2 core 3 peptide **98**. **B:** Analytical HPLC chromatogram after SPPS. **C**: Complete deacetylation after 20 h. (grad. in **B** and **C**: water/84% acetonitrile + 0.1% TFA (95:5) → (40:60), 55 min = 1% min^{-1}).

The MUC1 glycopeptide library was synthesized according to the examples shown above. MUC1(22mer) PAHGVTSAPDT*RPAPGS*T*APPA sequences **131**-**133** were synthesized according to the same SPPS protocol. The serine **211** and threonine **212** T_N-antigen building

blocks for **131**-**133** and **69**-**75** were kindly provided by Dr. Hui Cai. *Table 4.2* summarizes all synthesized MUC1 glycopeptides and relevant reaction parameters.

Table 4.2: Overview of all synthesized MUC1 sequences with relevant reaction parameters.

			SPPS		deacetylation		
Glycan	**#**	**Sequence**	eq glycosylated amino acid	min reaction time glycosylated amino acid (h)	pH	reaction time (h)	final overall yield (%)
T_N	**68**	PAHGVT*SAPDTRPAPGSTA	1.5	2	9.5	6	67
	69	PAHGVTSAPDT*RPAPGSTA	1.5	2	9.5	6	65
	70	PAHGVTSAPDTRPAPGST*A	1.5	2	9.5	6	70
	71	PAHGVT*SAPDT*RPAPGSTA	1.5	2	9.5	6	51
	72	PAHGVT*SAPDTRPAPGST*A	1.5	2	9.5	6	52
	73	PAHGVTSAPDT*RPAPGST*A	1.5	2	9.5	6	56
	74	PAHGVT*SAPDT*RPAPGST*A	1.5	2	9.5	6	40
T	**75**	PAHGVT*SAPDTRPAPGSTA	1.5	2	9.5	14	85
	76	PAHGVTSAPDT*RPAPGSTA	1.5	2	9.5	14	85
	77	PAHGVTSAPDTRPAPGST*A	1.5	2	9.5	14	76
	78	PAHGVT*SAPDT*RPAPGSTA	1.5	2	9.5	14	67
	79	PAHGVT*SAPDTRPAPGST*A	1.5	2	9.5	14	85
	80	PAHGVTSAPDT*RPAPGST*A	1.5	2	9.5	14	62
	81	PAHGVT*SAPDT*RPAPGST*A	1.5	2	9.5	14	59
	82	PGSTAPPAHGVTSAPDT*RPA	1.5	2	9.5	4	71
	83	APDT*RPAPGSTAPPAHGVTSA	1.5	2	9.5	4	54
	84	APDT*RPA	1.5	2	9.5	15	66
type-1 core 3	**85**	PAHGVT*SAPDTRPAPGSTA	1.5	5	9.5	16	60
	86	PAHGVTSAPDT*RPAPGSTA	1.5	5	9.5	16	74
	87	PAHGVTSAPDTRPAPGST*A	1.5	5	9.5	16	79
	88	PAHGVT*SAPDT*RPAPGSTA	1.5	5	9.5	16	37
	89	PAHGVT*SAPDTRPAPGST*A	1.5	5	9.5	16	75
	90	PAHGVTSAPDT*RPAPGST*A	1.5	5	9.5	24	47
	91	PAHGVT*SAPDT*RPAPGST*A	1.5	5	9.5	16	62
type-2 core 3	**92**	PAHGVT*SAPDTRPAPGSTA	1.5	5	9.5	18	70
	93	PAHGVTSAPDT*RPAPGSTA	1.5	5	9.5	18	65
	94	PAHGVTSAPDTRPAPGST*A	1.5	5	9.5	27	72
	95	PAHGVT*SAPDT*RPAPGSTA	1.5	5	9.5	20	74
	96	PAHGVT*SAPDTRPAPGST*A	1.5	5	9.5	21	62
	97	PAHGVTSAPDT*RPAPGST*A	1.5	5	9.5	21	62
	98	PAHGVT*SAPDT*RPAPGST*A	1.5	5	9.5	20	59
	99	PAHGVT*SA	1.5	5	9.5	15	24
type-1 core 1	**100**	PAHGVT*SAPDTRPAPGSTA	1.5	6	11.5	53	40
	101	PAHGVTSAPDT*RPAPGSTA	1.5	6	11.5	53	45
	102	PAHGVTSAPDTRPAPGST*A	1.5	6	11.5	80	43
	103	PAHGVT*SAPDT*RPAPGSTA	1.5	6	11.5	84	36
	104	PAHGVT*SAPDTRPAPGST*A	1.5	6	11.5	62	35
	105	PAHGVTSAPDT*RPAPGST*A	1.5	6	11.5	62	36
	106	PAHGVT*SAPDT*RPAPGST*A	1.5	6	11.5	84	27
type-2 core 1	**107**	PAHGVT*SAPDTRPAPGSTA	1.5	6	11.5	84	49
	108	PAHGVTSAPDT*RPAPGSTA	1.5	6	11.5	98	53
	109	PAHGVTSAPDTRPAPGST*A	1.5	6	11.5	108	45
	110	PAHGVT*SAPDT*RPAPGSTA	1.5	6	11.5	119	34
	111	PAHGVT*SAPDTRPAPGST*A	1.5	6	11.5	95	26
	112	PAHGVTSAPDT*RPAPGST*A	1.5	6	11.5	80	29
	113	PAHGVT*SAPDT*RPAPGST*A	1.5	6	11.5	127	17
	114	GST*APPAHGVTSAPDTRPA	1.5	6	11.5	78	14
	115	PDTRPAPGST*APPAHGVTSA	1.5	6	11.5	80	41
	116	PDTRPAPGSTAPPAHGVT*SA	1.5	6	11.5	60	40
type-1 core 2	**117**	PAHGVT*SAPDTRPAPGSTA	2.0	10	11.5	81	36
	118	PAHGVTSAPDT*RPAPGSTA	2.0	10	11.5	82	34

Glycan	#	Sequence	SPPS		deacetylation		
			eq glycosylated amino acid	min reaction time glycosylated amino acid (h)	pH	reaction time (h)	final overall yield (%)
	119	PAHGVTSAPDTRPAPGST*A	2.0	10	11.5	103	37
	120	PAHGVT*SAPDT*RPAPGSTA	2.0	10	11.5	82	22
	121	PAHGVT*SAPDTRPAPGST*A	2.0	10	11.5	138	20
	122	PAHGVTSAPDT*RPAPGST*A	2.0	10	11.5	82	20
	123	PAHGVT*SAPDT*RPAPGST*A	2.0	10	11.5	129	9
type-2 core 2	**124**	PAHGVT*SAPDTRPAPGSTA	2.0	10	11.5	90	43
	125	PAHGVTSAPDT*RPAPGSTA	2.0	10	11.5	120	33
	126	PAHGVTSAPDTRPAPGST*A	2.0	10	11.5	129	40
	127	PAHGVT*SAPDT*RPAPGSTA	2.0	10	11.5	96	18
	128	PAHGVT*SAPDTRPAPGST*A	2.0	10	11.5	129	8
	129	PAHGVTSAPDT*RPAPGST*A	2.0	10	11.5	149	14
	130	PAHGVT*SAPDT*RPAPGST*A	2.0	10	11.5	129	3
*type-2 core 1 + $^{\#}$2xTn	**131**	PAHGVTSAPDT*RPAPGS$^{\#}$T$^{\#}$APPA	3x1.5	9	11.5	106	16
*type-2 core 3 + $^{\#}$2xTn	**132**	PAHGVTSAPDT*RPAPGS$^{\#}$T$^{\#}$APPA	3x1.5	9	9.5	16	19
*type-2 core 2 + $^{\#}$2xTn	**133**	PAHGVTSAPDT*RPAPGS$^{\#}$T$^{\#}$APPA	1x2.0+2x1.5	9	11.5	106	8

4.2.4 Synthesis of a MUC5B glycopeptide library

MUC5B(13mer) glycopeptides were synthesized on *TentaGel*® R Trt-Pro Fmoc. The MUC5B tandem repeat consists of 29 amino acids of the sequence ATGSTATPSSTPGTTHTPPVLTTTATTPT with 16 potential serine or threonine glycosylation sites. The peptide sequence ATPSSTPGTTHTP was chosen for a feasible synthesis of a multivalent pattern, similar to the described MUC1 glycopeptides. This MUC5B(13mer) sequence contains seven *O*-glycosylation sites of which five could be addressed with the glycosylated threonine amino acids described herein.

Triglycosylated MUC5B peptide **161** with type-1 core 2 hexasaccharide glycan decoration was synthesized according to the SPPS protocol displayed in *figure 4.35*. The type-1 core 2 glycosylated amino acid building block **58** was incorporated at the threonine glycosylation sites T^2, T^6 and T^9 and was applied in 2 eq excess with elongated reaction times (minimum 10 h). *Figure 4.42* shows the analytical HPLC chromatogram of acetylated glycopeptide **161** after SPPS (*figure 4.42*, **B**) and the course of deacetylation to result in the final product sequence **161** (*figure 4.42*, **C**).

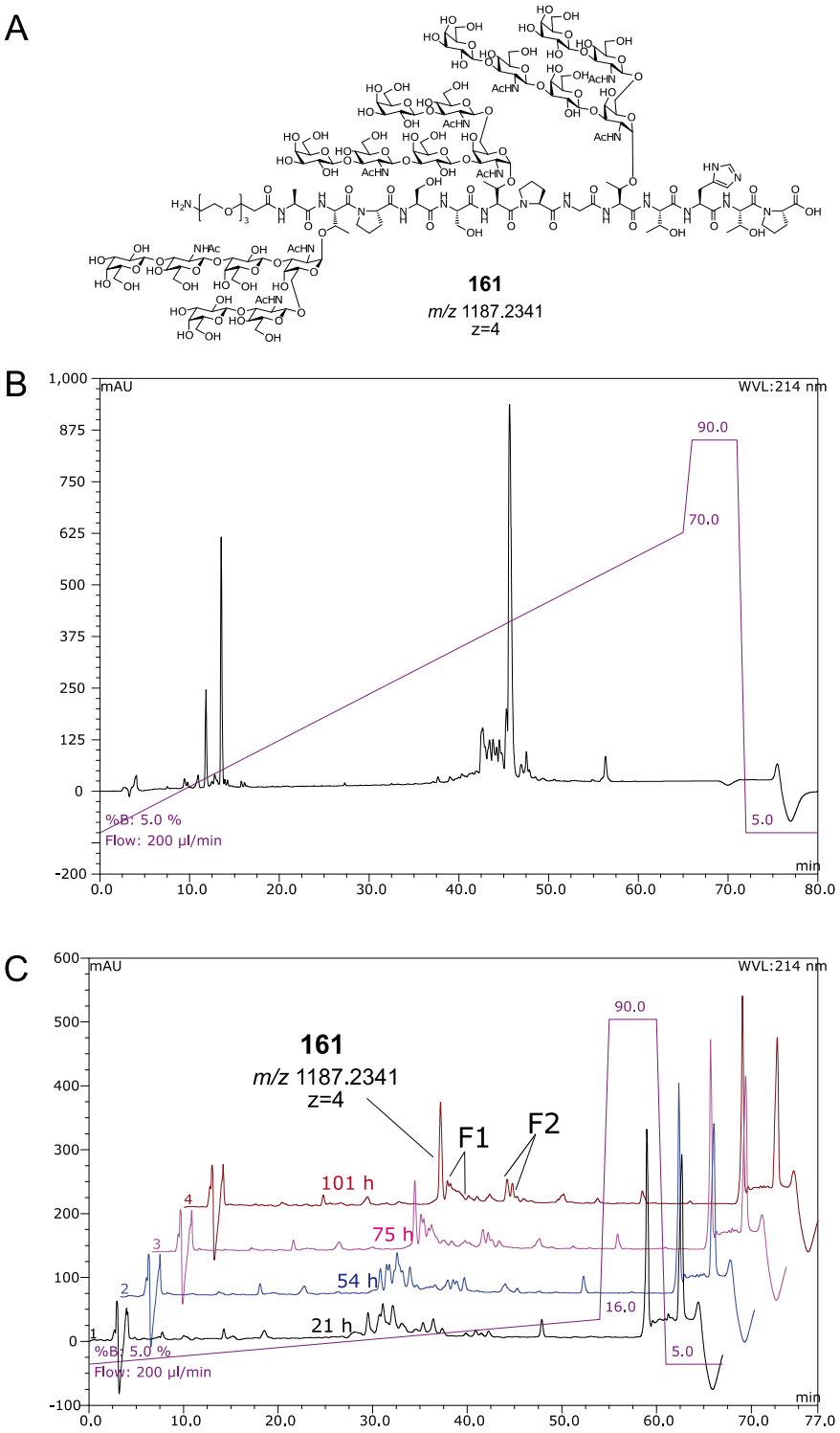


***Figure 4.42*: A:** Structure of deacetylated product **161**. **B:** Analytical HPLC of crude, acetylated glycopeptide **161** after SPPS (grad.: water/84% acetonitrile + 0.1% TFA (95:5) → (30:70), 65 min, 1% min^{-1}). **C:** Deacetylation of **161** (grad.: water/84% acetonitrile + 0.1% TFA (95:5) → (16:84), 54 min, 0.2% min^{-1}).

The main peak in the chromatogram (*figure 4.42*, **C**) represented the desired fully deacetylated product, the triglycosylated type-2 core 2 glycopeptide **161**. Fractions F1 and F2 were also isolated and analyzed my mass spectrometry (*figure 4.43*). Fraction F1

contains a mixture of product isomers, which might be peptide epimerization and product masses fitting for shorter, acetyl capped sequences (*figure 4.43*, **A**). The sequences were capped at the corresponding proline amino acids P^7 and P^3, after which a glycosylated amino acid was supposed to follow. Since the used glycosylated amino acid **58** was applied in the synthesis of various other MUC1 and MUC5B sequences, a contamination with acetic acid, leftover from silica column chromatography after *tert*-butyl ester removal, was excluded. It was assumed that intra- or intermolecular acetyl-migration took place during the coupling of the second and third glycosylated amino acid. The short distances between the selected glycosylation sites in the MUC5B sequence could have facilitated close proximities between terminal prolines and the previously installed glycan. Then, the elongated reaction times and slower reaction rates of the bulky hexasaccharide glycosylated amino acids could have promoted acetyl-migration. Fraction F2 contained a mixture of diglycosylated deletion sequences, missing a complete glycosylated amino acid and glycopeptides with a single β-elimination site, having a unsaturated homoalanine instead of a glycosylated threonine (*figure 4.43*, **B**).

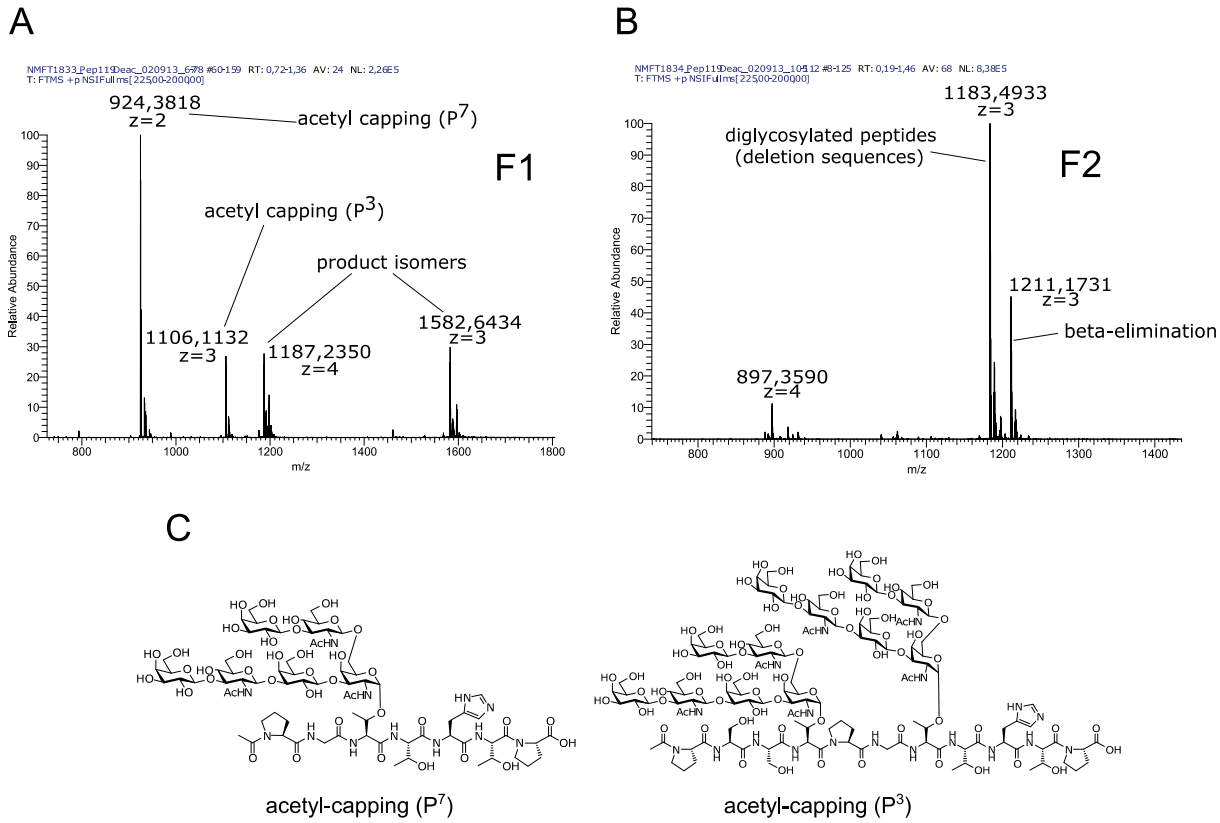


***Figure 4.43*:** Side products in SPPS of glycopeptide **161** in fractions F1 and F2.

However, the desired glycopeptide product was obtained as the main product (*figure 4.42*, **C**). Glycopeptide **161** was obtained in 10% yield overall after the final deacetylation, based on the amount of initially employed SPPS-resin. The synthesis of glycopeptide **161**

demonstrates the feasibility of SPPS of short peptide sequences (13 amino acids) with bulky glycans motifs, distributed on densely occupied glycosylation sites. The carbohydrate portion accounts for almost 70% of the overall weight of glycopeptide **161**. The MUC5B glycopeptide library was synthesized according to the general SPPS protocol also applied for MUC1 glycopeptide synthesis (*figure 4.35*). Relevant reaction parameters and overall yields for the synthesized MUC5B sequences are summarized in *table 4.3*.

***Table 4.3*:** Overview of all synthesized MUC5B sequences with relevant reaction parameters.

			SPPS		**deacetylation**		
Glycan	**#**	**Sequence**	equiv glycosylated amino acid	min reaction time for glycosylated amino acid/h	pH	time/h	final overall yield/%
T	**134**	ATPSSTPGT*THTP	1.5	2	9.5	14	78
	135	AT*PSSTPGTTHTP	1.5	2	9.5	8	72
	136	AT*PSSTPGT*THTP	1.5	2	9.5	14	79
	137	AT*PSST*PGTTHTP	1.5	2	9.5	14	62
	138	AT*PSST*PGT*THTP	1.5	2	9.5	14	61
	139	T*GSTAT*PSST*PGT*THTP	1.5	2	9.5	18	33
core 3 type1	**140**	ATPSSTPGT*THTP	1.5	5	9.5	12	72
	141	AT*PSSTPGTTHTP	1.5	5	9.5	12	70
	142	AT*PSSTPGT*THTP	1.5	5	9.5	16	61
	143	AT*PSST*PGTTHTP	1.5	5	9.5	12	54
	144	AT*PSST*PGT*THTP	1.5	5	9.5	21	62
	145	T*GSTAT*PSST*PGT*THTP	1.5	5	9.5	21	36
core 3 type2	**146**	ATPSSTPGT*THTP	1.5	5	9.5	12	67
	147	AT*PSSTPGTTHTP	1.5	5	9.5	18	75
	148	AT*PSSTPGT*THTP	1.5	5	9.5	20	62
	149	AT*PSST*PGTTHTP	1.5	5	9.5	12	67
	150	AT*PSST*PGT*THTP	1.5	5	9.5	20	65
	151	T*GSTAT*PSST*PGT*THTP	1.5	5	9.5	18	49
core 1 type 1	**152**	ATPSSTPGT*THTP	1.5	6	11.5	36	36
	153	AT*PSSTPGTTHTP	1.5	6	11.5	44	49
	154	AT*PSSTPGT*THTP	1.5	6	11.5	40	30
	155	AT*PSST*PGTTHTP	1.5	6	11.5	110	40
	156	AT*PSST*PGT*THTP	1.5	6	11.5	110	19
	157	T*GSTAT*PSST*PGT*THTP	1.5	6	11.5	110	10
core 1 type2	**158**	ATPSSTPGT*THTP	1.5	6	11.5	140	17
	159	AT*PSSTPGTTHTP	1.5	6	11.5	67	37
	160	AT*PSSTPGT*THTP	1.5	6	11.5	38	22
	161	AT*PSST*PGTTHTP	1.5	6	11.5	113	22
	162	AT*PSST*PGT*THTP	1.5	6	11.5	87	15
	163	T*GSTAT*PSST*PGT*THTP	1.5	6	11.5	87	9
core 2 type 1	**164**	ATPSSTPGT*THTP	2.0	10	11.5	185	20
	165	AT*PSSTPGTTHTP	2.0	10	11.5	175	28
	166	AT*PSSTPGT*THTP	2.0	10	11.5	146	16
	167	AT*PSST*PGTTHTP	2.0	10	11.5	143	21
	168	AT*PSST*PGT*THTP	2.0	10	11.5	101	10
core 2 type2	**169**	ATPSSTPGT*THTP	2.0	10	11.5	185	15
	170	AT*PSSTPGTTHTP	2.0	10	11.5	196	28
	171	AT*PSSTPGT*THTP	2.0	10	11.5	190	17
	172	AT*PSST*PGTTHTP	2.0	10	11.5	143	21
	173	AT*PSST*PGT*THTP	2.0	10	11.5	124	9

4.2.5 Synthesis of sialylated MUC1 glycopeptides

4.2.5.1 Enzymatic carbohydrate extension of MUC1 glycopeptides

Sialylation is a frequent glycan modification, terminating carbohydrate chains in *O*- and *N*-glycoproteins and glycolipids. In this context, glycans are usually decorated with *N*-acetylneuraminic acid (Neu5Ac) α2,3- or α2,6-linked to terminal galactose, α2,6-linked to GalNAc or sometimes α2,8- or α2,9-linked to a preceding Neu5Ac.[277] This type of glycosylation has significant effects on protein structure and is involved in a plethora of glycan-protein interactions.[278] Neu5Ac adds a negative charge to a glycoprotein, increasing the hydrophilicity and providing charge repulsion of cells, for example sialylated erythrocytes in circulation. The half-life of serum glycoproteins is influenced, since non-sialylated glycans are recognized by asialoglycoprotein receptors on hepatocytes and are then cleared from circulation. α2,3-Sialylation together with 1,3/4-fucosylation on Galβ1,3-GlcNAcβ and Galβ1,4-GlcNAcβ chains, forms sLe^a and sLe^x (sialyl Lewis) glycans. Lewis structures are essential for leukocyte recruitment to inflamed tissues, due to binding to the corresponding selectins on the endothelial tissue, blood platelets or to the leukocyte itself. Further, generation of sialyl Lewis motifs on mucins is increased in most carcinomas facilitating metastasis by aforementioned selectin interaction. The α2,3-sialylation is in particular responsible for the binding of the human adopted influenza virus. Also, α2,6-sialylation may have strong influence on apoptotic effects mediated by galectins. Cancer cells are suspected make to use of this “carbohydrate on-off switch”, in order to avoid cell death by extracellular galectin-induced apoptosis.[93]

A selection of the synthesized glycopeptides was enzymatically sialylated. Here, terminal sialylation of the T-antigen and extended core 1, 2 and 3 structures, represents the physiological glycosylation state of mucins and binding behavior of antibodies and lectins was evaluated with the sialylated glycopeptides.

A chemoenzymatic approach on the peptide level was chosen, since it offers the option to modify certain peptides after SPPS, instead of preparing an extra set of chemically sialylated glycosyl amino acids before incorporation in SPPS. Chemoenzymatic glycosylation at the amino acid stage before Fmoc-SPPS would not be convenient because of protecting group issues. Several chemical methods of sialylation of complex carbohydrates have been presented on the level of oligosaccharides and glycosylated amino acids. Despite all efforts, chemical sialylations are still very challenging.[279,280] The presence of an electron withdrawing carboxyl group at the anomeric center prompts elimination reactions towards 2,3-dehydro compounds, while the absence of a participating neighboring group at position 3 is unfavorable for the stereoselectivity of the glycosylation. Appropriate choices of anomeric C1

protecting groups and C2 leaving groups are essential for chemical sialylations. In contrast, enzymatic extension with sialyltransferases is reliable in terms of regio- and stereoselective sialylation outcome.[19] The only drawback is the fairly high price of the nucleotide donor sugar, here cytidine-5´-monophospho-*N*-acetylneuraminic acid (CMP-Neu5Ac). Alternatively, it is possible to generate CMP-Neu5Ac *in situ* from cytidine triphosphate (CTP) and Neu5Ac and Neu5Ac from *N*-acetylmannosamine in a one-pot multienzyme system (OPME).[281]

Three different commercial and one expressed (generously provided by Prof. Xi Chen) sialyltransferase were applied for glycopeptide sialylation in this work:

1. Rat 2,3-OST: α2,3-(O)-Sialyltransferase from rat, recombinant (*Spodoptera frugiperda*)[282]
2. PmST1 (*PM0188*): α2,3-(O)-Sialyltransferase from *Pasteurella multocida*, recombinant (*Escherichia coli*)[283]
3. PmST3 (*PM1174*): α2,3-(O)-Sialyltransferase from *Pasteurella multocida,* recombinant (*Escherichia coli*), from Prof. X. Chen.[284]
4. Pd2,6ST: α2,6-(O)-Sialyltransferase from *Photobacterium damsela*, recombinant (*Escherichia coli*)[285]

Rat 2,3-OST is a α2,3-sialyltransferase with a reported specificity for the T-antigen (Galβ1,3-GalNAcα) structure.[282] In test-reactions with MUC1, carrying LacNAc (Galβ1,3/4-GlcNAcβ) terminated glycopeptides, no reactivity with rat 2,3-OST was observed. PmST1 and PmST3 are a α2,3-O-sialyltransferases with reported reactivity towards LacNAc and lactose substrates and good reactivity with β-galactosides in general, and were used to α2,3-sialylate type-1 and type-2 LacNAc terminated glycopeptides.[283] Pd2,6ST is a α2,6-sialyltransferase with a reported substrate specificity for β-linked Gal and GalNAc residues, but low activity on α-linked substrates.[285] It was used to α2,6-sialylated the type-1 and type-2 terminated LacNAc glycopeptides (*figure 4.44*).

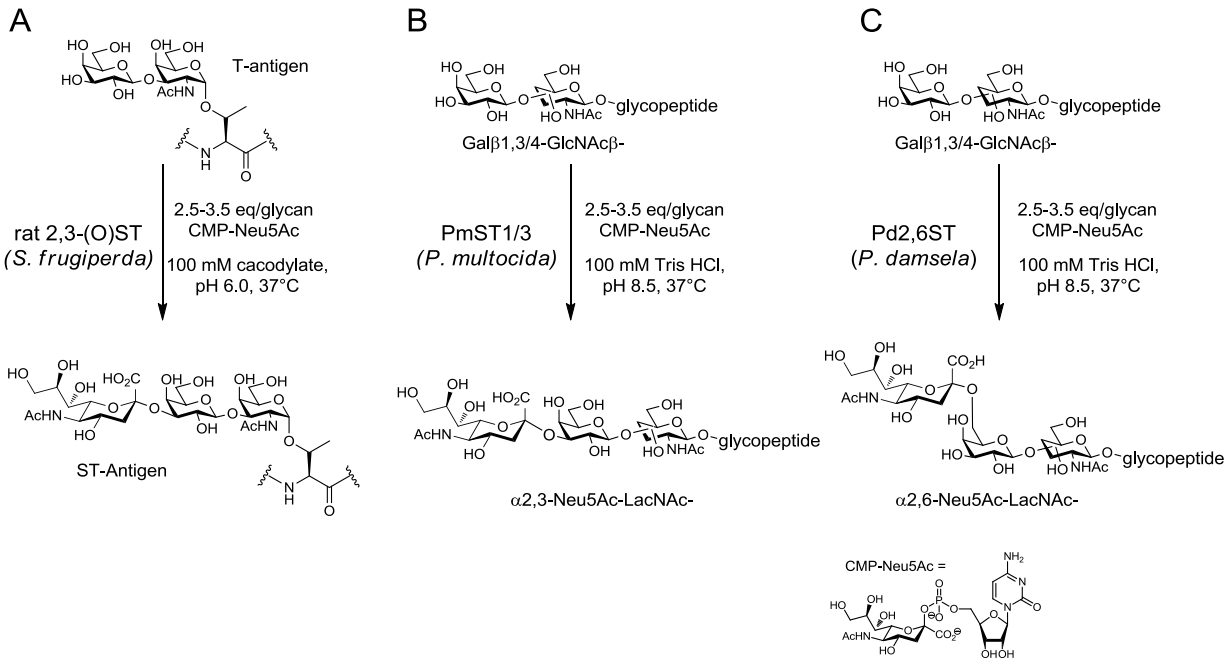


***Figure 4.44*:** Sialyltransferases used for enzymatic termination. **A**: Rat 2,3-OST for α2,3-sialylation of T-antigen. **B**: PmST1/3 for α2,3-sialylation of type-1 and type-2 LacNAc. **C**: Pd2,6ST for 2,6-sialylation of type-1 and type-2 LacNAc.

A selection of 21 MUC1 glycopeptides was chosen for in-solution enzymatic sialylation. The peptides (0.5-1.5 mg) were dissolved in an appropriate reaction buffer and CMP-Neu5Ac and the particular enzyme was added. Usually after 3-6 h an extra portion of donor sugar and enzyme was added and the reaction was left for shaking overnight at 37°C. Progress was monitored by HPLC-ESI-MS and upon complete or almost complete conversion of the glycopeptide the reaction was stopped by addition of cold acetonitrile (15-20% end-concentration). The reaction mixtures were desalted by solid phase extraction on a C18 cartridge. In order to ensure high peptide purity before immobilization on the microarray, the desired products were further purified by semi-preparative HPLC-MS. All analytical and preparative HPLC steps were performed with 0.1% of formic acid (FA) as a comparably weak acid for ion-pairing, instead of TFA. TFA did not provide significant differences in retention times for the non-sialylated starting materials and the sialylated products. For illustration, *figure 4.45* shows the enzymatic sialylation of glycopeptides **102** with Pd2,6ST (*figure 4.45*, **A**) and the preparative HPLC of the final product glycopeptide **191** (*figure 4.45*, **B**).

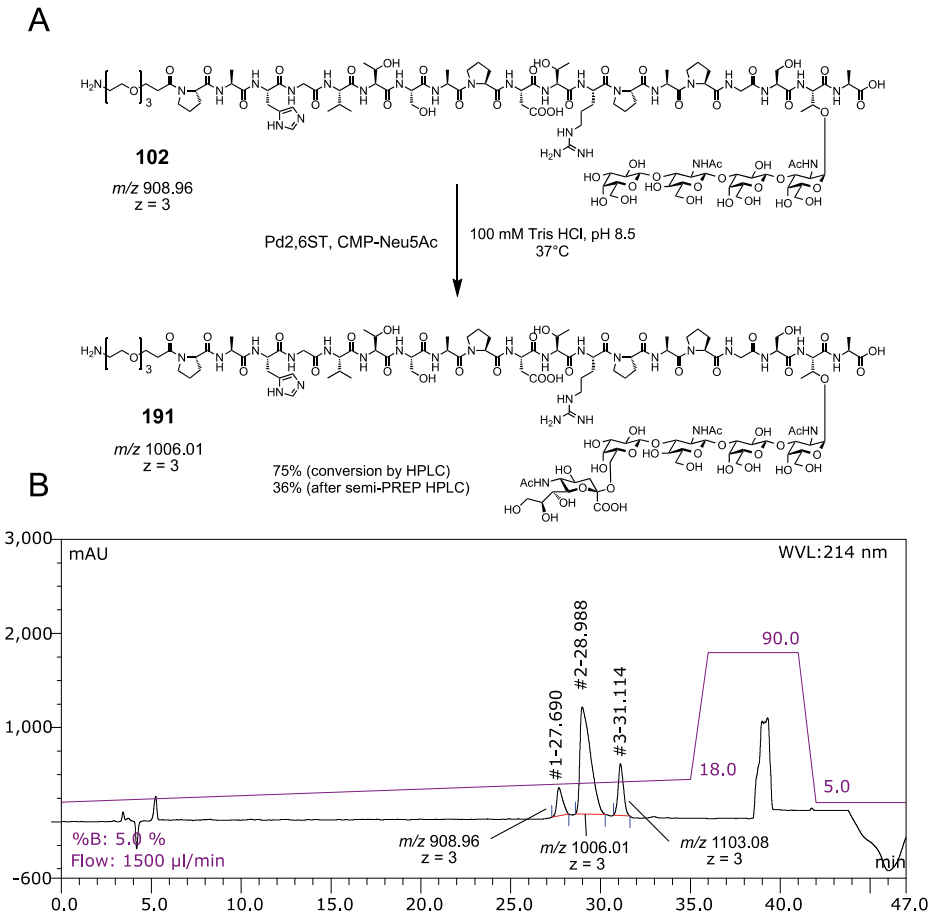


***Figure 4.45*: A:** Enzymatic α2,6-sialylation of glycopeptide **102** with Pd2,6ST to afford product **191**. **B**: Semi-preparative HPLC-MS of glycopeptide **191** with leftover starting material (*m/z* 908.96), monosialylated main product (*m/z* 1006.01) and disialylated product (*m/z* 1103.08).

Usually, treatment with Pd2,6ST resulted in pure monosialylated glycans (per equivalent LacNAc). The peptides with glycosylation in the GST*A site, also showed minor amounts of peptides with one extra Neu5Ac residue. It was recently reported that this bacterial sialyltransferase is also able to add Neu5Ac to internal galactose units of type-2 poly-LacNAc repeats.[286,287] The equivalent human enzyme, ST6Gal-I, is solely an exo-glycosyltransferase, adding a single α2,6-linked Neu5Ac only to the non-reducing end. However, here the internal Gal is not part of a poly-LacNAc chain, as reported, but of a core 1 (Galβ1,3-GalNAcα) substructure. Mass spectrometric HCD-MS2 fragmentation revealed additional sialylation on the T-antigen glycopeptide fragment ion (*m/z* 1325.6202, z=3) of the minor disialylated product, compared to the monosialylated main product (*figure 4.46*, **A** and **B**). This fragment included no information whether Neu5Ac was located on Gal or GalNAc. However, no 2,6-ST_N-antigen glycopeptide fragment ion could be observed, hinting that Neu5Ac is indeed located on Gal. Further HCD-MS3 fragmentation of ion *m/z* 1325.6202 showed ion *m/z* 454.03, representing a Neu5Ac-Gal ion (*figure 4.46*, **C**) and confirming sialylation of the

internal galactose. Additionally, HCD-MS2 of the three times sialylated core 2 structure of glycopeptide **204**, having the 6-OH position of the initial GalNAcα blocked with a lactosamine, also showed the fragment *m/z* 1325.6205 in HCD-MS2 and additionally disialylated core 1 hexasaccharide peptide *m/z* 1103.1065 (*figure 4.46*, **D**), further indicating sialylation on the galactose of the T-antigen substructure. Also, core 3 modified glycopeptides, lacking internal Gal, did not show any multiple sialylation products.

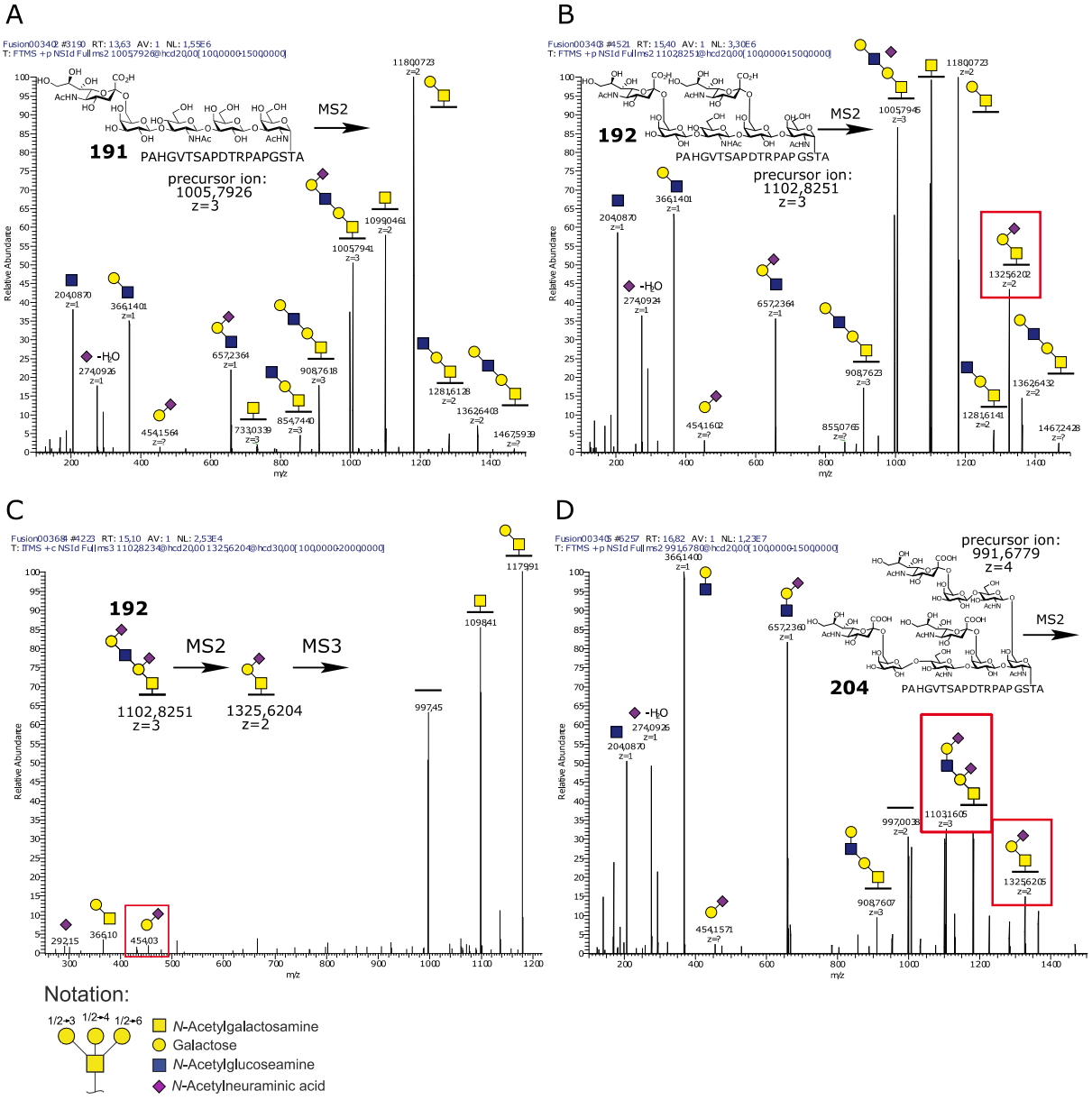


***Figure 4.46*: A:** HCD-MS2 fragmentation (20% NCE) of 2,6-monosialylated glycopeptide **191**. **B:** Disialylated product **192** with additional T-antigen sialylation (*m/z* 1325.6202). Neu5Ac with indistinct glycosylation position is shown to be “disconnected”. **C:** HCD-MS3 fragmentation (30% NCE) with Neu5AcGal oxonium ion *m/z* 454.03. **D:** HCD-MS2 fragmentation of monosaccharide glycopeptide **204** with additional T-antigen sialylation (*m/z* 1325.6202).

Minor internal sialylation by Pd2,6ST was observed on some glycopeptides with internal Gal, i.e. core 1 tetrasaccharide and core 2 hexasaccharides but only if the glycan was located in the GST*A region. Glycosylation in the VT*SA and PDT*R domains was not affected, at least not with the sialylation conditions applied here. The degree of internal sialylation is dependent on the conditions used. The applied conditions provided internal sialylation yields of 5-20%, except for glycopeptide **109** (type-2 core 1 on GST*A), yielding 50% of disialylated product (by HPLC peak integration). For three of the fourteen glycopeptides treated with Pd2,6ST, the minor by-product with a further α2,6-Neu5Ac residue was isolated (namely **192**, **198, 204,** *table 4.4*, *vide infra*), which were appended to the library.

The α2,3-sialyltransferases provided glycosylation products according to the expected specificities. Rat 2,3-OST generated exclusively sialylated T-antigen (Galβ1,3-GalNAcα) glycans. Test reactions with the T-antigen-specific rat 2,3-OST on type-1 and type-2 LacNAc extended peptides showed no activity, with the applied reaction conditions. Consequently, glycopeptide **174** with a core 2 tetrasaccharide, containing both, a type-2 LacNAc (Galβ1,4GlcNAcβ) branch and a T-antigen substructure (Galβ1,3GalNAcα), was only sialylated on the T-antigen motif to result in glycopeptide **205** (*figure 4.47*, reaction **A**). PmST3 was reported to sialylate terminal galactose and galactoside monosaccharides in general,[283] including T-antigen (Galβ1,3-GalNAcα/β) structures on smaller MUC1 peptide fragments.[288] However, no reactivity with the T-antigen MUC1 glycopeptides in test reaction was observed here. Accordingly, when employing PmST3 on core 2 glycopeptide **174**, only one equivalent of Neu5Ac was transferred to the glycopeptide resulting in glycopeptide **206** (*figure 4.47*, reaction **B**). No double sialylation was observed, although excess of CMP-Neu5Ac was applied (total 3.5 eq). It was concluded that glycopeptide **174** was selectively sialylated on the LacNAc branch.

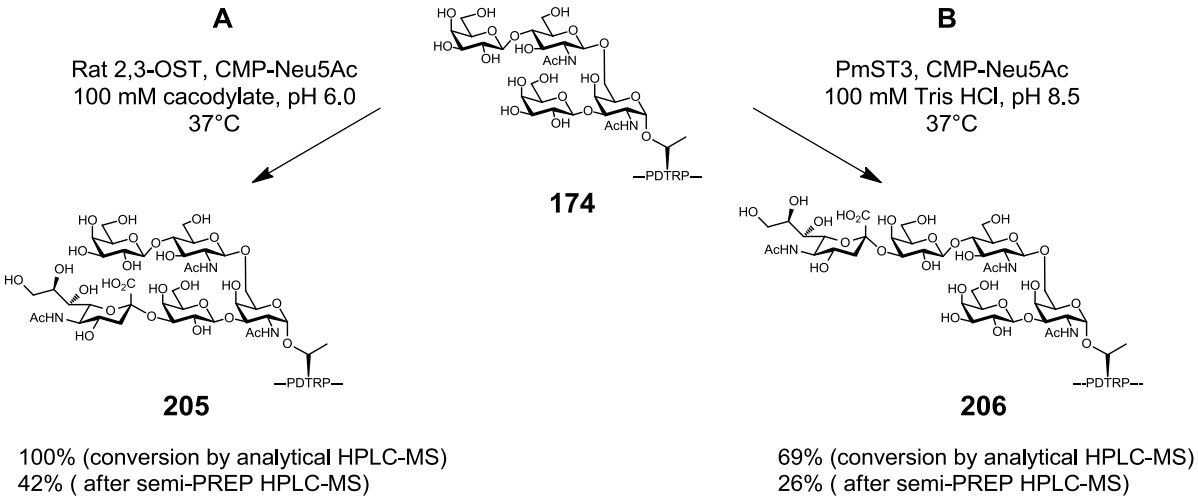


***Figure 4.47*:** Enzymatic α2,3-sialylation of glycopeptide **174** with rat 2,3-OST and PmST1 on different branches of the core 2 type-2 tetrasaccharide glycan, resulting in glycopeptides **205** and **206**.

The LacNAc sialylation on glycopeptide **206** by PmST3 was further verified by microarray screening. The binding of the LacNAc-specific lectin *Erythrina cristagalli* (ECA) was completely abolished after sialylation and sialic acid binding lectin wheat germ agglutinin (WGA) recognized only T-antigen sialylation, as in **205** (see *chapter 4.3.11.1*; *figure 4.75*, **A** and **E**). Also, mass spectrometric HCD-MS2 fragmentation of the precursor ions of **205** and **206**, revealed the ST-antigen glycopeptide fragment *m/z* 1352.6160 (z = 2) for **205** (*figure 4.48*, **A**), but not in the fragmentation spectrum of **206**, which instead showed the sialyl-LacNAc trisaccharide oxonium ion *m/z* 657.23, indicating LacNAc instead of T-antigen sialylation (*figure 4.48*, **B**).

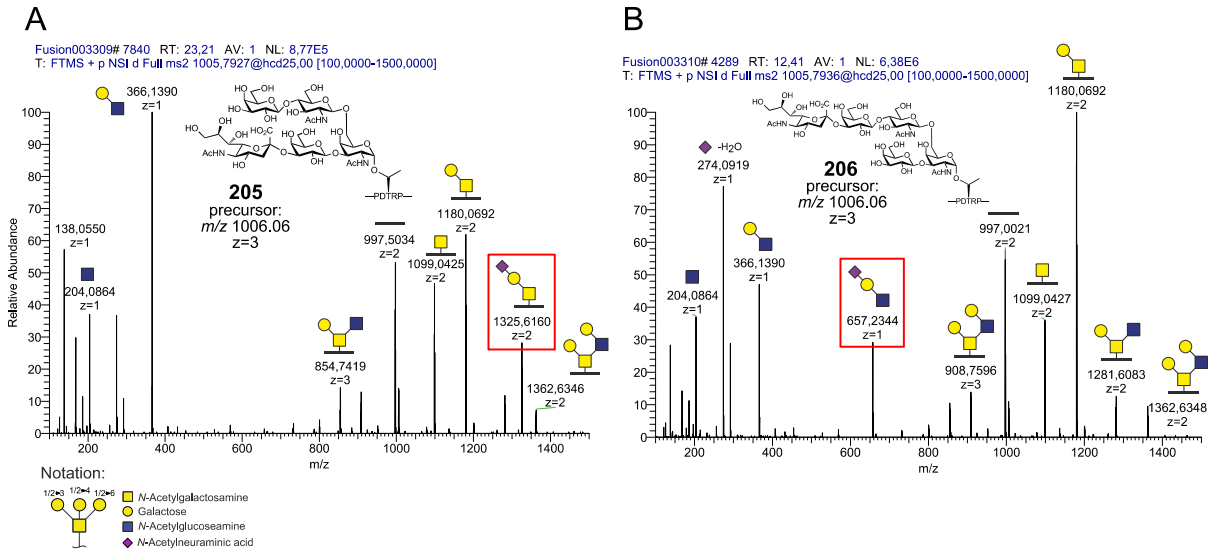


***Figure 4.48*:** HCD-MS2 fragmentation spectra of glycopeptides **205** (**A**) and **206** (**B**), both recorded at 25% NCE.

PmST3 was also not able to sialylate type-1 LacNAc containing glycopeptides, indicating low specificity for β1,3-linked disaccharides in general, in addition to the β1,3-linked T-antigen. Further sialylation reactions on type-1 LacNAc units and as well type-2 containing glycopeptides were performed using commercial PmST1, which had a similar activity towards both LacNAc structures. Significant reaction parameters for all enzymatic α2,3- and α2,6- sialylations are summarized in *table 4.4*.

***Table 4.4*:** α2,3-and α2,6-sialylated MUC1 sequences.

MUC1 Acceptor Peptide			Conditions		Product	
Glycan	#	Sequence	enzyme	reaction time (h)	#	conversion (%, by HPLC)
T	**75**	PAHGVT*SAPDTRPAPGSTA	rat2,3OST	20	**175**	97
	76	PAHGVTSAPDT*RPAPGSTA	rat2,3OST	36	**176**	82
	77	PAHGVTSAPDTRPAPGST*A	rat2,3OST	20	**177**	94
	78	PAHGVT*SAPDT*RPAPGSTA	rat2,3OST	15	**178**	89
	79	PAHGVT*SAPDTRPAPGST*A	rat2,3OST	15	**179**	100
	80	PAHGVTSAPDT*RPAPGST*A	rat2,3OST	15	**180**	100
	81	PAHGVT*SAPDT*RPAPGST*A	rat2,3OST	38	**181**	70
type-2 core 3	**92**	PAHGVT*SAPDTRPAPGSTA	PmST1	22	**182**	100
	93	PAHGVTSAPDT*RPAPGSTA	PmST1	24	**183**	88
	94	PAHGVTSAPDTRPAPGST*A	PmST1	22	**184**	94
	92	PAHGVT*SAPDTRPAPGSTA	Pd2,6ST	15	**185**	97
	93	PAHGVTSAPDT*RPAPGSTA	Pd2,6ST	24	**186**	88
	94	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	15	**187**	100
	98	PAHGVT*SAPDT*RPAPGST*A	Pd2,6ST	18	**188**	100
type-1 core 1	**101**	PAHGVTSAPDT*RPAPGSTA	PmST1	12	**189**	88
	102	PAHGVTSAPDTRPAPGST*A	PmST1	12	**190**	85
	102	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	12	**191**	75
	102	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	12	**192**	17
type-2 core 1	**107**	PAHGVT*SAPDTRPAPGSTA	PmST1	18	**193**	97
	108	PAHGVTSAPDT*RPAPGSTA	PmST1	39	**194**	100
	108	PAHGVTSAPDT*RPAPGSTA	Pd2,6ST	18	**195**	97
	109	PAHGVTSAPDTRPAPGST*A	PmST1	18	**196**	97
	109	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	12	**197**	55
	109	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	12	**198**	40
	113	PAHGVT*SAPDT*RPAPGST*A	PmST1	18	**199**	81
type-1 core 2	**119**	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	12	**200**	91
type-2 core 2 (hexa-saccharide)	**125**	PAHGVTSAPDT*RPAPGSTA	PmST1	36	**201**	66
	125	PAHGVTSAPDT*RPAPGSTA	Pd2,6ST	40	**202**	81
	126	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	12	**203**	84
	126	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	12	**204**	16
type-2 core 2 (tetra-saccharide)	**174**	PAHGVTSAPDT*RPAPGSTA	rat2,3OST	22	**205**	100
	174	PAHGVTSAPDT*RPAPGSTA	PmST3	40	**206**	69

All Neu5Ac modified glycopeptides were synthesized in 0.5-1.5 mg scale of starting peptide. In most cases the glycosylation proceeded to almost full conversion with the depicted conditions. Since high purity was desired for microarray spotting, leftover starting material and possible oversialylated byproducts were separated. The crude reaction mixture was first desalted by solid phase extraction on C18 cartridges and the main product was then isolated by semi-preparative HPLC-MS. The scale of the synthesis was small and gravimetric determination of the received reaction products was not considered to be reliable. Instead, the glycopeptides were dissolved in water and the glycopeptide concentration of an aliquot was determined by amino acid analysis (AAA). Briefly, the peptides were hydrolyzed with 6 N

HCl in the gas phase at 120°C for 22h. The obtained free amino acids were conjugated at the amino group with 6-aminoquinolyl-*N*-hydroxysuccinimidyl carbamate (AQC).[289] By HPLC with fluorescence detection, the signal intensity of each amino acid was compared with a likewise labeled norleucine internal standard. Considering the stoichiometric abundance of each amino acid, the overall peptide concentration in the glycopeptide solution was then deduced.

4.2.6 Tandem MS of sialylated mucin *O*-glycopeptides - oxonium ion pattern analysis

The analysis of protein glycosylation is a challenging task. One problem is the glycan heterogeneity at different or even the same glycosylation sites on a protein. The glycomic approach for saccharide identification and carbohydrate linkage elucidation usually involves chemical or enzymatic release from a protein, which prohibits information about glycosylation site occupation and site-specific glycan structures. Glycan removal itself can be problematic. *N*-glycan removal benefits from the existence of endoglycosydases PNGase F and A which enzymatically cleave the *N*-glycans at the linkage to the protein backbone. PNGases remove the glycan along with a change from the connecting asparagine to aspartate, indicating a former glycosylation site. In the case of *O*-glycans, no such endoglycosidases are known, except for *O*-glycanase from *Streptococcus pneumonia*, cleaving only the unmodified T-antigen. A stepwise exoglycosidase treatment must be used to digest the *O*-glycans. Chemical methods, such as β-elimination or hydrazinolysis, for the carbohydrate part and subsequent reaction of the dehydro-peptide with a *Michael*-donor can be employed for glycosylation site determination.[290] All this requires several wet lab manipulations on the precious samples. Furthermore, in order to elucidate the carbohydrate structure and linkage connectivity of the cleaved glycans, more modifications, such as glycan permethylation followed by hydrolysis, chemical reduction of the reducing end and sialic acid esterification are required, or the structure must be concluded by the digestion pattern of several glycosidases. In contrast to the glycomic approach, in a typical bottom-up proteome/glycoproteome analysis the (glyco-) proteins are digested by proteases and analyzed by mass spectrometry at the peptide level.[291,292] This approach simplifies detection of glycosylation sites and the identity of the glycosylated protein can be deduced from the obtained glycopeptides. However, analysis on the glycopeptide level still requires methodologies for linkage elucidation, determination of anomeric configuration and discrimination of isobaric saccharides. A promising method reported lately for the use on glycoconjugates, is ion mobility mass spectrometry (IM-MS). Analytes are separated by

differences of their rotationally averaged collisional cross-sections (CSS) during MS or after MS/MS. It has been reported that by this method it is possible to distinguish isomeric oligosaccharides and saccharides fragments from glycopeptides,[293,294] different glycosylation sites on otherwise identical glycopeptides,[295] epimeric methylglycosides[296] and may therefore be used in carbohydrate sequencing.

Another way to explore the structure of saccharide epimers and their connectivity in complex glycans attached to glycopeptides, is the analysis of the oxonium ion fragmentation patterns in tandem MS. MS/MS-fragmentation by collisional induced dissociation (CID) or higher-energy C-trap dissociation (HCD) allows stepwise detachment of the monosaccharides. HCD provides the advantage of simultaneous fragmentation of the peptide backbone and the carbohydrates, allowing analysis of glycopeptides in a single MS/MS run. Also high resolution MS2 spectra can be recorded.[297] Fragmentation of glycans, results predominantly in isobaric hexose (Gal, Man, Glc) and isobaric hexosamine (GalNAc, GlcNAc) oxonium ion fragments. The relative intensities of the isobaric ion fragments derived from the HexNAc and also Neu5Ac units, can be used as reporter ions in CID or HCD fragmentation (*table 4.5*). Comparison of the relative ion intensities gives then glycan structure specific fragment patterns.

***Table 4.5*:** Characteristic HexNAc and Neu5Ac oxonium ions in tandem MS experiments.

Fragment Ion	Monoisotopic Mass	Composition
126	126.055	$[HexNAc - C_2H_6O_3]^+$
138	135.055	$[HexNAc - CH_6O_3]^+$
144	144.065	$[HexNAc - C_2H_4O_2]^+$
168	168.066	$[HexNAc - 2H_2O]^+$
186	186.076	$[HexNAc - H_2O]^+$
204	204.087	$HexNAc^+$
274	274.092	$[Neu5Ac - H_2O]^+$
292	292.103	$Neu5Ac^+$
366	366.140	$HexHexNAc^+$
454	454.156	$Neu5AcHex^+$
657	657.235	$Neu5AcHexHexNAc^+$

For example, in cooperation with *J. Nilsson* (University of Gothenburg) it was demonstrated that isobaric αGalNAc and βGlcNAc glycan decoration can be distinguished by the HexNAc oxonium ion MS fragmentation pattern.[298] MS2 fragmentation of a MUC1 peptide with αGalNAc (T_N-antigen) glycosylation shows equally high intensities of fragments *m/z* 138, 144, 186 and 204. However, a βGlcNAc modification at the same glycosylation site shows strong intensities for 138 and low intensities for *m/z* 144 and *m/z* 186. The two epimeric $HexNAc^+$ oxonium ions of $GalNAc^+$ and $GlcNAc^+$, differing only in the configuration at C4,

have different fragmentation behaviors. This can be exploited by calculating the GlcNAc/GalNAc-ratio by the relative intensity ratio of *m/z* 138 + *m/z* 168 (GlcNAc) to *m/z* 126 + *m/z* 144 (GalNAc). Based on the developed method of calculating the GlcNAc/GalNAc-ratio, novel GlcNAcylated peptides, which were further extended on the GlcNAc unit while passing through the Golgi complex and loaded onto HLA I molecules, as well as normal mucin type GalNAc glycopeptides were recently identified.[299] In collaboration with *J. Nilsson*, the type-1 and type-2 core 3 MUC1 glycopeptides **85**-**87** and **92**-**94**, synthesized in this work, were also fragmented by HCD. The oxonium ion profile significantly differed between the type-1 and the type-2 LacNAc elongated glycans. In type-1 glycans strong intensities of *m/z* 204 and moderate intensities of *m/z* 186 were observed, while type-2 glycans resulted in low intensities of *m/z* 204 and *m/z* 186 ions. Like this, the connectivity of the common 1,3- and 1,4-linked LacNAc disaccharide units can be deduced on the level of intact glycopeptides.

In analogy to the experiments performed in collaboration with *J. Nilsson et al.*, the sialylated glycopeptides, synthesized with PmST1/3 and Pd2,6ST were analyzed, whether the terminal sialylation influences oxonium ion profiling of glycan LacNAc elongation. Thus, selected sialylated glycopeptides with α2,3- and α2,6-linked Neu5Ac on either type-1 or type-2 disaccharide units were subjected to HCD-MS2 fragmentation in an O*rbitrap Fusion* mass spectrometer. The normalized collision energy (NCE) was increased in 5% increments and relative intensities of relevant oxonium ions were recorded (*figure 4.49*, for fragment ion composition see *table 4.5*).

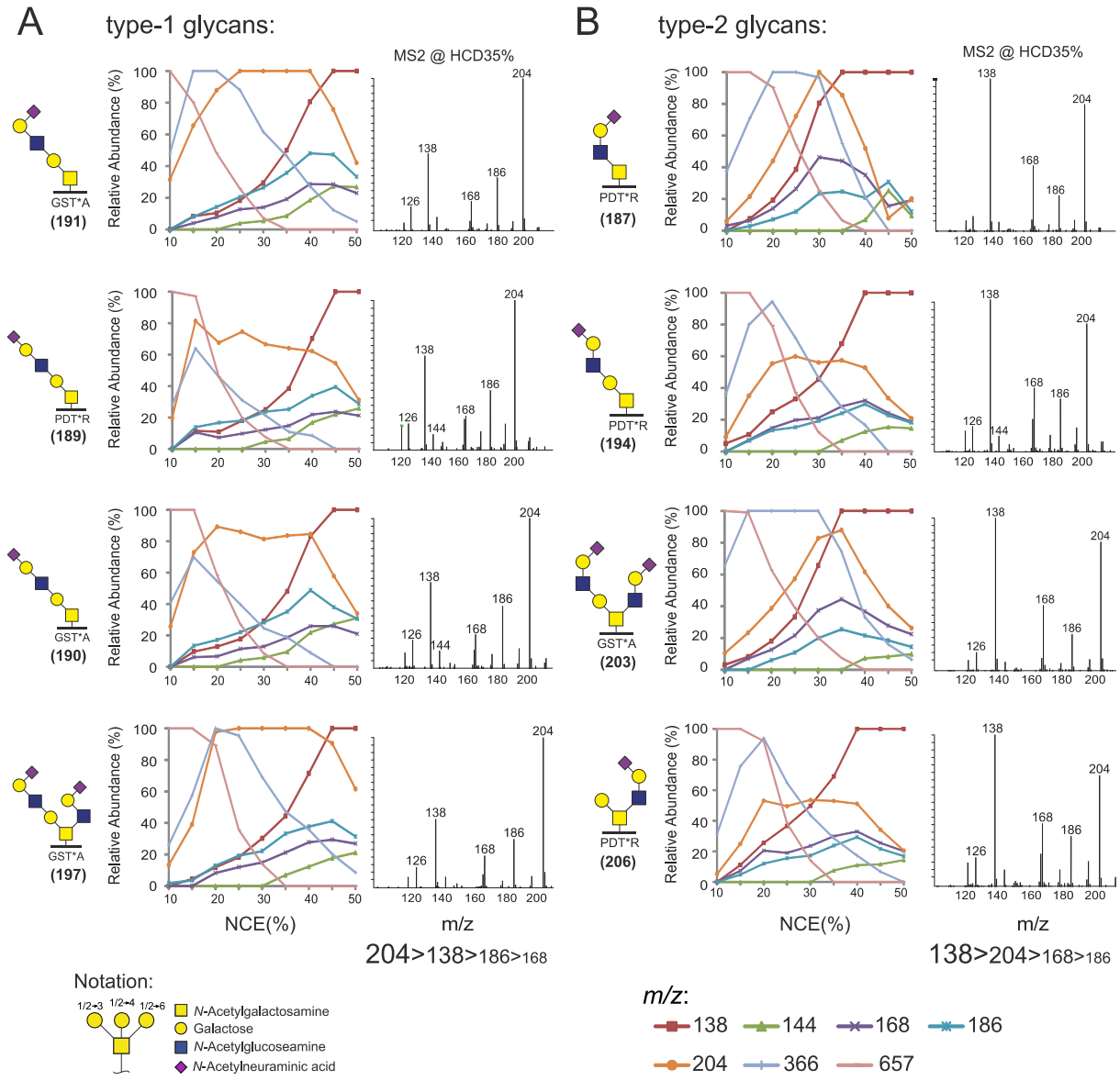


***Figure 4.49*:** Oxonium Ion profiles and spectra of various mucin core glycopeptides terminated with sialic acid. Oxonium ions were generated by HCD-MS2 in 5% increments of NCE from 10-50% on an *Orbitrap Fusion* mass spectrometer. Profiles show relative ion intensities at different NCE levels and spectra show *m/z* 100-220 at 35% NCE. **A:** $HexNAc^+$ fragmentation of type-1 glycans. **B**: $HexNAc^+$ fragmentation of type-1 glycans.

Type-1 disaccharide glycopeptides (**191**, **189**, **190**, **197**, *figure 4.49*, **A**) showed oxonium ion profiles with high intensities of fragment *m/z* 204 relative to *m/z* 138 at NCEs of 30-35%. Fragment *m/z* 138 surpassed *m/z* 204 only in the highest energy settings of 45-50%. The MS2 spectra from *m/z* 100-220 at 35% NCE showed the uniform ion patterns with intensities in the order *m/z* 204>138>186>168 for all type-1 elongated glycopeptides. In contrast, the type-2 elongated glycopeptides (**187**, **194**, **203**, **206,** *figure 4.49*, **B**) had similar intensities of ions *m/z* 138 and 204 in the area of 30-35% NCE. All spectra of the type-2 glycans at 35% NCE had again uniform oxonium ion levels in the order *m/z* 138>204>168>186, which was distinctly different to the type-1 glycopeptides. At NCE levels of 40%, type-1 glycopeptides

had approximately equal levels of *m/z* 138 and *m/z* 204, while in type-2 glycopeptides *m/z* 204 was significantly lower than *m/z* 138. This is in accordance with the results reported for the not sialylated type-1 and type-2 glycopeptides.[298] It was previously proposed that the lower levels of *m/z* 204 in type-2 glycans compared to type-1 could correspond to a higher acidity of H-5 in the GlcNA-ring, which is eliminated from the *m/z* 366 and possibly taken up by the galactose, leaving as neutral loss.[300,301] Therefore, generation of *m/z* 204 from type-2 $HexHexNAc^+$ oxonium ions is to a large extend skipped in favor of direct formation of *m/z* 186, which then loses another equivalent of water in order to aromatize to *m/z* 168 and finally gives away CH_2O to result in *m/z* 138 (*figure 4.50*).

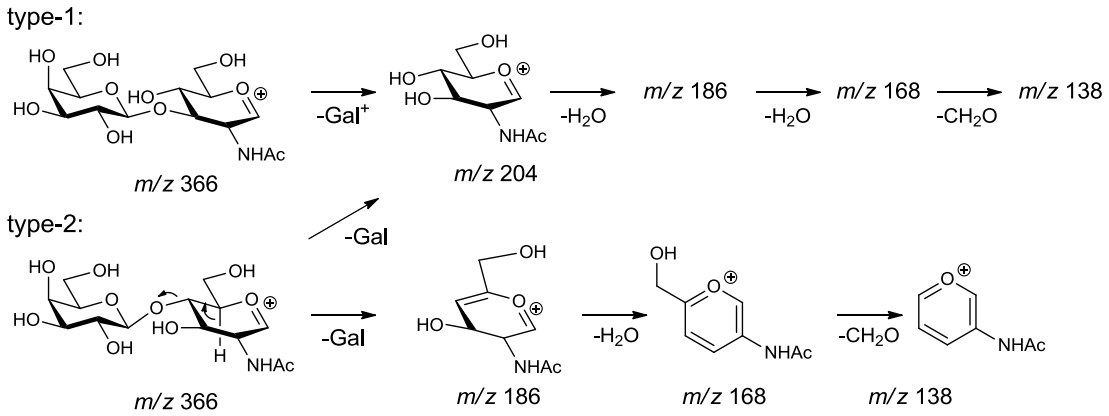


***Figure 4.50*:** Possible fragmentation mechanism of type-1 and type-2 derived oxonium ions.[300,301]

Compared to the already published results with *J. Nilsson et al.*, the oxonium ion profiles slightly differ in normalized collision energy levels. Such instrument specific variations were found when different mass spectrometers were used for HCD. In the already reported previous cases a *Q-Exactive* and a *LTQ-Orbitrap Velos*, both *Thermo Fisher Scientific*, were used. The instrument used for the here described measurements was an *Orbitrap Fusion*, also *Thermo Fisher Scientific*. However, the overall trends and relative intensities of the oxonium ions originating from the LacNAc disaccharides are according to the published results,[298] although the LacNAc structures are now further terminated by a sialic acid. Also, different glycosylation sites (here PDT*R or GST*A), the core structures (core 1, core 2, core 3) and the number of LacNAc units (core 2 → 2 eq LacNAc and core 1 or 3 → 1 equiv LacNAc) seemed to have no effects on these fragmentation differences. Some influence of of sialic acid attachment on the relative ion intensities of the $GlcNAc^+$ derived oxonium ions could be monitored. Apparently, all glycopeptides with α2,6-sialylation showed elevated levels of *m/z* 204 in relation to *m/z* 138 (e.g. **187** and **203** vs **194** and **206,** *figure 4.49*, **B**). It was assumed that sialylation on either 3- or 6-position of Gal differently influences the fragmentation behavior of $Neu5AcHexHexNAc^+$ ion *m/z* 657, which probably alters aforementioned fragmentation mechanism of ion *m/z* 366 (*figure 4.49*). Exact elucidation of

the fragmentation mechanisms is a matter of further research. However, conjugation with α2,3- and α2,6-sialic acid does not prevent discrimination of the type-1 and type-2 disaccharide structures by altered $GlcNAc^+$ oxonium ion pattern profiling.

Due to the modification with Neu5Ac, further characteristics of relative oxonium ion intensities were monitored in the area of *m/z* 250-660.

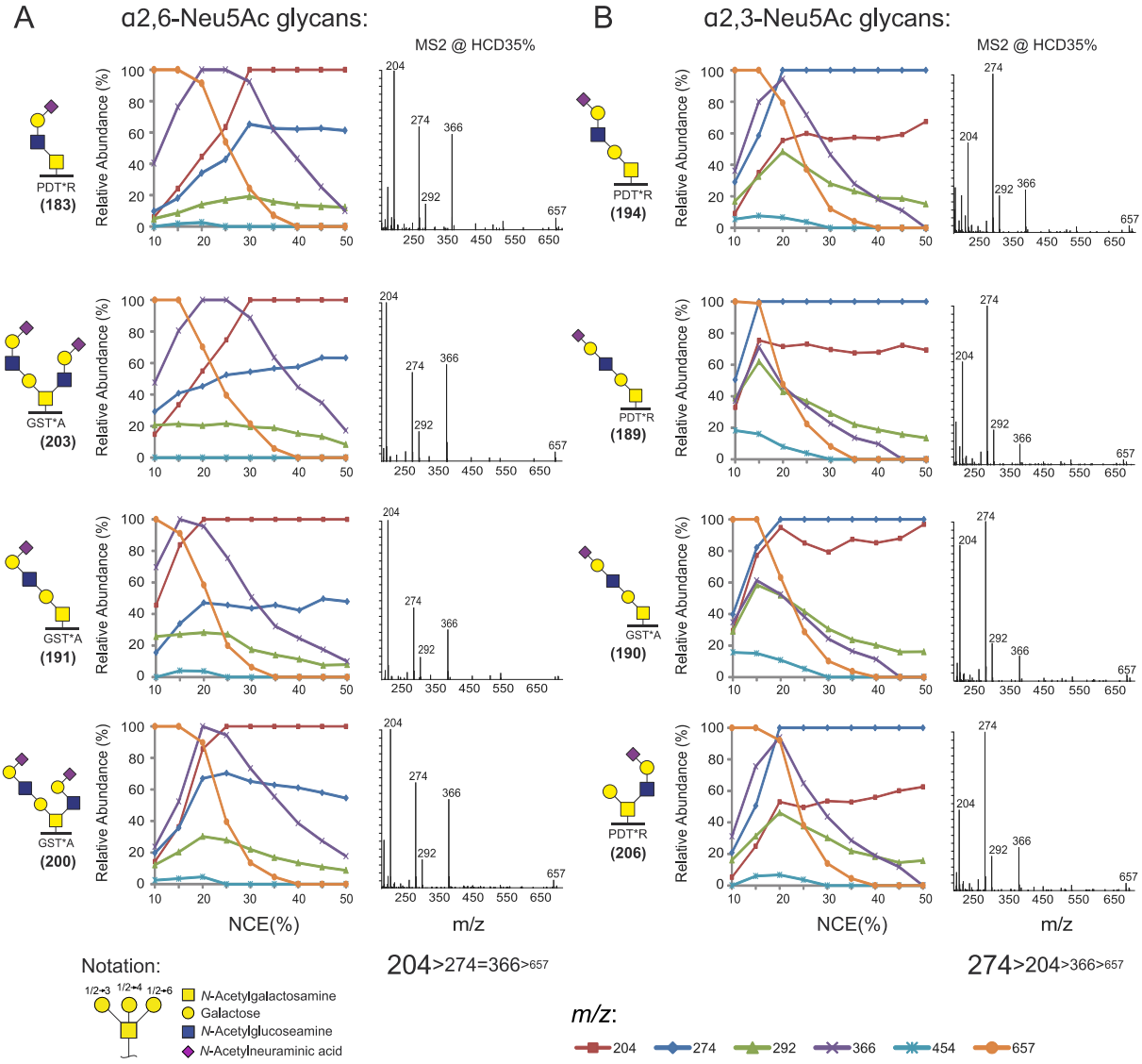


***Figure 4.51*:** Oxonium-ion profiles and spectra of various mucin core glycopeptides terminated with sialic acid. Profiles show relative ion intensities at different NCE levels and spectra show *m/z* 200-660 at 35% NCE. **A:** HexNAc and Neu5Ac fragmentation of α2,6-Neu5Ac glycans. **B**: HexNAc and Neu5Ac fragmentation of α2,6-Neu5Ac glycans.

In general, Neu5Ac was represented strongest by oxonium ion *m/z* 274 (*figure 4.51*), which corresponds to $Neu5Ac^+$ with loss of one equivalent of water (*table 4.5*). The $Neu5AcHex^+$ ion *m/z* 454 was seen in small intensities and the oxonium ion *m/z* 657 of the trisaccharide

Neu5AcHexHexNAc$^+$ ion was dominating the profile for low collisional energies (10-20%), but decays swiftly with increasing energy levels. α2,6-Neu5Ac modified glycopeptides showed a strong plateau of the HexNAc$^+$ ion *m/z* 204 at NCEs of 20-30% (*figure 4.51*, **A**). The Neu5Ac derived ion *m/z* 274 showed weaker intensities relative to *m/z* 204. Oxonium ion signal *m/z* 366 was high in relative intensities at NCEs around 20%. The spectra of all α2,6-Neu5Ac terminated glycans at 35% NCE showed similar fragment profiles with signal intensities in the order *m/z* 204 > 274 = 366 > 657. In contrast, α2,3-Neu5Ac glycosylated peptides showed Neu5Ac derived ion *m/z* 274 dominating the profile with highest ion intensities for NCEs ≥20% and HexNAc$^+$ ion *m/z* 204 being relatively weaker than *m/z* 274 (*figure 4.51*, **B**). Oxonium ion *m/z* 366 of HexHexNAc$^+$ was also weaker compared to *m/z* 274. The overall ionization pattern for α2,3-Neu5Ac modified glycopeptides at 35% NCE was found to be *m/z* 274>204>366>657. In fact, the relative intensity profiles for *m/z* 204 (HexNAc) and *m/z* 274 (Neu5Ac) showed reversed behavior for α2,6- and α2,3-Neu5Ac modified glycopeptides.

To conclude, in addition to oxonium ion discrimination of β1,3- and β1,4-linked LacNAc disaccharide glycans, assignment of the connectivity of the terminal sialylation was possible by comparing also the relative Neu5Ac derived oxonium ions. The two different types of terminal sialylation had an influence on the fragmentation behavior of HexHexNAc ion *m/z* 366. In case of α2,6-Neu5Ac glycopeptides, fragments of β1,3/4-LacNAc appear dominantly, namely *m/z* 204 and *m/z* 366. In the case of α2,3-Neu5Ac, the sialic acid derived ion *m/z* 274 had the strongest intensities. The mechanism by which the two kinds of sialic acids are fragmented are apparently different, generating higher levels of Neu5Ac oxonium ions in the case of α2,3-Neu5Ac glycosylation. Also, in the case of α2,3-Neu5Ac modification, higher intensities of the in general low abundant Neu5AcHex$^+$ *m/z* 454 were present, especially in combination with type-1 modified glycopeptides (peptides **189** and **190**, *figure 4.51*, **B**). This could be related to the lower intensities of HexHexNAc$^+$ ion *m/z* 366.

The demonstrated examples carried different mucin type core structures, which all have in common that the first monosaccharide attached to the peptide backbone is a αGalNAc. In the aforementioned fragmentation experiments of the sialylated glycopeptides, all HexNAc$^+$ oxonium ions m/z 204 were related to the GlcNAc in the LacNAc structures, ignoring the present GalNAc. By CID-MS3 experiments, *J. Nilsson et al.* demonstrated that the HexNAc$^+$ oxonium ions used for profiling, are merely derived from the terminal βGlcNAc of the elongating chains and not from the internal αGalNAc.[298] Similar results were observed by HCD-MS2 fragmentation. Additionally to these previous experiments, the same conclusion could be obtained here, by comparing core 1 (T-antigen) and branched core 2 MS2 fragmentation profiles (*figure 4.52*).

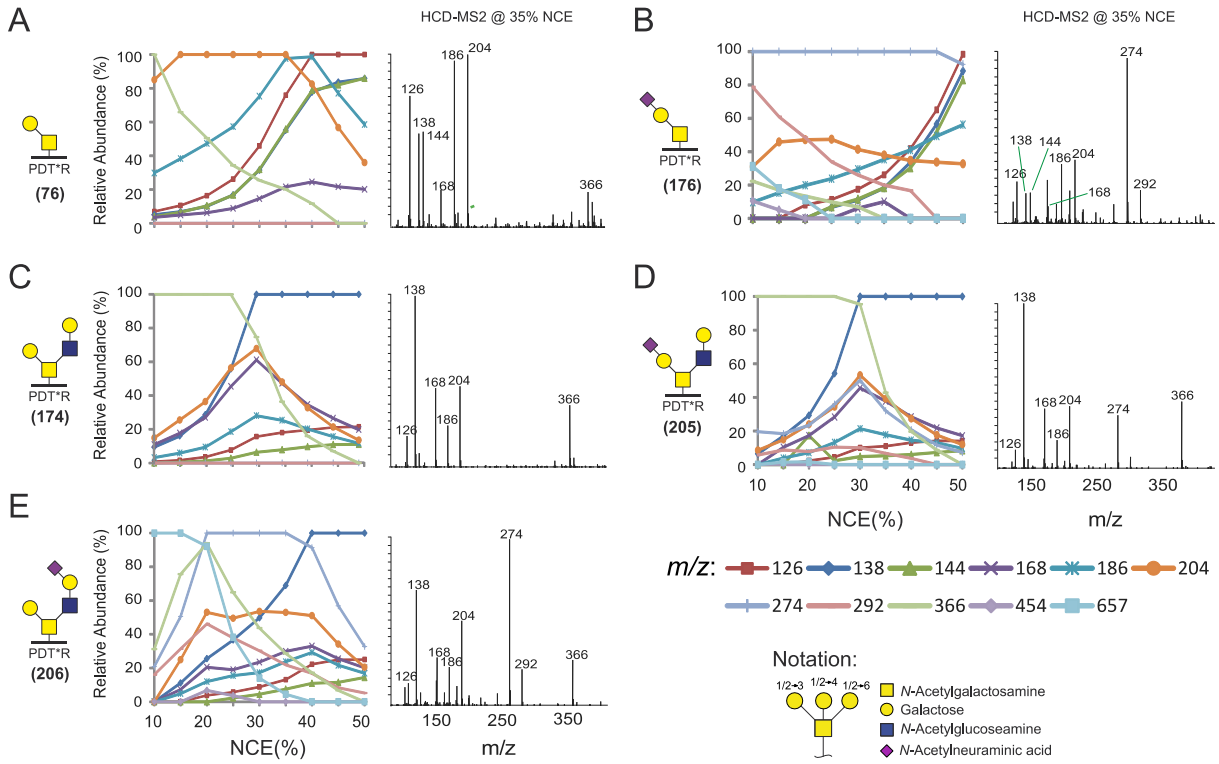


***Figure 4.52*:** Comparison of GalβGlcNAc- and GalβGalNAc-fragmentation. Profiles show relative ion intensities at different NCE levels and spectra show *m/z* 100-400 at 35% NCE. **A:** Core 1 (T-antigen) glycopeptide **76**. **B:** Core 1 (ST-antigen) glycopeptides **176**. **C**: Core 2 glycopeptide **174**. **D:** Core 2 glycopeptide **205** with "T-antigen-like" sialylation. **E**: Core 2 glycopeptide **206** with LacNAc-sialylation.

Glycopeptide **76** with T-antigen glycosylation showed a typical αGalNAc derived oxonium ion profile (*figure 4.52*, **A**).[298] Fragmentation of the sialyl-T (ST-antigen) glycopeptide **176** (*figure 4.52*, **B**) further produced Neu5Ac derived reporter ions, for relative comparison with the HexNAc (=GalNAc) derived oxonium ions. Fragmentation of the ST-antigen of glycopeptide **176** revealed oxonium ion *m/z* 274 ($[Neu5Ac - H_2O]^+$) as the most abundant ion in the profile over all collision energy settings, completely overwhelming the intensity of the αGalNAc ion profile, compared to not sialylated glycopeptide **76**. The oxonium ion pattern was clearly dominated by the Neu5Ac fragmentation. The $Neu5AcHexHexNAc^+$ ion *m/z* 657 and $HexHexNAc^+$ ion *m/z* 366 were only weak, compared to *m/z* 274. On the other hand, if a LacNAc is attached to the T-antigen in 6-position of the αGalNAc, like in the core 2 glycopeptide **205** (*figure 4.52*, **D**), the same Neu5Ac ion signal was itself overwhelmed by peak *m/z* 138, which must therefore mainly derive from the branched βGlcNAc. This demonstrates that, HexNAc fragmentation was almost independent from the internal αGalNAc. Further, the Galβ1,4GlcNAc ion *m/z* 366 was present in significant amounts for 10-35% NCE. The unsialylated core 2 glycopeptide equivalent **174** (*figure 4.52*, **C**) was virtually fragmenting according to **205** (*figure 4.52*, **D**), except without the Neu5Ac reporter ions *m/z* 274 and *m/z* 292. In contrast to **205** (*figure 4.52*, **D**), sialylated at the inner core

Galβ1,3GalNAc, the ion profile of the core 2 modified glycopeptide **206** (*figure 4.52*, **E**), sialylated at the LacNAc branch instead, showed strong Neu5Ac reporter ions *m/z* 274 and *m/z* 292. Compared to the ST-antigen peptide **176** (*figure 4.52*, **B**), glycopeptide **206** (*figure 4.52*, **E**) additionally showed *m/z* 366 and *m/z* 657 ions, which were absent or low in ST-antigen peptide **176** (*figure 4.52*, **B**). The low abundance of these di- and trisaccharides in **176** (*figure 4.52*, **B**) and their high profile in **205** (*figure 4.52*, **D**) and **206** (*figure 4.52*, **E**), can probably be explained by the higher stability of the αGalNAc (α-anomers are thermodynamically favored), directly connected to the peptide backbone.

To conclude, with the two common forms of sialylation (α2,3- and α2,6-Neu5Ac), terminating type-1 or type-2 LacNAc disaccharides, four isobaric trisaccharide units can be formed. The oxonium ion pattern analysis of the synthesized sialylated *O*-glycopeptides revealed that it is possible to distinguish between these four isobaric structures by the abundance of the HexNAc- and Neu5Ac-derived oxonium ions, in ion-trap HCD tandem MS experiments. The analysis on the tested MUC1 glycopeptides is independent from glycosylation site or core structure. Further does the initial αGalNAc not hamper with the GlcNAc fragmentation of the type-1 and -2 disaccharides. The isobaric Neu5AcHexHexNAc trisaccharides are common terminating sugar units, not only on mucin type *O*-glycosylation but also on α-*O*-mannosylated proteins, α-*O*-fucosylated proteins and of course on *N*-glycans of the complex and hybrid type. Lately, complex mucin type glycosylation was discovered on tyrosine residues on amyloid precursor proteins,[15] and afterwards on several (but still few in total numbers) other proteins, e.g. by *Simple Cell* technology.[16,302] This still unexplored type of glycosylation was shown to develop complex *O*-glycan motifs,[15] but tyrosine *O*-glycosylation is not amenable for chemical glycan cleavage by β-elimination and relies on glycoproteomic elucidation on the glycopeptide level. Discrimination of sialic acid linkage connectivities is commonly done by exoglycosidase digest and intermediate MS control. For example, α2,3-sialidase digest may be applied and if a α2,3-connected sialic acid is present, the corresponding mass reduction will be observed. This is laborious and in the case of β1,3/4-galactosidases, no such enzymatic discrimination is possible and requires even more laborious glycan removal, carbohydrate permethylation and hydrolysis steps combined with MS control. Comparison of relative oxonium ion intensities might facilitate glycan sequencing directly on the glycopeptide level. On the glycopeptide level, the information of glycosylation site occupancy and the identity of the correlating glycoprotein are retained. Therefore, knowledge about the fragmentation behavior of common glycan structures may help to elucidate the glycoproteome. This further underlines the importance of chemical and chemoenzymatic glycan synthesis strategies, for generation of defined glycan/glycopeptide reference molecules.

4.3 Part 3 – Microarray studies with glycopeptide binding proteins

4.3.1 Introduction on microarray development

Microarray based technologies were first applied in analysis of DNA-arrays as tools in genomic and transcriptomic research.[303] With the advent of high throughput molecular biology, the increasing interest in expression studies and whole genome sequencing, the ability of performing hundreds or thousands of reactions to study binding events in miniaturized experiments matched with the biological needs of the time. Oligonucleotide synthesis and polymerase chain reaction (PCR) technology enables scientists to build-up polynucleotide libraries and microarrays represented a methodology for fast screening of many samples in parallel with consumption of only minute amounts. In genetic analysis, microarrays are used for sequencing, to detect single nucleotide polymorphisms (SNPs) or mutations or for expression analysis.[304] In a typical expression analysis, messenger RNA (mRNA) strands are isolated and transferred to complementary DNA (cDNA) by the action of reverse transcriptase. The amplified product is fluorescently labeled and screened against an immobilized library of single stranded DNA. Hybridization reveals the existence and relative amount of a transcript, which allows drawing conclusions about the expression level of a certain protein. However, mRNA levels, protein expression and protein interactions do not always correlate. Protein splicing and post-translational modifications often change the protein functionality and dictate behavior in cellular pathways or interaction with binding partners, which cannot be explained solely by evaluation of protein expression. After sequencing the human genome and that of other organisms, the function and abundance of the expression products has attracted more attention. Thus, the development of protein microarray platforms for interaction studies was the logical continuation.[305] While DNA/RNA substrates are uniform, negatively charged and hydrophilic molecules, build from only four different nucleotides, proteins on the other hand consist of up to twenty amino acids creating macromolecules with varying hydrophilicity/hydrophobicity and differently folded conformations, that need to be conserved during microarray experiments. Protein microarrays are employed in studies of protein/protein-, antigen/antibody-, enzyme/substrate-, protein/DNA- and ligand/receptor-interactions and offer a high throughput alternative to commonly used enzyme-linked immunosorbent assays (ELISA) or western blots. The great versatility of protein arrays makes them useful as tools in clinical biomarker detection, drug discovery or elucidation of proteomic interaction networks.[306] Consequently, microarray technology was also transferred to carbohydrate analytics shortly after the first

protein applications. Like protein microarrays, carbohydrate microarrays benefit even more from the multivalent presentation of densely immobilized test molecules, since most carbohydrate binding proteins, like lectins or antibodies, are dependent on multivalent recognition for strong binding.[307] The *Consortium for Functional Glycomics* (CFG) is an international research initiative, providing the research community with a comprehensive collection of glycans on a glycoarray platform. The current array version (v5.2) contains 609 carbohydrate structures. Mentionable in this context, are the efforts of high throughput characterization of the receptor specificities of hemagglutinins from pandemic influenza viruses, for example from modern H5N1 or the 1918 H1N1 virus heamagglutinins.[308,309] Like this, mutations in the virus genome can be directly correlated with changes in binding preferences, enabling for example avian specific viruses to become infectious for human hosts. Carbohydrates can also be arrayed as glycoconjugates. Glycopeptide arrays offer the possibility to evaluate carbohydrate-protein binding events with carbohydrates presented on a natural scaffold.[310] The peptide backbone is sometimes included in part or completely in the binding epitope, whose binding is either modulated or prevented by the glycan.[311]

4.3.2 Immobilization methods for glycopeptide microarrays

Microarrays have been established on spherical polymer beads, gold surfaces, nitrocellulose membranes or coated glass slides, with the latter being the most used surface when applied with fluorescence detection. The functional groups of the surface coating determine the kind of connection established and must fit the criteria of the probes that are supposed to be immobilized. The binding methodology also determines whether a molecule is attached to the surface in a random orientation or oriented by site-specific interactions.[312,313] Methods for random orientation in protein immobilization often involve direct binding to the surface by amino acids residues. Lysine residues and the protein N-terminus can be linked to aldehyde, epoxy, isothiocyanate or *N*-hydroxysuccinimide (NHS) ester modified surfaces (*figure 4.53*, **A**). Cysteine residues can connect via a thioether or disulfide bond (*figure 4.53*, **B**). Alternatively, cysteines can adhere non-covalently to gold surfaces. The drawback of non-site-specific attachment is that the active site of proteins might on part be covered and activity is lost. Installation of remote reaction sites in the proteins, distant from the active site, helps to preserves activity. Proteins can be expressed with affinity tags for non-covalent and site-directed adhesion, i.e. hexahistidine (His_6-tag) sequences for nitrilotriacetic acid (NTA) surfaces or glutathione-S-transferase (GST-tag) for glutathione modified surfaces. Proteins can also be expressed as intein-fusion-proteins and then conjugated site-specifically with a biotinylated peptide by expressed protein ligation (EPL). Alternatively, bio-orthogonal chemical groups may be inserted into the probe, for instance azides, alkines,

cyclopentadienes, to feature site specific reactions, such as 1,3-dipolar cycloadditions, *Staudinger* ligations or *Diels-Alder* reactions (*figure 4.53*, **C**). Like this, probes will be linked with a uniform orientation.

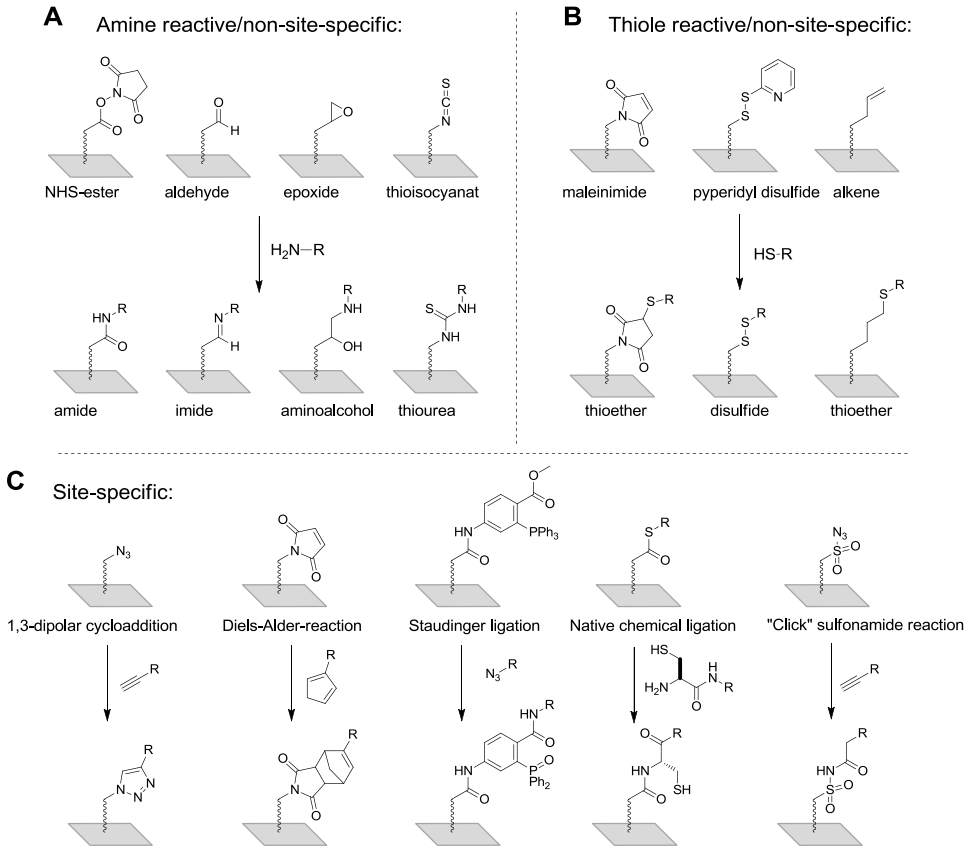


***Figure 4.53*:** Overview of common covalent non-site-(**A** and **B**) and site-specific (**C**) immobilization strategies for microarray application.

Site-directed and covalent immobilization techniques are often advantageous for protein activity.[312] Especially the selective, covalent immobilization of unpurified expressed proteins directly from cell lysates represents a highly efficient method of protein microarray fabrication. The group of *H. Waldmann* presented the immobilization of oxyamine-modified proteins from cell lysates.[314] The proteins of interest were overexpressed as fusion proteins with intein domains in *E.coli*. After cell lysis, the lysate was treated with 2-mercaptoethanesulfonate (MESNA) to generate a C-terminal thioesters on the expressed proteins, which was then transferred into the oxyamine by addition of bis(oxyamine)ethane. The high reactivity of the C-terminal oxyamine was used to form a stable oxime bond with a ketone functionalized surface. In another approach, proteins were expressed with a “CaaX-box”, a C-terminal tetrapeptide sequence which is recognized by protein farnesyltransferases. Farnesylation of the target proteins occurred *in vivo* when the protein expressing *E.coli* were additionally transfected with a farnesyltransferase encoding vector.

After cell lysis, the lysate was incubated on a thiol-functionalized microarray slide and irradiated with UV-light. This initiated a thiol-ene reaction with the olefinic isoprenoid groups and selectively attached the lipidated proteins to the surface.[315]

Commercial microarray slides are available for all kinds of application featuring various chemical immobilization strategies. *Nexterion® slide H* microarray slides (*Schott GmbH,* Mainz, Germany) are used in this work. The hydrogel coated slides have been used successfully for glycan and glycopeptide microarray serum analysis before.[59,60,310,316] The glass slides are coated with a cross-linked hydrophilic polymer. Hydrophilic spacers are crafted onto the polymer network with amine reactive NHS-esters at the ends (*figure 4.54*).

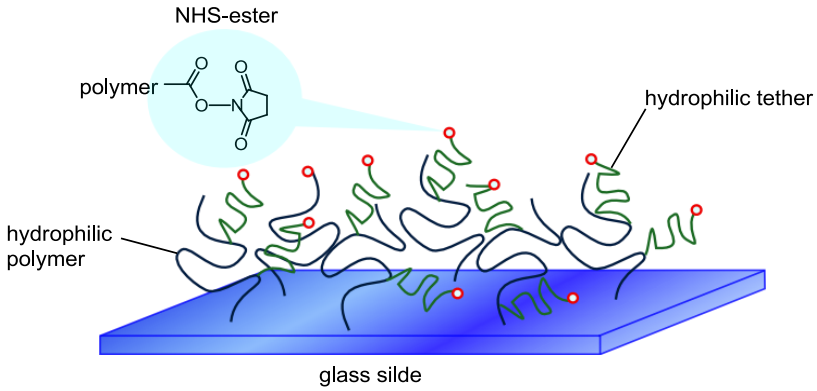


***Figure 4.54*:** Schematic representation of *Nexterion® slide H* composition.

The polymer and the spacer are supposed to generate spatial distance between the immobilized probes and the samples screened on the microarray. Sensible protein probes or samples may denature due to contact with the surface. Also, the hydrophilic network mimics a “in-solution” environment, supposed to support probe accessibility and maintain protein specificity and conformation. The absence of any free amine residue, except for the primary amine of the spacer amino acid at the N-terminal end of the MUC1 and MUC5B tandem repeat peptides, results in N-terminal attachment only and uniform orientation of the glycopeptides.

4.3.3 Principle of serum screening by glycopeptide microarrays

Glycopeptides were delivered to the NHS-functionalized microarray slides by piezo driven non-contact spotting. The slides were left for incubation at high humidity (70-100%), to complete the covalent coupling between the surface NHS-esters and the N-terminal amine of the glycopeptides. Unreacted NHS-ester groups were quenched with ethanolamine. A silicone superstructure was used to form different reaction wells. The antiserum was incubated on the array surface and the polyclonal antibodies bound with certain affinities to

the various glycopeptides. After incubation, the slides were washed with buffer to remove all unbound antibodies. Then, a secondary antibody with affinity for the F_c part of the primary serum antibodies was incubated on the slide. The secondary antibody was conveniently labeled, e.g. with a fluorophore for detection with a fluorescence scanner (*figure 4.55*).

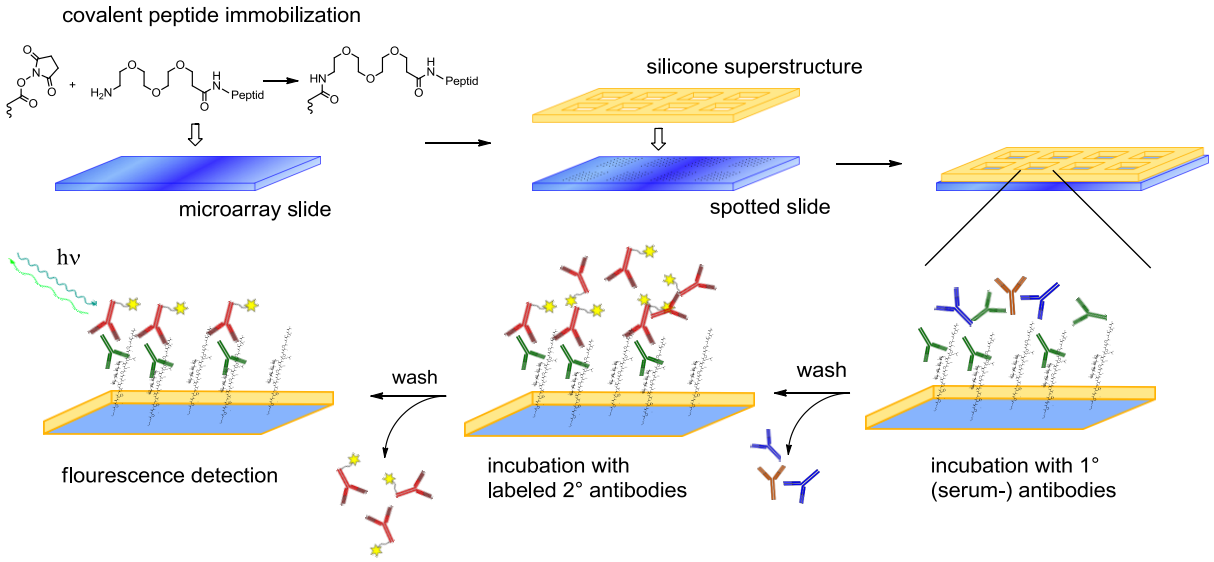


***Figure 4.55*:** Principle of glycopeptide microarray analysis with fluorescence detection.

4.3.4 Evaluation of mucin anti-tumor vaccines by glycopeptide microarray analysis of induced antibody responses

The MUC1 glycoprotein is a promising target for biomarker tumor diagnostics and a lead structure for the development of active anti-tumor vaccines.[317] The MUC1 is ubiquitously found on the surface of epithelial cells and is overexpressed by breast-, colon-, pancreas-, prostate- and ovary-carcinomas.[22] Other characteristics of the mucins on cancer cells are aberrant and truncated glycosylation causing loss of apical cell polarization, alteration of adhesion and anti-adhesion events and stimulation of downstream signaling, contributing to the establishment of a reactive tumor microenvironment. The aberrant glycosylation on tumor cells is partly caused by mutation of the Cosmc chaperone essential for T-synthase activity,[40] downregulation of the core 2 β-1,6-*N*-acetylglucosaminyltransferase-1 (C2GnT-1) and premature sialylation through increased sialyltransferase expression.[318,319] As a result, mucin type core 1 *O*-glycosylation, e.g. sialyl-T-, T-antigen and the shorter structures sialyl-T_N- and T_N-antigen, dominate over branched and extended core 2 structures. Thereby, unique peptide epitopes, previously shielded by the large core 2 glycans, are exposed to the immune system.[30,320,321]

Induction of humoral immune responses directed to these structurally different MUC1 glycopeptide epitopes, possibly accompanied with cytotoxic effects would be valuable assets for tumor immunotherapy. A synthetic vaccine has to meet high expectations. It needs to be safe in its application, elicit a strong immune response, overcome immune tolerance and create an immunological memory. On the other hand, it needs to be selective in eradication of tumor cells without creating autoimmune damages in clinical settings.[111,225] Several successful methodologies, which reliably increase antigenicity of mucin glycopeptides have been reported in the last years. MUC1 glycopeptide B-cell epitopes were conjugated to different immune stimulants like carrier proteins, T-cell epitope peptides and/or mitogens to form di- and tripartite vaccine candidates, or immune silent polymers for multivalent epitope presentation.[271,322] Thus, the question of how to break the immune tolerance to self antigens with concomitant strong immune stimulation is effectively addressed, but nevertheless, a deeper insight into quality and specificity of the raised polyclonal, humoral reactions are to date missing. This is due to the limited availability of glycopeptide probes for bioassays, as e.g. ELISA, microarray or SPR. As a consequence, only a few noteworthy examples have been reported for the extensive screening of tumor associated anti-MUC1 antibodies.[60,323,324]

In order to screen the specificity of polyclonal antibodies and humoral immune response upon stimulation with synthetic glycopeptide vaccine constructs, a unique MUC1 glycopeptide library is introduced for MUC1 glycopeptide microarray generation. Glycopeptide chip analysis enables screening of immune serum antibody specificity and cross-reactivity after induction by different MUC1 vaccine candidates.

The final MUC1 glycopeptide library features more than 130 synthetic MUC1 tandem repeat entries with tumor associated T_N- and T-antigen, as well as further elongated and branched, core 1, 2 and 3 structures in multivalent presentation mode. The extended carbohydrate core structures represent model glycopeptides that expect to resemble the glycosylation status in a non-diseased cells surface environment. The use of chemically defined glycopeptide structures, was aimed to answer questions about the different characteristics of the MUC1 tandem repeat glycosylation sites, VT*S*A, PDT*RP and GS*T*A, interactions between this domains and the different carbohydrate haptens. Influences on the antibody specificity due to the use of different immune stimulants were further visible.

4.3.5 Vaccine candidates for induction of immune response and antisera generation in mice

The sera of mice treated with several MUC1 vaccine candidates were screened on the MUC1 microarray chip (*figure 4.56*). The induced sera of the following vaccines were kindly provided and pre-evaluated (antibody titer, antibody isotype analysis, T-cell response and binding) by group team members and collaborators.

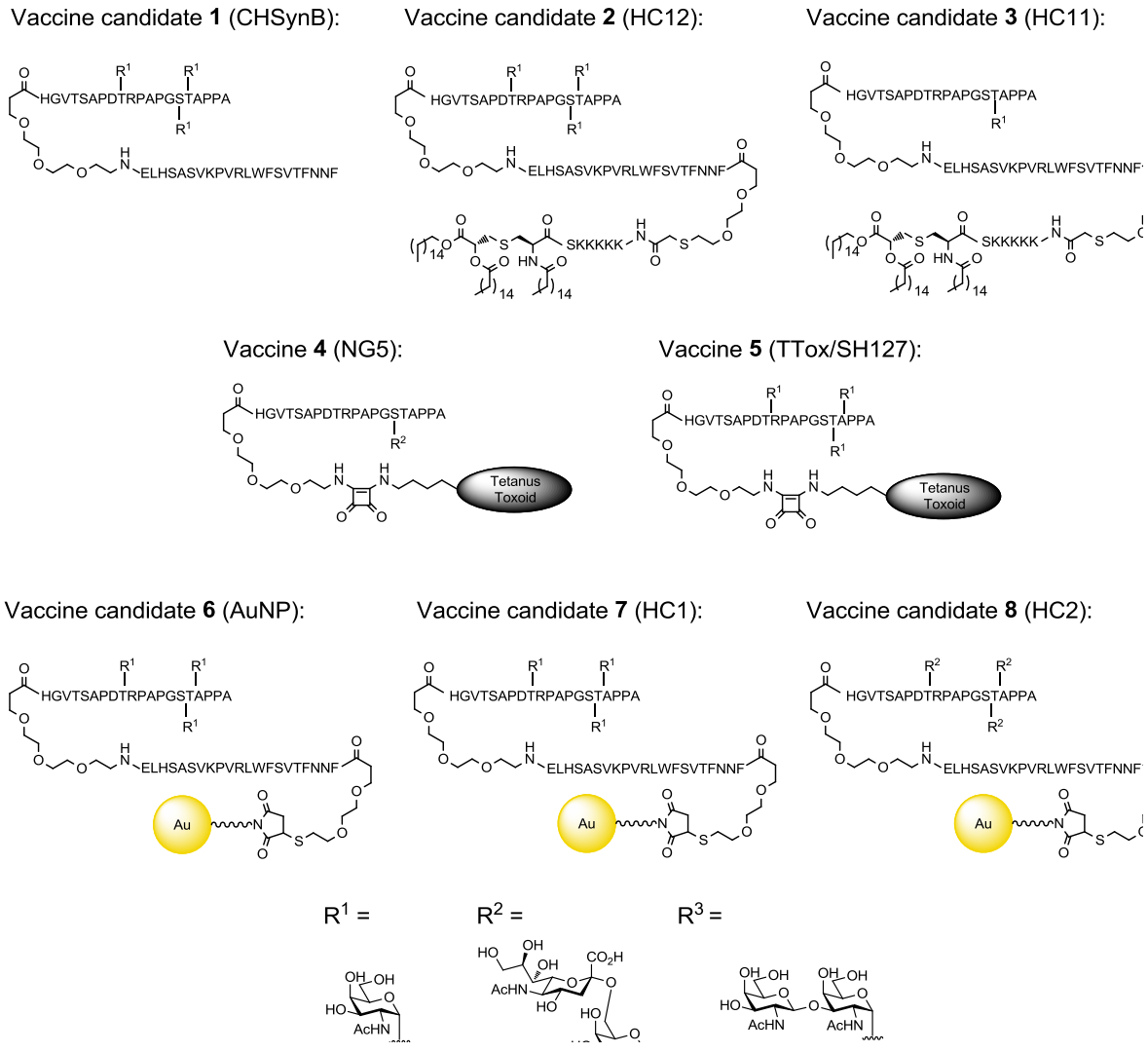


***Figure 4.56*:** Structures of the vaccine candidates used for antiserum generation in mice.

- Vaccine candidate **1** (CHSynB)

  Synthesized by Dr. H. Cai, Leibniz Institute for Analytical Sciences -ISAS-, work group Dr. Ulrika Westerlind. Two-component vaccine of a N-terminal MUC(20mer) B-cell epitope HGVTSAPDT*RPAPGAS*T*APPA with trivalent αGalNAc (T_N-antigen) glycosylation and P30 peptide FNNFTVSFWLRVPKVSASHLE as T-cell epitope.

Administration of the vaccine to three BALB/c mice was performed in the group of Prof. E. Schmitt, Institute of Immunology, Johannes Gutenberg University, Mainz.

- Vaccine candidate **2** (HC12)

  Synthesized by Dr. H. Cai, Leibniz Institute for Analytical Sciences -ISAS-, work group Dr. U. Westerlind. Three-component vaccine of a N-terminal MUC(20mer) B-cell epitope HGVTSAPDT*RPAPGAS*T*APPA with trivalent αGalNAc (T_N-antigen) glycosylation, a P30 peptide FNNFTVSFWLRVPKVSASHLE as T-cell epitope and a Pam_3CSK_4 lipopeptide as mitogen. Administration of the vaccine to three CL57B/6 mice was performed in the group of Prof. M. Lu, Institute of Virology, University Hospital of Essen.

- Vaccine candidate **3** (HC11)

  Synthesized by Dr. H. Cai, Leibniz Institute for Analytical Sciences -ISAS-, work group Dr. U. Westerlind. Three-component vaccine of a N-terminal MUC(20mer) B-cell epitope HGVTSAPDTRPAPGS*TAPPA with single αGalNAc (T_N-antigen) glycosylation, a P30 peptide FNNFTVSFWLRVPKVSASHLE as T-cell epitope and a C-terminal Pam_3CSK_4 lipopeptide as mitogen. Administration of the vaccine to three CL57B/6 mice was performed in the group of Prof. M. Lu, Institute of Virology, University Hospital of Essen.

- Vaccine candidate **4** (NG5)[126]

  Synthesized by Dr. N. Gaidzik at Johannes Gutenberg University Mainz, work group of Prof. Dr. H. Kunz. Two-component vaccine of a *N*-terminal MUC(22mer) B-cell epitope PAHGVTSAPDTRPAPGS*TAPPA with trivalent αGalNAc (ST_N-antigen) glycosylation, conjugated to a tetanus toxoid carrier protein. One serum (mouse 5) from administration of the vaccine to one BALB/c mouse, performed in the group of Prof. E. Schmitt, Institute of Immunology, Johannes Gutenberg University, Mainz.

- Vaccine candidate **5** (SH127)

  Synthesized by Dr. H. Cai, Leibniz Institute for Analytical Sciences -ISAS-, work group Dr. Ulrika Westerlind. Two-component vaccine of a N-terminal MUC(20mer) B-cell epitope HGVTSAPDT*RPAPGAS*T*APPA with trivalent αGalNAc (T_N-antigen) glycosylation, conjugated to a tetanus toxoid carrier protein. Administration of the vaccine to three Balb/c mice was performed in the group of Prof. E. Schmitt, Institute of Immunology, Johannes Gutenberg University, Mainz.

- Vaccine candidate **6** (AuNP)

  Synthesized by Dr. H. Cai, Leibniz Institute for Analytical Sciences -ISAS-, work group Dr. Ulrika Westerlind. Three-component vaccine of a N-terminal MUC(20mer) B-cell epitope HGVTSAPDT*RPAPGAS*T*APPA with trivalent αGalNAc (T_N-antigen) glycosylation, a P30 peptide FNNFTVSFWLRVPKVSASHLE as T-cell epitope and

coupled to gold nanoparticles. Administration of the vaccine to three Balb/c mice was performed in the group of Prof. E. Schmitt, Institute of Immunology, Johannes Gutenberg University Mainz.

- Vaccine candidate **7** (HC1)
  Synthesized by Dr. H. Cai, Leibniz Institute for Analytical Sciences -ISAS-, work group Dr. Ulrika Westerlind. Three-component vaccine of a N-terminal MUC(20mer) B-cell epitope HGVTSAPDT*RPAPGAS*T*APPA with trivalent αGalNAc (T_N-antigen) glycosylation, a P30 peptide FNNFTVSFWLRVPKVSASHLE as T-cell epitope and coupled to gold nanoparticles. Administration of the vaccine to three CL57B/6 mice was performed in the group of Prof. M. Lu, Institute of Virology, University Hospital of Essen.
- Vaccine candidate **8** (HC2)
  Synthesized by Dr. H. Cai, Leibniz Institute for Analytical Sciences –ISAS-, work group Dr. Ulrika Westerlind. Three-component vaccine of a N-terminal MUC(20mer) B-cell epitope HGVTSAPDT*RPAPGAS*T*APPA with trivalent Galβ1,3-GalNAcα (T-antigen) glycosylation, a P30 peptide, FNNFTVSFWLRVPKVSASHLE as T-cell epitope and coupled to gold nanoparticles. Administration of the vaccine to three CL57B/6 mice was performed in the group of Prof. M. Lu, Institute of Virology, University Hospital of Essen.

4.3.6 Utilized microarray platforms for antisera, plant lectin and galectin-3 screening

The MUC1 glycopeptide sequence library, synthesized in this work, was appended by MUC1 glycopeptides previously synthesized by Dr. U. Westerlind[325,310] and Dr. H. Cai[326]. Five different MUC1 microarrays were established and utilized for antisera, plant lectin and galectin-3 screening (for detailed spotting pattern information on the different microarray formats and a summary list of all containing peptides see *chapter 6.3.4*). Microarray 1 (**MA1**) consists of one of three MUC1 B-cell epitope sequences used for vaccine candidate development (glycopeptides **232**, **234** and **237**). This array format was initially used to determine antibody titer values of the induced antisera, in order to estimate the antibody concentration before proceeding with further experiments in larger glycopeptide microarray formats. The antisera were incubated on the wells in different dilutions series (*chapter 4.3.7*). Microarray format 2 (**MA2**) contained 85 MUC1 glycopeptides and was used for the evaluation of serum antibody specificities of the induced antisera. Microarray format 3 (**MA3**)

contained further 19 MUC1 glycopeptides for giving more specific insights into antibody specificity, after screening experiments with **MA2** (*chapter 4.3.8*). Microarray format 4 (**MA4**) contained 55 glycopeptides containing sialylated glycopeptides and the corresponding non-sialylated precursors, in order to evaluate the influence of complex glycan sialylation on antibody binding. **MA4** was further used to evaluate the *O*-glycan specificity of galectin-3 (*chapter 4.3.11.3*). Microarray format 5 (**MA5**) contained 132 MUC1 glycopeptides for screening the carbohydrate specificity of several plant lectins (*chapter 4.3.11.1*).

4.3.7 Antibody titer determination of antisera induced by MUC1 vaccine candidates

Before the sera were administered to the MUC1 microarray chips, adequate concentrations for incubations were determined. Therefore, only the antigenic MUC1 sequence of the corresponding vaccine was immobilized on the microarray slide in relative small wells (**MA1**, 3 x 3 mm, 5 µL incubation volume) and incubated with dilution series of the antisera. Peptide spotting concentration and spotting droplet size were equal to the spotting parameters of the full library (experimental details, *chapter 6.4*). Like this, saturation concentrations for binding to the B-cell epitope were monitored. Accordingly, further incubations of the complete MUC1 library were made using antibody concentrations close to the saturation limit, in order to also detect weak antibody recognition to bound MUC1 glycopeptides. Well replicates were also prepared with lower serum concentrations underneath the saturation level of the antigenic MUC1 glycopeptide. In extreme cases, the more dilute antibody sera only show reactivity to the antigenic B-cell epitope sequence and all other signals disappeared in the background. This may give the impression, that a polyclonal serum is very specific only for the antigen sequence. To generate a complete picture of weak and strong binders, several concentrations of serum must be tested. Due to differences in induced antibody titers between different individuals, the sera from immunized mice must be compared in terms of specificity, depending on the measured antibody concentrations. Similar to titer determination by enzyme linked immunosorbent assay (ELISA), the antibody binding concentrations towards the antigen can be directly evaluated on a microarray slide. In ELISA optical density (extinction) of a colorimetric reaction is measured and results are usually plotted in a half-logarithmic correlation. Here, fluorescence of a fluorophore tagged secondary antibody is detected and directly plotted against the concentration (*figure 4.57*).

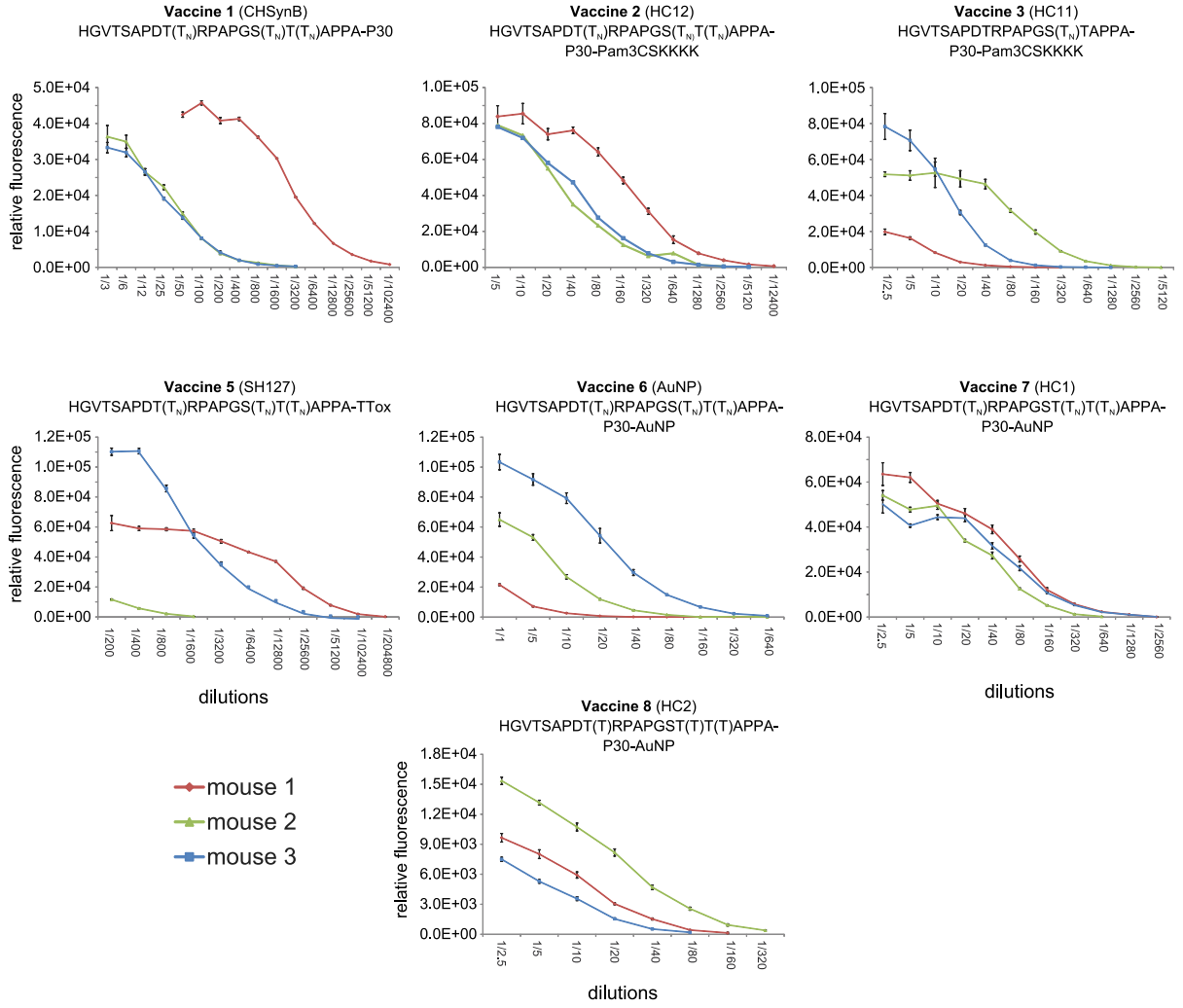
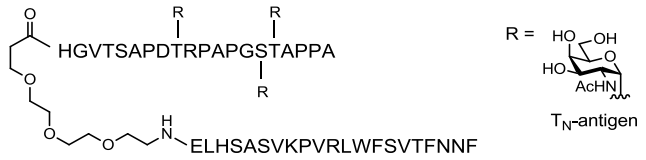


***Figure 4.57*:** Dilution series of all tested vaccine candidates on **MA1**.

The microarray results were in accordance with ELISA tests performed by Dr. Hui Cai. The sera of mouse 1 and 3 induced by the tetanus toxoid conjugated with vaccine **5** show by far the highest antibody titers, confirming the strong immune response elicited by tetanus toxoid conjugated vaccines.[128] Also, mouse 1 of vaccine candidate 1 gave a strong immune response. Based upon the dilution series, concentrations for incubations were estimated and compared among the vaccines.

4.3.8 Elucidation of antibody specificity from vaccination experiments using the MUC1 glycopeptide microarray

4.3.8.1 Vaccine candidate 1 ($3T_N$-MUC1-P30, CHSynB)



The Two component vaccine candidate **1** consisted of the B-cell epitope (HGVTSAPDT*RPAPGS*T*APPA) with T_N-antigen glycosylation in the PDT*R domain and in the GS*T*A domain, both on serine and threonine. The PDT*R domain is known to be immune relevant and glycosylation on this domain enhances cancer derived monoclonal antibody recognition.[327] The GSTA domain has been identified as an immunogenic domain due to findings of monoclonal antibodies and auto-antibodies from cancer patients. The second part of the vaccine consisted of the peptide T-cell epitope P30, a peptide sequence FNNFTVSFWLRVPKVSASHLE derived from the immunogenic tetanus toxoid protein. The vaccine was immunized in mice without extra addition of an adjuvant. The vaccine generated high antibody titers directed to the B-cell epitope in all sera, mouse 1 had a very strong immune response. IgG_1 was the dominating antibody isotype. Breast cancer tumor cells from the T47D and MCF-7 cell lines were recognized in all sera. In addition to the evaluation of the vaccine induced antibody response, antibody glycopeptide binding epitopes were evaluated by microarray analysis (**MA2**: *figure 4.58,* **MA3**: *figure 4.59*).

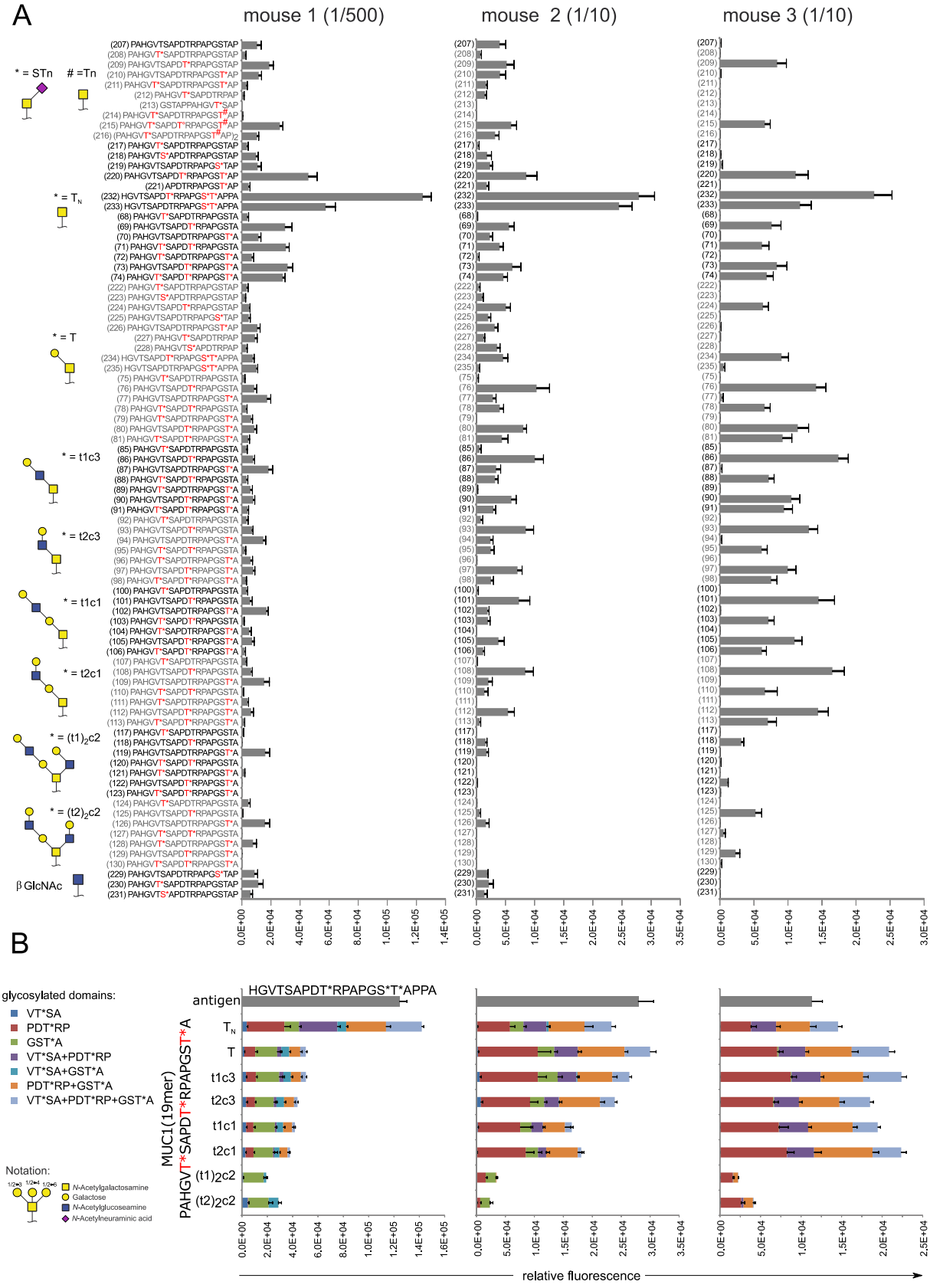


***Figure 4.58*:** Binding of vaccine candidate **1** (CHSynB) antisera (mouse 1-3) on microarray **MA2**. **A:** Full peptide list. **B:** MUC(19mer) sequences, plotted as cumulative binding depending on the glycosylation site.

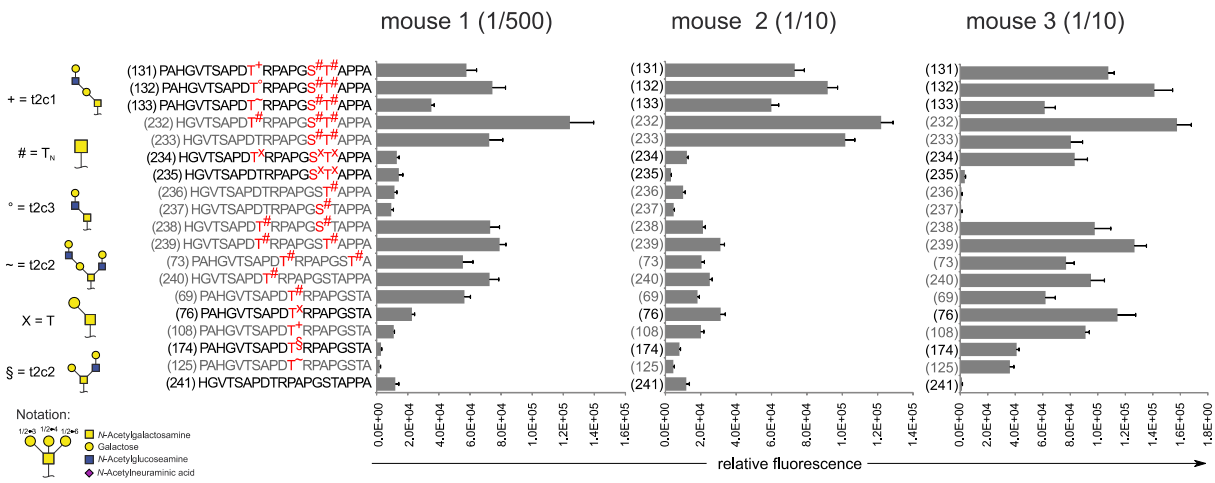


***Figure 4.59*:** Binding of vaccine candidate **1** (CHSynB) antisera (mouse 1-3) on microarray **MA3**.

The vaccine candidate **1** generated antisera from all three mice, showed low antibody specificity for the unglycosylated peptide sequences **207** and **241**, demonstrating that glycosylation was mandatory for strong antibody binding.

Mouse 1 serum (dilution: 1/500):

The mouse 1 serum antibodies had a mixed specificity for both glycosylation sites of the antigen structure, PDT*R and GS*T*A. Apart from strong binding to the B-cell epitope sequence **232**, monoglycosylated peptides with T_N-antigen glycosylation in PDT*R (**70**, **72**, **74**, **75,** *figure 4.58*) were well recognized. Deviations to different extended glycans other than T_N-antigen in PDT*R on the MUC1(19mer) resulted in no significant binding. Sole glycosylation with T_N-antigen in the GSTA region on the serine (**237**, *figure 4.59*) or threonine (**236**, *figure 4.59*) composed only weak epitopes for the antibodies (*figure 4.59*). However, glycopeptide **233** with clustered T_N-antigen glycosylation in the GS*T*A site was highly recognized by the antibodies, proposing that the GSTA domain must be glycosylated at both amino acids. On the other hand, clustered glycosylation in the GSTA domain with T–antigen alone (**235**) or with additional glycosylation in PDT*R (**234**) resulted in very weak antibody affinity. Thus, glycosylation with T_N-antigen exclusively in PDT*R as well as in GS*T*A contribute or add up to overall antibody binding (*figure 4.59*). Larger type-2 core 3, core 1 and core 2 glycans **131-133** in PDT*R reduce binding, but are tolerated to a certain extent, as long as GS*T*A carries clustered glycosylation with T_N-antigen. The equivalent MUC1(19mer) sequences without glycosylation in GS*T*A (**93**, **108** and **125**, *figure 4.58*) were not recognized. Peptides glycosylated with T_N-antigen in the VT*SA region alone, were not recognized, since this domain is not part of the binding epitope. The C-terminal PPA sequence, which is not present in the MUC1(19mer) (ending with GSTA) and some of the

MUC1(20mer) (ending with GSTAP) sequences, has only a minor effect on the overall antibody recognition, e.g. observed by comparing of the sequences **239**/**73** and **204**/**69** (*figure 4.59*). Among the rest of the glycopeptides, e.g. for the block of MUC1(19mer) **75**-**113**, a low basal antibody recognition was observed, that obliterates towards the larger hexasaccharide decorated sequences **117**-**130**, except for the core 2 hexasaccharides in GST*A (**119**, **126**), which still showed weak antibody reactivity.

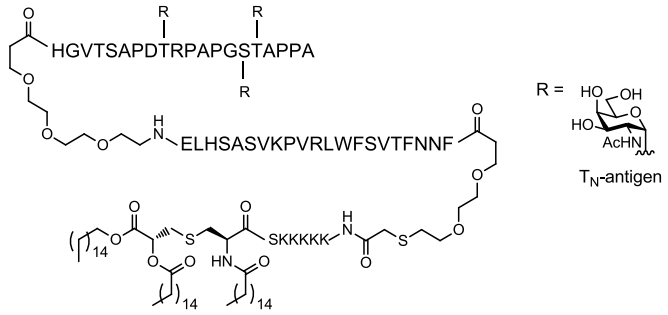
Mouse 2 and mouse 3 sera (dilution: 1/10):

Mouse 2 and 3 sera showed very similar binding specificity with preference for peptides glycosylated with T_N-antigen in PDT*R and GS*T*A. Mouse 2 serum antibodies had stronger reactivity towards clustered T_N-antigen GS*T*A epitopes, while mouse 3 serum antibodies preferred the glycosylated PDT*R domain. In accordance with mouse 1 serum, sequence **233** with double T_N-antigen glycosylation in GSTA was well recognized by both sera and antibody reactivity towards single glycosylation on serine or threonine (**236**, **237**) was weak. The equivalent sequence **235** with clustered T-antigen glycosylation was not detected (*figure 4.58*). Like the mouse 1 serum antibodies, clustered glycosylation with T_N-antigen on serine and threonine in GS*T*A was a prerequisite for contribution to antibody binding. In contrast to mouse 1 antibodies, glycosylation in PDTR was allowed to deviate to glycan structures other than T_N-antigen and strong antibody reactivity was observed on virtually all MUC1(19mer) peptides with glycosylation ranging from monosaccharide to core 1 tetrasaccharide structures in PDTR (e.g. **69** with T_N-antigen and **108** with type-2 core 1, *figure 4.58*). The GS*T*A epitope on the other hand, only contributes when carrying exclusively the T_N-antigen on both serine and threonine. Sequences with PDT*R and single threonine glycosylation in the GS*T*A site are recognized as well as PDT*R glycosylation alone. Single glycosylation in the VT*SA site was not recognized by the antibodies and glycosylation in the VT*SA combined with glycosylation in the PDT*R even reduced the antibody reactivity towards the PDT*R binding epitope. Glycopeptide sequences with glycans up to the size of linear tetrasaccharides in the PDTR site, were well recognized, binding to branched core 2 hexasaccharides were strongly reduced. In particular, serum of mouse 3 showed significant stronger fluorescence intensities for the linear type-2 core 1 tetrasaccharide glycopeptide **108** than for the branched type-2 core 2 tetrasaccharide glycopeptide **174** (*figure 4.59*). The branched type-2 core 2 hexasaccharide peptide **125**, sharing the same linear tetrasaccharide substructure with peptide **108**, showed the same antibody reactivity as the branched tetrasaccharide glycopeptide **205**. This demonstrates that the branching of the inner αGalNAc influences the antibody recognition and antibodies prefer core 1 over core 2 glycosylation, rather than being negatively influenced in antibody recognition by the glycan size in terms of numbers of saccharides. Also glycopeptides **131**

(tetra core 1) and **133** (hexa core 2) show this trend, albeit less pronounced, since antibodies in the immune sera compensate for this effect by recognition of the additional GS*T*A binding epitope glycosylated with clustered T_N-antigen (*figure 4.59*).

For all trivalent MUC1(19mer) glycopeptides (VT*SA+PDT*R+GST*A) with linear core 3 and core1 glycosylation (**74**, **81**, **91**, **98**, **106**, **113**, *figure 4.58*, **B**) no obvious shielding effect by the high glycan density influenced antibody recognition to the PTDR epitope. These peptides were all recognized by the antibody sera, as well as the equivalent divalent peptides with glycosylation in VT*SA+PDT*R or PDT*R+GST*A. The reactivity of the divalent peptides, glycosylated in VT*SA and PDT*R, was weaker than divalent peptides with PDT*R+GST*A glycosylation. Divalent glycopeptides with PDT*R+GST*A-glycosylation were recognized as effective as monovalent peptides with glycosylated PDT*R region alone. Hence, even for trivalent glycopeptides with three core 1 tetrasaccharides, no significant reduction in antibody binding was observed due to multivalent glycosylation. However, trivalent core 2 hexasaccharide glycopeptides (t1c2 and t2c2, VT*SA+PDT*R+GST*A, *figure 4.58*, **B**) were not recognized any more by the antibodies, probably because of steric reasons and the above mentioned unfavorable core 2 recognition. These peptides may therefore be regarded as mucin tandem repeats completely shielded by *O*-glycans, which are found to a large extend on healthy epithelial cells.

4.3.8.2 Vaccine candidate 2 ($3T_N$-MUC1-P30-Pam_3CSK_4, HC12)



The three component vaccine candidate **2** consisted of the same B-cell epitope as vaccine candidate **1**, (HGVTSAPDT*RPAPGS*T*APPA) with T_N-antigen glycosylation in the PDT*R domain and in the GS*T*A domain, both on serine and threonine. The B-cell epitope was additionally connected to the P30 T-cell epitope and the toll-like receptor-2 ligand Pam_3CysK_4, which functions as a build-in immunoadjuvant. The vaccine was immunized in mice without extra addition of an external adjuvant. The vaccine generated high antibody titers directed to the B-cell epitope in all sera. A mixed IgG response was generated, with IgG1 as the dominating antibody isotype and some IgG_{2b} and IgG_3 antibodies. Breast cancer

tumor cells from the T47D and MCF-7 cell lines were recognized in all sera, pancreas capan-1 and -2 tumor cells were here not recognized. Microarray analysis for evaluation of antibody glycopeptide binding epitopes is shown below (**MA2**: *figure 4.60,* **MA3**: *figure 4.61*).

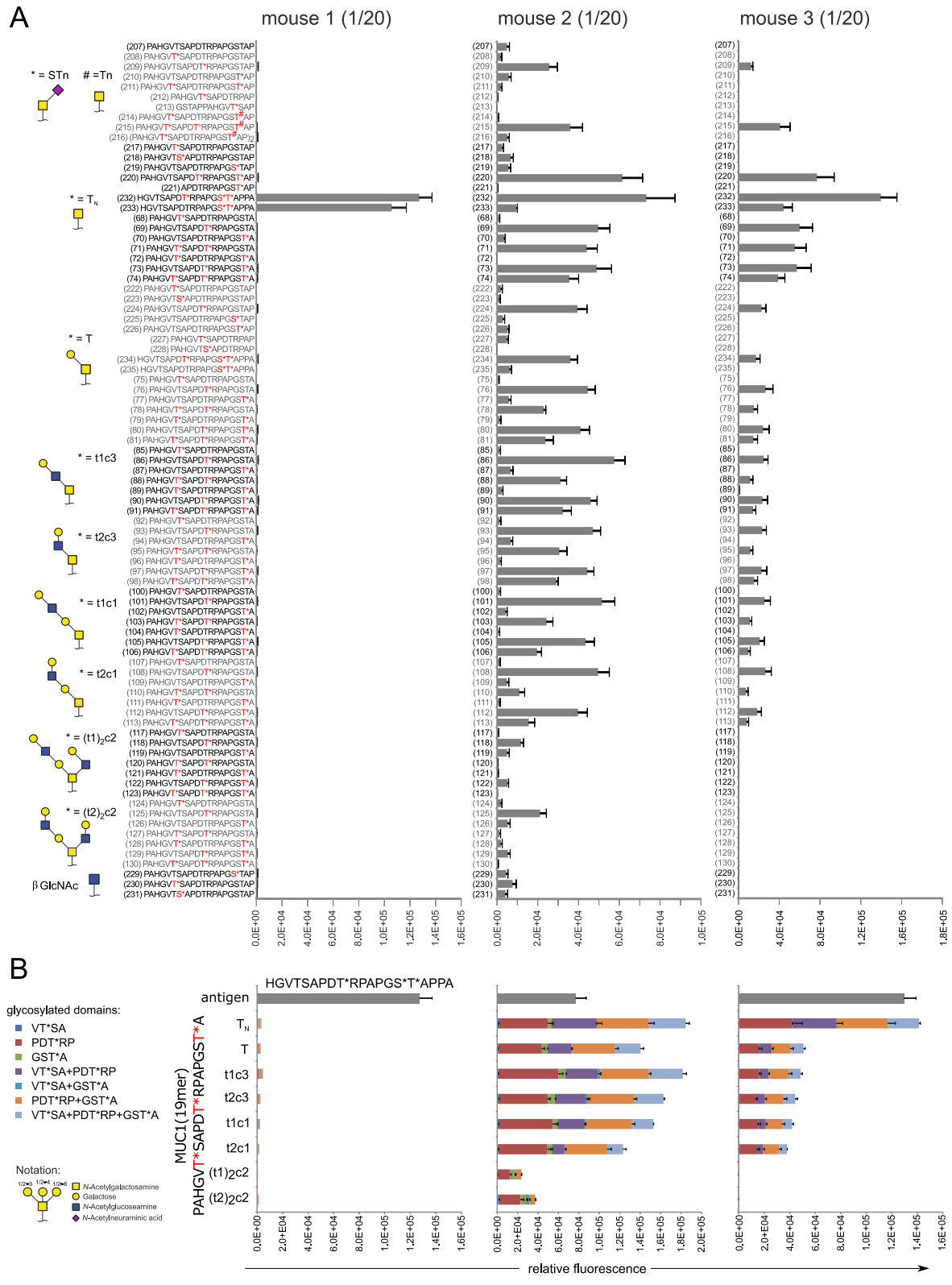


***Figure 4.60*:** Binding of vaccine candidate **2** (HC12) antisera (mouse 1-3) to microarray **MA2**. **A:** Full peptide list. **B:** MUC(19mer) sequences, plotted as cumulative binding depending on the glycosylation site.

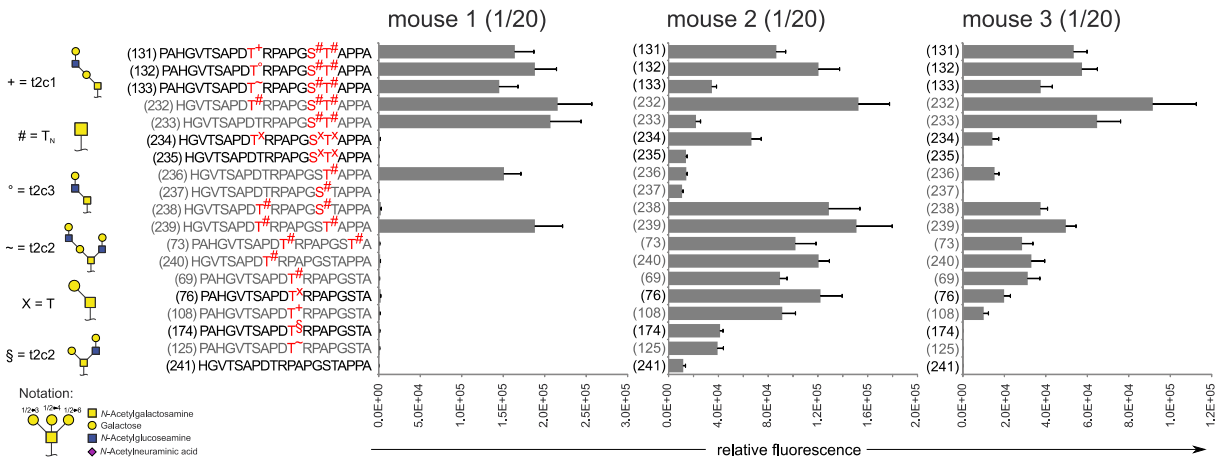


***Figure 4.61*:** Binding of vaccine candidate **2** (HC12) antisera (mouse 1-3) on microarray **MA3**.

All sera induced by vaccine candidate **2** showed no recognition of unglycosylated MUC1 sequences **207** and **241** (*figure 4.60*, *figure 4.61*). Glycosylation was mandatory. Mouse 2 and 3 had in principle similar affinities like mouse 2 and 3 of vaccine candidate **1**.

