isas

technische universität dortmund

"Chemical Tools for Protein Analysis: Development of an MS-cleavable Cross-linker and a New Approach for Glycopeptide Enrichment"

DISSERTATION

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

der Fakultät Chemie und Chemische Biologie

der Technischen Universität Dortmund

Vorgelegt von

M. Sc. Chem.

Vanessa Caixeta Pereira

aus Patos de Minas, Brasilien.

Dortmund 2018

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

Erstgutachter: Prof. Dr. Dr. Herbert Waldmann

Zweitgutachterin: Prof. Dr. Ulrika Westerlind

Publications:

A New Approach for Efficient Enrichment and Bead-Release of Intact Glycopeptides

Caixeta, V., Guldbrandsen, A., Vaudel, M., Zahedi, R. P., Larson, G., Berven, F., Nilsson, J., Westerlind, U., *Manuscript submitted.*

Impairment of angiogenesis by fatty acid synthase inhibition involves mTOR malonylation

Bruning, U., Morales-Rodriguez, F., Kalucka, J., Goveia, J., Taverna, F., Queiroz, K. C. S., Dubois, C., Cantelmo, A. R., Chen, R., Loroch, S., Timmerman, E., Caixeta, V., Bloch, K., Conradi, L. C., Treps, L., Staes, A., Gevaert, K., Tee, A., Dewerchin, M., Semenkovich, C. F., Impens, F., Schilling, B., Verdin, E., Swinnen, J. V., Meier, J. L., Kulkarni, R. A., Sickmann, A., Ghesquiere, B., Schoonjans, L., Li, X., Mazzone, M., Carmeliet, P., *Cell Metab.* **2018**, *In Press*.

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

Pett, C., Nasir, W., Sihlbom, C., Olsson, B.-M., Caixeta, V., Schorlemer, M., Zahedi, R. P., Larson, G., Nilsson, J., Westerlind, U., *Angew. Chem. Int. Ed.* **2018**, *57*, 9320.

Further participation in the following works:

Design and synthesis of *N*-acylated aza-goniothalamin derivatives and evaluation of their *in vitro* and *in vivo* antitumor activity

Barcelos, R. C., Pastre, J. C., Vendramini-Costa, D. B., Caixeta, V., Longato, G. B., Monteiro, P. A., de Carvalho, J. E., Pilli, R. A., *Chem. Med. Chem.* **2014**, *9*, 2725.

Synthesis of methoxylated goniothalamin, aza-goniothalamin and gamma-pyrones and their *in vitro* evaluation against human cancer cells

Barcelos, R. C., Pastre, J. C., Caixeta, V., Vendramini-Costa, D. B., de Carvalho, J. E., Pilli, R. A., *Biorg. Med. Chem.* **2012**, *20*, 3635.

Table of Contents

Abbreviations

AAA	amino acid analysis
Ac	acetyl
APP	amyloid precursor protein
Ay	acryloyl
azide-A-DBSO	azide-tagged, acid-clevable disuccinimidyl bis-sulfoxide
BabA	blood group-antigen binding adhesin
BID	*N*-benzylimidodiacetoyloxysuccinimid
Boc	*tert*-butyloxycarbonyl
BSA	bovine serum albumin
BS^3	bis (sulfo succinimidyl) surberate
Bu	butyl
calc.	calculated
cat.	catalytic
CDG	congenital disorders of glycosylation
CID	collision-induced dissociation
CFG	Consortium for Functional Glycomics
ConA	concanavalin A
CSF	cerebrospinal fluid
d	doublet
Da	dalton
DCM	dichloromethane
DHB	2,5-dihydroxy benzoic acid
DIC	*N*,*N*'-diisopropylcarbodiimide
DIPEA	diisopropylethylamine
DMAP	4-(dimethylamino(pyridine)
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DSG	disuccinimidyl glutarate
DSS	disuccinimidyl surberate
DTSSP	3,3'-dithiobis (sulfosuccinimidyl propionate)
DTT	1,4-dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
eq.	equivalents
ER	endoplasmic reticulum
ERLIC	electrostatic repulsion hydrophilic interaction chromatography
ESI	eletron spray ionization
ECD	electron capture dissociation
ETD	electron transfer dissociation
EThcD	electron-transfer and higher-energy collisional dissociation
FA	formic acid
Fmoc	*N*-(9H-fluoren-9-yl)-methoxycarbonyl
Fuc	*L*-fucose
Gal	*D*-galactose
GalNAc	*N*-acetyl-*D*-galactosamine
Glc	*D*-glucose
GlcA	glucuronic acid
GlcNAc	*N*-acetyl-*D*-glucosamine
Glu-C	endoproteinase Glu-C
GOase	galactose oxidase
h	hours
HBTU	*O*-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HCD	high-energy C-trap dissociation higher-energy collisional dissociation
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hex	hexose
HexNAc	*N*-acetylhexosamine
HILIC	hydrophilic interaction chromatography
HOBt	1-hydroxybenzotriazole
HPLC	high performance liquid chromatography
HR	high resolution
Hz	hertz
I	relative intensity
IAA	iodoacetamide
IdoA	iduronic acid
Ig	immunoglobulin
IMS	ion mobility spectrometry
iTRAQ	isobaric tags for relative and absolute quantitation
J	coupling constant
LC	liquid chromatography
LDS	lithium dodecyl sulfate
L-PHA	L-phytohemaggutinin
ls	large singlet
Lys-C	endoproteinase Lys-C
m	multiplet
M	molarity
MALDI	matrix assisted laser desorption ionization
Man	*D*-mannose

MeCN	acetonitrile
min	minutes
MOPS	3-(*N*-morpholino)propanesulfonic acid
MS	mass spectrometry
MS/MS MS^n	tandem mass spectrometry
MUC	mucin type peptide
m/*z*	ratio mass-charge
NCBI	National Center for Biotechnology Information
NCE	normalized collision energy
Neu5Ac	*N*-acetylneuraminic acid
Neu5Gc	*N*- glycolylneuraminic acid
NHS	*N*-hydroxysuccinimide
NMR	nuclear magnetic resonance
NOG	*N*-octyl-β-*D*-glucopyranose
ox	oxidized
PBS	phosphate buffered saline
PEG	polyethylene glycol
Pep	peptide
PGC	porous graphitized carbon
Ph	phenyl
pH	potentia hydrogenii
PNA	peanut agglutinin
PNGase F	peptide-*N*-glycosidase F
Pp	phosphonate
ppm	parts per million
PTM	post-translational modification
PXL	phosphonate cross-linker
q	quartet
qui	quintet
R_f	retention factor
RNase S	ribonuclease S
rpm	rotation per minute
r.t.	retention time
s	singlet
SA	sialic acid
SabA	sialic acid-binding adhesin
SDS	sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SLD	soft laser desorption
SPPS	solid phase peptide synthesis
Su	succinimide
SuDP	bisuccinimidyl-succinamyl-aspartyl-proline
sulfo-SBP	sulfosuccinimidyl benzophenone
sulfo-SDA	sulfosuccinimidyl 4,4'-azipentanoate
SXC	size exclusion chromatography
t	triplet
T	Thomsen-Friedenreich antigen
*t*Bu	*tert*-butyl
TCEP	*tris*(2-carboxyethyl)phosphine
TEAB	triethylammonium bicarbonate
tert	tertiary
TFA	trifluoroacetic acid
TIPS	triisopropylsilane
TLC	thin layer chromatography
TMT	tandem mass tag
T_n	T-antigen nouvelle
u	unit
u	atomic mass unit
UDP	Uridine diphosphate
UV	ultraviolet
XL	cross-link/er
Xyl	*D*-xylose
VVA	vicia villosa agglutinin
WGA	wheat germ agglutinin
ZIC-HILIC	zwitterionic HILIC
Å	angstrom
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

Abstract

The technological development in analytical methods, especially in chromatography and mass spectrometry, has allowed a deeper understanding of biological systems at a molecular level. Characterization and quantification of biomolecules and its modifications are crucial for comprehension of the entire biological machinery. However, additional bioanalytical methods are needed to gain more detailed information about protein structure, dynamic protein complexes and low abundant protein modifications. Herein chemical tools are used in addition to well-established shotgun proteomic approaches in order to enhance the knowledge of protein structure and information about glycosylation as a protein modification.

Cross-linkers are used as molecular rulers or molecular bridges to connect two amino acid residues at a specific distance via covalent bonds. These tools assist in the characterization of the three-dimensional structure of a protein or of protein complexes. In the first part of this work a new MS-cleavable phosphonate cross-linker was proposed, with the aim to perform a neutral elimination reaction during the tandem MS fragmentation. The new phosphonate cross-linker PXL **14** was synthesized in 3 steps from commercial sources. This linker was evaluated using synthetic peptides. The linkage reaction was optimized by adjusting the buffer system and pH value in order to avoid *in-situ* chemical cleavage. Distinct fragmentation techniques were applied to observe the linker fragmentation in tandem mass spectrometric experiments. Low-energy HCD fragmentation showed favorable cleavage of the cross-linker bonds instead of the peptide backbone fragmentation. The acryloyl fragment of the Ac-IEAE***K***GR peptide is the second most intense ion in HCD with 18% NCE. Bovine serum albumin was used as model protein to evaluate PXL in a more complex system. The cross-linker reaction with BSA was monitored using SDS-PAGE. After optimization of conditions for cross-linker reaction and proteolytic digestion, cross-linked peptides were analyzed using LC-MS/MS methods. Manual data analysis allowed the identification of 7 intra-linked peptides, including 2 non-previously reported; 3 loop linked peptides and 8 peptides derived from chemical cleavage.

Chemical tools were also developed in this work to improve glycopeptide enrichment in studies of glycan protein modifications. Glycosylation is one of the most common post-translational protein modifications and one of the most complex modifications due to the diversity in the monosaccharides building blocks, the linkage types among those monosaccharides and their extension. Besides micro- and macroheterogeneity, the low abundance of modified peptides after proteolytic protein degradation is a major obstacle for analysis of glycopeptides using the standard proteomic workflows. Therefore, an efficient glycopeptide enrichment strategy is required prior to LC-MS/MS analysis. In this work an improved enrichment methodology is proposed based on the chemical properties of glycopeptides/glycoproteins. A smart cleavable disulfide linker containing hydroxylamine groups was attached on the surface of beads. The linker **40** was synthesized in 6 steps from commercial source prior to the bead surface modification. The hydroxylamine group reacts with aldehydes leading to a covalent oxime bond. Reactive aldehyde groups were introduced on terminal glycan residues, e.g. via selective oxidation of sialic acid using sodium periodate or via enzymatic oxidation of terminal galactose by galactose oxidase. In this approach both, neutral and negatively charged, synthetic glycopeptides were enriched and released under reductive conditions by disulfide bond cleavage. This strategy allowed glycopeptide release independently of glycan nature and peptide backbone. The released glycopeptides presented only a minor modification of the chemical

structure in the terminal glycan residue. Moreover, we demonstrated that this approach is compatible with tandem mass tag labeling, which would enable quantification with current available TMT methodologies. Amino acid analysis demonstrated that the enrichment of synthetic sialylated glycopeptide had a yield of 32% and a yield of 40% for a synthetic galactosylated peptide.

The enrichment method was also evaluated using model proteins. In this case, fetuin and asialofetuin were applied using different oxidation methods: sodium periodate and GOase oxidation, respectively. For the fetuin sample, a total of 28 unique glycopeptides were identified. Among them, 7 were unique *O*-glycopeptides and 21 were unique *N*-glycopeptides. These peptides covered 2 *O*- and 2 *N*-glycosylation sites from fetuin A, and 2 *O*-glycosylation sites from fetuin B. On the other hand, 98 unique glycopeptides were identified for asialofetuin. Herein, 11 unique *O*-glycopeptides covered 1 fetuin A glycosylation site and 2 fetuin B glycosylation sites. For the *N*-glycopeptides 87 unique peptides were covered on 2 glycosylation sites from each fetuin A and B. Serum and CSF were used to expand the sample complexity scope of the improved enrichment methodology. In serum a total of 34 proteins and 162 unique sialylated glycopeptides were identified by our methodology. In a CSF sample 32 proteins and 129 unique sialylated glycopeptides were further identified.

The *N*-glycopeptide enrichment method using hydrazine beads is typically used for determination of *N*-glycosylation sites after removal of the glycans during enzymatic bead release. Clinical CSF samples were here also used for the comparison of *N*-glycopeptide enrichment, by using this method employing hydrazine beads and enzymatic release with PNGase F and the new method containing a cleavable linker on hydroxylamine beads with chemical and/or enzymatic PNGase F release. A total of 2 963 peptides were identified by samples enriched with hydrazine and hydroxylamine beads together and enzymatic peptide release. While 350 unique glycopeptides were identified from 80 different proteins in the samples enriched by hydroxylamine beads and mild chemical cleavage. Among these identifications, 251 were *O*-glycopeptide structures and 222 were *N*-glycopeptide structures. As a conclusion, the new enrichment method allows parallel analysis of *O*- and *N*-glycopeptides from both galactosylated and sialylated peptide classes under mild chemical bead release conditions. Besides leaving the glycan structure intact, the method enables labeling with a TMT for quantitative MS detection and is also suitable to be explored in further large scale glycoproteomic studies.

Zusammenfassung

Die technische Entwicklung analytischer Methoden, speziell der Chromatographie und Massenspektrometrie erlaubt ein tieferes Verständnis biologischer Systeme auf molekularer Ebene. Charakterisierung und Quantifizierung von Biomolekülen und ihrer Modifikationen sind entscheidend für das Verständnis biologischer Mechanismen. Zusätzliche bioanalytische Methoden sind notwendig um detaillierte Informationen über Proteinstrukturen, dynamische Proteinkomplexe oder Proteinmodifikationen mit geringer Häufigkeit zu erhalten. Hierfür werden chemische Werkzeuge zusammen mit bereits etablierten Shotgun-Ansätzen genutzt, um Erkenntnisse über Proteinstrukturen und Glykosylierungen zu erweitern.

Cross-Linker dienen als molekulare Lineale, die zwei Aminosäureseitenketten in einem definierten Abstand kovalent miteinander verknüpfen. Sie sind hilfreich bei der Charakterisierung dreidimensionaler Protein- bzw. Proteinkomplex-Strukturen. Im ersten Teil dieser Arbeit wird ein MS-spaltbarer Phosphonat-Cross-Linker (PXL) vorgestellt, der eine neutrale Eliminierung während der Tandem-MS-Fragmentierung ermöglichen soll. Der neue Phosphonat-Cross-Linker PXL **14** konnte ausgehend von kommerziell erhältlichen Edukten in drei Stufen dargestellt werden. Die Evaluierung wurde an synthetischen Peptiden vorgenommen. Die Optimierung der Reaktionsbedingungen der Verknüpfung erfolgte hinsichtlich des Puffer-Systems und pH-Werts, um eine *in-situ*-Freisetzung des Cross-Linkers zu vermeiden. Verschiedene Fragmentierungstechniken wurden zur Erforschung der Linker-Fragmentierung in Tandem-Massenspektrometrieexperimenten untersucht. Die gewünschte Abspaltung des Linkers konnte mittels *Low-energy-HCD*-Fragmentierung realisiert werden ohne dabei das Peptidrückgrat zu beschädigen. Das Acryloyl-Fragment des Ac-IEAE***K***GR-Peptids ist das zweitintensivste Ion nach HCD-Fragmentierung bei 18% NCE. Als Modellprotein zur Evaluierung der PXL in komplexeren Systemen fungierte Bovines Serumalbumin (BSA). Die Verknüpfung des Cross-Linkers mit BSA wurde mittels SDS-PAGE verfolgt. Nach Optimierung der Bedingungen der Cross-Linker-Kupplung und der proteolytischen Verdauung dienten LC-MS/MS-Methoden zur Analyse der cross-linker-verknüpften Peptide. Die manuelle Datenanalyse erlaubte die Identifizierung von sieben intramolekular verknüpften Peptiden, inklusive zweier bisher nicht dokumentierten; drei schleifen-verknüpften Peptiden und acht Peptiden als Resultat einer chemischen Spaltung.

Des Weiteren wurden chemische Methoden zur Verbesserung der Anreicherung von Glykopeptiden zur Untersuchung von Glykosylierungsmodifikationen entwickelt. Die Glykosylierung stellt eine der am häufigsten vorkommenden und auf Grund der hohen Diversität an Monosacchariden und an Möglichkeiten derer Verknüpfung auch komplexesten posttranslationalen Modifikationen von Peptiden dar. Neben der Mikro- und Makroheterogenität stellt einem vor allem die geringe Menge an modifizierten Peptiden nach erfolgtem proteolytischen Abbau in der Analytik von Glykopeptiden mit Hilfe von Standard-Proteomic-Verfahren vor Herausforderungen. Aus diesem Grund ist eine effiziente Strategie zur Anreicherung von Glykopeptiden für LC-MS/MS-Analysen erforderlich. In dieser Arbeit wird eine verbesserte Anreicherungsmethode basierend auf den chemischen Eigenschaften von Glykopeptiden/-proteinen vorgestellt. Ein abspaltbarer Disulfid-Linker mit Hydroxylamin-Funktionen wurde auf der Oberfläche von Beads angebracht. Linker **40** konnte in sechs Stufen ausgehend von kommerziell erhältlichen Edukten dargestellt werden. Über die Hydroxylamin-Gruppen des Linkers können kovalente Verknüpfungen mit den Aldehyd-Funktionen der terminalen Monosacchariden der Glykosylierungsmodifikation unter Ausbildung von

Oxim-Bindungen geschlossen werden. Die Einführung dieser Aldehyd-Funktionen kann z.B. mittels selektiver Oxidation von Sialinsäuren mit Natriumperiodat oder enzymatischer Oxidation von terminaler Galactose mit Galactose-Oxidase erreicht werden. Bei diesem Ansatz werden sowohl neutrale als auch negativ geladene, synthetische Glykopeptide angereichert und über die Spaltung der Disulfid-Bindung unter reduktiven Bedingungen freigesetzt. Diese Art der Freisetzung ist unabhängig von der Natur der Glykanstruktur und des Peptidrückgrates. Die freigesetzten Glykopeptide weisen dabei nur eine geringfügige chemische Modifikation in terminaler Position auf. Darüber hinaus konnte gezeigt werden, dass dieser Ansatz für Tandem Mass Tag (TMT) Labeling geeignet ist. Dies würde eine Quantifizierung mit derzeitigen TMT-Methoden ermöglichen. Aminosäureuntersuchungen konnten aufzeigen, dass die Ausbeute der Anreicherung von synthetischen sialylierten Glykopeptiden bei 32% und von synthetischen galactosylierten Glykopeptiden bei 40% liegt.

Die Anreicherungsmethode wurde ebenso mit Hilfe von Fetuin und Asialofetuin als Model-Proteinen evaluiert. Für die Oxidation wurde entsprechend Natriumperiodat bzw. Galactose-Oxidase eingesetzt. In der Fetuin-Probe konnten 28 einzigartige Glykopeptide identifiziert werden; darunter sieben *O*-Glykopeptide und 21 *N*-Glykopeptide. Diese korrespondieren mit zwei *O*- und zwei *N*-Glykosylierungsseiten des Fetuin A und zwei *O*-Glykosylierungsseiten des Fetuin B. In der Asialofetuin-Probe wurden 98 einzigartige Glykopeptide identifiziert. Die elf *O*-Glykopeptide korrespondieren mit einer *O*-Glykosylierungsposition des Fetuin A und zwei *O*-Glykosylierungsseiten des Fetuin B. Die 87 *N*-Glykopeptide korrespondieren jeweils mit zwei *N*-Glykosylierungspositionen des Fetuin A bzw. Fetuin B. Um die Anreicherungsmethode für die Anwendung auf komplexere Systeme untersuchen zu können, wurden Blutserum und CSF als Proben gewählt. Unsere Methode ermöglichte im Blutserum 34 Proteine und 162 einzigartige sialylierte Glykopeptide zu identifizieren. In der CSF-Probe wurden 32 Proteine und 129 einzigartige sialylierte Glykopeptide identifiziert.

Die Methode zur Anreicherung von *N*-Glykopeptiden mittels Hydrazin-Beads wird typischerweise zur Bestimmung von *N*-Glykosylierungsseiten nach enzymatischer Freisetzung der Oligosaccharide genutzt. Klinische CSF Proben wurden zum Vergleich der *N*-Glykopeptid-Anreicherung herangezogen; wobei Hydrazin-Beads für enzymatische Freisetzung mittels PNGase F unserer Methode mit abspaltbarem Linker auf Hydroxylamin-Beads für chemische oder enzymatische Abspaltung mit PNGase F gegenübergestellt werden sollten. 2 963 Peptide wurden mit Hilfe der Anreicherung sowohl von Hydrazin- als auch Hydroxylamin-Beads mit enzymatischer Freisetzung identifiziert. Bei Verwendung der Hydroxylamin-Beads in Kombination mit milder, chemischer Freisetzung war eine Identifikation von 350 einzigartigen Glykopeptiden der 80 verschiedenen Proteinen möglich. Unter diesen waren 251 *O*-Glykopeptidstrukturen und 222 *N*-Glykopeptidstrukturen zu beobachten. Zusammenfassend ermöglicht die neue Methode, basierend auf der chemischen Freisetzung der Beads unter milden Bedingungen, eine parallele Untersuchung von galactosylierten und/oder sialylierten *O*- und *N*-Glykopeptiden. Da die chemische Glykanstruktur nahezu unverändert bleibt, sind TMT Markierungen für quantitative Massenspektrometrie und weiterführende Studien auf dem Gebiet der Glykoproteomics im größeren Maßstab möglich.

Historical Aspects of Mass Spectrometry in Proteomic Research

3.1- General background

Mass spectrometry (MS) is an analytical method that provides information about the ratio mass/charge (*m*/*z*) of an atom or molecule. This technique requires that the atom or molecule is brought to gas phase using an ionization source; the ion beam is directed to the mass analyzer using vacuum where they are sorted according to their *m*/*z* and transmitted to the detector in order to obtain the signal (Figure 1). The standard unit for atomic mass recommended by International Union of Pure and Applied Chemistry (IUPAC) is the *unified atomic mass unit* (*u*), which is defined as 1/12 of ^{12}C mass. The *dalton* (Da) is often used in biological and biochemistry applications and it is identical to u.

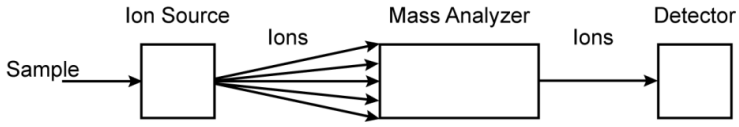


Figure 1. Diagram of a mass spectrometer system: first the sample is carried to the gas phase by an ionization source, later the ions are transmitted to the mass analyzer for separation according to the *m*/*z* and finally measured by a detector.

For a long time, the use of this technique was limited to compounds in gas phase. Thomson constructed in Cambridge the first mass spectrometer, called parabola spectrograph, at the beginning of the 20th century.[1] He obtained mass spectra of gases compounds as O_2, N_2, CO and CO_2 and observed charged ions with this instrument. Therefore this methodology was used for analysis of gas compounds and isotope separation. Biological applications began in the 40's when CO_2 production in animals was monitored using heavy stable isotopes traces.[2]

In the following years, applications in organic chemistry showed the potential of the MS for studies of organic compound properties[3] and chemical reactions[4]. By the 80's, analysis of organic molecules by mass spectrometry became routine, nevertheless, the technique was limited to small compounds. The first complete mass spectrum of bovine insulin (5 730Da) was published in 1982 by Haakansson *et. al.* also suggested that the MS limitations could be circumvented.[5] Although the concept of a suitable mass analyzer for huge molecules was proposed in the 40's[6] and 50's[7], analysis of biomolecules such as proteins, nucleic acids, and complex carbohydrates remained a challenge. The main issue was to bring those macromolecules to the gas phase without fragmentation or decomposition.

One breakthrough in this field happened when John Fenn in 1989 used electrospray (ESI) as ionization technique to analyze entire proteins such as cytochrome c (12 400Da), myoglobin (17 000Da), alcohol dehydrogenase (39 800Da) and conalbumin (76 000Da).[8] The electrospray is produced when the charged (acidic or basic) and very diluted solution of a compound pass through a capillary tube with a reduced flux to an electric field, the solvent in those droplets evaporates in

atmospheric pressure to a point that, due to columbic repulsions, the droplets break into unique ionized compound unities that can then passes through the mass analyzer (Figure 2a).

In parallel to John Fenn's work, Koichi Tanaka from Shimadzu used a mixture of glycerol and metal nanoparticles as matrix to ionize and vaporize macromolecules under direct laser pulse radiation without fragmenting their structure, a process later called soft laser desorption (SLD).[9] The detection of lysozyme (14 306Da) and chymotrypsinogen (25 717Da) together with polymers poly(propylene glycol) (4kDa) and poly(ethylene glycol) (20kDa) was reported in 1988.[10] Also in the 80's, Hillenkamp and collaborators, during the studies of laser desorption of small organic molecules, observed that the necessary energy to ionize both tryptophan and alanine together was smaller than the energy for alanine alone. In this way, they searched for chemical substances to work as matrix that allows the desorption of proteins, and developed, also in 1988, the MALDI: matrix-assisted laser desorption ionization (Figure 2b).[11] Spectra of proteins like porcine trypsin (23 463Da) and bovine albumin (67 000Da) in nicotinic acid matrix in the subnanograms range were reported in the following years.[12] For reasons including higher sensitivity compared to SLD, MALDI was the laser desorption ionization technique adopted by most of the mass spectrometric community. The importance of ionization techniques gave Tanaka and Fenn the Nobel Prize in Chemistry in 2002 for "the development of soft desorption ionization methods for mass spectrometric analyses of biological macromolecules".[13]

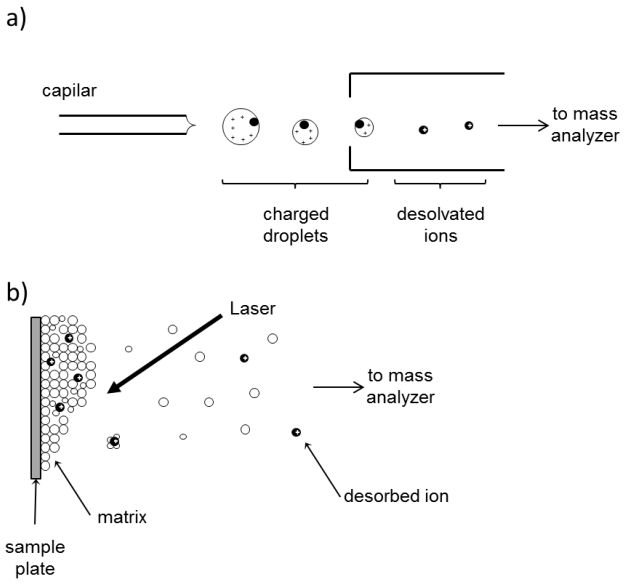


Figure 2. Common ionization methods: a) electrospray (ESI), the charged droplets has the solvent evaporated under atmospheric pressure and the desolvated ions are directed to the mass analyzer; b) matrix-assisted laser desorption (MALDI), pulsed laser desorbs the analyte from the solid mixture with a compound called matrix.

The successful development of suitable ionization methodologies for biomolecules encouraged further studies on protein structure by mass spectrometry. In 1990, Chowdhury *et. al.* presented the peptide mapping of apolipoprotein A1 (28kDa) after tryptic digestion and analyzation of the fragments using mass spectrometry without chromatographic fractionation or chemical

derivatization.[14] And just a few years later, Pappin *et. al.* developed the automated software MOWSE for determination of peptide masses compared to the fragment masses database.[15] With all technological developments in analytical chemistry and emerging of genome sequencing studies, the interest of understanding biological systems raised. The term proteome was used for the first time in 1995 as the entire proteins expressed by a genome, cell or tissue type.[16] Proteomic became then a field for systematic studying of protein properties, providing a detailed description of structure, function and control of biological systems.[17]

Within a short time, mass spectrometry took the place of Edman's degradation[18] as methodology for protein/peptide analysis, due to its sensitivity, speed (peptides can be fragmented in second instead of hours or days), requiring less purification of proteins and compatibility with protein modifications. The coupling between mass spectrometry and separation methods as high-performance liquid chromatography (HPLC)[19] in hands with the development of software for peptide analysis[20-21] allowed the high-throughput screening and identification of a huge number of proteins in shotgun proteomic experiments in the 90's. In these experiments, proteins from a cell culture or tissue are firstly fractionated by electrophoresis methodologies. In sequence the proteins are digested by one or more proteases (trypsin, Lys-C or Glu-C), the peptides are separated via HPLC and directly analyzed using tandem mass spectrometry methodologies. In the end, the mass spectra are compared with the theoretical (*in silico* digested) spectra by a software that matches the predicted to experimental fragments (Figure 3).

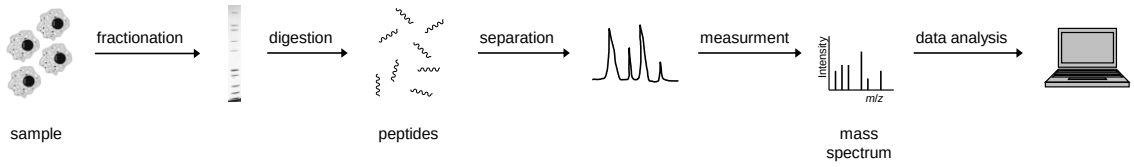


Figure 3. Steps of a proteomic shotgun experiment: a sample from a biological source is fractionated: separating proteins often in an SDS-PAGE. Proteins are then digested by one or more protease (trypsin) and the peptide mixture is separated off- or online using chromatographic methods (SXC, HPLC). Peptides are analyzed in the mass spectrometer and the obtained data are later searched against a database using specific programs (Mascot, Sequest) for identification and/or characterization of sample components.

Tandem mass spectrometry (MS/MS or MS^n) experiment comprises the selection of ion and fragmentation of those ions upon different techniques. The first time a protein sequence was analyzed by MS/MS was in 1986. Apolipoprotein B was digested with trypsin and modified with cyanogen bromide, fractionated by liquid chromatography and directly analyzed in the mass spectrometer.[22]

Precursor ions, selected ions with a certain *m*/*z*, are submitted to dissociation methods for formation and detection of new fragment ions or product ions. One type of dissociation is when an inert gas (He, N_2 or Ar) transfers its kinetic energy via collision to the molecular ion that converts it in vibrational energy promoting the ion bond cleavage. This is called collision-induced dissociation (CID)[23-24]. In the 70's, analysis of peptide fragments showed the preference of amide bond cleavage under CID dissociation, generating b- and y- fragment types (see Figure 4 for the explanation of

fragment ion types). For this reason, it is the main fragmentation technique used for sequencing of digested peptide mixture. Roepstorff and Fohlmann[25] proposed a common nomenclature for peptide fragment ions generated in mass spectrometry. This nomenclature was later optimized by Biermann[26] and adopted by the scientific community (Figure 4). However, CID dissociation suffers some limitations: large peptides (>15 amino acids) present incomplete fragmentation and limited scope for characterizing post-translational modifications (PTM) and their sites since a lot of those modifications are easily cleaved off. Additionally in relation to mass spectrometer setup, for a more efficient fragmentation, the ion trap does not allow the trapping of fragments below 28% of the precursor mass, subjecting this fragmentation to "low mass cutoff", which is known as "1/3 Rule". Therefore, potentially important ions as y1 and b1 are missed.

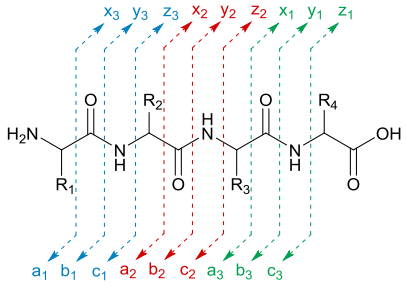


Figure 4. Designation for peptide fragment ions according to Roepstorff-Fohlmann-Biemann nomenclature.[25-26]

McLafferty and collaborators found accidentally that a low-energy electron transferred to protein cations cleaves C-N bonds forming ***c*** and ***z*** ions.[27] Some years later, Syka *et. al.* optimized the technique by changing the mass spectrometer setup in order to transfer just one electron to the peptide cation, promoting the same bond cleavage previously described and called it electron transfer dissociation (ETD).[28] This new fragmentation method requires ions in high charge state (>2) for efficient fragmentation, also low fragmentation efficiency and limitations for *m/z* range less than 1400 are reported.[29] Even with those limitations, ETD dissociation became popular and relevant especially for post-translational modification identification, such as glycosylation, phosphorylation and disulfide linkages, for retaining the modifications during peptide backbone fragmentation which allows the identification of protein modification sites.[30]

Matthias Mann's group in collaboration with Thermo Fisher adopted Orbitrap instruments in a way that the C-trap, a compartment used until then for ion storage before analysis in the Orbitrap, could also be used as a collision chamber for tandem mass spectrometry experiments.[31] The collision dissociation that happened in the C-trap was called high-energy C-trap dissociation (HCD). When those fragment ions are analyzed in the Orbitrap, high accuracy and full mass range are obtained for MS/MS spectra, without the limitation of the "1/3 Rule". Although this fragmentation is also known as "higher-energy collisional dissociation", the dissociation process is similar to CID dissociation: an inert gas transfers energy via collision to the selected ion that decomposes in fragments. While in the CID each collision is converted to vibrational energy, in the HCD each

collision promotes a bond fragmentation. However “higher-energy” in the original name refers to a higher radiofrequency voltage used in the C-trap to retain the maximum amount of fragment ions.[31]

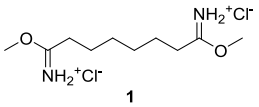
In order to optimize the efficiency of the ETD fragmentation, Swaney *et. al.* had the idea to include a supplemental collisional activation that targets species that has previously received an electron during the process for ETD dissociation.[32] However, the information generated by ETD is lost due to the isolation of the reduced charge ion for further fragmentation with a supplemental energy. To make a better use of all information obtained in all dissociation steps, Frese *et. al.* reported in 2012 a double fragmentation strategy implementing HCD dissociation after an electron transfer reaction from ETD in a way that all ions generated from ETD (c- and z-type) and HCD (b- and y-type) are detected and used for the precursor identification.[33] This fragmentation approach was named electron-transfer and higher-energy collision dissociation (EThcD) by the authors. This approach potential, especially for identification of label modification, was demonstrated by comparing the localization score of phosphorylation sites with the scores obtained from ETD and HCD alone. The spectra containing more fragments, as those obtained by EThcD, provided better confidence for PTM identification.[34]

The constant development of instrumentation and methods for protein analysis in mass spectrometry tremendously raised the number of publications. This phenomenon was clearly exemplified in the phosphoproteomic field: until the 50’s, there were just 13 publications regarding phosphorylation and in 2012 nearly 13 000 articles related to phosphorylation were published in NCBI-listed articles.[35] Mass spectrometry is a powerful instrument for analysis of not only proteins but other sorts of biomolecules as carbohydrates[36], lipids[37], nucleic acids[38] and metabolites[39] as well. Characterization of protein structures[40] and intact complexes[41] have been carried out using MS. Quantification between two or more physiological states[42] opens the possibilities of studying e.g. healthy and disease modifications in a molecular level of proteins[43]. Nevertheless, until now the full understanding of biological systems is far from being complete, therefore new analytical methodologies and instrument improvement are required to cover the limitations of current technology. Still, high accuracy and sensitivity of analysis in mass spectrometry hint that in the future shotgun proteomic analysis could become a medical diagnotic tool for personalized treatment.[44]

Development of an MS-Cleavable Cross-Linker

4.1- Background: The development of MS-cleavable cross-linkers

One of the main goals in current biological research is the comprehension of the protein networks at a molecular level. For this reason, it is crucial to study such interactions in their natural environment, since most of the time these interactions occur by weak forces as electrostatic interactions, hydrogen bonds, and van der Waals interactions. In 1970, Davies and Stark used dimethyl suberimidate **1** to cross-link the side chain amino groups of isolated proteins and native aspartate transcarbamylase complexes of *E. coli*, suggesting the possibility of structure and stoichiometry studies of protein complexes.[45] Since then, cross-linkers are used, together with mass spectrometry, as a powerful technique for understanding protein structure and protein complexes interactions.



1

Although considered a method for low resolution for solving structures, cross-linkers associated to mass spectrometry have the ability to overcome limitations of traditional techniques in protein structure determination. X-ray crystallography, nuclear magnetic resonance (NMR) and cryo-electron microscopy offer an atomic resolution in solving structures, however, all these methodologies are applied for proteins in an artificial environment and often are able to solve only partial structures. The X-ray crystallography is limited for highly purified proteins that form suitable crystals. Moreover, there is a lack of crytallography data of large protein complex. NMR structure resolution occurs in solution but is limited for small proteins and, as the X-ray crystallography, requires a protein with a high purity level. [46-48] While cryo-electron microscopy solves typically large, homogenous and rigid proteins.[49] On the other hand, the cross-linkers promote contrictions in the protein/complex structure providing important and accurate three-dimension information allied to a sensitive and rapid method for protein analysis. Besides that, the cross-linker diversity allows the method adaptation depending on the conditions required for the tagged protein/complex, e.g. membrane proteins requires more hydrophobic reaction conditions while cytosolic proteins might require hydrophilic conditions under physiological pH. These characteristics allow the cross-linking/mass spectrometry to study protein and protein complexes in their native environment.[50]

Cross-linkers reacts by forming a covalent bond between two spatial neighboring amino acid residues, within one protein or within a interacting protein complex. Nevertheless side reactions might occur by forming an undesired cross-linked product between two neigbouring amino acid residues in the protein sequence or by deactivation of one cross-linker side.[51] After protein cross-linking, the reaction products are proteolytically digested, isolated and analysed by nano-HPLC tandem MS. In Figure 5, possible reaction products from a general cross-link reaction between peptides of two different proteins (blue and red) are demonstrated.[52] Type 0 presents one

cross-linker side reacted with the peptide and the other side deactivated, leading to a “dead end”. Type 1 is also called loop-linked because both cross-link sides react with two close residues in close proximity within a protein sequence, which means that the two amino acid residues of the same peptide fragment are connected. Finally, the type 2 is subdivided into two categories: intra- and inter-linked residues. The intra-linked peptide is the consequence of the reaction of two amino acid residues spatially close, but sometimes far from each other in the peptide sequence, thus resulting in two different connected peptides of the same protein. On the other hand, the inter-linked peptides are obtained when cross-linking is successful between two different proteins, which often is the desired reaction product and provides the most important information regarding protein complexes. The described nomenclature for the different cross-linked products may vary within the published literature.

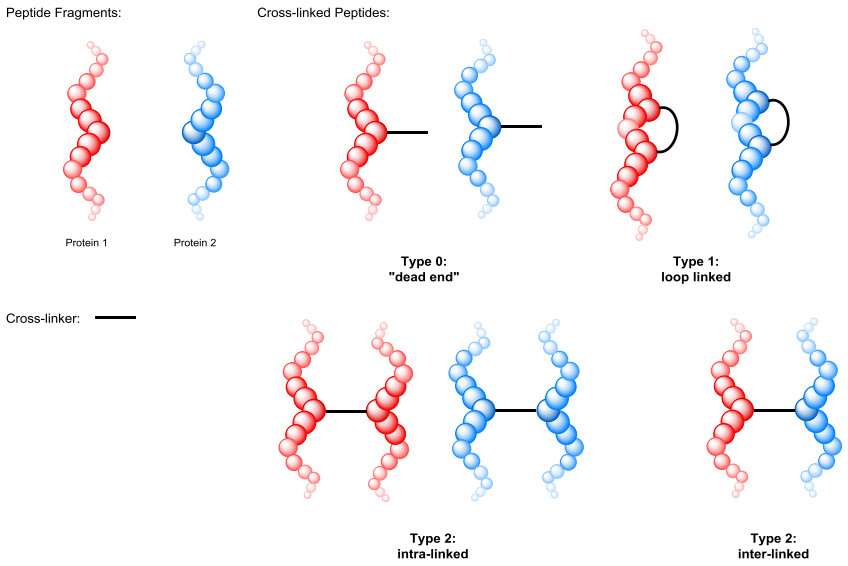


Figure 5. Products in the cross-link reaction between two peptides of different proteins (blue and red), according to Schilling *et. al.*[52]

The linkers can be classified according to their function (amino-reactive or thiol-reactive) and symmetry (homo- and heterobifunctional). The most commonly used are the homobifunctional cross-linkers addressed to primary amines (lysine), containing *N*-hydroxysuccinimidyl (NHS) ester groups. One example is the disuccinimidyl suberate **2** (DSS), which is a simple alkane chain with a *N*-hydroxysuccinimidyl ester in each end of the linker. The simplicity of the structure inspired the development of similar structures for different purposes: to increase hydrophilicity, bis(sulfo succinimidyl) suberate **3** (BS^3); to achieve shorter distances, disuccinimidyl glutarate **4** (DSG)[53]; or to facilitate MS identification and quantification, D_{12}-disuccinimidyl suberate **5**[54]. The accessibility of the primary amine groups in a hydrophilic environment and high prevalence of lysine in proteins (~6% of all residues) are the main reasons for targeting primary amine groups. In order to study hydrophobic protein regions, such as membrane protein complexes, or proteins with lower lysine content, photoactive cross-linkers are and atractive alternative. In general, photoactive linkers are heterobifunctional reagents, combining NHS ester groups with diaziridine **6** or benzophenones **7**,[55]

targeting on one side lysines and on the other side unspecific amino acid residues. However, data search and analysis of those cross-linked products might be more complicated due to unspecific reaction towards amino acid residues.

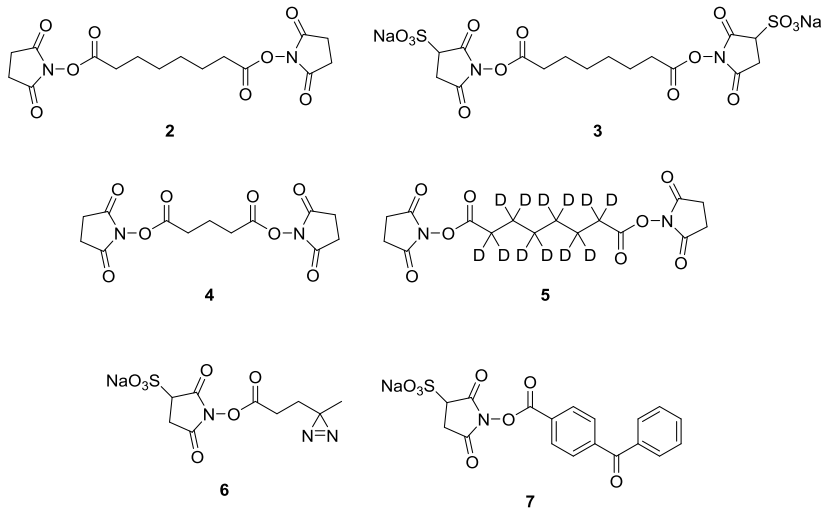
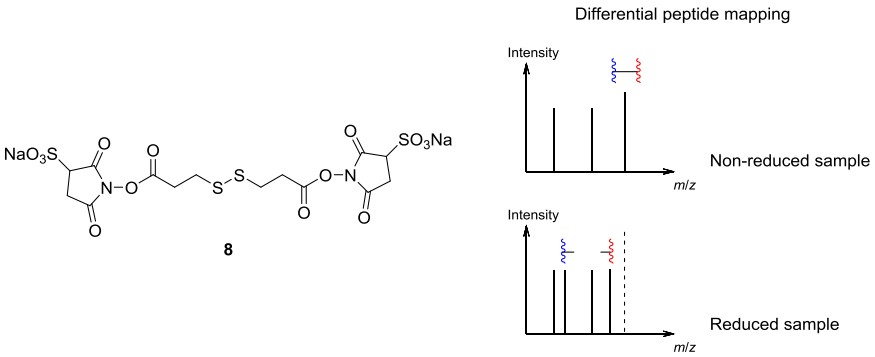


Figure 6. Structure of cross-linkers reagents: disuccinimidyl suberate **2** (DSS), bis(sulfo succinimidyl) suberate **3** (BS^3), disuccinimidyl glutarate **4** (DSG), D_{12}-disuccinimidyl suberate **5**, sulfosuccinimidyl 4,4'-azipentanoate **6** (sulfo-SDA), and sulfosuccinimidyl benzophenone **7** (sulfo-SBP).

Challenges in this research field comprise the complex fragment spectra of cross-linked peptides, low abundance of cross-linked species in a peptide mixture and heterogeneity of the products. Chromatographic separation techniques were also employed in order to reduce the complexity of cross-linked samples. Cross-linked peptides are generally larger and with higher charge state compared to proteolytic peptide fragments. For this reason, size exclusion[56] and strong-cation exchange[57] chromatography were used prior shotgun proteomic analysis. Usually, the searching engines use a theoretical spectrum library to compare against the experimental MS/MS. Therefore, from a bioinformatics point of view, the cross-linkers searches require a searching space that increases exponentially, which requires a huge computational power and time. In order to overcome this issue, various software have been designed for identification of cross-linked peptides, even if most of them do not present a graphic user interface with requires from the user informatics skills.[58]

Linker 3,3'-dithiobis(sulfosuccinimidyl propionate) **8** (DTSSP), quite close structurally similar to BS^3, was developed by Bennett *et. al.*[59] The unique characteristic of this linker is the cleavable disulfide bond that allows a comparison of MS spectra under reducing and non-reducing conditions. The signal of cross-linked peptides thereby disappears after the reducing conditions, which enables identification of cross-linked peptides even by using low-resolution mass spectrometers (Scheme 1). A year later Back *et. al.* reported a new strategy for identification of cross-linked peptide using a linker, *N*-benzyliminodiacetoyloxysuccinimid **9** (BID), that releases a characteristic benzyl cation under low-energy CID fragmentation.[60] But only in the middle 2000's the idea to use a cross-linker with a controllable and specific gas cleavage site was reported. Tang *et. al.* observed in 2005 that the RINK amide group, used in solid-phase peptide synthesis, contains a more labile bond under low CID

collision energy than peptide bonds.[61] In 2006, Soderblom and Goshe developed a cross-linker **10** containing a specific peptide bond that is able to fragment under lower CID collision energy independent of the peptide size.[62] The inspiration came from the observation of the cleavage preference of the amide bond between aspartic acid and proline compared to the other bonds. In this way, the two peptide fragments with a small additional mass could be analyzed by standard tandem mass spectrometry experiments at higher CID collision energies.[63]



Scheme 1. One of the first examples of chemically cleavable linker 3,3'-dithiobis(sulfosuccinimidyl propionate) **8** (DTSSP) and comparative analysis of non-reduced and reduced samples for cross-linked peptide identification.

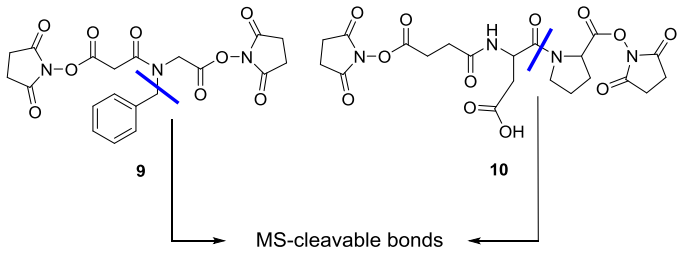


Figure 7. Chemical structure of the MS-cleavable cross-linkers *N*-benzyliminodiacetoyloxysuccinimid **9** (BID) and bisuccinimidyl-succinamyl-aspartyl-proline **10** (SuDP) with their cleavage sites (blue bold lines).

Since then different chemical bonds were explored and incorporated to cross-linker structures. MS-cleavable cross-linkers with label C-S bonds besides sulfoxide[64] and sulfonium ion[65], C-N bonds in urea[66] and quaternary amine[67] are reported. Although CID collision energy is the most popular fragmentation type used with cross-linkers, hydrazone N-N[68] and C-N[69] label bonds under ETD dissociation are also described.

The interest in combining different techniques and MS-cleavable cross-linkers applications rose in recent years. Bruce and collaborators have designed trifunctional cross-linkers for studies of RNase S complex formation. Later on, they improved the cross-linker adding a photo cleavable tag for the release of cross-linked peptides from the biotin-avidin complex (Figure 8).[70] The combination

of bifunctional linkers with a further extension arm allowed enrichment and purification of cross-linked peptides.

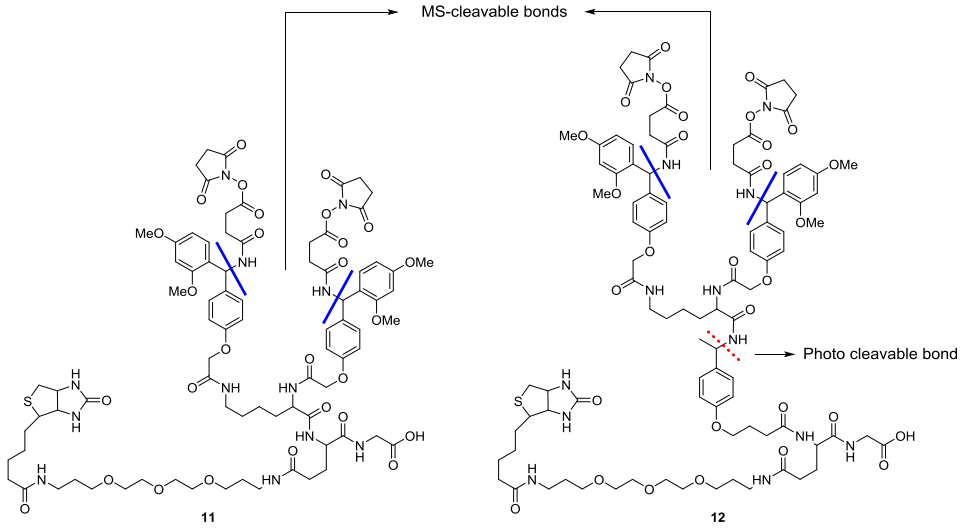
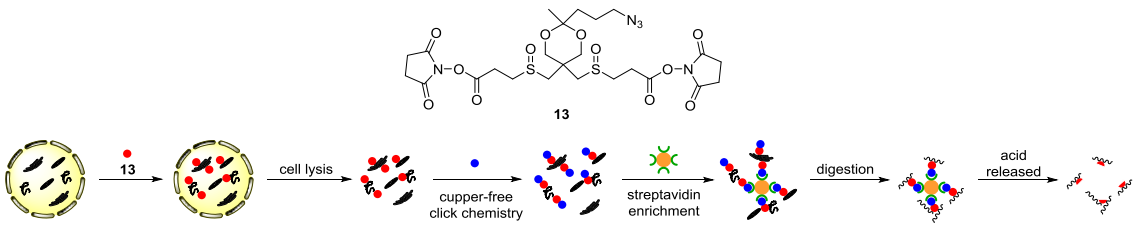


Figure 8. Structure of Bruce and coworkers cross-linkers **11** and **12**. MS-cleavable bonds are indicated with blue bold lines and photo cleavable bond is red dash line.[70]

Kaake *et. al.* developed a strategy to analyze *in vivo* cross-linked peptides from mammalian cells. Azide-tagged, acid-cleavable disuccinimidyl bis-sulfoxide (azide-A-DSBSO) **13** presents *N*-succinimidyl ester groups for targeting the lysine residues and azide to attach biotin for cross-linked species enrichment. Besides that, the structure presents an acetal group, acid label, enabling the elution of enriched cross-linked peptides, and sulfoxide groups that are cleaved under CID collision energy. Within this approach, 240 unique interlinked peptides from HEK 293 cells, and 85 protein-protein interactions among 54 proteins were reported.[71]



Scheme 2. Chemical structure of azide-tagged, acid-cleavable disuccinimidyl bis-sulfoxide (azide-A-DSBSO) **13** and scheme of labeling *in vivo* or *in vitro* using the linker developed by Kaake *et. al.*[71]

Although cleavable cross-linkers were developed to simplify and solve problems as previously mentioned, data analysis still remains as one of the main challenges in this research area. MeroX[72] and XLinkX[73] are the currently available software to handle data containing MS-cleavable

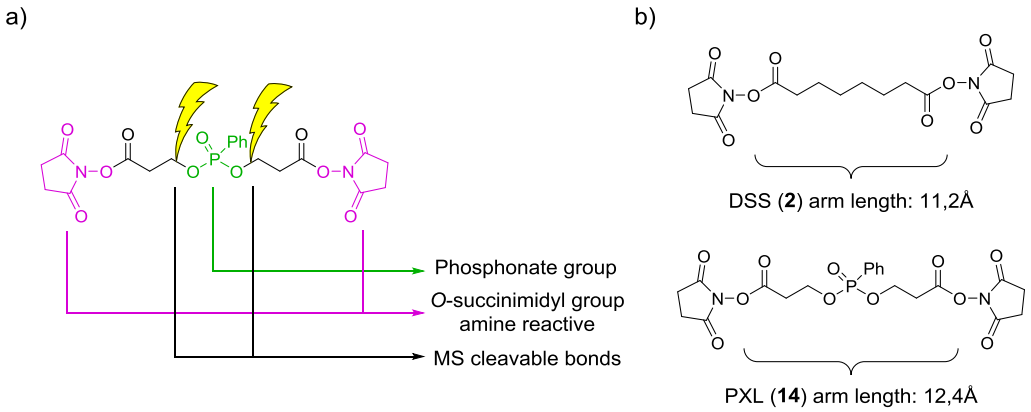
cross-linkers. MeroX uses the doublets generated from the linker cleavage (in this case the linker produces two peptide fragments with different masses) and compares with theoretical fragments to identify the candidates. XLinkX works in a quite similar manner compared to Merox, however, it was developed to handle proteome-wide data, while MeroX was created for more simple samples (ca. 5 000 proteins).

In spite of the high level of complexity in the cross-linker field, this strategy is already used in a number of studies for the elucidation of protein complexes[74-75], cell lysate[76] and *in vivo*[77] proteome analysis.

4.2- Aim

One way to determine the presence of phosphorylation as peptide modification via mass spectrometry analysis is using neutral loss scanning. In this methodology, the fragment ions are monitored when they present a specific neutral loss. In other words, when there is a change in mass but not in charge of the determined ion. For phosphopeptides, the neutral loss has 98Da, which indicates the loss of H_3PO_4 due to β-elimination.[78] The β-elimination is correlated with phosphopeptides under other circumstances. Chemical β-elimination of phosphoserine under basic conditions during solid-phase peptide synthesis is reported for over 20 years.[79] Moreover, chemical β-elimination is a tool used for tagging and labeling of phosphopeptides prior MS analysis.[80-81]

These reports inspired the design of a MS-cleavable cross-linker containing a phosphonate group (PXL) **14** that was expected to cleave as the phosphate groups under tandem mass spectrometer measurements. In the optimum case, the symmetrical linker cleavage leaves a small acryloyl modification in the peptide. The *O*-succinimidyl ester groups allow the reaction with lysine side chains and protein *N*-termini in accordance with well-established cross-linkers previously reported. Futhermore, the arm length of PXL, 12,4Å, is comparable in size to well-established cross-linkers, as disuccinimidyl surberate **2** (DSS), 11,2Å, calculated by Chem3D. The cross-linker arm length is responsible for the determination of the maximum distance between the cross-linked amino acid residues consequentely for the method resolution. For this reason, the distance among the residues cannot be too small that requires direct contact between the residues, neither too long that implies in the loss of resolution in a way that the conclusions regarding protein structure are dubious.[51, 82] The distance restrain promoted by DSS is already well characterized[83] and allows a higher number of confidentialy identified cross-linked peptides[51].