Mouse 1 serum (dilution 1/20):

Serum antibodies of mouse 1 were very specific for glycopeptides with single T_N-antigen glycosylation on threonine in GST*A. The glycopeptide 19mer **70** (ending with GST*A, *figure 4.60*) and the 20mer **237** with serine glycosylation in GS*TAPPA were not bound by the antibodies. The threonine glycosylated C-terminal T*APPA (e.g. in **236**) was the major antibody binding epitope (*figure 4.61)*.

Mouse 2 serum (dilution 1/20):

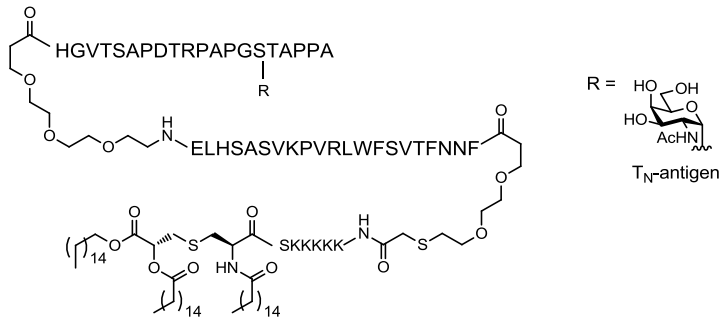
The antibodies were almost only dependent on the glycosylated PDT*R domain. Single or double glycosylation in the GSTA region, no matter of the glycan structure, only provided weak interaction with the antibodies. In accordance with the mouse 2 and 3 sera of vaccine candidate 1, antibodies were accepting various linear glycans in the PDT*R epitope, e.g. T_N-antigen, T-antigen, extended core 3 and core 1 (*figure 4.60*, **A**). Branched core 2 glycans in PDT*R provided again lower affinity to the antibodies than the core 1 glycans (*figure 4.61*). In multivalent glycopeptides, glycosylation in VT*SA had a negative effect on binding, while single glycosylation on threonine in GST*A was tolerated (*figure 4.60*, **B**). Strong reduction of binding was observed for glycopeptides with multivalent core 2 hexasaccharide glycosylation (*figure 4.60*, **B**).

Mouse 3 serum (dilution 1/20):

Mouse 3 serum antibodies had a mixed PDT*R and GS*T*A dependency. Glycosylation in the GS*T*A region must be clustered T_N-antigen glycosylation (compare **233** and **235**, *figure*

4.61). Glycosylation in PDT*R had strongest effects when T_N-antigen glycosylation was present, albeit other glycans in this domain were accepted with lower affinity (**131**, **132**, **132**, **232**, *figure 4.61*). Core 1 over core 2 specificity in PDT*R was present, but at a weak level due to overall weaker affinity to PDT*R, compared to GS*T*A.

4.3.8.3 Vaccine candidate 3 (T_N-MUC1-P30-Pam_3CSK_4, HC11)



The three component vaccine candidate **3** consisted of the same B-cell epitope peptide sequence as vaccine candidate **1** and **2**, but with T_N-antigen glycosylation only in the GS*TA domain (HGVTSAPDTRPAPGS*TAPPA). In accordance to vaccine candidate 2, the B-cell epitope was connected to the P30 T-cell epitope and the toll-like receptor-2 ligand Pam_3CysK_4. The vaccine candidate was administered to the mice without addition of an external adjuvant. The vaccine generated high antibody titers directed to the B-cell epitope in all sera. A mixed IgG response was generated, consisting of IgG_1, IgG_{2b} and IgG_3 antibodies, mouse serum 1 additionally had high levels of IgM antibodies. Strong recognition of breast cancer MCF-7 tumor cells and pancreas capan-1 and -2 tumor cells was mainly observed in the mouse 2 serum. The mouse 2 serum antibodies further recognized tumor cells from patient breast cancer tumor cells, evaluated on a tumor tissue array including 360 patient samples. Glycopeptide microarray analysis for evaluation of the induced antibody glycopeptide binding epitopes is shown below (**MA2**: *figure 4.62,* **MA3**: *figure 4.63).*

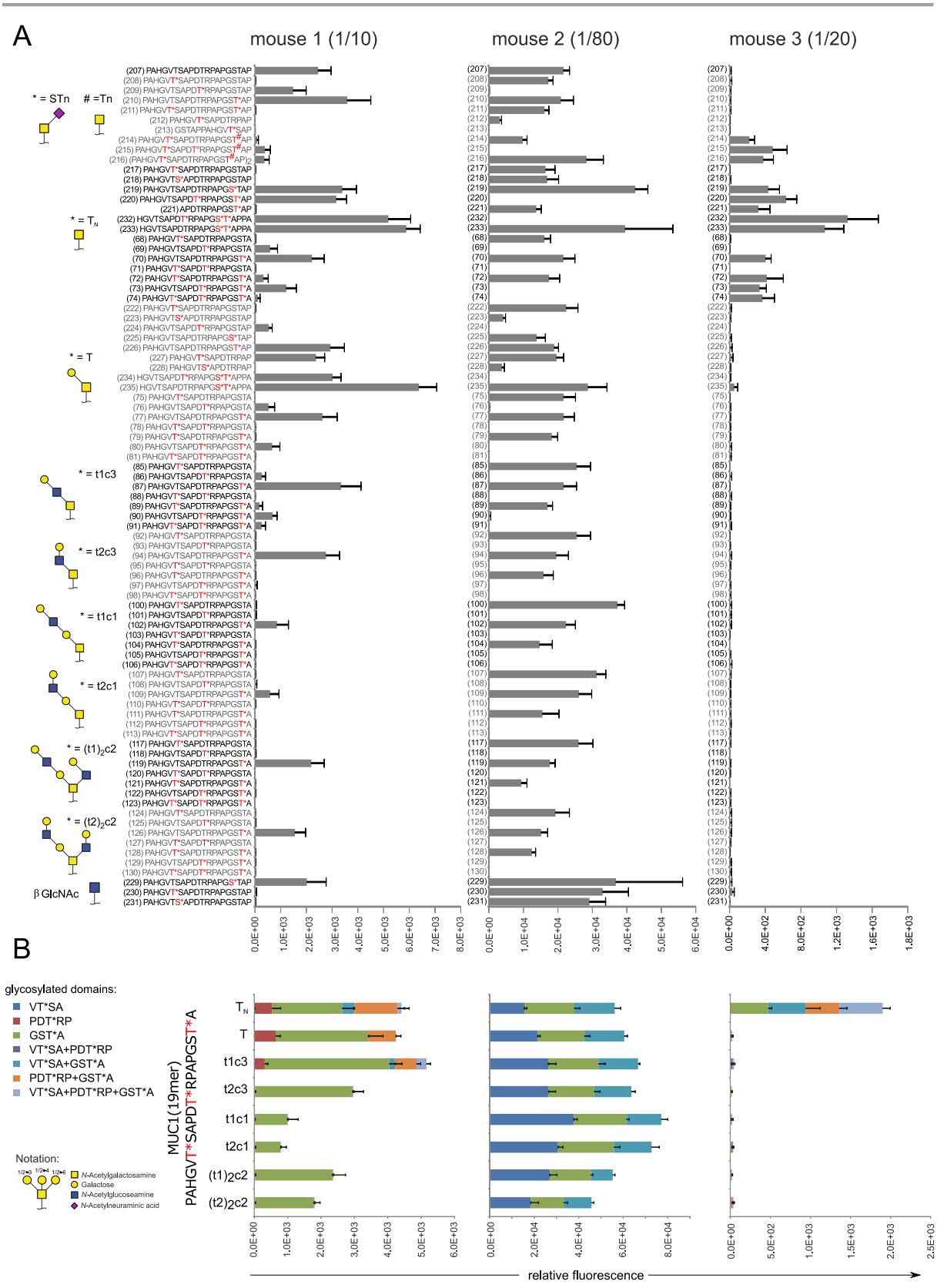


***Figure 4.62*:** Binding of vaccine candidate **3** (HC12) antisera (mouse 1-3) to microarray **MA2**. **A:** Full peptide list. **B:** MUC(19mer) sequences, plotted as cumulative binding depending on the glycosylation site.

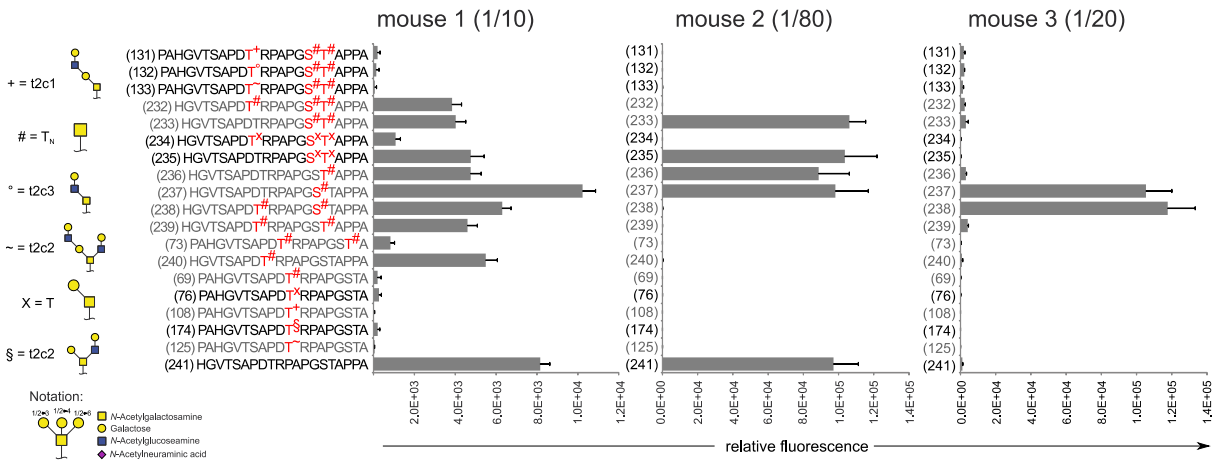


***Figure 4.63*:** Binding of vaccine candidate **3** (HC12) antisera (mouse 1-3) on microarray **MA3**.

The B-cell epitope sequence HGVTSAPDTRPAPGS*TAPPA with T_N-antigen glycosylation was not included in the MUC1 chip **MA2** (*figure 4.62*). Peptide HGVTSAPDTRPAPGS*TAPPA **237** was included in **MA3** (*figure 4.63*).

Mouse 1 serum (dilution 1/10):

The serum antibodies recognized unglycosylated **207** and **241** (*figure 4.62*, *figure 4.63*). Peptides **207** (unglycosylated), **208** (VT*SA), **209** (PDT*R) and **210** (GST*A) imply that VTSA is an epitope, but must not be glycosylated for binding (*figure 4.62*). Short glycopeptide **221** without VTSA domain is also not bound by the antibodies. Further, the C-terminal STAPPA sequence seemed to be the second peptide binding epitope, as seen when **241** is compared to the 19mer sequences (*figure 4.58*). T_N-antigen glycosylation on serine (**237**, *figure 4.63*), as in the B-cell epitope, slightly enhances antibody affinity. Additional glycosylation in PDT*R was tolerated for small T_N-antigen glycosylation, while larger glycans in that domain prevent antibody binding (**131**, **132**, **133**, **232**, *figure 4.62*). Also, peptides glycosylated on threonine in GST*A with ST_N-antigen, T-antigen and extended core 3, core 1 and core 2 were recognized. The monoglycosylated sequences **119** and **126**, with branched core 2 hexasaccharide glycosylation, also showed strong antibody recognition, even stronger than sequences **102** and **109** with the smaller, unbranched core 1 tetrasaccharide glycosylation (*figure 4.62*, **A**+**B**).

Mouse 2 serum (dilution: 1/80):

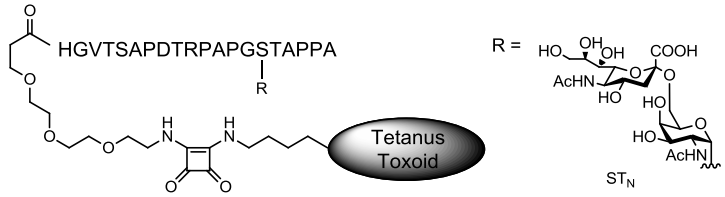
Unglycosylated MUC1 sequences **207** and **241** were recognized by the mouse 2 serum antibodies, proposing that the binding epitope is an unglycosylated peptide segment (*figure 4.62*, *figure 4.63*). Further, all glycopeptides with glycosylation in the VT*SA and/or GST*A regions were all accepted by the antibodies. Recognition was lost when the PDT*R domain was glycosylated, demonstrating that the non-glycosylated PDTR domain was highly

important for antibody recognition. In GSTA, glycosylation in serine as well as threonine or both was tolerated, in VTSA however, only glycosylation in threonine was accepted, while glycosylation on serine obliterates all binding (**222**/**223** and **227**/**228**, *figure 4.62*). The minimum peptide binding epitope might therefore be unglycosylated SAPDTR.

Mouse 3 serum (dilution: 1/20):

The unglycosylated peptides **207** and **241** were not recognized (*figure 4.62*, *figure 4.63*). This antibody serum exclusively recognizes peptides with T_N-antigen glycosylation in the GSTA region with remarkable specificity. No other glycans in this domain were accepted. The specificity was directed on the serine glycosylated with T_N-antigen of the B-cell epitope sequence **237**. The equivalent sequence with T_N-antigen glycosylation on threonine, sequence **236** (*figure 4.63*) was not a binding epitope. Interestingly, T_N-antigen glycosylation on threonine was accepted in the MUC(19mer) peptides with C-terminal GST*A (**70**, **72**, **73**, **74**, *figure 4.62*) and the MUC1(20mer) with C-terminal GST*AP (e.g. **215**, **220**, *figure 4.62*) in contrast to the GST*APPA terminated peptides.

4.3.8.4 Vaccine candidate 4 (T_N-MUC1-TT, NG5)



The vaccine candidate **4** consisted of the same B-cell epitope as vaccine candidate **3**, but with ST_N-antigen glycosylation instead of T_N-antigen in the GS*TA domain (HGVTSAPDTRPAPGS*TAPPA). The B-cell epitope was here conjugated to a tetanus toxoid immune carrier protein. The vaccine candidate was administered to mice with addition of *Freund's* adjuvant. The vaccine generated very high antibody titers directed to the B-cell epitope in all sera. IgG_1 was the dominating antibody isotype. Breast cancer MCF-7 cells were recognized by all sera. The mouse 5 serum antibodies were further evaluated for recognition of tumor cells from breast cancer patients, the antibody staining was gradually increased by evaluation of stage 1 to stage 3 cancer patient tissues. These polyclonal antibodies were therefore considered interesting for glycopeptide microarray analysis (**MA2**+**MA3**: *figure 4.64*).

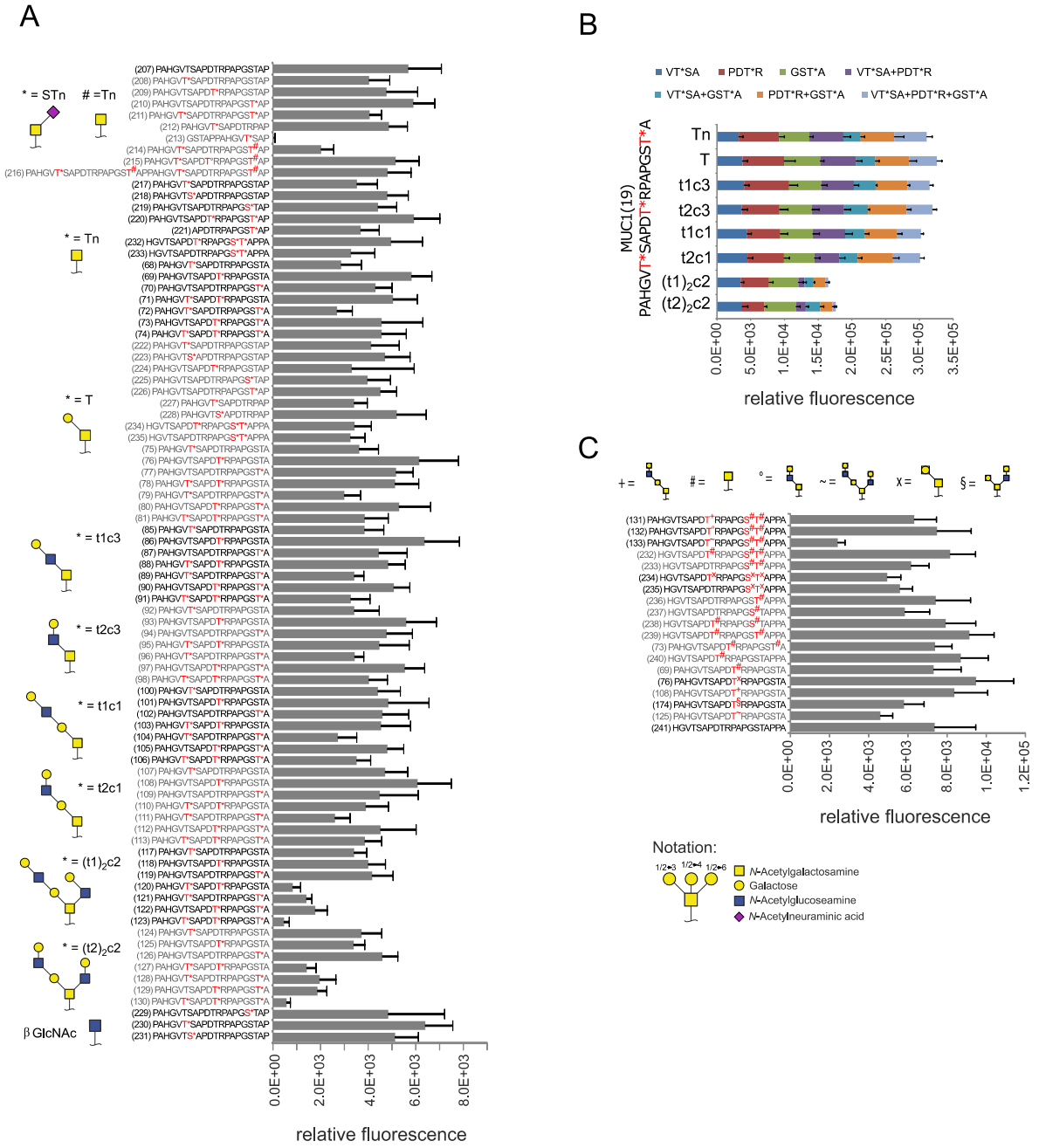


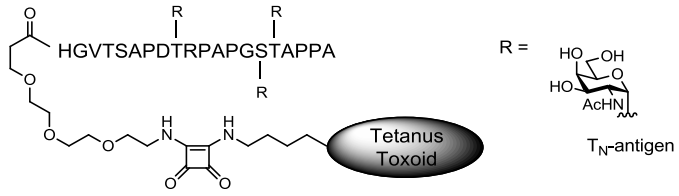
***Figure 4.64*: Binding of vaccine candidate **4** (NG5) antiserum on microarrays (dilution 1:100) **MA2** and **MA3**. **A: MA2**: full peptide list. **B: MA2**: MUC(19mer) sequences, plotted as cumulative binding depending on the glycosylation site. **C: MA3**: full peptide list.**

The B-cell epitope sequence HGVTSAPDTRPAPGS*TAPPA with ST_N-glycosylation was not included in either of the two MUC1 chips. However, peptide sequence HGVTSAPDTRPAPGS*TAP **210** (*figure 4.63*, **C**), shorter by two N-terminal amino acids, was strongly recognized.

The mouse 5 (NG5) antibody serum showed very broad glycopeptide recognition and almost all sequences were recognized by the antibodies including the unglycosylated peptides **207** and **241** (*figure 4.64*, **A**+**C**). No distinct preferences were present, except for certain

peptides glycosylated in the immunodominant PDTR region. The only peptide not recognized was **213** (*figure 4.64*, **A**), consisting of a peptide lacking the PDTR domain. Binding of core 2 hexasaccharide glycopeptides was less strong and vanishes completely for trivalent glycopeptides, which probably completely sterically hinder antibody recognition to peptide backbone epitopes.

4.3.8.5 Vaccine candidate 5 ($3T_N$-MUC1-TT, SH127)



The vaccine candidate **5** consisted of the same B-cell epitope as vaccine candidate **1** and **2**, (HGVTSAPDT*RPAPGS*T*APPA) with T_N-antigen glycosylation in the PDT*R domain and in the GS*T*A domain, both on serine and threonine. The B-cell epitope was here conjugated to the tetanus toxoid immune carrier protein. The vaccine candidate was administered to mice with addition of *Freund's* adjuvant. The vaccine candidate generated very high antibody titers directed to the B-cell epitope in all sera. IgG_1 was the dominating antibody isotype. Breast cancer MCF-7 and T47D cells were recognized in all sera. Microarray analysis for evaluation of the induced antibody glycopeptide binding epitopes is shown below (**MA2**: *figure 4.65,* **MA3**: *figure 4.66*).

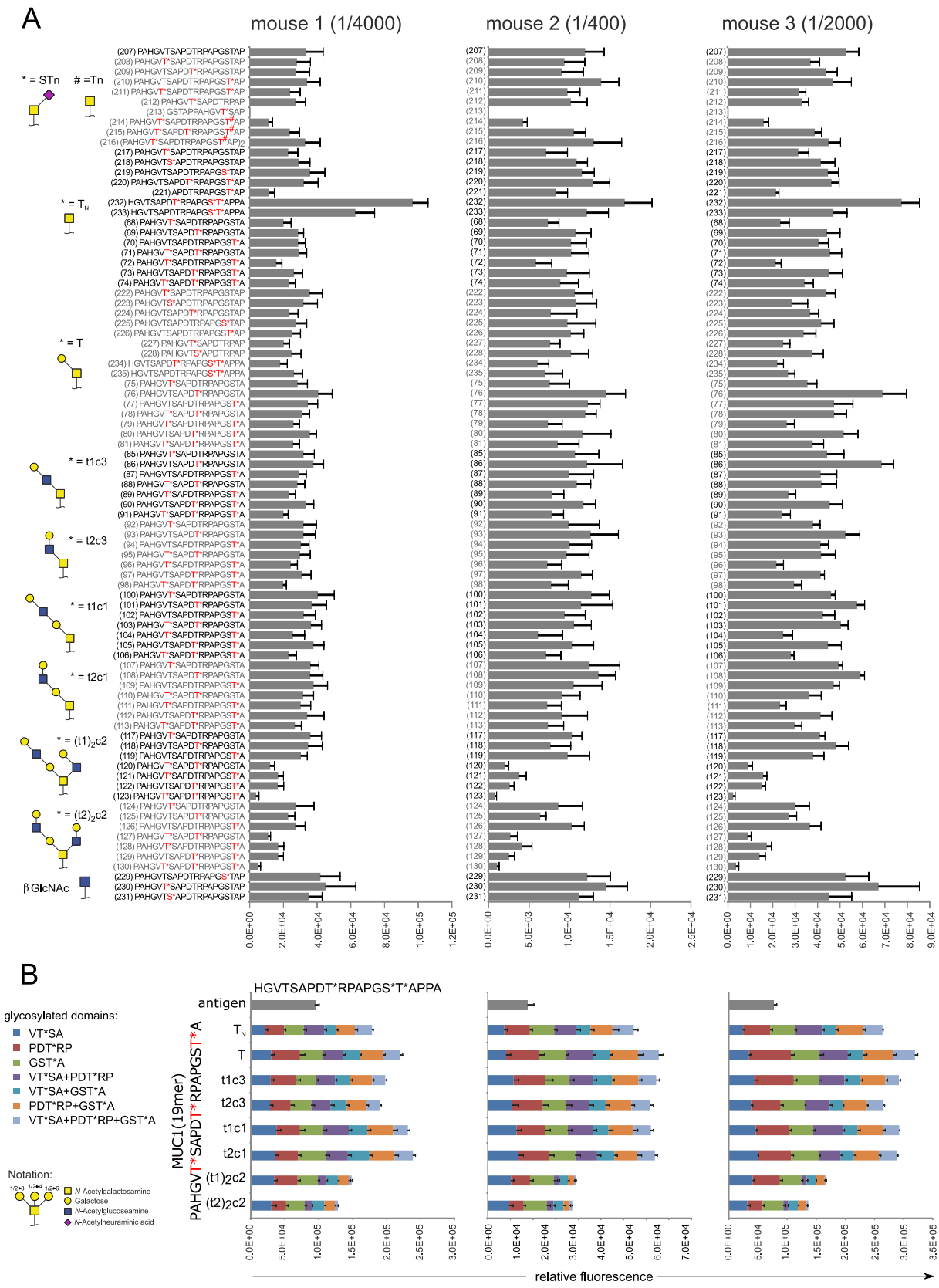


***Figure 4.65*:** Binding of vaccine candidate **5** (SH127) antisera (mouse 1-3) to microarray **MA2**. **A:** Full peptide list.

B: MUC(19mer) sequences, plotted as cumulative binding depending on the glycosylation site.

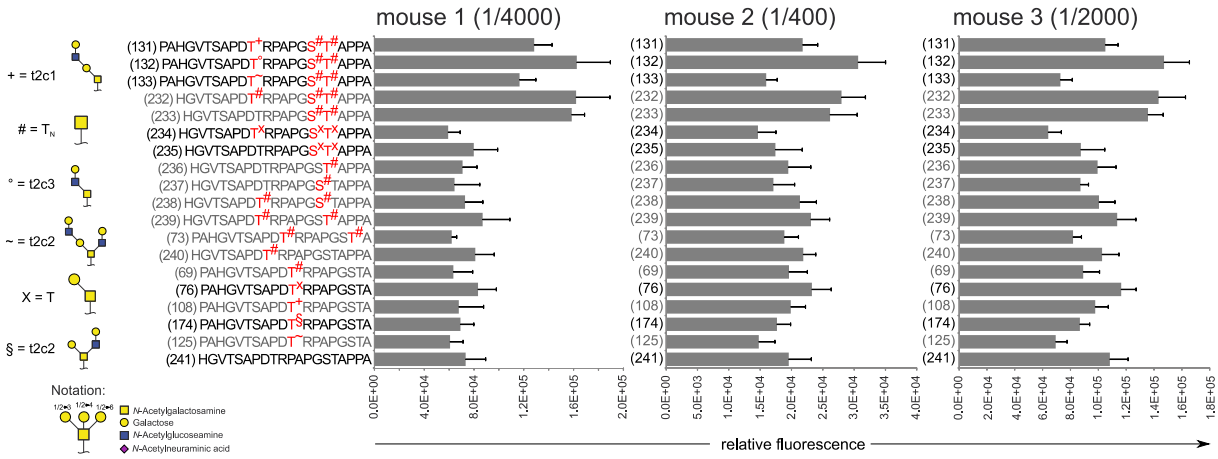
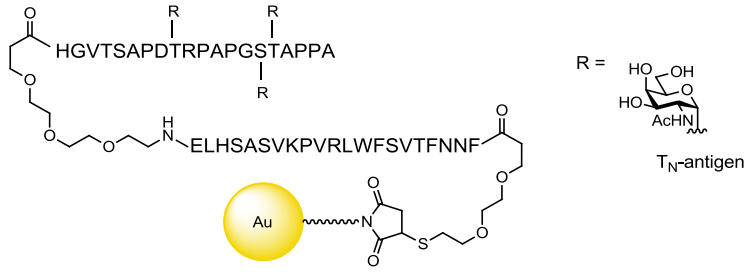


***Figure 4.66*:** Binding of vaccine candidate **5** (HC12) antisera (mouse 1-3) on microarray **MA3**.

Mouse 1 serum (dilution: 1/4000), mouse 2 serum (dilution: 1/400) and Mouse 3 serum (dilution: 1/2000):

The tetanus toxoid conjugate of B-cell epitope **232** induced three mouse antisera with very similar binding patterns. The B-cell epitope sequence **232** was detected strongest by the antibodies, but all other entries were also detected to a good extent. Peptides with glycosylation in PDT*R were slightly preferred by the antibodies of mouse serum 3 (*figure 4.64*). All sera have slight preference for sequences with clustered T_N-antigen glycosylation in GS*T*A (*figure 4.65*, **B**) but no further distinct preferences, neither for domains nor for specific glycan decoration existed. Unglycosylated peptides **207** and **241** were detected as well as most of the others. Only the short peptide **213**, which completely lacks the immunodominant PDTR domain was not bound by any of the three sera. A shielding effect of the large and branched core 2 hexasaccharide structures existed, but was weaker than in the previous examples. Only glycopeptides with hexasaccharide decoration on all three glycosylation sites are effectively shielded from antibody recognition (*figure 4.65*, **B**). The well replicates using higher dilutions of each serum, showed the same overall patterns. Saturation effects can therefore be precluded. Vaccine candidate **5** shares the same B-cell epitope as the above described vaccine candidate **1** and **2**, which shows that the choice of immune stimulant has a strong influence on the induced antibody specificity.

4.3.8.6 Vaccine candidate 6 ($3T_N$-MUC1-P30-AuNP, AuNP)



The three component vaccine candidate **6** consisted of the same B-cell epitope as vaccine candidate **1**, **2** and **5**, (HGVTSAPDT*RPAPGS*T*APPA) with T_N-antigen glycosylation in the PDT*R domain and in the GS*T*A domain, both on serine and threonine. The B-cell epitope was additionally connected to the P30 T-cell epitope and gold nanoparticles (AuNPs) as carriers. The AuNPs were considered to contribute with multivalent presentation of the antigen structure, resulting in increased uptake by antigen presenting cells. The vaccine was administered to mice with addition of *Freund's* adjuvant. High antibody titers directed to the B-cell epitope were generated in all sera, as depicted by ELISA. A mixed IgG response was present, consisting of IgG_1, IgG_{2a} and IgG_{2b} isotype antibodies. Breast cancer tumor cells from MCF-7 cell lines were stained by the mouse 2 and 3 antisera in FACS, but not by the very specific mouse 1 serum. Microarray analysis for evaluation of antibody glycopeptide binding epitopes is shown below (**MA2**: *figure 4.67,* **MA3**: *figure 4.68*).

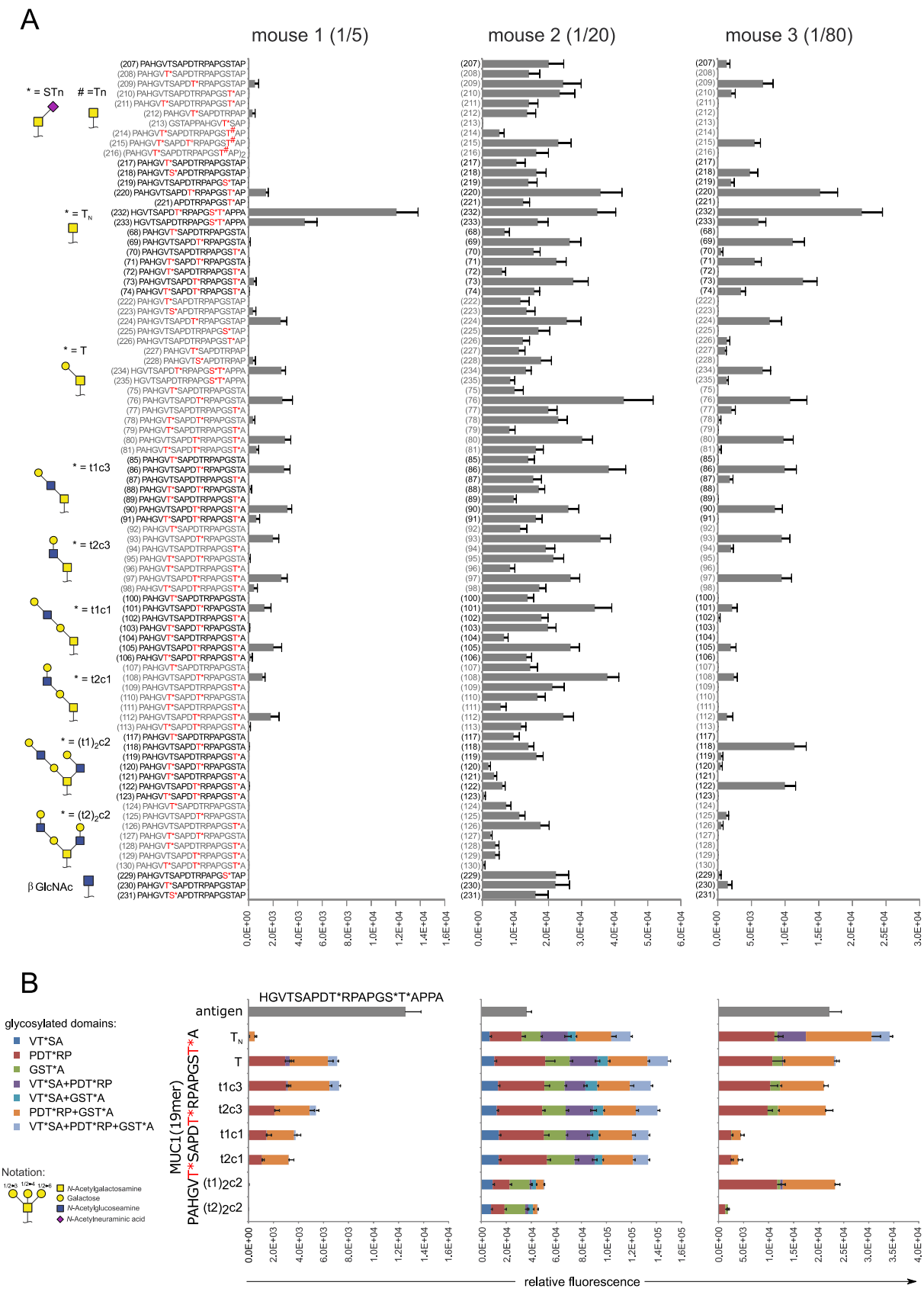


***Figure 4.67*:** Binding of vaccine candidate **6** (AuNP) antisera (mouse 1-3) to microarray **MA2**. **A:** Full peptide list. **B:** MUC(19mer) sequences, plotted as cumulative binding depending on the glycosylation site.

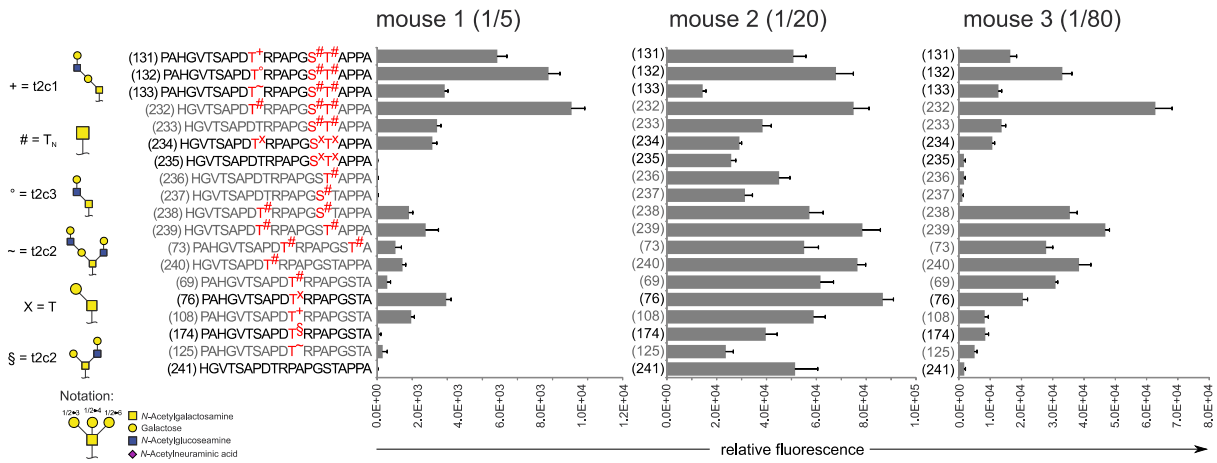


***Figure 4.68*:** Binding of vaccine candidate **6** (AuNP) antisera (mouse 1-3) on microarray **MA3**.

Mouse 1 serum (dilution: 1/5):

Serum from mouse 1 had no preferences for unglycosylated peptides **207** and **241** (*figure 4.67*, *figure 4.68*). The serum antibodies showed a very strong reactivity to the glycopeptide B-cell epitope sequence **232** (*figure 4.68*). As shown by peptide **233**, glycosylation in both PDT*R and GS*T*A contribute to recognition by the antibodies. Peptides **235**, **236** and **237** demonstrated that the clustered T_N-antigen glycosylation must be present in the GS*T*A binding epitope (*figure 4.68*). Interestingly, single glycosylation with T_N-antigen in PDT*R resulted in rather weak antibody binding for the MUC1 20mers (**240**, *figure 4.67*) as well as for the MUC1 19mers (**69**, *figure 4.68*). A stronger antibody affinity was observed for the T-antigen and the tetrasaccharide core 1 glycans, monoglycosylated in PDT*R. Again, antibodies universally accept larger glycans in the PDT*R binding epitope (**131**-**133**, *figure 4.68*). An antibody preference for monoglycosylated core 1 over core 2 structures in the PDT*R epitope was present (**76**, **108**, **174**, *figure 4.68*).

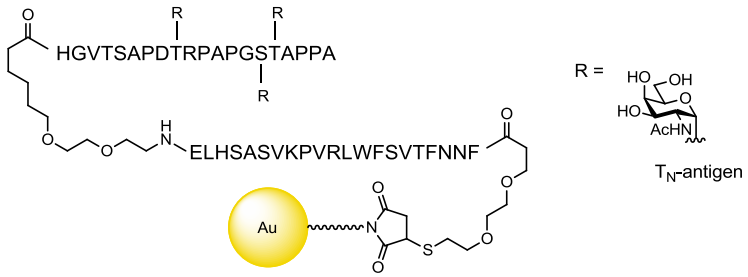
Mouse 2 serum (dilution: 1/20):

The mouse 2 serum antibodies showed reactivity against the unglycosylated MUC1 sequences **207** and **241** (*figure 4.67*, *figure 4.68*). A general binding to most of the glycopeptides on the chip was observed, with a preference for entries glycosylated in PDT*R. Only the short glycopeptide **213** (*figure 4.67)*, devoid of a PDT*R domain, showed no signal. Still, a preference for core 1 over core 2 glycans was present (**76**, **108**, **174** and **131**, **132**, **133**, *figure 4.68*) and a signal cut off for multivalent glycopeptides with large hexasaccharide glycans was observed.

Mouse 3 (dilution 1/80)

Mouse 3 serum antibodies showed no recognition of unglycosylated **207** and **241** (*figure 4.67*, *figure 4.68*). Antibodies mainly reacted on glycosylation in PDT*R, as observed also in other sera, with tolerance for various glycan structures. Glycosylation in the GSTA region was only found to be a minor binding epitope and was weakly recognized only when doubly glycosylated on the serine and threonine. T_N-antigen glycosylation in PDT*R resulted in the strongest antibody binding. Unexpectedly, serum antibodies had a relatively strong binding to mono- and divalent glycosylated type-1 core 2 hexasaccharide modified sequences **118** and **122** (*figure 4.67*). The antibodies showed only weak binding to the corresponding type-2 glycopeptides **125** and **129** as well as the linear, unbranched type-1 tetrasaccharide structures **101**. The preference of core 1 over core 2 glycans was here not present, as seen with glycopeptides **108**, **174** and **125** (*figure 4.68*).

4.3.8.7 Vaccine candidate 7 ($3T_N$-MUC1-P30-AuNP, HC1)



The vaccine candidate **7** is identical to vaccine **6** with the difference that the vaccine candidate was administered to the mice were without extra addition of an external adjuvant (*Freund´s* adjuvant in case of vaccine candidate **6**). The adjuvant liposome formation was believed to disturb the multivalent AuNP antigen presentation, with the risk of encapsulation of the nanoparticles. The vaccine generated high antibody titers directed to the B-cell epitope in all sera. IgG_1 was the dominating antibody isotype. Breast cancer T47D and MCF-7 cells and pancreas capan-1 and -2 tumor cells were recognized in all sera. The mouse 2 serum antibodies were selected and stained tumor cells from patient tissues, evaluated on a tumor tissue array including 360 patient samples. Glycopeptide microarray analysis for evaluation of antibody glycopeptide binding epitopes is shown below (**MA2**: *figure 4.69,* **MA3**: *figure 4.70*).

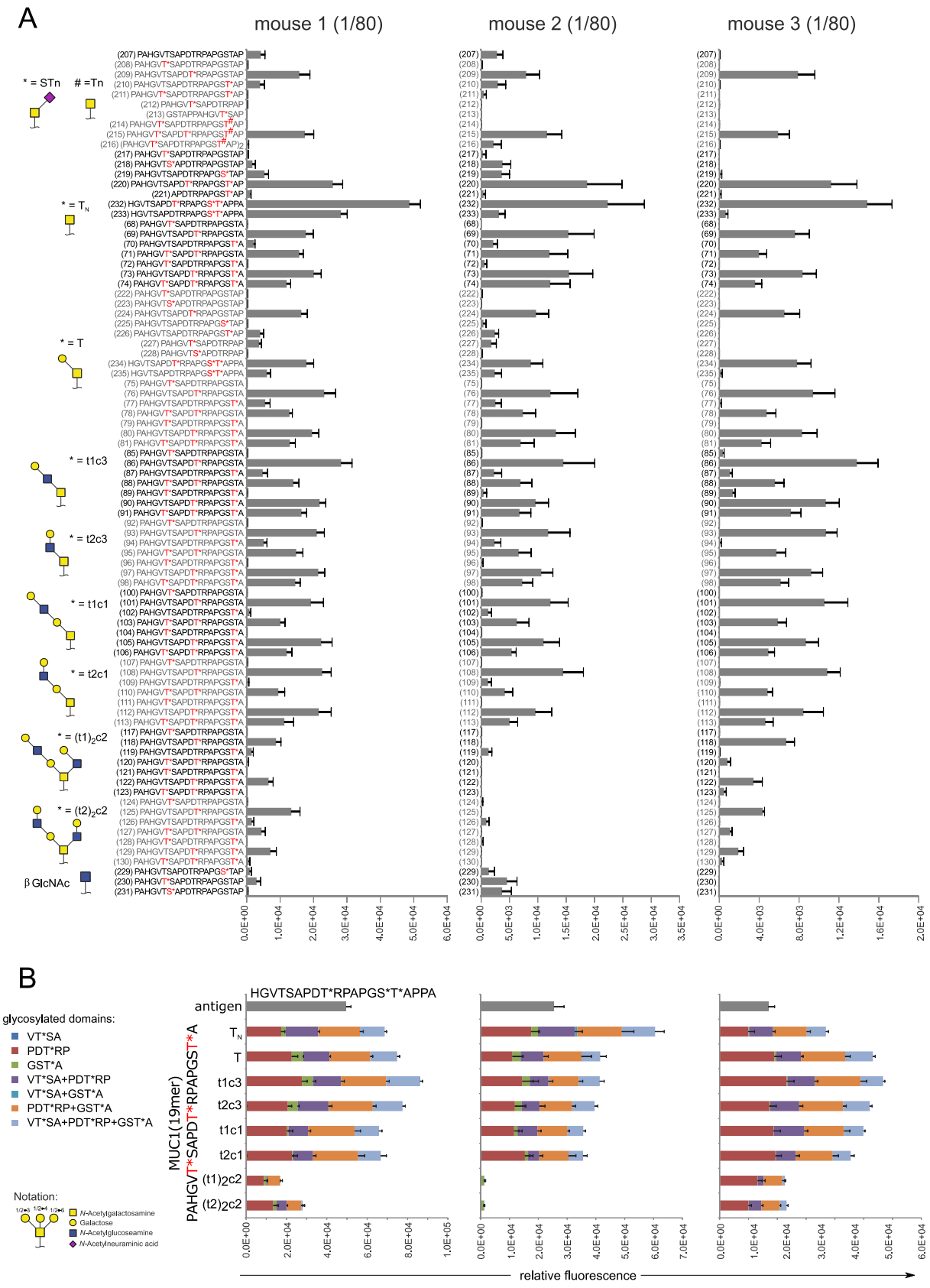


***Figure 4.69*: Binding of vaccine candidate **7** (HC1) antisera (mouse 1-3) to microarray **MA2**. **A:** Full peptide list. **B:** MUC(19mer) sequences, plotted as cumulative binding depending on the glycosylation site.**

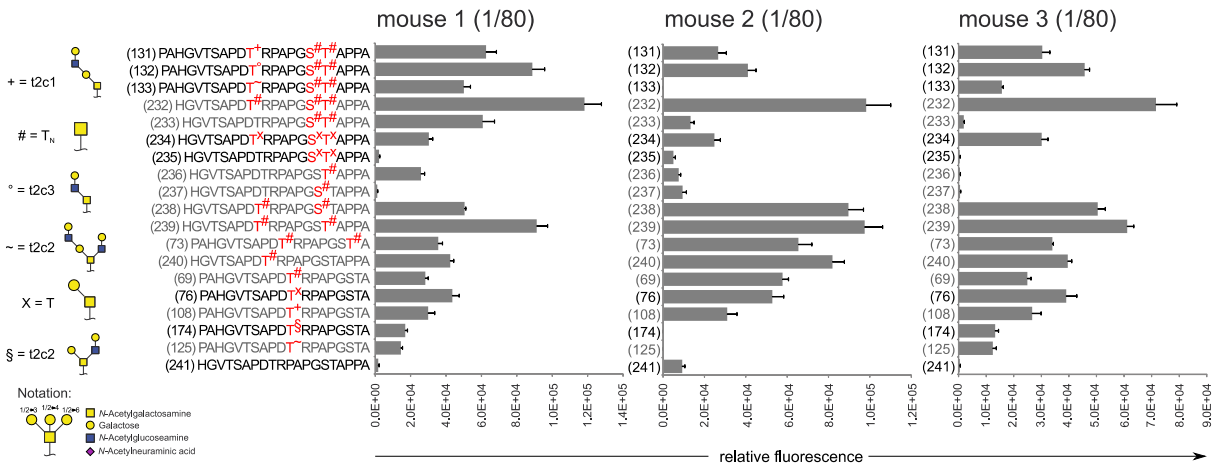
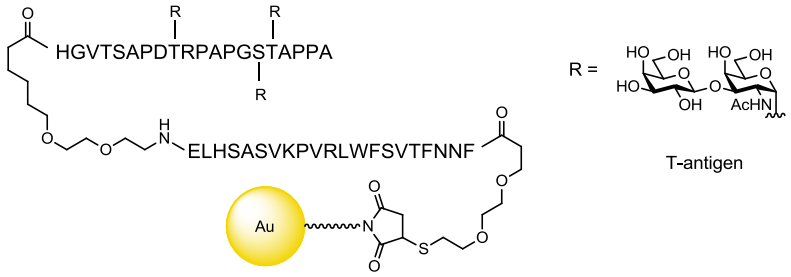


***Figure 4.70*:** Binding of vaccine candidate **7** (HC1) antisera (mouse 1-3) on microarray **MA3**.

Mouse 1 serum (dilution: 1/80), mouse 2 serum (dilution: 1/80) and mouse 3 serum (dilution: 1/80):

Unglycosylated peptides **207** and **241** were only weakly recognized by the antibody sera (*figure 4.69*, *figure 4.70*). The overall patterns for all three sera were similar and more homogenous compared to the aforementioned gold nanoparticle vaccine candidate **6** (AuNP) induced antisera. All sera from this vaccine candidate showed the same overall dependent binding on glycosylation in PDT*R. Dependency on glycosylation in GS*T*A for antibody recognition, was present in the mouse 1 serum, preferably with clustered T_N-antigen presentation. Mouse 2 and 3 serum antibodies had weak (mouse 2) or no interactions (mouse 3) with the glycosylated GS*T*A domains (**233**, **236** and **237**, *figure 4.69*, **A** and green bars in **B**, *figure 4.69*). In all three sera, weaker antibody affinity was observed for, the PDT*R core 2 glycopeptide peptide structures. As with core 2 tetrasaccharide **174** and the core 2 hexasaccharide **125**, weaker antibody affinity was observed to branched core 2 structures than to the linear core 1 structures, such as tetrasaccharide **108** (*figure 4.70*). According to the observations in the previous examples, preferences for core 1 over core 2 glycosylation in PDT*R were present with the serum antibodies having high affinity to this glycosylated domain (**76**, **108**, **174**, *figure 4.70*). Shielding of the peptide backbone with core 2 hexasaccharide triglycosylation was observed (*figure 4.69*, **B**), similar to the previous examples.

4.3.8.8 Vaccine candidate 8 (3T-MUC1-P30-AuNP, HC2)



The vaccine candidate **8** is identical to vaccine **7** with the difference that the B-cell epitope presents the T-antigen structure instead of T_N-antigen on the peptide tandem repeat. The vaccine candidate was administered to the mice without extra addition of an external adjuvant. The vaccine generated high antibody titers directed to the B-cell epitope in all sera. IgG_1 was the dominating antibody isotype. Breast cancer T47D and MCF-7 cells and pancreas capan-1 and -2 tumor cells were recognized with all sera. Glycopeptide microarray analysis for evaluation of antibody glycopeptide binding epitopes is shown below (**MA2**: *figure 4.71,* **MA3**: *figure 4.72*).

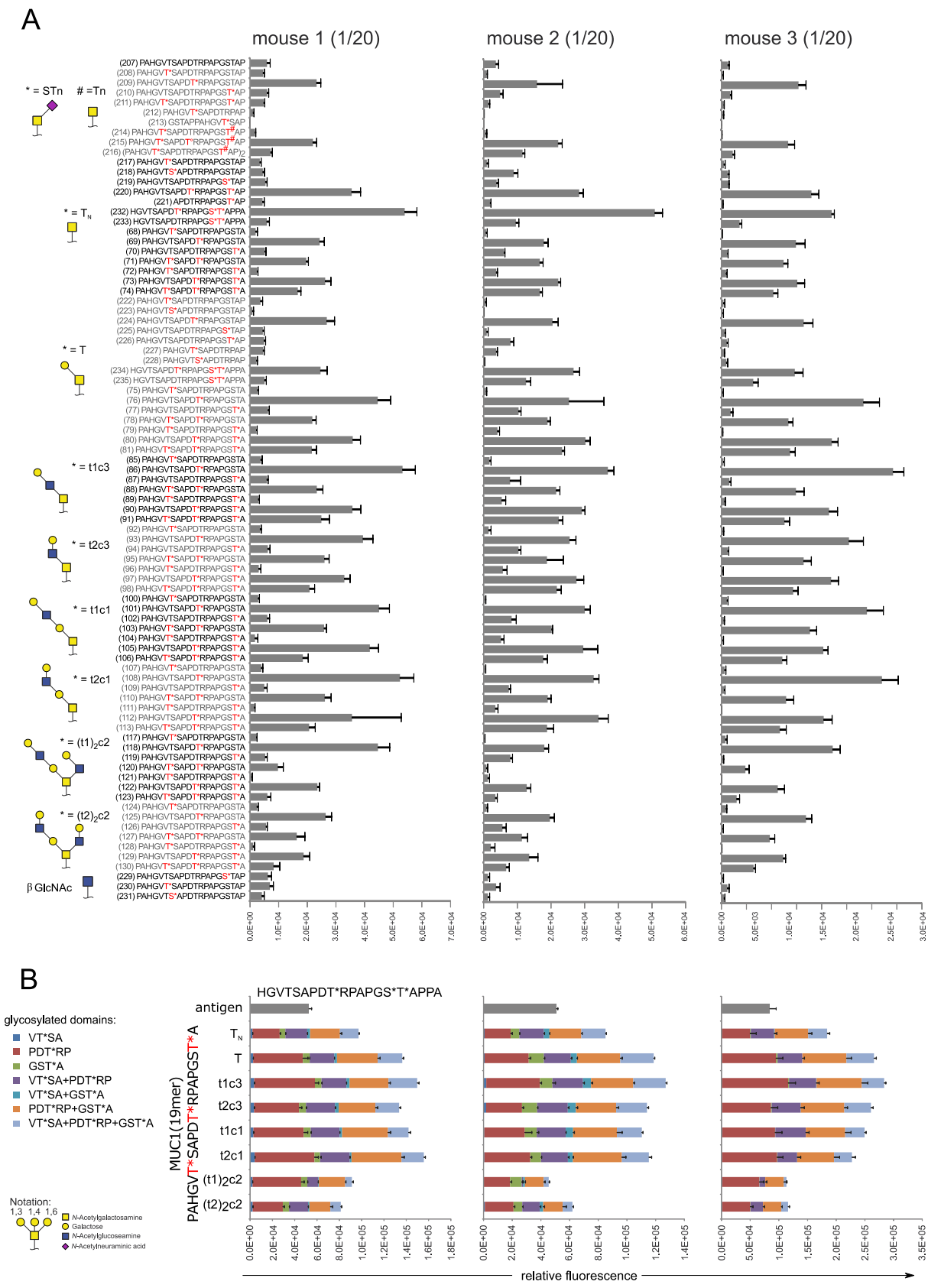


***Figure 4.71*:** Binding of vaccine candidate **8** (HC2) antisera (mouse 1-3) to microarray **MA2**. **A:** Full peptide list. **B:** MUC(19mer) sequences, plotted as cumulative binding depending on the glycosylation site.

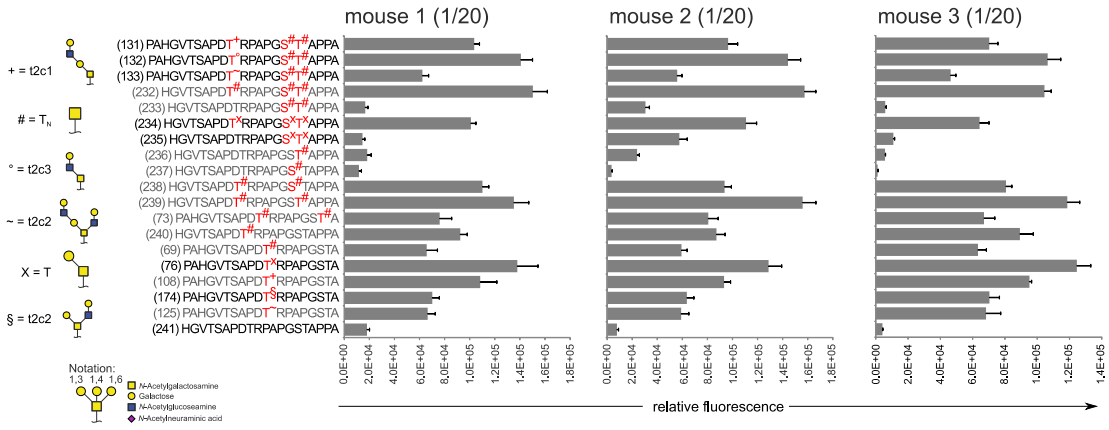


***Figure 4.72*:** Binding of vaccine candidate **8** (HC2) antisera (mouse 1-3) on microarray **MA3**.

Mouse 1 serum (dilution: 1/20), mouse 2 serum (dilution: 1/20) and mouse 3 serum (dilution: 1/20):

The unglycosylated peptides **207** and **241** were recognized with weak affinity by the antibodies (*figure 4.71*, *figure 4.72*). Like vaccine candidate **7**, the overall recognition pattern was rather homogenous for all three sera. The B-cell epitope sequence **234** was strongly recognized by the serum antibodies, but in contrast to the above described antibodies of vaccine candidate **7** (HC1), induced with the triple T_N-antigen glycosylated analogues, several other MUC1 sequences had similar affinity like the B-cell epitope towards the antibodies. Glycosylation in GSTA with T_N-antigen on the serine or threonine or (e.g. **225**, **226**, *figure 4.71*; **233**, **235**, **236**, **237**, *figure 4.72*) double glycosylation with T-antigen in GS*T*A (**235**, *figure 4.72*) or any other glycan on threonine in this region (*figure 4.71*, green bars in **B**), resulted in very weak antibody binding. Only the mouse 2 serum showed some reactivity against clustered glycosylation with T-antigen in the GS*T*A domain. However, if additionally glycosylation in PDT*R was present or if PDT*R was the only glycosylated site, than strong binding was observed, demonstrating that antibody reactivity was directed mainly to the PDT*R epitope. A broad antibody recognition against entries with glycosylation in the PDT*R domain was observed (*figure 4.71*).

Further, the change from T_N- to T-antigen glycosylation of the B-cell epitope (vaccine candidate **7** compared to vaccine candidate **8**), seemed to result in slightly higher antibody acceptance for larger glycans on the peptide backbone. The antibody cut-off for the core 2 hexasaccharide peptides was not as pronounced (*figure 4.71*, **B**) as in vaccine candidate **7** or the two-and three component vaccine candidates **1** and **2**, sharing the same T_N-antigen glycosylated B-cell epitope. Larger glycans in glycopeptide based cancer vaccines may therefore lead to a general higher tolerance for different glycan structures on the mucin peptide tandem repeats. Such vaccine constructs may therefore generate antibodies with lower specificity for small tumor-associated glycans.

Core 1 over core 2 glycan preference still existed, but the difference between glycopeptides **108**, **174** and **241** was not as pronounced as in the other vaccine candidate examples with T_N-glycosylation (*figure 4.72*).

4.3.9 Elucidation of antibody binding specificity towards sialylated MUC1 glycopeptides on a microarray

Sialylation at the terminal, non-reducing end is common for naturally occurring *N*- and *O*-glycans. Sialic acids are hydrophilic, bulky and negatively charged monosaccharides, which commonly serve as recognition domains, but may at the same time shield other carbohydrate epitopes from recognition. These events usually rely on interactions of carbohydrate binding proteins such as bacterial and viral lectins, selectins or siglecs with sialic acid containing glycoproteins or glycolipids.[328] Antibodies may further be directed to unique sialylated carbohydrate epitopes. Here, antibodies directed against tumor associated MUC1 glycopeptides were evaluated for potential cross-reactivity to sialylated glycopeptides. The synthesized sialylated mucin glycopeptides were compared with the unsialylated glycopeptides, to elucidate the potential effects on antibody recognition by the additional sialylation.

After the enzymatic glycosylation, yielding the synthesized sialylated glycopeptides, the low abundant products were quantified by amino acid analysis. To avoid any discrepancies between different methods of quantification of the glycopeptides used for microarray analysis, the corresponding unsialylated glycopeptides were as well determined by amino acid analysis instead of gravimetrical weighing. The unsialylated peptides and their sialylated counterparts were spotted on microarray format **MA4**.

The sera of vaccine candidate **1** (CHSynB) mouse 2, vaccine candidate **2** (HC12) mouse 2, vaccine candidate **7** (HC1) mouse 3 and vaccine candidate **8** (HC2) mouse 2 and 3 were chosen for antibody screening. The antisera were chosen based on the strong dependency on the glycosylated PDT*R binding epitope. Antibodies directed to the GSTA domain were usually highly restricted to double (clustered) T_N-antigen glycosylation and were therefore not selected for elucidation of antibody recognition to sialylated epitopes. Additionally, the commercial monoclonal IgG_1 antibody Ma552 was screened on this array (*figure 4.73*, **F**). Ma552 originated from immunization of mice with ZR75-1 breast cancer cells. A broad set of glycopeptides were recognized by this antibody, although, similar to the probed vaccine candidates, this mAb targets the PDTRPAPGS binding epitope of the MUC1 tandem repeat, preferably if this domain is glycosylated.[46,329,330]

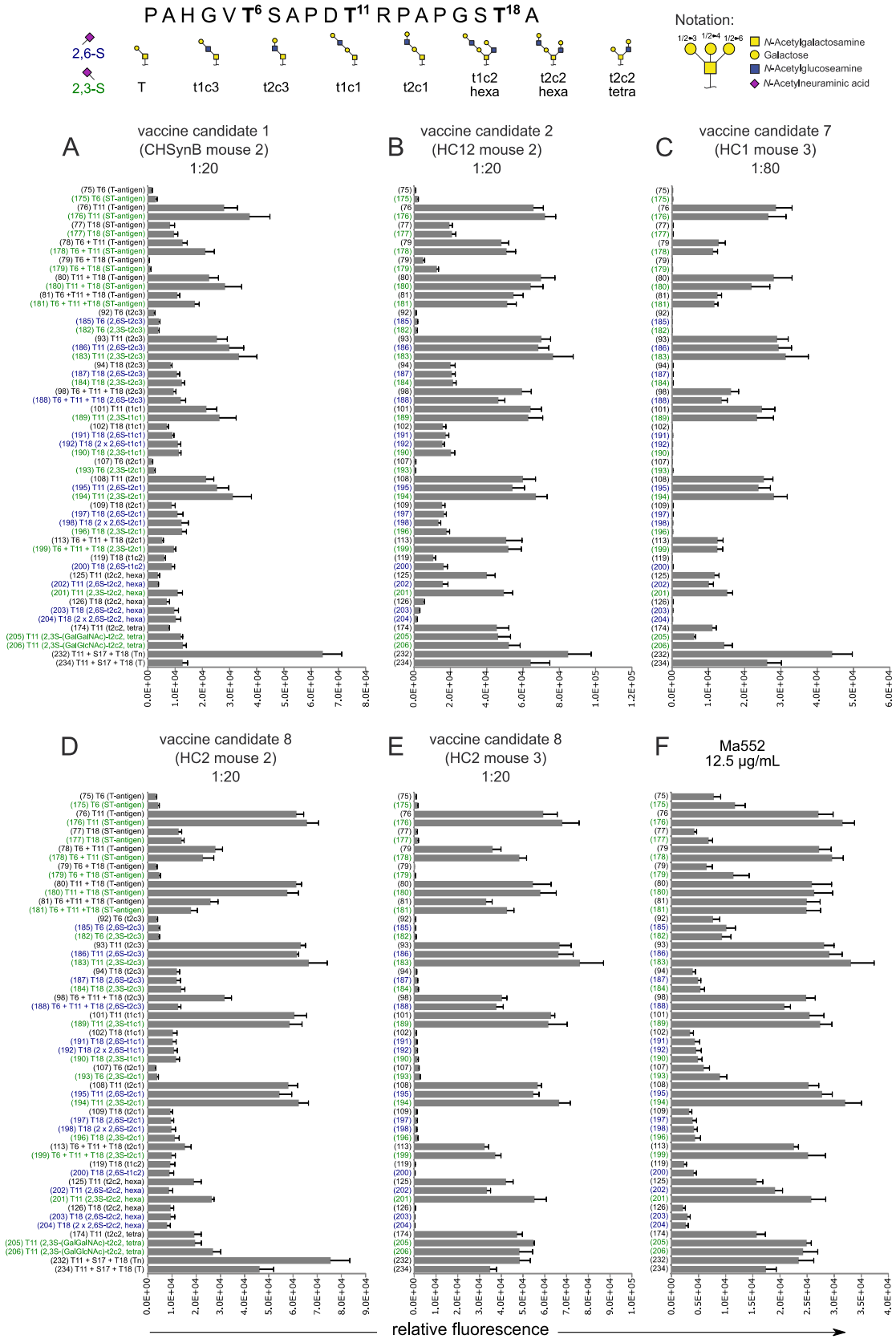


***Figure 4.73*:** Comparison of antibody binding to sialylated glycopeptides and to the unglycosylated counterparts. Unsialylated glycopeptides are depicted in black, α2,6-sialylation in blue and α2,3-sialylation in green. Entries **232** (**A**, **B**, **C**) and **234** (**D**, **E**) are the administered B-cell epitopes.

Sialylation with α2,3-Neu5Ac (in green) and α2,6-Neu5Ac (in blue) resulted in no significant changes in antibody reactivity compared with the unsialylated MUC1 glycopeptide epitopes (in black, *figure 4.73*). This means that the T-antigen and the ST-antigen, which are both tumor-associated antigens, as well as the linear sialylated extended core 1 and core 3 glycopeptide epitopes were well recognized by the induced antibodies. Antibody reactivity to monoglycosylated sialylated and unsialylated core 2 epitopes was also observed. However, on certain core 2 glycopeptides, small but distinct effects of the sialylation were found. Peptide entry **201** with a bulky α2,3-disialylated core 2 octasaccharide in PDT*R seemed to slightly increase antibody recognition compared to the unglycosylated counterpart (**125** vs **201**, *figure 4.73*). α2,6-Sialylation on the core 2 octasaccharide of glycopeptide **202** had a considerably decreasing effect for binding of antibodies, in particular observed in the antisera from vaccine candidate **2** mouse 2 (*figure 4.73*, **B**) and vaccine candidate **8** mouse 2 (*figure 4.73*, **D**).

The monoclonal antibody Ma552, induced from human breast cancer cells, generated a similar antibody recognition to the sialylated and to unsialylated MUC1 glycopeptides. The antibody cross-reactivity to different MUC1 glycopeptide structures was thus in general higher than for most of the antisera induced by the above described synthetic vaccines.

4.3.10 Discussion of MUC1 serum antibody specificity elucidation

4.3.10.1 The role of PDTR-glycosylation

The PDTR peptide domain has been identified as an immuno-dominant region in the MUC1 tandem repeat sequence. The ISOBM TD-4 workshop (San Diego, California, 1996) examined a comprehensive collection of 56 monoclonal antibodies directed against mucin glycoproteins or tumor associated carbohydrates with antigen structures derived from isolated mucins or tumor cells. About 34 antibodies were directed against the MUC1 peptide tandem repeat and 28 of them showed reactivity against epitopes including the core DTR. For the majority of most of the other antibodies, evidence for carbohydrate dependent epitopes were found. However, monoclonal antibody Ma552, which was also part of the ISOBM-investigation, was described to be PDTR-specific, although this investigation showed that Ma552 has merely a preference for PDTR and recognizes further MUC1 glycopeptides with lower affinity (*figure 4.73*, **F**). At the time of the ISOBM workshop, it was unclear if the immunogenicity of the PDTR domain solely relied on the peptide epitope and had a better accessibility when unglycosylated or if the glycosylation stabilized the formation of

immunogenic secondary turn structures. It was found that the majority of available monoclonal antibodies derived from tumor sources showed enhanced or even exclusive binding when the PDT*R motif was glycosylated with T_N- or T-antigen.[327,330] It was therefore proposed, that a MUC1 peptide vaccine would be more potent when glycosylated in the PDTR domain.[47]

Two- and three component vaccine candidates **1** and **2** and as well the gold-nanoparticle bound vaccine candidates **6** and **7** shared the same B-cell epitope, which include T_N-antigen glycosylation in the PDTR and GSTA region and vaccine candidate **8** contains an analogue B-cell epitope with T-antigen decoration. All of them showed strong binding to the B-cell epitope **232** and all mouse sera except one (vaccine candidate **2** (HC12), mouse 1) showed broad recognition dependent on PDT*R glycosylation (14 out of 15 sera). Glycosylation was not restricted to the T_N-antigen only, but also T-antigen and elongated core 3 trisaccharide and core 1 tetrasaccharide structures provide binding. To some degree even branched core 2 hexasaccharides were accepted in this domain. Thus, similar to the majority of monoclonal antibodies that have formerly been generated from tumor tissue or other MUC1 sources (e.g. human milk, urine), vaccines of this kind indeed induce specificity for glycosylation in PDTR. *Vice versa*, the specificity of the known monoclonal antibodies derived from PDTR glycosylated source material, suggests that the PDTR domain often is glycosylated in tumor cell MUC1 glycoproteins. It is known that all glycosylation sites of the MUC1 tandem repeat are potential targets for the polypeptide glycosyltransferases in cancer cells.[331] Also, some monoclonal antibodies were recently reported to have tolerance for T_N-, T-, and ST- glycoforms in PDT*R, screened on a MUC1 microarray system.[324] In the current thesis, a broad tolerance of glycoforms by the induced antibodies was shown not only for the T_N-, T- and ST_N-antigen structures, but to some extend for linear and to sometimes even branched type-1 and -2 elongated mucin core structures that are not classic tumor-associated antigens. Vaccine candidate **3** (HC11, *figure 4.62* and *4.63*), containing a B-cell epitope without glycosylation in PDTR induced three different antibody specificities in each of the three mice immune sera, of which only one recognized the unglycosylated PDTR domain (mouse 2), indicating that the glycosylated PDT*R indeed is more immunogenic than an unglycosylated PDTR and also more immunogenic than a single glycosylation on the serine of GS*TA.

4.3.10.2 Role of GSTA glycosylation

The results obtained in this work from the investigated murine antisera induced by vaccine candidates **1**, **2**, **6**, **7**, **8** containing a B-cell epitope with homogenous T_N- or T-antigen

glycosylation in PDT*R and GS*T*A, showed a very narrow and specific dependence on glycosylated GS*T*A. The two- and three component vaccines candidates **1** and **2** induced antibodies, which showed significant binding to clustered T_N-antigen glycosylation in the GS*T*A epitope (**233**). If the glycosylation in this domain was exchanged to clustered T-antigen glycosylation (**235**) antibody binding was lost. Double glycosylation in this domain was mandatory, since single glycosylation in GSTA, whether with T_N-antigen on serine or threonine (**236** and **237**) or a different core glycan on threonine, resulted in no or weak antibody binding. While glycosylation in the GS*T*A binding epitope needed to match the vaccine template, the PDT*R glycosylation often provided a broader spectrum of possible binding interactions and was therefore considered responsible for a major part of the antibody recognition. In contrast, vaccine candidate **3** (HC11) has a B-cell epitope only glycosylated on serine in the GS*TA region. The antibodies from two out of the three generated sera induced by vaccine candidate **3** (mouse serum 1 and 2) had peptide epitopes that neither relied on glycosylation nor on the GSTA domain. Only mouse 3 showed very specific reactivity for T_N-antigen on either serine or threonine in GSTA. The immunogenicity of the monoglycosylated GS*TA domain was inferior to that of PDT*R. In the di- and tripartite vaccines with clustered T_N-glycosylation in GSTA the immunogenicity was higher than monoglycosylation in GSTA of vaccine candidate **3** (HC11). The influence of GSTA domain glycosylation seems not to be very important in the case of the gold nanoparticle vaccine candidates **6**, **7** and **8**. Only 3 out of 9 mouse sera (vaccine **6** mouse 1, vaccine **7** mouse 1, vaccine **8** mouse 2) showed strong antibody interactions with double or monoglycosylated GSTA (e.g. **233**, **235**). Here, glycosylation in PDTR clearly dominates antibody binding in comparison to GSTA glycosylation, again highlighting the immunogenic relevance. In the AuNP based vaccines, the induced antibodies had a strong preference for the PDTR epitope in spite of the attached B-cell epitopes containing T_N-antigen in both the PDTR and GSTA region, proposing that the latter has a lower immunogenicity. *R. Cummings* and *D. Live* administered MUC1 tandem repeat sequences on KLH to breast cancer patients in remission, containing either T_N-antigen glycosylation in PDT*R or clustered in GS*T*A or in both epitopes.[332] They identified both epitopes as immunogenic, however, the vaccine candidates containing glycosylation in both epitopes, generated polyclonal immune sera with higher affinity to the glycosylated PDT*R domain.

It can be speculated, that the very narrow antibody specificity to clustered T_N-antigen glycosylation in the GSTA domain was the result of the conformational influence of the carbohydrate decoration. GS*T*A with consecutive glycosylation on serine and threonine was found to be locked in a rigid and extended structure. *F. Corzana et al*. conducted molecular dynamics (MD) simulation with a serine-threonine dipeptide with clustered T_N-antigen glycosylation (S*T*, as in the GS*T*A domain) and found that the glycans adopted

an orientation similar to the orientation of monoglycosylated serine-threonine dipeptides without vicinal glycosylation (ST* or S*T).[333] The vicinal glycans in diglycosylated peptides can occupy natural positions without sterical interferences. The superimposed MD structures showed low flexibility for this orientation, meaning that the glycans are fixed in their positions. Similar results were obtained by the groups *R. Cummings* and *D. Live*.[332] NOE signals observed through NMR spectroscopy of synthetic T_N-antigen glycosylated peptides were compared and peptides with vicinal, clustered T_N-antigen glycosylation showed fortified NOE interactions between the peptide backbone and the carbohydrate moieties, indicating more rigid structures for glycopeptide with clustered glycosylation. The group around *H. Kunz* showed that a PAPGS*T*APPA decapeptide with clustered T_N-glycosylation, adopts an extended “rod-like” structure.[334] The rigid peptide conformation with fixed and inflexible clustered T_N-antigen glycosylation in the GS*T*A domain of MUC1, may be recognized by the immune effector cells and translated into a very precise antibody design with high specificity for the template, but low flexibility for deviating structures.

4.3.10.3 Role of the VTSA domain

The VTSA domain of the vaccine candidates was not glycosylated and no antibody reactivity against the VTSA peptide domain, whether glycosylated or not, was observed here, with the exception of vaccine candidate **3** (HC11) mouse 1. The B-cell epitope of vaccine candidate **3** had one T_N-antigen in the GS*TA domain and produced different antibody specificities in all three mice.

In the cases of the other vaccine candidates, glycosylation on threonine of VT*SA was here shown to decrease antibody reactivity to the glycosylated PDT*R motif in multivalent glycopeptides. PDT*R-specific antibodies often tolerated additional glycosylation in GST*A, without (almost) any decrease in binding affinity. Additional VT*SA glycosylation, however, often reduced binding to PDT*R. *Hanisch et al.* described disturbed antibody binding of monoclonal antibodies directed to unglycosylated PDTR in the presence of glycosylation in the VTSA region.[335] *Dziadek et al* demonstrated that glycosylation with a α2,6-ST-antigen on Thr in VT*SA resulted in a long range effect on the conformation of the unglycosylated PDTR turn structure.[54] The results from this work are in accordance with these findings, but appending that a conformational influence from VT*SA also influences glycosylated PDT*R domains and thereby modify antibody recognition.

4.3.10.4 Influence of glycan size, core branching and sialylation

As stated above, elongated core 3 and core 1 structures in general do not obstruct antibody binding to the immuno-dominant PDTR motif and the antibodies directed to the PDTR epitope tolerate to some extend multivalent glycosylation on other glycosylations sites of MUC1, containing these glycans. A limitation in the recognition of the glycopeptides was set by the branched core 2 hexasaccharide structures. For direct comparison, glycopeptide **174**, modified with a type-2 core 2 tetrasaccharide in PDT*R was added to microarray **MA3** along with the corresponding linear core 1 tetrasaccharide sequence **108,** core 2 hexasaccharide sequence **125** and the core 1 T-antigen peptide **76** (diagrams for **MA3** in the corresponding *chapter 4.3.8*). Disaccharide decorated peptide **76** was usually recognized well and linear elongation with a lactosamine unit to form a core 1 tetrasaccharide as in glycopeptide **108** was detected as good or moderately weaker. If glycan elongation with the lactosamine was in 6-position of the αGalNAc instead, to form the equivalent branched core 2 tetrasaccharide, recognition dropped significantly. Furthermore, if the core 2 glycan was further elongated, to give branched core 2 hexasaccharide glycopeptide **125**, no further loss of recognition was accordingly observed. In fact, the core 2 tetrasaccharide, hexasaccharide and even disialylated octasaccharide on glycopeptide **201** (**MA4**, *figure 4.73*) decorated peptides showed approximately the same binding intensities by the induced antibodies (*figure 4.74*).

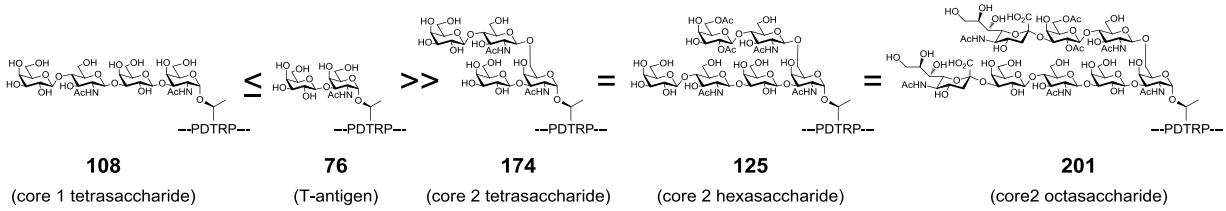


***Figure 4.74*:** General binding affinities of the PDT*R-specific serum antibodies depending on the *O*-glycosylation in this peptide domain.

This demonstrates that linear core 1 structures are preferred over branched core 2 glycans in the immunodominant PDTR domain. The antisera induced by T_N-antigen decorated vaccine candidates **1**, **2**, **6, 7, 8** show in general almost stable antibody recognition in terms of intensity for linear di-, tri- and tetrasaccharide glycan structures in the PDTR domain (diagrams for **MA2** section **B** in *chapter 4.3.8*), but a significant drop in antibody recognition for the core 2 hexasaccharides. The reduced antibody preference for β1,6-branched glycans could either be a steric effect due to the branching or a conformational alteration of the secondary structure of the peptide backbone induced by the glycan. A αGalNAc in the PDT*R motif results in a conformational change in the PDTR turn structure.[53,331] It is still not fully investigated how further core elongation may change the peptide conformation.

Previous NMR studies of extended glycosylation on non-MUC1 mucin glycopeptides showed that further glycan elongation, beyond the initial αGalNAc, does not have a great impact on the backbone conformation,[336] including elongation in a core 2 fashion.[337,57] On the other hand, *H. Kunz* showed, that peptides from LI-cadherin with various glycans attached, exhibited glycan dependent conformational changes in similar NMR experiments, when glycosylation was further extended on the αGalNAc monosaccharide.[338] Based on the coupling constants and NOE coupling in NMR experiments, the group of *S.-I. Nishimura* reported that distal sialylation on an extended core 2 glycan at the PDT*R domain in a MUC1 tandem repeat, had remote effects on the conformation in the domain, while sialylation of glycans in the VTSA and GSTA sites had no conformational effect on these glycosylation sites.[57] Conformational effects are thus most likely sequence dependent and can therefore be a plausible reason for the reduced antibody affinity to the core 2 glycosylated PDT*R motifs. It may indicate that additionally to the conformational change in the PDTR domain, caused by glycosylation with the basic αGalNAc, the βGlcNAc linked to the 6-position of αGalNAc is also influencing the peptide backbone conformation. In contrast, elongation in 3-position of the αGalNAc with β1,3Gal (core 1) or β1,3GalNAc (core 3) may not have a comparable conformational influence on the peptide, due to the similar binding affinities of the antibodies towards the core structures.

The findings are interesting, since several studies, for instance with breast carcinoma cells, indicate that core 1 glycan decoration is more abundant than core 2 glycosylation on tumor cells, while core 2 glycans decoration is often found on healthy cells.[30,321,339] The aberrant glycosylation on tumor cells has been associated with decreased expression of core 2 1,6-*N*-acetylglucosaminyltransferase 1 (C2GnT-1) and upregulated expression of human α2,3-sialyltransferase 1 (ST3Gal1) during oncogenesis.[320] Further, high contents of secreted and shedded MUC1 carrying core 1 based glycans have been found in the sera of advanced breast cancer patients.[34] Therefore it would be highly eligible, to produce antibodies that show higher affinity to tumor related core 1 glycoproteins on various adenocarcinomas.[340] To some extent, this feature appears to be inherent in the presented vaccine candidates. The aberrant glycosylation on tumor cells is sometimes also caused by mutation of the Cosmc chaperone essential for T-synthase activity resulting in presentation of T_N- and ST_N-antigen modified epitopes, structures which are also covered by the antibodies induced by the above described vaccines.

Vaccine candidates **7** and **8** share the same B-cell epitope sequence with T_N- and T-antigen as haptens, respectively, on a gold nanoparticle support. The three sera originating from the T-antigen glycosylated vaccine candidate **8** seemed to have a slightly higher tolerance for the sequences with the bulky core 2 hexasaccharides than the T_N-antigen vaccine **7** (*figures*

4.69 and *4.71*, **B**). It appears reasonable that the slightly bigger T-antigen hapten induced antibodies with specificity for somewhat larger glycans. This study clearly shows that there is a size tolerance for larger carbohydrates than the vaccinated haptens, at least in the immuno-dominant PDTR motif. This may in particular be problematic if the antibody discrepancy between core 1 and core 2 structures is lost with the risk of generating an immunological memory, also to normal healthy cells. The larger tumor-associated antigens, such as the ST- (Neu5Acα2,3-Galβ1,3-GalNAcα) trisaccharide or the di-ST-(Neu5Acα2,3-Galβ1,3-[Neu5Acα2,6]-GalNAcα) tetrasaccharide may therefore not be the optimal choice for vaccination. Due to the potentially higher immunogenicity of the negatively charged sialic acid residues in these structures and the fact that sialylation seems to influence the MUC1 peptide backbone conformation, these tumor-associated glycan structures may thus behave different from the above described T-antigen vaccine candidate **8**. There is no defined threshold yet, that marks the borderline between tumor-associated and healthy, regarding the size and valency of carbohydrates that obstruct antibody epitopes, such as the PDTR domain. Information of this kind is important because of possible autoimmunity, induced by an inadequate, potential therapeutic vaccine.

Finally, no significant differences in antibody recognition were monitored for type-1 or type-2 core elongation. In general, the type of carbohydrate backbone elongation by type-1 or type-2 LacNAc does not seem to have conformational influences on the tandem repeat peptide of MUC1. Only one antiserum from vaccine candidate **6** mouse 3 (*figure 4.67*, **A**+**B**) showed strong discrepancy for binding to peptides, glycosylated with core 2 type-1 hexasaccharides in PDT*R. Since the corresponding core 1 type-1 tetrasaccharide and the core 2 type-2 hexasaccharide were not recognized, the antibodies must have a preference for the peptide epitope conformation induced by core 2 glycosylation with further β1,3-linked elongation, or directly to the carbohydrate structure itself.

4.3.10.5 Antibody binding of multivalent peptides with glycosylation in different domains

Trivalent peptides with di- (basic core 1), tri- (extended core 3), tetra- (extended core 1) and pentasaccharides (sialylated, extended core 1) in all three glycosylation sites, still exhibit accessible epitopes for the induced antibodies. However, trivalent core 2 hexasaccharide decoration in all three glycosylation sites (HGVT*+PDT*R+GST*A, diagrams of **MA2**, section **B** in *chapter 4.3.8*) effectively prevents antibody binding and was therefore regarded as a model structure for a mucin peptide/protein backbone masked by extensive *O*-glycosylation.

These results suggest that multiple branched core 2 glycans are more effective in shielding the mucin proteins than linear core 1 or 3 glycans.

4.3.10.6 Influence of the T-cell-epitopes, mitogens and carriers on antibody binding

All vaccine candidates, except for number **3**, **4**, and **8**, share the same B-cell epitope sequence (HGVTSAPDT*RPAPGS*T*APPA, * = T_N-antigen). However, differences in the antibody binding patterns were observed. Most obvious was the general low antibody specificity of the tetanus toxoid conjugated candidates **4** and **5** (*chapter 4.3.8.4* and *chapter 4.3.8.5*). Basically all sequences were recognized to a good extent, no matter if glycosylated or not. Only sequence **213**, missing the PDTR domain, showed no reactivity with the antibodies. This clearly indicates that the tetanus toxoid vaccine induced antisera were specific for the PDTR sequence, and therefore also MUC1-specific. In this case, reduction of antibody binding to the large core 2 hexasaccharide peptides was observed, but was less pronounced. Still, trivalent hexasaccharide peptides were shielded by the glycans from antibody binding. To conclude, the tetanus toxoid vaccine candidates elicited uniform antibody specificity for MUC1 tandem repeats with an antibody binding epitope dependent on the PDTR domain and to a large extend independent from the glycosylation status. A cutoff for larger carbohydrate core 2 structures, which are found on healthy cells, was still observed. The glycan decoration of the B-cell epitope sequence, conjugated to the tetanus toxoid vaccine candidates, does not have great influence on accentuated antibody binding profiles with the glycopeptide library presented here, as can be seen by comparing vaccine candidate **4** and **5**. Carrier proteins are known to produce a larger spectrum of antibodies also against peptide epitopes on the carrier protein, which in the worst case can overrule the hapten specific immune response.[129,130] This should mainly influence the antibody titer of the immune response against a certain antigen. Here it seems that the strong immune response induced by the tetanus toxoid stimulant also induces less accentuated antibody specificity. The extraordinary strong immune response is combined with an expense of antibody binding specificity. Anyhow, as demonstrated by *Gaidzik et al.*, mouse 5 of tetanus toxoid conjugated vaccine candidate **4** (NG5), tested on the MUC1 microarray in this work, elicited IgG-antibodies in mice, that specifically showed high reactivity against advanced tumor cells (G3-phase) in mammary carcinoma tissue sections, but only weak recognition of tumor tissue in the early phase (G1-phase).[126] This could be the direct influence of the observed core 2 affinity cutoff (more normal core 2 glycans on early phase tumor cells) or just a result of the dramatically increased expression of the MUC1 glycoprotein on the advanced tumor cells or

a combination of both effects. Such an immune response, combined with the general high antibody titers observed, may just be right to create an adequate therapeutic immune reaction or might be borderline, regarding the risks of inducing autoimmune reactions against normal cell surface glycosylation.

The two- and three-compound vaccine candidates **1**, **2** and **3** produced sera that were not uniform in the induced antibody profile. A minority of the antibody sera showed very specific binding to a few sequences including site specific T_N-antigen glycosylation according to the induced antigen structure, while others showed showed broad recognition of differently glycosylated domains. The attached mitogen Pam_3CSK_4 in vaccine **2** and **3** does not have an obvious positive effect on the antibody profile compared to candidate **1**. Thus, a larger variance of isotype antibody classes are formed by immune stimulation with the build-in adjuvant Pam_3CSK_4 vaccines, which might be favorable for an efficient immune response directed to tumor cells.

The P30-AuNP conjugated vaccines **7** and **8** on the other hand, presented a more uniform induced antibody response in all of the mice. The immune sera were not only specific for the glycoform of the B-cell epitope, but showed in general broad acceptance of various PDTR glycosylated sequences. Generation of uniform immune reactions is an important feature of a safe vaccine.

4.3.11 Binding of plant lectins to MUC1 glycopeptides on a microarray

4.3.11.1 Glycopeptide microarray analysis with plant lectins

Plant lectins are commonly used in carbohydrate structure analysis to detect specific recognition elements on glycoproteins or glycolipids. Specific lectins are also used in glycobiology and glycoproteomics for enrichment of certain glycoproteins and glycopeptides by lectin weak affinity chromatography (LWAC).[341] The presented mucin glycopeptide microarray platform was used to evaluate interactions with lectins known to recognize specific carbohydrate binding elements. Effects on glycan presentation at different glycosylation sites, presentation in different core structures and influences by multivalent ligand effects in the lectin interactions were evaluated. A library of 132 MUC 1 glycopeptides was screened on microarray format **MA5**. *Figure 4.75* only shows the positive binders of the library (for complete peptide list see *chapter 6.3.4* and complete data *chapter 8.3.4*).

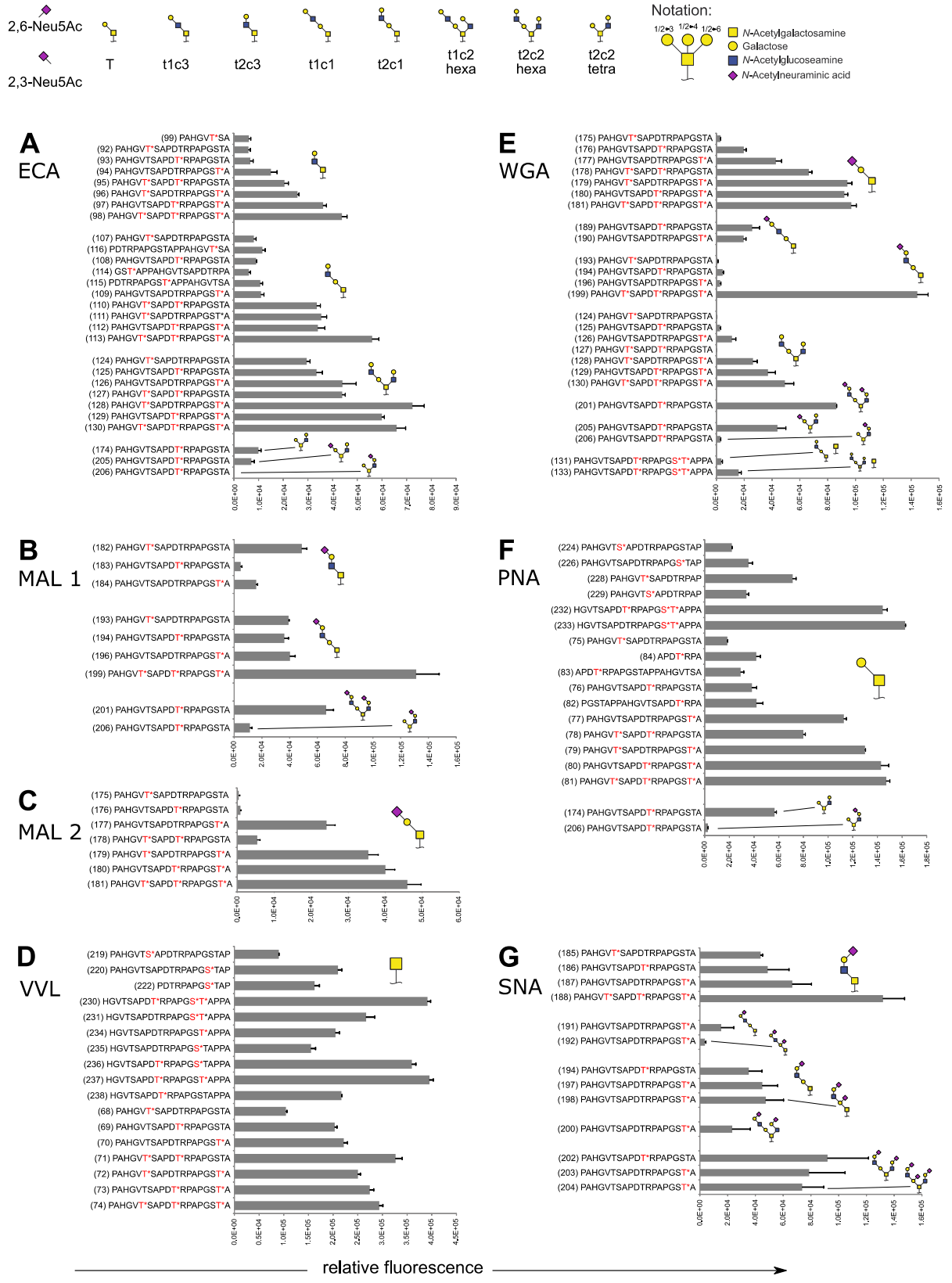


***Figure 4.75*:** Binding of different plant lectins on the glycopeptide array platform **MA5**. **A**: *Erythrina cristagalli* (ECA, 100 µg/mL), **B**: *Maackia amurensis* I (MAL I, 80 µg/mL), **C**: *Maackia amurensis* II (MAL II, 20 µg/mL), **D**: *Vicia villosa* (VVL, 50 µg/mL), **E**: Wheat germ agglutinin (WGA, 50 µg/mL), **F**: Peanut agglutinin (PNA, 10 µg/mL), **G**: *Sambucus nigra agglutinin* (SNA, 20 mg/mL).

All lectins recognized the glycopeptides according to their reported carbohydrate binding specificities (*figure 4.75*).[342] Additionally to HPLC and MS characterization, lectin binding validates the enzymatic α2,3- and α2,6-sialylation performed on glycopeptide level. The observed lectin binding specificities to the mucin core glycopeptides found on the microarray were as follows:

A *Erythrina cristagalli* (ECA)

ECA was reported to recognize terminal Galβ1,4-GlcNAc (type-2 LacNAc) structures and accordingly showed affinity to the type-2 terminated core 1, core 2 and core 3 glycans of the MUC1 glycopeptides. As seen best in the core 3 glycosylated peptides **92**-**99**, multivalent presentation of the glycans on the peptide backbone multiplies binding affinity. Branched core 2 glycans with two terminal LacNAc extensions are recognized exponentially stronger than the corresponding unbranched core 1 or core 3 glycans. α2,3- or α2,6-Sialylation blocks LacNAc recognition. Type-1 LacNAc terminated glycans were not recognized, as expected.

B *Maackia amurensis I* (MAL I)

MAL I was previously reported to recognize the terminal α2,3-sialylated type-2 LacNAc trisaccharide structures. Evaluation on the glycopeptide array showed that the sialyl-LacNAc core 3 was a particular good ligand when presented in the VT*SA domain of peptide **182**, compared to the other glycosylation sites, indicating a conformational interaction of glycan and peptide backbone. The α2,3-sialylated, type-2 extended core 1 glycans of peptides **193, 194** and **196,** which were not bound by WGA (**E**), were recognized by MAL-I, as expected. Furthermore, the sialylated LacNAc was recognized better, if attached in a linear mode, e.g. in an extended core 1 glycopeptide (**194**), than in a branched fashion as a core 2 glycopeptide (**206**).

C *Maackia amurensis II* (MAL II, MAH)

MAL II was reported to recognize α2,3-sialylated T-antigen structures and binds to the ST-antigen glycopeptide motives on the microarray as expected.

D *Vicia villosa* (VVL)

VVL was reported to recognize terminal GalNAcα as in the T_N-antigen and has been used in glycoproteomics to enrich GalNAcα glycopeptides from *Simple cell* lines. *Simple cell* lines were recently reported to be very useful in identification of *O*-glycosylation sites, which is rather difficult with standard glycomics and proteomics techniques.[16] Glycopeptide microarray analysis showed that the VVL lectin was very specific to the T_N-antigen glycopeptides. Minor preference for threonine over serine

glycosylation was seen, e.g. **234** (Thr) versus **235** (Ser), which is in accordance with enzyme-linked lectin assay (ELLA) results by *Madariaga et al.*, although the difference here on the microarray was not so pronounced.[343]

E Wheat germ agglutinin (WGA, *Triticum aestivum*)

WGA was reported to recognize terminal GlcNAcβ and sialic acid residues. The lectin is commonly used to enrich sialylated *O*-glycans, *N*-glycans and *O*-GlcNAcylated peptides from glycoproteomic tryptic samples. Here, the α2,3-sialylated T-antigen glycopeptides were very good ligands, better than terminally sialylated type-1 or -2 LacNAc glycan conjugates. In fact, the sialyl type-1 extended core 1 peptides **189** and **190** were weak to medium binders, while the type-2 analogs **194** and **196** show very low lectin binding. A 1,3-glycosidic bond between the Gal and the GlcNAc seems to be favored. The peptides with α2,3-Neu5Ac on the type-2 core 3 glycosylated peptides were not recognized. Even more surprising, the triglycosylated linear sialyl type-2 core 1 glycopeptide **199** was recognized very strong, although all the corresponding monovalent glycopeptides showed almost no affinity. This is an extreme example for lectin avidity. A significant lectin binding was also observed to multivalent, not sialylated type-2 extended core 2 hexasaccharide glycans on peptides **128**-**130**. Binding was probably mediated by multivalent presentation of internal GlcNAcβ of this glycoform, although LacNAc is reported to be a poor ligand. The analog type-1 hexasaccharide modified peptides were not recognized. Glycopeptide **133** with an unsialylated core 2 type-2 hexasaccharide and GalNAcα at another glycosylation site, was weakly recognized, hinting on a cooperative effect between these glycans, which did not result in lectin binding itself on different peptides. In summary, additionally to the known binding specificity for sialic acids and terminal GlcNAcβ, WGA also interacts with various core 2 glycoforms which do neither posses terminal sialic acids nor terminal GlcNAcβ. Further, the connectivity between a sialylated Gal and HexNAc influences the binding, for instance a single α2,3-sialylated type-2 LacNAc was a poor ligand, while the equivalent type-1 LacNAc showed significant binding. Also, only α2,3-linked Neu5Ac was recognized and no affinity for α2,6-Neu5Ac was observed.

F Peanut agglutinin (PNA, *Arachis hypogaea*)

PNA was reported to recognize T-antigen structures and binds to the array as expected. Threonine glycosylation was preferred over serine glycosylation, as shown in glycopeptides **228** and **229**. Core 2 glycosylation on the T-antigen, as in peptide

174 did not prohibit recognition. On the other hand, if the core 2 was α2,3-sialylated on the LacNAc side chain, as in glycopeptides **206**, lectin recognition was prohibited.

G *Sambucus nigra agglutinin* (SNA)

SNA was reported to recognize terminal Neu5Acα2,6-Gal and as expected bound to glycopeptides containing this structure on the microarray. α2,6-Sialylated type-2 LacNAc is thereby preferred over sialylated type-1 LacNAc (e.g. **191** vs. **197**). Also, internal α2,6-sialylation next to the type-1 prevents binding (**192**) , which is not the case when the internal sialylation is next to a type-2 structure (**204**). No binding to Neu5Acα2,6-GalNAcα or 2,3-sialylated structures was observed.

In summary, all relevant glycoforms were addressed by the corresponding lectins and unique fine specificities were sometimes observed.

By comparing the different glycosylation sites when decorated with the identical glycan, it became apparent that most lectins recognized the glycan stronger when attached to the glycosylation site at the free C-terminal end of the peptide and *vice versa* weaker if the glycan was attached to the glycosylation domain close to the N-terminal end, where the peptide was linked to the polymer of the hydrogel. In most of the glycopeptides, the glycosylation site at the C-terminal end was the GST*A domain and at the N-terminal end the VT*SA motif. This was seen very clearly in the case of ECA (*figure 4.75*, **A**, e.g. **124** vs. **125** vs. **126**), WGA (*figure 4.75*, **E**, e.g. **175** vs. **176** vs. **177**) or MAL II (*figure 4.75*, **C**, e.g. **175** vs. **176** vs. **177**). In fact, the only lectin that does not follow this trend is MAL I (*figure 4.75*, **B**, **184** vs. **185** vs. **186**). It was not fully clear though, if this was a result of a sterical effect and that the lectins could simply approach glycans at the loose end of the peptide better than at the surface bound end. Alternatively, it could be a real conformational effect by which glycans are presented in a favorable way by the GSTA domain. Peptides with the VT*SA or PDT*R glycosylation sites present at different distance from the N-terminal end, were compared, e.g. **107** and **116** (*figure 4.75*, **A**, ECA, VT*SA) or **76**, **82** and **83** (*figure 4.75*, **F**, PNA, PDT*R). An effect of reduced binding by the lectins if the glycan was closer to the immobilized N-terminal site was observed, although the difference in signal intensity is not large. To fully evaluate these effects, more peptides need to be synthesized.

4.3.12 Binding of galectin-3 to MUC1 glycopeptides on a microarray

4.3.12.1 Introduction on the binding specificities of galectin-3

Galectin-3 is involved in intracellular, carbohydrate-independent processes, which have in general anti-apoptotic effects and are therefore promoting cancer development. On the other hand, extracellular galectin-3 is recognized by protein receptors in a carbohydrate-dependent way and is suspected to have apoptotic effects on cells.[344] For example, T-cells were shown to become apoptotic upon contact with galectin-3. Various kinds of tumor cells overexpress galectin-3 and it is believed that cancer cells gain advantage from intracellular anti-apoptotic galectin-3 effects with simultaneous immune suppression by expressed extracellular galectin-3.[93] There is evidence that tumor cells protect themselves from the extracellular apoptotic effects by aberrant cell surface α2,6-sialylation. Furthermore, the characteristic surface glycosylation of cancer cells seems to be correlated with cancer cell progression.[79,83] The MUC1 glycoprotein is overexpressed on the surface of cancer cells and further exhibits tumor-associated glycosylation. Galectin-3 was observed to interact with the T-antigen structure on MUC1 and is probably involved in the extracellular, pro-metastatic effects, mediated by galectin-3.[79] The established MUC1 tandem repeat glycopeptide array provides a platform for elucidation of galectin-3 binding preferences, facing typical tumor-associated glycan motifs, differently elongated *O*-glycan core structures, sialic acid termination and multivalent presentation branching on the glycan core structure or glycan decoration on several glycosylation sites. As a future perspective, such a glycopeptide array may provide a useful platform in the search of potent glycomimetics, such as competitive galectin-inhibitors.[345,346]

Galectins share a common preference for β-galactose, due to conserved sequence homology in their carbohydrate recognition domain (CRD). X-ray crystallography of the conserved galectin-3 CRD, complexed with Lac/LacNAc or T-antigen disaccharides, identified a binding groove with additional space at the non-reducing end of the carbohydrates.[347,348] The suggested binding model comprises five subsites, namely A, B, C, D and E.[349] Sites A-D cover an area large enough for a tetrasaccharide, while E can hold further saccharide residues or parts from the connected protein or lipid. Lactose or LacNAc as the central binding motifs for galectin CRDs, are located in sites C-D. The C-site features the mandatory contact to the axial 4'-OH and the 6'-OH of the galactose.[347] The less restricted D-site houses the Glc, GlcNAc or GalNAc and contributes to overall binding (*figure 4.76*).

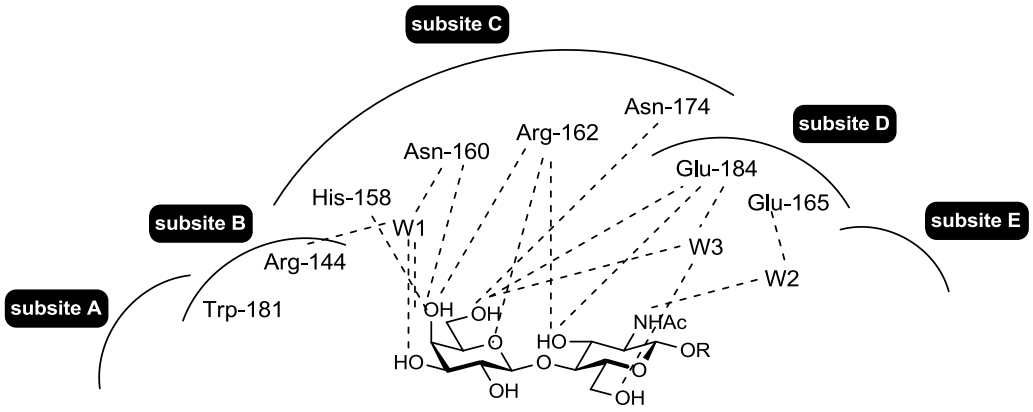


***Figure 4.76*:** Model of galectin-3 subsites with hydrogen bonding to *N*-acetyllactosamine (LacNAc). Sites C-D house the central disaccharide. Sites A, B and E are vacant and can be occupied by modifying residues. W1-3 represent water molecules, participating in the LacNAc binding.

The A, B and E sites vary in preference for the residues around the central Lac/LacNAc for the different members of the galectin family. Like this, the various galectins interact with different interaction partners and are able to fulfill distinct actions. Identifying the characteristic interaction partners and unique specificities of the galectins is therefore necessary to understand galectin functions.

Several methods have been applied to study lectin-carbohydrate interaction and affinity of galectins to various biologically relevant carbohydrate structures. Solution-phase assays, such as fluorescence polarization,[350] isothermal titration calorimetry (ITC),[351] frontal affinity chromatography (FAC)[67] and flow cytometry[352] have been used to evaluate or quantify binding affinities to a specific ligand in solution. On the other hand, immobilized ligands on a surface permit multivalent interactions, crucial for avidity in biological processes. Carbohydrate microarrays,[353,354,355,356] ELISA-like plate assays[357] and SPR[353,358] have been used to study galectin specificity. In the published literature, the data obtained from solution-phase or surface assays are partly inconsistent within different studies of galectins.[359,360] In the case of galectin-3 for example, the group around *R.D. Cummings* screened galectin-3 on a glycoarray by the *Consortium for Functional Glycomics* (CFG).[353,361] The array failed to detect type-1 and type-2 LacNAc, the sialylated versions thereof, type-2 LacNAc terminated core 2 and core 4 glycans and the T-antigen, although the LacNAc disaccharide has before already been identified as the central galectin binding motif by other methodologies. *H. Tetano et al*. presented a glycoarray with glycans attached to a polyacrylamid backbone immobilized on epoxy slides on which binding to type-1 but not to type-2 LacNAc was observed.[355] The *Seeberger* group immobilized glycans with a thiol-linker onto maleimide-functionalized slides and observed stronger recognition of Lac than of LacNAc,[356] although binding affinity was reported the other way around by several other studies. In general, the

way of presentation and the linker strategy, ligand concentrations, galectin preparation and detection method, may determine the result of galectin ligand profiling significantly.[350,353,356,362]

Ultimately, studying specificity of galectin-3 to carbohydrates linked to tandem repeat peptides of the natural binding partner MUC1 was consequently reasonable, since the glycans are presented in a more natural way on the peptide backbone and in a multivalent mode. Also, no detailed study of extended *O*-glycan core structures as ligands for galectin-3 was reported. The CFG glycoarray platform also features some shorter core structures, but the results with galectin-3 were not conclusive since LacNAc terminated *O*-glycans as well as the known ligand LacNAc itself failed to show binding in previous studies.[353,361]

4.3.12.2 Incubation of the MUC1 microarray with galectin-3

The MUC1 microarray **MA4** was incubated with recombinant (His-tagged) human galectin-3 expressed in *E.coli*. The lectin was detected by a secondary *Alexa Fluor 488* anti-mouse/human Mac-2/Galectin-3 antibody (dilution 1:500). Galectin-3 was applied in dilutions from 25 µg/mL (ca 0.89 µM) to 0.79 µg/mL (ca 28 nM) in PBST (0.2%) buffer (*figure 4.77*). Normal physiological serum levels of galectin-3 are around 10 ng/mL and around 100-500 ng/mL in colorectal and breast cancer patients.[95]

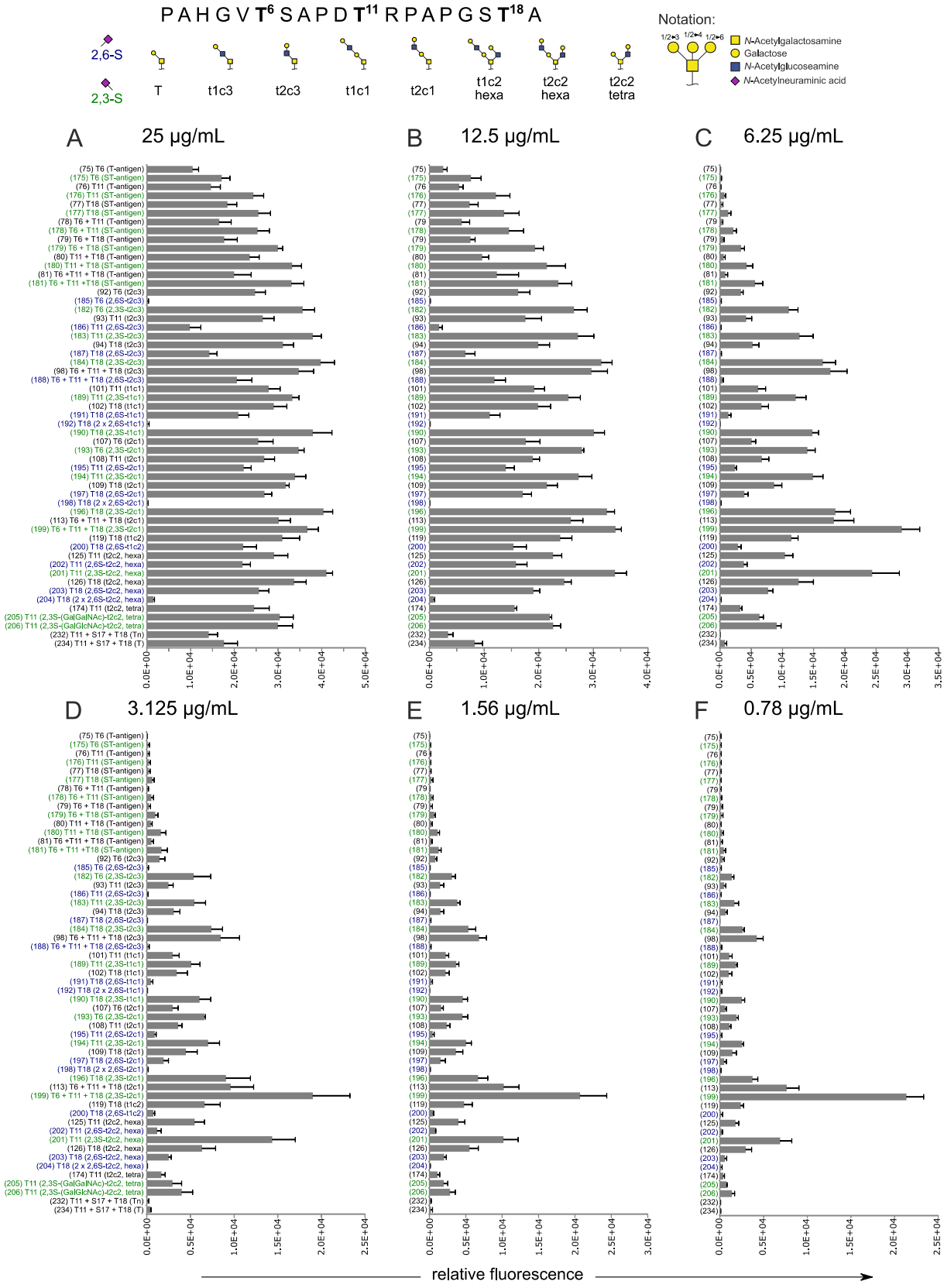


***Figure 4.77*:** Galectin-3 binding of MUC1 glycopeptide tandem repeats. Incubation with **A:** 25 µg/mL, **B:** 12.5 µg/mL, **C:** 6.25 µg/mL, **D**: 3.13 µg/mL, **E**: 1.56 µg/mL, **F**: 0.79 µg/mL of galectin-3 in PBST (0.2%) buffer. Unsialylated glycopeptides are depicted in black, α2,6-sialylation in blue and α2,3-sialylation in green.

4.3.12.3 Galectin-3 binding to unsialylated type-1 and type-2 LacNAc structures

At 25 µg/mL (*figure 4.77*, **A**) most of the surface bound glycopeptides were apparently saturated with galectin-3. Pronounced discrimination of binding specificities was best seen at concentration around 6.25 µg/mL (*figure 4.77*, **C**). At the lowest tested concentration of 0.78 µg/mL (*figure 4.77*, **F**), the three highest signals belong to glycopeptides with multivalent glycan presentation at the peptide backbone. Of those, entries **110** and **199** carry three LacNAc extended core 1 glycans with α2,3-sialylation on the latter and entry **201** presents an octasaccharide with double α2,3-Neu5Ac-LacNAc termination. Interestingly, at high galectin-3 concentrations (25 µg/mL and 12.5 µg/mL) glycopeptide **232**, decorated with three αGalNAc monosaccharides, was also recognized by galectin-3. The recognition of the sole T_N-antigen was not described before in the literature. Glycopeptide **234**, has the same glycosylation pattern as **232**, but with three T-antigen (Galβ1,3-GalNAc) structures instead, which both showed only medium recognition at 12.5 µg/mL. The tumor-associated T_N- and T-antigen glycans decorating the mucin glycoproteins, might therefore be low-affinity binding partners for galectin-3 in cancer tissue, were elevated levels of this structure can be found on overexpressed mucins.

The tumor relevant T_N- and T-antigen were as well recognized at high galectin-3 concentrations. However, LacNAc terminated glycans represent better binding partners, whereupon type-1 and type-2 LacNAc glycans were equally well bound. LacNAc being attached to the basic core 1 in a linear fashion (**108**) was preferred over LacNAc branched in a regioisomeric core 2 structure (**174**) (*figure 4.78*). The core 2 hexasaccharide (**125**) offering two terminal LacNAc units showed thus strongest binding, while T-antigen (**76**, basic core 1) was recognized only with high concentrations of galectin-3.

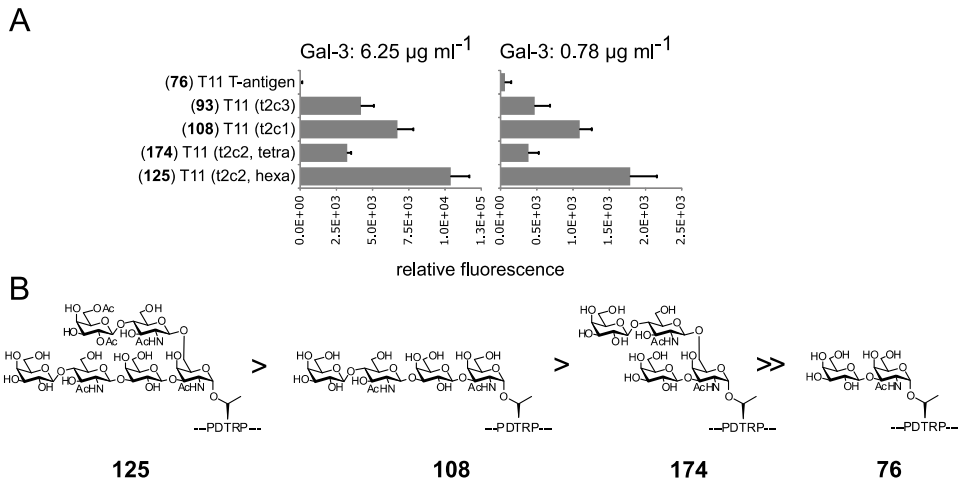


Figure 4.78: **A:** Extract from full array **MA4** (*figure 4.77*). Comparison of galectin-3 binding to LacNAc present in linear and branched *O*-glycan core structures at galectin concentrations of 6.25 µg/ml and 0.78 µg/ml. **B:** Structures of the compared glycans in order of recognition affinity by galectin-3.

Glycopeptide **174** presents two terminal galactose units, while glycopeptide **108** has one terminal galactose and one internal galactose unit. However, the long carbohydrate chain of **108** (extended core 1) showed better recognition by galectin-3, than **108** (branched core 2). Therefore, the long chain variant further stabilizes the binding to the galectin by filling the vacant pockets of the binding groove, either by filling the A-B sites with the terminal LacNAc and the internal Galβ1,3GalNAc in C-D or with LacNAc in C-D and the internal galactose in the E site.

Gabius et al. tested the binding of human galectin-3 to a selection of glycoproteins by their inhibitory potency on enzyme-linked lectinosorbent assay (ELLSA). It was found that LacNAc, as a terminal epitope in core 2 glycans on hog gastric mucin, possessed lower inhibitory potency relative to glycoproteins with terminal LacNAc in core 1 glycans, which is in accordance with the trend observed on the MUC1 glycopeptides.[363]

4.3.12.4 Influence of α2,3-sialylation of terminal LacNAc on galectin-3 binding

Further, direct influence of α2,3-sialylation was observed. According to the galectin binding model, modification on the terminal Gal residue in binding pocket B (*figure 4.76*) of the binding groove, was reported to have in many cases an affinity-enhancing effect on galectin-3 binding. Such enhancing effects were mediated by 3'-O-α/β-galactosylation and 2'-O-fucosylation of the galactose unit in pocket C.[353,355,356] These kinds of modifications are for instance found in the blood group antigens. The reports about α2,3-sialylation in this position are however varying. While in some cases no significant influence on galectin-3 binding was found, some reports document a decrease or even complete blocking of galectin-3 binding.[351,352,353,364,365] However, sialylation with terminal α2,3-Neu5Ac on type-1 or type-2 LacNAc or T-antigen structures on the MUC1 glycopeptide platform (*figure 4.77*, green entries), increases the binding affinity compared to the corresponding unsialylated glycopeptides (in black). Within the medium concentrations 6.25-1.56 µg/ml, the signal intensity doubled by α2,3-sialylation compared to the neutral glycans (e.g. **93** vs. **183**, **108** vs. **194**, **125** vs. **201** etc., *figure 4.77*). The triglycosylated, core 1 decorated peptide **199** with α2,3-sialylation was in particular well recognized, even in the lowest applied galectin-3 concentration (0.78 µg/ml). Clustered, α2,3-sialylated *O*-glycans, and probably also *N*-glycans, would form formidable epitopes for galectin-3 according to this binding behavior. The reason why α2,3-sialylation was previously found to have either none or affinity decreasing effects in the former reports is unknown, though different assays and ligand substrates have been applied. The groups of *R.D. Cummings* and *J. Hirabayashi* used glycoarrays to assess the specificity of galectin-3.[353,355] The former makes use of the glycan

array from the *Consortium for Functional Glycomics* (CFG) which failed to identify the basic well-known binding motif LacNAc and *O*-glycan core structures terminated with LacNAc even at high galectin-3 concentrations (1-10 µM).[361,353] The second study strangely showed binding only to type-1 LacNAc, but not to type-2 LacNAc as galectin-3 binding partner. Other studies made use of ITC with galectin-3 and oligosaccharides,[351] cell aggregation assays with galectin-3 and breast cancer cell lines,[365] flow cytometry with chinese hamster ovary (CHO) glycosylation mutants,[352] or co-immunoprecipitation of galectin-3 with MUC1 and EGFR.[364] In all previous examples, α2,3-sialylation was proposed to have none or even a minor reducing effect towards galectin-3 binding. The influence of sialylation on T-antigens on mucins, expressed by cancer cells, is in particular interesting, since this glycoform has been associated with galectin-3 assisted metastasis of tumor cells.[79] In this working model, overexpressed MUC1 on the cancer cell surface shields adhesion molecules needed for cell-cell interaction in metastasis. Galectin-3 binds the tumor-associated T-antigen epitopes on the mucins, and locally concentrates MUC1, which liberates the adhesion proteins in other areas for cell-cell-interaction. It was shown that interaction of galectin-3 with immunoprecipitated MUC1 from HT29 colon cancer cells lines vanished, when treated with *O*-glycanase (T-antigen specific) or became intensified when sialic acids were removed by treatment with an unspecific 2,3-, 2,6-, 2,8-sialidase (*A. ureafaciens*).[79] It was concluded that only unsubstituted T-antigen was responsible for binding. However, the presence of α2,6-Neu5Ac on the T-antigen could have been responsible for diminished binding (see next chapter), while the tumor-associated 2,3-ST-antigen, according to the results of this work, is in fact a better epitope for galectin-3. It is also not known, if a α2,6-Neu5Ac located on the αGalNAc of the T-antigen would prevent galectin-3 from binding. The glycopeptides on this array were chemically precisely defined and binding to the different carbohydrate structures was validated with the according lectins, which gives confidence regarding the found specificities.

4.3.12.5 Influence of α2,6-sialylation of terminal LacNAc on galectin-3 binding

In contrast to α2,3-sialylation, binding of galectin-3 to α2,6-sialylated glycans was considerably reduced (*figure 4.76*, blue entries). According to the galectin binding model, a free 6'-OH on galactose is necessary for hydrogen bonding. Whether type-1 or type-2 LacNAc was terminated by α2,6-sialic acid, did not make a difference. At high galectin-3 concentrations (25-12.5 µg/mL) also α2,6-sialylated glycans were recognized to some extent, the three entries with further internal α2,6-linked sialic acid on the basic core 1 galactose (peptides **192**, **198** and **204**) were completely unavailable for galectin-3 binding. Similarly, glycopeptide **182** with a type-2 LacNAc extended core 3 glycan is not at all recognized and

the same glycoforms with the glycan in PDT*R (**183**) or GST*A (**184**) are considerably weaker bound than the corresponding peptides with the LacNAc extended core 1 (**195** and **197**) (*figure 4.79*).

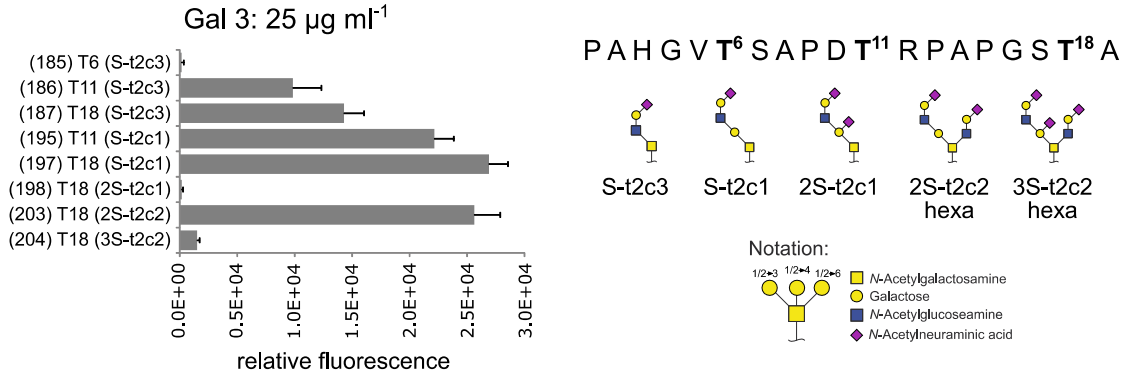


***Figure 4.79*:** Extract from array **MA4** (*figure 4.77*). Comparison of α2,6-sialylated glycopeptides for evaluation of internal galactose binding by galectin-3 (concentration: 25 µg/mL).

This shows that galectin-3 binds also the internal galactose of the core 1 motif in complex *O*-glycans. In contrast to other galectins, galectin-3 recognizes internal LacNAc units, which strengthens binding affinity to poly-LacNAc chains.[351] Glycopeptides with terminally α2,6-sialylated glycans show significantly reduced reactivity with galectin-3. Terminal α2,6-sialylation was also found to reduce galectin-3 binding in other assays.[67,351,363,364] The biological interactions are intensively studied. This type of sialylation, which blocks the binding epitopes for many galectins including galectin-3, has been correlated with the expression levels of ST6Gal-I, which mainly sialylates *N*-glycans, but also *O*-glycoproteins and to some extend *O*-glycosphingolipids and ST6GalNAc-I, which mainly sialylates *O*-glycans. ST6Gal-I is overexpressed on many cancer types,[366] including breast,[367] colon[368] or ovary.[369] In addition, also the levels of galectin-3 are elevated.[370] In contrast to its anti-apoptotic and carbohydrate-independent intracellular functions, extracellular galectin-3 has pro-apoptotic effects (see introduction, *chapter 1.3*). It was postulated that overexpression of ST6Gal-I and increase of extracellular sialylation counteracts the extracellular apoptotic mechanisms induced by galectin-3, additional to the simultaneous anti-apoptotic effects inside the cell. This would be advantageous for the survival of tumor cells.[371] For example, when the β1-subunit of integrins is α2,6-sialylated, epithelial colon cancer cells do not respond to galectin-3 induced apoptosis anymore.[93] Furthermore, sialylation of the β1-integrin unit promotes cell migration and therefore tumor cell metastasis.[372]

4.3.12.6 Influence of the peptide backbone multivalent presentation

Similar to the tested plant lectins described above, carbohydrate binding by galectin-3 was usually stronger, when the peptides were glycosylated at the C-terminal end (see *chapter 4.3.10.1*), as can be seen with the T-antigen glycopeptides **75**-**81** or ST-antigen glycopeptides **175**-**181** (*figure 4.77*). The glycans on the C-terminal end of the peptides were better accessible for the carbohydrate binding proteins than the glycans on the N-terminal site, which was connected site-specifically to the hydrogel microarray surface. A randomized orientation on the surface was not possible, at least not without further chemical modifications, such as implementation of another amine reactive spacer at the C-terminal end.

Some glycopeptides with multivalent glycan presentation were immobilized on **MA4**, such as the triglycosylated T- and ST-antigen peptides **81** and **181**, the triglycosylated type-2 core 3 and core 1 peptides **98**, **113**. Additionally, the core 2 hexa- or octasaccharides on peptides **125** and **201**, presented terminal LacNAc in a biantennary way. As expected, these multiglycosylated peptides were recognized better relative to the monoglycosylated equivalents. Mono- and trivalent variants of glycopeptides with the same glycan were compared by calculating the multivalency ratio, which is the signal intensity of trivalent conjugated peptides divided by the signal intensity of monovalent conjugated peptides. The multivalency ratio represents a factor for signal amplification due to multivalent binding modes (*figure 4.80*).

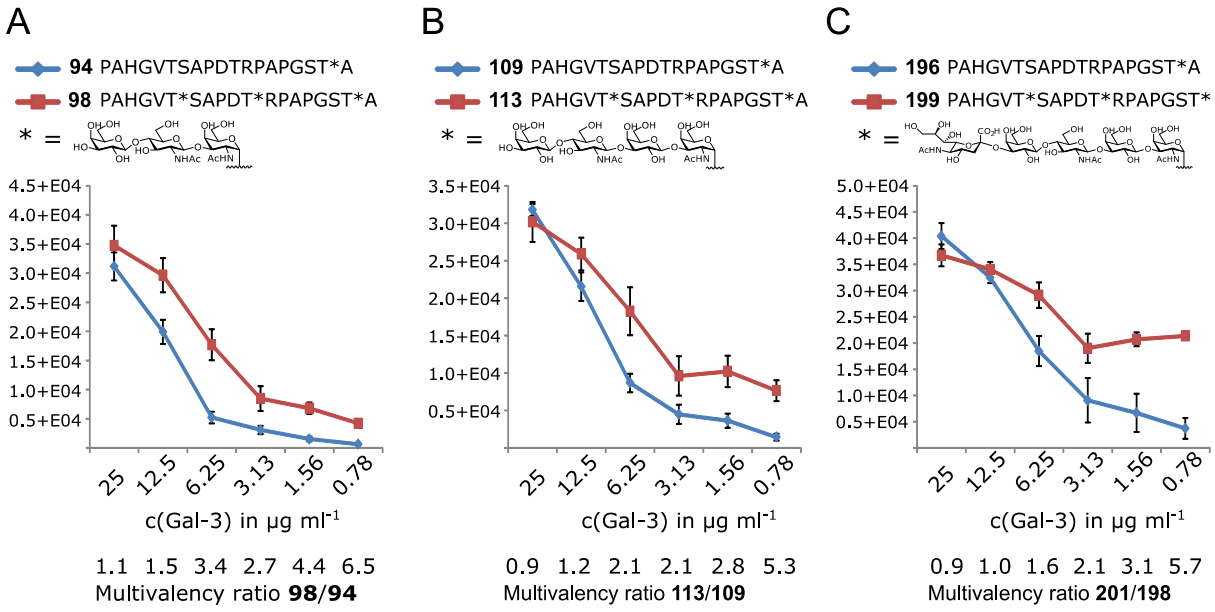


***Figure 4.80*:** Comparison of signal intensities at different galectin concentrations from mono- and triglycosylated peptides, represented by the multivalency ratio (= signal intensity [trivalent/monovalent] peptide). **A**: type-2 core 3 peptides **94** and **98**. **B**: type-2 core 1 peptides **109** and **113**. **C**: Sialyl type-2 core 1 peptides **201** and **198**.

Multivalent carbohydrate binding of galectin-3 did not follow a linear increase, since comparison of mono- and trivalent glycopeptide ligands should otherwise give a maximum multivalency ratio of 3, which would equal a three times higher affinity for triglycosylated peptides. At high concentrations of galectin-3 the ratio is around 1 because the system is saturated by the lectin. At concentrations around 1.5 µg/ml, approximately three times more galectin-3 is bound by the trivalent glycopeptides and at the lowest concentration the ratio showed a value around six, meaning that three equivalents of carbohydrate lead to an increase in binding by the factor of six at that galectin concentration. Galectin-3 therefore definitely recognized the carbohydrates by a multivalent binding mode, which increased in an exponential way if more carbohydrate structures were available for binding. It can be expected that the ratio would increase further if lower concentrations of galectin-3 were tested.

5 SUMMARY

Tumor cells differ from normal cells by the glycosylation pattern on their cell surface proteins. Aberrant glycosylation on the mucin glycoproteins is based on the altered expression of several glycosyltransferases. The downregulation of core 2 *N*-acetylglucosaminyltransferases and increased upregulation of sialyltransferases results in truncated and prematurely sialylated glycans. Mutations of the gene *COSMC* are responsible for lower C1GnT (T-synthase) levels and therefore hampered core 1 elongation. As a consequence, altered glycan structures are truncated and often prematurely sialylated. The smaller carbohydrate structures represent tumor-associated antigens, which make the peptide backbone accessible for tumor-specific antibodies. Serum auto-antibodies have been identified against these characteristic tumor structures and prove that a humoral immune response against altered tumor cell surface structures is feasible. Stimulation of the immune system with an adequate vaccine, directed against tumor-associated glycopeptide structures, would be a valuable therapeutic strategy to reduce the tumor burden, provide protection against metastasis and create long term protecting against relapses.

One of the most promising targets for active vaccination in cancer immunotherapy is the MUC1 glycoprotein. This membrane-bound protein is found on epithelial cell tissues and is highly overexpressed and aberrantly glycosylated on carcinoma cells. Organic synthesis is a reliable way to generate precisely defined vaccines, based on the MUC1 tandem repeat structure, for the potential use in immunotherapy. The main challenges in this undertaking, is to overcome the self-tolerance of the immune system against these endogenous structures and in relation to that, provoke a guaranteed specificity against the tumor-associated structures.

To study the induced antibody specificity of potential tumor-associated vaccine candidates, a MUC1 tandem repeat glycopeptide microarray platform was designed. For that purpose, a MUC1 glycopeptide library was prepared, with the premise to screen antibody sera for the binding specificities towards different glycopeptide epitopes. The antibody cross-reactivity of peptides, glycosylated with *O*-glycan core structures, such as extended core 1, 2 and 3 structures, which likely represent the elaborate glycosylation status of the mucins on healthy cells, was compared with antibody recognition of peptides with shorter, tumor-associated structures. The glycans were presented on the peptide backbone in a systematic and

multivalent pattern, to mimic the structures on natural mucin glycoproteins. The glycopeptides were synthesized in a convergent strategy with presynthesized and appropriately protected glycosylated amino acids. These amino acids were synthesized from a small pool of common building blocks, which were assembled in a uniform fashion to generate several glycosylated threonine amino acids for solid phase peptide synthesis.

The modular strategy involved the synthesis of type-1 (Galβ1,3-GlcNAcβ) and type-2 (Galβ1,3-GlcNAcβ) *N*-acetyllactosamine disaccharides **27** and **31** for flexible elongation of threonine T_N- and T-antigen acceptor amino acids. Trichloroacetimidate couplings were selected for the construction of the disaccharides, which themselves were further utilized in thioglycoside coupling reactions. The synthesis of the type-1 disaccharide, including a β1,3-linkage, turned out to be synthetically more challenging, since problems with low reactivity and orthoester formation occurred. By improving the coupling conditions, the β1,3-connected type-1 disaccharide was finally prepared in high yields at a multigram scale. The synthesized β1,4- and β1,3-linked LacNAc thiophenyl donors approved to be universal building blocks as donors for appropriately protected amino acid acceptors **12** and **44**. The products of the convergent synthesis approach were core 1, 2 and 3 mucin-type *O*-glycans, terminated by galactose. This strategy mimics the biosynthesis of complex *O*-glycans, which are enzymatically elongated by the alternating addition of *N*-acetylglucosamine followed by galactose *(figure 5.1)*.

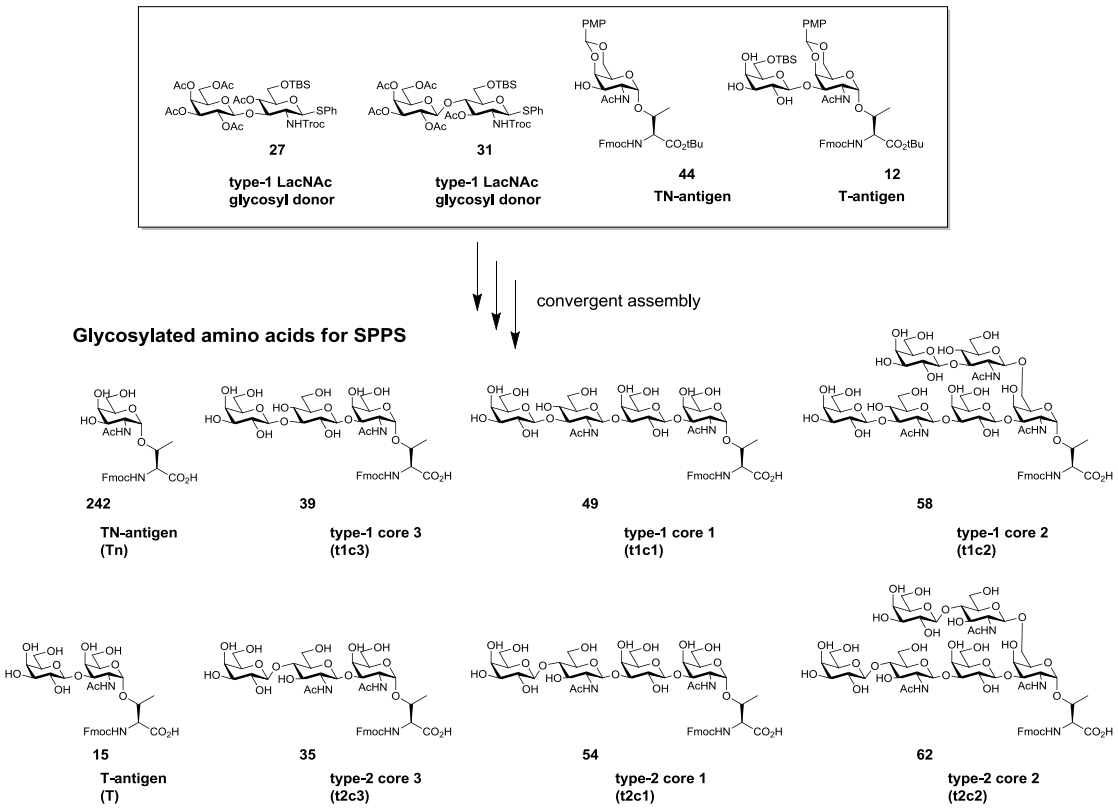


***Figure 5.1*:** Synthesized glycosylated amino acids for SPPS by convergent assembly of basic building blocks.

Glycosylation of the T-antigen acceptor **44** with either LacNAc donor **27** or **31** proceeded regioselectively on the more reactive 3-position of the galactose with full β-stereoselectivity. After changing the temporary protecting groups on the carbohydrate part to acetyl groups, the final type-1 core 1 and type-2 core 1 elongated amino acids **49** and **54**, were ready to be applied in SPPS. The intermediate products **47** and **52** were also used for further regioselective glycosylation on the 6-position of the *N*-acetylgalactosamine with the thioglycoside donors **27** and **31**, to result in the final core 2 hexasaccharide amino acids **58** and **62** (*figure 5.2*).

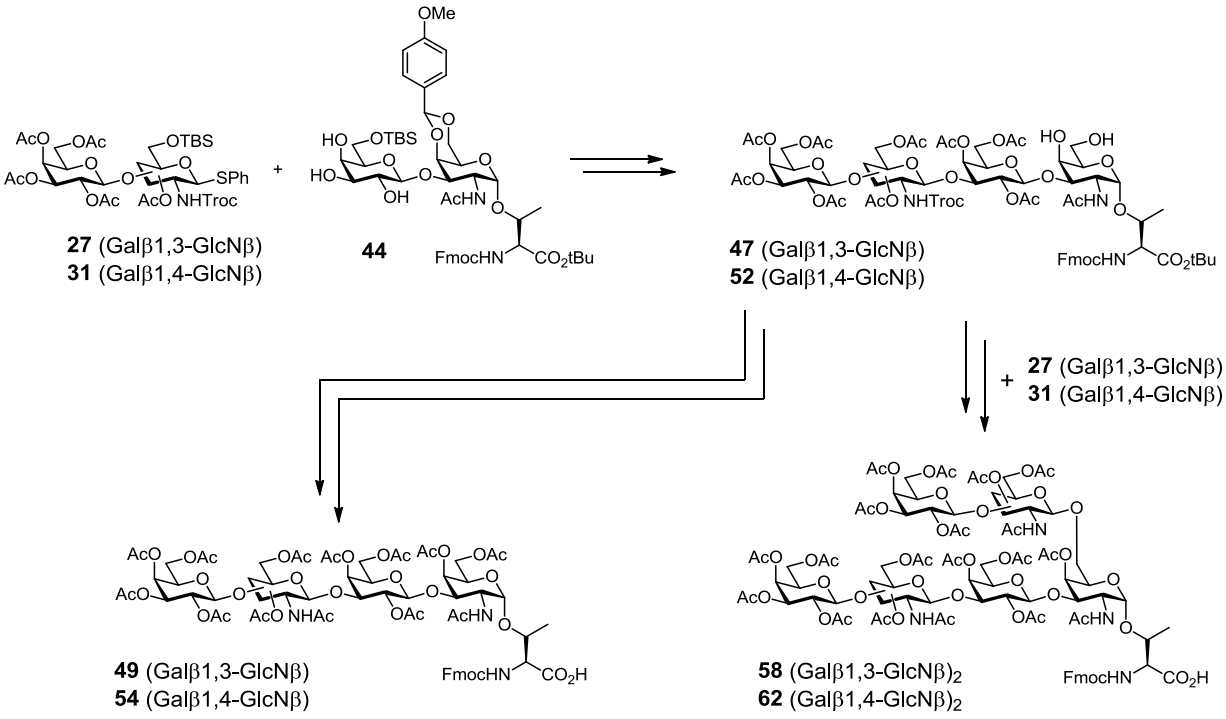


***Figure 5.2*:** Synthesis route of extended core 1 and 2 amino acid building blocks.

Glycosylations with the T_N-antigen acceptor **12** and the thioglycoside donors **27** or **31** to form the corresponding elongated core 3 amino acids, gave unexpected results. Reaction of the type-1 donor **31** with acceptor **12** gave an anomeric mixture of the glycosylation product (α/β = 1:1.2). The diastereomers were separated after the successive hydrolysis of the 4,6-*O*-*para*-methoxybenzylidene acetal and the 6'-*O*-*tert*-butyldimethylsilyl ether. Further protecting group manipulations provided the appropriately peracetylated type-1 core 3 amino acids for SPPS (*figure 5.3*, **A**). For formation of the type-2 core 3 trisaccharide amino acid, glycosylation with the type-2 donor **27** generated the desired β-anomer, but also formed a byproduct, with a phenyl sulfide group, attached to the nitrogen of the *N*-Troc carbamate (*figure 5.3*, **B**). Increased amounts of NIS as a promoter, were found to favor the formation of the phenyl sulfide adduct. Since the phenyl sulfide group could be removed in a later step along with *N*-Troc deprotection, increased amounts of NIS were applied, turning the

byproduct into the main product. After few protecting group manipulation steps the desired type-2 core 3 glycosylated amino acid was obtained, ready for SPPS.

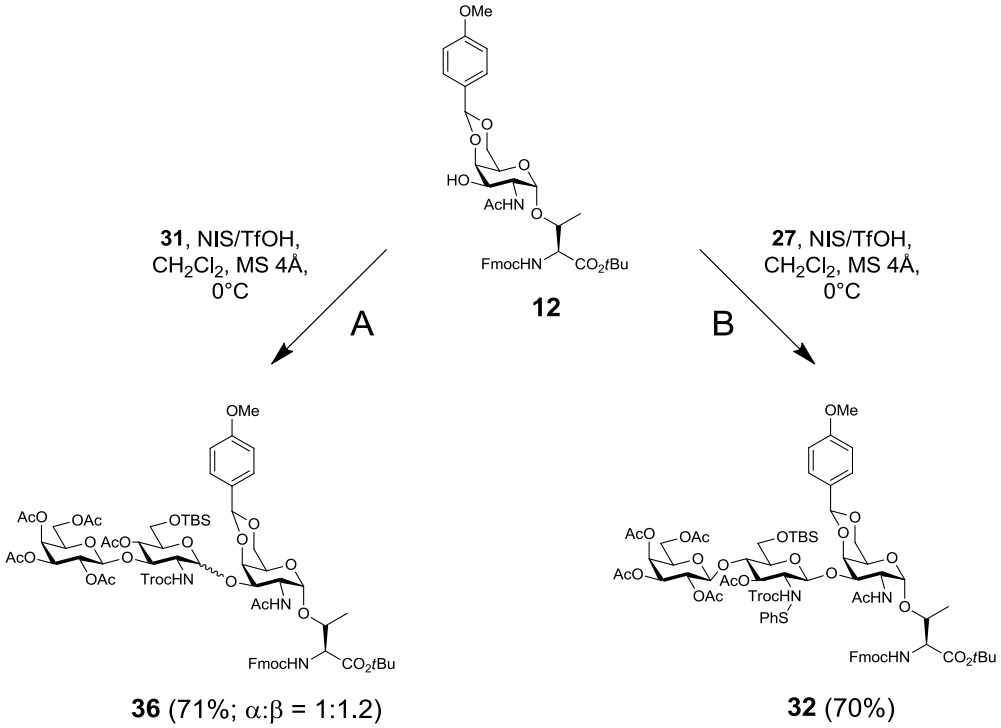
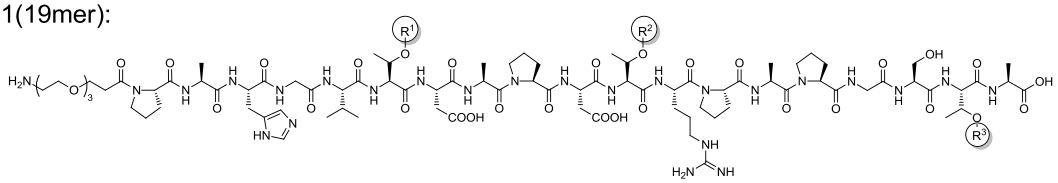


***Figure 5.3*:** Glycosylation reactions for synthesis of extended core 3 amino acids. **A**: Reaction of donor **31** with acceptor **12** resulted in anomeric product **36** with type-1 extension. **B**: Reaction of donor **27** with acceptor **12** resulted in product **32** with type-2 extension and *N*-phenyl sulfide group linked to *N*-Troc.

The glycosylated amino acids were globally *O*-acetylated for protection of the carbohydrate moieties during Fmoc-SPPS. MUC1 tandem repeat glycopeptides were synthesized on *Tentagel*®-Trityl-resins. The glycosylated amino acids were incorporated at all three threonine glycosylation sites of the MUC1 tandem repeat to produce several variants of mono-, di- and triglycosylated MUC1 peptides. The amino acid sequences were equipped with an amino triethylene glycol spacer (TEG) on the N-terminus, as a linker for later immobilization to microarray chips (*figure 5.4*).

MUC1(19mer):



68: R^1 = **Tn**, R^2 = H, R^3 = H
69: R^1 = H, R^2 = **Tn**, R^3 = H
70: R^1 = H, R^2 = H, R^3 = **Tn**
71: R^1 = **Tn**, R^2 = **Tn**, R^3 = H
72: R^1 = **Tn**, R^2 = H, R^3 = **Tn**
73: R^1 = H, R^2 = **Tn**, R^3 = **Tn**
74: R^1 = **Tn**, R^2 = **Tn**, R^3 = **Tn**

75: R^1 = **T**, R^2 = H, R^3 = H
76: R^1 = H, R^2 = **T**, R^3 = H
77: R^1 = H, R^2 = H, R^3 = **T**
78: R^1 = **T**, R^2 = **T**, R^3 = H
79: R^1 = **T**, R^2 = H, R^3 = **T**
80: R^1 = H, R^2 = **T**, R^3 = **T**
81: R^1 = **T**, R^2 = **T**, R^3 = **T**

85: R^1 = **t1c3**, R^2 = H, R^3 = H
86: R^1 = H, R^2 = **t1c3**, R^3 = H
87: R^1 = H, R^2 = H, R^3 = **t1c3**
88: R^1 = **t1c3**, R^2 = **t1c3**, R^3 = H
89: R^1 = **t1c3**, R^2 = H, R^3 = **t1c3**
90: R^1 = H, R^2 = **t1c3**, R^3 = **t1c3**
91: R^1 = **t1c3**, R^2 = **t1c3**, R^3 = **t1c3**

92: R^1 = **t2c3**, R^2 = H, R^3 = H
93: R^1 = H, R^2 = **t2c3**, R^3 = H
94: R^1 = H, R^2 = H, R^3 = **t2c3**
95: R^1 = **t2c3**, R^2 = **t2c3**, R^3 = H
96: R^1 = **t2c3**, R^2 = H, R^3 = **t2c3**
97: R^1 = H, R^2 = **t2c3**, R^3 = **t2c3**
98: R^1 = **t2c3**, R^2 = **t2c3**, R^3 = **t2c3**

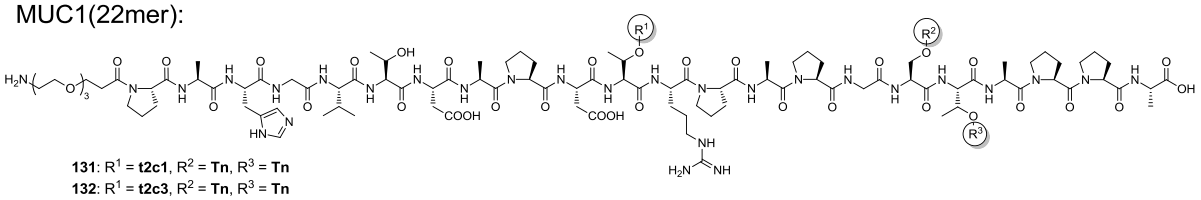
100: R^1 = **t1c1**, R^2 = H, R^3 = H
101: R^1 = H, R^2 = **t1c1**, R^3 = H
102: R^1 = H, R^2 = H, R^3 = **t1c1**
103: R^1 = **t1c1**, R^2 = **t1c1**, R^3 = H
104: R^1 = **t1c1**, R^2 = H, R^3 = **t1c1**
105: R^1 = H, R^2 = **t1c1**, R^3 = **t1c1**
106: R^1 = **t1c1**, R^2 = **t1c1**, R^3 = **t1c1**

107: R^1 = **t1c2**, R^2 = H, R^3 = H
108: R^1 = H, R^2 = **t1c2**, R^3 = H
109: R^1 = H, R^2 = H, R^3 = **t1c2**
110: R^1 = **t1c2**, R^2 = **t1c2**, R^3 = H
111: R^1 = **t1c2**, R^2 = H, R^3 = **t1c2**
112: R^1 = H, R^2 = **t1c2**, R^3 = **t1c2**
113: R^1 = **t1c2**, R^2 = **t1c2**, R^3 = **t1c2**

117: R^1 = **t1c2**, R^2 = H, R^3 = H
118: R^1 = H, R^2 = **t1c2**, R^3 = H
119: R^1 = H, R^2 = H, R^3 = **t1c2**
120: R^1 = **t1c2**, R^2 = **t1c2**, R^3 = H
121: R^1 = **t1c2**, R^2 = H, R^3 = **t1c2**
122: R^1 = H, R^2 = **t1c2**, R^3 = **t1c2**
123: R^1 = **t1c2**, R^2 = **t1c2**, R^3 = **t1c2**

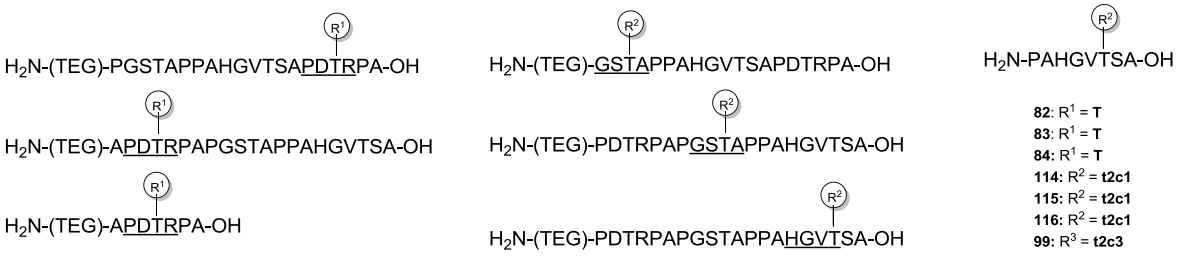
124: R^1 = **t2c2**, R^2 = H, R^3 = H
125: R^1 = H, R^2 = **t2c2**, R^3 = H
126: R^1 = H, R^2 = H, R^3 = **t2c2**
127: R^1 = **t1c2**, R^2 = **t2c2**, R^3 = H
128: R^1 = **t2c2**, R^2 = H, R^3 = **t2c2**
129: R^1 = H, R^2 = **t2c2**, R^3 = **t2c2**
130: R^1 = **t2c2**, R^2 = **t2c2**, R^3 = **t2c2**

MUC1(22mer):



131: R^1 = **t2c1**, R^2 = **Tn**, R^3 = **Tn**
132: R^1 = **t2c3**, R^2 = **Tn**, R^3 = **Tn**
133: R^1 = **t2c2**, R^2 = **Tn**, R^3 = **Tn**

MUC1 glycosylation site variants:

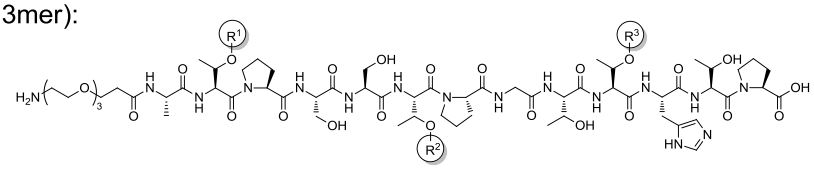


82: R^1 = **T**
83: R^1 = **T**
84: R^1 = **T**
114: R^2 = **t2c1**
115: R^2 = **t2c1**
116: R^2 = **t2c1**
99: R^3 = **t2c3**

***Figure 5.4*:** Synthesized MUC1 glycopeptide sequences.

Further, the glycosylated amino acids were introduced into a sequence of the MUC5B tandem repeat. Similar to the synthesis of the MUC1 glycopeptides, several multivalent MUC5B glycopeptides were constructed (*figure 5.5*). MUC5B is one major protein involved in mucus formation in the lungs and respiratory pathways. In future works, these glycopeptides will be used for binding studies, for example with from *Pseudomonas aeroginosa* bacterial strains as well as the with the isolated bacterial lectins *Lec A* and *Lec B*. *Pseudomonas aeroginosa* is a contributor to chronic infections and inflammatory environment causing exacerbations in COPD and Asthma airway disease.

MUC5B(13mer):



134: R^1 = H, R^2 = H, R^3 = **T**
135: R^1 = **T**, R^2 = H, R^3 = H
136: R^1 = **T**, R^2 = H, R^3 = **T**
137: R^1 = **T**, R^2 = **T**, R^3 = H
138: R^1 = **T**, R^2 = **T**, R^3 = **T**

140: R^1 = H, R^2 = H, R^3 = **c3t1**
141: R^1 = **c3t1**, R^2 = H, R^3 = H
142: R^1 = **c3t1**, R^2 = H, R^3 = **c3t1**
143: R^1 = **c3t1**, R^2 = **c3t1**, R^3 = H
144: R^1 = **c3t1**, R^2 = **c3t1**, R^3 = **c3t1**

146: R^1 = H, R^2 = H, R^3 = **c3t2**
147: R^1 = **c3t2**, R^2 = H, R^3 = H
148: R^1 = **c3t2**, R^2 = H, R^3 = **c3t2**
149: R^1 = **c3t2**, R^2 = **c3t2**, R^3 = H
150: R^1 = **c3t2**, R^2 = **c3t2**, R^3 = **c3t2**

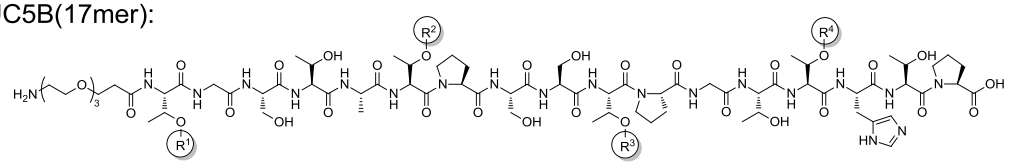
152: R^1 = H, R^2 = H, R^3 = **c1t1**
153: R^1 = **c1t1**, R^2 = H, R^3 = H
154: R^1 = **c1t1**, R^2 = H, R^3 = **c1t1**
155: R^1 = **c1t1**, R^2 = **c1t1**, R^3 = H
156: R^1 = **c1t1**, R^2 = **c1t1**, R^3 = **c1t1**

158: R^1 = H, R^2 = H, R^3 = **c1t2**
159: R^1 = **c1t2**, R^2 = H, R^3 = H
160: R^1 = **c1t2**, R^2 = H, R^3 = **c1t2**
161: R^1 = **c1t2**, R^2 = **c1t2**, R^3 = H
162: R^1 = **c1t2**, R^2 = **c1t2**, R^3 = **c1t2**

164: R^1 = H, R^2 = H, R^3 = **c2t1**
165: R^1 = **c2t1**, R^2 = H, R^3 = H
166: R^1 = **c2t1**, R^2 = H, R^3 = **c2t1**
167: R^1 = **c2t1**, R^2 = **c2t1**, R^3 = H
168: R^1 = **c2t1**, R^2 = **c2t1**, R^3 = **c2t1**

170: R^1 = H, R^2 = H, R^3 = **c2t2**
171: R^1 = **c2t2**, R^2 = H, R^3 = H
172: R^1 = **c2t2**, R^2 = H, R^3 = **c2t2**
173: R^1 = **c2t2**, R^2 = **c2t2**, R^3 = H
174: R^1 = **c2t2**, R^2 = **c2t2**, R^3 = **c2t2**

MUC5B(17mer):



139: R^1 = **T**, R^2 = **T**, R^3 = **T**, R^4 = **T**
145: R^1 = **c3t1**, R^2 = **c3t1**, R^3 = **c3t1**, R^4 = **c3t1**
151: R^1 = **c3t2**, R^2 = **c3t2,** R^3 = **c3t2**, R^4 = **c3t2**
157: R^1 = **c1t1**, R^2 = **c1t1**, R^3 = **c1t1**, R^4 = **c1t1**
163: R^1 = **c1t2**, R^2 = **c1t2**, R^3 = **c1t2**, R^4 = **c1t2**

***Figure 5.5*:** Synthesized MUC5B sequences.

After solid phase peptide synthesis, the carbohydrate acetyl protecting groups were cleaved off in solution under basic conditions. T_N-, T-antigen and core 3 structures were readily deprotected under *Zemplén* conditions. Elongated core 1 and core 2 glycopeptides required harsher conditions at pH of 11.5, in order to remove less the reactive 4'-*O*-acetyl group of the internal galactose unit. Certain β-elimination and some epimerization reactions were monitored by analytical HPLC and separated without problems from the desired products by preparative HPLC.

A selected fraction of the chemically synthesized MUC1 glycopeptides were further employed in enzymatic termination of the peptide glycan chains by the use of the three α2,3-sialyltransferases rat2,3OST (rat, recombinant, *S. frugiperda*), PmST1 and PmST3 (*P. multocida*) and the α2,6-sialyltransferase Pd2,6ST (*P. damsela*). The selection of peptides comprised representatives of various glycans with occupation of different glycosylation sites (*figure 5.6*). In continuing experiments, the variants of sialylated peptides will serve as standards for on-slide enzymatic modifications on glycopeptide microarray chips. During the course of this thesis, the sialylated MUC1 glycopeptides were already used in microarray evaluation of induced polyclonal antibody binding specificity against the sialylated MUC1 glycopeptides as well as with carbohydrate binding plant lectins and human galectin-3.

Further, the fragmentation patterns of the sialylated peptides were analyzed by HCD-MS fragmentation.

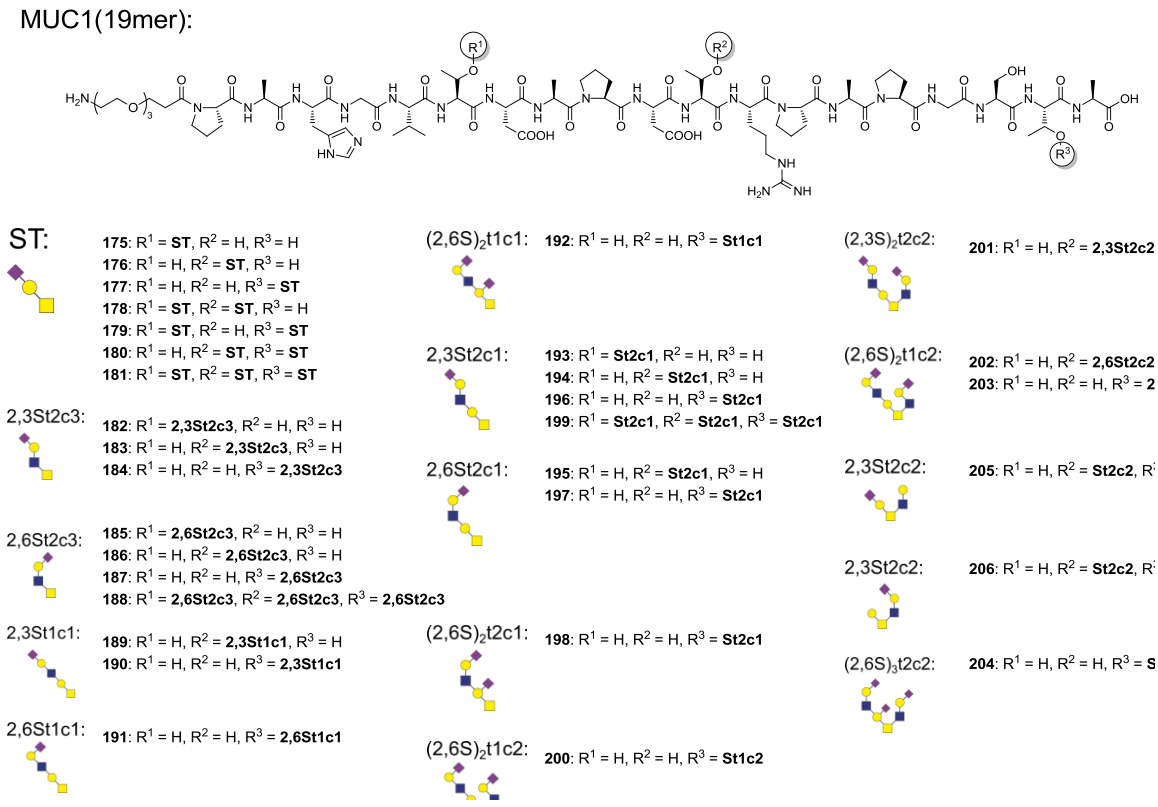
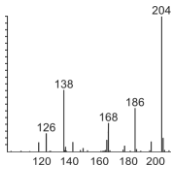
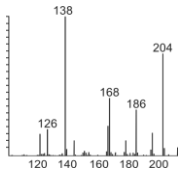
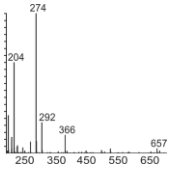
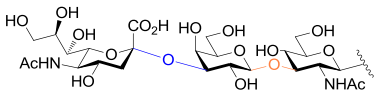
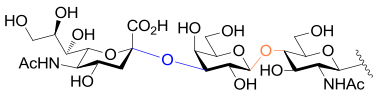
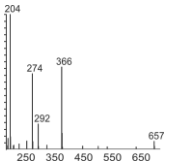
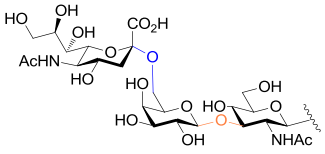
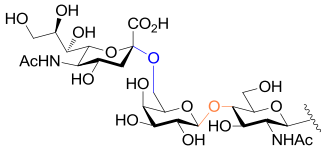


***Figure 5.6*:** Synthesized MUC1 glycopeptides with terminal α2,3- or α2,6-*N*-acetylneuraminic acid.

The α2,6-sialyltransferase Pd2,6ST has been reported to sialylate terminal as well as internal galactose units in poly-LacNAc oligosaccharide chains. Here, the enzyme was used to sialylate terminal type-2 and type-1 disaccharides and it was found capable to sialylate also internal galactose units of the fundamental core 1 substructure. However, the affinity to the internal core 1 galactose was lower and only peptides with glycosylation in the GSTA domain of the MUC1 tandem repeat were found to be partly sialylated (probably due to less sterical hindrance). The additional products, including an extra internal core 1 sialic acid, was collected if found in amounts significant enough for semi-preparative HPLC, as in peptides **192**, **198** and **204**. The products of the enzymatic conversions were finally analyzed by analytical HPLC and high-resolution mass spectrometry. The characteristic HexNAc and Neu5Ac oxonium ions, originating from the glycans of the different glycopeptides by applying higher energy C-trap dissociation (HCD) fragmentation, were analyzed, regarding their relative intensities. It was found, that the ions of the HexNAc-fragmentation almost exclusively originate from βGlcNAc of the type-1 and type-2 LacNAc residues and not from the internal αGalNAc. The observed diagnostic oxonium ions could therefore be further used

to discriminate spectrometrically between the β1,3- and β1,4-linked LacNAc disaccharides, by comparing the relative ion intensities of the fragments, derived from HexNAc fragmentation. At the same time, the connectivity of the α2,3- and α2,6-linked Neu5Ac could be differentiated by comparing the relative intensities of the $HexNAc^+$ and the $Neu5Ac^+$ ions. Higher intensities of *m/z* 204 ($HexNAc^+$) were found for peptides with α2,3-sialylation, while higher levels of *m/z* 274 ($Neu5Ac^+$-H_2O) were found in cases of α2,6-sialylation. These observations allowed distinguishing between the four isobaric trisaccharides residue variants (Neu5AcHexHexNAc), commonly terminating physiological glycan motifs on *N*- and *O*-glycoproteins (*table 5.1*).

***Table 5.1*:** Oxonium ion profiles for isobaric Neu5AcHexHexNAc trisaccharide residues in HCD-MS fragmentation experiments (at 35% NCE).

isobaric Neu5AcHexHexNAc	type-1	type-2
α2,3 or α2,6 → type-1 or type-2 →	*m/z*: 204>138>186>168 	*m/z*: 138>204>168>186 
2,3-Neu5Ac *m/z*: 274>204>366>657 		
2,6-Neu5Ac *m/z*: 204>274=366>657 		

The oxonium ion pattern analysis may be integrated into standard glycoproteomic workflows. MS-analysis of the glycans after proteolytic digestion of the samples on the glycopeptide level would omit laborious wet lab modifications, such as β-elimination, hydrazinolysis and permethylation. Further, structural analysis on the intact glycopeptide, would retain the information of its origin from the former glycoprotein in complex samples. Oxonium ion profile analysis with CID or HCD for glycan sequencing paired with ETD or ECD for glycosylation site determination on the level of the glycopeptides may become extremely helpful in glycome and glycoproteome elucidation. The whole approach also underlines the importance of organic and chemoenzymatic carbohydrate and glycopeptide synthesis, providing useful standards for structural glycoproteomic analysis.

The synthesized mucin glycopeptide library was thus employed to study antibody and lectin interactions by microarray analysis. All glycopeptides were synthesized with a triethylene glycol spacer amino acid at the N-terminus with a single amine group for site selective and covalent immobilization to the amine reactive *N*-hydroxysuccinimide coated microarray surface. The prepared glycopeptide arrays were used to screen for polyclonal antibody binding specificity of murine antisera, induced by different synthetic anti-tumor vaccine candidates. The induced antibodies were obtained by immunization of mice with synthetic vaccine candidates, synthesized in the work group and by the collaborator, containing MUC1 B-cell epitope glycopeptides and different immune stimulants (*figure 5.7*).

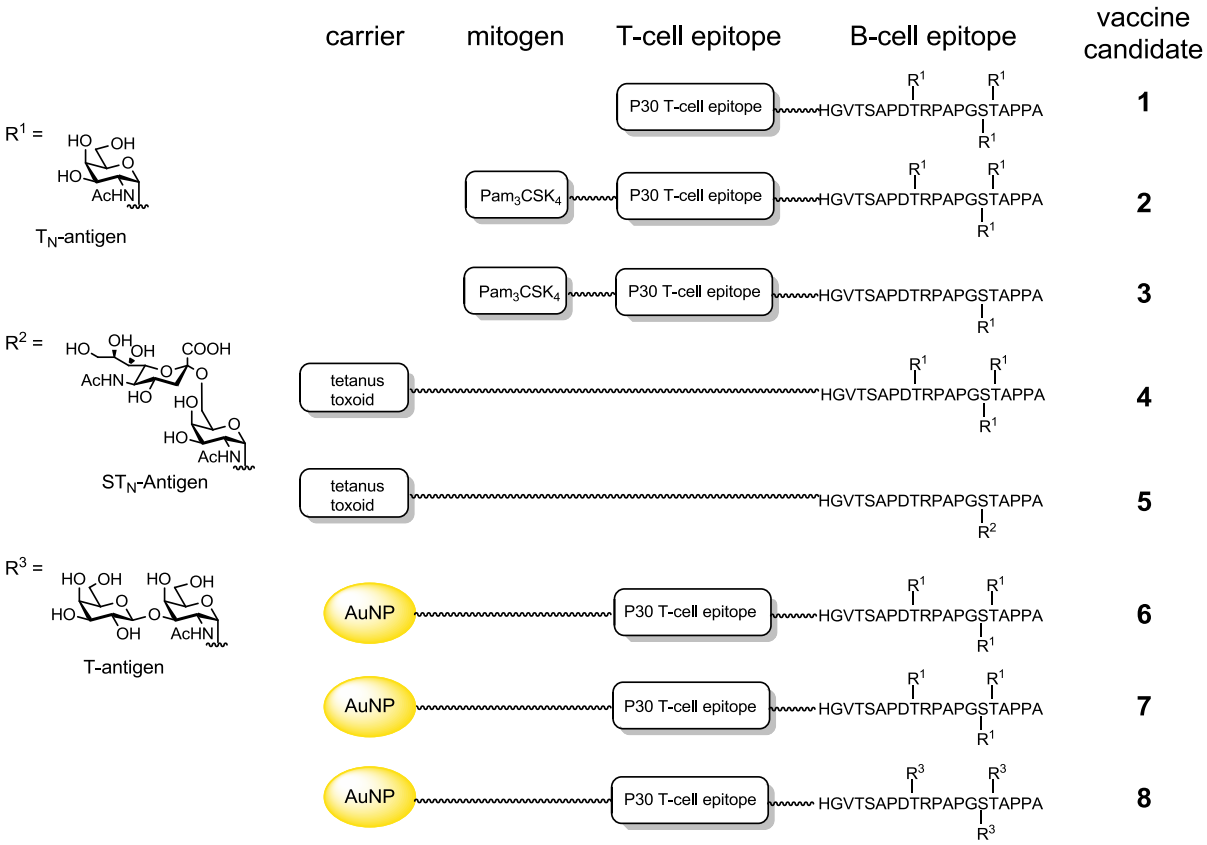


***Figure 5.7*:** Structures of synthetic vaccine candidates used for generation of polyclonal antibody immune sera in mice.

The antisera of the mice were screened in various concentrations on the MUC1 microarray slides. The major findings were:

- Vaccine candidates with T_N-antigen glycosylation in both the immunodominant PDT*R region and the GS*T*A epitope domain of the MUC1 peptide tandem repeat, elicited antibodies with reactivity against the glycosylated PDT*R epitope. PDT*R-specific antibodies, to a certain degree, tolerated extended core glycans in the

immunodominant domain in addition to the tumor-associated T_N- or T-antigen structure that was included in the B-cell epitope of the vaccine.

- Linear glycosylation (e.g. extended core 1 structures) in PDTR was preferred over branched core structures (e.g. extended core 2 structures). In the epithelial tissue of breast cells, core 2 glycans dominate cell surface glycosylation in the healthy state, compared to core 1 glycans in the tumor cells. This means that the induced serum antibodies discriminate to a great extent between healthy and cancerous cell surface glycosylation.
- Triglycosylated MUC1 tandem repeat sequences with extended core 2 glycosylation on all three threonine glycosylation sites were typically not recognized by the serum antibodies. Thus, this glycoform variant is regarded as a mimic for MUC1 glycosylation on healthy epithelial cells, which is entirely masking the protein epitopes. The induced antibodies are therefore expected to discriminate between extensive, healthy MUC1 cell glycosylation and malignant, truncated cancer glycosylation.
- Immune reactivity directed against clustered (vicinal diglycosylated) T_N-antigen glycosylation in the GS*T*A domain of the vaccine candidates, generated immune sera with restricted reactivity specifically for clustered T_N-antigen glycosylation. In contrast to the immune response induced by PDT*R glycosylated with T_N-antigen, no glycoform variation is tolerated by the antibodies. Hence, PDT*R and GS*T*A are fundamentally different classes of epitopes, according to the different reported peptide conformations.
- Single glycosylation with tumor associated T_N-antigen in GS*TA on a tripartite vaccine, elicited completely different antibody specificities in each of the three mice. Only one serum specificity was directed against glycosylation in GSTA. Single glycosylation in this domain appeared to be less immunogenic, compared to glycosylation in PDT*R.
- B-cell epitope conjugation with a P30 T-helper-cell epitope alone (two-component vaccine), combined with P30 and Pam_3CysSK_4 TLR-2 ligand (three-component vaccine) or together with P30 conjugated to gold nanoparticles, produced similar antibody specificity profiles in several mice. However, different antibody isotype pattern and variances in antibody titers were observed. These factors are also important for vaccine efficiency.
- B-cell epitope conjugation to tetanus toxoid resulted in very high antibody titers. The strong immune response appeared to create less glycoform and glycosylation site specific antibodies. However, the binding epitopes must contain the PDTR peptide sequence, hence the immune sera are still selective for the MUC1 tandem repeat

peptide sequence. Further, the antibody recognition was lost, when the MUC1 tandem repeat peptides were glycosylated with extended core 2 hexasaccharides in all three threonine glycosylation sites. Discrimination of extensive glycosylation on proteins of healthy cells and truncated glycosylation of proteins on diseased cells may therefore occur by exclusion of glycan size and structure.

In summary, the above described findings give new insights on antibody binding, elicited by synthetic vaccine constructs. Especially the contribution of the different glycosylation sites in MUC1 will help to understand the characteristics by which humoral immune responses react on tumor-associated MUC1 protein glycosylation and will aid with rationalizing vaccine designs. The presented glycoepeptide microarray system was for instance able to elucidate antibody cross-reactivity with various glycans at different sites, showing antibody discrimination between linear core 1 and branched core 2 structures.

The glycopeptide library was further used to evaluate binding to plant lectins, with certain known glycan binding specificity. It was found, that the glycopeptide microarray platform was able to produce results according to the expected lectin binding. Additionally, new binding specificities for further fine-tuning of the known binding specificities were observed, such as discrimination for the sialylation site, structure of *O*-glycan core elongation, amino acid preferences and glycosylation site specificity. The glycopeptide array was further utilized to explore the binding behavior of recombinant human galectin-3 towards MUC1 glycopeptides decorated with typical mucin-type *O*-glycan decoration. The physiological and pathophysiological properties of the galectin family are under intensive research since they are involved in several intra- and extracellular processes, promoting tumor development. For instance, there is evidence that extracellular galectin-3 is involved in adhesion of circulating tumor cells to secondary epithelial sites. This adhesion process is facilitated by binding to overexpressed MUC1 *via* the tumor-associated T-antigen, influencing the MUC1 membrane localization and increasing the likeliness of selectin binding between the circulating tumor cells and the endothelial cell surface.

Galectin-3 was incubated on the MUC1 glycopeptide chip at several concentrations. It was found that galectin-3 recognized the glycopeptides with tumor-associated T_N- and T-antigen glycosylation as weak affinity ligands. The affinity for these antigens was however lower than for the core 1, core 2 and core 3 *O*-glycans. Glycopeptides with terminating α2,3-*N*-acetylneuraminic acid showed increased recognition. This is interesting, since the influence of α2,3-sialylation is so far described as rather being tolerated by galectin-3 and in some reports even hindering interaction with glycoproteins. Previous studies have shown that modifications on the 2'- or 3'-position of *N*-acetyllactosamine with fucose, α-galactose or α-*N*-acetylgalactosamine residues (but not yet *N*-acetylneuraminic acid) enhance binding by

fitting into the B-site of the galectin binding groove. In this study, it is shown that α2,3-sialylation enhances galectin-3 affinity, supposedly by binding in this pocket. In accordance with already observed glycan recognition studies, α2,6-*N*-acetylneuraminic acid decorated glycopeptides showed decreased galectin-3 binding. Additional internal sialylation on the core 1 substructure further reduced galectin-3 binding, since galectin-3 is able to recognize internal galactose units. The α2,6-linked sialic acids have profound biological effects and are regarded as molecular switches, by which cells can regulate extracellular galectin-3 mediated apoptosis. Although this kind of sialylation is predominantly found on *N*-glycans, the α2,6-sialylation on *O*-glycoproteins on tumor cells appears the to be increased. The found galectin-3 preferences are summarized in *figure 5.8*.

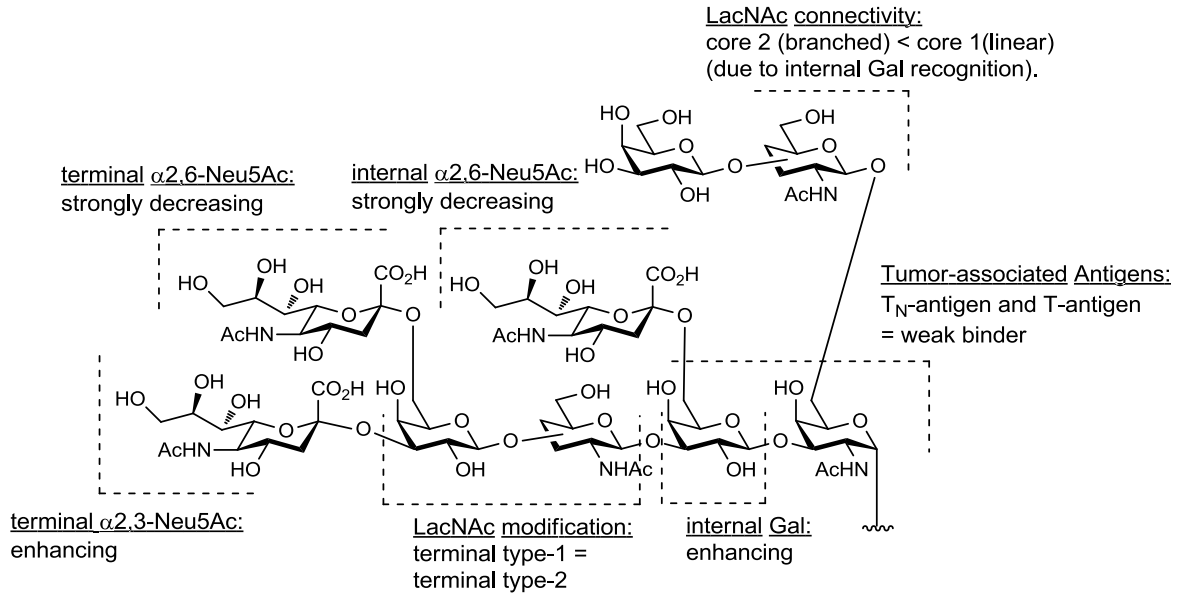


***Figure 5.8*:** Influence of *O*-glycan modifications on galectin-3 binding affinity.

The use of glycopeptides with numerous occupied glycosylation sites allowed the evaluation of multivalent binding effects. Since the binding intensities increased exponentially from mono- to triglycosylated peptide variants, it was concluded that the binding mode is multivalent and must therefore be accompanied with galectin-3 oligomerization.

In conclusion, the presented glycopeptide microarray platform is well equipped for evaluation of protein-glycan and protein-glycopeptide interactions. The glycopeptide library features a comprehensive selection of MUC1 peptides with various different core glycans presented in a multivalent fashion. Additionally, a MUC5B glycopeptide library was prepared, which can be used for further lectin binding studies. The constructed mucin glycopeptide library will also be modified by on-slide enzymatic glycosylation, in order to quickly produce more diverse glycan structures. Reactions with sialyl-, fucosyl-, α/β-galactosyl- or *N*-acetylgalactosaminyl-transferases can produce (sialyl-)Lewis antigens or blood group antigens. The synthesized

sialylated glycopeptides are a first step in this direction, since they will be used as standards for validation of on-slide enzymatic sialylation. In-solution chemoenzymatic syntheses were further used to optimize the enzyme reaction conditions and to test the enzyme substrate scope to different glycan structures. In addition to studies of human lectins, like the galectins, the constructed mucin glycopeptide array platform, will be used to study lectin interactions with viruses and bacteria, which e.g. are responsible for infections in airway diseases.

6 ZUSAMMENFASSUNG

Tumorzellen unterscheiden sich von gesunden Zellen durch eine veränderte Glycosylierung der Proteine an der Zelloberfläche. Verminderte Expression von Core 2 *N*-Acetylglucosaminyltransferase und erhöhte Expression von Sialyltransferasen, resultiert in verkürzten und vorzeitig sialylierten Glycanen. Mutationen des Gens *COSMC* sind verantwortlich für verminderte Level von C1Gn-T (T-synthase), wodurch die Ausbildung von Core 1 Glycanstrukturen unterbunden wird. Die verkürzten Kohlenhydrate repräsentieren Tumor-assoziierte Antigene und machen das Protein-Rückgrat zugänglich für Tumor-spezifische Antikörper. Serum Autoantikörper gegen diese charakteristischen Strukturen wurden identifiziert und zeigen, dass eine Immunantwort gegen diese veränderten Zelloberflächenstrukturen möglich ist. Stimulation des Immunsystems mit geeigneten Vakzinen, welche gegen diese Strukturen gerichtet sind, könnten eine wertvolle Ergänzung der therapeutischen Methoden darstellen, um die Tumorlast zu senken, Schutz vor Metastasierung aufzubauen und Langzeitschutz gegen Rezidiven zu bilden.

Ein vielversprechendes Ziel einer aktiven Vakzinierung in der Krebs-Immuntherapie, stellt das MUC1 Glycoprotein dar. Dieses membrangebundene Protein, welches sich an der Oberfläche von Epithelzellen befindet, wird auf Karzinomzellen überexprimiert und weist ein abnormes Glycosylierungsmuster auf. Organische Synthese ist eine zuverlässige Methode um definierte Vakzine darzustellen, welche auf der *tandem repeat* Peptidsequenz des MUC1-Proteins basieren und potentiell in der Immuntherapie eingesetzt werden können. Dabei besteht die Herausforderung darin, die Eigentoleranz des Immunsystems gegenüber diesen endogenen Strukturen zu überwinden, aber gleichzeitig eine garantierte Spezifität für die Tumor-assoziierten Strukturen zu erzeugen.

Um die erzeugte Antikörperspezifität von potentiellen Vakzinen gegen Tumor-assoziierte MUC1 Strukturen untersuchen zu können, wurde im Rahmen dieser Arbeit ein MUC1 Glycopeptid-Mikroarray System erstellt. Hierzu wurde eine Bibliothek von MUC1 Glycopeptiden synthetisiert, welche dazu diente, die Bindungsspezifitäten von induzierten Antiköperseren zu evaluieren. So konnte die Kreuzspezifität der Antikörper für die erweiterten Core 1, 2 und 3 *O*-Glycan-Strukturen, welche dem Glycosylierungsmuster auf gesunden Epithelzellen entsprechen, mit der Spezifität für Peptide mit Tumor-assoziierten Antigenen, wie dem T_N- und dem T-Antigen, verglichen werden. Dabei wurden die Glycane

auch in multivalenten Glycoformen der Peptide auf dem Array präsentiert, um die natürlichen Muster der zellulären Glycoproteine zu imitieren. Die Glycopeptide wurden in einer konvergenten Synthesestrategie, mit synthetischen glycosylierten Aminosäuren aufgebaut. Diese Aminosäuren wurden selbst aus einem Pool von wenigen Bausteinen synthetisiert, welcher es ermöglichte mehrere komplexe glycosylierte Aminosäuren zu erzeugen.

Die modulare Strategie beinhaltete die Synthese von Type-1 (Galβ1,3-GlcNAcβ) und Type-2 (Galβ1,3-GlcNAcβ) *N*-Acetyllactosamin Disacchariden **27** und **31**, mit denen die T_N- und T-Antigen Glycosylakzeptoren flexibel erweitert werden konnten. Trichloracetimidat-vermittelte Kupplungen wurden für die Disaccharid-Synthesen ausgewählt, welche selbst als Thioglycoside in weiteren Glycosylierungen verwendet wurden. Die Synthese der Type-1 Disaccharide war zunächst Aufgrund von Problemen mit niedrigen Reaktivitäten und Orthoester-Bildung synthetisch anspruchsvoller. Optimierung der Reaktionsbedingungen lieferte die β1,3-verknüpften Type-1 Disaccharide in hohen Ausbeuten auch in Multigramm-Ansätzen. Die synthetisierten β1,4- und β1,3-verknüpften LacNAc Thioglycoside erwiesen sich als universelle Donoren für die entsprechend geschützten Aminosäure-Akzeptoren **12** und **44**. Die Produkte der konvergenten Synthese entsprachen Core 1, Core 2 und Core 3 *O*-Glycanen mit terminalen Galactoseresten. Die verwendete Strategie imitierte die *O*-Glycanbiosynthese, bei der die grundlegenden Core Strukturen enzymatisch durch alternierende Hinzufügung von *N*-Acetylglucosamin- und Galactose-Einheiten verlängert werden (*Abbildung 5.*1).

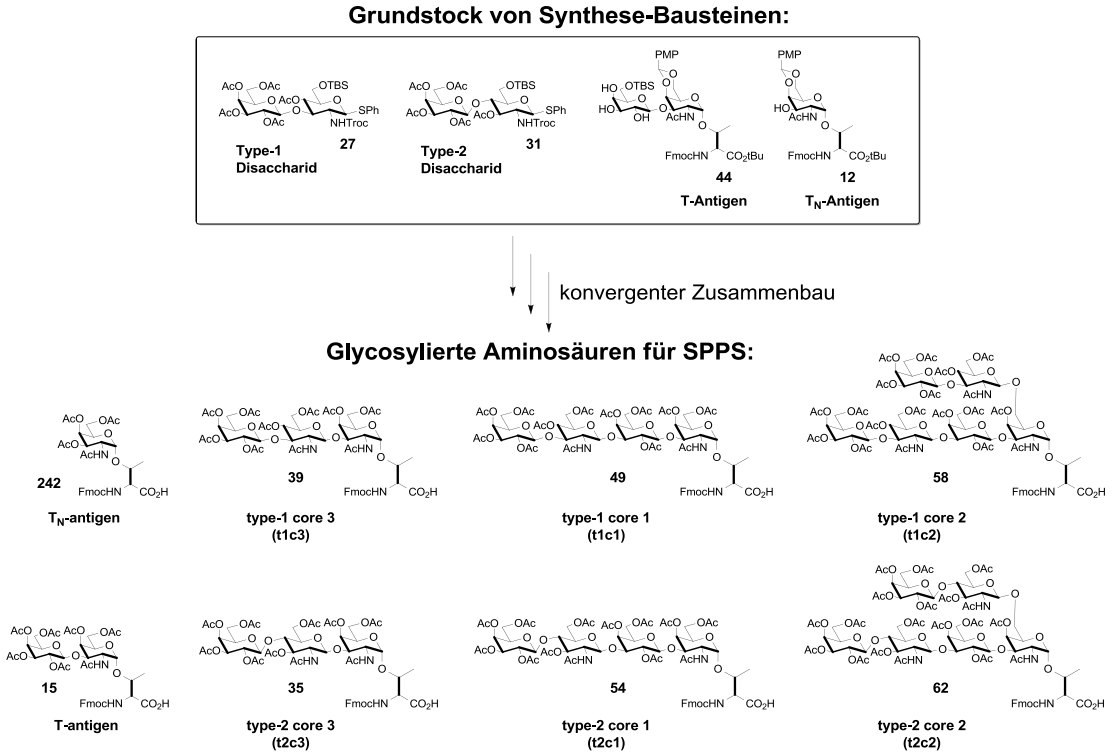


***Abbildung 6.1*:** Konvergente Synthese von glycosylierten Aminosäuren.

Glycosylierung des T-Antigen Akzeptors **44** mit den LacNAc-Donoren **27** oder **31** verlief regioselektiv an der reaktiveren 3‘-OH-Gruppe der Galactose mit voller β-Stereoselektivität. Nachdem die temporären Kohlenhydrat-Schutzgruppen gegen *O*-Acetylgruppen ausgetauscht wurden, konnten die finalen Type-1 und Type-2 Core 1 Aminosäuren **49** und **54** in der Peptid-Festphasensynthese eingesetzt werden. Die Zwischenprodukte **47** und **52** der Syntheseroute wurden für weitere Glycosylierungen in Position 6 des *N*-Acetylgalactosamins mit Thioglycosid-Donoren **27** und **31** eingesetzt um, die Core 2 Hexasaccharid-modifizierten Aminosäuren **58** und **62** darzustellen (*Abbildung 5.2*).

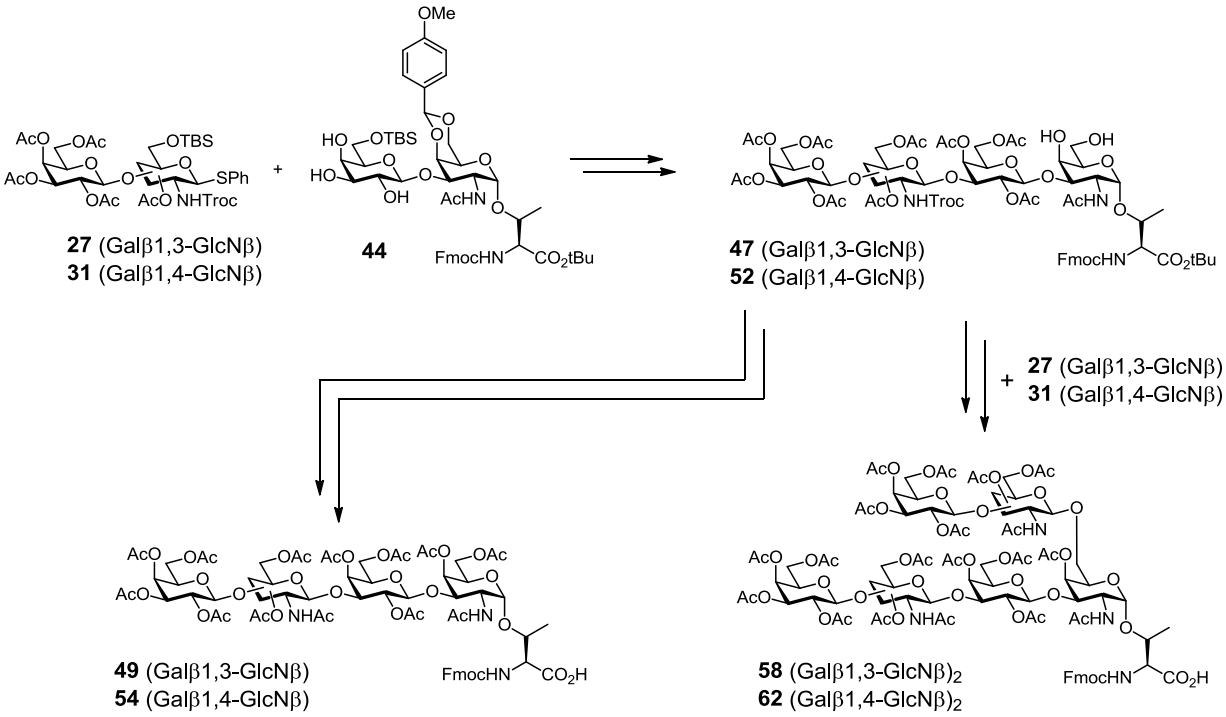


***Abbildung 6.2*:** Synthese der erweiterten core 1 und core 2 Aminosäuren.

Bei den Glycosylierungen von T_N-Antigen Akzeptor **12** mit den beiden Thioglycosid-Donoren zur Darstellung der erweiterten Core 3 Aminosäuren, wurden unerwartete Reaktionsprodukte erhalten. Die Reaktion des Type-1 Donors **31** mit Akzeptor **12** lieferte ein anomeres Gemisch des Glycosylierungsprodukts (α/β = 1:1.2). Die Diastereomere konnten nach der folgenden sauren Hydrolyse der 4,6-*O*-*para*-Methoxybenzylidengruppe und der 6‘-*O*-*tert*-Butyldimethylsilylethergruppe getrennt werden (*Abbildung 5.3*, **A**). Weitere Modifikationen des Schutzgruppenmusters ergab die peracetylierte Type-1 Core 3 Aminosäure für den Einsatz in der Peptid-Festphasensynthese. Zur Darstellung der entsprechenden Type-2 Core 3 Trisaccharid-Aminosäure, lieferte die Glycosylierung mit Type-2 Donor **27** das erwünschte β-Anomer, allerdings mit einem Nebenprodukt, bei dem die Phenylsulfid-Abgangsgruppe an den Stickstoff des Carbamats der Troc-Schutzgruppe gebunden wurde (*Abbildung 5.3*, **B**). Reaktion mit *N*-Iodosuccinimid in hohem Überschuss,

lieferte das Phenylsulfid-Nebenprodukt als Hauptprodukt. Nach Anpassung der temporären Schutzgruppen, wurde die peracetylierte Type-2 Core 3 Aminosäure für die Festphasensynthese erhalten.

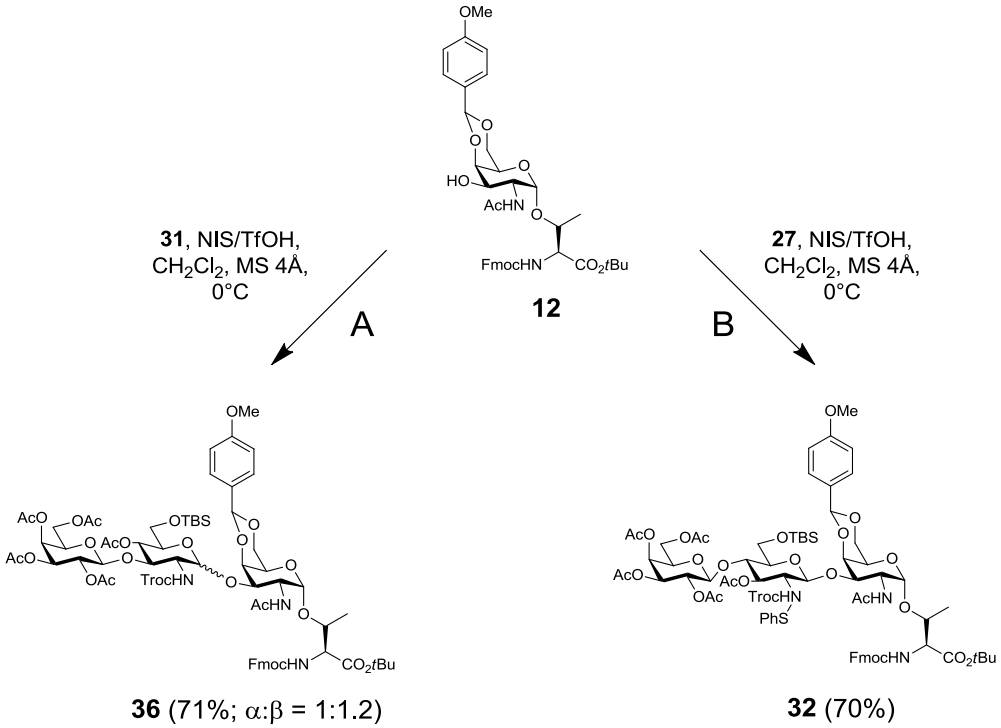
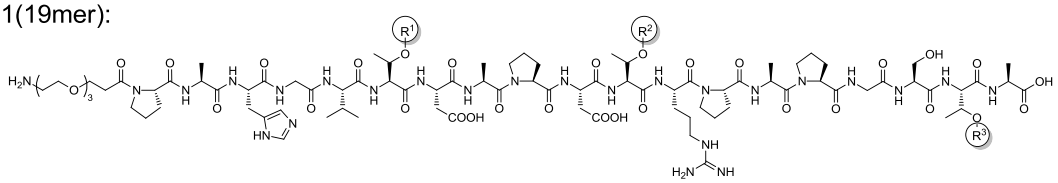


***Abbildung 6.3*:** Glycosylierungen zur Darstellung der Core 3 Aminosäure Bausteine. **A**: Reaktion des Donors **31** mit Akzeptor **12** und Bildung des anomeren Produktgemischs **36**. **B**: Reaktion von Donor **27** mit Akzeptor **12** ergab Produkt **32** als Phenylsulfid-Addukt.

Die MUC1 Peptidsequenzen wurden an *Tentagel*®-Trityl Harzen synthetisiert. Die Kohlenhydratreste waren während der Peptid-Festphasensynthese global mit *O*-Acetylgruppen geschützt. Die glycosylierten Aminosäuren wurden an den drei Threonin-Glycosylierungsstellen der Peptidsequenz eingebracht um verschiedene Glycoformen des mono-, di- und triglycosylierten MUC1 Peptids darzustellen. Um räumlichen Abstand zwischen der Mikroarray-Oberfläche und der Glycopeptidsequenz in folgenden Experimenten zu gewährleisten, wurden die Aminosäuresequenzen mit einer N-terminalen Triethylenglycol-Aminosäure (TEG) synthetisiert (*Abbildung 5.4*).

MUC1(19mer):



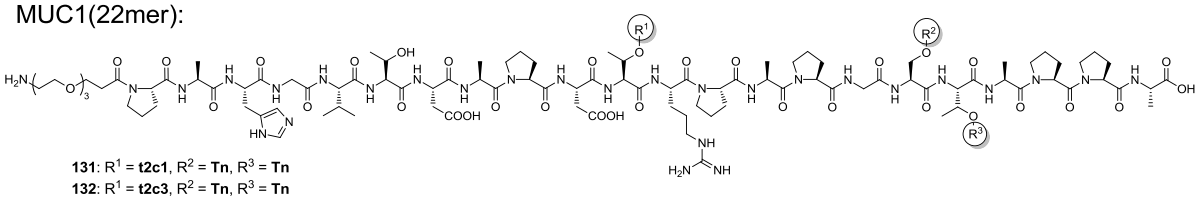
68: R^1 = **Tn**, R^2 = H, R^3 = H
69: R^1 = H, R^2 = **Tn**, R^3 = H
70: R^1 = H, R^2 = H, R^3 = **Tn**
71: R^1 = **Tn**, R^2 = **Tn**, R^3 = H
72: R^1 = **Tn**, R^2 = H, R^3 = **Tn**
73: R^1 = H, R^2 = **Tn**, R^3 = **Tn**
74: R^1 = **Tn**, R^2 = **Tn**, R^3 = **Tn**
75: R^1 = **T**, R^2 = H, R^3 = H
76: R^1 = H, R^2 = **T**, R^3 = H
77: R^1 = H, R^2 = H, R^3 = **T**
78: R^1 = **T**, R^2 = **T**, R^3 = H
79: R^1 = **T**, R^2 = H, R^3 = **T**
80: R^1 = H, R^2 = **T**, R^3 = **T**
81: R^1 = **T**, R^2 = **T**, R^3 = **T**

85: R^1 = **t1c3**, R^2 = H, R^3 = H
86: R^1 = H, R^2 = **t1c3**, R^3 = H
87: R^1 = H, R^2 = H, R^3 = **t1c3**
88: R^1 = **t1c3**, R^2 = **t1c3**, R^3 = H
89: R^1 = **t1c3**, R^2 = H, R^3 = **t1c3**
90: R^1 = H, R^2 = **t1c3**, R^3 = **t1c3**
91: R^1 = **t1c3**, R^2 = **t1c3**, R^3 = **t1c3**
92: R^1 = **t2c3**, R^2 = H, R^3 = H
93: R^1 = H, R^2 = **t2c3**, R^3 = H
94: R^1 = H, R^2 = H, R^3 = **t2c3**
95: R^1 = **t2c3**, R^2 = **t2c3**, R^3 = H
96: R^1 = **t2c3**, R^2 = H, R^3 = **t2c3**
97: R^1 = H, R^2 = **t2c3**, R^3 = **t2c3**
98: R^1 = **t2c3**, R^2 = **t2c3**, R^3 = **t2c3**

100: R^1 = **t1c1**, R^2 = H, R^3 = H
101: R^1 = H, R^2 = **t1c1**, R^3 = H
102: R^1 = H, R^2 = H, R^3 = **t1c1**
103: R^1 = **t1c1**, R^2 = **t1c1**, R^3 = H
104: R^1 = **t1c1**, R^2 = H, R^3 = **t1c1**
105: R^1 = H, R^2 = **t1c1**, R^3 = **t1c1**
106: R^1 = **t1c1**, R^2 = **t1c1**, R^3 = **t1c1**
107: R^1 = **t1c2**, R^2 = H, R^3 = H
108: R^1 = H, R^2 = **t1c2**, R^3 = H
109: R^1 = H, R^2 = H, R^3 = **t1c2**
110: R^1 = **t1c2**, R^2 = **t1c2**, R^3 = H
111: R^1 = **t1c2**, R^2 = H, R^3 = **t1c2**
112: R^1 = H, R^2 = **t1c2**, R^3 = **t1c2**
113: R^1 = **t1c2**, R^2 = **t1c2**, R^3 = **t1c2**

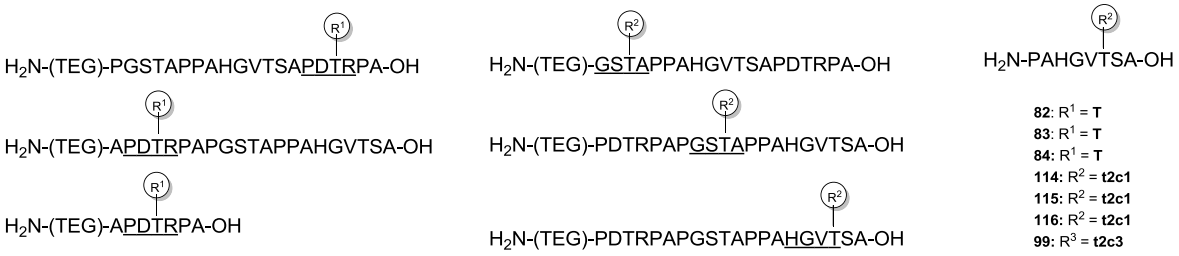
117: R^1 = **t1c2**, R^2 = H, R^3 = H
118: R^1 = H, R^2 = **t1c2**, R^3 = H
119: R^1 = H, R^2 = H, R^3 = **t1c2**
120: R^1 = **t1c2**, R^2 = **t1c2**, R^3 = H
121: R^1 = **t1c2**, R^2 = H, R^3 = **t1c2**
122: R^1 = H, R^2 = **t1c2**, R^3 = **t1c2**
123: R^1 = **t1c2**, R^2 = **t1c2**, R^3 = **t1c2**
124: R^1 = **t2c2**, R^2 = H, R^3 = H
125: R^1 = H, R^2 = **t2c2**, R^3 = H
126: R^1 = H, R^2 = H, R^3 = **t2c2**
127: R^1 = **t1c2**, R^2 = **t2c2**, R^3 = H
128: R^1 = **t2c2**, R^2 = H, R^3 = **t2c2**
129: R^1 = H, R^2 = **t2c2**, R^3 = **t2c2**
130: R^1 = **t2c2**, R^2 = **t2c2**, R^3 = **t2c2**

MUC1(22mer):



131: R^1 = **t2c1**, R^2 = **Tn**, R^3 = **Tn**
132: R^1 = **t2c3**, R^2 = **Tn**, R^3 = **Tn**
133: R^1 = **t2c2**, R^2 = **Tn**, R^3 = **Tn**

MUC1-Peptidsequenzen mit versetzten Glycosylierungsstellen:

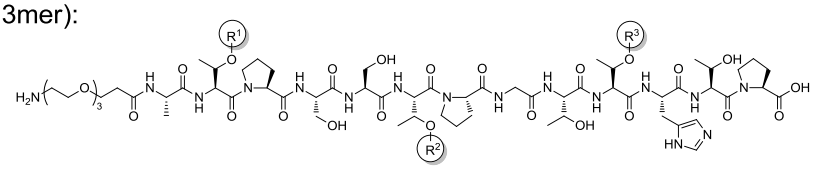


82: R^1 = **T**
83: R^1 = **T**
84: R^1 = **T**
114: R^2 = **t2c1**
115: R^2 = **t2c1**
116: R^2 = **t2c1**
99: R^3 = **t2c3**

***Abbildung 6.4*:** Synthetisierte MUC1 Glycopeptid-Sequenzen.

Weiterhin wurden die glycosylierten Aminosäuren in Peptid-Wiederholungssequenzen des Proteins MUC5B synthetisch eingebracht. Ähnlich der Synthese der MUC1 Glycopeptide, wurden multivalente Glycoformen der MUC5B-Sequenz hergestellt (*Abbildung 5.5*). MUC5B ist ein bedeutendes Glycoprotein, welches hauptsächlich für die Bildung von Mucus in den Atemwegen verantwortlich ist. In weiteren Studien, werden diese Glycopeptide für Bindungsstudien mit Atemwegs-infizierenden Bakterienstämmen, wie z.B. *Pseudomonas aeroginosa*, sowie isolierten bakteriellen Lektinen, wie z.B. *Lec A* und *Lec B* verwendet werden. Chronische Infektionen mit *Pseudomonas aeroginosa* und begleitenden Entzündungsprozessen stellen eine der Hauptfaktoren für schlechte Prognosen bei COPD (*chronic obstructive pulmonary disease*) und chronischen Asthmaerkrankungen dar.

MUC5B(13mer):



134: R^1 = H, R^2 = H, R^3 = **T**
135: R^1 = **T**, R^2 = H, R^3 = H
136: R^1 = **T**, R^2 = H, R^3 = **T**
137: R^1 = **T**, R^2 = **T**, R^3 = H
138: R^1 = **T**, R^2 = **T**, R^3 = **T**

140: R^1 = H, R^2 = H, R^3 = **c3t1**
141: R^1 = **c3t1**, R^2 = H, R^3 = H
142: R^1 = **c3t1**, R^2 = H, R^3 = **c3t1**
143: R^1 = **c3t1**, R^2 = **c3t1**, R^3 = H
144: R^1 = **c3t1**, R^2 = **c3t1**, R^3 = **c3t1**

146: R^1 = H, R^2 = H, R^3 = **c3t2**
147: R^1 = **c3t2**, R^2 = H, R^3 = H
148: R^1 = **c3t2**, R^2 = H, R^3 = **c3t2**
149: R^1 = **c3t2**, R^2 = **c3t2**, R^3 = H
150: R^1 = **c3t2**, R^2 = **c3t2**, R^3 = **c3t2**

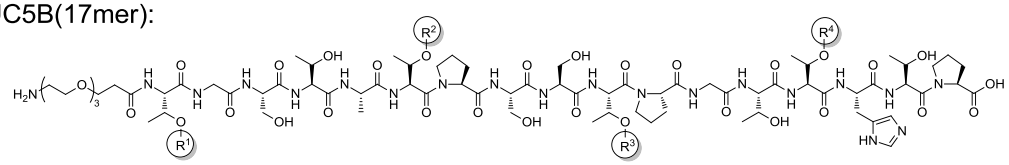
152: R^1 = H, R^2 = H, R^3 = **c1t1**
153: R^1 = **c1t1**, R^2 = H, R^3 = H
154: R^1 = **c1t1**, R^2 = H, R^3 = **c1t1**
155: R^1 = **c1t1**, R^2 = **c1t1**, R^3 = H
156: R^1 = **c1t1**, R^2 = **c1t1**, R^3 = **c1t1**

158: R^1 = H, R^2 = H, R^3 = **c1t2**
159: R^1 = **c1t2**, R^2 = H, R^3 = H
160: R^1 = **c1t2**, R^2 = H, R^3 = **c1t2**
161: R^1 = **c1t2**, R^2 = **c1t2**, R^3 = H
162: R^1 = **c1t2**, R^2 = **c1t2**, R^3 = **c1t2**

164: R^1 = H, R^2 = H, R^3 = **c2t1**
165: R^1 = **c2t1**, R^2 = H, R^3 = H
166: R^1 = **c2t1**, R^2 = H, R^3 = **c2t1**
167: R^1 = **c2t1**, R^2 = **c2t1**, R^3 = H
168: R^1 = **c2t1**, R^2 = **c2t1**, R^3 = **c2t1**

170: R^1 = H, R^2 = H, R^3 = **c2t2**
171: R^1 = **c2t2**, R^2 = H, R^3 = H
172: R^1 = **c2t2**, R^2 = H, R^3 = **c2t2**
173: R^1 = **c2t2**, R^2 = **c2t2**, R^3 = H
174: R^1 = **c2t2**, R^2 = **c2t2**, R^3 = **c2t2**

MUC5B(17mer):



139: R^1 = **T**, R^2 = **T**, R^3 = **T**, R^4 = **T**
145: R^1 = **c3t1**, R^2 = **c3t1**, R^3 = **c3t1**, R^4 = **c3t1**
151: R^1 = **c3t2**, R^2 = **c3t2,** R^3 = **c3t2**, R^4 = **c3t2**
157: R^1 = **c1t1**, R^2 = **c1t1**, R^3 = **c1t1**, R^4 = **c1t1**
163: R^1 = **c1t2**, R^2 = **c1t2**, R^3 = **c1t2**, R^4 = **c1t2**

***Abbildung 6.5*:** Synthetisierte MUC5B-Glycopeptidsequenzen.

Nach der Peptid-Festphasensynthese wurden die *O*-Acetylgruppen der Glycane unter basischen Bedingungen abgespalten. T_N- und T-Antigen sowie Core 3 dekorierte Glycopeptide wurden problemlos unter standardmäßigen *Zemplén* Bedingungen bei pH 8,5 entschützt. Erweiterte Core 1 und Core 2 Glycopeptide benötigten höhere pH-Werte (~11,5) um die reaktionsträgen 4‘-*O*-Acetylgruppe der internen Galactose zu hydrolsieren. Etwaige β-Eliminierungs- und Epimerisierungs-Produkte wurden mittels analytischer HPLC beobachtet und durch präparative HPLC abgetrennt.

Ausgewählte MUC1 Glycopeptide wurden weiterhin durch enzymatische Sialylierung an den Glycanketten terminiert. Zum Einsatz kamen die drei α2,3-Sialyltransferasen Rat2,3OST (rat, recombinant, *S. frugiperda*), PmST1 und PmST3 (*P. multocida*) und die α2,6-Sialyltransferase Pd2,6ST (*P. damsela*). Die Auswahl der Glycopeptide beinhaltete Vertreter mit verschiedenen Glycanen an unterschiedlichen Glycosylierungsstellen (*Abbildung 5.6*). In weiterführenden Experimenten können diese sialylierten Peptide als Standards für globale enzymatische Sialylierungen auf den Mikroarray Oberflächen dienen. Im Rahmen dieser Arbeit wurden die sialylierten MUC1 Peptide bereits verwendet um die Bindungsspezifitäten von polyklonalen Antikörpern aus Vakzinierungsexperimenten, sowie die Bindungsspezifitäten von Lektinen und humanem Galectin-3 zu evaluieren. Weiterhin

wurden die Fragmentierungsmuster der sialylierten Peptide durch HCD-MS Fragmentierung untersucht.

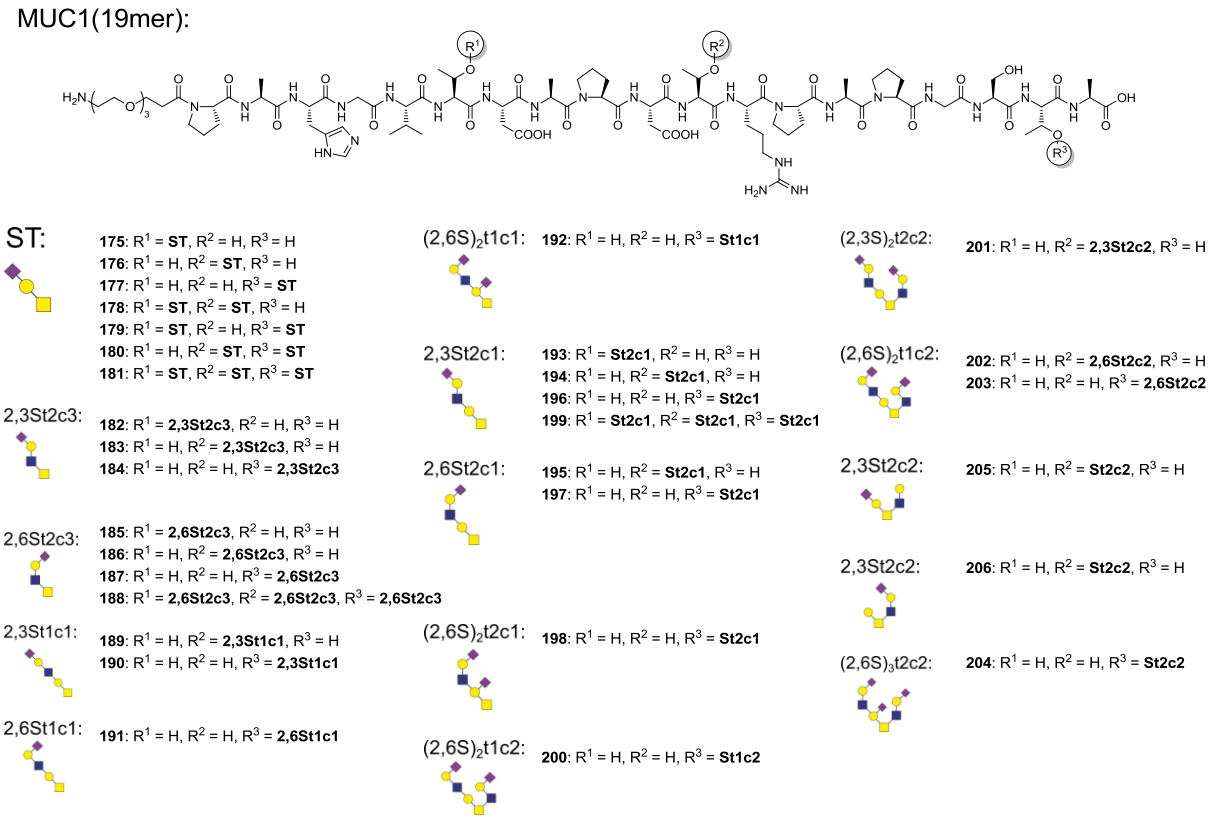
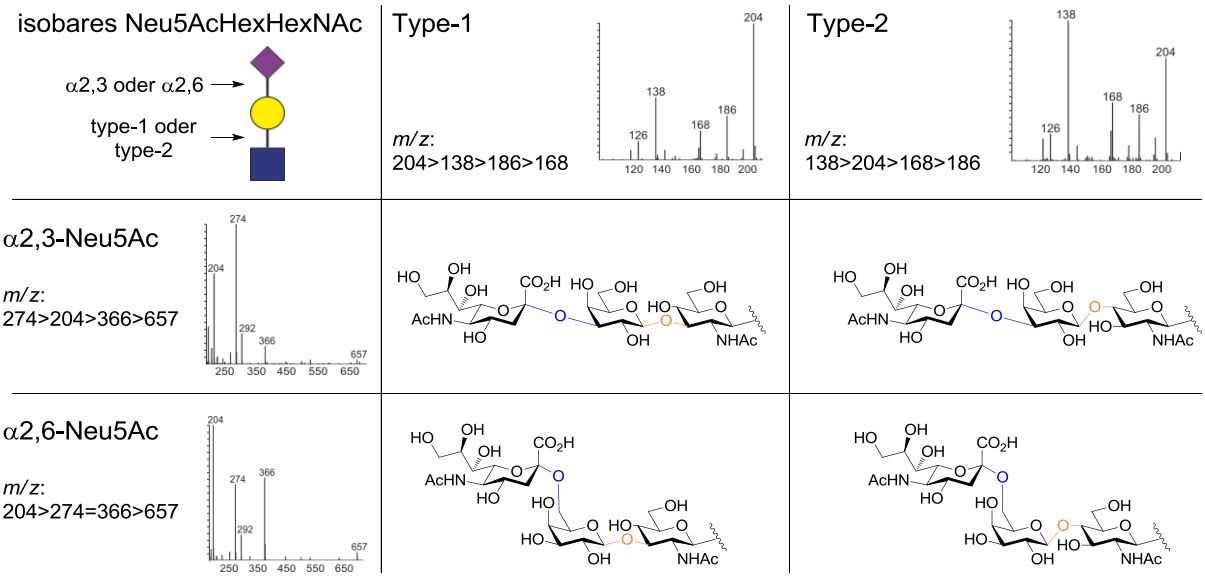


***Abbildung 6.6*:** MUC1 Glykopeptide mit terminaler α2,3- und α2,6-Sialylierung.

Die Sialyltransferase Pd2,6ST ist in der Lage terminale, als auch interne Galactose-Einheiten in Type-2 poly-LacNAc Oligosacchariden zu sialylieren. In dieser Arbeit wurde das Enzym verwendet, um Sialinsäurereste an terminale type-2 und type-1 LacNAc-Reste zu binden. Dabei wurde festgestellt, dass Pd2,6ST auch in der Lage ist, die interne Galactose der Core 1 Substruktur von erweiterten Core 1 und Core 2 Glycanen zu modifizieren. Die Affinität des Enzyms zu der internen Galactose war geringer als zu terminalen Galactoseresten und ausschließlich Peptide mit Glycosylierung in der GSTA Domäne des MUC1-Peptids wurden zum Teil zusätzlich sialyliert (vermutlich aufgrund geringerer sterischer Behinderung). Dieses weitere Reaktionsprodukt, mit zusätzlicher interner Core 1 α2,6-Sialylierung wurde nach Möglichkeit ebenfalls isoliert, wie es bei den Glycopeptidsequenzen **192**, **198** und **204** der Fall war. Die sialylierten Produkte aus den enzymatischen Reaktionen wurden mittels analytischer HPLC und hochauflösender Massenspektrometrie analysiert. Die charakteristischen HexNAc- und Neu5Ac-Oxoniumionen, welche unter HCD-Fragmentierungsbedingungen (*higher energy C-trap dissociation*) auftraten, wurden bezüglich ihrer relativen Intensitäten untersucht. Es stellte sich heraus, dass es sich bei den

auftretenden HexNAc-Fragmentionen fast ausschließlich um Fragmente der β-GlcNAc aus den Type-1 und Type-2 LacNAc-Resten handelt und nur sehr wenig von den internen α-GalNAc Einheiten stammt. Die beobachteten, diagnostischen HexNAc-Fragmentionen konnten daher benutzt werden um spektrometrisch zwischen den β1,3- und β1,4-verknüpften LacNAc Disacchariden zu unterscheiden. Zusätzlich war es möglich, die Konnektivität der α2,3 oder α2,6 gebundenen Sialinsäurereste zu bestimmen, indem die relativen Intensitäten der $HexNAc^+$-Ionen mit den $Neu5Ac^+$-Ionen verglichen wurden. Höhere Intensitäten von *m/z* 204 ($HexNAc^+$) wurden für α2,6-Sialylierung gefunden, wobei hohe Level von *m/z* 274 ($Neu5Ac^+$-H_2O) im Fall von α2,3-Sialylierung dominierten. Diese Beobachtungen ermöglichen es, die vier isobaren Trisaccharidreste, welche häufig physiologische *N*- und *O*-Glycane terminieren, massenspektrometrisch zu unterscheiden (*Tabelle 5.1*).

***Tabelle 6.1*:** Intensitätsprofile der Oxoniumionen für die vier isobaren Neu5AcHexHexNAc Trisaccharidreste in HCD-MS Fragmentierungsexperimenten (bei 35% NCE).



Die Analyse von Oxoniumion-Profilen könnte standardmäßig in glycoproteomische Arbeitsabläufe eingefügt werden. Massenspektrometrische Analyse der Glycane nach dem proteolytischen Verdau der Proben auf Glycopeptid-Level, könnte aufwendige, nass-chemische Methoden, wie β-Eliminierung, Hydrazinolyse und Permethylierung aussparen. Desweiteren würde die Strukturanalyse am intakten Glycopeptid die Information über seine Herkunft vom ursprünglichen Glycoprotein bewahren. Oxoniumionen-Analyse mittels CID oder HCD zur Glycansequenzierung, zusammen mit ETD oder ECD Fragmentierung zur Identifizierung der Glycosylierungsstellen auf Stufe des Glykopeptids, könnte ein hilfreiches Mittel sein in der Aufklärung des Glycoms und Glycoproteoms. Desweiteren unterstreicht

dieser Ansatz die Wichtigkeit, von organischer und chemo-enzymatischer Synthese von Kohlenhydraten und Glycopeptiden, um wertvolle Standards für glycoproteomische Untersuchungen zu liefern.

Die synthetische MUC1 Glycopeptidbibliothek wurde weiterhin benutzt um Wechselwirkungen mit Antikörpern und Lektinen durch Mikroarray-Analyse durchzuführen. Alle Glycopeptide wurden mit einer N-terminalen Triethylenglycol-Aminosäure als Linker synthetisiert, welche die einzige Aminogruppe des MUC1-Peptids beinhaltete. Somit wurden die Glycopeptide selektiv und gerichtet mit dem N-Terminus an *N*-Hydroxysuccinimidgruppen-beschichtete Mikroarrays gebunden. Die so erhaltenen Glycopeptid-Arrays wurden benutzt um die Bindungsspezifität von polyklonalen Antikörpern aus murinen Antiseren zu untersuchen. Diese Antiseren wurden durch Immunisierung von Mäusen mit synthetischen Vakzin-Kandidaten erhalten, welche in der Arbeitsgruppe oder durch Kooperationspartner synthetisiert wurden. Die Vakzin-Kandidaten beruhten auf MUC1 Glycopeptid-Strukturen als immunogene B-Zell-Epitope, konjugiert an diverse Immunstimulatoren. (*Abbildung 5.7*).

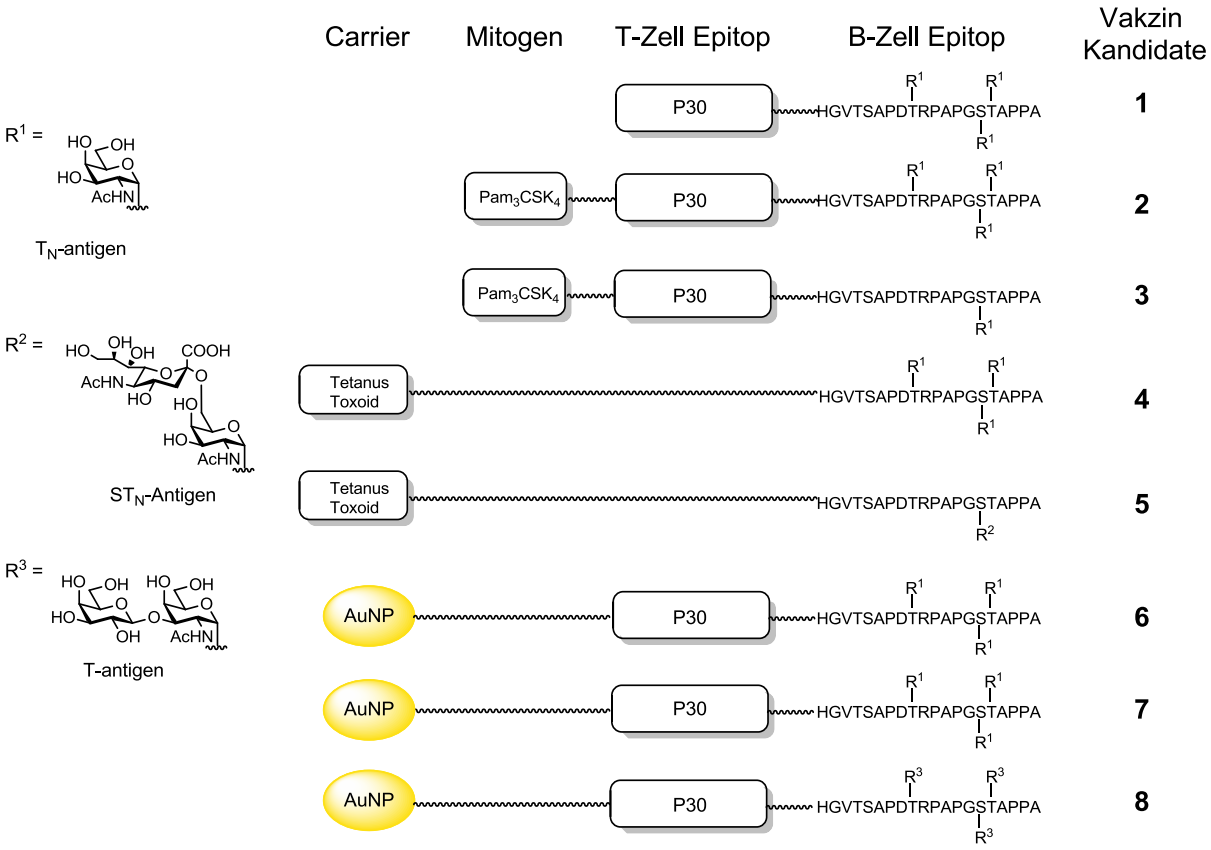


***Abbildung 6.7*:** Strukturen der synthetischen Vakzin-Kandidaten zur Erzeugung von Antiseren in Mäusen.

Die so erhaltenen Antiseren wurden in unterschiedlichen Konzentrationen auf den MUC1 Glykopeptid-Mikroarrays auf Bindung untersucht. Die Ergebnisse dieser Untersuchungen waren:

- Vakzin-Kandidaten mit T_N-Antigen Glycosylierung in der immundominanten PDT*R Region und im GS*T*A Epitop des MUC1 Peptids, generieren Antikörper mit Spezifität für das glycosylierte PDT*R Epitop. PDT*R-spezifische Antikörper, tolerieren zu einem gewissen Grad erweiterte Core Glycane in der immundominanten Domäne, zusätzlich zu den Tumor-assoziierten T_N- und T-Antigen Strukturen in den B-Zell Epitopen der Vakzin-Kandidaten.
- Lineare Glycosylierung (z.B. erweiterte Core 1 Strukturen) in der PDTR Sequenz, wurde bevorzugt von den erzeugten Antikörpern erkannt, im Gegensatz zu verzweigter Glycosylierung (erweitere Core 2 Strukturen). Im Fall von Brustgewebszellen, dominiert verzweigte Core 2 Glycosylierung auf der Oberfläche gesunder Zellen, während auf Mammakarzinomzellen unverzweigte Core 1 Strukturen überwiegen. Dies bedeutet, dass die erzeugten Antikörper zum großen Teil zwischen gesunder und krankhafter Zelloberflächenglycosylierung unterscheiden können.
- Triglycosylierte MUC1 Peptidsequenzen mit erweiterter Core 2 Glycosylierung an allen drei Threonin-Glycosylierungsstellen, wurden durch die Antikörper kaum bis gar nicht erkannt. Daher können diese Glycoformen als repräsentativ für MUC1 Glycosylierung auf gesunden Epithelzellen angesehen werden, welche die Proteinepitope komplett abschirmt. Es wird daher erwartet, dass die induzierten Antikörper zwischen extensiver MUC1 Glycosylierung auf gesunden Zellen und verkürzter Glycosylierung auf krankhaften Zellen unterscheiden können.
- Immunreaktivität, welche sich gegen vicinale (*clustered*) T_N-Antigen Glycosylierung in der GS*T*A Region der Vakzin-Kandidaten richtet, erzeugte Antiseren mit ausschließlicher Reaktivität der Antikörper für exakt diese vicinale T_N-Antigen Glycosylierung. Im Gegensatz zu der Immunreaktivität, welche sich gegen die PDT*R Domäne mit nur einem T_N-Antigen richtete, werden in der GS*T*A Sequenz keine anderen Glycoformen durch die Antikörper toleriert. Daher sind die beiden Domänen, einfach glycosylierte PDT*R Sequenz und vicinal glykosylierte GS*T*A Sequenz, als fundamental unterschiedliche Epitopklassen zu betrachten. Dies entspricht den in der Literatur beschriebenen Beobachtungen, dass die beiden glycoslierten Peptidsequenzen unterschiedliche Arten von Konformationen einnehmen.
- Einzelne Glycosylierung mit T_N-Antigen in der GS*TA Domäne eines Drei-Komponenten Vakzin-Kandidaten, rief eine jeweils unterschiedliche Antikörperspezifität in den drei Mäusen hervor. Nur eines der drei Seren, zeigte dabei

eine Antikörperspezifität für die glycosylierte GS*TA Region. Einfache Glycosylierung mit nur einem Glycan in dieser Region scheint daher weniger immunogen zu sein, als doppelte, vicinale Glycosylierung in GS*T*A oder als entsprechende einfache Glycosylierung in der PDT*R Sequenz.

- MUC1 B-Zell-Epitop, konjugiert mit einer P30 T-Zell-Epitop Peptidsequenz (Zwei-Komponenten-Vakzine), kombinierte Konjugation mit P30 und Pam_3CysSK_4 (Drei-Komponenten-Vakzine) oder Konjugation mit P30 an Goldnanopartikel, induzierte jeweils ähnliche Antikörper Spezifitäten in fast allen Mäusen. Dennoch unterschieden sich die verschiedenen Strategien in der Zusammensetzung der unterschiedlichen Antikörper-Isotypen und der Antikörpertiter. Hierbei handelt es ebenfalls um wichtige Parameter, welche die Effizienz eines Vakzins bestimmen.
- Konjugation der MUC1 B-Zell-Epitope an ein Tetanus Toxoid Protein, induzierte sehr hohe Antikörpertiter. Die starke Immunreaktion scheint jedoch auf Kosten der Spezifität für Glycoformen und Glycosylierungstellen zu gehen. Dennoch musste das Peptidepitop die immundominante PDTR Sequenz enthalten, womit die Antiseren weiterhin eine Selektivität für die MUC1 Peptidsequenz aufwiesen. Weiterhin wurde die Erkennung durch die Antikörper stark beeinflusst, wenn alle Threonin-Glycosylierungstellen mit verzweigten Core 2 Glycanen modifiziert waren. Eine Unterscheidung von MUC1 Proteinen auf gesunden oder erkrankten Zellen, scheint daher durch die ausgeprägten Glycanstrukturen zusammen mit einem hohen Glycosylierungsgrad möglich zu sein.

Zusammenfassend lässt sich sagen, dass die beschriebenen Effekte neue Informationen über Bindungspezifitäten von Antikörpern geben, welche durch synthetische Vakzin-Strukturen erzeugt wurden. Besonders deutlich wird die Beteiligung der unterschiedlichen Glycosylierungsstellen des MUC1 Peptids an der Erkennung durch die Antikörper. Dies trägt dazu bei, die Kriterien nach denen sich die humorale Immunreaktion gegen Tumor-assoziierte Strukturen auf MUC1 Proteinen richtet, besser zu verstehen und Vakzin-Designs zu optimieren. Das vorgestellte Glycopeptid-Mikroarray System, konnte so zum Beispiel Antikörper-Kreuzreaktivität gegenüber verschiedenen Glycoformen aufzeigen, wie die Unterschiede zwischen Core 1 und Core 2 Erkennung

Die Glycopeptidbibliothek wurde weiterhin dazu benutzt, um die Bindung von Lektinen zu evaluieren, deren Glycan-Spezifität bereits weitgängig bekannt ist. Der Glycopeptid-Mikroarray bestätigte dabei die zu erwartenden Lektin-Bindungen. Es wurden allerdings auch einige neue Spezifitäten bezüglich unterschiedlicher gebundener Sialylierung, erweiterten *O*-Glycan Core Strukturen und Aminosäure Präferenzen entdeckt, welche die bereits bekannten Spezifitäten weiter verfeinern. Der Glycopeptid-Mikroarray wurde auch benutzt

um das Bindungsverhalten von humanem Galectin-3 gegenüber *O*-Glycosylierung vom Mucin-Typ zu untersuchen. Die physiologischen und pathophysiologischen Funktionen der Galectine werden derzeit intensiv erforscht, besonders da sie an intra- und extrazellulären Prozessen beteiligt sind, welche die Entwicklung von Tumoren beeinflussen und fördern. Beispielsweise gibt es Hinweise darauf, dass extrazelluläres Galectin-3 an der Adhäsion von zirkulierenden Tumorzellen an epitheliale Oberflächen involviert ist. Diese Adhäsionsprozesse werden durch Bindung an Tumor-assoziiertes T-antigen auf überexprimiertem MUC1 begünstigt. Bindung von MUC1 durch Galectin-3 auf den zirkulierenden Tumorzellen führt vermutlich zu einer Lokalisierung des langen MUC1 Proteins auf der Zelloberfläche, wodurch Zelladhäsionsproteine wie Selectine und Selectinliganden freigelegt werden und somit den Kontakt zu den Epithelzellen herstellen.

Galectin-3 wurde auf einem MUC1 Glycopeptid-Array in unterschiedlichen Konzentrationen inkubiert. Dabei wurde beobachtet, dass sowohl die Tumor-assoziierten T-Antigene als auch die T_N-Antigene Galectin-Bindungspartner darstellen. Die Affinität zu diesen Antigenen ist jedoch schwächer als die Bindungsaffinitäten zu den erweiterten Core 1, Core 2 und Core 3 Glycanen. Glycopeptide mit terminierenden α2,3-*N*-Acetylneuraminsäureresten zeigten erhöhte Affinität zu Galectin-3. Dies ist interessant, da bisherige Studien α2,3-Sialylierung als Modifikation beschreiben, welche entweder ohne größeren Einfluss toleriert wird oder sogar die Affinität des Galectins zum Glycan senkt. Bisherige Studien zeigten weiterhin, dass Modifikation in 2‘- oder 3‘-Position von *N*-Acetyllactosamin mit Fucose, α-Galactose oder α-*N*-Acetylgalactosamin die Bindung zu Galectin-3 verstärken, indem sie die B-Stelle der Glycanbindungstasche des Galectins besetzen. In dieser Arbeit wurde nun gezeigt, dass α2,3-Sialylierung die Bindungsaffinität gegenüber Galectin-3 steigert, vermutlich ebenfalls durch Anlagerung an die erwähnte B-Stelle. In Übereinstimmung mit bisherigen Erkenntnissen wurde beobachtet, dass Glycopeptide mit α2,6-Sialylierung, verringerte Bindung durch Galectin-3 aufwiesen. Weiterhin wurde die Bindung stark verringert, wenn die *O*-Glycane zusätzliche interne α2,6-Sialylierung aufwiesen, da Galectin-3 ebenfalls Galactose Einheiten binden kann, welche sich intern in einem Glycan befinden. Die Modifikation von Glycanen mit α2,6-verknüpfter Sialinsäure hat bedeutende biologische Funktionen. Sie wird als ein molekularer Schalter betrachtet, mit dem Zellen induzierte Apoptose durch extrazelluläres Galectin-3 regulieren können. Die gefundenen Bindungspräferenzen von Galectin-3 mit den präsentierten *O*-Glycanen sind in *Abbildung 5.8* zusammengefasst.

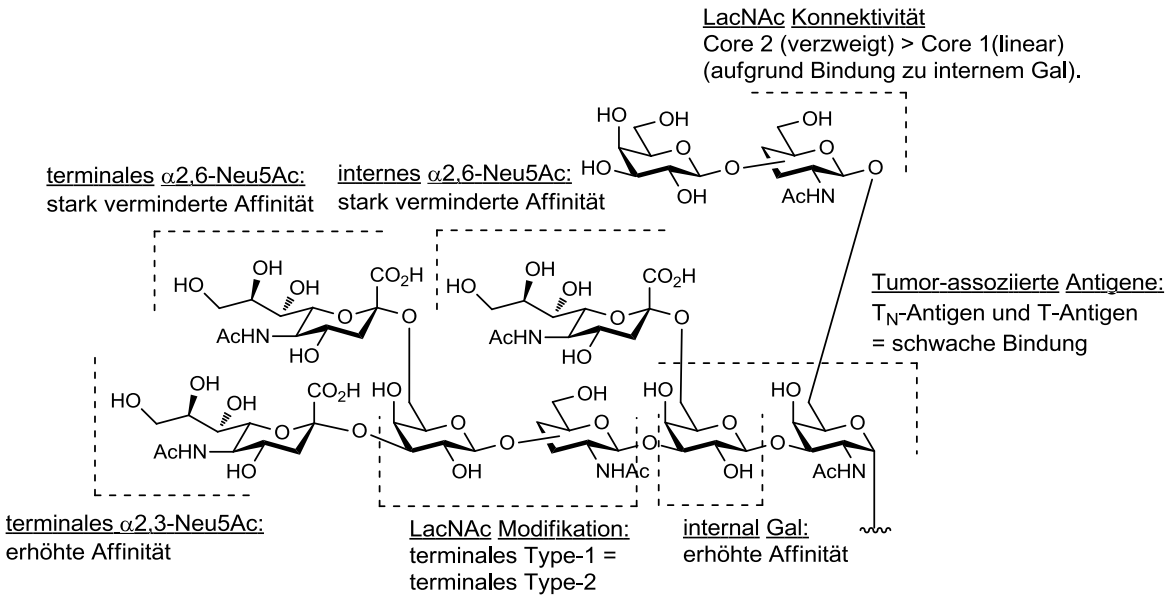


***Abbildung 6.8*:** Einfluß von *O*-Glycan-Modifikationen auf Galectin-3 Bindungsaffinität.

Die Benutzung von Glycopeptiden mit mehreren besetzten Glycosylierungsstellen, ermöglichte die Evaluierung von multivalenten Bindungseffekten. Da die Bindungsintensitäten exponentiell von mono- zu triglycosylierten Peptiden anstieg, kann gefolgert werden, dass es sich um einen multivalenten Bindungsmodus handelt, welcher Galectin-3 Oligomerisierung voraussetzt.

Das im Rahmen dieser Arbeit vorgestellte Mikroarray System erwies als geeignet um Protein-Glycan und Protein-Glycopeptid Interaktionen zu untersuchen. Die Glycopeptidbibliothek umfasst eine umfangreiche Sammlung von MUC1 Glycopeptiden mit verschiedenen *O*-Glycanen in multivalenter Präsentation. Weiterhin wurde eine MUC5B Glycopeptidbibliothek dargestellt, welche ebenso für weitere Bindingsstudien mit Lektinen verwendet werden kann. Die Diversität der hergestellten Mucin Glycopeptidbibliothek wird in Zukunft durch enzymatische Modifikation direkt auf dem Mikroarray Chip erweitert werden. Durch Reaktionen mit Sialyl-, Fucosyl-, α/β-Galactosyl- oder *N*-Acetylgalactosaminyltransferasen lassen sich beispielsweise die biologisch relevanten (Sialyl-) Lewis- oder Blutgruppenantigene darstellen. Die synthetisierten, sialylierten Glycopeptide stellen einen ersten Schritt in diese Richtung dar, da sie als interne Standards zur Validierung der enzymatischen Sialylierung auf den Mikroarrays benutzt werden können. Die in Lösung durchgeführten enzymatischen Sialylierungen dienten auch der Optimierung der Reaktionsbedingungen und zum Testen der Substratspezifität der verwendeten Sialyltransferasen bezüglich komplexen *O*-glycosylierten Peptiden. Zusätzlich zur Untersuchung von verschiedenen Lektinen, kann das Glykopeptid-Arrays System in Zukunft

auch verwendet werden um Interaktionen mit ganzen Viren oder Bakterien zu untersuchen, welche zum Beispiel verantwortlich für Infektionen in den Atemwegen sind.

7 EXPERIMENTAL

7.1 Syntheses of chapter 4.1

7.1.1 General

If not otherwise stated, all reactions were performed at room temperature. All distillations stated *in vacuo*, were performed at 40°C.

Solvents: Solvents were purchased in the quality grade *pro analysi* (p.a.) and not further purified, if not otherwise stated. Drying of dichloromethane, chloroform and acetonitrile was performed by standing over calcium hydride and subsequent distillation under argon atmosphere. Analogous, nitromethane was dried with calcium chloride. Dry methanol, toluene and diethyl ether were purchased in septum-sealed bottles with molecular sieves (*Acroseal*®) from *Acros Organics*.

Thin layer chromatography (TLC): Was performed using aluminum plates coated with silica gel (*Kieselgel 60 F_{254}*) from *Merck KGaA*, Darmstadt. Spots were detected by:

- UV light (254 nm)
- “Sugar-stain”: 10% sulfuric acid in ethanol. Visualized by heat exposure.
- Potassium permanganate solution (2 g $KMnO_4$, 13.2 g K_2CO_3, 165 mg NaOH in 200 mL H_2O). Visualized by heat exposure.

Optical rotation: Specific rotation ($[\alpha]_D^{20}$) was recorded on a *Polaritronic HH8* from *Schmidt + Heansch GmbH*, Berlin at λ = 598.5 (sodium D line). Solvents and concentrations are given with each compound.

Flash column chromatography: Compounds were purified by flash column chromatography on silica gel (*Kieselgel 60*, 0.04-0.063 nm) from *Carl Roth GmbH*, Karlsruhe. Solvents were redistilled for further use in chromatography.

Mass spectrometry: ESI-spectra were measured with a *Micromass LCT*-spectrometer from *Mircromass*, Eschborn. High-resolution ESI-spectra were measured with a *LTQ-FT ICR ultra mass* spectrometer and *Thermo Orbitrap Fusion Tribrid* spectrometer, both *Thermo Scientific*. HPLC-MS was performed with the described analytical HPLC system coupled with a *MSQ Plus* spectrometer from *Thermo Scientific*.

NMR-spectroscopy: ^{1}H and ^{13}C spectra were measured at the following instruments:

- Varian Mercury 400, *Agilent*: 400 MHz ^{1}H-NMR, 100.6 MHz ^{13}C-NMR, COSY, TOCSY, HSQC, HMBC.
- Advance DRX 500, *Bruker*: 500 MHz ^{1}H-NMR, 125.8 MHz ^{13}C-NMR, COSY, TOCSY, HSQC, HMBC.
- Ascend 600, *Bruker*: 600 MHz ^{1}H-NMR, 150.9 MHz ^{13}C-NMR, COSY, TOCSY, HSQC, HMBC.

All spectra were recorded at 295 K. The reported values for the chemical shifts δ (ppm) were calibrated to the residual proton or carbon resonance signal of the deuterated solvent, relatively correlated to the corresponding tetramethylsilane signal.[373] Signal multiplicity is assigned as follows: s = singlet, d = duplet, t = triplet, q = quartet, m = multiplet, br = broad. Elucidation of the ^{1}H and ^{13}C signals was performed by usage of the gCOSY, gTOCSY, gHSQC and gHMBC correlation experiments as stated at the compound. For interpretation of the carbohydrate ring systems, the ^{1}H and ^{13}C signals are distinguished by assignment with apostrophes (') as follows:

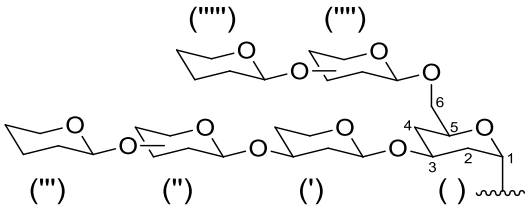


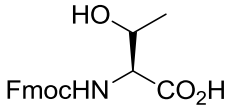
***Figure 7.1*:** Assignment of carbohydrate rings and atom positions.

- ( ): α-*N*-acetyl-galactosamine (T_N-Antigen)
- ('): β-1,3-galactose (core 1, 2) or β-1,3-*N*-acetyl-glucosamine (core 3)
- (''): β-1,3-*N*-acetyl-glucosamine (core 1, 2) or β-1,3/4-galactose (core 3)
- ('''): β-1,3/4-galactose (core 1, 2)
- (''''): β-1,6-*N*-acetyl-glucosamine (core 2)
- (''''): β-1,3/4-galactose (core 2)

Proton and carbon atoms of amino acids are assigned by indexed, greek letters.

7.1.2 Synthesis of the T_N-antigen glycosyl acceptor

***N*-(9H-fluoren-9-yl)-methoxycarbonyl-L-threonine**[237] **(2)**
(Fmoc-Thr-OH)



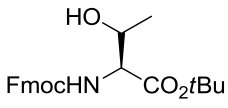
To a solution of L-threonine **1** (10.14 g, 84.8 mmol) and sodium bicarbonate (7.44 g, 88.1 mmol) in acetone/water (450 mL, 1:1) was added *N*-(9H-fluoren-9-yl)-methoxycarbonyl-succinimidyl carbonate (28.55 g, 84.5 mmol) in portions and the reaction was stirred for 24 h. Then the reaction was acidified with 12 N HCl to pH 2. The acetone was removed *in vacuo* and the aqueous phase was extracted three times with dichloromethane (200 mL) and the combined organic phases were then washed two times with of 1 N HCl and two times with water (100 mL). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo.* The residue was co-evaporated with toluene and dichloromethane (three times each).

Yield: 28.0 g (82.0 mmol, 97%), pale yellow, amorphous solid, R_f = 0.24 (EtOAc + 1vol% AcOH).

$C_{19}H_{19}NO_5$ (M = 341.36 g/mol) [341.13].

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.62 (d, 2H, H4-, H5-Fmoc, $J_{H4,H3}$ = $J_{H5,H6}$ = 7.4 Hz), 7.47 (m, 2H, H1-, H8-Fmoc), 7.27-7.16 (m, 4H, H2-, H3-, H6-, H7-Fmoc), 6.48 (s_{br}, 1H, OH), 5.99 (d, 1H, NH, $J_{NH,T\alpha}$ = 9.3 Hz), 4.37-4.26 (m, 4H, CH_2(Fmoc), T^{α}, T^{β}), 4.07 (t, 1H, H9(Fmoc), $J_{H9,CH2}$ = 7.1 Hz), 1.13 (d, 3H, T^{γ}, $J_{T\alpha,T\beta}$ = 6.2 Hz).

***N*-(9H-fluoren-9-yl)-methoxycarbonyl-L-threonine-*tert*-butylester**[238,239,240] **(3)**
(Fmoc-Thr-O*t*Bu)



A mixture of *N,N´*-dicyclohexylcarbodiimide (55.6 g, 270.0 mmol) and copper (I) chloride (650 mg, 6.6 mmol) in *tert*-butanol (25.53 g, 344.0 mmol) was stirred under argon

atmosphere in the dark for 3d. The mixture was then diluted with dry dichloromethane (30 mL) and cooled in an ice-bath. Fmoc-Thr-OH **2** (27.5 g, 80.6 mmol) was dissolved in dry dichloromethane (60 mL) and added dropwise to the stirred mixture. Stirring was continued for 45 min until TLC (toluene/acetone 4:1) indicated full consumption of Fmoc-Thr-OH **2**. Precipitated urea was removed by filtration over *Celite*® and washed with dichloromethane. The organic phase was washed with saturated sodium bicarbonate solution (four times), dried over sodium sulfate and concentrated *in vacuo*. The residue was taken up in ethyl acetate (100 mL) and stored overnight at -20°C. Further precipitated urea was filtered off over *Celite*®. The solvent was concentrated *in vacuo* and the residue was purified by flash column chromatography on silica (^{C}Hex/EtOAc 4:1 → 3:1) and subsequent recrystallization from diethyl ether/petrol ether.

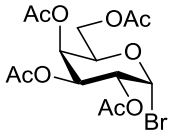
Yield: 19.1 g (48.1 mmol, 60%), colorless crystals, R_f = 0.41 (^{C}Hex/EtOAc 2:1).

$C_{23}H_{27}NO_5$ (M = 397.46 g/mol) [397.19].

ESI-MS (pos), m/z: 420.10 ($[M+Na]^+$, calc. 420.18), 363.96 ($[M-t\text{Bu}+H]^+$, calc. 364.12).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.76 (d, 2H, H4-, H5-Fmoc, $J_{H4,H3} = J_{H5,H6}$ = 7.5 Hz), 7.62 (d, 2H, H1-, H8-Fmoc, $J_{H1,H2} = J_{H8,H7}$ = 7.6 Hz), 7.40 (t, 2H, H3-, H6-Fmoc, $J_{H3,H2} = J_{H3,H4} = J_{H6,H5} = J_{H6,H7}$ = 7.4 Hz), 7.31 (m, 2H, H2-, H7-Fmoc), 5.62 (d, 1H, NH, $J_{NH,T\alpha}$ = 9.0 Hz), 4.42 (d, 2H, CH_2(Fmoc), $J_{CH2,H9}$ = 7.2 Hz), 4.30-4.22 (m, 3H, H9(Fmoc), T^{α}, T^{β}), 2.14 (s_{br}, 1H, OH), 1.49 (s, 9H, *t*Bu), 1.24 (d, 3H, T^{γ}, $J_{T\gamma,T\beta}$ = 6.4 Hz).

2,3,4,6-Tetra-*O*-acetyl-α-D-galactosylpyranosyl bromide (5)
(αAc$_4$GalBr)



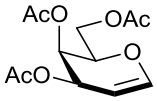
Acetic anhydride (600 mL, 6.4 mol) was cooled in an ice bath and perchloric acid (4 mL, 60%) was added. D-(+)-galactose **4** (130.25 g, 722.7 mmol) was added in portions (with a spoon, waiting for the previous portion to dissolve) and stirred for 1 h at room temperature. The solution was cooled in an ice bath and phosphorus tribromide (391 g, 1.4 mol) was added dropwise, followed by slow dropwise addition of water (90 mL, 5.1 mol) in a way to prevent the temperature from rising above 20°C. The ice bath was removed and the reaction was stirred for 3 h. The reaction was diluted with dichloromethane (800 mL) and poured into ice-water (400 mL). The mixture was stirred until the ice was completely molten. The aqueous phase was separated and the organic phase washed with saturated sodium bicarbonate solution (two times). The organic phase was dried over magnesium sulfate, filtered and concentrated *in vacuo*. The residue was crystallized from of diethyl ether (150 mL). The crystals were filtered off and washed with ice cold ether and cyclohexane.

Yield: 239.8 g (583.2 mmol, 81%), colorless crystals, R_f = 0.58 (Tol/EtOAc 3:2).

$C_{14}H_{19}BrO_9$ (M = 411.20 g/mol) [410.02].

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 6.69 (d, 1H, H1, $J_{H1,H2}$ = 4.0 Hz), 5.52 (dd, 1H, H4, $J_{H4,H5}$ = 1.4 Hz, $J_{H4,H3}$ = 3.3 Hz), 5.40 (dd, 1H, H3, $J_{H3,H4}$ = 3.3 Hz, $J_{H3,H2}$ = 10.6 Hz), 5.04 (dd, 1H, H2, $J_{H2,H1}$ = 4.0 Hz, $J_{H2,H3}$ = 10.6 Hz), 4.48 (m, 1H, H5), 4.14 (m, 2H, $H6_{ab}$), 2.15, 2.11, 2.05, 2.01 (4 x s; 4 x 3H, 4 x CH_3(Ac)).

3,4,6-Tri-*O*-acetyl-D-galactal[234,235] **(6)**



A mixture of water and acetic acid (700 mL) was cooled to -10°C. Zinc dust (112.73 g, 1.72 mol) and copper (II) sulfate pentahydrate (12.32 g, 49.3 mmol) were added. After gas generation had started, αAc_4GalBr **5** (110.11 g, 267.8 mmol) dissolved in dichloromethane (110 mL) was added dropwise. After addition, a second batch of zinc dust (55.09 g, 840.1 mmol) and copper (II) sulfate pentahydrate (6.73 g, 27.0 mmol) was added. The reaction was stirred for another 3 h. The zinc was filtered off and washed with water/acetic acid (1:1) and with dichloromethane. The filtrate was transferred into a separation funnel and extracted with dichloromethane (three times). The organic phase was washed with water and saturated sodium bicarbonate solution. The organic phase was dried over magnesium sulfate, filtered and concentrated *in vacuo*. The product was obtained after flash column chromatography on silica (^{C}Hex/EtOAc 4:1).

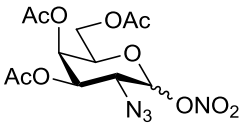
Yield: 58.7 g (215.6 mmol, 80%), colorless oil, R_f = 0.49 (Tol/EtOAc 3:2).

$C_{12}H_{16}O_7$ (M = 272.25 g/mol) [272.09].

ESI-MS (pos), m/z: 294.79 ($[M+Na]^+$, calc. 295.08).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 6.41 (dd, 1H, H1, $J_{H1,H2}$ = 7.1 Hz, $J_{H1,H3}$ = 1.7 Hz), 5.52-5.49 (m, 1H, H3), 5.38-5.36 (m, 1H, H4), 4.68 (ddd, 1H, H2, $J_{H2,H1}$ = 6.3 Hz, $J_{H2,H3}$ = 2.7 Hz, $J_{H2,H4}$ = 1.5 Hz), 4.30-4.14 (m, 3H, H5, $H6_{ab}$), 2.08, 2.03, 1.99 (3 x s, 3 x 3H, 3 x CH_3(Ac)).

3,4,6-Tri-*O*-acetyl-2-azido-2-desoxy-α/β-D-galactosylpyranosyl nitrate[236] (7) (α/βAc$_3$GalN$_3$-ONO$_2$)



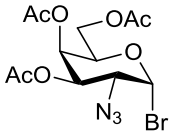
Cerium (IV) ammonium nitrate (114.75 g, 209.2 mmol) and sodium azide (7.03 g, 108.1 mmol (both thoroughly grinded in a mortar and vacuum-dried) were carefully added to dry acetonitrile, cooled to -20°C under argon atmosphere. Galactal **6** (18.66 g, 68.5 mmol) was dissolved in dry acetonitrile (150 mL) and added dropwise to the flask. The temperature was kept between -10°C and -20°C and the reaction was stirred until TLC (Tol/EtOAc 7:3) indicated full consumption of the starting material (after 5.5 h). The reaction mixture was diluted with diethyl ether (500 mL) and washed with ice cold water (three times, 300 mL each). The organic phase was dried over sodium sulfate and the solvent was concentrated *in vacuo.* The product was obtained from the residue after flash column chromatography on silica (^{C}Hex/EtOAc 4:1).

Yield: 15.03 g (α/β 1:1, 39.9 mmol, 58 %), light yellow, amorphous solid, R_f = 0.56 (Tol/EtOAc 7:3).

$C_{12}H_{16}N_4O_{10}$ (M = 376.28 g/mol) [376.09].

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 6.33 (d, 1H, H1$^\alpha$, $J_{H1\alpha,H2\alpha}$ = 4.2 Hz), 5.58 (d, 1H, H1$^\beta$, $J_{H1\beta,H2\beta}$ = 8.8 Hz), 5.49 (m, 1H, H4$^\alpha$), 5.38 (m, 1H, H4$^\beta$), 5.25 (dd, 1H, H3$^\alpha$, $J_{H3\alpha,H2\alpha}$ = 11.3 Hz, $J_{H3\alpha,H4\beta}$ = 3.2 Hz), 4.96 (dd, 1H, H3$^\beta$, $J_{H3\beta,H2\beta}$ = 10.6 Hz, $J_{H3\beta,H4\beta}$ = 3.3 Hz), 4.36 (m, 1H, H5$^\alpha$), 4.14-4.03 (m, 6H, H2$^\alpha$, H5$^\beta$, H6$^\alpha{}_{ab}$, H6$^\beta{}_{ab}$), 3.81 (dd, 1H, H2$^\beta$, $J_{H2\beta,H1\beta}$ = 8.8 Hz, $J_{H2\beta,H3\beta}$ = 10.6 Hz), 2.16 (s, 6H, CH_3(Ac$^\alpha$), CH_3(Ac$^\beta$)), 2.07, 2.06, 2.03, 2.02 (4 x s, 4 x 3H, CH_3(Ac$^\alpha$), CH_3(Ac$^\beta$)).

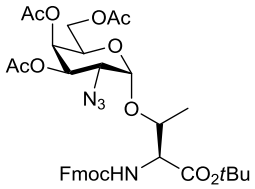
3,4,6-Tri-*O*-acetyl-2-azido-2-desoxy-α-D-galactosylpyranosyl bromide[236] (8) (αAc$_3$GalN$_3$-Br)



α/βAc$_3$GalN$_3$-ONO$_3$ **7** (14.52 g, 38.5 mmol) was dissolved in dry acetonitrile (300 mL). Lithium bromide (34.56 g, 397.2 mmol, vacuum-dried and grinded in a mortar before usage) was added and the suspension was stirred for 16 h. The reaction was diluted with dichloromethane (300 mL) and washed with brine (three times, 100 mL each). The organic phase was dried over sodium sulfate, concentrated *in vacuo* and the residue was purified by flash column chromatography on silica (^{C}Hex/EtOAc 4:1).

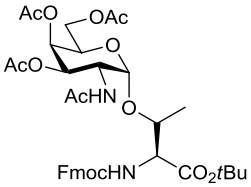
Yield: 12.83 g (32.7 mmol, 85%), colorless, amorphous solid, R_f = 0.64 (Tol/EtOAc 7:3).
$C_{12}H_{16}BrN_3O_7$ (M = 394.18 g/mol) [393.04].
^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 6.46 (d, 1H, H1, $J_{H1,H2}$ = 3.8 Hz), 5.49 (dd, 1H, H4, $J_{H4,H3}$ = 3.2 Hz, $J_{H4,H5}$ = 1.3 Hz), 5.33 (dd, 1H, H3, $J_{H3,H2}$ = 10.7 Hz, $J_{H3,H4}$ = 3.2 Hz), 4.47 (m, 1H, H5), 4.19-4.11 (m, 2H, $H6_{a,b}$), 3.97 (dd, 1H, H2, $J_{H2,H3}$ = 10.7 Hz, $J_{H2,H1}$ = 3.8 Hz), 2.15, 2.05, 2.04 (3 x s, 3 x 3H, CH_3(Ac)).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(3,4,6-tri-_O_-acetyl-2-azido-2-desoxy-α-D-galactosylpyranosyl)-L-threonine-_tert_-butylester[240,208,221,241] (9)
(Fmoc-Thr(αAc$_3$GalN$_3$)-O*t*Bu)



Fmoc-L-threonine-O*t*Bu **3** (22.24 g, 55.8 mmol) was dissolved in dry dichloromethane/toluene (300 mL, 1:1) and stirred with molecular sieves (20.3 g, 4Å) for 45 min under argon atmosphere. The mixture was cooled in an ice bath. Then silver carbonate (21.05 g, 76.1 mmol, vacuum dried) and silver perchlorate monohydrate (3.42 g,15.2 mmol, three times carefully co-evaporated with 20 mL dry toluene each) in dry toluene (40 mL) were added and the reaction was stirred for 30 min in the dark. Then αAc$_3$GalN$_3$-Br **8** (19.85 g, 49.9 mmol) dissolved in dry dichloromethane/toluene (200 mL, 1:1) was added dropwise and the reaction was stirred for 24 h in the dark while the reaction warmed up to room temperature. The molecular sieves were filtered off over *Celite*® and washed with dichloromethane (400 mL). The solution was washed with saturated sodium bicarbonate solution (three times) and with water and brine (once each). The organic phase was separated, dried over sodium sulfate, filtered and the concentrated *in vacuo*. The crude product was purified by flash column chromatography on silica (DCM/EtOAc 40:1).
Yield: 21.39 g (30.1 mmol, 60%), colorless, amorphous solid, R_f = 0.34 (DCM/EtOAc 20:1).
$C_{35}H_{42}N_4O_{12}$ (M = 710.73 g/mol) [710.28].
ESI-MS (pos), m/z: 710.80 ([M+H]$^+$, calc. 711.28); 733.13 ([M+Na]$^+$, calc. 733.72).
^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.76 (d, 2H, H4-, H5-Fmoc, $J_{H4,H3}$ = $J_{H5,H6}$ = 7.5 Hz), 7.63 (d, 2H, H1-, H8-Fmoc, $J_{H1,H2}$ = $J_{H8,H7}$ = 7.4 Hz), 7.42-7.31 (m, 4H, H2-, H7-, H3-, H6-Fmoc), 5.65 (d, 1H, NH-Fmoc, $J_{NH,T\alpha}$ = 9.4 Hz), 5.47 (m, 1H, H4), 5.34 (dd, 1H, H3, $J_{H3,H2}$ = 11.2 Hz, $J_{H3,H4}$ = 3.0 Hz), 5.11 (d, 1H, H1, $J_{H1,H2}$ = 3.6 Hz), 4.45-4.25 (m, 6H, H9-, CH_2(Fmoc), H5, T^{α}, T^{β}), 4.13-4.09 (m, 2H, $H6_{ab}$), 3.64 (dd, 1H, H2, $J_{H2,H3}$ = 11.2 Hz, $J_{H2,H1}$ = 3.7 Hz), 2.15, 2.07, 2.04 (3x s, 3 x 3H, 3 x CH_3(Ac)), 1.51 (s, 9H, *t*Bu), 1.35 (d, 3H, T^{γ}, $J_{T\gamma,T\beta}$ = 6.4 Hz).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-acetamindo-3,4,6-tri-_O_-acetyl-2-desoxy-α-D-galactosylpyranosyl)-L-threonine-_tert_-butylester[243] (10)
(Fmoc-Thr(αAc$_3$GalNAc)-O*t*Bu)



Fmoc-Thr(αAc$_3$GalN$_3$)-O*t*Bu **9** (21.33 g, 30.0 mmol) was dissolved in pyridine/thioacetic acid (100 mL, 1:1) and stirred at room temperature for 2 d. The precipitate was filtered off and washed with pyridine and toluene. The filtrate was co-evaporated with toluene. The residue was taken up with ethyl acetate (100 mL), further precipitate was filtered off, washed and the filtrate concentrated *in vacuo.* The product was purified by flash column chromatography on silica (^{C}Hex/EtOAc 1:1 → 1:2).

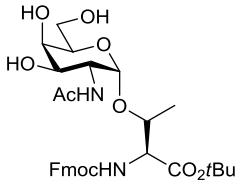
Yield: 16.61 g (22.8 mmol, 76%), colorless, amorphous solid, R_f = 0.27 (^{C}Hex/EtOAc 1:2).

$C_{37}H_{46}N_2O_{13}$ (M = 726.77 g/mol) [726.30].

ESI-MS (pos), m/z: 748.98 ([M+Na]$^+$, calc. 749.29), 727.01 ([M+H]$^+$, calc. 727.31), 670.80 ([M-*t*Bu+H]$^+$, calc. 671.24).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.77 (d, 2H, H4-, H5-Fmoc, $J_{H4,H3}$ = $J_{H5,H6}$ = 7.7 Hz), 7.64 (d, 2H, H1-, H8-Fmoc, $J_{H1,H2}$ = $J_{H8,H7}$ = 7.6 Hz), 7.42-7.33 (m, 4H, H2-, H7-, H3-, H6-Fmoc), 5.93 (d, 1H, NH-Fmoc, $J_{NH,T\alpha}$ = 9.8 Hz), 5.45 (d, 1H, NH-Ac, $J_{NH,H2}$ = 9.4 Hz), 5.39 (m, 1H, H4), 5.08 (dd, 1H, H3, $J_{H3,H2}$ = 11.3 Hz, $J_{H3,H4}$ = 3.2 Hz), 4.89 (d, 1H, H1, $J_{H1,H2}$ = 3.7 Hz), 4.61 (m, 1H, H2), 4.38 (m, 2H, CH_2(Fmoc)), 4.22-4.12 (m, 4H, H9(Fmoc), H5, T$^\alpha$, T$^\beta$), 2.16, 2.04, 2.00 (3 x s, 4 x 3H, 4 x CH_3(Ac)), 1.46 (s, 9H, *t*Bu), 1.26 (d, 3H, T$^\gamma$, $J_{T\gamma,T\alpha}$ = 6.4 Hz).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-acetamindo-2-desoxy-α-D-galactosylpyranosyl)-L-threonine-_tert_-butylester[208,237] (11)
(Fmoc-Thr(αGalNAc)-O*t*Bu)



Fmoc-Thr(αAc$_3$GalNAc)-O*t*Bu **10** (27.98 g, 38.5 mmol) was dissolved in methanol (250 mL) and a solution of sodium methoxide (1 wt%) in methanol was added slowly in a dropwise manner, until pH 9.5 was established (wet pH paper). After 24 h TLC showed the main deacetylated product plus product with cleavage of the Fmoc-group. The reaction was

neutralized with acidic cation exchange resin (*Dowex 50WX8*). The resin was filtered off and the solvent removed *in vacuo*. The residue was taken up with 1,4-dioxane (150 mL) and a solution of sodium bicarbonate (4.09 g, 48.7 mmol) in water (150 mL) was added. Then *N*-9-Fluorenylmethyloxycarbonyl-succinimidylcarbonate (6.89 g, 20.4 mmol) was added and the suspension was stirred for 16 h at room temperature. The majority of the 1,4-dioxane was removed *in vacuo* and of ethyl acetate (600 mL) was added. The organic phase was washed with saturated sodium bicarbonate, water and brine. The organic phase was dried over sodium sulfate, filtered and the concentrated *in vacuo*. The residue was purified by flash column chromatography on silica (EtOAc → EtOAc/MeOH 15:1).

Yield: 21.00 g (35.0 mmol, 91%), colorless, amorphous solid, R_f = 0.26 (EtOAc/MeOH 15:1).

$C_{31}H_{40}N_2O_{10}$ (M = 600.66 g/mol) [600.27].

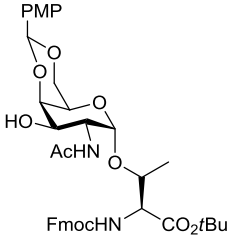
ESI-MS (pos), m/z: 601.00 ($[M+H]^+$, calc. 601.28).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.76 (d, 2H, H4-, H5-Fmoc, $J_{H4,H3} = J_{H5,H6}$ = 7.3 Hz), 7.60 (d, 2H, H1-, H8-Fmoc, $J_{H1,H2} = J_{H8,H7}$ = 7.6 Hz), 7.41-7.29 (m, 4H, H2-, H7-, H3-, H6-Fmoc), 6.86 (d, 1H, NH-Ac, $J_{NH,H2}$ = 7.8 Hz), 5.75 (d, 1H, NH-Fmoc, $J_{NH,T\alpha}$ = 9.4 Hz), 4.95 (s_{br}, 1H, OH), 4.87 (d, 1H, H1, $J_{H1,H2}$ = 3.0 Hz), 4.44-4.40 (m, 2H, CH_2(Fmoc)), 4.31-4.24 (m, 3H, H9(Fmoc), H2, T^{α}), 4.14 (m, 1H, T^{β}), 4.04 (m, 1H, H4), 3.92-3.82 (m, 5H, H3, H5, $H6_{ab}$, OH), 2.10 (s, 3H, CH_3(Ac)), 1.44 (s, 9H, *t*Bu), 1.29 (d, 3H, T^{γ}, $J_{T\gamma,T\beta}$ = 6.2 Hz).

^{13}C-NMR (100.6 MHz, $CDCl_3$), *δ (ppm)*: 171.29, 171.06 (C=O(Ac), COO*t*Bu), 156.66 (C=O(Fmoc)), 143.28 ($C1_a$-, $C8_a$-Fmoc), 141.45 ($C4_a$-, $C5_a$-Fmoc), 127.91 (C3-, C6-Fmoc), 127.22 (C2-, C7-Fmoc), 125.13 (C1-, C8-Fmoc), 120.14 (C4-, C5-Fmoc), 99.73 (C1), 83.44 (C_q(*t*Bu)), 76.51 (T^{β}), 71.01, 70.21, 69.92 (C3, C4, C5), 67.32 (CH_2(Fmoc)), 62.70 (C6), 59.12 (T^{α}), 50.99, 47.34 (C9(Fmoc), C2), 28.20 (*t*Bu), 23.06 (CH_3(Ac)), 21.18 (T^{α}).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-acetamindo-4,6-_O_-_para_-methoxybenzylidenacetal-2-desoxy-α-D-galactosylpyranosyl)-L-threonine-_tert_-butylester (12)

(Fmoc-Thr(PMP-GalNAc)-O*t*Bu)



A solution of Fmoc-Thr(αGalNAc)-O*t*Bu **11** (15.06 g, 25.8 mmol) and *para*-methoxybenzaldehyde dimethyl acetal (5.94 g, 32.7 mmol) in dry acetonitrile (150 mL) was adjusted to pH 4.5 by addition of *para*-toluenesulfonic acid monohydrate (careful addition

with the tip of a spatula). The reaction was stirred for 2.5 h and then neutralized with a few drops of *N,N*-diisopropylethylamine. The solvent was removed *in vacuo* and the residue was purified by flash column chromatography on silica (^{C}Hex/EtOAc 1:1 → 1:2).

Yield: 16.38 g (22.8 mmol, 91%), colorless, amorphous solid, R_f = 0.33 (Tol/EtOAc 1:2).

$C_{39}H_{46}N_2O_{11}$ (M = 718.79 g/mol) [718.31].

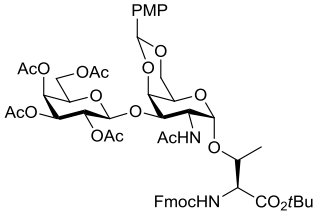
ESI-MS (pos), m/z: 718.93 ([M+H]$^+$, calc. 719.32), 741.13 ([M+Na]$^+$, calc. 741.30), 1436.80 ([2M+H]$^+$, calc. 1437.63).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.77 (d, 2H, H4-, H5-Fmoc, $J_{H4,H3} = J_{H5,H6}$ = 7.5 Hz), 7.61 (d, 2H, H1-, H8-Fmoc, $J_{H1,H2} = J_{H8,H7}$ = 8.0 Hz), 7.45-7.32 (m, 6H, H2-, H7-, H3-, H6-Fmoc, H2-, H6-PMP), 6.88 (d, 2H, H3-, H5-PMP, $J_{H2,H1} = J_{H3,H4}$ = 8.4 Hz), 6.50 (d, 1H, NH-Fmoc, $J_{NH,T\alpha}$ = 8.6 Hz), 5.60 (d, 1H, NH-Ac, $J_{NH,H2}$ = 9.5 Hz), 5.50 (s, 1H, CH(PMP)), 4.95 (d, 1H, H1, $J_{H1,H2}$ = 3.6 Hz), 4.52-4.43 (m, 3H, CH_2(Fmoc), H2), 4.27-4.01 (m, 6H, H9(Fmoc), H5, $H6_{a,b}$, T^{α}, T^{β}), 3.88-3.83 (m, 1H, H3), 3.79 (s, 3H, CH_3(PMP)), 3.68 (s, 1H, H4), 3.17 (s_{br}, 1H, OH), 2.10 (s, 3H, CH_3(Ac)), 1.46 (s, 9H, CH_3(*t*Bu)), 1.29 (d, 3H, T^{γ}, $J_{T\gamma,T\beta}$ = 5.9 Hz).

^{13}C-NMR (100.6 MHz, $CDCl_3$), *δ (ppm)*: 172.54, 170.84 (C=O(Ac), COO*t*Bu), 160.28 (C_{para}(PMP)), 156.59 (C=O(Fmoc)), 143.80 ($C1_a$-, $C8_a$-Fmoc), 141.45 ($C4_a$-, $C5_a$-Fmoc), 132.11 (C_{arom}(PMP)), 130.23 (C_{ipso}(PMP)), 127.90, 127.81, 127.12 (C3-, C6-Fmoc, C_{arom}(PMP)), 125.17 (C1-, C8-Fmoc), 120.14 (C4-, C5-Fmoc), 114.43, 113.69 (C_{arom}(PMP)), 101.27 (CH(PMP)), 100.58 (C1), 83.46 (C_q(*t*Bu), 76.46 (T^{β}), 75.63 (C5), 69.72 (C3), 69.27 (C6), 67.32 (CH_2(Fmoc)), 63.77 (C4), 59.08 (T^{α}), 55.41 (CH_3(PMP)), 47.33 (C9(Fmoc)), 28.21 (*t*Bu), 23.19 (CH_3(Ac)), 19.13 (T^{γ}).

7.1.3 Synthesis of the T-antigen amino acid

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-acetamindo-3-_O_-[2,3,4,6-tetra-_O_-acetyl-α-D-galactosylpyranosyl]-4,6-_O-para_-methoxybenzylidenacetal-2-desoxy-α-D-galactosylpyranosyl)-L-threonine-_tert_-butylester[220,244,245] (13)
(Fmoc-Thr(αAc$_4$Gal-(1→3)-αPMP-GalNAc)-O*t*Bu)



Fmoc-Thr(PMP-GalNAc)-O*t*Bu **12** (2.72 g, 3.78 mmol) was dissolved in a mixture of dry nitromethane and dry dichloromethane (105 mL, 2:1) under argon atmosphere. Activated molecular sieves (4 g, 4Å) and mercury (II) cyanide (2.87 g, 11.36 mmol) were added in the

dark. The mixture was stirred for 45 min and then αAc$_4$GalBr **5** (4.69 g, 11.41 mmol) dissolved in dry dichloromethane (5 mL) was added in portions (1 mL) over 40 min and the reaction was stirred in the dark. After 3 h another of portion of mercury (II) cyanide (950 mg, 3.76 mmol) was added and the reaction was stirred for 20 h in the dark. The molecular sieves were filtered off over *Celite*® and washed with dichloromethane. The dichloromethane phase was washed with saturated sodium hydrogen carbonate solution (three times, 100 mL each), saturated sodium iodide solution (two times, 80 mL each) and once with brine (100 mL). The organic phase was separated, dried over sodium sulfate, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography on silica (^{C}Hex/EtOAc 1:1 → 1:2).

Yield: 2.77 g (2.64 mmol, 70%), colorless, amorphous solid, $[\alpha]_D^{20}$ +74.76 (c = 0.49, $CHCl_3$), R_f = 0.48 (Tol/EtOAc 1:3).

$C_{53}H_{64}N_2O_{20}$ (M = 1048.41 g/mol) [1049.41].

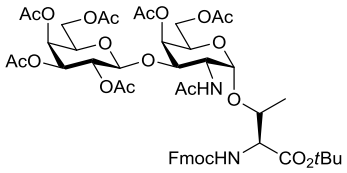
ESI-MS (pos), m/z: 1071.20 ([M+Na]$^+$, calc. 1071.39).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.71 (d, 2H, H4-, H5-Fmoc, $J_{H4,H3}$ = $J_{H5,H6}$ = 7.2 Hz), 7.54 (m, 2H, H1-, H8-Fmoc), 7.39 (d, 2H, H2-, H6-PMP, $J_{H2,H3}$ = $J_{H6,H5}$ = 8.6 Hz), 7.36-7.23 (m, 4H, H2-, H3-, H6-, H7-Fmoc), 6.81 (d, 2H, H3-, H5-PMP, $J_{H3,H2}$ = $J_{H5,H6}$ = 8.8 Hz), 5.75 (d, 1H, NH-Ac, $J_{NH,H2}$ = 7.3 Hz), 5.50-5.44 (m, 2H, NH-Fmoc, CH(PMP)), 5.31 (s, 1H, H4'), 5.12 (dd, 1H, H2', $J_{H2',H1'}$ = 7.9 Hz, $J_{H2',H3'}$ = 9.7 Hz), 4.91-4.88 (m, 2H, H1, H3'), 4.66-4.60 (m, 2H, H1', H2), 4.47-4.38 (m, 2H, CH_2(Fmoc)), 4.18-3.82 (m, 10H, H3, H5, H6$_{a,b}$, H5', H6'$_{a,b}$, H9(Fmoc), T^α, T^β), 3.72 (s, 3H, CH_3(PMP)), 3.59 (m, 1H, H4), 2.28, 2.06, 1.97, 1.94, 1.90 (5 x s, 5 x CH_3(Ac)), 1.39 (s, 9H, *t*Bu), 1.23-1.18 (m, 3H, T^γ).

^{13}C-NMR (100.6 MHz, HSQC, $CDCl_3$), *δ (ppm)*: 170.50, 170.39, 170.27, 169.80, 169.56 (C=O(Ac), COO*t*Bu), 160.07 (C_{para}(PMP)), 156.58 (C=O(Fmoc)), 143.78 (C1$_a$-, C8$_a$-Fmoc), 141.46 (C4$_a$-, C5$_a$-Fmoc), 130.31 (C_{arom}(PMP)), 129.13, 128.32, 127.94, 127.65, 127.19 (C3-, C6-, C2-, C7-Fmoc, 3 x C_{arom}(PMP)), 125.39 (C1-, C8-Fmoc), 120.19 (C4-, C5-Fmoc), 113.63 (C_{arom}(PMP)), 101.38 (C1), 100.58 (CH(PMP)), 100.51 (C1'), 83.30 (C_q(*t*Bu)), 76.46 (T^β), 75.69 (C5), 74.39 (C3), 71.02 (C3', C5'), 69.16 (C6), 68.86 (C2'), 67.14 (C4', CH_2(Fmoc)), 63.82 (C4), 61.44 (C6'), 59.21 (T^α), 55.40 (CH_3(PMP)), 47.96 (C2), 47.39 (C9(Fmoc)), 28.21 (*t*Bu), 23.60, 21.55, 20.82, 20.67 (CH_3(Ac)), 19.07 (T^γ).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-acetamindo-3-_O_-[2,3,4,6-tetra-_O_-acetyl-α-D-galactosylpyranosyl]-4,6-_O_-acetyl-2-desoxy-α-D-galactosylpyranosyl)-L-threonine-_tert_-butylester[220] (14)

(Fmoc-Thr(αAc$_4$Gal-(1→3)-αAc$_2$GalNAc)-O*t*Bu)



A solution of Fmoc-Thr(αAc$_4$Gal-(1→3)-αPMP-GalNAc)-O*t*Bu **13** (833 mg, 0.85 mmol) in 80% acetic acid (15 mL) was stirred at 50°C for 4 h. The reaction was co-evaporated with toluene. The residue was taken up with pyridine/acetic anhydride (15 mL, 2:1) and stirred for 16 h. The reaction was concentrated *in vacuo* and the residue was co-evaporated with toluene. The crude product was purified by flash column chromatography on silica (^{C}Hex/EtOAc 1:2).

Yield: 782 mg (0.77 mmol, 90%), colorless, amorphous solid, $[\alpha]_D^{20}$ +50.20 (c = 0.51, $CHCl_3$), R_f = 0.48 (Tol/EtOAc).

ESI-MS (pos), m/z: 1014.93 ($[M+H]^+$, calc. 1015.39).

$C_{49}H_{62}N_2O_{21}$ (M = 1015.02 g/mol) [1014.38].

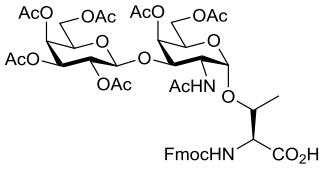
^{1}H NMR (400 MHz, 25°C, internal $CDCl_3$, *δ*(H) = 7.26 ppm): 7.78 (d, ^3J(H-4,H-3) = ^3J(H-5,H-6) = 7.6 Hz, 2H, H-4-, H-5-Fmoc), 7.61 (d, $^3J_{H1,H2}$ = $^3J_{H8,H7}$ = 7.0 Hz, 2H, H-1-, H-8-Fmoc), 7.42 (t, ^3J(H-3,H-4) = ^3J(H-3,H-2) = ^3J(H-6,H-5) = ^3J(H-6,H-7) = 7.5 Hz, 2H, H-3-, H-6-Fmoc), 7.33 (t, ^3J(H-2,H-3) = ^3J(H-2,H-1) = ^3J(H-7,H-8) = ^3J(H-7,H-6) = 7.5 Hz, 2H, H-2-, H-7-Fmoc), 5.99 (d, ^3J(NH,H-2) = 9.8 Hz, 1H, NH-Ac), 5.44 (d, ^3J(NH,H-2) = 8.9 Hz, 1H, NH-Fmoc), 5.37-5.34 (m, 2H, H-4, H-4'), 5.09 (dd, ^3J(H-2',H-1') = 7.7 Hz, ^3J(H-2', H-3') = 10.5 Hz, 1H, H-2'), 4.94 (dd, ^3J(H-3',H-2') = 10.5 Hz, ^3J(H-3',H-4') = 3.0 Hz), 4.83 (s, 1H, H-1), 4.59-4.49 (m, 4H, H-1', H-2, CH_{2ab}-Fmoc), 4.27-4.09 (m, 7H, H-6$_{ab}$, H-6'$_{ab}$, H-9-Fmoc, T$^\alpha$, T$^\beta$), 4.01-3.97 (m, 1H, H-5'), 3.88-3.85 (m, 1H, H-5'), 3.79 (d, ^3J(H-3,H-2) = 8.5 Hz), 2.16, 2.13, 2.06, 2.05, 2.04, 2.01, 1.97 (7 x s, 21 H, CH_3-(Ac)), 1.46 (s, 9H, CH_3-*t*Bu), 1.29 (d, ^3J(T$^\gamma$,T$^\beta$) = 5.5 Hz, 1H, T$^\gamma$).

^{13}C NMR (100.6 MHz, 25°C, internal $CDCl_3$, *δ*(H) = 77.16 ppm): 170.54, 170.28, 170.00, 169.66 (C=O-(Ac)), 156.04 (C=O-(Fmoc)), 143.79 (C1$_a$-, C8$_a$-Fmoc), 141.51 (C4$_a$-, C5$_a$-Fmoc), 128.02 (C2-, C7-Fmoc), 127.26 (C3-, C6-Fmoc), 125.01 (C1-, C8-Fmoc), 120.26 (C4-, C5-Fmoc), 101.08 (C-1, C-1'), 83.45 (C_{quart}-*t*Bu), 78.0 (C-3), 76.5 (T$^\beta$), 73.39, 70.89, 70.85, 69.27 (C-3', C-4, C-5, C-5'), 68.17 (C-2'), 66.90 (C-4', CH_2-Fmoc), 63.14 (C-6), 61.14 (C-6'), 59.18 (T$^\alpha$), 48.57 (C-2), 47.43 (C-9-Fmoc), 28.25 (CH_3-(tBu)), 20.87, 20.86, 20.84, 20.79, 20.70 (CH_3-(Ac)), 18.7 (T$^\gamma$).

***N*-(9H-fluoren-9-yl)-methoxycarbonyl-*O*-(2-acetamindo-3-*O*-[2,3,4,6-tetra-*O*-acetyl-α-D-galactosylpyranosyl]-4,6-*O*-acetyl-2-desoxy-α-D-galactosylpyranosyl)-L-threonine**[220]

(15)

(Fmoc-Thr(αAc$_4$Gal-(1→3)-αAc$_2$GalNAc)-OH)



Fmoc-Thr(αAc$_4$Gal-(1→3)-αAc$_2$GalNAc)-O*t*Bu **14** (750 mg, 0.74 mmol) was dissolved in dichloromethane (3 mL) and anisole (1 mL) and trifluoroacetic acid (9 mL) was added. The reaction was stirred for 5 h and then co-evaporated with toluene. The crude product was purified by flash column chromatography on silica (EtOAc→EtOAc/MeOH 10:1).

Yield: 638 mg (0.67 mmol, 90%), colorless, amorphous solid, $[\alpha]_D^{20}$ +75.17 (c = 0.99, $CHCl_3$) R_f = 0.27 (EtOAc/MeOH 4:1).

$C_{45}H_{54}N_2O_{21}$ (M = 958.91 g/mol) [958.32].

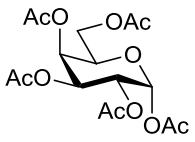
ESI-MS (pos), m/z: 958.93 ($[M+H]^+$, calc. 959.33), 1917.46 ($[2M+H]^+$, calc. 1917.65).