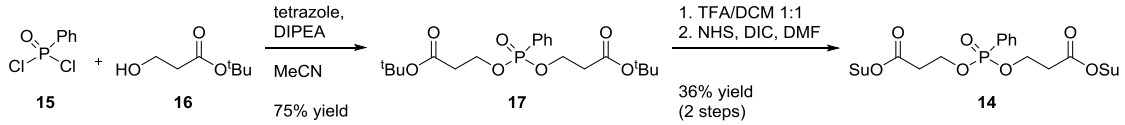


Scheme 3. a) Design of the phosphonate MS-cleavable cross-linker **14** (PXL): *O*-succinimidyl ester groups are specific for reaction with primary amines; phosphonate group follows two symmetric MS-cleavable bonds; b) comparison of arm length size of PXL and a reference cross-linker disuccinimidyl suberate (DSS), calculated by Chem3D.

During the identification of cross-linked protein-protein binding partners, the large size of cross-linked proteolytic peptide fragments is a major challenge for the MS-data analysis. The use of cleavable cross-linkers overcomes non-cleavable types by reducing the complexity and size of the linked peptides, which can be divided into two smaller peptides before or during MS analysis. In particular MS-cleavable cross-linkers outperform chemically or photo-cleavable linkers, which allow better correlation between the connected species and separated linked peptides. Moreover, a unique and specific linker cleavage can simplify the data analysis by providing an accurate identification of cross-linked peptides. This project aimed to synthesize and evaluate the MS-cleavable cross-linker PXL **14** under different collision energies and in different complexity levels initially using synthetic peptides followed by studies of a model protein.

4.3- Synthesis of the MS-cleavable cross-linker

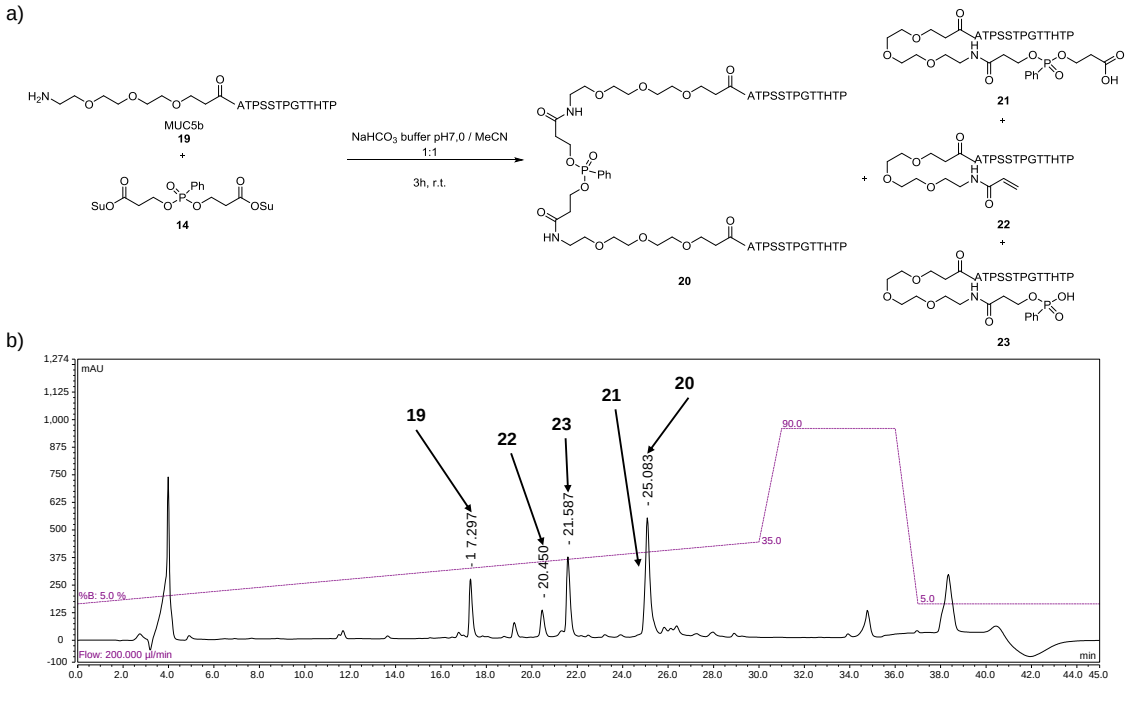
The designed MS-cleavable phosphonate cross-linker (PXL) was synthesized in 3 steps (Scheme 4). The commercial phenylphosphonic dichloride (**15**) and *tert*-butyl 3-hydroxypropionate (**16**) reacted under basic conditions to generate the intermediate **17** with a yield of 75%. Compound **17** had the *tert*-butyl groups removed with a solution of TFA/DCM 1:1. The free dicarboxylic acid was converted into the activated *O*-succinimidyl ester after reaction with *N*-hydroxysuccinimide mediated by diisopropylcarbodiimid. The final linker PXL **14** was obtained in 36% yield over two steps. The low yield, especially for the last step, might be due to the compound stability. In previous attempts to obtain the activated ester the elimination product, acryloyl *O*-succinimidyl ester, was isolated and characterized.



Scheme 4. Synthesis of MS-cleavable cross-linker.

4.4- Cross-linker evaluation using a peptide model

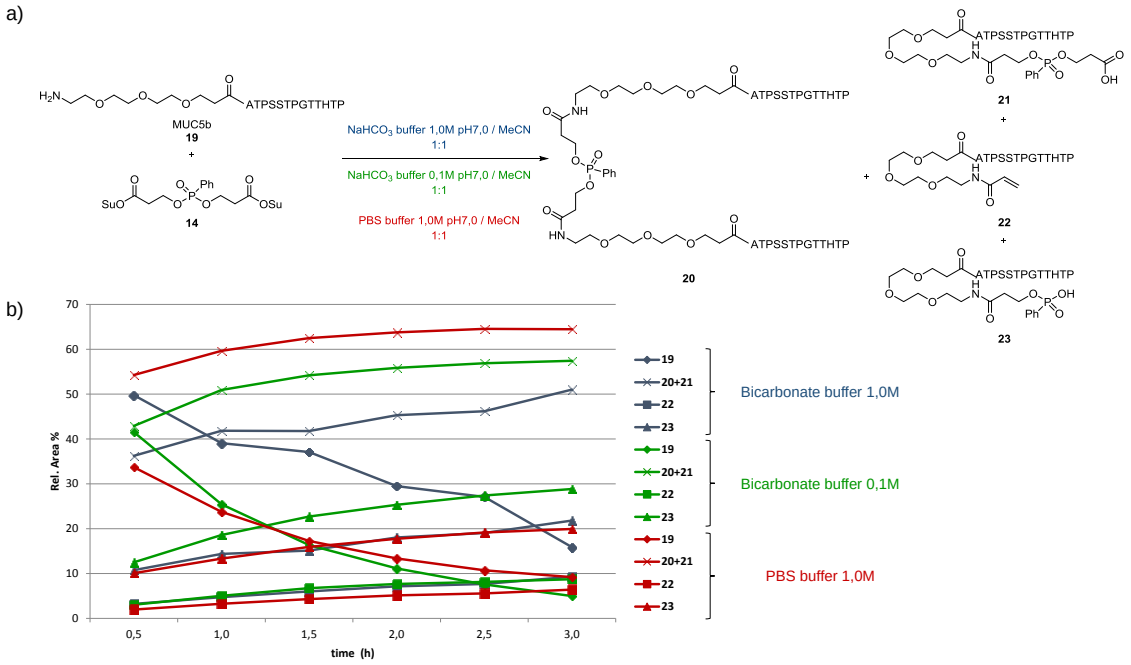
The first conjugation experiment was made by the reaction of a MUC5b model peptide (**19**) modified with a triethylene glycol amino acid spacer in the *N*-terminal. The cross-linker was conjugated with the peptide using 1M bicarbonate buffer pH7,0 and MeCN 1:1 in the proportion peptide/linker of 2:1. Analysis of the reaction mixture using LC-MS showed besides the cross-linked product **20**, the "dead-end" cross-linked peptide **21**, and both byproducts resulted from heterogenic cross-linker chemical cleavage: acryloyl-linked peptide **22** and phosphonate-linked peptide **23**.



Scheme 5. a) Reaction between MUC5b (**19**) and PXL **14** in bicarbonate buffer pH7,0 / MeCN and b) chromatogram of the reaction components after 3h.

In order to avoid the chemical elimination, the reaction was monitored in order to observe the formation of the intra-linked and elimination products upon different reaction conditions.

Bicarbonate buffer 1M, 0,1M, and PBS buffer 1M were evaluated. The reactions were monitored by HPLC at different time points, every 30 minutes over 3 hours. The normalized area for all the obtained product components were plotted against the reaction time (Scheme 6).

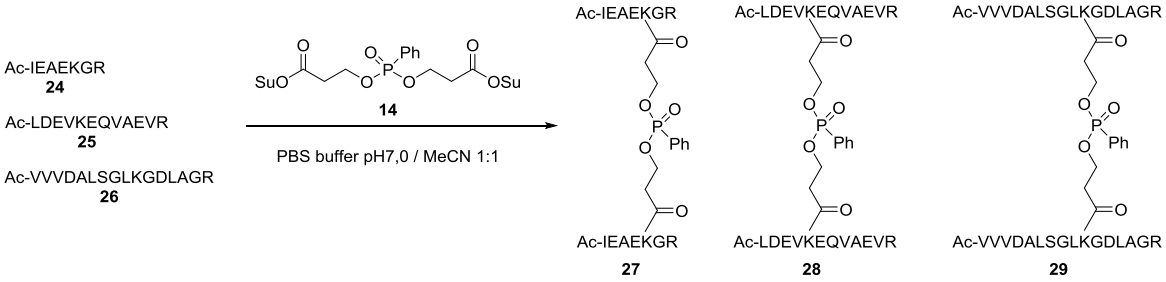


Scheme 6. a) Evaluation of PXL chemical cleavage during the coupling reaction with different buffers and b) components were monitored with LC for 3h reactions each half hour.

The diagram (Scheme 6b) shows the consumption of the initial peptide MUC5b **19** and the formation of the intra-linked peptide **20** and the "dead-end" linked peptide **21** (here both components are represented by one signal because there was no chromatographic separation), and the formation of the byproducts due to chemical cleavage, acryloyl- (**22**) and phosphonate-linked (**23**) peptides.

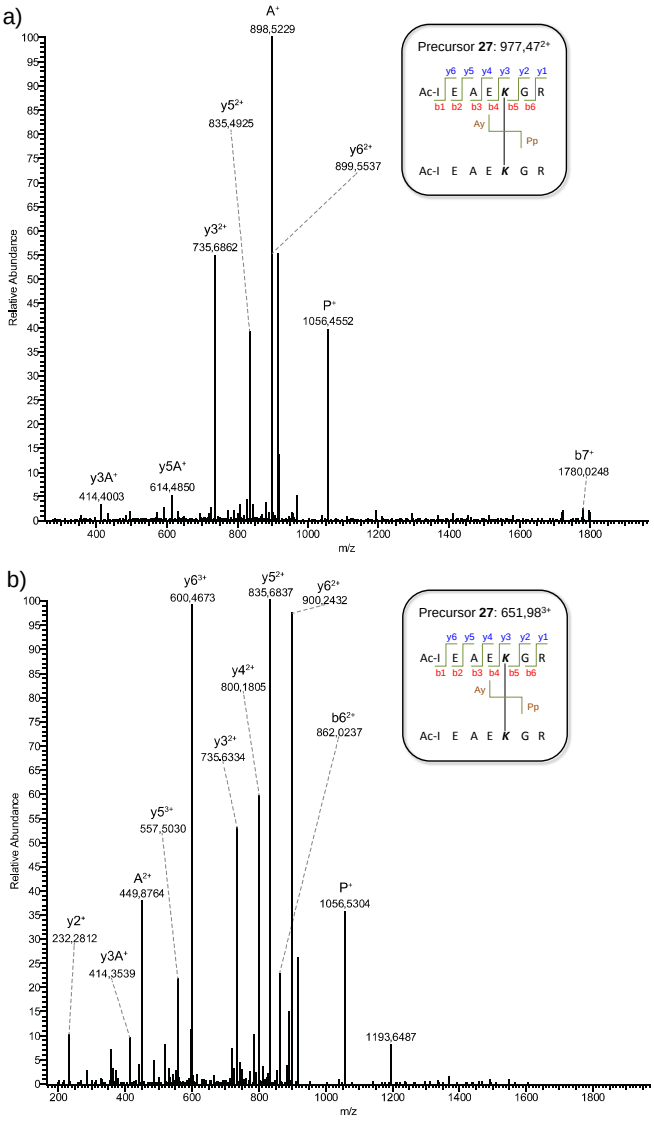
Herein, PBS buffer, represented by the red color, showed less decomposition by β-elimination than both concentrations of bicarbonate buffer. After 3h reaction time, the chromatographic peak area of the compounds **22** and **23** are smaller in PSB buffer than in other conditions. For this reason, PBS buffer was chosen for the following experiments.

In the interest of simulating the biological environment, tryptic peptides with a miscleavage and *N*-termini blocked by an acetyl group were synthesized using solid-phase peptide synthesis.[84] Sequences already used for cross-linker evaluation[64] or sequences identified in proteomic studies[85] with a nice fragmentation profile were used to simulate the reaction between PXL-peptide and fragmentation in a complex sample with short, medium or large peptides.

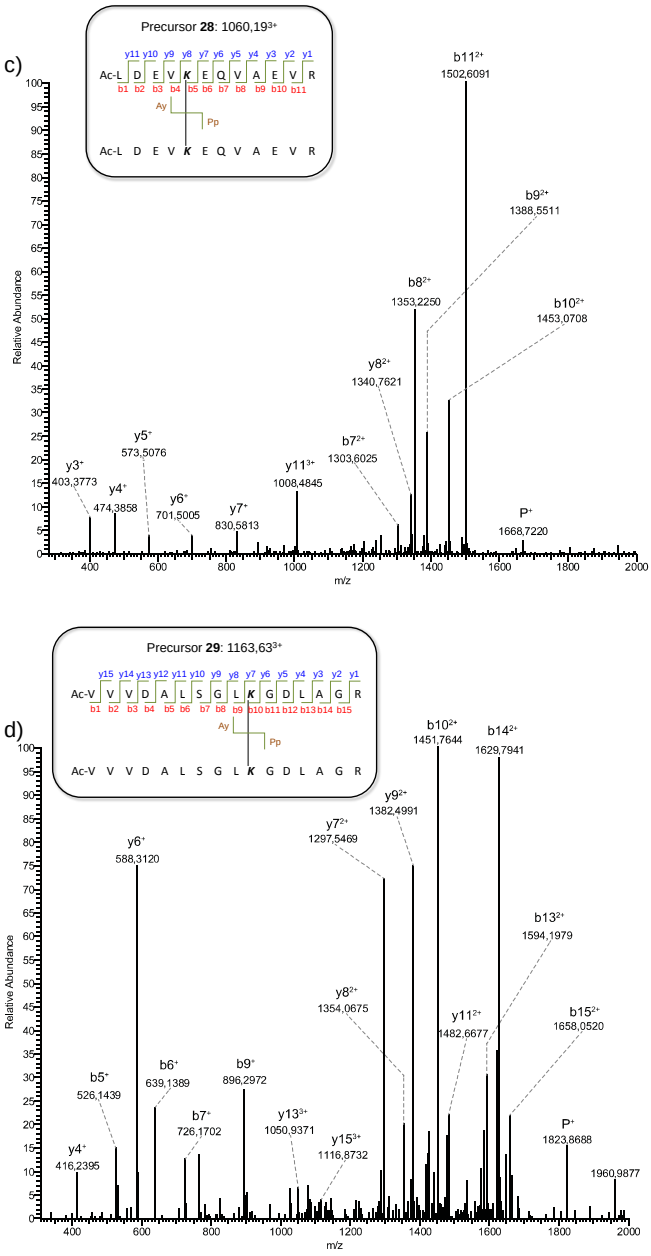


Scheme 7. Reaction between PXL (**14**) and acetylated peptides with tryptic miscleavage.

All obtained linked peptides were isolated using semi-preparative HPLC and submitted for MS/MS analysis. Collision induced dissociation with normalized collision energy of 35% was used for fragmentation.



Scheme 8. Fragmentation of cross-linked peptides with CID normalized collisional energy of 35%.



Scheme 8 (Continuation). Fragmentation of cross-linked peptides with CID normalized collisional energy of 35%.

Analysis of the fragmentation spectra of all linked peptides with the same charge state (3+) showed that there is a cleavage in the cross-linker bonds but the cleavage is not the dominant within the selected peptides (Scheme 8b-d). However, the analysis of the peptide Ac-IEAE***K***GR (**27**) spectra in the charge state 2+ hints that the energy for preferential fragmentation of the cross-linker could possibly be modulated. The spectrum in the Scheme 8a showed the acryloyl fragment (*m*/*z* 898,52^{1+}) as the most intense ion and the phosphonate fragment (*m*/*z* 1056,46^{1+}) as one of the 5 most intense fragment ions.

In this way, the cross-linked peptide Ac-IEAE***K***GR was submitted to HCD fragmentation analysis under different normalized collision energies (NCE). The 25 most intense peaks from the

precursor ion *m/z* 651,98^{3+} were obtained. The relative intensity of fragment ions of interest was plotted against the normalized collision energy. The diagram in Figure 9 showed for this peptide that the cross-linker cleavage apparently is preferable under lower collision energy. The acryloyl fragment Ay^{+}, corresponding to the cross-linker bond cleavage, is the second most intense fragment when the Ac-IEAE***K***GR peptide is submitted to HCD dissociation with 18% NCE.

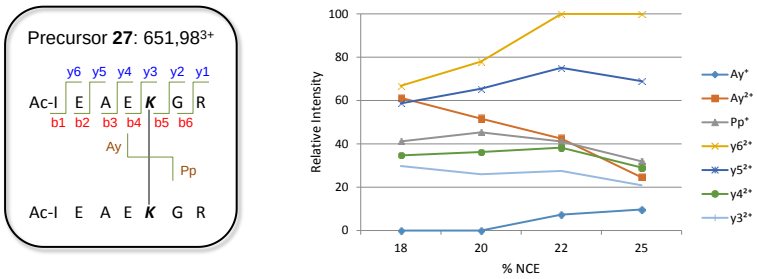
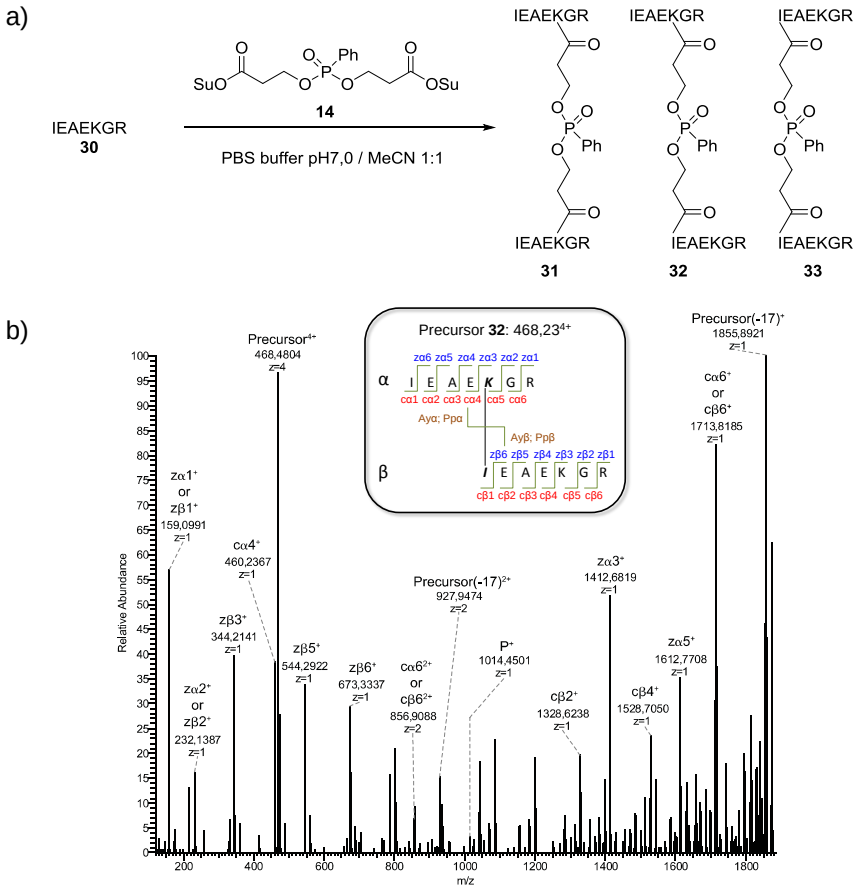


Figure 9. Distribution of fragment ions from Ac-IEAE***K***GR peptide **27**, precursor *m/z* 651,98^{3+}, according to different HCD normalized collision energies (NCE).

The cross-linker cleavage was also evaluated using ETD fragmentation. Therefore the peptide without acetylation was used in this case to increase the charge of the precursor and consequently increase the fragmentation efficiency, once ETD is optimum for fragment peptides with higher charge state (≥3+). During the preparation of the new cross-linked peptide, three possible linked isomers could be obtained (Scheme 9a). The linkage type of each product were distinguished depending on the fragment ions observed in the spectra (Table A 4).

Analysis of the cross-linked peptide in a Fusion mass spectrometer allowed the detection of ETD fragment ions in an Orbitrap with high resolution. This is important for the distinction between the species cα6 or cβ6 (*m/z* 856,92^{2+}) and Ayα or Ayβ (*m/z* 856,45^{1+}) that present close *m/z* (see theoretical mass Table A 4 on appendix). Herein, we assumed that the reaction product is cross-linked on both the lysine and *N*-terminal amine groups. The spectrum in Scheme 9b hints the linkage in the lysine (IEAE***K***GR) due to the presence of the fragment ions cα4 (*m/z* 460,24^{2+}) together with cα6 or cβ6 (*m/z* 1713,82^{1+} and *m/z* 856,92^{2+}) and zα3 (*m/z* 1412,68^{2+}). In the same way there is also linkage in the *N*-terminus (***I***EAEKGR) due to the signals zβ3 (*m/z* 344,21^{1+}), zβ5 (*m/z* 544,29^{1+}), zβ6 (*m/z* 673,33^{1+}), cβ2 (*m/z* 1328,62^{2+}) and cβ4 (*m/z* 1528,71^{2+}). However, the fragmentation of the linked peptide, which possesses *m/z* 468,23^{4+}, showed that the preferential cleavage occured in the peptide backbone and the most intense fragment ions were the precursor with ammonia loss (*m/z* 1855,89^{1+}) and the c6 fragment (*m/z* 1713,82^{1+}).

After fragmentation evaluation, the HCD with low collision energy presented the best option for the cleavage of the linker bonds.

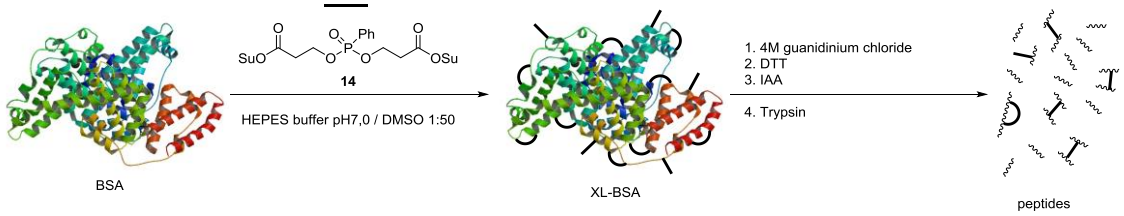


Scheme 9. All possible products from the reaction between peptide IEAEKGR **30** and PXL **14** (a) and the isolated product analyzed using ETD fragmentation (b).

4.5- Cross-linker evaluation using a model protein

Serum albumin is the most abundant protein in mammalian plasma. The mature form of bovine serum albumin (BSA) possesses 583 amino acid residues and a molecular weight of approximately 66kDa. Due to its stability, this protein is often used as a concentration standard.[86] Accessibility and high lysine residues content, 59 in total, make BSA a good model protein for cross-linker evaluation. Bovine serum albumin was already been further used as model protein in different studies. The first evaluation was reported by Huang *et. al.* using non-cleavable linkers. Bis (sulfosuccinimidyl) suberate **3** and disuccinimidyl suberate **2**, both with 11,4Å, were respectively used to access lysines in hydrophilic and hydrophobic regions of the protein, and disuccinimidyl glutarate **4**, 7,7Å, to access short distance lysines in order to obtain 3D information.[87] Bovine serum albumin was also used as model protein for evaluation of cleavable cross-linkers[62, 88-89] and enrichment method for cross-linked peptides[90].

Herein, BSA was cross-linked with PXL in a protein/linker ratio 1:50, for 30min and 1h (Scheme 10). The reaction between cross-linker and protein was firstly verified with SDS polyacrylamide gel electrophoresis – SDS-PAGE. This method is used to separate denatured proteins according to the molecular weight. In standard SDS-PAGE protocols, the protein is denatured in hot SDS solution, the disulfides bonds are reduced, and the unfolded protein remained as a polypeptide chain coated by the SDS anion. The solution with denatured protein is loaded onto the polyacrylamide gel and electrical field is applied. In this way, the negatively coated unfolded proteins migrate to the anode (positively charged electrode). The separation occurs because smaller proteins migrate through the gel porous faster than bigger proteins. The protein visualization is possible after staining the gel. Coomassie Brilliant Blue is one of the most common dye used for protein staining. Moreover, this methodology is optimal for controlling the cross-linker reaction on the protein level. Due to the linker, the protein cannot unfold properly in the denaturation step. As a consequence, differently cross-linked proteins present different forms and no longer a uniform unfolded protein chain. Therefore the cross-linked protein band in the gel is broader and more undefined when compared to the control. Figure 10 clearly shows the difference between the cross-linked and control reactions. In the control reaction (Figure 10 BSA-K), the BSA band is more intense and well-defined compared to both cross-linked protein bands (Figure 10 BSA-30min and BSA-1h), demonstrating that the cross-linker reaction on a protein level was successful. In this procedure, BSA was not reduced after denaturation, therefore the disulfide bonds impeded proper run of the protein and BSA ran until 55kDa instead of expected 66kDa according to the reference.



Scheme 10. Experimental test of PXL **14** employing BSA as a model protein.

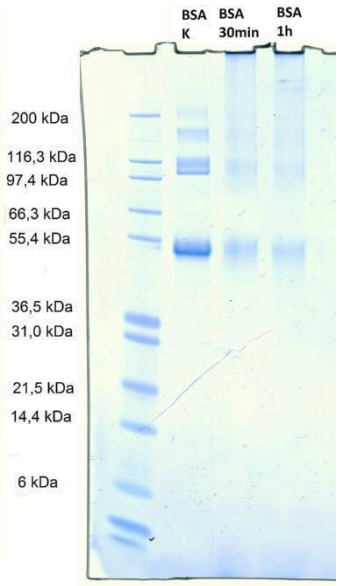


Figure 10. SDS-PAGE of BSA reacted with cross-linker PXL **14** for 30min and 1h compared to the reaction control (K). The reaction control (K) consists in the same cross-linker reaction conditions but without the linker PXL.

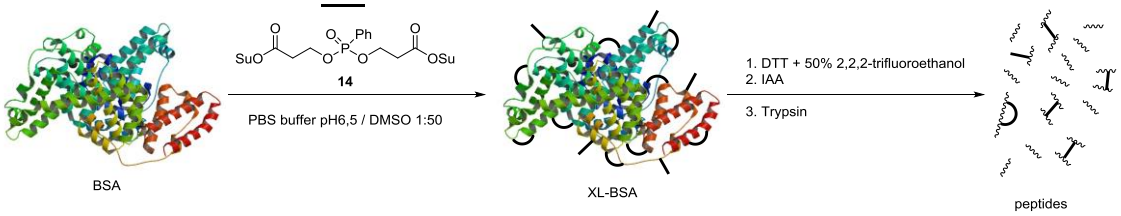
After checking the reaction, the cross-linked protein was proteolytically cleaved for further evaluation. The tryptic peptide mixture was submitted for analysis by nano-LC-MS/MS using an Orbitrap Velos Pro mass spectrometer. MS/MS experiments used CID with normalized collision energy of 20%. Data analysis was performed manually. According to calculations using Chem3D, disuccinimidyl suberate **2** has 11,2Å of arm length and PXL **14** has 12,4Å. Since their size is similar, the masses of identified cross-linked peptides using disuccinimidyl suberate **2** and bis (sulfosuccinimidyl) suberate **3**[87] were corrected for the PXL mass and a theoretical database was created (see appendix Table A 6). Peptides with the charge state 3+ and 4+ were searched manually with a deviation of ±0,5*m*/*z* in Xcalibur Qual Browser 2.2. A database with peptides originated from PXL chemical cleavage was also created (see appendix Table A 7). These peptides with charge state of 2+ were also searched in the same manner as the cross-linked peptides.

The results summarized in Table 1 showed that a maximum of 4 cross-linked peptides were identified in both experiments. Furthermore, a total of 19 different peptides were also identified by linker chemical cleavage after 30min or 1h cross-linking reaction time. In order to avoid chemical cleavage, the cross-linker procedure was optimized, changing the reactions conditions. In this way, the cross-linker reaction was carried out under weak acidic conditions, PBS buffer pH6,5, and the protein denaturation protocol was changed for acidic conditions: substitution of guanidinium chloride by 2,2,2-trifluoroethanol[91] (Scheme 11). This modification also aimed to avoid basic nitrogen from guanidinium chloride once the nitrogen electron pair can promote the β-elimination in the PXL linker.

Table 1. Peptides manually identified after reaction with PXL: cross-linked peptides (XL) and peptides resulted from PXL chemical cleavage acryloyl (Ay) and phosphonate (Pp).

Peptide code	Charge state	Peptide sequence	Theoretical mass	30min reaction*		1h reaction*	
				m/z	r.t. (min)	*m/z*	r.t. (min)
7XL	3+	$_{130}$K.DDSPDLPKL***K***PDPNTLCDEFK.A$_{152}$	808,7	808,7	15,79	808,7	15,72
	4+	$_{451}$R.SLG***K***VGTR.C$_{460}$	606,8	606,8	15,79	606,8	15,72
9XL	3+	$_{232}$R.AL***K***AWSVAR.L$_{242}$	761,0	761,0	12,95	n.i.	n.i.
	4+	$_{263}$K.VH***K***ECCHGDLLECADDR.A$_{281}$	571,0	n.i.	n.i.	n.i.	n.i.
10XL	3+	$_{436}$R.***K***VPQVSTPTLVEVSR.S$_{452}$	751,0	751,0	15,55	751,0	15,48
	4+	$_{557}$K.HKP***K***ATEEQL.K$_{569}$	563,5	563,5	15,55	563,5	15,48
17XL	3+	$_{489}$K.TPVSE***K***VT***K***CCTESLVNR.R$_{508}$	792,7	792,7	21,85	n.i.	n.i.
	4+		594,8	n.i.	n.i.	n.i.	n.i.
Ay-2	2+	$_{451}$R.SLG***K***VGTR.C$_{460}$	436,3	436,3	14,79	436,3	14,83
Ay-3	2+	$_{232}$R.AL***K***AWSVAR.L$_{242}$	528,3	528,3	22,71	528,3	22,8
Ay-5	2+	$_{436}$R.***K***VPQVSTPTLVEVSR.S$_{452}$	847,5	847,5	26,19	847,5	26,18
Ay-7	2+	$_{371}$R.LA***K***EYEATLEECCAK.D$_{387}$	934,9	934,9	22,34	934,9	22,39
Ay-8	2+	$_{222}$R.CASIQ***K***FGER.A$_{233}$	625,3	625,3	19,61	625,3	19,61
Ay-9	2+	$_{482}$R.LCVLHE***K***TPVSEK.V$_{496}$	797,4	797,4	18,24	797,4	18,36
Ay-10	2+	$_{197}$K.GACLLP***K***IETMR.E$_{210}$	721,9	n.i.	n.i.	721,9	28,83
Ay-11	2+	$_{489}$K.TPVSE***K***VTK.C$_{499}$	521,8	521,8	14,64	521,8	14,68
Ay-12	2+	$_{209}$R.E***K***VLASSAR.Q$_{219}$	536,8	536,8	9,86	536,8	10,04
Ay-13	2+	$_{459}$R.CCT***K***PESER.M$_{469}$	610,8	610,8	10,76	n.i.	n.i.
Ay-14	2+	$_{256}$K.LVTDLT***K***VH.K$_{267}$	604,4	604,4	18,57	604,4	18,72
Ay-15	2+	$_{241}$R.LSQ***K***FPK.A$_{249}$	458,8	458,7	14,11	458,7	14,18
Ay-16	2+	$_{547}$K.***K***QTALVELLK.H$_{558}$	598,9	598,9	28,25	598,9	28,2
Pp-6	2+	$_{557}$K.HKP***K***ATEEQL.K$_{569}$	788,4	788,4	21,79	788,4	21,86
Pp-8	2+	$_{222}$R.CASIQ***K***FGER.A$_{233}$	731,8	n.i.	n.i.	731,8	32,86
Pp-10	2+	$_{197}$K.GACLLP***K***IETMR.E$_{210}$	828,4	828,5	26,92	828,5	26,93
Pp-12	2+	$_{209}$R.E***K***VLASSAR.Q$_{219}$	643,3	n.i.	n.i.	643,3	26,72
Pp-13	2+	$_{459}$R.CCT***K***PESER.M$_{469}$	717,3	717,3	23,68	717,3	23,78
Pp-14	2+	$_{256}$K.LVTDLT***K***VH.K$_{267}$	710,9	710,8	14,99	710,8	14,98

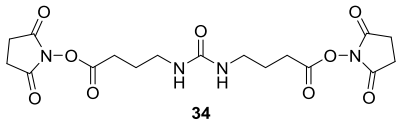
*n.i. – not identified



Scheme 11. Optimized conditions for BSA cross-linking with PXL (**14**).

The data presented in Table 2 demonstrated an improvement of identified cross-linked peptides compared to the previous condition. Herein, 7 intra-linked peptides were identified; including 2 cross-linked peptides not previously reported by Huang *et. al.*[87]. Like in the case of other identified cross-linked peptides, 20XL and 21XL were also found with co-eluted parent ions at different charge states, 3+ and 4+, and were considered for this reason as potential cross-linked peptides. However manual identification of these peptides was not possible. Also, 3 loop linked peptides were identified, which they were not observed in the first condition. And finally, only 8 peptides from linker chemical cleavage were identified. The increase in the number of XL-peptides

and the decrease in the number of peptides identified from chemical cleavage demonstrate the improvement in the experimental conditions. In order to improve the amount of peptide identifications, an automated search with the program MeroX 1.4.9 was performed, however, the identification score was not enough to have a high confidence identification even after changing searches parameters. Deeper adjustment in the program for searches using PXL as cross-linker might be required since this software was developed and optimized for the Sinz group MS-cleavable cross-linker BuUrBu (**34**).[72]



34

Table 2. Peptides manually identified after optimized reaction with PXL: cross-linked peptides (XL) and peptides resulted from PXL chemical cleavage acryloyl (Ay) and phosphonate (Pp).

Peptide code	Charge state	Peptide sequence*	Theoretical mass	30 min reaction*		1 h reaction*	
				m/z	r.t. (min)	*m/z*	r.t. (min)
8XL	3+	$_{232}$R.AL***K***AWSVAR.L$_{242}$	808,0	808,0	16,72	808,3	16,86
	4+	$_{256}$K.LVTDLT***K***VH.K$_{267}$	606,3	606,3	16,72	606,5	16,86
9XL	3+	$_{232}$R.AL***K***AWSVAR.L$_{242}$	761,0	760,6	22,90	n.i.	n.i.
	4+	$_{263}$K.VH***K***ECCHGDLLECADDR.A$_{281}$	571,0	n.i.	n.i.	n.i.	n.i.
10XL	3+	$_{436}$R.***K***VPQVSTPTLVEVSR.S$_{452}$	751,0	750,6	19,16	750,7	19,28
	4+	$_{557}$K.HKP***K***ATEEQL.K$_{569}$	563,5	n.i.	n.i.	563,2	19,28
11XL	3+	$_{209}$R.E***K***VLASSAR.Q$_{219}$	702,3	702,7	22,52	702,3	22,67
	4+	$_{241}$R.LSQ***K***FPK.A$_{249}$	527,0	n.i.	n.i.	n.i.	n.i.
12XL	3+	$_{209}$R.E***K***VLASSAR.Q$_{219}$	640,3	640,3	23,20	640,3	23,49
	4+	$_{241}$R.LSQ***K***FPK.A$_{249}$	480,5	n.i.	n.i.	n.i.	n.i.
15XL	3+	$_{256}$K.LVTDLT***K***VH***K***ECCHGDLLECADDR.A$_{281}$	1051,3	1051,1	27,86	n.i.	n.i.
	4+		788,8	788,8	27,86	n.i.	n.i.
17XL	3+	$_{489}$K.TPVSE***K***VT***K***CCTESLVNR.R$_{508}$	792,7	792,7	27,86	n.i.	n.i.
	4+		594,8	n.i.	n.i.	n.i.	n.i.
19XL	3+	$_{155}$K.***K***FWG***K***YLYEIAR.R$_{168}$	614,7	615,0	22,25	615,0	22,37
	4+		461,3	n.i.	n.i.	n.i.	n.i.
20XL	3+	n.d.	-	978,5	47,29	978,5	47,31
	4+		-	734,1	47,29	734,1	47,31
21XL	3+	n.d.	-	634,6	17,63	634,6	17,77
	4+		-	476,2	17,63	476,2	17,77
Ay-5	2+	$_{436}$R.***K***VPQVSTPTLVEVSR.S$_{452}$	847,5	847,4	24,12	847,9	22,69
Ay-8	2+	$_{222}$R.CASIQ***K***FGER.A$_{233}$	625,3	625,3	11,78	625,3	15,31
Ay-15	2+	$_{241}$R.LSQ***K***FPK.A$_{249}$	458,8	458,7	14,95	458,7	15,17
Pp-8	2+	$_{222}$R.CASIQ***K***FGER.A$_{233}$	731,8	731,8	33,9	731,8	34,19
Pp-10	2+	$_{197}$K.GACLLP***K***IETMR.E$_{210}$	828,4	828,3	24,51	828,3	24,76
Pp-12	2+	$_{209}$R.E***K***VLASSAR.Q$_{219}$	643,3	n.i.	n.i.	643,3	28,08
Pp-13	2+	$_{459}$R.CCT***K***PESER.M$_{469}$	717,3	717,3	24,79	717,3	25,06
Pp-17	2+	$_{557}$K.H***K***PK.A$_{562}$	393,7	393,2	12,41	393,2	12,59

*n.d. – not determined; n.i. – not identified

4.6- Summary and outlook

Neutral elimination in phosphopeptides observed in tandem MS analysis inspired the development of a MS-cleavable protein cross-linker (PXL). The linker enabled bilateral MS-cleavage generating peptides with a small acryloyl modification. Moreover, PXL was designed in a way to have an optimum size in order to maximize the number of identifications. The arm length 12,4Å is comparable to the size of DSS (**2**), 11,2Å, previously reported. The synthesis of the phosphonate MS-cleavable cross-linker PXL **14** was performed in 3 steps from commercial compounds in a yield of 27%. The lower yield can be explained partially by the chemical instability of the phosphonate in the presence of mild basic conditions and a elimination byproduct could additionally be observed during linker synthesis.

Synthetic peptides were used for evaluation of the generated PXL linker. PBS buffer at neutral pH was found optimal compared to bicarbonate buffer to avoiding chemical cleavage during the cross-linker reaction. Different fragmentation techniques, CID, ETD, and HCD, were employed to evaluate the cross-linker cleavage in MS. HCD fragmentation with low normalized collision energies favored the linker bond fragmentation in comparison to the peptide backbone fragmentation. The acryloyl fragment of the Ac-IEAE***K***GR **27** is the second most intense ion in HCD with 18% NCE.

The linker PXL (**14**) was also evaluated with BSA as a protein model. SDS-PAGE and LC-MS/MS analysis were used to examine the cross-linked protein. SDS-PAGE showed that the cross-linker reaction of the protein worked, which was observed by smearing of the BSA bands in the SDS-gel. Optimized conditions for the cross-linking reaction and proteolytic digestion showed the identification of 7 intra-linked peptides, including 2 non-previously reported, 3 loop linked peptides and 8 peptides derived from chemical cleavage. For the optimization of data analysis, spectra from LC-MS/MS experiments still needs to be analyzed with a better optimized software for the PXL linker. Nevertheless, these results indicate that optimized conditions in the experimental procedure on both peptide and protein levels minimized the chemical cleavage of the linker, and PXL (**14**) has the potential to be applied as a useful cross-linker in more complex systems such as protein complexes.

New Approach for Glycopeptide Enrichment

5.1- Background: Glycosylation as protein modification

Information about protein structure alone is not enough to understand the involvement of these biomolecules in diverse biological functions. Evolution also brought chemical modifications after protein translation to refine and modulate protein activities.[92] Nowadays, co-translational and post-translational modifications (PTM's) are known to interfere in cell regulation, protein function and interactions. Modified proteins work in interactive processes transmiting extracellular and intracellular signals important for cell function in a way that many diseases as cancer can be interpreted as pathological alterations in the network signalling.[93] Although PTM's are crucial for cell survival, their function, *modus-operandi* and regulation are still not fully comprehended (some modifications are described in Table 3 and illustrated in Figure 11). In most cases these modifications are not predictable based on the genome and they rely upon a fine and complex cellular mechanism of regulation system. With some exceptions, there are proteins that present a consensus sequence which indicates the presence of a PTM, e. g. glycosylation in asparagine residues requires a consensus sequence Asn-Xxx-Thr/Ser, where Xxx is any amino acid except proline. The development in the mass spectrometry field increased the possibility of identification and study of PTM's. Due to the large variety and status of those modifications it is still a challenge to understand deeply their role. Some PTM's are dynamic like phosphorylation; in contrast others like carbonylation that are irreversible.[94]

Table 3. Some common and important PTM's and their function.[95]

PTM type	Function
Acetylation	Protein stability, protection of *N*-terminus, regulation protein-DNA interaction
Acylation (fatty acid)	Cellular localization and signaling, mediator protein-protein interaction
Disulfide bond	Intra- and intermolecular cross linkage and protein stability
Glycosylation	Excretion, cell recognition, signaling, regulation
Hydroxylation	Protein stability, protein-ligand interactions
Methylation	Gene expression regulation
Nitration	Oxidative damage during inflammation
Phosphorylation	Signaling, modulation of enzymatic activity
Ubiquitination	Protein degradation

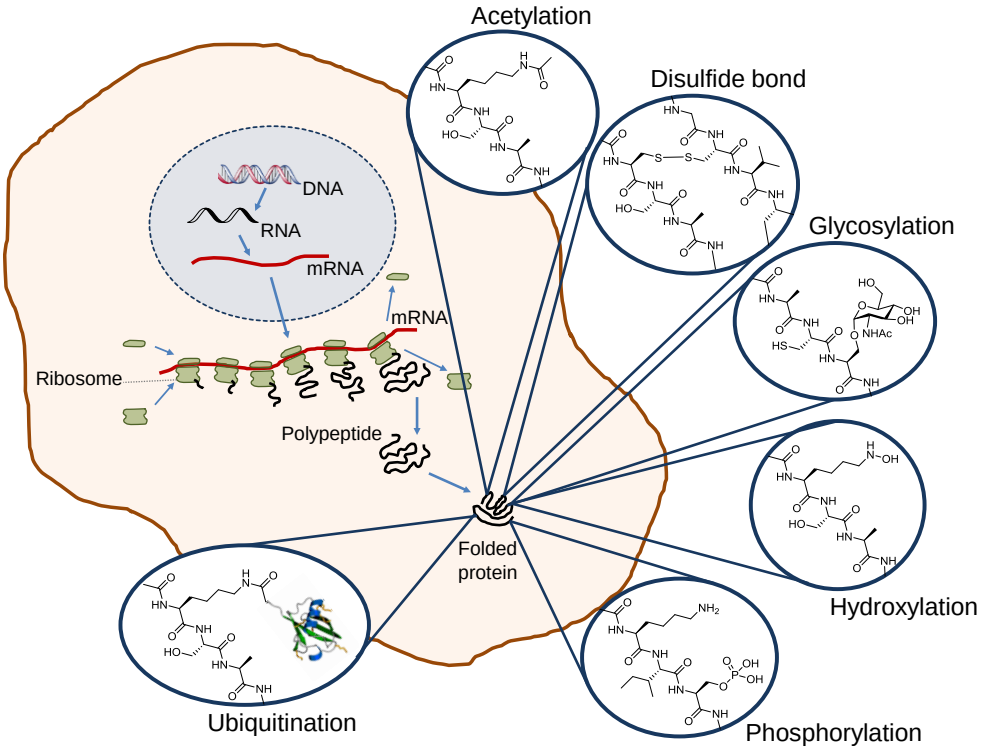


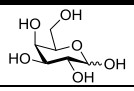
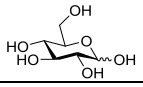
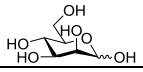
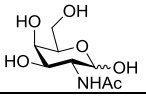
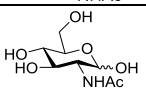
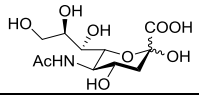
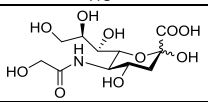
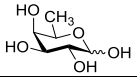
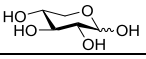
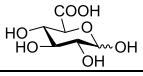
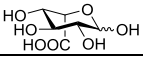
Figure 11. Protein synthesis: DNA ribbon is open and copied as RNA ribbon that is further modified to messenger RNA (mRNA). Once mRNA crosses the nucleus membrane, ribosomes attached to mRNA for the translation of the information into a polypeptide sequence. After translation, the polypeptide sequence can be chemically modified in order to help the protein folding and assume the active form. Chemical structures of some post-translational modifications are highlighted.

In 2013 Beltrao and collaborators reported 460 different amino acid modifications according to Uniprot database, in which different amino acids were counted individually, e. g. phosphoserine and phosphothreonine were different modifications.[92] Besides the complex cellular pathway that regulates where and when modifications take place, there are also variations within the group of PTMs regarding their chemical nature and function. *N*-glycosylation presents a rather stable C-N bond between the glycan and amino acid and takes place in many different cell functions as structural role as one example. While *O*-glycosylation (*O*-GlcNAcylation, mucin type, *O*-mannosylation, *O*-fucosylation and *O*-glucose modifications) consist of a more labile O-glycosidic bond in the form of an acetal between the glycan and amino acid residue. Among the O-glycans, O-GlcNAcylation has been found to be a dynamic modification that sometimes interplay with phosphorylation. Acylation, addition of palmitoyl or myristoyl groups to the amino acid, directs the protein to be allocated in the membrane[96] or may lead to apoptosis[97] respectively. This chapter focus on protein glycosylations, which will be further discussed in the following sections.

5.1.1 - Chemistry in glycosylation

The carbohydrates belongs to a very important class of biomolecules. Beyond being essential as source of energy, carbohydrates presented on lipids and proteins are main components of extracellular matrix and they are involved in many biological processes. Chemically, carbohydrates are defined as polyhydroxyaldehydes, polyhydroxyketones and their derivatives. Monosaccharides are the smallest units from the carbohydrate that cannot be hydrolyzed. When conjugated with other macromolecules the term glycan is often used to refer to a carbohydrate.

Table 4. Chemical structure, symbol nomenclature and exact mass of the main monosaccharide presented in mammalians.

Class	Name	Abreviation	Chemical Structure	Symbol	Exact mass
Hexose	Galactose	Gal			180,0634
	Glucose	Glc			180,0634
	Mannose	Man			180,0634
N-acetylhexosamine	*N*-acetylgalactosamine	GalNAc			221,0899
	N-acetylglucosamine	GlcNAc			221,0899
Sialic acid	*N*-acetylneuraminic acid	Neu5Ac			309,1060
	N-glycolylneuraminic acid	Neu5Gc			325,1009
Deoxyhexose	Fucose	Fuc			164,0685
Pentose	Xylose	Xyl			150,0528
Acidic hexose	Glucuronic acid	GlcA			194,0427
	Iduronic acid	IdoA			194,0427

Glycosylations as post-translational modifications are estimated to occur in at least half of all proteins.[98] This group of modifications are the most complex among all protein modifications. The complexity of a glycan structure is related to the presence of multiple stereocenters in the carbohydrate moiety, diversity of linkage connections between the carbohydrates itself and with other biomolecules like proteins and extension of the carbohydrate chains. In mammalians a number higher than 7 000 different glycan structures is estimated and all these structures are built-up from ten monosaccharides: galactose (Gal), glucose (Glc), *N*-acetylgalactosamine (GalNAc), *N*-

acetylglucosamine (GlcNAc), mannose (Man), sialic acid (SA), fucose (Fuc), xylose (Xyl), glucuronic acid (GlcA) and iduronic acid (IdoA).[99] The sialic acids consists of a family of more than 50 monosaccharides with nine carbons also known as neuraminic acids. The two most common sialic acid structures in vertebrates are *N*-acetylneuraminic acid (Neu5Ac) and *N*-glycolylneuraminic acid (Neu5Gc). However humans only synthesize Neu5Ac, other animals also synthesies Neu5Gc besides Neu5Ac.[100] Structure and symbol nomenclature used to represent each monosaccharide are represented in Table 4 as well as their exact mass. Within this thesis, the symbol nomenclature used to represent the glycosylation is the nomenclature described by Varki and coworkers.[101]

In the process of cyclization of the polyhydroxyaldehydes or polyhydroxyketones, a new stereocenter is formed at the ring carbon C1, which is called the anomeric center (Figure 12a and b). Due to its chemical nature, this position is the most susceptible to receive nucleophile attacks in order to form a glycosidic bond (Figure 12c). The glycosidic linkage is formed when a nucleophile, that in most of cases is a hydroxyl or amine group, from another glycan, peptide or lipid, within biological systems, attacks the anomeric center leading to the formation of a oligosaccharide (glycans) or a bioconjugate (glycans attached to peptides or lipids). Observing the Fischer projection, in case the substitution at the anomeric center is *cis* in relation to the most distant stereogenic center of the ring, the monosaccharide is defined as α anomer (Figure 12a). If the configuration of the anomeric center and the most distant stereogenic center is *trans*, the monosaccharide is defined as β anomer (Figure 12b).

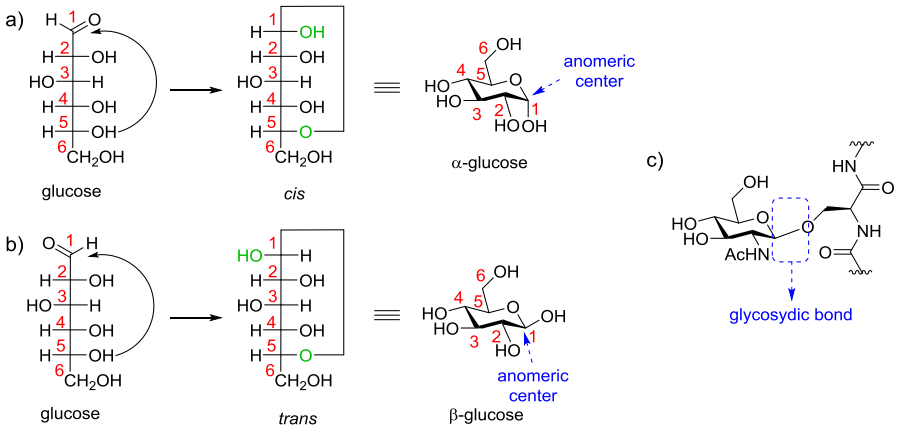


Figure 12. Highlights in the glycan chemistry: formation of α- (a) and β-glucose (b) anomers from Fischer projection, and example of glycosidic bond between *N*-acetylglucosamine and serine amino acid residue (c).

Glycan unique properties can completely change the characteristics of a protein, for instance influence protein folding or physicochemical properties of the protein. Due to the large complexity and variety, the glycosylation, as protein modification, is divided in some categories, according to the amino acid where the glycosylation is attached.

5.1.1.1- *N*-Glycosylation

A glycan linked through the nitrogen atom of the asparagine side chain of a protein is classified as *N*-glycosylation. This is one of the most common and stable glycan modifications and is found in both eukaryotes and prokaryotes and play crucial role in protein folding, stability, trafficking and orientation. Proteins with the consensus sequence Asn-Xxx-Thr/Ser, where Xxx is any amino acid except proline, are often modified with *N*-glycosylation. Rarely, proteins with sequence Asn-Xxx-Cys are also glycosylated[102-103] and even rarer, Asn-Xxx-Val and Asn-Gly motives were also reported glycosylated.[104]

The biosynthesis of *N*-glycans takes place in the endoplasmic reticulum (ER) – Golgi complex. The transfer of a phospho-GlcNAc to phospho-dolichol is the first step of the process, present in the cytoplasmic side of the ER membrane. In the following steps, glycosyltransferases transfer one extra GlcNAc and Man residues until the achievement of the oligosaccharide $GlcNAc_2Man_5$, when the glycan is transported to ER lumen for further elongation. Extra Man and Glc residues are used for further glycan elongation and only the oligossacharide $GlcNAc_2Man_9Glc_3$ is transferred to the peptide sequence with the consensus sequence, releasing dolichol di-phosphate. Once the glycosylation is located in the protein, there is a trimming process in which one glucose unit is cut off by a specific glycosydase in other to prepare the transport of the glycoprotein to the Golgi complex for further maturation. In the Golgi, glycosydases and glycosyltransferases modify the glycosylation in order to obtain the different types of *N*-glycosylation. This whole processes also contributes to the quality control of protein folding.

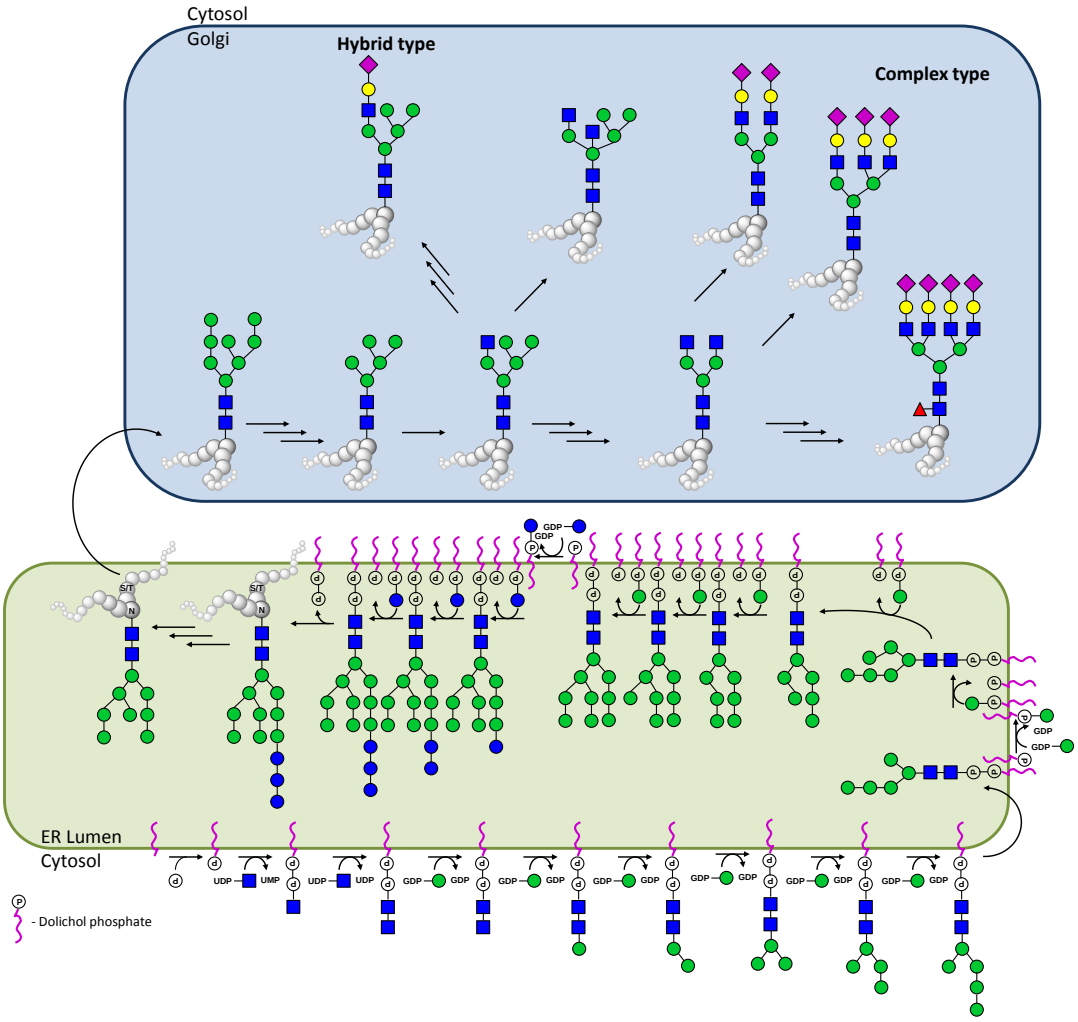


Figure 13. *N*-glycosylation biosynthetic pathway: dolichol attaches to ER membrane and is activated via phosphorylation in order to begin the synthesis of the *N*-glycan core (bottom left). Glycosyltransferases extend the glycan structure in the cytosol until the achievement of $GlcNAc_2Man_5$ and the glycan is transported to ER lumen for further elongation. The oligosaccharide $GlnNAc_2Man_9Glc_3$ is transferred to the peptide sequence that contains the motif N-X-S/T, where X is any amino acid except P. Glycosydases trim the glycan structure to prepare for the transport to Golgi complex where further elongation in the glycan structure occurs. The glycans can be elongated forming diverse types of *N*-glycosylation e.g. hybrid and complex type.