^{1}H-NMR (400 MHz, gCOSY, DMSO), *δ (ppm)*: 7.90 (d, 2H, H4-, H5-Fmoc, $J_{H4,H3} = J_{H5,H6}$ = 7.8 Hz), 7.74 (m, 2H, H1-, H8-Fmoc), 7.45-7.29 (m, 4H, H2-, H3-, H6-, H7-Fmoc), 5.30 (d, 1H, H4, $J_{H4,H3}$ = 3.3 Hz), 5.24 (d, 1H, H4', $J_{H4',H3'}$ = 3.6 Hz), 5.05 (dd, 1H, H3', $J_{H3',H2'}$ = 10.4 Hz, $J_{H3',H4'}$ = 3.7 Hz), 4.84 (dd, 1H, H2', $J_{H2',H1'}$ = 7.8 Hz, $J_{H2',H3'}$ = 10.2 Hz), 4.74 (d, 1H, H1', $J_{H1',H2'}$ = 8.0 Hz), 4.70 (d, 1H, H1, $J_{H1,H2}$ = 4.0 Hz), 4.54-4.44 (m, 2H, CH_2-Fmoc), 4.32-4.26 (m, 2H, H9-Fmoc, T^{β}), 4.13-4.02 (m, 6H, H2, H3, H5', $H6_a$, $H6_a$', T^{α}), 3.93-3.84 (m, 3H, H5, $H6_b$, $H6_b$'), 2.10, 2.04, 2.00, 1.98, 1.90, 1.83 (6 x s, 6 x 3H, 6 x CH_3-Ac), 1.15 (d, 3H, T^{γ}, $J_{T\gamma,T\beta}$ = 6.3 Hz).

^{13}C-NMR (100.6 MHz, HSQC, $CDCl_3$), *δ (ppm)*: 171.96, 171.69, 170.04, 169.90, 169.81, 169.75, 169.49 (C=O(Ac)), 156.81 (C=O(Fmoc)), 143.71 ($C1_a$-, $C8_a$-Fmoc), 140.79 ($C4_a$-, $C5_a$-Fmoc), 128.31 (C2-, C7-Fmoc), 127.78 (C3-, C6-Fmoc), 125.82 (C1-, C8-Fmoc), 120.92 (C4-, C5-Fmoc), 100.42 (C1'), 98.75 (C1), 84.78 (C_{quart}-*t*Bu), 75.29 (T^{γ}), 74.47 (C5), 70.91 (C3'), 70.24 (C4, C3), 69.14 (C2'), 68.01 (C5'), 67.79 (C4'), 66.23 (CH_2-Fmoc), 63.78 (C6), 61.49 (C6'), 59.19 (T^{α}), 48.50 (C2), 47.47 (C9-Fmoc), 31.28 (CH_3-*t*Bu), 22.72, 20.63, 20.48, 20.45, 20.36, 20.30 (CH_3-Ac), 18.41 (T^{γ}).

7.1.4 Syntheses of the type-1 *N*-acetyllactosamine glycosyl donor

1,2,3,4,6-Penta-*O*-acetyl-α-D-galactosylpyranoside (16)
(αAc_5Gal)



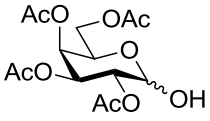
Perchloric acid (2 mL, 60%) was added to acetic anhydride (400 mL, 4.2 mol), cooled in an ice bath. D-(+)-Galactose **4** (70.82 g, 393 mmol) was added in portions (with a spoon, waiting for the previous portion to dissolve). After addition, the ice-bath was removed and the reaction was stirred for 4 h. Then the reaction was poured into ice water (1.2 L) and stirred for 1 h. The precipitate was filtered off and washed with water. The solid precipitate was dissolved in ethyl acetate (600 mL) and washed with saturated sodium hydrogen bicarbonate, water and brine (150 mL each). The organic phase was dried over sodium sulfate, filtered and the solvent removed in *vacuo*. The product was obtained by crystallization from ethyl acetate/cyclohexane.

Yield: 122.7 g (314.4 mmol, 80 %), colorless crystals, R_f = 0.44 (Tol/EtOAc 3:2).

$C_{16}H_{22}O_{11}$ (M = 390.34 g/mol) [390.12].

^{1}*H-NMR* (400 MHz, $CDCl_3$), *δ (ppm)*: 6.31 (d, 1H, H1, $J_{H1,H2}$ = 2.5 Hz), 5.44 (m, 1H, H4), 5.27 (m, 2H, H3, H2), 4.29 (m, 1H, H5), 4.06-4.02 (m, 2H, $H6_{a,b}$), 2.10, 1.98, 1.96, 1.94 (4 x s, 5 x 3H, 5 x CH_3-Ac).

2,3,4,6-Tetra-*O*-acetyl-α/β-D-galactosylpyranose (17)
($\alpha/\beta Ac_4Gal$-OH)



αAc_5Gal **16** (45.24 g, 115.8 mmol) was dissolved in dry *N,N*-dimethylformamid (200 mL) and hydrazine acetate (12.45 g, 138.2 mmol) was added. The reaction was stirred for 3 h at 50°C. Ethyl acetate (1 L) was added and washed with water (two times, 600 mL) and with brine (400 mL). The organic phase was dried over sodium sulfate, filtered and the solvent removed *in vacuo*. The product was purified by flash column chromatography on silica (Tol/EtOAc 2:1).

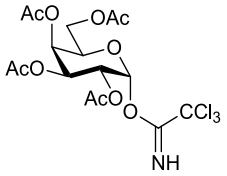
Yield: 36.24 g (α/β 3.2:1, 104.0 mmol, 90%), colorless, viscous oil, R_f = 0.23 (Tol/EtOAc).

$C_{14}H_{20}O_{10}$ (M = 348.30 g/mol) [348.11].

^{1}*H-NMR* (400 MHz, $CDCl_3$), *δ (ppm)*: 5.49 (t, 1H, H1, H1α, $J_{H1\alpha,H2\alpha} = J_{H1\alpha,OH\alpha}$ = 3.6 Hz), 5.45 (dd, 1H, H4α, $J_{H4\alpha,H3\alpha}$ = 3.4 Hz, $J_{H4\alpha,H5\alpha}$ = 1.2 Hz), 5.41-5.38 (m, 2H, H3α, H4β), 5.13 (m, 1H,

H2α), 5.07-5.05 (m, 2H, H2β, H3β), 4.71-4.67 (m, 1H, H1β), 4.47-4.44 (m, 1H, H5α), 4.14-4.06 (m, 4H, H6α, H6β), 3.98-3.94 (m, 2H, H5β, OHβ), 3.68-3.67 (m, 1H, OHα), 2.12, 2.07, 2.03, 1.97 (4 x s, 4 x 3H, 4 x CH_3α(Ac)), 2.14, 2.08, 2.02, 1.98 (4 x s, 4 x 3H, CH_3β(Ac)).

2,3,4,6-Tetra-*O*-acetyl-α-D-galactosylpyranosyl-trichloroacetimidate[246] (18) (αAc$_4$Gal-trichloroacetimidate)



To a solution of α/βAc$_4$Gal-OH **17** (25.29 g, 72.6 mmol) in dry dichloromethane (200 mL) was added trichloroacetonitrile (30.1 mL, 145.0 mmol) and the solution was cooled in an ice-bath before 1,8-diazabicyclo[5.4.0]undec-7-ene (2.17 mL, 14.5 mmol) was added. The reaction was stirred for 2 h in the ice-bath. The solvent was removed *in vacuo* and the product crystallized from ethyl acetate/cyclohexane.

Yield: 28.6 g (58.1 mmol, 82%), colorless crystals, $[\alpha]_D^{20}$ +100.60 (c = 0.50, $CHCl_3$), R_f = 0.40 (^{C}Hex/EtOAc 2:1).

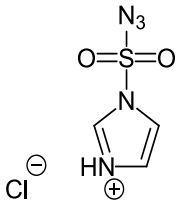
$C_{16}H_{20}Cl_3NO_{10}$ (M = 492.69 g/mol) [491.01].

ESI-MS (pos), m/z: 513.80 ($[M+Na]^+$, calc. 514.00).

^{1}H-NMR (400 MHz, $CDCl_3$), ***δ (ppm)***: 8.66 (s, 1H, NH), 6.59 (d, 1H, H1, $J_{H1,H2}$ = 3.4 Hz), 5.55 (dd, 1H, H4, $J_{H4,H3}$ = 3.1 Hz, $J_{H4,H5}$ = 1.2 Hz), 5.41 (dd, 1H, H3, $J_{H3,H4}$ = 3.1 Hz, $J_{H3,H2}$ = 10.9 Hz), 5.35 (dd, 1H, H2, $J_{H2,H1}$ = 3.5 Hz, $J_{H2,H3}$ = 10.8 Hz), 4.44-4.41 (m, 1H, H5), 4.15 (dd, 1H, $H6_a$, $J_{H6a,H6b}$ = 11.3 Hz, $J_{H6a,H5}$ = 6.6 Hz), 4.07 (dd, 1H, $H6_b$, $J_{H6b,H6a}$ = 11.3 Hz, $J_{H6b,H5}$ = 6.7 Hz), 2.15, 2.01, 2.00 (3 x s, 4 x 3H, 4 x CH_3-Ac).

^{13}C-NMR (100.6 MHz, $CDCl_3$), ***δ (ppm)***: 170.42, 170.22, 170.20, 170.09 (C=O(Ac)), 161.10 (C=NH), 93.71 (C1), 90.94 (CCl_3), 69.16 (C3'), 67.67 (C5'), 67.54 (C2'), 67.07 (C4'), 61.41 (C6), 20.80, 20.77, 20.74, 20.68 (CH_3-Ac).

Imidazole-1-sulfonyl azide hydrochloride[248] (19)



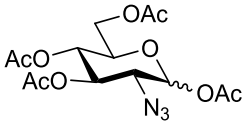
A suspension of sodium azide (35.18 g, 541 mmol) in acetonitrile (650 mL) was cooled in an ice-bath. Then sulfuryl chloride (72.63 g, 537 mmol) was added dropwise and stirred for 18 h. Then, imidazole (69.63 g, 1.02 mol) was added in portions and stirred for 4 h. The mixture was diluted with ethyl acetate (1.2 L) and washed with water (two times, 400 mL) and with saturated sodium bicarbonate solution (400 ml). The organic phase was dried over sodium sulfate, filtered and cooled in an ice-bath. The cooled mixture was treated with hydrochloric acid in methanol (65 mL acetyl chloride previously dropped into 225 mL of ice-cooled dry methanol). The mixture was stored overnight at -20°C. The precipitated product was filtered off and washed with ethyl acetate. The mother liquor must not be concentrated after filtration (risk of explosion).

Yield: 68.80 g (328.2 mmol, 61%), colorless crystals.

$C_3H_4ClN_5O_2S$ (M = 209.61 g/mol) [208.98].

^{1}H-NMR (400 MHz, D_2O), *δ (ppm)*: 7.65 (m, 1H, H4), 8.06 (m, 1H, H5), 9.42 (m, 1H, H2).

1,3,4,6-Tetra-*O*-acetyl-2-azido-2-deoxy-α/β-D-glucopyranose[248] (21)
(α/βAc$_4$GlcN$_3$)



To a suspension of potassium carbonate (80.72 g, 583 mmol) and (694 mg, 2.8 mmol) copper (II) sulfate pentahydrate in methanol (800 mL) were added first glucosamine hydrochloride **20** (60.88 g, 282 mmol) and then imidazole-1-sulfonyl azide hydrochloride **19** (64.17 g, 306 mmol). The suspension was stirred with a mechanical stirrer for 4 h. The solvent was removed *in vacuo* and co-evaporated with toluene. The residue was dissolved in pyridine (600 mL) and cooled in an ice bath. Then acetic anhydride (300 mL, 3.2 mol) was added and the reaction was stirred for 18 h at room temperature. The reaction volume was split in two halves and water (500 mL each) was added. Each aqueous phase was extracted with ethyl acetate (four times, 500 mL each). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The black residue was purified by flash column chromatography on silica (^{C}Hex/EtOAc 3:1 → 2:1).

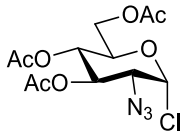
Yield: 80.3 g (α/β 1:2, 215.1 mmol, 77%), colorless oil, $[\alpha]_D^{20}$ +45.19 (c = 1.01, $CHCl_3$), R_f = 0.21 (^{C}Hex/EtOAc 3:1).

$C_{14}H_{19}N_3O_9$ (M = 373.32 g/mol) [373.11].

ESI-MS (pos), m/z: 396.07 ($[M+Na]^+$, calc. 396.10).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 6.28 (d, 1H, $H1^{\alpha}$, $J_{H1\alpha,H2\alpha}$ = 3.6 Hz), 5.54 (d, 1H, $H1^{\beta}$, $J_{H1\beta,H2\beta}$ = 8.6 Hz), 5.43 (dd, 1H, $H3^{\alpha}$, $J_{H3\alpha,H2\alpha}$ = 9.4 Hz, $J_{H3\alpha,H4\alpha}$ = 10.4 Hz), 5.12-5.10 (m, 3H, $H3^{\beta}$, $H4^{\alpha}$, $H4^{\beta}$), 4.31-4.26 (m, 2H, $H6^{\alpha}{}_a$, $H6^{\beta}{}_a$), 4.08-4.02 (m, 3H, $H5^{\alpha}$, $H6^{\alpha}{}_b$, $H6^{\beta}{}_b$), 3.79 (m, 1H, $H5^{\beta}$), 3.67-3.62 (m, 2H, $H2^{\alpha}$, $H2^{\beta}$), 2.17, 2.09, 2.06, 2.03 (4 x s, 4 x CH_3-Ac^{α}), 2.17, 2.08, 2.06, 2.01 (4 x s, 4 x CH_3-Ac^{β}).

3,4,6-Tri-*O*-acetyl-2-azido-2-deoxy-α-D-glucopyranosyl chloride[236,251,252] (22) (αAc$_3$GlcN$_3$-Cl)



A solution of α/βAc$_4$GlcN$_3$ **21** (80.2 g, 214 mmol) in dry chloroform (600 mL) was cooled in an ice-bath. Then titanium (IV) chloride (47.1 mL, 429.6 mmol) was added dropwise. The mixture turned yellow and a solid started to precipitate. The temperature was allowed to reach room temperature and was stirred for 45 min. Then the temperature was raised until reflux of the solvent. The solid dissolved and the solution turned brown. After 4.5 h stirring at reflux, the solution was diluted with dichloromethane and washed with saturated sodium bicarbonate solution (two times) and two times with brine. The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography on silica (^{C}Hex/EtOAc 3:1).

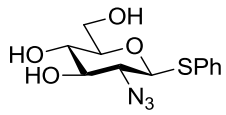
Yield: 69.1 g (197.4 mmol, 92 %), yellow oil, R_f = 0.52 (Tol/EtOAc 3:1).

$C_{12}H_{16}ClN_3O_7$ (M = 349.72 g/mol) [349.07]

ESI-MS (pos), m/z: 350.32 ($[M+H]^+$, calc. 350.08)

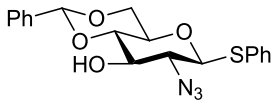
^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 6.11 (d, 1H, H1, $J_{H1,H2}$ = 3.8 Hz), 5.50 (dd, 1H, H3, $J_{H3,H2}$ = 10.2 Hz, $J_{H3,H4}$ = 9.7 Hz), 5.10 (m, 1H, H4), 4.33-4.29 (m, 2H, H5, $H6_a$), 4.10 (m, 1H, $H6_b$), 3.85 (dd, 1H, H2, $J_{H2,H3}$ = 10.3 Hz, $J_{H2,H1}$ = 3.8 Hz), 2.09, 2.08, 2.04 (3 x s, 3 x CH_3-Ac).

Phenyl 2-azido-2-deoxy-1-thio-β-D-glucopyranoside[253] (23)
(βGlcN$_3$-SPh)



Thiophenol (22.2 mL, 217 mmol) and potassium hydroxide (12.25 g, 217 mmol) was added to ethanol (400 mL). Then αAc$_3$GlcN$_3$-Cl **22** (69.18 g, 197.4 mmol) was dissolved in chloroform (400 mL) and added. The reaction was stirred for 3 h. The solution was diluted with dichloromethane and washed with saturated sodium bicarbonate (two times). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The residue was dissolved in 0.25 M sodium methoxide in methanol (400 mL) and stirred for 2 h. The solution was neutralized with acidic cation exchange resin (*Dowex 50WX8*). The resin was filtered off and the filtrate concentrated *in vacuo.* The residue was taken up with dichloromethane and washed with saturated sodium bicarbonate solution and with water (once each). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography on silica (Tol/EtOAc 8:7).
Yield: 28.3 g (95.2 mmol, 48%), colorless solid, R_f = 0.33 (100% EtOAc).
$C_{12}H_{15}N_3O_4S$ (M = 297.33 g/mol) [297.08]
ESI-MS (pos), m/z: 298.30 ([M+H]$^+$, calc. 298.09), 320.20 ([M+Na]$^+$, calc. 320.07).
^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.53-7.51 (m, 2H, H2-, H6-SPh), 7.32-7.30 (m, 3H, H3-, H4-, H5-SPh), 4.53 (d, 1H, H1, $J_{H1,H2}$ = 10.1 Hz), 3.91-3.81 (m, 2H, H6), 3.59 3.49 (2 x t, 2H, H3, H4), 3.34-3.29 (m, 2H, H2, H5).
^{13}C-NMR (100.6 MHz, $CDCl_3$), *δ (ppm)*: 133.00 (C2-, C6-SPh), 131.90 (C1-SPh), 129.39 (C3-, C5-SPh), 128.63 (C4-SPh), 87.02 (C1), 79.51 (C5), 77.18, 69.80 (C3, C4), 65.38 (C2), 62.10 (C6).

Phenyl 2-azido-2-deoxy-4,6-*O*-benzylidene-1-thio-β-D-glucopyranoside[254] (24)
(βBn-GlcN$_3$-SPh)



A solution of βGlcN$_3$-SPh **23** (28.20 g 94.8 mmol) and benzylidene dimethyl acetal (21.4 mL, 142.3 mmol) in dry acetonitrile (300 mL) set to pH 4.5 by addition of *para*-toluene sulfonic acid monohydrate (careful addition with the tip of a spatula). The solution was stirred at 50°C for 1.5 h. The reaction was neutralized with a few drops of *N,N*-diisopropylethylamine (wet pH paper). The reaction was concentrated *in vacuo* and the residue taken up with ethyl

acetate and washed with saturated sodium bicarbonate solution and with brine (once each). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The product was obtained by recrystallization from ethyl acetate/cyclohexane.

Yield: 27.0 g (70.0 mmol, 74%), pale yellow solid, $[\alpha]_D^{20}$ -72.13 (c = 0.50, $CHCl_3$), R_f = 0.60 (^{C}Hex/EtOAc).

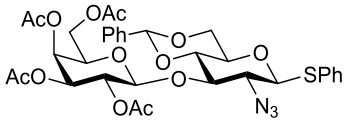
$C_{19}H_{19}N_3O_4S$ (M = 385.44 g/mol) [385.11].

ESI-MS (pos), m/z: 386.42 ($[M+H]^+$, calc. 386.12), 408.26 ($[M+Na]^+$, calc. 408.10).

^{1}H-NMR (400 MHz, $CDCl_3$), ***δ (ppm)***: 7.59-7.56 (m, 2H, H2-, H6-Bn), 7.48-7.45 (m, 2H, H2-, H6-SPh), 7.38-7.35 (m, 6H, H3-, H4-, H5-Bn, H3-, H4-, H5-SPh), 5.53 (s, 1H, CH-Bn), 4.54 (d, 1H, H1, $J_{H1,H2}$ = 10.2 Hz), 4.38 (dd, 1H, $H6_a$, $J_{H6a,H6b}$ = 10.4 Hz, $J_{H6a,H5}$ = 4.7 Hz), 3.79-3.74 (m, 2H, H3, $H6_b$), 3.50-3.44 (m, 2H, H4, H5), 3.35 (dd, 1H, H2, $J_{H2,H1}$ = 10.1 Hz, $J_{H2,H3}$ = 9.0 Hz), 2.79 (s_{br}, 1H, OH).

^{13}C-NMR (100.6 MHz, $CDCl_3$), ***δ (ppm)***: 133.83 (C2-, C6-SPh), 129.57, 129.29, 128.84, 128.54 (C3-, C4-, C5-Bn, C3-, C4-, C5-SPh), 102.10 (CH-Bn), 86.99 (C1), 80.36 (C5), 74.26 (C3), 70.43 (C4), 68.57 (C6), 65.35 (C2).

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[255] (25)
(βAc$_4$Gal-(1→3)-βBn-GlcN$_3$-SPh)



βBn-GlcN$_3$-SPh **24** (12.14 g, 31.5 mmol) and αAc$_4$Gal-trichloroacetimidate **18** (15.54 g, 31.5 mmol) were dissolved in dry diethyl ether (250 mL). A solution of trimethylsilyl trifluoromethansulfonate in diethyl ether (6.3 mL, 0.5 M, 3.15 mmol) was added dropwise. After 30 min of stirring, the reaction was neutralized by dropwise addition of triethylamine (wet pH paper). The mixture was diluted with diethyl ether (150 mL) and washed with saturated sodium bicarbonate solution, water and brine (once each). The organic phase was dried over sodium sulfate, filtered and the solvent removed *in vacuo*. The crude product was purified by flash column chromatography on silica (^{C}Hex/EtOAc 2:1).

Yield: 19.22 (26.9 mmol, 85%), colorless, amorphous solid, $[\alpha]_D^{20}$ -26.11 (c = 1.01, MeOH), R_f = 0.52 (Tol/EtOAc 3:1).

$C_{33}H_{37}N_3O_{13}S$ (M = 715.72 g/mol) [715.20].

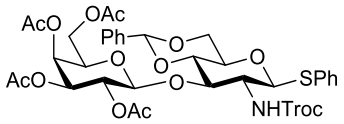
ESI-MS (pos), m/z: 733.07 ($[M+NH_4]^+$, calc. 733.24), 738.20 ($[M+Na]^+$, calc. 738.19).

HR-ESI-MS (pos), m/z: 733.2387 ($[M+NH_4]^+$, calc. 733.2391).

^{1}H NMR (400 MHz, $CDCl_3$, 25°C, internal $CDCl_3$ δ(H)=7.26 ppm): δ= 7.59-7.56 (m, 2H, H-2-, H-6-Bn), 7.45-7.38 (m, 2H, H-2-, H-6-SPh), 7.38-7.35 (m, 6H, H-3-, H-4-, H-5-Bn, H-3-, H-4-, H-5-SPh), 5.54 (s, 1H, CH-Bn), 5.29 (d, ^{3}J(H-4',H-3') = 3.4 Hz, 1H, H-4'), 5.22 (dd, ^{3}J(H-2',H-1') = 8.0 Hz, ^{3}J(H-2',H-3') = 10.4 Hz, 1H, H-2'), 4.95 (dd, ^{3}J(H-3',H-2') = 10.4 Hz, ^{3}J(H-3',H-4') = 3.4 Hz, 1H, H-3'), 4.77 (d, ^{3}J(H-1',H-2') = 8.0 Hz, 1H, H-1'), 4.50 (d, ^{3}J(H-1,H-2) = 10.1 Hz, 1H, H-1), 4.37 (dd, ^{2}J(H-6_a,H-6_b) = 10.6 Hz, ^{3}J (H-6_a,H-5) = 4.9 Hz 1H, H-6_a), 4.03 (dd, ^{2}J(H-$6'_b$,H-$6'_a$) = 7.7 Hz, ^{3}J(H-$6'_b$,H-5') = 11.1 Hz 1H, H-$6'_b$), 3.86-3.59 (m, 5H, H-6_b, H-6_b', H-5', H-4, H-3), 3.44-3.33 (m, 2H, H-5, H-2), 2.10, 2.05, 1.96, 1.91 (4 x s, 4 x 3H, 4 x CH_3-Ac).

^{13}C NMR (150.9 MHz, $CDCl_3$, 30°C, internal $CDCl_3$ δ(C)=77.16 ppm): δ = 170.36, 170.26, 169.58 (C=O-(Ac)), 136.92 (C-1-SPh), 134.20 (C-2-, C-6-SPh), 130.11 (C-1-Bn), 129.42, 129.34, 129.10 (C-3-, C-4-, C-5-Bn), 128.47 (C-3-, C-4-, C-5-SPh), 126.07 (C-2-, C-6-Bn), 101.44 (CH-Bn), 101.39 (C-1'), 87.12 (C-1), 81.48 (C-3), 79.38 (C-4), 71.09 (C-3'), 70.92 (C-5'), 70.62 (C-5), 69.53 (C-2'), 68.49 (C-6), 66.92 (C-4'), 64.53 (C-2), 61.05 (C-6'), 20.85, 20.75, 20.70, 20.67 (CH_3-(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-trichlorooxycarbonyl)-β-D-glucopyranoside (26)
(βAc_4Gal-(1→3)-βBn-GlcNHTroc-SPh)



βAc_4Gal-(1→3)-βBn-GlcN_3-SPh **25** (19.20 g, 26.9 mmol) was dissolved in mixture of 1,4-dioxane and acetic acid (220 mL, 10:1) and cooled in an ice bath. Zinc dust was activated by treatment with 1 N hydrochloric acid for a few minutes followed by washing with water, methanol and finally diethyl ether. Zinc dust (10.60 g, 161.4 mmol) was added to the solution and the reaction was stirred overnight, while the temperature rose to room temperature. The zinc dust was filtered off, washed with ethyl acetate (400 mL) and the filtrate was further diluted with ethyl acetate (400 mL). The organic phase was washed with saturated sodium bicarbonate solution (two times), water and brine (500 mL each). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The residue was dissolved in 1,4-dioxane (200 mL) and sodium bicarbonate (5.20 g, 61.8 mmol) dissolved in water was added (100 mL). The stirred suspension was cooled in an ice bath and a solution of 2,2,2-trichloroethyl chloroformate (6.84 g, 32.3 mmol) in 1,4-dioxane (40 mL) was added dropwise. The ice bath was removed and the reaction was stirred further for 3 h. The 1,4-dioxane was removed *in vacuo* and the aqueous phase was extracted with ethyl

acetate (500 mL). The ethyl acetate phase was washed with water (two times, 150 mL each) and with brine (once, 150 mL). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The product was purified by flash column chromatography on silica (toluene/EtOAc 3:1).

Yield: 16.81 (19.4 mmol, 72%), colorless, amorphous solid, $[\alpha]_D^{20}$ -10.0° (c = 1.01, $CHCl_3$); R_f = 0.37 (Tol/EtOAc 3:1).

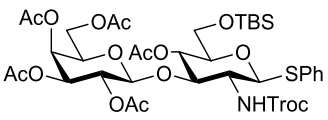
$C_{36}H_{40}Cl_3NO_{15}S$ (M = 865.12 g/mol) [863.12].

HR-ESI-MS (pos), m/z: 864.1261 ($[M+H]^+$, calc. 864.1257), 881.1525 ($[M+NH_4]^+$, calc. 881.1522), 886.1073 ($[M+Na]^+$, calc. 886.1076).

^{1}H NMR (400 MHz, $CDCl_3$, 25°C, internal $CDCl_3$ *δ*(C)=7.26 ppm): ***δ*=** 7.47-7.45 (m, 4H, H-2-, H-6-SPh, H-2-, H-6-Bn), 7.37-7.36 (m, 3H, H-3-, H-4-, H-5-Bn), 7.32-7.31 (m, 3H, H-3-, H-4-, H-5-SPh), 5.53 (s, 1H, CH-(Bn)), 5.32 (d, ^3J(NH,H-2) = 7.9 Hz, 1H, NH-Troc), 5.29 (d, ^3J(H-4',H-3') = 3.2 Hz, 1H, H-4'), 5.16 (m, 2H, H-2', H-1), 4.92 (dd, ^3J(H-3',H-2') = 10.4 Hz, ^3J(H-3',H-4') = 3.2 Hz, 1H, H-3'), 4.82 (d, $^2J(CH_{2a},CH_{2b})$ = 12.1 Hz, 1H, CH_{2a}-(Troc)), 4.72-4.68 (m, 2H, H-1', CH_{2b}-(Troc)), 4.36 (dd, $^3J(H\text{-}6_a,H\text{-}6_b)$ = 10.6 Hz, $^3J(H\text{-}6_a,H\text{-}5)$ = 4.8 Hz, 1H, H-6_a), 4.29 (t, ^3J(H-3,H-2) = ^3J(H-3,H-4) = 8.9 Hz, 1H, H-3), 4.04 (m, 1H, H-6_a'), 3.82-3.77 (m, 2H, H-6_b, H-6_b'), 3.67 (t, ^3J(H-4,H-3) = ^3J(H-4,H-5) = 9.2 Hz, 1H, H-4), 3.62 (m, 1H, H-5'), 3.56-3.52 (m, 1H, H-5), 3.36 (m, 1H, H-2), 2.10, 1.99, 1.95, 1.94 (4 x s, 4 x 3H, CH_3-(Ac)).

^{13}C NMR (150.9 MHz, $CDCl_3$, 30°C, internal $CDCl_3$ *δ*(H)=77.16 ppm): ***δ* =** 170.34, 170.23, 169.50 (C=O-(Ac)), 153.76 (C=O-(Troc)), 137.08 (C-1-Bn), 133.03 (C-2-, C-6-SPh), 131.84 (C-1-SPh), 129.41, 129.27 (C-3-, C-4-, C-5-Bn), 128.54, 128.45 (C-3-, C-4-, C-5-SPh), 126.15 (C-2-, C-6-Bn), 101.48 (CH-(Bn)), 100.87 (C-1'), 95.41 (CCl_3-(Troc)), 86.21 (C-1), 79.96 (C-4), 79.18 (C-3), 74.61 (CH_2-(Troc)), 71.04 (C-3'), 70.66 (C-5'), 70.52 (C-5), 69.41 (C-2'), 68.67 (C-6), 66.87 (C-4'), 60.96 (C-6'), 56.53 (C-2), 20.80, 20.75, 20.71, 20.67 (CH_3-(Ac)).

Phenyl 3-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-4-*O*-acetyl-6-*O*-*tert*-butyldimethysilyl-2-deoxy-1-thio-2-*N*-(2,2,2-trichloroethoxycarbonyl)-β-D-glucopyranoside (27)
(βAc$_4$Gal-(1→3)-βAc-TBS-GlcNHTroc-SPh)



βAc$_4$Gal-(1→3)-βBn-GlcNHTroc-SPh **26** (8.08 g, 9.3 mmol) was dissolved in 80% acetic acid (80 mL) and stirred at 55°C for 7 h. The reaction was diluted and co-evaporated with toluene.

The residue was dissolved in dry *N,N*-dimethylformamide (80 mL). Imidazole (1.58 g, 23.3 mmol) and *tert*-butyldimethylsilyl chloride (1.67 g, 11.1 mmol) were added. After stirring for 3 h the reaction was diluted with ethyl acetate (400 mL) and washed with water (two times, 400 mL each) and with brine (once, 200 mL). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The residue was dissolved in pyridine/acetic anhydride (150 mL, 2:1) and stirred for 16 h. Then the solvent was removed by co-evaporation with toluene *in vacuo*. The crude was purified by column chromatography on silica (^{C}Hex/EtOAc 2:1).

Yield: 7.82 g (8.4 mmol, 90%), colorless, amorphous solid, $[\alpha]_D^{20}$ +4.3° (c = 0.99, MeOH); R_f = 0.34 (^{C}Hex/EtOAc 2:1).

$C_{37}H_{52}Cl_3NO_{16}SSi$ (M = 933,32 g/mol) [931.18].

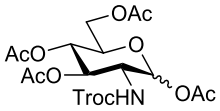
HR-ESI-MS (pos), m/z: 932.1924 ($[M+H]^+$, calc. 932.1914), 949.2182 ($[M+NH_4]^+$, calc. 949.2180), 954.1733 ($[M+Na]^+$, calc. 954.1734).

^{1}H NMR (400 MHz, $CDCl_3$, 25°C, internal $CDCl_3$ δ(H)=7.26 ppm): δ= 7.52-7.49 (m, 2H, H-2-, H-6-SPh), 7.29-7.28 (m, 3H, H-3-, H-4-, H-5-SPh), 5.33 (d, ^{3}J(H-4',H-3') = 3.4 Hz, 1H, H-4'), 5.14 (d, ^{3}J(NH,H-2) = 8.4 Hz, 1H, NH-Troc), 5.06 (dd, ^{3}J(H-2',H-1') = 7.8 Hz, ^{3}J(H-2',H-3') = 10.3 Hz, 1H, H-2'), 4.98 (d, ^{3}J(H-1,H-2) = 10.4 Hz, 1H, H-1), 4.92 (dd, ^{3}J(H-3',H-2') = 10.3 Hz, ^{3}J(H-3',H-4') = 3.4 Hz, 1H, H-3'), 4.86-4.80 (m, 2H, H-4, CH_{2a}-(Troc)), 4.73 (d, $^{2}J(CH_{2a},CH_{2b})$ = 12.1 Hz, 1H, CH_{2b}-(Troc)), 4.62 (d, ^{3}J(H-1',H-2') = 7.8 Hz, 1H, H-1'), 4.19 (t, ^{3}J(H-3,H-2) = ^{3}J(H-3,H-4) = 9.4 Hz, 1H, H-3), 4.13-4.04 (m, 2H, H-6'$_{ab}$), 3.84 (t, ^{3}J(H-5',H-6'$_{ab}$) = 6.8 Hz, 1H, H-5'), 3.70-3.68 (m, 2H, H-6$_{ab}$), 3.59-3.55 (m, 1H, H-5), 3.44-3.37 (m, 1H, H-2), 2.14, 2.06, 2.03, 1.95 (s, 5 x 3H, CH_3-(Ac)), 0.89 (s, 9H, *t*Bu-TBS), 0.06, 0.04 (s, 2 x 3H, CH_3-(TBS)).

^{13}C NMR (150.9 MHz, $CDCl_3$, 30°C, internal $CDCl_3$ δ(C)=77.16 ppm): δ = 170.52, 170.29, 170.18, 169.41, 169.32 (5 x C=O-(Ac)), 153.82 (C=O-(Troc)), 132.97 (C-1-SPh), 132.16 (C-2-, C-6-SPh), 129.16 (C-3-, C-5-SPh), 128.03 (C-4), 100.96 (C-1'), 95.43 (CCl_3-(Troc)), 85.92 (CH-Bn), 79.39 (C-5), 79.03 (C-3), 74.62 (CH_2-(Troc)), 71.05 (C-3'), 70.74 (C-5'), 69.41 (C-4), 69.23 (C-2'), 66.99 (C-4'), 63.09 (C-6), 61.21 (C-6'), 56.96 (C-2), 26.00 (*t*Bu-TBS), 21.00, 20.87, 20.82, 20.78, 20.66 (5 x CH_3-(Ac)), 18.47 (C_q-TBS), -5.15, -5.31 (CH_3-(TBS)).

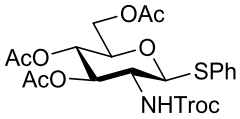
7.1.5 Syntheses of the type-2 *N*-acetyllactosamine glycosyl donor

1,3,4,6-Tetra-*O*-acetyl-2-*N*-(2,2,2-trichloroethoxycarbonyl)-2-deoxy-α/β-D-glucopyranoside (28)
(α/βAc$_4$-GlcNTroc)



To a solution of *N*-glucosamine hydrochloride **20** (10.11 g, 46.8 mmol) in water (100 mL), was added sodium bicarbonate (8.79 g, 104.7 mmol). A solution of 2,2,2-trichloroethyl chloroformate (8.1 mL, 59.9 mmol) in 1,4-dioxane (50 mL) was added dropwise to the stirred solution. After 1.5 h the 1,4-dioxane was removed *in vacuo*. A white solid precipitated from the aqueous solution. The flask was stored for 5 h at 4°C. The precipitate was filtered off and washed with cold water and then with dichloromethane. The solid was dissolved in pyridine/acetic anhydride (120 mL, 2:1) and 4-(dimethylamino)pyridine (61 mg, 0.5 mmol) was added. After 4 h of stirring the reaction was co-evaporated es with toluene. The crude was purified by flash column chromatography on silica (^{C}Hex/EtOAc 2:1).
Yield: 19.6 g (α/β 8.3:1, 37.5 mmol, 80%), colorless solid, $[\alpha]_D^{20}$ +67.13 (c = 0.50, $CHCl_3$), R_f = 0.22 (^{C}Hex/EtOAc).
$C_{17}H_{22}ClNO_{11}$ (M = 522.72 g/mol) [521.03].
^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 6.23 (d, 1H, H1$^\alpha$, $J_{H1\alpha,H2\alpha}$ = 3.6 Hz), 5.73 (d, 1H, H1$^\beta$, 8.8 Hz), 5.29-5.15 (m, 5H, NH$^\alpha$, H3$^\alpha$, H4$^\alpha$, H3$^\beta$, H4$^\beta$), 4.81 (d, 1H, $CH_{2a}{}^\alpha$(Troc), $J_{CH2a,CH2b}$ = 12.0 Hz), 4.76-4.71 (m, 2H, $CH_{2ab}{}^\beta$(Troc)), 4.62 (d, 1H, $CH_{2b}{}^\alpha$(Troc), $J_{CH2b,CH2a}$ = 12.0 Hz), 4.29-3.99 (m, 6H, H2$^\alpha$, H6$^\alpha$, H6$^\beta$, H5$^\alpha$), 3.95-3.93 (m, 1H, H2$^\beta$), 3.84-3.80 (m, 1H, H5$^\beta$), 2.19, 2.08, 2.04, 2.03 (4 x s, 4 x 3H, 4 x CH_3(Ac)).
^{13}C-NMR (100.6 MHz, $CDCl_3$), *δ (ppm)*: 171.39, 170.76, 169.29, 168.71 (C=O(Ac)), 154.17 (C=O(Troc)), 90.57 (CCl_3(Troc)), 74.83 (C1), 70.54, 69.89, 67.69 (C3, C4, C5), 61.62 (C6), 53.38 (C2), 21.05, 20.82, 20.79, 20.69 (CH_3(Ac)).

Phenyl 3,4,6-tetra-*O*-acetyl-2-deoxy-2-*N*-(2,2,2-trichloroethoxycarbonyl)-1-thio-β-D-glucopyranoside[262] (29)
(βAc$_3$GlcNTroc-SPh)



Thiophenol (12.0 g, 109.0 mmol) was added to a solution of αβAc$_4$GlcNTroc **28** (19.0 g, 36.3 mmol) in dry dichloromethane (200 mL) under argon atmosphere. Boron trifluoride diethyl

etherate (48%, 7.4 mL, 60.0 mmol) was added and the reaction stirred for 3 d. The reaction was neutralized with triethylamine and the solvent removed *in vacuo*. The residue was purified by flash column chromatography on silica (^{C}Hex/EtOAc 2:1).

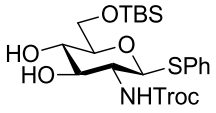
Yield: 19.3 g (33.6 mmol, 93%), colorless solid, $[\alpha]_D^{20}$ -3.33 (c = 1.00, $CHCl_3$), R_f = 0.34 (^{C}Hex/EtOAc 2:1).

$C_{21}H_{24}Cl_3NO_9S$ (M = 572.84 g/mol) [571.02].

ESI-MS (pos), m/z: 462.73 ([M-SPh]$^+$, calc. 462.01), 571.67 ([M+H]$^+$, calc. 572.03), 593.93 ([M+Na]$^+$, calc. 594.01).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.52-7.49 (m, 2H, H2-, H6-SPh), 7.31-7.29 (m, 3H, H3-, H4-, H5-SPh), 5.43 (d, 1H, NH, $J_{NH,H2}$ = 9.1 Hz), 5.28 (t, 1H, H3, $J_{H3,H2}$ = $J_{H3,H4}$ = 9.5 Hz), 5.01 (t, 1H, H4, $J_{H4,H3}$ = $J_{H4,H5}$ = 9.6 Hz), 4.86 (d, 1H, H1, $J_{H1,H2}$ = 10.4 Hz), 4.79 (d, 1H, CH_{2a}(Troc), $J_{CH2a,CH2b}$ = 12.1 Hz), 4.71 (d, 1H, CH_{2b}(Troc), $J_{CH2b,CH2a}$ = 12.1 Hz), 4.21-4.13 (m, 2H, H6), 3.74-3.69 (m, 2H, H2, H5), 2.07, 2.00, 1.99 (3 x s, 3 x 3H, 3 x CH_3(Ac)).

Phenyl 6-*O*-*tert*-butyldimethylsilyl-2-deoxy-2-*N*-(2,2,2-trichloroethoxycarbonyl)-1-thio-β-D-glucopyranoside (30)
(βTBS-GlcNTroc-SPh)



Acetyl chloride (20 mL) was dropped into dry methanol (200 mL), cooled in an ice bath. After 15 min, a solution of βAc_3GlcNTroc-SPh **29** (19.24 g, 33.6 mmol) in dichloromethane (50 mL) was added dropwise. After complete addition the ice bath was removed and the reaction stirred for 6 h. Then the solvent was removed *in vacuo* and the residue taken up in ethyl acetate (400 mL). The ethyl acetate phase was washed with water, saturated sodium bicarbonate, water again and brine (150 mL each). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The gel-like residue was dried at reduced pressure (10^{-2}-10^{-1} mbar), until a white, amorphous solid was left. The solid was dissolved in 160 mL of dry *N,N*-dimethylformamide under argon atmosphere. Imidazole (5.72 g, 84.0 mmol) and *tert*-butyldimethylsilyl chloride (5.57 g, 37.0 mmol) were added. The reaction was stirred for 3 h. Then it was diluted with ethyl acetate (500 mL) and washed with water (two times, 150 mL) and with brine (once, 150 mL each). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The crude was purified by flash column chromatography on silica (DCM→DCM/MeOH 15:1).

Yield: 15.03 (26.8 mmol, 80%), pale yellow, amorphous solid, $[\alpha]_D^{20}$ +20.50 (c = 1.01, $CHCl_3$), R_f = 0.15 (Tol/EtOAc 3:1), 0.52 (DCM/MeOH 10:1).

$C_{21}H_{32}Cl_3NO_6SSi$ (M = 560.99 g/mol) [559.08].

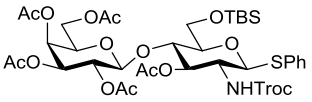
ESI-MS (pos), m/z: 559.80 ($[M+H]^+$, calc. 560.08).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.49-7.47 (m, 2H, H2-, H6-SPh), 7.28-7.26 (m, 3H, H3-, H4-, H5-SPh), 5.63 (d, 1H, NH, $J_{NH,H2}$ = 8.0 Hz), 4.81-4.78 (m, 2H, H1, CH_{2a}(Troc)), 4.70 (d, 1H, CH_{2b}(Troc), $J_{CH2a,CH2b}$ = 12.0 Hz), 3.90 (dd, 1H, H6, $J_{H6a,H6b}$ = 4.7 Hz, $J_{H6a,H5}$ = 1.6 Hz), 3.74 (t, 1H, H3, $J_{H3,H4}$ = $J_{H3,H2}$ = 9.2 Hz), 3.54 (t, 1H, H4, $J_{H4,H3}$ = $J_{H4,H5}$ = 9.4 Hz), 3.47-3.40 (m, 2H, H2, H5), 0.91 (s, 9H, *t*Bu(TBS)), 0.10, 0.09 (2 x s, 2 x 3H, 2 x Me(TBS).

^{13}C-NMR (100.6 MHz, $CDCl_3$), *δ (ppm)*: 154.97 (C=O(Troc)), 132.33 (C2-, C6-SPh), 129.08 (C3-, C4-, C5-SPh), 127.97 (C1-SPh), 95.55 (CCl_3(Troc)), 86.45 (C1), 78.74 (C5), 76.02 (C3), 74.85 (CH_2(Troc)), 72.97 (C6), 64.48 (C2), 56.58 (C2), 26.01 (*t*Bu(TBS)), 18.39 (C_{quart}(TBS)), -5.26, -5.29 (Me(TBS)).

Phenyl 3-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-6-*O*-*tert*-butyldimethylsilyl-2-deoxy-1-thio-2-*N*-(2,2,2-trichloroethoxycarbonyl)-β-D-glucopyranoside (31)

(βAc$_4$Gal-(1→4)-βAc-TBS-GlcNTroc-SPh)



βTBS-GlcNTroc-SPh **30** (14.98 g, 26.70 mmol) and αAc$_4$Gal-trichloroacetimidate **18** (13.15 g, 26.70 mmol) were dissolved in dry dichloromethane (200 mL). Activated molecular sieves (15 g, 4Å) were added under argon atmosphere. The suspension was stirred for 1 h at room temperature and then cooled to -50°C. Trimethylsilyl trifluoromethanesulfonate (482 µL, 593 mg, 2.67 mmol) in a suspension with dry dichloromethane (5 mL) was added dropwise *via* a syringe. The reaction was stirred at -50°C for 1.5 h and then quenched with triethylamine (5 mL). The molecular sieves were filtered off over *Celite*® and washed with dichloromethane (200 mL). The filtrate was washed with saturated sodium bicarbonate, water and brine (150 mL each). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The residue was dissolved in pyridine/acetic anhydride (225 mL, 2:1) and stirred for 18 h. The reaction was co-evaporated with toluene *in vacuo* and the residue purified by flash column chromatography on silica (^{C}Hex:EtOAc 2.5:1).

Yield: 19.15 g (20.5 mmol, 77%), colorless, amorphous solid, $[\alpha]_D^{20}$ -12.05 (c = 1.00, $CHCl_3$), R_f = 0.34 (^{C}Hex/EtOAc 2:1).

$C_{37}H_{52}Cl_3NO_{16}SSi$ (M = 933.32 g/mol) [931.18].

ESI-MS (pos), m/z: 932.80 ($[M+H]^+$, calc.932.19); 948.67 ($[M+NH_4]^+$, calc. 949.22); 1888.33 ($[2M+Na]^+$, calc. 1887.36).

HR-ESI-MS (pos), m/z: 932.1922 ($[M+H]^+$, calc. 932.1920), 822.1728 ($[M-SPH]^+$, calc. 822.1730).

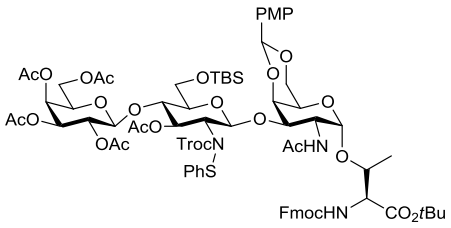
^{1}H NMR (400 MHz, 25°C, internal $CDCl_3$, *δ*(H)=7.26 ppm): ***δ*=** 7.50-7.47 (m, 2H, H-2-, H-6-SPh), 7.29-7.27 (m, 3H, H-3-, H-4-, H-5-SPh), 5.37 (d, ^3J(NH,H-2) = 9.8 Hz, 1H, NH-Ac), 5.33 (dd, ^3J(H-4',H-3') = 3.4 Hz, ^3J(H-4',H-5') = 0.9 Hz, 1H, H-4'), 5.12-5.03 (m, 2H, H-3, H-2'), 4.91 (dd, ^3J(H3',H2') = 10.4 Hz, ^3J(H-3',H-4') = 3.6 Hz, 1H, H-3'), 4.79-4.70 (m, 3H, H-1', CH_2-(Troc)), 4.66 (d, ^3J(H-1,H-2) = 9.5 Hz, 1H, H-1), 4.09 (d, ^2J(H-6_a',H-6_b') = 6.7 Hz, 2H, H-6_{ab}'), 3.93-3.79 (m, 4H, H-4, H-5', H-6_{ab}), 3.70 (q, ^3J(H-2,H-1) = ^3J(H-2,H-3) = ^3J(H-2,NH) = 9.8 Hz, 1H, H-2), 3.37-3.33 (m, 1H, H-5), 2.10, 2.04, 2.02, 2.00, 1.95 (5 x s, 15H, CH_3-(Ac)), 0.93 (s, 9H, *t*Bu-(TBS)), 0.13, 0.10 (2 x s, 6H, Me-(TBS)).

^{13}C NMR (100.6 MHz, 25°C, internal $CDCl_3$, *δ*(C)=77.16 ppm): ***δ*=** 170.78, 170.42, 170.27, 170.20, 168.84 (C=O-(AcO)), 154.27 (C=O-(Troc)), 132.80 (C-1-SPh), 128.99, 128.03 (C-2-, C-3-, C-4-, C-5-, C-6-SPh), 100.36 (C-1'), 95.62 (CCl_3-(Troc)), 87.03 (C-1), 79.68 (C-5), 74.64 (CH_2-(Troc)), 74.06 (C-4), 73.75 (C-3), 71.24 (C-3'), 70.78 (C-5'), 69.37 (C-2'), 67.01 (C-4'), 61.20 (C-6), 55.19 (C-2), 26.01 (*t*Bu-(TBS)), 20.92, 20.88, 20.75, 20.70, 20.66 (CH_3-(Ac)), 18.38 (C_{quart}-(TBS)), -4.87, -5.16 (CH_3-(TBS)).

7.1.6 Synthesis of the extended type-2 core 3 amino acid

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-{2-acetamido-2-deoxy-4,6-_O_-_para_-methoxybenzylidene-3-_O_-[3-_O_-acetyl-6-_O_-_tert_-butyldimethylsilyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-2-_N_-phenylsulfide-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-α-D-galactosylpyranosyl}-L-threonine-_tert_-butylester (32)

(Fmoc-Thr(βAc_4Gal-(1→4)-βAc-TBS-GlcN(SPh)Troc(1→3)-αPMP-GalNAc)-O*t*Bu)



Fmoc-Thr(αPMP-GalNAc)-O*t*Bu **12** (1.50 g, 2.09 mmol) and βAc_4Gal-(1→4)-βGlcNTroc-SPh **31** (2.54 g, 2.72 mmol) were dissolved in dry dichloromethane (20 mL). Activated molecular sieves (2.0 g, 4Å) were added under argon atmosphere. The suspension was stirred for 1 h and then was cooled in an ice bath. Then *N*-iodosuccinimide (1.18 g, 5.23 mmol) was added followed by the dropwise addition of a suspension of trimethylsilyl trifluoromethanesulfonic

acid (55.5 µmol 94.1 mg, 0.63 mmol) in dry dichloromethane (300 µL) *via* a syringe. The reaction was stirred for 4.5 h and then diluted with dichloromethane (20 mL). The molecular sieves were filtered off and washed with dichloromethane. More dichloromethane was added to give a total volume of 150 mL. The dichloromethane phase was washed with 0.5 M sodium thiosulfate solution, saturated sodium bicarbonate, water and brine (60 mL each). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography on silica (Tol/EtOAc 4:1).

Yield: 2.41 g (1.46 mmol, 70%), colorless, amorphous solid, $[\alpha]_D^{20}$ +51.73 (c = 0.52, $CHCl_3$), R_f = 0.53 (Tol/EtOAc 1:1).

$C_{76}H_{96}Cl_3N_3O_{27}SSi$ (M = 1650.09 g/mol) [1647.48].

ESI-MS (pos), m/z: 1666.00 ($[M+NH_4]^+$, calc. 1665.51).

HR-ESI-MS (pos), m/z: 1648.4844 ($[M+H]^+$, calc. 1648.4865).

^{1}H NMR (600 MHz, 25°C, [D_4]MeOD, internal MeOH, *δ*(H)=4.87 ppm): ***δ*=** 7.86 (m, 4H, H-4-, H-5-Fmoc, N-H-Ac, N-H-Ac*), 7.71-7.69 (m, 2H, H-1-, H-8-Fmoc), 7.44-7.38 (m, 4H, H-2-, H-7-Fmoc, H-2-, H-6-PMP), 7.35-7.30 (m, 4H, H-3-, H-6-Fmoc, 2 H_{arom}-SPh), 7.25-7.19 (m, 3H, 3 H_{arom}-SPh), 6.72-6.89 (m, 2H, H-3-, H-5-PMP), 6.84-6.81 (m, 2H, H-3-, H-5-PMP*), 5.53 (s, 1H, CH-(PMP)), 5.42-5.29 (m, 3H, H-4'', H-4''*, CH_{2a}*-(Troc)), 5.24 (t, ^3J(H-3',H-2') = 9.8 Hz, 1H, H-3'), 5.09-5.07 (m, 1H, H-3'*), 5.00-4.95 (m, 2H, H-2'', H-3''), 4.93-4.86 (m, 3H, H-1''*, H-2''*, H-3''*), 4.82-4.68 (m, 5H, H-1''*, H-1, H-1*, H-1', H-1'*), 4.63-4.54 (m, 5H, H-2*, H-2'*, CH_{2ab}-(Fmoc), CH_{2b}-(Troc), CH^*_{2b}-(Troc)), 4.53-4.41 (m, 5H, H-2, H-2', H-4, H-4*, T^{β}), 4.32-4.29 (m, 1H, H-9-Fmoc), 4.18-4.03 (m, 5H, H-6_{ab}, H-6_{ab}'', T^{α}), 4.00-3.85 (m, 6H, H-3, H-4', H-5'', H-5''*, H-$6'_{ab}$), 3.81-3.76 (m, 5H, CH_3-(PMP), H-5, T^{α}*), 3.49-3.45 (m, 1H, H-5'), 2.16-1.91 (m, CH_3-(Ac)), 1.45 (s, 9H, *t*Bu-(Thr)), 1.23-1.22 (m, 3H, T^{γ}), 0.91 (s, 9H, *t*Bu-(TBS)), 0.18-0.13 (m, 12H, 2 Me-(TBS)).

^{13}C NMR (150.9 MHz, 25°C, [D_4]MeOD, internal MeOH, *δ*(C)=49.00 ppm): ***δ*=** 172.02, 171.94, 171.91, 171.88, 171.83, 171.73, 171.39, 171.36, 171.03, 170.91, 170.89, 170.83, 170.82, 170.78, 170.62, 170.46 (C=O-(Ac), C=O-(*t*Bu)), 161.50 (C-1-PMP), 159.13 (C=O-(Fmoc)), 156.62 (C=O-(Troc)), 145.27, 145.15 (C-1_a-, C-8_a-Fmoc), 142.71 (C-4_a-, C-5_a-Fmoc), 138.31 (C-1-SPh), 131.79, 131.66 (C-4-PMP), 130.20 (C_{arom}-SPh), 128.84 (C-2-, C-7-Fmoc), 128.22 (C-3-, C-6-Fmoc), 126.03 (C-1-, C-8-Fmoc), 121.06, 121.04 (C-4-, C-5-Fmoc), 114.36 (C-3-, C-5-PMP), 103.90 (C-1', C-1'*), 101.95 (CH-PMP), 101.37 (C-1'', C-1''*), 100.91 (C-1, C-1*), 96.31, 95.97 (C_{quart}-(Troc)), 83.54 (C_{quart}-(*t*Bu)), 77.59 (C-3, CH_{2ab}-(Troc)), 77.23 (C-4), 77.14 (CH_{2ab}-(Troc*)), 76.44 (C-5'), 76.18 (C-4'), 75.50 (T^{γ}), 72.54 (C-3''), 72.04, 71.88 (C-5'', C-5''*), 71.55 (C-3'), 70.60 (C-2'', C-2''*), 70.23 (C-6_{ab}), 68.68, 68.56 (C-4'', C-4''*), 67.45 (CH_{2ab}-(Fmoc)), 64.80, 64.65 (C-5), 63.55 (C-2'), 62.82, 62.50, 62.35 (C-6', C-6''), 60.75, 60.72 (T^{α}, T^{α}*), 55.69 (CH_3-PMP), 48.95 (C-$2^{\#}$), 48.68 (C-9-(Fmoc)$^{\#}$), 28.40,

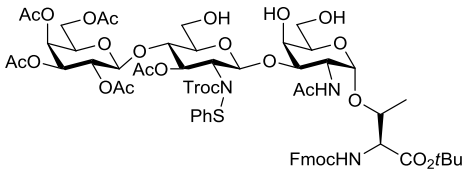
28.36 (*t*Bu-(Thr), *t*Bu-(Thr)*), 26.57 (*t*Bu-(TBS)), 24.44, 24.00, 21.67-20.45 (CH_3-(Ac)), 20.05 (*t*Bu), 19.21, 19.13 (T^{γ}).

(*, rotational-isomers, if identifiable)

($^{\#}$,congruent with solvent peak, shift from HSQC)

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-{2-acetamido-2-deoxy-3-_O_-[3-_O_-acetyl-2-deoxy-2-(2,2,2-trichloroethoxycarbonyl)-2-_N_-phenylsulfide-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-α-D-galactosylpyranosyl}-L-threonine-_tert_-butylester (33)

(Fmoc-Thr(βAc$_4$Gal-(1→4)-βAcGlcN(SPh)Troc(1→3)-GalNAc)-O*t*Bu)



Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc-TBS-GlcN(SPh)Troc(1→3)-αPMP-GalNAc)-O*t*Bu) **32** (2.35 g, 1.42 mmol) was dissolved in 80 vol% acetic acid (25 mL) and stirred for 18 h at 50°C. The solvent was removed *in vacuo* and the residue was co-evaporated with toluene. The crude was purified by flash column chromatography on silica (EtOAc→EtOAc/MeOH 25:1).

Yield: 1.67 g (1.17 mmol, 82%), colorless, amorphous solid, $[\alpha]_D^{20}$ +26.08 (c = 0.51, $CHCl_3$), R_f = 0.43 (EtOAc/MeOH 25:1).

$C_{62}H_{76}Cl_3N_3O_{26}S$ (M = 1417.70 g/mol) [1415.35].

ESI-MS (pos), m/z: 1416.53 ($[M+H]^+$, calc. 1416.36), 1433.80 ($[M+NH_4]^+$, calc. 1433.38).

HR-ESI-MS (pos), m/z: 1416.3584 ($[M+H]^+$, calc. 1416.3582).

^{1}H NMR (600 MHz, 25°C, [D$_4$]MeOD, internal MeOH, *δ*(H)=4.87 ppm): ***δ*=** 7.84-7.78 (m, 3H, H-4-, H-5-Fmoc, N-H-Fmoc), 7.71-7.68 (m, 2H, H-1-, H-8-Fmoc), 7.60-7.56 (s_{br}, 1H, N-H-Ac), 7.43-7.24 (m, 9H, H-2-, H-3-, H-6-, H-7-Fmoc, C_{arom}-SPh), 5.37-5.30 (m, 3H, H-4'', H-4''*, CH_{2a}-(Troc)), 5.25 (d, $^2J(CH_{2a},CH_{2b})$ = 12.4 Hz, 1H, CH_{2b}-(Troc)), 5.12-5.02 (m, 1H, H-3'')$^{\#}$, 5.02-4.95 (m, 1H, H-3'), 4.94-4.88 (m, 2H, H-1', H-2''), 4.78-4.72 (m, 1H, H-1), 4.70-4.62 (m, 4H, H-1'', CH_{2a}-(Fmoc), CH_{2b}-(Troc), CH_{2b}*-(Troc)), 4.42-4.33 (m, 2H, H-2, T^{β}), 4.29 (s_{br}, 1H, H-9-Fmoc), 4.25-4.15 (m, 2H, H-5, H-5*), 4.14-4.00 (m, 4H, H-5'', H-6'', T^{α}), 3.91-3.82 (m, 3H, H-4, H-4', H-6$_a$), 3.80-3.64 (m, 4H, H-3, H-6$_b$, H-6'$_{ab}$), 3.48-3.41 (m, 1H, H-5''), 2.13-1.75 (m, 18H, CH_3-(Ac)), 1.43 (s, 9H, *t*Bu), 1.25 (m, 3H, T^{γ}).

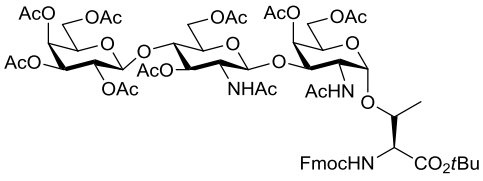
^{13}C NMR (150.9 MHz, 25°C, [D$_4$]MeOD, internal MeOH, *δ*(C)=49.00 ppm): ***δ*=**173.44, 171.84, 171.37, 170.92, 170.73 (C=O-(Ac), C=O-(*t*Bu)), 159.36, 159.23, (CH_{2ab}-(Fmoc), CH_{2ab}*-(Fmoc)), 158.65, 158.33 (CH_{2ab}-(Troc), CH_{2ab}*-(Troc)), 145.30, 145.13 (C-1$_a$-, C-8$_a$-Fmoc), 142.69 (C-4$_a$-, C-5$_a$-Fmoc), 138.08 (C-1-SPh), 130.27 (C_{arom}-SPh), 128.84 (C-2-, C-7-Fmoc),

128.22 (C-3-, C-6-Fmoc), 126.03 (C-1-, C-8-Fmoc), 121.04 (C-4-, C-5-Fmoc), 103.49, 103.00 (C-1', C-1'*), 101.69, 101.45 (C-1, C-1*), 100.52, 100.44 (C-1", C-1"*), 96.52, 96.38 (C_{quart}-(Troc)), 83.40 (C_{quart}-(*t*Bu)), 80.91, 80.41 (C-3, C-3*), 77.50, 77.21 (CH_{2ab}-, CH_{2ab}-*(Troc)), 76.83 (C-4), 76.18 (C-5'), 75.77 (T^{β}), 72.40 (C-3", C-4'), 71.65 (C-2", C-3'), 69.98, 69.35 (C-5", C-5"*), 68.51 (C-4", C-4"*), 67.47 (CH_{2ab}-(Fmoc)), 63.52, 63.04 (C-2', C-2'*), 62.65, 62.60 (C-6_{ab}'), 62.29, 62.23 (C-6_{ab}"), 60.76 (T^{α}, C-6_{ab}), 48.77 (C-2$^{\#}$), 48.64 (C-9-(Fmoc)$^{\#}$), 24.87, 24.52, 24.09, 20.77-20.47 (CH_3-(Ac), CH_3-(Ac)*), 19.84 (T^{γ}).

(*, rotational-isomers, if identifiable)

($^{\#}$, congruent with solvent peak, shift from HSQC)

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-{2-acetamido-4,6-di-_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]-α-D-galactosylpyranosyl}-L-threonine-*tert*-butylester (34)
(Fmoc-Thr(βAc_4Gal-(1→4)-βAc_2GlcNAc-(1→3)-αAc_2GalNAc)-O*t*Bu)



Zinc dust was preactivated by treating it for a few minutes with 1 N hydrochloric acid and then washed with water, methanol and diethyl etherr. Fmoc-Thr(βAc4Gal-(1→4)-βGlcN(SPh)Troc(1→3)-GalNAc)-OtBu **33** (1.60 g, 1.13 mmol) was dissolved in acetic acid and preactivated zinc (1.11 g, 16.8 mmol) was added. The reaction was stirred at room temperature for 16 h. Then the same amount of zinc was additionally added and the reaction stirred for 20 h. The zinc powder was filtered off and washed with acetic acid. The filtrate was co-evaporated with toluene *in vacuo*. The residue was dissolved in pyridine/acetic anhydride (22.5 mL, 2:1) and stirred for 18 h. The solvent was removed *in vacuo* and the crude purified by flash column chromatography on silica (EtOAc→EtOAc/MeOH 25:1).

Yield: 1.11 g (0.85 mmol, 75%), pale yellow, amorphous solid, $[\alpha]_D^{20}$ +47.45 (c = 0.51, $CHCl_3$), R_f = 0.23 (EtOAc/MeOH 25:1).

$C_{61}H_{79}N_3O_{28}$ (M = 1302.28 g/mol) [1301.49].

ESI-MS (pos), m/z: 1301.73 ($[M+H]^+$, calc. 1302.49), 1318.30 ($[M+NH_4]^+$, calc. 1319.52).

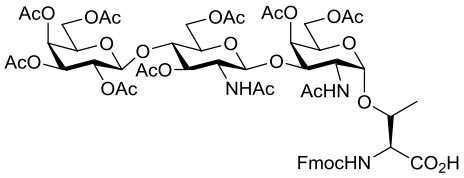
HR-ESI-MS (pos), m/z: 1302.4927 ($[M+H]^+$, calc. 1302.4928).

^{1}H NMR (400 MHz, 25°C, $CDCl_3$, internal $CHCl_3$, *δ*(H)=7.26 ppm): ***δ*=** 7.76 (d, ^3J(H-4,H-3) = ^3J(H-5,H-6) = 7.4 Hz, 2H, H-4-, H-5-Fmoc), 7.62 (d, ^3J(H-1,H-2) = ^3J(H-8,H-7) = 7.3 Hz, 2H, H-1-, H-8-Fmoc), 7.42-7.26 (m, 4H, H-2-, H-7-, H-3-, H-6-Fmoc), 6.46 (d, ^3J (NH,H-2) = 9.3

Hz, 1H, NH-Ac), 6.03 (d, ^3J(NH,T$^{\alpha}$) = 9.5 Hz, 1H, NH-Fmoc), 5.92 (d, ^3J(NH',H-2') = 8.4 Hz, 1H, NH'-Ac), 5.33 (s_{br}, 1H, H-4''), 5.29 (s, 1H, H-4), 5.16-5.07 (m, 2H, H-3', H-2''), 4.97 (dd, ^3J(H-3'',H-2'') = 10.4 Hz, ^3J(H-3'',H-4'') = 3.3 Hz, 1H, H-3''), 4.82 (s_{br}, 1H, H-1), 4.73-4.70 (m, 2H, H-1', H-6$_a$'), 4.55-4.51 (m, 2H, H-1'', CH_{2a}-(Fmoc)), 4.46-4.42 (m, 2H, H-2, CH_{2b}-(Fmoc)), 4.26-4.19 (m, 3H, T$^{\alpha}$,T$^{\beta}$, H-9-Fmoc), 4.11-3.97 (m, H5, H-6$_{ab}$, H-6$_b$', H-6$_{ab}$''), 3.85-3.77 (m, 3H, H-3, H-4', H-5''), 3.68 (q, ^3J(H-2',H-1') = ^3J(H-2',NH') = ^3J(H-2',H-3') = 8.4 Hz, 1H, H-2'), 3.56 (m, 1H, H-5'), 2.13-1.95 (m, 30H, CH_3-(Ac)), 1.44 (s, 9H, *t*Bu), 1.28 (d, ^3J(T$^{\gamma}$,T$^{\beta}$) = 5.4 Hz, 3H, T$^{\gamma}$).

^{13}C NMR (100.6 MHz, 25°C, $CDCl_3$, internal $CHCl_3$, *δ*(H)=77.16 ppm): ***δ*=** 170.41-169.26 (C=O-(Ac), C=O-(*t*Bu)), 157.00 (C=O-(Fmoc)), 143.97, 143.85 (C-1$_a$-, C-8$_a$-Fmoc), 141.45 (C-4$_a$-, C-5$_a$-Fomc), 127.90 (C-2-, C-7-Fmoc), 127.24 (C-3-, C-6-Fmoc), 125.21 (C-1-, C-8-Fmoc), 120.15 (C-4-, C-5-Fmoc), 101.29 (C-1''), 99.66 (C-1, C-1'), 83.19 (C_{quart}-(*t*Bu)), 76.10 (T$^{\beta}$), 75.62 (C-4'), 73.41 (C-5'), 72.75 (C-3'), 72.28 (C-3), 70.98 (C-3''), 70.77 (C-5''), 69.29 (C-4, C-2''), 67.89 (C-5), 67.10 (CH_{2ab}-(Fmoc)), 66.67 (C-4''), 62.69 (C-6$_{ab}$), 61.39 (C-6'$_{ab}$), 60.74 (C-6''$_{ab}$), 59.24 (T$^{\alpha}$), 54.24 (C-2'), 48.82 (C-2), 47.39 (C-9-Fmoc), 28.22 (CH_3-(*t*Bu)), 23.38, 21.02-20.66 (CH_3-(Ac)), 19.08 (T$^{\gamma}$).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-{2-acetamindo-4,6-di-_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]-α-D-galactosylpyranosyl}-L-threonine (35)
(Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc$_2$GlcNAc-(1→3)-αAc$_2$GalNAc)-OH)



Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc$_2$GlcN-(1→3)-αAc$_2$GalNAc)-O*t*Bu **34** (1.39 g, 1.07 mmol) was dissolved in dichloromethane (7.5 mL) and anisole (3 mL). Then trifluoroacetic acid (22.5 mL) was added and the solution was stirred for 2 h. The reaction was co-evaporated with toluene. The crude was purified by flash column chromatography on silica (EtOAc→EtOAc/MeOH 25:1→10:1).

Yield: 1.27 g (1.02 mmol, 95%), colorless, amorphous solid, $[\alpha]_D^{20}$ +53.60 (c = 0.99, $CHCl_3$), R_f = 0.41 (EtOAc/MeOH 3:1).

$C_{57}H_{71}N_3O_{28}$ (M = 1246.18 g/mol) [1245.42].

ESI-MS (pos), m/z: 1245.80 ([M+H]$^+$, calc. 1246.43), 1262.47 ([M+NH$_4$]$^+$, calc. 1263.46), 1268.07 ([M+Na]$^+$, calc. 1268.41).

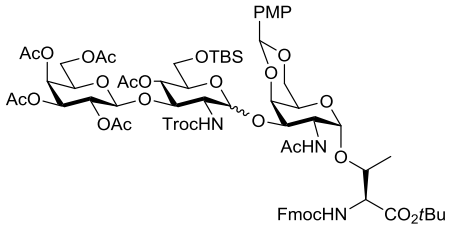
HR-ESI-MS (pos), m/z: 1246.4297 ($[M+H]^+$, calc. 1246.4302).

^{1}H NMR (600 MHz, 25°C, [D_6]DMSO, internal DMSO, *δ*(H)=2.50 ppm): ***δ*=** 12.83 (s_{br}, 1H, COOH), 7.91, 7.90 (2 x d, $J_{H\text{-}4,H\text{-}3} = J_{H\text{-}5,H\text{-}6}$ = 7.6 Hz, 2H, H-4-, H-5-Fmoc), 7.81 (d, $J_{NH,H\text{-}2'}$ = 8.3 Hz, 1H, N-H'-Ac), 2 x 7.73 (2 x d, $J_{H\text{-}1,H\text{-}2} = J_{H\text{-}8,H\text{-}7}$ = 7.4 Hz, 2H, H-1-, H-8-Fmoc), 7.44-7.41 (m, 2H, H-2-, H-7-Fmoc), 7.37 (d, $J_{NH,T\alpha}$ = 9.6 Hz, 1H, N-H-Fmoc), 7.34-7-31 (m, 2H, H-3-, H-6-Fmoc), 7.21 (d, $J_{NH,H2}$ = 9.2 Hz, 1H, N-H-Ac), 5.28 (d, $J_{H\text{-}4,H\text{-}3}$ = 3.5 Hz, 1H, H-4), 5.22 (d, $J_{H\text{-}4'',H\text{-}3''}$ = 3.6 Hz, 1H, H-4'), 5.17 (dd, $J_{H\text{-}3'',H\text{-}4''}$ = 3.6 Hz, $J_{H\text{-}3'',H\text{-}2''}$ = 10.3 Hz, 1H, H-2''), 5.09 (t, $J_{H\text{-}3',H\text{-}4'} = J_{H\text{-}3',H\text{-}2'}$ = 9.4 Hz, 1H, H-3'), 4.85-4.81 (m, 2H, H-1', H-2''), 4.75 (d, $J_{H\text{-}1'',H\text{-}2''}$ = 8.0 Hz, 1H, H1''), 4.74 (d, $J_{H\text{-}1,H\text{-}2}$ = 4.0 Hz, 1H, H-1), 4.51-4-47 (m, 2H, CH_{2ab}-Troc), 4.31 (t, $J_{H\text{-}9,CH2}$ =6.7 Hz, 1H, H-9-Fmoc), 4.27 (m, 1H, H-6'$_a$), 4.22-4.20 (m, 2H, H-5'', T^{β}), 4.12 (dd, $J_{T\alpha,T\beta}$ = 2.1 Hz, $J_{T\alpha,NH\text{-}Fmoc}$ = 9.6 Hz), 4.09-4.02 (m, 4H, H-2, H-5, H-6$_a$, H-6'$_b$, 4.00-3.97 (m, 2H, H-6''$_{ab}$), 3.86-3.80 (m, 2H, H-3, H-6$_b$), 3.65 (t, $J_{H\text{-}4',H\text{-}3'} = J_{H\text{-}4',H\text{-}5'}$ = 9.4 Hz, 1H, H-4'), 3.57 (m, 1H, H5'), 3.25 (m, 1H, H-2'), 2.09, 2.04, 2.03, 2.01, 2.00, 1.99, 1.95, 1.90, 1.83, 1.69 (each s, 30H, CH_3-(Ac)), 1.14 (d, $J_{T\gamma,T\beta}$ = 6.5 Hz, 1H, T^{γ}).

^{13}C NMR (150.9 MHz, 25°C, [D_6]DMSO, internal DMSO, *δ*(H)=39.52 ppm): ***δ*=** 171.54, 170.24, 170.12, 2 x 169.87, 169.53, 169.52, 169.38, 169.24, 169.18, 169.18, 169.17 (C=O-(Ac), 156.79 (C=O-(Fmoc)), 143.77, 143.74 (C-1$_a$-, C-8$_a$-Fmoc), 140.82, 140.79 (C-4$_a$-, C-5$_a$-Fmoc), 127.69, 127.65 (C-2-, C-7-Fmoc), 127.07 (C-3-, C-6-Fmoc), 125.16, 125.08 (C-1-, C-8-Fmoc), 120.23, 120.18 (C-4-, C-5-Fmoc), 100.09 (C-1''), 100.00 (C-1'), 98.93 (C-1), 76.67 (C-4'), 75.02 (T^{β}), 73.02 (C-3), 72.77 (C-3'), 71.49 (C-5'), 70.22 (C-3''), 69.91 (C-4), 69.57 (C-5''), 68.99 (C-2''), 67.03 (C-4'', C-5), 65.50 (CH_{2ab}-(Fmoc)), 62.75 (C-6$_{ab}$), 61.98 (C-6'$_{ab}$), 60.72 (C-6''$_{ab}$), 58.57 (T^{β}), 54.43 (C-2'), 47.96 (C-2), 46.79 (C-9-Fmoc), 22.72, 22.65, 20.78-20.33 (CH_3-(Ac)), 18.72 (T^{γ}).

7.1.7 Synthesis of the extended type-1 core 3 amino acid

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-_N_-acetamido-2-deoxy-4,6-_O_-para-methoxybenzylideneacetal-3-_O_-{4-_O_-acetyl-6-_O_-_tert_-butyldimethysilyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-α/β-D-glucopyranosyl}-α-D-galactopyranosyl)-L-threonine-tert-butylester (36) (βAc$_4$Gal-(1→3)-α/βAc-TBS-GlcNHTroc-(1→3)-αPMP-GalNAc)-Thr-OtBu)



Fmoc-Thr(αPMP-GalNAc)-O*t*Bu **12** (1.51 g, 2.10 mmol) and βAc$_4$Gal-(1→3)-βGlcNTroc-SPh **27** (2.55 g, 2.73 mmol) were dissolved in dry dichloromethane (30 mL). Activated molecular sieves 4Å (2.1 g) were added under argon atmosphere. The suspension was stirred for 1 h and then cooled in an ice bath. *N*-iodosuccinimide (614 mg, 2.73 mmol) was added to the suspension, followed by the dropwise addition of a suspension of trifluoromethanesulfonic acid (37.2 µL, 63.0 mg, 0.42 mmol) in dry dichloromethane (300 µL) *via* a syringe. After 2 h the reaction was treated with another addition of **27** (255 mg, 0.27 mmol) and *N*-iodosuccinimide (61 mg, 0.27 mmol). The reaction was stirred for 3 h under argon atmosphere and ice-cooling. The reaction was diluted with dichloromethane (20 mL). The molecular sieves were filtered off and washed with dichloromethane. The filtrate was further diluted with dichloromethane (total volume 180 mL). The dichloromethane phase was washed with a solution of 0.5 M sodium thiosulfate, saturated sodium bicarbonate, water and brine (60 mL each). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography on silica (^{C}Hex/EtOAc 2:1→1:1), to give a mixture of the α- and β-anomer.

Yield: 2.29 g (α/β (1:1.2), 1.48 mmol, 71%), colorless, amorphous solid, $[\alpha]_D^{20}$ +55.40 (c = 1.00, $CHCl_3$), R_f = 0.47 (^{C}Hex/EtOac 1:1).

$C_{70}H_{92}Cl_3N_3O_{27}Si$ (M = 1541.93 g/mol) [1539.48].

ESI-MS (pos), m/z: 1540.00 ($[M+H]^+$, calc. 1540.48), 1562.20 ($[M+Na]^+$, calc. 1562.47).

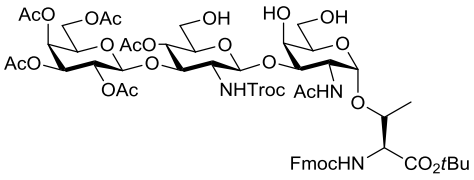
HR-ESI-MS (pos), m/z: 1540.4824 ($[M+H]^+$, calc. 1540.4831), 1562.4653 ($[M+Na]^+$, calc. 1562.4651).

^{1}H NMR (600 MHz, [D$_6$]DMSO, 30°C, internal DMSO *δ*(H)=2.50 ppm): ***δ*=** 7.92-7.88 (m, 4H, H-4$_{αβ}$-, H-5$_{αβ}$-Fmoc), 7.75-7.69 (m, 5H, H-1$_{αβ}$-, H-8$_{αβ}$-Fmoc, NH'$_β$-Troc), 7.56 (d, ^3J(NH$_α$,T^α) = 9.7 Hz, 1H, NH$_α$-Fmoc), 7.51 (d, ^3J(NH$_β$,T^α) = 9.7 Hz, 1H, NH$_β$-Fmoc), 7.44-7.29 (m, 14H,

$NH_{\alpha\beta}$-Ac, $H2_{\alpha\beta}$-, $H7_{\alpha\beta}$-Fmoc, $H\text{-}3_{\alpha\beta}$-, $H\text{-}6_{\alpha\beta}$-Fmoc, $H\text{-}2_{\alpha\beta}$-, $H\text{-}6_{\alpha\beta}$-SPh), 7.07 (d, 1H, NH'_{α}-Troc), 6.96-6.93 (m, 4H, $H\text{-}3_{\alpha\beta}$-, $H\text{-}5_{\alpha\beta}$-SPh), 5.70 (s, 1H, $CH\text{-}PMP_{\alpha}$), 5.45 (s, 1H, $CH\text{-}PMP_{\beta}$), 5.26 (d, $^{3}J(H\text{-}4'_{\alpha},H\text{-}3'_{\alpha})$ = 3.4 Hz, $H\text{-}4'_{\alpha}$), 5.22 (d, $^{3}J(H\text{-}4'_{\beta},H\text{-}3'_{\beta})$ = 3.2 Hz, $H\text{-}4'_{\beta}$), 5.03 (d, $^{3}J(H\text{-}1'_{\alpha},H\text{-}2'_{\beta})$ = 3.2 Hz, $H\text{-}1'_{\alpha}$), 4.99-4.96 (m, 1H, $H\text{-}3'_{\beta}$), 4.94-4.82 (m, 3H, $H4'_{\alpha}$, $H\text{-}3''_{\beta}$, CH_{2a}-Troc), 4.80-4.72 (m, 7H, $H\text{-}1_{\alpha}$, $H\text{-}1_{\beta}$, $H\text{-}1''_{\alpha}$, $H\text{-}2''_{\alpha}$, $H\text{-}2''_{\beta}$), CH_{2a}-Troc, CH_{2b}-Troc), 4.63-4.57 (m, 4H, $H\text{-}1'_{\beta}$, $H\text{-}1''_{\beta}$, $H\text{-}4'_{\beta}$, CH_{2b}-Troc), 4.51-4.38 (m, 5H, $H4_{\alpha}$, $CH_{2\alpha\beta}$-Fmoc), 4.33-4.24 (m, 6H, $H\text{-}2_{\beta}$, $H\text{-}4_{\beta}$, $H\text{-}9_{\alpha\beta}$-Fmoc, $T^{\beta}_{\alpha\beta}$), 4.21-4.16 (m, 1H, $H\text{-}2_{\alpha}$), 4.14-3.90 (m, 14H, $H\text{-}3'_{\beta}$, $H\text{-}5''_{\alpha\beta}$, $H\text{-}6_{ab\alpha\beta}$, $H\text{-}6''_{ab\alpha\beta}$), 3.82-3.74 (m, 8H, $H\text{-}3_{\alpha}$, $H\text{-}3_{\beta}$, $H\text{-}3'_{\alpha}$, $CH_3\text{-}PMP_{\alpha\beta}$), 3.69-3.54 (m, 9H, $H\text{-}2'_{\alpha}$, $H\text{-}5_{\alpha}$, $H\text{-}5_{\beta}$, $H\text{-}5'_{\alpha}$, $H\text{-}6'_{ab\alpha\beta}$), 3.37-3.34 (m,1H, $H\text{-}5'_{\beta}$), 3.18-3.12 (m, 1H, $H\text{-}2'_{\beta}$), 2.12, 2.08, 2.07, 2.02, 1.99, 1.98, 1.96, 1.95, 1.92, 1.91, 1.90, 1.88, 1.87, 1.86 (m, 36H, $CH_3\text{-}(Ac)_{\alpha\beta}$), 1.36 (s, 18H, $CH_3\text{-}(t\text{Bu})_{\alpha\beta}$), 1.17-1.13 (m, 6H, $T^{\gamma}_{\alpha\beta}$), 0.87, 0.83 (2 x s, 18H, $CH_3\text{-}TBS_{\alpha\beta}$), 0.06, -0.02, -0.03 (3 x s, $Me\text{-}TBS_{\alpha\beta}$).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO δ(C)=39.51 ppm): δ= 169.93, 169.88, 169.87, 169.84, 169.84, 169.55, 169.46, 169.41, 169.37, 169.31, 169.21, 169.19, 169.16, 169.12, 169.10, 169.07, 168.98, 168.46 (C=O-(Ac)), 159.55, 159.50 ($C\text{-}1\text{-}PMP_{\alpha\beta}$) 156.79, 156,75 ($C{=}O\text{-}(Fmoc)_{\alpha\beta}$), 154.36, 153.64 ($C{=}O\text{-}(Troc)_{\alpha\beta}$), 143.85, 143.72,143.68, 143.65 ($C\text{-}1_a$-, $C\text{-}8_b$-$Fmoc_{\alpha\beta}$), 140.80, 140.77, 140.73 ($C\text{-}4_a$-, $C\text{-}5_b$-$Fmoc_{\alpha\beta}$), 130.74, 130.67 ($C\text{-}4\text{-}PMP_{\alpha\beta}$), 127.69, 127.66, 127.64, 127.60, 127.47, 127.02 (C-2-, C-7-$Fmoc_{\alpha\beta}$, C-2-, C-6-$PMP_{\alpha\beta}$), 125.22, 125.18, 125.16, 125.13, 124.85 (C-3-, C-6-$Fmoc_{\alpha\beta}$, C-1-, C-8-$Fmoc_{\alpha\beta}$), 120.21, 120.17, 120.09 (C-4- ,C-5-$Fmoc_{\alpha\beta}$), 113.4, 113.35 (C-3-, C-5-$PMP_{\alpha\beta}$), 101.93 ($C\text{-}1'_{\beta}$), 100.09 ($CH\text{-}PMP_{\beta}$), 99.99 ($C\text{-}1''_{\beta}$), 99.43 ($CH\text{-}PMP_{\alpha}$), 99.24, 99.07 ($C\text{-}1_{\alpha}$, $C\text{-}1_{\beta}$, $C\text{-}1''_{\alpha}$), 97.14 ($C\text{-}1'_{\alpha}$), 96.09, 95.93, 95.74 (CCl_3-$Troc_{\alpha\beta}$), 81.42, 81.37 (C_q-*t*Bu), 76.79 ($C\text{-}3'_{\beta}$), 76.16 ($C\text{-}3_{\beta}$), 75.80 ($C\text{-}4_{\beta}$), 75.39 ($C\text{-}3'_{\alpha}$), 74.02 ($C\text{-}3_{\alpha}$), 73.90 (T^{β}_{β}), 73.82 (CH_2-Troc), 73.61 (T^{β}_{α}), 73.18 ($C\text{-}5'_{\beta}$), 73.52 (CH_2-Troc), 72.74 ($C\text{-}4_{\alpha}$), 70.33 ($C\text{-}3''_{\alpha}$), 70.29 ($C\text{-}3''_{\beta}$), 69.99 ($C\text{-}5'_{\alpha}$), 69.77 ($C\text{-}5''_{\alpha}$), 69.72 ($C\text{-}5''_{\beta}$), 68.89 ($C\text{-}4'_{\alpha}$), 68.75 ($C\text{-}2''_{\alpha}$), 68.49, 68.46 ($C\text{-}6_{\alpha\beta}$), 68.39 ($C\text{-}2''_{\beta}$), 67.84 ($C\text{-}4'_{\alpha}$), 67.20, 67.14 ($C\text{-}4''_{\alpha\beta}$), 65.76, 65.57 ($CH_2$-$Fmoc_{\alpha\beta}$), 62.76 ($C\text{-}5_{\alpha}$), 62.71 ($C\text{-}5_{\beta}$), 62.18 ($C\text{-}6'_{\beta}$), 61.09, 60.97 ($C\text{-}6''_{\alpha\beta}$), 60.84 ($C\text{-}6'_{\alpha}$), 59.39, 59.24 ($T^{\alpha}_{\alpha\beta}$), 57.31 ($C\text{-}2'_{\beta}$), 55.54 ($C\text{-}2'_{\alpha}$), 55.10, 55.06 ($CH_3\text{-}PMP_{\alpha\beta}$), 47.28, 47.11 ($C\text{-}2_{\alpha\beta}$), 46.78, 46.75 (C-9-$Fmoc_{\alpha\beta}$), 27.58, 25.74 ($CH_3$-(*t*Bu)), 23.04, 22.97, 20.84, 20.75, 20.61, 20.53, 20.49, 20.44, 20.41, 20.33, 20.27, 20.21 (CH_3-(Ac)), 19.18, 19.17 ($T^{\gamma}_{\alpha\beta}$), -5.35, -5.39, -5.61, -5.70 ($Me\text{-}TBS_{\alpha\beta}$).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-acetamido-2-deoxy-3-_O_-{4-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-α/β-D-glucopyranosyl}-α-D-galactopyranosyl)-L-threonine-_tert_-butylester (37)
(βAc$_4$Gal-(1→3)-βAcGlcNHTroc-(1→3)-αGalNAc)-Thr-OtBu)



Fmoc-Thr(βAc4Gal-(1→3)-α/βAc-TBS-GlcNHTroc(1→3)-αPMP-GalNAc)-O*t*Bu **36** (2.03 g, 1.32 g mmol) was dissolved in 80 vol% acetic acid (25 mL) and stirred 18 h at 40°C. The solvent was removed *in vacuo* by co-evaporation with toluene. The crude was purified by flash column chromatography on silica (^{C}Hex/EtOAc 1:2→EtOAc→EtOAc/MeOH 25:1).

Yield: β-Anomer: 861 mg (0.66 mmol, 50%), colorless, amorphous solid, $[\alpha]_D^{20}$ +29.93 (c = 0.50, $CHCl_3$), R_f = 0.30 (EtOAc/MeOH 20:1).

α-Anomer: 502 mg (0.38 mmol, 29%), colorless, amorphous solid, $[\alpha]_D^{20}$ +56.47 (c = 0.51, $CHCl_3$), R_f = 0.61 (EtOAc/MeOH 20:1).

$C_{56}H_{72}Cl_3N_3O_{26}$ (M = 1309.53 g/mol) [1307.35].

β-Anomer:

ESI-MS (pos), m/z: 1307.93 ($[M+H]^+$, calc. 1308.35).

HR-ESI-MS (pos), m/z: 1308.3546 ($[M+H]^+$, calc. 1308.3548).

^{1}H NMR (600 MHz, [D$_6$]DMSO, 30°C, internal DMSO *δ*(H)=2.50 ppm): ***δ*=** 7.91-7.89 (m, 2H, H-4-, H-5-Fmoc), 7.75 (m, 2H, H-1-, H-8-Fmoc), 7.57 (d, ^3J(NH,H2') = 9.2 Hz, 1H, NH-Troc), 7.52 (d, 1H, ^3J(NH,T$^\alpha$) = 9.8 Hz, NH-Fmoc), 7.43-7.41 (m, 2H, H2-, H7-Fmoc), 7.36-7.29 (m, 3H, NH-Ac, H-3-, H-6-Fmoc), 5.24 (d, ^3J(H-4",H-3") = 3.8 Hz, 1H, H-4"), 4.90 (dd, ^3J(H3",H4") = 3.7 Hz, ^3J(H3",H2") = 10.5 Hz, 1H, H-3"), 4.86 (d, ^2J(CH$_{2a}$,CH$_{2b}$) = 12.3 Hz, 1H, CH$_{2a}$-(Troc)), 4.78 (dd, ^3J(H-2",H-3") = 10.2 Hz, ^3J(H-2",H-1") = 8.0 Hz, 1H, H-2"), 4.73 (t, ^3J(OH-6',H-6'$_{ab}$) = 5.9 Hz, 1H, OH-6'), 4.65-4.63 (m, 2H, H-1", OH-6), 4.59-4.53 (m, 4H, H-1, H-1', H-4', CH$_{2b}$-(Troc)), 4.49 (dd, ^2J(CH$_{2a}$,CH$_{2b}$) = 11.0 Hz, ^3J(CH$_{2a}$,H-9) = 6.9 Hz, 1H, CH$_{2a}$-(Fmoc)), 4.43 (dd, ^2J(CH$_{2b}$,CH$_{2a}$) = 11.0 Hz, ^3J(CH$_{2b}$,H-9) = 4.6 Hz, 1H, CH$_{2b}$-(Fmoc)), 4.32-4.30 (m, 2H, H-9-Fmoc, OH-4), 4.21-4.17 (m, 2H, H-2, T$^\beta$), 4.12-4.09 (m, 1H, H-6'$_a$), 4.07-4.02 (m, 2H, H-5'', T$^\alpha$), 3.97-3.95 (m, 3H, H-4, H-3', H-6"$_b$), 3.69-3.67 (m, 1H, H-5), 3.57 (dd, ^3J(H-3,H-4) = 2.9 Hz, ^3J(H-3,H-2) = 11.1 Hz, 1H, H-3), 3.51-3.45 (m, 3H, H-6'$_a$, H-6$_{ab}$), 3.37-3.34 (m, 3H, H-2', H-6'$_b$, H-5'), 2.09, 2.00, 1.98, 1.90, 1.88 (s, 6 x 3H, CH$_3$-(Ac)), 1.34 (s, 9H, *t*Bu), 1.17 (d, ^3J(T$^\gamma$,T$^\beta$) = 4.7 Hz, 3H, T$^\gamma$).

^{13}C NMR (150.9 MHz, [D$_6$]DMSO, 30°C, internal DMSO *δ*(C)=39.51 ppm): ***δ*=** 172.00, 169.91, 169.86, 169.37, 169.32, 169.20, 169.14, 169.01 (C=O-(Ac), C=O-(*t*Bu)), 156.83

(C=O-(Fmoc)), 153.70 (C=O-(Troc)), 143.71, 143.69 (C-1_a-, C-8_a-Fmoc), 140.81, 140.76 (C-4_a-, C-5_a-Fmoc), 127.71, 127.66 (C-2-, C-7-Fmoc), 127.07, 127.03 (C-3-,C-6-Fmoc), 125.31, 125.16 (C-1-,C-8-Fmoc), 120.20, 120.15 (C-4-, C-5-Fmoc), 101.77 (C-1'), 99.98 (C-1''), 98.96 (C-1), 95.82 (CCl_3-(Troc)), 81.22 (C_q-*t*Bu), 77.18 (C-3'), 76.48 (C-3), 73.74 (CH_2-(Troc), 73.66 (C-5'), 73.53 (T^{β}), 71.52 (C-5), 70.18 (C-3''), 69.71 (C-5''), 69.20 (C-4'), 68.41 (C-2''), 67.86 (C-4), 67.25 (C-4''), 65.57 (CH_2-(Fmoc)), 61.06 (C-6''), 60.94 (C-6), 60.54 (C-6'), 59.42 (T^{α}), 57.40 (C-2'), 47.22 (C-2), 46.76 (C-9-Fmoc), 27.62 (CH_3-(*t*Bu)), 22.93, 20.74, 20.59, 20.51, 20.42, 20.36, 20.23 (CH_3-(Ac)), 19.08 (T^{γ}).

α-Anomer:

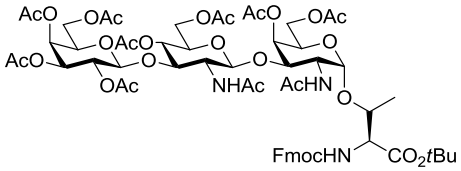
ESI-MS (pos), m/z: 1308.00 ($[M+H]^+$, calc. 1308.35).

HR-ESI-MS (pos), m/z: 1308.3558 ($[M+H]^+$, calc. 1308.3548).

^{1}H NMR (600 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm): ***δ*=** 7.91-7.89 (m, 2H, H-4-, H-5-Fmoc), 7.78-7.71 (m, 3H, NH-Troc, H-1-, H-8-Fmoc), 7.60 (d, ^3J(NH,T^{α}) = 9.8 Hz, 1H, NH-Fmoc), 7.55 (d, ^3J(NH,H-2) = 10.1 Hz, 1H, NH-Ac), 7.42-7.40 (m, 2H, H-2-, H-7-Fmoc), 7.33-7.29 (m, 2H, H-3-, H-6-Fmoc), 5.25 (d, ^3J(H-4'',H-3'') = 3.5 Hz, 1H, H-4''), 5.05-5.02 (m, 2H, H-3'', CH_{2a}-(Troc)), 4.88 (t, ^3J(H-4',H-3') = ^3J(H-4',H-3') = 9.6 Hz, 1H, H-4'), 4.86 (d, ^{3}J(H-1',H-2') = 3.4 Hz, 1H, H-1'), 4.81 (dd, ^3J(H-2'',H-1'') = 8.0 Hz, ^3J(H-2'',H-3'') =10.1 Hz, 1H, H-2''), 4.76 (d, ^3J(H-1'',H-2'') = 8.0 Hz, 1H, H-1''), 4.68-4.65 (m, 2H, OH-6, CH_{2b}-(Troc)), 4.61 (d, ^3J(H-1,H-2) = 4.0 Hz, 1H, H-1), 4.53 (t, ^3J(OH-6',H-6'$_{ab}$) = 5.2 Hz, 1H, OH-6'), 4.48 (d, ^3J(OH-4,H-4) = 3.3 Hz, 1H, OH-4), 4.44 (dd, ^2J(CH_{2a},CH_{2b}) = 10.7 Hz, ^3J(CH_{2a},H-9) = 6.9 Hz, 1H, CH_{2a}-(Fmoc)), 4.37 (dd, ^2J(CH_{2b},CH_{2a}) = 10.7 Hz, ^3J(CH_{2b},H-9) = 6.9 Hz, 1H, CH_{2b}-(Fmoc)), 4.29-4.25 (m, 3H, H-2, H-9-(Fmoc), T^{β}), 4.11-4.09 (m, 2H, H-5'', H-6''$_a$), 4.04-4.01 (m, 2H, T^{α}, H-6''$_b$), 3.87-3.83 (m, 2H, H-4, H-3'), 3.70-3.65 (m, 2H, H-5, H-5'), 3.59-3.51 (m, 5H, H-3, H-2', H-6'$_a$, H-6$_{ab}$), 3.42-3.40 (m, 1H, H-6'$_b$), 2.10, 2.01, 1.99, 1.94, 1.90, 1.84 (6 x s, 6 x 3H, CH_3-(Ac)), 1.35, (s, 9H, CH_3-(*t*Bu)), 1.19 (d, ^3J(T^{γ},T^{β}) = 5.5 Hz, 3H, T^{γ}).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(C)=39.51 ppm): ***δ*=** 169.89, 169.80, 169.39, 169.24, 169.19, 169.03, 168.88 (C=O-(Ac), C=O-(*t*Bu)), 156.86 (C=O-(Fmoc)), 154.57 (C=O-(Troc)), 143.71 (C-1_a-, C-8_a-Fmoc), 140.70 (C-4_a-, C-5_a-Fmoc), 127.62 (C-2-, C-7-Fmoc), 127.00 (C-3-, C-6-Fmoc), 125.34 (C-1-, C-8-Fmoc), 120.07 (C-4-, C-5-Fmoc), 99.99 (C-1''), 98.64 (C-1), 97.96 (C-1'), 95.87 (CCl_3-(Troc)), 81.22 (C_q-*t*Bu), 78.37 (C-3), 75.82 (C-3'), 73.52 (CH_2-(Troc)), 73.11 (T^{β}), 71.33 (C-5), 70.30 (C-5'), 70.20 (C-3''), 69.59 (C-5''), 68.81 (C-2''), 68.17 (C-4'), 67.07 (C-4''), 66.42 (C-4), 65.69 (CH_2-(Fmoc)), 61.03, 60.87 (C-6, C-6''), 59.53 (C-6'), 59.26 (T^{α}), 55.40 (C-2'), 47.17, 46.69 (C-2, C-9-Fmoc), 27.58 (CH_3-(*t*Bu)), 22.92, 20.60, 20.48, 20.32, 20.26, 20.19, (CH_3-(Ac)), 19.12 (T^{γ}).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-_N_-acetamido-4,6-_O_-acetyl-2-deoxy-3-_O_-{2-_N_-acetyl-2-deoxy-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl}-α-D-galactopyranosyl)-L-threonine-_tert_-butylester (38)
(Fmoc-Thr(βAc$_4$Gal-(1→3)-βAc$_2$GlcNAc-(1→3)-αAc$_2$GalNAc)-OtBu)



Zinc dust was activated by treatment with 1 N hydrochloric acid for a few minutes followed by washing with water, methanol and finally diethyl ether. Fmoc-Thr(βAc$_4$-Gal-(1→3)-βAc$_2$GlcNTroc-(1→3)-αGalNAc)-O*t*Bu **37** (861 mg, 0.66 mmol) was dissolved in acetic acid and activated zinc was added. The reaction was stirred at room temperature for 16 h. Then more of the activated zinc (430 mg, 6.6 mmol) was added and the reaction stirred for 22 h. The zinc dust was filtered off and washed with acetic acid. The filtrate was co-evaporated with toluene *in vacuo*. The residue was dissolved in pyridine/acetic anhydride (15 mL, 2:1) and stirred overnight. The solvent was removed *in vacuo* by co-evaporation with toluene and the crude product was purified by flash column chromatography on silica (100% EtOAc).

Yield: 724 mg (0.56 mmol, 85%), colorless, amorphous solid, $[\alpha]_D^{20}$ +56.39 (c = 0.49, $CHCl_3$), R_f = 0.24 (EtOAc).

$C_{61}H_{79}N_3O_{28}$ (M = 1302.28 g/mol) [1301.49].

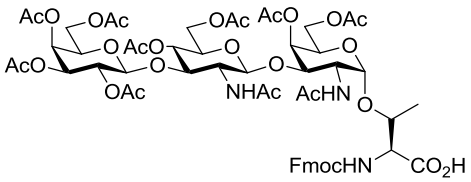
ESI-MS (pos), m/z: 1302.00 ($[M+H]^+$, calc. 1302.49), 1324.40 ($[M+Na]^+$, calc. 1324.47).

HR-ESI-MS (pos), m/z: 1302.4928 ($[M+H]^+$, calc. 1302.4928).

^{1}H NMR (600 MHz, $CDCl_3$, 30°C, internal $CDCl_3$ *δ*(H)= 7.26 ppm): ***δ*=** 7.75 (d, 2H, ^3J(H-4,H-3) = ^3J(H-5,H-6) = 7.6 Hz, H-4-, H-5-Fmoc), 7.61-7.60 (m, 2H, H-1-, H-8-Fmoc), 7.39 (t, ^3J(H-2,H-1) = ^3J(H-2,H-3) = ^3J(H-7,H-6) = ^3J(H-7,H-8) = 7.4 Hz, 2H, H-2-, H-7-Fmoc), 7.31-7.28 (m, 2H, H-3-, H-6-Fmoc), 6.25-6.21 (m, 2H, NH-Ac, NH'-Ac), 6.17 (d, ^3J(NH,T$^\alpha$) = 9.6 Hz, 1H, NH-Fmoc), 5.33 (d, ^3J(H-4'',H-3'') = 3.5 Hz, 1H, H-4''), 5.27 (d, ^3J(H-4,H-3) = 2.7 Hz, 1H, H-4), 5.06 (d, ^3J(H-1',H-2') = 8.4 Hz, 1H, H-1'), 5.04 (dd, ^3J(H-2'',H-3'') = 10.5 Hz, ^3J(H-2'',H-1'') = 8.1 Hz, 1H, H-2''), 4.94-4.90 (m, 2H, H-3'', H-4'), 4.83 (d, ^3J(H-1,H-2) = 3.7 Hz, 1H, H-1), 4.69 (t, ^3J(H-3',H-2') = ^3J(H-3',H-4') = 9.6 Hz, 1H, H-3'), 4.55-4.49 (m, 2H, H-6'$_a$, CH_{2a}-(Fmoc)), 4.46 (d, ^3J(H-1',H-2') = 8.0 Hz, 1H, H-1'), 4.44-4.37 (m, 2H, H-2, CH_{2b}-(Fmoc)), 4.28-4.22 (m, 3H, T$^\alpha$, T$^\beta$, H-9-Fmoc), 4.13-4.05 (m, 4H, H-5, H-6$_a$, H-6''), 4.01-3.97 (m, 2H, H-6$_b$, H-6'$_b$), 3.89 (dd, ^3J(H-3,H-2) = 10.9 Hz, ^3J(H-3,H-4) = 2.6 Hz, 1H, H-3), 3.85 (m, 1H, -H5''), 3.68 (m, 1H, H-5'), 2.88 (q, ^3J(H-2',H-1') = ^3J(H-2',H-3') = ^3J(H-2',NH) = 8.4 Hz, 1H, H-2'), 2.13, 2.08, 2.05, 2.04, 2.03, 1.99, 1.95 (m, 30 H, CH_3-(Ac)), 1.44 (s, 9H, *t*Bu), 1.30 (d, ^3J(T$^\gamma$,T$^\beta$) = 6.4 Hz, 3H, T$^\gamma$).

^{13}C NMR (150.9 MHz, $CDCl_3$, 30°C, internal $CDCl_3$ *δ*(C)=77.16 ppm): ***δ*=** 171.94, 171.78, 170.98, 170.66, 170.54, 170.51, 170.36, 170.28, 169.90, 169.38, 169.30 (C=O-(Ac), C=O-(*t*Bu)), 157.06 (C=O-(Fmoc)), 143.94, 143.89 (C-1_a-, C-8_a-Fmoc), 141.47 (C-4_a-, C-5_a-Fmoc), 127.92, 127.90 (C-2-, C-7-Fmoc), 127.22, 127.18 (C-3-, C-6-Fmoc), 125.09 (C-1-, C-8-Fmoc), 120.15 (C-4-, C-5-Fmoc), 101.07 (C-1''), 99.85 (C-1), 97.74 (C-1'), 83.19 (C_q-*t*Bu), 76.46 (C-3'), 76.00 (T^{β}), 72.49 (C-5'), 71.18 (C-3''), 70.73 (C-5''), 70.67 (C-3), 69.95 (C-4), 69.45 (C-2''), 68.77 (C-4'), 68.68 (C-2''), 67.67 (C-5), 67.03 (CH_2-(Fmoc)), 66.99 (C-4''), 62.62 (C-6''), 61.82 (C-6'), 61.13 (C-6), 59.24 (T^{α}), 58.79 (C-2'), 48.33 (C-2), 47.38 (C-9-Fmoc), 28.24 (*t*Bu), 23.87, 23.47, 21.11, 21.00, 20.98, 20.92, 20.82, 20.80, 20.76, 20.68 (CH_3-(Ac)), 19.08 (T^{γ}).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-_N_-acetamido-4,6-_O_-acetyl-2-deoxy-3-_O_-{2-_N_-acetamido-4,6-_O_-acetyl-2-deoxy-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl}-α-D-galactopyranosyl)-L-threonine (39)
(Fmoc-Thr(βAc$_4$Gal-(1→3)-βAc$_2$GlcNAc-(1→3)-αAc$_2$GalNAc)-OH)



Fmoc-Thr(βAc$_4$Gal-(1→3)-βAc$_2$GlcNAc-(1→3)-αAc$_2$GalNAc)-O*t*Bu (1.27 g, 0.97 mmol) **38** was dissolved in dichloromethane (5 mL) and anisole (1 mL). Then trifluoroacetic acid (15 mL) was added and the solution was stirred for 2.5 h. The reaction was co-evaporated with toluene. The crude was purified by flash column chromatography on silica (EtOAc→EtOAc/MeOH 20:1).

Yield: 1.06 g (0.85 mmol, 87 %), colorless, amorphous solid, $[\alpha]_D^{20}$ +64.90 (c = 1.00, $CHCl_3$), R_f = 0.35 (EtOAc/MeOH/AcOH/H_2O 50:3:3:2).

$C_{57}H_{71}N_3O_{28}$ (M = 1246.18 g/mol) [1245.42].

HR-ESI-MS (pos), m/z: 1246.4301 ($[M+H]^+$, calc. 1246.4302).

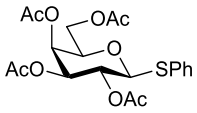
^1H NMR (500 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm): ***δ*=** 7.91-7.89 (m, 2H, H-4-, H-5-Fmoc), 7.74 (d, ^3J(H-1,H-2) = ^3J(H-8,H-7) = 7.4 Hz, 2H, H-1-, H-8-Fmoc), 7.69 (d, ^3J(NH,H-2') = 7.1 Hz, 1H, NH'-Ac), 7.45-7.30 (m, 5H, H-2-, H-3-, H-6-, H-7-Fmoc, NH-Fmoc), 7.26 (d, ^3J(NH,H-2) = 9.4 Hz, 1H, NH-Ac), 5.27 (d, ^3J(H-4,H-3) = 3.4 Hz, 1H, H-4), 5.24 (d, ^3J(H-4'',H-3'') = 3.9 Hz, 1H, H-4''), 5.08 (dd, ^3J(H-3'',H-2'') = 10.3 Hz, ^3J(H-3'',H-4'') = 3.6 Hz, 1H, H-3''), 4.77 (dd, ^3J(H-2'',H-3'') = 10.2 Hz, ^3J(H-2'',H-1'') = 7.8 Hz, 1H, H-2''), 4.74 (d, ^{3}J(H-1,H-2) = 4.1 Hz, 1H, H-1), 4.64 (d, ^3J(H-1'', H-2'') = 8.1 Hz, 1H, H-1''), 4.63-4.60 (m, 2H, H-1', H-4'), 4.53-4.44 (m, 2H, CH_2-(Fmoc)), 4.31 (t, ^3J(H-9,CH_{2a}) = ^3J(H-9,CH_{2b}) = 6.6 Hz, 1H, H-9-

Fmoc), 4.26-4.24 (m, 1H, T^{β}), 4.15-3.97 (m, 9H, H-2, H-5, H-5'', H-$6_{a,b}$, H-$6'_{a,b}$, H-$6''_{a}$, T^{α}), 3.88-3.85 (m, 1H, H-$6''_{b}$), 3.80 (dd, ^{3}J(H-3,H-2) = 11.2 Hz, ^{3}J(H-3,H-4) = 3.4 Hz, 1H, H-3), 3.65-3.62 (m, 1H, H-5'), 3.38 (m, 1H, H-2'), 2.09, 2.04, 2.01, 2.00, 1.99, 1.98, 1.89, 1.88, 1.98 (s, 30 H, CH_3-(Ac)), 1.16-1.14 (m, 3H, T^{γ}).

^{13}C NMR (125.8 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(C)=39.51 ppm): ***δ=*** 171.55, 170.24, 170.00, 169.99, 169.82, 169.60, 169.40, 169.35, 169.29, 169.16, 169.12 (C=O-(Ac)), 156.74 (C=O-(Fmoc)), 143.73, 143.68 (C-1_a-, C-8_a-Fmoc), 140.78, 140.73 (C-4_a-, C-5_a-Fmoc), 127.64, 127.59 (C-2-, C-7-Fmoc), 127.00 (C-3-, C-6-Fmoc), 125.16, 125.07 (C-1-, C-8-Fmoc), 120.13, 120.07 (C-4-, C-5-Fmoc), 100.56 (C-1'), 99.87 (C-1''), 98.82 (C-1), 76.80 (C-3'), 74.65 (T^{β}), 72.62 (C-3), 70.49, 70.38 (C-3'', C-5'), 69.74 (C-4), 69.47 (C-5), 68.85 (C-4'), 68.61 (C-2''), 67.14 (C-4''), 67.00 (C-5''), 66.51 (CH_2-(Fmoc)), 62.70 (C-6''), 61.68, 61.07 (C-6, C-6'), 58.49 (T^{α}), 47.09 (C-2), 46.78 (C-9-Fmoc), 22.83, 22.76, 20.68, 20.58, 20.48, 20.43, 20.39, 20.27, 20.22 (CH_3-(Ac)), 18.34 (T^{γ}).

7.1.8 Synthesis of the T-antigen glycosyl acceptor building block

Phenyl 2,3,4,6-tetra-*O*-acetyl-1-thio-β-D-galactosylpyranoside (40) (βAc_4Gal-SPh)



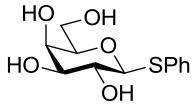
αAc_5Gal **16** (60.0 g, 153.7 mmol) and thiophenol (62.7 mL, 615 mmol) were dissolved in dry dichloromethane (400 mL). Boron trifluoride diethyl etherate (31.3 mL, 48%, 254.1 mmol) was added dropwise. The reaction was stirred for 5 d and then neutralized with triethylamine. The solvent was removed *in vacuo*. The crude was purified by flash column chromatography on silica (^{C}Hex/EtOAc 2:1).

Yield: 55.0 g (124.9 mmol, 81%), colorless, amorphous solid, R_f = 0.48 (Tol/EtOAc 3:1).

$C_{20}H_{24}O_9S$ (M = 440.46 g/mol) [440.11].

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.53-7.50 (m, 2H, H2-, H6-SPh), 7.32-7.31 (m, 3H, H3-, H4-, H5-SPh), 5.41 (d, 1H, H4, $J_{H4,H3}$ = 2.7 Hz), 5.24 (t, 1H, H2, $J_{H2,H3}$ = $J_{H2,H1}$ = 9.9 Hz), 5.05 (dd, 1H, H3, $J_{H3,H2}$ = 9.9 Hz, $J_{H3,H4}$ = 3.3 Hz), 4.72 (d, 1H, H1, $J_{H1,H2}$ = 10.0 Hz), 4.21-4.09 (m, 2H, H6), 3.94 (t, 1H, H5, $J_{H5,H6ab}$ = 7.1 Hz), 2.12, 2.09, 2.04, 1.97 (4 x s, 4 x 3H, 4 x CH_3(Ac)).

Phenyl 1-thio-β-D-galactosylpyranoside (41)
(βGal-SPh)



βAc4Gal-SPh **40** (55.05 g, 124.9 mmol) was added to a solution of sodium methoxide (1.08 g, 20.0 mmol) in methanol (400 mL) and the reaction was stirred for 3 h. It was neutralized with acidic cation exchange resin (*Dowex 50WX8*). After filtration of the resin, the filtrate was concentrated *in vacuo* and the residue purified by flash column chromatography on silica (DCM/MeOH 7:1).

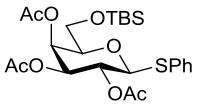
Yield: 33.1 g (121.6 mmol, 97%), colorless solid, $[\alpha]_D^{20}$ -31.29 (c = 1.01, $CHCl_3$), R_f = 0.13 (DCM/MeOH 7:1).

$C_{12}H_{16}O_5S$ (M = 272.32 g/mol) [272.07].

^{1}H-NMR (400 MHz, MeOH-d_4, gCOSY, gHSQC), *δ (ppm)*: 7.61-7.58 (m, 2H, H2-, H6-SPh), 7.35-7.27 (m, 3H, H3-, H4-, H5-SPh), 4.58 (d, 1H, H1, $J_{H1,H2}$ = 9.7 Hz), 3.95 (dd, 1H, H4, $J_{H4,H3}$ = 3.3 Hz, $J_{H4,H5}$ = 0.8 Hz), 3.84-3.3.74 (m, 2H, $H6_{ab}$), 3.68-3.60 (m, 2H, H2, H5), 3.55 (dd, 1H, H3, $J_{H3,H2}$ = 9.2 Hz, $J_{H3,H4}$ = 3.3 Hz).

^{13}C-NMR (100.6 MHz, MeOH-d_4, gHSQY), *δ (ppm)*: 136.02, 132.09, 129.85, 127.97 (PhS), 90.25 (C1), 80.58 (H5), 76.32 (C3), 70.99 (C2), 70.40 (C4), 62.60 (C6).

Phenyl 2,3,4-*O*-acetyl-6-*tert*-butyldimethylsilyl-1-thio-β-D-galactosylpyranoside[266] (42)
(βAc$_3$TBS-Gal-SPh)



Imidazole (10.0 g, 146.8 mmol) and *tert*-butyldimethylsilyl chloride (12.2 g, 80.8 mmol) was added to a solution of βGal-SPh **41** (20.1 g, 73.6 mmol) in dry *N,N*-dimethylformamide (150 mL). The reaction was stirred for 3 h at room temperature. The reaction was diluted with ethyl acetate (600 mL) and washed with water (600 mL), saturated sodium bicarbonate (200 mL) and brine (200 mL). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The residue was dissolved in pyridine and acetic anhydride (300 mL, 2:1). The reaction was stirred for 18 h and then co-evaporated with toluene *in vacuo*. The crude was purified by flash column chromatography on silica (^{C}Hex/EtOAc 4:1).

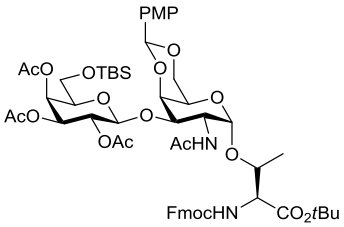
Yield: 27.3 g (53.2 mmol, 73%), colorless, amorphous solid, $[\alpha]_D^{20}$ +5.10 (c = 1.00, $CHCl_3$), R_f = 0.4 (^{C}Hex/EtOAc 4:1).

$C_{24}H_{36}O_8SSi$ (M = 512.69 g/mol) [512.19].

^{1}H-NMR (400 MHz, $CDCl_3$, gCOSY, gHSQC), *δ (ppm)*: 7.50-7.48 (m, 2H, H2-, H6-SPh), 7.30-7.28 (m, 3H, H3-, H4-, H5-SPh), 5.48 (d, 1H, H4, $J_{H4,H3}$ = 3.2 Hz), 5.25-5.20 (m, 1H, H2), 5.07-5.04 (m, 1H, H3), 4.74-4.71 (m, 1H, H1), 3.76-3.71 (m, 2H, H5, $H6_a$), 3.63-3.58 (m, 1H, $H6_b$), 2.10, 2.07, 1.96 (3 x s, 3 x 3H, 3 x CH_3(Ac)), 0.84 (s, 9H, *t*Bu(TBS)), 0.01, 0.00 (2 x s, 2 x 3H, Me(TBS)).

^{13}C-NMR (100.6 MHz, $CDCl_3$, gHSQY), *δ (ppm)*: 170.19, 170.14, 169.62 (C=O(Ac)), 133.28 (C1-SPh), 132.11 (C2-, C6-SPh), 129.05 (H3-, H5-SPh), 127.98 (C4-SPh), 87.04 (C1), 77.55 (C5), 72.48 (C3), 67.76 (C2), 67.39 (C4), 60.99 (C6), 21.00, 20.84, 20.77 (CH_3(Ac)), 18.11 (*t*Bu(TBS)), -5.40,-5.56 (Me(TBS)).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-[2-acetamido-2-deoxy-4,6-_O_-_para_-methoxybenzylidene-3-_O_-(2,3,4-_O_-acetyl-6-_tert_-butyldimethylsilyl-β-D-galactopyranosyl)-α-D-galactopyranosyl]-L-threonine-_tert_-butylester (43) (Fmoc-Thr(βAc$_3$TBS-Gal-(1→3)-αPMP-GalNAc)-O*t*Bu)



Fmoc-Thr(PMP-GalNAc)-O*t*Bu **12** (13.93 g, 19.4 mmol) and βAc$_3$TBS-Gal-SPh **42** (13.91 g, 27.2 mmol) were dissolved in dry dichloromethane (200 mL) under argon atmosphere. Activated molecular sieves (15 g, 4Å) were added and the suspension was stirred for 1 h at room temperature before the mixture was cooled in an ice bath. *N*-iodosuccinimide (6.07 g, 27.1 mmol) was added, followed by (336 µL, 570 mg, 3.8 mmol) trifluoromethanesulfonic acid in dichloromethane (4 mL). The reaction was stirred for 3 h in the ice bath and then 0.5 h without cooling. The reaction was filtered over *Celite*® and washed with dichloromethane (200 mL). The dichloromethane phase was washed with 0.5 M sodium thiosulfate solution, sodium bicarbonate, water and brine (150 mL each). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The crude was purified by flash column chromatography on silica (^{C}Hex/EtOAc 1:1).

Yield: 16.08 g (14.3 mmol, 74%), colorless, amorphous solid, $[\alpha]_D^{20}$ +57.90 (c = 1.00, $CHCl_3$), R_f = 0.31 (^{C}Hex/EtOAc 2:3).

$C_{57}H_{76}N_2O_{19}Si$ (M = 1121.3 g/mol) [120.48].

ESI-MS (pos), m/z: 1003.13 ([M-PMP+H]$^+$, calc. 1003.44), 1121.20 ([M+H]$^+$, calc. 1121.49).

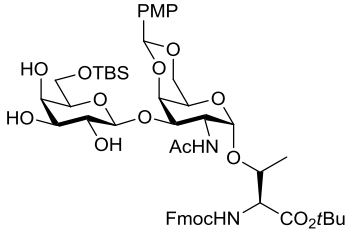
HR-ESI-MS (pos), m/z: 1121.4875 ([M+H]$^+$, calc. 1121.4890).

^{1}H NMR (400 MHz, [D_6]DMSO, 30°C, internal DMSO δ(H)= 2.50 ppm): δ= 7.90 (d, ^{3}J(H-3,H-4) = ^{3}J(H-5,H-6) = 7.4 Hz, 2H, H-4-, H-5-Fmoc), 7.73 (d, ^{3}J(H-1,H-2) = ^{3}J(H-8,H-7) = 7.6 Hz, 2H, H-1-, H-8-Fmoc), 7.48-7.29 (m, 7H, H-2-, H-7-, H-3-, H-6-Fmoc, NH-Ac, NH-Fmoc, H-3-, H-5-PMP), 6.94 (d, ^{3}J(H-2,H-3) = ^{3}J(H-6,H-5) = 9.0 Hz, 2H, H-2-, H-6-PMP), 5.53 (s, 1H, CH-(PMP)), 5.32 (d, ^{3}J(H-4',H-3') = 3.6Hz, 1H, H-4'), 5.00 (dd, ^{3}J(H-3',H-4') = 3.5 Hz, ^{3}J(H-3',H-2') = 10.4 Hz, 1H, H-3'), 4.91 (m, 1H, H-2'), 4.76 (d, ^{3}J(H-1',H-2') = 8.0 Hz, 1H, H-1'), 4.72 (d, ^{3}J(H-1,H-2) = 3.8 Hz, 1H, H-1), 4.51-4.49 (m, 2H, CH_2-(Fmoc)), 4.31-4.22 (m, 4H, H-2, H-5, H-9-Fmoc, T^{β}), 4.11-4.00 (m, 3H, H-6_{ab}', T^{α}), 3.93 (t, ^{3}J(H-5',H-6a') = ^{3}J(H-5',H-6b') = 6.8Hz, 1H, H-5'), 3.80 (dd, ^{3}J(H-3,H-4) = 3.3 Hz, ^{3}J(H-3,H-2) = 11.4 Hz, 1H, H-3), 3.76 (s, 3H, CH_3-(PMP)), 3.66-3.64 (m, 2H, H-4, H-6_a), 3.56-3.54 (m, 1H, H-6_b), 2.09, 1.98, 1.89, 1.85 (each s, 3H, CH_3-(Ac)), 1.37 (s, 9H, *t*Bu-(Thr)), 1.13 (d, ^{3}J(T^{γ},T^{β}) = 6.3 Hz, 3H, T^{γ}), 0.83 (s, 9H, *t*Bu-(TBS)), 0.02, 0.00 (each s, 3H, Me-(TBS)).

^{13}C NMR (100.6 MHz, [D_6]DMSO, 30°C, internal DMSO δ(C)= 39.52 ppm): δ= 169.82, 169.55, 168.13, 168.93, 168.86 (C=O-(Ac), C=O-(*t*Bu)), 159.53 (C-1-PMP), 156.75 (C=O-(Fmoc)), 143.74, 143.60 (C-1_a-,C-8_b-Fmoc), 140.82, 140.79 (C-4_a-, C-5_b-Fmoc), 130.74 (C-4-PMP), 127.72, 127.69, 127.63, 126.98 (C-2-, C-6-PMP, C-2-,C-7-, C-3-, C-6-Fmoc), 125.17, 125.11 (C-1-, C-8-Fmoc), 120.22, 120.16 (C-4-, C-5-Fmoc), 113.38 (C-3-, C-5-PMP), 101.49 (C-1'), 99.73 (CH-(PMP)), 99.21 (C-1), 81.45 (C_{quart}-(*t*Bu)), 75.19 (T^{β}), 74.55 (C-3), 73.73 (C-5), 72.42, 72.35 (C-5')*, 70.74 (C-3'), 68.32 (C-2', C-6_{ab}), 66.89 (C-4'), 65.51 (CH_2-(Fmoc)), 62.89 (C-4), 60.43 (C-6_{ab}'), 59.30 (T^{α}), 55.13 (CH_3-(PMP)), 47.04 (C-2), 46.79 (C-9-(Fmoc)), 27.59 (*t*Bu-(Thr)), 25.64 (*t*Bu-(TBS)), 22.97, 20.49, 20.41, 20.34 (CH_3(Ac)), 19.89 (T^{γ}), 18.84 (C_{quart}(TBS)), -6.09, -6.12 (Me(TBS)).

*signal doubling due to conformers

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-[2-acetamido-2-deoxy-4,6-_O_-_para_-methoxybenzylidene-3-_O_-(6-_tert_-butyldimethylsilyl-β-D-galactopyranosyl)-α-D-galactopyranosyl]-L-threonine-_tert_-butylester (44)
(Fmoc-Thr(βTBS-Gal-(1→3)-αPMP-GalNAc)-O*t*Bu)



To a solution of Fmoc-Thr(βAc_3TBS-Gal-(1→3)-βPMP-GalNAc)-O*t*Bu **43** (14.74 g, 13.1 mmol) in methanol (150 mL) was added small portions of a solution of sodium methoxide (1 wt%) in methanol, to keep the pH at 9.5. After 36 h the solution was neutralized with acidic

cation exchange resin (*Dowex 50 WX8*). The resin was filtered off and washed with methanol. The filtrate was concentrated *in vacuo*. The residue was dissolved in 1,4-dioxane (150 mL) and a solution of sodium bicarbonate (1.50 g, 17.8 mmol) in water (150 mL) was added. *N*-9-Fluorenylmethyloxycarbonyl-succinimidylcarbonate (3.69 g, 10.9 mmol) was added and stirred for 3 h. The 1,4-dioxane was removed *in vacuo*, the aqueous phase diluted with water (150 mL) and extracted with ethyl acetate (500 mL and 200 mL). The combined organic phases were washed with water and brine, dried over sodium sulfate, filtered and concentrated *in vacuo*. The crude was purified by flash column chromatography on silica (^{C}Hex/EtOAc 1:2→ EtOAc).

Yield: 9.42 g (9.47 mmol, 72%), colorless, amorphous solid, $[\alpha]_D^{20}$ +80.20 (c = 1.00, $CHCl_3$), R_f = 0.19 (EtOAc).

$C_{51}H_{70}N_2O_{16}Si$ (M = 995.19 g/mol) [994.45].

ESI-MS (pos), m/z: 995.20 ($[M+H]^+$, calc. 995.46), 1017.40 ($[M+Na]^+$, calc. 1017.44).

HR-ESI-MS (pos), m/z: 995.4569 ($[M+H]^+$, calc. 995.4573), 881.3706 ($[M\text{-}TBS+H]^+$, calc. 881.3708).

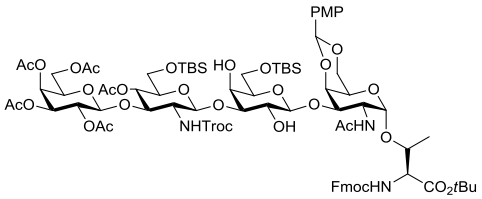
^{1}H NMR (400 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm): ***δ*=** 7.90 (d, ^{3}J(H-3,H-4) = ^{3}J(H-5,H-6) = 6.7 Hz, 2H, H-4-, H-5-Fmoc), 7.72 (d, ^{3}J(H-1,H-2) = ^{3}J(H-8,H-7) = 7.5 Hz, 2H, H-1-, H-8-Fmoc), 7.59 (d, ^{3}J(NH,T^{α}) = 9.8 Hz, 1H, NH-Fmoc), 7.50 (d, ^{3}J(NH,H-2) = 9.8 Hz, 1H, NH-Ac), 7.43 (t, ^{3}J(H-2,H-1) = ^{3}J(H-2,H-3) = ^{3}J(H-7,H-6) = ^{3}J(H-7,H-8) = 7.1 Hz, 2H, H2-, H-7-Fmoc), 7.37 (d, ^{3}J(H-2,H-3) = ^{3}J(H-6,H-5) = 8.7 Hz, 2H, H-2-, H-6-PMP), 7.31 (t, ^{3}J(H-3,H-2) = ^{3}J(H-3,H-4) = ^{3}J(H-6,H-5) = ^{3}J(H-6,H-7) = 7.1 Hz, 2H, H-3-, H-6-Fmoc), 6.94 (d, ^{3}J(H-3,H-4) = ^{3}J(H-3,H-2) = ^{3}J(H-6,H-5) = ^{3}J(H-6,H-7) = 8.7 Hz, 2H, H-3-, H-6-PMP), 5.51 (s, 1H, CH-(PMP)), 4.78-4.76 (m, 2H, H-1, OH-2'), 4.48-4.46 (m, 3H, CH_{2ab}-(Fmoc), OH-4'), 4.33-4.26 (m, 5H, H-1', H-2, H-5, H-9-Fmoc, T^{β}), 4.13-4.10 (m, 1H, T^{α}), 4.07-3.96 (m, 2H, H-6_{ab}), 3.81 (dd, ^{3}J(H-3,H-4) = 2.4 Hz, ^{3}J(H-3,H-2) = 11.3 Hz, 1H, H-3), 3.76 (s, 3H, CH_3-(PMP)), 3.73-3.67 (m, 2H, H-6_{ab}'), 3.63-3.59 (m, 2H, H-4, H-4'), 3.40 (t, ^{3}J(H-5',H-6a'), ^{3}J(H-5',H-6b') = 5.7 Hz, 1H, H-5'), 3.31-3.24 (m, 3H, H-2', H-3', OH-3'), 1.82 (s, 3H, CH_3-(Ac)), 1.37 (s, 9H, *t*Bu-(Thr)), 1.12 (d, ^{3}J(T^{γ},T^{β}) = 6.2 Hz, 3H, T^{γ}), 0.87 (s, 9H, *t*Bu-(TBS)), 0.05 (s, 6H, Me(TBS)).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(C)= 39.52 ppm): ***δ*=** 169.44, 169.02 (C=O-(Ac), C=O-(*t*Bu)), 159.55 (C-1-PMP), 156.81 (C=O-(Fmoc)), 143.76, 143.58 (C-1_a-,C-8_b-Fmoc), 140.80, 140.78 (C-4_a-, C-5_b-Fmoc), 130.84 (C-4-PMP), 127.73, 126.67 (C-2-, C-6-PMP, C-2-, C-7-, C-3-, C-6-Fmoc), 125.18, 125.15 (C-1-,C-8-Fmoc), 120.22, 120.16 (C-4-, C-5-Fmoc), 113.31 (C-3-, C-5-PMP), 105.78 (C-1'), 99.85 (CH-(PMP)), 99.62 (C-1), 81.36 (C_{quart}-(*t*Bu)), 75.87 (T^{β}), 75.44 (C-5'), 74.62 (C-3), 73.86 (C-5), 72.28 (C-3'), 70.13 (C-2'), 68.43 (C-6_{ab}), 68.24 (C-4'), 65.53 (CH_{2ab}-(Fmoc)), 63.07 (C-4), 62.77 (C-6_{ab}'), 59.77 (T^{α}),

55.14 (CH_3-(PMP)), 47.36 (C-2), 46.80 (C-9-Fmoc), 27.61 (*t*Bu-(Thr)), 25.80 (*t*Bu-(TBS)), 23.01 (CH_3-(N-H-Ac)), 19.29 (T^{γ}), 17.95 (C_{quart}-(TBS)), -5.29 (2 x Me-(TBS)).

7.1.9 Synthesis of the extended type-1 core 1 amino acid

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-_N_-acetamido-2-deoxy-4,6-_O_-_para_-methoxybenzylideneacetal-3-_O_-{6-_O_-_tert_-butyldimethylsilyl-3-_O_-[4-_O_-acetyl-6-_O_-_tert_-butyldimethysilyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonyl)-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-β-D-glucopyranosyl}-α-D-galactopyranosyl)-L-threonine-_tert_-butylester (45)
(Fmoc-Thr(βAc_4Gal-(1→3)-βAc-TBS-GlcNAc-(1→3)-βTBS-Gal-(1→3)-αPMP-GalNAc)-OtBu)



Fmoc-Thr(βTBS-Gal-(1→3)-βPMP-GalNAc)-O*t*Bu **44** (4.96 g, 4.99 mmol) and βAc_4Gal-(1→3)-βAc-TBS-GlcNHTroc-SPh **27** (5.91 g, 6.33 mmol) were dissolved in dry dichloromethane (100 mL) under argon atmosphere. Activated molecular sieves 4Å (6.31 g) were added and the suspension was stirred for 1 h before it was cooled in an ice bath. *N*-iodosuccinimide (1.46 g, 6.33 mmol) was added followed by the dropwise addition of a suspension of trifluoromethanesulfonic acid (88 µL, 150 mg, 1.0 mmol) in dry dichloromethane (1 mL) *via* a syringe. After 2 h the reaction was treated with another addition of **27** (1.07 g, 1.0 mmol) and *N*-iodosuccinimide (225 mg, 1.0 mmol). The reaction was stirred for 1.5 h under and then diluted with dichloromethane (100 mL). The molecular sieves were filtered off and washed with dichloromethane (total volume 300 mL). The dichloromethane phase was washed with 0.5 M sodium thiosulfate solution, saturated sodium bicarbonate, water and brine (100 mL each). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography on silica (^{C}Hex/EtOAc 2:1→1:1).

Yield: 5.81 g (3.2 mmol, 64%), colorless, amorphous solid, $[\alpha]_D^{20}$ +49.04 (c = 1.01, $CHCl_3$), R_f = 0.30 (^{C}Hex/EtOAc).

$C_{82}H_{116}Cl_3N_3O_{32}Si_2$ (M = 1818.33 g/mol) [1815.61].

ESI-MS (pos), m/z: 1816.87 ($[M+H]^+$, calc. 1816.62). 1834.00 ($[M+NH_4]^+$, calc. 1833.65), 1838.00 ($[M+Na]^+$, calc. 1838.60).

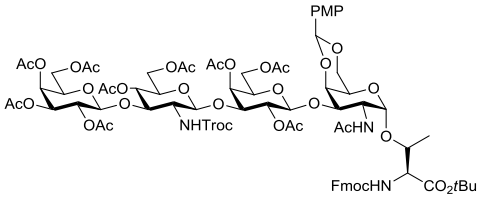
HR-ESI-MS (pos), m/z: 1816.6234 ($[M+H]^+$, calc. 1816.6224).

^{1}H NMR (600 MHz, $[D_6]$DMSO, 30°C, internal DMSO δ(H)= 2.50 ppm): δ= 7.90-7.88 (m, 2H, H-4-, H-5-Fmoc), 7.75-7.72 (m, 2H, H-1-, H-8-Fmoc), 7.62 (d, ^3J(NH,H-2'') = 9.0 Hz, 1H, NH-Troc), 7.52 (d, ^3J(NH,T$^\gamma$) = 9.8 Hz, 1H, NH-Fmoc), 7.48 (d, ^3J(NH,H-2) = 9.6 Hz, 1H, NH-Ac), 7.42 (t, ^3J(H-2,H-7) = ^3J(H-2,H-1) = ^3J(H-7,H-6) = ^3J(H-7,H-8) = 7.5 Hz, 2H, H-2-, H-7-Fmoc), 7.36 (d, ^3J(H-3,H-2) = ^3J(H-5,H-6) = 8.7 Hz, 2H, H-3-, H-5-PMP), 7.31 (t, ^3J(H-3,H-2) = ^3J(H-3,H-4) = ^3J(H-6,H-7) = ^3J(H-6,H-5) =7.6 Hz, 2H, H-3-, H-6-Fmoc), 6.94 (d, ^3J(H-2,H-3) = ^3J(H-6,H-5) = 8.8 Hz, 2H, H-2-, H-6-PMP), 5.53 (s, 1H, CH-PMP), 5.21 (d, ^3J(H-4'',H-3'') = 3.6 Hz, 1H, H-4''), 4.97 (dd, ^3J(H-3'',H-4'') = 3.6 Hz, J(H-3'',H-2'') = 10.3 Hz, 1H, H-3''), 4.83-4.81 (d, ^3J(CH_{2a},CH_{2b}) = 12.2 Hz, 1H, (CH_{2a}-Troc), 4.75 (m, 3H, H-1, H-1'', H-2'''), 4.70 (d, ^3J(CH_{2b},CH_{2a}) = 12.2 Hz, CH_{2b}-(Troc)), 4.67 (d, ^{3}J(H-1'',H-2'') = 8.0 Hz, 1H, H-1'''), 4.56 (t, ^3J(H4'',H3'') = ^3J(H-4'',H-5'') = 9.4 Hz, 1H, H-4''), 4.47-4.44 (m, 3H, OH-2', CH_2-(Fmoc)), 4.38 (d, ^3J(OH-4',H-4') = 4.9 Hz, 1H, OH-4'), 4.34-4.27 (m, 5H, H-1', H-2, H-5, T$^\gamma$, H-9-Fmoc), 4.13 (d, ^3J(T$^\alpha$,T$^\beta$) = 11.1 Hz, 1H, T$^\alpha$), 4.09-4.02 (m, 3H, H-5''', H-6_a, H-6'''$_a$), 4.00-3.94 (m, 3H, H-3'', H-6_b, H-6'''$_b$), 3.83 (dd, ^3J(H-3,H-4) = 2.0 Hz, ^3J(H-3,H-2) = 11.2 Hz, 1H, H-3), 3.78-3.75 (m, 4H, CH_3-(PMP), H-4'), 3.72-3.65 (m, 2H, H-6''$_{ab}$), 3.64 (s, 1H, H-4), 3.61-3.51 (m, 2H, H-6'$_{ab}$), 3.45-3.42 (m, 1H, H-2'), 3.39-3.33 (m, 5H, H-2', H-3', H-5', H-2'', H-5''), 2.09, 2.00, 1.99, 1.98, 1.87, 1.84 (m, 18H, CH_3-(Ac)), 1.38 (*t*Bu-(Thr)), 1.12 (d, ^3J(T$^\gamma$,T$^\beta$) = 7.2 Hz, 3H, T$^\gamma$), 0.86, 0.78 (s, 2 x 9H, *t*Bu-TBS), 0.05, -0.02, -0.03 (s, 4 x 3H, CH_3-TBS).

^{13}C NMR (150.9 MHz, $[D_6]$DMSO, 30°C, internal DMSO δ(C)=39.51 ppm): δ= 169.84, 169.48, 169.37, 169.27, 169.07, 169.04 (C=O-(Ac), C=O-(*t*Bu)), 159.49 (C-1-PMP), 156.66 (C=O-(Fmoc)), 154.34 (C=O-(Troc)), 143.80, 143.51 (C-1_a-, C-8_b-Fmoc), 140.79 (C-4_a-, C-5_b-Fmoc), 130.81 (C-4-PMP), 127.68, 127.64, 127.60 (C-2-, C-7-Fmoc, C-2-, C-6-PMP), 126.99 (C-3-, C-6-Fmoc), 125.19 (C-1-, C-8-Fmoc), 120.16, 120.10 (C-4- ,C-5-Fmoc), 113.26 (C-3-, C-5-PMP), 105.56 (C-1'), 101.39 (C-1''), 100.05 (C-1'''), 99.65 (CH-(PMP)), 99.60 (C-1''), 95.93 (CCl_3-(Troc)), 81.40 (C_q-*t*Bu(Thr)), 80.59 (C-3'), 77.55 (C-3''), 75.77 (C-5), 75.14 (C-5'), 74.88 (C-3), 74.09 (T$^\beta$), 73.80 (C-5''), 73.66 (CH_2-(Troc)), 70.36 (C-3'''), 69.78 (C-2'), 69.61 (C-5'''), 69.15 (C-4''), 68.64 (C-2'''), 68.40 (C-6), 67.28 (C-4'), 67.15 (C-4'''), 65.59 (CH_2-(Fmoc)), 63.10 (C-4), 62.35 (C-6'), 62.21 (C-6''), 60.87 (C-6'''), 59.23 (T$^\alpha$), 57.13 (C-2''), 55.12 (CH_3-(PMP)), 47.41 (C-2), 46.78 (C-9-Fmoc), 27.61 (*t*Bu-Thr), 25.78, 25.62 (*t*Bu-TBS), 23.14, 20.60, 20.48, 20.41, 20.35, 20.27 (CH_3-(Ac)), 19.27 (T$^\gamma$), 17.95, 17.91 (C_q-TBS), -5.33, -5.37, -5.40, -5.44 (CH_3-(TBS)).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-_N_-acetamido-2-deoxy-4,6-_para_-methoxybenzylideneacetal-3-_O_-{2,4,6-_O_-acetyl-3-_O_-[4,6-_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]-β-D-galactopyranosyl}-α-D-galactopyranosyl)-L-threonine-_tert_-butylester (46)

(Fmoc-Thr(βAc_4Gal-(1→3)-βAc_2GlcNHTroc-(1→3)-βAc_3Gal-(1→3)-αPMP-GalNAc)-OtBu)



A solution of Fmoc-Thr(βAc_4Gal-(1→3)-βAc-TBS-GlcNHTroc-(1→3)-βTBS-Gal-(1→3)-βPMP-GalNAc)-O*t*Bu **45** (6.98 g, 3.84 mmol) in tetrahydrofuran (100 mL) was cooled in ice bath. A solution of tetra-n-butylammoniumfluoride trihydrate (14.54 g, 46.1 mmol) and acetic acid (7.69 mL, 8.07 g, 134.4 mmol) in tetrahydrofuran (40 mL) was added. After 1 h the ice bath was removed and the reaction stirred for 8 h until TLC (EtOAc/MeOH 25:1) indicated complete consumption. The reaction was diluted with ethyl acetate (600 mL) and the organic phase was washed with saturated bicarbonate solution and brine (two times, 600 mL, 1:1,) and with brine (once, 400 mL). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The residue was dissolved in pyridine (120 mL). Then *N,N*-(dimethylamino)pyridine (45 mg, 0.37 mmol) was added, followed by addition of acetic anhydride (60 mL) and the reaction was stirred for 18 h. The solvent was removed by co-evaporation with toluene *in vacuo*. The crude product was purified by flash column chromatography on silica (^{c}Hex/EtOAc 1:2).