Although all *N*-glycan modifications present a common core, $GlcNAc_2Man_3$, they are classified in 3 different types: high mannose, complex and hybrid. The high mannose type consists of the common pentasaccharide core structure elongated only with mannose residues. The mannose residues number can vary from 5 up to 9. Complex type *N*-glycan modifications presents the common core modified with a more diverse range of structures including alternating Gal, GlcNAc, Neu5Ac, Fuc and Xyl residues. And the hybrid type modifications presents a mixture of the other two modifications for instance presenting one branch of the pentasaccharide core elongated with mannose residues only and another branch modified with the complex type structures. The complex type can also be subdivided according to the branching. When each final mannose from the common core presents linear elongation of the two terminal mannose arms with GlcNAc or alternating GlcNAc and Gal residues (lactosamine chains, LacNAc), this structure is called biantennary. On the other

hand, when one or both of the terminal mannose residues in the pentasaccharide core structure are further branched with additional GlcNAc or LacNAc residues triantennary and tetraantennary structures are formed (see Figure 14).[105] Furthermore organisms mature the *N*-glycosylation in different ways by addition of diverse monossacharides and linkage connections to the terminal LacNAc residues or to specific positions of the pentasaccharide core, e.g. yeasts present mainly highy mannose type[106], and plants present Xyl in their *N*-glycan structure[107], while humans and other animals often have their structure terminated with sulfatation, fucosylation and sialylation. Animals can be modified with the sialic acids Neu5Gc and Neu5Ac[100], in constrast human sialic acid termination only presents Neu5Ac.

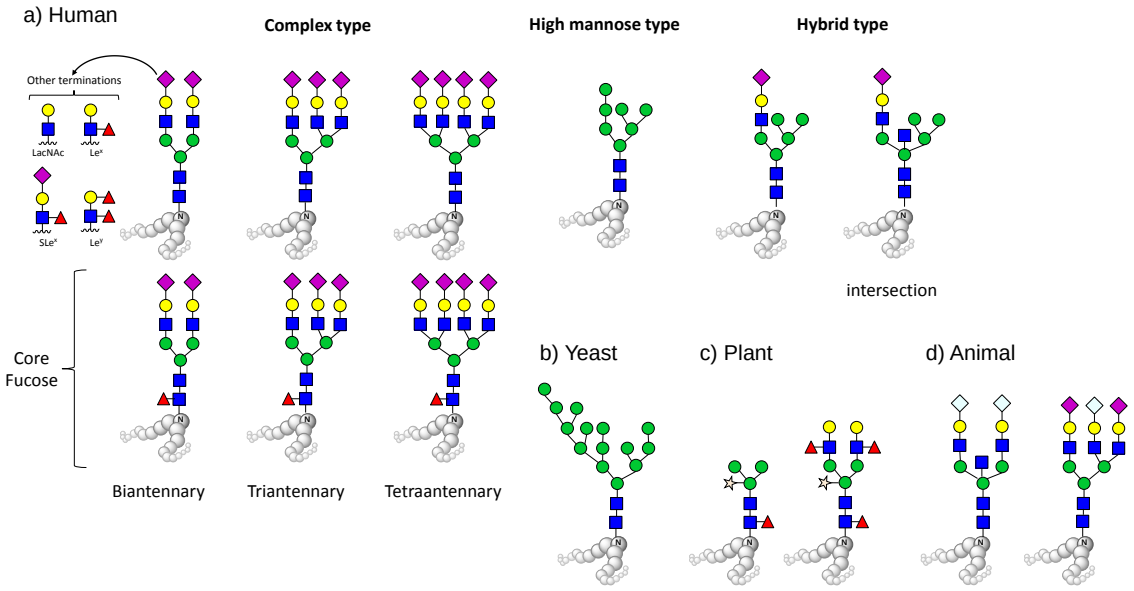


Figure 14. Structural diversity of *N*-glycosylation. Although variations in the structure, *N*-glycans are divided into 3 main types: high mannose, hybrid and complex. The complex type can be also subdivided according to the amount of extension in the pentasaccharide core $GlnNAc_2Man_3$. Structural differences, in both linkage type and monosaccharide building block, occurs for distinct organisms: a) human, b) yeast, c) plant and d) other animals.

5.1.1.2- *O*-Glycosylation

In protein *O*-glycosylation the glycan residue is linked to the oxygen of a Thr, Ser or Tyr amino acid side chain forming an *O*-glycosidic bond (an acetal). Although *O*-glycosylation of hydroxyproline in plans[108] and hydroxylisine in collagen[109] are also reported. Seven different monosaccharides can be introduced directly to the amino acid residue: α-GalNAc, β-GlcNAc, α-Man, α-Fuc, β-Xyl, β-Glc and β-Gal.[110] Due to the large structure diversity, *O*-glycosylation plays a role in different biological functions: mucose protection, immumogenic response, cell signalling, fertilization, cell growth, differentiation and apoptosis. Although mucin like and epidermal growth factor like proteins present tandem repeat motif for glycosylation, there is no consensus sequence that indicates *O*-glycosylation.

Mucins are highly glycosylated proteins, discovered over 100 years ago, present in the epithelial surfaces of the body. Mucin-type glycosylation is the most common *O*-glycan modification and is initiated by introduction of an α-GalNAc residue that is subdivided into 8 different core structures by elongation with Gal, GlcNAc and GalNAc saccharides (Figure 15). Some *O*-GalNAc glycans present part of their structures, known as epitopes, where the antibody binds. The biological role of antibodies is discussed in the following section. Figure 15 has also the chemical structures of some epitopes. Other important type of *O*-glycosylation is one present in the epidermal growth factor (EGF) which plays a crucial role in the Notch signaling pathway, resposible for the cell fate. The modification is also present in the cell surface and, in contrast to mucin-type glycosylation, is initiated by introduction of α-Fuc, β-Glc or β-GlcNAc.

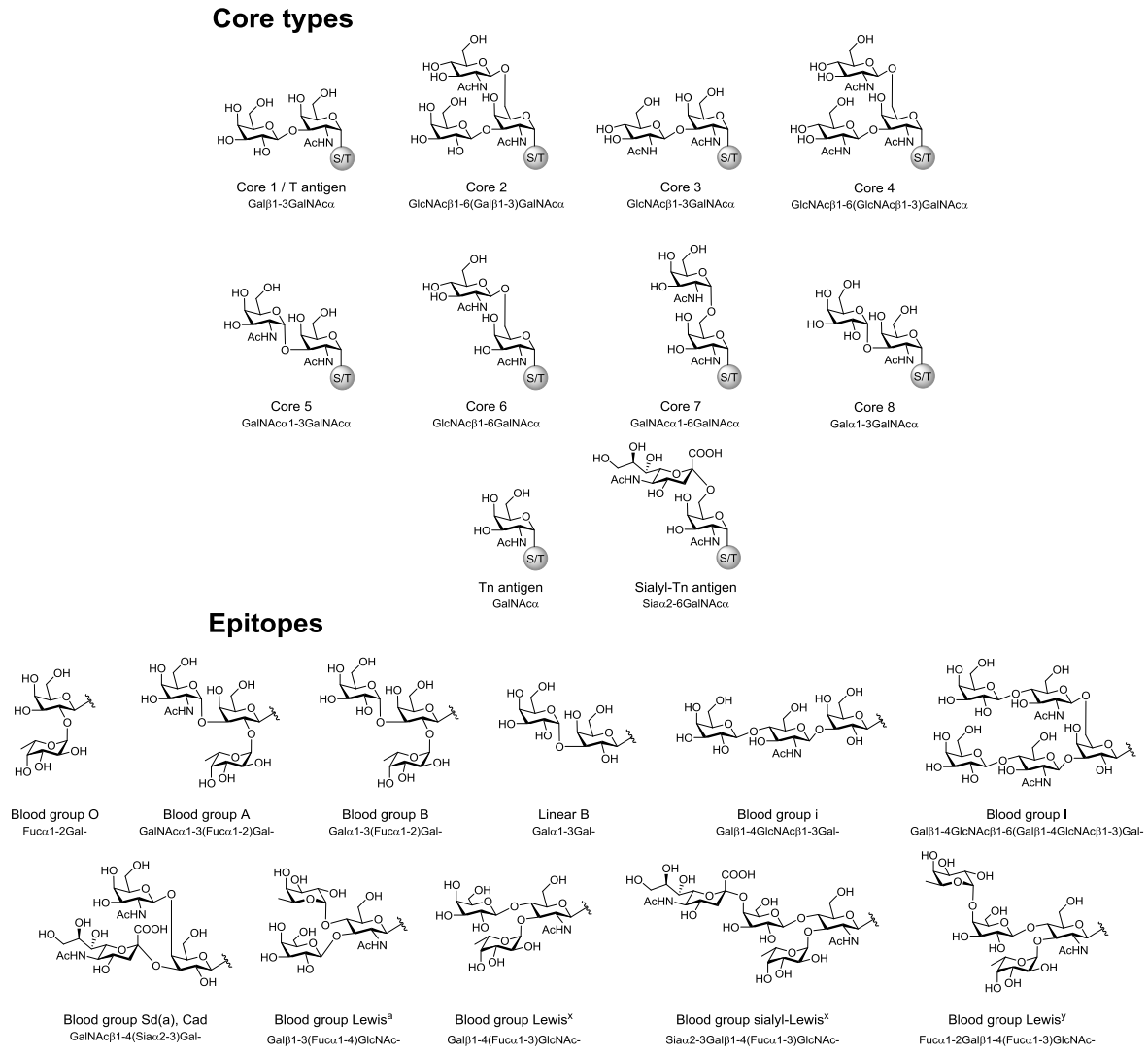


Figure 15. Chemical structure of the *O*-glycan mucin core types and epitopes.

Specific glycosyltransferases link the glycan unit to the substrate and elongate the glycan modification. The biosynthesis of *O*-glycosylation modification happens in the ER-Golgi complex with glycosyltransferases that attach each glycan unit and elongate them directly in the protein backbone

without the intermidiate of the phospho-isoprene activator, like in the *N*-glycosylation. Termination of the *O*-glycosylation normally proceeds by sialylation, fucosylation or sulfonylation. Figure 16 illustrates the biosynthesis of the cores 1-4. The exception is the β-*O*-GlcNAcylation modification that usually is presented without further extension. β-*O*-GlcNAcylation occurs exclusively in the cytoplasm and nucleus, which is a dynamic modificantion that competes with phosphorylation.[111]

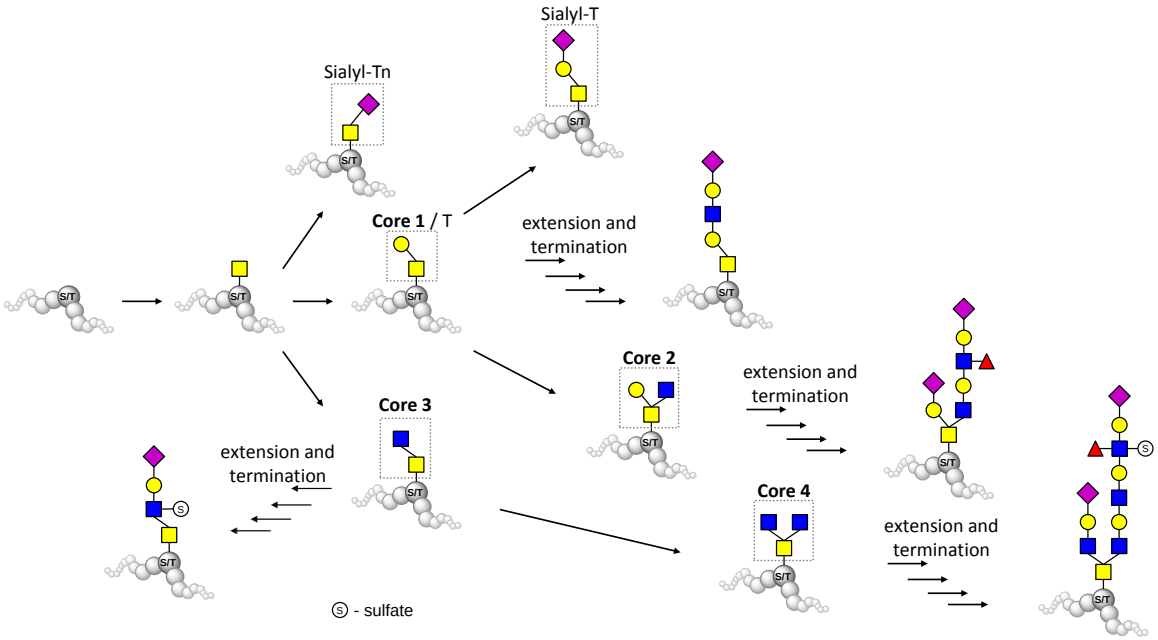


Figure 16. Biosynthesis of mucin-type *O*-glycoproteins. Herein the potential *O*-glycoproteins are modified and elongated complex by glycosyltransferases direct in the glycosylation site (S/T) in the ER-Golgi complex. The formation of the main cores 1-4, sialyl-T, sialyl-T_n and examples of extension and termination are demonstrated.

5.1.1.3- *C*-Glycosylation

The binding of a mannose unit at the indole C2 carbon of a tryptophan is a special glycosylation modification and not fully understood (Figure 17). The modification occurs in the first tryptophan of the consensus sequence Trp-Xxx-Xxx-Trp, where Xxx is any other amino acid, by a specific mannosyltransferase. Although its role is not yet clarified, this glycosylation is involved in protein folding, targeting of substrate proteins and modulating functions.[112]

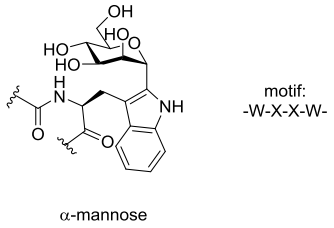


Figure 17. *C*-mannosylation of tryptophan (W).

5.1.2- Biological implications: Glycosylation role in diseases

When conjugated with other macromolecules like lipids and proteins (Figure 18), carbohydrates participate in cell-cell recognition, cell differentiation, signaling events and pathways regulation.[113-114] In this way glycosylation dysregulation is associated with cancer, inflammation, congenital, metabolic, autoimmune and neurodegenerative diseases.

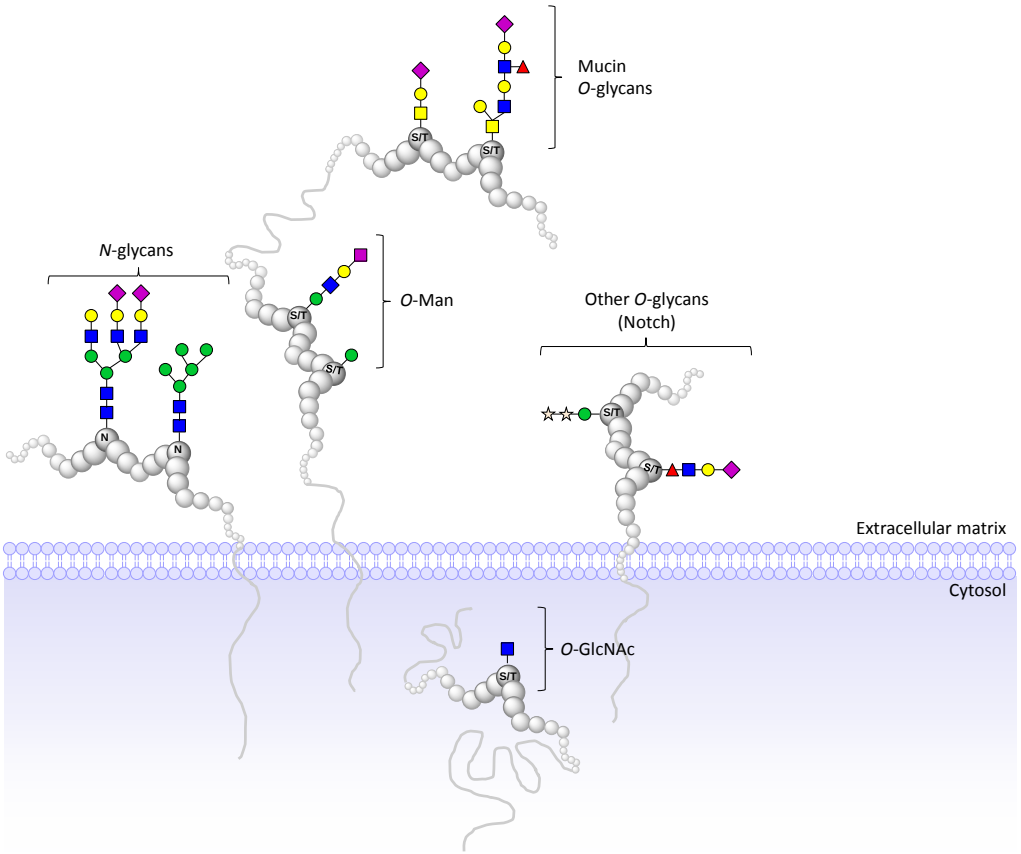


Figure 18. Glycosylation as protein modification.

Defects in glycosylation pathways for instance leads to a group of diseases known as congenital disorders of glycosylation (CDG). These disorders were first characterized by genetic mutations of proteins responsible for the synthesis of *N*-glycoproteins, and recently genetic defects on proteins involved in *O*-glycosylation were included in this group as well.[115] Depending on the mutated enzyme, CDG are classically divided in two groups[116] although a new nomenclature was proposed.[117-118] Type 1 involves all defects regarding the lipid-linked oligosaccharide chain, while type 2 coves defects in the process of the recent synthesized glycoprotein (see Figure 13). Most of the CDG presents symptoms as psychomotor retardation, hyptonia, strabismus, cerebellar underdevelopment, microcephaly, dismorphy, liver disease, muscular dystrophy, among others.[115, 119] The diagnosis of classic cases occurs via analysis of serum transferrin by isoelectrofocusing. Transferrin is one protein abundant in serum and presents two glycosylation sites with biantennary structure, with four sialic acid residues in total. Type 1 pattern involves an increase of di- and asialotransferrin form, and type 2 pattern, besides the type 1, increases of tri- but not always monosialotransferrin.[120] Although classic CDG cases are related to mutations of enzymes that participate on the *N*-glycoprotein synthesis, mutations of proteins responsible for *O*-fucosylation[121], *O*-xylosylation[122], *O*-GalNAclation[123], *O*-mannosylation[124-125] and other glycosylation pathways are also reported.

Infection starts when the pathogen invades the host cell organism. The first contact pathogen-host occurs via interaction ligand-receptor and glycosylation structures are frequently targeted during this initial stage. Hemagglutinin, one protein of Influenza A virus membrane, binds to the sialic acid glycans on the respiratory tract cell surface and is responsible for the first contact of virus to the host cell plasma membrane.[126] In this process, hemaglutinin also supports the fusion of the infecting virus with the host cell membrane in the initial invasion first step.[127] The Gram-negative bacterium *Helicobacter pylori* is responsible for a silent infection in the gastric epithelia that can lead to further complications as gastric ulcers and gastric carcinoma.[128] This bacterium reveals to be adaptable to changes in the host binding enviroment: one of its main studied adhesins *H. pylori* BabA (blood group-antigen binding adhesin) binds either to Lewis b or H-type 1 structures of healthy epithelial cells, while *H. pylori* SabA (sialic acid-binding adhesin) binds to inflamed gastric mucosa via sialylated structures as sialyl-Le^a and sialyl-Le^x.[129] *Escherichia coli* is also a Gram-negative bacterium part of human intestinal flora, however it is also responsible for most of the cases of urinary tract infection. FimH adhesin present in *E. coli* fimbria is one of the best characterized bacterial adhesin and known for being a mannose binding-specific protein.[130-131] A recent study perfomed by Altarac and coworkes compared the treatment efficacy in women urinary tract infection by admistration of *D*-mannose with nitrofurantoin. Although this study was performed in a small scale, both treatments presented similar efficacy with significantly lower risk of side effects for the *D*-mannose patient group.[132]

The interaction between immune system and different harmful agents, such as pathogens and damaged cells, also occurs via glycosylation receptor-ligand interaction. Glycosylation is so important in these interactions that glycan complexity has changed according to evolution. The use of mannose linkages decreased while the use of different monosaccharides increased from lower eukaryotes to vertebrates.[92] These differences led to the creation of epitopes that are able to distinguish cell lines, or indicate diseases and prognosis.[133-134] The expression of lectins responsible for the specific recognition depends on the cell location and surroundings. E- and P-selectins, expressed in endothelial cell and platelets respectively, recognize sialyl-Lewis X and has its expression

induced during inflammation. L-lectins are expressed in leukocytes and recognize 6-sulpho-sialyl-Lewis X.[135] These lectins play crucial roles in leukocytes attachment to endothelial cells, platelets and other leukocytes. In this way there is an interaction between the E- and P-lectins on the surface of endothelial cells and platelets with the glycoproteins in the cell surface of leukocytes and mutually L-selectins in the leukocytes interact with glycoproteins in the surface of platelets and endothelial cells. This process is known as leukocyte rolling.[136] Interactions between leukocyte-leukocyte and leukocyte-platelet play a major role in the recruitment of leukocytes, amplifying the signal.[137-140] Immunoglobulin (Ig), also known as antibody, is a very specific glycoprotein family responsible for recognition of foregin material. The current glycosylation of immunoglobulins is critical for activation and deactivation signaling of the immune system. Alterations in the glycan compositions, presence and location are linked to autoimmune diseases. In rheumatoid arthritis, the glycosylation on IgG is an important biomarker for the diagnosis. Levels of terminal galactosylation (less terminal sialylation) in IgG are related to the disease severity.[141] An evidence that supports this correlation is that patients with rheumatoid arthritis during pregnancy present a temporary increase of galactosylation that provokes a brief disease improvement due to the temporary changes in the immune system to accommodate the fetus.[142] In Henoch-Schönlein purpura, an autoimmune nephropathy, patients presented under *O*-galactosylation levels of IgA1 compared to controls. On the other hand, hypersialylation of IgA leads to an accumulation of IgA in the serum which is correlated with Sjögren's syndrome.[143] Genetic modifications of enzymes that promotes glycosylation can also induce autoimmune diseases. Studies with knockout mice for α-mannosidase IIa1, an enzyme in the *N*-glycosylation pathway, revealed a systemic autoimmune disease similar to human systemic lupus erythematosus.[144]

Cancer is often associated with changes in the glycosylation of proteins and lipids at the cell surface. Comparison between *O*-glycans in healthy and cancer cells shows that the cancer associated glycans are often truncated, presenting T_n and T antigens, high sialylated and less sulfated structures.[145] T-cells, one type of lymphocytes, are reported to recognize *O*-glycan mono- and disaccharides, although there is not deep knowledge about the recognition details.[146] In this way, as previously mention, the immune system not only protects against pathogens but also keeps the organism clearing malignant cells. However, the failure of the immune system to recognize tumor cells as foregin in cancer is a major challenge.[147] For this reason extensive efforts are made in the development of antitumor vaccines in order to active the immune system.[148] Alteration in the genetic expression of glycosyltransferases leads to premature glycan termination generating truncated aberrant membrane bound mucins. Changes in the glycan structure in proteins and cell membrane are associated with all cancers.[149] In breast cancer, as example, truncate structures of *O*-glycans, as well as β1-6 branching of *N*-glycans, and presence or over-expression of Lewis-type epitopes were identified.[150] The increase of sialylation, especially of Lewis epitopes, was related with metastasis, reincidence and poor prognosis in colon cancer.[151] Alteration in the glycoform of α-fetoprotein, one of few approved biomarkers, shows a good indicator of disease progression even distinguishing hepatocellular cancer from other chronic liver diseases.[152] Analysis of α-fetoprotein glycoforms shows that in cases of hepatocellular cancer the protein carries a monosialo biantennary form either with or without fucosylation. Meanwhile the benign glycoform of α-fetoprotein appears to be a disialo biantennary structure.[153] Recently, Lubman and coworkers have described a possible panel of glycoprotein biomarkers for pancreatic cancer. They report the combination of

glycoproteins, α-1-antichymotrypsin, thrombospondin-1 and haptoglobin, outperformed the well-known marker CA 19-9 in distinguishing pancreatic cancer from normal controls.[154]

Neurodegenerative diseases are also related to defects in glycosylation. Alzheimer is the most common neurodegenerative disease. Alteration on protein sialylation is often reported in this disease.[155-156] Moreover alteration in glycosylation of specific proteins, as amyloid precursor protein (APP), tau, acetylcholinesterase and other proteins, has a huge impact on neuronal malfunction.[157] Precursor of amyloid-β, APP is involved in synapse regulation[158], neural plasticity[159] and iron export[160]. Abnormal APP glycosylation results in accumulation of the protein in perinuclear region and interferes with APP sorting and secretion[161-164] and those problems are often associated with brain damage.[165] Aggregation of hyperphosphorylated tau, protein responsible for microtubes stabilization, is commonly related as primary marker of Alzheimer disease.[166-167] Unusual tau glycosylation makes the protein more susceptible to phosphorylation and less to desphosphorylation leading to an accumulation pair helical filament in the cell interfering in the microtubule and axonal transport.[168-169] On Parkinson's disease, the protein agglomerate, named Lewy bodies, composed mainly of α-synuclein is a pathogenic symptom.[170] Parkinson's patients present highly glycated proteins and glycated Lewy bodies have higher immunoreactivity.[171] Although the disease mechanism is not completely understood, the increase of resistance to proteolytic degradation is supposed to influence the accumulation of aggregate and neuronal death in the end of the process.[172] Due to its high abundance in the Lewy bodies, high lysine content in its sequence and ability to bind to DNA[173], α-synuclein is a presumed target for glycation in Parkinson's disease. Glycation is a glycosylation process without the assistance of an enzyme in organisms. Multiple Sclerosis is a neurodegenerative and autoimmune disease affected by the genetics and environment.[174-177] Although the isolation of both components is a challenge, Demetriou and collaborators showed that the dysregulation in *N*-glycosylation pathway is involved in the molecular mechanism of the disease. Up- or downregulation of mannosyl transferase MGAT-1 interferes in the glycosylation of T-cells in which reduced *N*-glycosylation leads to a loss of self-tolerance, increasing the risk of multiple sclerosis. On the other hand, vitamin D3 signaling increases *N*-glycosylation branching, enhances the immunoreactivity by alteration of CTLA-4, IL-2 and IL-7 levels which directs to a decrease of multiple sclerosis risk.[178]

5.1.3- Analysis of glycoproteins

Due to its importance, glycoproteins are a potential source for biomarker discovery in order to detect diseases at early stage and to increase patient survival. Efficient methodology can monitor disease progression that is often related to pathway dysregulation including changes of the glycosylation extension as well as the carbohydrate structure.[179] At this point two major approaches are used to characterize and understand the role of glycans in the biological system. Glycomics is the study of the structure and function of the glycans, attached or not to other macromolecules, expressed in a cell or organism (glycome). This study includes the identification of genes responsible for encoding proteins related to glycan synthesis and glycoproteins. On the other hand, glycoproteomics is the identification and characterization of proteins which possess glycans as post-translational modifications including glycosylation sites, micro- and macroheterogeneity, using a

mix of techniques applied to glycomics and proteomics. There are levels of analysis proposed for better understanding of the entire glycome. First level is glycoprofiling, which comprises the analysis of individual glycans structures, creating a glycan fingerprint of the system. Second level deals with the glycan variability and/or degree of occupancy of a given glycosylation site in proteins. And the third level consists in the comprehension of cellular communication and the influence of the glycome in different cells and tissues. Although proteomics is a solid research field with robust and well-established techniques, glycomics and glycoproteomics are emergent research areas of a great interest in order to understand more complex biological processes.[180-181]

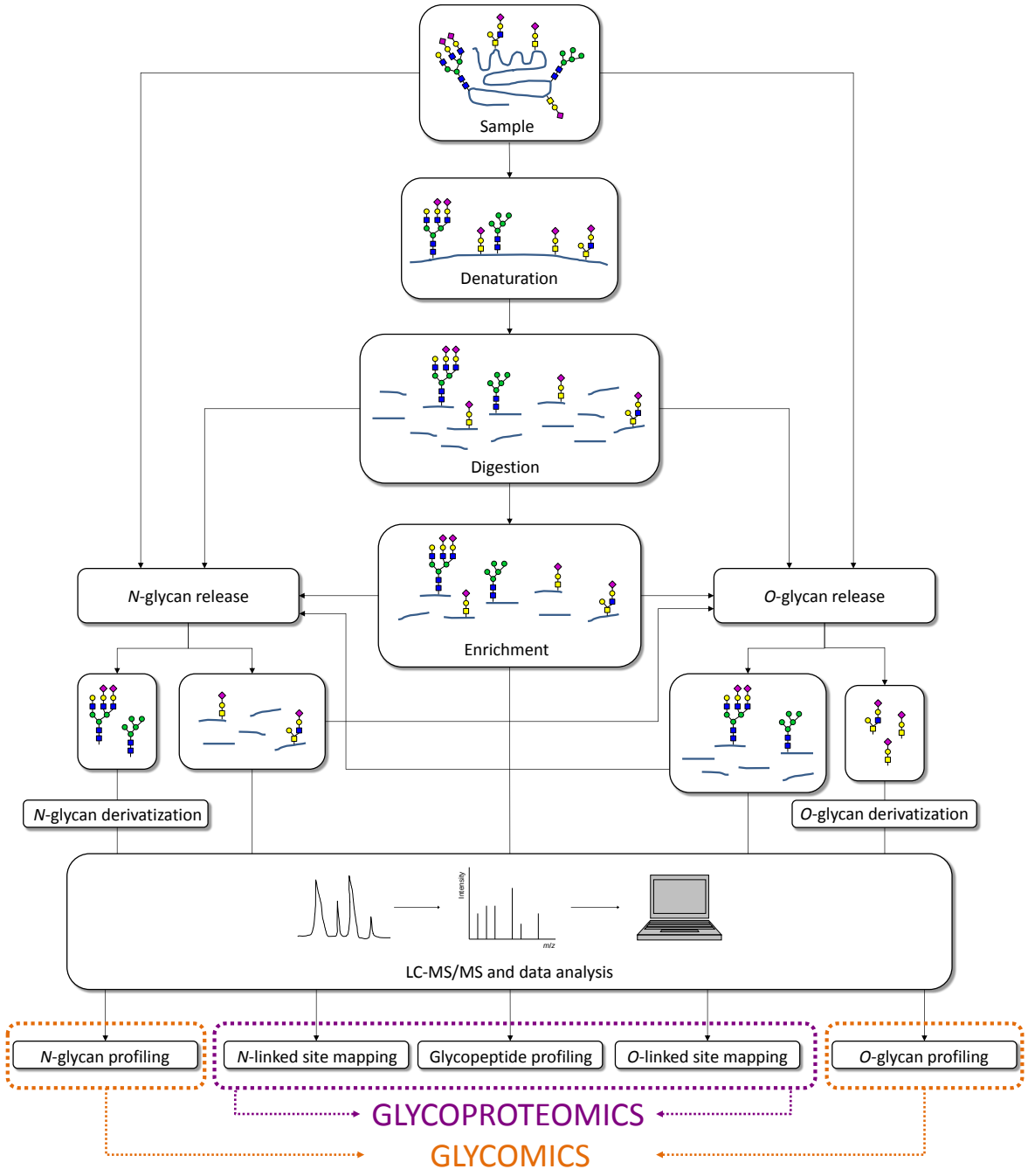
Glycosylation and glycan analysis is a broad subject and here the focus is on the discussion of glycosylation as a protein post-translational modification. Advances in separation and mass spectrometry methods make analysis of glycoproteins possible, including site identification and characterization of glycan structures even in complex samples. Different approaches have been created in order to obtain a better knowledge of protein glycosylation: analysis of intact glycoproteins or glycopeptides, analysis of glycosylation sites by glycan release, and analysis of released glycans. Analysis of entire proteins is known inside the proteomic field as “top-down” approach. In glycoproteomic “top-down” approach has the advantage of providing information of the protein glycosylation in its native form, avoiding modification losses due to wet-lab steps for sample preparation. Besides that, this approach also opens the possibility of observation of glycosylation interactions in protein complex. On the other hand, the use of this approach requires high specialized expertise, extensive protein purification prior MS analysis and challenging data analysis due to the combinatorial arrangement of possible isomers.[182] “Bottom-up” approach in proteomics is the analysis of peptide mixture after proteolysis instead. With this technique the identification of thousands peptides from digested proteins is possible. For this reason this is by far the most used approach in the field. However, the number of glycosylated peptides is a minor portion compared to the total number of peptides in a digested mixture, which leads to the main drawback of this approach in glycoproteomic field. Therefore enrichment of glycopeptides is necessary to obtain a detailed coverage of the protein glycosylation profile. A single glycoprotein can have a variable number of glycosylation sites (macroheterogeneity) as well as variable glycan structures at each glycosylation site (microheterogeneity).[183] There is also an intermediate approach called “middle-down”, applied for analysis of IgG antibodies, where the antibody is fragmented into 3 subunits enabling the differentiation of close related protein isoforms.[184]

5.1.3.1- Glycomics vs. Glycoproteomics: From sample preparation until data analysis

Depending on the approach, there are different steps in the workflow for the analysis in glycomic or glycoproteomic field, resumed at Scheme 12. A crucial step of the glycomic approach is the release of the glycans. The current strategy uses enzymes in order to cleave *N*-glycans while *O*-glycans are chemically cleaved. The most commonly used enzyme for *N*-glycan cleavage is the peptide-*N*-glycosidase F (PNGase F), which has a broad specificity, cleaving the bond between the first GlcNAc residue and asparagine, converting the asparagine into aspartic acid during the process and adding a mass shift of 0,9840Da to the original peptide. PNGase F does not cleave *N*-glycans that possess a α-1,3-Fuc attached to the first GlcNAc residue and is inefficient releasing microbial *N*-glycans.[185] Although there are alternative glycosidases to circumvent PNGase F limitations, as

PNGase A, they do not present the same efficiency. For the release of *O*-glycans, there is no glycosidase as efficient as PNGase F. Endo-α-*N*-acetylgalactosaminidase, known as *O*-glycanase, cleaves only Gal-GalNAc structures and besides that, the enzyme does not tolerate any extension in the disaccharide motif restricting the application.[186] Therefore the chemical procedures for *O*-glycan release that comprise reductive β-elimination[187], ammonia-based non-reductive β-elimination[188] and hydrazinolysis[189] are used instead of enzymatic cleavage. In order to increase the sensitivity, the released glycans are derivatized either by permethylation or labelling of the reducing end with chromophores, e.g. 2-aminobenzamide or 2-aminopryridine. Permethylation is etherification of free hydroxyl groups of the native glycan and esterification of the carboxyl group in the sialic acid. Traditionally, iodomethane is used as methylating agent and the reaction is carried under basic conditions. This is the most common derivatization and has as advantages increase of MS sensitivity by reducing glycan hydrophilicity, provides information about the monosaccharides sequence and enhance cross-ring fragmentation during tandem MS. Although the glycan release approach permits a deep analysis of the glycan structure, the attachment site and occupancy rates information are lost.

During sample wet-lab procedures, the deamidation of asparagine can happen spontaneously, which might lead to false *N*-glycosylation site determination.[190] One attempt to bypass the *N*-glycosylation site problem is via analysis of labelled deglycopeptides. *N*-deglycosylation is performed under "heavy" water ($H_2{}^{18}O$), in this way ^{18}O-labelled deglycosylated asparagine presents a peptide mass shift of 2,9882Da.[191] However difficulties to control the deamidation step of asparagine are reported in the literature.[190] The SimpleCell approach was designed by Steentoft *et. al.* as a completely new strategy to identify *O*-glycosylation sites.[192] Herein, genetic tools are used to silence further elongation of the initial GalNAc-*O*-glycosylation in different human cancer cell lines (Capan1, Jurkat, K562, Colo205 and T3M4). Later, *Vicia villosa* agglutinin (VVA) is employed for enrichment of truncated *O*-glycopeptides and after LC-MS/MS more than 350 GalNAc-*O*-glycosylation sites in over 100 proteins were identified, including *O*-GalNAc-Tyr residues. Although all the information regarding the *O*-glycosylation elongation is lost, the investigation of further *O*-glycosylation sites with this technique was also applied in posterior work as the mapping of human GalNAc-*O*-glycoproteome[193] and *O*-mannose glycosylation sites[194].



Scheme 12. Glycomics vs. glycoproteomics: simplified workflow for glycoprofiling assays. Samples can be either analyzed with standard proteomic procedures for glycopeptide profiling or deglycosylated before LC-MS/MS experiments for the site mapping or glycan profiling. The analysis *N*- and/or *O*-linked site mapping and glycopeptide profiling comprise the glycoproteomic area. While the *N*- and/or *O*-glycan profiling without the peptide backbone comprehends the glycomic field.

Even if a complex mixture of oligosaccharides can be analyzed by direct MS injection, the use of separation methods, liquid chromatography or capillary electrophoresis, prior MS analysis is preferable because retention or migration times can be combined with precursor mass and fragmentation spectra information in order to better characterize and distinguish glycan isomers. Online database as UniCarb-DB[195] already uses retention time as an identification parameter. HPLC

coupled online with MS has proven to be the main technique for characterization, identification and quantification, especially after improvements in sensitivity and resolution, not only in glycomics, but in all omics areas. The use of porous graphitized carbon (PGC) columns in HPLC shows a high separation power for structural isomer oligosaccharides[196], because the retention of analytes not only depends on hydrophobicity, it also depends on the interaction of polarized groups with graphite.[197] Hydrophilic interaction chromatography (HILIC) is a popular technique to separate polar compounds, which include native oligosaccharides. Details regarding this technique are described in the enrichment methods section. Another separation technique used for unmodified saccharides is high-performance anion-exchange chromatography coupled with online desalting and MS.[198] Besides all the improvement of HPLC and ultra-high performance liquid chromatography (UHPLC) allowed the increase of the separation power of the complex mixtures. Apart from separation chromatography methods, a gas-phase separation method ion mobility spectrometry (IMS) got a lot of interest after demonstration of partially separation of GalNAc and GlcNAc oxonium ions.[199] Compounds with the same mass but different structures can present different collisional cross sections; therefore different isomers differ in drift time, allowing the separation of isobaric species in gas-phase when coupled on-line with a mass spectrometer.[200] However, the absence of reference compound to create a reliable database and software for IMS data analysis rise difficulties for the use of this methodology.

The development in the mass spectrometry instrumentation opened a possibility of advanced accurate mass analysis with high resolution. The standard ionization techniques, MALDI and ESI, are used to analyze glycans. The tolerance to salt contaminants and automation capabilities are the main advantages of MALDI.[201] On the other hand ESI ionization is reported to be gentler than MALDI. Besides that ESI is coupled easier with separation methods and is a better ionization technique for hydrophobic compounds, which can be nicely combined with permethylated glycans. However ESI sensitivity decreases with the increase of glycan mass.[202] In the opposite direction of peptide analysis, where tryptic peptides are preferably analyzed in positive mode, the use of negative mode is quite attractive for glycan analysis, particularly for analysis of sialylated glycans.[203-204]

Even though a lot of information in the glycan analysis is obtained, the analysis of protein glycosylation in peptide level has a huge importance. Determination of glycosylation site and site occupancy are the main information obtained in the glycoproteomic field. The first step for glycopeptide analysis is the proteolytic digestion of the proteins. Depending on sample complexity, protein separation on gel or other separations techiniques might also be required prior digestion. Trypsin is the most used protease for protein digestion, it has a specific cleavage site mainly at the carboxyl side of the amino acids lysine or arginine, except for residues followed by proline. This cleavage gives the peptide a minimum of one positively charged amino acid side chain, important for the MS analysis in positive mode. The trypsin specificity simplifies the calculations for data analysis in the steps after data acquisition. However, glycosylation in proteins is often used to hinder proteolytic cleavages, which leads to large and heavily glycosylated peptides that are difficult to be analyzed by MS.[205] The use of broad-spectrum proteases such proteinase K[206], pronase[207] or subtilisin[208] can overcome such problems, generating smaller glycopeptides with accessible modification sites suitable for MS analysis. Another recurrent problem in glycopeptide analysis comes due to glycosylation micro- and macroheterogeneity: glycosylated peptides normally occur in substoichiometry compared to unmodified peptides and their glycan diversity contribute to the

reduction of the abundance of each glycopeptide species. Enrichment methods are here required to separate glycopeptides prior analysis. These methods are discussed in a separated section with more details.

Depending on the glycan size and MS/MS approach different information is obtained regarding peptide and glycan fragmentation. A nomenclature system for glycan fragmentation was proposed by Domon and Costello in 1988 and the labels were inspired in the peptide fragmentation system.[25-26] Herein, A_i, B_i and C_i are annotated to terminal (non-reducing end) glycan units, while X_j, Y_j and Z_j represent the ions that contains the reducing end units. Relative positions are indicated with subscriptions and the cleavages within the carbohydrate ring are indicated with superscriptions. Bond numbering that indicates ring cleavage are also indicated in the Figure 19.[209]

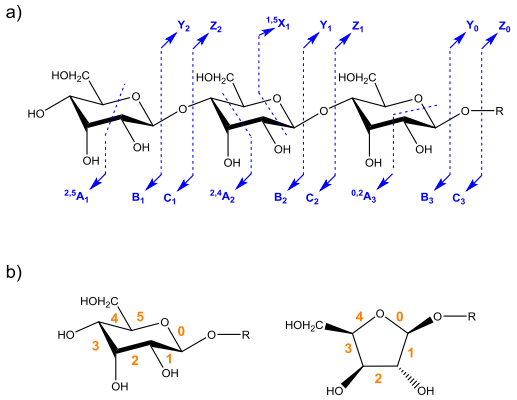


Figure 19. Representation of glycan fragment nomenclature (a) and bond numbering (b) for ring cleavage according to Domon and Costello rules.[209]

Although collision-induced dissociation (CID) is the most common fragmentation technique for peptides in proteomics, when it comes to glycopeptide fragmentation the peptide backbone remains intact and the glycosidic bonds are broken during the process, generating B- and Y-ions. Mainly the glycan sequence, which means the order of Hex and HexNAc, can be deduced from this analysis. No information regarding the structural isomers, linkage position nor configuration is obtained. Moreover the "one-third" *m*/*z* cutoff rule in CID fragmentation might impede the detection of oxonium ions and/or ring cleavage fragments if the peptide has a huge glycan modification. Nevertheless, *N*-glycosidic bonds (pep-HexNAc or Y_1-ion) are more stable under CID fragmentation than *O*-glycosidic bonds for this reason only *N*-glycosylation sites may be confirmed after CID-MS^3 analysis with peptide sequencing of the Y_1-ion, $[pep\text{-}HexNAc]^+$.[206, 210] The combination of CID with other fragmentation techniques is also explored to overcome its limitations. Halim and collaborators combined CID with electron capture dissociation (ECD) to investigate glycoproteins in human urine. Over 100 *N*- and *O*-glycopeptides were identified using CID-ECD and CID-MS^3 approach.[211]

A particular good alternative for glycosylation site determination is the use of electron transfer dissociation (ETD) or ECD fragmentation. An election from a radical anion is transferred to the multiple charged glycopeptide that cleaves the peptide N-C_α bond. Unlikely in CID fragmentation, even labile modifications, as *O*-glycosylation, are generally retained during peptide backbone

fragmentation. Halim and coworkers used ECD to report for the first time a mammalian extracellular protein with tyrosine *O*-glycosylation. The Tyr_{681} containing $Neu5Ac_2$-Hex-HexNAc was found in amyloid beta from CSF samples, furthermore an upregulation of this glycosylation was observed in Alzheimer patients when compared to control.[212] Only ETD was used to identify over 2 500 unique *N*- and *O*-glycopeptides on 453 proteins enriched by wheat germ agglutinin (WGA) in large-scale study leaded by Trinidad and collaborators in murine synaptosome.[213] On the other hand, no information regarding the glycan structure is obtained besides the oligosaccharide monoisotopic mass, once there is no cleavage on glycosidic bonds using ETD/ECD. In addition to that, the necessity of charges ≥2+, preferentially ≥3+, for observing the precursor is a drawback.[214] Low fragmentation efficiency, around 20%, particularly for precursors with *m*/*z* higher than 850 is also a disadvantage for trace analysis in biological samples.[215-216]

Higher-energy collisional dissociation (HCD) has its popularity rising related to glycoproteomic analysis. Due to the instrumental design, high resolution and full *m*/*z* 100-2000 MS/MS are possible to obtain and permit the detection of small oxonium ions and ring cleavage fragments.[31] Both glycan moiety and peptide backbone are fragmented simultaneously with HCD, which allow peptide sequencing and glycan characterization in a single MS^2 step. Characterization of different fragmentation pattern of *N*- and *O*-glycopeptides generated from a peptide mixture of standard glycoproteins according to different normalized collisional energy (NCE) values were evaluated by Cao and coworkes.[217] This study suggests that HCD at lower NCE values prefers to cleave glycosidic bonds enabling glycan structure characterization, while HCD at higher NCE values are prefered for peptide bond cleavage providing confident information for peptide sequencing. Segu and Mechref reported for the first time that Y_1 ions of *N*-glycopeptides are the most abundant ions in HCD spectra and they used this ion for further fragmentation, in a MS^3 experiment, in order to obtain peptide sequencing of *N*-glycopeptides.[218] In 2014 MS was used for distinguishing glycan isomers, in this work, Halim and coworkers in collaboration with our group (Synthetic Biomolecules), observed that GalNAc and GlcNAc oxonium ions from *O*-glycopeptides produce different fragmentation pattern under HCD and CID methods and this pattern could be used to differentiate both structural isomers.[219] Experiments with deuterated substituted glycans suggest that GalNAc and GlcNAc fragementations undergo preferentially through different decomposition pathways independently of further substitution or fragmentation technique, which allowed the isomers distinction (Figure 20).[220] A GlcNAc/GalNAc ratio was defined as:

$$\frac{GlcNAc}{GalNAc} = \frac{(I_{138}+I_{168})}{(I_{126}+I_{144})}$$

where I is the relative intensity of fragment ions *m*/*z* 126, 138, 144, 168 (Figure 20). For synthetic *O*-GalNAc glycopeptides, the observed GlcNAc/GalNAc ratio values fit in a range of 0,4-0,6, while the observed GlcNAc/GalNAc ratio values for synthetic *O*-GlcNAc glycopeptides are in a range of 2,0-2,5.[219]

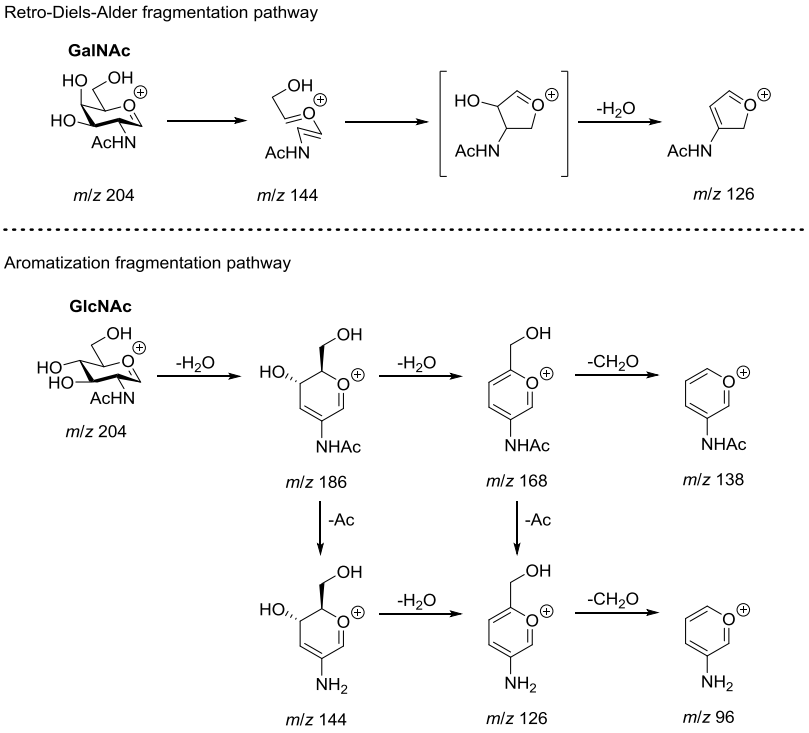


Figure 20. Proposed fragmentation pathway for HexNAc oxonium ions. Those ions can decompose under tandem mass experiments via two different fragmentation mechanisms: Retro-Diels-Alder pathway is the favored way observed for GalNAc oxonium ions (upper) and aromatization pathway is preferential for GlcNAc oxonium ions (lower), which leads to different proportion of fragment ions allowing both isomers distinction.[220]

The relative abundance intensities of oxonium ions for type-1 (Gal-β1,3-GlcNAc) and type-2 (Gal-β1,4-GlcNAc) disaccharides glycopeptides were also different in MS measurements. Synthetic glycopeptides containing different type-1 and type-2 elongations were analyzed using HCD fragmentation and different NCE in order to monitor the relative intensities of generated oxonium ions. Although the intensities profiles slightly change under different NCE levels, at NCE 35%, the oxonium profile for type-1 glycopeptides follows the order *m*/*z* 204>138>186>168, at the same time as type-2 glycopeptides oxonium profiles has the order *m*/*z* 138>204>168>186.[219] Later on, an analogous observation from our group regarding different relative oxonium ion intensity from Neu5Ac-α2,3-Gal and Neu5Ac-α2,6-Gal was also observed.[208] Herein, Neu5Ac-α2,3-Gal oxonium ion profile follows the order *m*/*z* 274>204>366>675, while Neu5Ac-α2,6-Gal oxonium ion profile follows the order *m*/*z* 204>274=366>675 (Figure 21). Those studies are examples of recent development in glycan structure elucidation using mass spectrometry techniques.

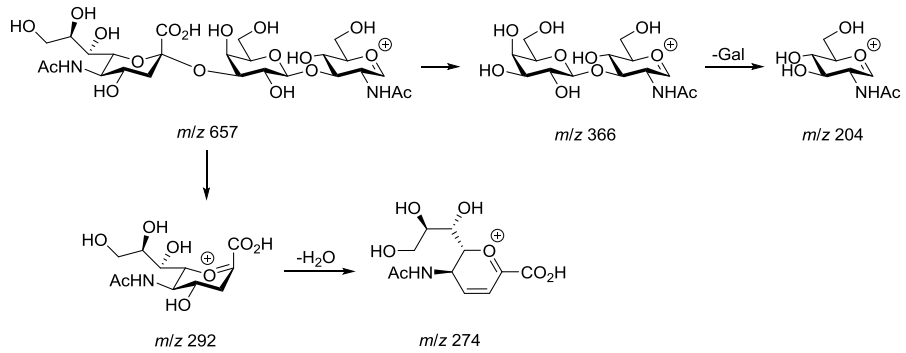


Figure 21. Fragmentation species of trisaccharide Neu5Ac-Hex-HexNAc.

Besides the use of combined fragmentation methods in order to obtain a larger coverage of protein modification, as CID-ECD or HCD-ETD, the use of the hybrid dissociation method proposed by Frese and collaborators electron-transfer and higher-energy collision dissociation (EThcD)[33] has showed to be useful in glycoproteomics. Herein, all ions derived from ETD fragmentation receives a supplementary energy (HCD) in order to obtain a more information per spectra in which a single spectrum presents both fragmentation characteristics and pair of ions, b/y-ions and c/z-ions. Yu and collaborators used EThcD for analysis of fetuin and *N*-glycopeptides from human plasma and rat carotid arteries.[221] Their analysis of human plasma *N*-glycopeptides showed that EThcD outperformed HCD fragmentation, 184 glycoforms were identified using EThcD fragmentation while HCD presented 106 identifications. Rat carotid arteries *N*-glycopeptides were analyzed exclusively with EThcD and over 1 000 glycopeptides were identified. Even more recently, Zhang and collaborators explored the *O*-glycosylation in human serum using HCD and EThcD fragmentation and they point out that EThcD provides more complete information specially regarding to glycosylation site localization than HCD moreover EThcD is a good technique for identification of glycopeptides with multiple glycosylation sites.[222] A total of 499 non-redundant *O*-glycopeptides were identified, including 121 novel *O*-glycosylation sites.

The complexity of glycan and glycoprotein structures as well as the probability of isobaric compositions makes the data analysis a big challenge. Manual analysis of the data set requires significant expertise and demands a lot of time. For this reason, softwares and database are developed to assist in prediction/interpretation of glycan and glycopeptides spectra. However the development of such tools demands time and the lack computational solutions for both *N*- and *O*-glycopeptide composition, structure and site make this an open field for further studies and improvements when compared with consolidated fields as genomics and proteomics.

Standard proteomics software packages[21, 223] are able to handle simple protein modifications such phosphorylation and single glycosylation, but they are not able to search and solve the structures of complex glycans. There were some attempts in using consolidate proteomic software as Mascot, still the resolutions and annotation of the glycan fragment signals should be done manually[211], which demands expertise and time, or using programing language to reprogram the assignment system[224], that might not be accessible to most of scientists.

Table 5. Bioinformatic tools for glycans and glycopeptide annotation.

Software	Description	Application[a]	Analysis level[b]	Type[c]	Ref.
GlycoWorkbench	Semi-automated analysis and annotation of glycan spectra, supports glycan drawing and mass matching.	Gc	MS^n	Java GUI, Linux, Windows, MAC	[225]
SysBioWare	Analysis of raw spectra online, proposed structures based on their own database.	Gc	MS	Web	[226]
Cartoonist	Automatic annotation of *N*-glycans based on MALDI TOF spectra with a confidence score.	Gc	MS	Commercial Windows GUI	[227]
SimGlycan	Annotate fragment spectra with its own database with score.	Gc	MS^n	Commercial Windows GUI	[228]
GlycoSearchMS	Compare spectra with candidates on SweeDB[229] and annotate the fragments according to Domon and Costello nomenclature with GlycoFragment.	Gc	MS^n	Web	[230]
Glyco-PeakFinder	*De novo* determination of *N*-glycan composition.	Gc	MS^n	Web	[231]
GlycoMod	Search for possible composition to a given experimental mass.	Gc,GP	MS	Web	[232]
GlycoPep ID	Identify peptide portion of negatively charged glycopeptides.	GP	MS^n	Web	[233]
Glycopeptide ID	Identify intact *N*-glycopeptide matching MS/MS spectra against protein and glycan database preloaded in the web server. Also include de novo glycan search and score for both peptide and glycan.	GP	MS^n	Web	[234]
Byonic™	Score, rank and identify *N*- and *O*-glycopeptides using separated protein and glycan database. Provide top-down and bottom-up searches and Wildcard for unknown searches. Support all fragmentation types. It is integrated to Proteome Discover.	GP	MS^n	Commercial Windows GUI	[235]
SweetNET	Identify chondroitin sulfated, *N*- and *O*-glycopeptides and cluster related glycoforms. Use the oxonium ions for scoring and gives automatically GlcNAc/GalNAc ratio for all spectra.	GP	MS^n	Windows GUI	[236]
pGlyco 2.0	Designed for *N*-glycopeptide identification. Combine stepped collision energy spectra. Program is validated using metabolically labelled samples.	GP	MS^n	Linux CLI	[237]

Abbreviations used for tool classification: a) Gc: glycans, GP: glycopeptides; b) MS: no fragmentation information used for matching spectra, MS^n: fragmentation information used for matching spectra; c) GUI: graphical user interface, Web: web-based interface, CLI: command line interface.

The first improvements were related to the release of software packages to assign free glycans compositions; some are based on intact mass searches, other in tandem mass fragments. Regarding glycopeptides, the difficulty from the programing perspective is due to the complexity of the molecules. Because of their large size, large number of monomers units (5 glycan unique masses for mammalian organisms and 19 amino acid unique masses – leucine and isoleucine are isobaric

species), a large computational power is required, particularly when low error range is demanded. Nevertheless nowadays there are user-friendly solutions, the description of some software are listed in Table 5.

Table 6. Description and address of some database and repository for glycomics and glycoproteomics.

Database		Description	Web link
General databases that contain glycosylation information	Uniprot	Universal protein database for protein sequence and annotation with modification and biological information. Contain predicted and confirmed glycosylation sites.	www.uniprot.org
	PDB	Database for 3D structures of large biological molecules.	www.rcsb.org
	ExPASy	Portal that provides tools and database in different scientific area, including glycomics.	www.expasy.org
	Human Protein Atlas	Database that contains information about human body in different levels as cell, tissues and organs obtained via "omys" and system biology studies.	www.proteinatlas.org
	PhosphoSitePlus	Portal that contains database and tools for protein PTM's. Contain information about *O*-GlcNAc and *O*-GalNAc sites.	www.phosphosite.org
Glycomic and Glycoproteomic databases	UniCarbKB	Contain structural, experimental and functional information for glycomics and glycobiology.	www.unicarbkb.org
	Unipep	Library of *N*-glycopeptide and proteins. Provide *in silico* tryptic digestion and predict if the glycosylation site is or not detectable via MS^n experiment.	www.unipep.org
	UniCarb-DB	Glycomic MS database and repository.	https://unicarb-db.expasy.org/
	GlycoFish	Database of *N*-glycosylated proteins characterized from *Danio rerio* (zebrafish) embryo by MS.	http://betenbaugh.jhu.edu/GlycoFish/
	GlycoFly	Database of *N*-glycosylated proteins characterized from *Drosophila melanogaster* nervous system by MS.	http://betenbaugh.jhu.edu/GlycoFly/
	GlycoProtDB	Database of N-glycosylated proteins characterized from *Caenorhabditis elegans* and mouse.	https://jcggdb.jp/rcmg/gpdb/
	Lectin Frontier Database	Contain affinity constants for lectin-glycans interactions measured by frontal affinity chromatogram with fluorescence detection.	http://acgg.asia/lfdb2/
	GlycoMob	Collision cross section database for ion mobility mass spectrometry.	www.glycomob.org

Due to all high-throughput techniques and methods quality improvement there is an increase of data generation in glycomic and glycoproteomic field, reflecting in the rising of number of glycan-related databases created in the last decade.[238] Databases can be used for identification or cross-validation of recognized structures. Besides that, they can offer valuable biological information such as binding partner and connection with pathways. In this direction there is a huge importance in the multi-institution project and consortium as EUROCarbDB[239], Consortium for Functional Glycomics[240] and JCGGDB[241] to combine efforts and create accessible tools and database to expand knowledge. Tools like GlycoWorkBench and a huge part of the data available nowadays at UniCarbKB database came from EUROCarbDB project as well as GlycoProtDB database came from JCGGDB

initiative. Some database and repository with valuable information for glycomic and glycoproteomic field are listed in Table 6.

5.1.4- Glycopeptide enrichment

In previous sections, some challenges of intact glycopeptide analysis were disclosed. The low number of glycosylated peptide fragments compared to non-glycosylated in a proteolytic mixture, as well as micro and macroheterogeneicity are main issues in the detection of glycan protein modifications. For this reason, enrichment methods have been developed to address these problems of separating glycosylated peptides prior MS-analysis. Enrichment methods applied for glycosylated peptides and/or proteins are generally based on lectins, chromatographic separations, metabolic labelling and chemical properties of glycans.

5.1.4.1- Methods based on lectins

Lectins are glycan-binding proteins and are found in the most diverse sources like plants, such as legumes and cereals, and animals; lectins are used for a long time in the detection and isolation of glycopeptides and glycoproteins as well as assisting in the characterization of the glycan moiety. In general, hydrogen bonds with van der Waals and hydrophobic interactions are the main forces acting in the glycan-lectin binding.[242] Although the diversity among the lectins is huge, their specificity for certain glycan motif can be really high even though binding often relies on multivalent binding interactions and the binding affinity to an individual glycan residue/epitope is not always strong. The importance of the amino acid sequence in the carbohydrate-recognition domain for lectin specificity is demonstrated by Drickramer. In his work, genetic modification of the lectin domain that recognizes mannose was changed from Glu_{185} and Asn_{187} to Gln_{185} and Asp_{187} residues. This alteration reduces the affinity for mannose and increases the affinity for galactose.[243]

For enrichment proposes, lectins can be used in solution[104, 244] or can be conjugated with solid support as agarose/sepharose[245], silica[246] or polymers[247], used in spin devices, low-pressured LC and HPLC chromatography. After separation, these enriched glycopeptides can be released using a specific saccharide solution[248], generally mono- or disaccharides in higher concentration that compete for the binding, or released under acidic conditions[249], in which the acidic solution breaks the glycan-lectin interaction. Moreover glycopeptides can also be released by deglycosylation using PNGase F[104, 244]. In this case all the glycan information is lost and it can be applied only for investigation of *N*-glycosytion sites.

In general lines, there are a huge number of lectins used for enrichment proposes[250-251], nevertheless, due to their glycan binding groups and availability, some lectins are of more importance and are here highlighted: concanavalin A, wheat germ agglutinin, peanut agglutinin and jacalin. Concanavalin A (Con A) is employed for the enrichment of high-mannose, hybrid and bianntenary *N*-glycopeptides by the recognition of α-mannose and α-glucosamine linkages[252]. Wheat germ agglutinin (WGA) is in turn used to enrich complex type sialylated *N*-glycopeptides, due

to its affinity for sialic acid and *N*-acetylglucosamine.[253-254] For enrichment of *O*-glycopeptides jacalin, which recognizes the core disaccharide Gal-GalNAc linked to Ser/Thr, is often used. Peanut agglutinin (PNA) has specificity similar to jacalin but sialic acid attached to galactose, common in O-glycopeptides, affects its affinity.[255] The enrichment procedure lays on the lectin affinity for a specific glycan motif which restricts the coverage of the glycosylation evaluation, even if the opposite has been demonstrated[256]. In order to overcome this problem, Cummings and Kornfeld proposed the use of multiple lectins in a serial lectin-agarose affinity chromatography.[257] They used concanavalin A-sepharose, pea lectin-sepharose and L-phytohemagluttinin (L-PHA)-agarose columns to study *N*-glycopeptides in mouse lymphoma cells. Using mannose labeled with tritium, they observed that around 85% labelled *N*-glycopeptide were retained in one or more columns. Almost twenty years later, Yang and Hancock developed multi-affinity column by mixing equivalent weights of Con A-agarose, WGA-agarose and jacalin-agarose for the enrichment of *N*- and *O*-glycopeptides from human serum.[258] After depletion (from six most abundant proteins), they observed that 50% (w/w) of the remaining serum proteins were glycosylated, though the authors state that this number might be even bigger, reflecting the lectin limitations.

5.1.4.2- Methods based on chromatography and solid-phase extractions

The term HILIC for hydrophilic interaction chromatography was introduced by Alpert in 1990 to describe a chromatographic methodology in which hydrophilic column is eluted with a hydrophobic mobile phase.[259] The method employs a water-miscible organic solvent, typically acetonitrile as mobile phase, and as stationary phase a hydrophilic solid support that is coated with a water layer in a way that the retention occurs, most probably, via partitioning mechanism. Analogous to liquid-liquid extraction and the retention increases with solute hydrophilicity. In his work, Alpert suggested this technique for the separation of highly polar compounds such as peptides, nucleic acids and other polar compounds even when the methodology was previously used for saccharides analysis and separation of different glycoforms.[260-262] Due to distinct characteristics of glycopeptides and other analytes, different stationary phases have been developed in order to optimize the separation and enrichment.[250, 263-264] Although this method also suits to a separation techniques as described in previous sections, HILIC for solid-phase extraction can be used as an additional separation dimension prior LC-MS, in order to separate unmodified peptides from glycopeptides.[265-266] Moreover, Zauner and collaborators were able to separate a mixture of glycopeptides from fetuin and asialofetuin generated after treatment with Proteinase K between *O*- and *N*-glycopeptides using HILIC[206] even if the same separation for more complex samples was not reported yet.

Besides conventional HILIC, variations of this methodology enhance the separation power of the technique, especially those which include charges in the stationary phase. In the zwitterionic-HILIC (ZIC-HILIC), a stationary phase modified with sulfobetaine presents at neutral pH both charges, positive and negative, which may lead to charge-charge and charge-dipole interactions. These electrostatic interactions contribute in particular to a better separation of sialylated glycopeptides. Nevertheless, charged unmodified peptides might interfere during the separation. Modifications in the mobile phase might also influence in the separation capacity, as the change of the acid content FA for TFA present in the mobile phase enhance the separation between glycopeptides and non-glycosylated peptides.[267] In this same work, a comparison within ZIC-HILIC

and strong cation exchange-reverse phase chromatography showed that the first one is more efficient for glycopeptide enrichment. Wohlgemuth *et. al.* evaluated quantitatively *N*-glycopeptide enrichment by 3 different HILIC stationary phases (cellulose, ZIC-HILIC and amide), TiO_2 and hydrazine beads. Their experiments demonstrate that ZIC-HILIC and amide stationary phase (TSKgel Amide-80) presented the best performance regarding general *N*-glycopeptide enrichment, while TiO_2 was efficient only for sialylated *N*-glycopeptides by the analysis of labelled released *N*-glycans.[268] Takegawa and collaborators demonstrated that *N*-glycopeptides of α_1-acid glycoprotein can be separated from non-glycosylated. Furthermore even larger isomeric tri- and tetraantennary glycopeptides can be also separated by ZIC-HILIC.[269]

Combination of HILIC with ion-exchange chromatography was proposed by Alpert in 2008.[270] Electrostatic repulsion hydrophilic interaction chromatography (ERLIC) proposes analyte separation according to the charge and polarity simultaneously. Although the separation model is not fully understood, the hypothesis indicates that hydrophilic interactions dominates over electrostatic repulsions under low water content which provides better species separations. Lewandrowski *et. al.* used ERLIC for enrichment of *N*-glycopeptides from human platelet, identifying in total 125 glycosylation sites on 66 proteins.[271] Some years later, Hao *et. al.* optimized ERLIC methodology for separation of glyco- and phosphopeptides in a shotgun proteomic study of rat kidney tissue.[272]

There is an effort in creating different materials for glycopeptide enrichment, as zwitterionic nanopolymers[273], kapok fiber[274] and bioinspired saccharide-saccharide interaction polymer[275], recently published, however they should be further evaluated for high throughput protocols.

5.1.4.3- Methods based on metabolic labelling

The use of glycan biosynthetic pathway for the incorporation of unnatural substrates *in vivo* was demonstrated by Kayser *et. al.* in the beginning of the 90's.[276] Feeding rats with chemically modified *N*-acylmannosamine and *N*-acylglucosamine, proteins modified with *N*-acyl sialic acid were detected in most organs. This and following studies suggested the possibility of the introduction of biorthogonal groups in glycoproteins with the intention of comprehend and visualize cell and organs. In 2011, Boyce and Bertozzi elegantly define Bioorthogonal Chemistry as "chemical transformations among abiotic reactants that can proceed in living systems without interfering with, or interference from, the surrounding biological milieu". Those abiotic groups can be completely absent from living organisms, such as azide and alkyne, or groups that are absent in certain cell compartment, like aldehydes in cell surfaces.[277]

Labelling *in vivo* and *ex vivo* were performed after administration of azido-ManNAc via Staudinger ligation and analysis of mouse splenocytes, indicating the feasibility of biorthogonal chemical reactions in a living system without interference in the organism.[278] Copper(I)-catalyzed azide-alkyne cycloaddition, or mostly known as "click chemistry", was also used for labelling of Jurkat and Hep3B cell lines using both, azido- and alkyl-, modified fucose.[279] In order to overcome the copper cytotoxicity Bertozzi group developed a "copper-free click chemistry" version employing cyclooctynes[280] and applied this approach in the imaging of glycan in live zebrafish embryo for spatiotemporal analysis of glycan expression and trafficking during embryo development.[281]

Metabolic incorporation of azido-GlcNAc in *Helicobacter pylori* was used by Champasa *et. al.* to label bacterial glycoproteins and phosphine-FLAG probe was used, after cell lysate, via Staudinger ligation to enrich those proteins together with anti-FLAG antibody and affinity chromatography. After β-elimination and LC-MS/MS analysis, in total, 125 glycoproteins were identified including some proteins linked to bacterial pathogenesis.[282] Chen's group developed a liposome able to deliver unnatural azido-sialic acid to specific targeted cells.[283] They used this approach to perform tissue-selective metabolic labeling *in vivo* in mice with melanoma. Alkyne biotin tag was used for further glycoprotein labeling by classical "click chemistry" reaction followed by streptavidin-beads enrichment and shotgun proteomics analysis in the isolated tumor tissue. A total of 204, 180 and 156 *N*-glycoproteins were identified in the tumor-day 14, 18 and 21 respectively.[284]

Despite of notable applications, metabolic labeling face some limitations as unnatural monosaccharides might perturb the cell system by non-recognition with the same affinity by the natural binding partners; and epimerase activity can interconvert monosaccharides, reducing the availability of the desired glycan for incorporation.

5.1.4.4- Method based on chemoenzymatic tagging

When metabolic incorporation cannot be carried out, glycans can be further functionalized with enzymes that bind a modified glycan with a bioorthogonal function that can be afterwards enriched. Griffin *et. al.* develop a methodology for O-GlcNAc mapping with a partical cleavable linker.[285] An azido-GalNAc unit was transferred to the GlcNAc using UDP-azido-GalNAc and the specific glycosyltransferase Y289L GalT.[286] Copper-catalyzed click-chemistry was used to attach cleavable biotin tag for enrichment with streptavidin beads, which were cleaved under hydrazine treatment prior LC-MS/MS analysis. Proteins α-crystallin and O-GlcNAc transferase were used to validate the method. Recently, Wang and co-workers used chemoenzymatic approach to label sialyl-T with a biotin-sialic acid to detect this low abundant epitope in different living cell lines.[287] Regarding chemoenzymatic tagging, the complication relays on identification of enzymes that accept unnatural monosaccharides as subtracts and the preparation of those modified glycan derivatives, as UDP-glycans.

5.1.4.5- Methods based on chemical properties

Exploration of chemical properties of glycans, for instance the presence of *cis*-diols, for formation of reversible adducts or conversion to aldehydes for covalent binding to solid materials is desired due to its simplicity, effectiveness and broader scope, in particular when compared to lectin approach.

Boronic acid when reacted with *cis*-diols generates stable cyclic esters under high pH and the reaction can be reversed under low pH (Figure 22).[288] In this way, boronic acid can be used for enrichment of glycopeptides containing Man, Gal, Fuc and Neu5Ac which comprehends a larger range of both *N*- and *O*-glycopeptides. Xu *et. al.* used a phenylboronic acid resin to enrich glycoproteins from saliva for a proteomic study.[289] A column containing boronic acid gel was used by

Krisp *et. al.* to enrich glycopeptides from human wound fluids in which 104 glycoproteins in total were identified.[290] Chen *et. al.* compared the enrichment by lectins (Con A and WGA) and boronic acid magnetic beads of *N*-glycopeptides from yeast cell lysate and observed that boronic acid enrichment was slightly better than the lectins. Consequently they obtained 816 *N*-glycosylation sites in 332 yeast proteins combining boronic acid enrichment with LC-MS/MS analysis.[291] Recently, Xiao *et. al.* designed a strategy that enhances the interaction between boronic acid and glycopeptides.[292] Beads containing dendrimer modified with benzoboroxole was used to create multiples interaction points between glycopeptide and boro motif. The method was applied in yeast, mouse brain tissue and human cell lines. However the complexity of biological samples is reported to interfere in the binding matrix which might compromise the enrichment efficiency.[250]

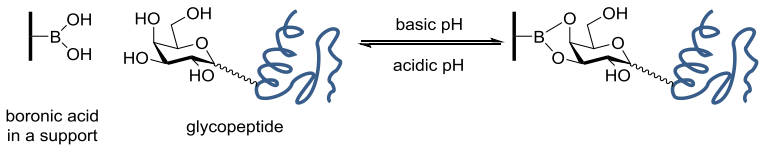


Figure 22. Cyclic boronate ester formation between boronic acid and *cis*-diols.

Titanium dioxide beads are commonly used for enrichment of phosphopeptides in proteomic protocols.[293] However Larsen *et. al.* proposed their use for enrichment of sialylated peptides.[294] They observed that sialic acid, due to its negative charge and hydroxyl groups, can also bind to the titanium dioxide beyond phosphate groups. So the enrichment of sialylated glycopeptides from a mixture of standard glycoproteins, human plasma and saliva was demonstrated after sample treatment with phosphatases to avoid binding competition. During this study, 192 glycosylation sites were identified in human plasma and 97 in saliva. Wohlgemuth *et. al.* also observed that titanium dioxide beads were efficient for enrichment of sialylated *N*-glycopeptides from a glycoprotein mixture in a comparative study with lectins and hydrazine beads.[268] On the other hand, problem for elution of glycopeptides containing polysialylation and oligosaccharides is reported.[295]

Inspired by classical oxidation of *cis*-diols using periodate[296-297], Aebersold and co-workers implemented a new methodology for glycopeptide enrichment.[298] Sodium periodate (15mM, 1h, at room temperature) was used for glycan oxidation in glycoproteins, afterwards the aldehyde reacted with a hydrazine group attached to a solid support. Unbound proteins were washed out prior trypsin digestion. Unbound digested peptide fragments were washed out before the release of the peptide part of the covalently linked N-glycopeptides with PNGase F, followed by LC-MS/MS of the released peptides. They identified 145 *N*-glycopeptides in 57 glycoproteins in human serum and 104 *N*-glycopeptides in 64 glycoproteins in the surface of prostate epithelial cancer cell line. Later Zhang and collaborators published an optimized protocol for enrichment of *N*-glycopeptides, where the sample was digested with trypsin or other protease before the oxidation with periodate and capture (Figure 23).[299] This procedure is claimed to increase the specificity once during proteolytic cleavage the glycosylation modification becomes more accessible to the beads upgrading the enrichment efficiency[300], although protein-level enrichment indicates to achieve higher identifications[301]. Theoretically, both *N*- and *O*-glycopeptides are enriched by this methodology, which gives information about glycosylation sites of the peptides, but not the glycan structure. However, due to

the lack of *O*-glycosidase as efficient as PNGase F, only *N*-glycopeptides are release end evaluated with this strategy. The chemical method for release employing hydroxylamine was used by Klement *et. al.* for the study of *O*-GlcNAcylation in *Drosophila*.[302] For this study the oxidation conditions were changed with higher temperature and oxidant concentration, and longer reaction time (20mM $NaIO_4$, 6h, 37°C) in order to oxidize the *trans*-diols present in the GlcNAc. Although side reactions were not explored, harsher conditions make more susceptible the oxidation in the peptide moiety, including cysteines, methionines and *N*-terminal serine and threonine which might lead to false-positive identification.[296, 303] Because of the simplicity of the procedure, the methodology was widely applied in diverse scope of biological samples[250, 299] with reproducibility required for large proteomic studies[301].

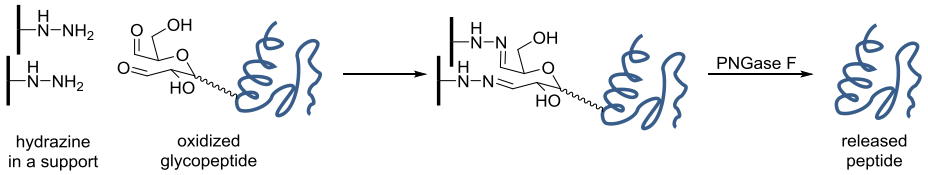


Figure 23. Capture of oxidized glycopeptide by hydrazine beads and enzymatic peptide release.

Variant of this methodology in order to overcome its weak points was recently published. Zou and collaborators proposed dimethylation of peptide *N*-terminal to avoid side reactions in the peptide backbone[304]. Moreover hydroxylamine assisted PNGase F release improved the deglycosylation step by bringing the glycopeptides to a homogeneous solution for the glycosylation cleavage[305]. But the main disadvantage of this strategy refers to *N*-glycopeptides enrichment exclusively. Nilsson *et. al.* partially addressed to this issue by optimizing the periodate oxidation selectively to the glycerol moiety of sialic acid.[306] The use of cold periodic acid for a short time (2mM, 0°C, 10 min) and acidic conditions (formic acid 0,1M, 80°C, 1h) for releasing after capture with hydrazine allows the enrichment of both, *N*- and *O*-glycopeptides. Here the labile sialic acid glycosidic bonds are cleaved, remaining the major part of the glycosylation still attached to the peptide backbone. This approach was applied to study cerebrospinal fluid and 36 *N*-glycosylation and 44 *O*-glycosylation sites were identified. The same group applied this methodology later to study glycosylation modification in human urine samples.[211] Nevertheless this methodology targets exclusively sialylated glycopeptides and during hydrolysis, information regarding sialylation and partially fucosylation are lost. Besides that peptides bond after asparagine residues can be occasionally hydrolyzed. Similar approach, named reverse glycoblotting, was published by Kurogoshi *et. al.* where sialylated glycopeptides were mildly oxidized (1mM $NaIO_4$, 0°C, 15 min), enriched using a polymer functionalized with hydroxylamine and asialoglycopeptides were released after acidic treatment (3% TFA, 100°C, 1h).[307] In following studies they changed the releasing strategy for a cold acid treatment (1M HCl, 0°C) to avoid further hydrolysis problems and labelled the released glycopeptides with 2-aminopyridin using reductive amination conditions to analyze mouse serum samples.[308]

Although chemical methods for introducing an aldehyde are the most explored, there is also enzymatic methodology for oxidizing specific glycan motifs. Galactose oxidase is an enzyme that

converts the hydroxyl group at C6 position of terminal galactose into aldehyde.[309-310] Taga *et. al.* studied *O*-glycosylation in collagen using galactose oxidase to generate the aldehyde, hydrazine beads for enrichment and acidic treatment for glycopeptide release.[311] The strategy was used not only for identification of glycosylation site, but also for discrimination between bovine collagen type I (Gal) and type II (Glc-Gal) on hydroxylysine. Later on Paulson and co-workers used both galactose oxidase and periodate oxidation to explore the glycosylation on the surface of B-cell lines.[312] After oxidation, glycoproteins were labeled with aminooxy-biotin for enrichment with streptavidin beads and *N*-glycopeptides were release using PNGase F prior LC-MS/MS analysis. From a total of 108 glycoproteins identified, 48 were identified using both oxidation methods, 40 glycoproteins were identified exclusively by periodate oxidation and remaining 20 glycoproteins identified were exclusively obtained using enzymatic oxidation. Recently, Zheng *et. al.* used enzymatic oxidation to enrich T_n antigen glycoproteins.[313] Captured glycopeptides were released from hydrazine beads using methoxylamine before LC-MS/MS analysis. Using this approach, 96 glycoproteins were identified in Jurkat cells.

5.2- Aim

The identification and characterization of glycans and glycoproteins from biological samples such as tissues, cell culture, urine, blood and CSF is still a major challenge compared to characterization work in other omics areas. The complexity and microheterogenity of glycans that modify proteins is one major reason for these challenges. There is a need of efficient glycopeptide enrichment methods in order to enable characterization of the diverse glycan modifications of proteins. A great number of different enrichment strategies are previously reported as discussed above. However, the currently available methodologies still have problems to cover the heterogeneity of glycosylation and the enrichment often focus on a specific group of modifications. Some glycopeptide enrichment methods also includes complete or partial release of glycans, resulting in loss of glycan information upon peptide/glycopeptide characterization. Reliable methodology for glycopeptide quantification is another major challenge in glycoproteomic analysis.

Lectin weak affinity chromatography has over the years shown to be a very powerful methodology for glycopeptide enrichment, in particular if the focus is to analyze a specific group of glycan modifications/terminal epitopes. For instance, impressive work with analytical quantitative and qualitative reproducibility for Con A is reported[314]. However, certain limitations remain with this methodology, interactions between lectins and glycan motives are often relatively weak and are naturally often relying on multivalent interactions. The presentation of the glycans on different peptide backbones or the glycan valency on the peptide may interfere and have an impact on the efficacy of enrichment. Due to the weak lectin binding some glycopeptides may not even be enriched. Human IgG, which typically presents biantennary *N*-glycan structures, is one classical example of how glycan presentation in the protein structure influence recognition/enrichment by Con A.[315-316] Most of the IgG glycosylation are located in the Fc domain and that means low solvent accessible and consequently no lectin recognition. Low glycan specificity is sometimes also a problem using lectin affinity enrichment.[256]

HILIC have shown to be powerful for *N*-glycopeptide analysis[317], but this approach requires a highly organic mobile phase (more than 70% MeCN) during loading of the peptide mix on the column, for this reason precipitation of glycopeptides, especially for large and heavily glycosylated structures, might occur in particular during the beginning of gradient leading to sample material losses. In addition some column materials may lose their binding capacity over time[318] and significant variation of retention times have been observed for different systems.

Methods based on metabolic labelling and chemoenzymatic tagging have shown to be very efficient for glycan and glycoprotein analysis, but both methods make use of unnatural monosaccharides, which might interfere with the biological system. For instance the natural enzymes are not always efficiently accepting unnatural monosaccharide donors as subtrates. Further, both methodologies are limited to study cell cultures or simple systems and not suitable for diagnostics and analysis of patient samples.

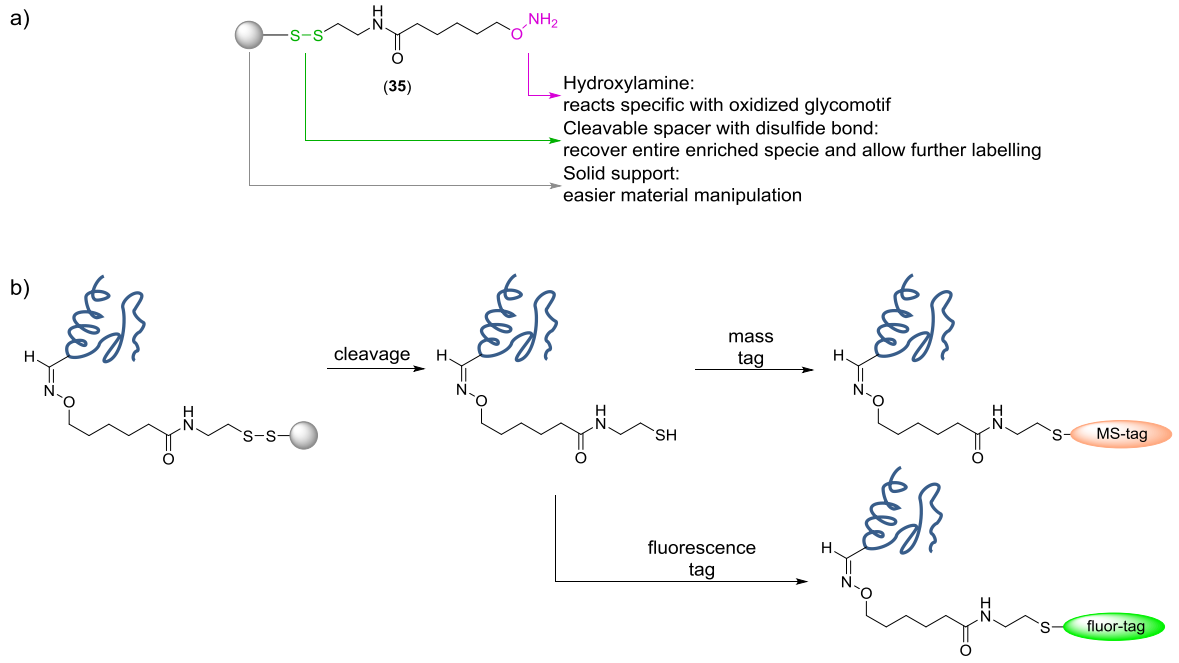
Methods using hydrazine beads are one of the most common for glycopeptide enrichment.[298] Vicinal *cis*-diols are oxidized and the generated aldehydes react with hydrazine forming a hydrazone bond between the solid support and the glycopeptide/glycoprotein. The covalent bond allows harsh washing conditions improving the enrichment procedure. For later peptide release from the solid-support, this enrichment methodology is mainly used in combination with PNGase F. This approach is very useful for *N*-glycosylation site mapping, but the lack of *O*-glycosidades restricts the methodology and additionaly information about the glycan structures that modify the sites is lost. Moreover conversion of asparagine into aspartic acid is often observed in shotgun proteomic procedures.[190] For this reason false positive identification might be obtained if only deamidation is observed in the peptide sequence.

To circumvent the main drawbacks of the hydrazine-bead enrichment methodology, Nilsson and collaborators used mild oxidation conditions and acidic conditions during peptide bead release to cleave the most labile sialic acid glycosidic bonds.[306] In this way the enrichment of *N*- and *O*-sialylated glycopeptides retain a major part of the glycan structure attached to the peptide backbone. However, information of the number of sialic acids and linkage connections on the glycopeptides are lost. Further, fucose residues may partially be cleaved during the peptide bead release conditions and non-sialylated glycopeptides will not be enriched, as also is the case by using titanium dioxide beads.

Changes of terminal glycan epitopes, involving sialylation and fucosylation, are known to be involved many biological processes as discussed in the former section. The analysis of intact glycopeptides containg the entire glycan structure and simultaneously give the full information regarding glycosylation sites, micro and macroheterogeneity, is critical to understand the role of glycosylation in diverse diseases and to identify new biomarkers. For this reason the demands are high for development of methodology that can enrich both *O*- and *N*-glycospecies in a quantitative and qualitative manner to be incorporated in shotgun proteomic experiment procedures.

This project aimed to improve the previously reported enrichment methodology using hydrazine beads by keeping the complete glycopeptide structures intact during bead release (Scheme 13a). Bead enrichment based on glycan chemistry properties is preferable in this case. This approach involves the formation of a covalent bond between the beads and glycan motif, which is important for proper washing steps leading to higher enrichment yields. Oxime formation is chosen

instead of the traditional hydrazone because oxime bonds are reported to be more stable than hydrazones and imines.[319] Moreover the proposed bead enrichment is in this case independent of the peptide backbone and can be applied on molecules with different complexity by performing enrichment on the glycopeptide and glycoprotein level. In addition, beads with cleavable linkers are able to recover the entire glycopeptide/glycoprotein structure independent of the glycosylation classes: *N*- or *O*-glycosylated, sialylated or galactosylated. Furthermore the smart linker design enables, besides enrichment, *in situ* labeling with tags compatible with standard proteomic procedures for quantitative detection of desirable enriched glycopeptides (Scheme 13b).

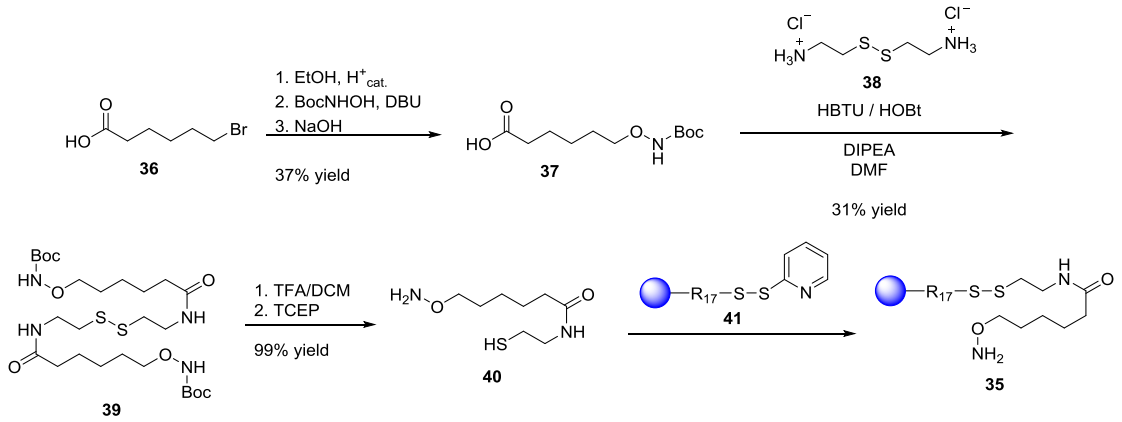


Scheme 13. a) Design of the cleavable beads (**35**) for glycopeptide/glycoprotein enrichment. b) Molecules containing oxidized glycan motifs are chemically attached to the solid support. The release is independent to the sample due to disulfide group present in the spacer. After cleavage the thiol can be capped with iodoacetamide or labelled with fluorescence or mass tags for identification and/or quantification purpose.

5.3- Linker synthesis and magnetic bead modification

The cleavable linker was synthesized in a few steps before attachment to thiol-activated beads (Scheme 14). The commercial bromo acid **36** was protected as an ethyl ester, reacted with Boc-hydroxylamine and converted into the hydroxylamine **37** after saponification within a yield of 37% for 3 steps. Intermediate **37** was coupled to cystamine (**38**) using HBTU/HOBt leading to the key dimer **39** in 31% yield. The Boc-group was cleaved using a 1:1 TFA/DCM mixture and the disulfide bond was reduced using TCEP leading to the desired spacer **40** in quantitative yields in both steps.

Finally spacer **40** was coupled to commercial magnetic beads from Bioclone® (**41**) leading to modified beads **35**. Magnetic beads were chosen due to the facility of the enrichment in the following steps.



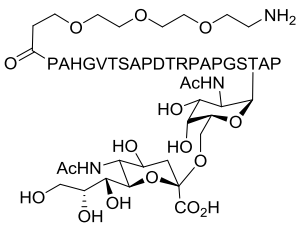
Scheme 14. Synthesis of the smart linker and modification of magnetic beads surface in order to obtain the cleavable hydroxylamine beads for glycan enrichment.

The reaction between **40** and the 2-pyridinesulfenyl thiol activated magnetic beads (**41**) was followed by analytical HPLC observing the release of 2-mercaptopyridine. This compound presents high absorbance under 343nm which facilitates to follow its release. The release of 2-mercaptopyridine by the coupling reaction was compared with its release under TCEP 0,1M solution. It was observed that the reaction achieved equilibrium after 24h incubation time.

5.4- Evaluation of glycopeptide enrichment

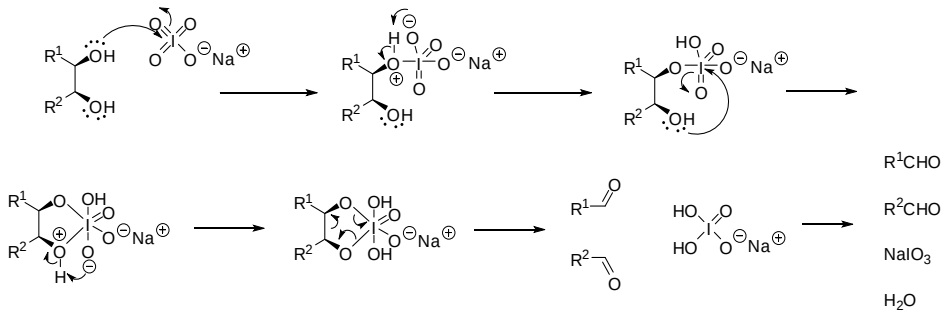
5.4.1- Sialylated peptide enrichment using magnetic beads

At first our modified beads were evaluated using model glycopeptide synthesized using methodology developed by our group[84]: a 20-amino acid MUC1 peptide type (**42**) with containing NeuAc-GalNAc motif in the Thr_{18}. For instance, the sialic acid of the synthetic peptide was oxidized with sodium periodate solution.



42

Sodium periodate is a well-known reactant selective for the oxidation of vicinal *cis*-diols. One of the hydroxyl groups attacks the periodate anion. After a proton transfer, only a *cis* hydroxyl group has access to attack the periodate in order to form a cyclic iodate ester intermediate. At this stage, the cyclic intermediate breaks down, reducing the iodine(VII) to iodine(V), cleaving a C-C σ-bond and forming two C-O π-bond (Scheme 15).[320]



Scheme 15. General oxidation mechanism of vicinal *cis*-diols by sodium periodate.

Under proper conditions, this reaction can be used in order to oxidize different monosaccharides as mannose, galactose and sialic acid.[296-297] The last one is a good target for this approach because it presents 3 easily accessible vicinal *cis* hydroxyl groups on the glyceryl portion (C7, C8 and C9). For this reason mild oxidation conditions, using diluted solution, low temperature and short reaction time, were reported[306-307] in order to oxidize exclusively the glycerol motif in the sialic acid. Therefore an experiment for evaluating and analyzing all obtained species during the oxidation was performed. The oxidation was stopped in different time points and the proportion of the species generated was analyzed via mass spectrometry. Figure 24 summarizes all the species found in each time point for the oxidation of a synthetic glycopeptide, reacted with 2mM sodium periodate solution in water at 4°C. This figure indicates that there is a mixture of aldehydes and respective ketals from sialic acid C7 and C8 during the first 15 minutes. After 30 minutes reaction, only the aldehyde and ketal derivate from C7 is present. For this reason, 30 minutes reaction time was used in the following experiments.

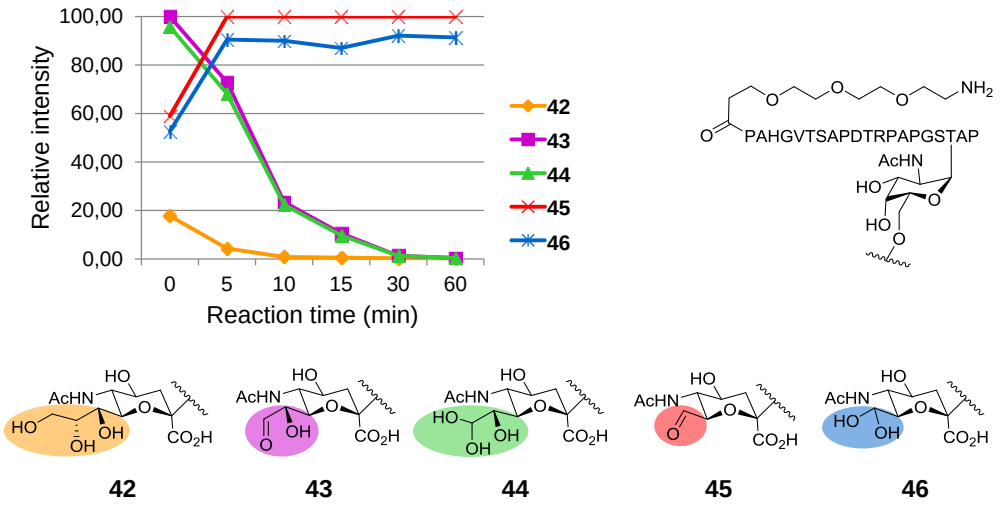
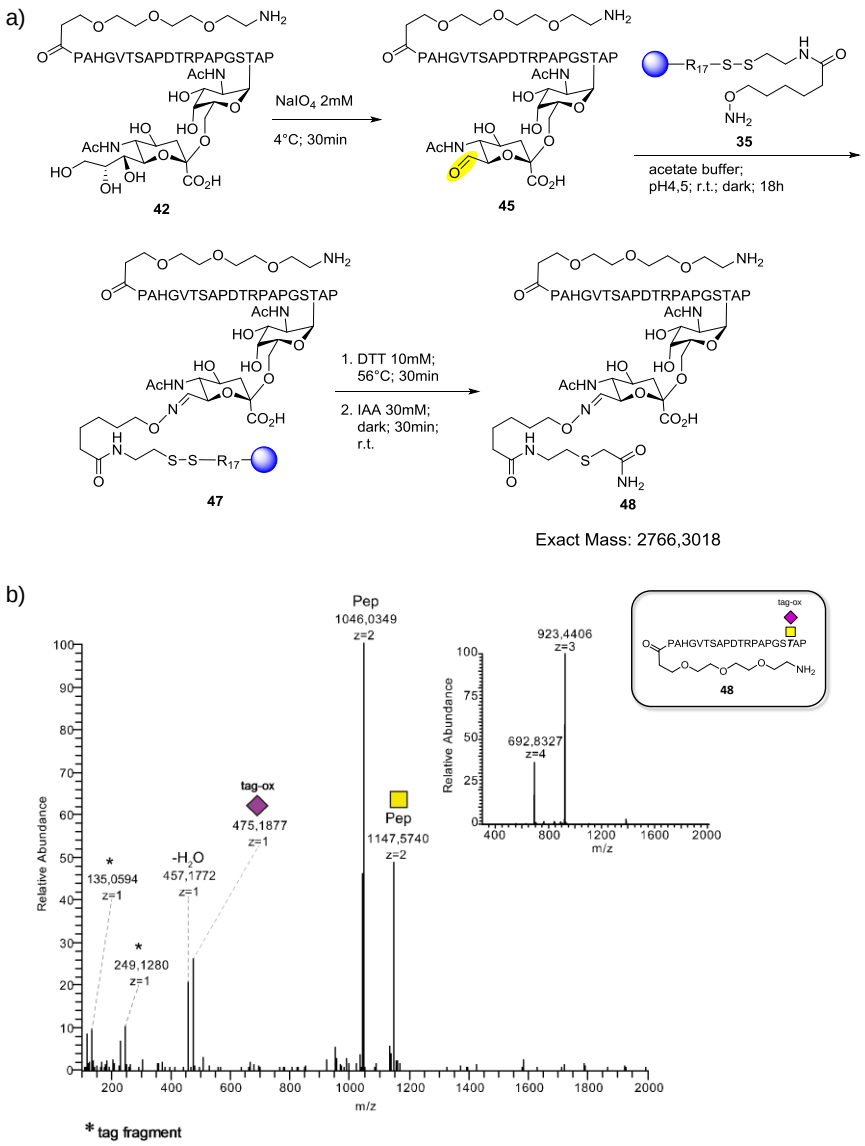


Figure 24. Species distribution of Neu5Ac-Gal-peptide oxidized with sodium periodate (2mM) at 4°C in different time points.

The oxidized peptide **45** was captured on our beads through incubation in acetate buffer overnight with a ratio peptide/beads 1:10 in molarity. After a washing step, the glycopeptide was released using dithiothreitol (10mM) and the free thiol was capped with iodoacetamide (30mM) (Scheme 16). The released sample was analyzed by nano-LC-MS/MS using an Orbitrap Velos Pro mass spectrometer. MS/MS experiment used HCD with normalized collision energy of 30%. The full MS scan showed the desired peptide **48** (*m/z* 923,44^{3+}). Fragment ions obtained by using HCD confirmed its structure: tag-Neu5Ac (*m/z* 475,19^{1+}), GalNAc-peptide (*m/z* 1 147,57^{2+}) and peptide (*m/z* 1 045,54^{2+}) were observed. Furthermore linker fragmentation was also observed (*m/z* 135,06^{1+} and 249,13^{1+}).



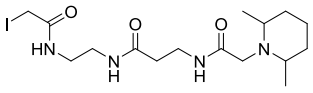
Scheme 16. a) Enrichment of a negative charged glycopeptide using hydroxylamine beads. b) MS/MS spectrum of the enriched Neu5Ac-GalNAc-peptide. Full MS is showed in the upper right corner.

A careful look in the masses obtained for linker fragmentation (or tag fragments) allowed the identification of the chemical structure of these fragments. The description of a more detailed assignment is found in appendix Table A 8.

Although the oxime bond is a more stable bond than hydrazone and imine[319], we observed that this reaction can be reversible depending on the time and storage conditions. The reversible reaction leading to the correspondent aldehyde and the hydroxylamine linker was observed in small proportions for the samples stored over longer periods of time and in acidic solutions (FA or TFA in water). This can be prevented by storage of dry samples or samples dissolved in water.

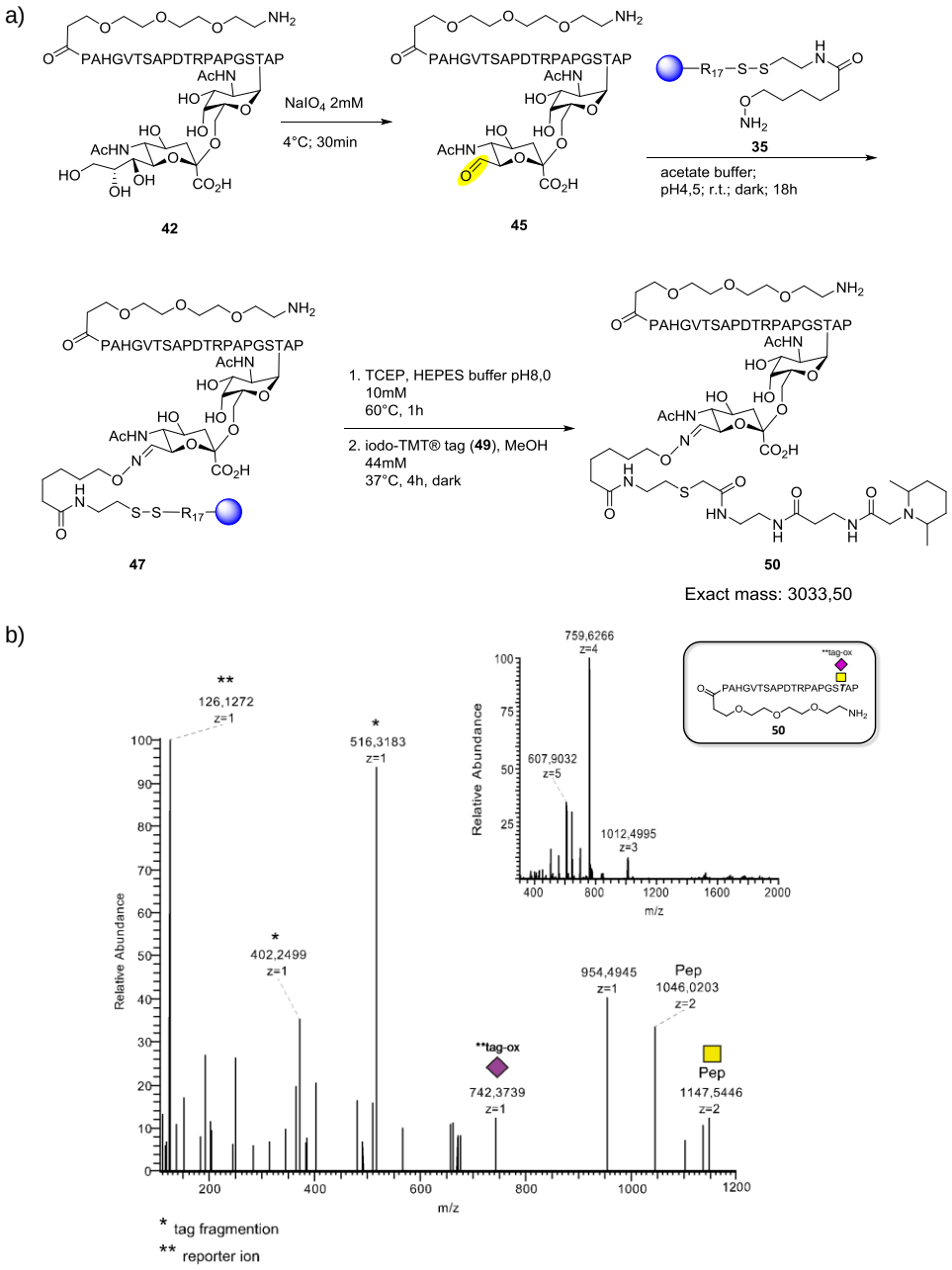
5.4.1.1- TMT® labeling of released glycopeptides

We also evaluated the use of modified hydroxylamine beads in combination with tandem mass tag (TMT) labeling. None of the previously described methods specific for enrichment of glycopeptides has been combined with mass tag quantification. The use of a glyco-TMT® is described[321], however the cleavage of the glycan from the peptide backbone is required, which makes this strategy possible only for N-glycans, but then the information about the protein modified and site occupied is lost. The use of mass tag labeling in proteomic approaches enables identification and multiplexed quantification of proteins in different samples at the same time. The most well-known and used reagent is the amine-specific iTRAQ. In addition to reagents mention before, specific tags for thiol groups are also commercially available. Iodo-TMT® tag (**49**) from Thermo was the labeling tag chosen for capping the free thiol group generated after cleavage the peptides from the magnetic beads solid support.



Iodo-TMT®
(**49**)

The Neu5Ac-GalNAc-peptide model was oxidized with $NaIO_4$ and enriched with hydroxylamine beads. After reductive peptide release, the obtained free thiol was capped with iodo-TMT®. After some capping problems due to the consumption of iodo-TMT® by dithiothreitol, we decided to change the reducing agent for the peptide cleavage from the beads. When the peptide release was compared using dithiothreitol and *tris*(2-carboxyethyl)phosphine, the peptide signal in the mass spectrometry measurement increased using the later reducing strategy, under the same reaction conditions. From this moment on, the protocol was optimized and the peptide release was made with *tris*(2-carboxyethyl)phosphine solution. Then, the model glycopeptide was labelled successfully with iodo-TMT® tag under the optimized protocol conditions (Scheme 17).



Scheme 17. a) Model glycopeptide enrichment and labelling with the quantification iodo-TMT® tag. b) MS/MS spectra of the enriched and labeled Neu5Ac-GalNAc-peptide. Full MS is showed in the upper right corner.

The tagged peptide was analyzed by nano-LC-MS/MS using an Orbitrap Velos Pro mass spectrometer. MS/MS experiment used HCD with normalized collision energy of 40%. The tagged glycopeptide was observed in the full MS scan (*m/z* 759,63^{4+}). The report ion from TMT-tag (*m/z* 126,13^{1+}) was observed in MS/MS analysis in combination with fragments of a TMT-linker species (*m/z* 402,25^{1+} and 516,32^{1+}), TMT-tag-Neu5Ac (*m/z* 742,37^{1+}), peptide (*m/z* 1 046,02^{2+}) and GalNAc-peptide (*m/z* 1 147,54^{2+}) which confirm the glycopeptide structure.

5.4.1.2- Enrichment quantification using magnetic beads

In order to evaluate the quantification of the peptide recover with a parallel methodology, the concentration of the peptide in the released solution was measured with amino acid analysis (AAA). In this procedure, the peptide is degraded under acidic conditions to amino acids which are labeled with a fluorophore and the composition is determined using HPLC analysis.[18, 322] Therefore only peptides traces were found in the solution after the release from the beads. For this reason the coupling reaction was evaluated using different beads concentrations 10, 25 and 50 molar equivalents of beads in relation to peptide. Monitoring the reaction using LC-MS indicated equilibrium after 18h.

Table 7. Peptide distribution in coupling and release solution according to different bead/peptide ratio.

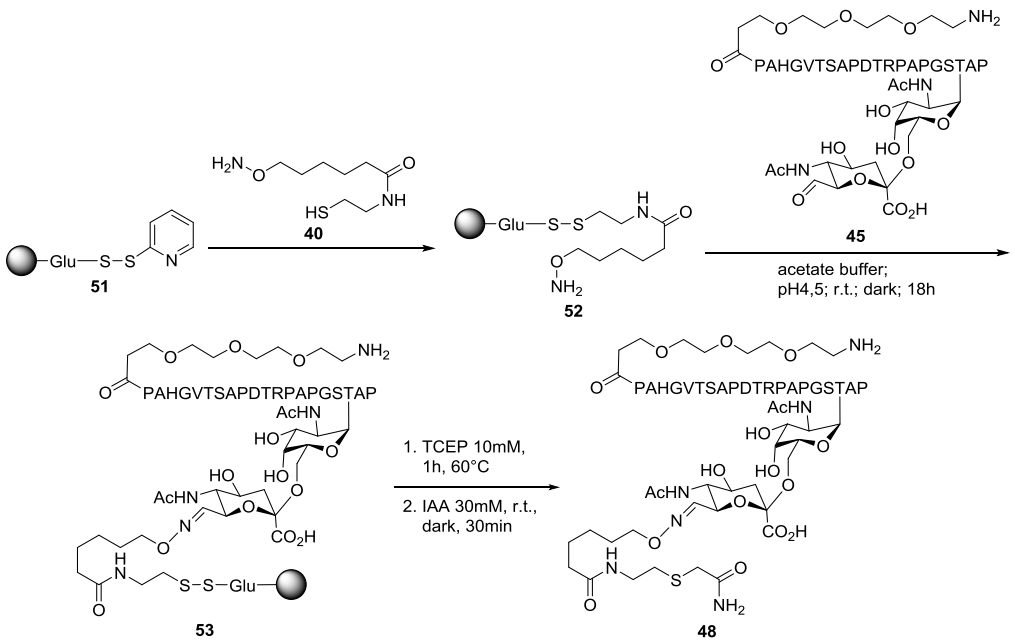
Experiment (bead/peptide ratio)		Concentration (mmol/L)	Yield (%)
Peptide standard		0,310	
Coupling solution[1]	10	0,100	32
	25	0,080	26
	50	0,060	19
Released solution[2]	10	0,003	1
	25	0,010	3
	50	0,020	6

[1] Remaining peptide concentration in the coupling solution after the reaction with the beads.
[2] Released peptide concentration after cleavage from the beads.

Herein Table 7 indicates that there is no significant changes in terms of peptide capture with an increase of beads amount. The enrichment efficiency was not satisfactory because the highest peptide concentration remained in the coupling solution and only peptide traces were found after release from the beads. Furthermore, PEG contamination originated from the magnetic beads was observed in the released samples.

5.4.2- Sialylated peptide enrichment with agarose beads

In order to overcome the problems observed with the magnetic beads, another solid support material was used: Activated Thiol Sepharose® 4B (**51**), which contains activated cysteines. The solid support used was agarose beads with amodified surface containing glutathione spacer arms. Its cysteines were activated with 2-mercaptopyridine. These beads were then functionalized with our linker **40** under the same conditions as described for the magnetic beads. Enrichment with the oxidized Neu5Ac-glycopeptide model was made as a proof of concept. Coupling and release were performed according to the conditions previously described.



Scheme 18. Agarose beads coating with the smart linker (**40**) and enrichment of synthetic glycopeptide for the new beads evaluation.