Yield: 4.96 g (2.82 mmol, 73%), colorless, amorphous solid, $[\alpha]_D^{20}$ +45.66 (c = 1.00, $CHCl_3$), R_f = 0.34 (Tol/EtOAc 1:3).

$C_{78}H_{96}Cl_3N_3O_{36}$ (M = 1757.95 g/mol) [1755.48].

ESI-MS (pos), m/z: 1756.80 ($[M+H]^+$, calc. 1756.49), 1773.87 ($[M+NH_4]^+$, calc. 1773.52).

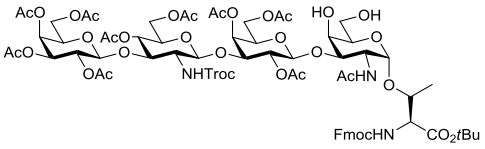
HR-ESI-MS (pos), m/z: 1756.4936 ($[M+H]^+$, calc. 1756.4917), 897.7234 ($[M+K+H]^{2+}$, calc. 897.7277).

^{1}H NMR (600 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm): ***δ*=** 7.92-7-90 (m, 2H, H-4-,H-5-Fmoc), 7.75 (m, 2H, H-1-, H-8-Fmoc), 7.70 (d, ^3J(NH,H-2'') = 9.3 Hz, 1H, NH-Troc), 7.44-7.41 (m, 4H, H-2-, H-7-Fmoc, NH-Fmoc, NH-Ac), 7.34-7.30 (m, 4H, H-3-, H-6-Fmoc, H-2-, H-6-PMP), 6.93 (d, ^3J(H-3,H-2) = ^3J(H-5,H-4) = 8.8 Hz, 2H, H-3-, H-5-PMP), 5.46 (s, 1H, CH-(PMP)), 5.30 (d, ^3J(H-4'',H-3'') = 3.9 Hz, 1H, H-4''), 5.23 (d, ^3J(H-4''',H-3''') = 3.7 Hz, 1H, H-4'''), 4.92 (dd, ^3J(H-3''',H-4''') = 3.6 Hz, ^3J(H-3''',H-2''') = 10.3 Hz, 1H, H-3'''), 4.86-4.80 (m, 2H, H-2', CH_{2a}-(Troc)), 4.76 (dd, ^3J(H-2''',H-1''') = 7.9 Hz, ^3J(H-2''',H-3''') = 10.3 Hz, H-2'''),

4.71 (d, ^3J(H-1,H-2) = 3.7 Hz, 1H, H-1), 4.65-4.60 (m, 3H, H-1', H-1''', H-4''), 4.55-4.44 (m, 4H, H-1'', CH_{2b}-(Troc), CH_2-(Fmoc)), 4.31 (t, ^3J(H-9,CH_2-(Fmoc)) = 6.6 Hz, 1H, H-9-Fmoc), 4.27-4.24 (m, 2H, H-5, T^{β}), 4.19 (dt, ^3J(H-2,H-1) = 3.5 Hz, ^3J(H-2,NH) = ^3J(H-2,H-3) = 10.4 Hz, 1H, H-2), 4.13-3.90 (m, 12H, H-5', H-3'', H-5''', H-6_{ab}, H-6'$_{ab}$, H-6''$_{ab}$, H-6'''$_{ab}$, T^{α}), 3.83 (dd, ^3J(H-3',H-2') = 10.3 Hz, ^3J(H-3',H-4') = 3.7 Hz, 1H, H-3'), 3.77-3.75 (m, 4H, H-3, CH_3-(PMP)), 3.68-3.63 (m, 2H, H-4, H-5''), 3.29-3.23 (q, ^3J(H-2'',H1'') = ^3J(H-2'',H-3'') = ^3J(H-2'',NH) = 9.2 Hz, 1H, H-2''), 2.09, 2.05, 2.04, 2.02, 2.01, 1.99, 1.97, 1.90, 1.88, 1.83 (m, 30H, CH_3-(Ac)), 1.36 (s, 9H, *t*Bu), 1.14 (d, $^3J(T^{\gamma},T^{\beta})$ = 6.4 Hz, 3H, T^{γ}).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(C)=39.51 ppm): ***δ*=** 170.08, 169.98, 169.88, 169.82, 169.31, 169.24, 169.18, 169.13, 168.73, 168.58 (C=O-(Ac), C=O-(*t*Bu)), 159.48 (CH-PMP), 156.73 (C=O-(Fmoc)), 153.80 (C=O-(Troc)), 143.70, 143.67 (C-1_a-, C-8_a-Fmoc), 140.81, 140.77 (C-4_a-, C-5_a-Fmoc), 130.74 (C-4-PMP), 127.71, 127.66 (C-2-, C-7-Fmoc), 127.42 (C-3-, C-6-Fmoc), 127.03, 127.02 (C-3-, C-5-PMP), 125.24, 125.12 (C-1-, C-8-Fmoc), 120.20, 120.15 (C-4-, C-5-Fmoc), 113.37 (C-2-, C-6-PMP), 101.38 (C-1'''), 101.14 (C-1''), 99.79 (C1'), 99.72 (CH-(PMP)), 99.01 (C-1), 95.54 (CCl_3-(Troc)), 81.44 (C_q-(*t*Bu)), 77.10 (C-3'), 76.54 (C-3''), 74.98 (T^{β}), 74.75 (C-3), 73.96 (CH_2-(Troc)), 73.41 (C-5), 70.55 (C-5'), 70.19 (C-5''), 70.14 (C-3'''), 69.74 (C-5'''), 69.48, 69.44 (C-2', C-4'), 68.50 (C-6, C-2'''), 68.29 (C-4''), 67.22 (C-4'''), 65.60 (CH_2-Fmoc), 62.88 (C-4), 62.25 (C-6'), 61.52 (C-6''), 60.93 (C-6'''), 59.28 (T^{α}), 56.95 (C-2''), 55.10 (CH_3-(PMP)), 46.91 (C-2), 46.77 (C-9-Fmoc), 27.58 (*t*Bu), 22.91, 20.76, 20.64, 20.60, 20.51, 20.47, 20.42, 20.34, 20.22 (CH_3-(Ac)), 19.28 (T^{γ}).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-_N_-acetamido-2-deoxy-3-_O_-{2,4,6-_O_-acetyl-3-_O_-[4,6-_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]-β-D-galactopyranosyl}-α-D-galactopyranosyl)-L-threonine-*tert*-butylester (47)
(Fmoc-Thr(βAc_4Gal-(1→3)-βAc_2GlcNHTroc-(1→3)-Ac_3Gal-(1→3)-α-GalNAc)-OtBu)



A solution of Fmoc-Thr(βAc_4Gal-(1→3)-βAc_2GlcNHTroc-(1→3)-βAc_3Gal-(1→3)-βPMP-GalNAc)-O*t*Bu **46** (4.84 g, 2.75 mmol) in 80% acetic acid (60 mL) was stirred for 2 h at 40°C. The solvent was removed by co-evaporation with toluene *in vacuo*. The crude product was purified by flash column chromatography on silica (^{C}Hex/EtOAc 1:3→EtOAc).

Yield: 4.05 g (2.47 mmol, 90%), colorless, amorphous solid, $[\alpha]_D^{20}$ +31.22 (c = 0.49, $CHCl_3$), R_f = 0.19 (EtOAc).

$C_{70}H_{90}Cl_3N_3O_{35}$ (M = 1639.82 g/mol) [1637.44].

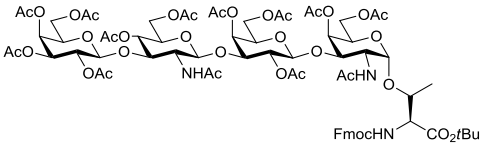
ESI-MS (pos), m/z: 1638.00 ($[M+H]^+$, calc. 1638.45).

HR-ESI-MS (pos), m/z: 1638.4506 ($[M+H]^+$, calc. 1638.4499), 838.7021 ($[M+K+H]^{2+}$, calc. 838.7068).

^{1}H NMR (600 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm): ***δ*=** 7.92-7.90 (m, 2H, C-4-, C-5-Fmoc), 7.76-7.71 (m, 3H, C-1-, C-8-Fmoc, NH-Troc), 7.44-7.41 (m, 4H, C-2-, C-7-Fmoc, NH-Ac, NH-Fmoc), 7.33-7.31 (m, 2H, C-3-, C-6-Fmoc), 5.29 (d, ^3J(H-4',H-3') = 3.8 Hz, 1H, H-4'), 5.24 (d, ^3J(H-4''',H-3''') = 3.7 Hz, 1H, H-4'''), 4.93 (dd, ^3J(H-3''',H-2''') = 10.5 Hz, J(H-3''',H-4''') = 3.6 Hz, 1H, H-3'''), 4.90-4.85 (m, 2H, H-2',CH_{2a}-(Troc)), 4.77 (dd, ^3J(H-2''',H-1''') = 7.9 Hz, ^3J(H-2''',H-3''') = 10.4 Hz, 1H, H-2'''), 4.66-4.55 (m, 5H, H-1, H-1', H-1''', H-4''', CH_{2b}-(Troc)), 4.51-4.42 (m, 3H, H-1'', CH_2-(Fmoc)), 4.36 (d, ^3J(OH-4,H-4) = 5.4 Hz, 1H, OH-4), 4.32 (t, ^3J(H-9,CH_{2ab}) = 6.8 Hz, 1H, H-9-Fmoc), 4.24-4.20 (m, 2H, H-2, T^{β}), 4.14-3.87 (m, 11H, H-3'', H-4, H-5', H-5''', H-6'$_{ab}$, H-6''$_{ab}$, H-6'''$_{ab}$, T^{α}), 3.83 (dd, ^3J(H-3',H-4') = 3.3 Hz, ^3J(H-3',H-2') = 9.9 Hz, 1H, H-3'), 3.67-3.61 (m, 2H, H5, H5''), 3.54 (dd, 1H, H3, $J_{H3,H4}$ = 2.2 Hz, $J_{H3,H2}$ = 11.0 Hz), 3.48-3.46 (m, 2H, H-6), 3.28-3.26 (m, 1H, H-2''), 2.09, 2.06, 2.04, 2.01, 1.99, 1.98, 1.88, 1.85 (m, 30H, CH_3-(Ac)), 1.35 (s, 9H, *t*Bu), 1.16-1.15 (m, 3H, T^{γ}).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(C)=39.51 ppm): ***δ*=** 170.08, 169.97, 169.88, 169.82, 169.65, 169.32, 169.23, 169.13, 168.11, 168.64 (C=O-(Ac), (C=O)-*t*Bu), 156.77 (C=O-(Fmoc)), 153.81 (C=O-(Troc)), 143.71, 143.65 (C-1_a-, C-8_a-Fmoc), 140.81, 140.76 (C4_a-, C5_a-Fmoc), 127.71, 126.66 (C-2-, C-7-Fmoc), 127.04, 127.01 (C-3-, C-6-Fmoc), 125.29, 125.15 (C-1-, C-8-Fmoc), 120.21, 120.15 (C-4-, C-5-Fmoc), 101.39 (C-1'), 101.14 (C-1''), 99.77 (C-1'''), 98.91 (C-1), 95.54 (CCl_3-Troc), 81.33 (C_q-*t*Bu), 77.44 (C-3'), 77.23 (C-3), 76.59 (C-3''), 73.99 (CH_2-(Troc)), 73.48 (T^{β}), 71.67 (C-5), 70.58 (C-5'), 70.21 (C-5''), 70.15 (C-3'''), 69.75 (C-5'''), 69.40 (C-2', C-4'), 68.52 (C-2'''), 68.25 (C-4''), 67.43 (C-4), 67.22 (C-4'''), 65.58 (CH_{2a}-(Fmoc)), 62.05 (C-6'''), 61.54 (C-6''), 60.92 (C-6'), 60.53 (C-6), 59.35 (T^{α}), 56.93 (C-2''), 47.04 (C-2), 46.77 (C-9-Fmoc), 27.59 (*t*Bu), 22.94, 20.76, 20.62, 20.53, 20.46, 20.43, 20.34, 20.24 (m, 30H, CH_3-(Ac)), 19.05 (T^{γ}).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-_N_-acetamido-4,6-_O_-acetyl-2-deoxy-3-_O_-{2,4,6-_O_-acetyl-3-_O_-[2-_N_-acetamido-4,6-_O_-acetyl-2-deoxy-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-β-D-galactopyranosyl}-α-D-galactopyranosyl)-L-threonine-_tert_-butylester (48)
(Fmoc-Thr(βAc$_4$Gal-(1→3)-βAc$_2$GlcNAc-(1→3)-Ac$_3$Gal-(1→3)-αAc$_2$GalNAc)-O*t*Bu)



Zinc dust was activated by treatment with 1 N hydrochloric acid for a few minutes followed by washing with water, methanol and diethyl ether. Fmoc-Thr(βAc$_4$Gal-(1→3)-βAc$_2$GlcNHTroc-(1→3)-βAc$_3$Gal-(1→3)-βGalNAc)-O*t*Bu **47** (1.36 g, 0.83 mmol) was dissolved in acetic acid (20 mL) and pre-activated zinc (1.09 g, 16.6 mmol) was added. The reaction was stirred at 40°C for 18 h. The zinc dust was filtered off and washed with acetic acid. The filtrate was co-evaporated with toluene *in vacuo*. The residue was dissolved in pyridine/acetic anhydride (30 mL, 2:1). Then 4-(dimethylamino)pyridine (10.1 mg, 0.08 mmol) was added and the solution stirred for 20 h. The solvent was removed by co-evaporation with toluene *in vacuo* and the crude product purified by flash column chromatography on silica (100% EtOAc).

Yield: 1.23 g (0.77 mmol, 93%), colorless, amorphous solid, $[\alpha]_D^{20}$ +49.60 (c = 0.50, $CHCl_3$), R_f = 0.26 (EtOAc).

$C_{73}H_{95}N_3O_{36}$ (M = 1590.53 g/mol) [1589.57].

ESI-MS (pos), m/z: 1590.13 ($[M+H]^+$, calc. 1590.58).

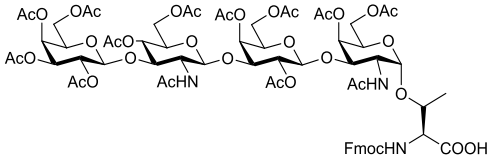
HR-ESI-MS (pos), m/z: 1590.5772 ($[M+H]^+$, calc. 1590.5774).

^{1}H NMR (600 MHz, [D$_6$]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm): ***δ*=** 7.92-7.90 (m, 2H, C-4-, C-5-Fmoc), 7.76-7.73 (m, 3H, C-1-, C-8-Fmoc), 7.69 (d, ^3J(NH'',H2'') = 9.2 Hz, 1H, NH''-Ac), 7.50 (d, ^3J(NH,H2) = 9.8 Hz, 1H, NH-Ac), 7.44-7.40 (m, 3H, C-2-, C-7-Fmoc, NH-Fmoc), 7.34-7.30 (m, 2H, C-3-, C-6-Fmoc), 5.26 (m, 2H, H-4, H-4'), 5.23 (d, ^3J(H-4''',H-3''') = 3.7 Hz, 1H, H-4'''), 5.08 (dd, ^3J(H-3''', H-2''') = 10.3 Hz, ^3J(H-3''',H-4''') = 3.5 Hz, 1H, H-3'''), 4.77-4.71 (m, 2H, H-2', H-2'''), 4.68 (d, ^3J(H-1,H-2) = 4.0 Hz, 1H, H-1), 4.66 (d, ^3J(H-1''',H-2''') = 8.0 Hz, 1H, H-1'''), 4.62 (t, ^3J(H-4'',H-3'') = ^3J(H-4'',H-2'') = 9.7 Hz, 1H, H-4''), 4.58 (d, ^3J(H-1',H-2') = 7.9 Hz, 1H, H-1'), 4.52-4.43 (m, 3H, CH_2-(Fmoc), H-1''), 4.31 (t, ^3J(H-9,CH_{2ab}) = 6.7 Hz, 1H, H-9-Fmoc), 4.22-4.15 (m, 2H, H-2, T$^\beta$), 4.11-3.84 (m, 11H, H-5, H-5', H-5''', H-6, H-6', H-6'', H-6''', T$^\alpha$), 3.82-3.79 (m, 3H, H-3, H-3', H-3''), 3.70-3.68 (m, 1H, H-5''), 3.56-3.47 (m, 1H, H-2''), 2.09, 2.05, 2.02, 2.01, 2.00, 1.99, 1.98, 1.98, 1.89, 1.84, 1.82 (m, 39H, CH_3-(Ac)), 1.35 (s, 9H, *t*Bu), 1.16 (d, ^3J(T$^\gamma$,T$^\beta$) = 6.7 Hz, 3H, T$^\gamma$).

^{13}C NMR (150.9 MHz, [D$_6$]DMSO, 30°C, internal DMSO *δ*(C)=39.51 ppm): ***δ*=** 170.08, 170.07, 169.95, 169.90, 169.88, 169.70, 169.52, 169.48, 169.37, 169.14, 169.08, 168.93,

168.79, 168.72 (C=O-(Ac), C=O-(*t*Bu)), 156.74 (C=O-(Fmoc)), 143.71, 143.64 (C-1_a-, C-8_a-Fmoc), 140.80, 140.77 (C-4_a-, C-5_a-Fmoc), 127.72, 127.67 (C-2-, C-7-Fmoc), 127.04, 127.01 (C-3, C-6-Fmoc), 125.22, 125.10 (C-1-, C-8-Fmoc), 120.21, 120.16 (C-4-, C-5-Fmoc), 100.67 (C-1''), 100.45 (C-1'), 99.90 (C-1'''), 98.69 (C-1), 81.48 (C_q-*t*Bu), 77.15 (C-3''), 76.40 (C-3'), 74.02 (T^{β}), 73.01 (C-3), 70.46 (C-3'''), 70.28 (C-5), 70.24 (C-5''), 69.72 (C-4', C-2'), 69.47 (C-5'''), 68.99 (C-4), 68.55 (C-2'''), 68.46 (C-4''), 67.27 (C-5'), 67.15 (C-4'''), 65.61 (CH_2-(Fmoc)), 63.03, 61.67, 61.46, 61.10 (C-6, C-6', C-6'', C-6'''), 59.30 (T^{α}), 54.20 (C-2''), 47.62 (C-2), 46.75 (C-9-Fmoc), 27.58 (CH_3-(*t*Bu)), 22.82, 20.66, 20.58, 20.50, 20.47, 20.37, 20.33, 20.29 (CH_3-(Ac)), 18.74 (T^{γ}).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-_N_-acetamido-4,6-_O_-acetyl-2-deoxy-3-_O_-{2,4,6-_O_-acetyl-3-_O_-[2-_N_-acetamido-4,6-_O_-acetyl-2-deoxy-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-β-D-galactopyranosyl}-α-D-galactopyranosyl)-L-threonine (49)
(Fmoc-Thr(βAc$_4$Gal-(1→3)-βAc$_2$GlcNAc-(1→3)-Ac$_3$Gal-(1→3)-αAc$_2$GalNAc)-OH)



Fmoc-Thr(βAc$_4$Gal-(1→3)-βAc$_2$GlcNAc-(1→3)-βAc$_3$Gal-(1→3)-βAc$_2$GalNAc)-O*t*Bu **48** (1.20 g, 0.75 mmol) was dissolved in dichloromethane (5 mL) and anisole (0.5 mL). Trifluoroacetic acid (15 mL) was added and the reaction was stirred for 90 min. The solvent was removed by co-evaporation with toluene. The crude product was purified by flash column chromatography on silica (EtOAc→EtOAc/MeOH 25:1).

Yield: 933 mg (0.61 mmol, 81%), colorless, amorphous solid, $[\alpha]_D^{20}$ +57.92 (c = 1.01, $CHCl_3$), R_f = 0.24 (EtOAc/MeOH/AcOH/H_2O 50:3:3:2).

$C_{69}H_{87}N_3O_{36}$ (M = 1534.43 g/mol) [1533.51].

MALDI-TOF-MS (dhb, pos), m/z: 1556.5294 ($[M+Na]^+$, calc. 1556.4967), 1572.4962 ($[M+K]^+$, calc. 1572.4706).

HR-ESI-MS (pos), m/z: 1534.5143 ($[M+H]^+$, calc. 1534.5148).

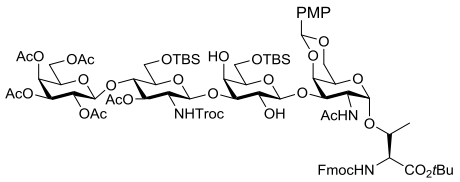
^{1}H NMR (600 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm): ***δ*=** 12.96 (CO_2H), 7.92-7.90 (m, 2H, H-4-, H-5-Fmoc), 7.76-7.71 (m, 3H, H-1-, H-8-Fmoc, NH''-Ac), 7.47-7.30 (m, 6H, H-2-, H-3-, H-5-, H-6-Fmoc, NH-Ac, NH-Fmoc), 5.27-5.25 (m, 2H, H-4, H-4'), 5.23 (d, ^3J(H-4''',H-3''') = 3.6 Hz, 1H, H-4'''), 5.07 (dd, ^3J(H-3''',H-4''') = 3.6 Hz, ^3J(H-3''',H-2''') = 10.3 Hz, 1H, H-3'''), 4.77-4.73 (m, 2H, H-2', H-2'''), 4.71 (d, ^3J(H1,H2) = 3.9 Hz, 1H, H-1), 4.67 (d, ^3J(H-1''',H-2''') = 7.9 Hz, 1H, H1'''), 4.64-4.58 (m, 2H, H-1', H-4''), 4.52-4.43 (m, 3H, H-1'',

CH_2-Fmoc), 4.31 (t, ^3J(H-9,CH_2) = 6.7 Hz, 1H, H-9-Fmoc), 4.27-4.25 (m, 1H, T^{β}), 4.14-4.03 (m, 7H, H-2, H-5', H-5''', H-6'$_a$, H-6''$_a$, H-6'''$_a$, T^{α}), 4.01-3.96 (m, 2H, H-6''$_a$, H-6'''$_a$), 3.93-3.78 (m, 7H, H-3, H-3', H-3'', H-5, H-6$_{ab}$, H-6'$_b$), 3.69-3.67 (m, 1H, H-5''), 3.54-3.52 (m, 1H, H-2''), 2.09, 2.04, 2.02, 2.01, 2.00, 1.99, 1.97, 1.89, 1.84, 1.81 (s, 39H, CH_3-(Ac)), 1.14 (d, ^3J(T^{γ},T^{β}) = 6.3 Hz, 3H, T^{γ}).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO δ(C)=39.51 ppm): δ= 171.71, 170.10 170.00, 169.92, 169.73, 169.55, 169.51, 169.51, 169.39, 169.16, 169.02, 168.95, 168.69 (C=O-(Ac)), 156.78 (C=O-(Fmoc)), 143.76, 143.71 (C-1$_a$-, C-8$_a$-Fmoc), 140.77, 140.81, 140.76 (C-4$_a$-, C-5$_a$-Fmoc), 127.72, 127.66 (C-2-, C-7-Fmoc), 127.05 (C-3-, C-6-Fmoc), 125.26, 125.13 (C-1-, C-8-Fmoc), 120.21, 120.16 (C-4-, C-5-Fmoc), 100.67 (C-1''), 100.56 (C-1'), 99.91 (C-1'''), 98.75 (C-1), 77.19 (C-3''), 79.46 (C-3'), 74.44 (T^{β}), 73.29 (C-3), 70.48 (C-3'''), 70.30 (C-5''), 70.23 (C-5), 69.64 (C-2', C-4'), 69.48 (C-5'''), 69.00 (C-4), 68.56 (C-2'''), 68.46 (C-4''), 67.26 (C-5'), 67.16 (C-4'''), 65.57 (CH_2-(Fmoc)), 63.05 (C-6'), 61.77 (C-6), 61.48 (C-6''), 61.11 (C-6'''), 58.45 (T^{α}), 54.20 (C-2''), 47.77 (C-2), 46.77 (C-9-Fmoc), 22.83, 22.79, 20.66, 20.62, 20.58, 20.51, 20.39, 20.35, 20.31 (CH_3-(Ac)), 18.43 (T^{γ}).

7.1.10 Synthesis of the extended type-2 core 1 amino

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-acetamido-2-deoxy-4,6-_O_-_para_-methoxybenzylidenacetal-3-_O_-{6-_tert_-butyldimethylsilyl-3-_O_-[3-_O_-acetyl-2-deoxy-6-_tert_-butyldimethyldimethylsilyl-2-_N_-(2,2,2-trichloroethoxycarbonyl)-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-β-D-galactopyranosyl}-α-D-galactopyranosyl)-L-threonine-_tert_-butylester (50)
(Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc-TBS-GlcNHTroc-(1→3)-βTBS-Gal-(1→3)-αPMP-GalNAc)-O_t_Bu)



Fmoc-Thr(βTBS-Gal-(1→3)-βPMP-GalNAc)-O*t*Bu **39** (3.98 g, 4.00 mmol) and βAc$_4$Gal-(1→4)-βAc-TBS-GlcNHTroc-SPh **31** (4.88 g, 5.23 mmol) were dissolved in dry dichloromethane (60 mL) under argon atmosphere. Activated molecular sieves (5 g, 4Å) were added. The suspension was stirred for 1 h before it was cooled in an ice bath. *N*-iodosuccinimide (1.18 g, 5.23 mmol) was added followed by the dropwise addition of a suspension of trifluoromethanesulfonic acid (53 µL, 91 mg, 0.61 mmol) in dry dichloromethane (1 mL) *via* a syringe. After 2.5 h the reaction was treated with another

addition of βAc$_4$Gal-(1→4)-βAc-TBS-GlcNHTroc-SPh **31** (1.12 g, 1.20 mmol) and *N*-iodosuccinimide (270 mg, 1.2 mmol). The reaction was stirred further for 1.5 h under argon atmosphere and ice-cooling. The reaction was diluted with dichloromethane (40 mL). The molecular sieves were filtered off and washed with dichloromethane (total volume of 200 mL). The dichloromethane phase was washed with 0.5 M sodium thiosulfate solution, saturated sodium bicarbonate solution, water and brine (80 mL each). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography on silica (^{C}Hex/EtOAc 2:1→3:2→1:1).

Yield: 4.71 g (2.59 mmol, 65%), colorless, amorphous solid, $[\alpha]_D^{20}$ +42.22 (c = 0.99, $CHCl_3$), R_f = 0.45 (^{C}Hex/EtOAc 2:3).

$C_{82}H_{116}Cl_3N_3O_{32}Si_2$ (M = 1818.33 g/mol) [1815.61].

ESI-MS (pos), m/z: 1817.00 ($[M+H]^+$, calc. 1816.62). 1833.93 ($[M+NH_4]^+$, calc. 1833.65), 1838.47 ($[M+Na]^+$, calc. 1838.60).

HR-ESI-MS (pos), m/z: 1816.6224 ($[M+H]^+$, calc. 1816.6224), 1702.5239 ($[M-TBS+H]^+$, calc. 1702.5360).

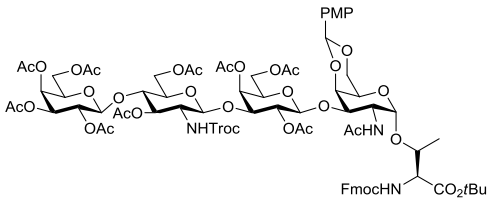
^{1}H NMR (600 MHz, [D$_6$]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm): ***δ*=** 7.90-7.88 (m, 2H, H-4-, H-5-Fmoc), 7.76-7.73 (m, 2H, H-1-, H-8-Fmoc), 7.63 (d, ^3J(N-H,H-2'') **=** 8.3 Hz, 1H, N-H-Troc,), 7.56 (d, ^3J(N-H,T$^\alpha$) = 9.8 Hz, 1H, N-H-Fmoc), 7.47 (d, ^3J(N-H,H-2) = 9.7 Hz, 1H, N-H-Ac), 7.42 (t, ^3J(H-2,H-1) = ^3J(H-2,H-3) = ^3J(H-7,H-6) = ^3J(H-7,H-8) = 7.4 Hz, 2H, H-2-, H-7-Fmoc), 7.37 (d, ^3J(H-2,H-3) = ^3J(H-6,H-5) = 8.8 Hz, 2H, H-2-, H-6-PMP), 7.31 (t, ^3J(H-3,H-4) = ^3J(H-3,H-2) = ^3J(H-6,H-7) = ^3J(H-6,H-5) = 7.4 Hz, 2H, H-3-, H-6-Fmoc), 6.93 (d, ^3J(H-3,H-2) = ^3J(H-5,H-6) = 8.8 Hz, 2H, H-3-, H-5-PMP), 5.52 (s, 1H, CH-(PMP)), 5.24 (d, ^3J(H-4''',H-3''') = 3.6 Hz, 1H, H-4'''), 5.11 (dd, ^3J(H-3''',H-4''') = 3.5 Hz, ^3J(H-3''',H-2''') = 10.3 Hz, 1H, H-3'''), 4.98 (d, ^3J(H-1'',H-2'') = 8.5 Hz, 1H, H-1''), 4.93 (t, ^3J(H-3'',H-2'') = ^3J(H-3'',H-4'') = 6.6 Hz, 1H, H-3''), 4.89 (d, ^2J(CH$_{2a}$,CH$_{2b}$) = 12.4 Hz, 1H, CH$_{2a}$-(Troc)), 4.83 (m, 1H, H-2'''), 4.77 (d, ^3J(H-1,H-2) = 3.3 Hz, 1H, H-1), 4.72 (d, ^3J(H-1''',H-2''') = 7.9 Hz 1H, H-1'''), 4.68-4.65 (m, 2H, CH$_{2b}$-(Troc), OH-2'), 4.48-4.40 (m, 3H, CH$_{2ab}$-(Fmoc), OH-4'), 4.36-4.26 (m, 5H, H-2, H-5, H-1', H-9-Fmoc, T$^\beta$), 4.15-4.10 (m, 2H, H-5''', T$^\alpha$), 4.06-3.97 (m, 4H, H-6$_{ab}$, H-6$_{ab}$'''), 3.83-3.62 (m, 11H, H-3, H-4, H-4', H-4'', H-6$_{ab}$', H-6$_{ab}$'', CH$_3$-PMP), 3.45-3.41 (m, 3H, H-2', H-2'', H-5''), 3.36-3.30 (m, 2H, H-5', H-3'), 2.09-1.86 (m, 18H, CH$_3$-(Ac)), 1.39 (s, 9H, *t*Bu-(Thr)), 1.12 (d, ^3J(T$^\gamma$,T$^\beta$) = 6.2 Hz, 3H, T$^\gamma$), 0.87, 0.83 (each s, 9H, *t*Bu-(TBS)), 0.05, 0.04, 0.03 (each s, 3H, Me(TBS)).

^{13}C NMR (150.9 MHz, [D$_6$]DMSO, 30°C, internal DMSO *δ*(C)=39.52 ppm): ***δ*=** 169.85, 169.46, 169.34, 169.33, 169.01 (C=O-(Ac), C=O-(*t*Bu)), 159.50 (C-1-PMP), 156.69 (C=O-(Fmoc)), 154.39 (C=O-(Troc)), 143.83, 143.53 (C-1$_a$-, C-8$_a$-Fmoc), 140.77 (C-4$_a$-, C-5$_a$-Fmoc), 130.82 (C-4-PMP), 127.62 (C-2-, C-7-Fmoc, C-2-, C-6-PMP), 126.99 (C-3-, C-6-Fmoc), 125.21 (C-1-, C-8-Fmoc), 120.16, 120.10 (C-4-, C-5-Fmoc), 113.29 (C-3-, C-5-Pmb),

105.39 (C-1'), 100.26 (C-1''), 2 x 99.75, 99.67 (C-1, C-1''', CH-(PMP)), 96.17 (C_{quart}-(*t*Bu-(TBS)), 81.39 (C_{quart}-(*t*Bu-(Thr))), 80.43 (C-5''), 75.78 (T^{β}), 75.61 (C-4''), 75.08, 75.03, 74.99 (C-3, C-3', C-5'), 74.16 (C-5), 73.50 (C-3''), 73.39 (CH_{2ab}-(Troc)), 70.27 (C-3'''), 69.94 (C-2'), 69.77 (C-5''), 69.04 (C-2'''), 68.42 (C-6_{ab}), 67.13 (C-4'), 67.03 (C-4'''), 65.59 (CH_{2ab}-(Fmoc)), 63.10 (C-4), 62.09 (C-6'$_{ab}$), 61.58 (C-6''$_{ab}$), 60.89 (C-6'''$_{ab}$), 59.25 (T^{α}), 55.82 (C-2''), 55.12 (CH_3-(PMP)), 47.28, 46.79 (C-2, C-9-(Fmoc)), 27.61 (*t*Bu-(Thr)), 25.79, 25.62 (2 *t*Bu-(TBS)), 23.22, 20.59, 2 x 20.45, 20.37, 20.33 (6 CH_3-(Ac)), 19.31 (T^{γ}), 17.96, 17.90 (2 C_q(TBS)), -5.25, -5.34, -5.42 (4 Me(TBS)).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-acetamido-2-deoxy-4,6-_O_-_para_-methoxybenzylidene-3-_O_-{2,4,6-tri-_O_-acetyl-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]-β-D-galactosylpyranosyl}-α-D-galactopyranosyl)-L-threonine-_tert_-butylester (51)

(Fmoc-Thr(βAc_4Gal-(1→4)-βAc_2GlcNHTroc-(1→3)-βAc_3Gal-(1→3)-αPMP-GalNAc)-O*t*Bu)



A solution of Fmoc-Thr(βAc_4Gal-(1→4)-βAc-TBS-GlcNHTroc-(1→3)-βTBS-Gal-(1→3)-βPMP-GalNAc)-O*t*Bu (3.80 g, 2.09 mmol) **50** in tetrahydrofuran (60 mL) was cooled in ice bath. A solution of tetra butylammoniumfluoride trihydrate (6.59 g, 20.9 mmol) and acetic acid (2.39 mL, 41.8 mmol) in tetrahydrofuran (20 mL) was added. After 1 h the ice bath was removed and the reaction was further stirred for 16 h. The reaction was diluted with ethyl acetate (400 mL) and the organic phase was washed with saturated sodium bicarbonate solution/brine (two times, 450 mL each, 1:1) and once with brine (200 mL). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The residue was dissolved in pyridine (60 mL). Then 4-(dimethylamino)pyridine (25 mg, 0.20 mmol) was added followed by acetic anhydride (30 mL) and the reaction was stirred for 24 h. The solvent was removed by co-evaporation with toluene *in vacuo*. The crude was purified by flash column chromatography on silica (^{C}Hex/EtOAc 1:2).

Yield: 2.50 g (1.44 mmol, 68%), colorless, amorphous solid, $[\alpha]_D^{20}$ +56.94 (c = 0.49, $CHCl_3$), R_f = 0.43 (Tol/EtOAc 1:3).

$C_{78}H_{96}Cl_3N_3O_{36}$ (M = 1757.95 g/mol) [1755.48].

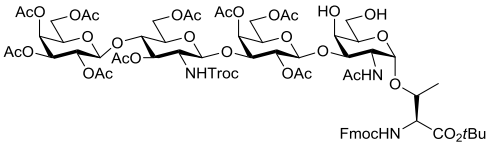
ESI-MS (pos), m/z: 1756.93 ($[M+H]^+$, calc. 1756.49), 1773.73 ($[M+NH_4]^+$, calc. 1773.52).

HR-ESI-MS (pos), m/z: 1756.4922 ($[M+H]^+$, calc. 1756.4917), 898.7227 ($[M+K+H]^{2+}$, calc. 897.7277).

^{1}H NMR (600 MHz, [D_6]DMSO, 30°C, internal DMSO δ(H)= 2.50 ppm): **δ=** 7.92-7.90 (m, 2H, H-4-, H-5-Fmoc), 7.81 (d, ^3J(N-H,H-2) = 9.1 Hz, 1H, N-H-Troc), 7.77-7.34 (m, 1H, H-1, H-8-Fmoc), 7.43-7.41 (m, 4H, H-2-, H-7-Fmoc, N-H-Ac, N-H-Fmoc), 7.34-7.30 (m, 4H, H-2-, H-6-PMP, H-3-, H-6-Fmoc), 6.93 (d, 2H, H-3-, H-5-PMP, ^3J(H-3,H-2) = ^3J(H-5,H-4) = 8.8 Hz), 5.44 (s, 1H, CH-(PMP)), 5.33 (m, 1H, H-4'), 5.22 (d, ^3J(H-4''',H-3''') = 3.2 Hz, 1H, H-4'''), 5.16 (dd, ^3J(H-3''',H-2''') = 10.3 Hz, ^3J(H-3''',H-4''') = 3.3 Hz, 1H, H-3'''), 4.98-4.97 (m, 2H, H-3'', CH_{2a}-(Troc)), 4.84-4.80 (m, 2H, H-2', H-2'''), 4.71-4.69 (m, 2H, H-1, H-1'''), 4.63-4.61 (m, 2H, H-1', H-1''), 4.53-4.45 (m, 3H, CH_{2ab}-(Fmoc), CH_{2b}-(Troc)), 4.32 (t, ^3J(H-9,CH_{2ab}) = 6.7 Hz, 1H, H-9-Fmoc), 4.28-4.26 (m, 3H, H-5, H-6_a'', T^{α}), 4.23-4.17 (m, 2H, H-2, H-5'''), 4.09-3.99 (m, 6H, H-6_{ab}, H-6_a', H-6_{ab}''', T^{α}), 3.93-3.89 (m, 3H, H-5', H-6_b', H-6_b''), 3.97 (dd, ^3J(H-3',H-4') = 4.0 Hz, ^3J(H-3',H-2') = 10.2 Hz, 1H, H-3'), 3.75-3.72 (m, 4H, H-3, CH_3-PMP), 3.70-3.65 (m, 2H, H-4, H-4''), 3.54-3.51 (m, 1H, H-5''), 3.29 (q, ^3J(H-2'',H-1'') = ^3J(H-2'',N-H'') = ^3J(H-2'',H-3'') = 9.1 Hz 1H, H-2''), 2.09-1.83 (m, 30H, CH_3-(Ac)), 1.36 (s, 9H, *t*Bu), 1.15 (d, ^3J(T^{γ},T^{β}) = 6.2 Hz, 3H, T^{γ}).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO δ(C)=39.52 ppm): **δ=** 170.30, 170.02, 169.87, 169.54, 169.50, 169.24, 169.19, 169.09, 168.71 (C=O-(Ac), C=O-(*t*Bu)), 159.51 (C1-PMP), 156.78 (C=O-(Fmoc)), 153.99 (C=O-(Troc)), 143.73, 143.70 (C-1_a-, C-8_a-Fmoc), 140.83, 140.79 (C-4_a-, C-5_a-Fmoc), 130.75 (C-4-PMP), 127.67 (C-2-, C-6-PMP), 127.04 (C-3-, C-6-Fmoc, C-2-, C-7-Fmoc), 125.25, 125.14 (C-1-, C-8-Fmoc), 120.21, 120,16 (C-4-, C-5-Fmoc), 113.39 (C-3-, C-5-PMP), 101.46 (C-1'), 100.59 (C-1''), 100.02 (C-1'''), 99.83 (CH-(PMP)), 99.06 (C-1), 96.08 (C_{quart}-(Troc)), 81.46 (C_{quart}-(*t*Bu)), 78.13 (C-3'), 75.98 (C-4''), 75.05, 75.00 (T^{β}, C-3), 73.40 (C-5, CH_{2ab}-(Troc)), 72.75 (C-3''), 71.51 (C-5''), 70.67 (C-5), 70.27 (C-3'''), 69.61 (C-5''), 69.42 (C-4'), 69.05, 68.92 (C-2', C-2''), 68.59 (C-6_{ab}), 67.04 (C-4'''), 65.61 (CH_{2ab}-(Fmoc)), 62.89 (C-4), 62.47 (C-6_{ab}'), 61.79 (C-6_{ab}''), 60.81 (C-6_{ab}'''), 59.29 (T^{α}), 55.78 (C-2''), 55.11 (CH_3-(PMP)), 46.86, 46.79 (C-2, C-9-Fmoc), 27.59 (*t*Bu), 22.92, 20.76, 20.69, 20.64, 20.49, 20.47, 20.46, 20.36, 20.33, 20.31 (CH_3-(Ac)), 19.32 (T^{γ}).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-acetamido-2-deoxy-3-_O_-{2,4,6-tri-_O_-acetyl-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]-β-D-galactopyranosyl}-α-D-galactopyranosyl)-L-threonine-_tert_-butylester (52)
(Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc$_2$GlcNHTroc-(1→3)-βAc$_3$Gal-(1→3)-αGalNAc)-O_t_Bu)



A solution of Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc$_2$GlcNHTroc-(1→3)-βAc$_3$Gal-(1→3)-βPMP-GalNAc)-O*t*Bu **51** (2.45 g, 1.39 mmol) in 80% acetic acid (50 mL) was stirred for 1 h at 40°C. The solvent was removed by co-evaporation with toluene *in vacuo*. The crude product was purified by flash column chromatography on silica (100% EtOAc).

Yield: 2.05 g (1.25 mmol, 90%), colorless, amorphous solid, $[\alpha]_D^{20}$+36.86 (c = 0.51, $CHCl_3$), R_f = 0.27 (EtOAc).

$C_{70}H_{90}Cl_3N_3O_{35}$ (M = 1639.82 g/mol) [1637.44].

ESI-MS (pos), m/z: 1638.93 ($[M+H]^+$, calc. 1638.45).

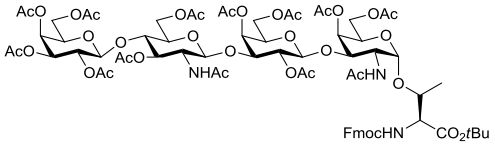
HR-ESI-MS (pos), m/z: 1638.4480 ($[M+H]^+$, calc. 1638.4499).

^{1}H NMR (600 MHz, [D$_6$]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm): ***δ*=** 7.92-7.90 (m, 2H, H-4-, H-5-Fmoc), 7.81 (d, ^3J(N-H,H-2'') = 9.1 Hz, 1H, N-H-Troc), 7.75 (t, ^3J(H-1,H-2) = ^3J(H-8,H-7) = 7.7 Hz, 2H, H-1-, H-8-Fmoc), 7.46-7.41 (m, 4H, H-2-, H-7-Fmoc, N-H-Fmoc, N-H-Ac), 7.34-7.29 (m, 2H, H-3-, H-6-Fmoc), 5.31 (d, ^3J(H-4',H-3') = 3.5 Hz, 1H, H-4'), 5.21 (d, ^3J(H-4''',H-3''') = 3.6 Hz, 1H, H-4'''), 5.16 (dd, ^3J(H-3''',H-2''') = 10.3 Hz, ^3J(H-3''',H-4''') = 3.5 Hz, 1H, H-3'''), 4.99-4.97 (m, 2H, H-3'', CH_{2a}-(Troc)), 4.86 (t, ^3J(H-2',H-3') = ^3J(H-2',H-1') = 8.2 Hz, 1H, H-2'), 4.82 (dd, ^3J(H-2''',H-3''') = 10.3 Hz, ^3J(H-2''',H-1''') = 7.9 Hz, 1H, H-2'''), 4.70-4.58 (m, 4H, H-1, H-1', H-1'', H-1'''), 4.52-4.49 (m, 2H, CH_{2b}-(Troc), CH_{2a}-(Fmoc)), 4.46-4.30 (m, 1H, CH_{2b}-(Fmoc)), 4.39 (d, ^3J(OH-4,H-4) = 5.4 Hz, 1H, OH-4,), 4.32 (t, ^3J(H-9,CH_{2ab}) = 6.7 Hz, 1H, H-9-Fmoc), 4.26-4.20 (m, 4H, H-2, H-5''', H-6$_a$'', T$^\beta$), 4.08-4.06 (m, 1H, T$^\alpha$), 4.01-3.91 (m, 6H, H-5', H-6$_{ab}$', H-6$_b$'', H-6$_{ab}$'''), 3.89-3.87 (m, 1H, H-4), 3.80-3.78 (m, 1H, H-3'), 3.68 (t, ^3J(H-4'',H-3'') = ^3J(H-4'',H-5'') = 9.5 Hz 1H, H-4''), 3.63-3.61 (m, 1H, H-5), 3.55-3.82 (m, 2H, H-3, H-5''), 3.48-3.46 (m, 1H, H-6$_{ab}$), 3.33-3.29 (m, 2H, OH-6, H-2''), 2.09- 1.85 (m, 30H, CH_3-(Ac)), 1.35 (s, 9H, *t*Bu), 1.16-1.15 (m, 3H, T$^\gamma$).

^{13}C NMR (150.9 MHz, [D$_6$]DMSO, 30°C, internal DMSO *δ*(C)= 39.52 ppm): ***δ*=** 170.30, 170.01, 170.00, 169.87, 169.59, 169.51, 169.23, 169.11, 169.08, 168.77, 168.64 (C=O-(Ac), C=O-(*t*Bu)), 156.80 (C=O-(Fmoc)), 154.00 (C=O-(Troc)), 143.73, 143.66 (C-1$_a$-, C-8$_a$-Fmoc), 140.82, 140.77 (C-4$_a$-, C-5$_a$-Fmoc), 127.73, 127.67 (C-2-, C-7-Fmoc), 127.03 (C-3-, C-6-Fmoc), 125.30, 125.17 (C-1-, C-8-Fmoc), 120.22, 120.16 (C-4-, C-5-Fmoc), 101.42 (C-1'),

100.55 (C-1''), 99.99 (C-1'''), 98.94 (C-1), 96.07 (C_{quart}-(Troc)), 81.36 (C_{quart}-(*t*Bu)), 78.33 (C-3'), 77.43 (C-3), 75.91 (C-4''), 73.49 (T^{β}), 73.38 (CH_{2ab}-(Troc)), 72.80 (C-3''), 71.67 (C-5), 71.51 (C-5''), 70.68 (C-5'), 70.26 (C-3'''), 69.59 (C-5'''), 69.31 (C-4'), 69.05 (C-2'), 68.91 (C-2'''), 67.43 (C-4), 67.02 (C-4'''), 65.58 (CH_{2ab}-(Fmoc)), 62.22 (C-6_{ab}), 61.80 (C-6_{ab}''), 60.81 (C-6_{ab}'''), 60.54 (C-6_{ab}), 59.36 (T^{α}), 55.75 (C-2''), 46.99 (C-2), 46.79 (C-9-Fmoc), 27.59 (*t*Bu), 22.95, 20.76, 20.66, 20.64, 20.49, 20.45, 20.43, 20.36, 20.32, 20.30 (CH_3-(Ac)), 19.09 (T^{γ}).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-acetamido-4,6-_O_-acetyl-2-deoxy-3-_O_-{2,4,6-tri-_O_-acetyl-3-_O_-[2-acetamido-3,6-_O_-acetyl-2-deoxy-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-β-D-galactosylpyranosyl}-α-D-galactosylpyranosyl)-L-threonine-*tert*-butylester (53)
(Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc$_2$GlcNAc-(1→3)-βAc$_3$Gal-(1→3)-αAc$_2$GalNAc)-O*t*Bu)



Zinc dust was activated by treatment with 1 N hydrochloric acid for a few minutes followed by washing with water, methanol and diethyl ether. Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc$_2$GlcNHTroc-(1→3)-βAc$_3$Gal-(1→3)-βGalNAc)-O*t*Bu **52** (492 mg, 0.29 mmol) was dissolved in acetic acid (6 mL) and preactivated zinc dust (393 mg, 6.0 mmol) was added. The reaction was stirred at 40°C for 18 h. The zinc dust was filtered off and washed with acetic acid. The filtrate was co-evaporated several times with toluene *in vacuo*. The residue was dissolved in pyridine/acetic anhydride (9 mL, 2:1). Then 4-(dimethylamino)pyridine (3.7 mg, 0.03 mmol) was added and the reaction was stirred for 20 h. The solvent was removed by co-evaporation with toluene *in vacuo* and the crude product purified by flash column chromatography on silica (100% EtOAc).

Yield: 389 mg (0.25 mmol, 82%), colorless, amorphous solid, $[\alpha]_D^{20}$ +41.40 (c = 0.50, $CHCl_3$), R_f = 0.21 (EtOAc/MeOH 20:1).

$C_{73}H_{95}N_3O_{36}$ (M = 1590.53 g/mol) [1589.57].

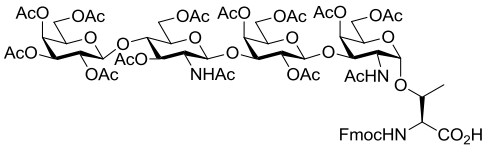
HR-ESI-MS (pos), m/z: 1590.5763 ($[M+H]^+$, calc. 1590.5774), 814.7661 ($[M+K+H]^{2+}$, calc. 814.7706).

^{1}H NMR (500 MHz, [D$_6$]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm): ***δ*=** 7.77 (d, ^3J(H-4,H-3) = ^3J(H-5,H-6) = 9.1 Hz, 2H, H-4-, H-5-Fmoc,), 7.62-7.59 (m, 2H, H-1-, H-8-Fmoc), 7.49-7.46 (m, 2H, H-2-, H-7-Fmoc), 7.39-7.35 (m, 2H, H-3-, H-6-Fmoc), 6.22 (d, ^3J(N-H,H-2) = 9.1 Hz, 1H, N-H-Ac), 5.92 (d, ^3J(N-H,T^{α}) = 9.1 Hz, 1H, N-H-Fmoc), 5.60-5.54 (m, 1H, N-H''-Ac), 5.32-5.29 (m, 3H, H-4, H-4', H-4'''), 5.15 (t, ^3J(H-3'', H-2'') = ^3J(H-3'',H-4'') = 9.5 Hz, 1H, H-3''),

5.07 (dd, ^3J(H-2''',H-1''') = 8.0 Hz, ^3J(H-2''',H-3''') = 10.3 Hz, 1H, H-2'''), 4.98-4.95 (m, 2H, H-2', H-3'''), 4.83 (s_{br}, 1H, H-1), 4.76-4.68 (m, 2H, H-1'', H-6_a''), 4.57 (d, ^3J(H-1',H-2') = 7.8 Hz, 1H, H-1'), 4.51-4.49 (m, 2H, H-1''', H-2), 4.43 (m, 1H, CH_{2a}-(Fmoc)), 4.35-4.31 (m, 1H, CH_{2b}-(Fmoc)), 4.25-4.17 (m, 3H, T^{β}, T^{α}, H-9-Fmoc), 4.12-4.02 (m, 6H, H-5, H-6_{ab}, H-6_a', H-6_{ab}'''), 3.99-3.90 (m, 3H, H-3, H-6_b', H-6_b'''), 3.86-3.50 (m, 3H, H-3', H-5, H-5'), 3.77 (t, ^3J(H-4'',H-3'') = ^3J(H-4'',H-5'') = 9.3 Hz, 1H, H-4''), 3.58-3.52 (m, 2H, H-2'', H-5''), 2.11-1.87 (m, 39H, CH_3(Ac)), 1.43 (s, 9H, *t*Bu), 1.30 (m, 3H, T^{γ}).

^{13}C NMR (125.6 MHz, [D_6]DMSO, 30°C, internal DMSO δ(C)= 39.20 ppm: **δ=** 170.89, 170.62, 170.50, 170.46, 170.27, 170.22, 170.16, 170.15, 170.00, 169.67, 169.26 (C=O-(Ac), C=O(*t*Bu)), 156.63 (C=O-(Fmoc)), 143.82 (C-1_a-, C-8_a-Fmoc), 141.49 (C-4_a-, C-5_a-Fmoc), 127.99 (C-2-, C-7-Fmoc), 127.27 (C-3-, C-6-Fmoc), 125.10 (C-1-, C-8-Fmoc), 120.24 (C-4-, C-5-Fmoc), 101.23 (C-1'''), 100.60 (C-1'), 100.45 (C-1''), 100.09 (C-1), 83.56 (C_{quart}-(*t*Bu)), 76.6 (T^{β}), 76.13 (C-5), 75.94 (C-4''), 72.68 (C-5''), 71.92 (C-3, C-3''), 71.33 (C-5'), 71.01 (C-3'''), 70.83 (C-3'), 70.70 (C-2'), 69.29 (C-4, C-4'), 69.19 (C-2'''), 68.15 (C-5'''), 67.20 (CH_{2ab}-(Fmoc)), 66.75 (C-4'''), 63.16 (C-6_{ab}'), 61.91 (C-6_{ab}'''), 60.90 (C-6_{ab}, C-6_{ab}''), 59.20 (T^{α}), 55.30 (C-2''), 48.70 (C-2), 47.36 (C-9-(Fmoc)), 28.24 (*t*Bu), 23.30-20.68 (CH_3-(Ac)), 18.76 (T^{γ}).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-acetamido-4,6-_O_-acetyl-2-deoxy-3-_O_-{2,4,6-tri-_O_-acetyl-3-_O_-[2-acetamido-3,6-_O_-acetyl-2-deoxy-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-β-D-galactosylpyranosyl}-α-D-galactosylpyranosyl)-L-threonine (54)
(Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc$_2$GlcNAc-(1→3)-βAc$_3$Gal-(1→3)-αAc$_2$GalNAc)-OH)



To a solution of Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc$_2$GlcNAc-(1→3)-βAc$_3$Gal-(1→3)-βAc$_2$GalNAc)-O*t*Bu **53** (718 mg, 0.45 mmol) in dichloromethane (3 mL) and anisole (0.5 mL) was added trifluoroacetic acid (9 mL) and the reaction was stirred for 2.5 h. The solvent was removed by co-evaporation with toluene. The crude product was purified by flash column chromatography on silica (EtOAc→EtOAc/MeOH 25:1).

Yield: 616 mg (0.40 mmol, 89%), colorless, amorphous solid, $[\alpha]_D^{20}$ +48.57 (c = 1.00, $CHCl_3$)
R_f = 0.21 (EtOAc/MeOH/AcOH/H_2O 50:3:3:2).

$C_{69}H_{87}N_3O_{36}$ (M = 1534.43 g/mol) [1533.51].

ESI-MS (pos), m/z: 1534.00 ([M+H]$^+$, calc. 1534.51), 1556.33 ([M+Na]$^+$, calc. 1556.50).

HR-ESI-MS (pos), m/z: 1534.5139 ([M+H]$^+$, calc. 1534.5148).

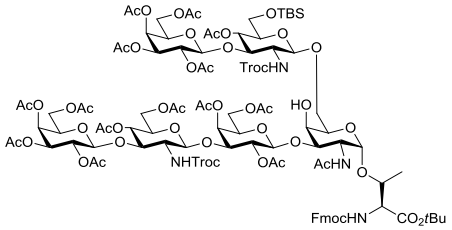
^{1}H NMR (600 MHz, [D_6]DMSO, 30°C, internal DMSO δ(H)= 2.50 ppm): δ= 12.90 (COOH), 7.91, 7.90 (2 x d, $J_{H\text{-}4,H\text{-}3} = J_{H\text{-}5,H\text{-}6}$ = 7.4 Hz, 2 x 1H, H-4-, H-5-Fmoc), 7.83 (d, $J_{N\text{-}H'',H\text{-}2''}$ = 9.0 Hz, 1H, N-H''-Ac), 7.74 (m, 2H, H-1-, H-8-Fmoc), 7.45-7.30 (m, 6H, H-2-, H-3-, H-6-, H-7-Fmoc, N-H-Fmoc, N-H-Ac), 5.28 (d, $J_{H\text{-}4',H\text{-}3'}$ = 3.7 Hz, 1H, H-4'), 5.27 (d, $J_{H4,H3}$ = 3.4 Hz, 1H, H-4), 5.22 (d, $J_{H\text{-}4''',H\text{-}3'''}$ = 3.5 Hz, 1H, H-4'''), 5.17 (dd, $J_{H\text{-}3''',H\text{-}4'''}$ = 3.5 Hz, $J_{H\text{-}3''',H\text{-}2'''}$ = 10.3 Hz, 1H, H-3'''), 5.00 (t, $J_{H\text{-}3'',H\text{-}4''} = J_{H\text{-}3'',H\text{-}2''}$ = 9.5 Hz, 1H, H-3''), 4.83 (dd, $J_{H\text{-}2''',H\text{-}3'''}$ = 10.3 Hz, $J_{H\text{-}2''',H\text{-}1'''}$ = 8.0 Hz, 1H, H-2'''), 4.74-4.69 (m, 4H, H-1, H-2, H-1'', H-1'''), 4.59 (d, $J_{H\text{-}1',H\text{-}2'}$ = 8.0 Hz, 1H, H-1'), 4.52-4.43 (m, 2H, CH_{2ab}-(Fmoc)), 4.32-4.30 (m, 2H, H-9-(Fmoc), H-6_a''), 4.27-4.25 (m, 1H, T^{β}), 4.21 (t, $J_{H\text{-}5''',H\text{-}6'''}$ = 6.8 Hz, 1H, H5'''), 4.15-4.08 (m, 4H, H-2, H-5, H-6_a', T^{α}), 4.00-3.98 (m, 2H, H-6_{ab}'''), 3.94-3.79 (m, 6H, H-3, H-3', H-5', H-6_{ab}, H-6_b', H-6_b''), 3.65 (t, $J_{H\text{-}4'',H\text{-}3''} = J_{H\text{-}4'',H\text{-}5''}$ = 9.5 Hz, 1H, H-4''), 3.57-3.54 (m, 1H, H-5''), 3.41-3.37 (m, 1H, H-2''), 2.09-1.70 (m, 39H, CH_3-(Ac)).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO δ(C)= 39.20 ppm: δ= 170.29, 170.11, 170.01, 169.87, 169.73, 169.51, 169.43, 169.39, 169.13, 169.02, 168.82 (C=O-(Ac)), 156.78 (C=O-(Fmoc)), 143.76, 143.72 (C-1_a-, C-8_a-Fmoc), 140.81, 140.77 (C-4_a-, C-5_a-Fmoc), 127.71, 127.67 (C-2-, C-7-Fmoc), 127.06 (C-3, C-6-Fmoc), 125.27, 125.14 (C-1-, C-8-Fmoc), 120.22, 120.16 (C-4-, C-5-Fmoc), 100.55 (C-1'), 100.00, 98.74 (C-1, C-1'', C-1'''), 76.87 (C-3'), 76.14 (C-4''), 74.45 (T^{β}), 73.43 (C-3), 72.71 (C-3''), 71.44 (C-5''), 70.25 (C-5',C-3'''), 69.67 (C-4), 69.56 (C-5'''), 69.47 (C-2'), 69.04 (C-4'), 68.94 (C-2'''), 67.26 (C-5), 67.03 (C-4''), 65.56 (CH_{2ab}-(Fmoc)), 63.08 (C-6_{ab}'), 61.77, 61.69 (C-6_{ab}, C-6_{ab}''), 60.77 (C-6_{ab}'''), 58.47 (T^{α}), 53.98 (C-2''), 47.71 (C-2), 46.78 (C-9-Fmoc), 22.80, 2 x 22.63, 20.64-20.38 (CH_3-(Ac)), 18.43 (T^{γ}).

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

***N*-(9H-fluoren-9-yl)-methoxycarbonyl-*O*-(2-*N*-acetamido-2-deoxy-3-*O*-{2,4,6-*O*-acetyl-3-*O*-[4,6-*O*-acetyl-2-deoxy-2-*N*-(2,2,2-trichloroetoxycarbonyl)-3-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-β-D-galactopyranosyl}-6-*O*-{4-*O*-acetyl-6-*O*-*tert*-butyldimethylsilyl-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 (55)**

(Fmoc-Thr(βAc$_4$Gal-(1→3)-βAc$_2$-TBS-GlcNHTroc-(1→3)-Ac$_3$Gal-(1→3)-{βAc$_4$Gal-(1→3)-βAc$_2$-TBS-GlcNHTroc-(1→6)}-αGalNAc)-OtBu)



Fmoc-Thr(βAc$_4$Gal-(1→3)-βAc$_2$GlcNHTroc-(1→3)-βAc$_3$Gal-(1→3)-βGalNAc)-O*t*Bu **47** (1.25 g, 0.76 mmol) and βAc$_4$Gal-(1→3)-βAc-TBS-GlcNHTroc-SPh **27** (933 mg, 1.00 mmol) were dissolved in dry dichloromethane (15 mL) under argon atmosphere. Activated molecular sieves 4Å (1 g) were added and the suspension was stirred for 1 h before it was cooled in an ice bath. *N*-iodosuccinimide (226 mg, 1.00 mmol) was added followed by the dropwise addition of a suspension of trifluoromethanesulfonic acid (13.5 µL, 23 mg, 0.15 mmol) in dry dichloromethane (100 µL) *via* a syringe. After 1.5 h the reaction was treated with another addition of **27** (95 mg, 0.1 mmol) and *N*-iodosuccinimide (28 mg, 0.1 mmol). The reaction was stirred for 2 h under argon atmosphere and ice-cooling. The reaction was diluted with dichloromethane (20 mL) and the molecular sieves were filtered off and washed with dichloromethane. The filtrate was diluted with dichloromethane (total volume of 100 mL). The dichloromethane phase was washed with 0.5 M sodium thiosulfate solution, saturated sodium bicarbonate solution, water and brine (30 mL each). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography on silica (^{c}Hex/EtOAc 1:1→1:2).

Yield: 1.47 g (0.597 mmol, 78%), colorless, amorphous solid, $[\alpha]_D^{20}$ +19.20 (c = 0.49, $CHCl_3$), R_f = 0.41 (Tol/EtOAc 1:3).

$C_{101}H_{136}Cl_6N_4O_{51}Si$ (M = 2462.96 g/mol) [2458.61].

MALDI-TOF-MS (dhb, pos), m/z: 2386.41 ([M-TBS+K]$^+$, calc. 2386.48), 2500.40 ([M+K]$^+$, calc. 2500.57).

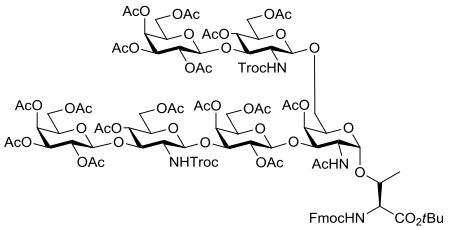
HR-ESI-MS (pos), m/z: 1250.7860 ([M+K+H]$^{2+}$, calc. 1250.7896).

^{1}H NMR (600 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm): ***δ*=** 7.91-7.90 (m, 2H, H-4-, H-5-Fmoc), 7.76-7.70 (m, 4H, H-1-, H-8-Fmoc, NH''-Troc, NH''''-Troc), 7.45-7.30 (m, 6H, H-2-, H-7-Fmoc, C-3-, C-6-Fmoc, NH-Ac, NH-Fmoc), 5.30 (m, 1H, H-4'), 5.24-5.22 (m, 2H, H-4''', H-4''''), 5.01 (dd, ^{3}J(H-3'''',H-4'''') = 3.5 Hz, ^{3}J(H-3'''',H-2'''') = 10.5 Hz, 1H, H-3''''), 4.93 (dd, ^{3}J(H3''',H4''') = 3.5 Hz, ^{3}J(H-3''',H-2''') = 10.4 Hz, 1H, H-3'''), 4.88-4.85 (m, 2H, H-2', CH_{2a}-(Troc)), 4.81-4.75 (m, 3H, H-2''', H-2'''', CH_{2a}-(Troc)), 4.72-4.59 (m, 5H, H-1''', H-1'''', H-4'', H-4'''', CH_{2b}-(Troc)), 4.57-4.54 (m, 3H, H-1, H-1', CH_{2b}-(Troc)), 4.50-4.44 (m, 3H, H-1'', CH_2-(Fmoc)), 4.39-4.38 (d, ^{3}J(H1'''',H2'''') = 8.4 Hz, 1H, H-1''''), 4.32 (t, ^{3}J(H-9,CH_2) = 6.7 Hz, 1H, H-9-Fmoc), 4.21-4.17 (m, 1H, H-2), 4.12-3.91 (m, 12H, H-3'', H-5''', H-5'''', H-6'$_{ab}$, H-6''$_{ab}$, H-6'''$_{ab}$, H-6''''$_{ab}$, T^{α}), 3.88-3.81 (m, 4H, H-3', H-3'''', H-5', H-6$_a$), 3.78-3.73 (m, 2H, H-4, H-5), 3.68-3.63 (m, 2H, H-6''''$_a$, H-5''), 3.60-3.49 (m, 3H, H-3, H-6$_b$, H-6''''$_b$), 3.47-3.45 (m, 1H, H-5''''), 3.38 (q, ^{3}J(H2'''',H3'''') = ^{3}J(H2'''', H1'''') = ^{3}J(H2'''',NH) = 9.3 Hz, 1H, H-2''''), 3.29-3.24 (m, 1H, H-2''), 2.09, 2.04, 2.03, 2.01, 2.00, 1.99, 1.98, 1.97, 1.96, 1.88, 1.84 (m, 45H, 15 x CH_3-(Ac)), 1.36 (s, 9H, *t*Bu-(Thr)), 1.13 (d, ^{3}J(T^{γ},T^{β}) = 6.0 Hz, 3H, T^{γ}), 0.85 (s, 9H, *t*Bu-(TBS)), 0.01, 0.00 (2 x s, 2 x 3H, 2 x CH_3-(TBS)).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(C)=39.51 ppm): ***δ*=**170.07, 170.02, 169.88, 169.82, 169.73, 169.69, 169.64, 169.53, 169.53, 169.48, 169.38, 169.31, 169.25, 169.23, 169.20, 169.13, 169.09, 169.04, 169.86, 168.66 (C=O-(Ac), C=O-(*t*Bu)), 156.75 (C=O-(Fmoc)), 154.08, 153.80 (2 x C=O-(Troc)), 143.68, 143.67 (C-1$_a$-, C-8$_a$-Fmoc), 140.80, 140.76 (C-4$_a$-, C-5$_a$-Fmoc), 127.70, 127.66 (C-2-, C-7-Fmoc), 125.23, 125.12 (C-1-, C-8-Fmoc), 120.21, 120.16 (C-4-, C-5-Fmoc), 101.44 (C-1'), 101.13 (C-1''), 100.61 (C-1''''), 99.96 (C-1''''), 99.77 (C-1'''), 99.05 (C-1), 95.89, 95.52 (2 x CCl_3-(Troc)), 81.39 (C_q-*t*Bu), 77.92 (C-3''''), 77.31 (C-3'), 76.70 (C-3), 76.59 (C-3''), 74.00 (T^{β}, CH_2-(Troc)), 73.55 (C-5''''), 73.49 (CH_2-(Troc)), 70.48 (C-5'), 70.33 (C-3''''), 70.20 (C-5''), 70.15 (C-3'''), 70.06 (C-5), 69.75 (C-5''', C-6), 69.60 (C-5''''), 69.37 (C-2'), 69.24 (C-4'), 68.81 (C-4''''), 68.75 (C-2'''), 68.53 (C-2''''), 68.43 (C-4), 68.25 (C-4''), 67.22, 67.18 (C-4''', C-4''''), 65.60 (CH_2-(Fmoc)), 62.03 (C-6''''), 61.70 (C-6''), 61.50 (C-6'), 60.92 (C-6'''), 60.84 (C-6''''), 59.57 (T^{α}), 56.93 (C-2''), 56.62 (C-2''''), 46.92 (C-2), 46.77 (C-9-Fmoc), 27.62 (C_q-*t*Bu(Thr)), 25.74 (C_q-TBS), 22.95, 20.62, 20.60, 20.50, 20.47, 20.42, 20.34, 20.27, 20.23 (CH_3-(Ac)), 19.03 (T^{γ}), 18.02 (CH_3-(*t*Bu-TBS)), -5.42, -5.44 (CH_3-(TBS)).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(4-_O_-acetyl-2-_N_-acetamido-2-deoxy-3-_O_-{2,4,6-_O_-acetyl-3-_O_-[4,6-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroetoxycarbonyl)-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-β-D-galactopyranosyl}-6-_O_-{4,6-_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 (56)

(Fmoc-Thr(βAc$_4$Gal-(1→3)-βAc$_2$GlcNHTroc-(1→3)-Ac$_3$Gal-(1→3)-{βAc$_4$Gal-(1→3)-βAc$_2$GlcNHTroc-(1→6)}-αGalNAc)-OtBu)



A solution of Fmoc-Thr(βAc$_4$Gal-(1→3)-βAc$_2$-TBS-GlcNHTroc-(1→3)-βAc$_3$Gal-(1→3)-[βAc$_4$Gal-(1→3)-βAc$_2$-TBS-GlcNHTroc-(1→6)]-αGalNAc)-O*t*Bu **55** (2.91 g, 1.18 mmol) in 80% acetic acid (30 mL) was stirred for 17 h at 35°C. The solvent was removed by co-evaporation with toluene *in vacuo*. The residue was dissolved in pyridine (20 mL) and cooled in an ice bath. Then 4-(dimethylamino)pyridine (15 mg, 0.12 mmol) and acetic anhydride (10 mL) were added and the solution was stirred for 6.5 h. The crude product was purified by flash column chromatography on silica (^{C}Hex/EtOAc 1:3).

Yield: 2.74 g (1.12 mmol, 95%), colorless, amorphous solid, $[\alpha]_D^{20}$ +26.70 (c = 0.50, $CHCl_3$), R_f = 0.36 (Tol/EtOAc 1:4).

$C_{99}H_{126}Cl_6N_4O_{53}$ (M = 2432.77 g/mol) [2428.54].

MALDI-TOF-MS (dhb, pos), m/z: 2471.00 ([M+K]$^+$, calc. 2470.51).

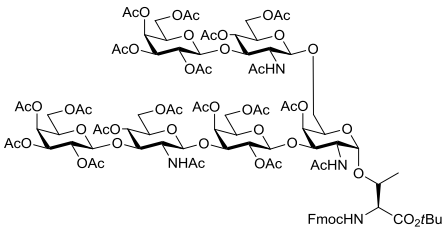
HR-ESI-MS (pos), m/z: 1235.7530 ([M+K+H]$^{2+}$, calc. 1235.7569).

^{1}H NMR (600 MHz, [D$_6$]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm): ***δ*=** 7.91-7.90 (m, 2H, H-4-, H-5-Fmoc), 7.79 (d, ^3J(NH,H-2"") = 9.7 Hz, 1H, NH""-Troc,), 7.76-7.73 (m, 2H, H-1-, H-8-Fmoc), 7.68 (d, ^3J(NH,H2") = 9.4 Hz, 1H, NH"-Troc), 7.49 (d, ^3J(NH,H-2) = 9.8 Hz, 1H, NH-Ac), 7.43-7.41 (m, 3H, H-2-, H-7-Fmoc, NH-Fmoc), 7.34-7.29 (m, 2H, H-3-, H-6-Fmoc), 5.27 (s_{br}, 1H, H-4'), 5.23-5.22 (m, 2H, H-4''', H-4''''), 5.21 (s_{br}, 1H, H-4), 5.03 (dd, ^3J(H-3'''',H-2'''') = 10.3 Hz, ^3J(H-3'''',H-4'''') = 3.6 Hz, 1H, H-3''''), 4.93 (dd, ^3J(H-3''',H-2''') = 3.5 Hz, ^3J(H-3''',H-4''') = 10.4 Hz, 1H, H-3'''), 4.87-4.85 (d, ^2J(CH_{2a},CH_{2b}) = 12.1 Hz, 1H, CH_{2a}-(Troc)), 4.82-4.75 (m, 3H, CH_{2a}-(Troc), H-2''', H-2''''), 4.73-4.67 (m, 4H, H-1"", H-2, H-4"", CH_{2b}-(Troc)), 4.66-4.58 (m, 3H, H-1, H-1''', H-4"), 4.56-4.50 (m, 3H, H-1', CH_{2b}-(Troc), CH_{2a}-(Fmoc)), 4.48-4.39 (m, 3H, H-1", H-1"", CH_{2b}-(Fmoc)), 4.31 (t, ^3J(H-9,CH_2) = 6.7 Hz, 1H, H-9-Fmoc), 4.14-4.01 (m, 9H, H-2, H-6'$_a$, H-6'''$_a$, H-6""$_a$, H-5, H-5''', H-5'''', T$^\beta$), 3.99-3.93 (m, 5H, H-6'$_b$, H-6'''$_b$, H-6""$_b$, H-6''''$_b$, T$^\alpha$), 3.92-3.87 (m, 3H, H-3", H-6"$_{ab}$), 3.86-3.81 (m, 3H, H-3', H-3"", H-5'), 3.80-

3.73 (m, 2H, H-3, H-6_a), 3.70-3.69 (m, 1H, H-5''''), 3.66-3.64 (m, 1H, H-5''), 3.43-3.35 (m, 2H, H-6_b, H-2''''), 3.26 (q, ^3J(H-2'', H-3'') = ^3J(H-2'',H-1'') = ^3J(H-2'',NH) = 9.5 Hz, 1H, H-2''), 2.09, 2.06, 2.02, 2.01, 2.00, 1.99, 1.98, 1.97, 1.96, 1.88 (m, 51H, CH_3-(Ac)), 1.35 (s, 9H, *t*Bu), 1.15 (d, $^3J(T^{\gamma},T^{\beta})$ = 6.3 Hz, 3H, T^{γ}).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO δ(C)=39.52 ppm): δ = 170.07, 170.06, 169.92, 169.89, 169.83, 169.59, 169.39, 169.35, 169.32, 169.29, 169.24, 169.14, 169.07, 168.76, 168.70 (C=O-(Ac), C=O-(*t*Bu)), 156.77 (C=O-(Fmoc)), 154.05 (C=O-(Troc)), 153.82 (C=O-(Troc)), 143.71, 143.67 (C-1_a-, C-8_a-Fmoc), 140.82, 140.77 (C-4_a-, C-5_a-Fmoc), 127.69, 127.67 (C-2-, C-7-Fmoc), 127.03, 127.02 (C-3-, C-6-Fmoc), 125.23, 125.11 (C-1-, C-8-Fmoc), 120.19, 120.14 (C-4-, C-5-Fmoc), 101.21 (C-1''), 100.52 (C-1'), 100.50 C-1''''), 99.91 (C-1''''), 99.81 (C-1'''), 98.74 (C-1), 95.84 (CCl_3-Troc), 95.55 (CCl_3-Troc), 81.48 (C_q-*t*Bu), 77.70 (C-3'), 76.81 (C-3''), 76.58 (C-3''''), 74.18 (T^{β}), 73.93 (CH_2-(Troc)), 73.64 (CH_2-(Troc)), 73.30 (C-3), 70.60 (C-5''), 70.24 (C-3''''), 70.15 (C-3''', C-4, C-5', C-5''''), 69.74 (C-5''''), 69.61 (C-2', C-5'''), 69.21 (C-4'), 68.98 (C-6), 68.86 (C-5, C-4''''), 68.75, 68.51 (C-2''', C-2''''), 68.32 (C-4''), 67.23, 67.14 (C-4''', C-4''''), 65.57 (CH_2-Fmoc), 62.00 (C-6''''), 61.57 (C-$6''_{ab}$), 61.45 (C-$6'_{ab}$), 60.93 (C-$6'''_{ab}$), 60.80 (C-$6''''_{ab}$), 59.57 (T^{α}), 56.96 (C-2''''), 56.62 (C-2''), 47.62 (C-2), 46.78 (C-9-Fmoc), 27.61 (*t*Bu), 22.82, 20.64, 20.60, 20.56, 20.51, 20.47, 20.44, 20.41, 20.34, 20.27, 20.23 (CH_3-(Ac)), 19.03 (T^{γ}).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(4-_O_-acetyl-2-_N_-acetamido-2-deoxy-3-_O_-{2,4,6-_O_-acetyl-3-_O_-[2-_N_-acetamido-4,6-_O_-acetyl-2-deoxy-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-β-D-galactopyranosyl}-6-_O_-{2-_N_-acetamido-4,6-_O_-acetyl-2-deoxy-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl}-α-D-galactopyranosyl)-L-threonine-*tert*-butylester (57)
(Fmoc-Thr(βAc$_4$Gal-(1→3)-βAc$_2$GlcNAc-(1→3)-Ac$_3$Gal-(1→3)-{βAc$_4$Gal-(1→3)-βAc$_2$GlcNAc-(1→6)}-αGalNAc)-OtBu)



Zinc dust was activated by treatment with 1 N hydrochloric acid for a few minutes followed by washing with water, methanol and finally diethyl ether. Fmoc-Thr(βAc$_4$Gal-(1→3)-βAc$_2$GlcNHTroc-(1→3)-βAc$_3$Gal-(1→3)-[βAc$_4$Gal-(1→3)-βAc$_2$GlcNHTroc-(1→6)]-αAcGalNAc)-O*t*Bu **56** (2.68 g, 1.10 mmol) was dissolved in acetic acid (35 mL) and activated zinc dust (2.16 g, 16.5 mmol) was added. The reaction was stirred at 35°C for 36 h. The zinc

was filtered off and washed with acetic acid. The filtrate was co-evaporated with toluene *in vacuo*. The residue was dissolved in pyridine/acetic anhydride (30 mL, 2:1). Then 4-(dimethylamino)pyridine was added (7 mg, 0.06 mmol) and stirred for 17 h. The solvent was removed by co-evaporation with toluene *in vacuo* and the crude product purified by flash column chromatography on silica (EtOAc→EtOAc/MeOH 40:1).

Yield: 1.91 g (0.88 mmol, 80%), colorless, amorphous solid, $[\alpha]_D^{20}$ +26.53 (c = 0.49, $CHCl_3$), R_f = 0.30 (EtOAc/MeOH 25:1).

$C_{97}H_{128}N_4O_{51}$ (M = 2166.05 g/mol) [2164.75].

MALDI-TOF-MS (dhb, pos), m/z: 2187.62 ($[M+Na]^+$, calc. 2187.74), 2203.58 ($[M+K]^+$, calc. 2203.72).

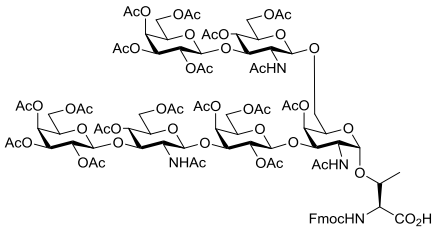
HR-ESI-MS (pos), m/z: 1083.8874 ($[M+2H]^{2+}$, calc. 1083.8868), 1102.8611 ($[M+K+H]^{2+}$, calc. 1102.8647).

^{1}H NMR (600 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm): ***δ*=** 7.91-7.88 (m, 2H, H-4-, H-5-Fmoc), 7.79 (d, ^3J(NH'''',H-2'''') = 7.6 Hz, 1H, NH''''-Ac), 7.77-7.73 (m, 2H, H-1-, H-8-Fmoc), 7.69 (d, ^3J(NH'',H-2'') = 9.3 Hz, 1H, NH''-Ac), 7.56 (d, ^3J(NH,H-2) = 10.0 Hz, 1H, NH-Ac), 7.44-7.39 (m, 3H, H-2-, H-7-Fmoc, NH-Fmoc), 7.33-7.30 (m, 2H, H-3-, H-6-Fmoc), 5.25 (s_{br}, 1H, H-4'), 5.23 (m, 1H, H-4''', H-4''''), 5.21 (s_{br}, 1H, H-4), 5.09-5.06 (m, 2H, H-3''', H3''''), 4.77-4.74 (m, 2H, H-2''', H-2''''), 4.71-4.70 (m, 2H, H-1'''', H-2'), 4.66-4.60 (m, 4H, H-1, H-1''', H-4'', H-4''''), 4.57 (d, ^3J(H-1',H-2') = 8.0 Hz, 1H, H-1'), 4.52-4.51 (m, 1H, CH_{2a}-(Fmoc)), 4.48-4.43 (m, 2H, H-1'', CH_{2b}-(Fmoc)), 4.38-4.37 (m, 1H, H-1''''), 4.31 (t, ^3J(H-9,CH_2) = 6.7 Hz, 1H, H-9-Fmoc), 4.18-3.96 (m, 13H, H-2, T^{α} ,T^{β}, H-5, H-5''', H-5'''', H-6'$_{ab}$, H-6$_a$'', H-6'''$_{ab}$, H-6''''$_{ab}$), 3.90-3.89 (m, 2H, H-6$_b$''), 3.85-3.83 (m, 1H, H-5'), 3.80-3.73 (m, 7H, H-2'''', H-3, H-3', H-3'', H-3'''', H-5'''', H-6$_a$), 3.69-3.67 (m, 1H, H-5''), 3.52-3.50 (m, 1H, H-2''), 3.34-3.32 (m, 1H, H-6$_b$), 2.09, 2.05, 2.02-1.98, 1.91, 1.90, 1.88, 1.85, 1.83, 1.81 (m, 57H, CH_3-(Ac)), 1.34 (s, 9H, *t*Bu), 1.13 (d, ^3J(T^{γ},T^{β}) = 6.1 Hz, 3H, T^{γ}).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(C)=39.51 ppm): ***δ*=** 170.10, 170.04, 169.99, 169.92, 169.89, 169.63, 169.50, 169.39, 169.27, 169.16, 168.93, 168.86, 168.81, 168.75 (C=O-(Ac), C=O-(*t*Bu)), 156.80 (C=O-(Fmoc)), 143.73, 143.67 (C-1$_a$-, C-8$_a$-Fmoc), 140.82, 140.78 (C-4$_a$, C-5$_a$-Fmoc), 127.73, 127.70 (C-2-, C-7-Fmoc), 127.06, 127.03 (C-3-, C-6-Fmoc), 125.28, 125.12 (C-1-, C-8-Fmoc), 120.21, 120.15 (C-4-, C-5-Fmoc), 100.97 (C-1''''), 100.70 (C-1''), 100.53 (C-1'), 99.97 (C-1''''), 99.92 (C-1'''), 98.33 (C-1), 81.42 (C_q-tBu), 77.85 (C-3''''), 77.19 (C-3'), 76.47 (C-3''), 73.58 (T^{β}), 73.30 (C-3), 70.57 (C-5''''), 70.52, 70.48 (C-3''', C-3''''), 70.29, 70.16 (C-5'', C-5'), 69.92 (C-4), 69.61 (C-2'), 69.49 (C-5''', C-5''''), 69.16 (C-6), 68.98 (C-4'), 68.68 (C-5), 68.63, 68.56 (C-2'''', C-4''''), 68.47, 68.46 (C-2''', C-4''), 67.17 (C-4''', C-4''''), 65.60 (CH_2-(Fmoc)), 62.03 (C-6''''$_{ab}$), 61.66 (C-6''$_{ab}$), 61.46 (C-6'$_{ab}$), 61.12 (C-6'''$_{ab}$), 61.05 (C-6''''$_{ab}$), 59.51 (T^{α}), 54.22 (C-2''), 53.62 (C-2''''), 47.72 (C-2),

46.78 (C-9-Fmoc), 27.57 (*t*Bu), 22.91, 22.82, 22.81, 20.76, 20.67, 20.58, 20.52, 20.47, 20.43, 20.39, 20.35, 20.31 (CH_3-(Ac)), 19.17 (T^{γ}).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(4-_O_-acetyl-2-_N_-acetamido-2-deoxy-3-_O_-{2,4,6-_O_-acetyl-3-_O_-[2-_N_-acetamido-4,6-_O_-acetyl-2-deoxy-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-β-D-galactopyranosyl}-6-_O_-{2-_N_-acetamido-4,6-_O_-acetyl-2-deoxy-3-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl}-α-D-galactopyranosyl)-L-threonine (58)
(Fmoc-Thr(βAc$_4$Gal-(1→3)-βAc$_2$GlcNAc-(1→3)-Ac$_3$Gal-(1→3)-{βAc$_4$Gal-(1→3)-βAc$_2$GlcNAc-(1→6)}-αGalNAc-OH)



Trifluoroacetic acid (15 mL) was added to a solution of Fmoc-Thr(βAc$_4$Gal-(1→3)-βAc$_2$GlcNAc-(1→3)-βAc$_3$Gal-(1→3)-[βAc$_4$Gal-(1→3)-βAc$_2$GlcNAc-(1→6)]-αAcGalNAc)-O*t*Bu **57** (1.82 g, 0.84 mmol) and anisole (2 mL) in dichloromethane (5 mL) and the reaction mixture was stirred for 3 h. The solvent was removed by co-evaporation with toluene. The crude product was purified by flash column chromatography on silica (EtOAc→EtOAc/MeOH/AcOH/H_2O 50:3:3:2).

Yield: 1.63 g (0.77 mmol, 92%), colorless, amorphous solid, $[\alpha]_D^{20}$ +34.36 (c = 1.00, $CHCl_3$), R_f = 0.12 (EtOAc/MeOH/AcOH/H_2O 50:3:3:2).

$C_{93}H_{120}N_4O_{51}$ (M = 2109.94 g/mol) [2108.69].

MALDI-TOF-MS (dhb, pos), m/z: 2132.06 ($[M+Na]^+$, calc. 2132.93), 2148.01 ($[M+K]^+$, calc. 2148.66).

HR-ESI-MS (pos), m/z: 1055.8560 ($[M+2H]^{2+}$, calc. 1055.8555), 1074.8302 ($[M+K+H]^{2+}$, calc. 1074.8334).

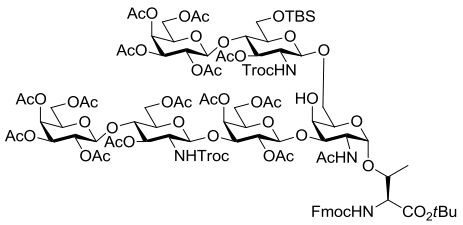
^{1}H NMR (600 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm): ***δ*=** 12.20 (s, 1H, COOH), 7.91 (m, 2H, H-4-, H-5-Fmoc), 7.81-7.79 (m, 1H, NH""-Ac), 7.77-7.73 (m, 3H, H-1-, H-8-Fmoc, NH''-Ac), 7.45-7.41 (m, 3H, H-2-, H-7-Fmoc, NH-Ac), 7.34-7.30 (m, 3H, H-3-, H-6-Fmoc, NH-Fmoc), 5.25 (d, ^3J(H4',H3') = 3.5 Hz, 1H, H-4'), 5.23-5.22 (m, 2H, H-4''', H-4''''), 5.21 (s, 1H, H-4), 5.09-5.06 (m, 2H, H-3''', H-3''''), 4.77-4.73 (m, 2H, H-2''', H-2''''), 4.72-4.69 (m, 2H, H-1'''', H-2'), 4.67-4.60 (m, 4H, H-1, H-1''', H-4'', H-4''''), 4.56 (d, ^3J(H1',H2') = 8.1 Hz, 1H, H-1'), 4.52 (dd, ^2J(CH_{2a},CH_{2b}) = 10.8 Hz, ^3J(CH_{2a},H-9) = 6.7 Hz, 1H, CH_{2a}-(Fmoc)), 4.47 (m, 1H, H-1''), 4.43 (dd, ^2J(CH_{2b},CH_{2a}) = 11 Hz, ^3J(CH_{2b},H-9) = 6.9 Hz, 1H, CH_{2b}-Fmoc), 4.38-

4.37 (m, 1H, H-1''''), 4.31 (t, ^3J(H-9,CH_2) = 6.8 Hz, 1H, H-9-Fmoc), 4.24-4.22 (m, 1H, T^{β}), 4.15-4.04 (m, 8H, H-2, T^{α}, H-5''', H-5'''', H-6'$_a$, H-6''$_a$, H-6'''$_a$, H-6''''$_a$), 4.03-3.95 (m, 5H, H-5, H-6'$_b$, H-6''$_b$, H-6'''$_b$, H-6''''$_b$), 3.91-3.84 (m, 3H, H-5', H-6''''$_{ab}$) 3.82-3.71 (m, 7H, H-2'''', H-6$_a$, H-5'''', H-3, H-3', H-3'', H-3''''), 3.70-3.66 (m, 1H, H-5''), 3.51 (s_{br}, 1H, H-2''), 3.32-3.30 (m, 1H, H-6$_b$), 2.09, 2.04, 2.01, 2.00, 1.99, 1.98, 1.97, 1.91, 1.88, 1.85, 1.81 (m, 57H, CH_3-(Ac)), 1.12 (d, ^3J(T^{γ},T^{β}) = 6.2 Hz, 3H, T^{γ}).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(C)=39.51 ppm): ***δ*=** 170.09, 170.05, 169.99, 169.92, 169.87, 169.59, 169.50, 169.38, 169.26, 169.16, 169.08, 169.08, 168.93, 168.92, 168,71 (C=O-(Ac)), 156.76 (C=O-(Fmoc)), 143.76, 143.70 (C-1$_a$-, C-8$_a$-Fmoc), 140.81, 140.75 (C-4$_a$, C-5$_a$-Fmoc), 127.70, 127.67 (C-2-, C-7-Fmoc), 127.06 (C-3-, C-6-Fmoc), 125.29, 125.14 (C-1-, C-8-Fmoc), 120.19, 120.13 (C-4-, C-5-Fmoc), 100.90 (C-1''''), 100.68 (C-1''), 100.57 (C-1'), 99.94, 99.91 (C-1''', C-1''''), 98.33 (C-1), 77.81 (C-3''''), 77.20 (C-3''), 76.49 (C-3'), 73.92 (T^{β}), 73.46 (C-3), 70.60, 70.56 (C-3''', C-3'''') 70.51 (C-5''''), 70.30, 70.18 (C-4, C-5'), 69.62 (C-2'), 69.47 (C-5''', C-5''''), 68.96 (C-4', C-6), 68.70 (C-5), 68.62, 68.55, (C-2''', C-2'''', C-4''''), 68.46 (C-4'') 67.15 (C-4''', C-4''''), 65.58 (CH_2-(Fmoc)), 62.03 (C-6''''), 61.71 (C-6''), 61.48 (C-6'), 61.10 (C-6'''), 61.03 (C-6''''), 58.52 (T^{α}), 54.20 (C-2''), 53.66 (C-2''''), 47.85 (C-2), 46.77 (C-9-Fmoc), 22.86, 22.82, 22.76, 20.76, 20.57, 20.51, 20.47, 20.45, 20.43, 20.38, 20.34, 20.30 (CH_3-(Ac)), 18.77 (T^{γ}).

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

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-acetamido-2-deoxy-3-_O_-{2,4,6-tri-_O_-acetyl-3-_O_-[3,6-_O_-acetyl-2-_N_-(2,2,2-trichloroethoxycarbonylamino)-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-2-deoxy-β-D-glucopyranosyl]-β-D-galactosylpyranosyl}-6-_O_-[3-_O_-acetyl-2-deoxy-6-_O_-_tert_-butyldimethylsilyl-2-_N_-(2,2,2-trichloroethoxycarbonylamino)-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-α-D-galactosylpyranosyl)-L-threonine-_tert_-butylester (59) (Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc$_2$-TBS-GlcNHTroc-(1→3)-βAc$_3$Gal-(1→3)-[βAc$_4$Gal-(1→4)-βAc$_2$-TBS-GlcNHTroc-(1→6)]-αGalNAc)-O*t*Bu)



Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc$_2$-TBS-GlcNHTroc-(1→3)-βAc$_3$Gal-(1→3)-βGalNAc)-O*t*Bu **52** (694 mg, 0.42 mmol) and βAc$_4$Gal-(1→4)-βAc-TBS-GlcNTroc-SPh **31** (515 mg, 0.55 mmol)

were dissolved in dry dichloromethane (15 mL) and 0.6 g of activated molecular sieves 4Å were added under argon atmosphere. The suspension was stirred for 1 h before and then cooled in an ice bath. *N*-iodosuccinimide (124 mg, 0.55 mmol) was added followed by the dropwise addition of a suspension of trifluoromethanesulfonic acid (7.5 µmol, 0.09 mmol) in dry dichloromethane (100 µL) *via* a syringe. After 3 h the reaction was treated with another addition of βAc_4Gal-(1→4)-βAc-TBS-GlcNTroc-SPh **31** (119 mg, 0.13 mmol) and *N*-iodosuccinimide (28 mg, 0.13 mmol). The reaction was stirred further for 3.5 h under argon atmosphere and ice-cooling. The molecular sieves were filtered off and washed with dichloromethane. More dichloromethane was added (total volume of 120 mL). The dichloromethane phase was washed with of 0.5 M sodium thiosulfate solution, saturated sodium bicarbonate, water and brine (30 mL each). The organic phase was dried over sodium sulfate, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography on silica (^{C}Hex/EtOAc 1:1→1:3).

Yield: 722 mg (0.29 mmol, 69%), colorless, amorphous solid, $[\alpha]_D^{20}$ +20.27 (c = 0.49, $CHCl_3$), R_f = 0.39 (Tol/EtOAc 1:3).

$C_{101}H_{136}Cl_6N_4O_{51}Si$ (M = 2462.96 g/mol) [2458.61].

MALDI-TOF-MS (dhb, pos), m/z: 2500.72 ($[M+K]^+$, calc. 2500.57).

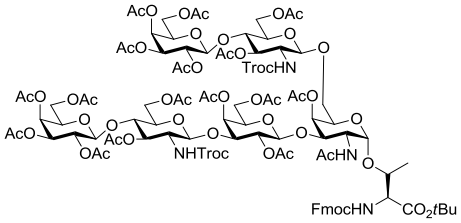
HR-ESI-MS (pos), m/z: 1250.7862 ($[M+K+H]^{2+}$, calc. 1250.7896).

^{1}H NMR (600 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm: ***δ*=** 7.92-7.88 (m, 2H, H-4-, H-5-Fmoc), 7.81-7.73 (m, 4H, H-1-, H-8-Fmoc, 2 N-H-Troc), 7.46-7.38 (m, 4H, H-2-, H-7-Fmoc, N-H-Ac, N-H-Fmoc), 7.34-7.30 (m, 2H, H-3-, H-6-Fmoc), 5.32 (s_{br}, 1H, H-4'), 5.25 (d, $J_{H\text{-}4'''',H\text{-}3''''}$ = 3.6 Hz, 1H, H-4''''), 5.22 (d, $J_{H\text{-}4''',H\text{-}3'''}$ = 3.6 Hz, 1H, H4'''), 5.16 (dd, $J_{H\text{-}3''',H\text{-}4'''}$ = 3.6 Hz, $J_{H\text{-}3''',H\text{-}2'''}$ = 10.2 Hz, 1H, H3'''), 5.08 (dd, $J_{H\text{-}3'''',H\text{-}4''''}$ = 3.6 Hz, $J_{H3'''',H2''''}$ = 10.3 Hz, 1H, H-3''''), 4.98-4.88 (m, 4H, H-3'', H-3'''', CH_{2a}''-(Troc), CH_{2a}''''-(Troc)), 4.86-4.80 (m, 3H, H-2', H-2''', H-2''''), 4.72-4.69 (m, 2H, H-1''', H-1''''), 4.63-4.59 (m, 2H, H-1'', CH_{2b}''''-(Troc)), 4.57-4.54 (m, 2H, H-1, H-1'), 4.52-4.44 (m, 4H, H-1'''', CH_{2b}''-(Troc), CH_{2ab}-(Fmoc)), 4.32 (t, $J_{H\text{-}9,CH2ab}$ = 6.6 Hz, 1H, H-9-(Fmoc)), 4.27-4.25 (m, 1H, H-6_a''), 4.23-4.19 (m, 2H, H-2, H-5'''), 4.08 (m, 10H, H-5'''', H-6_{ab}', H-6_b'', H-6_{ab}''', H-6_{ab}'''', T^{α}, T^{β}), 3.88-3.85 (m, 1H, H-5''), 3.84-3.73 (m, 8H, H-3', H-4, H-4'''', H-5, H-5', H-6_a, H-6_{ab}''''), 3.68 (t, $J_{H\text{-}4'',H\text{-}3''}$ = $J_{H\text{-}4'',H\text{-}2''}$ = 9.3 Hz, 1H, H-4''), 3.56-3.48 (m, 3H, H-3, H-5'', H-6_b), 3.42-3.37 (m, 1H, H-2''''), 3.33-3.28 (m, 2H, H-2'', H-5''''), 2.10-1.84 (m, 45H, CH_3-(Ac)), 1.35 (s, 9H, *t*Bu-(Thr)), 1.09 (d, $J_{T\gamma,T\beta}$ = 6.1 Hz, 3H, T^{γ}), 0.88 (s, 9H, *t*Bu-(TBS)), 0.07, 0.06 (each s, 3H, 2 Me-(TBS)).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(C)= 39.52 ppm): ***δ*=** 170.30, 169.86, 169.68, 169.40, 169.51, 169.45, 169.27, 169.07, 169.00, 168.80, 168.66 (C=O-(Ac), C=O-(*t*Bu)), 156.76 (C=O-(Fmoc)), 154.13, 154.01 (2 C=O-(Troc)), 143.71, 143.66 (C-1_a-, C-8_a-Fmoc), 140.83, 140.79 (C-4_a-, C-5_a-Fmoc), 127.66 (C-2-, C-7-Fmoc), 127.03 (C-3-, C-6-Fmoc), 125.24, 125.14 (C-1-, C-8-Fmoc), 120.23, 120.17 (C-4-, C5-Fmoc), 101.48 (C-1'),

100.53 (C-1''), 100.00 (C-1''', C-1''''), 99.63 (C-1'''''), 99.10 (C-1), 96.16, 96.07 (C_{quart}-(Troc)), 81.40 (C_{quart}-(tBu)), 78.22 (C-3'), 76.78 (C-3), 75.89 (C-4''), 74.75 (C-4''''), 74.52 (C-5''''), 74.21 (T^{β}), 73.40 (CH_{2ab}''-(Troc)), 73.31 (CH_{2ab}''''-(Troc)), 72.80 (C-3'', C-3''''), 71.53 (C-5''), 70.65 (C-5'), 70.26 (C-3''', C-3''''', C-5), 70.04 ($C\text{-}6_{ab}$), 69.94 (C-5'''''), 69.60 (C-5'''), 69.20 (C-4'), 68.99, 68.92 (C-2', C-2''', C-2'''''), 68.58 (C-4), 67.10 (C-4'''''), 67.03 (C-4'''), 65.57 (CH_{2ab}-(Fmoc)), 61.92 ($C\text{-}6_{ab}$'), 61.76 ($C\text{-}6_{ab}$''), 61.06 ($C\text{-}6_{ab}$'''', $C\text{-}6_{ab}$'''''), 60.82 ($C\text{-}6_{ab}$'''), 59.57 (T^{α}), 55.78 (C-2''), 55.60 (C-2''''), 46.94 (C-2), 46.80 (C-9-(Fmoc)), 27.64 (*t*Bu-(Thr)), 25.76 (*t*Bu-(TBS)), 22.99, 20.69-20.33 (CH_3-(Ac)), 18.86 (T^{γ}), 17.96 (C_{quart}-(TBS), -5.23, -5.31 (2 Me(TBS)).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-acetamido-4-_O_-acetyl-2-deoxy-3-_O_-{2,4,6-tri-_O_-acetyl-3-_O_-[3,6-_O_-acetyl-2-deoxy-2-_N_-(2,2,2-trichloroethoxycarbonylamino)-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-β-D-galactopyranosyl}-6-_O_-[3,6-_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 (60)
(Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc$_2$GlcNHTroc-(1→3)-βAc$_3$Gal-(1→3)-[βAc$_4$Gal-(1→4)-βAc$_2$-GlcNHTroc-(1→6)]-αAcGalNAc)-O*t*Bu)



A solution of Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc2-TBS-GlcNHTroc-(1→3)-βAc$_3$Gal-(1→3)-[βAc$_4$Gal-(1→4)-βAc$_2$-TBS-GlcNHTroc-(1→6)]-αGalNAc)-O*t*Bu **59** (1.66 g, 0.67 mmol) in of 80% acetic acid (18 mL) was stirred for 20 h at 40°C. The solvent was removed by co-evaporation with toluene *in vacuo*. The residue was dissolved in pyridine (20 mL) and cooled in an ice bath. Then 4-(dimethylamino)pyridine (8 mg, 0.07 mmol) and acetic anhydride (10 mL) were added and the solution was stirred 18 h. The crude product was purified by flash column chromatography on silica (^{C}Hex/acetone 3:2→4:3).

Yield: 1.48 g (0.61 mmol, 91%), colorless, amorphous solid, $[\alpha]_D^{20}$ +24.30 (c = 1.00, $CHCl_3$), R_f = 0.48 (Tol/EtOAc 1:2).

$C_{99}H_{126}Cl_6N_4O_{53}$ (M = 2432.77 g/mol) [2428.54].

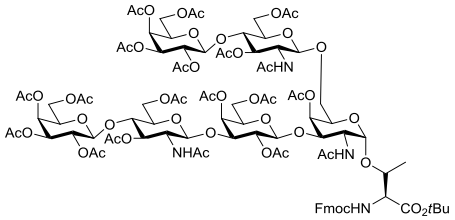
MALDI-TOF-MS (dhb, pos), m/z: 2455.04 ($[M+Na]^+$, calc. 2455.76), 2471.01 ($[M+K]^+$, calc. 2471.87).

HR-ESI-MS (pos), m/z: 1216.7809 ($[M+2H]^{2+}$, calc. 1216.7790).

^{1}H NMR (600 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm: ***δ*=** 7.91-7.89 (m, 2H, H-4-, H-5-Fmoc), 7.76-7.72 (m, 4H, H-1-, H-8-Fmoc, 2 N-H-Troc), 7.45-7.41 (m, 4H, H-2-, H-7-Fmoc, N-H-Ac, N-H-Fmoc), 7.33-7.31 (m, 2H, C-3-, C-6-Fmoc), 5.29 (S_{br}, 1H, H-4'), 5.23-5.21 (m, 3H, H-4, H-4''', H-4''''), 5.17-5.14 (m, 2H, H-3''', H-3''''), 4.97-4.89 (m, 3H, H-3'', H-3'''', CH_{2a}''-(Troc)), 4.91-4.89 (m, 1H, CH_{2a}''''-(Troc)), 4.83-4.80 (m, 2H, H-2''', H-2''''), 4.73-4.70 (m, 3H, H-1''', H-1'''', H-2'), 4.63-4.55 (m, 4H, H-1, H-1', H-1'', CH_{2b}''''-(Troc)), 4.51-4.47 (m, 4H, H-1'''', CH_{2b}''-(Troc), CH_{2ab}-(Fmoc)), 4.32-4.27 (m, 3H, H-9-(Fmoc), H-6_a'', H-6_a''''), 4.22-4.19 (m, 2H, H-5''', H-5''''), 4.14-4.09 (m, 3H, H-2, H-6_b'', T^{β}), 4.05-3.99 (m, 6H, H-5'''', H-6_{ab}''', H-6_{ab}'''', T^{α}), 3.94-3.88 (m, 3H, H-6_{ab}', H-6_b''''), 3.81-3.76 (m, 3H, H-3, H-3', H-6_a), 3.71-3.65 (m, 2H, H-4'', H-4''''), 3.61-3.58 (m, 1H, H-5''), 3.53-3.50 (m, 1H, H-5''''), 3.44 (q, $J_{H\text{-}2'''',H\text{-}3''''} = J_{H\text{-}2'''',H\text{-}4''''} = J_{H\text{-}2'''',N\text{-}H}$ = 11.5 Hz, 1H, H-2''''), 3.35-3.27 (m, 2H, H-2'', H-6_b), 2.09-1.82 (m, 51H, CH_3(Ac)), 1.35 (s, 9H, *t*Bu), 1.13 (d, $J_{T\gamma,T\beta}$ = 7.7 Hz, 3H, T^{γ}).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(C)= 39.52 ppm): ***δ*=** 170.28, 169.97, 169.90, 169.87, 169.60, 169.51, 169.30, 169.23, 169.19, 169.09, 168.80, 168.75 (C=O-(Ac), C=O-(*t*Bu)), 156.80 (C=O-(Fmoc)), 154.15, 153.99 (2 C=O-(Troc)), 143.75, 143.65 (C-1_a-, C-8_a-Fmoc), 140.82, 140.78 (C-4_a-, C-5_a), 127.72, 127.67 (C-2-, C-7-Fmoc), 127.03 (C-3-, C-6-Fmoc), 125.24, 125.13 (C-1-, C-8-Fmoc), 120.22, 120.17 (C-4-, C-5-Fmoc), 100.63 (C-1''), 100.55 (C-1'), 100.05 (C-1''''), 99.99, 99.89 (C-1''', C-1''''), 99.76 (C-1), 96.09, 96.02 (2 C_{quart}-(Troc)), 81.49 (C_{quart}-(*t*Bu)), 77.62 (C-3'), 76.37 (C-4''''), 75.93 (C-4''), 74.28 (T^{β}), 73.37, 73.34 (C-3, CH_{2ab}''-(Troc)), 73.34 (CH_{2ab}''''-(Troc)), 72.72 (C-3'', C-3''''), 71.75 (C-5''), 71.46 (C-5''''), 70.22 (C-4, C-3''', C-3'''', C-5'), 69.66, 69.60 (C-5''', C-5''''), 69.28 (C-2'), 69.20 (C-4', C-6_{ab}), 68.91 (C-2''', C-2''''), 68.87 (C-5), 67.04 (C-4''', C-4''''), 65.62 (CH_{2ab}-(Fmoc)), 62.19 (C-6_{ab}''), 61.81, 61.69 (C-6_{ab}', C-6_{ab}''''), 60.87, 60.81 (C-6_{ab}''', C-6_{ab}''''), 59.53 (T^{α}), 55.79 (C-2''), 55.64 (C-2''''), 47.57 (C-2), 46.77 (C-9-Fmoc), 27.63 (*t*Bu), 22.84, 20.64-20.32 (CH_3-(Ac)), 18.85 (T^{γ}).

***N*-(9H-fluoren-9-yl)-methoxycarbonyl-*O*-(2-acetamido-4-*O*-acetyl-2-deoxy-3-*O*-{2,4,6-tri-*O*-acetyl-3-*O*-[2-acetamido-3,6-*O*-acetyl-2-deoxy-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-β-D-galactopyranosyl}-6-*O*-[2-acetamido-3,6-*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 (61)**
(Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc$_2$GlcNAc-(1→3)-βAc$_3$Gal-(1→3)-[βAc$_4$Gal-(1→4)-βAc$_2$-GlcNAc-(1→6)]-αAcGalNAc)-O*t*Bu)



Zinc dust was activated by treatment with 1 N hydrochloric acid for a few minutes followed by washing with water, methanol and finally diethyl ether. Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc$_2$GlcNHTroc-(1→3)-βAc$_3$Gal-(1→3)-[βAc$_4$Gal-(1→4)-βAc$_2$-GlcNHTroc-(1→6)]-αAcGalNAc)-O*t*Bu) **60** (2.20 g, 0.90 mmol) was dissolved in acetic acid (30 mL) and preactivated zinc (1.18 g, 18.0 mmol) was added. The reaction was stirred at 40°C for 36 h. The solid was filtered off and washed with acetic acid. The filtrate was co-evaporated with toluene *in vacuo*. The residue was dissolved in pyridine/acetic anhydride (30 mL, 2:1). Then 11 mg (0.09 mmol) 4-(dimethylamino)pyridine was added and stirred for 18 h. The solvent was removed by co-evaporation with toluene *in vacuo* and the crude product purified by flash column chromatography on silica (EtOAc→EtOAc/MeOH 20:1).

Yield: 1.46 g (0.68 mmol, 75%), colorless, amorphous solid, $[\alpha]_D^{20}$ +21.05 (c = 0.51, $CHCl_3$), R_f = 0.21 (EtOAc/MeOH 15:1).

$C_{97}H_{128}N_4O_{51}$ (M = 2166.05 g/mol) [2164.75].

MALDI-TOF-MS (dhb, pos), m/z: 2204.37 ($[M+K]^+$, calc. 2204.15).

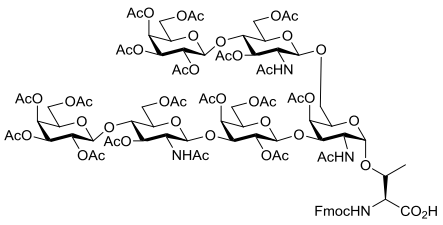
HR-ESI-MS (pos), m/z: 1083.8872 ($[M+2H]^{2+}$, calc. 1083.8868).

^{1}H NMR (600 MHz, [D$_6$]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm: ***δ*=** 7.91-7.89 (m, 2H, H-4-, H-5-Fmoc), 7.81-7.78 (m, 2H, N-H''-Ac, N-H''''-Ac), 7.76-7.72 (m, 2H, H-1-, H-8-Fmoc), 7.52 (d, $J_{N\text{-}H,H\text{-}2}$ = 9.7 Hz, 1H, N-H-Ac), 7.45-7.41 (m, 2H, H-2-, H-7-Fmoc), 7.40-7.38 (m, 1H, N-H-Fmoc), 7.34-7.29 (m, 2H, H-3-, H-6-Fmoc), 5.27 (d, ^3J(H4',H3') = 3.7 Hz, 1H, H-4'), 5.23-5.21 (m, 2H, H-4''', H-4''''), 5.19 (d, ^3J(H-4,H-3) = 1.9 Hz, 1H, H-4), 5.18-5.15 (m, 2H, H-3''', H-3''''), 5.00 (t, ^3J(H-3'',H-4'') = ^3J(H-3'',H-2'') = 9.3 Hz, 1H, H-3''), 4.91 (t, ^3J(H-3'''',H-4'''') = ^3J(H-3'''',H-2'''') = 9.3 Hz, 1H, H-3''''), 4.85-4.81 (m, 2H, H-2''', H-2''''), 4.73-4.67 (m, 4H, H-1'', H-1''', H-1'''', H-2'), 4.60 (d, ^3J(H-1,H-2) = 4.1 Hz 1H, H-1), 4.56 (d, ^3J(H-1',H-2') = 8.0 Hz, 1H, H-1'), 4.52-4.45 (m, 3H, H-1'''', CH_{2ab}-(Fmoc)), 4.33-4.26 (m, 3H, H-6$_a$'', H-6$_a$'''', H-9-Fmoc), 4.22-4.19 (m, 2H, H-5''', H-5''''), 4.15-4.11 (m, 2H, H-2, T$^\beta$), 4.08-4.05 (m, 1H, H-6$_b$''), 4.02-

3.98 (m, 6H, H-5'''', H-6_{ab}''', H-6_{ab}'''', T^{α}), 3.94-3.89 (m, 3H, H-6_{ab}', H-6_b''''), 3.84 (t, ^3J(H-5',H-6') = 5.9 Hz, 1H, H-5'), 3.80-3.74 (m, 3H, H-3', H-3'''', H-6_a), 3.70-3.63 (m, 4H, H-2'''', H-4'', H-4'''', H-5''), 3.56-3.54 (m, 1H, H-5''''), 3.39 (q, ^3J(H-2'',H-1'') = ^3J(H-2'',H-3'') = ^3J(H-2'',N-H) = 9.3 Hz, 1H, H-2''), 3.31-3.27 (m, 1H, H-6_b), 2.09-1.69 (m, 57H, CH_3-(Ac)), 1.34 (s, 1H, *t*Bu), 1.11 (d, $J_{T\gamma,T\beta}$ = 6.3 Hz, 3H, T^{γ}).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(C)= 39.52 ppm): ***δ*=** 170.28, 170.24, 170.00, 169.90, 169.87, 169.59, 169.51, 169.42, 169.42, 169.37, 169.34, 169.12, 169.10, 169.01, 168.01, 168.87, 168.83 (C=O-(Ac), C=O-(*t*Bu)), 156.80 (C=O-(Fmoc)), 143.73, 143.66 (C-1_a-, C-5_a-Fmoc), 140.82, 140.78 (C-4_a-, C-5_a-Fmoc), 127.73, 127.69 (C-2-, C-7-Fmoc), 127.03 (C-3-, C-6-Fmoc), 125.25, 125.12 (C-1-, C-8-Fmoc), 120.22, 120.17 (C-4-, C-5-Fmoc), 100.51 (C-1'), 100.42 (C-1''''), 100.02, 99.99, 99.91 (C-1'', C-1''', C-1''''), 98.44 (C-1), 81.43 (C_{quart}-(*t*Bu)), 76.81 (C-3'), 76.49 (C-4''''), 76.13 (C-4''), 73.77 (T^{β}), 73.37 (C-3, C-3''''), 72.72 (C-3''), 71.68 (C-5''), 71.43 (C-5''''), 70.22, 70.16, 70.05 (C-3''', C-3'''', C-4, C-5'), 69.62, 69.56 (C-5''', C-5''''), 69.45 (C-2'), 69.13 (C-6_{ab}), 68.99 (C-4'), 68.93, 68.91 (C-2''', C-2''''), 68.59 (C-5), 67.03, 67.02 (C-4''', C-4''''), 65.60 (CH_{2ab}-(Fmoc)), 62.23 (C-6_{ab}''), 61.56 (C-6_{ab}', C-6_{ab}''''), 60.82, 60.76 (C-6_{ab}''', C-6_{ab}''''), 59.48 (T^{α}), 53.96 (C-2''), 53.08 (C-2''''), 47.67 (C-2), 46.77 (C-9-(Fmoc)), 27.57 (*t*Bu), 22.81, 22.68, 22.62, 20.64-20.33 (CH_3-(Ac)), 19.00 (T^{γ}).

_N_-(9H-fluoren-9-yl)-methoxycarbonyl-_O_-(2-acetamido-2-deoxy-4-_O_-acetyl-3-_O_-{2,4,6-tri-_O_-acetyl-3-_O_-[2-acetamido-3,6-_O_-acetyl-2-deoxy-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-β-D-galactopyranosyl}-6-_O_-[2-acetamido-3,6-_O_-acetyl-2-deoxy-4-_O_-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-α-D-galactopyranosyl)-L-threonine (62)
(Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc$_2$GlcNAc-(1→3)-βAc$_3$Gal-(1→3)-[βAc$_4$Gal-(1→4)-βAc$_2$-GlcNAc-(1→6)]-αAcGalNAc)-OH)



To a solution of Fmoc-Thr(βAc$_4$Gal-(1→4)-βAc$_2$GlcNAc-(1→3)-βAc$_3$Gal-(1→3)-[βAc$_4$Gal-(1→4)-βAc$_2$-GlcNAc-(1→6)]-αAcGalNAc)-O*t*Bu **61** (1.82 g, 0.84 mmol) in dichloromethane (5 mL) and anisole (2 mL) was added trifluoroacetic acid (15 mL) and the solution was stirred for 3 h. The solvent was removed by co-evaporation with toluene. The crude product was

purified by flash column chromatography on silica (EtOAc→EtOAc/MeOH/AcOH/H_2O 50:3:3:2).

Yield: 1.629 mg (0.77 mmol, 92%), colorless, amorphous solid, $[\alpha]_D^{20}$ +27.60 (c = 1.00, $CHCl_3$), R_f = 0.23 (EtOAc/MeOH/AcOH/H_2O 50:3:3:2).

HR-ESI-MS (pos), m/z: 1055.8560 ($[M+2H]^{2+}$, calc. 1055.8555), 1074.8311 ($[M+K+H]^{2+}$, calc. 1074.8334).

^{1}H NMR (600 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(H)= 2.50 ppm: ***δ*=** 7.91-7.90 (m, 2H, H-4-, H-5-Fmoc), 7.83-7.79 (m, 2H, NH'''-Ac, NH''''-Ac), 7.77-7.73 (m, 2H, H-1-, H-8-Fmoc), 7.45-7.41 (m, 3H, NH-Ac, H-2-, H-7-Fmoc), 7.34-7.31 (m, 3H, NH-Fmoc, H-3-, H-6-Fmoc), 5.28 (d, ^3J(H-4',H-3') = 3.5 Hz, 1H, H-4'), 5.23-5.22 (m, 1H, H-4''', H-4''''), 5.20 (d, 1H, H-4), 5.18-5.15 (m, 2H, H-3''', H-3''''), 5.00 (t, ^3J(H-3'',H-2'') = ^3J(H-3'',H-4''), 1H, H-3'') = 9.5 Hz), 4.91 (t, ^3J(H-3'''',H-2'''') = ^3J(H-3'''',H-4'''') = 9.5 Hz, 1H, H-3''''), 4.85-4.82 (m, 2H, H-2''', H-2''''), 4.73-4.68 (m, 4H, H-1'', H-1''', H-1'''', H-2'), 4.64 (d, ^3J(H-1,H-2) = 4.1 Hz, 1H, H-1), 4.56 (d, ^3J(H-1',H-2') = 7.9 Hz 1H, H-1'), 4.52-4.43 (m, 3H, H-1'''', CH_{2ab}-(Fmoc)), 4.32-4.30 (m, 2H, H-9-Fmoc, H-6''''$_a$), 4.28-4.26 (m, 1H, H-6''$_a$), 4.23-4.20 (m, 3H, H-5''', H-5'''', T^{β}), 4.12-4.05 (m, 3H, H-2, H-6''$_b$, T^{β}), 4.02-3.98 (m, 4H, H-6'''$_{ab}$, H-6''''$_{ab}$), 3.94-3.85 (m, 3H, H-6'$_{ab}$, H-6''''$_b$, H-5'), 3.81-3.77 (m, 2H, H-3, H-3'), 3.75-3.73 (m, 1H, H-6$_a$), 3.71-3.63 (m, 4H, H-2'''', H-4'', H-4'''', H-5''), 3.56-3.54 (m, 1H, H-5''''), 3.41-3.36 (m, 1H, H-2''), 3.31-3.27 (m, 1H, H-6$_b$), 2.30, 2.09, 2.08, 2.07, 2.02, 2.01, 2.00, 1.99, 1.98, 1.95, 1.94, 1.91, 1.90, 1.89, 1.84, 1.74, 1.69 (m, 57H, CH_3-(Ac)), 1.10 (d, 3H, T^{γ}, ^3J(T^{γ},T^{β}) = 6.4 Hz).

^{13}C NMR (150.9 MHz, [D_6]DMSO, 30°C, internal DMSO *δ*(C)= 39.52 ppm): ***δ*=** 172.02, 171.80, 170.26, 170.23, 170.00, 169.89, 169.85, 169.55, 169.50, 169.41, 169.37, 169.35, 169.11, 169.09, 169.02, 168.95, 168.83 (C=O-(Ac)), 156.80 (C=O-(Fmoc)), 143.77, 143.72 (C-1$_a$-, C-8$_a$-Fmoc), 140.82, 140.77 (C-4$_a$-, C-5$_a$-Fmoc), 127.71, 127.67 (C-2-, C-7-Fmoc), 127.06 (C-3-, C-6-Fmoc), 125.27, 125.14 (C-1-, C-8-Fmoc), 120.19, 120.14 (C-4-, C-5-Fmoc), 100.60 (C-1'), 100.37 (C-1''''), 100.03, 100.00, 99.93 (C-1'', C-1''', C-1''''), 98.46 (C-1''''), 76.82 (C-3'), 76.54 (C-4''''), 76.14 (C-4''), 74.08 (T^{β}), 73.47 (C-3), 73.37 (C-3''''), 72.72 (C-3''), 71.70 (C-5''), 71.47 (C-5''''), 70.25, 70.19, 70.13 (C-4, C-3''', C-3'''', C-5'), 69.65, 69.59, 69.50 (C-2', C-5''', C-5''''), 69.00, 68.96, 68.93 (C-4', C-2''', C-2'''', C-6$_{ab}$), 68.57 (C-5), 67.06 (C-4''', C-4''''), 65.58 (CH_{2ab}-(Fmoc)), 62.27 (C-6''$_{ab}$), 61.70 (C-6', C-6''''$_{ab}$), 60.84, 60.77 (C-6'''$_{ab}$, C-6''''$_{ab}$), 58.59 (T^{α}), 53.99 (C-2''), 53.12 (C-2''''), 47.83 (C-2), 46.80 (C9-Fmoc), 22.78, 22.65, 22.61, 21.04, 20.62, 20.61, 20.57, 20.54, 20.52, 20.49, 20.45, 20.35, 20.32, 20.30 (CH_3-(Ac)), 18.68 (T^{γ}).

7.2 Syntheses of chapter 4.2

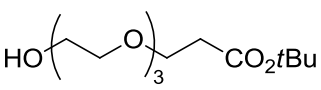
7.2.1 General

Solid phase peptide synthesis: SPPS of peptides and glycopeptides was carried out on a *Syro I* peptide synthesizer by *Multisyntech GmbH*. Solid phase resins were purchased from *Rapp Polymere GmbH*. Protected amino acid building blocks for Fmoc-SPPS (*Novabiochem*®) were purchased from *Merck KGaA*, and *Merck Schuchardt OHG*. *N*,*N*-dimethylformamide, *N*-methypyrrolidone, trifluoroacetic acid and piperidine were purchased from *Biosolve Chimie SARL*. Coupling reagents HBTU and HATU (*Novabiochem*®) were purchased from Merck KGaA. HOBt monohydrate was purchased from *Sigma-Aldrich GmbH* and recrystallized from absolute ethanol and dried at reduced pressure. HOAt was purchased from *GL Biochem*, *Shanghai*.

RP-HPLC: Analytic RP-HPLC was performed on a Dionex U-3000 (*Thermo Scientific*) system (DR-3600 six channel degasser, LPG-3x00 pump, TCC-3100 column compartment, DAD-3000 UV/VIS diode array detector, WPS-3000 autosampler). A *Luna C18(2)* column (3 µm, 100Å, 150 x 2.0 mm) from *Phenomenex* was applied for analytical HPLC.
Semi-preparative RP-HLPC was performed on the same Dionex HPLC setup as for analytical RP-HPLC with a *InertSustain C18* (5 µm, 250 x 6.0 mm) from *GL Sciences Inc*.
Preparative RP-HPLC was performed on a Dionex U-3000 system (HPG-3200P, VWD-3400 UV/VIS detector, AFC-3000 sampler). A *Luna C18(2)* column (10 µm, 100Å, 250 x 21.2 mm) from *Phenomenex* was applied for preparative HPLC. The flow rate was set to 20 mL/min.
(Glyco-)Peptides were detected at λ = 214 nm. Both systems were operated and chromatograms analyzed with *Dionex Chromeleon* (version 6.80DU10a Build 2826(171948)).

7.2.2 Synthesis of the triethylene glycol spacer amino acid

12-Hydroxy-4,7,10-trioxa-dodecanate-*tert*-butylester[275] (63)
(HO-{TEG}-COO*t*Bu)



Triethylene glycol (33.6 g, 223.7 mmol) was given into dry THF under argon atmosphere and sodium (50 mg, 2.2 mmol) was added in small portions. After the sodium had completely dissolved, *tert*-butyl-acrylate (9.3 g, 72.6 mmol) was added and stirred for 20 h. The reaction

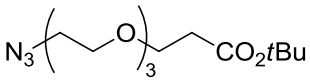
was neutralized with 1 N hydrochloric acid (2 mL) and the solvent was removed *in vacuo*. The residue was taken up with brine (70 mL) and was then extracted with ethyl acetate (four times, 50 mL each). The combined organic phases were washed with brine (30 mL), dried over magnesium sulfate and filtered. The solvent was removed *in vacuo* and the product was used further without any more purification.

Yield: 15.56 g (55.9 mmol, 77%), pale yellow liquid, R_f = 0.28 (EtOAc).

$C_{13}H_{26}O_6$ (M = 278.34 g/mol) [278.17].

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 3.67-3.53 (m, 14H, OCH_2), 2.96 (s, 1H, OH), 2.44 (t, 1H, 2-CH_2, $J_{CH2,CH2}$ = 6.5 Hz), 1.38 (s, 9H, *t*Bu).

12-Azido-4,7,10-trioxa-dodecanate-*tert*-butylester[276] (64)
(N_3-{TEG}-COO*t*Bu)



A solution of 12-hydroxy-4,7,10-trioxa-dodecansäure-*tert*-butylester **63** (15.12 g, 54.3 mmol) and triethylamine (20.1 mL, 145.3 mmol) in 20 mL dry dichloromethane was cooled in an ice-bath and methanesulfonyl chloride (10.1 mL, 132.3 mmol) was added dropwise. After stirring for 5 h, the mixture was filtered over *Celite*® and washed with dichloromethane (100 mL). The filtrate was washed with ice-water (60 mL) and with brine (two times, 50 mL). The organic phase was dried over magnesium sulfate, filtered and the solvent removed *in vacuo*. The residue was taken up with *N,N*-dimethylformamide (20 mL) and sodium azide (21.72 g, 333.8 mmol) was added. The reaction was stirred at 60°C for 18 h. The solvent was removed *in vacuo* and the residue was taken up with water (80 mL). The aqueous phase was extracted with diethyl ether (three times). The combined ether phases were dried over magnesium sulfate, filtered and the solvent removed *in vacuo*. The crude product was purified by flash column chromatography on silica (^{C}Hex/EtOAc 4:1).

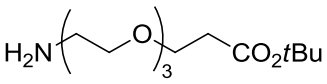
Yield: 10.92 g (36.1 mmol, 67%) colorless oil, R_f = 0.23 (^{C}Hex/EtOAc).

$C_{13}H_{25}N_3O_5$ (M = 303.35 g/mol) [303.18].

ESI-MS (pos), m/z: 326.07 ($[M+H]^+$, calc. 326.34).

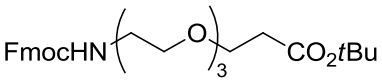
^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 3.68-3.57 (m, 12H, 6 x OCH_2), 3.36 (t, 2H, 12-OCH_2), 2.47 (t, 2H, 2-CH_2), 1.42 (s, 9H, *t*Bu).

12-Amino-4,7,10-trioxa-dodecanate-*tert*-butylester[276] (65)
(H_2N-{TEG}-COO*t*Bu)



To a suspension of a nickel-aluminum alloy (*Raney-Nickel*, 8.5 g) in water (250 mL), were given sodium hydroxide pellets until no further gas generation could be observed. The water was decanted and the solid washed with water until the supernatant remained neutral. Then it was washed several times with 2-propanol. The *Raney-Nickel* was given to a solution of N_3-{TEG}-COO*t*Bu **64** (10.81 g, 35.6 mmol) in 2-propanol (80 mL). The reaction flask was evacuated several times via a membrane pump in order to remove dissolved air, before the flask was flushed with hydrogen and stirred for 16 h. The *Raney-Nickel* was filtered off over *Celite*® and the solvent was removed *in vacuo*. The crude product was purified by flash column chromatography on silica ($Et_2O \rightarrow Et_2O$/MeOH 1:1).
Yield: 6.69 g (24.1 mmol, 68%), yellow oil.
$C_{13}H_{27}NO_5$ (M = 277.36 g/mol) [277.19].
ESI-MS (pos), m/z: 278.07 ($[M+H]^+$, calc. 278.36).
^{1}H-NMR (400 MHz, $CDCl_3$), ***δ (ppm)***: 3.63 (t, 2H, 3-CH_2, $J_{CH2,CH2}$ = 6.5 Hz), 3.59-3.52 (m, 8H, 4 x OCH_2), 3.43 (t, 2H, OCH_2, $J_{CH2,CH2}$ = 5.3 Hz), 2.79 (t, 2H, OCH_2, $J_{CH2,CH2}$ = 5.2 Hz), 2.42 (t, 2H, OCH_2, $J_{CH2,CH2}$ = 6.5 Hz), 1.71 (s_{br}, 2H, NH_2), 1.36 (s, 9H, *t*Bu).

***N*-(9H-Fluoren-9-yl)-methoxycarbonyl-amido-4,7,10-trioxa-dodecanate-*tert*-butylester[240] (66)**
(FmocHN-{TEG}-CO_2*t*Bu)

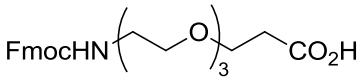


To a solution of H_2N-{TEG}-CO_2*t*Bu **65** (6.38 g, 23.0 mmol) in a mixture of water and acetone (160 mL, 1:1) was added sodium bicarbonate (2.10 g, 25.0 mmol). Fmoc-OSu (8.00 g, 23.7 mmol) was added in portions to the solution and the reaction was stirred for 18 h. Then, the reaction was acidified to pH 6 using 6 N hydrochloric acid. The acetone was removed *in vacuo* and the remaining aqueous phase was extracted with dichloromethane (four times, 100 mL each). The combined organic phases were dried over magnesium sulfate, filtered and the solvent removed in vacuo. The crude product was purified by flash column chromatography on silica (^{C}Hex/EtOAc 2:1).
Yield: 9.75 g (19.5 mmol, 85%), yellow oil, R_f = 0.26 (^{C}Hex/EtOAc 1:1).

$C_{28}H_{37}NO_7$ (M = 499.60 g/mol) [499.26].

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 7.76 (d, 2H, H4-, H5-Fmoc, $J_{H4,H3}$ = $J_{H5,H6}$ = 6.7), 7.60 (d, 2H, H1-, H8-Fmoc, $J_{H1,H2}$ = $J_{H8,H7}$ = 7.6 Hz), 7.39 (t, 2H, H2-, H7-Fmoc, $J_{H2,H3}$ = $J_{H2,H1}$ = $J_{H7,H6}$ = $J_{H7,H8}$ = 7.5 Hz), 7.33-7.29 (m, 2H, H3-, H6-Fmoc), 5.43 (s_{br}, 1H, NH), 4.40 (d, 2H, CH_2(Fmoc), $J_{CH2,H9}$ = 7.0 Hz), 4.22 (t, 1H, H9(Fmoc), $J_{H9,CH2}$ =7.0 Hz), 3.69 (t, 2H, 11-CH_2, $J_{CH2,CH2}$ = 6.5 Hz), 3.63-3.57 (m, 10, 5 x OCH_2), 3.41-3.39 (m, 2H, 12-CH_2), 2.49 (t, 2H, 2-CH_2, $J_{CH2,CH2}$ = 6.5 Hz), 1.43 (s, 9H, *t*Bu).

***N*-(9H-Fluoren-9-yl)-methoxycarbonyl-amido-4,7,10-trioxa-dodecanate[240] (67) (FmocHN-{TEG}-CO_2H)**



FmocHN-{TEG}-CO_2*t*Bu **66** (9.51 g, 19.0 mmol) was dissolved in a mixture of trifluoroacetic acid and water (22 mL, 10:1) and stirred for 2 h. The solvent was removed by co-evaporation with toluene and dichloromethane. The crude product was purified by flash column chromatography on silica (DCM/MeOH/AcOH 100:2:1). The product was co-evaporated, with toluene and dichloromethane (three times each).

Yield: 8.4 g (24.3 mmol, quant.), pale yellow oil, R_f = 0.46 (DCM/MeOH/AcOH 100:2:1).

$C_{24}H_{29}NO_7$ (M = 443.49 g/mol) [443.19].

ESI-MS (pos), m/z: 466.31 ($[M+Na]^+$, calc. 466.18).

^{1}H-NMR (400 MHz, $CDCl_3$), *δ (ppm)*: 9.69 (s_{br}, 1H, COOH), 7.76 (d, 2H, H4-, H5-Fmoc, $J_{H4,H3}$ = $J_{H5,H6}$ = 7.6), 7.60 (m, 2H, H1-, H8-Fmoc), 7.40 (t, 2H, H2-, H7-Fmoc, $J_{H2,H3}$ = $J_{H2,H1}$ = $J_{H7,H6}$ = $J_{H7,H8}$ = 7.4 Hz), 7.31 (t, 2H, H3-, H6-Fmoc, $J_{H3,H2}$ = $J_{H3,H4}$ = $J_{H6,H5}$ = $J_{H6,H7}$ = 7.4Hz), 5.57 (s_{br}, 1H, NH), 4.43-4.22 (m, 2H, CH_2(Fmoc)), 4.24-4.22 (m, 1H, H9(Fmoc)), 3.75-3.72 (m, 2H, 11-CH_2), 3.62-3.56 (m, 10H, 5 x OCH_2), 3.40-3.39 (m, 2H, 12-CH_2), 2.60 (t, 2H, 2-CH_2, $J_{CH2,CH2}$ = 6.2 Hz).

^{13}C-NMR (100.6 MHz, $CDCl_3$), *δ (ppm)*: 175.68 (COOH), 156.81 (C=O(Fmoc), 144.03 ($C1_a$-, $C8_a$-Fmoc), 141.36 ($C4_a$-, $C5_a$-Fmoc), 127.72 (C3-, C6-Fmoc), 127.11 (C2-, C7-Fmoc), 125.13 (C1-, C8-Fmoc), 120.00 (C4-, C5-Fmoc), 70.57, 70.43, 70.33, 70.12 (4 x OCH_2), 66.72 (CH_2(Fmoc)), 66.43 (11-CH_2), 47.30 (C9(Fmoc)), 40.98 (12-CH_2), 34.84 (2-CH_2).

7.2.3 General protocol for MUC1 and MUC5B solid phase glycopeptide synthesis

Automated steps of the solid phase peptide synthesis: The peptide syntheses were carried out automatically on a peptide synthesizer following a standard protocol for Fmoc solid phase peptide synthesis until the peptides were cleaved from the resin. The reservoir bottles of the peptide synthesizer were loaded with 0.5 M Fmoc-aa-OH in DMF, 0.45 M (each) HBTU and HOBt in DMF and 2 M DIPEA in NMP. The preloaded *TentaGel*-Fmoc-aa-Trt resins (batch size shown with each peptide below) were given into 2 mL synthesis reactors equipped with an filter frit and swollen with of DCM (0.5-1 mL each) for 30 min. The Fmoc-protecting group of the preloaded resin amino acid was initially cleaved through a triple addition of 20 vol% piperidine in DMF (46.2 µL per 1 µmol batch size; 2 x 3 min + 1 x 9 min; 15s vortex, 45 s break). After the deprotection, the resin was washed with DMF (46.2 µL per 1 µmol batch size; 6 x 1 min, 15 s vortex; 45 s break). In automated reaction cycles the corresponding Fmoc-aa-OH (8 eqiuv.), HBTU (7.6 eq), HOBt (8 eq) and DIPEA (16 eq) were added automatically and the reaction was shaken by vortex (40 min reaction time, 15 s vortex; 2.75 min break). A reaction cycle was concluded by Fmoc-deprotection before the next cycle was carried out.

Manual coupling steps with glycosylated amino acids: The glycosylated amino acids were dissolved in DMF (23 µL per 1 µmol batch size) in an external vessel and pre-activated with the coupling reagents HATU, HOAt and DIPEA. The glycosylated amino acids **15**, **35**, **39**, **49**, **54**, **242** were applied in 1.5 equivalents excess with 1.45 eq HATU/HOAt and 3 eq DIPEA. The glycosylated amino acids **58** and **62** were applied in 2.0 equivalents excess with 1.9 eq HATU/HOAt and 4 eq DIPEA. After 2-3 min of pre-activation the reaction mixtures were manually pipetted into the synthesis reactor and shaken by vortex (15 s vortex; 2.45 min break). The minimum reaction times for glycosylated amino acids were: **15**, **242**: 2h; **35**, **39**: 5h; **49**, **54**: 6h; **58**, **62**: 10h. All coupling reactions were performed at room temperature. After the reaction, the resin was washed with DMF (46.2 µL per 1 µmol batch size; 6 x 1 min, 15 s vortex; 45 s break). After manual coupling steps with a glycosylated amino acids the two following standard amino acids were coupled via a double coupling.

Manual coupling steps with spacer amino acid 67: The triethylene glycol spacer amino acid **67** was N-terminally linked to the mucin tandem repeat peptide sequences. Amino acid **67** was dissolved in DMF (23µL per 1 µmol batch size) in an external vessel and pre-activated with HBTU/HOBt and DIPEA. The spacer amino acid **67** was applied in 3 eq excess with 2.8 eq HBTU/HOBt and 6 eq DIPEA. After 2-3 min of pre-activation time the

reaction mixture was added to the synthesis reactor. The synthesis was occasionally shaken by vortex (15 s vortex; 2.45 min break) for 2 h at room temperature. Afterwards, the resin was washed with DMF (6 x 1 min 6 x 500 µL, 15 s vortex; 45 s break). The Fmoc-group was removed by triple addition of 20 vol% piperidine in DMF (500 µL each, 3 x 3 x 9 min; 15s vortex, 45 s break) and the resin was washed with DMF (6 x 1 min 6 x 500 µL, 15 s vortex; 45 s break).

Release of the peptides from the solid phase resin: After peptide synthesis the resin was washed with dichloromethane, isopropanol and diethyl ether (5 x 500 µL each; 5 x 1 min; 15 s vortex; 45 s break). The resin was dried in an airstream for 30 min and then transferred from the synthesis reactor into a 2 mL syringe clogged by a filter frit. The glycopeptides were cleaved from the resin by three additions of TFA/TIPS/H_2O 15:1:1 (1 x 120 min + 2 x 10 min; 15 s vortex; 2.45 min break). The combined filtrates were co-evaporated with toluene *in vacuo*. The residue was taken up with water (2 x 5 mL), 30% buffer B, 50% buffer B and 70% buffer B (5 mL each; in this order) and passed over a C18-column (*Waters Sep-Pak® Vac 6cc (1 g)*). The fractions containing buffer B were combined, the acetonitrile evaporated *in vacuo* and the aqueous residue was lyophilized to give the crude glycopeptide product.

Removal of the carbohydrate acetyl protecting groups: The crude glycopeptides carrying glycosylated amino acids **15**, **35**, **39**, **242** were dissolved in methanol (10 mL) and were treated with portions of 1% sodium methoxide in methanol until a pH of 9.5 was reached (wet pH paper). The reaction was stirred at room temperature and the reaction was followed by analytical HPLC (6-18 h).
Crude glycopeptide carrying glycans from glycosylated amino acids **49**, **54**, **58**, **62** were dissolved in water (4.5 mL) and methanol (3.9 mL) and a solution sodium methoxide in water (600 µL, 150 mM) was added. The reaction is treated carefully with injections (100-300 µL) of the sodium methoxide solution to slowly raise the pH to 11.5 (wet pH paper) and the reaction was followed by analytical HPLC (36-168 h). The reaction was acidified with 20 µL of acetic acid and the solvents evaporated *in vacuo*. The crude glycopeptide was purified by preparative HPLC and the product was lyophilized from water.

7.2.4 Synthesis of MUC1 glycopeptides

All MUC1-glycopeptides **68**-**133** were synthesized following the general procedure reported in chapter 6.3.2.1. For MUC1 sequences, *TentaGel* R Fmoc-Ala-Trt resin (loading: 0.17 mmol/g; *Rapp Polymere, Tübingen*) was used. HPLC eluents were composed of gradients of

modifier A: water + 0.1% TFA and modifier B: 84% acetonitrile + 0.1% TFA. Peptide bonds were detected at λ = 214 nm during HPLC. The threonine T_N-antigen building block **242** in glycopeptides **68**-**74** and **131**-**133** and the serine building block **243** in glycopeptides **131**-**133** were kindly provided by Dr. H. Cai.

MUC1(19mer) T_N-antigen: H_2N-(TEG)-PAHGVT*SAPDTRPAPGSTA-OH (68)

The glycopeptide was synthesized on 58.5 mg (10 µmol) of resin.

Analytical HPLC R_t = 17.85 min (A/B, (95:5) → (70:30), 200 µL/min, 25 min); Preparative HPLC R_t = 15.27 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 733.0323 ($[M+3H]^{3+}$, calc. 733.0319); *Yield*: 67% (14.6 mg, 6.7 µmol).

MUC1(19mer) T_N-antigen: H_2N-(TEG)-PAHGVTSAPDT*RPAPGSTA-OH (69)

The glycopeptide was synthesized on 58.5 mg (10 µmol) of resin.

Analytical HPLC R_t = 18.30 min (A/B, (95:5) → (70:30), 200 µL/min, 25 min); Preparative HPLC R_t = 15.91 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 733.0320 ($[M+3H]^{3+}$, calc. 733.0319); *Yield*: 65% (14.3 mg, 6.5 µmol).

MUC1(19mer) T_N-antigen: H_2N-(TEG)-PAHGVTSAPDTRPAPGST*A-OH (70)

The glycopeptide was synthesized on 58.5 mg (10 µmol) of resin.

Analytical HPLC R_t = 18.95 min (A/B, (95:5) → (70:30), 200 µL/min, 25 min); Preparative HPLC R_t = 16.12 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 733.0324 ($[M+3H]^{3+}$, calc. 733.0319); *Yield*: 70% (15.4 mg, 7.0 µmol).

MUC1(19mer) T_N-antigen: H_2N-(TEG)-PAHGVT*SAPDT*RPAPGSTA-OH (71)

The glycopeptide was synthesized on 58.5 mg (10 µmol) of resin.

Analytical HPLC R_t = 17.52 min (A/B, (95:5) → (70:30), 200 µL/min, 25 min); Preparative HPLC R_t = 14.72 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 800.7258 ($[M+3H]^{3+}$, calc. 800.7251); *Yield*: 51% (12.2 mg, 5.1 µmol).

MUC1(19mer) T_N-antigen: H_2N-(TEG)-PAHGVT*SAPDTRPAPGST*A-OH (72)

The glycopeptide was synthesized on 58.5 mg (10 µmol) of resin.

Analytical HPLC R_t = 17.85 min (A/B, (95:5) → (70:30), 200 µL/min, 25 min); Preparative HPLC R_t = 15.08 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 800.7260 ($[M+3H]^{3+}$, calc. 800.7251); *Yield*: 52% (12.5 mg, 5.3 µmol).

MUC1(19mer) T_N-antigen: H_2N-(TEG)-PAHGVTSAPDT*RPAPGST*A-OH (73)

The glycopeptide was synthesized on 58.5 mg (10 µmol) of resin.

Analytical HPLC R_t = 18.46 min (A/B, (95:5) → (70:30), flow: 200 µL/min, 25 min); Preparative HPLC R_t = 15.64 min (A/B, (95:5) → (70:30), flow: 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 800.7260 ($[M+3H]^{3+}$, calc. 800.7251); *Yield*: 56% (13.4 mg, 5.6 µmol).

MUC1(19mer) T_N-antigen: H_2N-(TEG)-PAHGVT*SAPDT*RPAPGST*A-OH (74)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 17.34 min (A/B, (95:5) → (70:30), 200 µL/min, 25 min); Preparative HPLC R_t = 14.52 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 868.4197 ($[M+3H]^{3+}$, calc. 868.4182), 651.5677 ($[M+4H]^{4+}$, calc. 651.5655); *Yield*: 40% (10.4 mg, 4.0 µmol).

MUC1(19mer) T-antigen: H_2N-(TEG)-PAHGVT*SAPDTRPAPGSTA-OH (75)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 16.42 min (A/B, (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 14.49 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 787.0485 ($[M+3H]^{3+}$, calc. 787.0495), 600.0255 ($[M+4H]^{4+}$, calc. 600.0279); *Yield*: 85% (26.2 mg, 11.1 µmol).

MUC1(19mer) T-antigen: H_2N-(TEG)-PAHGVTSAPDT*RPAPGSTA-OH (76)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 17.33 min (A/B, (95:5) → (50:50), 200 µL/min, 45 min); Preparative HPLC R_t = 15.69 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 787.0495 ($[M+3H]^{3+}$, calc. 787.0485); *Yield*: 85% (25.9 mg, 11.0 µmol).

MUC1(19mer) T-antigen: H_2N-(TEG)-PAHGVTSAPDTRPAPGST*A-OH (77)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 17.46 min (A/B, (95:5) → (50:50), 200 µL/min, 50 min); Preparative HPLC R_t = 15.98 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 787.0486 ($[M+3H]^{3+}$, calc. 787.0495); *Yield*: 76% (23.4 mg, 9.9 µmol).

MUC1(19mer) T-antigen: H_2N-(TEG)-PAHGVT*SAPDT*RPAPGSTA-OH (78)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 15.6 min (A/B, (95:5) → (65:35), 200 µL/min, 35 min); Preparative HPLC R_t = 13.35 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 908.7603 ($[M+3H]^{3+}$, calc. 908.7603); *Yield*: 67% (23.8 mg, 8.7 µmol).

MUC1(19mer) T-antigen: H_2N-(TEG)-PAHGVT*SAPDTRPAPGST*A-OH (79)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 15.19 min (grad.: water/84% acetonitrile + 0.1% TFA (95:5) → (65:35), 200 µL/min, 35 min, wavelength = 214 nm); Preparative HPLC R_t = 13.67 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 908.7610 ($[M+3H]^{3+}$, calc. 908.7603); *Yield*: 76% (27.0 mg, 9.9 µmol).

MUC1(19mer) T-antigen: H_2N-(TEG)-PAHGVTSAPDT*RPAPGST*A-OH (80)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 16.39 min (A/B, (95:5) → (65:35), 200 µL/min, 35 min); Preparative HPLC R_t = 14.63 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 908.7598 ($[M+3H]^{3+}$, calc. 908.7603) 691.3084 ($[M+3H+K]^{4+}$, calc. 691.3110), 553.2481 ($[M+4H+K]^{5+}$, calc553.2502); *Yield*: 62% (22.1 mg, 8.1 µmol).

MUC1(19mer) T-antigen: H_2N-(TEG)-PAHGVT*SAPDT*RPAPGST*A-OH (81)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 14.51 min (A/B, (95:5) → (65:35), flow: 200 µL/min, 35 min); Preparative HPLC R_t = 12.78 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1030.4723 ($[M+3H]^{3+}$, calc 1030.4723); *Yield*: 58% (23.6 mg, 7.6 µmol).

MUC1(20mer) T-antigen: H_2N-(TEG)-PGSTAPPAHGVTSAPDT*RPA-OH (82)

The glycopeptide was synthesized on 66.7 mg (10 µmol) of resin.

Analytical HPLC R_t = 18.95 min (A/B, (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 15.64 min (A/B, (95:5) → (65:35), 20 mL/min, 30 min); *HR-ESI-MS*, *m/z*: 819.4007 ($[M+3H]^{3+}$, calc. 819.4005), *Yield*: 71% (16.7 mg, 7.1 µmol).

MUC1(21) T-antigen: H_2N-(TEG)-APDT*RPAPGSTAPPAHGVTSA-OH (83)

The glycopeptide was synthesized on 66.7 mg (10 µmol) of resin.

Analytical HPLC R_t = 18.88 min (A/B, (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 15.67 min (A/B, (95:5) → (65:35), 20 mL/min, 30 min); *HR-ESI-MS*, *m/z*: 843.0798 ($[M+3H]^{3+}$, calc.843.0795), *Yield*: 54% (13.1 mg, 5.3 µmol).

MUC1(19mer) T-antigen: H_2N-(TEG)-APDT*RPA-OH (84)

The glycopeptide was synthesized on 46.7 mg (7 µmol) of resin.

Analytical HPLC R_t = 15.68 min (A/B, (95:5) → (79:21), flow: 200 µL/min, 54 min); Preparative HPLC R_t = 11.89 min (A/B, (95:5) → (75:25), flow: 20 mL/min, 20 min); *HR-ESI-MS*, *m/z*: 648.3143 ($[M+2H]^{2+}$, calc. 648.3154); *Yield*: 66% (5.9 mg, 4.6 µmol).

MUC1(19mer) type-1 core 3: H_2N-(TEG)-PAHGVT*SAPDTRPAPGSTA-OH (85)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 18.65 min (A/B, (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 14.89 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1099.0464 ([M-{Galβ(1,3)-GlcNac}+2H]$^{2+}$, calc. 1099.0443); 854.7428 ([M+3H]$^{3+}$, calc. 854.7427); 650.7965 ([M+K+3H]$^{4+}$, calc. 650.7978); *Yield*: 60% (20.0 mg, 7.81 µmol).

MUC1(19mer) type-1 core 3: H_2N-(TEG)-PAHGVTSAPDT*RPAPGSTA-OH (86)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 19.66 min (A/B, (95:5) → (70:30), 200 µL/min, 30 min); Preparative HPLC R_t = 15.35 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1099.0464 ([M-{Galβ(1,3)-GlcNac}+2H]$^{2+}$, calc. 1099.0443); 854.7429 ([M+3H]$^{3+}$, calc. 854.7427); 650.7963 ([M+K+3H]$^{4+}$, calc. 650.7978); *Yield*: 74% (24.2 mg, 9.6 µmol).

MUC1(19mer) type-1 core 3: H_2N-(TEG)-PAHGVTSAPDTRPAPGST*A-OH (87)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 20.34 min (A/B, (95:5) → (70:30), 200 µL/min, 30 min); Preparative HPLC R_t = 15.87 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1098.5453 ([M-{Galβ(1,3)-GlcNac}+2H]$^{2+}$, calc. 1099.0443); 862.0710 ([M+Na+2H]$^{3+}$, 862.0700); 854.7430 ([M+3H]$^{3+}$, calc. 854.7427); 646.8051 ([M+K+3H]$^{4+}$, calc. 646.8043); *Yield*: 79% (26.4 mg, 10.3 µmol).

MUC1(19mer) type-1 core 3: H_2N-(TEG)-PAHGVT*SAPDT*RPAPGSTA-OH (88)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 33.36 min (A/B, (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 12.98 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS (Orbitrap Fusion)*, *m/z*: 1044.1473 ([M+3H]$^{3+}$, calc. 1044.1465); 783.3623 ([M+4H]$^{4+}$, calc. 783.3617); *Yield*: 37% (15.1 mg, 4.8 µmol).

MUC1(19mer) type-1 core 3: H_2N-(TEG)-PAHGVT*SAPDTRPAPGST*A-OH (89)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 34.70 min (A/B, (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 13.44 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1044.1491 ([M+3H]$^{3+}$, calc. 1044.1465); 783.3648 ([M+4H]$^{4+}$, calc. 783.3617); *Yield*: 75% (30.5 mg, 9.7 µmol).

MUC1(19mer) type-1 core 3: H_2N-(TEG)-PAHGVTSAPDT*RPAPGST*A-OH (90)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 38.93 min (A/B, (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 14.49 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1044.1489 ($[M+3H]^{3+}$, calc. 1044.1465); 783.3647 ($[M+4H]^{4+}$, calc. 783.3617); *Yield*: 47% (14.8 mg, 4.7 µmol).

MUC1(19mer) type-1 core 3: H_2N-(TEG)-PAHGVT*SAPDT*RPAPGST*A-OH (91)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 31.50 min (A/B, (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 12.52 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1233.5506 ($[M+3H]^{3+}$, calc. 1233.5504), 925.4160 ($[M+4H]^{4+}$, calc. 925.4146); *Yield*: 62% (29.8 mg, 8.1 µmol).

MUC1(19mer) type-2 core 3: H_2N-(TEG)-PAHGVT*SAPDTRPAPGSTA-OH (92)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 15.91 min (A/B, (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 14.03 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1099.0472 ($[M-\{Gal\beta(1,4)\text{-}GlcNac\}+2H]^{2+}$, calc. 1099.0443); 997.0067 ($[M-\{Gal\beta(1,4)\text{-}GlcNac\beta(1,3)\text{-}GalNAc\}+2H]^{2+}$, calc. 997.0029); 854.7432 ($[M+3H]^{3+}$, calc. 854.7427), 366.1399 ($[\beta Gal(1,4)\text{-}\beta GlcNac]^{+}$, calc. 366.1400); *Yield*: 70% (23.2 mg, 9.1 µmol).

MUC1(19mer) type-2 core 3: H_2N-(TEG)-PAHGVTSAPDT*RPAPGSTA-OH (93)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 16.80 min (A/B, (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 14.94 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1099.0455 ($[M-\{Gal\beta(1,4)\text{-}GlcNac\}+2H]^{2+}$, calc. 1099.0443); 997.0057 ($[M-\{Gal\beta(1,4)\text{-}GlcNac\beta(1,3)\text{-}GalNAc\}+2H]^{2+}$, calc. 997.0029); 854.7432 ($[M+3H]^{3+}$, calc. 854.7427); 650.7965 ($[M+K+3H]^{4+}$, calc. 650.7978), 366.1398 ($[\beta Gal(1,4)\text{-}\beta GlcNac]^{+}$, calc. 366.1400); *Yield*: 65% (21.4 mg, 8.4 µmol).

MUC1(19mer) type-2 core 3: H_2N-(TEG)-PAHGVTSAPDTRPAPGST*A-OH (94)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 20.34 min (A/B, (95:5) → (70:30), 200 µL/min, 30 min); Preparative HPLC R_t = 15.87 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1099.0450 ($[M-\{Gal\beta(1,4)\text{-}GlcNac\}+2H]^{2+}$, calc. 1099.0443); 997.5067 ([M-{Galβ(1,4)-

GlcNacβ(1,3)-GalNAc}+2H]$^{2+}$, calc. 997.5046); 854.7431 ([M+3H]$^{3+}$, calc. 854.7427); *Yield*: 72% (23.7 mg, 9.3 µmol).

MUC1(19mer) type-2 core 3: H_2N-(TEG)-PAHGVT*SAPDT*RPAPGSTA-OH (95)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 14.76 min (A/B, (95:5) → (65:35), 200 µL/min, 35 min); Preparative HPLC R_t = 12.87 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1383.1496 ([M-{Galβ(1,4)-βGlcNac}+2H]$^{2+}$, calc. 1383.1500); 1200.5839 ([M-{Galβ(1,4)-GlcNac}+2H]$^{2+}$, calc. 1200.5839); 1044.1468 ([M+3H]$^{3+}$, calc. 1044.1465), 922.4361 ([M-{Galβ(1,4)-GlcNac}+3H]$^{3+}$, calc. 922.4358), 366.1395 ([Galβ(1,4)-GlcNac]$^{+}$, calc. 366.1400); *Yield*: 74% (30.4 mg, 9.7 µmol).

MUC1(19mer) type-2 core 3: H_2N-(TEG)-PAHGVT*SAPDTRPAPGST*A-OH (96)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 15.63 min (A/B, (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 13.55 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1044.1503 ([M+3H]$^{3+}$, calc. 1044.1465); 783.3658 ([M+4H]$^{4+}$, calc. 783.3617); *Yield*: 89% (44.0 mg, 11.6 µmol).

MUC1(19mer) type-2 core 3: H_2N-(TEG)-PAHGVTSAPDT*RPAPGST*A-OH (97)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 15.98 min (A/B, (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 14.47 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1044.1502 ([M+3H]$^{3+}$, calc. 1044.1465); 783.3655 ([M+4H]$^{4+}$, calc. 783.3617); *Yield*: 89% (44.0 mg, 11.6 µmol).

MUC1(19mer) type-2 core 3: H_2N-(TEG)-PAHGVT*SAPDT*RPAPGST*A-OH (98)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 14.13 min (A/B, (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 12.25 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1233.8912 ([M+3H]$^{3+}$, calc. 1233.8848), 1111.8482 (M-{Galβ(1,4)-GlcNac}+3H]$^{3+}$, calc. 1111.8396), 1124.4968 (M-{Galβ(1,4)-GlcNac}+Na+2H]$^{2+}$, calc. 1124.4916), 944.3933 ([M+2K+2H])$^{4+}$, calc. 944.3925), 935.1564 ([M+K+3H])$^{4+}$, calc. 935.1544), 925.4198 ([M+4H]$^{4+}$, calc. 925.4146), 755.9159 ([M+2K+3H]$^{5+}$, calc. 755.9162), 748.1260 ([M+5H]$^{5+}$, calc. 748.1243), 366.1402 ([Galβ(1,4)-GlcNac]$^{+}$, calc. 366.1400); *Yield*: 62% (21.0 mg, 8.1 µmol).

MUC1(19mer) type-2 core 3: H_2N-(TEG)-PAHGVT*SA-OH (99)

The glycopeptide was synthesized on 46.7 mg (7 µmol) of resin.

Analytical HPLC R_t = 14.826 min (A/B, (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 12.37 min (A/B, (95:5) → (75:25), 20 mL/min, 20 min); *HR-ESI-MS*, *m/z*: 755.8554 ($[M+2H]^{2+}$, calc. 755.8540); *Yield*: 24% (2.51 mg, 1.7 µmol).

MUC1(19mer) type-1 core 1: H_2N-(TEG)-PAHGVT*SAPDTRPAPGSTA-OH (100)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 36.32min (A/B, (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 14.03 min (A/B, (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 908.7662 ($[M+3H]^{3+}$, calc. 908.7603); *Yield*: 40% (14.3 mg, 5.2 µmol).

MUC1(19mer) type-1 core 1: H_2N-(TEG)-PAHGVTSAPDT*RPAPGSTA-OH (101)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 40.72 min (A/B, (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 46.12 min (A/B, (95:5) → (82:18), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 908.7633 ($[M+3H]^{3+}$, calc. 908.7603); 681.8250 ($[M+4H]^{4+}$, calc. 681.8220); *Yield*: 45% (16.2 mg, 5.9 µmol).

MUC1(19mer) type-1 core 1: H_2N-(TEG)-PAHGVTSAPDTRPAPGST*A-OH (102)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 43.20 min (A/B, (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 41.25 min (A/B, (95:5) → (79:21), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 908.7615 ($[M+3H]^{3+}$, calc. 908.7603); 681.8239 ($[M+4H]^{4+}$, calc. 681.8220); *Yield*: 43% (15.2 mg, 5.6 µmol).

MUC1(19mer) type-1 core 1: H_2N-(TEG)-PAHGVT*SAPDT*RPAPGSTA-OH (103)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 35.04 min (A/B, (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 29.07 min (A/B, (95:5) → (79:21), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1152.1842 ($[M+3H]^{3+}$, calc. 1152.1817); 864.3899 ($[M+4H]^{4+}$, calc. 864.3881); *Yield*: 36% (16.1 mg, 4.7 µmol).

MUC1(19mer) type-1 core 1: H_2N-(TEG)-PAHGVT*SAPDTRPAPGST*A-OH (104)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 35.88 min (A/B, (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 32.64 min (A/B, (95:5) → (79:21), flow: 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*:

1152.1831 ([M+3H]$^{3+}$, calc. 1152.1817); 864.3892 ([M+4H]$^{4+}$, calc. 864.3881); *Yield*: 35% (15.4 mg, 4.5 µmol).

MUC1(19mer) type-1 core 1: H_2N-(TEG)-PAHGVTSAPDT*RPAPGST*A-OH (105)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 38.81 min (A/B, (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 34.15 min (A/B, (95:5) → (79:21), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1152.1839 ([M+3H]$^{3+}$, calc. 1152.1817); 864.3893 ([M+4H]$^{4+}$, calc. 864.3881); *Yield*: 36% (16.3 mg, 4.7 µmol).

MUC1(19mer) type-1 core 1: H_2N-(TEG)-PAHGVT*SAPDT*RPAPGST*A-OH (106)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 32.87 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 27.15 min (A/B: (95:5) → (79:21), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1423.2668 ([M+2Na+K]$^{3+}$, calc. 1423.2442 ([M+3H]$^{3+}$, calc. 1395.9376), 1067.7021 ([M+2Na+K+H]$^{4+}$, calc. 1067.6850), 1051.4628 ([M+NH_4+3H]$^{4+}$, calc. 1051.4617), 1047.2058 ([M+4H]$^{4+}$, calc. 1047.2051); *Yield*: 27% (14.5 mg, 3.5 µmol).

MUC1(19mer) type-2 core 1: H_2N-(TEG)-PAHGVT*SAPDTRPAPGSTA-OH (107)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 38.86 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 33.31 min (A/B: (95:5) → (79:21), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 908.7606 ([M+3H]$^{3+}$, calc. 908.7603); 691.3096 ([M+K+3H]$^{4+}$, calc. 691.3110); *Yield*: 49% (17.4 mg, 6.4 µmol).

MUC1(19mer) type-2 core 1: H_2N-(TEG)-PAHGVTSAPDT*RPAPGSTA-OH (108)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 41.53 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 38.52 min (A/B: (95:5) → (79:21), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 908.7621 ([M+3H]$^{3+}$, calc. 908.7603); 681.8247 ([M+4H]$^{4+}$, calc. 681.8220); *Yield*: 53% (18.9 mg, 6.9 µmol).

MUC1(19mer) type-2 core 1: H_2N-(TEG)-PAHGVTSAPDTRPAPGST*A-OH (109)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 42.95 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 38.39 min (A/B: (95:5) → (79:21), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*:

908.76199 ([M+3H]$^{3+}$, calc. 908.7603); 681.8245 ([M+4H]$^{4+}$, calc. 681.8220); *Yield*: 45% (15.8 mg, 5.8 µmol).

MUC1(19mer) type-2 core 1: H_2N-(TEG)-PAHGVT*SAPDT*RPAPGSTA-OH (110)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 31.33 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 38.84 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1152.1847 ([M+3H]$^{3+}$, calc. 1152.1817); 869.8854 ([M+4H]$^{4+}$, calc. 869.8836), 864.3898 ([M+4H]$^{4+}$, calc. 864.3881); *Yield*: 34% (15.3 mg, 4.4 µmol).

MUC1(19mer) type-2 core 1: H_2N-(TEG)-PAHGVT*SAPDTRPAPGST*A-OH (111)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 35.57 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 33.41 min (A/B: (95:5) → (82:18), 20 mL/min, 49 min); *HR-ESI-MS*, *m/z*: 1152.5247 ([M+3H]$^{3+}$, calc. 1152.5247); 864.6428 ([M+4H]$^{4+}$, calc. 864.6390); *Yield*: 26% (11.6 mg, 4.5 µmol).

MUC1(19mer) type-2 core 1: H_2N-(TEG)-PAHGVTSAPDT*RPAPGST*A-OH (112)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 38.80 min ((95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 36.41 min (A/B: (95:5) → (79:21), flow: 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1152.1841 ([M+3H]$^{3+}$, calc. 1152.1817); 864.3894 ([M+4H]$^{4+}$, calc. 864.3881); *Yield*: 29% (12.8 mg, 3.7 µmol).

MUC1(19mer) type-2 core 1: H_2N-(TEG)-PAHGVT*SAPDT*RPAPGST*A-OH (113)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 32.26 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 28.57 min (A/B: (95:5) → (79:21), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1395.9435 ([M+3H]$^{3+}$, calc. 1395.9376), 1047.2065 ([M+4H]$^{4+}$, calc. 1047.2051); *Yield*: 17% (9.0 mg, 2.2 µmol).

MUC1(19mer) type-2 core 1: H_2N-(TEG)-GST*APPAHGVTSAPDTRPA-OH (114)

The glycopeptide was synthesized on 53.3 mg (10 µmol) of resin.

Analytical HPLC R_t = 37.31 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 42.17 min (A/B: (95:5) → (82:18), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 908.7609 ([M+3H]$^{3+}$, calc. 908.7603); *Yield*: 14% (3.1 mg, 1.1 µmol).

MUC1(20mer) type-2 core 1: H_2N-(TEG)-PDTRPAGST*APPAHGVTSA-OH (115)

The glycopeptide was synthesized on 53.3 mg (8 µmol) of resin.

Analytical HPLC R_t = 39.38 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 45.54 min (A/B: (95:5) → (82:18), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 941.1122 ($[M+3H]^{3+}$, 941.1112 calc.); *Yield*: 41% (9.4 mg, 3.3 µmol).

MUC1(20mer) type-2 core 1: H_2N-(TEG)-PDTRPAGSTAPPAHGVT*SA-OH (116)

The glycopeptide was synthesized on 53.3 mg (8 µmol) of resin.

Analytical HPLC R_t = 39.02 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 44.37 min (A/B: (95:5) → (82:18), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 941.1125 ($[M+3H]^{3+}$, 941.1112 calc.); *Yield*: 40% (9.1 mg, 3.2 µmol).

MUC1(19mer) type-1 core 2: H_2N-(TEG)-PAHGVT*SAPDTRPAPGSTA-OH (117)

The glycopeptide was synthesized on 47.1 mg (8 µmol) of resin.

Analytical HPLC R_t = 33.24 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 34.12 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1030.4716 ($[M+3H]^{3+}$, calc. 1030.4710); 777.3636 ($[M+NH_4+3H]^{4+}$, calc. 777.3617); *Yield*: 36% (9.1 mg, 2.9 µmol).

MUC1(19mer) type-1 core 2: H_2N-(TEG)-PAHGVTSAPDT*RPAPGSTA-OH (118)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 37.54 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 36.80 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1030.4734 ($[M+3H]^{3+}$, calc. 1030.4710); 773.1071 ($[M+4H]^{4+}$, calc. 773.1051); *Yield*: 34% (10.6 mg, 3.4 µmol).

MUC1(19mer) type-1 core 2: H_2N-(TEG)-PAHGVTSAPDTRPAPGST*A-OH (119)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 40.91 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 41.18 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1030.4717 ($[M+3H]^{3+}$, calc. 1030.4710); 773.1062 ($[M+4H]^{4+}$, calc. 773.1051); *Yield*: 37% (11.5 mg, 3.7 µmol).

MUC1(19mer) type-1 core 2: H_2N-(TEG)-PAHGVT*SAPDT*RPAPGSTA-OH (120)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 29.84 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 27.75 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*:

1395.9406 ([M+3H]$^{3+}$, calc. 1395.9391); 1047.2064 ([M+4H]$^{4+}$, calc. 1047.2062); *Yield*: 22% (9.3 mg, 2.2 µmol).

MUC1(19mer) type-1 core 2:H$_2$N-(TEG)-PAHGVT*SAPDTRPAPGST*A-OH (121)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 31.83 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 30.32 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1395.9406 ([M+3H]$^{3+}$, calc. 1395.9439); 1047.2073 ([M+4H]$^{4+}$, calc. 1047.2062); *Yield*: 20% (8.4 mg, 2.0 µmol).

MUC1(19mer) type-1 core 2: H$_2$N-(TEG)-PAHGVTSAPDT*RPAPGST*A-OH (122)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 34.49 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 32.69 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1395.9427 ([M+3H]$^{3+}$, calc. 1395.9439); 1047.2072 ([M+4H]$^{4+}$, calc. 1047.2062); *Yield*: 20% (8.3 mg, 2.0 µmol).

MUC1(19mer) type-1 core 2: H$_2$N-(TEG)-PAHGVT*SAPDT*RPAPGST*A-OH (123)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 27.69 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 24.15 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1321.0576 ([M+4H]$^{4+}$, calc. 1321.0542), 1057.2493 ([M+5H]$^{5+}$, calc. 1057.2455); *Yield*: 9% (4.8 mg, 0.9 µmol).

MUC1(19mer) type-2 core 2: H$_2$N-(TEG)-PAHGVT*SAPDTRPAPGSTA-OH (124)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 36.45 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 31.71 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1057.7999 ([M+Na+2H]$^{3+}$, calc. 1057.7969), ([M+3H]$^{3+}$, calc. 1030.4710); 773.1058 ([M+4H]$^{4+}$, calc. 773.1051); *Yield*: 43% (13.2 mg, 4.3 µmol).

MUC1(19mer) type-2 core 2: H$_2$N-(TEG)-PAHGVTSAPDT*RPAPGSTA-OH (125)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 38.36 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 36.59 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1030.4716 ([M+3H]$^{3+}$, calc. 1030.4710); 773.1068 ([M+4H]$^{4+}$, calc. 773.1051); *Yield*: 34% (10.6 mg, 3.4 µmol).

MUC1(19mer) type-2 core 2: H_2N-(TEG)-PAHGVTSAPDTRPAPGST*A-OH (126)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 40.94 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 40.43 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1030.4726 ($[M+3H]^{3+}$, calc. 1030.4710); 773.1067 ($[M+4H]^{4+}$, calc. 773.1051); *Yield*: 40% (12.4 mg, 4.0 µmol).

MUC1(19mer) type-2 core 2: H_2N-(TEG)-PAHGVT*SAPDT*RPAPGSTA-OH (127)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 30.19 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 27.62 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1051.4627 ($[M+NH_4+H]^{4+}$, calc. 1051.4606), 1395.9391 ($[M+3H]^{3+}$, calc. 1395.9391); 1047.2065 ($[M+4H]^{4+}$, calc. 1047.2062); *Yield*: 18% (7.5 mg, 2.2 µmol).

MUC1(19mer) type-2 core 2: H_2N-(TEG)-PAHGVT*SAPDTRPAPGST*A-OH (128)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 32.06 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 30.09 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1395.9425 ($[M+3H]^{3+}$, calc. 1395.9439); 1047.2076 ($[M+4H]^{4+}$, calc. 1047.2062); 845.5567 ($[M+K+4H]^{5+}$, calc. 845.5567); *Yield*: 8% (3.4 mg, 0.8 µmol).

MUC1(19mer) type-2 core 2: H_2N-(TEG)-PAHGVTSAPDT*RPAPGST*A-OH (129)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 33.61 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 33.09 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1395.9433 ($[M+3H]^{3+}$, calc. 1395.9439); 1047.2074 ($[M+4H]^{4+}$, calc. 1047.2062); *Yield*: 14% (5.8 mg, 1.4 µmol).

MUC1(19mer) type-2 core 2: H_2N-(TEG)-PAHGVT*SAPDT*RPAPGST*A-OH (130)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 26.71 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 24.59 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1321.3103 ($[M+4H]^{4+}$, calc. 1321.3050), 1057.0485 ($[M+5H]^{5+}$, calc. 1057.0448); *Yield*: 3% (1.7 mg, 0.3 µmol).

MUC1(22mer) type-2 core 1 (Thr^{11}) + 2xTn (Ser^{15}, Thr^{16}):

H_2N-(TEG)-PAHGVTSAPDT*RPAPGS*T*APPA-OH (131)

The glycopeptide was synthesized on 66.7 mg (10 µmol) of resin (0.15 mmol/g).

Analytical HPLC R_t = 40.42 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 44.71 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1132.5323 ($[M+3H]^{3+}$, calc. 1132.5274), 849.6495 ($[M+4H]^{4+}$, calc. 849.6474); *Yield*: 16% (5.3 mg, 1.6 µmol).

MUC1(22mer) type-2 core 3 (Thr^{11}) + 2xTn (Ser^{15}, Thr^{16})

H_2N-(TEG)-PAHGVTSAPDT*RPAPGS*T*APPA-OH (132)

The glycopeptide was synthesized on 66.7 mg (10 µmol) of resin (0.15 mmol/g).

Analytical HPLC R_t = 39.31 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 44.72 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1078.5148 ($[M+3H]^{3+}$, calc. 1078.5098), 809.1357 ($[M+4H]^{4+}$, calc. 809.1342); *Yield*: 19% (6.1 mg, 1.9 µmol).

MUC1(22mer) type-2 core 2 (Thr^{11}) + 2xTn (Ser^{15}, Thr^{16})

H_2N-(TEG)-PAHGVTSAPDT*RPAPGS*T*APPA-OH (133)

The glycopeptide was synthesized on 66.7 mg (10 µmol) of resin (0.15 mmol/g).

Analytical HPLC R_t = 36.18 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 41.24 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1254.2446 ($[M+3H]^{3+}$, calc. 1254.2381), 940.9311 ($[M+4H]^{4+}$, calc. 940.9304); *Yield*: 8% (2.9 mg, 0.77 µmol).

7.2.5 Synthesis of MUC5B glycopeptides

The MUC5B glycopeptides **134**-**174** were synthesized following the general procedure reported in chapter 6.3.2.1. For MUC5B sequences, *TentaGel* R Fmoc-Pro-Trt resin (loading: 0.19 mmol/g) was used.

MUC5B(13mer) T-antigen: H_2N-(TEG)-ATPSSTPGT*THTP-OH (134)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 15.11 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 13.65 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 911.9250 ($[M+2H]^{2+}$, calc. 911.9257), 608.2862 ($[M+3H]^{3+}$, calc. 608.2862); *Yield*: 78% (18.4 mg, 10.1 µmol).

MUC5B(13mer) T-antigen: H_2N-(TEG)-AT*PSSTPGTTHTP-OH (135)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 15.15 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 13.85 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 911.9263 ($[M+2H]^{2+}$, calc. 911.9257), 608.2869 ($[M+3H]^{3+}$, calc. 608.2862); *Yield*: 72% (17.0 mg, 9.9 µmol).

MUC5B(13mer) T-antigen: H_2N-(TEG)-AT*PSSTPGT*THTP-OH (136)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 14.24 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 12.63 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1094.9937 ($[M+2H]^{2+}$, calc. 1094.9934), 730.3307 ($[M+3H]^{3+}$, calc. 730.3314), 557.4867 ($[M+K+3H]^{4+}$, calc. 557.4893); *Yield*: 79% (22.3 mg, 10.2 µmol).

MUC5B(13mer) T-antigen: H_2N-(TEG)-AT*PSST*PGTTHTP-OH (137)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 14.22 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 12.86 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1094.9930 ($[M+2H]^{2+}$, calc. 1094.9934), 911.9259 ($[M-\{Galβ(1,4)-GalNac\}+2H]^{2+}$, calc. 911.9353), 366.1395 ($[Galβ(1,4)-GalNac]^{+}$, calc. 366.1395); *Yield*: 62% (17.7 mg, 8.1 µmol).

MUC5B(13mer) T-antigen: H_2N-(TEG)-AT*PSST*PGT*THTP-OH (138)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 13.79 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 13.82 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1277.5626 ($[M+2H]^{2+}$, calc. 1277.5595), 852.0418 ($[M+3H]^{3+}$, calc. 852.0421), 648.7709 ($[M+K+3H]^{4+}$, calc. 648.7724); *Yield*: 61% (20.2 mg, 8.1 µmol).

MUC5B(17mer) T-antigen:H_2N-(TEG)-T*GSTAT*PSST*PGT*THTP-OH (139)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 13.16 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 11.71 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1089.1364 ($[M+3H]^{3+}$, calc. 1089.1358), 826.5911 ($[M+K+3H]^{4+}$, calc. 826.5926); *Yield*: 33% (18.1 mg, 4.2 µmol).

MUC5B(13mer) type-1 core 3: H_2N-(TEG)-ATPSSTPGT*THTP-OH (140)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 16.41 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 13.37 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1013.9677 ($[M+2H]^{2+}$, calc. 1013.9670), 683.3074 ($[M+3H]^{3+}$, calc. 683.3066); *Yield*: 72% (19.0 mg, 9.4 µmol).

MUC5B(13mer) type-1 core 3: H_2N-(TEG)-AT*PSSTPGTTHTP-OH (141)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 16.34 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 13.35 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1013.9695 ($[M+2H]^{2+}$, calc. 1013.9670), 683.3086 ($[M+3H]^{3+}$, calc. 683.3066); *Yield*: 70% (18.3 mg, 9.0 µmol).

MUC5B(13mer) type-1 core 3: H_2N-(TEG)-AT*PSSTPGT*THTP-OH (142)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 15.03 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 12.32 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1298.0817 ($[M+2H]^{2+}$, calc. 1298.0728), 871.3967 ($[M+NH_4+2H]^{3+}$, calc. 871.3931), 865.7193 ($[M+3H]^{3+}$, calc. 865.7176); *Yield*: 61% (20.6 mg, 7.9 µmol).

MUC5B(13mer) type-1 core 3: H_2N-(TEG)-AT*PSST*PGTTHTP-OH (143)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 15.26 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 12.63 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1298.0780 ($[M+2H]^{2+}$, calc. 1298.0728), 871.3954 ($[M+NH_4+2H]^{3+}$, calc. 871.3931), 865.7190 ($[M+3H]^{3+}$, calc. 865.7176); *Yield*: 54% (18.1 mg, 7.0 µmol).

MUC5B(13mer) type-1 core 3: H_2N-(TEG)-AT*PSST*PGT*THTP-OH (144)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 14.58 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 11.79 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1055.1250 ($[M+2H]^{2+}$, calc. 1055.1215); *Yield*: 62% (25.6 mg, 8.1 µmol).

MUC5B(17mer) type-1 core 3: H_2N-(TEG)-T*GSTAT*PSST*PGT*THTP-OH (145)

The glycopeptide was synthesized on 41.2 mg (7 µmol) of resin.

Analytical HPLC R_t = 37.98 min (A/B: (95:5) → (84:16), 200 µL/min, 54 min); Preparative HPLC R_t = 30.06 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1365.9219 ($[M+NH_4+3H]^{4+}$, calc. 1365.9183), 1360.2442 ($[M+3H]^{3+}$, calc. 1360.2427), 1029.9224 ($[M+K+3H]^{4+}$, calc. 1029.9228); *Yield*: 36% (10.2 mg, 2.5 µmol).

MUC5B(13mer) type-2 core 3: H_2N-(TEG)-ATPSSTPGT*THTP-OH (146)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 14.73 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 13.15 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1013.9674 ($[M+2H]^{2+}$, calc. 1013.9670), 830.8999 ($[M-\{Galβ(1,4)-GlcNAc\}+2H]^{2+}$, calc. 830.8992), 366.1398 ($[Galβ(1,4)-GlcNAc]^{+}$, calc 366.1395); *Yield*: 66% (17.5 mg, 8.6 µmol).

MUC5B(13mer) type-2 core 3: H_2N-(TEG)-AT*PSSTPGTTHTP-OH (147)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 14.96 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 13.17 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1013.9673 ($[M+2H]^{2+}$, calc. 1013.9670), 830.8999 ($[M-\{Galβ(1,4)-GlcNAc\}+2H]^{2+}$, calc. 830.8992), 688.6289 ($[M+K+2H]^{3+}$, calc. 688.6313), 366.1397 ($[Galβ(1,4)-GlcNAc]^{+}$, calc 366.1395); *Yield*: 75% (17.6 mg, 9.7 µmol).

MUC5B(13mer) type-2 core 3: H_2N-(TEG)-AT*PSSTPGT*THTP-OH (148)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 13.71 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 11.99 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1298.0751 ($[M+2H]^{2+}$, calc. 1298.0728), 1115.5072 ($[M-\{βGal(1,4)-GlcNAc-OH\}+2H]^{2+}$, calc. 1115.5056), 865.7179 ($[M+3H]^{3+}$, calc. 865.7176), 366.1396 ($[βGal(1,4)-GlcNAc-OH]^{+}$, calc 366.1395); *Yield*: 62% (21.0 mg, 8.1 µmol).

MUC5B(13mer) type-2 core 3:H_2N-(TEG)-AT*PSST*PGTTHTP-OH (149)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 13.96 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 12.28 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1298.0784 ($[M+2H]^{2+}$, calc. 1298.0728), 878.3700 ($[M+K+2H]^{3+}$, calc. 878.3696), 871.3954 ($[M+NH_4+2H]^{3+}$, calc. 871.3931), 865.7190 ($[M+3H]^{3+}$, calc. 865.7176), 659.0281 ($[M+K+3H]^{4+}$, calc. 659.0279); *Yield*: 67% (22.5 mg, 8.7 µmol).

MUC5B(13mer) type-2 core 3: H_2N-(TEG)-AT*PSST*PGT*THTP-OH (150)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 13.28 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 11.81 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1399.6147 ([M-{Galβ(1,4)-GlcNAc}+2H]$^{2+}$, calc. 1399.6125), 1062.4496 ([M+Na+2H]$^{3+}$, calc. 1062.4488), 1055.1222 ([M+3H]$^{3+}$, calc. 1055.1215); *Yield*: 65% (26.6 mg, 8.4 µmol).

MUC5B(17mer) type-2 core 3: H_2N-(TEG)-T*GSTAT*PSST*PGT*THTP-OH (151)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 12.91 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 10.94 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1360.2452 ([M+3H]$^{3+}$, calc. 1360.2427), 1020.4353 ([M+4H]$^{4+}$, calc. 1029.4339); *Yield*: 49% (26.1 mg, 6.4 µmol).

MUC5B(13mer) type-1 core 1: H_2N-(TEG)-ATPSSTPGT*THTP-OH (152)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 16.01 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 13.21 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1094.9953 ([M+2H]$^{2+}$, calc. 1094.9934), 730.3317 ([M+3H]$^{3+}$, calc. 730.3314); *Yield*: 36% (8.0 mg, 3.6 µmol).

MUC5B(13mer) type-1 core 1: H_2N-(TEG)-AT*PSSTPGTTHTP-OH (153)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 16.23 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 13.66 min (A/B: (95:5) → (70:30), 20 mL/min, 25 min); *HR-ESI-MS*, *m/z*: 1094.9991 ([M+2H]$^{2+}$, calc. 1094.9934), 730.3324 ([M+3H]$^{3+}$, calc. 730.3314); *Yield*: 49% (8.3 mg, 3.8 µmol).

MUC5B(13mer) type-1 core 1: H_2N-(TEG)-AT*PSSTPGT*THTP-OH (154)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 14.11 min (A/B: (95:5) → (70:30), 200 µL/min, 25 min); Preparative HPLC R_t = 29.06 min (A/B: (95:5) → (82:18), 20 mL/min, 49 min); *HR-ESI-MS*, *m/z*: 973.7534 ([M+3H]$^{3+}$, calc. 973.7528); *Yield*: 30% (11.5 mg, 3.9 µmol).

MUC5B(13mer) type-1 core 1: H_2N-(TEG)-AT*PSST*PGTTHTP-OH (155)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 31.77 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 29.54 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 979.4302 ($[M+NH_4+2H]^{3+}$, calc. 979.4284) 973.7534 ($[M+3H]^{3+}$, calc. 973.7528); *Yield*: 40% (15.3 mg, 5.2 µmol).

MUC5B(13mer) type-1 core 1: H_2N-(TEG)-AT*PSST*PGT*THTP-OH (156)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 29.77 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 28.28 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1222.8532 ($[M+NH_4+2H]^{3+}$, calc. 1222.8498), 1217.1760 ($[M+3H]^{3+}$, calc.1217.1728); *Yield*: 19% (9.1 mg, 2.5 µmol).

MUC5B(17mer) type-1 core 1: H_2N-(TEG)-T*GSTAT*PSST*PGT*THTP-OH (157)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 28.56 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 31.41 min (A/B: (95:5) → (82:18), 20 mL/min, 65 min); *HR-ESI-MS*, *m/z*: 1581.9901 ($[M+NH_4+2H]^{3+}$, calc. 1581.9887), 1576.3150 ($[M+3H]^{3+}$, calc. 1576.3132), 1186.7444 ($[M+NH_4+3H]^{4+}$, calc. 1186.7633), 1182.4878 ($[M+4H]^{4+}$, calc. 1182.4867); *Yield*: 9% (5.0 mg, 1.1 µmol).

MUC5B(13mer) type-2 core 1: H_2N-(TEG)-ATPSSTPGT*THTP-OH (158)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 34.41 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 33.94 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1094.9950 ($[M+2H]^{2+}$, calc. 1094.9934), 730.3325 ($[M+3H]^{3+}$, calc. 730.3314); *Yield*: 17% (3.7 mg, 1.7 µmol).

MUC5B(13mer) type-2 core 1: H_2N-(TEG)-AT*PSSTPGTTHTP-OH (159)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 34.21 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 31.85 min (A/B: (95:5) → (82:18), 20 mL/min, 49 min); *HR-ESI-MS*, *m/z*: 1094.9938 ($[M+2H]^{2+}$, calc. 1094.9934), 730.3318 ($[M+3H]^{3+}$, calc. 730.3314); *Yield*: 37% (10.4 mg, 4.8 µmol).

MUC5B(13mer) type-2 core 1: H_2N-(TEG)-AT*PSSTPGT*THTP-OH (160)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 14.67 min (A/B: (95:5) → (65:35), 200 µL/min, 30 min); Preparative HPLC R_t = 28.02 min (A/B: (95:5) → (82:18), 20 mL/min, 49 min); *HR-ESI-MS*, *m/z*: 979.4328 ($[M+NH_4+2H]^{3+}$, calc. 979.4284), 973.7551 ($[M+3H]^{3+}$, calc. 973.7528); *Yield*: 22% (8.4 mg, 2.9 µmol).

MUC5B(13mer) type-2 core 1: H_2N-(TEG)-AT*PSST*PGTTHTP-OH (161)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 32.66 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 25.87 min (A/B: (95:5) → (79:21), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 973.7535 ($[M+3H]^{3+}$, calc. 973.7528); *Yield*: 22% (8.4 mg, 2.9 µmol).

MUC5B(13mer) type-2 core 1: H_2N-(TEG)-AT*PSST*PGT*THTP-OH (162)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 29.55 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 25.67 min (A/B: (95:5) → (79:21), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1244.8405 ($[M+2Na+K]^{3+}$, calc. 1222.8153), ($[M+NH_4+2H]^{3+}$, calc. 1222.8498), 1217.1770 ($[M+3H]^{3+}$, calc. 1217.1728), 913.3859 ($[M+4H]^{4+}$, calc. 913.3834); *Yield*: 15% (7.3 mg, 2.0 µmol).

MUC5B(17) type-2 core 1: H_2N-(TEG)-T*GSTAT*PSST*PGT*THTP-OH (163)

The glycopeptide was synthesized on 76.5 mg (13 µmol) of resin.

Analytical HPLC R_t = 28.36 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 25.67 min (A/B: (95:5) → (79:21), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1576.3150 ($[M+3H]^{3+}$, calc. 1576.3132), 1186.7467 ($[M+NH_4+3H]^{4+}$, calc. 1186.7633), 1182.4892 ($[M+4H]^{4+}$, calc. 1182.4867); *Yield*: 9% (5.0 mg, 1.1 µmol).

MUC5B(13mer) type-1 core 2: H_2N-(TEG)-ATPSSTPGT*THTP-OH (164)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 40.52 min (A/B: (95:5) → (84:16), 200 µL/min, 54 min); Preparative HPLC R_t = 30.09 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1277.5636 ($[M+2H]^{2+}$, calc. 1277.5595), 852.0428 ($[M+3H]^{3+}$, calc. 852.0421); *Yield*: 20% (5.2 mg, 2.0 µmol).

MUC5B(13mer) type-1 core 2: H_2N-(TEG)-AT*PSSTPGTTHTP-OH (165)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 30.78 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 30.23 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*:

1277.5644 ([M+2H]$^{2+}$, calc. 1277.5595), 852.0431 ([M+3H]$^{3+}$, calc. 852.0421); *Yield*: 28% (7.2 mg, 2.8 µmol).

MUC5B(13mer) type-1 core 2: H_2N-(TEG)-AT*PSSTPGT*THTP-OH (166)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 25.40 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 23.29 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1222.8551 ([M+NH_4+2H]$^{3+}$, calc. 1222.8498), 1217.1780 ([M+3H]$^{3+}$, calc. 1217.1743); *Yield*: 16% (5.9 mg, 1.6 µmol).

MUC5B(13mer) type-1 core 2: H_2N-(TEG)-AT*PSST*PGTTHTP-OH (167)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 24.17 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 21.78 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1217.1779 ([M+3H]$^{3+}$, calc. 1217.1743); *Yield*: 21% (7.8 mg, 2.1 µmol).

MUC5B(17mer) type-1 core 2: H_2N-(TEG)-AT*PSST*PGT*THTP-OH (168)

The glycopeptide was synthesized on 47.1 mg (8 µmol) of resin.

Analytical HPLC R_t = 27.35 min (A/B: (95:5) → (84:16), 200 µL/min, 54 min); Preparative HPLC R_t = 24.57 min (A/B: (95:5) → (86:14), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1582.6435 ([M+3H]$^{3+}$, calc. 1582.6410), 1196.7215 ([M+K+3H]$^{4+}$, calc. 1196.7215), 1191.4909 ([M+NH_4+3H]$^{4+}$, calc. 1191.4892), 1187.2341 ([M+4H]$^{4+}$, calc. 1187.2325); *Yield*: 10% (3.7 mg, 0.8 µmol).

MUC5B(13mer) type-2 core 2: H_2N-(TEG)-ATPSSTPGT*THTP-OH (169)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 41.00 min (A/B: (95:5) → (84:16), 200 µL/min, 54 min); Preparative HPLC R_t = 30.49 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1277.5644 ([M+2H]$^{2+}$, calc. 1277.5595), 852.0429 ([M+3H]$^{3+}$, calc. 852.0421); *Yield*: 28% (7.3 mg, 2.8 µmol).

MUC5B(13mer) type-2 core 2: H_2N-(TEG)-AT*PSSTPGTTHTP-OH (170)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 30.90 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 29.55 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1277.5644 ([M+2H]$^{2+}$, calc. 1277.5595), 852.0433 ([M+3H]$^{3+}$, calc. 852.0421); *Yield*: 28% (7.2 mg, 2.8 µmol).

MUC5B(13mer) type-2 core 2: H_2N-(TEG)-AT*PSSTPGT*THTP-OH (171)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 25.40 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 23.27 min (A/B: (95:5) → (80:20), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1222.8551 ($[M+NH_4+2H]^{3+}$, calc. 1222.8498), 1217.1769 ($[M+3H]^{3+}$, calc. 1217.1743), 922.6223 ($[M+4H]^{4+}$, calc. 922.6215); *Yield*: 17% (6.2 mg, 1.7 µmol).

MUC5B(13mer) type-2 core 2: H_2N-(TEG)-AT*PSST*PGTTHTP-OH (172)

The glycopeptide was synthesized on 58.8 mg (10 µmol) of resin.

Analytical HPLC R_t = 24.63 min (A/B: (95:5) → (79:21), 200 µL/min, 54 min); Preparative HPLC R_t = 24.27 min (A/B: (95:5) → (82:18), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1217.1770 ($[M+3H]^{3+}$, calc. 1217.1743); *Yield*: 21% (7.7 mg, 2.1 µmol).

MUC5B(13mer) type-2 core 2: H_2N-(TEG)-AT*PSST*PGT*THTP-OH (173)

The glycopeptide was synthesized on 47.1 mg (8 µmol) of resin.

Analytical HPLC R_t = 28.32 min (A/B: (95:5) → (84:16), 200 µL/min, 54 min); Preparative HPLC R_t 25.57 = min (A/B: (95:5) → (86:14), 20 mL/min, 60 min); *HR-ESI-MS*, *m/z*: 1582.6453 ($[M+3H]^{3+}$, calc. 1582.6410), 1196.7216 ($[M+K+3H]^{3+}$, calc. 1196.7215), 1191.4916 ($[M+NH_4+3H]^{4+}$, calc. 1191.4892), 1187.2346 ($[M+4H]^{4+}$, calc. 1187.2325); *Yield*: 9% (3.4 mg, 0.7 µmol).

7.2.6 Synthesis of α2,3- and α2,6-sialylated MUC1 glycopeptides

7.2.6.1 General procedure for enzymatic sialylations

Utilized Sialyltransferases:

1. For α2,3-sialylation of T-antigen containing peptides:
   Rat 2,3-OST: α2,3-(O)-Sialyltransferase from rat, recombinant (*Spodoptera frugiperda*)[282], (*Calbiochem*) *Merck KGaA*.
2. For α2,3-sialylation of type-1 and -2 LacNAc containing peptides:
   PmST1: α2,3-(O)-Sialyltransferase from *Pasteurella multocida*, recombinant (*Escherichia coli* BL21(DE3))[283], *Sigma Aldrich Inc*.
3. For α2,3-sialylation of type-2 LacNAc containing peptide (only **174**):
   PmST3: α2,3-(O)-Sialyltransferase from *Pasteurella multocida*, recombinant (*Escherichia coli* BL21(DE3)), Prof. Dr X. Chen, University of California.[284]

4. For α2,6-sialylation of LacNAc structures:
   Pd2,6ST: α2,6-(O)-Sialyltransferase from *Photobacterium damsela*, recombinant (*Escherichia coli* BL21(DE3))[285], *Sigma Aldrich Inc*.

The MUC1 glycopeptide (0.5-1.5 mg) was transferred into a 500 µL reaction vial and dissolved in corresponding reaction buffer (Rat 2,3-OST: 100 mM sodium cacodylate, pH 6.0; PmST1 and Pd2,6ST: 100 mM tris(hydroxymethyl)aminomethane hydrochloride (TRIS HCl), pH 8.5). Citidin-5´-monophospho-*N*-actylneuraminc acid (CMP-Neu5Ac) was dissolved in reaction buffer and added to the glycopeptide (2.5 eq per carbohydrate substrate) in the reaction vial. Then the sialyltransferase in reaction buffer was added to the reaction mixture (see *table 6.1* for quantity in mU). The final glycopeptide concentration in the reaction was 4 mM. The reaction vial was shaken overnight at 37°C. The reaction was followed by analytic HPLC and eventually readjusted by further addition of CMP-Neu5Ac and sialyltransferase (see *table 6.1*). The reaction was stopped by addition of cold acetonitrile (20% end concentration). The solvent was evaporated *in vacuo* and the residue loaded on a solid phase extraction (SPE) column for desalting. The SPE column (*Agilent Spec* 3 mL C18AR 15mg) was previously equilibrated with acetonitrile (3 x 500 µL) and 0.05% FA (3 x 500 µL). Crude glycopeptide was taken up with 0.05% FA (3 x 100 µL), loaded onto the SPE cartridge and washed with 0.05% FA (3 x 300 µL). The glycopeptide was eluted with 30% acetonitrile + 0.05% FA (2 x 300 µL) and 50% acetonitrile + 0.05% FA (1 x 300 µL). The combined acetonitrile fractions were concentrated *in vacuo* and the finally lyophilized from water. The lyophilisate was purified by semi-preparative HPLC-MS (see *chapter 6.2.1* for used HPLC systems). The amount of sialylated glycopeptide product was determined by amino acid analysis. *Table 6.1* summarizes relevant reaction parameter.

***Table 7.1*:** Summary of reaction parameters for enzymatic α2,3- and α2,6-sialylation.

MUC1 Acceptor Peptide			**Conditions**					**Product**	
Glycan	#	Sequence	enzyme	eq CMP-Neu5Ac	reaction time (h)	enzyme activity (mU)	buffer	#	Neu5Ac
T	**75**	PAHGVT*SAPDTRPAPGSTA	rat2,3OST	2.0	20	5.6	1	**175**	α2,3
	76	PAHGVTSAPDT*RPAPGSTA	rat2,3OST	2.0+0.5	36	5.0+1.4	1	**176**	α2,3
	77	PAHGVTSAPDTRPAPGST*A	rat2,3OST	2.0	20	5.6	1	**177**	α2,3
	78	PAHGVT*SAPDT*RPAPGSTA	rat2,3OST	4.0+2.0	15	4.2+1.4	1	**178**	α2,3
	79	PAHGVT*SAPDTRPAPGST*A	rat2,3OST	4.0+2.0	15	4.2+1.4	1	**179**	α2,3
	80	PAHGVTSAPDT*RPAPGST*A	rat2,3OST	4.0+2.0	15	4.2+1.4	1	**180**	α2,3
	81	PAHGVT*SAPDT*RPAPGST*A	rat2,3OST	6.0	38	10.5	1	**181**	α2,3
type-2 core 3	**92**	PAHGVT*SAPDTRPAPGSTA	PmST1	2.5+1.0	22	5+2	2	**182**	α2,3
	93	PAHGVTSAPDT*RPAPGSTA	PmST1	2.5+1.0	24	5+2	2	**183**	α2,3
	94	PAHGVTSAPDTRPAPGST*A	PmST1	2.5+1.0	22	5+2	2	**184**	α2,3
	92	PAHGVT*SAPDTRPAPGSTA	Pd2,6ST	2.5+1.0	15	20+12.5	2	**185**	α2,6
	93	PAHGVTSAPDT*RPAPGSTA	Pd2,6ST	2.5+1.0	24	20+12.5	2	**186**	α2,6
	94	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	2.5+1.0	15	20+12.5	2	**187**	α2,6
	98	PAHGVT*SAPDT*RPAPGST*A	Pd2,6ST	7.5+3.0	18	30+15	2	**188**	α2,6

MUC1 Acceptor Peptide			Conditions					Product	
Glycan	#	Sequence	enzyme	eq CMP-Neu5Ac	reaction time (h)	enzyme activity (mU)	buffer	#	Neu5Ac
type-1 core 1	**101**	PAHGVTSAPDT*RPAPGSTA	PmST1	2.0+1.0	12	6	2	**189**	α2,3
	102	PAHGVTSAPDTRPAPGST*A	PmST1	2.0+1.0	12	6	2	**190**	α2,3
	102	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	2.0+1.0	12	25	2	**191**	α2,6
	102	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	2.0+1.0				**192**	2x(α2,6)
type-2 core 1	**110**	PAHGVT*SAPDTRPAPGSTA	PmST1	2.5+1.0	18	5+2	2	**193**	α2,3
	111	PAHGVTSAPDT*RPAPGSTA	PmST1	2.0+1.0	39	6	2	**194**	α2,3
	111	PAHGVTSAPDT*RPAPGSTA	Pd2,6ST	2.5+1.0	18	20	2	**195**	α2,6
	112	PAHGVTSAPDTRPAPGST*A	PmST1	2.0+1.0	18	6	2	**196**	α2,3
	112	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	2.0+1.0	12	25	2	**197**	α2,6
	112	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	2.0+1.0	12	25	2	**198**	2x(α2,6)
	116	PAHGVT*SAPDT*RPAPGST*A	PmST1	7.5+3.0	18	10+4	2	**199**	α2,3
type-1 core 2	**119**	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	4.0+2.0	12	25	2	**200**	α2,6
type-2 core 2 (hexa-saccharide)	**125**	PAHGVTSAPDT*RPAPGSTA	PmST1	4.0+2.0	36	4.2+1.4	2	**201**	α2,3
	125	PAHGVTSAPDT*RPAPGSTA	Pd2,6ST	4.0+2.0	40	30	2	**202**	α2,6
	126	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	4.0+2.0	12	25	2	**203**	α2,6
	126	PAHGVTSAPDTRPAPGST*A	Pd2,6ST	4.0+2.0	12	25	2	**204**	3x(α2,6)
type-2 core 2 (tetra-saccharide)	**174**	PAHGVTSAPDT*RPAPGSTA	rat2,3OST	2.5+1.0	22	4.2+1.4	1	**205**	α2,3
	174	PAHGVTSAPDT*RPAPGSTA	PmST3	2.5+1.0	40	19 µg	2	**206**	α2,3

buffers: 1 = 100 mM cacodylate pH 6.0
2 = 100 mM TRIS HCl pH 8.5

7.2.6.2 Sialylated MUC1 glycopeptides **175**-**206**

Yields are represented by glycopeptide conversion observed by analytical HPLC and after SPE and semi-preparative HPLC quantified by amino acid analysis HPLC eluents were composed of gradients of modifier A: water + 0.1% FA and modifier B: 84% acetonitrile + 0.1% FA. Peptide bonds were detected at λ = 214 nm during HPLC. In order to prevent overloading the HPLC column, only a fraction of the crude glycopeptide was purified and yields from amino acid analysis (AAA) were correlated to the amount of starting material.

MUC1(19mer) α2,3-sialyl T: H_2N-(TEG)-PAHGVT*SAPDTRPGSTA-OH (175)

Glycopeptide **75** (1.34 mg, 568 nmol) was sialylated with rat 2,3-OST.

Analytical HPLC R_t = 26.81 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 27.83 min (A/B: (95:5) → (82:18), 1.5 mL/min, 35 min); *HR-ESI-MS*, *m/z*: 884.0827 ($[M+3H]^{3+}$, calc. 884.0813), 663.3131 ($[M+4H]^{4+}$, calc. 663.3128); *Yield*: 97 % (conversion by HPLC), 45% (253 µmol, AAA).

MUC1(19mer) α2,3-sialyl T: H_2N-(TEG)-PAHGVTSAPDT*RPGSTA-OH (176)

Glycopeptide **76** (1.62 mg, 686 nmol) was sialylated with rat 2,3-OST.

Analytical HPLC R_t = 29.76 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t= 27.83 min (A/B: (95:5) → (78:22), 1.5 mL/min, 45 min); *HR-ESI-MS*, *m/z*: 884.0823 ($[M+3H]^{3+}$, calc. 884.0813), 663.3131 ($[M+4H]^{4+}$, calc. 663.3128); *Yield*: 82 % (conversion by HPLC), 47% (322 nmol, AAA).

MUC1(19mer) α2,3-sialyl T: H_2N-(TEG)-PAHGVTSAPDTRPGST*A-OH (177)

Glycopeptide **77** (1.58 mg, 669 nmol) was sialylated with rat 2,3-OST.

Analytical HPLC R_t = 30.74 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t= 27.27 min (A/B: (95:5) → (78:22), 1.5 mL/min, 45 min); *HR-ESI-MS*, *m/z*: 884.0824 ($[M+3H]^{3+}$, calc. 884.0813), 663.3131 ($[M+4H]^{4+}$, calc. 663.3128); *Yield*: 94 % (conversion by HPLC), 45% (303 nmol, AAA).

MUC1(19mer) α2,3-sialyl T: H_2N-(TEG)-PAHGVT*SAPDT*RPGSTA-OH (178)

Glycopeptide **78** (0.80 mg, 294 nmol) was sialylated with rat 2,3-OST.

Analytical HPLC R_t = 28.45 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t= 26.08 min (A/B: (95:5) → (78:22), 1.5 mL/min, 45 min); *HR-ESI-MS*, *m/z*: 1102.8255 ($[M+3H]^{3+}$, calc. 1102.8239), 827.3724 ($[M+4H]^{4+}$, calc. 827.3697); *Yield*: 89% (conversion by HPLC), 28% (81 nmol, AAA).

MUC1(19mer) α2,3-sialyl T: H_2N-(TEG)-PAHGVT*SAPDTRPGST*A-OH (179)

Glycopeptide **79** (0.99 mg, 364 nmol) was sialylated with rat 2,3-OST.

Analytical HPLC R_t = 29.54 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t= 26.84 min (A/B: (95:5) → (78:22), 1.5 mL/min, 45 min); *HR-ESI-MS*, *m/z*: 1102.8251 ($[M+3H]^{3+}$, calc. 1102.8239), 1102.8251 ($[M+4H]^{4+}$, calc. 827.3697); *Yield*: 100% (conversion by HPLC), 26% (94 nmol, AAA).

MUC1(19mer) α2,3-sialyl T: H_2N-(TEG)-PAHGVTSAPDT*RPGST*A-OH (180)

Glycopeptide **80** (0.78 mg, 286 nmol) was sialylated with rat 2,3-OST.

Analytical HPLC R_t = 32.60 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t= 28.84 min (A/B: (95:5) → (78:22), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1102.8249 ($[M+3H]^{3+}$, calc. 1102.8239), 827.3717 ($[M+4H]^{4+}$, calc. 827.3697); *Yield*: 100 % (conversion by HPLC), 41% (118 nmol, AAA).

MUC1(19mer) α2,3-sialyl T: H_2N-(TEG)-PAHGVT*SAPDT*RPGST*A-OH (181)

Glycopeptide **81** (1.52 mg, 492 µmol) was sialylated with rat 2,3-OST.

Analytical HPLC R_t = 30.85 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t= 25.44 min (A/B: (95:5) → (78:22), 1.5 mL/min, 45 min); *HR-ESI-MS*, *m/z*: 1321.9009 ($[M+3H]^{3+}$, calc. 1321.5664), 991.6781 ($[M+4H]^{4+}$, calc. 991.6764); *Yield*: 70 % (conversion by HPLC), 55% (272 nmol, AAA).

MUC1(19mer) α2,3-sialyl type-2 core 3: H_2N-(TEG)-PAHGVT*SAPDTRPGSTA-OH (182)

Glycopeptide **92** (0.73 mg, 285 nmol) was sialylated with PmST1.

Analytical HPLC R_t = 26.68 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t= 26.39 min (A/B: (95:5) → (78:22), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 951.7759 ($[M+3H]^{3+}$, calc. 951.7745), 714.0831 ($[M+4H]^{4+}$, calc. 714.0827); *Yield*: 100% (conversion by HPLC), 48% (136 nmol, AAA).

MUC1(19mer) α2,3-sialyl type-2 core 3: H_2N-(TEG)-PAHGVTSAPDT*RPGSTA-OH (183)

Glycopeptide **93** (0.70 mg, 273 nmol) was sialylated with PmST1.

Analytical HPLC R_t = 29.794 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t= 28.05 min (A/B: (95:5) → (78:22), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 951.7756 ($[M+3H]^{3+}$, calc. 951.7745), 714.0832 ($[M+4H]^{4+}$, calc. 714.0827); *Yield*: 88% (conversion by HPLC), 43% (118 nmol, AAA).

MUC1(19mer) α2,3-sialyl type-2 core 3: H_2N-(TEG)-PAHGVTSAPDTRPGST*A-OH (184)

Glycopeptide **94** (0.62 mg, 242 nmol) was sialylated with PmST1.

Analytical HPLC R_t = 30.89 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t= 29.55 min (A/B: (95:5) → (82:18), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 951.7758 ($[M+3H]^{3+}$, calc. 951.7745), 714.0833 ($[M+4H]^{4+}$, calc. 714.0827); *Yield*: 94% (conversion by HPLC), 46% (110 nmol, AAA).

MUC1(19mer) α2,6-sialyl type-2 core 3: H_2N-(TEG)-PAHGVT*SAPDTRPGSTA-OH (185)

Glycopeptide **92** (0.66 mg, 258 nmol) was sialylated with Pd2,6ST.

Analytical HPLC R_t = 26.51 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t= 25.62 min (A/B: (95:5) → (82:18), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 951.7759 ($[M+3H]^{3+}$, calc. 951.7745), 714.0834 ($[M+4H]^{4+}$, calc. 714.0827); *Yield*: 97% (conversion by HPLC), 57% (147 nmol, AAA).

MUC1(19mer) α2,6-sialyl type-2 core 3: H_2N-(TEG)-PAHGVTSAPDT*RPGSTA-OH (186)

Glycopeptide **93** (0.79 mg, 273 nmol) was sialylated with Pd2,6ST.

Analytical HPLC R_t = 29.14 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t= 28.17 min (A/B: (95:5) → (82:18), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 951.7754

([M+3H]$^{3+}$, calc. 951.7745), 714.0831 ([M+4H]$^{4+}$, calc. 714.0827); *Yield*: 88% (conversion by HPLC), 57% (176 nmol, AAA).

MUC1(19mer) α2,6-sialyl type-2 core 3: H_2N-(TEG)-PAHGVTSAPDTRPGST*A-OH (187)

Glycopeptide **94** (0.64 mg, 250 nmol) was sialylated with Pd2,6ST.

Analytical HPLC R_t = 30.56 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 29.09 min (A/B: (95:5) → (82:18), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 951.7755 ([M+3H]$^{3+}$, calc. 951.7745), 714.0833 ([M+4H]$^{4+}$, calc. 714.0827); *Yield*: 100% (conversion by HPLC), 67% (166 nmol, AAA).

MUC1(19mer) α2,3-sialyl type-2 core 3: H_2N-(TEG)-PAHGVT*SAPDT*RPGST*A-OH (188)

Glycopeptide **98** (0.80 mg, 216 nmol) was sialylated with Pd2,6ST.

Analytical HPLC R_t = 29.65 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 24.97 min (A/B: (95:5) → (82:18), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1143.9884 ([M+4H]$^{4+}$, calc.1143.9870), 915.3914 ([M+5H]$^{5+}$, calc. 915.3911); *Yield*: 100% (conversion by HPLC), 10% (23 nmol, AAA).

MUC1(19mer) α2,3-sialyl type-1 core 1: H_2N-(TEG)-PAHGVTSAPDT*RPGSTA-OH (189)

Glycopeptide **101** (0.83 mg, 304 µmol) was sialylated with PmST1.

Analytical HPLC R_t = 28.97 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 25.98 min (A/B: (95:5) → (82:18), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1005.7932 ([M+3H]$^{3+}$, calc. 1005.7921), 754.5964 ([M+4H]$^{4+}$, calc. 754.5959); *Yield*: 88% (conversion by HPLC), 38% (133 nmol, AAA).

MUC1(19mer) α2,3-sialyl type-1 core 1: H_2N-(TEG)-PAHGVTSAPDTRPGST*A-OH (190)

Glycopeptide **102** (0.69 mg, 253 nmol) was sialylated with PmST1.

Analytical HPLC R_t = 30.13 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 27.17 min (A/B: (95:5) → (78:22), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1005.7968 ([M+3H]$^{3+}$, calc. 1005.7921), 754.5996 ([M+4H]$^{4+}$, calc. 754.5959); *Yield*: 85% (conversion by HPLC), 44% (111 nmol, AAA).

MUC1(19mer) α2,6-sialyl type-1 core 1: H_2N-(TEG)-PAHGVTSAPDTRPGST*A-OH (191) + (192)

Glycopeptide **102** (0.67 mg, 246 nmol) was sialylated with Pd2,6ST.

191 (monosialylated main product): Analytical HPLC R_t = 29.34 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 28.99 min (A/B: (95:5) → (82:18), 1.5 mL/min,

min); *HR-ESI-MS*, *m/z*: 1005.7938 ([M+3H]$^{3+}$, calc. 1005.7921), 754.5972 ([M+4H]$^{4+}$, calc. 754.5959); *Yield*: 75% (conversion by HPLC), 36% (89 nmol, AAA).

192 (disialylated by-product): Analytical HPLC R_t = 33.88 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 31.11 min (A/B: (95:5) → (82:18), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1102.8253 ([M+3H]$^{3+}$, calc. 1102.8239), 827.3719 ([M+4H]$^{4+}$, calc. 827.3697); *Yield*: 17% (conversion by HPLC), 15% (36 nmol, AAA).

MUC1(19mer) α2,3-sialyl type-2 core 1: H_2N-(TEG)-PAHGVT*SAPDTRPGSTA-OH (193)

Glycopeptide **110** (0.75 mg, 304 nmol) was sialylated with PmST1.

Analytical HPLC R_t = 25.54 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 25.73 min (A/B: (95:5) → (82:18), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1005.7947 ([M+3H]$^{3+}$, calc. 1005.7921), 754.5963 ([M+4H]$^{4+}$, calc. 754.5959); *Yield*: 97% (conversion by HPLC), 42% (115 nmol, AAA).

MUC1(19mer) α2,3-sialyl type-2 core 1: H_2N-(TEG)-PAHGVTSAPDT*RPGSTA-OH (194)

Glycopeptide **111** (0.95 mg, 349 nmol) was sialylated with PmST1.

Analytical HPLC R_t = 29.23 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 26.99 min (A/B: (95:5) → (78:22), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1005.7937 ([M+3H]$^{3+}$, calc. 1005.7921), 754.5963 ([M+4H]$^{4+}$, calc.754.5959); *Yield*: 100% (conversion by HPLC), 58% (202 nmol, AAA).

MUC1(19mer) α2,6-sialyl type-2 core 1: H_2N-(TEG)-PAHGVTSAPDT*RPGSTA-OH (195)

Glycopeptide **111** (0.80 mg, 294 nmol) was sialylated with Pd2,6ST.

Analytical HPLC R_t = 28.82 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 27.83 min (A/B: (95:5) → (82:18), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1005.7931 ([M+3H]$^{3+}$, calc. 1005.7921), 754.5963 ([M+4H]$^{4+}$, calc. 754.5959); *Yield*: 97% (conversion by HPLC), 38% (110 nmol, AAA).

MUC1(19mer) α2,3-sialyl type-2 core 1: H_2N-(TEG)-PAHGVTSAPDTRPGST*A-OH (196)

Glycopeptide **112** (0.69 mg, 253 nmol) was sialylated with PmST1.

Analytical HPLC R_t = 30.26 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 28.37 min (A/B: (95:5) → (78:22), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1005.7936 ([M+3H]$^{3+}$, calc. 1005.7921), 754.5965 ([M+4H]$^{4+}$, calc. 754.5959); *Yield*: 97% (conversion by HPLC), 38% (95 nmol, AAA).

MUC1(19mer) α2,6-sialyl type-2 core 1: H_2N-(TEG)-PAHGVTSAPDTRPGST*A-OH (197) + (198)

Glycopeptide **112** (0.63 mg, 231 nmol) was sialylated with Pd2,6ST.

197 (monosialylated main product): Analytical HPLC R_t = 29.75 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 27.06 min (A/B: (95:5) → (78:22), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1005.7934 ($[M+3H]^{3+}$, calc. 1005.7921), 754.5966 ($[M+4H]^{4+}$, calc. 754.5959); *Yield*: 55% (conversion by HPLC), 20% (47 nmol, AAA).

198 (disialylated by-product): Analytical HPLC R_t = 34.00 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 28.95 min (A/B: (95:5) → (78:22), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1102.8250 ($[M+3H]^{3+}$, calc. 1102.8239), 827.3714 ($[M+4H]^{4+}$, calc. 827.3697); *Yield*: 40% (conversion by HPLC), 18% (38 nmol, AAA).

MUC1(19mer) α2,3-sialyl type-2 core 1: H_2N-(TEG)-PAHGVT*SAPDT*RPGST*A-OH (199)

Glycopeptide **116** (0.63 mg, 231 nmol) was sialylated with PmST1.

Analytical HPLC R_t = 28.42 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 24.00 min (A/B: (95:5) → (82:18), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1265.5250 ($[M+4H]^{4+}$, calc. 1265.5266), 1012.6191 ($[M+5H]^{5+}$, calc. 1012.6228); *Yield*: 81% (conversion by HPLC), 42% (63 nmol, AAA).

MUC1(19mer) α2,6-sialyl type-1 core 2: (H_2N-(TEG)-PAHGVTSAPDTRPGST*A-OH (200)

Glycopeptide **119** (0.64 mg, 151 nmol) was sialylated with Pd2,6ST.

Analytical HPLC R_t = 31.39 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 26.26 min (A/B: (95:5) → (78:22), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1224.5352 ($[M+3H]^{3+}$, calc. 1224.5346), 918.6557 ($[M+4H]^{4+}$, calc. 918.6528); *Yield*: 91% (conversion by HPLC), 68% (141 nmol, AAA).

MUC1(19mer) α2,3-sialyl type-2 core 2: H_2N-(TEG)-PAHGVTSAPDT*RPGSTA-OH (201)

Glycopeptide **125** (0.69 mg, 223 nmol) was sialylated with PmST1.

Analytical HPLC R_t = 30.68 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 26.19 min (A/B: (95:5) → (78:22), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1224.5350 ($[M+3H]^{3+}$, calc. 1224.5346), 918.6555 ($[M+4H]^{4+}$, calc. 918.6528); *Yield*: 66% (conversion by HPLC), 52% (116 nmol, AAA).

MUC1(19mer) α2,6-sialyl type-2 core 2: H_2N-(TEG)-PAHGVTSAPDT*RPGSTA-OH (202)

Glycopeptide **125** (0.78 mg, 396 µmol) was sialylated with Pd2,6ST.

Analytical HPLC R_t = 29.63 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 26.29 min (A/B: (95:5) → (78:22), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1224.5352 ($[M+3H]^{3+}$, calc. 1224.5346), 918.6553 ($[M+4H]^{4+}$, calc. 918.6528); *Yield*: 81% (conversion by HPLC), 15% (61 nmol, AAA).

MUC1(19mer) α2,6-sialyl type-2 core 2: H_2N-(TEG)-PAHGVTSAPDTRPGST*A-OH (203) + (204)

Glycopeptide **126** (0.61 mg, 198 µmol) was sialylated with Pd2,6ST.

203 (monosialylated main product): Analytical HPLC R_t = 32.08 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 26.79 min (A/B: (95:5) → (78:22), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1224.5350 ($[M+3H]^{3+}$, calc. 1224.5346), 918.6553 ($[M+4H]^{4+}$, calc. 918.6528); *Yield*: 84% (conversion by HPLC), 76% (150 nmol, AAA).

204 (disialylated by-product): Analytical HPLC R_t = 36.31 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 28.05 min (A/B: (95:5) → (78:22), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1321. 9008 ($[M+3H]^{3+}$, calc. 1321.9009), 991.6779 ($[M+4H]^{4+}$, calc. 991.6775); *Yield*: 16% (conversion by HPLC), 10% (21 nmol, after semi-prep HPLC).

MUC1(19mer) α2,3-sialyl type-2 core 2: H_2N-(TEG)-PAHGVTSAPDTRPGSTA-OH (205)

Glycopeptide **174** (0.64 mg, 235 nmol) was sialylated with rat 2,3-OST.

Analytical HPLC R_t = 26.82 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 26.46 min ((95:5) → (78:22), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1005.7929 ($[M+3H]^{3+}$, calc. 1005.7921), 754.5967 ($[M+4H]^{4+}$, calc. 754.5959); *Yield*: 100% (conversion by HPLC), 42% (99 nmol, AAA).

MUC1(19mer) α2,6-sialyl type-2 core 2: H_2N-(TEG)-PAHGVTSAPDTRPGSTA-OH (206)

Glycopeptide **174** (0.73 mg, 268 nmol) was sialylated with PmST3.

Analytical HPLC R_t = 27.13 min (A/B: (95:5) → (85:15), 200 µL/min, 40 min); Preparative HPLC R_t = 26.58 min (A/B: (95:5) → (78:22), 1.5 mL/min, min); *HR-ESI-MS*, *m/z*: 1005.7935 ($[M+3H]^{3+}$, calc. 1005.7921), 754.5966 ($[M+4H]^{4+}$, calc. 754.5959); *Yield*: 69% (conversion by HPLC), 26% (69 nmol, AAA).

7.3 Microarray experiments with immobilized MUC1 glycopeptides

7.3.1 General

Microarray Spotting: Microarray chips were spotted with an *iTwo 400* spotter from *M2-Automation* equipped with a humidity control unit. The *Nexterion H*® microarray slides were purchased from *Schott GmbH*, Mainz. Glycopeptide substrates for spotting were pipetted into *Nunc*® 384-well plates from *Thermo Scientific* and loaded into the microarray spotter.

Microarray Scanning: Scanning of the microarray slides was performed with a *Typhoon Trio+* by *Amersham Biosciences (GE Healthcare)*. The photo-multiplier tube voltage was set to 480 V or 500 V and it was scanned at 10 µm resolution. Excitation of *Alexa Fluor 488* conjugates (Ex_{max}: 495 nm, Em_{max}: 519 nm) was performed with a blue laser (488 nm) and a 520 nm band-pass emission filter. Excitation of *Cy5* conjugates (Ex_{max}: 650 nm, Em_{max}: 670 nm) was performed with a red laser (633 nm) and a 670 nm band-pass emission filter.

Microarray data processing: Data was obtained with *Amersham Typhoon Array* software. Background was substracted by *spot edge average* background substraction. Data was processed with *Microsoft Excel*. Given values in the diagrams represent the mean of the spot replicates with standard deviation.

Antibodies and lectins:

Secondary goat-anti-mouse antibody used for detection of primary serum antibodies and primary anti-MUC1 Ma552 antibody:

- *Alexa Fluor*® 488 goat-anti-mouse IgG (H+L, 2 mg/mL) was purchased from *LifeTechnologies (Thermo Fisher Scientific)*, Eugene, OR, USA. Experimental dilution: 1:3000 (0.67 µg/ml).
- Mouse monoclonal IgG_1 antibody NCL-MUC1 core glycoprotein Ma552, human breast cancer cell line ZR75-1, 25 µg/ml, *Leica Biosystems*, Newcastle Ldt, UK. Experimental dilution: 12.5 µg/ml.

Biotinylated plant lectins were purchased from *Vector Laboratories Inc*:

- Biotinylated *Erythrina cristagalli* (ECA), 5mg/ml.
- Biotinylated Wheat germ agglutinin (WGA, *Triticum aestivum*), 5mg/ml.

- Biotinylated *Maackia amurensis* I (MAL I), 2 mg/ml.
- Biotinylated *Maackia amurensis* II (MAL II), 1 mg/ml.
- Biotinylated Peanut agglutinin (PNA, *Arachis hypogaea*), 5mg/ml.
- Biotinylated *Vicia villosa* (VVL), 2 mg/ml.
- Biotinylated *Sambucus nigra agglutinin* (SNA), 2 mg/ml.

Streptavidin-Cy5-conjugate for detection of the biotinylated lectin:

- Cy5-Streptavidin conjugate (ZyMax™ Grade), 1mg/ml, *Invitrogen (Thermo Scientific)*. Experimental dilution 1:1500 (0.67 µg/ml).

Galectin-3 and anti-galectin-3 antibody:

- Human, His-tagged, recombinant (*E.coli*) galectin-3 (1-250aa), *ATGen* Ltd, South Korea.
- *Alexa Fluor*® 488 anti-mouse/human Mac-2 (Galectin-3), 0.5 mg/mL, *Biolegends Inc*, San Diego, USA. Experimental dilution: 1:500 (1 µg/ml).

7.3.2 *Nexterion slide H* microarray loading capacity

The standard protocol for NHS-coated *Nexterion H*® microarray slides from *Schott* recommends surface loading with concentrations of 0.1-1 mg/mL for proteins (antibodies). The smaller MUC1 glycopeptides contain one reactive amine group on the N-terminus for covalent surface bonding. MUC1(19mer) **69** [$(TEG)_3$-PAHGVTSAPDT(T_N)RPAPGSTA] was spotted on two *Nexterion* slides, incubated with vaccine candidate 5 mouse 1 (SH127, 1:2000) and detected by addition with either goat-anti-mouse-IgG-Cy5 or goat-anti-mouse-IgG-AlexaFluor488 in concentrations ranging from 5 to 250 µM was performed to evaluate best spotting conditions (*figure 6.1*).

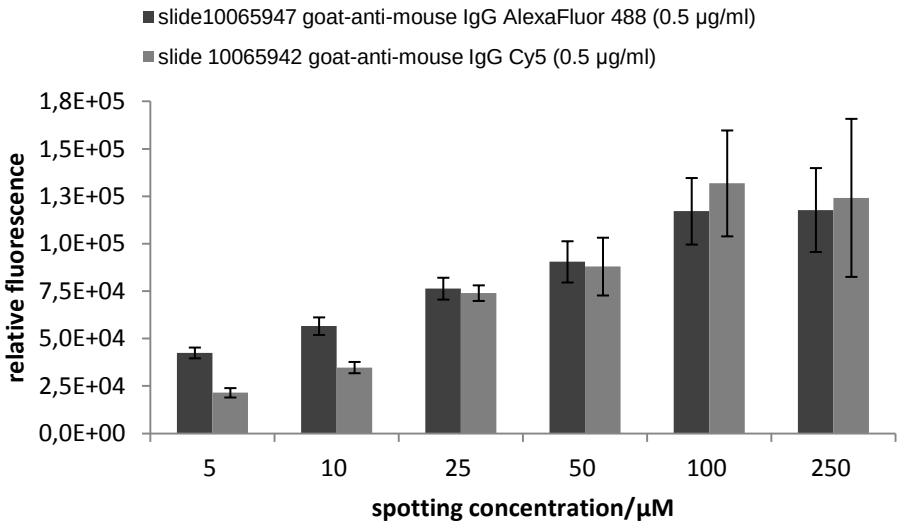


***Figure 7.2*:** Loading capacity of *Nexterion slide H*.

Each bar represents the mean of 672 spots per well. Slide saturation is reached between 50 and 100 µM, according a previous publication on a glycoarray using *Nexterion H*® slides.[316] Higher spotting concentrations only increased standard deviation. Best spotting conditions in terms of best signal to standard deviation ratios are found between 25 and 50 µM. For highest signals intensity all further microarray spotting was carried out in concentrations of 50 µM which equals 0.11-0.26 mg/mL, depending on the molecular weight of each glycopeptide.

7.3.3 General spotting conditions

Reagents:

- Spotting buffer: 150 mM NaH_2PO_4/Na_2HPO_4 buffer (pH 8.5)

Glycopeptides were dissolved and diluted to a spotting concentration of 50 mM in spotting buffer. A volume of 20 µL of each glycopeptide solution was pipetted into a *Nunc*® 384-well plate. The 384-well plate and the *Nexterion H*® slides were loaded in the spotter. Droplets were generated by piezo-driven droplet generation. The spotter settings were adjusted to generate substrate spots from single droplets of 100 pL ± 3 pL onto the microarray slides. During spotting, humidity in the spotting chamber was kept between 50-60%. After the spotting process, microarray slides were kept at 90-99% relative humidity overnight, to complete surface immobilization. Not immediately used slides are stored below -20°C, until further use.

7.3.4 Spotting of MUC1 microarrays MA1, MA2, MA3, MA4, MA5

Five different microarray formats (**MA1**-**MA5**) were spotted for the different applications and with growing size of the MUC1 glycopeptide library:

- **MA1**: Serum antibody titer microarray (*chapter 4.3.7*)
- **MA2**: Serum antibody microarray 1 (*chapter 4.3.8*)
- **MA3**: Serum antibody microarray 2 (*chapter 4.3.8*)
- **MA4**: Serum antibody microarray 3 and galectin microarray (*chapter 4.3.9*; *chapter 4.3.11.3*)
- **MA5**: Plant lectin microarray (*chapter 4.3.11.1*)

The glycopeptide library synthesized in the course of this work was appended by glycopeptides previously synthesized by M.Sc. M.Schorlemer, Leibniz-Institut für Analytische Wissenschaften -ISAS- e.V., (entry **174**), Dr. U. Westerlind and Dr. H. Cai Leibniz-Institut für Analytische Wissenschaften-ISAS-e.V., Dortmund and Johannes-Gutenberg-Universität Mainz, laboratory of Prof. Dr. H. Kunz (entries **207**-**241**). *Table 6.2* lists all glycopeptides utilized for microarray spotting and the according microarray format.

Table 7.2: Full list of MUC1 glycopeptides utilized for microarray spotting of the microarray formats **MA1**, **MA2**, **MA3**, **MA4**, **MA5** with positions of the glycopeptide entries on the arrays. (-) indicates absence of the according glycopeptide.

Peptide number	MUC1 sequence	Glycan	Origin	Peptide position on Microarray				
				MA1	MA2	MA3	MA4	MA5
68	PAHGVT*SAPDTRPAPGSTA	T_N	This thesis	-	27	-	-	21
69	PAHGVTSAPDT*RPAPGSTA	T_N	This thesis	-	28	14	-	22
70	PAHGVTSAPDTRPAPGST*A	T_N	This thesis	-	29	-	-	23
71	PAHGVT*SAPDT*RPAPGSTA	T_N	This thesis	-	30	-	-	24
72	PAHGVT*SAPDTRPAPGST*A	T_N	This thesis	-	31	-	-	25
73	PAHGVTSAPDT*RPAPGST*A	T_N	This thesis	-	32	12	-	26
74	PAHGVT*SAPDT*RPAPGST*A	T_N	This thesis	-	33	-	-	27
75	PAHGVT*SAPDTRPAPGSTA	T	This thesis	-	34	-	1	37
76	PAHGVTSAPDT*RPAPGSTA	T	This thesis	-	35	15	3	38
77	PAHGVTSAPDTRPAPGST*A	T	This thesis	-	36	-	5	39
78	PAHGVT*SAPDT*RPAPGSTA	T	This thesis	-	37	-	7	40
79	PAHGVT*SAPDTRPAPGST*A	T	This thesis	-	38	-	9	41
80	PAHGVTSAPDT*RPAPGST*A	T	This thesis	-	39	-	11	42
81	PAHGVT*SAPDT*RPAPGST*A	T	This thesis	-	40	-	13	43
82	PGSTAPPAHGVTSAPDT*RPA	T	This thesis	-	-	-	-	34
83	APDT*RPAPGSTAPPAHGVTSA	T	This thesis	-	-	-	-	35

Peptide number	MUC1 sequence	Glycan	Origin	Peptide position on Microarray				
				MA1	MA2	MA3	MA4	MA5
84	APDT*RPA	T	This thesis	-	-	-	-	36
85	PAHGVT*SAPDTRPAPGSTA	type-1 core 3	This thesis	-	41	-	-	52
86	PAHGVTSAPDT*RPAPGSTA	type-1 core 3	This thesis	-	42	-	-	53
87	PAHGVTSAPDTRPAPGST*A	type-1 core 3	This thesis	-	43	-	-	54
88	PAHGVT*SAPDT*RPAPGSTA	type-1 core 3	This thesis	-	44	-	-	55
89	PAHGVT*SAPDTRPAPGST*A	type-1 core 3	This thesis	-	45	-	-	56
90	PAHGVTSAPDT*RPAPGST*A	type-1 core 3	This thesis	-	46	-	-	57
91	PAHGVT*SAPDT*RPAPGST*A	type-1 core 3	This thesis	-	47	-	-	58
92	PAHGVT*SAPDTRPAPGSTA	type-2 core 3	This thesis	-	48	-	15	59
93	PAHGVTSAPDT*RPAPGSTA	type-2 core 3	This thesis	-	49	-	18	60
94	PAHGVTSAPDTRPAPGST*A	type-2 core 3	This thesis	-	50	-	29	61
95	PAHGVT*SAPDT*RPAPGSTA	type-2 core 3	This thesis	-	51	-	-	62
96	PAHGVT*SAPDTRPAPGST*A	type-2 core 3	This thesis	-	52	-	-	63
97	PAHGVTSAPDT*RPAPGST*A	type-2 core 3	This thesis	-	53	-	-	64
98	PAHGVT*SAPDT*RPAPGST*A	type-2 core 3	This thesis	-	54	-	32	65
99	PAHGVT*SA	type-2 core 3	This thesis	-	-	-	-	73
100	PAHGVT*SAPDTRPAPGSTA	type-1 core 1	This thesis	-	55	-	-	74
101	PAHGVTSAPDT*RPAPGSTA	type-1 core 1	This thesis	-	56	-	35	75
102	PAHGVTSAPDTRPAPGST*A	type-1 core 1	This thesis	-	57	-	37	76
103	PAHGVT*SAPDT*RPAPGSTA	type-1 core 1	This thesis	-	58	-	-	77
104	PAHGVT*SAPDTRPAPGST*A	type-1 core 1	This thesis	-	59	-	-	78
105	PAHGVTSAPDT*RPAPGST*A	type-1 core 1	This thesis	-	60	-	-	79
106	PAHGVT*SAPDT*RPAPGST*A	type-1 core 1	This thesis	-	61	-	-	80
107	PAHGVT*SAPDTRPAPGSTA	type-2 core 1	This thesis	-	62	-	41	85
108	PAHGVTSAPDT*RPAPGSTA	type-2 core 1	This thesis	-	63	16	43	86
109	PAHGVTSAPDTRPAPGST*A	type-2 core 1	This thesis	-	64	-	46	87
110	PAHGVT*SAPDT*RPAPGSTA	type-2 core 1	This thesis	-	65	-	-	88
111	PAHGVT*SAPDTRPAPGST*A	type-2 core 1	This thesis	-	66	-	-	89
112	PAHGVTSAPDT*RPAPGST*A	type-2 core 1	This thesis	-	67	-	-	90
113	PAHGVT*SAPDT*RPAPGST*A	type-2 core 1	This thesis	-	68	-	50	91
114	GST*APPAHGVTSAPDTRPA	type-2 core 1	This thesis	-	-	-	-	94
115	PDTRPAPGST*APPAHGVTSA	type-2 core 1	This thesis	-	-	-	-	93
116	PDTRPAPGSTAPPAHGVT*SA	type-2 core 1	This thesis	-	-	-	-	92
117	PAHGVT*SAPDTRPAPGSTA	type-1 core 2 (hexa)	This thesis	-	69	-	-	107
118	PAHGVTSAPDT*RPAPGSTA	type-1 core 2 (hexa)	This thesis	-	70	-	-	108
119	PAHGVTSAPDTRPAPGST*A	type-1 core 2 (hexa)	This thesis	-	71	-	53	109
120	PAHGVT*SAPDT*RPAPGSTA	type-1 core 2 (hexa)	This thesis	-	72	-	-	110
121	PAHGVT*SAPDTRPAPGST*A	type-1 core 2 (hexa)	This thesis	-	73	-	-	111
122	PAHGVTSAPDT*RPAPGST*A	type-1 core 2 (hexa)	This thesis	-	74	-	-	112
123	PAHGVT*SAPDT*RPAPGST*A	type-1 core 2 (hexa)	This thesis	-	75	-	-	113
124	PAHGVT*SAPDTRPAPGSTA	type-2 core 2 (hexa)	This thesis	-	76	-	-	115
125	PAHGVTSAPDT*RPAPGSTA	type-2 core 2 (hexa)	This thesis	-	77	17	55	116
126	PAHGVTSAPDTRPAPGST*A	type-2 core 2 (hexa)	This thesis	-	78	-	58	117
127	PAHGVT*SAPDT*RPAPGSTA	type-2 core 2 (hexa)	This thesis	-	79	-	-	118
128	PAHGVT*SAPDTRPAPGST*A	type-2 core 2 (hexa)	This thesis	-	80	-	-	119

Peptide number	MUC1 sequence	Glycan	Origin	Peptide position on Microarray				
				MA1	MA2	MA3	MA4	MA5
129	PAHGVTSAPDT*RPAPGST*A	type-2 core 2 (hexa)	This thesis	-	81	-	-	120
130	PAHGVT*SAPDT*RPAPGST*A	type-2 core 2 (hexa)	This thesis	-	82	-	-	121
131	PAHGVTSAPDT*RPAPGS$^{\#}$T$^{\#}$APPA	*type-2 core 1 + $^{\#}$2xTn	This thesis	-	-	1	-	130
132	PAHGVTSAPDT*RPAPGS$^{\#}$T$^{\#}$APPA	*type-2 core 3 + $^{\#}$2xTn	This thesis	-	-	2	-	131
133	PAHGVTSAPDT*RPAPGS$^{\#}$T$^{\#}$APPA	*type-2 core 2 + $^{\#}$2xTn	This thesis	-	-	3	-	132
174	PAHGVTSAPDT*RPAPGSTA	type-2 core 2 (tetra)	M. Sc. M. Schorlemer	-	-	18	61	127
175	PAHGVT*SAPDTRPAPGSTA	ST	This thesis	-	-	-	2	44
176	PAHGVTSAPDT*RPAPGSTA	ST	This thesis	-	-	-	4	45
177	PAHGVTSAPDTRPAPGST*A	ST	This thesis	-	-	-	6	46
178	PAHGVT*SAPDT*RPAPGSTA	ST	This thesis	-	-	-	8	47
179	PAHGVT*SAPDTRPAPGST*A	ST	This thesis	-	-	-	10	48
180	PAHGVTSAPDT*RPAPGST*A	ST	This thesis	-	-	-	12	49
181	PAHGVT*SAPDT*RPAPGST*A	ST	This thesis	-	-	-	14	50
182	PAHGVT*SAPDTRPAPGSTA	2,3-sialyl type-2 core 3	This thesis	-	-	-	17	66
183	PAHGVTSAPDT*RPAPGSTA	2,3-sialyl type-2 core 3	This thesis	-	-	-	24	67
184	PAHGVTSAPDTRPAPGST*A	2,3-sialyl type-2 core 3	This thesis	-	-	-	31	68
185	PAHGVT*SAPDTRPAPGSTA	2,6-sialyl type-2 core 3	This thesis	-	-	-	16	70
186	PAHGVTSAPDT*RPAPGSTA	2,6-sialyl type-2 core 3	This thesis	-	-	-	19	71
187	PAHGVTSAPDTRPAPGST*A	2,6-sialyl type-2 core 3	This thesis	-	-	-	30	72
188	PAHGVT*SAPDT*RPAPGST*A	2,6-sialyl type-2 core 3	This thesis	-	-	-	34	69
189	PAHGVTSAPDT*RPAPGSTA	2,3-sialyl type-1 core 1	This thesis	-	-	-	36	83
190	PAHGVTSAPDTRPAPGST*A	2,3-sialyl type-1 core 1	This thesis	-	-	-	40	84
191	PAHGVTSAPDTRPAPGST*A	2,6-sialyl type-1 core 1	This thesis	-	-	-	38	81
192	PAHGVTSAPDTRPAPGST*A	2,6-sialyl type-1 core 1	This thesis	-	-	-	39	82
193	PAHGVT*SAPDTRPAPGSTA	2,3-sialyl type-2 core 1	This thesis	-	-	-	42	98
194	PAHGVTSAPDT*RPAPGSTA	2,3-sialyl type-2 core 1	This thesis	-	-	-	45	99
195	PAHGVTSAPDT*RPAPGSTA	2,6-sialyl type-2 core 1	This thesis	-	-	-	44	95
196	PAHGVTSAPDTRPAPGST*A	2,3-sialyl type-2 core 1	This thesis	-	-	-	49	100
197	PAHGVTSAPDTRPAPGST*A	2,6-sialyl type-2 core 1	This thesis	-	-	-	47	96
198	PAHGVTSAPDTRPAPGST*A	2,6-sialyl type-2 core 1	This thesis	-	-	-	48	97
199	PAHGVT*SAPDT*RPAPGST*A	2,3-sialyl type-2 core 1	This thesis	-	-	-	51	102
200	PAHGVTSAPDTRPAPGST*A	2,6-sialyl type-1 core 2 (hexa)	This thesis	-	-	-	54	114
201	PAHGVTSAPDT*RPAPGSTA	2,3-sialyl type-2 core 2 (hexa)	This thesis	-	-	-	57	126
202	PAHGVTSAPDT*RPAPGSTA	2,6-sialyl type-2 core 2 (hexa)	This thesis	-	-	-	56	122
203	PAHGVTSAPDTRPAPGST*A	2,6-sialyl type-2 core 2 (hexa)	This thesis	-	-	-	59	123
204	PAHGVTSAPDTRPAPGST*A	2,6-sialyl type-2 core 2 (hexa)	This thesis	-	-	-	60	124
205	PAHGVTSAPDT*RPAPGSTA	2,3-sialyl (core1) type-2 core 2 (tetra)	This thesis	-	-	-	62	128
206	PAHGVTSAPDT*RPAPGSTA	2,3-sialyl(LacNAc) type-2 core 2 (tetra)	This thesis	-	-	-	63	129

Peptide number	MUC1 sequence	Glycan	Origin	Peptide position on Microarray				
				MA1	MA2	MA3	MA4	MA5
207	PAHGVTSAPDTRPAPGSTAP	-	Dr. U. Westerlind	-	1	-	-	-
208	PAHGVT*SAPDTRPAPGSTAP	ST_N	Dr. U. Westerlind	-	2	-	-	2
209	PAHGVTSAPDT*RPAPGSTAP	ST_N	Dr. U. Westerlind	-	3	-	-	3
210	PAHGVTSAPDTRPAPGST*AP	ST_N	Dr. U. Westerlind	-	4	-	-	4
211	PAHGVT*SAPDTRPAPGST*AP	ST_N	Dr. U. Westerlind	-	5	-	-	5
212	PAHGVT*SAPDTRPAP	ST_N	Dr. U. Westerlind	-	6	-	-	6
213	GSTAPPAHGVT*SAP	* ST_N, ° T_N	Dr. U. Westerlind	-	7	-	-	7
214	PAHGVT*SAPDTRPAPGST°AP	* ST_N, ° T_N	Dr. U. Westerlind	-	8	-	-	8
215	PAHGVT*SAPDT°RPAPGST*AP	* ST_N, ° T_N	Dr. U. Westerlind	-	9	-	-	9
216	PAHGVT*SAPDTRPAPGST°APPA HGVT*SAPDTRPAPGST°AP	* ST_N, ° T_N	Dr. U. Westerlind	-	10	-	-	10
217	PAHGVT*SAPDTRPAPGSTAP	T_N	Dr. U. Westerlind	-	11	-	-	-
218	PAHGVTS*APDTRPAPGSTAP	T_N	Dr. U. Westerlind	-	12	-	-	11
219	PAHGVTSAPDTRPAPGS*TAP	T_N	Dr. U. Westerlind	-	13	-	-	12
220	PAHGVTSAPDT*RPAPGS*TAP	T_N	Dr. U. Westerlind	-	14	-	-	-
221	APDTRPAPGST*AP	T_N	Dr. U. Westerlind	-	15	-	-	13
222	PAHGVT*SAPDTRPAPGSTAP	T	Dr. U. Westerlind	-	16	-	-	-
223	PAHGVTS*APDTRPAPGSTAP	T	Dr. U. Westerlind	-	17	-	-	28
224	PAHGVTSAPDT*RPAPGSTAP	T	Dr. U. Westerlind	-	18	-	-	-
225	PAHGVTSAPDTRPAPGS*TAP	T	Dr. U. Westerlind	-	19	-	-	29
226	PAHGVTSAPDTRPAPGST*AP	T	Dr. U. Westerlind	-	20	-	-	-
227	PAHGVT*SAPDTRPAP	T	Dr. U. Westerlind	-	21	-	-	30
228	PAHGVTS*APDTRPAP	T	Dr. U. Westerlind	-	22	-	-	31
229	PAHGVTSAPDTRPAPGS*TAP	αGlcNAc	Dr. U. Westerlind	-	83	-	-	-
230	PAHGVT*SAPDTRPAPGSTAP	αGlcNAc	Dr. U. Westerlind	-	84	-	-	-
231	PAHGVTS*APDTRPAPGSTA	αGlcNAc	Dr. U. Westerlind	-	85	-	-	-
232	HGVTSAPDT*RPAPGS*T*APPA	T_N	Dr. H. Cai	1	23	4	64	14
233	HGVTSAPDTRPAPGS*T*APPA	T_N	Dr. H. Cai	-	24	5	65	15
234	HGVTSAPDT*RPAPGS*T*APPA	T	Dr. H. Cai	1	25	6	-	32
235	HGVTSAPDTRPAPGS*T*APPA	T	Dr. H. Cai	-	26	7	-	33
236	HGVTSAPDTRPAPGST*APPA	T_N	Dr. H. Cai	-	-	8	-	16
237	HGVTSAPDTRPAPGS*TAPPA	T_N	Dr. H. Cai	1	-	9	-	17
238	HGVTSAPDT*RPAPGS*TAPPA	T_N	Dr. H. Cai	-	-	10	-	18
239	HGVTSAPDT*RPAPGST*APPA	T_N	Dr. H. Cai	-	-	11	-	19
240	HGVTSAPDT*RPAPGSTAPPA	T_N	Dr. H. Cai	-	-	13	-	20
241	HGVTSAPDTRPAPGSTAPPA	-	Dr. H. Cai	-	-	19	-	1

7.3.4.1 Format of microarray **MA1**

Microarray **MA1** was utilized for antibody titer determination (chapter 4.3.7). **MA1** was spotted homogenously with each of the corresponding antigen peptides of the vaccine candidates **229**, **231** or **234** (*table 6.1*). The glycopeptide was spotted in 39 arrays with 7 x 7 = 49 replicates. The pitch was set to 450 µM in each direction (*figure 6.2*).

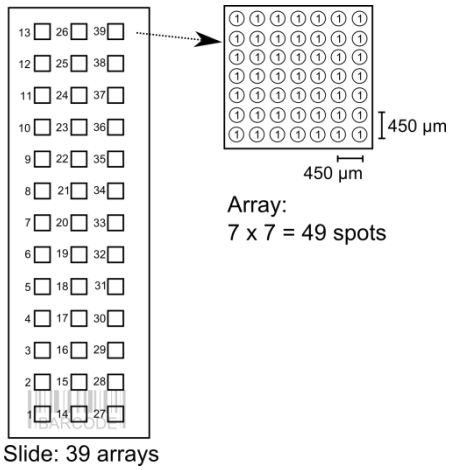


***Figure 7.3*:** Format of microarray **MA1**.

The murine serum antibodies were detected with *Alexa Fluor® 488* goat-anti-mouse IgG (1:3000) as described in the general procedure for array incubation (*chapter 6.3.5*). Central nine replicates were chosen for data acquisition.

7.3.4.2 Format of microarray **MA2**

Microarray **MA2** was utilized for serum antibody screening (*chapter 4.3.8*). Glycopeptides were arranged in seven rectangular arrays per slide. The barcode area was left unspotted. Each array consisted of four identical sub-blocks of eight columns and 22 rows, resulting in maximum 176 spots per sub-block and 4 x 176 = 704 possible spots per array. In total 85 MUC1 glycopeptides were spotted two times into each sub-block (2 x 85 = 170 spots; 176 (max) - 170 (used) = 6 spots unused), resulting in eight spot replicates per array. Peptides were spotted in the order according to *table 6.1*. Pitch spacing was set to 335 µm in x-direction and 327 µm y-direction (*figure 6.3*).

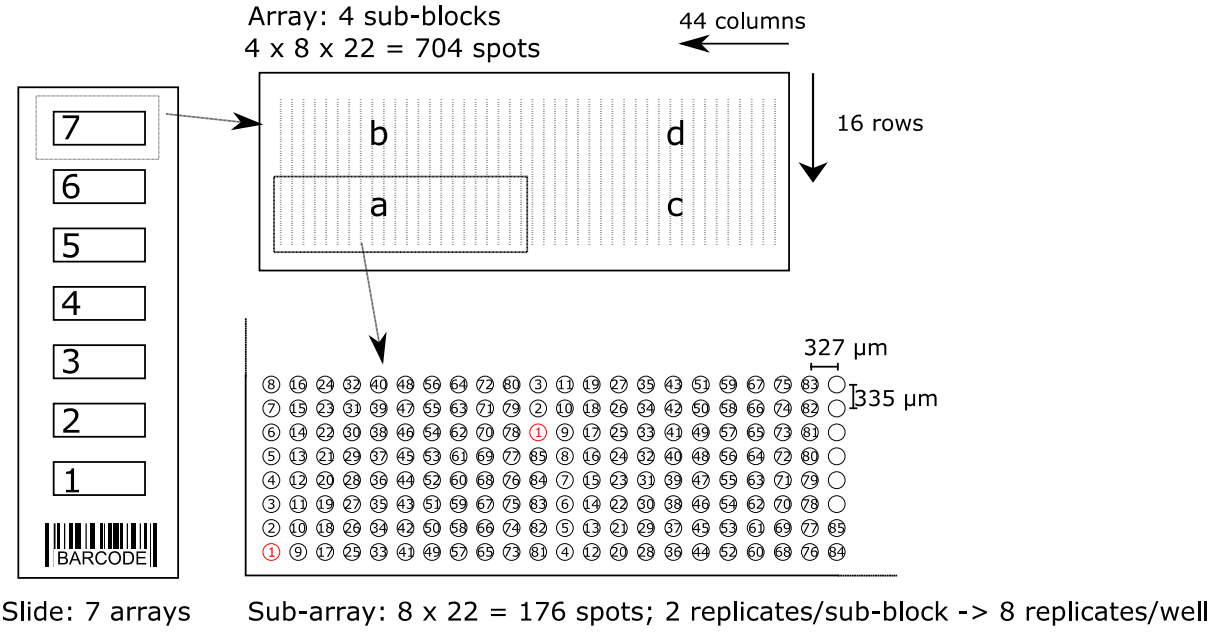


***Figure 7.4*:** Format of microarray **MA2**.

The murine serum antibodies from synthetic vaccine candidates **1**-**8** were detected with *Alexa Fluor® 488* goat-anti-mouse IgG (1:3000) as described in the general procedure for array incubation (*chapter 6.3.5)*.

7.3.4.3 Format of microarray **MA3**

Microarray **MA3** was utilized in serum antibody screening (*chapter 4.3.8*). Glycopeptides were arranged in 16 wells on the microarray slide. Each array was spotted in a 12 x 12 = 144 spots pattern. The glycopeptides were spotted with 7 replicates each (*figure 6.3*). Pitch spacing was set to 380 µm in x-direction and 410 µm y-direction (*figure 6.4*)

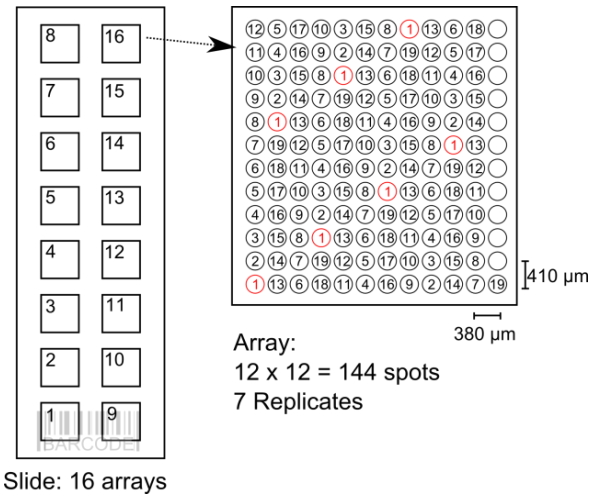


***Figure 7.5*:** Format of microarray **MA3**.

The murine serum antibodies from synthetic vaccine candidates **1**-**8** were detected with *Alexa Fluor® 488* goat-anti-mouse IgG (1:3000) as described in the general procedure for array incubation (*chapter 6.3.5*).

7.3.4.4 Format of microarray **MA4**

Microarray **MA4** was utilized in serum antibody screening and galectin-3 screening (*chapter 4.3.9* and *chapter 4.3.11.3*). Glycopeptides were arranged in eight arrays per slide. Each array consisted of two identical sub-blocks of 12 columns and 17 rows, resulting in maximum 204 spots per sub-block and 2 x 204 = 408 spots per array. In total 65 MUC1 glycopeptides were spotted three times into each sub-block (3 x 65 = 195 spots; 204 (max) - 195 (used) = 9 spots unused), resulting in seven spot replicates per array. Peptides were spotted in the order according to *table 6.1*. Pitch spacing was set to 400 µm in x-direction and 400 µm y-direction (*figure 6.5*).

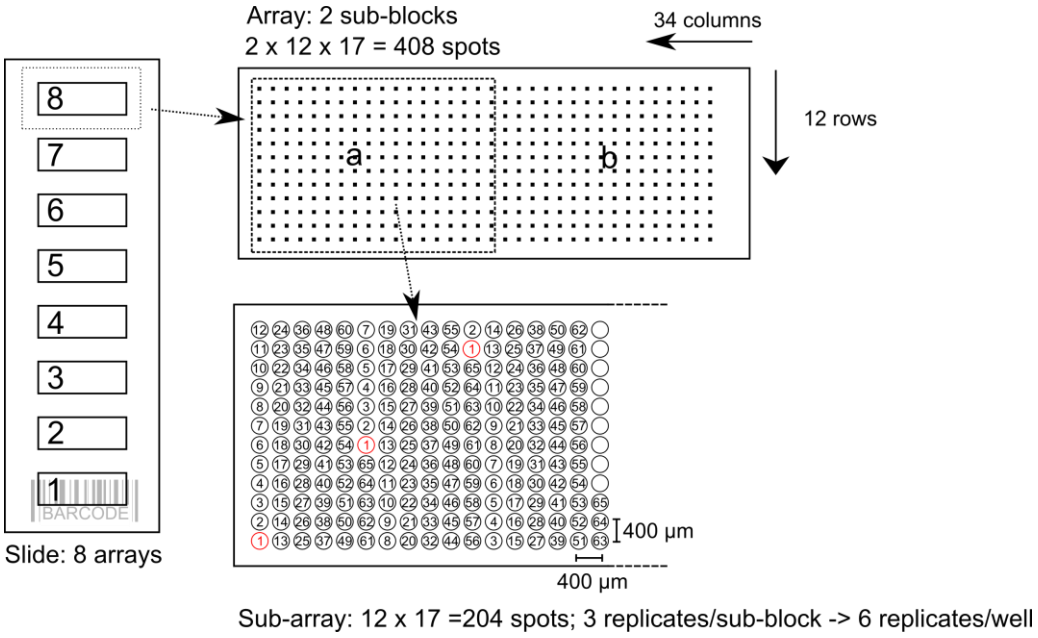


***Figure 7.6*:** Format of microarray **MA4**.

The murine serum antibodies from synthetic vaccine candidates **1**-**8** were detected with *Alexa Fluor®* 488 goat-anti-mouse IgG (1:3000) and galectin-3 was detected with *Alexa Fluor®* 488 anti-mouse/human MAC2/galectin-3 (1:500) an as described in the general procedure for array incubation (*chapter 6.3.5*).

7.3.4.5 Format of microarray **MA5**

Microarray **MA5** was utilized in plant lectin screening (*chapter 4.3.11.1*). Glycopeptides were arranged in eight arrays per slide. Each array contained spots arranged in 16 columns and 42 rows, resulting in maximum 672 spots per array. In total 132 MUC1 glycopeptides were spotted five spot replicates per array (5 x 132 = 660 spots; 672 (max) - 660 (used) = 12 spots unused). Peptides were spotted in the order according to *table 6.1*. Pitch spacing was set to 340 µm in x-direction and 343 µm in y-direction (*figure 6.6*).

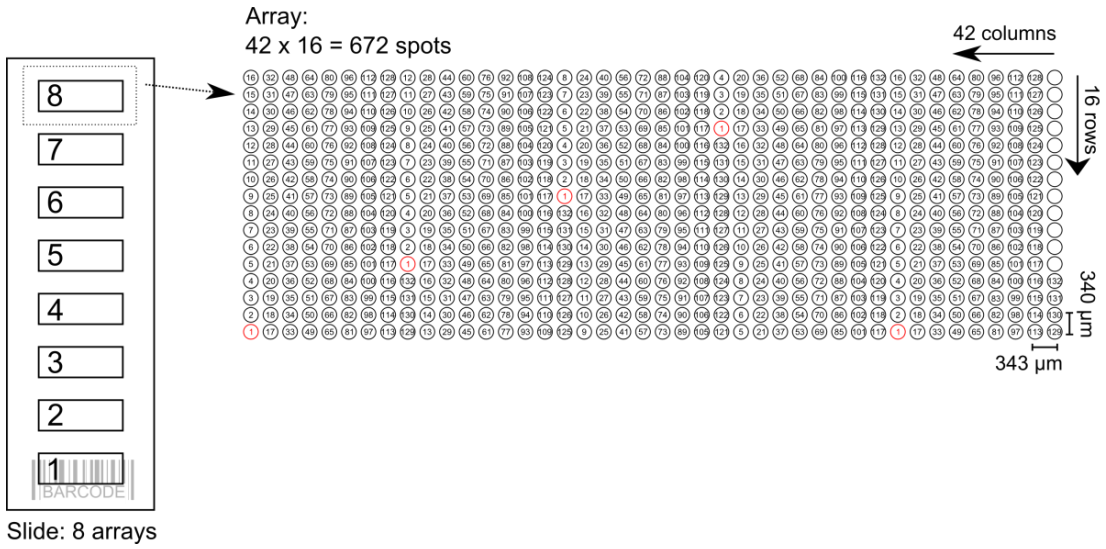


***Figure 7.7*:** Format of Microarray **MA5**.

The biotinylated plant lectins were detected with Cy5-Streptavidin Conjugate (1:1500) as described in the general procedure for array incubation (*chapter 6.3.5*).

7.3.5 General protocol for microarray assays with antiserum or lectin samples

Reagents:

- Block-buffer: 25 mM ethanolamine in 100 mM sodium tetraborate buffer (pH 9.0)
- Incubation-buffer: PBST-buffer (0.2% Tween-20): 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4, 0.2% Tween-20.
- Wash-buffer 1: PBST-buffer (0.05%)
- Wash-buffer 2: PBS-buffer

The volumes for all incubation and washing steps were depended on the corresponding well size of the microarray:

- **MA1**: 3 x 3 mm → 5 µL
- **MA2**, **MA4**, **MA5**: 6.59 x 15.58 mm → 100 µL
- **MA3**: 6.59 x 6.59 mm → 50 µL

General incubation procedure for microarray assay:

Blocking:

1. The slides are immersed in a bath of block-buffer. No further blocking with a protein component (i.e. BSA) is required.
2. The slides are rinsed three times with water and then spin dried in a centrifuge.
3. The slide is mounted into a slide holder with a well-forming silicone superstructure.
4. Incubation with antiserum or lectin:
   The antiserum/lectin is diluted in incubation-buffer and pipetted into the wells. The slide holder is positioned into a humidity chamber (70% RH) and gently shaken for 60 min.
5. The wells are washed two times wash buffer 1 and once with wash buffer 2 for 15 min each.
6. Incubation with secondary antibody:
   Wash-buffer is carefully pipetted out of the wells and of secondary antibody diluted in incubation-buffer is added. The slide holder is positioned into a humidity chamber (70% RH) and gently shaken for 60 min.
7. Washing according to step 5.
8. Slides are removed from the slide holder, the silicone superstructure is removed and the slides are generously rinsed with water and spin-dried.
9. Slides are scanned for fluorescence.

8 REFERENCES

1. J. Dekker, J. W. Rossen, H. A. Büller, A. Einerhand, *Trends in Biochemical Sciences* **2002**, 27, 126.
2. D. J. Thornton, J. K. Sheehan, *Proc. Am. Thorac. Soc.* **2004**, 1, 54.
3. S. J. Gendler, A. P. Spicer, *Annu. Rev. Physiol.* **1995**, 57, 607.
4. J. Taylor-Papadimitriou, J. A. Peterson, J. Arklie, J. Burchell, R. L. Ceriani, W. F. Bodmer, *Int. J. Cancer* **1981**, 28, 17.
5. J. Hilkens, Ligtenberg, Marjolÿn J. L. H. L. Vos, S. V. Litvinov, *Trends in Biochemical Sciences* **1992**, 17, 359.
6. S. Zrihan-Licht, A. Baruch, O. Elroy-Stein, I. Keydar, D. H. Wreschner, *FEBS Letters* **1994**, 356, 130.
7. P. K. Singh, M. A. Hollingsworth, *Trends in Cell Biology* **2006**, 16, 467.
8. D. J. Thornton, K. Rousseau, M. A. McGuckin, *Annu. Rev. Physiol.* **2008**, 70, 459.
9. S. Gendler, J. Taylor-Papadimitriou, T. Duhig, J. Rothbard, J. Burchell, *J. Biol. Chem.* **1988**, 263, 12820.
10. J.-L. Desseyn, V. Guyonnet-Duperat, N. Porchet, J.-P. Aubert, A. Laine, *J. Biol. Chem.* **1997**, 272, 3168.
11. S. E. Baldus, K. Engelmann, F.-G. Hanisch, *Crit. Rev. Clin. Lab. Sci.* **2004**, 41, 189.
12. N. Jentoft, *Trends in Biochemical Sciences* **1990**, 15, 291.
13. F. Levitin, O. Stern, M. Weiss, C. Gil-Henn, R. Ziv, Z. Prokocimer, N. I. Smorodinsky, D. B. Rubinstein, D. H. Wreschner, *J. Biol. Chem.* **2005**, 280, 33374.
14. C. L. Kinlough, R. J. McMahan, P. A. Poland, J. B. Bruns, K. L. Harkleroad, R. J. Stremple, O. B. Kashlan, K. M. Weixel, O. A. Weisz, R. P. Hughey, *J. Biol. Chem.* **2006**, 281, 12112.
15. A. Halim, G. Brinkmalm, U. Rüetschi, A. Westman-Brinkmalm, E. Portelius, H. Zetterberg, K. Blennow, G. Larson, J. Nilsson, *Proc. Natl. Acad. Sci. U.S.A.* **2011**, 108, 11848.

16. C. Steentoft, S. Y. Vakhrushev, M. B. Vester-Christensen, Schjoldager, Katrine T-B G, Y. Kong, E. P. Bennett, U. Mandel, H. Wandall, S. B. Levery, H. Clausen, *Nat. Methods* **2011**, 8, 977.

17. I. Brockhausen. in *Comprehensive Glycoscience,* edited by H. Kamerling (Elsevier, Oxford, 2007), pp. 33–59.

18. T. Iwai, N. Inaba, A. Naundorf, Y. Zhang, M. Gotoh, H. Iwasaki, T. Kudo, A. Togayachi, Y. Ishizuka, H. Nakanishi, H. Narimatsu, *J. Biol. Chem.* **2002**, 277, 12802.

19. Y. Li, X. Chen, *Appl. Microbiol. Biotechnol.* **2012**, 94, 887-905.

20. R. McEver, *Glycoconj. J.* **1997**, 14, 585-591.

21. Jan-Willem Van Klinken, B. A. Einerhand, H. A. Büller, J. Dekker, *Analytical Biochemistry* **1998**, 265, 103.

22. J. Taylor-Papadimitriou, J. Burchell, D. W. Miles, M. Dalziel, *Biochim. Biophys. Acta Molecular Basis of Disease* **1999**, 1455, 301.

23. P. D. Vermeer, L. A. Einwalter, T. O. Moninger, T. Rokhlina, J. A. Kern, J. Zabner, M. J. Welsh, *Nature* **2003**, 422, 322.

24. K. Shin, V. C. Fogg, B. Margolis, *Annu. Rev. Cell Dev. Biol.* **2006**, 22, 207.

25. Y. Li, W.-h. Yu, J. Ren, W. Chen, L. Huang, S. Kharbanda, M. Loda, D. Kufe, *Mol. Cancer Res.* **2003**, 1, 765.

26. M. Yamamoto, A. Bharti, Y. Li, D. Kufe, *J. Biol. Chem.* **1997**, 272, 12492.

27. Ligtenberg, Marjolijn J. L. F. Buijs, H. L. Vos, J. Hilkens, *Cancer Res.* **1992**, 52, 2318.

28. J. Wesseling, *J. Cell Biol.* **1995**, 129, 255.

29. I. Brockhausen, J.-M. Yang, J. Burchell, C. Whitehouse, J. Taylor-Papadimitriou, *Eur. J. Biochem.* **1995**, 233, 607.

30. M. Dalziel, C. Whitehouse, I. McFarlane, I. Brockhausen, S. Gschmeissner, T. Schwientek, H. Clausen, J. M. Burchell, J. Taylor-Papadimitriou, *J. Biol. Chem.* **2001**, 276, 11007.

31. R. Sewell, M. Bäckström, M. Dalziel, S. Gschmeissner, H. Karlsson, T. Noll, J. Gätgens, H. Clausen, G. C. Hansson, J. Burchell, J. Taylor-Papadimitriou, *J. Biol. Chem.* **2006**, 281, 3586.

32. A. Solatycka, T. Owczarek, F. Piller, V. Piller, B. Pula, L. Wojciech, M. Podhorska-Okolow, P. Dziegiel, M. Ugorski, *Glycobiology* **2012**, 22, 1042.

33. F. Dall'Olio, M. Chiricolo, *Glycoconj. J.* **2001**, 18, 841.

34. S. J. Storr, L. Royle, C. J. Chapman, Hamid, Umi M Abd, J. F. Robertson, A. Murray, R. A. Dwek, P. M. Rudd, *Glycobiology* **2008**, 18, 456.

35. N. Matsuura, T. Narita, N. Hiraiwa, M. Hiraiwa, H. Murai, T. Iwase, H. Funahashi, T. Imai, H. Takagi, R. Kannagi, *Int. J. Oncol.* **1998**,

36. G. An, B. Wei, B. Xia, J. M. McDaniel, T. Ju, R. D. Cummings, J. Braun, L. Xia, *J. Exp. Med.* **2007**, 204, 1417.

37. T. Iwai, T. Kudo, R. Kawamoto, T. Kubota, A. Togayachi, T. Hiruma, T. Okada, T. Kawamoto, K. Morozumi, H. Narimatsu, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, 102, 4572.

38. P. Radhakrishnan, P. M. Grandgenett, A. M. Mohr, S. K. Bunt, F. Yu, S. Chowdhury, M. A. Hollingsworth, *Int. J. Cancer* **2013**, 133, 2824.

39. T. Ju, R. D. Cummings, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, 99, 16613.

40. T. Ju, G. S. Lanneau, T. Gautam, Y. Wang, B. Xia, S. R. Stowell, M. T. Willard, W. Wang, J. Y. Xia, R. E. Zuna, Z. Laszik, D. M. Benbrook, M. H. Hanigan, R. D. Cummings, *Cancer Res.* **2008**, 68, 1636.

41. Y. Wang, T. Ju, X. Ding, B. Xia, W. Wang, L. Xia, M. He, R. D. Cummings, *Proc. Natl. Acad. Sci. U.S.A.* **2010**, 107, 9228.

42. T. Koike, N. Kimura, K. Miyazaki, T. Yabuta, K. Kumamoto, S. Takenoshita, J. Chen, M. Kobayashi, M. Hosokawa, A. Taniguchi, T. Kojima, N. Ishida, M. Kawakita, H. Yamamoto, H. Takematsu, A. Suzuki, Y. Kozutsumi, R. Kannagi, R. Kanangi, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, 101, 8132.

43. K. Kumamoto, Y. Goto, K. Sekikawa, S. Takenoshita, N. Ishida, M. Kawakita, R. Kannagi, *Cancer Res.* **2001**, 61, 4620.

44. G. Springer, *Science* **1984**, 224, 1198.

45. B. J. Campbell, I. A. Finnie, E. F. Hounsell, J. M. Rhodes, *J. Clin. Invest.* **1995**, 95, 571.

46. M. R. Price, P. D. Rye, E. Petrakou, A. Murray, K. Brady, S. Imai, S. Haga, Y. Kiyozuka, D. Schol, Meulenbroek, M. F. A. Snijdewint, F. G. M. S. von Mensdorff-Pouilly, R. A. Verstraeten, K. Kenemans, A. Blockzjil, N. Nilsson, O. Nilsson, R. Reddish, M. R. Suresh, K. Koganty, S. Fortier, B. Baronic, A. Berg, M. B. Longenecker, H. Hilkens, M. Boer, K. Karanikas, McKenzie, I. F. C. G. Galanina, L. A. Simeoni, A. G. Ter-Grigoryan, I. M. Belyanchikov, N. V. Bovin, Y. Cao, U. Karsten, J. Dai, W. J. Allard, G. Davis, K. K. Yeung, F.-G. Hanisch, K. O. Lloyd, V. Kudryashov, R. Sikut, A. Sikut, K. Zhang, D. Baeckström, G. C. Hansson, C. A. Reis, H. Hassan, E. P. Bennett, H. Claussen, L. Norum, T. Varaas, B. Kierulf, K. Nustad, P. Ciborowski, W. M. Konitzki, J. Magarian-Blander, O. J. Finn, J. Hilgers, *Tumor Biol.* **1998**, 19(suppl 1), 1.

47. S. von Mensdorff-Pouilly, E. Petrakou, P. Kenemans, K. van Uffelen, A. A. Verstraeten, Snijdewint, Frank G. M. van Kamp, Gerard J. D. J. Schol, C. A. Reis, M. R. Price, P. O. Livingston, J. Hilgers, *Int. J. Cancer* **2000**, 86, 702.

48. J. Taylor-Papadimitriou, J. Burchell, T. Plunkett, R. Graham, I. Correa, D. Miles, M. Smith, *J. Mammary Gland. Biol. Neoplasia* **2002**, 7, 209-221.

49. A. Danielczyk, R. Stahn, D. Faulstich, A. Löffler, A. Märten, U. Karsten, S. Goletz, *Cancer Immunol. Immunother.* **2006**, 55, 1337-1347.

50. M. J. Scanlon, S. D. Morley, D. E. Jackson, M. R. Price, S. J. Tendler, *Biochem. J.* **1992**, 284 (Pt 1), 137.

51. J. D. Fontenot, S. V. Mariappan, P. Catasti, N. Domenech, O. J. Finn, G. Gupta, *J. Biomol. Struct. Dyn.* **1995**, 13, 245.

52. O. J. Finn, K. R. Jerome, R. A. Henderson, G. Pecher, N. Domenech, J. Magarian-Blander, S. M. Barratt-Boyes, *Immunol. Rev.* **1995**, 145, 61.

53. J. Schuman, A. P. Campbell, R. R. Koganty, B. M. Longenecker, *J. Pept. Res.* **2003**, 61, 91.

54. S. Dziadek, C. Griesinger, H. Kunz, U. M. Reinscheid, *Chem. Eur. J.* **2006**, 12, 4981.

55. J. S. Grinstead, R. R. Koganty, M. J. Krantz, B. M. Longenecker, A. P. Campbell, *Biochemistry* **2002**, 41, 9946.

56. L. Kirnarsky, O. Prakash, S. M. Vogen, M. Nomoto, M. A. Hollingsworth, S. Sherman, *Biochemistry* **2000**, 39, 12076.

57. T. Matsushita, N. Ohyabu, N. Fujitani, K. Naruchi, H. Shimizu, H. Hinou, S.-I. Nishimura, *Biochemistry* **2013**, 52, 402.

58. M. A. Tarp, A. L. Sorensen, U. Mandel, H. Paulsen, J. Burchell, J. Taylor-Papadimitriou, H. Clausen, *Glycobiology* **2007**, 17, 197.

59. H. H. Wandall, O. Blixt, M. A. Tarp, J. W. Pedersen, E. P. Bennett, U. Mandel, G. Ragupathi, P. O. Livingston, M. A. Hollingsworth, J. Taylor-Papadimitriou, J. Burchell, H. Clausen, *Cancer Res.* **2010**, 70, 1306.

60. O. Blixt, E. Cló, A. S. Nudelman, K. K. Sørensen, T. Clausen, H. H. Wandall, P. O. Livingston, H. Clausen, K. J. Jensen, *J. Proteome Res.* **2010**, 9, 5250.

61. O. Blixt, D. Bueti, B. Burford, D. Allen, S. Julien, M. Hollingsworth, A. Gammerman, I. Fentiman, J. Taylor-Papadimitriou, J. M. Burchell, *Breast Cancer Res.* **2011**, 13, R25.

62. J. W. Pedersen, O. Blixt, E. P. Bennett, M. A. Tarp, I. Dar, U. Mandel, S. S. Poulsen, A. E. Pedersen, S. Rasmussen, P. Jess, H. Clausen, H. H. Wandall, *Int. J. Cancer* **2011**, 128, 1860.

63. F.-G. Hanisch, T. Stadie, K. Boßlet, *Cancer Res.* **1995**, 55, 4036.

64. M. Yamamoto, V. P. Bhavanandan, S. Nakamori, T. Irimura, *Cancer Sci.* **1996**, 87, 488.

65. F.-G. Hanisch, T. Schwientek, Von Bergwelt-Baildon, Michael S, J. L. Schultze, O. Finn, *Eur. J. Immunol.* **2003**, 33, 3242.

66. F.-G. Hanisch, *Biochem. Soc. Trans* **2005**, 33, 705.

67. J. Hirabayashi, T. Hashidate, Y. Arata, N. Nishi, T. Nakamura, M. Hirashima, T. Urashima, T. Oka, M. Futai, Muller, Werner E. G, F. Yagi, K.-i. Kasai, *Biochim. Biophys. Acta General Subjects* **2002**, 1572, 232.

68. M. Viguier, T. Advedissian, D. Delacour, F. Poirier, F. Deshayes, *Tissue Barriers* **2014**, 2, e29103.

69. H. Ahmed, P. Guha, E. Kaptan, G. Bandyopadhyaya, *Trends in carbohydrate research* **2011**, 3, 13.

70. G. Elad-Sfadia, R. Haklai, E. Balan, Y. Kloog, *J. Biol. Chem.* **2004**, 279, 34922.

71. R. Shalom-Feuerstein, T. Cooks, A. Raz, Y. Kloog, *Cancer Res.* **2005**, 65, 7292.

72. Y. Takenaka, T. Fukumori, T. Yoshii, N. Oka, H. Inohara, H.-R. C. Kim, R. S. Bresalier, A. Raz, *Mol. Cell Biol.* **2004**, 24, 4395.

73. T. Shimura, Y. Takenaka, T. Fukumori, S. Tsutsumi, K. Okada, V. Hogan, A. Kikuchi, H. Kuwano, A. Raz, *Cancer Res.* **2005**, 65, 3535.

74. S. Song, N. Mazurek, C. Liu, Y. Sun, Q. Q. Ding, K. Liu, M.-C. Hung, R. S. Bresalier, *Cancer Res.* **2009**, 69, 1343.

75. T. Shimura, Y. Takenaka, S. Tsutsumi, V. Hogan, A. Kikuchi, A. Raz, *Cancer Res.* **2004**, 64, 6363.

76. J. Dumic, S. Dabelic, M. Flögel, *Glycoproteomics* **2006**, 1760, 616.

77. R. S. Bresalier, J. C. Byrd, L. Wang, A. Raz, *Cancer Res.* **1996**, 56, 4354.

78. S. Senapati, P. Chaturvedi, W. G. Chaney, S. Chakraborty, V. S. Gnanapragassam, A. R. Sasson, S. K. Batra, *Clin. Cancer Res.* **2011**, 17, 267.

79. L.-G. Yu, N. Andrews, Q. Zhao, D. McKean, J. F. Williams, L. J. Connor, O. V. Gerasimenko, J. Hilkens, J. Hirabayashi, K. Kasai, J. M. Rhodes, *J. Biol. Chem.* **2007**, 282, 773.

80. V. V. Glinsky, M. E. Huflejt, G. V. Glinsky, S. L. Deutscher, T. P. Quinn, *Cancer Res.* **2000**, 60, 2584.

81. V. V. Glinsky, G. V. Glinsky, K. Rittenhouse-Olson, M. E. Huflejt, O. V. Glinskii, S. L. Deutscher, T. P. Quinn, *Cancer Res.* **2001**, 61, 4851.

82. S. K. Khaldoyanidi, V. V. Glinsky, L. Sikora, A. B. Glinskii, V. V. Mossine, T. P. Quinn, G. V. Glinsky, P. Sriramarao, *J. Biol. Chem.* **2003**, 278, 4127.

83. Q. Zhao, X. Guo, G. B. Nash, P. C. Stone, J. Hilkens, J. M. Rhodes, L.-G. Yu, *Cancer Res.* **2009**, 69, 6799.

84. Q. Zhao, M. Barclay, J. Hilkens, X. Guo, H. Barrow, J. M. Rhodes, L.-G. Yu, *Mol. Cancer* **2010**, 9, 154.

85. P. Matarrese, O. Fusco, N. Tinari, C. Natoli, F.-T. Liu, M. L. Semeraro, W. Malorni, S. Iacobelli, *Int. J. Cancer* **2000**, 85, 545.

86. S. M. Frisch, R. A. Screaton, *Curr. Opin. Cell Biol.* **2001**, 13, 555.

87. T. Fukumori, Y. Takenaka, T. Yoshii, H.-R. C. Kim, V. Hogan, H. Inohara, S. Kagawa, A. Raz, *Cancer Res.* **2003**, 63, 8302.

88. Y. Suzuki, T. Inoue, T. Yoshimaru, C. Ra, *Biochim Biophys Acta Molecular Cell Research* **2008**, 1783, 924.

89. P. Guha, E. Kaptan, G. Bandyopadhyaya, S. Kaczanowska, E. Davila, K. Thompson, S. S. Martin, D. V. Kalvakolanu, G. R. Vasta, H. Ahmed, *Proc. Natl. Acad. Sci. U.S.A.* **2013**, 110, 5052.

90. W. Peng, H. Y. Wang, Y. Miyahara, G. Peng, R.-F. Wang, *Cancer Res.* **2008**, 68, 7228.

91. O. Suzuki, M. Abe, *Oncol. Rep.* **2008**, 19, 743.

92. B. N. Stillman, D. K. Hsu, M. Pang, C. F. Brewer, P. Johnson, F.-T. Liu, L. G. Baum, *J. Immunol.* **2006**, 176, 778.

93. Y. Zhuo, R. Chammas, S. L. Bellis, *J. Biol. Chem.* **2008**, 283, 22177.

94. J. Xue, X. Gao, C. Fu, Z. Cong, H. Jiang, W. Wang, T. Chen, Q. Wei, C. Qin, *FEBS Lett.* **2013**, 587, 3986.

95. H. Barrow, X. Guo, H. H. Wandall, J. W. Pedersen, B. Fu, Q. Zhao, C. Chen, J. M. Rhodes, L.-G. Yu, *Clinical Cancer Research* **2011**, 17, 7035.

96. S. Ramasamy, S. Duraisamy, S. Barbashov, T. Kawano, S. Kharbanda, D. Kufe, *Mol. Cell* **2007**, 27, 992.

97. Graves, C R L, Robertson, J F R, A. Murray, M. R. Price, C. J. Chapman, *J. Pept. Res.* **2005**, 66, 357.

98. J. Li, Z. Zhu, *Acta Pharmacol. Sin.* **2010**, 31, 1198.

99. J. G. Elvin, R. G. Couston, van der Walle, Christopher F. *Int. J. Pharm.* **2013**, 440, 83.

100. Hwang, William Ying Khee, J. Foote, *Methods* **2005**, 36, 3.

101. O. J. Finn, *J. Immunol.* **2008**, 181, 1589.

102. J. Neefjes, H. Ovaa, *Nat. Chem. Biol.* **2013**, 9, 769.

103. F. R. Carbone, P. A. Gleeson, *Glycobiology* **1997**, 7, 725.

104. C. M. Cabrera, P. Jiménez, T. Cabrera, C. Esparza, F. Ruiz-Cabello, F. Garrido, *Tissue Antigens* **2003**, 61, 211.

105. F. Garrido, I. Algarra. in *Advances in Cancer Research* (Academic Press2001), pp. 117–158.

106. H. T. Khong, N. P. Restifo, *Nat. Immunol.* **2002**, 3, 999.

107. B. Agrawal, M. J. Krantz, M. A. Reddish, B. M. Longenecker, *Nat. Med.* **1998**, 4, 43.

108. A. K. Chan, D. C. Lockhart, W. v. Bernstorff, R. A. Spanjaard, H.-G. Joo, T. J. Eberlein, P. S. Goedegebuure, *Int. J. Cancer* **1999**, 82, 721.

109. E. C. Beuvery, F. van Rossum, J. Nagel, *Infec. Immun.* **1982**, 37, 15.

110. B. M. Longenecker, M. Reddish, R. Koganty, G. MacLean. in *Antigen and Antibody Molecular Engineering in Breast Cancer Diagnosis and Treatment,* edited by R. Ceriani (Springer US1994), pp. 105-124.

111. G. Ragupathi, S. F. Slovin, S. Adluri, D. Sames, I. J. Kim, H. M. Kim, M. Spassova, W. G. Bornmann, K. O. Lloyd, H. I. Scher, P. O. Livingston, S. J. Danishefsky, *Angew. Chem. Int. Ed.* **1999**, 38, 563.

112. E. Kagan, G. Ragupathi, S. Yi, C. Reis, J. Gildersleeve, D. Kahne, H. Clausen, S. Danishefsky, P. Livingston, *Cancer Immunol. Immunother.* **2005**, 54, 424-430.

113. S. F. Slovin, G. Ragupathi, C. Musselli, K. Olkiewicz, D. Verbel, S. D. Kuduk, J. B. Schwarz, D. Sames, S. Danishefsky, P. O. Livingston, H. I. Scher, *J. Clin. Oncol.* **2003**, 21, 4292.

114. T. Gilewski, S. Adluri, G. Ragupathi, S. Zhang, T.-J. Yao, K. Panageas, M. Moynahan, A. Houghton, L. Norton, P. O. Livingston, *Clin. Cancer Res.* **2000**, 6, 1693.

115. M. A. Reddish, G. D. MacLean, R. R. Koganty, J. Kan-Mitchell, V. Jones, M. S. Mitchell, B. M. Longenecker, *Int. J. Cancer* **1998**, 76, 817.

116. M. M. Soares, V. Mehta, O. J. Finn, *J. Immunol.* **2001**, 166, 6555.

117. B. Acres, V. Apostolopoulos, J.-M. Balloul, D. Wreschner, P.-X. Xing, D. Ali-Hadji, N. Bizouarne, M. P. Kieny, McKenzie, Ian F. C. *Cancer Immunol. Immunother.* **2000**, 48, 588-594.