The sample obtained after the enrichment was cleaner and in higher concentration as those samples obtained previously using magnetic beads. To demonstrate the difference between both materials, the model peptide was enriched using both beads in a proportion of 1:10 peptide/beads in molarity. Coupling as well as release solutions were monitored using HPLC. Figure 25 contains the chromatograms of coupling (line 1) and releasing (line 2) solutions of peptide reaction and release respectively from magnetic beads (blue lines) and agarose beads (gray lines). For magnetic beads, the glycopeptide remained in the coupling solution. In contrast, for agarose beads most of the peptide was as desired found in the bead release solution. LC-MS analysis confirmed the peptide structure in both cases.

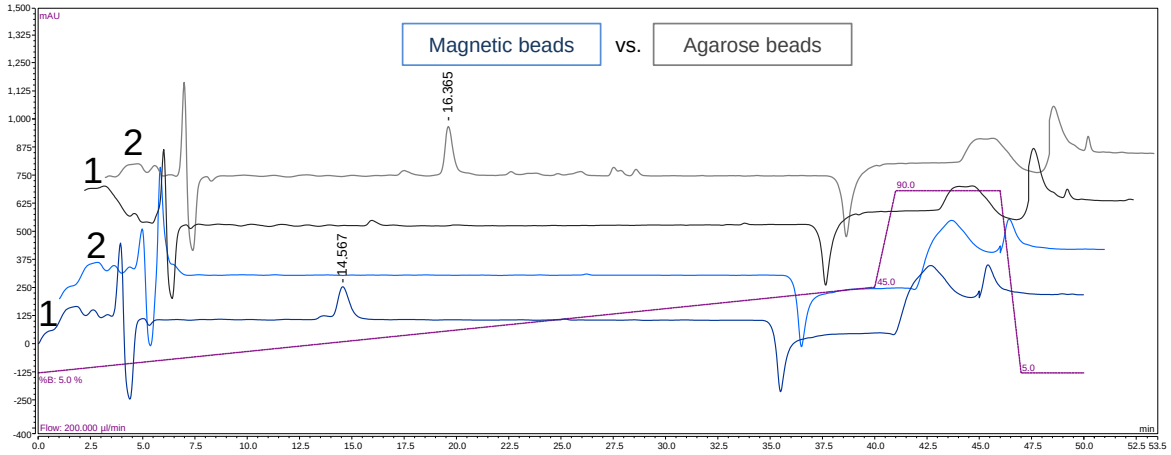


Figure 25. Chromatograms comparing the coupling (line 1) and release (line 2) solutions from the enrichment of Neu5Ac terminal glycopeptide with magnetic (blue lines) and agarose (gray lines) beads coated with hydroxylamine groups.

A significant difference in terms of peptide recovery depending on the beads material was observed; therefore the magnetic beads were substituted with agarose beads in all following enrichment experiments. The evaluation of the peptide recovery quantification was performed to validate the enrichment process. An experiment was made using 3 replicates that contain the same amount of peptide and a proportion of 1:10 peptide/beads in molarity to obtain the yield of the agarose bead enrichment.

Table 8. Peptide distribution in coupling and release solution using agarose beads for enrichment.

Sample	Replicate	Concentration (mmol/L)	Yield (%)
Peptide standard		0,080	
Coupling solution[1]	1	0,005	6
	2	n.d.[3]	n.d.[3]
	3	n.d.[3]	n.d.[3]
Released solution[2]	1	0,010	13
	2	0,020	25
	3	0,010	13

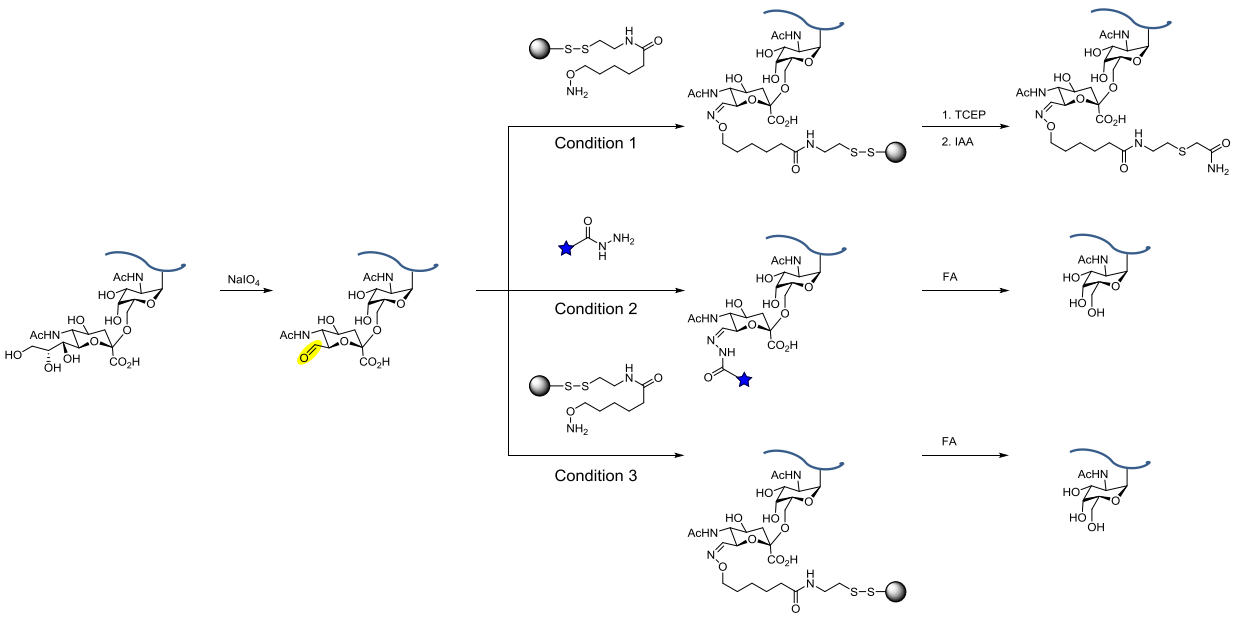
[1] Remaining peptide concentration in the coupling solution after the reaction with the beads.
[2] Released peptide concentration after cleavage from the beads.
[3] Concentration not detectable.

Although the amount of peptide recovered after the enrichment was small compared to the initial peptide amount, it is clear that most of the glycopeptide material is in the solution after the release from the beads. Further experiments were made in order to evaluate in which step from the enrichment the peptide is lost. We evaluated the washing and no relevant peptide amount was found. We also evaluated if the peptide was binding to the agarose instead of being enriched. Additional washing step with 4M urea solution, previous incubation of the beads with non-oxidized synthetic glycopeptide as well as addition of sucrose to the coupling buffer were evaluated. However none of these changes in the experimental procedure influenced the peptide recovery. For this

reason enrichment procedure was used without further modifications during the following experiments.

5.4.2.1- Comparison between bead enrichments and cleavage strategies using a peptide model

Next evaluation was the comparison of the method we developed with the bead enrichment method previously described by Nilsson et. al.[306] We compared the enrichment of the synthetic peptide using our beads containing a cleavable hydroxylamine linker with the commercial agarose beads coated with hydrazine groups (Scheme 19). The enrichment using hydrazine coated beads also required the oxidation of terminal sialic acid residues. However, for bead release formic acid 0,1M solution was used, which cleaves the labeled sialic acid glycosidic bond.[306] In this case the glycopeptides enriched with the commercial hydrazine beads were recovered without sialylation. In order to evaluate if the cleavage methodology had an impact on the amount of retrieved peptides, a third condition was added in which the peptide was enriched using our linker modified hydroxylamine beads and with the glycopeptide released using the acidic conditions (0,1M FA).



Scheme 19. Comparison of glycopeptide enrichment methodologies: cleavable hydroxylamine beads (condition 1) vs. commercial hydrazine beads (condition 2). A mixed strategy using hydroxylamine beads and acidic cleavage was also included to evaluate the release method (condition 3).

Table 9. Comparison between enrichment using agarose beads coated with hydroxylamine groups and commercial agarose beads coated with hydrazine groups. The peptide concentration here described refers to the concentration after peptide release from the beads.

Condition/Replicate		Concentration (mmol/L)	Avarage	Sd	Yield (%)
Peptide standard		0,085			
Condition 1: Hydroxylamine beads – TCEP release	1	0,025	0,027	0,0012	32
	2	0,027			
	3	0,027			
Condition 2: Hydrazine beads – acid release	1	0,039	0,039	0,0021	46
	2	0,042			
	3	0,038			
Condition 3: Hydroxylamine beads – acid release	1	0,024	0,024	0,0006	28
	2	0,025			
	3	0,024			

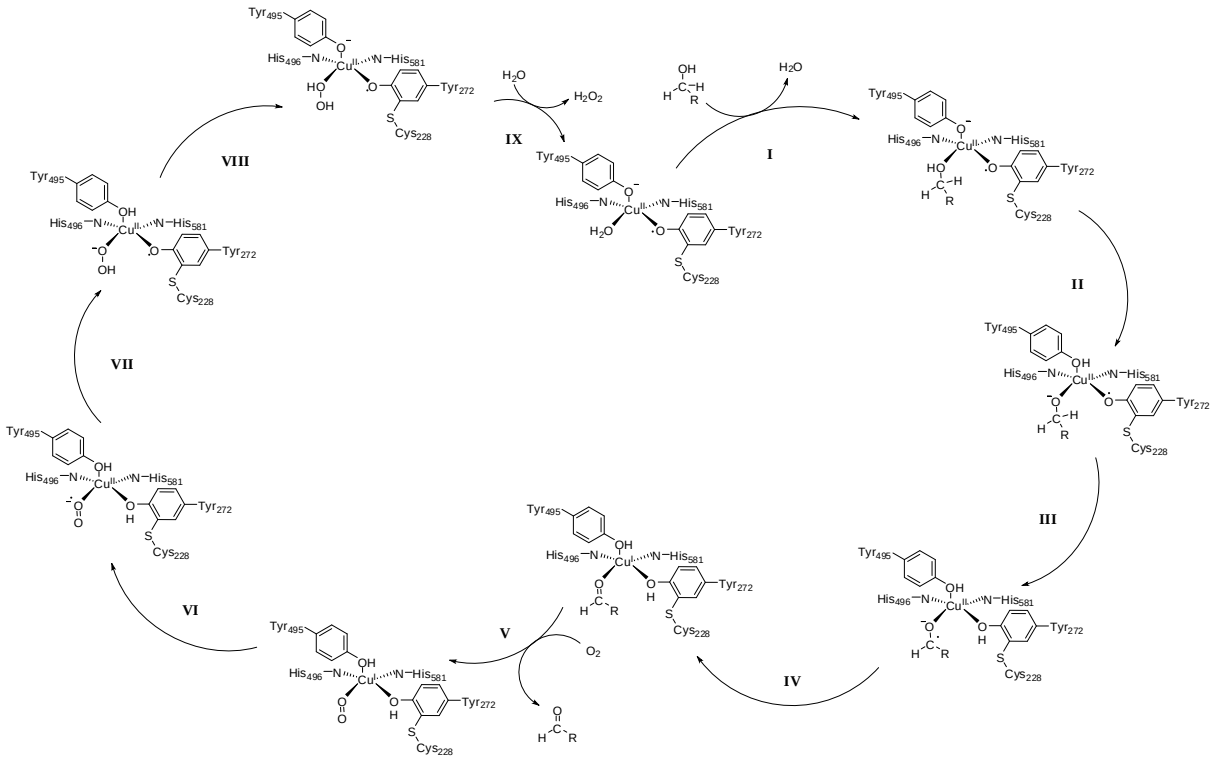
Observing the data in Table 9, the enrichment using hydrazine beads was slightly better in terms of glycopeptide recovery. Therefore the concentration of the reactive group of both beads were compared, the hydrazine beads have minimum 10µmol of activated groups/mL drained medium while the hydroxylamine beads have 1µmol of activated groups/mL drained medium. Besides the efficiency of the enrichment compared with the hydrazine beads, the influence of the release strategy on the enrichment was also evaluated. The model peptide was enriched with hydroxylamine beads and the peptide was released under acidic conditions. In this case the yield of the peptide recovery using acid was in average similar to the conditions using TCEP release rather to the enrichment with hydrazine beads. These results suggest that the bead release strategy is not the key point that influences the enrichment yield. After all, the enrichment of glycopeptides using hydroxylamine beads was satisfactory.

5.4.3- Galactose terminated peptide enrichment

Theoretically neutral (galactose terminated) glycopeptides can be enriched using hydrazine beads, as the sialylated glycopeptides. Terminal galactose is selectively oxidized in the C-6 position using galactose oxidase[309-310], generating the aldehyde necessary for the attachment to the beads. However an enrichment method for analysis of neutral glycopeptides faced problems in the release step for a long time because the galactose glycosidic bond is not as label as the sialic acid glycosidic bond. Last year methoxylamine was elegantly used to release T_n antigen glycopeptides from hydrazine beads.[313] Acid treatment[311], which might cleave glycosidic bonds, and PNGase F[312], specific for *N*-glycopeptides, were recently also used in combination with galactose oxide for determination of peptide glycosylation sites. Since the hydroxylamine beads with the disulfide cleavable linker permit release independent of the enriched glycopeptide structure, evaluation of Gal-terminated glycopeptide enrichment was also performed using our methodology.

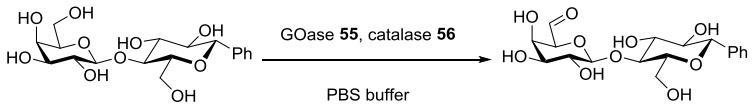
Galactose oxidase (GOase) is an enzyme from the alcohol oxidoreductase class. This enzyme employs copper in its catalytic core and oxygen molecular as cosubstrate.[323] The active pocket presents a Tyr_{495} anion that acts as a base and a radical $Tyr_{272}Cys_{228}$ specie that binds in the

coordination sphere of copper(II) (Scheme 20). There, the oxygen atom from the position C6 binds to the copper(II) as the first step of the catalytic cycle. Tyr_{495} subtracts the proton from the hydroxyl group. A hydrogen atom from the C6 is transferred to the radical $Tyr_{272}Cys_{228}$ and an electron from the carbohydrate reduces the copper generating the aldehyde and copper(I). The second step consists in the regeneration of the copper(II). As the aldehyde leaves the copper coordination, a molecule of O_2 takes place, an electron is transferred from the copper(I) leading copper oxidation. A hydrogen atom is then transferred from the $Tyr_{272}Cys_{228}$ specie, regenerating the radical active specie, and formation of a superoxide intermediate. This last one takes the proton from Tyr_{495}, and close the cycle of regeneration of the enzyme catalytic pocket.[310, 324]



Scheme 20. Catalytic cycle of the enzymatic oxidation mediated by galactose oxidase.

Phenyl lactoside was used for optimization of GOase reaction conditions. In the initial conditions described by Ramya and collaborators[312] the reaction was proceeding too slow (see Scheme 21, first entry). During the oxidation hydrogen peroxide is formed as byproduct, which might inhibit the enzyme in high concentrations. For this reason, addition of catalase was required in order to breakdown the peroxide formed during the process. Catalase is one enzyme presented in almost all aerobic systems. The enzyme uses iron as cofactor for converting hydrogen peroxide into water. After catalase addition the oxidation proceed faster comparing entry 1 and 2 (Scheme 21). However the condition on entry 3 was the best and guarantees the total conversion to the aldehyde after 18h reaction time.



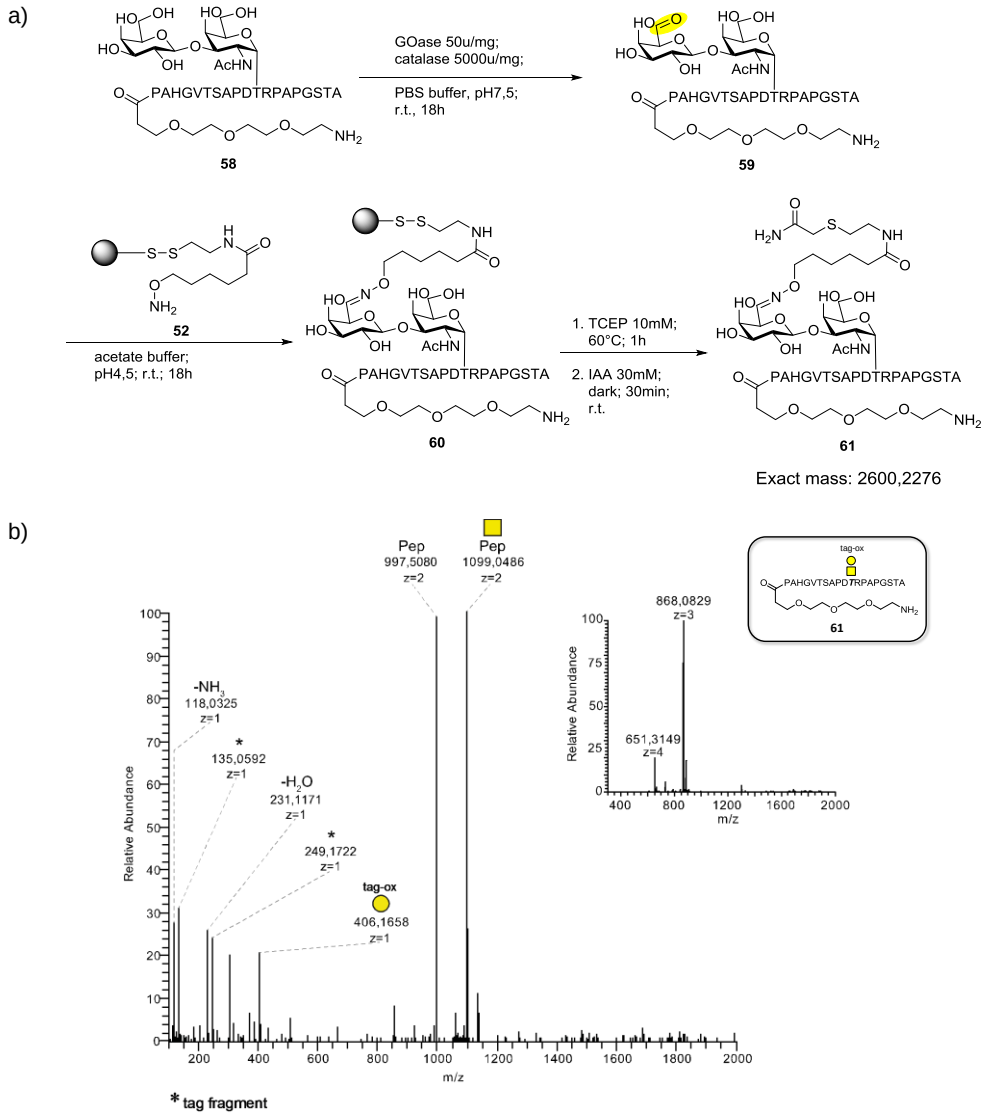
54 57

Entry	Condition **54** (mg) : **55** (units): **56** (units)	Ratio **54** : **57**
1	1 : 5: 0	87 : 13[a]
2	1 : 5 : 5000	26 : 74[b]
3	1 : 50 : 5000	0 : 100[b]

a) after 24h reaction.
b) after 18h reaction.

Scheme 21. Condition optimization of oxidation with galactose oxidase.

For further proof of concept, a neutral synthetic peptide[84], a MUC5b peptide type **58** containing 19 amino acids and the Thr_{11} decorated with Gal-GalNAc motif, was oxidized with galactose oxidase in presence of catalase. The peptide was conjugated and released from our beads (Scheme 22). The sample was analyzed by nano-LC-MS/MS using an Orbitrap Velos Pro mass spectrometer. MS/MS experiment used HCD with normalized collision energy of 30%. The desired glycopeptide (*m/z* 868,08^{3+}) was identified in the full MS scan. Fragments of the linker (*m/z* 135,061^{1+} and 249,18^{1+}), tag-Gal (*m/z* 406,17^{1+}), peptide sequence (*m/z* 997,51^{2+}) and GalNAc-peptide (*m/z* 1 099,05^{2+}) were observed by MS/MS analysis. The observed fragment ions confirmed the glycopeptide structure.



Scheme 22. a) Enrichment of a neutral charged glycopeptide using hydroxylamine beads. b) MS/MS spectra of the enriched Gal-GalNAc-peptide. Full MS is showed in the upper right corner.

Amino acid analysis was used to quantify the amount of peptide recovered after enrichment. Due to the lack of other available methodologies to enrich and release glycopeptides with terminal galactose when those experiments were executed. No direct comparison with other methodologies was performed, as in the sialylated case. The average yield obtained for the Neu5Ac terminal glycopeptide (Table 9) and for the Gal terminal glycopeptide (Table 10) were found to be of the same magnitude. The enrichment of both glycopeptide classes was considered satisfactory with a yield of 32% for sialylated glycopeptides (periodate oxidation) and 40% for non-sialylated (GOase oxidation).

Table 10. Quantification by amino acid analysis of neutral glycopeptide enriched with agarose beads coated with hydroxylamine groups.

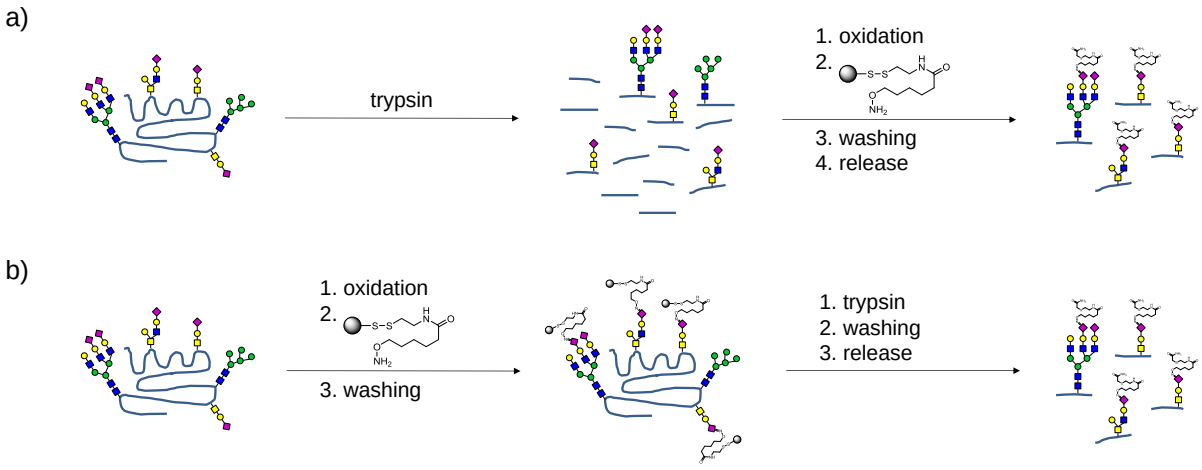
Condition/Replicate		Concentration (mmol/L)	Avarage	Sd	Yield (%)
Peptide standard		0,083			
Hydroxylamine beads – TCEP release	1	0,033	0,033	0,0021	40
	2	0,036			
	3	0,032			

5.5- Protein model evaluation

5.5.1- Enrichment of glycopeptides from fetuin

Fetuin was identified for the first time in 1944 as the most abundant protein of bovine fetal serum.[325] This protein presents 9 reported glycosylation sites: Asn_{99}, Asn_{156}, Asn_{176}, Ser_{271}, Thr_{280}, Ser_{282}, Ser_{296}, Thr_{334} and Ser_{341}.[216, 326] Alpha 2-HS glycoprotein is the fetuin human homolog. The protein is produced by adult liver and highly accumulated in the bones.[327] Due to its commercial availability, it is one of the glycoproteins largely used as model in method development in the glycosylation field. Furthermore fetuin is also available commercially in the asialo form. For these reasons, fetuin was selected as our model protein for testing our new enrichment method in a higher level of complexity than by using synthetic glycopeptides.

There are two different approaches in order to enrich glycosylated peptides from a protein and/or complex mixture. In the first approach the enrichment can be performed in the peptide level. Here the sample is treated with trypsin in order to obtain a peptide mixture prior the glycan motif oxidation and conjugation to the beads (Scheme 23a). On the other hand, the enrichment can be done also in the protein level. In this case the oxidation of the glycan motif and the chemical ligation to the beads occur before the trypsinization step (Scheme 23b). The later approach can be useful when the matrix is complex because all unbounded proteins are washed away before trypsinization.



Scheme 23. Enrichment workflow for peptide (a) and protein (b) levels.

For the following experiments, the enrichment in the peptide level was selected. The tryptic peptide mixture was then oxidized using sodium periodate and sialylated glycopeptides were immobilized through a covalent linkage to the beads. After the release the obtained peptides were analyzed by nano-LC-MS/MS using an Orbitrap Velos Pro mass spectrometer. MS/MS experiments used HCD with normalized collision energy of 30 and 40. For data analysis, the raw files of the sample oxidized were searched against the Uniprot sequence of bovine alpha 2-HS glycoprotein or fetuin-A (P12763) and bovine fetuin-B (Q58D62) using Byonic™ as search engine. Byonic™ has the advantage upon other search engines to have an user-friendly interface, the possibility to add special chemically modified monosaccharides (upon developers request) and being easy to edit of the glycan list. For our data analysis, Byonic™ included in their monosaccharide database 4 extra modified saccharides (2 Neu5Ac and 2 Gal, each in an oxidized version and with tag chemically attached). The respective chemical composition, monoisotopic mass and Byonic™ code for the each glycan assignment is described in Table 11 (Table A 8 contains their chemical structures). The software offers 3 defaults O-glycans list and 4 defaults *N*-glycans list. The Byonic™ default list with 70 human *O*-glycans was modified in order to obtain 27 different *O*-glycans with sialyl motives. And also the Byonic™ default list with 164 human *N*-glycans without sodium and multiple fucose was modified in order to obtain 68 *N*-glycans with sialyl motif (see appendix Table A 9 for O-glycan list and Table A 10 for N-glycan list). The search files were further manually filtered in order to obtain the number of unique glycopeptides identified. Herein a unique glycopeptide means that both peptide sequence and glycan motif are different for each hit. Glycosylation hits are these identified peptides that present "|Log Prob|" <1 and "Delta Mod" >10. According to the Byonic™ user guide, "|Log Prob|" is the absolute value of logarithm with base 10 of posterior error probability of each identification. Posterior error probability is a Byonic parameter that takes into consideration 10 features including score, delta, precursor error, digestion specificity. "Delta Mod" is a value that indicates the probability of the glycosylation site is rightly assigned. In total 28 unique glycopeptides that cover 2 *O*-glycosylation sites and 2 *N*-glycosylation sites of bovine fetuin-A and 2 *O*-glycosylation sites of bovine fetuin-B were identified (Table 12).

Table 11. Code, composition and monoisotopic mass of each monosaccharide unit for glycopeptide identification according Byonic™.

Glycan code	Composition	Monoisotopic Mass
Hex	$C_6H_{11}O_5$	163,0601
HexNAc	$C_8H_{14}NO_5$	204,0866
Fuc	$C_6H_{11}O_4$	147,0652
NeuAc	$C_{11}H_{18}NO_8$	292,1027
Sialic474	$C_{19}H_{31}N_4O_8S$	475,1857
Sialic229	$C_9H_{12}NO_6$	230,0659
Gal405	$C_{16}H_{28}N_3O_7S$	406,1642
Gal160	$C_6H_9O_5$	161,0444

Table 12. Glycopeptides identified from commercial fetuin sample oxidized with periodate and enriched with hydroxylamine beads. (Entries in black are *O*-glycopeptides and in blue are *N*-glycopeptides.)

Protein Name	Accession Number (UniprotKB)	Peptide Number	Peptide	Glycans	Glycosylation site	Observed m/z	z	Delta Mod	|Log Prob|
Fetuin-A	P12763	1	S.AAGPPVASVVVGPS[+839,310]VVAVPLPLHR.A	HexNAc(1)Hex(1)Sialic474(1)	S296	1043,555	3	43,1	3,78
Fetuin-A	P12763	2	S.VVVGPS[+839,310]VVAVPLPLHR.A	HexNAc(1)Hex(1)Sialic474(1)	S296	826,775	3	687,3	8,82
Fetuin-A	P12763	2	S.VVVGPS[+594,191]VVAVPLPLHR.A	HexNAc(1)Hex(1)Sialic229(1)	S296	745,070	3	742,3	9,82
Fetuin-A	P12763	3	S.VVVGPS[+1068,369]VVAVPLPLHR.A	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S296	903,129	3	697,3	8,82
Fetuin-A	P12763	4	K.TPIVGQPS[+839,310]IPGGPVR.L	HexNAc(1)Hex(1)Sialic474(1)	S341	772,053	3	123,8	11,26
Fetuin-A	P12763	5	R.KLC[+57,021]PDC[+57,021]PLLAPLN[+2325,819]DSR.V	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N156	1398,916	3	178,8	2,49
Fetuin-A	P12763	6	R.KLC[+57,021]PDC[+57,021]PLLAPLN[+2216,773]DSR.V	HexNAc(5)Hex(6)Sialic229(1)	N156	1022,181	4	239,6	1,96
Fetuin-A	P12763	7	R.KLC[+57,021]PDC[+57,021]PLLAPLN[+2445,832]DSR.V	HexNAc(5)Hex(6)Sialic229(2)	N156	1438,926	3	153,9	1,91
Fetuin-A	P12763	8	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2570,938]DSR.V	HexNAc(4)Hex(5)Sialic474(2)	N156	1437,928	3	126,6	2,08
Fetuin-A	P12763	8	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2325,819]DSR.V	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N156	1356,220	3	357,5	4,02
Fetuin-A	P12763	8	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2080,700]DSR.V	HexNAc(4)Hex(5)Sialic229(2)	N156	1274,513	3	461,4	5,06
Fetuin-A	P12763	9	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2096,760]DSR.V	HexNAc(4)Hex(5)Sialic474(1)	N156	1279,870	3	290,4	3,44
Fetuin-A	P12763	10	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2283,779]DSR.V	HexNAc(5)Hex(5)Sialic229(2)	N156	1341,874	3	286,5	3,31
Fetuin-A	P12763	11	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2216,773]DSR.V	HexNAc(5)Hex(6)Sialic229(1)	N156	1319,870	3	351,4	3,89
Fetuin-A	P12763	12	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2445,832]DSR.V	HexNAc(5)Hex(6)Sialic229(2)	N156	1395,889	3	308,3	3,44
Fetuin-A	P12763	12	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2690,951]DSR.V	HexNAc(5)Hex(6)Sialic474(1)Sialic229(1)	N156	1108,450	4	312,3	2,93
Fetuin-A	P12763	13	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2674,891]DSR.V	HexNAc(5)Hex(6)Sialic229(3)	N156	1472,577	3	174,7	2,60
Fetuin-A	P12763	13	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2920,010]DSR.V	HexNAc(5)Hex(6)Sialic474(1)Sialic229(2)	N156	1165,712	4	177,1	1,99
Fetuin-A	P12763	13	K.LC[+57,021]PDC[+57,021]PLLAPLN[+3165,129]DSR.V	HexNAc(5)Hex(6)Sialic474(2)Sialic229(1)	N156	1227,245	4	248,5	2,43
Fetuin-A	P12763	14	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2632,917]DSR.V	HexNAc(6)Hex(5)Fuc(1)Sialic229(2)	N156	1094,194	4	194,0	2,62
Fetuin-A	P12763	15	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2861,976]DSR.V	HexNAc(6)Hex(5)Fuc(1)Sialic229(3)	N156	1151,708	4	161,9	1,47
Fetuin-A	P12763	16	L.APLN[+3676,319]DSR.V	HexNAc(6)Hex(7)Fuc(1)Sialic474(2)Sialic229(1)	N156	1483,922	3	36,6	1,64
Fetuin-A	P12763	17	L.APLN[+3447,260]DSR.V	HexNAc(6)Hex(7)Fuc(1)Sialic474(2)	N156	1407,226	3	50,1	1,38
Fetuin-A	P12763	18	A.PLN[+3514,266]DSR.V	HexNAc(6)Hex(6)Fuc(1)Sialic474(2)Sialic229(1)	N156	1405,885	3	91,3	1,38
Fetuin-A	P12763	19	L.N[+3597,332]DSR.V	HexNAc(6)Hex(5)Fuc(1)Sialic474(3)	N156	1363,852	3	25,9	1,64
Fetuin-A	P12763	20	L.N[+4436,642]DSR.V	HexNAc(7)Hex(6)Fuc(1)Sialic474(4)	N156	1232,714	4	170,7	1,35
Fetuin-A	P12763	21	A.LATFNAESN[+2920,010]GSYLQLVEISR.A	HexNAc(5)Hex(6)Sialic474(1)Sialic229(2)	N176	1283,786	4	181,7	1,79
Fetuin-A	P12763	22	A.LATFNAESN[+2445,832]GSYLQLVEISR.A	HexNAc(5)Hex(6)Sialic229(2)	N176	1552,983	3	44,7	1,66
Fetuin-A	P12763	23	F.NAESN[+2920,010]GSYLQLVEISR.A	HexNAc(5)Hex(6)Sialic474(1)Sialic229(2)	N176	1175,727	4	141,9	1,48
Fetuin-A	P12763	24	F.NAESN[+2445,832]GSYLQLVEISR.A	HexNAc(5)Hex(6)Sialic229(2)	N176	1409,240	3	327,9	3,38
Fetuin-A	P12763	25	S.N[+3285,142]GSYLQLVEISR.A	HexNAc(6)Hex(7)Sialic474(1)Sialic229(2)	N176	1166,466	4	155,5	1,49
Fetuin-B	Q58D62	26	K.TEELQQQNT[+839,310]APTNSPTK.A	HexNAc(1)Hex(1)Sialic474(1)	T295	909,413	3	6,7	1,47
Fetuin-B	Q58D62	27	K.TEELQQQNTAPT[+839,310]NSPTK.A	HexNAc(1)Hex(1)Sialic474(1)	T295	909,412	3	4,4	5,80
Fetuin-B	Q58D62	28	R.PANPSKTEELQQQNTAPT[+839,310]NSPTK.A	HexNAc(1)Hex(1)Sialic474(1)	T292	830,889	4	0,9	4,18

By closer analysis of Table 12, the *O*-glycosylations (entries in black) at the sites Ser_{296} and Ser_{341} in fetuin-A and Thr_{292} and Thr_{295} in fetuin-B were identified as simple HexNAc-Hex structures with one Neu5Ac (see Figure 26a for example of *O*-glycopeptide identified), except in one example at Ser_{296}, 2 Neu5Ac were observed attached to HexNAc-Hex core. However, the largest glycan variety was found in the *N*-glycosylation motif (Table 12, entries in blue), especially for the Asn_{156} site, bi-, tri- and tetraantennary modifications were identified, and at Asn_{176} site bi- and triantennary structures were observed. (see Figure 26b for an example of *N*-glycopeptide identified).

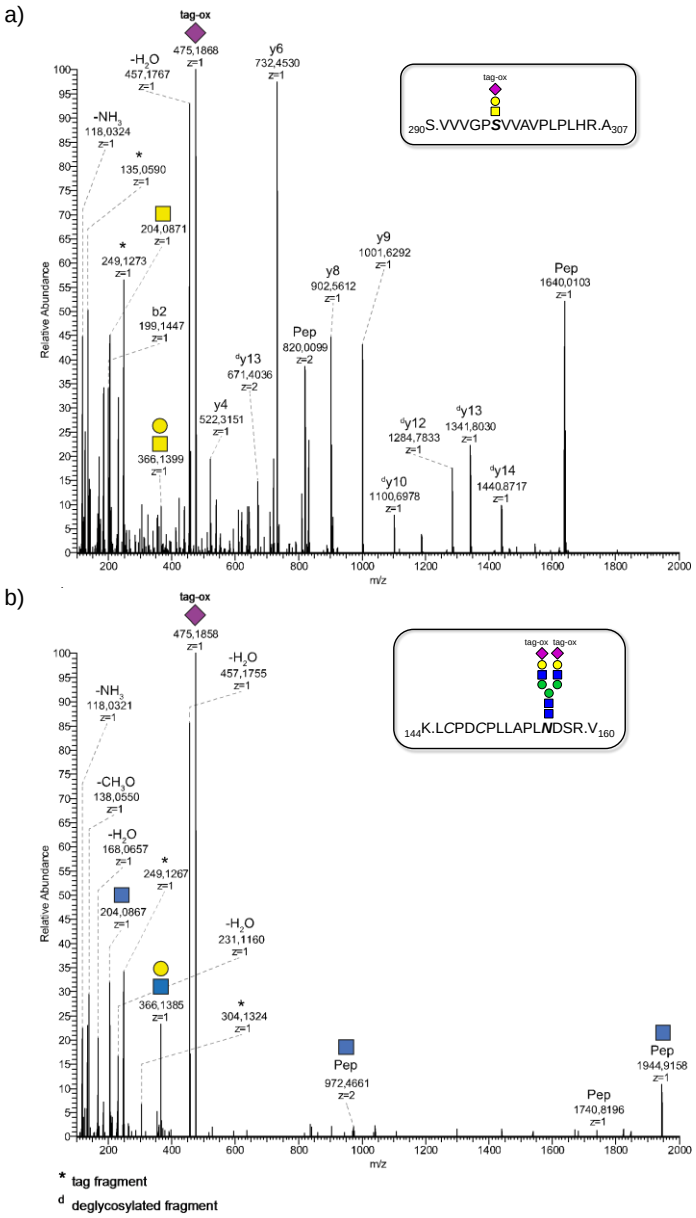


Figure 26. MS/MS spectra of *O*- (a) and *N*-glycopeptides (b) identified from fetuin sample using the new enrichment method protocol: oxidation with sodium periodate, enrichment with hydroxylamine beads and chemical cleavage for glycopeptide release.

5.5.2- Enrichment of glycopeptides from asialofetuin

With the success of enrichment of sialylated glycopeptides from protein sample, the enrichment of neutral glycopeptides from a model protein was also evaluated. For this reason, the commercial version of non-sialylated fetuin, asialofetuin, was chosen as model. A tryptic mixture of asialofetuin was oxidized using galactose oxidase in presence of bovine catalase. Glycopeptides were then chemically attached to the hydroxylamine beads for enrichment. After release, the peptides were analyzed using nano-LC-MS/MS methods described in the previous section. The obtained files were searched against the Uniprot sequence of bovine fetuin-A (P12763) and bovine fetuin-B (Q58D62) using Byonic™, as previously described. For the neutral enriched peptides, the same Byonic™ lists described in the previous session were modified for our searches. Therefore the lists contain 29 different *O*-glycans with at least one free terminal galactose and 36 *N*-glycans from bi- until tetraantennary structures with at least one free terminal galactose (see appendix Table A 11 for *O*-glycan list and Table A 12 for *N*-glycan list). In this case more complex structures were excluded from the list because they were not found in searches with the sialylated glycopeptides. The search files were further manually filtered in order to obtain the number of unique identified glycopeptides, using the same criteria as described before in the above section. In total 98 unique glycopeptides were found that cover 1 *O*-glycosylation site and 2 *N*-glycosylation sites of bovine fetuin-A and 2 *O*-glycosylation sites and 2 *N*-glycosylation sites of bovine fetuin-B (Table 13). Furthermore the software assigned 3 *O*-glycosylation sites that were not described in the protein database: Thr_{19}, Ser_{297} and Thr_{299}. However, for none of these new sites there were fragment oxonium ions containing the glycosylation in these positions in order to confirm them. For this reason the confirmation of these glycosylation sites should be validated using additional methodologies.

Nevertheless, a large diversity of glycosylation species in the non-sialylated sample is observed compared to the sialylated one. Heterogeneity that comes from the commercial source as well as from the oxidation process might lead to the largest differences. According to Sigma-Aldrich, asialofetuin is obtained from the acid treatment of fetuin, therefore sialic acids from the *N*-glycans might not be all removed, as observed, as well as loss of some *O*-glycans might also happen, since *O*-glycosylation is more labile under acidic conditions used to produce asialo fetuin from fetuin. Besides that, oxidation conditions using GOase and catalase were not optimized in a way that not all terminal galactose residues were oxidized increasing even more the complexity of the structures generated during the oxidation and enrichment process. Although the enrichment of neutral glycopeptide using hydroxylamine beads was successful, further improvement of the reaction conditions in order to reduce the sample complexity was considered necessary.

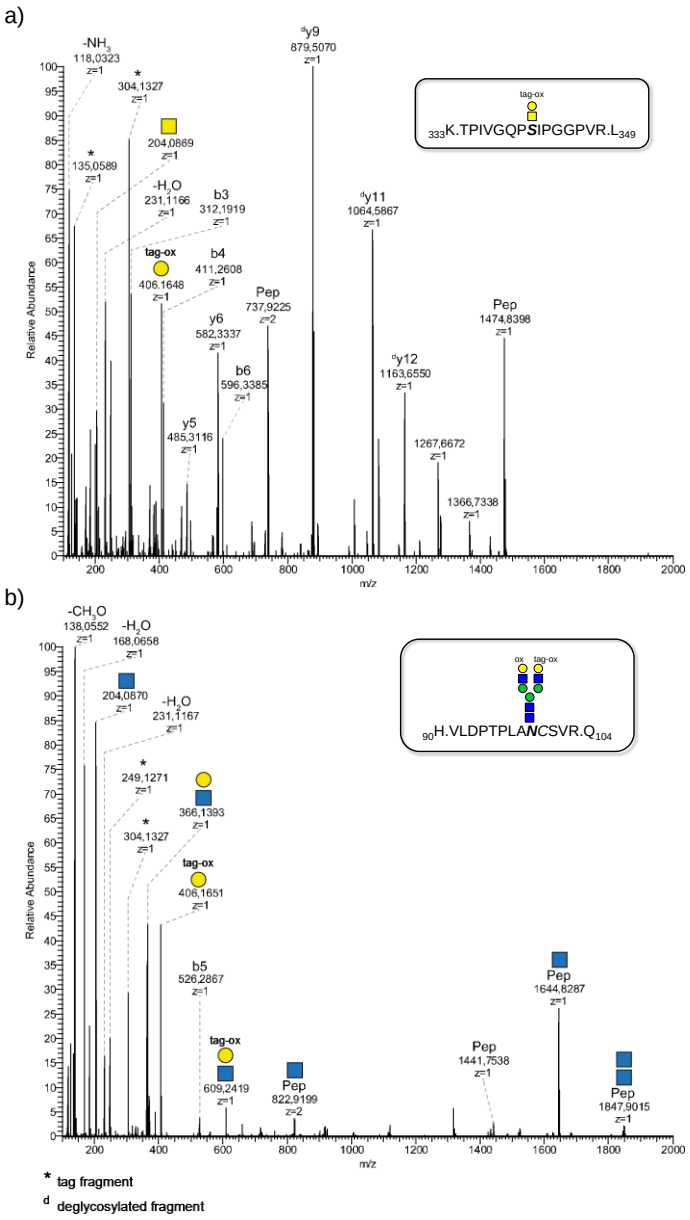


Figure 27. MS/MS spectra of *O*- (a) and *N*-glycopeptides (b) identified from asialofetuin sample using the new enrichment method protocol: oxidation with galactose oxidase, enrichment with hydroxylamine beads and chemical cleavage for glycopeptide release.

Table 13. Glycopeptides identified from commercial asialofetuin sample oxidized with galactose oxidase and enriched with hydroxylamine beads. (Entries in black are *O*-glycopeptides and in blue are *N*-glycopeptides.)

Protein Name	Accession Number (UniprotKB)	Peptide number	Peptide	Glycans	Glycosylation site	Observed m/z	z	Delta Mod	|Log Prob|
Fetuin-A	P12763	1	A.N[+2431,882]C[+57,021]SVR.Q	HexNAc(6)Hex(4)Gal160(1)Gal405(1)	N99	1023,066	3	46,0	2,84
Fetuin-A	P12763	2	A.N[+3373,254]C[+57,021]SVR.Q	HexNAc(6)Hex(4)NeuAc(1)Gal405(3)	N99	1002,644	4	18,7	2,62
Fetuin-A	P12763	3	C.HVLDPTPLAN[+2276,793]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(1)Gal160(1)	N99	1285,872	3	57,6	2,84
Fetuin-A	P12763	4	C.PDC[+57,021]PLLAPLN[+2230,818]DSR.V	HexNAc(5)Hex(5)Gal405(1)	N156	1233,518	3	34,8	2,98
Fetuin-A	P12763	5	C.PDC[+57,021]PLLAPLN[+2274,778]DSR.V	HexNAc(5)Hex(4)NeuAc(1)Gal160(2)	N156	936,387	4	55,3	2,33
Fetuin-A	P12763	6	D.C[+57,021]PLLAPLN[+1985,698]DSR.V	HexNAc(5)Hex(5)Gal160(1)	N156	1081,118	3	24,2	3,33
Fetuin-A	P12763	7	D.C[+57,021]PLLAPLN[+2129,740]DSR.V	HexNAc(5)Hex(4)Fuc(1)Gal160(2)	N156	847,096	4	66,9	3,01
Fetuin-A	P12763	8	D.C[+57,021]PLLAPLN[+2276,793]DSR.V	HexNAc(5)Hex(5)NeuAc(1)Gal160(1)	N156	1178,149	3	40,0	1,42
Fetuin-A	P12763	9	D.PTPLAN[+1911,661]C[+57,021]SVR.Q	HexNAc(4)Hex(4)NeuAc(1)Gal160(1)	N99	1009,412	3	91,1	4,36
Fetuin-A	P12763	9	D.PTPLAN[+2156,781]C[+57,021]SVR.Q	HexNAc(4)Hex(4)NeuAc(1)Gal405(1)	N99	1091,121	3	131,1	4,74
Fetuin-A	P12763	10	D.PTPLAN[+1823,645]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal160(1)	N99	980,076	3	147,3	3,34
Fetuin-A	P12763	10	D.PTPLAN[+2068,765]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal405(1)	N99	1061,781	3	167,5	5,96
Fetuin-A	P12763	11	D.PTPLAN[+2457,927]C[+57,021]SVR.Q	HexNAc(5)Hex(3)Fuc(1)Gal405(2)	N99	1191,503	3	77,5	1,47
Fetuin-A	P12763	12	D.PTPLAN[+2226,787]C[+57,021]SVR.Q	HexNAc(5)Hex(3)Gal160(2)Gal405(1)	N99	836,096	4	170,0	5,36
Fetuin-A	P12763	12	D.PTPLAN[+1981,667]C[+57,021]SVR.Q	HexNAc(5)Hex(3)Gal160(3)	N99	1032,754	3	66,3	2,99
Fetuin-A	P12763	12	D.PTPLAN[+2228,802]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal160(1)Gal405(1)	N99	1115,129	3	28,6	3,32
Fetuin-A	P12763	12	D.PTPLAN[+1985,698]C[+57,021]SVR.Q	HexNAc(5)Hex(5)Gal160(1)	N99	775,821	4	159,5	5,40
Fetuin-A	P12763	13	D.PTPLAN[+2129,740]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Fuc(1)Gal160(2)	N99	1082,103	3	134,0	4,01
Fetuin-A	P12763	13	D.PTPLAN[+2274,778]C[+57,021]SVR.Q	HexNAc(5)Hex(4)NeuAc(1)Gal160(2)	N99	1130,457	3	196,6	5,02
Fetuin-A	P12763	14	D.PTPLAN[+2276,793]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(1)Gal160(1)	N99	1131,127	3	76,5	3,80
Fetuin-A	P12763	14	D.PTPLAN[+2521,913]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(1)Gal405(1)	N99	1212,832	3	166,0	5,05
Fetuin-A	P12763	15	D.PTPLAN[+2567,889]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(2)Gal160(1)	N99	921,365	4	195,1	5,23
Fetuin-A	P12763	16	D.PTPLAN[+2641,926]C[+57,021]SVR.Q	HexNAc(6)Hex(6)NeuAc(1)Gal160(1)	N99	1252,838	3	74,1	4,23
Fetuin-A	P12763	17	D.TLETTC[+57,021]HVLDPTPLAN[+1911,661]C[+57,021]SVR.Q	HexNAc(4)Hex(4)NeuAc(1)Gal160(1)	N99	1049,696	4	64,2	3,16
Fetuin-A	P12763	18	D.TLETTC[+57,021]HVLDPTPLAN[+1967,688]C[+57,021]SVR.Q	HexNAc(5)Hex(3)Fuc(1)Gal160(2)	N99	1063,454	4	248,3	5,42
Fetuin-A	P12763	19	D.TLETTC[+57,021]HVLDPTPLAN[+1983,682]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal160(2)	N99	1067,953	4	138,0	4,42
Fetuin-A	P12763	19	D.TLETTC[+57,021]HVLDPTPLAN[+2230,818]C[+57,021]SVR.Q	HexNAc(5)Hex(5)Gal405(1)	N99	1505,643	3	117,1	5,67
Fetuin-A	P12763	20	D.TLETTC[+57,021]HVLDPTPLAN[+2276,793]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(1)Gal160(1)	N99	912,986	5	56,0	2,91
Fetuin-A	P12763	21	E.TTC[+57,021]HVLDPTPLAN[+1983,682]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal160(2)	N99	981,911	4	49,5	3,87
Fetuin-A	P12763	21	E.TTC[+57,021]HVLDPTPLAN[+1985,698]C[+57,021]SVR.Q	HexNAc(5)Hex(5)Gal160(1)	N99	982,412	4	118,9	5,07
Fetuin-A	P12763	22	E.TTC[+57,021]HVLDPTPLAN[+2276,793]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(1)Gal160(1)	N99	1055,186	4	230,4	6,21
Fetuin-A	P12763	23	H.VLDPTPLAN[+1766,624]C[+57,021]SVR.Q	HexNAc(4)Hex(4)Fuc(1)Gal160(1)	N99	1070,126	3	151,0	4,93
Fetuin-A	P12763	24	H.VLDPTPLAN[+1911,661]C[+57,021]SVR.Q	HexNAc(4)Hex(4)NeuAc(1)Gal160(1)	N99	1118,473	3	39,1	2,84
Fetuin-A	P12763	24	H.VLDPTPLAN[+2156,781]C[+57,021]SVR.Q	HexNAc(4)Hex(4)NeuAc(1)Gal405(1)	N99	900,388	4	0,0	4,14
Fetuin-A	P12763	25	H.VLDPTPLAN[+1823,645]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal160(1)	N99	1089,133	3	12,7	1,48
Fetuin-A	P12763	26	H.VLDPTPLAN[+2228,802]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal160(1)Gal405(1)	N99	918,392	4	23,6	2,28
Fetuin-A	P12763	26	H.VLDPTPLAN[+1985,698]C[+57,021]SVR.Q	HexNAc(5)Hex(5)Gal160(1)	N99	1143,151	3	139,4	1,73
Fetuin-A	P12763	27	H.VLDPTPLAN[+2274,778]C[+57,021]SVR.Q	HexNAc(5)Hex(4)NeuAc(1)Gal160(2)	N99	929,885	4	5,0	3,04

Fetuin-A	P12763	28	H.VLDPTPLAN[+2131,756]C[+57,021]SVR.Q	HexNAc(5)Hex(5)Fuc(1)Gal160(1)	N99	1191,832	3	38,6	1,39
Fetuin-A	P12763	28	H.VLDPTPLAN[+2376,876]C[+57,021]SVR.Q	HexNAc(5)Hex(5)Fuc(1)Gal405(1)	N99	1273,540	3	42,3	1,52
Fetuin-A	P12763	29	H.VLDPTPLAN[+2276,793]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(1)Gal160(1)	N99	930,389	4	104,9	3,08
Fetuin-A	P12763	30	H.VLDPTPLAN[+2641,926]C[+57,021]SVR.Q	HexNAc(6)Hex(6)NeuAc(1)Gal160(1)	N99	1021,423	4	90,6	3,57
Fetuin-A	P12763	31	I.DTLETTC[+57,021]HVLDPTPLAN[+1985,698]C[+57,021]SVR.Q	HexNAc(5)Hex(5)Gal160(1)	N99	1096,963	4	115,8	3,57
Fetuin-A	P12763	32	K.LC[+57,021]PDC[+57,021]PLLAPLN[+1458,513]DSR.V	HexNAc(4)Hex(3)Gal160(1)	N156	1067,122	3	207,9	5,34
Fetuin-A	P12763	33	K.LC[+57,021]PDC[+57,021]PLLAPLN[+1911,661]DSR.V	HexNAc(4)Hex(4)NeuAc(1)Gal160(1)	N156	1218,171	3	354,4	6,40
Fetuin-A	P12763	33	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2156,781]DSR.V	HexNAc(4)Hex(4)NeuAc(1)Gal405(1)	N156	975,163	4	398,1	7,27
Fetuin-A	P12763	34	K.LC[+57,021]PDC[+57,021]PLLAPLN[+1967,688]DSR.V	HexNAc(5)Hex(3)Fuc(1)Gal160(2)	N156	1236,853	3	401,0	7,12
Fetuin-A	P12763	35	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2374,860]DSR.V	HexNAc(5)Hex(4)Fuc(1)Gal160(1)Gal405(1)	N156	1029,929	4	302,8	5,36
Fetuin-A	P12763	36	K.LC[+57,021]PDC[+57,021]PLLAPLN[+1823,645]DSR.V	HexNAc(5)Hex(4)Gal160(1)	N156	1188,833	3	469,8	6,41
Fetuin-A	P12763	36	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2068,765]DSR.V	HexNAc(5)Hex(4)Gal405(1)	N156	953,158	4	440,3	7,22
Fetuin-A	P12763	37	K.LC[+57,021]PDC[+57,021]PLLAPLN[+1983,682]DSR.V	HexNAc(5)Hex(4)Gal160(2)	N156	931,888	4	203,0	5,00
Fetuin-A	P12763	37	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2473,922]DSR.V	HexNAc(5)Hex(4)Gal405(2)	N156	1054,448	4	186,4	5,89
Fetuin-A	P12763	37	K.LC[+57,021]PDC[+57,021]PLLAPLN[+1985,698]DSR.V	HexNAc(5)Hex(5)Gal160(1)	N156	932,390	4	234,8	5,27
Fetuin-A	P12763	37	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2230,818]DSR.V	HexNAc(5)Hex(5)Gal405(1)	N156	993,670	4	472,8	6,90
Fetuin-A	P12763	38	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2274,778]DSR.V	HexNAc(5)Hex(4)NeuAc(1)Gal160(2)	N156	1004,914	4	84,2	3,32
Fetuin-A	P12763	38	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2765,018]DSR.V	HexNAc(5)Hex(4)NeuAc(1)Gal405(2)	N156	1127,221	4	370,6	6,30
Fetuin-A	P12763	39	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2131,756]DSR.V	HexNAc(5)Hex(5)Fuc(1)Gal160(1)	N156	1291,533	3	193,0	5,38
Fetuin-A	P12763	40	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2276,793]DSR.V	HexNAc(5)Hex(5)NeuAc(1)Gal160(1)	N156	1339,880	3	54,8	3,57
Fetuin-A	P12763	40	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2521,913]DSR.V	HexNAc(5)Hex(5)NeuAc(1)Gal405(1)	N156	1066,444	4	5,0	3,54
Fetuin-A	P12763	41	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2567,889]DSR.V	HexNAc(5)Hex(5)NeuAc(2)Gal160(1)	N156	1436,910	3	6,1	2,04
Fetuin-A	P12763	41	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2813,009]DSR.V	HexNAc(5)Hex(5)NeuAc(2)Gal405(1)	N156	1139,219	4	264,3	6,77
Fetuin-A	P12763	42	K.LC[+57,021]PDC[+57,021]PLLAPLN[+1864,672]DSR.V	HexNAc(6)Hex(3)Gal160(1)	N156	902,136	4	296,0	6,86
Fetuin-A	P12763	43	K.LC[+57,021]PDC[+57,021]PLLAPLN[+2188,778]DSR.V	HexNAc(6)Hex(5)Gal160(1)	N156	983,158	4	447,6	6,93
Fetuin-A	P12763	44	K.TPIVGQPS[+363,117]IPGGPVR.L	HexNAc(1)Gal160(1)	S341	613,322	3	242,5	6,96
Fetuin-A	P12763	44	K.TPIVGQPS[+608,236]IPGGPVR.L	HexNAc(1)Gal405(1)	S341	695,029	3	376,7	7,16
Fetuin-A	P12763	45	L.AN[+2156,781]C[+57,021]SVR.Q	HexNAc(4)Hex(4)NeuAc(1)Gal405(1)	N99	955,046	3	515,1	7,41
Fetuin-A	P12763	46	L.AN[+2226,787]C[+57,021]SVR.Q	HexNAc(5)Hex(3)Gal160(2)Gal405(1)	N99	978,381	3	305,3	5,89
Fetuin-A	P12763	46	L.AN[+2228,802]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal160(1)Gal405(1)	N99	979,053	3	276,9	6,40
Fetuin-A	P12763	46	L.AN[+1985,698]C[+57,021]SVR.Q	HexNAc(5)Hex(5)Gal160(1)	N99	898,018	3	228,5	6,50
Fetuin-A	P12763	46	L.AN[+2230,818]C[+57,021]SVR.Q	HexNAc(5)Hex(5)Gal405(1)	N99	979,723	3	242,8	6,56
Fetuin-A	P12763	47	L.AN[+2521,913]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(1)Gal405(1)	N99	807,820	4	133,2	4,99
Fetuin-A	P12763	48	L,APLN[+2228,802]DSR.V	HexNAc(5)Hex(4)Gal160(1)Gal405(1)	N156	1001,074	3	298,1	6,41
Fetuin-A	P12763	48	L.APLN[+1983,682]DSR.V	HexNAc(5)Hex(4)Gal160(2)	N156	919,367	3	88,8	4,36
Fetuin-A	P12763	48	L.APLN[+1985,698]DSR.V	HexNAc(5)Hex(5)Gal160(1)	N156	920,039	3	0,6	3,70
Fetuin-A	P12763	49	L.APLN[+2276,793]DSR.V	HexNAc(5)Hex(5)NeuAc(1)Gal160(1)	N156	1017,071	3	311,5	6,80
Fetuin-A	P12763	49	L.APLN[+2521,913]DSR.V	HexNAc(5)Hex(5)NeuAc(1)Gal405(1)	N156	1098,772	3	128,0	5,27
Fetuin-A	P12763	50	L.APLN[+3519,312]DSR.V	HexNAc(6)Hex(4)Fuc(1)NeuAc(1)Gal405(3)	N156	1430,908	3	36,9	3,17
Fetuin-A	P12763	51	L.APLN[+2933,021]DSR.V	HexNAc(6)Hex(6)NeuAc(2)Gal160(1)	N156	1235,802	3	24,0	3,82
Fetuin-A	P12763	52	L.C[+57,021]PDC[+57,021]PLLAPLN[+1985,698]DSR.V	HexNAc(5)Hex(5)Gal160(1)	N156	1205,154	3	54,1	3,29
Fetuin-A	P12763	53	L.DPTPLAN[+1983,682]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal160(2)	N99	1071,766	3	310,2	5,57
Fetuin-A	P12763	54	L.DPTPLAN[+2276,793]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(1)Gal160(1)	N99	1169,469	3	3,2	3,89