118. G. J. Rowse, R. M. Tempero, M. L. VanLith, M. A. Hollingsworth, S. J. Gendler, *Cancer Res.* **1998**, 58, 315.

119. C. Butts, M. A. Socinski, P. L. Mitchell, N. Thatcher, L. Havel, M. Krzakowski, S. Nawrocki, T.-E. Ciuleanu, L. Bosquée, J. M. Trigo, A. Spira, L. Tremblay, J. Nyman, R. Ramlau, G. Wickart-Johansson, P. Ellis, O. Gladkov, J. R. Pereira, Eberhardt, Wilfried Ernst Erich, C. Helwig, A. Schröder, F. A. Shepherd, *The Lancet Oncol.* **2014**, 15, 59.

120. Merck. Press release. Available at http://www.merckgroup.com/en/media/extNewsDetail.html?newsId=8475BA17A3F51470C1257D50006901B4&newsType=1.

121. V. Lakshminarayanan, P. Thompson, M. A. Wolfert, T. Buskas, J. M. Bradley, L. B. Pathangey, C. S. Madsen, P. A. Cohen, S. J. Gendler, G.-J. Boons, *Proc. Natl. Acad. Sci. U.S.A.* **2012**, 109, 261.

122. H. Cai, Z.-H. Huang, L. Shi, P. Zou, Y.-F. Zhao, H. Kunz, Y.-M. Li, *Eur. J. Org. Chem.* **2011**, 2011, 3685.

123. H. Cai, Z.-H. Huang, L. Shi, Z.-Y. Sun, Y.-F. Zhao, H. Kunz, Y.-M. Li, *Angew. Chem. Int. Ed.* **2012**, 51, 1719.

124. A. L. Sorensen, *Glycobiology* **2005**, 16, 96.

125. J. Zhu, Q. Wan, D. Lee, G. Yang, M. K. Spassova, O. Ouerfelli, G. Ragupathi, P. Damani, P. O. Livingston, S. J. Danishefsky, *J. Am. Chem. Soc.* **2009**, 131, 9298.

126. N. Gaidzik, A. Kaiser, D. Kowalczyk, U. Westerlind, B. Gerlitzki, H. P. Sinn, E. Schmitt, H. Kunz, *Angew. Chem. Int. Ed.* **2011**, 50, 9977.

127. A. Hoffmann-Röder, A. Kaiser, S. Wagner, N. Gaidzik, D. Kowalczyk, U. Westerlind, B. Gerlitzki, E. Schmitt, H. Kunz, *Angew. Chem. Int. Ed.* **2010**, 49, 8498.

128. A. Kaiser, N. Gaidzik, U. Westerlind, D. Kowalczyk, A. Hobel, E. Schmitt, H. Kunz, *Angew. Chem. Int. Ed.* **2009**, 48, 7551.

129. T. L. McCool, C. V. Harding, N. S. Greenspan, J. R. Schreiber, *Infect. Immun.* **1999**, 67, 4862.

130. L. A. Herzenberg, T. Tokuhisa, L. A. Herzenberg, *Nature* **1980**, 285, 664.

131. S. Dziadek, A. Hobel, E. Schmitt, H. Kunz, *Angew. Chem. Int. Ed.* **2005**, 44, 7630.

132. S. Ingale, M. A. Wolfert, T. Buskas, G.-J. Boons, *ChemBioChem* **2009**, 10, 455.

133. J. Alexander, M.-F. del Guercio, A. Maewal, L. Qiao, J. Fikes, R. W. Chesnut, J. Paulson, D. R. Bundle, S. DeFrees, A. Sette, *J. Immunol.* **2000**, 164, 1625.

134. H. Cai, M.-S. Chen, Z.-Y. Sun, Y.-F. Zhao, H. Kunz, Y.-M. Li, *Angew. Chem. Int. Ed.* **2013**, 52, 6106.

135. H. Cai, Z.-Y. Sun, M.-S. Chen, Y.-F. Zhao, H. Kunz, Y.-M. Li, *Angew. Chem. Int. Ed.* **2014**, 53, 1699.

136. A. Lahiri, P. Das, D. Chakravortty, *Vaccine* **2008**, 26, 6777.

137. A. O. Aliprantis, *Science* **1999**, 285, 736.

138. K. R. Jerome, D. L. Barnd, K. M. Bendt, C. M. Boyer, J. Taylor-Papadimitriou, McKenzie, Ian F. C. R. C. Bast, O. J. Finn, *Cancer Res.* **1991**, 51, 2908.

139. K. R. Jerome, N. Domenech, O. J. Finn, *J. Immunol.* **1993**, 151, 1654.

140. N. Doménech, R. A. Henderson, O. J. Finn, *J. Immunol.* **1995**, 155, 4766.

141. P. Brossart, K. S. Heinrich, G. Stuhler, L. Behnke, V. L. Reichardt, S. Stevanovic, A. Muhm, H.-G. Rammensee, L. Kanz, W. Brugger, *Blood* **1999**, 93, 4309.

142. V. Apostolopoulos, E. Yuriev, P. A. Ramsland, J. Halton, C. Osinski, W. Li, M. Plebanski, H. Paulsen, McKenzie, Ian F C, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, 100, 15029.

143. D. Stepensky, E. Tzehoval, E. Vadai, L. Eisenbach, *Clin. Exp. Immunol.* **2006**, 143, 139.

144. T. Ninkovic, L. Kinarsky, K. Engelmann, V. Pisarev, S. Sherman, O. J. Finn, F.-G. Hanisch, *Mol. Immunol.* **2009**, 47, 131.

145. A. R. Hearn, L. de Haan, A. J. Pemberton, T. R. Hirst, A. J. Rivett, *J. Biol. Chem.* **2004**, 279, 51315.

146. U. Seifert, C. Marañón, A. Shmueli, J.-F. Desoutter, L. Wesoloski, K. Janek, P. Henklein, S. Diescher, M. Andrieu, de la Salle, Henri, T. Weinschenk, H. Schild, D. Laderach, A. Galy, G. Haas, P.-M. Kloetzel, Y. Reiss, A. Hosmalin, *Nat. Immunol.* **2003**, 4, 375.

147. L. Shen, L. J. Sigal, M. Boes, K. L. Rock, *Immunity* **2004**, 21, 155.

148. C. Napoletano, A. Rughetti, Agervig Tarp, Mads P, J. Coleman, E. P. Bennett, G. Picco, P. Sale, K. Denda-Nagai, T. Irimura, U. Mandel, H. Clausen, L. Frati, J. Taylor-Papadimitriou, J. Burchell, M. Nuti, *Cancer Res.* **2007**, 67, 8358.

149. E. Saeland, S. van Vliet, M. Bäckström, van den Berg, VeniceC. M. T. H. Geijtenbeek, G. Meijer, Y. van Kooyk, *Cancer Immunol. Immunother.* **2007**, 56, 1225-1236.

150. M. C. Rose, J. A. Voynow, *Physiol. Rev.* **2006**, 86, 245.

151. V. Venkatakrishnan, N. H. Packer, M. Thaysen-Andersen, *Expert Rev Respir Med* **2013**, 7, 553.

152. A. Livraghi, S. H. Randell, *Toxicol. Pathol.* **2007**, 35, 116.

153. N. Manri, Y. Takegawa, N. Fujitani, A. Kaneko, A. Hirabayashi, S.-I. Nishimura, T. Sakamoto, *Anal. Sci.* **2012**, 28, 723.

154. F.-G. Hanisch, S. Müller, *Glycobiology* **2000**, 10, 439.

155. L. M. Likhosherstov, O. S. Novikova, V. A. Derevitskaja, N. K. Kochetkov, *Carbohydrate Res.* **1986**, 146, C1.

156. S. T. Anisfeld, P. T. Lansbury, *J. Org. Chem.* **1990**, 55, 5560.

157. S. T. Cohen-Anisfeld, P. T. Lansbury, *J. Am. Chem. Soc.* **1993**, 115, 10531.

158. Rising, Thomas W D F, C. D. Heidecke, Moir, James W B, Z. Ling, A. J. Fairbanks, *Chem. Eur. J.* **2008**, 14, 6444.

159. K. M. Koeller, Smith, Mark E. B. R.-F. Huang, C.-H. Wong, *J. Am. Chem. Soc.* **2000**, 122, 4241.

160. P. H. Jensen, D. Kolarich, N. H. Packer, *FEBS J.* **2010**, 277, 81.

161. R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, 85, 2149.

162. R. B. Merrifield, *J. Am. Chem. Soc.* **1964**, 86, 304.

163. L. A. Carpino, G. Y. Han, *J. Org. Chem.* **1972**, 37, 3404.

164. E. Atherton, H. Fox, D. Harkiss, C. J. Logan, R. C. Sheppard, B. J. Williams, *J. Chem. Soc. Chem. Commun.* **1978**, , 537.

165. C.-D. Chang, J. Meienhofer, *Int. J. Pep. Prot. Res.* **1978**, 11, 246.

166. J. C. Sheehan, G. P. Hess, *J. Am. Chem. Soc.* **1955**, 77, 1067.

167. J. Coste, D. Le-Nguyen, B. Castro, *Tetrahedron Lett.* **1990**, 31, 205.

168. V. Dourtoglou, J.-C. Ziegler, B. Gross, *Tetrahedron Lett.* **1978**, 19, 1269.

169. R. Knorr, A. Trzeciak, W. Bannwarth, D. Gillessen, *Tetrahedron Lett.* **1989**, 30, 1927.

170. W. König, R. Geiger, *Chem. Ber.* **1970**, 103, 788.

171. L. A. Carpino, *J. Am. Chem. Soc.* **1993**, 115, 4397.

172. L. A. Carpino, A. El-Faham, F. Albericio, *Tetrahedron Lett.* **1994**, 35, 2279.

173. I. Abdelmoty, F. Albericio, L. Carpino, B. Foxman, S. Kates, *Lett. Pept. Sci.* **1994**, 1, 57-67.

174. L. A. Carpino, H. Imazumi, A. El-Faham, F. J. Ferrer, C. Zhang, Y. Lee, B. M. Foxman, P. Henklein, C. Hanay, C. Mügge, H. Wenschuh, J. Klose, M. Beyermann, M. Bienert, *Angew. Chem. Int. Ed.* **2002**, 41, 441.

175. M. Meldal, T. Bielfeldt, S. Peters, K. J. Jensen, H. Paulsen, K. Bock, *Int. J. Pep. Prot. Res.* **1994**, 43, 529.

176. J. Kihlberg, T. Vuljanic, *Tetrahedron Lett.* **1993**, 34, 6135.

177. E. Pedroso, A. Grandas, de las Heras, Xavier, R. Eritja, E. Giralt, *Tetrahedron Lett.* **1986**, 27, 743.

178. Y. Yang, W. V. Sweeney, K. Schneider, S. Thörnqvist, B. T. Chait, J. P. Tam, *Tetrahedron Lett.* **1994**, 35, 9689.

179. J. Lauer, C. Fields, G. Fields, *Lett. Pept. Sci.* **1995**, 1, 197-205.

180. P. Sjölin, M. Elofsson, J. Kihlberg, *J. Org. Chem.* **1996**, 61, 560.

181. A. Michael, *Am. Chem. J.* **1879**, 1, 305.

182. E. Fischer, *Ber. Dtsch. Chem. Ges.* **1893**, 26, 2400.

183. G. H. Veeneman, van Boom, J. H, *Tetrahedron Lett.* **1990**, 31, 275.

184. G. H. Veeneman, van Leeuwen, S. H. van Boom, J. H. *Tetrahedron Lett.* **1990**, 31, 1331.

185. K. C. Nicolaou, S. P. Seitz, D. P. Papahatjis, *J. Am. Chem. Soc.* **1983**, 105, 2430.

186. P. Fügedi, P. J. Garegg, *Carbohydrate Res.* **1986**, 149, C9.

187. P. Fügedi, P. Garegg, H. Lönn, T. Norberg, *Glycoconj. J.* **1987**, 4, 97-108.

188. R. R. Schmidt, J. Michel, *Angew. Chem. Int. Ed.* **1980**, 19, 731.

189. J. R. Pougny, J. C. Jacquinet, M. Nassr, D. Duchet, M. L. Milat, P. Sinay, *J. Am. Chem. Soc.* **1977**, 99, 6762.

190. L. K. Mydock, A. V. Demchenko, *Org. Biomol. Chem.* **2010**, 8, 497.

191. D. R. Mootoo, P. Konradsson, U. Udodong, B. Fraser-Reid, *J. Am. Chem. Soc.* **1988**, 110, 5583.

192. B. Fraser-Reid, Z. Wu, U. E. Udodong, H. Ottosson, *J. Org. Chem.* **1990**, 55, 6068.

193. M. N. Kamat, A. V. Demchenko, *Org. Lett.* **2005**, 7, 3215.

194. L. K. Mydock, A. V. Demchenko, *Org. Lett.* **2008**, 10, 2107.

195. C. McDonnell, O. López, P. Murphy, Fernández Bolaños, José G, R. Hazell, M. Bols, *J. Am. Chem. Soc.* **2004**, 126, 12374.

196. B. Fraser-Reid, Z. Wu, C. W. Andrews, E. Skowronski, J. P. Bowen, *J. Am. Chem. Soc.* **1991**, 113, 1434.

197. E. Juaristi, G. Cuevas, *Tetrahedron* **1992**, 48, 5019.

198. F. Barresi, O. Hindsgaul, *J. Am. Chem. Soc.* **1991**, 113, 9376.

199. A. Ishiwata, Y. Munemura, Y. Ito, *Eur. J. Org. Chem.* **2008**, 2008, 4250.

200. J.-H. Kim, H. Yang, J. Park, G.-J. Boons, *J. Am. Chem. Soc.* **2005**, 127, 12090.

201. T. Fang, K.-F. Mo, G.-J. Boons, *J. Am. Chem. Soc.* **2012**, 134, 7545.

202. J. P. Yasomanee, A. V. Demchenko, *J. Am. Chem. Soc.* **2012**, 134, 20097.

203. J. P. Yasomanee, A. V. Demchenko, *Angew. Chem. Int. Ed. Engl.* **2014**, 53, 10453.

204. J. P. Yasomanee, A. V. Demchenko, *Chem. Eur. J.* **2015**, 21, 6572.

205. P. v. d. Steen, P. M. Rudd, R. A. Dwek, G. Opdenakker, *Crit. Rev. Biochem. Mol. Biol.* **1998**, 33, 151.

206. G. Lamblin, S. Degroote, J.-M. Perini, P. Delmotte, A. Scharfman, M. Davril, J.-M. Lo-Guidice, N. Houdret, V. Dumur, A. Klein, P. Rousse, *Glycoconj. J.* **2001**, 18, 661-684.

207. B. Liebe, H. Kunz, *Tetrahedron Lett.* **1994**, 35, 8777.

208. B. Liebe, H. Kunz, *Angew. Chem. Int. Ed.* **1997**, 36, 618.

209. H. Kunz, C. Unverzagt, *Angew. Chem. Int. Ed.* **1988**, 27, 1697.

210. C. Unverzagt, H. Kunz, *Bioorg. Med. Chem.* **1994**, 2, 1189.

211. G. Zemplén, A. Kunz, *Ber. dtsch. Chem. Ges. A/B* **1923**, 56, 1705.

212. M. Schelhaas, H. Waldmann, *Angew. Chem. Int. Ed.* **1996**, 35, 2056.

213. G.-J. Boons, K. Hale. *Organic synthesis with carbohydrates* (Sheffield Academic Press; Blackwell Science, Sheffield, England, Malden, MA U.S.A. 2000).

214. U. Ellervik, G. Magnusson, *Carbohydrate Res.* **1996**, 280, 251.

215. S. Dziadek, C. Brocke, H. Kunz, *Chem. Eur. J.* **2004**, 10, 4150.

216. M. R. Pratt, C. R. Bertozzi, *Org. Lett.* **2004**, 6, 2345.

217. H. Paulsen, M. Paal, D. Hadamczyk, K.-M. Steiger, *Carbohydrate Res.* **1984**, 131, C1.

218. U. Ellervik, H. Grundberg, G. Magnusson, *J. Org. Chem.* **1998**, 63, 9323.

219. U. Ellervik, G. Magnusson, *J. Org. Chem.* **1998**, 63, 9314.

220. C. Brocke, H. Kunz, *Synthesis* **2004**, , 525.

221. H. Paulsen, J.-P. Hölck, *Carbohydrate Res.* **1982**, 109, 89.

222. S. Dziadek, D. Kowalczyk, H. Kunz, *Angew. Chem. Int. Ed.* **2005**, 44, 7624.

223. T. Reipen, H. Kunz, *Synthesis* **2003**, , 2487.

224. E. Meinjohanns, M. Meldal, A. Schleyer, H. Paulsen, K. Bock, *J. Chem. Soc. Perkin Trans. 1* **1996**, , 985.

225. P. W. Glunz, S. Hintermann, L. J. Williams, J. B. Schwarz, S. D. Kuduk, V. Kudryashov, K. O. Lloyd, S. J. Danishefsky, *J. Am. Chem. Soc.* **2000**, 122, 7273.

226. Y. Takano, M. Habiro, M. Someya, H. Hojo, Y. Nakahara, *Tetrahedron Lett.* **2002**, 43, 8395.

227. Y. Nakahara, C. Ozawa, E. Tanaka, K. Ohtsuka, Y. Takano, H. Hojo, Y. Nakahara, *Tetrahedron* **2007**, 63, 2161.

228. K. Baumann, D. Kowalczyk, T. Gutjahr, M. Pieczyk, C. Jones, M. K. Wild, D. Vestweber, H. Kunz, *Angew. Chem. Int. Ed.* **2009**, 48, 3174.

229. T. Matsushita, H. Hinou, M. Fumoto, M. Kurogochi, N. Fujitani, H. Shimizu, S.-I. Nishimura, *J. Org. Chem.* **2006**, 71, 3051.

230. T. Matsushita, R. Sadamoto, N. Ohyabu, H. Nakata, M. Fumoto, N. Fujitani, Y. Takegawa, T. Sakamoto, M. Kurogochi, H. Hinou, H. Shimizu, T. Ito, K. Naruchi, H. Togame, H. Takemoto, H. Kondo, S.-I. Nishimura, *Biochemistry* **2009**, 48, 11117.

231. R. Kaifu, T. Osawa, *Carbohydrate Res.* **1977**, 58, 235.

232. B. Ferrari, A. A. Pavia, *Carbohydrate Res.* **1980**, 79, C1.

233. W. Koenigs, E. Knorr, *Ber. Dtsch. Chem. Ges.* **1901**, 34, 957.

234. J. Broddefalk, U. Nilsson, J. Kihlberg, *J. Carb. Chem.* **1994**, 13, 129.

235. A. P. Kozikowski, J. Lee, *J. Org. Chem.* **1990**, 55, 863.

236. R. U. Lemieux, R. M. Ratcliffe, *Can. J. Chem.* **1979**, 57, 1244.

237. A. Paquet, *Can. J. Chem.* **1982**, 60, 976.

238. E. Schmidt, F. Moosmüller, *Justus Liebigs Ann. Chem.* **1955**, 597, 235.

239. E. Däbritz, *Angew. Chem.* **1966**, 78, 483.

240. S. Dziadek, *Dissertation, Johannes Gutenberg Univerität Mainz* **2005**,

241. B. Liebe, H. Kunz, *Helv. Chim. Acta* **1997**, 80, 1473.

242. T. Rosen, I. M. Lico, Chu, Daniel T. W. *J. Org. Chem.* **1988**, 53, 1580.

243. R. V. Kolakowski, N. Shangguan, R. R. Sauers, L. J. Williams, *J. Am. Chem. Soc.* **2006**, 128, 5695.

244. B. Helferich, K.-F. Wedemeyer, *Justus Liebigs Ann. Chem.* **1949**, 563, 139.

245. B. Helferich, K. Weis, *Chem. Ber.* **1956**, 89, 314.

246. R. R. Schmidt, M. Stumpp, *Liebigs Ann. Chem.* **1983**, 1983, 1249.

247. R. R. Schmidt, J. Michel, *Tetrahedron Lett.* **1984**, 25, 821.

248. E. D. Goddard-Borger, R. V. Stick, *Org. Lett.* **2007**, 9, 3797.

249. E. D. Goddard-Borger, R. V. Stick, *Org. Lett.* **2011**, 13, 2514.

250. N. Fischer, E. D. Goddard-Borger, R. Greiner, T. M. Klapötke, B. W. Skelton, J. Stierstorfer, *J. Org. Chem.* **2012**, 77, 1760.

251. H. Paulsen, A. Richter, V. Sinnwell, W. Stenzel, *Carbohydrate Res.* **1978**, 64, 339.

252. De Silva, Ravindra A. Q. Wang, T. Chidley, D. K. Appulage, P. R. Andreana, *J. Am. Chem. Soc.* **2009**, 131, 9622.

253. B. Lüning, T. Norberg, J. Tejbrant, *Glycoconj. J.* **1989**, 6, 5.

254. M. Vasan, M. A. Wolfert, G.-J. Boons, *Org. Biomol. Chem.* **2007**, 5, 2087.

255. A. Toepfer, R. R. Schmidt, *J. Carb. Chem.* **1993**, 12, 809.

256. D. Cato, T. Buskas, G. Boons, *J. Carb. Chem.* **2005**, 24, 503.

257. M. Wilstermann, G. Magnusson, *Carbohydrate Res.* **1995**, 272, 1.

258. G. Wulff, W. Schmidt, *Carbohydrate Res.* **1977**, 53, 33.

259. Z. Yang, W. Lin, B. Yu, *Carbohydrate Res.* **2000**, 329, 879.

260. H. H. Jensen, L. U. Nordstrøm, M. Bols, *J. Am. Chem. Soc.* **2004**, 126, 9205.

261. E. J. Corey, A. Venkateswarlu, *J. Am. Chem. Soc.* **1972**, 94, 6190.

262. F. Yan, S. Mehta, E. Eichler, W. W. Wakarchuk, M. Gilbert, M. J. Schur, D. M. Whitfield, *J. Org. Chem.* **2003**, 68, 2426.

263. U. Ellervik, G. Magnusson, *Tetrahedron Lett.* **1997**, 38, 1627.

264. T. B. Windholz, Johnston, David B. R. *Tetrahedron Lett.* **1967**, 8, 2555.

265. J. D. Codée, van den Bos, Leendert J, R. E. Litjens, H. S. Overkleeft, van Boeckel, Constant A.A, van Boom, Jacques H, van der Marel, Gijs A, *Tetrahedron* **2004**, 60, 1057.

266. S. Nambiar, J. F. Daeuble, R. J. Doyle, K. Grant Taylor, *Tetrahedron Lett.* **1989**, 30, 2179.

267. J. H. Clark, *Chem. Rev.* **1980**, 80, 429.

268. M. Ueki, M. Amemiya, *Tetrahedron Lett.* **1987**, 28, 6617.

269. G. Höfle, W. Steglich, H. Vorbrüggen, *Angew. Chem. Int. Ed.* **1978**, 17, 569.

270. A. H. Haines. in *Advances in Carbohydrate Chemistry and Biochemistry,* edited by R. Stuart Tipson and Derek Horton (Academic Press1976), pp. 11–109.

271. N. Gaidzik, U. Westerlind, H. Kunz, *Chem. Soc. Rev.* **2013**, 42, 4421.

272. O. C. Grant, Smith, Hannah M K, D. Firsova, E. Fadda, R. J. Woods, *Glycobiology* **2014**, 24, 17.

273. D. M. Lewallen, D. Siler, S. S. Iyer, *ChemBioChem* **2009**, 10, 1486.

274. S. Park, M.-R. Lee, S.-J. Pyo, I. Shin, *J. Am. Chem. Soc.* **2004**, 126, 4812.

275. M. Gewehr, H. Kunz, *Synthesis* **1997**, 1997, 1499.

276. S. Keil, C. Claus, W. Dippold, H. Kunz, *Angew. Chem. Int. Ed.* **2001**, 40, 366.

277. C. A. Bush, L. Deng, X. Chen, A. Varki, *Biopolymers* **2013**, 99, 650.

278. A. Varki, *Trends in Molecular Medicine* **2008**, 14, 351.

279. G.-J. Boons, A. V. Demchenko, *Chem. Rev.* **2000**, 100, 4539.

280. X. Chen, A. Varki, *ACS Chem. Biol.* **2010**, 5, 163.

281. H. Yu, H. A. Chokhawala, S. Huang, X. Chen, *Nat. Protoc.* **2006**, 1, 2485.

282. Y. C. Lee, N. Kojima, E. Wada, N. Kurosawa, T. Nakaoka, T. Hamamoto, S. Tsuji, *J. Biol. Chem.* **1994**, 269, 10028.

283. H. Yu, H. Chokhawala, R. Karpel, H. Yu, B. Wu, J. Zhang, Y. Zhang, Q. Jia, X. Chen, *J. Am. Chem. Soc.* **2005**, 127, 17618.

284. V. Thon, Y. Li, H. Yu, K. Lau, X. Chen, *Appl Microbiol Biotechnol* **2012**, 94, 977-985.

285. H. Yu, S. Huang, H. Chokhawala, M. Sun, H. Zheng, X. Chen, *Angew. Chem. Int. Ed.* **2006**, 45, 3938.

286. C. M. Nycholat, W. Peng, R. McBride, A. Antonopoulos, de Vries, Robert P, Z. Polonskaya, M. G. Finn, A. Dell, S. M. Haslam, J. C. Paulson, *J. Am. Chem. Soc.* **2013**, 135, 18280.

287. W.-T. Chien, C.-F. Liang, C.-C. Yu, C.-H. Lin, S.-P. Li, I. Primadona, Y.-J. Chen, Mong, Kwok Kong T, C.-C. Lin, *Chem. Commun.* **2014**, 50, 5786.

288. H. Malekan, G. Fung, V. Thon, Z. Khedri, H. Yu, J. Qu, Y. Li, L. Ding, K. S. Lam, X. Chen, *Carbohydrate-Processing Enzymes* **2013**, 21, 4778.

289. S. A. Cohen, D. P. Michaud, *Anal. Biochem.* **1993**, 211, 279.

290. G. Zauner, R. P. Kozak, R. A. Gardner, D. L. Fernandes, A. M. Deelder, M. Wuhrer, *Biol. Chem.* **2012**, 393, 687.

291. Y. Zhang, B. R. Fonslow, B. Shan, M.-C. Baek, J. R. Yates, *Chem. Rev.* **2013**, 113, 2343.

292. W. R. Alley, B. F. Mann, M. V. Novotny, *Chem. Rev.* **2013**, 113, 2668.

293. H. Li, B. Bendiak, W. F. Siems, D. R. Gang, H. H. Hill, *Anal. Chem.* **2013**, 85, 2760.

294. P. Both, A. P. Green, C. J. Gray, R. Sardzík, J. Voglmeir, C. Fontana, M. Austeri, M. Rejzek, D. Richardson, R. A. Field, G. Widmalm, S. L. Flitsch, C. E. Eyers, *Nat. Chem.* **2014**, 6, 65.

295. A. J. Creese, H. J. Cooper, *Anal. Chem.* **2012**, 84, 2597.

296. H. Li, K. Giles, B. Bendiak, K. Kaplan, W. F. Siems, H. H. Hill, *Anal. Chem.* **2012**, 84, 3231.

297. J. Nilsson, A. Halim, A. Grahn, G. Larson, *Glycoconj. J.* **2013**, 30, 119.

298. A. Halim, U. Westerlind, C. Pett, M. Schorlemer, U. Rüetschi, G. Brinkmalm, C. Sihlbom, J. Lengqvist, G. Larson, J. Nilsson, *J. Proteome Res.* **2014**, 13, 6024.

299. F. Marino, M. Bern, Mommen, Geert P M, A. C. Leney, van Gaans-van den Brink, Jacqueline A M, Bonvin, Alexandre M J J, C. Becker, van Els, Cécile A C M, Heck, Albert J R, *J. Am. Chem. Soc.* **2015**, 137, 10922.

300. Y. Ohashi, T. Ii, M. Kubota, S. Nunomura, H. Niwa, M. Ohashi, T. Ogawa, Y. Nagai, *J. Mass Spectrom. Soc. Jpn.* **1998**, 46, 45.

301. Y. Ohashi, *J. Mass Spectrom. Soc. Jpn.* **2007**, 55, 311.

302. J. C. Trinidad, R. Schoepfer, A. L. Burlingame, K. F. Medzihradszky, *Mol. Cell Proteomics* **2013**, 12, 3474.

303. M. Schena, D. Shalon, R. W. Davis, P. O. Brown, *Science* **1995**, 270, 467.

304. M. F. Templin, D. Stoll, M. Schrenk, P. C. Traub, C. F. Vöhringer, T. O. Joos, *Trends in Biotechnology* **2002**, 20, 160.

305. G. MacBeath, S. L. Schreiber, *Science* **2000**, 289, 1760.

306. D. Weinrich, P. Jonkheijm, C. M. Niemeyer, H. Waldmann, *Angew. Chem. Int. Ed.* **2009**, 48, 7744.

307. M. Mammen, S.-K. Choi, G. M. Whitesides, *Angew. Chem. Int. Ed.* **1998**, 37, 2754.

308. J. Stevens, O. Blixt, T. M. Tumpey, J. K. Taubenberger, J. C. Paulson, I. A. Wilson, *Science* **2006**, 312, 404.

309. J. Stevens, O. Blixt, L. Glaser, J. K. Taubenberger, P. Palese, J. C. Paulson, I. A. Wilson, *J. Mol. Biol.* **2006**, 355, 1143.

310. U. Westerlind, H. Schröder, A. Hobel, N. Gaidzik, A. Kaiser, C. M. Niemeyer, E. Schmitt, H. Waldmann, H. Kunz, *Ang. Chem. Int. Ed.* **2009**, 48, 8263.

311. O. Blixt, U. Westerlind, *Curr. Opin. Chem. Biol.* **2014**, 18, 62.

312. L. S. Wong, F. Khan, J. Micklefield, *Chem. Rev.* **2009**, 109, 4025.

313. T. Govindaraju, P. Jonkheijm, L. Gogolin, H. Schroeder, Becker, Christian F W, C. M. Niemeyer, H. Waldmann, *Chem. Commun.* **2008**, , 3723.

314. L. Yi, Y.-X. Chen, P.-C. Lin, H. Schröder, C. M. Niemeyer, Y.-W. Wu, R. S. Goody, G. Triola, H. Waldmann, *Chem. Commun.* **2012**, 48, 10829.

315. D. Weinrich, P.-C. Lin, P. Jonkheijm, Nguyen, Uyen T T, H. Schröder, C. M. Niemeyer, K. Alexandrov, R. Goody, H. Waldmann, *Angew. Chem. Int. Ed.* **2010**, 49, 1252.

316. O. Blixt, S. Head, T. Mondala, C. Scanlan, M. E. Huflejt, R. Alvarez, M. C. Bryan, F. Fazio, D. Calarese, J. Stevens, N. Razi, D. J. Stevens, J. J. Skehel, I. van Die, D. R. Burton, I. A. Wilson, R. Cummings, N. Bovin, C.-H. Wong, J. C. Paulson, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, 101, 17033.

317. D. H. Dube, C. R. Bertozzi, *Nat. Rev. Drug Discov.* **2005**, 4, 477.

318. J. Burchell, R. Poulsom, A. Hanby, C. Whitehouse, L. Cooper, H. Clausen, D. Miles, J. Taylor-Papadimitriou, *Glycobiology* **1999**, 9, 1307.

319. M. N. Christiansen, J. Chik, L. Lee, M. Anugraham, J. L. Abrahams, N. H. Packer, *Proteomics* **2013**,

320. J. Burchell, A. Mungul, J. Taylor-Papadimitriou, *J. Mammary Gland. Biol. Neoplasia* **2001**, 6, 355-364.

321. K. O. Lloyd, J. Burchell, V. Kudryashov, Yin, B. W. T. J. Taylor-Papadimitriou, *J. Biol. Chem.* **1996**, 271, 33325.

322. R. M. Wilson, S. J. Danishefsky, *J. Am. Chem. Soc.* **2013**, 135, 14462.

323. F. Garcia-Martin, T. Matsushita, H. Hinou, S.-I. Nishimura, *Chem. Eur. J.* **2014**, 20, 15891.

324. T. Matsushita, W. Takada, K. Igarashi, K. Naruchi, R. Miyoshi, F. Garcia-Martin, M. Amano, H. Hinou, S.-I. Nishimura, *Biochim. Biophys. Acta General Subjects* **2014**, 1840, 1105.

325. U. Westerlind, A. Hobel, N. Gaidzik, E. Schmitt, H. Kunz, *Angew. Chem. Int. Ed.* **2008**, 47, 7551.

326. H. Cai, Z.-Y. Sun, Z.-H. Huang, L. Shi, Y.-F. Zhao, H. Kunz, Y.-M. Li, *Chem. Eur. J.* **2013**, 19, 1962.

327. U. Karsten, C. Diotel, G. Klich, H. Paulsen, S. Goletz, S. Müller, F.-G. Hanisch, *Cancer Res.* **1998**, 58, 2541.

328. R. Schauer, *Curr. Opin. Struct. Biol.* **2009**, 19, 507.

329. D. Baeckström, O. Nilsson, M. R. Price, L. Lindholm, G. C. Hansson, *Cancer Res.* **1993**, 53, 755.

330. U. Karsten, N. Serttas, H. Paulsen, A. Danielczyk, S. Goletz, *Glycobiology* **2004**, 14, 681.

331. L. Kinarsky, G. Suryanarayanan, O. Prakash, H. Paulsen, H. Clausen, F.-G. Hanisch, M. A. Hollingsworth, S. Sherman, *Glycobiology* **2003**, 13, 929.

332. A. Borgert, J. Heimburg-Molinaro, X. Song, Y. Lasanajak, T. Ju, M. Liu, P. Thompson, G. Ragupathi, G. Barany, D. F. Smith, R. D. Cummings, D. Live, *ACS Chem. Biol.* **2012**, 7, 1031.

333. F. Corzana, J. H. Busto, García de Luis, Marisa, J. Jiménez-Barbero, A. Avenoza, J. M. Peregrina, *Chem. Eur. J.* **2009**, 15, 3863.

334. P. Braun, G. M. Davies, M. R. Price, P. M. Williams, Tendler, Saul J. B, H. Kunz, *Bioorg. Med. Chem.* **1998**, 6, 1531.

335. Stadie, Tanja R. E. W. Chai, A. M. Lawson, Byfield, Peter G. H. F.-G. Hanisch, *Eur. J. Biochem.* **1995**, 229, 140.

336. D. M. Coltart, A. K. Royyuru, L. J. Williams, P. W. Glunz, D. Sames, S. D. Kuduk, J. B. Schwarz, X.-T. Chen, S. J. Danishefsky, D. H. Live, *J. Am. Chem. Soc.* **2002**, 124, 9833.

337. R. Hashimoto, N. Fujitani, Y. Takegawa, M. Kurogochi, T. Matsushita, K. Naruchi, N. Ohyabu, H. Hinou, X. D. Gao, N. Manri, H. Satake, A. Kaneko, T. Sakamoto, S.-I. Nishimura, *Chem. Eur. J.* **2011**, 17, 2393.

338. A. Kuhn, H. Kunz, *Angew. Chem. Int. Ed.* **2007**, 46, 454.

339. K. Engelmann, C. L. Kinlough, S. Müller, H. Razawi, S. E. Baldus, R. P. Hughey, F.-G. Hanisch, *Glycobiology* **2005**, 15, 1111.

340. N. Ohyabu, H. Hinou, T. Matsushita, R. Izumi, H. Shimizu, K. Kawamoto, Y. Numata, H. Togame, H. Takemoto, H. Kondo, S.-I. Nishimura, *J. Am. Chem. Soc.* **2009**, 131, 17102.

341. X. Dan, W. Liu, T. B. Ng, *Med. Res. Rev.* **2015**, , 1.

342. J. Hirabayashi, M. Yamada, A. Kuno, H. Tateno, *Chem. Soc. Rev.* **2013**, 42, 4443.

343. D. Madariaga, N. Martínez-Sáez, V. J. Somovilla, L. García-García, M. Á. Berbis, J. Valero-Gónzalez, S. Martín-Santamaría, R. Hurtado-Guerrero, J. L. Asensio, J. Jiménez-Barbero, A. Avenoza, J. H. Busto, F. Corzana, J. M. Peregrina, *Chem. Eur. J.* **2014**, 20, 12616.

344. S. Di Lella, V. Sundblad, J. P. Cerliani, C. M. Guardia, D. A. Estrin, G. R. Vasta, G. A. Rabinovich, *Biochemistry* **2011**, 50, 7842.

345. D. Giguère, S. André, M.-A. Bonin, M.-A. Bellefleur, A. Provencal, P. Cloutier, B. Pucci, R. Roy, H.-J. Gabius, *Bioorg. Med. Chem.* **2011**, 19, 3280.

346. S. André, D. Giguère, T. K. Dam, F. Brewer, H.-J. Gabius, R. Roy, *New J. Chem.* **2010**, 34, 2229.

347. J. Seetharaman, A. Kanigsberg, R. Slaaby, H. Leffler, S. H. Barondes, J. M. Rini, *J. Biol. Chem.* **1998**, 273, 13047.

348. C.-F. Bian, Y. Zhang, H. Sun, D.-F. Li, D.-C. Wang, *PLoS ONE* **2011**, 6, e25007.

349. H. Leffler, S. Carlsson, M. Hedlund, Y. Qian, F. Poirier, *Glycoconj. J.* **2004**, 19, 433-440.

350. P. Sörme, B. Kahl-Knutsson, M. Huflejt, U. J. Nilsson, H. Leffler, *Anal. Biochem.* **2004**, 334, 36.

351. C. F. Brewer, *Glycoconj. J.* **2002**, 19, 459-465.

352. S. K. Patnaik, B. Potvin, S. Carlsson, D. Sturm, H. Leffler, P. Stanley, *Glycobiology* **2006**, 16, 305.

353. S. R. Stowell, C. M. Arthur, P. Mehta, K. A. Slanina, O. Blixt, H. Leffler, D. F. Smith, R. D. Cummings, *J. Biol. Chem.* **2008**, 283, 10109.

354. X. Song, B. Xia, S. R. Stowell, Y. Lasanajak, D. F. Smith, R. D. Cummings, *Chem. Biol.* **2009**, 16, 36.

355. H. Tateno, A. Mori, N. Uchiyama, R. Yabe, J. Iwaki, T. Shikanai, T. Angata, H. Narimatsu, J. Hirabayashi, *Glycobiology* **2008**, 18, 789.

356. T. Horlacher, M. A. Oberli, D. B. Werz, L. Kröck, S. Bufali, R. Mishra, J. Sobek, K. Simons, M. Hirashima, T. Niki, P. H. Seeberger, *ChemBioChem* **2010**, 11, 1563.

357. S. R. Stowell, M. Dias-Baruffi, L. Penttilä, O. Renkonen, A. K. Nyame, R. D. Cummings, *Glycobiology* **2004**, 14, 157.

358. van den Berg, T. K. H. Honing, N. Franke, A. van Remoortere, Schiphorst, W. E. C. M. F.-T. Liu, A. M. Deelder, R. D. Cummings, C. H. Hokke, I. van Die, *J. Immunol.* **2004**, 173, 1902.

359. A. Leppänen, S. Stowell, O. Blixt, R. D. Cummings, *J. Biol. Chem.* **2005**, 280, 5549.

360. E. M. Rapoport, N. V. Bovin, *Biochemistry Moscow* **2015**, 80, 846-856.

361. http://www.functionalglycomics.org/.

362. C. T. Oberg, S. Carlsson, E. Fillion, H. Leffler, U. J. Nilsson, *Bioconjug. Chem.* **2003**, 14, 1289.

363. M. Krzeminski, T. Singh, S. André, M. Lensch, A. M. Wu, Bonvin, Alexandre M. J. J. H.-J. Gabius, *Biochim. Biophys. Acta General Subjects* **2011**, 1810, 150.

364. J.-T. de Oliveira, A.-J. de Matos, A. L. Santos, R. Pinto, J. Gomes, V. Hespanhol, R. Chammas, A. Manninen, E. S. Bernardes, C. Albuquerque Reis, G. Rutteman, F. Gärtner, *Int. J. Dev. Biol.* **2011**, 55, 823.

365. T.-W. Lin, H.-T. Chang, C.-H. Chen, C.-H. Chen, S.-W. Lin, T.-L. Hsu, C.-H. Wong, *J. Am. Chem. Soc.* **2015**, 137, 9685.

366. A. F. Swindall, A. I. Londoño-Joshi, M. J. Schultz, N. Fineberg, D. J. Buchsbaum, S. L. Bellis, *Cancer Res.* **2013**, 73, 2368.

367. M.-A. Recchi, A. Harduin-Lepers, Y. Boilly-Marer, A. Verbert, P. Delannoy, *Glycoconj. J.* **1998**, 15, 19-27.

368. F. D. Olio, N. Malagolini, G. Di Stefano, F. Minni, D. Marrano, F. Serafini-Cessi, *Int. J. Cancer* **1989**, 44, 434.

369. P.-H. Wang, W.-L. Lee, C.-M. Juang, Y.-H. Yang, W.-H. Lo, C.-R. Lai, S.-L. Hsieh, C.-C. Yuan, *Gynecologic Oncology* **2005**, 99, 631.

370. F.-T. Liu, G. A. Rabinovich, *Nat. Rev. Cancer* **2005**, 5, 29.

371. Y. Zhuo, S. L. Bellis, *J. Biol. Chem.* **2011**, 286, 5935.

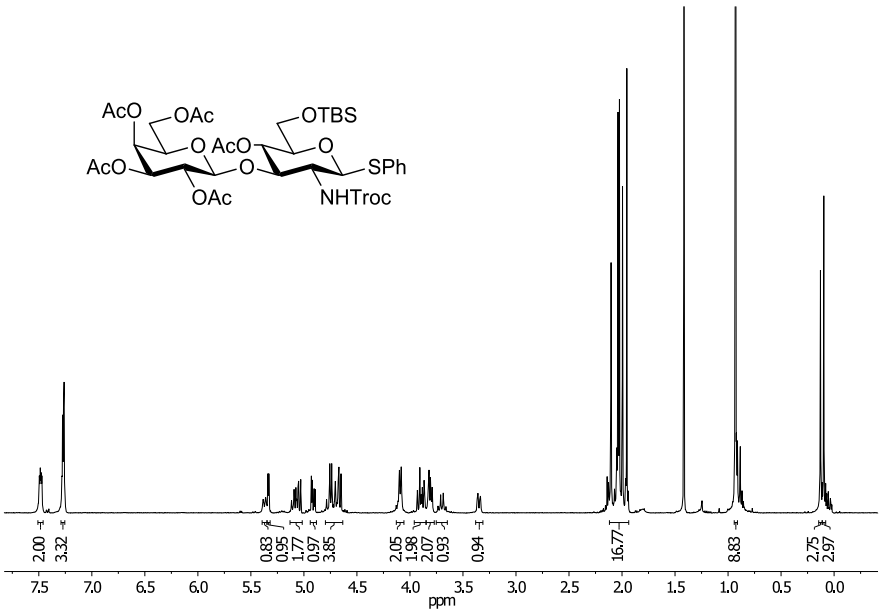
372. F. M. Shaikh, E. C. Seales, W. C. Clem, K. M. Hennessy, Y. Zhuo, S. L. Bellis, *Exp. Cell Res.* **2008**, 314, 2941.

373. H. E. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem.* **1997**, 62, 7512.

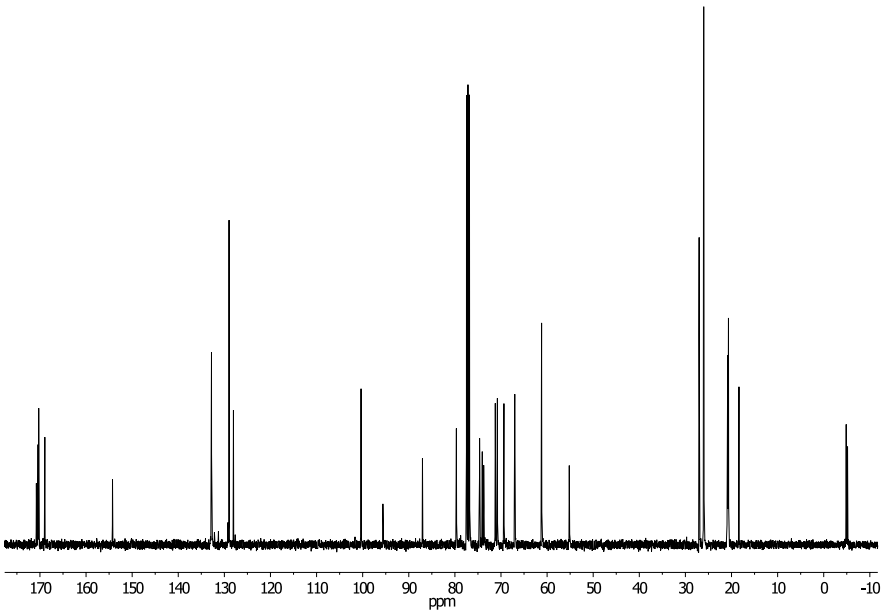
9 APPENDIX

9.1 Spectroscopic data

Compound **27**

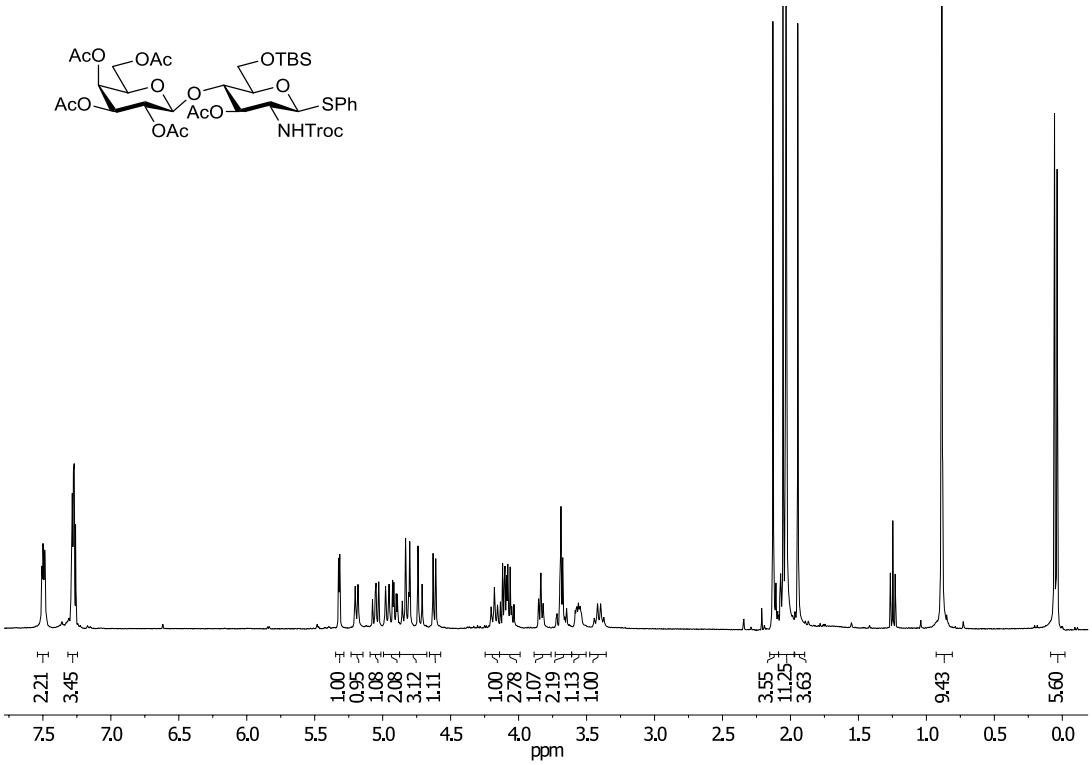


^{1}H NMR (400 MHz, $CDCl_3$)

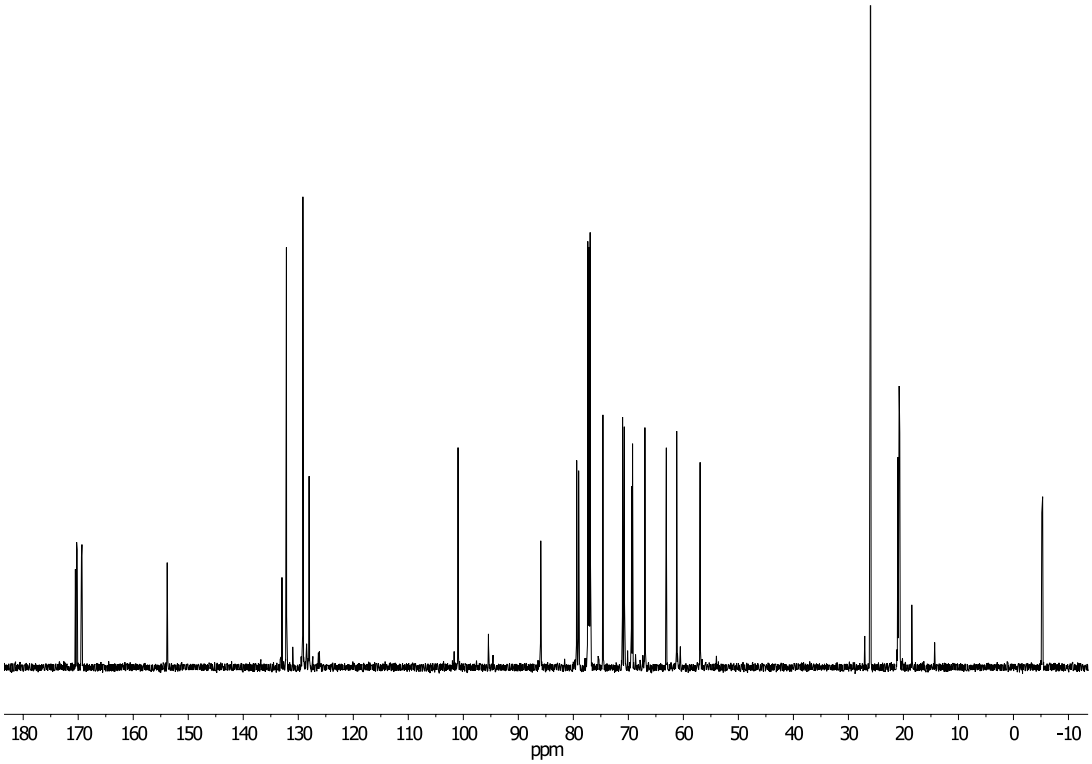


^{13}C NMR (100.6 MHz, $CDCl_3$)

Compound **31**

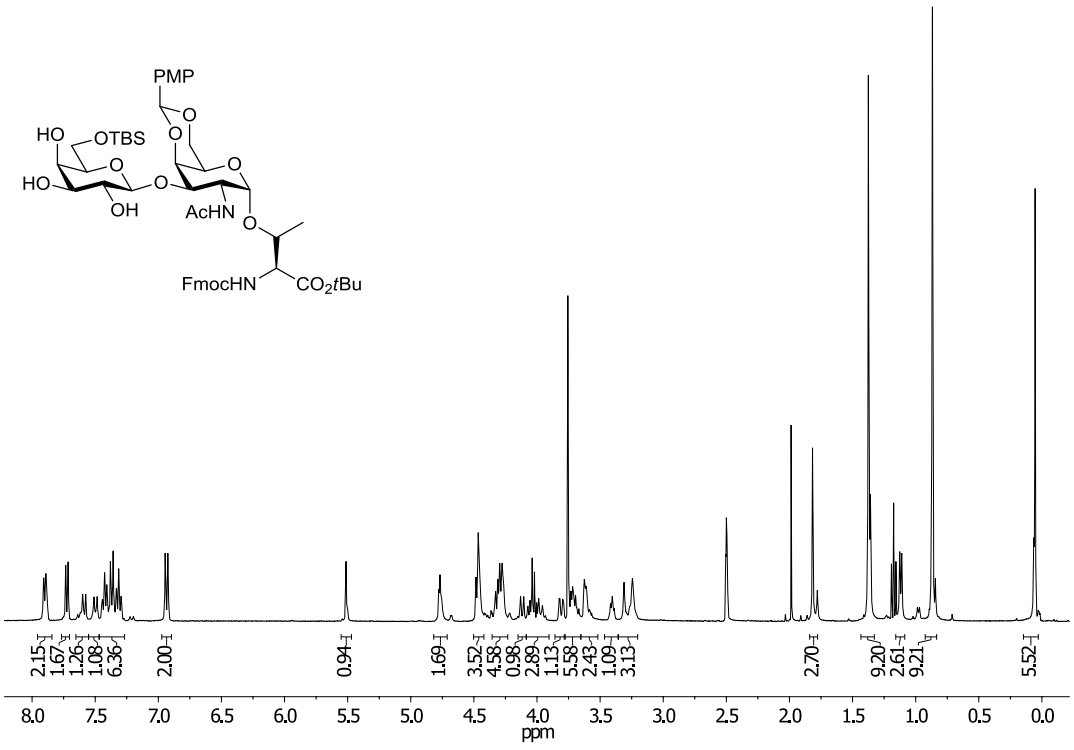


[1]H NMR (600 MHz, DMSO-d_6)

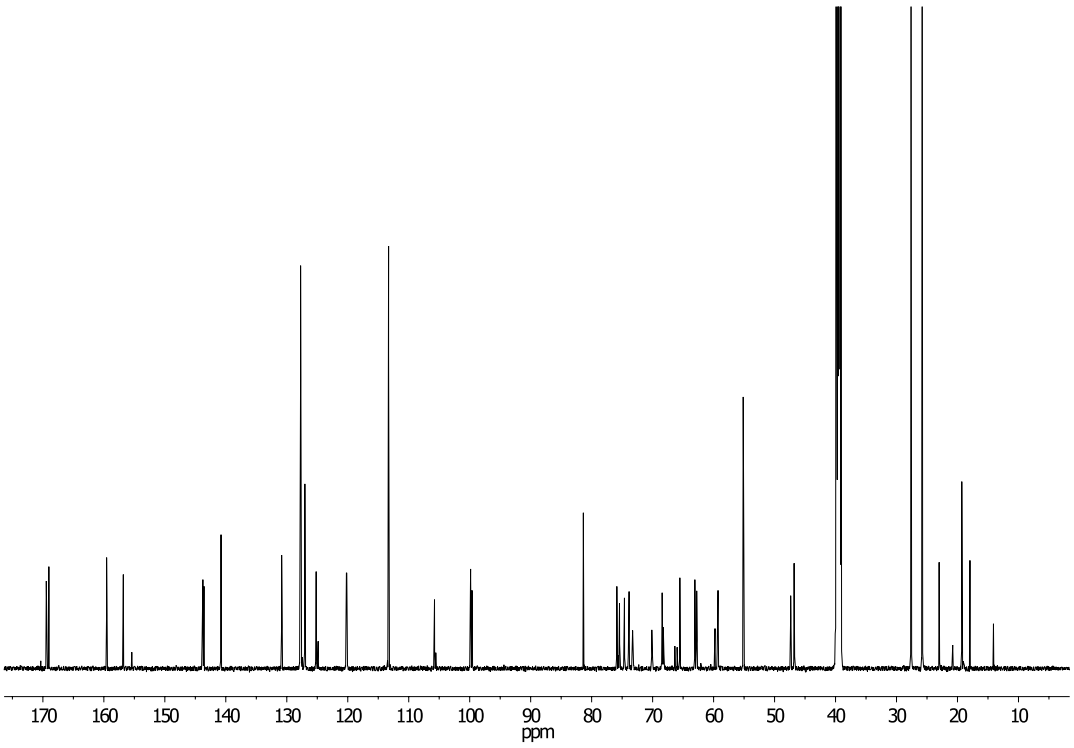


[13]C NMR (150.9 MHz, $CDCl_3$)

Compound **44**

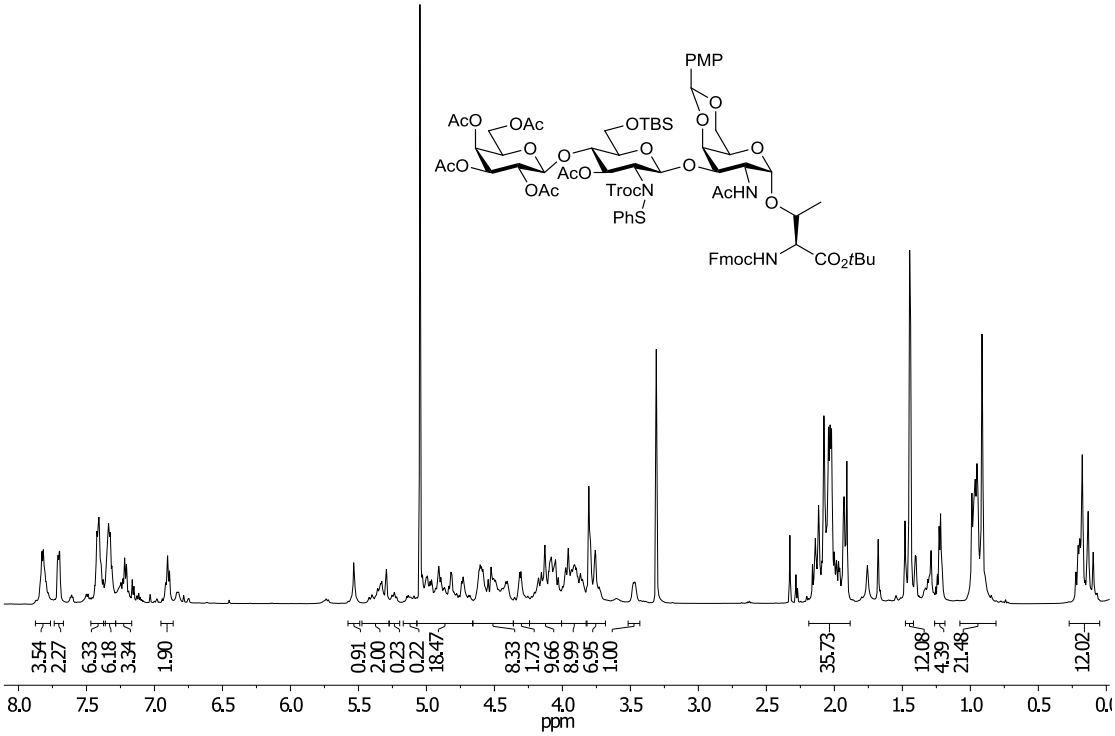


^{1}H NMR (400 MHz, DMSO-d_6)

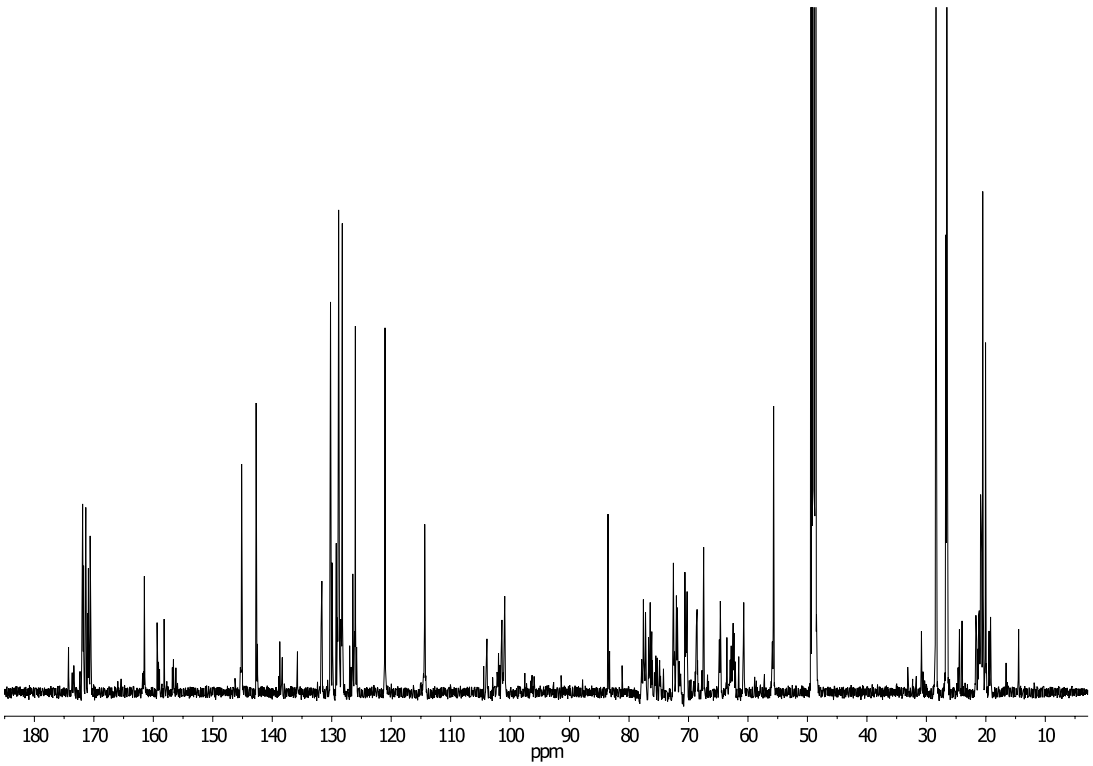


^{13}C NMR (100.6 MHz, DMSO-d_6)

Compound **32**

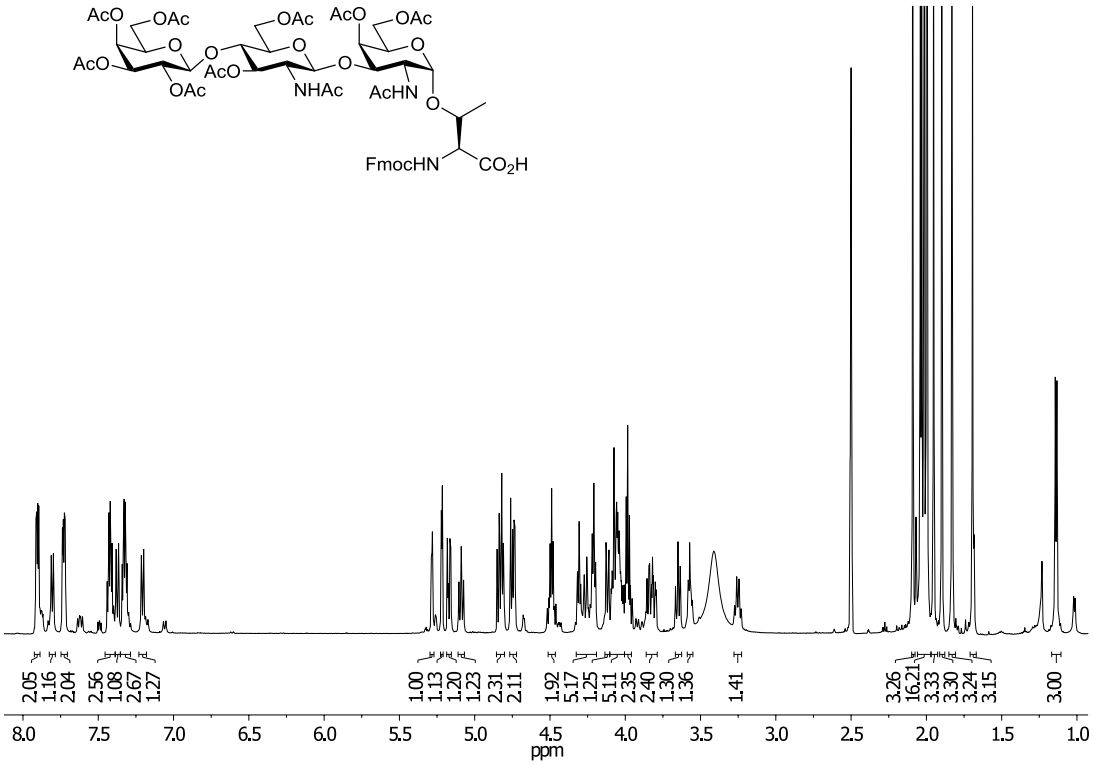


^{1}H *NMR* (600 MHz, MeOD-d_4)

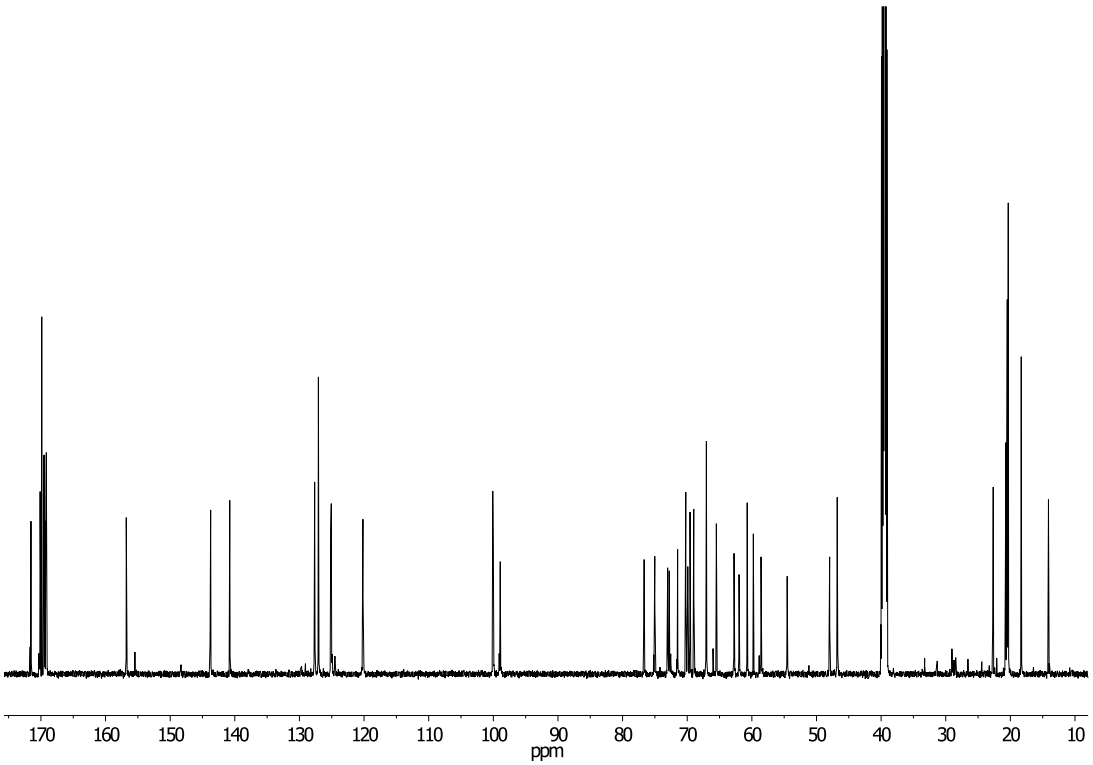


^{13}C NMR (150.9 MHz, MeOD-d_4)

Compound **35**

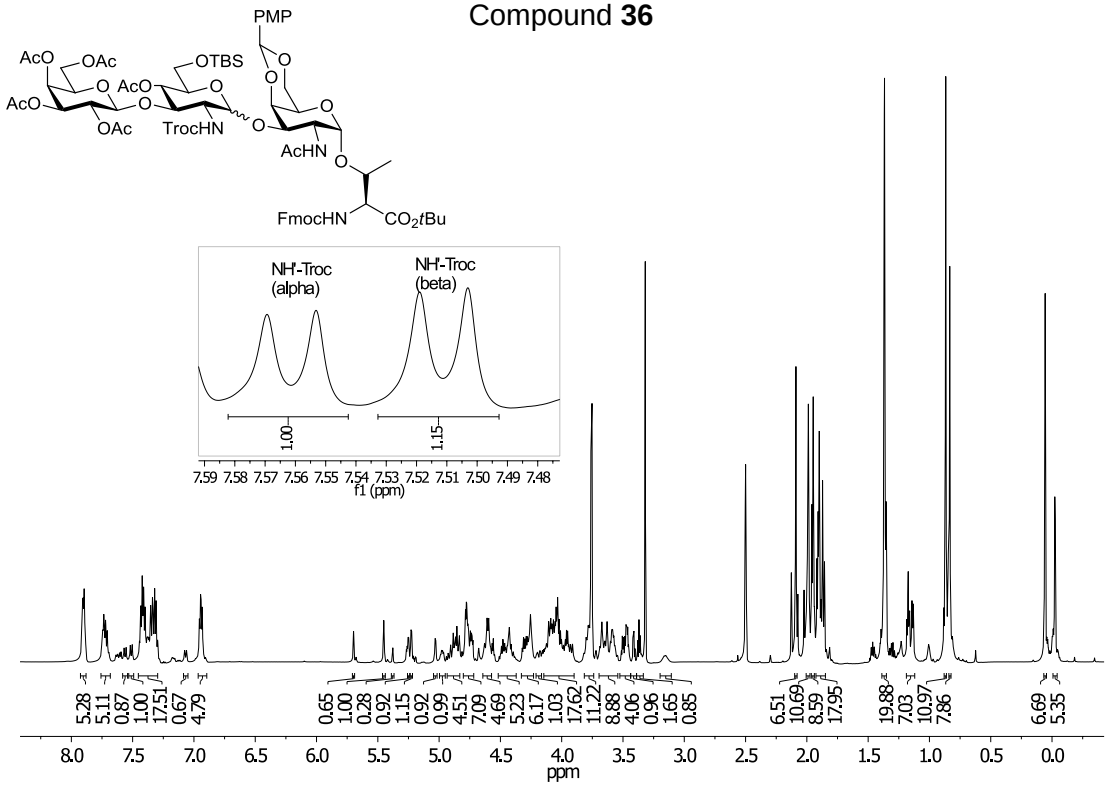


[1]H NMR (600 MHz, MeOD-d_4)

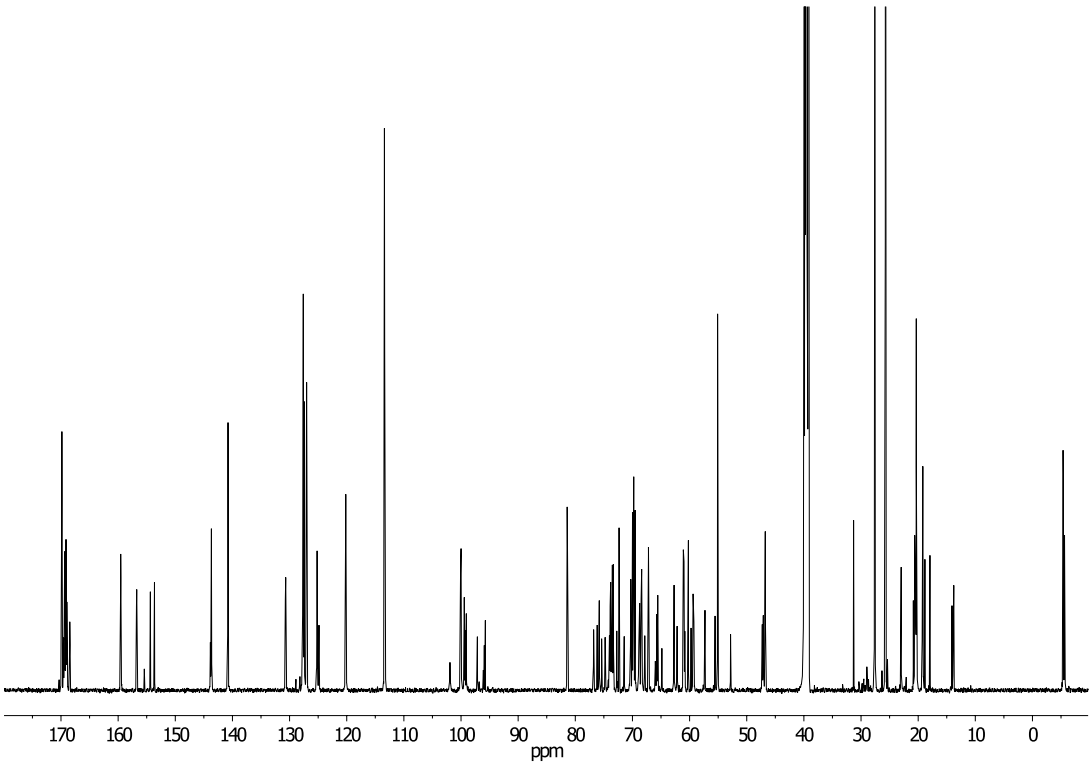


[13]C NMR (150.9 MHz, DMSO-d_6)

Compound **36**

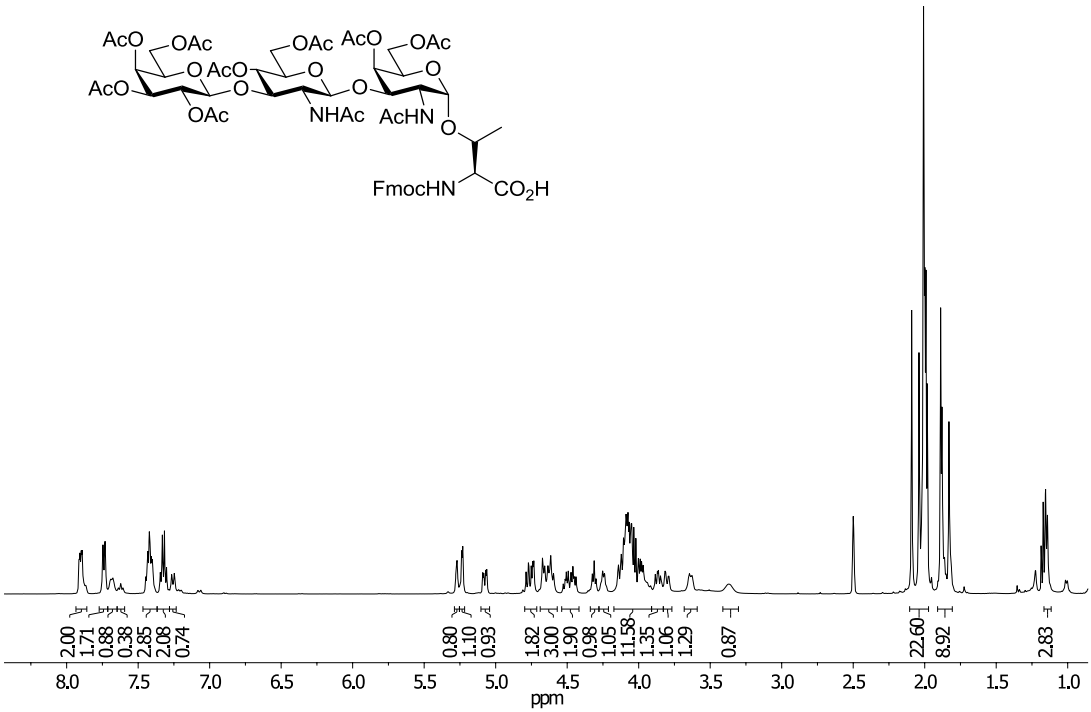


^{1}H *NMR* (600 MHz, DMSO-d_6)

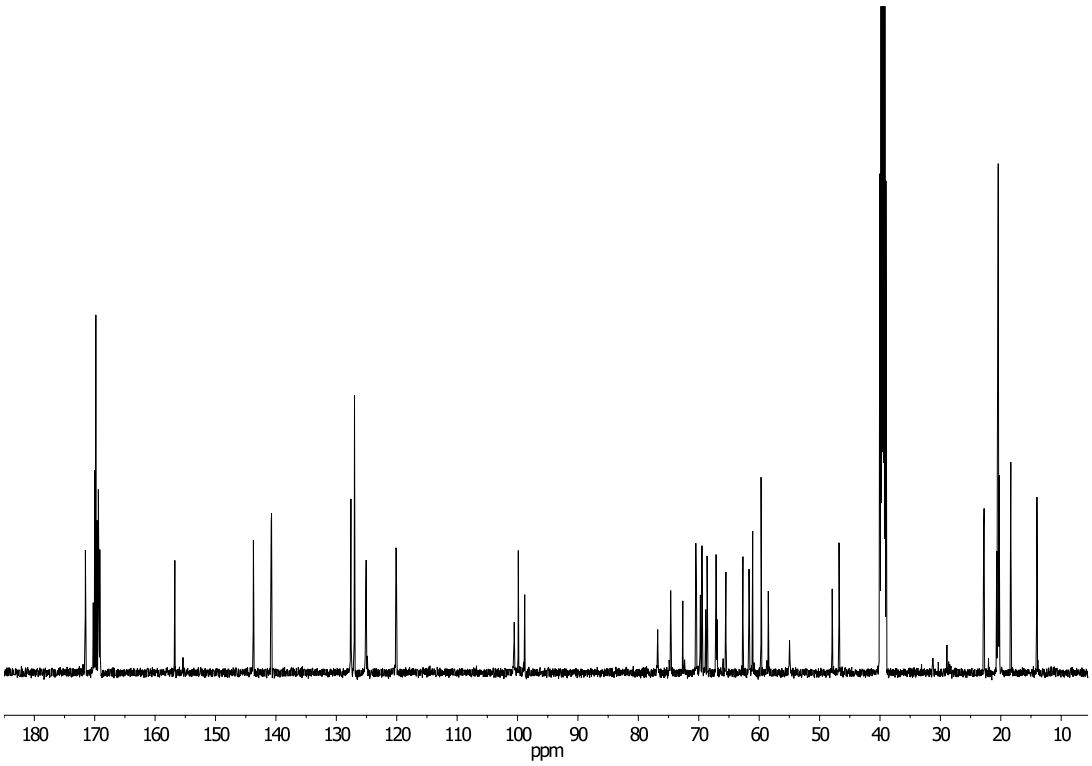


^{13}C NMR (150.9 MHz, DMSO-d_6)

Compound **39**

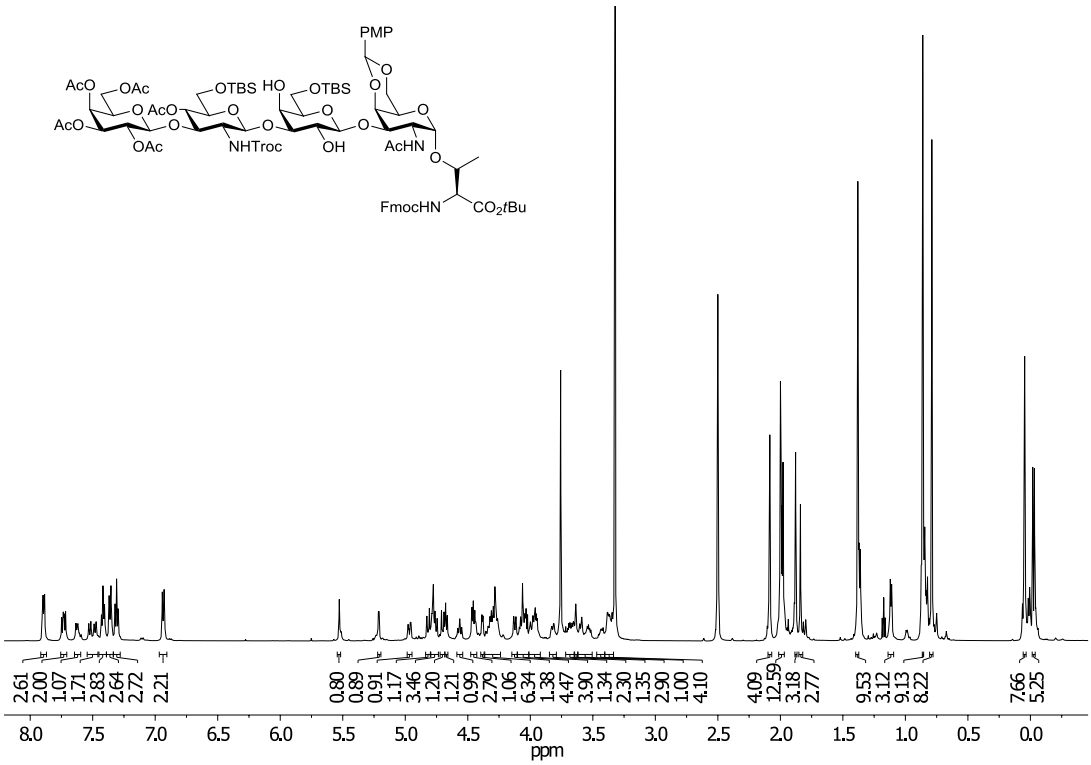


^{1}H *NMR* (600 MHz, DMSO-d_6)

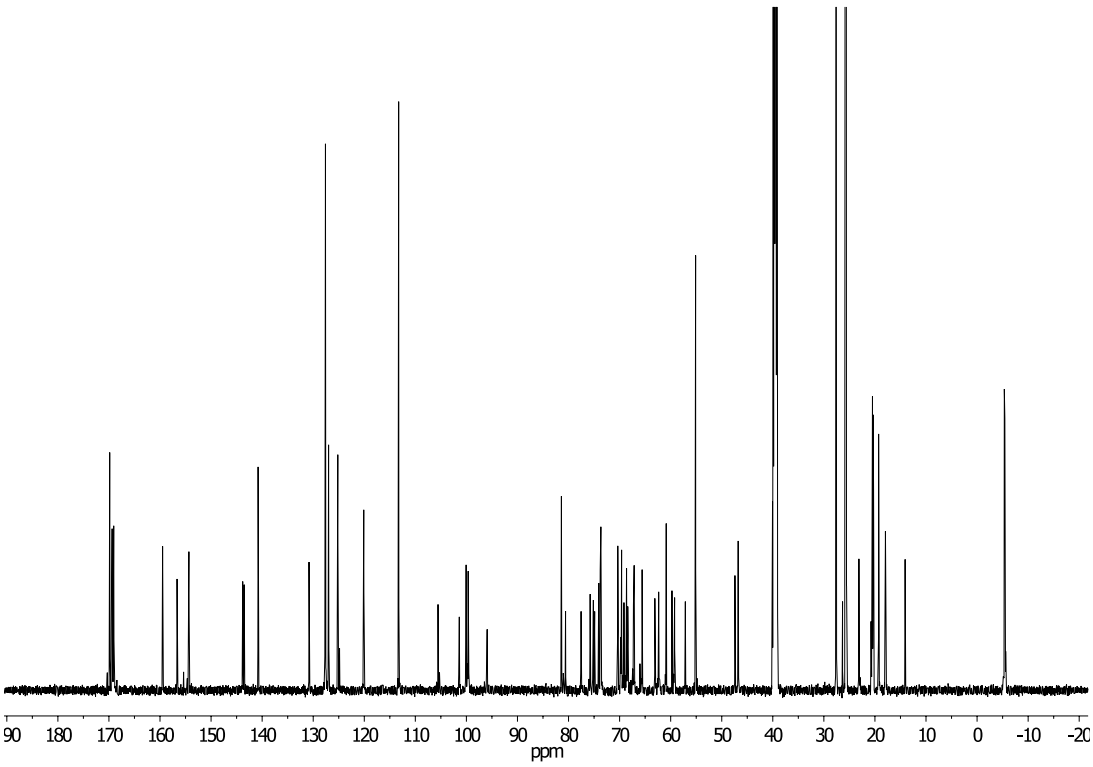


^{13}C NMR (150.9 MHz, DMSO-d_6)

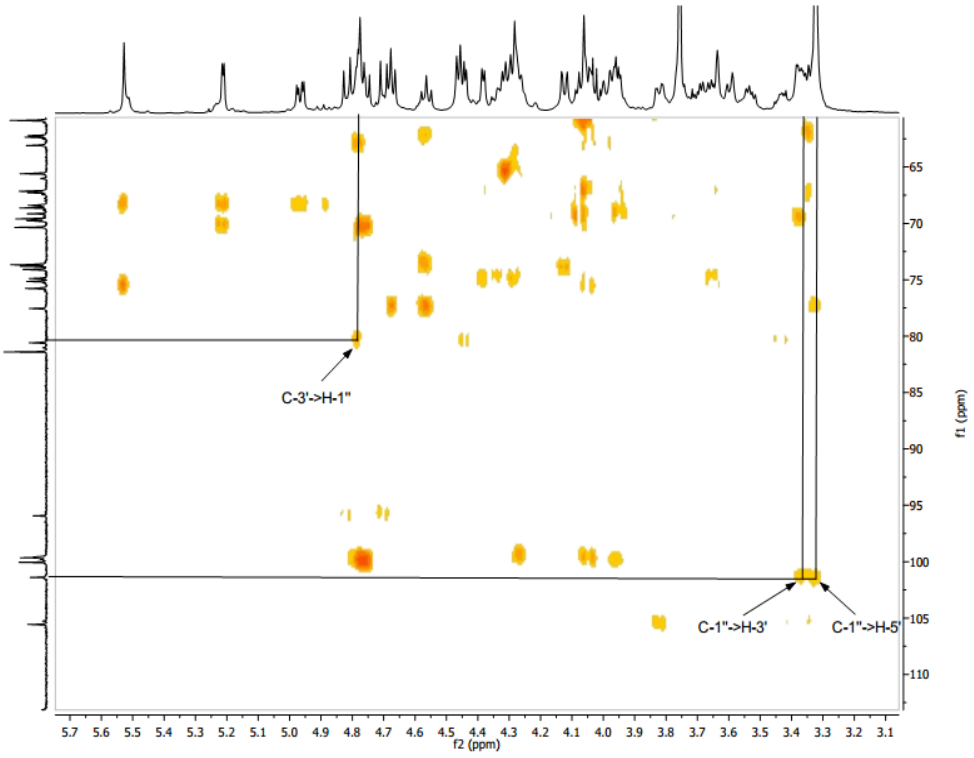
Compound **45**



^{1}H *NMR* (600 MHz, DMSO-d_6)

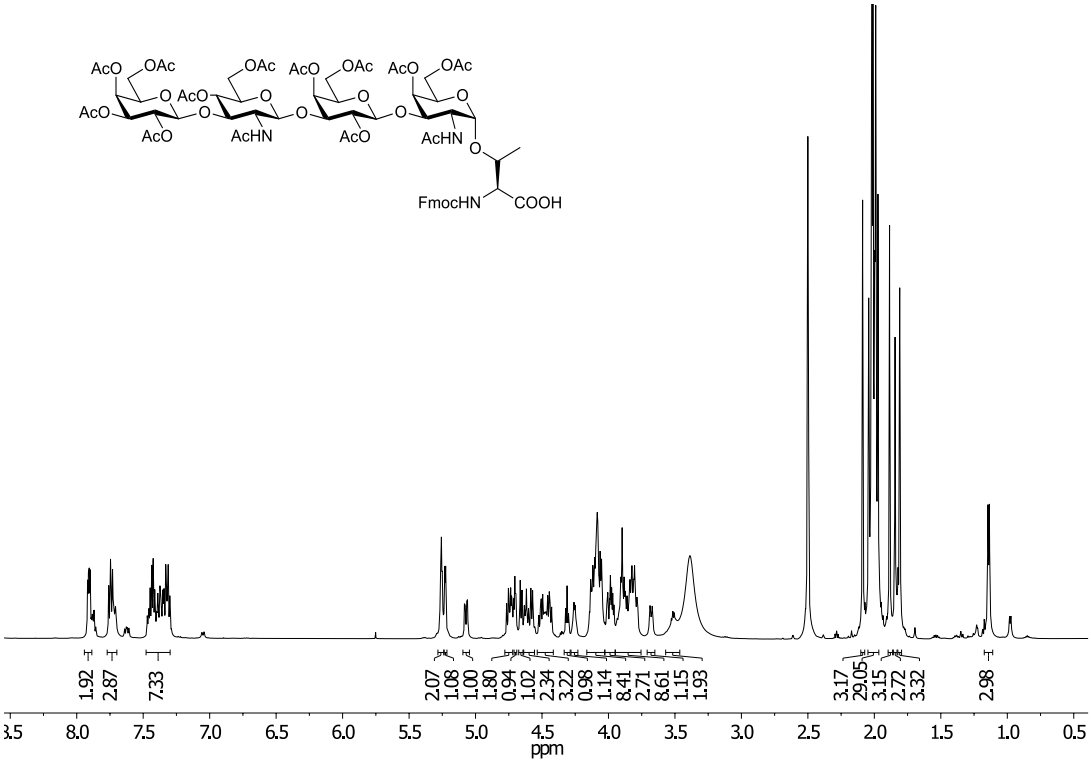


^{13}C NMR (150.9 MHz, DMSO-d_6)

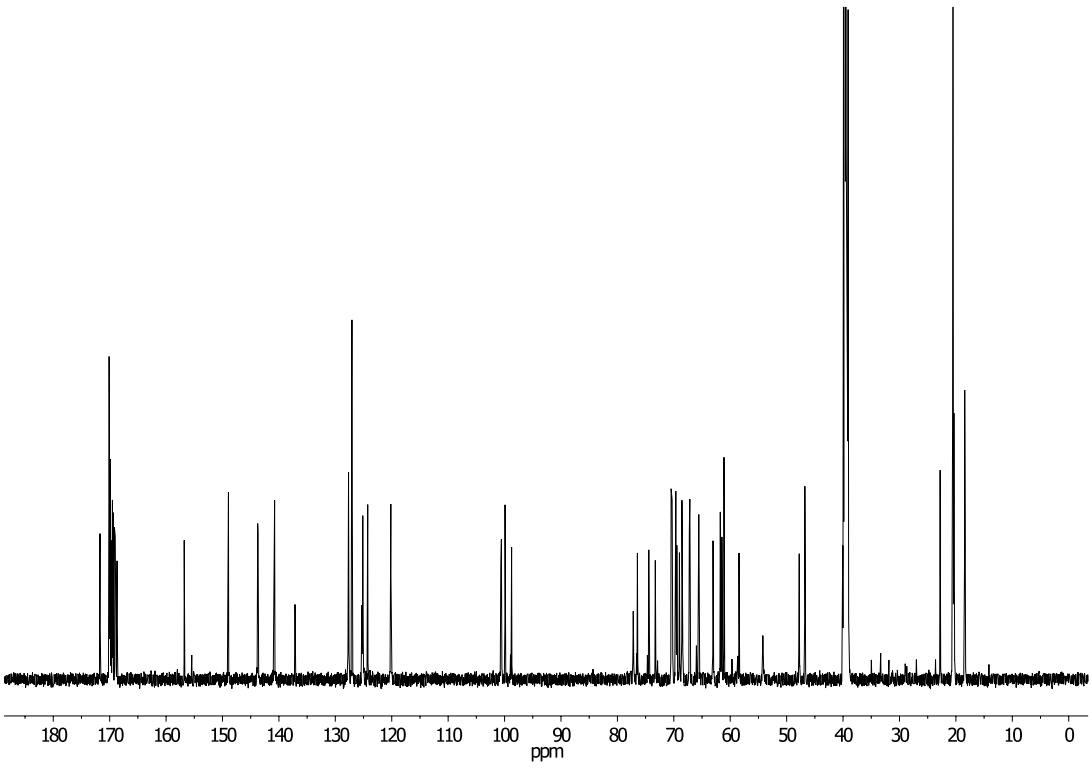


^{1}H-^{13}C-HMBC

Compound **49**

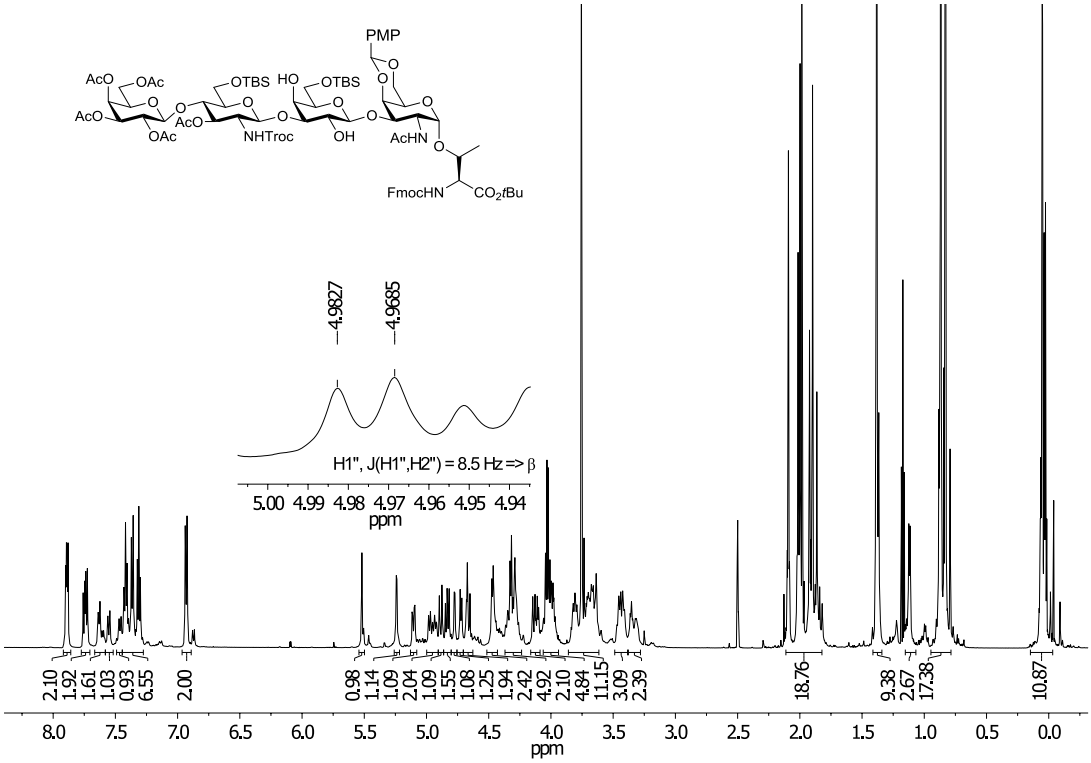


[1]H NMR (600 MHz, DMSO-d_6)

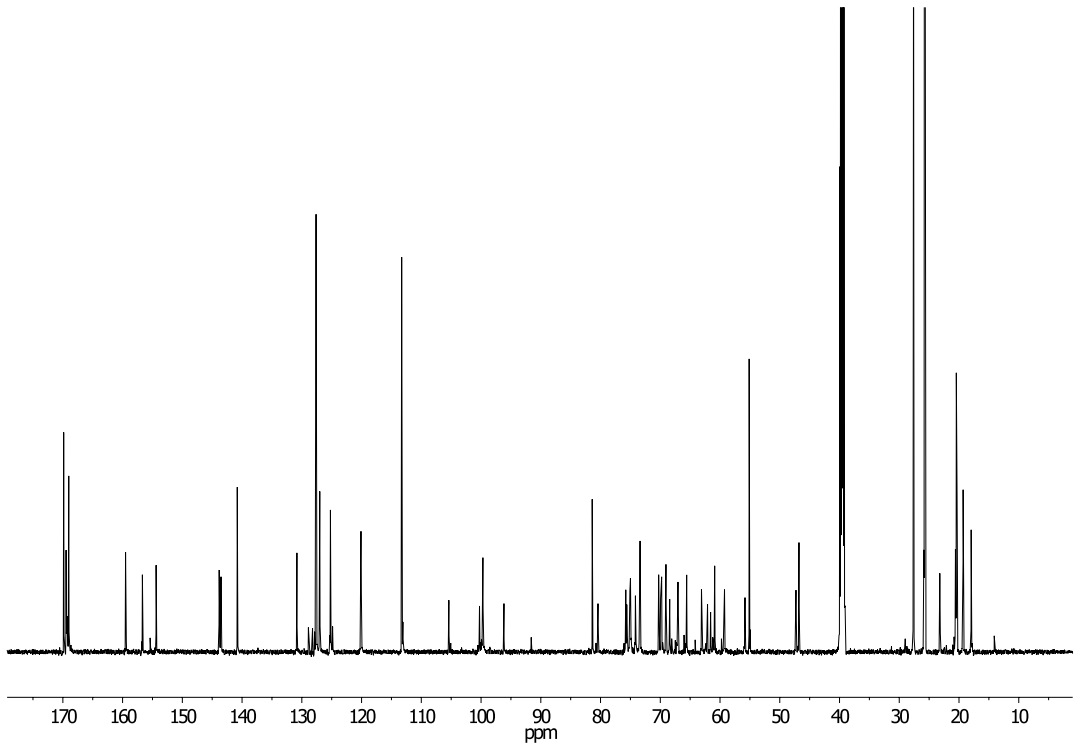


[13]C NMR (150.9 MHz, DMSO-d_6)

Compound **50**

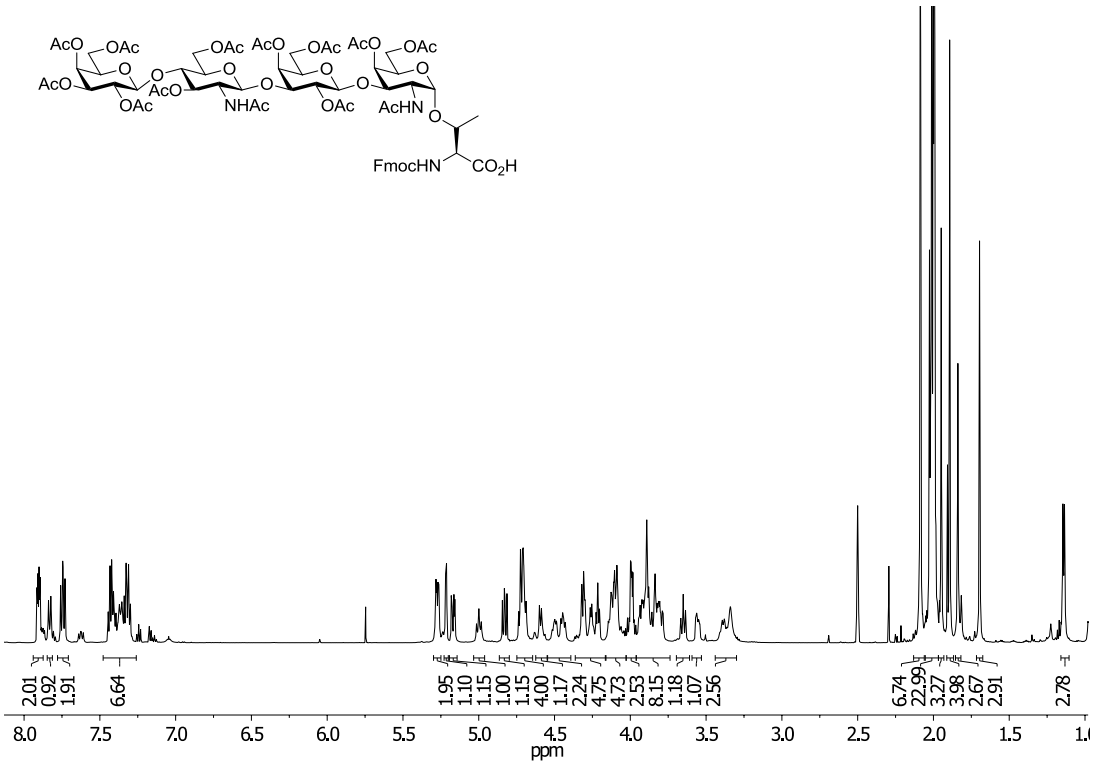


^{1}H *NMR* (600 MHz, DMSO-d_6)

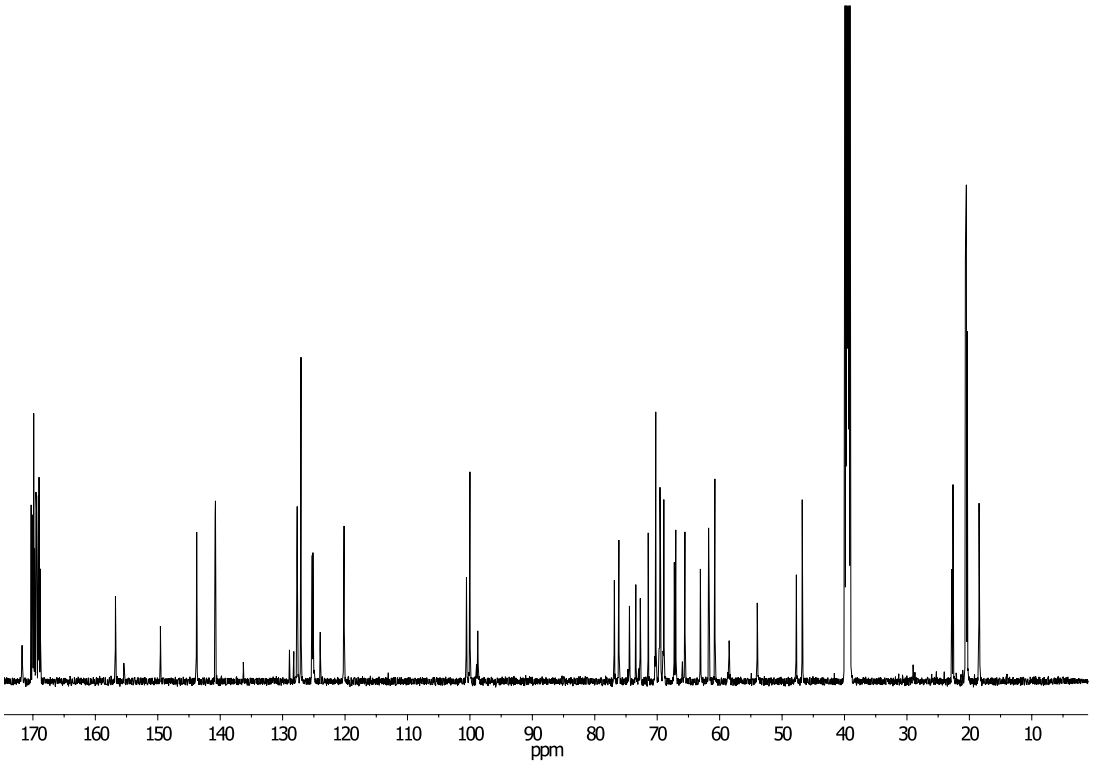


^{13}C NMR (150.9 MHz, DMSO-d_6)

Compound **54**

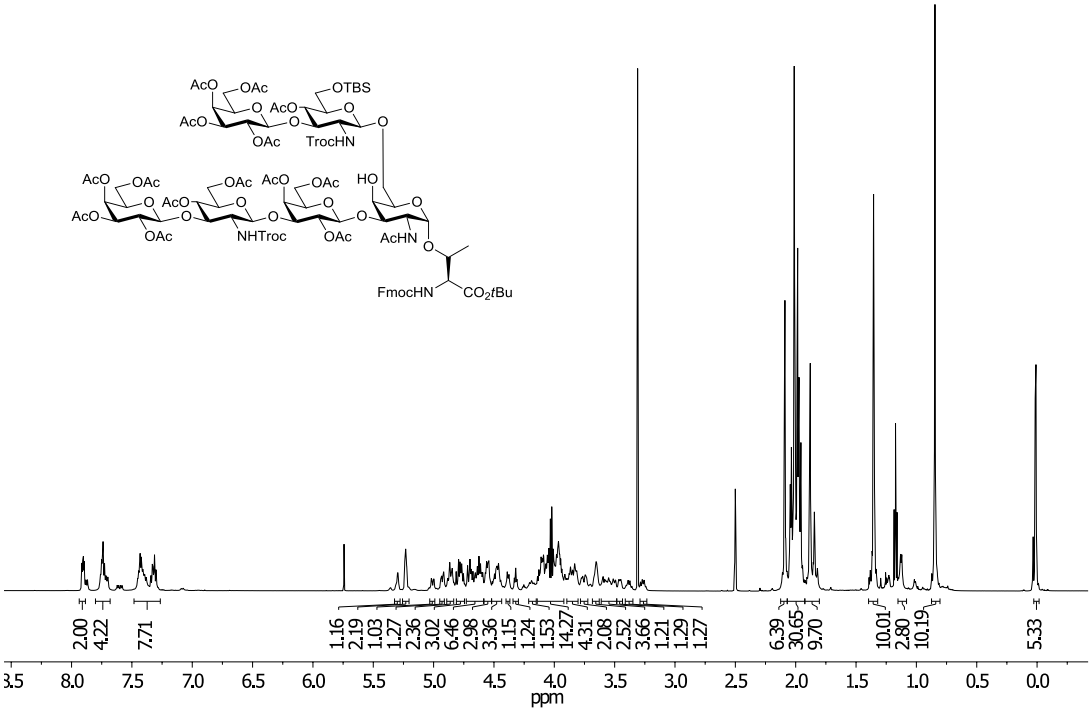


[1]H NMR (600 MHz, DMSO-d_6)

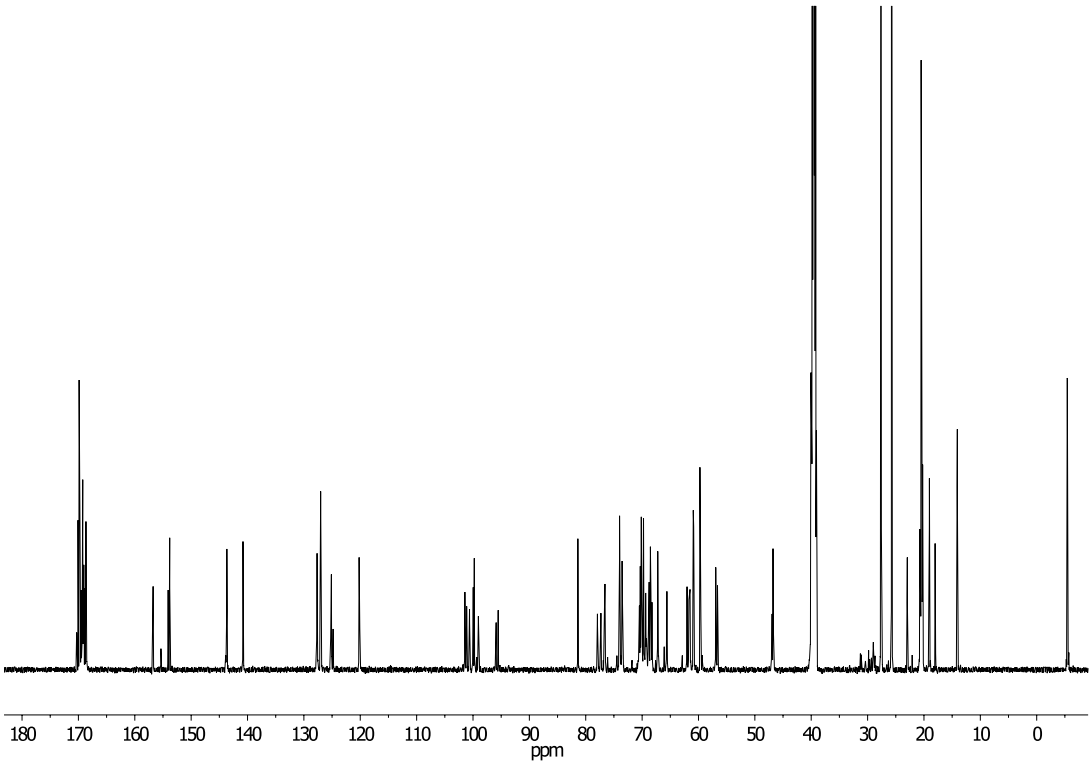


[13]C NMR (150.9 MHz, DMSO-d_6)

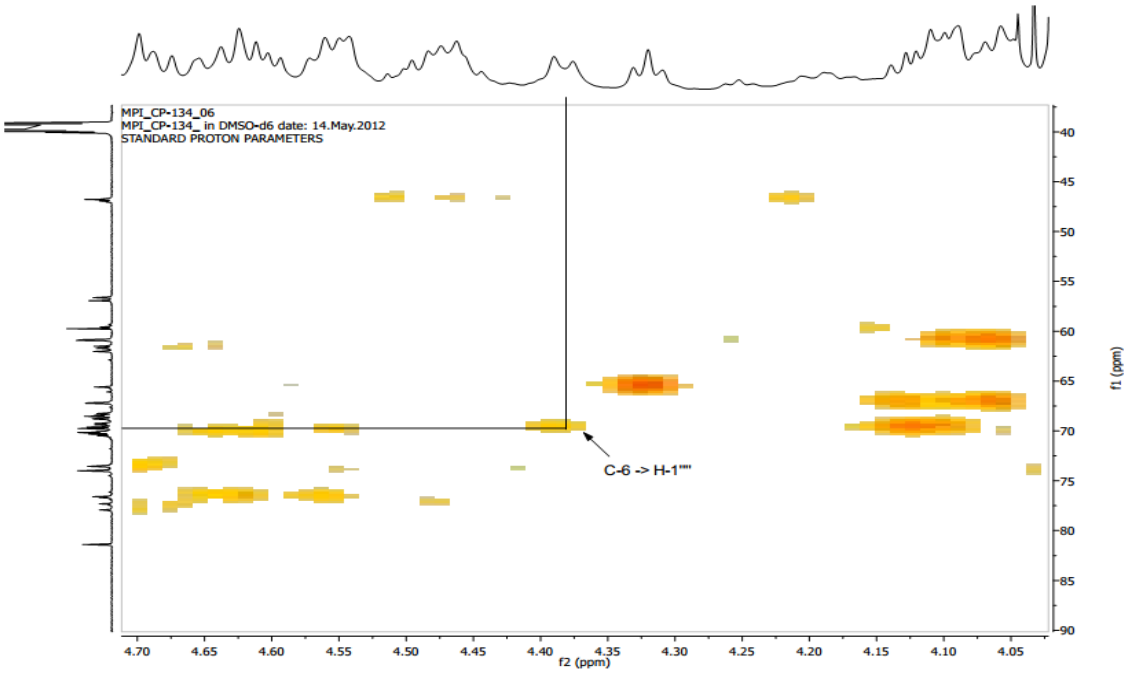
Compound **55**



[1]H NMR (600 MHz, DMSO-d_6)

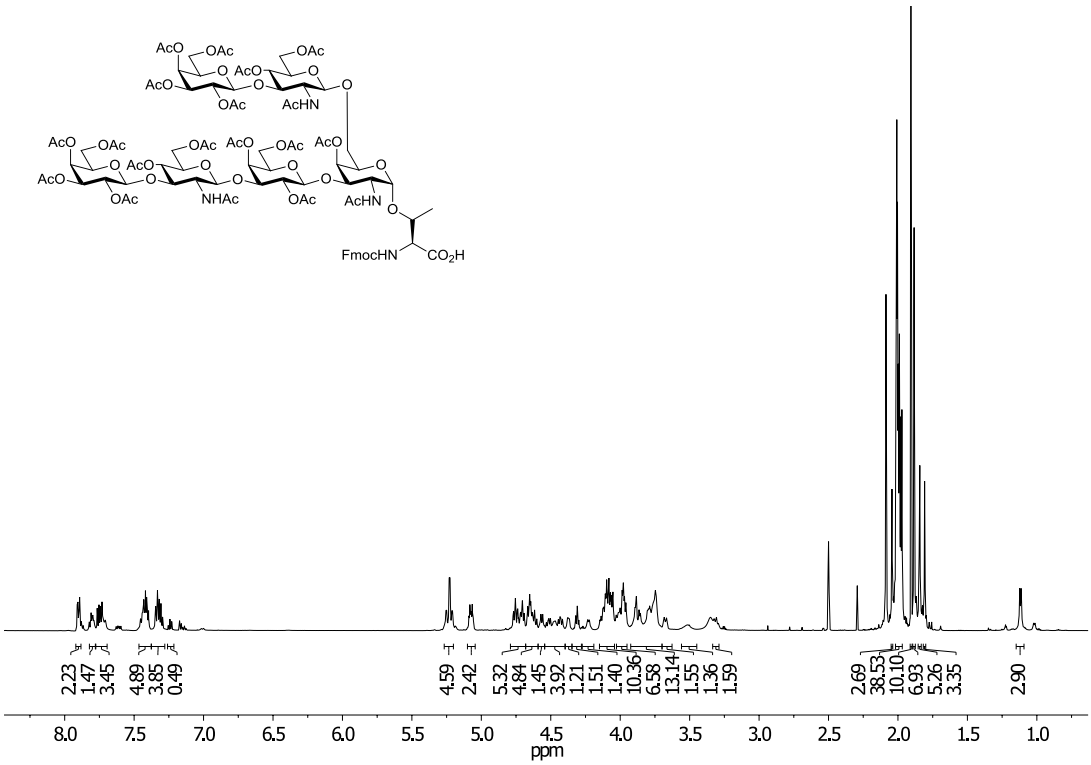


[13]C NMR (150.9 MHz, DMSO-d_6)

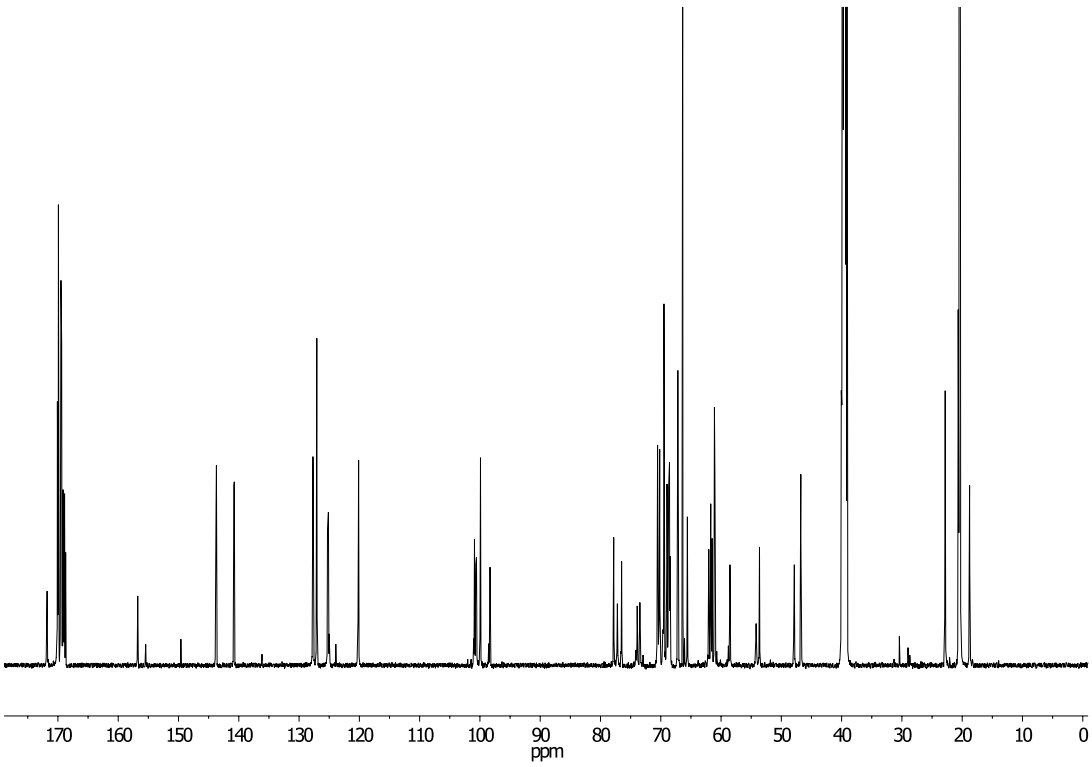

MPI_CP-134_06
MPI_CP-134_ in DMSO-d6 date: 14.May.2012
STANDARD PROTON PARAMETERS
C-6 -> H-1''''
f2 (ppm)
f1 (ppm)

^{1}H-^{13}C-HMBC

Compound **58**

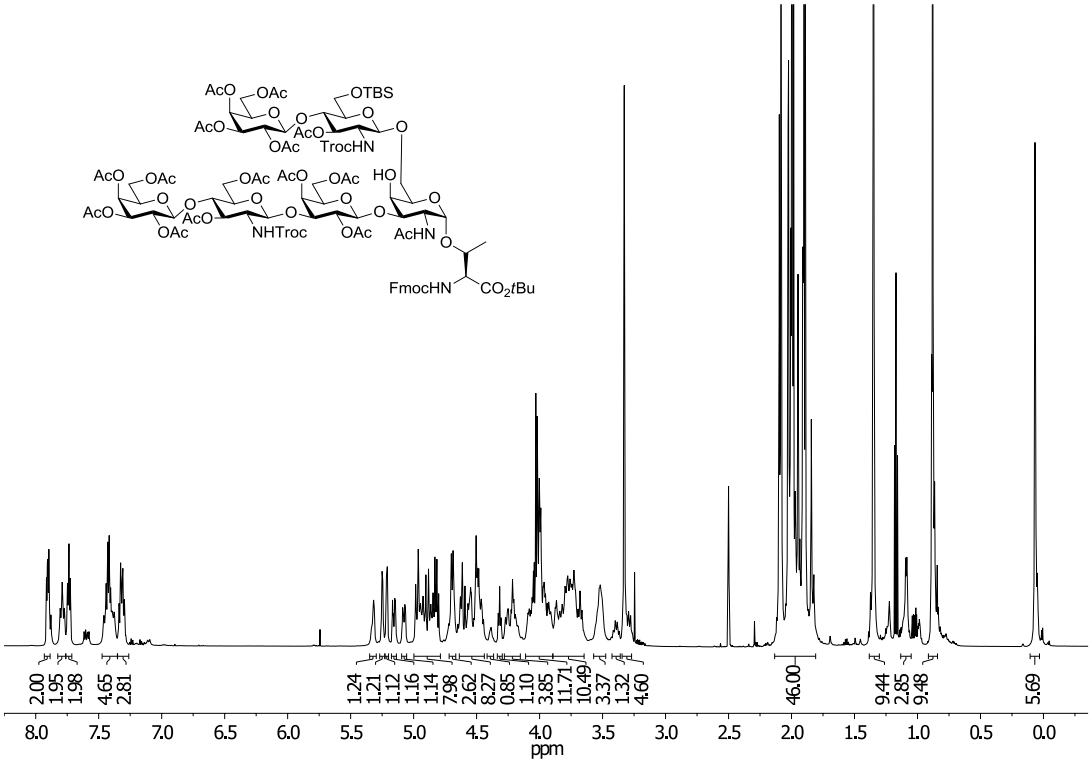


[1]H NMR (600 MHz, DMSO-d_6)

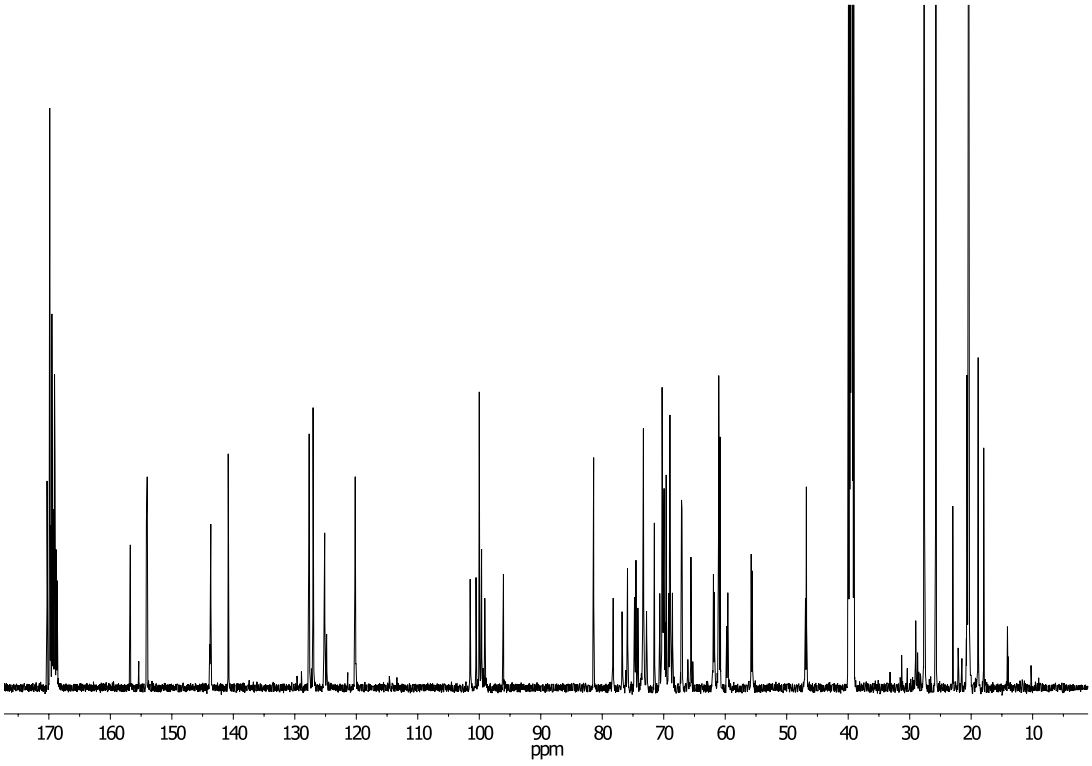


[13]C NMR (150.9 MHz, DMSO-d_6)

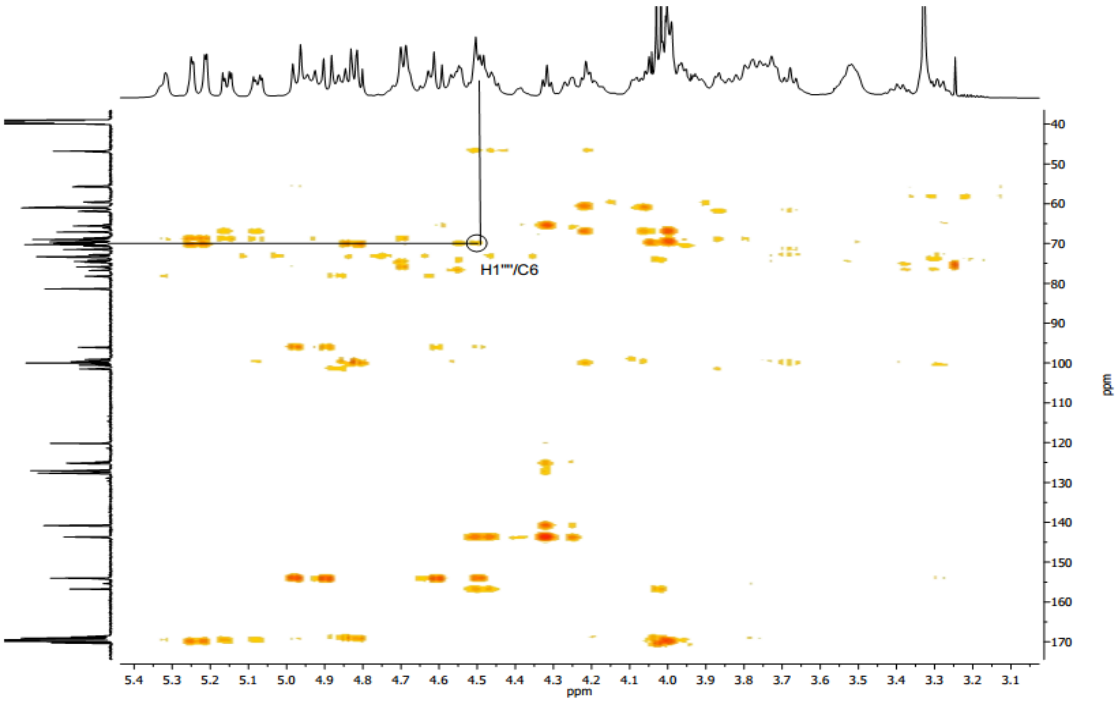
Compound **59**



^{1}H *NMR* (600 MHz, DMSO-d_6)

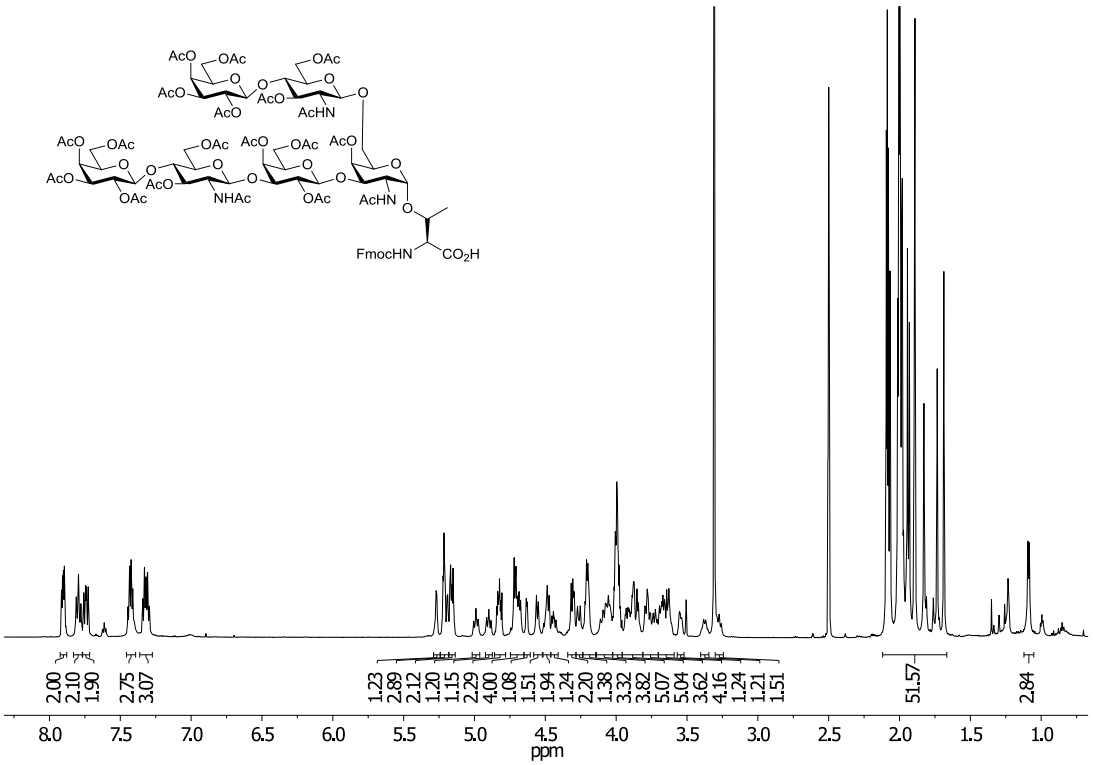


^{13}C NMR (150.9 MHz, DMSO-d_6)

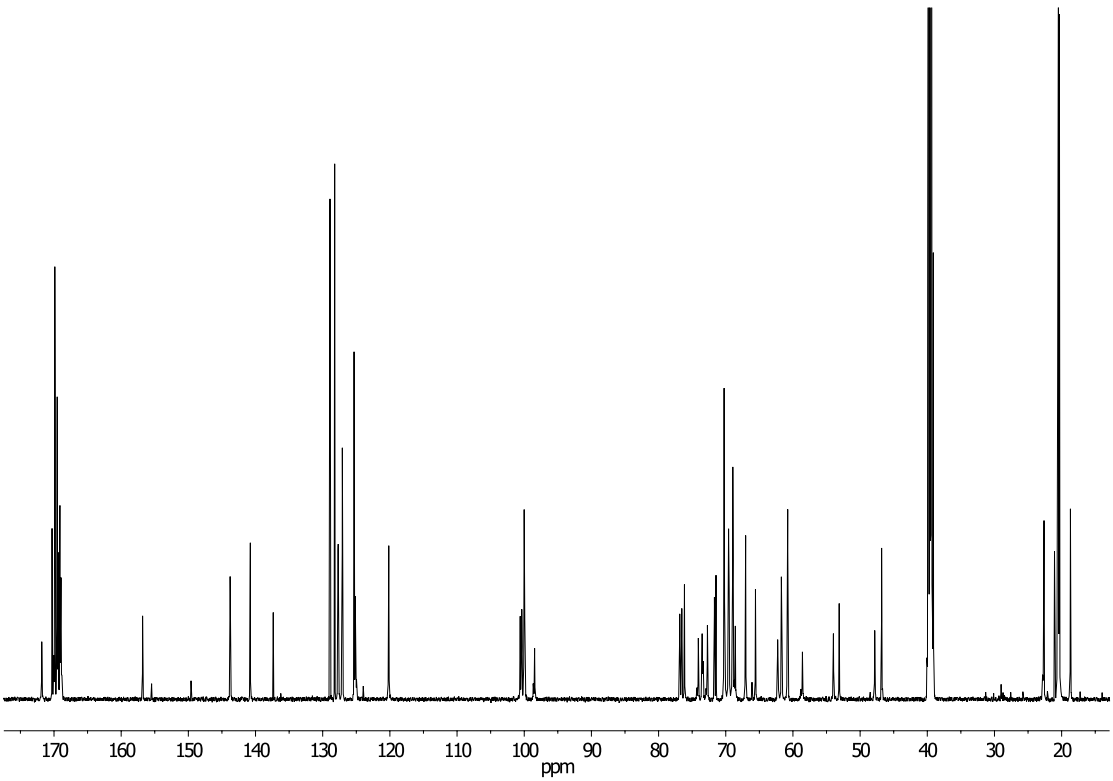


^{1}H-^{13}C-HMBC

Compound **62**



^{1}H *NMR* (600 MHz, DMSO-d_6)

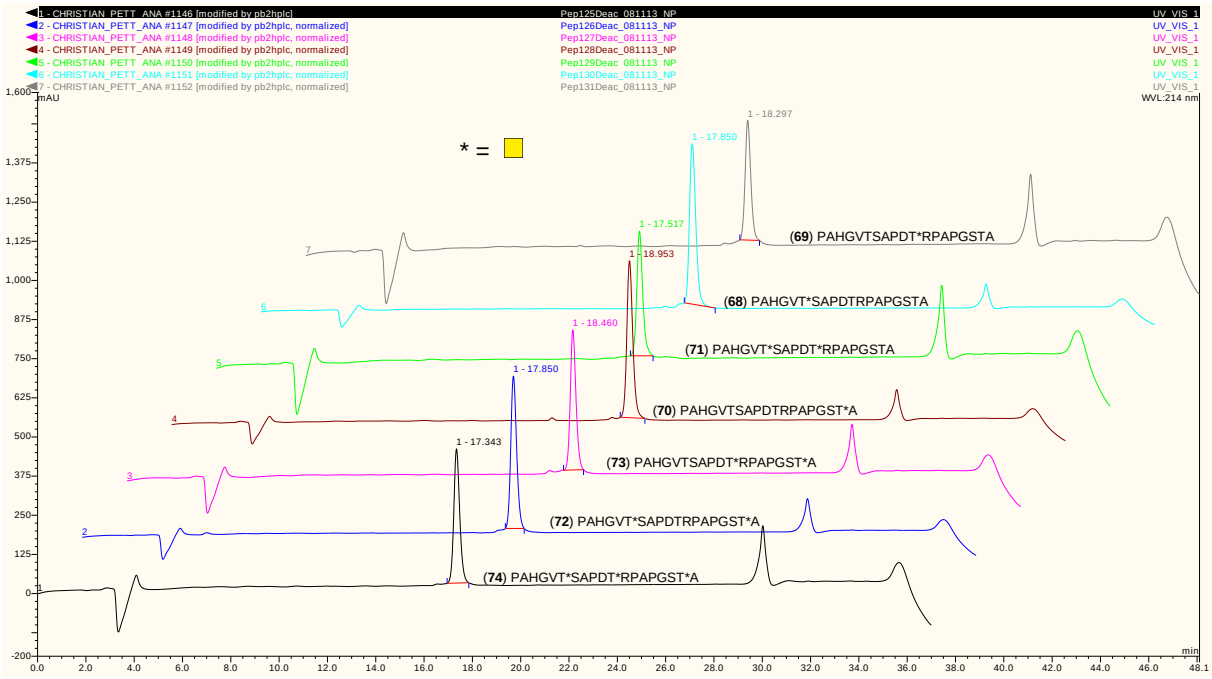


^{13}C NMR (150.9 MHz, DMSO-d_6)