Fetuin-A	P12763	54	L.DPTPLAN[+2521,913]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(1)Gal405(1)	N99	1251,175	3	9,9	4,05
Fetuin-A	P12763	55	L.ETTC[+57,021]HVLDPTPLAN[+1911,661]C[+57,021]SVR.Q	HexNAc(4)Hex(4)NeuAc(1)Gal160(1)	N99	996,165	4	180,6	5,80
Fetuin-A	P12763	55	L.ETTC[+57,021]HVLDPTPLAN[+2156,781]C[+57,021]SVR.Q	HexNAc(4)Hex(4)NeuAc(1)Gal405(1)	N99	1057,443	4	323,5	6,38
Fetuin-A	P12763	56	L.ETTC[+57,021]HVLDPTPLAN[+1967,688]C[+57,021]SVR.Q	HexNAc(5)Hex(3)Fuc(1)Gal160(2)	N99	1009,918	4	207,9	5,75
Fetuin-A	P12763	57	L.ETTC[+57,021]HVLDPTPLAN[+1823,645]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal160(1)	N99	974,162	4	280,1	5,49
Fetuin-A	P12763	57	L.ETTC[+57,021]HVLDPTPLAN[+2068,765]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal405(1)	N99	1035,441	4	195,0	1,70
Fetuin-A	P12763	58	L.ETTC[+57,021]HVLDPTPLAN[+1983,682]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal160(2)	N99	1351,887	3	240,0	3,51
Fetuin-A	P12763	58	L.ETTC[+57,021]HVLDPTPLAN[+2473,922]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal405(2)	N99	1136,729	4	62,6	9,20
Fetuin-A	P12763	58	L.ETTC[+57,021]HVLDPTPLAN[+1985,698]C[+57,021]SVR.Q	HexNAc(5)Hex(5)Gal160(1)	N99	1352,225	3	114,4	10,28
Fetuin-A	P12763	59	L.ETTC[+57,021]HVLDPTPLAN[+2114,741]C[+57,021]SVR.Q	HexNAc(5)Hex(4)NeuAc(1)Gal160(1)	N99	1046,933	4	38,0	2,79
Fetuin-A	P12763	60	L.ETTC[+57,021]HVLDPTPLAN[+2376,876]C[+57,021]SVR.Q	HexNAc(5)Hex(5)Fuc(1)Gal405(1)	N99	1112,462	4	27,8	2,98
Fetuin-A	P12763	61	L.ETTC[+57,021]HVLDPTPLAN[+2276,793]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(1)Gal160(1)	N99	1087,444	4	43,2	2,79
Fetuin-A	P12763	61	L.ETTC[+57,021]HVLDPTPLAN[+2521,913]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(1)Gal405(1)	N99	1531,298	3	36,9	2,62
Fetuin-A	P12763	62	L.ETTC[+57,021]HVLDPTPLAN[+2813,009]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(2)Gal405(1)	N99	1221,499	4	71,5	2,98
Fetuin-A	P12763	63	L.LAPLN[+1911,661]DSR.V	HexNAc(4)Hex(4)NeuAc(1)Gal160(1)	N156	933,054	3	77,9	3,84
Fetuin-A	P12763	64	L.LAPLN[+2228,802]DSR.V	HexNAc(5)Hex(4)Gal160(1)Gal405(1)	N156	1038,765	3	1,1	2,62
Fetuin-A	P12763	64	L.LAPLN[+1983,682]DSR.V	HexNAc(5)Hex(4)Gal160(2)	N156	957,062	3	68,8	2,83
Fetuin-A	P12763	64	L.LAPLN[+1985,698]DSR.V	HexNAc(5)Hex(5)Gal160(1)	N156	957,733	3	115,0	4,20
Fetuin-A	P12763	65	L.LAPLN[+2274,778]DSR.V	HexNAc(5)Hex(4)NeuAc(1)Gal160(2)	N156	1054,095	3	116,4	3,59
Fetuin-A	P12763	66	L.N[+2228,802]DSR.V	HexNAc(5)Hex(4)Gal160(1)Gal405(1)	N156	907,349	3	147,4	4,72
Fetuin-A	P12763	66	L.N[+2473,922]DSR.V	HexNAc(5)Hex(4)Gal405(2)	N156	989,057	3	107,3	3,37
Fetuin-A	P12763	66	L.N[+2230,818]DSR.V	HexNAc(5)Hex(5)Gal405(1)	N156	908,021	3	103,5	4,20
Fetuin-A	P12763	67	L.N[+2519,898]DSR.V	HexNAc(5)Hex(4)NeuAc(1)Gal160(1)Gal405(1)	N156	1004,382	3	138,2	4,55
Fetuin-A	P12763	67	L.N[+2521,913]DSR.V	HexNAc(5)Hex(5)NeuAc(1)Gal405(1)	N156	1005,053	3	2,8	2,45
Fetuin-A	P12763	68	L.N[+2959,067]DSR.V	HexNAc(5)Hex(5)Fuc(1)NeuAc(2)Gal405(1)	N156	1150,778	3	7,5	1,66
Fetuin-A	P12763	69	L.N[+2821,044]DSR.V	HexNAc(6)Hex(3)Fuc(1)Gal160(1)Gal405(2)	N156	828,829	4	204,6	5,93
Fetuin-A	P12763	69	L.N[+2577,940]DSR.V	HexNAc(6)Hex(4)Fuc(1)Gal160(1)Gal405(1)	N156	1023,725	3	127,7	4,56
Fetuin-A	P12763	69	L.N[+2334,835]DSR.V	HexNAc(6)Hex(5)Fuc(1)Gal160(1)	N156	942,684	3	143,7	5,39
Fetuin-A	P12763	70	L.N[+2674,986]DSR.V	HexNAc(6)Hex(3)Gal160(1)Gal405(2)	N156	1056,070	3	86,5	4,12
Fetuin-A	P12763	71	R.KLC[+57,021]PDC[+57,021]PLLAPLN[+1911,661]DSR.V	HexNAc(4)Hex(4)NeuAc(1)Gal160(1)	N156	945,905	4	5,6	2,29
Fetuin-A	P12763	71	R.KLC[+57,021]PDC[+57,021]PLLAPLN[+2156,781]DSR.V	HexNAc(4)Hex(4)NeuAc(1)Gal405(1)	N156	1007,184	4	188,0	4,97
Fetuin-A	P12763	72	R.KLC[+57,021]PDC[+57,021]PLLAPLN[+1983,682]DSR.V	HexNAc(5)Hex(4)Gal160(2)	N156	1284,877	3	275,6	5,24
Fetuin-A	P12763	72	R.KLC[+57,021]PDC[+57,021]PLLAPLN[+1985,698]DSR.V	HexNAc(5)Hex(5)Gal160(1)	N156	1285,548	3	180,0	5,12
Fetuin-A	P12763	72	R.KLC[+57,021]PDC[+57,021]PLLAPLN[+2230,818]DSR.V	HexNAc(5)Hex(5)Gal405(1)	N156	1367,253	3	174,3	3,35
Fetuin-A	P12763	73	R.KLC[+57,021]PDC[+57,021]PLLAPLN[+2276,793]DSR.V	HexNAc(5)Hex(5)NeuAc(1)Gal160(1)	N156	1037,185	4	132,2	4,06
Fetuin-A	P12763	73	R.KLC[+57,021]PDC[+57,021]PLLAPLN[+2521,913]DSR.V	HexNAc(5)Hex(5)NeuAc(1)Gal405(1)	N156	1098,466	4	6,0	1,89
Fetuin-A	P12763	74	R.KLC[+57,021]PDC[+57,021]PLLAPLN[+2567,889]DSR.V	HexNAc(5)Hex(5)NeuAc(2)Gal160(1)	N156	1109,960	4	289,8	5,17
Fetuin-A	P12763	74	R.KLC[+57,021]PDC[+57,021]PLLAPLN[+2813,009]DSR.V	HexNAc(5)Hex(5)NeuAc(2)Gal405(1)	N156	1171,240	4	50,3	3,55
Fetuin-A	P12763	75	T.C[+57,021]HVLDPTPLAN[+2230,818]C[+57,021]SVR.Q	HexNAc(5)Hex(5)Gal405(1)	N99	993,169	4	253,6	5,08
Fetuin-A	P12763	76	T.C[+57,021]HVLDPTPLAN[+2276,793]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(1)Gal160(1)	N99	1004,411	4	19,8	3,06
Fetuin-A	P12763	76	T.C[+57,021]HVLDPTPLAN[+2519,898]C[+57,021]SVR.Q	HexNAc(5)Hex(4)NeuAc(1)Gal160(1)Gal405(1)	N99	1065,689	4	127,8	3,61
Fetuin-A	P12763	77	T.LETTC[+57,021]HVLDPTPLAN[+2228,802]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal160(1)Gal405(1)	N99	1103,719	4	33,1	2,84
Fetuin-A	P12763	78	T.LETTC[+57,021]HVLDPTPLAN[+2521,913]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(1)Gal405(1)	N99	1176,993	4	158,8	4,61

Fetuin-A	P12763	79	T.PLAN[+1985,698]C[+57,021]SVR.Q	HexNAc(5)Hex(5)Gal160(1)	N99	968,064	3	166,4	4,15
Fetuin-A	P12763	80	T.TC[+57,021]HVLDPTPLAN[+1983,682]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal160(2)	N99	956,651	4	198,6	5,05
Fetuin-A	P12763	81	T.TC[+57,021]HVLDPTPLAN[+2276,793]C[+57,021]SVR.Q	HexNAc(5)Hex(5)NeuAc(1)Gal160(1)	N99	1029,918	4	4,3	2,23
Fetuin-A	P12763	82	V.LDPTPLAN[+1983,682]C[+57,021]SVR.Q	HexNAc(5)Hex(4)Gal160(2)	N99	1109,462	3	7,0	5,03
Fetuin-B	Q58D62	83	A.T[+608,236]SPPQPAAR.P	HexNAc(1)Gal405(1)	T19	511,581	3	116,3	3,70
Fetuin-B	Q58D62	84	E.ELQQQNTAPT[+363,117]NSPTK.A	HexNAc(1)Gal160(1)	T295	673,986	3	137,4	4,73
Fetuin-B	Q58D62	85	E.LQQQNTAPTNS[+608,236]PTK.A	HexNAc(1)Gal405(1)	S297	712,677	3	1,6	2,92
Fetuin-B	Q58D62	86	E.LQQQNTAPTNSPT[+363,117]K.A	HexNAc(1)Gal160(1)	T299	630,972	3	196,7	5,30
Fetuin-B	Q58D62	87	E.N[+2415,887]ATVNQR.P	HexNAc(6)Hex(3)Fuc(1)Gal160(1)Gal405(1)	N271	1073,436	3	166,9	5,35
Fetuin-B	Q58D62	88	E.N[+2625,931]ATVNQR.P	HexNAc(6)Hex(5)Fuc(1)NeuAc(1)Gal160(1)	N271	1143,459	3	171,1	5,24
Fetuin-B	Q58D62	89	E.N[+2917,026]ATVNQR.P	HexNAc(6)Hex(5)Fuc(1)NeuAc(2)Gal160(1)	N271	1240,493	3	3,8	2,87
Fetuin-B	Q58D62	90	K.TEELQQQNT[+608,236]APTNSPTK.A	HexNAc(1)Gal405(1)	T292	1248,076	2	61,5	1,32
Fetuin-B	Q58D62	91	K.TEELQQQNT[+728,249]APTNSPTK.A	HexNAc(2)Hex(1)Gal160(1)	T292	872,392	3	2,3	1,56
Fetuin-B	Q58D62	92	K.TEELQQQNTAPT[+363,117]NSPTK.A	HexNAc(1)Gal160(1)	T295	1125,515	2	69,4	2,17
Fetuin-B	Q58D62	92	K.TEELQQQNTAPT[+608,236]NSPTK.A	HexNAc(1)Gal405(1)	T295	832,389	3	124,8	3,67
Fetuin-B	Q58D62	93	K.TEELQQQNTAPT[+728,249]NSPTK.A	HexNAc(2)Hex(1)Gal160(1)	T295	872,394	3	116,9	3,71
Fetuin-B	Q58D62	93	K.TEELQQQNTAPT[+973,369]NSPTK.A	HexNAc(2)Hex(1)Gal405(1)	T295	954,096	3	12,2	2,62
Fetuin-B	Q58D62	94	K.TEELQQQNTAPTNS[+363,117]PTK.A	HexNAc(1)Gal160(1)	S297	750,683	3	9,5	3,33
Fetuin-B	Q58D62	94	K.TEELQQQNTAPTNS[+608,236]PTK.A	HexNAc(1)Gal405(1)	S297	832,386	3	4,9	2,62
Fetuin-B	Q58D62	95	N.TAPT[+608,236]NSPTK.A	HexNAc(1)Gal405(1)	T295	762,861	2	96,4	3,14
Fetuin-B	Q58D62	95	Q.NTAPT[+608,236]NSPTK.A	HexNAc(1)Gal405(1)	T295	819,881	2	3,9	3,10
Fetuin-B	Q58D62	96	R.GEN[+1985,698]ATVNQRPANPSK.T	HexNAc(5)Hex(5)Gal160(1)	N271	1190,171	3	73,3	3,80
Fetuin-B	Q58D62	96	R.GEN[+2230,818]ATVNQRPANPSK.T	HexNAc(5)Hex(5)Gal405(1)	N271	954,160	4	114,6	3,14
Fetuin-B	Q58D62	97	R.GEN[+2521,913]ATVNQRPANPSK.T	HexNAc(5)Hex(5)NeuAc(1)Gal405(1)	N271	1026,934	4	31,7	2,47
Fetuin-B	Q58D62	98	Y.N[+2567,889]C[+57,021]TLR.P	HexNAc(5)Hex(5)NeuAc(2)Gal160(1)	N137	1077,733	3	63,8	2,54
Fetuin-B	Q58D62	98	Y.N[+2813,009]C[+57,021]TLR.P	HexNAc(5)Hex(5)NeuAc(2)Gal405(1)	N137	869,839	4	67,2	4,53

5.6- Clinical sample evaluation

In order to verify the applicability of the methodology in complex systems, our enrichment method was applied for the enrichment of glycopetides in clinical samples (serum and cerebrospinal fluid - CSF) provided by our collaborator Dr. Jonas Nilsson from Sahlgrenska University Hospital in Gothenburg.

5.6.1- Enrichment of sialylated serum glycopeptides

Tryptic mixture of serum peptides was oxidized using periodic acid and the sialylated glycopeptides were then immobilized through a covalent linkage to the beads. After the bead release the peptides obtained were analyzed by nano-LC-MS/MS using a Fusion Lumos Tribrid mass spectrometer. MS/MS experiments used HCD with normalized collision energy of 30% and ETD supplemented by 33% HCD activation (EthCD). For data analysis, the raw files of the enriched sample were searched against the human Uniprot sequence using Byonic™ as search engine. The Byonic™ default lists with 27 modified sialylated *O*-glycans and 68 *N*-glycans were used as glycan database, as previously described for the fetuin analysis (see both lists in the appendix Table A 9 and Table A 10). The search files were further manually filtered in order to obtain the number of unique glycopeptides identified as described before. A glycosylation hit was defined as the peptides that presented "|Log Prob|" <1, "Delta Mod" >10. Glycosylation sites were further ranked according to Byonic™ parameter "Delta Mod" and manual inspection of the spectrum. MS/MS spectra that presented more than one fragment ion confirming the glycosylation site and "Delta Mod" >10 were assigned as high confidence site. Spectra with only one fragment ion confirming glycosylation site and "Delta Mod" >10 were assigned as good confidence site. These spectra with only one fragment ion confirming the glycosylation site or "Delta Mod" >10 were assigned as medium confidence site. Finally spectra with neither one fragment ion confirming glycosylation site nor "Delta Mod" >10 were assigned as poor confidence site.

In serum a total of 34 proteins and 162 unique glycopeptides were identified by enrichment of sialylated glycopeptides. From all these proteins, 20 have only *O*-glycosylation identified while 7 have only *N*-glycosylation, and 7 proteins presented both, *O*- and *N*-glycosylations (Table 14, and for detailed overview of identified glycopeptides: appendix Table A 13).

Although there are more *N*-glycosylation sites compared to the *O*-glycosylation sites, the analysis of *N*-glycopeptides simultaneously with *O*-glycopeptides is still a challenge. Fragmentation of the peptide backbone for *O*-glycopeptides is favored compare to *N*-glycopetides using the chosen methodology due to the size of the glycosylation and the stability of the glycosidic bond. *O*-glycosylation are in general smaller in size than the *N*-glycosylation modification and, for this reason the probability of peptide fragmentation compared to the glycosylation fragmentation is higher in *O*-glycopeptides.

Table 14. Identified proteins from serum sample oxidized with periodate and enriched using hydroxylamine beads.

Protein number	Protein	Accession number (UniprotKB)	Glycosylation Type
1	Ig lambda-7 chain C region	A0M8Q6	O
2	Ceruloplasmin	P00450	N
3	Haptoglobin	P00738	N
4	Plasminogen	P00747	O
5	Coagulation factor XII	P00748	O
6	Alpha-1-antitrypsin	P01009	N
7	Kininogen-1	P01042	O
8	Ig gamma-1 chain C region	P01857	O and N
9	Ig gamma-3 chain C region	P01860	O
10	Ig alpha-1 chain C region	P01876	O and N
11	Ig alpha-2 chain C region	P01877	N
12	Ig delta chain C region	P01880	O
13	Apolipoprotein E	P02649	O
14	Apolipoprotein A-II	P02652	O
15	Apolipoprotein C-III	P02656	O
16	Complement component C9	P02748	O
17	Beta-2-glycoprotein 1	P02749	O and N
18	Fibronectin	P02751	O
19	Alpha-2-HS-glycoprotein	P02765	O and N
20	Serum albumin	P02768	O
21	Serotransferrin	P02787	N
22	Hemopexin	P02790	O and N
23	Vitronectin	P04004	N
24	Apolipoprotein B-100	P04114	O and N
25	Histidine-rich glycoprotein	P04196	O and N
26	Plasma protease C1 inhibitor	P05155	O and N
27	Complement factor H	P08603	N
28	Complement C4-A	P0C0L4	O and N
29	Coagulation factor V	P12259	O
30	Inter-alpha-trypsin inhibitor heavy chain H1	P19827	O
31	Hepatocyte growth factor activator	Q04756	O
32	Apolipoprotein F	Q13790	O
33	Inter-alpha-trypsin inhibitor heavy chain H4	Q14624	O
34	Proteoglycan 4	Q92954	O

Comparing the protein list (Table 14 and Table A 13) with databases, nearly all identified proteins were previously reported as glycoproteins. Moreover, the presence of proteins that are related to coagulation processes as coagulation factor V and XII, as well as proteins described to be present in plasma[328] as alpha 2-HS glycoprotein, apolipoprotein C-III, complement C4-A, hemopexin, inter-alpha-trypsin inhibitor heavy chain H4 and kininogen-1 supports the applicability of the hydroxylamine beads.

One of the identified proteins, Ig lambda-7 chain C region (A0M8Q6), was not previously reported to be glycosylated. The sequence $_{88}$R.V***T***HEGSTVEK.T$_{99}$ (*m/z* 840,86^{2+} and 963,42^{2+}) was observed modified with mucin-type HexNAc-Hex-Neu5Ac motif. The fragment ions from the MS^2 scan of the precursor ion *m/z* 963,42^{2+} showed almost the complete b and/or c ions with the glycosylation modification while y ions appear without glycosylation (see Figure 28a). In the same sample, proteins that are already reported to be glycosylated with new glycosylation sites were also identified. One new possible glycosylation site belongs to plasminogen (P00747). This glycoprotein

has 3 glycosylations sites reported: Ser_{268}, Asp_{308} and Thr_{365}. The following peptides $_{261}$R.*C*TTPPP***S***SGPTYQ*C*LK.G_{278} (*m/z* 796,68^{3+}) and $_{348}$K.IPS*C*DSSPVSTEQLAP***T***APPELTPVVQD*C*YHGDG QSYR.G_{387} (*m/z* 1189,03^{4+}) were identified with previously described sites. Although manual spectra verification does not hint for the glycosylation site, the software assignment is in agreement with literature. Moreover Thr_{264} was also pointed as a possible glycosylation site (Figure 28b). From the same peptide $_{261}$R.*C*T***T***PPPSSGPTYQ*C*LK.G_{278} (*m/z* 796,68^{3+}), indirect glycosylation site was assigned due to fragment ions c2 (*m/z* 279,11^{1+}), c7 (*m/z* 1 352,49^{1+}), b7 (*m/z* 1335,67^{1+}), y12 (*m/z* 1334,72^{1+}), y13 (*m/z* 1431,81^{1+}) and y13^{2+} (*m/z* 716,35^{2+}). Nevertheless further validation should be done via other methodologies in order to confirm the new glycosylation sites even though data strongly hints to the new modification.

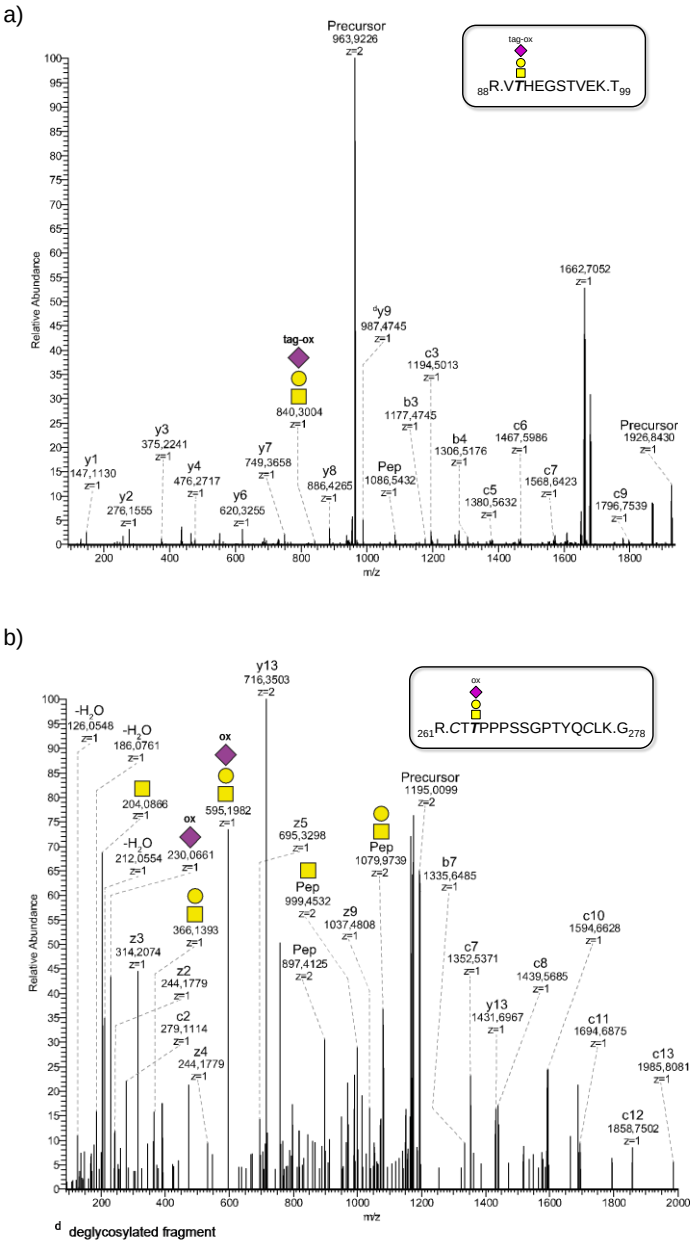


Figure 28. Fragmentation of peptides from Ig lambda-7 chain C region (a) and plasminogen (b) enriched from serum sample with cleavable hydroxylamine beads.

5.6.2- Enrichment of sialylated CSF glycopeptides

Tryptic mixture of CSF peptides were treated and analyzed under the same conditions described in the previous session for serum sample. In CSF 32 proteins and 129 unique glycopeptides were identified. From all identified proteins, 20 were found with only *O*-glycosylation sites while 8 had only *N*-glycosylation sites, and 4 proteins presented both, *O*- and *N*-glycosylation sites (Table 15, and for detailed overview of identified glycopeptides: Table A 14).

Table 15. Identified proteins from CSF sample oxidized with periodate and enriched using hydroxylamine beads.

Protein number	Protein	Accession number (UniprotKB)	Glycosylation Type
1	Ig lambda-7 chain C region	A0M8Q6	O
2	Neural cell adhesion molecule L1-like protein	O00533	N
3	Neurosecretory protein VGF	O15240	O
4	Prothrombin	P00734	N
5	Haptoglobin	P00738	N
6	Plasminogen	P00747	O
7	Alpha-1-antitrypsin	P01009	N
8	Complement C3	P01024	O
9	Kininogen-1	P01042	O
10	Ig gamma-3 chain C region	P01860	O
11	Apolipoprotein E	P02649	O
12	Apolipoprotein C-III	P02656	O
13	Fibrinogen alpha chain	P02671	O
14	Complement component C9	P02748	O
15	Beta-2-glycoprotein 1	P02749	N
16	Fibronectin	P02751	O
17	Alpha-2-HS-glycoprotein	P02765	O and N
18	Serum albumin	P02768	O
19	Serotransferrin	P02787	O and N
20	Hemopexin	P02790	O and N
21	Vitronectin	P04004	N
22	Histidine-rich glycoprotein	P04196	N
23	Complement component C7	P10643	O
24	Chromogranin-A	P10645	O
25	Inter-alpha-trypsin inhibitor heavy chain H1	P19827	O
26	Prostaglandin-H2 D-isomerase	P41222	O and N
27	Fractalkine	P78423	O
28	Hepatocyte growth factor activator	Q04756	O
29	Apolipoprotein F	Q13790	O
30	Inter-alpha-trypsin inhibitor heavy chain H4	Q14624	O
31	Dickkopf-related protein 3	Q9UBP4	O
32	ProSAAS	Q9UHG2	O

Herein proteins exclusively from CSF were identified in contrast to serum as neural cell adhesion molecule L1-like protein, neurosecretory protein VGF, chromogranin-A, fractalkine and proSAAS. Alpha 2-HS glycoprotein was here also identified. Alpha 2-HS glycoprotein (P02765), known as fetuin, has previously been reported with 5 glycosylation sites, according to UniprotKB: Asp_{156},

Asp_{176}, Thr_{256}, Thr_{270} and Ser_{346}. Two out of five glycosylation sites were identified. One of these sites has the sequence $_{340}$R.TVVQP***S***VGAAAGPVVPP***C***PGR.I$_{362}$ (*m/z* 952,47^{3+}) and fragment ions as b4 (*m/z* 428,25^{1+}), c5 (*m/z* 524,27^{1+}), c7 (*m/z* 1567,64^{1+}), b8$_{-18}$ (*m/z* 1 589,85^{1+}) and c8 (*m/z* 1 624,81^{1+}) confirm the glycosylation site (see Figure 29a). Another identified site belongs to the sequence $_{165}$K.AALAAFNAQN***N***GSNFQLEEISR.A$_{187}$ (*m/z* 1 173,25^{4+}). In this case the Figure 29b showed partially fragmentation of the peptide backbone for this *N*-glycopeptide, which was rarely observed.

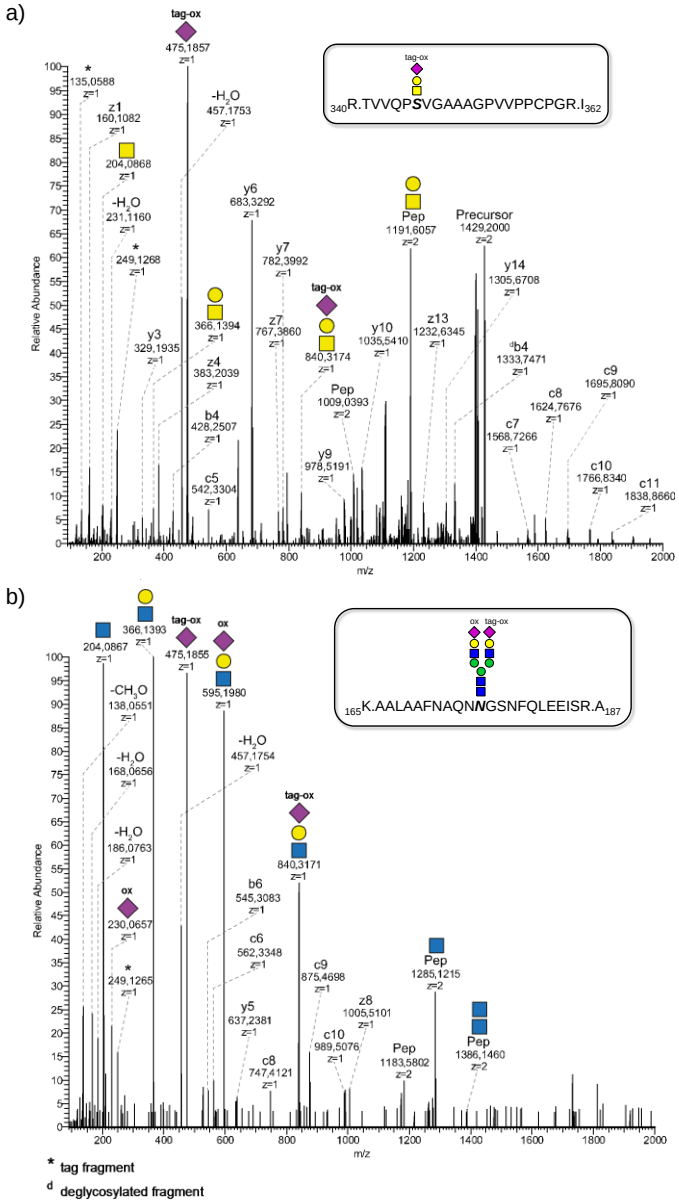


Figure 29. MS/MS spectra of two peptides from enriched alpha 2-HS glycoprotein (P02765) that belong to CSF sample presenting *O*-glycosylation (a) and *N*-glycosylation (b).

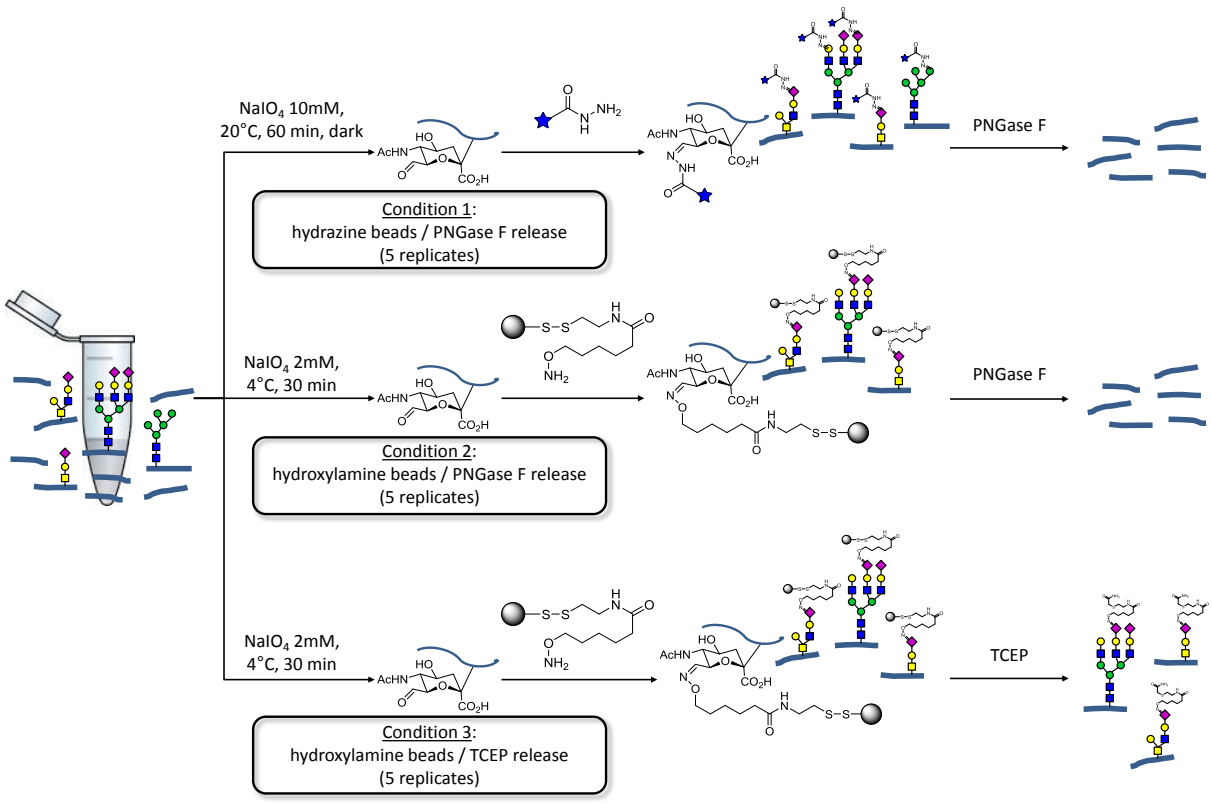
Comparing the enrichment from both, serum and CSF, samples most of the glycoproteins identified were found in both samples, in total 21. Only in serum, 13 unique proteins, involved in

coagulation process and present in plasma, were identified. And on the other hand 11 unique proteins were found exclusively in CSF.

5.6.3- Analysis of intact glycopeptides compared with methodologies for determination of glycosylation sites

The scope of the glycopeptide enrichment method using hydroxylamine beads was investigated in comparison with the hydrazine bead capture-release methodology. Neurological samples (CSF) were used to evaluate the potential use of the new methodology for high throughput glycoproteomics. This work was done in collaboration with Dr. Astrid Guldbrandsen and Prof. Dr. Frode Berven from the Proteomics Unit at the University of Bergen (PROBE), Norway.

Here, the hydroxylamine beads enrichment in CSF was quantitatively and qualitatively evaluated in replicates. A first set of samples was enriched using the standard approach with commercial hydrazine beads and PNGase F treatment for bead release of the enriched glycopeptides. A second set of samples was enriched with hydroxylamine beads and treated with PNGase F for peptide bead release. A third and final set of samples was enriched again with hydroxylamine beads, but were instead released from the beads by reduction of a disulfide linker to obtain glycopeptides leaving the complete glycan structure intact (Scheme 24). In this way it was possible compare the efficacy of glycopeptide capture, release and MS-analysis using the different methods using a large diversity of glycopeptide structures.



Scheme 24. Experimental conditions for evaluation of cleavable hydroxylamine beads in comparison with commercial hydrazine beads using a larger scope of glycopeptides from clinical CSF samples.

During the experiments, conditions were further optimized and the amount of hydroxylamine beads as well as the washing procedure were modified. As standard condition (condition 1), the previously described hydrazide-bead enrichment protocol by Guldbrandsen and coworkers[329] was used. This procedure used 10mM $NaIO_4$ at 20°C for 60 minutes in the dark as oxidation conditions, capture was done using commercial microporous hydrazine beads and the glycopeptides were released in the peptide form after treatment with PNGase F. Since the control of oxidation conditions is important for the posterior data analysis, for oxidation before capture with cleavable hydroxylamine beads, our optimized conditions were used as described above for the synthetic glycopeptides: 2mM $NaIO_4$ at 4°C for 30 minutes. The release of (glyco-)peptides were here evaluated in a larger scope. Oxidized glycopeptides captured by hydroxylamine beads were partially released using PNGase F treatment (condition 2), as described for the standard method. Another part of the peptides were released by reduction of the disulfide linker attached to the beads by treatment with TCEP in order to obtain the entire glycopeptide structure (condition 3).

In order to increase the number of identifications, the samples released with PNGase F were labeled with 10-plex TMT® and further fractionated by a mixed-mode reversed-phase anion exchange chromatography before LC-MS/MS analysis. Fractions were analyzed by nano-LC-MS/MS using an Orbitrap QExactive HF mass spectrometer. MS/MS experiments used HCD with normalized collision energy of 32%. For data analysis, the raw files of the fractions were searched against the human Uniprot reviewed database (SwissProt, January 2017) using SearchGUI[330] and the search engines X!Tandem[331], MyriMatch[332] and Comet[333]. Search results were summarized in

PeptideShaker[334] and peptide quantitative results were explored and exported using Reporter[335]. For each peptide, theTMT reporter abundance was normalized by scaling factors derived from the median abundances of all channels over all peptides. All peptides without values in at least one of the channels, without a consensus sequence Asn-Xxx-Thr/Ser, Xxx≠Pro and contaminants were excluded from the downstream analysis after the normalization.

The identification of deglycosylated peptides could for instance give false positive results. The deamidation in aspagarine with *N*-glycosylation consensus sequence strongly indicated the presence of glycosylation prior the treatment with PNGase F in those samples. Because of the uncertain degree in the glycosite localization and identification, scoring algorithms as D-score[336] and Probabilistic PTM score[329] were used to test the probability of a modification location in the peptide based on the spectrum information. Herein, the peptides in which the deamidation modification in consensus sequence asparagine (N-X-T/S) was doubtful assigned together with a confident assignment at asparagine that is not in a consensus sequence motif were also excluded.

A total of 2 963 peptides were identified using those criteria mentioned above. An F-test was applied in order to investigate the equality variance between the enrichment using commercial hydrazine beads and hydroxylamine beads. From all 2 963 identified peptides, 2 619 peptides did not have a significant F-test value (p>0,05), which means that approximately 88% of all identified peptides present an equal variance between both enrichments: using hydrazine and hydroxylamine beads with PNGase F release. A Student's *t*-test with equal variance was applied to the complete dataset and 1 777 peptides exhibited a p-value less or equal to 0,05. Due to the large number of identified peptides, a Bejamini-Hochberg multiple hypothesis test[337-338] was applied to the dataset in order to further refine the data's statistical significance. Using a FDR of 1%, 1 138 peptides, or 38% of all identified peptides, proved to be significant. Furthermore, the ratio of peptides enriched with hydrazine beads by peptides enriched with hydroxylamine beads presents a standard deviation of 0,76. All these results indicated that both enrichment methods, using hydrazine and hydroxylamine beads, produced too diverse results to efficiently be compared. The first difference might be occasioned by the oxidation method. The glycopeptide oxidation condition used for hydrazine bead enrichment (10mM $NaIO_4$, 60min, 20°C) was harsher compared to the condition used for glycopeptide oxidation for hydroxylamine beads (2mM $NaIO_4$, 30min, 4°C). Other *cis*-diols present in the glycan moiety, as mannose and galactose, could be oxidized besides the sialic acid. Meanwhile the oxidation used for hydroxylamine beads targets only the glycerol motif of the sialic acid, as previously discussed. For this reason, the targeted glycopeptides in both enrichment protocols are different depending on the oxidation conditions. The nature of the beads as well as the binding capacity could also influence the enrichment. Although there were no effects of unspecific binding observed in the previous experiment for hydroxylamine beads in the test with synthetic glycopeptides and the harsh washing protocol for the experiment, the solid support of both beads are different: the hydrazine beads have hydrophilic macroporous polymers[339] as solid support while the hydroxylamine beads are made from agarose[340]. Moreover, for both bead types, the reactive group concentration were different: the hydrazine beads, according to the vendor, present a minimum of 10µmol of activated groups/mL drained medium, while the hydroxylamine beads have 1µmol of activated groups/mL drained medium. Although the difference of active groups was compensated for the experiment, the amount of active groups in both bead types could not be exactly the same. In the case of hydrazine beads, a minimum amount of 250nmol of active groups for

300µg proteins was used and 150nmol of active groups of hydroxylamine beads for 300µg proteins were used in parallel.

Analyzing both conditions separately, for the 5 replicates enriched using the hydrazine beads, the average standard deviation was 0,18 and the average coefficient of variance was 16%. For the 5 replicates enriched using the hydroxylamine beads, the average standard deviation was 0,22 and the average coefficient of variance was 19%. The values of the standard deviation of the ratio between both enrichment methods (0,76) was compared with the average standard deviation for each method separately (0,18 and 0,22). The individual average standard deviation values are drastically lower compared to the former standard deviation, corroborating how different both enrichment methods are in fact.

The putative hypothesis that other glycan residues besides sialic acid could be oxidized and enriched was evaluated in a small scale. Peptides identified exclusively in both conditions were evaluated manually, once there was 1 208 peptides significantly ($p<0,05$) enriched by the hydrazine beads while 569 were peptides significantly enriched by hydroxylamine beads. Peptides were considered exclusively enriched by each beads when the hits were found in at least three of five replicates in each condition (Table 16). The glycosylation of these peptides were analyzed using two different databases, Uniprot and UniCarb KB. In case the hypothesis was true, that additional glycan residues were oxidized, then the *N*-glycosylation of peptides found exclusively by using hydrazine beads should include glycans also without sialic acids.

Table 16. List of identified peptides considered exclusive from hydrazine or hydroxylamine beads enrichment.

Glycopeptides from hydrazine beads				Glycopeptides from hydroxylamine beads		
Protein	Sequence	Glycan	Composition	Protein	Sequence	Glycan
Q06828	VPNNALEGLENLTALYLQHNEIQEVGSSMR	no info		Q6UXH1	NANGSYTCEECDSSCVGCTGEGPGNCK	no info
Q92823	SSAVYQCNASNEYGYLLANAFVNVLAEPPR	no info		P05556	KENSSEICSNNGECVCGQCVCR	no info
Q96KG7	WGPNCSLPCYCK	no info		O60462	TGSEDCSKNFTSPNGTIESPGFPEK	no info
P08294	AVVVHAGEDDLGRGGNQASVENGNAGR	no info		P07942	CVCNYLGTVQEHCNGSDCQCDK	no info
Q7Z3B1	DYSLQIQNVDVTDDGPYTCSVQTQHTPR	no info		Q07954	QSGDVTCNCTDGR	no info
P01210	DAEEDDSLANSSDLLKELLETGDNR	no info		Q8N2S1	CTPACDPGYQPTPGGGCQDVDECRNR	no info
Q14669	MADPESNQEAVNSSAAR	no info		Q14508	TGVCPELQADQNCTQECVSDSECADNLK	no info
Q9P0K1	LFEFSLDDLPSEFQQVNITPSK	no info				
P42263	VMHGGANITGFQIVNNENPMVQQFIQR	no info				
Q15818	GGPRNDTEER	no info				
Q7Z7M0	AGNCSEAACGAADCEQCTR	no info				
Q8IWV2	AYNSAGTGPSSATVNVTTR	no info				
Q14515	KAENSSNEEETSSEGNMR	no info				
P01009	QLAHQSNSTNIFFSPVSIATAFAM	complex [306, 341]	Hex5HexNAc4			
P05155	ASSNPNATSSSSQDPESLQDR	complex [211, 342]	Hex5HexNAc4; dHex1Hex5HexNAc4; Hex6HexNAc5; dHex1Hex6HexNAc5			
P00450	ELHHLQEQN	complex [342-345]	HexNAc5Hex6NeuAc3Fuc1; HexNAc5Hex6NeuAc3Fuc2; HexNAc6Hex7NeuAc4Fuc1; HexNAc4Hex5NeuAc2; HexNAc5Hex6NeuAc3; HexNAc4Hex5NeuAc2Fuc1; HexNAc6Hex7NeuAc4			

The number of peptides considered exclusive for each bead type according to the criteria explained above was not statistically significant compared to the entire number of identified peptides: 17 peptides were exclusively found using hydrazine beads while 7 peptides were

exclusively identified using hydroxylamine beads out of 2 963 total peptides identified by both enrichment methods together. However, information of previously identified sites in the two databases (Uniprot and UniCarb KB) here used were in principle absend (only 3 known glycosylation sites) regarding glycosylation sites identified in this study. Furthermore, the peptides with information found regarding the glycosylation, all of them presented sialic acid in the structure. These glycosylation sites were identified after hydrazine enrichment or lectin enrichment and released using PNGase F, with exception of the of the Nilsson and coworker procedure[306]. This observation reinforces the need of more methodologies to study the complete glycopeptide structures.

Samples released under TCEP treatment were analyzed by nano-LC-MS/MS using a Fusion Lumos Tribrid mass spectrometer. MS/MS experiments used HCD with normalized collision energy of 30%, ETD supplemented by 33% HCD activation (EThcD) in *m*/*z* range 90-2 000, and HCD with normalized collision energy of 40% in *m*/*z* range 800-2 000. The 3 most intense ions from this last scan were selected for further fragmentation using HCD with normalized collision energy of 33%. For data analysis, the raw files were converted into mgf files, using Proteome Discovery 2.1, retrieving the information individually for the MS/MS experiments HCD 30% NCE and EThcD. Mgf files were searched against the Uniprot sequence of human using Byonic™ as search engine and the following data analysis was executed as described in previous sessions. Searches with the *O*- and *N*-glycan database were performed separately for each one of the 5 replicates and files were then combined and processed in order to get the unique number of glycopeptides. A summary of the identification per sample is described in Table 17 and Table 18, furthermore detailed information about each identified glycopeptide is described in the Table A 15.

Table 17. Number of peptide and glycopeptides identified for all replicates for enrichment with hydroxylamine beads and release of glycopeptide structures.

Replicate	Unique proteins	Total of unique glycopeptides[a]	Total of glycopep-tide structures[b]	*O*-glycopeptides structures	*N*-glycopeptides structures
1	59	213	274	140	134
2	43	163	211	116	95
3	42	158	215	122	93
4	46	166	217	128	89
5	47	164	208	122	86
Average	47	173	225	126	99
Total	80	350	473	251	222

a) number of unique glycopeptides identified without considering the difference of the Neu5Ac form, oxidized with or without tag.

b) number of glycopeptide structure considering the difference of the Neu5Ac form, oxidized with or without tag.

Table 18. Protein assession number from Uniprot and unique peptide sequence of glycopeptides identified in all replicates. Peptides in black were identified as *O*-glycopeptides and peptides in blue were identified as *N*-glycopeptides.

Protein number	Peptide Sequence	Protein number	Peptide Sequence	Protein number	Peptide Sequence
A0M8Q6	VTHEGSTVEK	P02768	RPCFSALEVDETYVPK	P41222	SVVAPATDGGLNLTSTFLR
A5YKK6	VQEVIFGLALLNSSSSDLR	P02768	PADLPSLAADFVESK	P41222	SVVAPATDGGLNLTSTFLRK
O00459	INRTQAEEMLSGK	P02768	NDEMPADLPSLAADFVESK	P41222	NLTSTFLR
O00533	NATKLR	P02787	CGLVPVLAENYNK	P41222	AAPEAQVSVQPNFQQDK
O15240	APLPPPAPSQFQAR	P02787	QQQHLFGSNVTDCSGNFCLFR	P41222	APATDGGLNLTSTFLR
O60888	QVTESVSDSITVLP	P02787	APNHAVVTR	P41222	VAPATDGGLNLTSTFLR
O94804	SDPPTLLTPSK	P02787	QQHLFGSNVTDCSGNFCLFR	P41222	CKSVVAPATDGGLNLTSTFLR
P00450	ENLTAPGSDSAVFFEQGTTR	P02790	TPLPPTSAHGNVAEGETKPDPDVTER	P41222	LSMCKSVVAPATDGGLNLTSTFLR
P00734	GHVNITR	P02790	ALPQPQNVTSLLGCTH	P51689	ALQWNAGSGGLPENETTFAR
P00738	VVLHPNYSQVDIGLIK	P02790	LPPTSAHGNVAEGETKPDPDVTER	P51693	GVEYVCCPPPGTPDPSGTAVGDPSTR
P00747	IPSCDSSPVSTEQLAPTAPPELTPVVQDCYHGDGQSYR	P02790	PAVGNCSSALR	P78423	LGVLITPVPDAQAATR
P01009	NATAIFFLPDEGK	P02790	PPTSAHGNVAEGETKPDPDVTER	P78423	LGVLITPVPDAQAATRR
P01009	YLGNATAIFFLPDEGK	P02790	SWPAVGNCSSALR	Q00013	PLNTVTEDMYTNGSPAPGSPAQVK
P01009	LGNATAIFFLPDEGK	P04004	NGSLFAFR	Q02127	FTSLGLLPR
P01042	DIPTNSPELEETLTHTITK	P04196	SSTTKPPFKPHGSR	Q12766	AIEVSSPLPDVLNATEPLSTAQR
P01344	DVSTPPTVLPDNFPR	P04217	SSTSPDR	Q13822	AIIANLTCK
P01857	TKPREEQYNSTYR	P04632	GGGGGGGGGGLGGGLGNVLGGLISGAGGGGGG GGGGGGGGGGGGGGTAMR	Q14515	VHENENIGTTEPGEHQEAK
P01857	EEQYNSTYR	P04632	GGGGGGGGGGLGGGLGNVLGGLISGAGGGGG GGGGGGGGGGGGGGTAMR	Q14515	EHANSKQEEDNTQSDDILEESDQPTQV SK
P01859	TKPREEQFNSTFR	P05060	GEAGAPGEEDIQGPTKADTEK	Q14515	EKVHENENIGTTEPGEHQEAK
P01859	EEQFNSTFR	P05067	GLTTRPGSGLTNIK	Q14624	IPKPEASFSPR
P01860	SCDTPPPCPR	P05090	ADGTVNQIEGEATPVNLTEPAKLEVK	Q14624	QTPAPIQAPSAILPLPGQSVER
P01860	TKPREEQYNSTFR	P05154	VEDLHVGATVAPSSR	Q29983	RTVPPMVNVTR
P01861	TKPREEQFNSTYR	P05155	VATTVISK	Q5FBB7	LSLSPK
P01876	PALEDLLLGSEANLTCTLTGLR	P05452	EPPTQKPK	Q5VU65	PQSIVHFSISNQTVAVVNR
P01876	LSLHRPALEDLLLGSEANLTCTLTGLR	P07339	YSQAVPAVTEGPIPEVLK	Q7Z2W9	LEKVLLVGADNFTLLGK
P02649	KVEQAVETEPEPELR	P0C0L4	GSVTGSQSNAVSPTPAPR	Q7Z7D3	TGAFSMPEVNVDYNASSETLR
P02649	VQAAVGTSAAPVPSDNH	P10643	ENPLTQAVPK	Q8TAG5	THSTSSPQVVAK
P02649	AATVGSLAGQPLQER	P10645	AEGNNQAPGEEEEEEEEATNTHPPASLPSQK	Q8WVZ7	NNTASSWLMK
P02649	QQTEWQSGQR	P10645	EEEEEEEEEAEAGEEAVPEEEGPTVVLNPHPSLGYK	Q8WXD2	LFPAPSEK
P02649	VRAATVGSLAGQPLQER	P10909	EIRHNSTGCLR	Q8WZ92	VILFMIILSGNLSIIILIR
P02649	AGLVEKVQAAVGTSAAPVPSDNH	P10909	EPQDTYHYLPFSLPHR	Q92520	SALDTAARSTKPPR
P02649	VGSLAGQPLQER	P10909	HNSTGCLR	Q92823	FNHTQTIQQK
P02656	DLDPEVRPTSAVAA	P14209	DGGFDLSDALPDNENKKPTAIPK	Q96LW1	LAPSQR
P02679	VDKDLQSLEDILHQVENK	P19827	RTFVLSALQPSPTHSSSNTQR	Q96RD1	NFSILEISFTTVSIPK
P02749	VYKPSAGNNSLYR	P19827	TFVLSALQPSPTHSSSNTQR	Q9BYH1	DALPEGDASPLGPYLLPSGAPER
P02763	NEEYNK	P20929	NATQILNEK	Q9H1Z8	EAPVPTKTK
P02765	AALAAFNAQNNGSNFQLEEISR	P20929	NATDIASQIK	Q9P265	RTELTAATGER
P02765	VCQDCPLLAPLNDTR	P24592	ESKPQAGTARPQDVNR	Q9UH17	LAEFLSEHPNVTLTISAAR
P02765	TVVQPSVGAAAGPVVPPCPGR	P24592	ESKPQAGTARPQDVNRR	Q9UHG2	RLETPAPQVPAR
P02768	DLPSLAADFVESK	P25311	FGCEIENNR	Q9UHG2	GLSAASPPLAETGAPR
P02768	MPADLPSLAADFVESK	P28906	ASVNRGAQENGTGQATSR	Q9UHG2	LETPAPQVPAR
P02768	PSLAADFVESK	P36955	QNPASPPEEGSPDPDSTGALVEEEDPFFKVPVNK	Q9UHG2	LETPAPQVPARR
P02768	DKSLHTLFGDK	P41222	APEAQVSVQPNFQQDK	Q9UHI8	LATAAPGEKPPAPLQFHLLR
P02768	VENDEMPADLPSLAADFVESK	P41222	KSVVAPATDGGLNLTSTFLR	Q9Y6R7	VVTVAALGTNISIHKDEIGK
P02768	ADLPSLAADFVESK				

An average of 173 unique glycopeptides was identified among an average of 47 proteins. The number of identified *O*- and *N*-glycopeptides has a similar proportion as the identified hits in the previous section for the serum and CSF sample. MS limitations regarding the fragmentation of *N*-glycosylated peptides seemed to have a big impact in its identification. Therefore the experiment with MS^3 was set up together in order to increase the *N*-glycopeptides identification however a reliable workflow for data analysis is still under development.