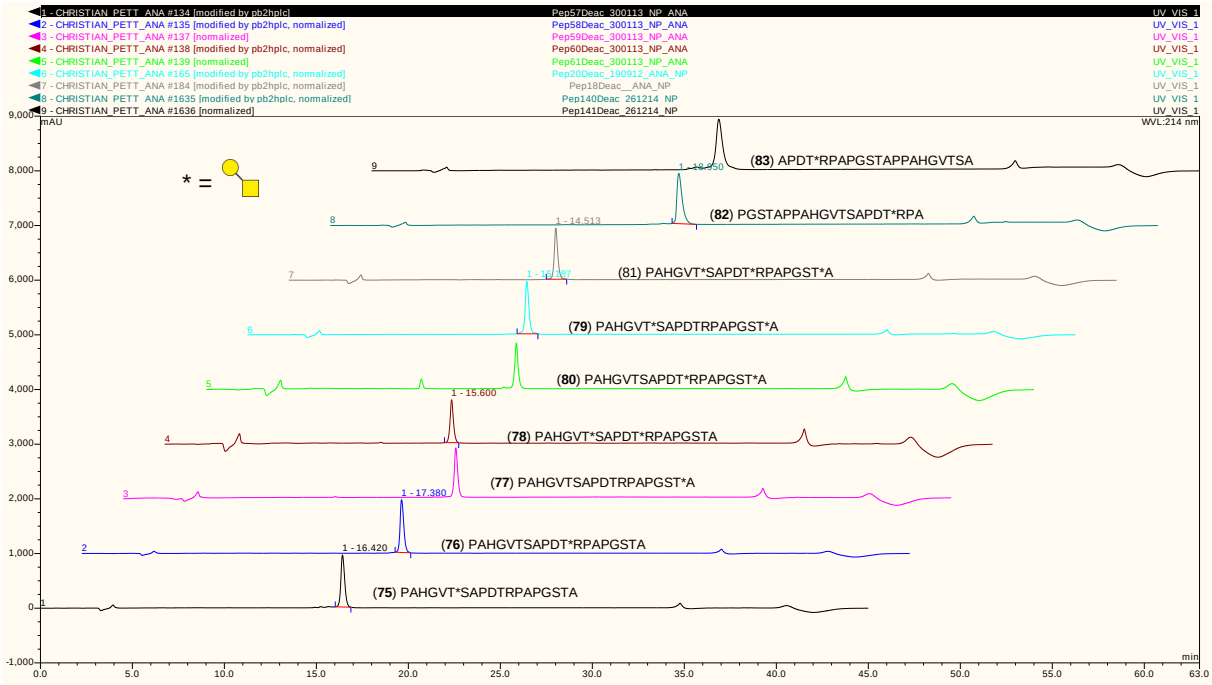
9.2 HPLC chromatograms

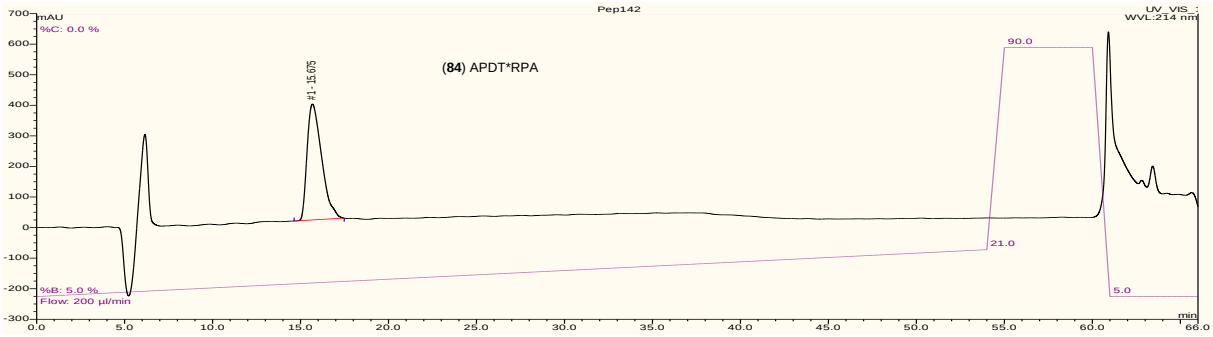
9.2.1 MUC1 glycopeptides

9.2.1.1 T_N-antigen glycopeptides

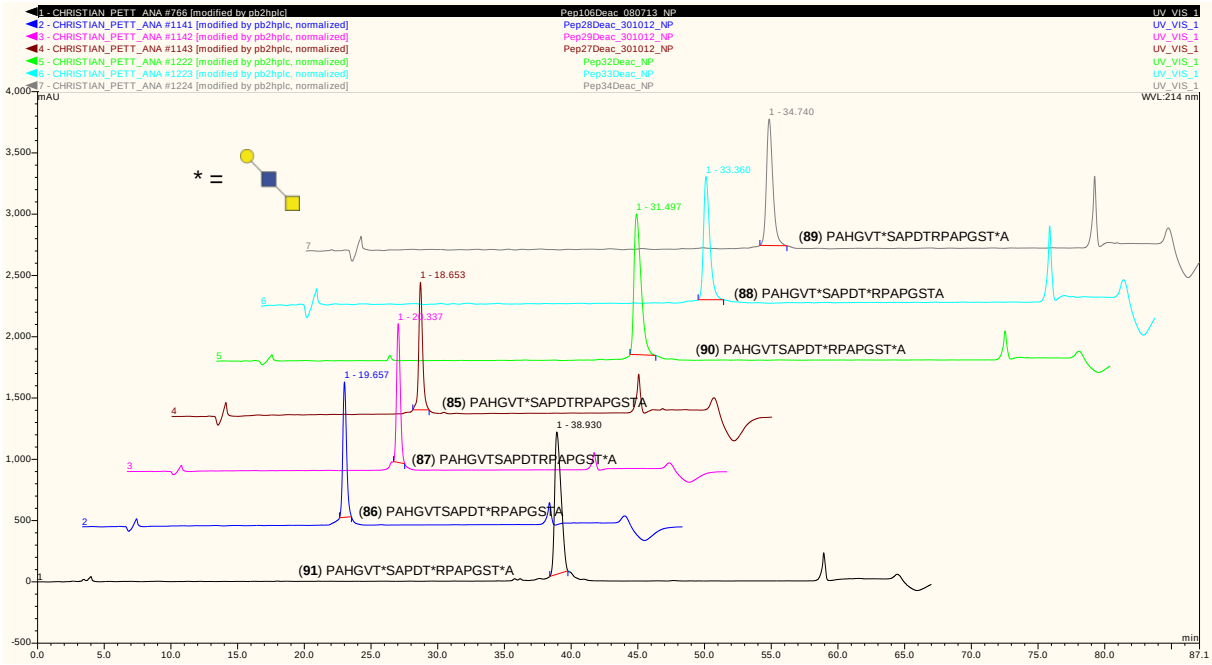


9.2.1.2 T-antigen glycopeptides

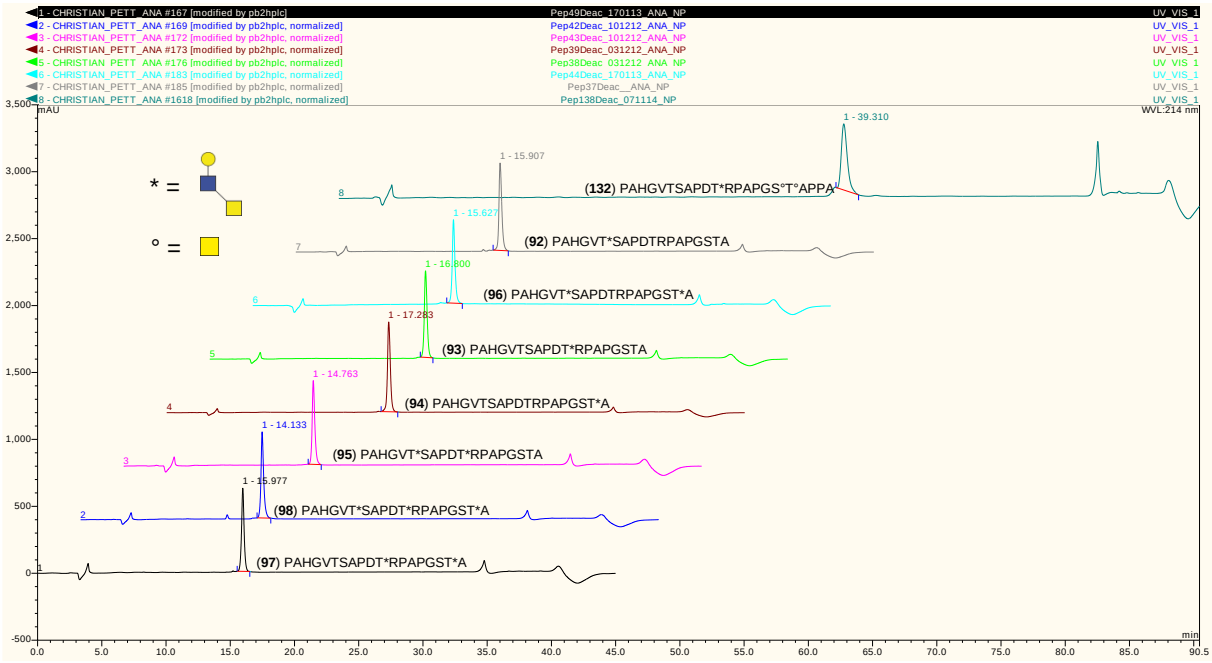


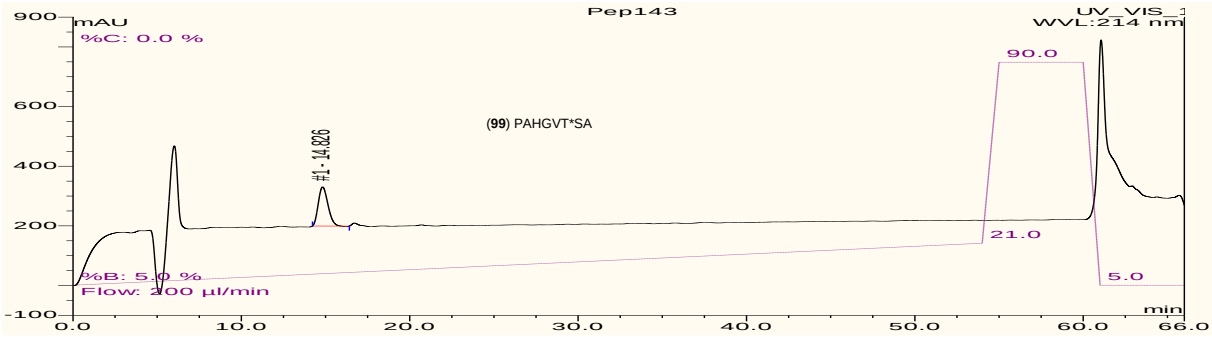


9.2.1.3 Core 3 type-1 glycopeptides

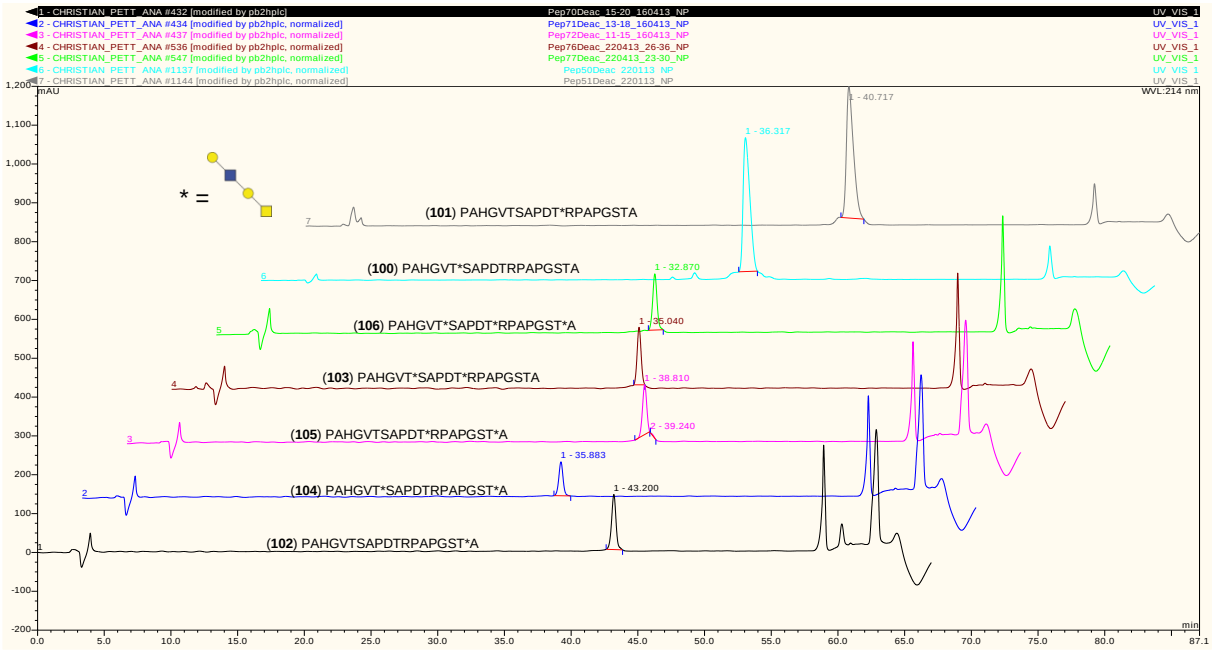


9.2.1.4 Core 3 type-2 glycopeptides

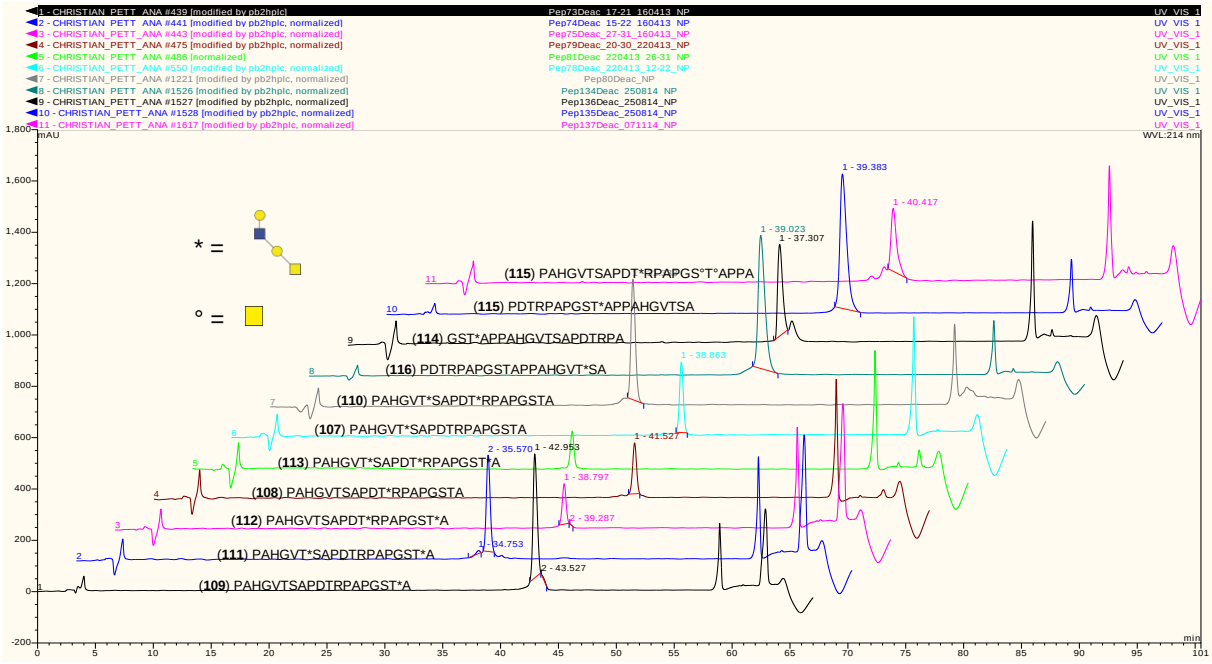




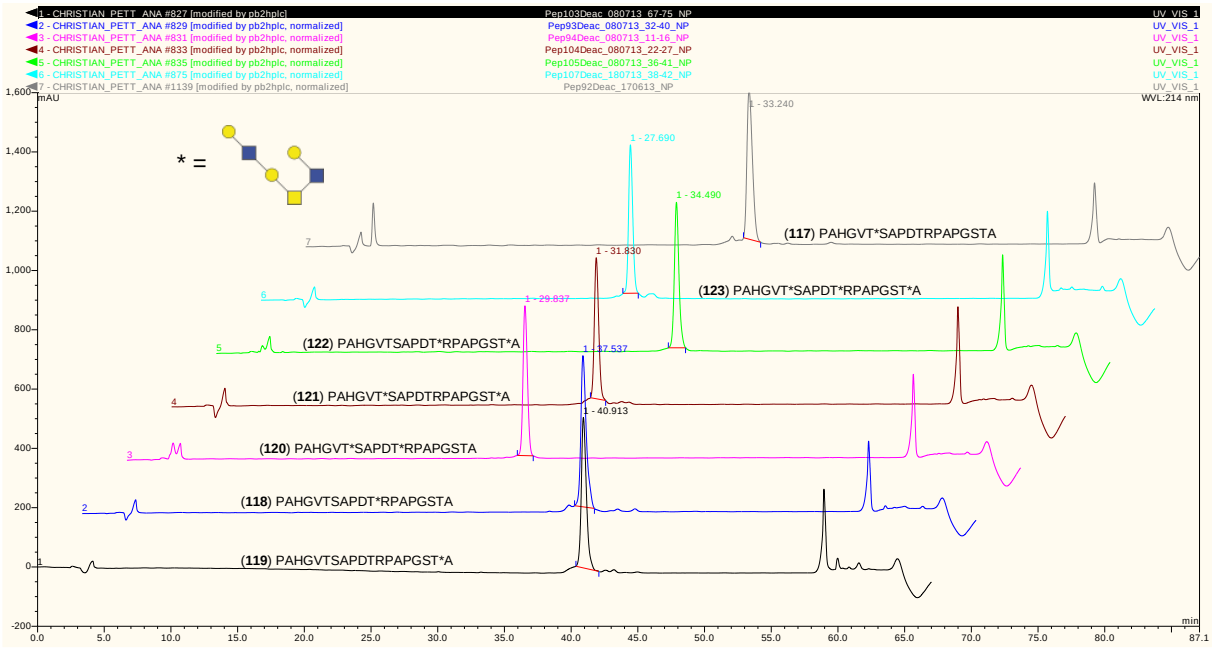
9.2.1.5 Core 1 type-1 glycopeptides



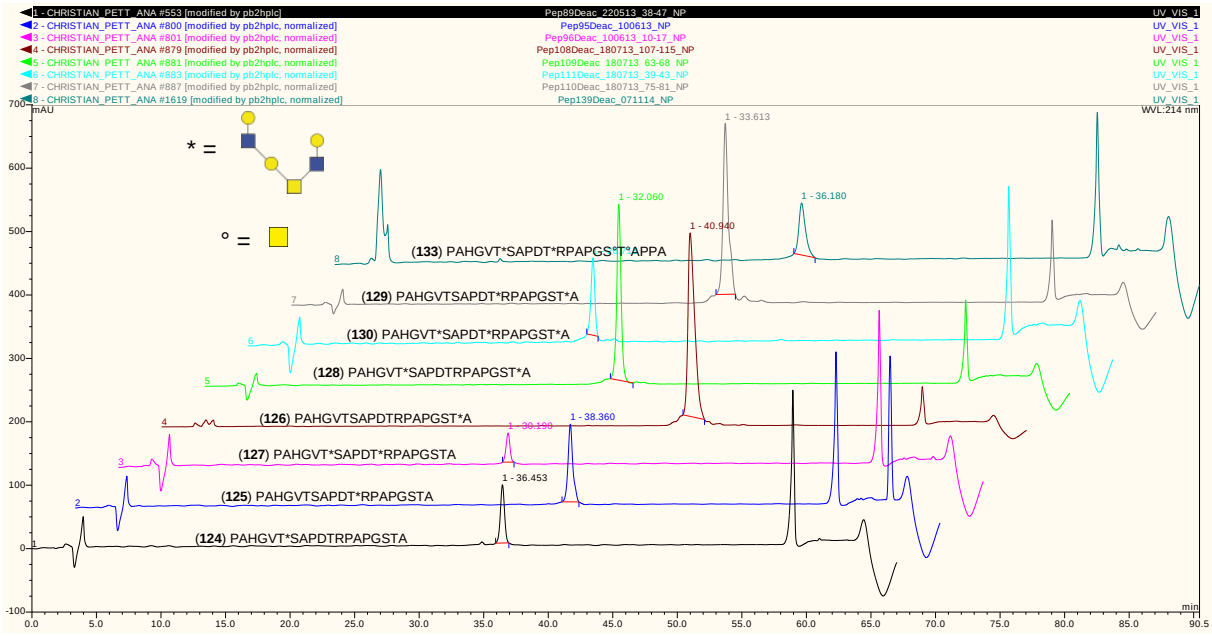
9.2.1.6 Core 1 type-2 glycopeptides



9.2.1.7 Core 2 type-1 glycopeptides

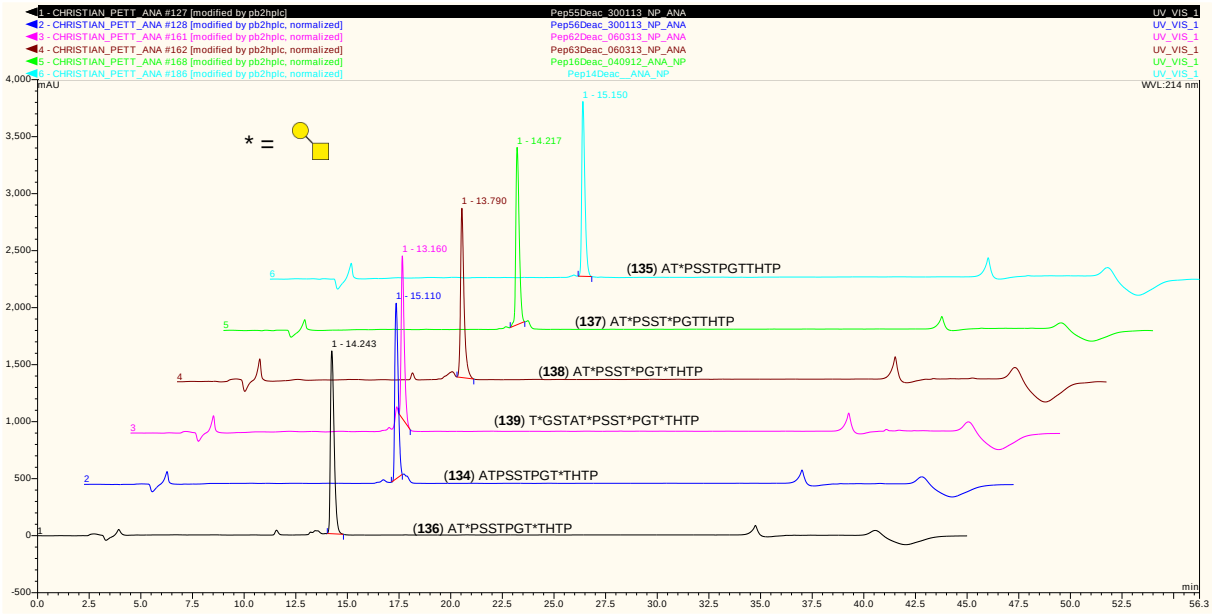


9.2.1.8 Core 2 type-2 glycopeptides

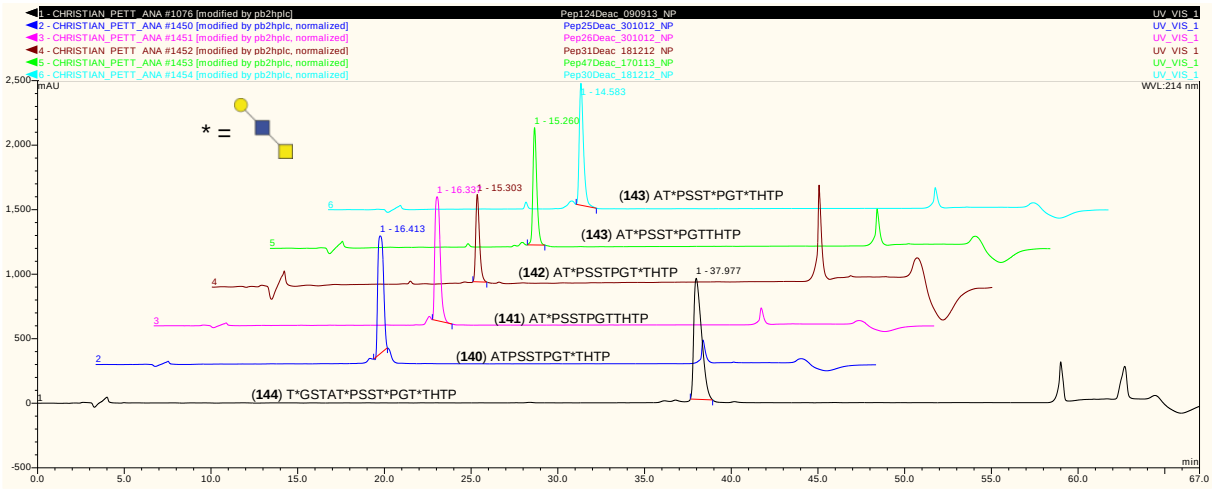


9.2.2 MUC5B glycopeptides

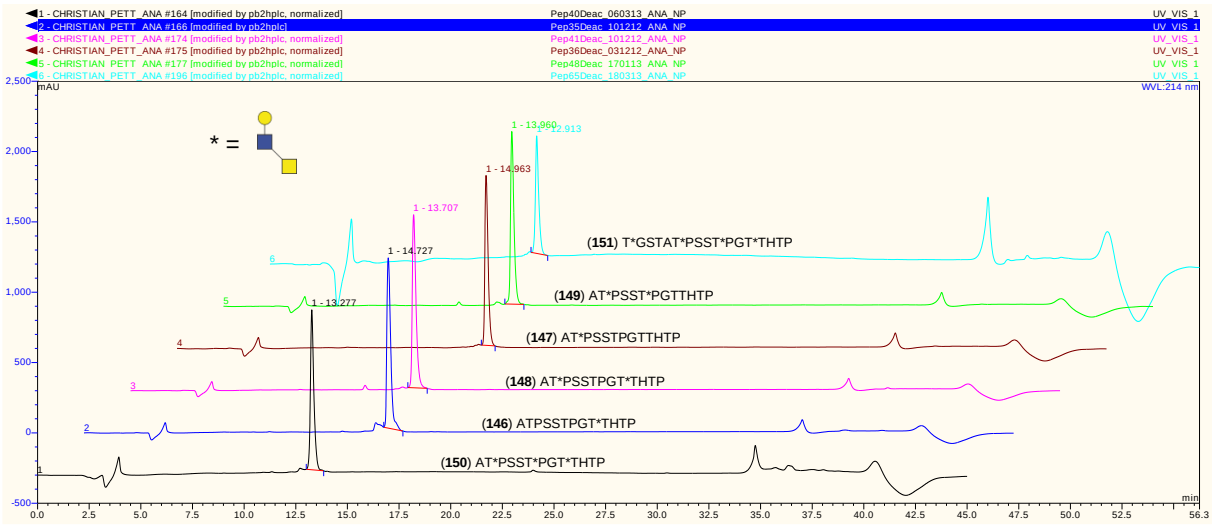
9.2.2.1 T-antigen glycopeptides



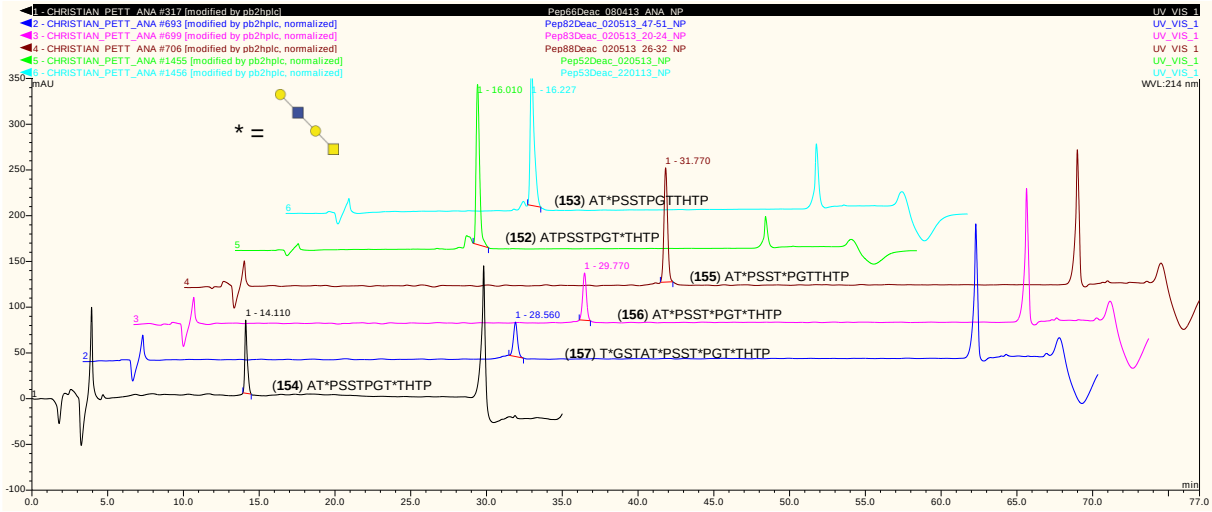
9.2.2.2 Core 3 type-1 glycopeptides



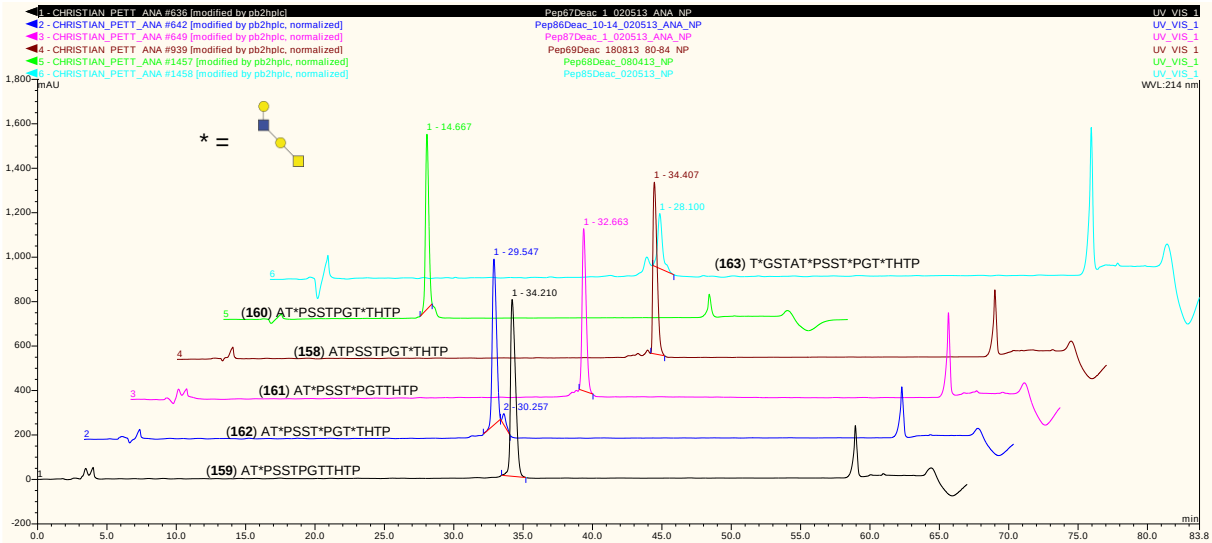
9.2.2.3 Core 3 type-2 glycopeptides



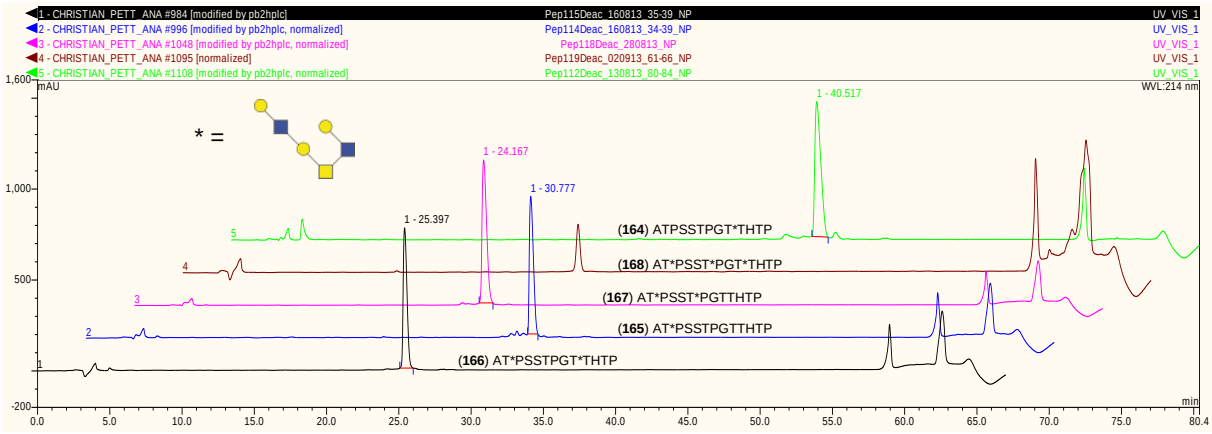
9.2.2.4 Core 1 type-1 glycopeptides



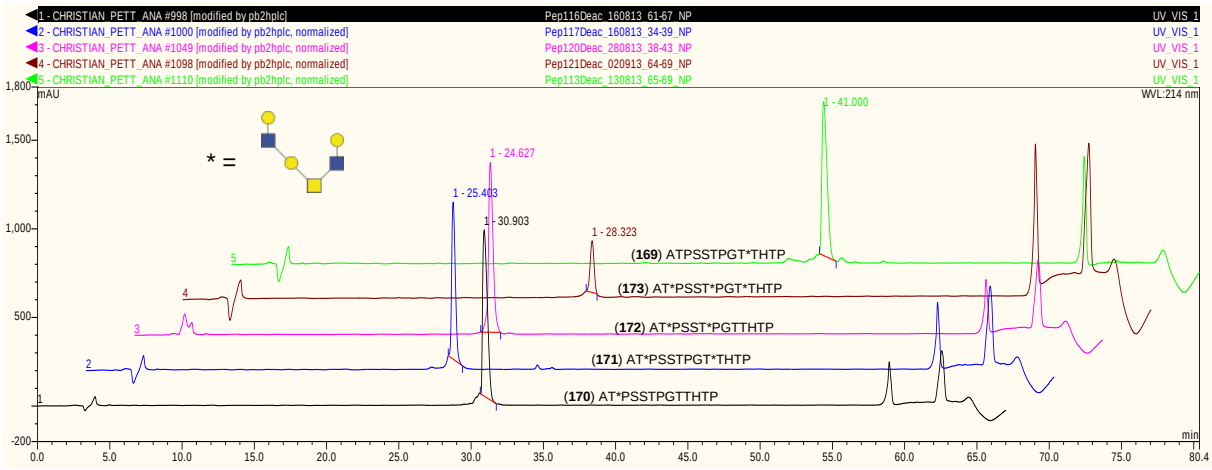
9.2.2.5 Core 1 type-2 glycopeptides



9.2.2.6 Core 2 type-1 glycopeptides

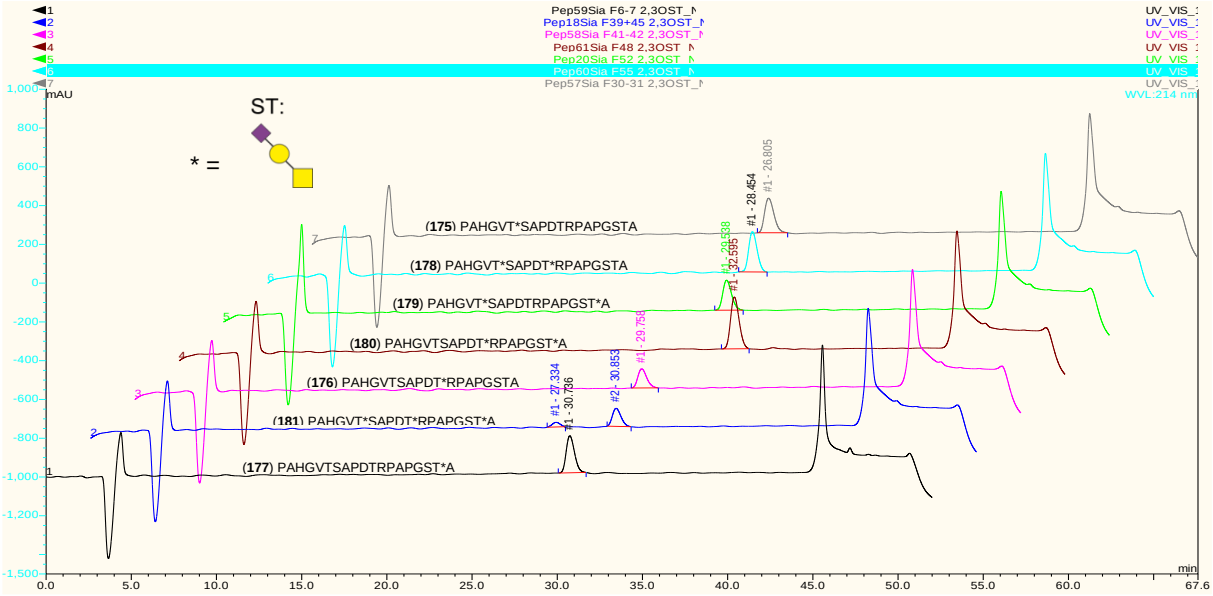


9.2.2.7 Core 2 type-2 glycopeptides

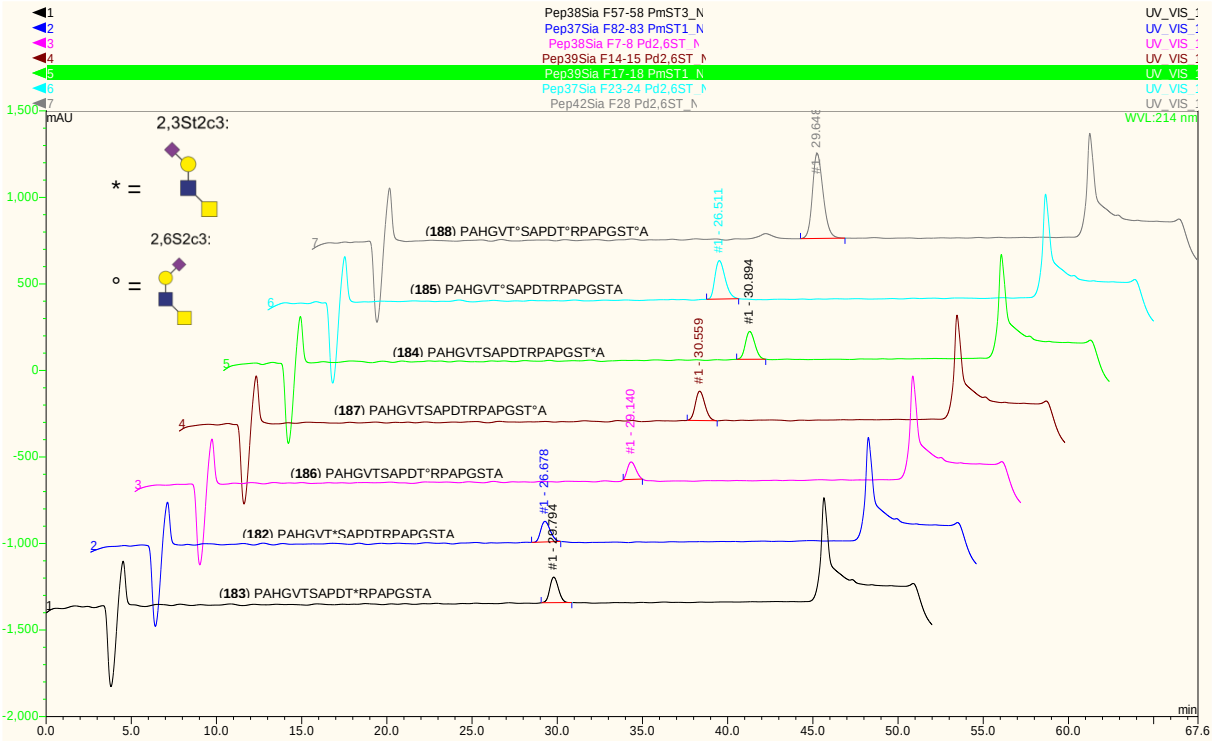


9.2.3 Sialylated MUC1 glycopeptides

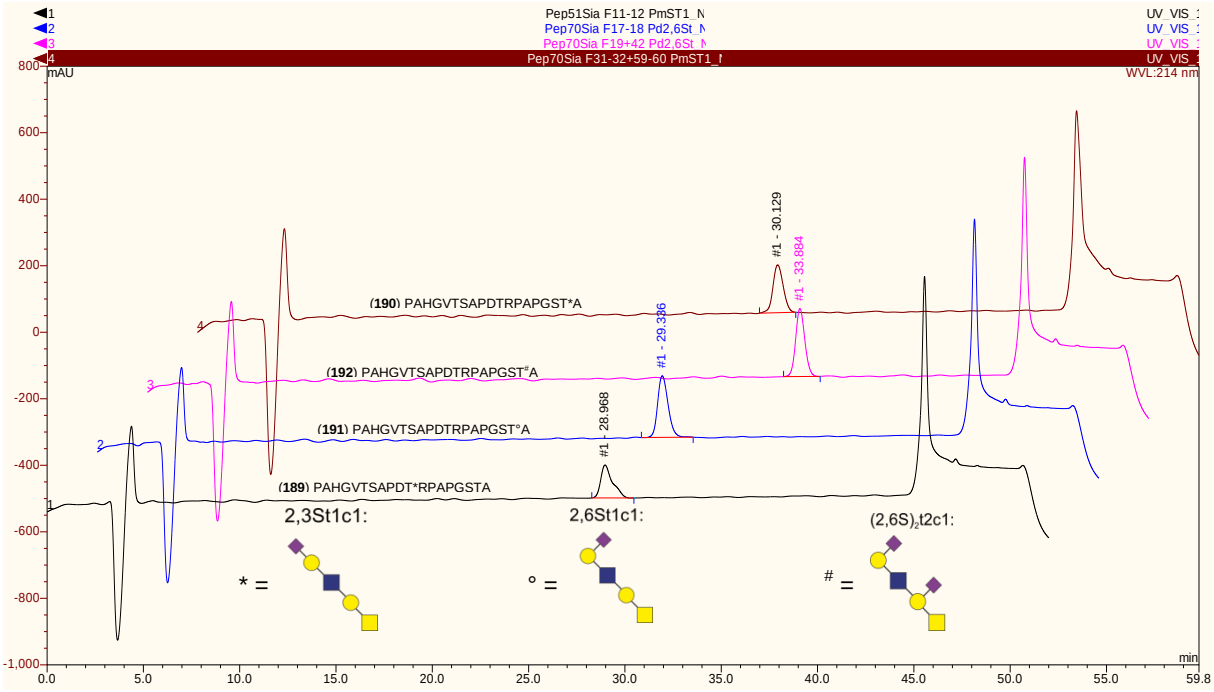
9.2.3.1 2,3-ST-antigen glycopeptides



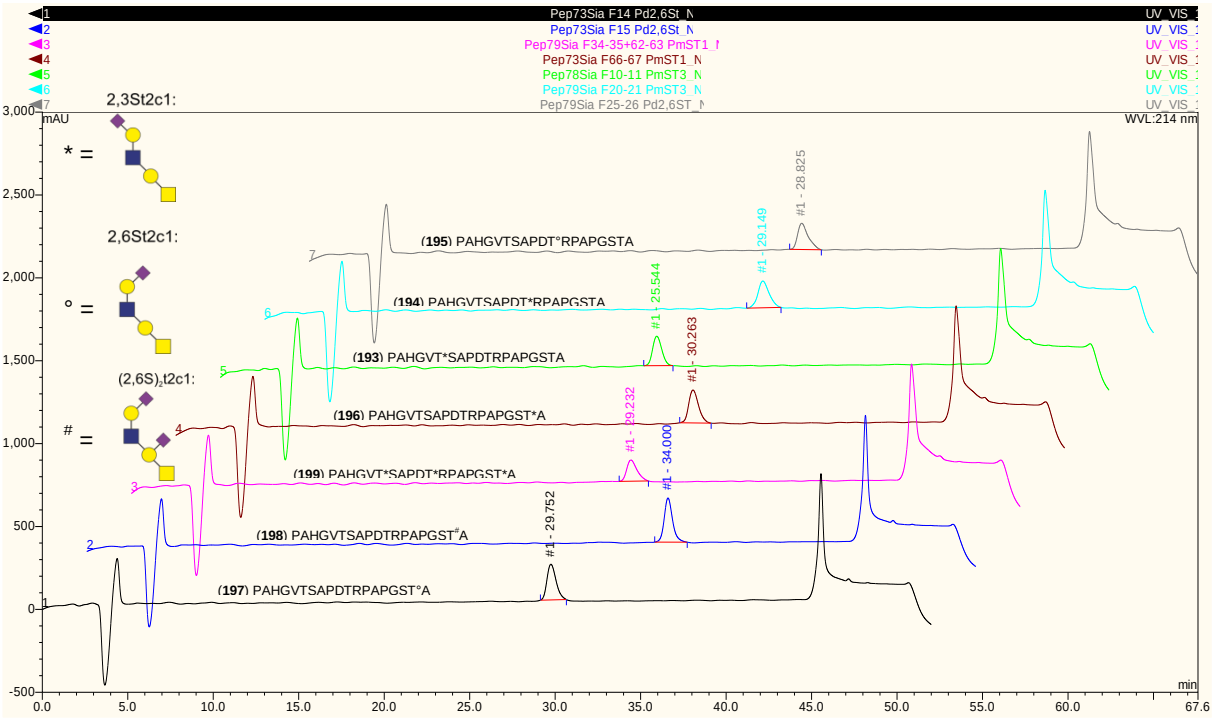
9.2.3.2 Sialylated core 3 type-2 glycopeptides



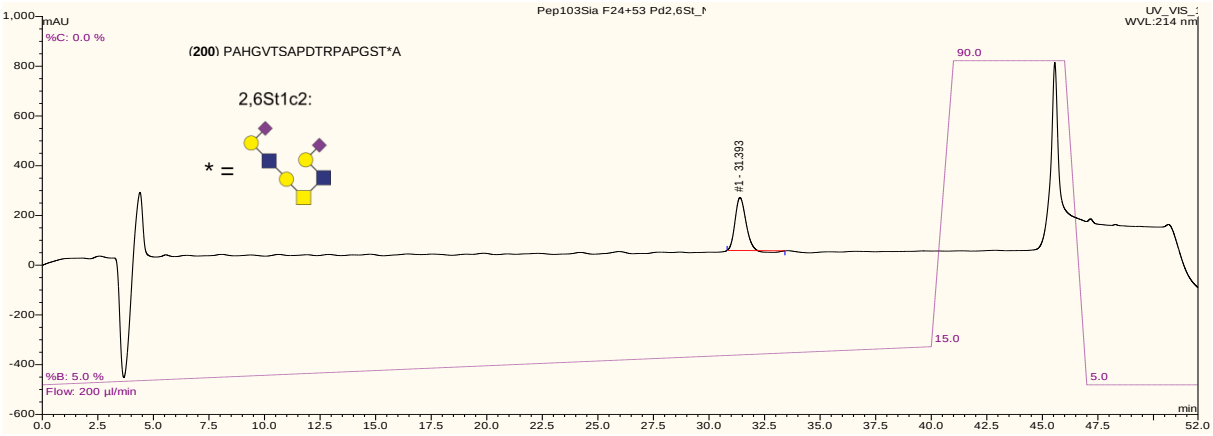
9.2.3.3 Sialylated core 1 type-1 glycopeptides



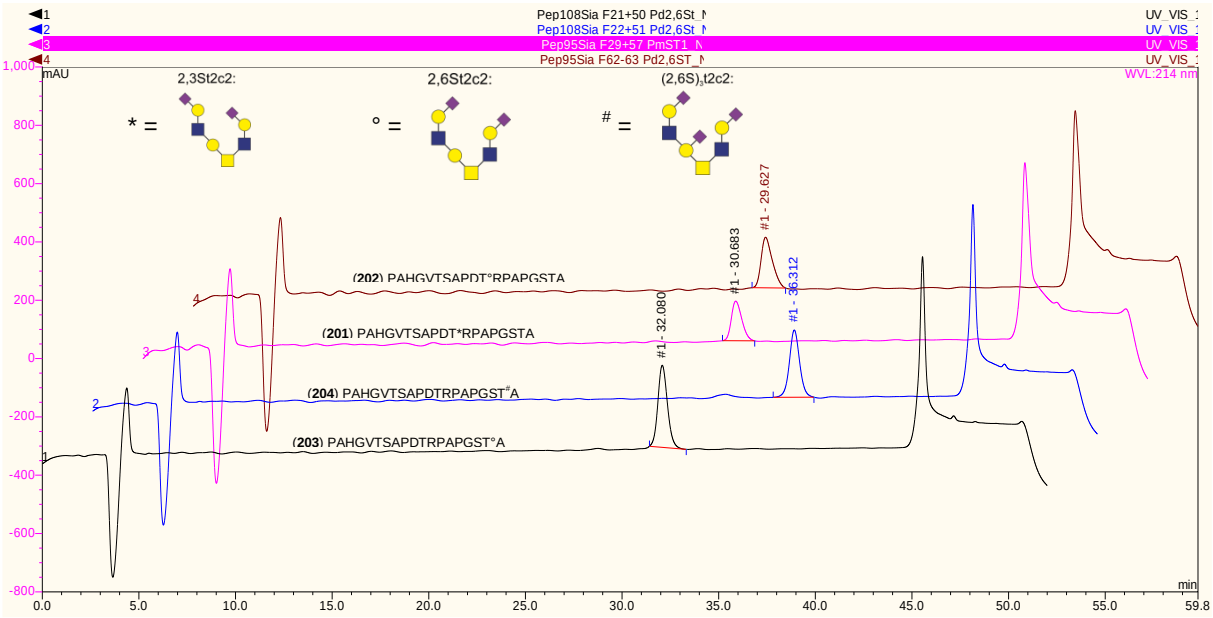
9.2.3.4 Sialylated core 1 type-2 glycopeptides



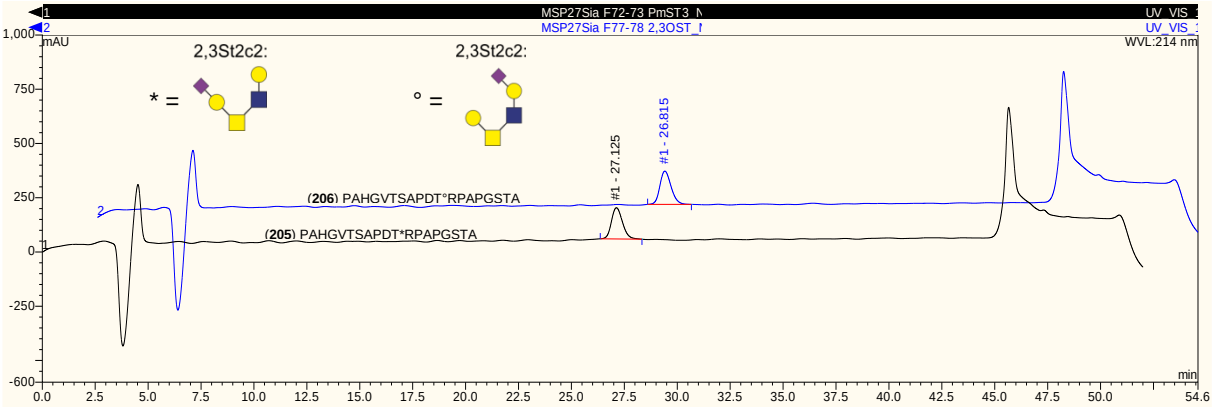
9.2.3.5 Sialylated core 2 type-1 glycopeptides



9.2.3.6 Sialylated core 2 type-2 hexasaccharide glycopeptides



9.2.3.7 Sialylated core 2 type-2 tetrasaccharide glycopeptides



9.3 Microarray data

9.3.1 Microarray format 1 (MA1)

dilutions	CHSynB mouse1		CHSynB mouse2		CHSynB mouse3	
	MEAN	SD	MEAN	SD	MEAN	SD
1/50	54946	3310				
1/100	53555	4396				
1/200	51492	6835				
1/400	40443	4856				
1/800	36667	1125				
1/1600	26935	1258				
1/3200	16533	1234				
1/6400	6657	393				
1/12800	2088	228				
1/25600	405	45				
1/51200	42	29				
1/2.5			43349	1473	42988	1744
1/5			38635	955	43571	934
1/10			33123	990	40318	1125
1/20			23765	664	30123	479
1/40			14461	349	21954	1445
1/80			7101	341	11775	708
1/160			2398	154	5780	164
1/320			793	59	1770	120
1/640			225	56	418	59
1/1280			13	18	90	47
1/2560					7	9
	HC12 mouse 1		HC12 mouse 2		HC12 mouse 3	
1/2.5	64563	2516	49077	2142	63379	3050
1/5	66021	1736	43833	1444	66021	1736
1/10	65771	3095	43975	954	50004	1244
1/20	66284	3071	37130	973	32381	2388
1/40	60681	1472	25189	997	27774	447
1/80	51283	2650	14614	462	11048	894
1/160	26617	715	7508	658	4561	317
1/320	12247	664	2684	177	1447	147
1/640	2589	200	566	65		
1/1280	137	33				
	HC11 mouse 1		HC11 mouse 2		HC11 mouse 3	
1/2.5	19840	1509	51786	1310	78344	7167
1/5	16338	698	51147	2626	70605	5704
1/10	8226	294	52594	8161	54613	3986
1/20	2988	92	49326	4533	30509	1261
1/40	1251	83	46371	2668	12502	525
1/80	424	42	31703	949	3910	186
1/160	163	29	19683	1188	1244	67
1/320	44	21	9051	245	315	21
1/640			3540	113	145	30
1/1280			1176	56	7	10
1/2560			255	29		
1/5120			55	21		

dilutions	SH127 mouse 1		SH127 mouse 2		SH127 mouse 3	
	MEAN	SD	MEAN	SD	MEAN	SD
1/200	62590	5033	11584	418	110210	2384
1/400	59026	1334	5723	222	110609	1827
1/800	58486	1001	2117	243	85695	2251
1/1600	57341	1196	302	109	54822	2209
1/3200	50479	1217			35390	1341
1/6400	43329	426			19952	413
1/12800	37002	505			10754	635
1/25600	19144	750			3439	249
1/51200	7800	389			583	74
1/102400	1913	144			45	25
1/204800	134	121				
	AuNP mouse 1		AuNP mouse 2		AuNP mouse 3	
1/1	21494	619	64996	4537	103300	5189
1/5	7099	399	53141	1939	91688	3829
1/10	2568	89	26825	1411	79229	3538
1/20	684	108	11868	420	54220	4860
1/40	80	36	4570	265	29669	1997
1/80	38	12	1544	138	14841	521
1/160	20	19	7	8	6674	377
1/320	6	8	319	118	2267	270
1/40	4	11	62	16	827	86
1/1280			8	11	132	32
1/2560			6	11	58	24
1/5120			4	8	26	27
	HC1 mouse1		HC1 mouse2		HC1 mouse3	
1/2.5	63527	5046	54111	2154	50186	3917
1/5	62061	2206	47775	1067	40651	731
1/10	50415	1463	49493	1610	44398	1059
1/20	46051	2159	33947	619	43962	1607
1/40	38956	1901	27338	1554	31643	1434
1/80	25783	1341	12514	447	21719	1111
1/160	12102	896	5187	154	10735	252
1/320	5732	329	1266	37	5295	236
1/640	2317	122	211	45	2155	108
1/1280	1058	165			1032	73
1/2560	21	15			96	35
	HC2 mouse 1		HC2 mouse 2		HC2 mouse 3	
1/2.5	9654	422	15349	360	7498	206
1/5	8013	438	13160	236	5283	180
1/10	5924	315	10729	386	3557	196
1/20	3043	90	8176	360	1540	88
1/40	1523	103	4699	219	523	66
1/80	423	37	2556	156	191	82
1/160	128	71	933	113		
1/320			382	57		

9.3.2 Microarray format 2 (MA2)

Vaccine candidate 1 (CHSynB):

Peptides	mouse1 (1/500)		mouse1 (1/2500)		mouse1 (1/10000)		mouse2 (1/10)		mouse2 (1/50)		mouse2 (1/200)		mouse3 (1/10)		mouse3 (1/50)		mouse3 (1/200)	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
231	6010	1416	842	108	33	34	1387	484	125	57	14	6	15	17	1	2	4	7
230	11369	3288	1875	339	66	50	2162	743	220	120	9	10	41	48	2	4	7	7
229	8767	2074	1546	221	2	4	1924	88	179	28	4	5	21	12	2	4	1	2
130	2	4	2	4	10	8	6	8	0	0	1	2	324	239	2	4	1	1
129	25	33	7	9	4	4	13	24	1	1	1	2	4641	1213	2141	646	420	91
128	7718	2380	963	230	30	28	16	13	3	4	1	2	7	11	1	2	4	6
127	10	12	2	4	2	4	10	17	0	0	1	2	1107	501	207	222	118	67
126	15996	3144	3707	321	181	83	1646	539	290	51	7	7	26	23	1	3	0	0
125	449	246	5	9	5	8	585	153	0	0	2	3	10493	1793	4351	410	1574	388
124	4627	1108	421	69	2	3	62	53	11	11	2	2	2	3	3	4	4	9
123	3	4	7	7	2	2	0	0	0	0	4	5	127	81	0	1	16	11
122	11	13	2	2	5	6	102	99	0	1	0	0	2321	150	699	490	44	58
121	1909	252	27	31	0	0	2	4	0	0	0	0	6	6	0	0	3	3
120	3	3	3	6	0	0	9	15	17	2	3	4	195	126	0	0	1	2
119	16138	2989	3230	393	51	57	1769	312	151	30	1	2	61	46	1	2	2	3
118	42	30	1	2	4	6	1583	246	57	54	1	2	6325	691	2124	430	500	195
117	1038	270	32	39	8	8	6	10	0	0	3	4	14	20	7	7	0	0
113	1835	177	207	192	13	13	599	150	59	14	4	5	14137	2481	6412	539	2630	409
112	6427	1603	853	208	38	41	5447	1080	558	125	10	11	28876	2971	13074	1292	6403	907
111	4044	551	358	75	22	31	4	6	2	2	1	2	15	26	1	2	2	4
110	1052	226	12	15	0	0	1408	600	138	59	0	1	13287	3574	5090	794	1301	576
109	15335	3620	2585	512	113	42	2084	660	173	103	2	4	49	44	8	10	0	0
108	6378	826	733	161	21	36	8416	1375	1443	240	20	22	33052	3498	13834	1183	6165	406
107	2801	590	191	52	5	12	131	37	6	9	1	2	17	22	2	2	8	9
106	2109	433	148	109	7	11	1168	230	34	30	5	4	12375	1296	7915	532	2781	547
105	7120	1379	951	239	14	22	3802	982	552	113	32	14	21963	2167	13186	3696	5915	618
104	5109	1383	619	193	5	9	18	14	3	4	2	2	17	22	0	0	4	6
103	1427	232	89	84	4	5	1982	385	134	51	1	2	14304	1646	6385	748	2210	183
102	16976	1144	3096	310	75	64	1931	242	202	122	5	6	134	69	1	1	1	2
101	5755	1393	657	208	16	22	7318	1866	795	438	19	17	29024	4648	12716	1696	5198	370
100	3282	710	205	89	16	17	242	106	8	14	1	1	20	28	6	8	2	3
98	2794	477	240	44	3	5	2535	397	310	71	7	7	15090	1669	8031	510	2764	277
97	7863	1123	1123	151	8	15	7083	769	684	83	35	17	19981	2369	12812	818	6530	486
96	6057	1371	670	235	0	1	59	52	4	5	3	4	3	5	9	7	3	4
95	2137	593	98	80	12	14	2501	550	489	119	2	4	12303	1606	7403	582	3131	271
94	14767	1690	2723	233	30	39	2437	434	261	118	0	0	452	113	1	2	0	0
93	6995	677	927	122	14	30	8499	1328	1396	334	21	15	26195	2553	13519	639	5643	588
92	3293	1057	293	109	17	22	771	326	135	94	1	2	15	26	0	1	3	5
91	3974	668	238	153	4	9	2933	388	315	77	5	9	18856	2492	8406	934	3592	238
90	7894	1014	1308	122	16	17	6030	811	919	94	34	29	21028	2451	13039	829	6817	477
89	5915	1328	861	190	23	16	188	62	7	8	0	0	21	23	3	4	0	0
88	3308	842	310	79	8	11	3231	500	406	42	16	13	14385	1548	8525	836	3739	393
87	18379	2837	4480	463	88	75	3400	776	308	59	14	12	423	205	0	0	1	2
86	7820	964	1142	128	14	34	10067	1451	1629	175	17	37	34835	2865	17083	1241	7097	513
85	3216	559	296	71	20	30	555	262	60	64	3	6	13	18	1	2	1	2
81	4391	952	430	161	2	3	4405	1026	561	227	1	2	18446	2709	8953	823	4113	657
80	8638	1525	1262	219	16	18	8070	554	1503	227	38	33	22901	3166	12935	1505	7092	656
79	6374	1028	511	87	4	10	49	29	20	5	0	0	16	18	2	4	1	3
78	3129	471	160	81	0	0	4027	645	445	55	2	6	13150	1623	8024	625	2837	228
77	17397	2514	3135	677	11	20	2840	544	431	54	8	8	646	322	0	0	0	0
76	8554	1688	1353	294	0	0	10346	2211	1239	197	68	34	28294	2854	12731	1697	6036	655
75	1928	313	168	28	6	7	251	126	2	3	1	2	8	13	5	5	3	4
235	10150	786	1866	152	50	59	517	74	14	17	2	4	1094	363	0	0	0	0
234	7632	877	1256	119	3	8	4598	802	408	54	2	4	18117	2034	9722	880	5465	607
228	3051	684	343	41	1	3	3549	526	369	157	9	14	9	7	3	8	2	4
227	9076	1495	1553	314	11	20	1293	153	106	32	10	12	137	52	0	0	17	5
226	10731	1846	1270	360	12	21	3165	547	189	47	21	8	223	116	1	3	0	1
225	4834	994	373	131	51	36	2030	455	207	42	2	4	15	16	5	7	3	8
224	5104	489	528	81	0	0	5080	761	615	163	4	4	12681	1579	8922	504	4438	537
223	2281	377	170	89	3	4	1030	176	23	15	1	3	13	14	1	2	0	0
222	3456	874	283	68	17	12	478	154	19	11	4	4	16	17	3	5	0	0
74	28170	1568	10374	1098	2551	260	4644	707	588	191	11	15	13788	1913	6683	713	2215	444
73	31497	3484	11156	1603	2924	738	6198	1447	685	288	0	0	16760	2936	8984	993	4794	556
72	6916	1249	847	220	8	20	355	133	6	9	4	5	13	13	3	6	13	12
71	30311	2240	9646	830	2410	276	4000	609	436	99	21	24	12379	2043	7260	431	2353	355
70	11487	1649	2126	196	11	22	2328	465	363	46	0	0	94	74	4	6	1	2
69	29778	4701	10354	1156	2810	214	5628	857	598	47	7	14	15223	2730	7373	420	3883	513
68	3797	850	266	113	39	35	156	76	2	4	1	3	9	6	5	8	4	5
233	57695	6613	27832	4021	8541	947	24507	2227	15729	1063	4329	718	23589	3268	8603	671	3543	652
232	124412	5754	55900	2607	13999	618	27929	2653	18701	1646	6538	633	45354	5204	22225	1117	11450	377
221	4762	832	594	57	50	52	1783	343	163	46	16	12	32	28	0	0	0	0
220	45824	5949	15433	2753	4285	359	8581	1821	966	403	79	31	22310	3606	12588	614	6638	453
219	10912	2427	2493	444	42	47	2359	411	328	137	4	7	668	147	0	0	9	9
218	9828	1437	1591	309	0	0	1842	746	198	66	3	5	152	104	0	0	6	6
217	3621	784	359	72	16	38	315	162	0	0	6	8	6	7	5	6	4	4
216	10447	1170	1652	245	80	62	3226	662	385	53	16	16	27	32	6	8	2	4
215	26176	2031	10181	694	2299	617	5988	914	770	160	23	20	13231	1580	9296	1068	3943	735
214	347	252	4	6	3	6	28	23	2	3	2	4	6	7	2	3	3	6
213	2	2	0	0	2	4	0	0	1	2	2	4	12	14	1	2	7	8
212	1232	418	148	82	14	14	1466	290	194	7	12	9	16	13	2	3	11	11
211	3212	577	224	118	5	8	1665	218	177	145	13	10	8	8	0	1	4	4
210	11650	1846	1989	300	11	19	4071	963	352	164	20	19	192	102	10	11	6	7
209	18990	2697	5887	1203	1303	789	5209	1256	465	103	11	9	16848	2706	8841	1103	4244	648
208	2490	370	170	52	0	0	724	146	120	43	15	11	14	19	0	0	3	6
207	10539	2774	1616	375	7	14	4030	1015	457	71	22	15	238	167	0	0	9	12

Vaccine candidate 2 (HC12):

Peptides	mouse1 (1/80)		mouse2 (1/20)		mouse2 (1/80)		mouse2 (1/320)		mouse3 (1/20)		mouse3 (1/80)		mouse3 (1/320)	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
231	285	295	4140	1061	957	292	280	148	61	72	74	30	11	14
230	515	249	7671	1798	2352	296	553	97	209	106	3	7	6	16
229	517	194	4513	888	1209	450	246	133	400	144	69	74	20	30
130	79	103	509	191	17	26	16	40	12	20	0	0	9	23
129	41	41	5360	1114	2821	1133	747	158	18	0	115	55	37	42
128	32	48	1911	672	442	159	200	65	50	91	84	18	7	17
127	52	60	1347	436	400	81	0	0	81	73	44	65	65	68
126	87	112	5278	1136	1627	352	337	119	139	109	5	7	6	15
125	190	147	21159	3045	9723	1637	3162	627	110	145	6	12	40	47
124	5	12	1830	590	287	161	8	19	#DIV/0!	#DIV/0!	11	11	44	60
123	77	88	111	144	20	38	38	54	0	0	2	6	0	0
122	27	40	5250	447	1716	171	428	143	39	53	8	16	50	54
121	71	77	520	179	68	84	1	2	71	93	17	27	25	45
120	38	62	335	85	90	94	10	26	13	25	0	0	0	0
119	32	42	5007	994	1440	353	401	113	16	21	9	15	30	52
118	45	50	11781	1216	3979	666	913	192	60	71	0	0	45	58
117	18	#DIV/0!	666	195	80	82	0	0	38	67	20	33	59	61
113	5	7	15563	2886	5564	1386	1978	413	10117	2317	2312	430	534	85
112	182	142	39780	4491	21732	4387	7805	1605	23033	4963	5674	1330	1290	188
111	28	50	1152	393	148	112	41	43	39	74	2	6	0	0
110	73	112	10984	2613	4899	1441	1749	687	9252	2565	2731	144	437	108
109	9	0	4842	962	1269	375	238	173	2	4	53	72	2	5
108	216	76	49570	5503	32235	6459	12460	2592	33136	7091	7156	1371	1690	378
107	24	32	1297	301	191	78	17	44	29	67	41	41	14	19
106	103	87	19665	2156	8241	1217	2643	403	12026	2473	2671	569	496	115
105	283	127	43452	4252	27105	5753	9525	2362	26141	5597	7650	1301	1837	316
104	21	38	924	417	18	52	14	33	0	0	2	5	24	41
103	15	#DIV/0!	24290	3118	11334	1237	3797	405	13973	2304	3240	548	660	113
102	67	78	4289	717	1140	250	142	127	36	79	8	19	20	28
101	256	135	51469	6287	31766	6502	12296	1443	31948	7253	6470	417	1842	140
100	56	71	1421	377	34	53	31	51	0	0	0	0	27	51
98	3	6	29085	900	17214	2901	6190	825	19430	3939	4545	669	1161	324
97	305	131	44296	3214	31058	6052	11889	2466	28230	6415	6582	1320	1836	211
96	48	72	1776	390	226	127	15	22	0	0	17	26	64	58
95	131	123	30585	3740	16491	3355	6120	1195	14640	2927	3749	447	776	106
94	35	61	6812	903	2085	381	463	90	0	0	0	0	6	14
93	255	167	47111	3694	32655	5930	11927	1898	29076	5130	7255	1651	1730	227
92	76	81	1265	657	146	169	8	14	0	0	7	13	10	25
91	458	142	32310	4165	18388	3088	6749	790	17814	3450	3962	262	1066	178
90	780	170	45941	3114	34396	6338	13343	2303	29378	6202	7449	1461	1832	517
89	254	164	2435	500	496	153	59	84	718	278	83	59	3	7
88	212	191	31226	2905	17738	3109	6215	1219	14406	3272	3347	691	635	130
87	318	107	6697	1127	2070	410	582	150	445	240	53	73	0	0
86	519	116	57473	5293	38106	5438	15094	2569	30977	5280	7845	1335	1507	262
85	87	78	1364	379	51	72	4	9	0	0	0	0	0	0
81	75	56	23830	3812	13930	2806	5346	637	18137	5275	3809	274	1068	156
80	374	242	40973	4364	30298	6147	12106	1943	30103	7526	6966	517	1711	311
79	81	92	1351	579	151	150	18	29	0	0	0	0	0	0
78	14	16	22916	1042	13990	1328	4823	553	18844	4578	4389	646	875	281
77	31	31	5897	1002	1703	455	487	116	0	0	0	0	6	17
76	366	212	44528	3635	30909	5700	11396	2212	33250	9052	8340	2048	1727	332
75	23	25	853	214	154	229	20	32	0	0	0	0	1	3
235	36	0	6263	817	1704	350	461	110	66	90	6	13	20	35
234	277	255	36034	3457	27116	5715	10297	1630	21610	4818	6124	724	1121	190
228	0	0	0	0	0	0	0	0	0	0	0	0	0	0
227	86	97	4896	681	1504	346	488	104	33	55	0	0	8	20
226	82	93	5245	628	1556	333	390	132	0	0	0	0	0	0
225	45	57	2905	858	625	339	82	101	23	44	33	53	13	23
224	268	159	39462	4672	29860	6533	11219	2750	28321	5469	8419	1145	1657	327
223	59	68	1193	427	66	92	26	47	0	0	0	0	7	13
222	0	0	1934	693	349	155	5	9	0	0	0	0	1	2
74	482	162	35506	4509	22892	3290	9119	1453	48502	8424	14789	2817	2665	394
73	0	0	48848	7328	36080	5806	17037	1835	71673	17470	19680	4186	4951	723
72	0	0	0	0	0	0	0	0	0	0	0	0	0	0
71	25	42	44014	5145	28797	5009	12019	1387	69022	13907	20336	2730	4190	733
70	401	193	3459	553	780	287	190	145	0	0	0	0	0	0
69	74	68	49615	5734	34132	6039	13678	2278	75144	15585	25984	4313	4639	1108
68	62	67	1172	213	162	149	35	51	8	22	0	0	0	0
233	80240	9915	9613	408	2905	1050	1127	247	55347	10674	41623	1902	12489	1193
232	99975	8144	73219	14051	52716	10248	25362	3627	174478	19804	83163	7569	24558	2462
221	0	0	214	313	7	18	13	35	0	0	0	0	0	0
220	777	295	61437	9992	45267	8929	21971	4495	96356	20976	32171	4204	7197	1363
219	25	43	5620	1048	1807	459	451	71	0	0	0	0	0	0
218	75	78	6826	1182	2302	360	586	53	140	155	18	24	18	44
217	49	68	2556	641	519	382	51	72	0	0	64	86	4	6
216	591	256	5063	957	1155	499	263	149	0	0	0	0	25	39
215	1007	232	35987	5917	23597	4716	9671	1115	51160	12135	18211	2618	3939	779
214	149	127	552	281	74	103	40	62	17	38	0	0	2	6
213	49	118	9	15	8	11	145	109	8	23	17	23	51	70
212	28	45	295	207	18	51	0	0	0	0	0	0	0	0
211	45	93	1970	462	417	155	123	82	73	94	0	0	36	50
210	99	107	5857	1073	1607	457	349	98	22	39	0	0	14	35
209	419	187	25605	4027	12394	3054	4894	596	15751	1904	7731	772	1603	520
208	58	94	1865	474	411	234	51	46	89	104	16	19	0	0
207	76	89	4905	1200	1273	360	434	79	0	0	0	0	27	32

Vaccine candidate 3 (HC11):

Peptides	mouse1 (1/10)		mouse1 (1/40)		mouse2 (1/80)		mouse2 (1/640)		mouse3 (1/20)		mouse3 (1/80)	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
(231)	13	13	3	4	29194	4645	7165	747	4	8	1	2
(230)	32	26	12	14	32993	7474	9463	1283	40	25	4	5
(229)	2018	750	316	218	36824	19343	13071	3662	16	16	5	5
(130)	9	11	11	13	0	1	6	6	12	15	5	8
(129)	12	13	10	10	6	10	6	12	13	11	1	1
(128)	12	12	4	6	13134	2147	2128	653	7	11	7	9
(127)	4	9	8	10	9	12	7	15	8	9	3	6
(126)	1553	415	187	117	15160	1880	4016	913	11	15	4	6
(125)	23	17	5	7	11	17	5	11	13	16	6	6
(124)	8	12	8	12	19320	4034	4059	582	11	18	12	9
(123)	2	3	3	6	2	2	6	11	9	9	8	8
(122)	13	13	2	4	4	8	4	7	8	10	2	5
(121)	10	10	7	10	9441	1601	1489	528	4	9	0	0
(120)	1	2	8	8	9	13	2	6	5	7	0	0
(119)	2190	499	295	128	17694	1507	4788	1306	7	10	0	1
(118)	22	22	4	4	10	11	7	7	1	3	2	4
(117)	16	13	7	9	26021	4143	5191	648	9	9	0	0
(113)	12	14	7	9	5	10	12	11	7	10	5	5
(112)	10	14	11	14	5	12	7	7	12	11	0	1
(111)	3	6	3	4	15546	4804	2915	1073	7	12	2	3
(110)	10	15	3	5	6	10	9	11	10	9	10	11
(109)	591	343	59	39	26147	3666	5852	1399	11	14	14	7
(108)	40	35	10	15	21	47	4	10	8	6	10	13
(107)	8	13	5	6	31235	2675	7071	1293	8	13	5	6
(106)	7	12	1	2	8	15	5	6	15	17	1	3
(105)	9	13	6	8	8	17	5	13	9	9	2	4
(104)	16	16	6	5	14706	3675	3025	977	3	6	0	0
(103)	3	5	5	6	8	16	5	7	2	3	1	3
(102)	855	447	102	97	22349	2730	4795	1200	9	16	4	5
(101)	22	30	1	2	4	5	6	12	10	14	6	10
(100)	61	153	11	14	37276	2080	7776	1929	11	18	2	3
(98)	7	10	5	8	11	14	9	13	7	13	1	3
(97)	36	50	5	13	2	4	8	8	8	10	1	2
(96)	12	15	5	8	15875	2810	4210	765	5	7	7	10
(95)	11	9	6	11	4	8	8	11	7	9	3	4
(94)	2762	527	574	149	19597	3471	5783	1411	13	15	2	5
(93)	18	14	3	6	6	9	5	13	5	7	0	0
(92)	14	24	9	14	25536	3944	5691	1423	5	9	2	3
(91)	264	144	48	29	16	19	9	6	15	11	14	13
(90)	686	173	221	405	147	387	3	9	6	9	8	10
(89)	298	367	22	29	17044	1411	4284	623	8	8	1	1
(88)	8	9	14	17	15	19	6	7	15	14	0	1
(87)	3349	775	678	204	21710	3698	5474	1210	6	11	0	0
(86)	272	138	52	32	8	14	7	12	13	15	5	3
(85)	6	8	1	4	25502	3963	6232	1195	1	3	3	4
(81)	8	11	10	15	20	25	3	7	9	14	1	2
(80)	669	300	95	39	13	15	5	6	13	15	0	0
(79)	14	11	8	12	18230	1709	4841	513	5	8	1	2
(78)	11	9	8	12	2	5	5	5	5	9	5	7
(77)	2627	569	460	144	21699	3079	6130	1118	8	6	2	3
(76)	529	236	53	43	55	130	17	24	7	9	2	3
(75)	1	4	3	4	21663	3450	5413	1088	4	7	8	7
(235)	6376	689	1638	337	28699	5535	8133	1225	65	48	102	77
(234)	3017	329	474	136	1	3	7	8	6	9	122	30
(228)	5	7	13	12	3794	715	562	255	6	11	144	81
(227)	2375	336	379	114	19610	2065	5988	788	26	21	1	2
(226)	2931	546	493	112	18026	2790	5787	968	15	20	87	42
(225)	2	3	17	22	13846	2530	5222	771	9	15	1	1
(224)	545	127	9	14	2	5	1	2	1	3	4	4
(223)	5	8	9	12	3715	1409	600	329	10	9	15	4
(222)	10	14	9	12	20001	7630	5027	1087	11	12	9	11
(74)	111	89	13	13	10	13	5	6	461	168	391	60
(73)	1216	392	192	84	20	15	6	11	369	172	478	59
(72)	333	181	80	36	17446	3069	5134	578	522	230	6	5
(71)	9	13	8	15	13	14	4	7	0	0	10	11
(70)	2217	473	373	135	21623	3321	5458	889	505	77	7	9
(69)	591	287	81	62	29	46	8	11	1	3	5	6
(68)	14	18	11	12	16095	1789	4432	651	8	13	4	4
(233)	5535	1107	1455	392	39520	13835	12917	2944	1180	507	1	2
(232)	5189	861	1376	355	12	19	10	19	1658	428	2	4
(221)	20	23	5	6	13758	1392	5112	653	406	161	109	41
(220)	3164	398	510	124	48	64	3	6	793	150	215	74
(219)	3399	547	533	183	42443	3655	11112	2243	542	158	161	32
(218)	5	7	5	6	16928	3291	5003	897	10	8	6	7
(217)	9	12	12	13	16423	2750	4802	683	2	3	5	7
(216)	370	184	60	62	28335	4874	7721	900	474	140	110	36
(215)	391	206	97	46	6	9	4	7	606	200	73	72
(214)	81	54	18	21	8622	3662	2718	560	279	72	87	6
(213)	4	7	11	12	3	5	0	0	3	4	3	5
(212)	2	4	3	7	3207	469	375	294	4	5	6	5
(211)	12	13	4	6	16153	1251	4479	625	9	11	3	5
(210)	3584	920	675	241	20917	3603	6301	773	9	10	2	3
(209)	1486	503	232	64	56	107	11	14	8	12	2	4
(208)	9	12	3	7	17259	1349	5224	621	14	17	9	8
(207)	2453	501	434	126	21645	1729	6190	884	13	10	4	6

Vaccine candidate 4 (NG5):

Peptides	NG5 mouse 5 (1/100)	
	MEAN	SD
231	5130	977
230	6400	1155
229	4846	2365
130	572	162
129	1873	387
128	1968	677
127	1423	389
126	4599	649
125	3394	466
124	3725	844
123	467	217
122	1778	503
121	1410	222
120	830	328
119	4167	877
118	4003	726
117	3410	529
113	3863	702
112	4520	1495
111	2606	634
110	3903	947
109	4491	1611
108	6081	1414
107	4717	943
106	3515	592
105	4809	685
104	2726	793
103	4547	1231
102	4602	1104
101	4849	1708
100	4407	946
98	4032	778
97	5541	829
96	3435	383
95	4477	1248
94	4779	1078
93	5601	1268
92	3420	1042
91	3270	807
90	5077	671
89	3415	404
88	4842	715
87	4454	1181
86	6373	1458
85	3854	804
81	3859	985
80	5305	1316
79	3002	679
78	5131	997
77	5175	711
76	6150	1645
75	3629	805
235	3256	610
234	3431	700
228	5202	1237
227	3412	552
226	4523	671
225	3975	962
224	3318	2612
223	4708	1052
222	4125	1176
74	4570	1031
73	4570	1724
72	2698	636
71	5046	1019
70	4310	694
69	5823	856
68	2863	854
233	3283	996
232	4967	1315
221	3685	777
220	5925	1089
219	4413	783
218	4811	876
217	3528	857
216	4822	981
215	5156	980
214	2024	526
213	30	59
212	4873	774
211	4065	495
210	5902	896
209	4776	1303
208	4043	864
207	5691	1386

Vaccine candidate 5 (SH127):

Peptides	mouse1 (1/4000)		mouse1 (1/12000)		mouse1 (1/50000)		mouse2 (1/400)		mouse2 (1/800)		mouse3 (1/2000)		mouse3 (1/4000)		mouse3 (1/12000)	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
231	35080	7913	29892	5585	4686	845	10046	3619	8281	2592	45083	10129	28889	5921	10733	2479
230	45055	17929	41075	8165	6659	1043	13027	4900	10408	2733	83811	50026	43405	9376	16840	4344
229	41988	11650	35275	9803	6825	649	12194	2871	10948	3057	52428	10408	35236	7172	14191	3459
130	5130	1342	2235	884	624	598	1040	255	2465	735	3603	1277	1740	579	331	225
129	16992	3011	9909	1659	2954	723	2517	683	3737	1034	14137	2509	8253	1374	3022	513
128	16955	3311	12877	3883	2973	794	4147	1210	4948	771	17350	2023	11071	2145	4262	852
127	11009	1376	8180	2055	1859	124	2689	854	4030	906	8783	1562	5516	1037	2107	502
126	27194	5630	23147	6028	4918	708	9377	2898	6940	1751	36601	4902	23837	3876	9810	2205
125	22840	3954	16773	2865	4081	598	5720	1990	5507	1199	27329	3330	17561	2211	6558	1163
124	27266	10875	23949	3887	4186	848	8615	2994	6450	1361	30156	6135	20574	3620	8137	1931
123	3788	1688	1189	766	26	51	788	190	1408	890	2349	855	885	609	169	217
122	16690	3501	11157	1146	3106	300	2601	536	3892	734	15358	1479	9514	1089	3562	537
121	16885	3146	11921	1379	2696	301	3800	844	5070	765	14929	2865	9792	1568	3742	591
120	12114	2531	7644	1817	742	539	1986	504	2840	465	8908	1977	5511	845	1863	306
119	30382	3752	26965	4071	4862	555	9769	2725	8028	1474	38027	4822	26288	2607	10966	1827
118	34750	8355	24829	3694	5547	833	7645	2506	8605	1955	48064	5796	27345	3157	10370	1694
117	36034	6587	26370	4852	4803	413	9451	2625	10013	1951	42352	4163	27699	3851	9377	2591
113	23729	9471	20527	2991	4190	525	7341	1931	6263	1099	29739	3210	18709	1987	7280	1432
112	34128	9960	27667	7819	5630	728	9043	3177	8963	1626	41318	5064	25451	3843	10241	1852
111	30236	5959	22073	2990	3221	367	7238	1742	6597	1544	23288	2698	14970	1653	6095	640
110	31662	6262	25075	4526	5288	619	9055	2249	8363	1295	36156	5434	23493	3137	9314	1426
109	37920	8264	33350	7468	5559	978	10507	3502	8069	1695	49456	6837	33226	5189	13566	1559
108	35924	7497	38868	5785	8035	886	12429	3789	9961	1408	60325	8062	40097	5251	16723	2751
107	39545	10548	37802	5854	5605	398	12490	3724	10401	2609	50845	4472	32518	3702	12488	1961
106	23108	4688	17190	4283	3943	502	7111	1829	5572	1285	28970	2680	18351	1693	7114	1055
105	33531	13468	30425	5221	5824	764	10306	2679	6844	1189	44642	5943	30145	4292	11875	2164
104	25641	6933	17290	7827	4113	650	6109	3045	6559	1295	24653	4226	13523	6078	5383	1369
103	36422	6205	30001	4771	5370	820	10568	2118	8120	1617	50429	3172	32375	4068	12116	1489
102	32106	6799	28107	5039	5917	835	9392	2547	7498	1670	42418	5241	26132	4481	10931	1933
101	36891	8751	33942	7130	6747	1108	11432	3912	8453	1363	60605	8933	40366	5764	16529	1933
100	40439	9748	36606	6583	4593	545	11427	4160	8058	1127	48084	6157	32863	4574	13703	2532
98	21324	4414	18868	3207	4937	707	7741	2095	6141	1062	29496	3328	19903	2604	8298	1459
97	30872	5506	28499	4073	4476	561	10699	2492	5624	1083	42683	3973	29097	3110	11652	1830
96	24385	3906	19644	3153	3864	438	7266	1771	5651	819	21626	3142	15524	2645	6460	1120
95	29850	6054	24022	4934	5229	574	9655	2755	5594	694	41569	6310	28682	5230	11441	1943
94	30415	4847	26773	5469	4124	805	10019	2699	6451	1466	41337	3533	28153	4183	11160	1663
93	31940	6982	33943	4382	5360	955	12615	3422	8448	1186	52420	6417	34116	4690	14753	2353
92	31994	7504	28199	7078	4177	515	9885	3793	6697	1835	40059	6537	26983	4604	12176	2531
91	20003	2736	17592	4276	4688	850	6933	2839	5732	1008	24227	3625	18099	2436	8013	1354
90	33516	4514	28720	5891	4266	712	10858	2800	5448	926	45419	5727	30699	3921	12288	2301
89	23392	3816	20262	3534	4096	867	7858	1437	5289	951	26962	3325	18080	2461	7082	1105
88	28329	4364	24050	4285	5126	832	10896	1725	6229	1049	41709	6788	27827	4332	11562	1794
87	29482	4038	24508	5683	4729	585	9913	3074	5745	798	41391	7188	27777	5111	11799	2247
86	37796	6087	36351	5711	7216	1018	12175	4363	8459	942	68600	5297	41511	5932	17737	3682
85	31940	6396	31205	5627	4156	316	10687	2925	6484	772	44241	7522	30095	5243	12850	2278
81	25776	3778	22152	4340	4874	585	8536	2630	5533	899	37798	4899	24543	3986	10880	2421
80	35760	3795	32133	6842	5235	649	11589	3525	7906	2227	51638	6383	35793	5686	16176	3517
79	25951	3607	19275	4458	3983	521	7339	1754	4582	650	26216	3388	16574	2770	7204	1796
78	31059	4409	27205	4867	4998	594	11208	2457	6080	1546	47391	5454	32410	5577	13993	3189
77	34530	5775	30482	4118	4461	709	12271	1493	6369	1782	47363	8398	34073	6927	14645	3579
76	40739	8142	36737	5535	5911	1739	13078	4587	7817	1580	68936	10651	45091	8177	19366	4802
75	28428	5803	21959	5514	3385	612	7564	2442	3996	880	35472	4229	24161	4432	10032	2292
235	26268	5083	18800	3746	3783	426	6954	2171	4480	943	26852	2996	19179	3749	7452	1531
234	18104	4175	15757	3075	3756	798	6047	1389	5045	1270	22047	2732	15765	2540	6623	1557
228	24843	5574	21009	4043	3426	290	10179	2168	4938	1492	37705	4861	29076	5462	12447	2480
227	20217	3468	17532	2894	4029	430	7610	1218	4291	562	24639	3010	19141	3700	7577	1776
226	25255	4869	21902	3876	4734	559	10162	1606	5552	1029	33785	5459	27030	5981	10847	2678
225	27641	6240	23953	4939	4000	469	9749	3498	5133	1286	41629	5673	30378	7354	12217	2509
224	23556	5109	21369	4188	3919	323	7680	3230	4793	1043	36640	3836	24136	4174	11419	5486
223	31943	8275	25258	5354	3680	344	10833	2533	5196	1107	28483	7084	24100	5576	11604	2934
222	35714	7341	27647	5109	4256	1021	10647	2216	5562	1033	43982	3827	30242	6199	12139	2527
74	23300	3780	21581	3802	4781	443	8866	2253	5670	1176	34180	3905	23786	2813	10998	2318
73	26291	5041	26892	6630	4204	289	9691	2750	5937	1315	45004	6087	32081	6232	14013	2982
72	15872	3275	14331	4326	3641	452	5861	1943	4195	672	21323	2393	14522	3653	6626	2124
71	29456	4374	25749	5010	4464	496	10218	2184	4679	858	45604	5087	31304	6733	13812	3400
70	28786	4245	26242	5569	4152	650	10195	1877	5458	1346	40422	4367	29007	6421	12678	2848
69	28788	3221	27512	5862	4602	1083	10755	1904	6615	1794	44228	5797	32327	7246	14175	3228
68	20100	4437	16687	5105	3156	438	6579	2487	4272	793	23446	3869	18082	4908	7122	2197
233	62741	11405	50046	9740	13449	2027	12148	2658	9153	2130	47069	6057	34425	7015	15581	4078
232	96702	9112	75487	13115	18032	4031	16823	3374	10451	2851	77477	7961	52849	7689	23513	4841
221	11561	3245	11271	2173	4287	516	8295	1464	4714	866	22597	3426	16494	3065	7498	1768
220	32181	8393	28696	3350	4678	556	12886	2085	6872	1669	46245	3353	33020	4381	13007	2603
219	35886	8709	28528	2459	4197	537	11538	1518	5330	1317	44659	4620	31190	5075	12440	1757
218	29054	6691	23530	3142	3770	371	10879	1336	4617	1077	41445	6301	29938	6196	11723	1948
217	22873	5701	19828	3992	3935	410	7130	2616	4330	520	31373	4702	21712	5017	8088	1547
216	32973	8843	30484	4022	5684	733	12996	3439	6572	749	44881	5359	28870	3668	13652	2555
215	23810	5851	20850	4689	4692	530	10537	1463	5619	867	38740	3041	27034	5906	11183	1934
214	11250	2289	9669	1660	2540	610	4267	494	3327	771	15945	2180	10567	1667	4295	969
213	0	0	0	0	6	13	0	0	5	8	0	0	0	0	0	0
212	27166	5994	22093	2636	3923	919	10156	2015	4965	1250	33099	2912	24612	4597	10651	1782
211	24039	5893	19678	2269	4804	1110	9742	1548	6011	855	31882	2863	22914	2951	9755	1567
210	34100	7690	28567	3430	4838	1093	13914	2180	6905	939	46952	8087	34013	4180	14307	2555
209	27499	7939	23595	5036	4614	1002	9033	2671	5422	956	43796	4964	27802	4277	10869	2021
208	28028	7993	25672	3929	3468	494	9381	2505	4347	557	36839	4093	25547	4066	10899	2137
207	33647	9920	30117	6265	3934	632	11925	2343	6515	2249	52751	5600	36013	5624	14217	3612

Vaccine candidate 6 (AuNP):

Peptides	mouse1 (1/5)		mouse2 (1/10)		mouse2 (1/20)		mouse2 (1/80)		mouse3 (1/80)		mouse3 (1/160)	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
231	4	7	23575	6331	16119	3755	4172	855	4	5	2	4
230	4	5	31990	11851	22119	4212	5149	1032	1498	610	3274	1669
229	6	7	32995	7952	22269	3802	6054	957	273	223	1383	524
130	8	11	658	307	442	359	18	29	4	7	6	4
129	9	10	4847	1181	3943	1121	690	300	6	11	1448	548
128	8	11	6297	2517	3958	1058	914	229	2	5	7	16
127	5	13	2498	822	2517	434	451	153	9	13	4	10
126	5	9	20928	3742	17620	2554	5224	683	521	278	1921	485
125	5	5	14002	3274	11200	1682	3163	532	1286	328	4108	1077
124	9	12	14971	3976	7233	1391	1529	548	2	3	11	13
123	5	8	576	314	609	322	26	24	6	11	505	600
122	20	19	7326	1409	6081	817	1495	204	9985	1580	15752	1641
121	3	5	4431	1424	3558	791	669	205	4	7	3	5
120	10	16	1926	727	1884	565	249	140	425	247	2287	761
119	9	11	19653	3229	16405	1950	4711	538	525	236	1889	401
118	18	20	24144	5237	13964	1588	3591	453	11410	1738	17837	2361
117	2	3	17212	4015	9449	1707	1623	278	2	4	4	7
113	60	85	13164	2196	11745	1414	2861	1198	6	12	0	0
112	1811	666	31609	6228	24526	3005	7527	1016	1379	853	4572	1472
111	4	6	11021	2951	5570	1602	1315	349	7	13	9	10
110	3	6	23125	4093	16672	2334	4378	897	3	5	2	4
109	13	12	27265	6332	21210	3523	5741	1071	8	17	673	454
108	1138	203	53168	8914	37834	3417	11706	1511	2425	475	6009	1011
107	2	6	27722	7070	14498	2151	2965	446	10	14	5	8
106	151	144	12523	1882	13378	1505	3264	442	2	4	2	4
105	2035	658	27027	2713	26688	2631	8245	1182	1935	777	5701	1568
104	3	5	10544	3510	6578	1208	1352	267	4	8	3	5
103	44	45	23543	3678	19868	2396	5078	688	2	5	5	6
102	6	10	27016	5186	17892	1955	5134	1092	173	178	1231	503
101	1294	527	45692	9933	34008	5106	10477	1432	2147	730	5359	1470
100	7	13	22928	6840	13700	1812	3243	563	11	10	7	12
98	443	269	16666	2301	17265	1946	5429	518	4	8	1047	506
97	2680	465	26111	2600	26693	2740	8561	759	9517	1447	14302	3043
96	4	8	13449	2884	8386	1428	2101	322	6	7	8	12
95	54	76	23057	4493	21480	3119	6549	887	3	4	1410	730
94	10	13	25458	4246	19214	2668	5826	744	1977	333	4711	1416
93	1991	465	41826	5835	35767	2948	10962	1922	9531	1186	14342	1689
92	6	7	22318	6250	11474	1965	2954	370	3	5	7	13
91	625	269	20258	3664	16203	1848	4749	571	5	13	1085	279
90	3176	337	26665	2978	26012	3137	7814	910	8510	1084	14083	1999
89	9	14	11631	2122	10009	1540	2392	230	14	25	0	0
88	139	103	21549	2915	17073	1774	4911	660	1	3	488	556
87	10	8	26690	5704	15415	2433	4838	596	1806	452	4412	1077
86	2899	472	51347	7490	38250	5085	11539	1859	9968	1708	15735	2241
85	9	10	19477	3185	13898	1817	3647	805	15	14	1	2
81	650	164	24896	4345	16235	2185	5131	1236	264	255	1707	626
80	2970	465	34276	4864	30197	3091	9302	1261	9795	1417	14934	2101
79	4	6	11891	2691	8235	1640	1742	261	15	18	8	10
78	354	152	24749	2476	22987	2669	6741	651	239	236	1902	411
77	9	15	24438	3180	19977	2611	5877	645	2066	545	4860	997
76	2761	813	53133	10514	42777	8846	13392	1634	10778	2467	15946	2789
75	4	6	17727	4148	9774	2603	2454	347	5	10	10	11
235	7	7	11572	2917	8446	1341	2092	429	1347	218	3419	586
234	2673	331	18392	2976	13175	1598	4491	760	6697	1156	10777	1339
228	370	181	23446	4212	17769	3108	5673	964	5	7	4	9
227	3	5	14892	2713	11100	1784	3468	454	1095	179	3029	619
226	6	8	21031	2948	12279	1983	4079	627	1378	413	3512	523
225	7	9	18193	4030	16997	3380	5630	1091	7	11	7	14
224	2619	488	23315	1629	25565	4256	7962	1823	7744	1736	11412	1211
223	352	239	20618	5070	13396	2500	3823	983	6	10	5	8
222	10	11	16804	4581	11561	2582	3369	769	10	12	5	11
74	17	19	24341	2703	15696	1613	4809	843	3461	652	6652	721
73	437	169	33410	6868	27513	4482	8836	1473	12679	2072	17433	1950
72	6	14	11260	3912	6003	1031	1531	293	7	10	4	6
71	16	26	26035	4482	22408	2912	7255	1073	5496	974	9738	1221
70	7	12	17724	2665	15523	1931	4783	676	506	266	1854	659
69	67	71	31572	4502	26429	3352	8110	1101	11162	1720	15877	1858
68	4	5	12149	3785	6759	1420	1650	256	3	6	200	559
233	4589	987	25340	4763	16765	3081	5623	883	6071	1075	10537	1368
232	12076	1747	35971	7291	34799	5527	10789	1454	21420	3051	29896	2440
221	1	4	12649	1812	12398	1871	4415	657	10	16	4	7
220	1426	173	41949	8916	35763	6448	11504	2090	15230	2571	20941	2750
219	1	4	25378	5615	13933	2601	4873	830	1990	456	4628	756
218	11	8	20691	4436	16384	2871	5118	1026	4797	1148	9091	1441
217	7	11	14288	4304	10341	2581	2894	546	0	0	4	9
216	11	13	28711	4666	16348	3627	5389	1436	11	12	8	13
215	12	13	25998	4116	22988	3848	8463	2722	5478	862	9567	1382
214	6	14	7153	1800	5198	1380	1493	331	3	4	3	5
213	10	11	4	7	5	7	8	12	7	7	4	6
212	317	209	18378	3943	13560	2564	4597	791	3	6	2	3
211	1	2	17113	3869	14090	2696	4634	700	7	11	3	9
210	10	11	25201	5184	23432	4501	7423	1434	2053	547	4954	1334
209	520	291	28085	7220	24431	5429	8055	1881	6731	1489	10761	1975
208	9	8	14052	3335	14111	3240	4289	1040	5	9	5	6
207	8	6	30499	8588	20103	4493	6341	1300	1345	457	3242	1052

Vaccine candidate 7 (HC1):

Peptides	mouse1 (1/80)		mouse1 (1/230)		mouse2 (1/80)		mouse2 (1/320)		mouse3 (1/80)		mouse3 (1/320)	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
231	283	119	54	35	3701	1636	1370	391	3	7	534	163
230	2952	1229	185	130	4547	1773	1699	630	50	107	1193	352
229	927	456	45	54	1356	997	349	395	5	5	298	117
130	615	319	63	62	6	9	10	13	370	192	6093	997
129	7122	1743	902	359	11	16	11	18	2386	644	12341	472
128	92	34	30	25	196	143	6	17	10	13	68	53
127	4427	1113	346	230	8	10	11	17	1511	442	9655	990
126	1507	610	90	44	919	462	31	43	16	20	319	75
125	13360	2582	2703	873	8	11	5	9	5655	652	16879	1143
124	127	68	23	19	131	170	55	84	14	14	931	146
123	61	31	27	24	6	9	8	15	601	263	3033	591
122	6546	1343	630	233	6	9	20	22	4299	1093	11248	1329
121	25	25	6	12	18	20	14	20	3	6	52	36
120	340	216	9	12	9	10	6	10	1021	392	4797	725
119	1488	458	106	67	1323	570	201	230	52	44	362	99
118	8766	1488	1527	456	9	11	46	43	8396	1030	22204	1488
117	79	56	20	21	45	55	22	30	12	17	798	337
113	11265	2775	1775	592	5018	1422	1470	592	5792	961	11681	1260
112	21569	3701	6012	1592	9594	2857	4421	985	10548	2505	20417	1673
111	42	41	13	23	16	22	11	17	12	19	24	26
110	9452	1935	1121	682	4186	1382	841	322	5708	1158	12921	1509
109	416	251	27	26	1216	595	138	217	20	22	560	144
108	22594	2667	6169	1153	14468	3567	4104	681	13549	1610	32028	3239
107	171	71	39	33	42	52	2	3	3	5	661	240
106	12005	1516	2030	417	4937	1480	1417	364	6164	765	12153	887
105	22308	3237	6396	1491	11015	2822	4362	1123	10858	1586	21050	2197
104	82	72	25	27	52	60	36	32	2	4	270	375
103	10079	1378	2230	468	6313	2114	1142	758	7345	1085	17681	1314
102	845	361	50	43	1260	569	149	223	8	15	505	103
101	19230	3740	5889	1066	12230	3112	3471	758	13189	2925	29030	3283
100	111	69	40	34	51	86	26	37	8	8	1120	189
98	14535	1474	3492	503	7276	1833	2248	485	7709	975	14313	971
97	21403	1937	7003	969	10621	2036	5065	1046	11534	1424	21888	1454
96	190	67	46	30	149	176	8	14	5	13	267	100
95	14800	2005	3722	675	6626	2186	2026	748	7174	1151	16404	1529
94	5109	935	301	84	2364	1108	409	597	140	152	1266	199
93	20968	2265	6956	1585	11831	3834	5072	954	13376	1392	25454	2883
92	174	69	43	34	90	92	30	50	7	10	340	130
91	16408	1531	4768	932	6776	1997	2048	487	8986	1250	12604	997
90	21838	1814	7289	841	9611	2268	5341	867	13349	1684	21477	1688
89	151	80	35	31	434	530	0	0	1697	280	641	141
88	13960	1646	4047	654	6952	1995	2382	663	6986	1123	14884	1608
87	4803	1431	280	144	2273	1348	314	381	1344	277	1572	331
86	28277	3235	9037	1261	14506	5530	4910	751	17253	2642	36287	6115
85	140	110	23	33	37	89	23	43	467	135	402	171
81	12921	1620	3388	522	6980	2375	2514	557	5308	1132	12677	3066
80	19573	1992	6562	1285	13189	3438	6074	943	10388	1862	22041	1234
79	48	35	7	11	38	72	5	9	4	5	200	84
78	12792	861	3280	498	7364	2244	2391	382	5948	1150	13368	891
77	5556	1412	378	90	2542	1017	574	618	163	126	1841	452
76	23215	3400	6759	1121	12247	4798	5547	838	11758	2747	28347	3193
75	76	72	27	32	29	41	25	32	8	8	260	117
235	6070	1042	3390	514	2425	1168	1	3	218	149	6350	968
234	17884	2139	6989	1021	8782	2109	3280	809	9764	1729	13503	3561
228	122	111	39	40	103	78	0	0	4	8	996	258
227	3722	668	4231	472	1819	858	0	0	5	8	581	49
226	4037	1007	80	60	2437	646	319	355	2	4	1090	251
225	219	134	5945	613	415	445	4	8	4	10	694	153
224	16351	1765	32	11	9697	2216	4297	681	8168	1918	16422	1834
223	144	125	521	139	18	19	4	4	13	7	271	130
222	162	109	6314	914	72	105	7	11	3	3	405	271
74	11947	1212	11702	1966	12202	3485	6462	783	4500	823	10380	892
73	20096	2183	21911	3152	15507	4184	7852	1422	10444	1721	13553	4693
72	174	123	13	16	589	405	0	0	6	9	943	187
71	15708	1206	207	105	12059	3225	7644	1093	4987	957	12411	845
70	2103	380	269	75	2142	723	183	287	0	0	1251	84
69	17732	2174	77	48	15402	4523	8054	1311	9503	1780	14868	1814
68	138	60	5737	947	0	0	71	50	3	7	160	47
233	28215	1846	41	35	3147	1070	1367	414	816	269	3183	1319
232	48720	3281	52	51	22339	6452	12638	2218	18557	3094	22629	1861
221	1147	155	81	45	411	346	0	0	123	107	269	64
220	25778	2947	9558	1672	18708	6134	10008	1642	14027	3222	17970	1460
219	5310	1201	449	47	3635	1370	1602	408	177	138	1413	129
218	1690	897	114	37	3789	1442	1994	341	3	5	1291	237
217	294	183	35	35	318	540	8	19	2	4	508	235
216	381	215	71	46	2201	1362	24	46	49	65	2269	332
215	17412	2642	5064	947	11635	2589	8533	1399	7375	1405	13320	1292
214	77	76	32	13	10	14	13	16	14	14	61	87
213	10	16	11	18	7	11	17	18	4	7	12	11
212	125	134	27	17	10	13	11	15	12	15	357	105
211	186	158	48	41	482	420	45	85	8	13	513	54
210	3996	1310	293	78	2959	1371	948	596	29	39	1773	255
209	15785	3124	4570	1077	7939	2373	5025	1008	9874	2097	15392	1481
208	209	172	46	33	82	97	44	82	6	12	278	117
207	4241	1208	277	159	2794	1042	781	516	44	54	1293	249

Vaccine candidate 8 (HC2):

Peptides	mouse1 (1/20)		mouse1 (1/80)		mouse2 (1/20)		mouse2 (1/80)		mouse3 (1/20)		mouse3 (1/80)	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
231	4160	809	925	417	1281	495	392	152	534	163	141	97
230	6955	1214	1844	557	3694	1162	1193	466	1193	352	292	170
229	5572	2299	1550	594	1258	500	265	171	298	117	36	51
130	8271	2124	1723	631	6695	900	2717	664	6093	997	2217	671
129	18712	2138	8428	1960	13532	2582	6265	1263	12341	472	6869	988
128	1080	510	263	110	2115	1227	599	196	68	53	15	28
127	16329	2836	4866	1131	11349	1694	4737	984	9655	990	4537	765
126	5279	1062	1749	207	5579	1095	1955	347	319	75	90	77
125	26430	2070	13825	1291	19689	1285	10234	1427	16879	1143	10249	1467
124	2429	444	704	233	891	258	205	76	931	146	231	72
123	5345	2352	1376	702	3124	1133	860	336	3033	591	851	282
122	23053	1550	11262	2629	12786	1220	5319	1034	11248	1329	6691	1046
121	613	178	173	70	1224	489	252	89	52	36	19	21
120	8690	3502	2033	513	798	394	164	62	4797	725	1830	668
119	5322	684	1735	175	7937	616	3361	454	362	99	444	1068
118	44652	4132	19263	475	18005	1270	6858	887	22204	1488	12596	1813
117	2324	135	654	119	299	73	55	44	798	337	200	40
113	20655	2105	10104	1580	18825	1961	10474	1058	11681	1260	6222	1206
112	35612	17186	27148	7177	31599	7619	15500	5907	20417	1673	11456	4300
111	1517	251	410	169	3460	779	793	253	24	26	43	59
110	26158	2146	12573	1520	19017	985	8609	809	12921	1509	6866	1255
109	4362	1991	1506	349	6530	2678	2665	751	560	144	146	66
108	52376	4837	30489	2319	32691	1556	19731	1398	32028	3239	19960	2766
107	3947	549	1042	141	552	64	116	56	661	240	140	50
106	18465	1844	10400	305	17715	1204	10497	1214	12153	887	6974	783
105	41853	3019	26481	2662	29571	4299	17401	1407	21050	2197	13713	2038
104	1747	997	645	187	4673	1597	1347	659	270	375	26	29
103	25373	1589	12932	1890	19925	811	11645	1167	17681	1314	10599	1171
102	6067	781	1758	395	8306	1346	3272	427	505	103	93	66
101	40022	14642	26679	2022	24858	10682	17123	4427	29030	3283	16649	5622
100	2879	401	748	269	497	145	131	48	1120	189	241	82
98	20811	1760	13819	1397	21734	1288	14487	870	14313	971	9515	992
97	31811	3690	23767	466	27592	2131	19441	1037	21888	1454	15593	1611
96	3079	711	941	130	5685	1237	1815	319	267	100	53	46
95	23808	6400	15538	1998	18802	4860	12622	1574	16404	1529	10104	1249
94	5953	782	1978	345	10343	697	4774	441	1266	199	275	74
93	39446	3477	24103	1271	25551	1840	16467	1222	25454	2883	17126	2390
92	3191	1056	1128	221	1458	714	415	140	340	130	203	260
91	24909	2808	16119	801	22307	1176	13732	1225	12604	997	8353	1149
90	35780	2905	24545	910	29193	886	20304	1220	21477	1688	15766	2107
89	2913	333	953	130	5361	1191	2156	523	641	141	157	71
88	23322	2131	15430	852	21523	1113	13926	987	14884	1608	9776	1355
87	5242	1913	1942	360	7839	3147	4019	1470	1572	331	328	91
86	53256	4463	32892	1430	36883	1792	23689	1958	36287	6115	22615	2559
85	3800	454	1034	336	1701	456	420	129	402	171	56	61
81	21613	1610	14911	1208	20458	8002	13428	4088	12677	3066	7240	4454
80	35894	2677	24409	1029	30238	1375	20623	971	22041	1234	15731	1993
79	2274	222	594	164	4091	630	1436	240	200	84	42	45
78	21763	1313	13037	413	18965	803	12018	1022	13368	891	8914	1073
77	6307	412	2265	312	10350	742	5167	397	1841	452	531	181
76	40997	10966	26331	2084	25363	10352	19065	2722	28347	3193	18458	2213
75	2678	315	826	213	745	163	142	74	260	117	67	43
235	5421	1053	1608	174	12574	1380	6384	788	6350	968	2260	290
234	24624	2312	21709	7390	28681	5798	18520	1075	13503	3561	11923	1684
228	2291	314	628	119	190	91	41	30	996	258	281	150
227	4578	405	1438	213	3762	320	1159	128	581	49	74	29
226	4937	515	1820	188	7979	951	3527	718	1090	251	236	53
225	4016	1436	1404	199	892	425	186	116	694	153	162	86
224	26818	2723	18345	958	20470	1643	13288	5558	16422	1834	12885	1846
223	887	375	204	97	62	58	31	21	271	130	84	41
222	3675	751	1007	312	518	259	128	92	405	271	114	94
74	16713	1085	9444	728	16714	730	10802	799	10380	892	6794	1031
73	26247	2015	16233	1435	19517	7514	13951	3439	13553	4693	8803	5415
72	2489	316	785	123	3751	341	1263	282	943	187	254	53
71	20076	1485	11036	704	16677	963	10969	859	12411	845	8115	948
70	5180	397	1655	211	5873	612	2215	361	1251	84	299	114
69	24341	1583	14187	1003	17887	1238	12213	1046	14868	1814	10237	1534
68	1959	631	557	159	690	329	122	53	160	47	22	24
233	5882	891	1959	392	8366	3477	4434	1850	3183	1319	752	474
232	54024	4232	30433	1156	50857	2285	29449	1725	22629	1861	16851	2226
221	4387	594	1430	273	1934	218	535	111	269	64	54	49
220	35511	3107	22399	1028	28382	1203	17446	1129	17970	1460	13268	1615
219	5401	566	1709	341	3848	495	1370	226	1413	129	205	96
218	4500	573	1441	171	8954	1197	4576	958	1291	237	336	83
217	3024	1211	1021	232	933	425	237	91	508	235	141	42
216	7220	558	2496	239	11560	661	5039	632	2269	332	562	111
215	21976	1286	13483	623	22146	1160	14674	630	13320	1292	8603	1435
214	1814	212	449	206	625	297	139	70	61	87	43	44
213	8	13	10	18	2	5	3	8	12	11	9	9
212	1097	267	309	83	117	43	28	30	357	105	258	418
211	4839	297	1645	173	1460	337	356	84	513	54	123	49
210	6017	513	2082	270	4912	849	2035	443	1773	255	467	109
209	21286	5776	11320	1251	15857	7549	8885	1827	15392	1481	10978	1334
208	4578	426	1459	192	949	160	207	93	278	117	72	60
207	5954	1080	1992	200	3614	810	1186	227	1293	249	263	200

9.3.3 Microarray format 3 (MA3)

Vaccine candidate 1 (CHSynB):

Peptides	mouse1 (1/500)		mouse2 (1/10)		mouse3 (1/10)	
	MEAN	SD	MEAN	SD	MEAN	SD
131	57882	6416	73120	5436	103813	10930
132	74588	8265	91741	5622	141171	13257
131	35286	1682	59962	4098	61466	7631
232	124478	15093	122011	6687	157653	10346
233	72519	8612	97213	12853	80421	8515
234	13089	1352	12167	742	83237	9233
235	14344	2500	3123	250	3289	611
236	11591	1422	10038	1116	965	352
237	9556	1175	4688	424	679	607
238	73126	5870	21275	1244	97831	11661
239	79197	3946	31088	2258	126685	8678
73	55546	6345	20448	1582	72588	12727
240	72725	5859	25113	1325	95025	9781
69	56557	3723	18187	965	61983	7120
76	22764	1660	31217	2618	114389	13153
108	11131	337	20132	1628	89257	5513
174	2899	388	7977	607	39184	5024
125	2064	457	4511	523	36106	2908
241	12112	2129	11834	1708	1270	351

Vaccine candidate 2 (HC12):

Peptides	mouse1 (1/20)		mouse1 (1/80)		mouse2 (1/20)		mouse2 (1/80)		mouse3 (1/20)		mouse3 (1/80)	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
131	164069	22915	132231	15283	86560	7519	35836	2825	53449	6425	35836	2825
132	188231	25918	150696	12036	120432	16975	49358	4561	57456	7389	49358	4561
131	145285	22504	114453	10958	34814	3682	12519	1523	37529	5569	12519	1523
232	215669	41099	178899	24204	152563	24984	79916	12567	91559	20960	79916	12567
233	206897	36939	173373	33755	21974	3689	7272	1013	64730	11339	7272	1013
234	1426	297	381	417	66632	7691	34380	3390	14239	3040	34380	3390
235	99	89	18	30	14643	2124	4241	437	26	68	4241	437
236	150875	20431	101993	11001	15026	2226	3442	300	15243	2132	3442	300
237	281	268	18	49	11550	1947	2702	272	0	0	2702	272
238	2423	395	263	115	128820	24936	68268	5448	34934	7377	68268	5448
239	188153	33369	126531	12177	151022	28596	80588	5828	49621	4924	80588	5828
73	908	330	117	147	101923	16539	53625	5585	28554	5236	53625	5585
240	1695	271	202	87	120544	8654	63508	3728	33014	6468	63508	3728
69	777	256	129	137	89533	5637	45929	2449	31275	5880	45929	2449
76	2281	201	389	158	122083	17351	52682	4953	19883	3013	52682	4953
108	1367	210	86	108	91320	10647	38125	3786	10024	2291	38125	3786
174	1112	142	168	174	43074	5041	16376	1463	0	0	16376	1463
125	1037	172	204	206	39360	4627	14784	1129	5	14	14784	1129
241	86	78	42	49	11813	1793	2832	509	15	41	2832	509

Vaccine candidate 3 (HC11):

Peptides	mouse1 (1/10)		mouse2 (1/80)		mouse3 (1/20)	
	MEAN	SD	MEAN	SD	MEAN	SD
131	200	121	0	0	1950	600
132	140	131	0	0	2188	935
131	66	87	43	57	1419	382
232	3845	464	0	0	2435	309
233	4021	496	106152	9251	3047	1340
234	1079	238	2	6	140	202
235	4761	654	103567	18278	247	145
236	4754	513	88708	17379	3049	457
237	10230	635	98174	18693	105321	14728
238	6302	426	93	123	117705	15619
239	4598	483	34	50	4740	1660
73	834	200	0	0	280	164
240	5492	569	139	118	809	506
69	203	171	0	0	133	130
76	273	112	29	51	108	123
108	24	36	0	0	23	62
174	207	104	0	0	33	82
125	57	123	8	14	78	146
241	8495	986	97155	14076	1100	237

Vaccine candidate 5 (SH127):

Peptides	mouse1 (1/4000)		mouse2 (1/400)		mouse3 (1/2000)	
	MEAN	SD	MEAN	SD	MEAN	SD
131	128338	14333	21782	2404	105023	9068
132	162584	26962	30644	4371	147196	18220
131	116628	13062	15968	1742	72784	8482
232	162063	27150	27982	3816	143294	19326
233	158482	10402	26156	4328	135604	10927
234	59276	9528	14654	2868	64077	9330
235	79853	19167	17444	4274	87442	17171
236	70784	11619	19485	3602	99445	13230
237	64402	20246	17120	3385	82826	12676
238	72857	14295	21350	2617	100452	11497
239	86948	21886	23107	2988	113567	13409
73	58846	9303	18870	2245	81870	5966
240	81047	15151	21879	2003	102670	12115
69	63402	15611	19604	2924	89155	11699
76	83213	14775	23261	3071	116405	10740
108	67808	19929	19885	2354	97728	9360
174	69127	10759	17684	2179	86765	7022
125	60733	10443	14783	2579	69345	8168
241	73298	16092	19572	3596	108315	12974

Vaccine candidate 6 (AuNP):

Peptides	mouse1 (1/5)		mouse2 (1/20)		mouse3 (1/80)	
	MEAN	SD	MEAN	SD	MEAN	SD
131	6855	544	50840	4994	16422	2055
132	9775	652	67905	6959	33075	3050
131	3853	185	14364	1320	12630	1057
232	11089	752	74968	6173	62812	5343
233	3416	225	38421	3472	13654	1328
234	3149	258	29992	2365	10677	764
235	30	70	25815	1845	1622	415
236	28	43	45073	4319	1626	373
237	56	112	31369	2974	1132	290
238	1822	214	57253	5491	35485	2330
239	2769	727	78499	7209	47882	3091
73	1338	819	55129	5721	27876	2137
240	1453	184	73573	8370	38397	3787
69	581	185	61664	5252	28458	4989
76	3940	275	86737	4209	20452	1489
108	1945	177	58947	4621	8249	1131
174	196	204	39685	4552	8447	1058
125	317	255	23645	3016	5016	773
241	20	26	51457	9125	1720	377

Vaccine candidate 7 (HC1):

Peptides	mouse1 (1/80)		mouse1 (1/320)		mouse2 (1/80)		mouse2 (1/320)		mouse3 (1/80)		mouse3 (1/320)	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
131	62600	5878	16228	2088	26662	3931	5660	530	30368	2727	6783	653
132	88754	6996	25862	3322	41043	3921	8695	1417	44403	3851	10572	694
131	50080	3700	12692	1456	0	0	9	24	15253	1199	3571	292
232	118232	9563	29439	3610	98205	11874	19597	3016	71566	7544	18944	2156
233	60763	6652	14538	2005	13321	1714	3632	732	1882	175	383	122
234	30313	2182	7496	767	24868	2826	5256	884	30155	2267	7232	521
235	2183	433	189	141	5015	935	956	158	388	157	26	34
236	25845	2086	5383	622	7581	967	1399	195	494	145	84	84
237	1177	303	182	121	9487	1789	1932	441	536	249	134	159
238	49457	2842	11281	1388	89653	7263	18904	1693	50435	2725	13547	926
239	91194	6223	22790	2210	97478	8706	21176	3283	61149	2387	16399	1263
73	35600	2412	7997	934	65439	6502	15526	2130	33550	1271	9386	740
240	42424	1949	9994	852	81902	5758	17345	2254	39588	1451	9817	650
69	28352	1560	7009	565	57757	2901	14492	1247	24952	1315	6976	281
76	43614	3804	9447	884	52855	5495	10637	1647	39056	3816	8920	962
108	29910	3708	6998	623	31060	4700	6567	1754	26675	3234	6100	980
174	16989	1253	3784	173	0	0	22	51	13274	1208	3106	518
125	14528	948	3124	322	51	136	66	78	12421	1246	2800	234
241	1696	462	284	184	9286	1275	1914	226	454	124	51	74

Vaccine candidate 8 (HC2):

Peptides	mouse1 (1/20)		mouse1 (1/80)		mouse2 (1/20)		mouse2 (1/80)		mouse3 (1/20)		mouse3 (1/80)	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
131	106588	8739	50155	7518	96477	7665	40498	5452	70262	5367	26820	2098
132	140794	9337	74318	13406	144198	10160	65560	9663	106447	8141	42146	4998
131	62314	4945	20596	3744	55932	3790	20125	2539	46326	3295	14947	1789
232	150340	11567	71911	13504	157447	8972	72032	11861	104507	4116	40914	7027
233	16765	2096	4335	1018	30611	3137	9423	1779	5794	598	1313	156
234	100920	4040	50846	6075	110827	8407	49692	8359	64324	5783	27260	2904
235	14567	1846	3534	480	57616	6285	16580	2522	10895	815	2544	367
236	18258	3145	4597	793	23941	1806	5782	669	5673	374	1181	237
237	11700	1928	2811	821	3641	864	993	308	1117	266	200	189
238	110176	5031	52842	9650	93693	5387	39075	6261	80679	3637	34050	3777
239	135253	11877	69781	14076	155948	10564	67836	10756	118626	7699	52367	6037
73	76074	9580	34465	8066	80761	7824	37144	4908	67026	6688	29038	2953
240	97484	14127	45224	8870	87212	6895	36378	4753	89285	8219	37705	4265
69	65749	8511	24893	3850	59338	4424	24594	4317	63303	5144	24349	3017
76	137940	16497	66770	9485	128750	10481	58564	11644	124577	8596	52015	5875
108	108355	13308	47653	7497	93281	5223	40412	8521	96129	2812	39273	3168
174	70158	5605	20924	3397	63657	5575	22364	4456	70379	6210	23921	2707
125	66499	6012	19246	2747	59138	6074	20673	4694	68300	9096	21408	3232
241	18108	2047	4268	924	8865	2736	1836	385	4062	338	857	146

9.3.4 Microarray format 4 (MA4)

Serum antibody microarray:

Peptides	CHSynB mouse1 (1/20)		HC12 mouse 2 (1/20)		HC1 mouse 3 (1/80)		HC2 mouse 2 (1/20)		HC2 mouse 3 (1/20)		Ma552 (25µg/mL)	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
75	1402	263	3356	326	80	69	786	108	830	233	7826	1249
175	2993	422	4297	443	67	104	1960	355	1567	161	11716	1876
76	28044	4891	61421	2991	28628	4435	65868	5268	59313	6407	27128	2681
176	37426	7411	65827	4764	26557	4850	72185	5807	68085	7655	31576	2181
77	8194	1612	12771	1209	169	142	19482	1574	1258	179	4245	405
177	9644	1346	13986	995	188	121	20911	2008	1644	338	6879	713
78	12814	1649	27934	2939	12853	1823	48124	3958	36037	3829	27231	2202
178	21144	3254	22809	4457	11328	1224	50945	5026	48300	3348	29640	2013
79	443	149	3446	389	23	36	5053	757	42	58	6482	1046
179	923	196	4726	458	49	111	12515	955	518	155	11428	2901
80	22504	3408	61417	2105	28042	5037	69995	7576	54579	8338	25907	3618
180	28380	6064	57695	4557	21899	5064	64289	6746	58010	7234	26328	3405
81	10850	863	25906	3104	12551	1094	54807	5248	33112	2686	24960	2447
181	17278	1553	17833	2644	11702	926	51334	4957	42642	3202	24910	2593
92	2284	335	3761	278	8	20	1013	184	584	170	7680	1229
185	4045	377	4490	472	17	27	1668	271	622	107	10142	1646
182	3846	308	4489	318	18	29	1488	243	660	260	9330	1636
93	25412	3744	63252	1945	28943	3075	70204	4997	66851	5246	28154	1823
186	29995	5232	61639	708	29389	3544	68500	5735	66348	6724	29135	2335
183	33515	6539	66533	7642	31329	6246	76575	11097	76180	10782	33111	4257
94	8413	470	12031	1067	130	129	20133	2434	1116	122	3938	486
187	10736	925	11949	1468	206	101	20807	1873	1404	211	4922	503
184	12518	974	13803	1488	189	131	21464	1680	1715	293	5362	723
98	9497	785	31753	2812	16264	2076	59448	5037	40358	2260	24848	1650
188	12141	1721	12543	1089	13695	1519	46458	3510	37812	3094	20843	1085
101	21504	3793	60618	5030	24779	3530	64217	5980	62922	1664	25489	2616
189	26331	6114	58618	5094	23439	4500	62972	7923	61648	8648	27428	2152
102	6990	569	10343	1737	58	50	15835	1654	894	202	3439	619
191	9053	536	10312	1365	109	88	17411	1686	1144	199	4336	868
192	11027	1070	10813	1375	80	79	15579	1069	1310	185	4581	906
190	11422	749	11706	1369	91	52	20333	2041	1562	366	4893	742
107	1503	271	3102	169	52	50	721	239	2116	151	5976	1032
193	2408	299	3866	508	153	59	805	65	2664	176	8942	1259
108	21411	2814	58120	3782	25308	2500	59947	6985	56781	1669	25287	1880
195	25455	4229	54414	5067	23849	3245	54255	6584	54753	2616	27781	1866
194	31264	6772	62395	4016	28091	3631	67152	6165	66597	5105	32037	2995
109	8806	1219	9347	856	173	119	15436	1559	1230	99	3301	459
197	10922	1956	9637	1063	132	149	16502	1209	1296	168	3930	687
198	12401	2448	9784	1631	93	105	13474	1197	1135	169	4182	582
196	12626	1511	11151	1766	74	97	17909	1608	1587	185	4371	994
113	5315	375	15371	2473	12566	1499	50786	8682	32367	1813	22597	815
199	9566	665	9895	1459	12495	1471	51981	7096	37311	2592	25201	3194
119	6031	454	9326	1817	56	63	10553	1243	353	150	2344	435
200	8812	1034	8971	1697	157	91	16275	2056	303	185	4117	483
125	3660	634	19094	2928	11773	1145	40045	4540	42156	3301	15720	1117
202	3766	163	8784	1460	10152	1204	15830	2771	33459	1659	19112	1385
201	11014	1702	26530	850	15233	1444	49464	4835	55301	5406	25783	2663
126	6974	956	9400	1288	142	148	5400	345	448	117	2158	409
203	9735	1484	9532	1577	108	93	2914	274	214	72	2987	488
204	10300	1795	7991	1235	54	84	1635	144	243	144	2647	448
174	7819	134	19242	2823	11155	1023	45566	6602	47324	2363	15731	1577
205	12167	604	19558	2530	6119	407	46288	6816	54816	356	24939	837
206	12886	1217	26919	3188	14412	2193	52254	6187	48411	5982	24276	2704
232	64309	6944	75609	7913	44169	5505	84912	12634	48522	4819	23459	2770
234	12770	1802	46217	5793	26256	3880	64302	10446	34915	2878	17442	1877

Galectin-3 microarray:

Peptides	Gal-3 (25 µg/mL)		Gal-3 (12.5 µg/mL)		Gal-3 (6.25 µg/mL)		Gal-3 (3.13 µg/mL)		Gal-3 (1.56 µg/mL)		Gal-3 (0.78 µg/mL)	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
75	10468	1367	2524	697	35	66	28	41	7	17	72	50
175	17134	1877	7615	1821	138	113	147	144	90	88	70	78
76	14667	2136	5453	699	58	89	144	132	96	101	60	86
176	24358	2352	12175	2544	705	205	236	179	92	103	53	99
77	18426	2127	7343	1564	262	199	222	157	117	106	54	82
177	25532	2746	13667	2741	1315	435	636	198	291	200	72	71
78	16563	2728	5918	1435	335	126	115	89	36	42	68	78
178	25320	2706	14565	2666	2155	475	518	251	329	178	97	105
79	17686	2981	7518	841	528	121	258	167	188	214	177	130
179	29966	1085	19332	1510	3381	534	970	304	660	147	188	157
80	23525	2183	9650	1201	596	248	465	151	248	136	48	94
180	33222	2152	21487	3413	4266	976	1610	569	1128	269	260	144
81	19950	3894	12347	3979	852	384	542	221	270	94	134	149
181	33103	2810	23600	2460	5618	1227	1687	665	1283	336	454	193
92	24762	2345	16242	2128	3351	373	1474	599	807	215	316	176
185	158	201	87	71	101	121	77	122	56	63	70	96
182	35660	2698	26460	2401	11037	1460	5374	1949	3092	483	1368	235
93	26523	2557	17552	2962	4194	894	2479	547	1476	489	471	211
186	9854	2475	1763	584	66	92	49	97	62	56	40	97
183	37977	1929	27187	2894	12743	2179	5449	1272	3839	368	1655	490
94	31144	2395	19921	2072	5222	1014	3090	699	1535	429	652	184
187	14288	1749	6591	1736	90	146	27	50	85	132	0	0
184	39784	3123	31472	1962	16479	2022	7409	1257	5365	992	2594	187
98	34742	3408	29654	2939	17697	2660	8476	2139	6824	1004	4230	716
188	20592	3415	11911	1989	401	145	161	121	94	134	104	106
101	27882	2617	19195	1862	6123	1186	2965	731	2247	399	1082	316
189	33287	1510	25458	2161	12125	1649	5094	990	3694	356	1837	152
102	29040	2964	19880	2287	6649	1122	3437	1211	2218	500	1011	362
191	20932	2383	11049	1871	1354	387	463	232	279	143	111	141
192	222	277	61	72	11	22	31	49	10	20	94	117
190	37973	4384	30124	1931	14809	973	6038	1281	4591	647	2514	315
107	25499	3370	17645	2578	5069	660	2970	627	1616	297	595	144
193	34728	1248	27923	326	14002	1253	6596	113	4538	733	1901	196
108	26860	2349	18922	1223	6719	1097	3589	437	2382	433	1092	166
195	22159	1687	13991	1520	2377	247	903	174	464	190	88	137
194	33899	2447	27322	2418	14887	1619	7031	1306	5031	736	2530	177
109	31774	791	21522	1911	8672	1237	4488	1290	3645	948	1436	451
197	26902	1650	17082	1598	3909	583	1928	559	1521	674	495	236
198	167	118	39	44	67	96	109	63	55	70	59	69
196	40403	2096	32493	1400	18481	2430	9072	2786	6696	1317	3724	609
113	30168	2671	25885	2202	18246	3198	9608	2625	10211	2095	7644	1411
199	36728	2479	34065	1070	29124	2872	19015	4241	20703	3668	21353	1976
119	30991	3937	23931	2144	11447	1058	6632	1762	4758	1128	2393	315
200	22027	3053	15391	2342	2899	480	752	147	487	107	119	188
125	29081	3161	22638	1581	10397	1286	5478	1123	3973	887	1788	369
202	21885	1800	15854	2018	3831	514	1183	459	792	80	197	96
201	41111	1374	33978	2109	24385	4347	14364	2646	10188	1998	6898	1334
126	33682	2752	24694	1270	12609	2357	6335	1534	5522	1207	2952	679
203	25628	2262	19043	1142	7703	786	2508	280	2016	301	562	212
204	1508	220	769	175	85	110	36	48	51	84	117	129
174	24522	3475	15584	393	3254	263	1675	417	1104	314	385	141
205	30399	3092	22067	313	6349	598	2952	1029	1967	542	746	71
206	29988	3344	22658	1348	9085	713	3996	1257	2843	713	1376	305
232	14145	2031	3395	908	57	2	99	124	170	75	62	17
234	17631	3109	8294	1425	776	308	414	72	202	218	32	55

9.3.5 Microarray format 5 (MA5)

Plant lectin microarray:

Peptides	ECA (0.1 mg/mL)		MAL I (0.08 mg/mL)		MAL I (0.2 mg/mL)		PNA (0.01 mg/mL)		PNA (0.05 mg/mL)		WGA(0.02 mg/mL)		VVL (0.05 mg/mL)		SNA (0.01 mg/mL)		MAL II (0.02 mg/mL)		MAL II (0.04 mg/mL)	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
238	0	0	0	0	20	28	1	2	0	0	0	0	19	26	0	0	8	18	0	0
208	0	0	31	46	24	43	16	23	3	7	0	0	33	48	0	0	32392	72431	0	0
209	0	0	83	136	0	0	8	10	14	32	0	0	3309	462	0	0	450	1006	849	1898
210	0	0	21	29	25	46	7	15	1	2	0	0	1665	441	1	1	160	358	0	0
211	0	0	1	2	21	47	55	74	0	0	0	0	763	129	2	2	2	4	0	0
212	29	34	21	48	31	69	12	26	0	0	873	1951	24	50	0	1	0	0	0	0
213	35	34	28	33	24	27	15	21	22	49	0	0	387	135	0	0	0	0	0	0
214	0	0	90	20	94	46	27	29	43	37	7	16	126463	7432	0	0	0	0	0	0
215	0	0	133	50	200	34	7	16	0	0	1816	2013	311410	5651	0	0	0	0	0	0
216	26	37	112	60	141	97	24	39	0	0	43621	5531	321832	22256	18	39	0	0	0	0
218	9	11	16	16	89	128	40	48	3	6	0	0	95403	16207	0	0	0	0	0	0
219	0	0	49	33	54	70	38	40	0	0	0	0	209901	7765	1	2	0	0	0	0
221	13	21	1	1	36	38	16	23	0	0	0	0	162420	9806	0	0	0	0	0	0
229	0	0	360	389	214	39	17	39	12	26	0	0	391474	6195	0	0	15	33	0	0
230	40	47	96	58	168	109	26	32	23	34	0	0	266799	17576	0	0	0	0	0	0
233	0	0	57	45	262	358	0	0	0	0	506	695	204769	8114	0	0	1	2	0	0
234	0	0	136	198	41	66	0	0	9	21	0	0	155377	9169	0	0	2940	6570	0	0
235	10	13	60	23	67	43	12	26	0	0	0	0	359530	8347	0	0	0	0	0	0
236	0	0	139	46	158	63	4	8	14	31	0	0	395243	7838	0	0	0	0	8	17
237	0	1	50	52	103	132	4	8	0	0	0	0	210199	9057	0	0	0	0	3	7
68	0	0	70	65	41	54	48	48	122	274	980	2190	100512	8717	0	0	0	0	0	0
69	18	27	42	35	21	26	12	22	2	4	0	0	203227	4688	0	0	0	0	0	0
70	25	26	82	54	139	109	15	21	21	23	0	0	221546	7616	0	0	0	0	1	1
71	4	5	59	54	112	131	2	4	27	42	0	0	326556	13399	1	1	120	268	0	0
72	29	23	173	83	448	115	7	15	16	23	55	118	250424	4922	0	0	0	0	3	8
73	29	28	183	75	318	39	7	14	10	23	64	144	274312	8069	5	9	0	0	0	0
74	40	44	303	50	485	77	0	0	0	0	14	30	293559	7352	1	2	0	0	0	0
223	0	0	22	22	41	71	22293	2156	43512	4627	0	0	57	78	2	4	0	0	0	0
225	36	39	32	29	0	0	35583	3141	58511	4602	39	86	19	42	0	0	0	0	0	0
227	18	19	6	11	0	0	69884	4029	87404	4260	0	0	0	0	0	0	0	0	0	0
228	20	19	0	1	15	33	33815	1839	59174	2049	20	45	20	45	0	0	9	21	4	10
231	0	0	18	40	0	0	144346	3638	165444	5302	0	0	53	99	0	0	8	19	0	0
232	0	0	33	32	1	1	162176	759	205987	29492	0	0	11	24	0	0	0	0	0	0
82	27	27	4	9	30	42	42065	5115	81476	3888	0	0	223	185	0	0	0	0	1	2
83	0	0	15	22	27	59	29131	2563	55157	1423	11972	26769	0	0	0	0	0	0	0	0
84	0	0	13	18	25	30	41871	3221	62466	2033	0	0	7	17	0	0	0	0	224	502
75	48	50	295	625	15	34	16966	2078	40901	3745	100	224	101	146	0	1	2	4	0	0
76	0	0	4	8	11	19	38338	3758	72944	3206	0	0	10	23	0	0	0	0	0	0
77	0	0	14	19	14	15	110741	4872	140381	6589	0	0	0	0	0	0	123	276	0	0
78	10	10	12	27	114	238	79986	1597	134500	5583	715	1599	0	0	0	0	0	0	0	0
79	0	0	56	72	12	26	129919	1972	168728	4154	0	0	38	85	0	0	218	488	0	0
80	15	22	14	21	54	89	143042	6180	173478	6313	0	0	29	30	0	0	0	0	0	0
81	25	23	4	8	31	52	149370	5365	170956	3871	0	0	158	117	5	10	9	21	0	0
175	0	0	24	33	31	42	43	39	49	49	3446	1090	15	31	1	2	446	263	851	70
176	0	0	26	40	51	115	0	0	14	32	19688	1667	53	52	0	0	885	130	1764	90
177	24	27	23	39	56	39	0	0	11	25	42712	4142	671	1473	0	0	24173	2293	38883	2271
178	0	0	13	23	48	48	0	0	46	104	63847	6216	2	4	0	0	5441	736	9734	972
179	0	0	53	48	35	66	14	16	0	0	92101	5365	78	120	0	0	35455	2651	52153	3344
180	24	28	50	49	48	50	0	0	26	59	93971	4922	170	184	0	0	40098	2499	59695	2171
181	0	0	71	50	76	55	356	157	1046	232	109496	3629	100	141	0	0	45996	3692	66923	3007
85	44	15	25	34	6	13	14	32	11	25	0	0	10	23	0	0	9	20	0	0
86	0	0	31	42	33	31	12	21	4	6	32	72	1	1	0	0	6	13	0	0
87	36	31	24	24	4	9	43	50	13	28	0	0	74	122	1	1	0	0	0	0
88	79	70	11	14	1	2	12	17	18	40	0	0	20	19	0	0	194	434	25	55
89	101	55	20	20	13	19	22	41	0	0	0	0	44	51	0	0	0	0	0	0
90	185	149	41	42	18	19	9	16	309	691	0	0	51	105	0	0	0	0	4	9
91	323	229	24	18	0	0	16	24	0	0	0	0	135	104	0	0	20	45	24	53
92	5939	744	281	90	369	118	33	72	7	14	0	0	47	50	0	0	858	1815	170	356
93	6641	1191	22	32	36	63	29	62	1	3	3029	6772	42	93	2	4	22	31	0	0
94	15030	2441	14	31	46	79	6	9	38	31	0	0	4562	10158	112	87	0	0	2	3
95	20513	1585	421	48	651	90	9	20	0	0	0	0	30	51	38	45	14	31	20	44
96	25775	598	656	79	1106	171	20	25	0	0	0	0	111	126	427	241	25	57	0	0
97	36176	1352	38	24	31	44	0	1	8	17	0	0	72	161	454	204	39	88	0	0
98	43758	2103	1044	163	1817	146	27157	60675	12	16	0	0	104	130	1119	270	652	1363	0	0
182	0	0	14	21	29	29	16	25	3	6	0	0	18	40	50529	9384	49	109	0	0
183	0	0	23	25	25	37	12	23	4443	9914	0	0	2	4	48869	15337	0	0	0	0
184	0	0	47	54	77	68	32	51	27	60	0	0	14	31	66615	13788	4	8	0	0
188	0	0	87	53	138	92	20	19	10	22	0	0	84	117	132000	15488	0	0	0	0
185	17	17	48751	3636	50902	5206	20	27	31	37	0	0	78	69	4	6	1089	285	1812	376
186	0	0	4751	805	5914	1995	29	27	17	38	46	71	60	113	6	10	52	55	285	113
187	0	0	14896	2495	16591	5305	148	191	32	39	665	188	81	83	3	6	150	103	309	199
99	6093	780	217	59	290	97	6	13	154	291	0	0	71	67	162	120	33	34	70	102
100	2	3	30	42	300	410	19	42	34	48	0	0	48	72	1	1	26	58	10	23
101	3	4	0	0	26	36	56	34	23	32	516	1154	59	59	1	1	31	70	39	74
102	56	55	27	29	11	23	55	17	17	38	111	248	19	20	0	1	0	0	12	20
103	96	59	1	1	19	33	69	45	6	13	0	0	81	127	2	3	82	166	6	14
104	154	97	38	55	55	123	123	218	81	56	0	0	2567	5497	0	0	0	0	0	0
105	217	15	12	15	67	116	16	35	9	20	0	0	81	97	6	13	29	64	46	78
106	552	76	11	15	136	305	21	22	36	47	0	0	141	170	0	1	68	138	0	0
191	0	0	55	37	41	53	36	36	23	22	0	0	6	13	15357	9072	8	18	7	17
192	2	3	17	25	387	795	14	32	5	10	818	1830	74	91	9492	8982	0	0	0	0
189	16	16	4	10	19	32	36	50	30	41	25726	5337	37	53	0	0	0	0	9	21
190	0	0	30	32	2	4	27	31	66	147	19554	1790	39	48	0	0	0	0	0	0
107	8113	918	100	35	150	106	9	19	160	291	244	546	58	51	23	22	5	7	0	0
108	8994	248	15	22	3	6	50	39	13	12	309	454	38	54	82	64	0	0	40	41
109	11025	1195	23	31	46	43	6	13	8	19	40	87	104	47	458	205	24	55	0	0
110	33612	1458	405	59	731	111	2	3	45	78	555	290	498	869	164	95	17	38	0	0
111	35445	2152	348	49	481	284	877	1910	20	30	291	367	56	51	584	381	0	0	7	16
112	34019	2843	168	59	368	95	8	15	0	1	651	371	127	86	1011	641	183	350	0	0
113	56040	2686	1063	51	2127	184	1	2	21	22	1584	1028	138	58	1027	294	0	0	0	0
116	11541	1137	331	54	380	226	182	59	123	43	48	68	460	84	319	226	78	97	52	44
115	10694	956	46	44	46	72	35	24	51	63	7395	16515	82	96	15	34	13	30	7	16
114	6073	620	47	42	92	115	632	1202	102	171	191	378	85	107	12	16	91	190	34	48
195	0	0	48	44	31	62	24	34	24	33	0	0	85	70	35234	9529	31	47	22	49
197	0	0	15	34	50	54	57	47	15	34	0	0	66	94	44965	11090	86	193	0	0
198	0	0	28	36	58	50	16	23	10	14	0	0	30	59	47423	12883	0	0	0	0
193	0	0	38562	3257	35347	12939	11	24	5	11	849	318	39	88	1	2	778	252	1340	259
194	6	8	36143	3007	35557	8592	35	37	35	69	4456	636	54	56	0	0	683	323	1048	295
196	0	0	40235	3733	38200	9417	18	25	30	44	2660	571	23	51	0	0	893	826	898	615
199	0	0	131226	16508	144436	6836	15	22	69	74	144420	7746	0	0	0	0	3908	1933	5435	1213
117	157	49	46	52	22	26	715	1520	38	64	42	93	36	42	1	1	5	12	15	24

Peptides	ECA (0.1 mg/mL)		MAL I (0.08 mg/mL)		MAL I (0.2 mg/mL)		PNA (0.01 mg/mL)		PNA (0.05 mg/mL)		WGA(0.02 mg/mL)		VVL (0.05 mg/mL)		SNA (0.01 mg/mL)		MAL II (0.02 mg/mL)		MAL II (0.04 mg/mL)	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
118	85	55	40	38	12	21	72	46	5525	12218	116	212	65	65	0	0	33	46	3	6
119	296	79	38	54	48	108	30	34	41	45	37	83	64	98	32	57	62	138	23	34
120	707	116	4	8	189	362	98	143	198	192	0	0	197	111	1	1	0	0	24	43
121	3101	512	12	18	22	31	164	91	573	318	0	0	341	577	88	59	0	0	0	0
122	2315	425	17	39	26	41	111	88	228	131	10	22	10	20	69	86	87	195	0	0
123	5202	1593	25	55	9	9	220	38	1036	334	15	34	64	62	63	53	12	26	6	13
200	0	0	28	38	38	79	19	22	510	1039	42	94	28	48	23355	12956	36	71	0	0
124	29499	1345	329	98	496	86	26	31	0	0	131	146	0	0	102	103	25	55	0	0
125	33534	2314	111	81	222	125	10	15	67	53	2545	489	0	0	49	54	16	35	8	17
126	43986	5539	76	58	44	45	3	7	41	46	11244	2716	72	82	3791	6814	16	36	21	39
127	43891	1219	882	236	1775	133	52	59	19	33	318	699	22	48	505	599	0	0	5	12
128	72387	4720	1331	325	2595	278	16	22	284	545	26340	3019	138	57	1887	999	13	28	0	0
129	59980	944	736	138	1508	64	161	313	0	1	37044	5313	162	75	4723	4934	2	5	7	16
130	65921	3674	3407	362	5852	702	20	35	14	13	49210	6387	167	145	2821	1730	49	104	0	0
202	0	0	41	48	256	472	13	14	37	46	42	63	72	98	91985	29520	0	0	6	14
203	0	0	31	36	68	58	17	28	118	180	187	419	39	36	78743	25824	8	19	198	444
204	0	0	59	48	45	43	14	23	17	27	0	0	132	111	73735	15447	598	1337	0	0
201	8	10	66275	5354	74780	6251	4	7	348	675	83394	4029	4	9	2	3	843	385	2094	222
174	9967	1069	3	6	34	33	54124	5719	88028	2766	201	141	134	136	6	13	0	0	576	1149
205	7065	1214	44	53	30	59	65	49	95	58	43927	6005	64	81	0	0	416	242	790	221
206	5	5	11388	1480	11958	3915	2082	788	10047	3663	2836	359	0	0	0	0	308	100	552	74
131	8068	1417	165	47	241	173	36	36	83	176	3563	774	397756	19688	19	26	0	0	37	63
132	9728	1639	204	29	384	387	32	45	68	47	497	190	626925	7760	0	0	0	0	12	27
133	20246	2681	241	53	510	116	7	12	18	27	15987	1935	485644	38920	69	101	24	35	11	23

ACKNOWLEDGEMENTS

First of all, I would like to express my gratitude to Dr. Ulrika Westerlind for enabling me to work on this project, for scientific guidance, supervision of this thesis, various discussions and critical review of this manuscript. Thank you for creating this opportunity. I am thankful to Prof. Dr. Herbert Waldmann for providing an interdisciplinary environment and making it possible to work on this thesis. Also, I am grateful to Prof. Dr. Albert Sickmann for welcoming us and integrating our group into ISAS. The institute has been a stimulating place, which allowed insights into techniques and methodologies which lay outside the frame of this thesis.

I would like to thank Prof. Dr. Christian Hedberg for many ideas regarding chemistry, but also practical support, bypassing various shortages. Further, I am most grateful to him for being part of the PhD committee.

I also want to thank Prof. Dr. Horst Kunz for all the helpfulness and cooperation, especially for providing us with antiserum from vaccination experiments. Here, I am especially grateful to Dr. Nikola Gaidzik. I would like to acknowledge our collaborators Prof. Dr. Edgar Schmitt and Prof. Dr. Mengji Lu for vaccination experiments, Dr. Jonas Nilsson and his coworkers for the collaboration on the glycopeptide mass spectrometry project and Prof. Dr. Xi Chen, for providing us with PmST3 sialyltransferase.

Also I am grateful to the scientific and technical staff of ISAS: Helma Geltenpoth and Rita Fobbe for measuring HR-ESI-MS, Susanne Krois for performing amino acid analysis, Ingo Feldmann and Rolf Bandur for technical support on chromatography and various different laboratory equipment, the workshop and the house service of ISAS for help with installation of technical devices, the team of the ISAS IT department, Dr. Jörg Lambert for all the support around the NMR-instrument, the technical service of the TU Dortmund for measuring various NMR samples and Dr. Petra Janning from the Max-Plank Institut für Molekulare Physiologie for support on ESI-MS.

Further I am thankful for the financial support from the Verband der Chemischen Industrie for a Liebig PhD scholarship.

Finally I am very grateful to my fellow labmates. Dr. Hui Cai for all the collaboration on the vaccine project and good comradeship and also to my other labmates Vanessa, Jin and

Manuel for discussion, help and creating a nice working atmosphere. Many thanks to all the other members of ISAS for scientific exchange and as well for several distracting events around the work that made the time at the institute especially worthwhile.

My greatest gratitude, however, belongs to Martha for all the support and patience, no matter what happened.

THANK YOU!

EIDESSTATTLICHE ERKLÄRUNG

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe. Die Arbeit wurde weder bisher im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt. Es haben bisher keine Promotionsverfahren stattgefunden.

Dortmund, den 26.01.2016

_____

Christian Pett