Table 19. List of unique peptide sequences of identified as *N*-glycopeptides obtained by hydroxylamine beads enrichment and chemical release compared to the sequence obtained via enzymatic release and literature.

Protein number	Peptide sequence	Glycosylation site	List of 2963 peptides	Literature (Uniprot and UniCarb KB)
A5YKK6	VQEVIFGLALLNSSSSDLR	N135	no	no
O00459	INRTQAEEMLSGK	N628	no	no
O00533	NATKLR	N513	no	no
P00450	ENLTAPGSDSAVFFEQGTTR	N397	yes	yes
P00734	GHVNITR	N121	no	yes
P00738	VVLHPNYSQVDIGLIK	N241	yes	yes
P01009	LGNATAIFFLPDEGK	N271	yes	yes
P01009	NATAIFFLPDEGK	N271	yes	yes
P01009	YLGNATAIFFLPDEGK	N271	yes	yes
P01857	EEQYNSTYR	N180	yes	yes
P01857	TKPREEQYNSTYR	N180	yes	yes
P01859	EEQFNSTFR	N176	yes	yes
P01859	TKPREEQFNSTFR	N176	yes	yes
P01860	TKPREEQYNSTFR	N227	yes	yes
P01861	TKPREEQFNSTYR	N177	yes	yes
P01876	LSLHRPALEDLLLGSEANLTCTLTGLR	N144	yes	yes
P01876	PALEDLLLGSEANLTCTLTGLR	N144	no	yes
P02679	VDKDLQSLEDILHQVENK	N78	no	yes
P02749	VYKPSAGNNSLYR	N162	yes	yes
P02763	NEEYNK	N56	no	yes
P02765	AALAAFNAQNNGSNFQLEEISR	N176	yes	yes
P02765	VCQDCPLLAPLNDTR	N156	yes	yes
P02787	CGLVPVLAENYNK	N432	yes	yes
P02787	QQHLFGSNVTDCSGNFCLFR	N630	yes	yes
P02787	QQQHLFGSNVTDCSGNFCLFR	N630	yes	yes
P02790	ALPQPQNVTSLLGCTH	N453	yes	yes
P02790	PAVGNCSSALR	N187	yes	yes
P02790	SWPAVGNCSSALR	N187	yes	yes
P04004	NGSLFAFR	N169	yes	yes
P05090	ADGTVNQIEGEATPVNLTEPAKLEVK	N98	yes	yes
P10909	EIRHNSTGCLR	N291	yes	yes
P10909	HNSTGCLR	N291	yes	yes
P20929	NATDIASQIK	N6001	no	no
P20929	NATQILNEK	N6039	no	no
P25311	FGCEIENNR	N128	yes	yes
P28906	ASVNRGAQENGTGQATSR	N362	no	no
P41222	APATDGGLNLTSTFLR	N78	yes	yes
P41222	CKSVVAPATDGGLNLTSTFLR	N78	no	yes
P41222	KSVVAPATDGGLNLTSTFLR	N78	yes	yes
P41222	LSMCKSVVAPATDGGLNLTSTFLR	N78	no	yes
P41222	NLTSTFLR	N78	no	yes
P41222	SVVAPATDGGLNLTSTFLR	N78	yes	yes
P41222	SVVAPATDGGLNLTSTFLRK	N78	yes	yes
P41222	VAPATDGGLNLTSTFLR	N78	yes	yes
P51689	ALQWNAGSGGLPENETTFAR	N128	no	yes
Q00013	PLNTVTEDMYTNGSPAPGSPAQVK	N50	no	no
Q12766	AIEVSSPLPDVLNATEPLSTAQR	N918	no	no
Q13822	AIIANLTCK	N411	yes	yes
Q29983	RTVPPMVNVTR	N210	no	yes
Q5VU65	PQSIVHFSISNQTVAVVNR	N1120	no	no
Q7Z2W9	LEKVLLVGADNFTLLGK	N140	no	no
Q7Z7D3	TGAFSMPEVNVDYNASSETLR	N160	no	no
Q8WVZ7	NNTASSWLMK	N11	no	no
Q8WZ92	VILFMIILSGNLSIIILIR	N38	no	no
Q92823	FNHTQTIQQK	N223	yes	yes
Q96RD1	NFSILEISFTTVSIPK	N63	no	no
Q9UH17	LAEFLSEHPNVTLTISAAR	N113	no	no
Q9Y6R7	VVTVAALGTNISIHKDEIGK	N1743	yes	yes

The lists of peptides obtained in both bead release methods, enzymatically with PNGase F and chemically with TCEP, were compared. A total of 58 unique identified peptide sequences as *N*-glycopeptides were compared against 2 963 peptides obtained via enzymatic cleavage (Table 19). Among them, 34 peptides were found in both identification lists. Most of the sequences that were not identified between both release strategies were also not reported in the databases, Uniprot and UniCarb KB. In this way the analysis of chemically released samples can be validated with the analysis of released enzymatically peptides. Nevertheless, this study showed the versatility of the enrichment method using cleavable hydroxylamine beads. The use of two different bead release methods enables both the classical study of glycosylation sites with enzymatic PNGase F cleavage and/or the identification of glycan heterogeneity on intact glycopeptides by using bead release under reducing conditions.

5.7- Summary and outlook

In this project, the methodology based on glycan chemical properties for enrichment of glycopeptides/glycoproteins was improved by introducing a cleavable disulfide linker on the bead surface (compounds **35** and **52**). Through oxime ligation the hydroxylamine on the linker beads reacted with aldehyde groups obtained by selective oxidation of sialylated glycopeptides. Alternatively, terminal galactose and *N*-acetylgalactose residues were oxidized enzymatically by galactose oxidase prior enrichment. After covalent attachment and washing, the synthetic model glycopeptides were released under reductive conditions cleaving the disulfide bond followed by *in-situ* capping of the free thiol or labeling with a tag for quantitative detection.

This approach has the advantage to recover the entire glycan structure on the peptide backbone containing a small linker modification on the terminal monosaccharide residue (either sialic acid or galacose). Modified hydroxylamine beads were evaluated with synthetic glycopeptides and the methodology was compared against the well-established hydrazine enrichment methodology often used for determination of *N*-glycosylation sites. Furthermore the approach is compatible with addition of a tandem mass tag as presented herein. TMT-labeling is aimed to be later applied in proteomic studies using multiple samples. The enrichment of glycopeptides from isolated proteins, fetuin and asialofetuin, demonstrated the applicability of our methodology using model protein samples. For the fetuin sample, a total of 28 unique glycopeptides were identified. Among them, 7 were unique *O*-glycopeptides and 21 were unique *N*-glycopeptides. These peptides covered 2 *O*- glycosylation sites from fetuin A and B individually, and 2 *N*-glycosylation sites from fetuin A. On the other hand, 98 unique peptides were identified for asialofetuin. Herein, 11 unique *O*-glycopeptides covers 1 fetuin A glycosylation site and 2 fetuin B glycosylation sites. While 87 unique *N*-glycopeptides covers 2 glycosylation sites from each fetuin A and B. The larger number of peptides identified in asialofetuin sample could be explained by an observed incomplete removal of sialic acid residues in the procedure from the vendor with a more structurally diverse sample as a result. Some remaining sialylated *N*-glycopeptide were observed during the data analysis. Moreover the incomplete terminal oxidation by the enzymatic protocol generated different unique

glycopeptide hits, especially among *N*-glycopeptides. These reasons lead to an increase of the sample complexity.

Analysis using complex samples as serum and CSF were also evaluated. Initially, in collaboration with Dr. Jonas Nilsson from Gothenburg University, the enrichment of sialylated glycoproteins using cleavable hydroxylamine beads was demonstrated. In serum a total of 34 proteins and 162 unique glycopeptides were identified by the new methodology. While in CSF 32 proteins and 129 unique glycopeptides were identified. Later on a more complex comparison between the enrichment methods protocols were executed in collaboration with Dr. Astrid Guldbrandsen and Prof. Dr. Frode Berven from the Proteomics Unit at the University of Bergen. In these experiment clinical CSF samples with 5 replicates per condition was used in the comparison between the standard enrichment method using commercial hydrazine beads and enzymatic release with PNGase F with the method using cleavable hydroxylamine beads and chemical release. Statistical analysis of enzymatically cleaved samples showed that both enrichment protocols (using hydrazine and hydroxylamine beads) are too different to be compared: oxidation procedure, quantity of active groups and solid support in each beads type might influence in the overall enrichment protocol. On the other hand the comparison of different cleavage procedures indicated that both methods are complementary. Important and different information from both approaches can be obtained: enzymatic PNGase peptide release gives an overview about the glycosylation sites and chemical bead release by TCEP reduction identifies glycan heterogeneity of the identified glycopeptides. Since a lack of information regarding glycan heterogeneity of the identified glycosylation sites was observed in the available databases, the combined methods for bead peptide enrichment and relase also allowed cross-validation of identified glycopeptide hits within the same experiment.

Enrichment of non-sialylated peptides using the disulfide linker cleavage approach was also demonstrated using asialofetuin as model protein. Although optimization in the oxidation procedure is further required in order to reduce sample complexity for data analysis, the evaluation of complex samples as serum and CSF are currently under analysis.

Improvement especially regarding analysis and identification of entire *N*-glycopeptides is still necessary: nowadays multiple techniques are required for determination of peptide sequence, glycan structure and glycosylation site. Even so, the enrichment with cleavable hydroxylamine beads method presented the versatility of being suitable for the release of both structures: entire intact *N*-glycopeptides or enzymatic glycan relase for identification of peptide glycosylation sites. According to this approach, the *N*-glycopeptide analysis could be validated in two different levels in parallel. The new approach also allows structural analysis of intact *O*-glycopeptides, in this case no suitable enzyme is available for efficient *O*-glycan release and *O*-glycan site determination in the manner that can be made for *N*-glycopeptides using PNGase F. Furthermore, besides the compatibility with available tandem mass tags, which allows quantification, this approach has the possibility to increase the glycan selectivity scope if applied with mutant enzymes specific for oxidation of terminal mannose and *N*-acetylglucosamine[346].

Experimental Procedures

6.1- General methods

Reagents were purchased with high commercial quality and all solvents were purchased as quality grade *pro analysi* (p.a.). Dichoromethane was dried by refluxing over CaH_2 and subsequent distillation under argon. Dimethylformamide was purchased anhydrous with a purity of ≥99,8%. No further purification unless otherwise stated. Solvent concentrations were performed at a reduced pressure (bath temperature <50°C). Synthesis of all compounds is described herein.

TLC was performed on Silica *Kieselgel 60 F_{254}* (Merck, Darmstadt, Germany) with detection by UV-light (254nm) and by staining with $KMnO_4$ or *p*-anisaldehyde solution and visualized by heat exposure. Column chromatography was performed on silica *Kieselgel 60* (0,04–0,063nm) from *Carl Roth GmbH* (Karlsruhe).

NMR spectra were recorded at 295K with the following instruments:

- Varian Mercury 400 (*Agilent*);

- Advance DRX 500 (*Bruker*);

- Avance III HD 700 (*Bruker*).

The residual solvent signals were used to calibrate the spetra[347]:

- $CDCl_3$ (internal $CHCl_3$, δ_H 7,26 ppm; δ_C 77,23 ppm);

- DMSO-d_6 (internal DMSO, δ_H 2,50 ppm; δ_C 39,51 ppm).

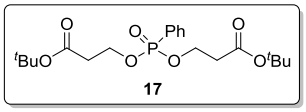
The HR-ESI-MS spectra were recorded with an *LTQ-FT ICR* ultra mass spectrometer or an *LTQ Orbitrap XL* mass spectrometer (both *Thermo Scientific*).

Analytical reverse phase HPLC was performed on a Dionex ultimate 3000 (*Thermo Scientific*) system (DR-3600 six channel degasser, LPG-3x00 pump, TCC-3100 column compartiment, DAD-3000 UV/VIS diode array detector, WPS-3000 autosampler) and *Luna C18(2)* column (3µm, 100Å, 150x2,0mm) from *Phenomenex*. Dionex ultimate 3000 LC system (HPG-3200P, VWD-3400, UV/Vis detector, AFC-3000 sampler) was used for semi-preparative and preparative measurements. *InertSustain C18* (5µm, 250x6,0mm) column from *GL Sciences Inc.* was used for semi-preparative measurements and *Luna C18(2)* column (10µm, 100Å, 250x21,2mm) from *Phenomenex* was used for preparative measurements.

6.2- Development of MS-cleavable cross-linker

6.2.1- Synthesis of MS-cleavable cross-linker

6.2.1.1- Preparation of di-*tert*-butyl 3,3'-((phenylphosphoryl)bis(oxy))dipropionate (17)



17

tert-Butyl 3-hydroxypropionate **16** (192,0µL, 190,0mg, 1,30mmol) was dissolved in a solution containing DIPEA (226,0µL, 168,0mg, 1,30mmol) and 3% tetrazole in acetronitrile (1,04mL). Phenylphosphonic dichloride **15** (73,7µL, 101,4mg, 0,52mmol) was added slowly to the previous solution and the reaction mixture was stirred overnight at room temperature. The solvent was removed, the product was purified by flash chromatography (ethyl acetate/cyclohexane 1:1) and 162mg of a colorless oil was obtained.

Yield: 75%

R_f (ethyl acetate:cyclohexane – 1:2): 0,3

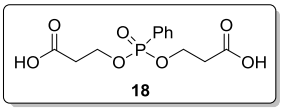
^{1}H-NMR (400 MHz, $CDCl_3$): δ = δ = 7,81 – 7,75 (m, 2H); 7,57 – 7,53 (m, 1H); 7,47 – 7,42 (m, 2H); 4,32 – 4,22 (m, 4H); 2,60 (t, *J* = 6,6 Hz, 4H); 1,42 (s, 18H).

^{13}C-NMR (101 MHz, $CDCl_3$): δ = 169,6; 132,7; 132,0; 131,9; 128,7; 128,5; 126,8; 81,2; 62,0; 61,9; 37,0; 36,9; 28,2.

^{31}P-NMR (162 MHz, $CDCl_3$): δ = 20,10.

HRMS (ESI/+): *m*/*z* calculated for **$C_{20}H_{31}O_7P$** $[MH]^+$: 415,1885; obtained $[MH]^+$: 415,1884; error: 0,2ppm.

6.2.1.2- Preparation of 3,3'-((phenylphosphoryl)bis(oxy))dipropionic acid (18)

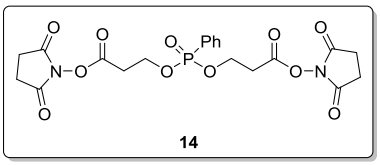


18

Compound **17** (136,0mg, 0,33mmol) was dissolved in a mixture 1:1 TFA/DCM (3,3mL). When the starting material was completely consumed, the solvent was removed under vacuum and the remaining TFA was removed by coevaporation with toluene. The obtained oil was used without further purification for next reaction.

R_f (ethyl acetate:methanol:acetic acid:water – 40:3:3:2): 0,6

6.2.1.3- Preparation of bis(2,5-dioxopyrrolidin-1-yl) 3,3'-((phenylphosphoryl)bis(oxy)) dipropionate (14)



Compound **18** (160mg, 0,53mmol) and *N*-hydroxysuccinimide (183mg, 1,60mmol) were dissolved in 1,6mL of DMF and the solution was cooled to 0°C. DIC, *N,N'*-diisopropylcarbodiimide, (246,0µL, 200,4mg, 1,60mmol) was added carefully to the cold solution. The mixture was stirred for 1h at 0°C. After this time the temperature was raised to room temperature and the mixture was stirred overnight. The solvent was removed, the product was purified by flash chromatography (ethyl acetate:methanol:acetic acid:water – 200:3:3:2) and 95,2mg of a viscous colorless oil **14** was obtained.

Yield: 36% (over 2 steps)

R_f (ethyl acetate:methanol:acetic acid:water – 100:3:3:2): 0,6

^{1}H-NMR (700 MHz, DMSO): δ = 7,77 – 7,73 (m, 2H); 7,67 – 7,64 (m, 1H); 7,55 – 7,52 (m, 2H); 4,28 – 4,25 (m, 4H); 3,14 – 3,12 (m, 4H); 2,82 (ls, 8H).

^{13}C-NMR (176 MHz, DMSO): δ = 170,0; 166,6; 132,9; 131,4; 131,3; 128,8; 128,7; 127,5; 126,4; 60,7; 32,1; 32,0; 25,4; 25,2.

^{31}P-NMR (283 MHz, $CDCl_3$): δ = 18,18.

HRMS (ESI/+): *m/z* calculated for **$C_{20}H_{22}N_2O_{11}P$** $[MH]^+$: 497,0956; obtained $[MH]^+$: 497,0955; error: 0,2ppm.

6.2.2- Linker test – Peptide level

6.2.2.1- Preparation of peptides

Solid-phase peptide synthesis, SPPS, followed the protocol published by Pett *et. al.*[84] The peptides were synthesized by stepwise SPPS by using the Fmoc strategy starting with a preloaded Fmoc-proline, Fmoc-alanine or Fmoc-arginine trityl resin (13mmol scale per peptide). Fmoc-protected amino acids were coupled automatically on a Multisyntech peptide synthesizer by using eight equivalents of the amino acid and HBTU/HOBt. Fmoc deprotection was performed according to standard conditions, 20% piperidine in DMF. Acetylation of *N*-terminus was performed before the final wash and resin cleavage with a solution of 500mM Ac_2O, 15mM HOBt and 125mM DIPEA in DMF. The peptides were then released from the resin, and all acid-sensitive side chain protecting groups were simultaneously removed by treatment with TFA/TIPS/H_2O (90:5:5) followed by solvent concentration, lyophilization, and purification by using a C-18 cartridge (1 g of C-18 material, Waters).

IEAEKGR (30**):**

Analytical HPLC r.t. : 12,95min (A/B: (95:5) → (70:30), 200µL/min, 50min)

Preparative HPLC r.t. : 7,09min (A/B: (95:5) → (55:45), 20mL/min, 32min)

HRMS (ESI/+): *m*/*z* calculated for $[M+2H]^{2+}$: 401,7245; obtained $[M+2H]^{2+}$: 401,7243; error: 0,5ppm.

Ac-IEAEKGR (24**):**

Analytical HPLC r.t. : 15,72min (A/B: (95:5) → (40:60), 200µL/min, 55min)

Preparative HPLC r.t. : 10,49min (A/B: (95:5) → (55:45), 20mL/min, 32min)

HRMS (ESI/+): *m*/*z* calculated for $[M+2H]^{2+}$: 422,7298; obtained $[M+2H]^{2+}$: 422,7297; error: 0,2ppm.

Ac-LDEVKEQVAEVR (25**):**

Analytical HPLC r.t. : 28,48min (A/B: (95:5) → (40:60), 200µL/min, 55min)

Preparative HPLC r.t. : 16,36min (A/B: (95:5) → (55:45), 20mL/min, 32min)

HRMS (ESI/+): *m*/*z* calculated for $[M+2H]^{2+}$: 728,8857; obtained $[M+2H]^{2+}$: 728,8858; error: 0,1ppm.

Ac-VVVDALSGLKGDLAGR (26**):**

Analytical HPLC r.t. : 39,30min (A/B: (95:5) → (40:60), 200µL/min, 55min)

Preparative HPLC r.t. : 21,57min (A/B: (95:5) → (55:45), 20mL/min, 32min)

HRMS (ESI/+): *m*/*z* calculated for $[M+2H]^{2+}$: 806,4569; obtained $[M+2H]^{2+}$: 806,4572; error: 0,4ppm.

6.2.2.2- Cross-linked peptide preparation

A solution 33mM peptide in PBS buffer pH7,0 was mixed with a solution 20mM PXL **14** in MeCN in a ratio 1:1 v/v. After 3 hours, the reaction was purified using semi-preparative HPLC. The obtained solution with the cross-linked peptide was diluted 1:10 000 for further LC-MS/MS measurement.

Cross-linked Ac-IEAEKGR (27**):**

Analytical HPLC r.t. : 27,56min (A/B: (95:5) → (40:60), 200µL/min, 55min)

Semi-preparative HPLC r.t. : 35,15min (A/B: (95:5) → (60:40), 5mL/min, 60min)

HRMS (ESI/+): *m*/*z* calculated for $[M+3H]^{3+}$: 651,9821; obtained $[M+3H]^{3+}$: 651,9824; error: 0,5ppm.

Cross-linked Ac-LDEVKEQVAEVR (28**):**

Analytical HPLC r.t. : 36,65min (A/B: (95:5) → (40:60), 200µL/min, 55min)

Semi-preparative HPLC r.t. : 35,15min (A/B: (95:5) → (50:50), 5mL/min, 70min)

HRMS (ESI/+): *m*/*z* calculated for $[M+3H]^{3+}$: 1060,5245; obtained $[M+3H]^{3+}$: 1060,5260; error: 1,4ppm.

Cross-linked Ac-VVVDALSGLKGDLAGR (29**):**

Analytical HPLC r.t. : 52,20min (A/B: (95:5) → (40:60), 200µL/min, 55min)

Semi-preparative HPLC r.t. : 35,15min (A/B: (95:5) → (70:30), 5mL/min, 5min, A/B: (70:30) → (35:65), 5mL/min, 60min)

HRMS (ESI/+): *m*/*z* calculated for $[M+3H]^{3+}$: 1163,9526; obtained $[M+3H]^{3+}$: 1163,9550; error: 2,1ppm.

Cross-linked IEAEKGR (32**):**

Analytical HPLC r.t. : 44,10min (A/B: (95:5) → (70:30), 200µL/min, 50min)

Semi-preparative HPLC r.t. : 37,95min (A/B: (95:5) → (60:40), 5mL/min, 70min)

HRMS (ESI/+): *m*/*z* calculated for $[M+3H]^{3+}$: 623,9751; obtained $[M+3H]^{3+}$: 623,9750; error: 0,2ppm.

6.2.2.3- LC-MS analysis

Cross-linked Ac-IEAEKGR, Ac-LDEVKEQVAEVR and Ac-VVVDALSGLKGDLAGR peptides were measured with direct injection using robotic nanoflow ion source TriVersa NanoMate (Advion BioSciences) into a LTQ-FT ICR ultra mass (Thermo Scientific). Samples were measured in data-dependent acquisition mode acquiring full MS scans for 5min at a resolution of 100 000 and no dynamic exclusion was used. The 5 most abundant ions (Top5) were selected for further MS/MS

fragmentation in ion trap using CID with normalized collision energy of 35%. First mass was set to 250*m*/*z* for full MS.

15µL of cross-linked Ac-IEAEKGR solution 33nM was also analyzed by nano-LC-MS/MS using an Orbitrap Elite mass spectrometer online-coupled to a nano RSLC HPLC system (both *Thermo Scientific*). LC separation was performed using a linear gradient from 5-50% of B in 72min. Samples were measured in data-dependent acquisition mode acquiring full MS scans at a resolution of 30 000 and dynamic exclusion of 5s. The 5 most abundant ions (Top5) were selected for further fragmentation at MS/MS resolution of 15 000. Target values were set to $1x10^6$ for MS and $5x10^4$ for MS/MS scans. Maximum injection times were 10 and 100ms, respectively. Each selected ion was fragmented using CID with normalized collision energy of 16%, HCD with normalized collision energy of 18%, 20%, 22% and 25%. First mass was set to 100 and 500*m*/*z*, respectively.

15µL of cross-linked IEAEKGR solution 33nM was also analyzed by nano-LC-MS/MS using an Orbitrap Fusion Tribid mass spectrometer online-coupled to a nano RSLC HPLC system (both *Thermo Scientific*). LC separation was performed using a linear gradient from 3-50% of B in 86min. Samples were measured in data-dependent acquisition mode acquiring full MS scans at a resolution of 120 000 and dynamic exclusion of 5s. The ions that presented intensity greater than 5 000 were further fragmented with ETD and activation time of 200ms and CID with normalized collision energy of 35%. Ions obtained in the ETD fragmentation with intensity greater than 5000 were fragmented in a MS^3 level with HCD with normalized collision energy of 32%, first mass is 120*m*/*z*. Target values were set to $5x10^4$ for MS, $1x10^5$ for ETD and $4x10^5$ for CID MS/MS scans, and $5x10^4$ for MS^3.

6.2.2.4- Data analysis

Peptide fragments were calculated using NIST Mass and Fragment Calculator v1.3. The list of the theoretical fragments is indexed in the appendix (Table A 1Table **A 4**). The raw files were visualized via program Xcalibur Qual Browser 2.2. The mass of each expected peptide was manually searched in the full spectrum with a mass range of ±0,5*m*/*z*.

6.2.3- Linker test – Protein level

6.2.3.1- Cross-linker reaction

Bovine serum albumin (BSA) solution 5µM in HEPES buffer 50mM pH7,0 was prepared. A solution of 25µM PXL (**14**) in DMSO was also prepared. Both solutions were mixed following the protein/linker ratio of 1:50 in molarity. 500µL reaction solution were used for monitoring the reaction in two different time points: 30min and 1h. After 30min and after 1h, 2% (10µL) of the reaction mixture were aliquoted for SDS-PAGE analysis and remain reactions were frozen with liquid nitrogen, the solvent was removed with speed-vac and stored at -40°C freezer. Control was also prepared with the same volume and conditions of the reaction except by the presence of linker.

6.2.3.2- SDS-PAGE analysis

NuPAGE™ 4-12% bis-tris protein gel, 1.0mm, 12-well (Invitrogen™ - Thermo Fisher Scientific) was used at neutral pH. 4,5µL of reaction (approximately 1,5µg protein) were mixed with 2,5µL NuPAGE™ 4x LDS buffer (Invitrogen™ - Thermo Fisher Scientific) and water was added to achieve a total volume of 10µL. The samples were heated at 85°C for 15min and cooled back to room temperature. The cooled samples were each filled into gel wells. One well was filled with 3µL protein marker (Mark12™ Unstained Standard - Invitrogen™ - Thermo Fisher Scientific) for the molecular weight determination of separated proteins.

The chamber used is an Xcell SureLock MiniCell Electrophoresis System (Thermo Fisher Scientific) with power supply EPS3501 XL (Pharmacia Biotech). MOPS buffer solution (50mL NuPAGE™ 20x MOPS buffer (Invitrogen™ - Thermo Fisher Scientific) and 950mL water) was used to fill the running chamber. The gel ran at 4°C with a two-step voltage: firstly 50 volts for 20 minutes and later 200 volts for 45-70min – or until the blue band almost reached the end of the gel. Afterward, the gel was incubated for 1h in a solution of 30% ethanol and 10% acetic acid in water to fix the proteins. Before the staining, the gel was briefly washed twice with water.

Coomassie staining solution was used for protein visualization. This solution was prepared with 18,4g (0,3mol) of aluminum sulfate hexadecahydrate firstly dissolved in 60ml of water. 20mL of ethanol were added and mixed very well. Subsequently, 40mg of Coomassie Brilliant Blue G-250 (Merk) were added to form a suspension. Afterward, 4ml of phosphoric acid (85%) were added and the mixture was diluted with water up to 200ml solution. The gel was incubated in 100mL of solution for several hours with gentle shaking until the bands were clearly visible. If necessary, the gel could be decolorized with 8% acetic acid solution.

6.2.3.3- Digestion protocol

Dried reaction mixtures were resolubilized with 50µL 4M guanidinium chloride in 50mM HEPES buffer pH7,0. DTT (5,56µL) 100mM stock solution in water was added and the mixture reacted for 30 minutes at 57°C. As soon as the mixture achieved room temperature, IAA (6,20µL) 200mM stock solution in water was added and the mixture reacted for more 30 minutes at room temperature in the dark. The solution was then diluted with 939µL water and 0,1µL $CaCl_2$ 1mM were added. Trypsin from Sigma (8,2µg) was added to the solution and it was incubated for 16h at 37°C. 10% TFA in water was added in order to decrease the pH solution to 2,0-2,5. Peptide mixture was purified with Agilent C18 Cartridge 4mg prewashed with acetonitrile (3x100µL) and 0,1% TFA in water (3x100µL). After loading the sample, the cartridge was washed with 0,1% TFA in water (3x100µL). The product was eluted using 60% acetonitrile/water (2x50µL) and dried under vacuum.

6.2.3.4- Optimized cross-linker reaction and digestion protocol

Bovine serum albumin (BSA) solution 5µM in PBS buffer pH6,5 was prepared. A solution of 25µM PXL (**14**) in DMSO was also prepared. Both solutions were mixed following the protein/linker ratio of 1:50 in molarity. 500µL reaction solution were used for monitoring the reaction in two different time points: 30min and 1h. After 30min and after 1h, 2% (10µL) of the reaction mixture were aliquoted and remain reactions were frozen with liquid nitrogen, the solvent was removed with speed-vac and stored at -40°C freezer.

Dried reaction mixtures were resolubilized with 20µL 200mM DTT in water and 20µL 2,2,2-trifluoroethanol. The mixture reacted for 20 minutes at 95°C. As soon as the mixture achieved room temperature, IAA (15µL) 200mM stock solution in water was added and the mixture reacted for more 1h at room temperature in the dark. The solution was then diluted with 465µL water. The pH was adjusted to 6 with 2M NaOH prior the addition of 1µL $CaCl_2$ 1mM. Trypsin from Sigma (4,1µg) was added to the solution and it was incubated for 16h at 37°C. 10% TFA in water was added in order to decrease the pH solution to 2,0-2,5. Peptide mixture was purified with Agilent C18 Cartridge 4mg prewashed with acetonitrile (3x100µL) and 0,1% TFA in water (3x100µL). After loading the sample, the cartridge was washed with 0,1% TFA in water (3x100µL). The product was eluted using 60% acetonitrile/water (2x50µL) and dried under vacuum.

6.2.3.5- LC-MS analysis

Samples were resolubilized in 0,1% TFA solution and an aliquot was diluted to 6,7nM solution. 15µL of the last solution were analyzed by nano-LC-MS/MS using a Velos Orbitrap mass spectrometer online-coupled to a nano RSLC HPLC system (both *Thermo Scientific*). LC separation was performed using a linear gradient from 3-45% of B in 90 minutes. Samples were measured in data-dependent acquisition mode acquiring full MS scans at a resolution of 100 000 and dynamic exclusion of 10s. The 5 most abundant ions (Top5) were selected for further MS/MS fragmentation in ion trap using CID with normalized collision energy of 20%. Target values were set to $1x10^6$ for MS and $2x10^5$ for MS/MS scans. First mass was set to 300*m*/*z* for full MS.

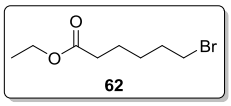
6.2.3.6- Data analysis

Cross-linked BSA peptides identified by Huang and coworkers[87] (table 1 and 2 from the reference) were used as peptide database for the search. The cross-linked peptide masses were adjusted for PXL linker (delta mass +129.9820) and charge states +3 and +4 were calculated (list attached in the appendix Table A 6). The masses of peptide resulted from chemical cleavage were also calculated with charge state 2+ (Table A 7). The raw files were visualized via program Xcalibur Qual Browser 2.2. The masses of each expected peptide was manually searched in the full spectrum with a mass range of ±0,5*m*/*z*.

6.3- New approach for glycopeptide enrichment

6.3.1- Synthesis of modified beads

6.3.1.1- Preparation of ethyl 6-bromohexanoate (62)



6-Bromohexanoic acid **36** (5,0g, 25,63mmol) was dissolved in 52mL of ethanol and 0,8mL of sulfuric acid was added. The mixture was refluxing for 1 hour. After this time, the reaction was cooled to room temperature and the ethanol was removed. The mixture was diluted with 70mL of ether, washed with $NaHCO_3$ 5% (1x 40mL) and followed by brine (1x 40mL). The organic layer was dried with Na_2SO_4 and the solvent was removed. Colorless oil was obtained in 5,2g and the product was used as starting material for next reaction without further purification.

Yield: 92%

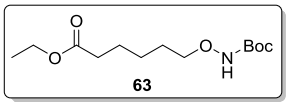
R_f (ethyl acetate:cyclohexane – 1:5): 0,6

^{1}H-NMR (400MHz, $CDCl_3$): δ = 4,12 (q, *J* = 7,2 Hz, 2H); 3,39 (t, *J* = 7,1 Hz, 2H); 2,30 (t, *J* = 7,1 Hz, 2H); 1,86 (qui, *J* = 7,1 Hz, 2H); 1,64 (qui, *J* = 7,1 Hz, 2H); 1,50 – 1,43 (m, 2H); 1,24 (t, *J* = 7,2 Hz, 3H).

^{13}C-NMR (101MHz, $CDCl_3$): δ = 173,6; 60,4; 34,2; 33,6; 32,5; 27,8; 24,2; 14,4.

^{1}H NMR (lit., $CDCl_3$)[348]: δ 4,13 (q, J= 7,1 Hz, 2H); 3,41 (t, J= 6,8 Hz, 2H); 2,32 (t, J= 7,4 Hz, 2H); 1,94 – 1,79 (m, 2H); 1,62 (s, 2H); 1,48 (d, J= 31,5 Hz, 2H); 1,26 (t, J= 7,1 Hz, 3H).

6.3.1.2- Preparation of ethyl 6-(*tert*-butoxycarbonylaminooxy)hexanoate (63)



Boc-hydroxylamine (617,0mg, 4,64mmol) was dissolved in 1,5mL of DMF and DBU (0,75mL, 0,77g, 4,96mmol) was added. After 5 minutes, compound **62** (739,0mg, 3,31mmol) was added and the reaction mixture was stirred overnight at room temperature. When the reaction was over, monitored by TLC, the solvent was removed. The product was purified by flash chromatography (ethyl acetate/cyclohexane 1:5) and 720mg of off-white oil were obtained.

Yield: 79%

R_f (ethyl acetate:cyclohexane – 1:5): 0,2

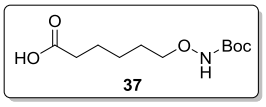
^{1}H-NMR (400 MHz, $CDCl_3$): δ = 7,14 (ls, 1H); 4,11 (q, *J* = 7,1 Hz, 2H); 3,83 (t, *J* = 6,9 Hz, 2H); 2,29 (t, *J* = 6,9 Hz, 2H); 1,68 – 1,60 (m, 4H); 1,51 – 1,38 (m, 12H); 1,24 (t, *J* = 7,1 Hz, 3H).

^{13}C-NMR (101 MHz, $CDCl_3$): δ = 173,8; 157,0; 81,7; 76,6; 60,4; 34,3; 28,4; 27,8; 25,6; 24,9; 14,7.

^{1}H-NMR (lit., $CDCl_3$)[349]: δ = 7,28 (s, 1H); 4,12 (q, 2H); 3,83 (t, 2H); 2,30 (t, 2H); 1,65 (m, 4H); 1,50 (s, 9H); 1,42 (m, 2H); 1,25 (t, 3H).

^{13}C-NMR (lit., $CDCl_3$)[349]: δ = 173,8; 157,1; 81,7; 76,6; 60,4; 34,3; 28,4; 27,8; 25,6; 24,9; 14,4.

6.3.1.3- Preparation of 6-(*tert*-butoxycarbonylaminooxy)hexanoic acid (37)



37

Compound **63** (700,0mg, 2,54mmol) was dissolved in 5mL of ethanol and 10mL of NaOH 16,0M was added. The mixture was stirred until consumption of the starting material. Ethanol was removed and HCl 1,0M was added to get a solution with pH2,0. The aqueous layer was extracted with chloroform (5x 40mL), dried with Na_2SO_4 and the solvent was removed. Viscous brown oil was obtained (323mg). This product was used as starting material for next reaction without further purification.

Yield: 51%

R_f (ethyl acetate:cyclohexane – 1:1): 0,1

^{1}H-NMR (400 MHz, DMSO): δ = 9,91 (ls, 1H); 3,65 (t, *J* = 6,5 Hz, 2H); 2,07 (t, *J* = 7,3 Hz, 2H); 1,51 – 1,43 (m, 4H); 1,39 (s, 9H); 1,32 – 1,26 (m, 2H).

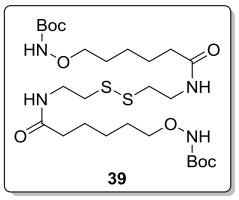
^{13}C-NMR (101 MHz, DMSO): δ = 175,7; 156,1; 79,3; 75,2; 35,4; 28,1; 27,5; 25,3; 25,1.

HRMS (ESI/+): *m/z* calculated for **$C_{11}H_{21}NO_5Na$** $[MNa]^+$: 270,1312; obtained $[MNa]^+$: 270,1315; error: 1,1ppm.

^{1}H-NMR (lit., $CDCl_3$)[350]: δ = 7,20 (s, 1H); 3,85 (t, 2H); 2,37 (t, 2H); 1,66 (m, 4H); 1,48 (s, 9H); 1,44 (m, 2H).

^{13}C-NMR (lit., $CDCl_3$)[350]: δ = 179,3; 157,5; 82,0; 76,6; 34,0; 28,4; 27,8; 25,5; 24,6.

6.3.1.4- Preparation of *tert*-butyl (6,6'-(2,2'-disulfanediylbis(ethane-2,1-diyl)bis(azanediyl)) bis(6-oxohexane-6,1-diyl))bis(oxy)dicarbamate (39)



39

A solution of **37** (195,0mg, 0,79mmol), HBTU (299,0mg, 0,79mmol) and HOBt (120,7mg, 0,79mmol) in 2mL of DMF was prepared. DIPEA (412µL, 305,7mg, 2,37mmol) was added and the solution became yellow. Cystamine dihydrochloride **38** (88,8mg, 0,39mmol) was added and the reaction was stirred overnight. When the reaction was over, monitored by TLC, the solvent

was removed. The product was purified by flash chromatography (ethyl acetate) and 150 mg of yellow oil were obtained.

Yield: 31%

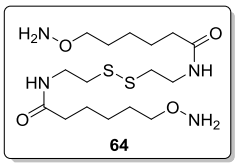
R_f (ethyl acetate): 0,3

^{1}H-NMR (400 MHz, DMSO): δ = 9,86 (ls, 2H); 7,96 (t, *J* = 6,3 Hz, 2H); 3,65 (t, *J* = 6,4 Hz, 4H); 3,31 (q, *J* = 6,4 Hz, 4H); 2,76 (t, *J* = 6,4 Hz, 4H); 2,06 (t, *J* = 7,4 Hz, 4H); 1,51 – 1,45 (m, 8H); 1,39 (s, 18H); 1,30 – 1,26 (m, 4H).

^{13}C-NMR (101 MHz, DMSO): δ = 172,2; 156,1; 79,4; 75,1; 37,9; 37,3; 35,2; 28,0; 27,4; 25,1; 25,0.

HRMS (ESI/+): *m/z* calculated for **$C_{26}H_{51}N_4O_8S_2$** $[MH]^+$: 611,3143; obtained $[MH]^+$: 611,3134; error: 1,5ppm.

6.3.1.5- Preparation of *N,N'*-(2,2'-disulfanediylbis(ethane-2,1-diyl))bis(6-(aminooxy) hexanamide) (64)



64

Compound **39** (95,0mg, 0,16mmol) was dissolved in 900µL of a mixture TFA/dichloromethane 1:1 at room temperature. When the reaction was over, monitored by TLC, the solvent was removed and the residual TFA was coevaporated with toluene (7x). 62,8mg of viscous oil were obtained and used as starting material for next reaction.

Yield: quantitative

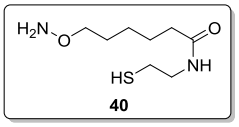
R_f (ethyl acetate:methanol:acetic acid:water – 20:3:3:2): 0,2

^{1}H-NMR (500 MHz, DMSO): δ = 8,02 (t, J = 5,6 Hz, 2H); 3,82 (t, J = 6,5 Hz, 4H); 3,34 – 3,27 (m, 4H); 2,75 (t, J = 6,8 Hz, 4H); 2,06 (t, J = 7,4 Hz, 4H); 1,58 – 1,44 (m, 8H); 1,31 – 1,21 (m, 4H).

^{13}C-NMR (126 MHz, DMSO): δ = 172,2; 74,2; 37,9; 37,3; 35,2; 27,1; 24,9.

HRMS (ESI/+): *m/z* calculated for **$C_{16}H_{35}N_4O_4S_2$** $[MH]^+$: 411,2088; obtained $[MH]^+$: 411,2091; error: 0,7ppm.

6.3.1.6- Preparation of 6-(aminooxy)-*N*-(2-mercaptoethyl)hexanamide (40)



40

Compound **64** (5,0mg, 12,2µmol) was dissolved in 51µL of acetonitrile and 254µL of *tris*(2-carboxyethyl)phosphine (TCEP) 0,2M in water was added. After 2 hours, the product was purified using preparative HPLC and used direct

in the next step.

6.3.1.7- Preparation of magnetic beads (41)

Beads (150mg) were suspended with 5mL of 20% ethanol in water solution. It was dispersed by vigorous vortexing and 1mL (or 30mg) was transferred to an eppendorf tube.

The tube was placed on a magnetic separator for 1-3 minutes. The supernatant was removed while the tube was remaining on the separator. The beads were suspended with 1mL coupling buffer (0,1M sodium phosphate; pH7,0; 5mM EDTA) and vortexed for 30 seconds. This procedure was repeated twice.

6.3.1.8- Preparation of modified magnetic beads (35)

Compound **40** (5,0mg, 24,2µmol) was dissolved in a mixture of 0,2mL acetonitrile and 1,0mL coupling buffer. This solution was incubated with washed beads **41** (30mg) for 2 hours at room temperature with gentle rotation. The beads were washed with 1mL of coupling buffer 4 times, like described in the preparation. The excess of reactive groups were blocked by reaction with 1mL cysteine hydrochloride solution (8mg/L in coupling buffer) for 30 minutes at room temperature with gentle rotation. The beads were washed with 1mL washing buffer (NaCl 1,0M; 0,05% m/v NaN_3) 4 times, after the reaction. Finally, the beads were suspended in 1mL storage buffer (PBS buffer; 0,05% m/v NaN_3) and stored at 4°C.

6.3.1.9- Preparation of modified agarose beads (52)

Lyophilized beads from Aldrich Thiol-Sepharose® 4B **51** (250mg, ca. 1µmol active groups) were swelled with 1mL of water with vigorous vortexing.

The beads were washed with 800µL water (5x). In each washing step, the beads were washed with water, centrifuged for 3min at 1 000rpm and the supernatant was removed. They were further washed (3x) with 800µL PBS buffer, 5mM EDTA degassed in He for 30 minutes (coupling buffer).

Compound **40** (6,0mg, 29,0µmol) was dissolved in a mixture of 100µL acetonitrile and 900µL coupling buffer. This solution was incubated with washed beads (250mg) overnight at room temperature with gentle rotation. The beads were washed with 5x 500µL of water and stored with 20% EtOH/water solution at 4°C.

6.3.2- Beads test – Peptide level

6.3.2.1- Negative charged peptide

Sialylated glycopeptide **42** (0,03mg; 11,9nmol) was oxidized using 87µL $NaIO_4$ solution 20mM for 30 minutes at 4°C. The product was purified with an Agilent C18 Cartridge 4mg prewashed with acetonitrile (3x100µL) and 0,1% TFA in water (3x100µL). After loading the sample, the cartridge was washed with 0,1% TFA in water (3x100µL) and the product was eluted using 60% acetonitrile/water (2x50µL). The solvent was removed in the speed-vac.

Beads (120µL; ca 120nmol active groups) were transferred to an eppendorf tube and washed (5x300µL) with reaction buffer - 0,1M acetate buffer pH4,5; 0,1M NaCl. In each washing step, the beads were washed with buffer, centrifuged for 3min at 1 000rpm and the filtrate was removed. Oxidized peptide was dissolved in 250µL reaction buffer and the solution was transferred to the beads. The mixture was gentle stirred overnight at room temperature.

The solution was separated from the beads via centrifugation for 3min at 1 000rpm and the filtrate was removed. Beads were washed then with the next reaction buffer – HEPES buffer 50mM, EDTA 10mM, pH8,0 (5x 200µL). *Tris*(2-carboxyethyl)phosphine (TCEP) solution 10mM in buffer (200µL) was added to the beads and the reaction was stirred for 1 hour at 60^{o}C. The reaction mixture was cooled to room temperature and 200µL iodoacetoamide solution 30mM in buffer was added. The mixture was standing for 30 minutes in the dark at room temperature. The solution was separated from the beads via centrifugation for 3min at 1 000rpm and the filtrate was removed. Beads were washed then with water (3x 200µL). Solutions were combined and the product was purified with an Agilent C18 Cartridge 4mg prewashed with acetonitrile (3x 100µL) and 0,1% TFA in water (3x 100µL). After loading the sample, the cartridge was washed with 0,1% TFA in water (3x 100µL) and the product was eluted using 60% acetonitrile/water (2x 50µL).

6.3.2.2- Negative charged peptide labelled with iodo-TMT®

Sialylated glycopeptide **42** (0,06mg; 23,8nmol) was oxidized using 174µL $NaIO_4$ solution 20mM for 30 minutes at 4°C. The product was purified with an Agilent C18 Cartridge 4mg prewashed with acetonitrile (3x 100µL) and 0,1% TFA in water (3x 100µL). After loading the sample, the cartridge was washed with 0,1% TFA in water (3x 100µL) and the product was eluted using 60% acetonitrile/water (2x 50µL). The solvent was removed in the speed-vac.

Magnetic beads **35** (50µL; ca. 247nmol active groups) were transferred to an eppendorf tube and washed (3x 100µL) with reaction buffer - 0,1M acetate buffer pH 4,5; 0,1M NaCl. In each washing step, the beads were washed with buffer, supernatant was removed while the tube was remaining on the magnetic separator and the solution was removed. Oxidized peptide was dissolved in 50µL reaction buffer and the solution was transferred to the beads. The mixture was gentle stirred overnight at room temperature.

The solution was separated from the beads while the tube was remaining on the magnetic separator and the filtrate was removed. Beads were washed then with the next reaction buffer – HEPES buffer 50mM, EDTA 10mM, pH8,0 (3x 100µL). *Tris*(2-carboxyethyl)phosphine (TCEP) solution 10mM in buffer (24µL) was added to the beads and the reaction was stirred for 1 hour at 60°C. The reaction mixture was cooled to room temperature and 10µL iodo-TMT® **49** solution 44mM in methanol was added. The mixture was standing for 4 hours in the dark at 37°C. The solution was separated from the beads using the magnetic support and the filtrate was removed. Beads were washed then with water (3x 50µL). Solutions were combined and the product was purified with an Agilent C18 Cartridge 4mg prewashed with acetonitrile (3x 100µL) and 0,1% TFA in water (3x 100µL). After loading the sample, the cartridge was washed with 0,1% TFA in water (3x 100µL) and the product was eluted using 60% acetonitrile/water (2x 50µL).

6.3.2.3- Neutral peptide

Commercial lyophilized galactose oxidase containing buffer salts and stabilizer (GOase – *Sigma-Aldrich* G7907) was diluted with water to 50units/mL. Catalase from bovine liver (2000-5000units/mg – *Sigma-Aldrich* C9322) was solubilized in PBS buffer pH7,5 to 0,01mg/µL.

0,1mg (0,04µmol) of glycopeptide **58**, 5 units of GOase and 500 units of catalase were diluted to 100µL (total volume) water and final pH7,5. The reaction was followed using analytical HPLC. After overnight, the mixture was acidified with 1µL acetic acid.

Beads **52** (130µL; ca 130nmol active groups) were transferred to an eppendorf tube and washed (5x 300µL) with reaction buffer - 0,1M acetate buffer pH4,5; 0,1M NaCl. In each washing step, the beads were washed with buffer, centrifuged for 3min at 1 000rpm and the filtrate was removed. 30µL of the previous solution were diluted with buffer up to 250µL and the new solution was transferred to the beads. The mixture was gentle stirred overnight at room temperature.

The solution was separated from the beads via centrifugation for 3min at 1 000rpm and the filtrate was removed. Beads were washed then with the next reaction buffer – HEPES buffer 50mM, EDTA 10mM, pH8,0 (5x 200µL). *Tris*(2-carboxyethyl)phosphine (TCEP) solution 10mM in buffer (200µL) was added to the beads and the reaction was stirred for 1 hour at 60°C. The reaction mixture was cooled to room temperature and 200µL iodoacetoamide solution 30mM in buffer was added. The mixture was standing for 30 minutes in the dark at room temperature. The solution was separated from the beads via centrifugation for 3min at 1 000rpm and the filtrate was removed. Beads were washed then with water (3x 200µL). Solutions were combined and the product was purified with an Agilent C18 Cartridge 4mg prewashed with acetonitrile (3x 100µL) and 0,1% TFA in water (3x 100µL). After loading the sample, the cartridge was washed with 0,1% TFA in water (3x 100µL) and the product was eluted using 60% acetonitrile/water (2x 50µL).

6.3.3- Beads test – Protein level

6.3.3.1- Digestion protocol

200µg of protein were dissolved in 25µL 4M guanidinium chloride in water. DTT (2,78µL) 100mM stock solution in water was added and the mixture reacted for 30 minutes at 56°C. As soon as the mixture achieved room temperature, IAA (3,10µL) 200mM stock solution in water was added and the mixture reacted for more 30 minutes at room temperature in the dark. The solution was then diluted with 469µL 1mM $CaCl_2$/ammonium bicarbonate buffer 50mM. Trypsin from Sigma (10µg) was added to the solution and it was incubated for 16h at 37°C. 10% TFA in water (100µL) was added in order to decrease the pH solution to 2,0-2,5. Peptide mixture was purified with Agilent C18 Cartridge 4mg prewashed with acetonitrile (3x 100µL) and 0,1% TFA in water (3x 100µL). After loading the sample, the cartridge was washed with 0,1% TFA in water (3x 100µL). The product was eluted using 60% acetonitrile/water (2x 50µL) and dried under vacuum.

6.3.3.2- Periodate oxidation

200µL of $NaIO_4$ 2mM in water was added to the peptide mixture. The mixture reacted for 30 minutes at 4°C. Sample was then desalted using Agilent C18 Cartridge 4mg prewashed with acetonitrile (3x 100µL) and 0,1% TFA in water (3x 100µL). After loading the sample, the cartridge was washed with 0,1% TFA in water (3x 100µL). The product was eluted using 60% acetonitrile/water (2x 50µL) and dried under vacuum.

6.3.3.3- Galactose oxidation

10U/mg of galactose oxidase solution 1U/µL and 1000U/mg of catalase bovine solution 20U/µl were added to the sample. The mixture were diluted with PBS buffer to 200µL/mg, the pH was corrected to 7,0 with NaOH 6M and the mixture reacted overnight at room temperature.

6.3.3.4- Beads conjugation

Beads (0,5µL/µg) were transferred to an Eppendorf tube and washed (5x 300µL) with reaction buffer - 0,1M acetate buffer pH 4,5; 0,1M NaCl. In each washing step, the beads were washed with buffer, centrifuged for 3min at 1 000rpm and the filtrate was removed. Oxidized peptide was diluted or dissolved in 1µL/µg reaction buffer and the solution was transferred to the beads. The mixture was gentle stirred overnight at room temperature.

6.3.3.5- Cleavage of peptide level enrichment

The solution was separated from the beads via centrifugation for 3 minutes at 1 000rpm and the filtrate was removed. Beads were washed (3x) with 1mL by vortexing for 5 minutes with the following solutions[301, 329] and centrifuged 3 minutes at 1 000rpm:

a) 80% MeCN in 0,1% TFA/water;
b) denaturation buffer (8M urea, 0,4M ammonium bicarbonate, 0,1% SDS);
c) DMF;
d) reaction buffer – HEPES buffer 50mM, EDTA 10mM, pH8,0.

Tris(2-carboxyethyl)phosphine (TCEP) solution 10mM in buffer (200µL) was added to the beads and the reaction was stirred for 1 hour at 60^{o}C. The reaction mixture was cooled to room temperature and iodoacetoamide solution 30mM in buffer was added (200µL). The mixture was standing for 30 minutes in the dark at room temperature. The solution was separated from the beads via centrifugation for 3 minutes at 1 000rpm. Sample was then desalted using Agilent C18 Cartridge 4mg prewashed with acetonitrile (3x 100µL) and 0,1% TFA in water (3x 100µL). After loading the sample, the cartridge was washed with 0,1% TFA in water (3x 100µL). The product was eluted using 60% acetonitrile/water (2x 50µL).

6.3.3.6- LC-MS analysis

Samples were solubilized in 100µL water. 1µL of this solution was further diluted to 15µL and it was analyzed by nano-LC-MS/MS using a Velos Orbitrap mass spectrometer online-coupled to a nano RSLC HPLC system (both *Thermo Scientific*). LC separation was performed using a linear gradient from 3-42% of B in 98 minutes. Samples were measured in data-dependent acquisition mode acquiring full MS scans at a resolution of 60 000 and dynamic exclusion of 10s. The 7 most abundant ions (Top7) were selected for further fragmentation at MS/MS resolution of 7 500. Target values were set to 1x10^{6} for MS and 2x10^{5} for MS/MS scans. Maximum injection times were 10 and 100ms, respectively. Each selected ion was fragmented using HCD with normalized collision energy of 30% and 40%. First mass was set to 100 and 500*m*/*z*, respectively.

6.3.4- Beads test – Serum and CSF

6.3.4.1- Preparation

This experiment was performed on de-identified serum and CSF samples supplied by Sahlgrenska University Hospital. The CSF was centrifuged at 1800x*g* for 10min. Both samples were kept at -80°C until use.

6.3.4.2- Digestion protocol

Sample was solubilized in 20µL mixture 1:1 protease MAX® (*Promega*) surfactant 0,2% and reaction buffer – ammonium bicarbonate buffer 50mM. Once the sample was in solution, it was diluted with 73,5µL buffer prior addition of 1µL dithiothreitol solution 0,5M in buffer. The mixture was incubated for 20 minutes at 56°C. Afterwards 2,7µL of 0,55M iodoacetamide solution in buffer were added and the mixture reacted for 15 minutes in the dark. Trypsin from Promega was solubilized to 1µg/µL with Promega standard solution and 2,5µL of this solution were added to the sample solution. The mixture reacted for 16h at 37°C. The reaction was stopped with 10% TFA until the pH reached 2. Sample was then desalted using Pierce C18 spin column pre-washed 3x 200µL 50% MeCN in 0,5% TFA solution and centrifuged for 1 minute at 1 500x*g*. The sample was loaded three times to the resin, centrifuged for 1 minute at 300x*g* and washed 2x 200µL 50% MeCN in 0,5% TFA solution with centrifugation of 1 minute at 1 500x*g*. The sample was finally eluted with 2x 20µL 70% MeCN in water, centrifuged for 1 minute at 1 500x*g* and dried under vacuum.

6.3.4.3- Periodate oxidation

0,7µL/µg of H_5IO_6 solution 2mM in water were added to the sample. The mixture reacted for 30 minutes at 4°C. Sample was then desalted using Pierce C18 spin column pre-washed 3x 200µL 50% MeCN in 0,5% TFA solution and centrifuged for 1 minute at 1 500x*g*. The sample was loaded three times to the resin, centrifuged for 1 minute at 300x*g* and washed 2x 200µL 50% MeCN in 0,5% TFA solution with centrifugation of 1 minute at 1 500x*g*. The sample was finally eluted with 2x 20µL 70% MeCN in water, centrifuged for 1 minute at 1 500x*g* and dried under vacuum.

6.3.4.4- Beads conjugation

Beads (0,5µL/µg) were transferred to an Eppendorf tube and washed (5x 300µL) with reaction buffer - 0,1M acetate buffer pH 4,5; 0,1M NaCl. In each washing step, the beads were washed with buffer, centrifuged for 3min at 1 000rpm and the filtrate was removed. Oxidized sample was diluted or dissolved in 1µL/µg reaction buffer and the solution was transferred to the beads. The mixture was gentle stirred overnight at room temperature.

6.3.4.5- Cleavage of peptides after enrichment

The solution was separated from the beads via centrifugation for 3 minutes at 1 000rpm and the filtrate was removed. Beads were washed with three times 300µL with the following solutions and centrifuged 3 minutes at 1 000rpm:

a) 80% MeCN in water;
b) 20% MeCN in water;
c) reaction buffer – HEPES buffer 50mM, EDTA 10mM, pH8,0.

Tris(2-carboxyethyl)phosphine (TCEP) solution 10mM in buffer (0,3µL/mg) was added to the beads and the reaction was stirred for 1 hour at 60°C. The reaction mixture was cooled to room temperature and 0,3µL/mg iodoacetoamide solution 30mM in buffer was added. The mixture was standing for 30 minutes in the dark at room temperature. The solution was separated from the beads via centrifugation for 3 minutes at 1 000rpm. Sample was then desalted using Pierce C18 spin column pre-washed 3x 200µL 50% MeCN in 0,5% TFA solution and centrifuged for 1 minute at 1 500x*g*. The sample was loaded three times to the resin, centrifuged for 1 minute at 300x*g* and washed 2x 200µL 50% MeCN in 0,5% TFA solution with centrifugation of 1 minute at 1 500x*g*. The sample was finally eluted with 2x 20µL 70% MeCN in water, centrifuged for 1 minute at 1 500x*g* and dried under vacuum.

6.3.4.6- LC-MS analysis

Samples were solubilized in 90µL water. 1µL of this solution was further diluted to 15µL and it was analyzed by nano-LC-MS/MS using a LUMOS mass spectrometer online-coupled to a nano RSLC HPLC system (both *Thermo Scientific*). LC separation was performed using a linear gradient from 3-42% of B in 90 minutes. Samples were measured in data-dependent acquisition mode acquiring full MS scans at a resolution of 120 000 and dynamic exclusion of 12s. The 20 most abundant ions (Top20) were selected for further fragmentation at MS/MS resolution of 30 000. Target values were set to $2x10^5$ for MS and $5x10^4$ for MS/MS scans. Maximum injection times were 50 and 100ms, respectively. Each selected ion was fragmented using HCD with normalized collision energy of 40% and ETD with a supplementary activation of 30%. First mass was set to 90*m*/*z* in both fragmentation types.

6.3.4.7- Data analysis: Byonic™ searches

In the case of synthetic glycopeptides, raw files were evaluated manually using the program Xcalibur Qual Browser 2.2. Futher generated raw data were searched against the Uniprot sequence of bovine α2-HS glycoprotein or fetuin-A (P12763) and bovine fetuin-B (Q58D62) (September, 2016), in the case of bovin fetuin and asialofetuin samples, and a human Uniprot database (November, 2015), in case of serum and CSF samples. Byonic™ software was used as search engine for peptide identification. Enzyme specificity was set to semi-specific N-ragged for trypsin with one allowed missed cleavage. Carbamidomethylation of cysteine (+57,0214Da) was selected as fixed, oxidation of methionine (+15,9949Da) and deamidation of asparagine (+0,9840Da) as variable modifications. Searches for *N*- and *O*-glycopeptides were also selected as variable modifications and performed separately for each file (glycan list in the appendix). Mass tolerances were set to 10ppm for MS and 0,3Da for MS/MS. Protein FDR cutoff was set to 2%.

After excluding all the protein hits from the first reverse sequence, both peptide lists (*N*- and *O*-glycopeptides searches) were merged in a single list. All non-glycosylated peptides identified as well as those glycopeptides that have "|Log Prob|" value <1 were excluded. Multiple identification of the same peptide was excluded based on their "Delta Mod" and "|Log Prob|" values, in order to obtain the number of unique glycopeptides. According to the Byonic™ user guide, "|Log Prob|" is the

absolute value of the p-value logarithm. “Delta Mod” is a value that indicates the probability of the glycosylation site being rightly assigned. A “Delta Mod”>10 has indicates a big high probabilities of a right glycopeptide assignment. Those peptides were considered true hits if the *O*-glycopeptide spectrum presented at least 3 b/c and y/z ions as well as one oxonium ion and the glycan fragment ions coincides with the glycan structure. In some rare cases, peptides with a tryptic cleavage and less than 3 b/c and y/z ions were also accepted as true hits. For *N*-glycopeptides spectrum, only spectra containing at least one chemically modified oxonium ion were considered a true hit. The glycosylation sites were ranked according to Byonic parameter “Delta Mod” and manual inspection of the spectrum. MS/MS spectra that presented more than one fragment ion confirming the glycosylation site and “Delta Mod”>10 were assigned as high confidence sites. Spectra with only one fragment ion confirming glycosylation site and “Delta Mod”>10 were assigned as good confidence sites. Those spectra with only one fragment ion confirming glycosylation site or “Delta Mod”>10 were assigned as medium confidence sites. Finally spectra with neither one fragment ion confirming glycosylation site nor “Delta Mod”>10 were assigned as poor confidence sites.

6.3.5- Beads test – Intact glycopeptides compared with methodologies for determination of glycosylation sites

6.3.5.1- CSF preparation

This experiment was performed on de-identified CSF samples supplied by Haukeland University Hospital. CSF samples from patients were combined in a single pool and separated in aliquots containing 300µg proteins. Samples were kept at -80°C until their use. Protein concentrations were measured using a Qubit™ fluorometer kit (*Invitrogen*, Carlsbrad, CA) according to the vendor’s instructions. CSF was purified, concentrated to 50µL using 3kDa ultracentrifugation filters pre-coated with 1mL 0.1% *N*-octyl-β-D-glucopyranose (NOG) and dried completely before the following steps.

6.3.5.2- Digestion protocol

Sample was solubilized by the addition of 150µL denaturation buffer (8M urea, 0,4M ammonium bicarbonate and 0,1% SDS in water). 12,5µL of 120mM TCEP solution in water were added prior the incubation for 60 minutes at 37°C. Afterwards 12,2µL of 160mM iodoacetamide solution in water were added and the mixture reacted for 30 minutes at 20°C in the dark. Trypsin from Promega was solubilized to 0,5µg/µL with Promega standard solution and 12µL of this solution were added to the sample solution. The mixture reacted for 16h at 37°C. The reaction was stopped by the addition of 12µL 10% FA in water. Samples were then purified using Oasis® plates, procedure described in the next session.

6.3.5.3- Oasis® HLB µElution cleanup

Reverse phase Oasis® HLB µElution plates 30µm (*Waters Corp.*, Milford, MA, USA) were used to desalt and clean up samples. The Oasis plate was washed once with 500µL 80% acetonitrile/0,1% FA in water, and twice with 500µL 0,1% FA in water prior sample addition. The washing was done by addition of the solution to the well and then spinning down the well plate using a Universal 320 R centrifuge (Hettich centrifuge) at 200x*g* for 1 minute. The sample was then added and the plate was centrifuged at 150x*g* for 3 minutes, followed by three more washing steps with 500µL 0,1% FA in water at 200x*g* for 1 minute. Finally, the sample was eluted with between 150µL 80% acetonitrile/0,1% FA in water, washed two times through the well at 200x*g* for 1 minute each time. All the centrifuging steps were done at 4°C. The sample was then collected from the collection plate, transferred to new eppendorf tubes and dried under vacuum prior the next step.

6.3.5.4- Periodate oxidation

6.3.5.4.1- Mild oxidation

Peptide mixture was diluted with 200µL water and 4µL of $NaIO_4$ 100mM in water were added to solution. The mixture reacted for 30 minutes at 4°C. Samples were acidified with 200µL 0,1% TFA in water before purification using Oasis® plates.

6.3.5.4.2- Harsh oxidation

Peptide mixture was diluted with 400µL 0,1% TFA in water and 45µL of $NaIO_4$ 100mM in water were added to solution. The mixture reacted for 60 minutes at 20°C in the dark. Samples were then purified using Oasis® plates.

6.3.5.5- Beads conjugation

6.3.5.5.1- Hydroxylamine beads

Beads (0,5µL/µg) were transferred to an Eppendorf tube and washed (3x 300µL) with reaction buffer - 0,1M acetate buffer pH 4,5; 0,1M NaCl. In each washing step, the beads were washed with buffer, centrifuged for 3 min at 1 000rpm and the supernatant was removed. Oxidized peptide was diluted or dissolved in 1µL/µg reaction buffer and the solution was transferred to the beads. The mixture was stirred gently overnight at room temperature.

6.3.5.5.2- Hydrazine beads

Beads (0,17µL/µg) were transferred to an Eppendorf tube and washed with 1mL deionized water, centrifuged for 1 min at 3 000x*g* and the supernatant was removed. Oxidized peptide was diluted in 300µL 0,1% TFA in water and the solution was transferred to the beads. The mixture was stirred overnight at 20°C.

6.3.5.6- Cleavage of peptide level enrichment

The solution was separated from the beads via centrifugation for 3 minutes at 1 000rpm and the filtrate was removed. Beads were washed with 3x 1mL by vortexing for 5 minutes with the following solutions and centrifuged 3 minutes at 1 000rpm:

a) 80% MeCN in 0,1% TFA/water;
b) denaturation buffer (8M urea, 0,4M ammonium bicarbonate, 0,1% SDS);
c) DMF;
d) reaction buffer (ammonium bicarbonate 100mM for hydrazine beads and HEPES buffer 50mM, EDTA 10mM, pH8,0 for hydroxylamine beads).

6.3.5.6.1- Hydroxylamine beads – chemical cleavage

Tris(2-carboxyethyl)phosphine (TCEP) solution 10mM in buffer (150µL) was added to the beads and the reaction was stirred for 1 hour at 60^{o}C. The reaction mixture was cooled to room temperature and iodoacetoamide solution 30mM in buffer was added (150µL). The mixture was standing for 30 minutes in the dark at room temperature. The solution was separated from the beads via centrifugation for 3 minutes at 1 000rpm. Samples were acidified before purification using Oasis® plates.

6.3.5.6.2- Hydroxylamine beads – enzymatic cleavage

800µL reaction buffer – HEPES buffer 50mM, EDTA 10mM, pH8,0 were added to each sample and 2 units of PNGase F were added to each sample. They were incubated overnight at 37°C. Samples were acidified before purification using Oasis® plates.

6.3.5.6.3- Hydrazine beads – enzymatic cleavage

1 000µL ammonium bicarbonate 100mM were added to each sample and 2 units of PNGase F were added to each sample. They were incubated overnight at 37°C. Samples were acidified before purification using Oasis® plates.

6.3.5.7- TMT® labeling and fractionation – enzymatically cleaved samples

5% acetonitrile/0,1% FA in water was used to solubilize the samples obtained via enzymatic cleavage (20µL for hydroxylamine beads and 40µL for hydrazine beads). 2µL of this solution were separated. The remaining solution was dried and resuspended in triethylammonium bicarbonate (TEAB) 100mM solution (50µL for hydroxylamine beads and 100µL for hydrazine beads). TMT® 10-plex label reagents were randomly assigned to the samples (see Table 20) and the labeling was performed according to the vendor procedure. After the labelling, the samples were pooled into one Eppendorf tube, 2% of the mixture were separated as control and the samples were dried prior purification using Oasis® plates.

Table 20. Assignment of enzymatic cleaved samples with TMT® 10-plex.

TMT® label reagent	Replicate
TMT10-126	Hydroxylamine beads 3
TMT10-127N	Hydrazine beads 4
TMT10-127C	Hydroxylamine beads 2
TMT10-128N	Hydroxylamine beads 4
TMT10-128C	Hydrazine beads 2
TMT10-129N	Hydrazine beads 1
TMT10-129C	Hydroxylamine beads 1
TMT10-130N	Hydroxylamine beads 5
TMT10-130C	Hydrazine beads 5
TMT10-131	Hydrazine beads 3

The TMT-labeled peptides were fractionated in 15 aliquots using mixed-mode chromatography on a Sielc Promix MP column (MP- 10.250.0530, 1,0 × 250mm, 5µm, 300Å, Sielc Technologies) and an Agilent 1260 series LC system (*Agilent Technologies*). The peptides were reconstituted in buffer A (20mM ammonium formate, 3% acetonitrile) and loaded on the column using 15% buffer B (2mM ammonium formate, 80% acetonitrile, pH 3.0). Details for the gradient and fraction collection are in Table 21.

Table 21. LC-conditions for TMT® labeled samples

LC-gradient	Time (min)	0	10	45	55	60	65	80
	% B	15	15	60	100	100	15	15
Fraction collection	Time (min)	0	2	9	70			
	Slice (min)	2	7	5	Off			

6.3.5.8- LC-MS analysis

6.3.5.8.1- Fractions – enzymatically cleaved samples

Dried fractions were resolubilized in 2% acetonitrile, 1% FA. Approximately 0,5µg of peptides from each fraction was injected into an Ultimate 3000 RSLC system (*Thermo Scientific*) connected to a Q-Exactive HF equipped with an EASY-spray ion source (*Thermo Scientific*). The samples were loaded and desalted on a precolumn (Acclaim PepMap 100, 2cmx75µm i.d. nanoViper column, packed with 3µm C18 beads) at a flow rate of 3µL/min for 5min with 0,1% TFA. The peptides were separated during a biphasic acetonitrile gradient from two nanoflow UPLC pumps (flow rate of 0,200 µL/min) on a 50cm analytical column (PepMap RSLC, 50cmx75µm i.d. EASY-spray column, packed with 2µm C18 beads (*Thermo Scientific*). Solvent A was 0,1% FA in water, and solvent B was 100% acetonitrile. The gradient started with 5 minutes of 5% solvent B. In 0,5min the solvent B content was increased to 7% and a gradient of 0,25% solvent B per minute was used until the minute 65 when the gradient changed for 0,60% solvent B per minute for more 22 minutes. Then fractions were measured in data-dependent acquisition mode acquiring full MS scans at a resolution of 60 000 and dynamic exclusion of 30s. The 15 most abundant ions (Top15) were selected for further fragmentation at MS/MS resolution of 30 000. Target values were set to $3x10^6$ for MS and $1x10^5$ for

MS/MS scans. Maximum injection times were 50 and 45ms, respectively. Each selected ion was fragmented using HCD with normalized collision energy of 32%.

6.3.5.8.2- Chemically cleaved samples

Samples were solubilized in 90µL water. 1µL of this solution was further diluted to 15µL and analyzed by nano-LC-MS/MS using a LUMOS mass spectrometer online-coupled to a nano RSLC HPLC system (both *Thermo Scientific*). LC separation was performed using a linear gradient from 3-35% of B in 90 minutes. Samples were measured in data-dependent acquisition mode acquiring full MS scans at a resolution of 120 000, dynamic exclusion of 30s, target values were set to $2x10^5$ and injection time of 50ms. The 7 most abundant ions (Top7) were selected for further fragmentation under 3 different methods: HCD with normalized collision energy of 30% and 40%, and ETD with a supplementary activation of 32%. Fragments obtained via HCD with NCE 30% and EThcD were analyzed in the ion trap at normal scan speed (33Da/s), target values were set to $1x10^4$, injection time of 100ms. First mass was set to 90*m*/*z* in both fragmentation types. Fragments obtained via HCD with NCE 40% were analyzed in the orbitrap at a MS/MS resolution of 30 000, target values were set to $5x10^4$ and injection time of 100ms. First mass was set to 800*m*/*z*. The 3 most abundant ions (Top3) were selected for further fragmentation. Fragments selected for MS^3 were analyzed in the obitrap at a resolution of 15 000, target values were set to $5x10^4$, injection time of 200ms and first mass was set to 90*m*/*z*. Each selected ion for MS^3 analysis was fragmented using HCD with normalized collision energy of 32%.

6.3.5.9- Data analysis

6.3.5.9.1- PeptideShaker searches – enzymatically cleaved samples

The raw files were converted to mgf using Proteowizard. Files were searched against the human Uniprot reviewed database (20 129 entries, downloaded January 2017), with the addition of 48 non-human contaminants from the cRAP database[351] using SearchGUI v.3.2.5 and the search engines X!Tandem, MyriMatch and Comet. The search settings were: carbamidomethylation of cysteine, TMT®-10plex labeling on lysine and N-terminal as fixed modifications; oxidation of methionine and deamidation of asparagine as variable modification; trypsin as enzyme with a maximum of two missed cleavages; precursor charge 2-5; peptide length 6-30; precursor and fragment mass tolerance 10ppm. Protein FDR cutoff was set to 1%. All other settings were left to default. Search results were summarized in PeptideShaker v.1.15.0 and peptide quantitative results were explored and exported using Reporter v.0.7.9. For each peptide, the TMT reporter abundance was normalized by scaling factors derived from the median abundances of all channels over all peptides. Previously *N*-glycosylated peptides were identified as those peptides containing a deamidation and an asparagine (N) residue in the *N*-glyco consensus motif N-X^P-S/T where X is any amino acid except proline (P) and S is serine and T is threonine. Amino acids C-terminal to the peptide were also considered when checking for the *N*-glyco consensus motif. Peptides without reported values in any of the TMT channels were removed after the normalization.

6.3.5.9.2- Byonic™ searches – chemically cleaved samples

Generated raw data were searched against a human Uniprot database (November, 2015). Byonic™ software was used as search engine for peptide identification. Enzyme specificity was set to semi-specific N-ragged for trypsin with one allowed missed cleavage. Carbamidomethylation of cysteine (+57,0214Da) was selected as fixed, oxidation of methionine (+15,9949Da) and deamidation of asparagine (+0,9840Da) as variable modifications. Searches for *N*- and *O*-glycopeptides were also selected as variable modifications and performed separately for each file (glycan list in the appendix). Mass tolerances were set to 10ppm for MS and 0,3Da for MS/MS. Protein FDR cutoff was set to 2%.

After excluding all the protein hits from the first reverse sequence, both peptide lists (*N*- and *O*-glycopeptides searches) were merged in a single list. All non-glycosylated peptides identified as well as those glycopeptides that have "|Log Prob|" value <1 were excluded. Multiple identification of the same peptide was excluded based on their "Delta Mod" and "|Log Prob|" values, in order to obtain the number of unique glycopeptides. According to the Byonic™ user guide, "|Log Prob|" is the absolute value of the p-value logarithm. "Delta Mod" is a value that indicates the probability of the glycosylation site being rightly assigned. A "Delta Mod">10 has indicates a big high probabilityities of a right glycopeptide assignment. Those peptides were considered true hits if the *O*-glycopeptide spectrum presented at least 3 b/c and y/z ions as well as one oxonium ion and the glycan fragment ions coincides with the glycan structure. In some rare cases, peptides with a tryptic cleavage and less than 3 b/c and y/z ions were also accepted as true hits. For *N*-glycopeptides spectrum, only spectra containing at least one chemically modified oxonium ion were considered a true hit. The glycosylation sites were ranked according to Byonic parameter "Delta Mod" and manual inspection of the spectrum. MS/MS spectra that presented more than one fragment ion confirming the glycosylation site and "Delta Mod">10 were assigned as high confidence sites. Spectra with only one fragment ion confirming glycosylation site and "Delta Mod">10 were assigned as good confidence sites. Those spectra with only one fragment ion confirming glycosylation site or "Delta Mod">10 were assigned as medium confidence sites. Finally spectra with neither one fragment ion confirming glycosylation site nor "Delta Mod">10 were assigned as poor confidence sites.

References

1. Thomson, J. J., *Rays of positive electricity and their application to chemical analyses*, Longmans, Green, London; New York, **1913**.
2. Laeter, J. R. d., *Applications of inorganic mass spectrometry*, John Wiley & Sons, New York, **2001**.
3. Conrad, R., *Transactions of the Faraday Society* **1934**, *30*, 215.
4. Mclafferty, F. W., *Anal. Chem.* **1959**, *31*, 82.
5. Haakansson, P., Kamensky, I., Sundqvist, B., Fohlman, J., Peterson, P., McNeal, C. J., Macfarlane, R. D., *J. Am. Chem. Soc.* **1982**, *104*, 2948.
6. Stephens, W. E., *Physical Review* **1946**, *69*, 691.
7. Paul, W., Reinhard, H. P., Vonzahn, U., *Zeitschrift Fur Physik* **1958**, *152*, 143.
8. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., Whitehouse, C. M., *Science* **1989**, *246*, 64.
9. Japan Patent S60035097 **1985**
10. Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., Matsuo, T., *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151.
11. Karas, M., Bachmann, D., Hillenkamp, F., *Anal. Chem.* **1985**, *57*, 2935.
12. Karas, M., Hillenkamp, F., *Anal. Chem.* **1988**, *60*, 2299.
13. "The Nobel Prize in Chemistry 2002". Nobelprize.org. Nobel Media AB 2014. Web. 7 Jun 2018.
    <http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2002/index.html>
14. Chowdhury, S. K., Katta, V., Chait, B. T., *Biochem. Biophys. Res. Commun.* **1990**, *167*, 686.
15. Pappin, D. J. C., Hojrup, P., Bleasby, A. J., *Curr. Biol.* **1993**, *3*, 327.
16. Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D., Humphery-Smith, I., Hochstrasser, D. F., Williams, K. L., *Biotechnol. Genet. Eng. Rev.* **1996**, *13*, 19.
17. Patterson, S. D., Aebersold, R. H., *Nat. Genet.* **2003**, *33 Suppl*, 311.
18. Edman, P., Begg, G., *Eur. J. Biochem.* **1967**, *1*, 80.
19. Arpino, P., Baldwin, M. A., McLafferty, F. W., *Biomed. Mass Spectrom.* **1974**, *1*, 80.
20. Mann, M., Wilm, M., *Anal. Chem.* **1994**, *66*, 4390.
21. Eng, J. K., McCormack, A. L., Yates, J. R., *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976.
22. Hunt, D. F., Yates, J. R., Shabanowitz, J., Winston, S., Hauer, C. R., *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 6233.
23. Mclafferty, F. W., Bryce, T. A., *Chem. Commun.* **1967**, 1215.
24. Jennings, K. R., *International Journal of Mass Spectrometry and Ion Physics* **1968**, *1*, 227.
25. Roepstorff, P., Fohlman, J., *Biol. Mass Spectrom.* **1984**, *11*, 601.
26. Biemann, K., *Annu. Rev. Biochem* **1992**, *61*, 977.
27. Zubarev, R. A., Kelleher, N. L., McLafferty, F. W., *J. Am. Chem. Soc.* **1998**, *120*, 3265.
28. Syka, J. E. P., Coon, J. J., Schroeder, M. J., Shabanowitz, J., Hunt, D. F., *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9528.
29. Mechref, Y., *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]* **2012**, *0 12*, Unit.
30. Sarbu, M., Ghiulai, R. M., Zamfir, A. D., *Amino Acids* **2014**, *46*, 1625.
31. Olsen, J. V., Macek, B., Lange, O., Makarov, A., Horning, S., Mann, M., *Nat Methods* **2007**, *4*, 709.

32. Swaney, D. L., McAlister, G. C., Wirtala, M., Schwartz, J. C., Syka, J. E., Coon, J. J., *Anal. Chem.* **2007**, *79*, 477.
33. Frese, C. K., Altelaar, A. F., van den Toorn, H., Nolting, D., Griep-Raming, J., Heck, A. J., Mohammed, S., *Anal. Chem.* **2012**, *84*, 9668.
34. Frese, C. K., Zhou, H. J., Taus, T., Altelaar, A. F. M., Mechter, K., Heck, A. J. R., Mohammed, S., *J. Proteome Res.* **2013**, *12*, 1520.
35. Loroch, S., Dickhut, C., Zahedi, R. P., Sickmann, A., *Electrophoresis* **2013**, *34*, 1483.
36. Wuhrer, M., *Glycoconj. J.* **2013**, *30*, 11.
37. Hu, T., Zhang, J. L., *J. Sep. Sci.* **2018**, *41*, 351.
38. Schurch, S., *Mass Spectrom. Rev.* **2016**, *35*, 483.
39. Wang, Y., Liu, S., Hu, Y., Li, P., Wan, J.-B., *RSC Advances* **2015**, *5*, 78728.
40. Artigues, A., Nadeau, O. W., Rimmer, M. A., Villar, M. T., Du, X., Fenton, A. W., Carlson, G. M., *Adv. Exp. Med. Biol.* **2016**, *919*, 397.
41. Heck, A. J., Van Den Heuvel, R. H., *Mass Spectrom. Rev.* **2004**, *23*, 368.
42. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., Kuster, B., *Analytical and Bioanalytical Chemistry* **2007**, *389*, 1017.
43. Burkhart, J. M., Vaudel, M., Gambaryan, S., Radau, S., Walter, U., Martens, L., Geiger, J., Sickmann, A., Zahedi, R. P., *Blood* **2012**, *120*, e73.
44. Duarte, T. T., Spencer, C. T., *Proteomes* **2016**, *4*.
45. Davies, G. E., Stark, G. R., *Proc. Natl. Acad. Sci. U. S. A.* **1970**, *66*, 651.
46. Schneider, M., Belsom, A., Rappsilber, J., *Trends Biochem. Sci* **2018**, *43*, 157.
47. Nadeau, O. W., Carlson, G. M., *Allostery: Methods and Protocols* **2012**, *796*, 117.
48. Alber, F., Forster, F., Korkin, D., Topf, M., Sali, A., *Annu. Rev. Biochem* **2008**, *77*, 443.
49. Bai, X. C., McMullan, G., Scheres, S. H. W., *Trends Biochem. Sci* **2015**, *40*, 49.
50. Zhang, H., Tang, X., Munske, G. R., Tolic, N., Anderson, G. A., Bruce, J. E., *Molecular & Cellular Proteomics* **2009**, *8*, 409.
51. Leitner, A., Walzthoeni, T., Kahraman, A., Herzog, F., Rinner, O., Beck, M., Aebersold, R., *Mol. Cell. Proteomics* **2010**, *9*, 1634.
52. Schilling, B., Row, R. H., Gibson, B. W., Guo, X., Young, M. M., *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 834.
53. Haniu, M., Narhi, L. O., Arakawa, T., Elliott, S., Rohde, M. F., *Protein Sci.* **1993**, *2*, 1441.
54. Seebacher, J., Mallick, P., Zhang, N., Eddes, J. S., Aebersold, R., Gelb, M. H., *J. Proteome Res.* **2006**, *5*, 2270.
55. Belsom, A., Mudd, G., Giese, S., Auer, M., Rappsilber, J., *Anal. Chem.* **2017**, *89*, 5319.
56. Leitner, A., Reischl, R., Walzthoeni, T., Herzog, F., Bohn, S., Forster, F., Aebersold, R., *Mol. Cell. Proteomics* **2012**, *11*, M111 014126.
57. Rinner, O., Seebacher, J., Walzthoeni, T., Mueller, L. N., Beck, M., Schmidt, A., Mueller, M., Aebersold, R., *Nat Methods* **2008**, *5*, 315.
58. Yu, C., Huang, L., *Anal. Chem.* **2018**, *90*, 144.
59. Bennett, K. L., Kussmann, M., Bjork, P., Godzwon, M., Mikkelsen, M., Sorensen, P., Roepstorff, P., *Protein Sci.* **2000**, *9*, 1503.
60. Back, J. W., Hartog, A. F., Dekker, H. L., Muijsers, A. O., de Koning, L. J., de Jong, L., *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 222.
61. Tang, X., Munske, G. R., Siems, W. F., Bruce, J. E., *Anal. Chem.* **2005**, *77*, 311.
62. Soderblom, E. J., Goshe, M. B., *Anal. Chem.* **2006**, *78*, 8059.

63. Soderblom, E. J., Bobay, B. G., Cavanagh, J., Goshe, M. B., *Rapid Commun. Mass Spectrom.* **2007**, *21*, 3395.
64. Kao, A., Chiu, C. L., Vellucci, D., Yang, Y., Patel, V. R., Guan, S., Randall, A., Baldi, P., Rychnovsky, S. D., Huang, L., *Mol. Cell. Proteomics* **2011**, *10*, M110 002212.
65. Lu, Y., Tanasova, M., Borhan, B., Reid, G. E., *Anal. Chem.* **2008**, *80*, 9279.
66. Muller, M. Q., Dreiocker, F., Ihling, C. H., Schafer, M., Sinz, A., *Anal. Chem.* **2010**, *82*, 6958.
67. Clifford-Nunn, B., Showalter, H. D., Andrews, P. C., *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 201.
68. Gardner, M. W., Brodbelt, J. S., *Anal. Chem.* **2010**, *82*, 5751.
69. Trnka, M. J., Burlingame, A. L., *Mol. Cell. Proteomics* **2010**, *9*, 2306.
70. Chowdhury, S. M., Munske, G. R., Tang, X., Bruce, J. E., *Anal. Chem.* **2006**, *78*, 8183.
71. Kaake, R. M., Wang, X., Burke, A., Yu, C., Kandur, W., Yang, Y., Novtisky, E. J., Second, T., Duan, J., Kao, A., Guan, S., Vellucci, D., Rychnovsky, S. D., Huang, L., *Mol. Cell. Proteomics* **2014**, *13*, 3533.
72. Gotze, M., Pettelkau, J., Fritzsche, R., Ihling, C. H., Schafer, M., Sinz, A., *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 83.
73. Liu, F., Rijkers, D. T., Post, H., Heck, A. J., *Nat Methods* **2015**, *12*, 1179.
74. Herzog, F., Kahraman, A., Boehringer, D., Mak, R., Bracher, A., Walzthoeni, T., Leitner, A., Beck, M., Hartl, F. U., Ban, N., Malmstrom, L., Aebersold, R., *Science* **2012**, *337*, 1348.
75. Liu, H., Huang, R. Y., Chen, J., Gross, M. L., Pakrasi, H. B., *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 18536.
76. Yang, B., Wu, Y. J., Zhu, M., Fan, S. B., Lin, J., Zhang, K., Li, S., Chi, H., Li, Y. X., Chen, H. F., Luo, S. K., Ding, Y. H., Wang, L. H., Hao, Z., Xiu, L. Y., Chen, S., Ye, K., He, S. M., Dong, M. Q., *Nat Methods* **2012**, *9*, 904.
77. Zhong, X., Navare, A. T., Chavez, J. D., Eng, J. K., Schweppe, D. K., Bruce, J. E., *J. Proteome Res.* **2017**, *16*, 720.
78. Schweppe, R. E., Haydon, C. E., Lewis, T. S., Resing, K. A., Ahn, N. G., *Acc. Chem. Res.* **2003**, *36*, 453.
79. Lacombe, J. M., Andriamanampisoa, F., Pavia, A. A., *Int. J. Pept. Protein Res.* **1990**, *36*, 275.
80. Meyer, H. E., Hoffmannposorske, E., Korte, H., Heilmeyer, L. M. G., *FEBS Lett.* **1986**, *204*, 61.
81. Karty, J. A., Reilly, J. P., *Anal. Chem.* **2005**, *77*, 4673.
82. Green, N. S., Reisler, E., Houk, K. N., *Protein Sci.* **2001**, *10*, 1293.
83. Merkley, E. D., Rysavy, S., Kahraman, A., Hafen, R. P., Daggett, V., Adkins, J. N., *Protein Sci.* **2014**, *23*, 747.
84. Pett, C., Schorlemer, M., Westerlind, U., *Chemistry* **2013**, *19*, 17001.
85. Mossmann, D., Vogtle, F. N., Taskin, A. A., Teixeira, P. F., Ring, J., Burkhart, J. M., Burger, N., Pinho, C. M., Tadic, J., Loreth, D., Graff, C., Metzger, F., Sickmann, A., Kretz, O., Wiedemann, N., Zahedi, R. P., Madeo, F., Glaser, E., Meisinger, C., *Cell Metab* **2014**, *20*, 662.
86. Majorek, K. A., Porebski, P. J., Dayal, A., Zimmerman, M. D., Jablonska, K., Stewart, A. J., Chruszcz, M., Minor, W., *Mol. Immunol.* **2012**, *52*, 174.
87. Huang, B. X., Kim, H. Y., Dass, C., *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1237.
88. Calabrese, A. N., Wang, T., Bowie, J. H., Pukala, T. L., *Rapid Commun. Mass Spectrom.* **2013**, *27*, 238.
89. Gutierrez, C. B., Yu, C., Novitsky, E. J., Huszagh, A. S., Rychnovsky, S. D., Huang, L., *Anal. Chem.* **2016**, *88*, 8315.
90. Tinnefeld, V., Venne, A. S., Sickmann, A., Zahedi, R. P., *J. Proteome Res.* **2017**, *16*, 459.

91. Horth, P., Miller, C. A., Preckel, T., Wenz, C., *Mol. Cell. Proteomics* **2006**, *5*, 1968.
92. Beltrao, P., Bork, P., Krogan, N. J., van Noort, V., *Mol. Syst. Biol.* **2013**, *9*, 714.
93. Choudhary, C., Mann, M., *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 427.
94. Karve, T. M., Cheema, A. K., *J Amino Acids* **2011**, *2011*, 207691.
95. Mann, M., Jensen, O. N., *Nat. Biotechnol.* **2003**, *21*, 255.
96. Rocks, O., Peyker, A., Kahms, M., Verveer, P. J., Koerner, C., Lumbierres, M., Kuhlmann, J., Waldmann, H., Wittinghofer, A., Bastiaens, P. I., *Science* **2005**, *307*, 1746.
97. Martin, D. D., Beauchamp, E., Berthiaume, L. G., *Biochimie* **2011**, *93*, 18.
98. Apweiler, R., Hermjakob, H., Sharon, N., *Biochimica Et Biophysica Acta-General Subjects* **1999**, *1473*, 4.
99. Cummings, R. D., *Molecular BioSystems* **2009**, *5*, 1087.
100. Angata, T., Varki, A., *Chem. Rev.* **2002**, *102*, 439.
101. Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Marth, J. D., Bertozzi, C. R., Hart, G. W., Etzler, M. E., *Proteomics* **2009**, *9*, 5398.
102. Titani, K., Kumar, S., Takio, K., Ericsson, L. H., Wade, R. D., Ashida, K., Walsh, K. A., Chopek, M. W., Sadler, J. E., Fujikawa, K., *Biochemistry* **1986**, *25*, 3171.
103. Miletich, J. P., Broze, G. J., *J. Biol. Chem.* **1990**, *265*, 11397.
104. Zielinska, D. F., Gnad, F., Wisniewski, J. R., Mann, M., *Cell* **2010**, *141*, 897.
105. Stanley, P., Taniguchi, N., Aebi, M., in *Essentials of Glycobiology*, Cold Spring Harbor (NY), **2015**, pp. 99.
106. Wildt, S., Gerngross, T. U., *Nature Reviews Microbiology* **2005**, *3*, 119.
107. Strasser, R., *Glycobiology* **2016**, *26*, 926.
108. Taylor, C. M., Karunaratne, C. V., Xie, N., *Glycobiology* **2012**, *22*, 757.
109. Yamauchi, M., Sricholpech, M., *Lysine-Based Post-Translational Modification of Proteins* **2012**, *52*, 113.
110. Levery, S. B., Steentoft, C., Halim, A., Narimatsu, Y., Clausen, H., Vakhrushev, S. Y., *Biochim. Biophys. Acta* **2015**, *1850*, 33.
111. Hart, G. W., *Annu. Rev. Biochem* **1997**, *66*, 315.
112. Ihara, Y., Inai, Y., Ikezaki, M., *Trends in Glycoscience and Glycotechnology* **2011**, *23*, 1.
113. Varki, A., *Glycobiology* **1993**, *3*, 97.
114. Dwek, R. A., *Chem. Rev.* **1996**, *96*, 683.
115. Grunewald, S., *Early Hum. Dev.* **2007**, *83*, 825.
116. Aebi, M., Helenius, A., Schenk, B., Barone, R., Fiumara, A., Berger, E. G., Hennet, T., Imbach, T., Stutz, A., Bjursell, C., Uller, A., Wahlström, J. G., Briones, P., Cardo, E., Clayton, P., Winchester, B., Cormier-Dalre, V., de Lonlay, P., Cuer, M., Dupré, T., Seta, N., de Koning, T., Dorland, L., de Loos, F., Kupers, L., Fabritz, L., Hasilik, M., Marquardt, T., Niehues, R., Freeze, H., Grünewald, S., Heykants, L., Jaeken, J., Matthijs, G., Schollen, E., xKeir, G. K., Kjaergaard, S., Schwartz, M., Skovby, F., Klein, A., Roussel, P., Körner, C., Lübke, T., Thiel, C., von Figura, K., Koscielak, J., Krasnewich, D., Lehle, L., Peters, V., Raab, M., Saether, O., Schachter, H., Van Schaftingen, E., Verbert, A., Vilaseca, A., Wevers, R., Yamashita, K., *Glycoconj. J.* **1999**, *16*, 669.
117. Jaeken, J., Hennet, T., Freeze, H. H., Matthijs, G., *J. Inherit. Metab. Dis.* **2008**, *31*, 669.
118. Morava, E., Lefeber, D., *J. Inherit. Metab. Dis.* **2011**, *34*, 847.
119. Grunewald, S., Matthijs, G., Jaeken, J., *Pediatr. Res.* **2002**, *52*, 618.
120. Jaeken, J., Carchon, H., Stibler, H., *Glycobiology* **1993**, *3*, 423.

121. Lubke, T., Marquardt, T., Etzioni, A., Hartmann, E., von Figura, K., Korner, C., *Nat. Genet.* **2001**, *28*, 73.
122. Wuyts, W., Van Hul, W., *Hum. Mutat.* **2000**, *15*, 220.
123. Topaz, O., Shurman, D. L., Bergman, R., Indelman, M., Ratajczak, P., Mizrachi, M., Khamaysi, Z., Behar, D., Petronius, D., Friedman, V., Zelikovic, I., Raimer, S., Metzker, A., Richard, G., Sprecher, E., *Nat. Genet.* **2004**, *36*, 579.
124. Beltran-Valero de Bernabe, D., Currier, S., Steinbrecher, A., Celli, J., van Beusekom, E., van der Zwaag, B., Kayserili, H., Merlini, L., Chitayat, D., Dobyns, W. B., Cormand, B., Lehesjoki, A. E., Cruces, J., Voit, T., Walsh, C. A., van Bokhoven, H., Brunner, H. G., *Am. J. Hum. Genet.* **2002**, *71*, 1033.
125. Yoshida, A., Kobayashi, K., Manya, H., Taniguchi, K., Kano, H., Mizuno, M., Inazu, T., Mitsuhashi, H., Takahashi, S., Takeuchi, M., Herrmann, R., Straub, V., Talim, B., Voit, T., Topaloglu, H., Toda, T., Endo, T., *Dev. Cell* **2001**, *1*, 717.
126. Wiley, D. C., Skehel, J. J., *Annu. Rev. Biochem* **1987**, *56*, 365.
127. Russell, R. J., Kerry, P. S., Stevens, D. J., Steinhauer, D. A., Martin, S. R., Gamblin, S. J., Skehel, J. J., *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17736.
128. Correa, P., Houghton, J., *Gastroenterology* **2007**, *133*, 659.
129. Magalhaes, A., Reis, C. A., *Braz. J. Med. Biol. Res.* **2010**, *43*, 611.
130. Krogfelt, K. A., Bergmans, H., Klemm, P., *Infect. Immun.* **1990**, *58*, 1995.
131. Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S. J., Knight, S. D., *Science* **1999**, *285*, 1061.
132. Kranjcec, B., Papes, D., Altarac, S., *World J. Urol.* **2014**, *32*, 79.
133. Hakomori, S., *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 10231.
134. Freeze, H. H., *Nature Reviews Genetics* **2006**, *7*, 537.
135. McEver, R. P., *Glycoconj. J.* **1997**, *14*, 585.
136. Leick, M., Azcutia, V., Newton, G., Luscinskas, F. W., *Cell Tissue Res.* **2014**, *355*, 647.
137. Bargatze, R. F., Kurk, S., Butcher, E. C., Jutila, M. A., *J. Exp. Med.* **1994**, *180*, 1785.
138. Jutila, M. A., Kurk, S., *J. Immunol.* **1996**, *156*, 289.
139. Simon, S. I., Rochon, Y. P., Lynam, E. B., Smith, C. W., Anderson, D. C., Sklar, L. A., *Blood* **1993**, *82*, 1097.
140. Fuhlbrigge, R. C., Alon, R., Puri, K. D., Lowe, J. B., Springer, T. A., *J. Cell Biol.* **1996**, *135*, 837.
141. Ohmi, Y., Ise, W., Harazono, A., Takakura, D., Fukuyama, H., Baba, Y., Narazaki, M., Shoda, H., Takahashi, N., Ohkawa, Y., Ji, S., Sugiyama, F., Fujio, K., Kumanogoh, A., Yamamoto, K., Kawasaki, N., Kurosaki, T., Takahashi, Y., Furukawa, K., *Nat Commun* **2016**, *7*, 11205.
142. Bondt, A., Selman, M. H., Deelder, A. M., Hazes, J. M., Willemsen, S. P., Wuhrer, M., Dolhain, R. J., *J. Proteome Res.* **2013**, *12*, 4522.
143. Dueymes, M., Bendaoud, B., Pennec, Y. L., Youinou, P., *Clin. Exp. Rheumatol.* **1995**, *13*, 247.
144. Chui, D., Sellakumar, G., Green, R., Sutton-Smith, M., McQuistan, T., Marek, K., Morris, H., Dell, A., Marth, J., *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 1142.
145. Brockhausen, I., *Biochim. Biophys. Acta* **1999**, *1473*, 67.
146. Galli-Stampino, L., Meinjohanns, E., Frische, K., Meldal, M., Jensen, T., Werdelin, O., Mouritsen, S., *Cancer Res.* **1997**, *57*, 3214.
147. Mapara, M. Y., Sykes, M., *J. Clin. Oncol.* **2004**, *22*, 1136.
148. Gaidzik, N., Westerlind, U., Kunz, H., *Chem. Soc. Rev.* **2013**, *42*, 4421.

149. Christiansen, M. N., Chik, J., Lee, L., Anugraham, M., Abrahams, J. L., Packer, N. H., *Proteomics* **2014**, *14*, 525.
150. Cazet, A., Julien, S., Bobowski, M., Burchell, J., Delannoy, P., *Breast Cancer Res.* **2010**, *12*, 204.
151. Nakamori, S., Kameyama, M., Imaoka, S., Furukawa, H., Ishikawa, O., Sasaki, Y., Kabuto, T., Iwanaga, T., Matsushita, Y., Irimura, T., *Cancer Res.* **1993**, *53*, 3632.
152. Oda, K., Ido, A., Tamai, T., Matsushita, M., Kumagai, K., Mawatari, S., Saishoji, A., Kure, T., Ohno, K., Toyokura, E., Imanaka, D., Moriuchi, A., Uto, H., Oketani, M., Hashiguchi, T., Tsubouchi, H., *Oncol. Rep.* **2011**, *26*, 1227.
153. Johnson, P. J., Poon, T. C., Hjelm, N. M., Ho, C. S., Blake, C., Ho, S. K., *Br. J. Cancer* **2000**, *83*, 1330.
154. Nie, S., Lo, A., Wu, J., Zhu, J., Tan, Z., Simeone, D. M., Anderson, M. A., Shedden, K. A., Ruffin, M. T., Lubman, D. M., *J. Proteome Res.* **2014**, *13*, 1873.
155. Maguire, T. M., Gillian, A. M., O'Mahony, D., Coughlan, C. M., Dennihan, A., Breen, K. C., *Neurobiol. Aging* **1994**, *15*, 99.
156. Maguire, T. M., Breen, K. C., *Dementia* **1995**, *6*, 185.
157. Schedin-Weiss, S., Winblad, B., Tjernberg, L. O., *FEBS J* **2014**, *281*, 46.
158. Priller, C., Bauer, T., Mitteregger, G., Krebs, B., Kretzschmar, H. A., Herms, J., *J. Neurosci.* **2006**, *26*, 7212.
159. Turner, P. R., O'Connor, K., Tate, W. P., Abraham, W. C., *Prog. Neurobiol.* **2003**, *70*, 1.
160. Duce, J. A., Tsatsanis, A., Cater, M. A., James, S. A., Robb, E., Wikhe, K., Leong, S. L., Perez, K., Johanssen, T., Greenough, M. A., Cho, H. H., Galatis, D., Moir, R. D., Masters, C. L., McLean, C., Tanzi, R. E., Cappai, R., Barnham, K. J., Ciccotosto, G. D., Rogers, J. T., Bush, A. I., *Cell* **2010**, *142*, 857.
161. Akasaka-Manya, K., Manya, H., Sakurai, Y., Wojczyk, B. S., Spitalnik, S. L., Endo, T., *Glycoconj. J.* **2008**, *25*, 775.
162. Kitazume, S., Tachida, Y., Kato, M., Yamaguchi, Y., Honda, T., Hashimoto, Y., Wada, Y., Saito, T., Iwata, N., Saido, T., Taniguchi, N., *J. Biol. Chem.* **2010**, *285*, 40097.
163. Tomita, S., Kirino, Y., Suzuki, T., *J. Biol. Chem.* **1998**, *273*, 19304.
164. Jacobsen, K. T., Iverfeldt, K., *Biochem. Biophys. Res. Commun.* **2011**, *404*, 882.
165. Masters, C. L., Bateman, R., Blennow, K., Rowe, C. C., Sperling, R. A., Cummings, J. L., *Nat Rev Dis Primers* **2015**, *1*, 15056.
166. Shin, R. W., Iwaki, T., Kitamoto, T., Tateishi, J., *Lab. Invest.* **1991**, *64*, 693.
167. Iqbal, K., Liu, F., Gong, C. X., *Nat. Rev. Neurol.* **2016**, *12*, 15.
168. Liu, F., Zaidi, T., Iqbal, K., Grundke-Iqbal, I., Merkle, R. K., Gong, C. X., *FEBS Lett.* **2002**, *512*, 101.
169. Liu, F., Zaidi, T., Iqbal, K., Grundke-Iqbal, I., Gong, C. X., *Neuroscience* **2002**, *115*, 829.
170. Spillantini, M. G., Schmidt, M. L., Lee, V. M., Trojanowski, J. Q., Jakes, R., Goedert, M., *Nature* **1997**, *388*, 839.
171. Castellani, R., Smith, M. A., Richey, P. L., Perry, G., *Brain Res.* **1996**, *737*, 195.
172. Goers, J., Manning-Bog, A. B., McCormack, A. L., Millett, I. S., Doniach, S., Di Monte, D. A., Uversky, V. N., Fink, A. L., *Biochemistry* **2003**, *42*, 8465.
173. Hegde, M. L., Jagannatha Rao, K. S., *Arch. Biochem. Biophys.* **2003**, *418*, 169.
174. Steinman, L., *Nat. Immunol.* **2001**, *2*, 762.
175. Smolders, J., Damoiseaux, J., Menheere, P., Hupperts, R., *J. Neuroimmunol.* **2008**, *194*, 7.
176. International Multiple Sclerosis Genetics, C., Hafler, D. A., Compston, A., Sawcer, S., Lander, E. S., Daly, M. J., De Jager, P. L., de Bakker, P. I., Gabriel, S. B., Mirel, D. B., Ivinson, A. J.,

Pericak-Vance, M. A., Gregory, S. G., Rioux, J. D., McCauley, J. L., Haines, J. L., Barcellos, L. F., Cree, B., Oksenberg, J. R., Hauser, S. L., *N. Engl. J. Med.* **2007**, *357*, 851.

177. Filippi, M., Rocca, M. A., *J. Neurol.* **2005**, *252 Suppl 5*, v16.
178. Mkhikian, H., Grigorian, A., Li, C. F., Chen, H. L., Newton, B., Zhou, R. W., Beeton, C., Torossian, S., Tatarian, G. G., Lee, S. U., Lau, K., Walker, E., Siminovitch, K. A., Chandy, K. G., Yu, Z., Dennis, J. W., Demetriou, M., *Nat Commun* **2011**, *2*, 334.
179. Lai, Z. W., Nice, E. C., Schilling, O., *Proteomics* **2013**, *13*, 512.
180. Rudd, P., Karlsson, N. G., Khoo, K. H., Packer, N. H., in *Essentials of Glycobiology*, Cold Spring Harbor (NY), **2015**, pp. 653.
181. Shajahan, A., Heiss, C., Ishihara, M., Azadi, P., *Anal Bioanal Chem* **2017**, *409*, 4483.
182. Yang, Y., Franc, V., Heck, A. J. R., *Trends Biotechnol.* **2017**, *35*, 598.
183. Rudd, P. M., Dwek, R. A., *Crit. Rev. Biochem. Mol. Biol.* **1997**, *32*, 1.
184. Fornelli, L., Ayoub, D., Aizikov, K., Beck, A., Tsybin, Y. O., *Anal. Chem.* **2014**, *86*, 3005.
185. Tretter, V., Altmann, F., Marz, L., *Eur. J. Biochem.* **1991**, *199*, 647.
186. Endo, Y., Kobata, A., *J. Biochem.* **1976**, *80*, 1.
187. Alley, W. R., Jr., Novotny, M. V., *Annu Rev Anal Chem (Palo Alto Calif)* **2013**, *6*, 237.
188. Turyan, I., Hronowski, X., Sosic, Z., Lyubarskaya, Y., *Anal. Biochem.* **2014**, *446*, 28.
189. Kozak, R. P., Royle, L., Gardner, R. A., Fernandes, D. L., Wuhrer, M., *Anal. Biochem.* **2012**, *423*, 119.
190. Palmisano, G., Melo-Braga, M. N., Engholm-Keller, K., Parker, B. L., Larsen, M. R., *J. Proteome Res.* **2012**, *11*, 1949.
191. Kuster, B., Mann, M., *Anal. Chem.* **1999**, *71*, 1431.
192. Steentoft, C., Vakhrushev, S. Y., Vester-Christensen, M. B., Schjoldager, K. T., Kong, Y., Bennett, E. P., Mandel, U., Wandall, H., Levery, S. B., Clausen, H., *Nat Methods* **2011**, *8*, 977.
193. Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T., Lavrsen, K., Dabelsteen, S., Pedersen, N. B., Marcos-Silva, L., Gupta, R., Bennett, E. P., Mandel, U., Brunak, S., Wandall, H. H., Levery, S. B., Clausen, H., *EMBO J.* **2013**, *32*, 1478.
194. Vester-Christensen, M. B., Halim, A., Joshi, H. J., Steentoft, C., Bennett, E. P., Levery, S. B., Vakhrushev, S. Y., Clausen, H., *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 21018.
195. www.unicarb-db.com
196. Ruhaak, L. R., Deelder, A. M., Wuhrer, M., *Anal Bioanal Chem* **2009**, *394*, 163.
197. West, C., Elfakir, C., Lafosse, M., *J. Chromatogr. A* **2010**, *1217*, 3201.
198. Bruggink, C., Wuhrer, M., Koeleman, C. A., Barreto, V., Liu, Y., Pohl, C., Ingendoh, A., Hokke, C. H., Deelder, A. M., *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2005**, *829*, 136.
199. Both, P., Green, A. P., Gray, C. J., Sardzik, R., Voglmeir, J., Fontana, C., Austeri, M., Rejzek, M., Richardson, D., Field, R. A., Widmalm, G., Flitsch, S. L., Eyers, C. E., *Nat Chem* **2014**, *6*, 65.
200. Pagel, K., Harvey, D. J., *Anal. Chem.* **2013**, *85*, 5138.
201. Rodrigues, J. A., Taylor, A. M., Sumpton, D. P., Reynolds, J. C., Pickford, R., Thomas-Oates, J., *Adv. Carbohydr. Chem. Biochem.* **2007**, *61*, 59.
202. Kailemia, M. J., Ruhaak, L. R., Lebrilla, C. B., Amster, I. J., *Anal. Chem.* **2014**, *86*, 196.
203. Harvey, D. J., *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2011**, *879*, 1196.
204. Qian, J., Liu, T., Yang, L., Daus, A., Crowley, R., Zhou, Q., *Anal. Biochem.* **2007**, *364*, 8.
205. Christiansen, M. N., Kolarich, D., Nevalainen, H., Packer, N. H., Jensen, P. H., *Anal. Chem.* **2010**, *82*, 3500.
206. Zauner, G., Koeleman, C. A., Deelder, A. M., Wuhrer, M., *J. Sep. Sci.* **2010**, *33*, 903.

207. Nwosu, C. C., Seipert, R. R., Strum, J. S., Hua, S. S., An, H. J., Zivkovic, A. M., German, B. J., Lebrilla, C. B., *J. Proteome Res.* **2011**, *10*, 2612.
208. Pett, C., Nasir, W., Sihlbom, C., Olsson, B. M., Caixeta, V., Schorlemer, M., Zahedi, R. P., Larson, G., Nilsson, J., Westerlind, U., *Angew. Chem. Int. Ed. Engl.* **2018**, *57*, 9320.
209. Domon, B., Costello, C. E., *Biochemistry* **1988**, *27*, 1534.
210. Wu, S. W., Liang, S. Y., Pu, T. H., Chang, F. Y., Khoo, K. H., *J. Proteomics* **2013**, *84*, 1.
211. Halim, A., Nilsson, J., Ruetschi, U., Hesse, C., Larson, G., *Mol. Cell. Proteomics* **2012**, *11*, M111 013649.
212. Halim, A., Brinkmalm, G., Ruetschi, U., Westman-Brinkmalm, A., Portelius, E., Zetterberg, H., Blennow, K., Larson, G., Nilsson, J., *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 11848.
213. Trinidad, J. C., Schoepfer, R., Burlingame, A. L., Medzihradszky, K. F., *Mol. Cell. Proteomics* **2013**, *12*, 3474.
214. Liu, X., Wesdemiotis, C., *Eur J Mass Spectrom (Chichester)* **2015**, *21*, 713.
215. Alley, W. R., Jr., Mechref, Y., Novotny, M. V., *Rapid Commun. Mass Spectrom.* **2009**, *23*, 495.
216. Darula, Z., Medzihradszky, K. F., *Mol. Cell. Proteomics* **2009**, *8*, 2515.
217. Cao, L., Tolic, N., Qu, Y., Meng, D., Zhao, R., Zhang, Q., Moore, R. J., Zink, E. M., Lipton, M. S., Pasa-Tolic, L., Wu, S., *Anal. Biochem.* **2014**, *452*, 96.
218. Segu, Z. M., Mechref, Y., *Rapid Commun. Mass Spectrom.* **2010**, *24*, 1217.
219. Halim, A., Westerlind, U., Pett, C., Schorlemer, M., Ruetschi, U., Brinkmalm, G., Sihlbom, C., Lengqvist, J., Larson, G., Nilsson, J., *J. Proteome Res.* **2014**, *13*, 6024.
220. Yu, J., Schorlemer, M., Gomez Toledo, A., Pett, C., Sihlbom, C., Larson, G., Westerlind, U., Nilsson, J., *Chemistry* **2016**, *22*, 1114.
221. Yu, Q., Wang, B., Chen, Z., Urabe, G., Glover, M. S., Shi, X., Guo, L. W., Kent, K. C., Li, L., *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 1751.
222. Zhang, Y., Xie, X., Zhao, X., Tian, F., Lv, J., Ying, W., Qian, X., *J. Proteomics* **2018**, *170*, 14.
223. Perkins, D. N., Pappin, D. J. C., Creasy, D. M., Cottrell, J. S., *Electrophoresis* **1999**, *20*, 3551.
224. Bollineni, R. C., Koehler, C. J., Gislefoss, R. E., Anonsen, J. H., Thiede, B., *Sci. Rep.* **2018**, *8*, 2117.
225. Ceroni, A., Maass, K., Geyer, H., Geyer, R., Dell, A., Haslam, S. M., *J. Proteome Res.* **2008**, *7*, 1650.
226. Vakhrushev, S. Y., Dadimov, D., Peter-Katalinic, J., *Anal. Chem.* **2009**, *81*, 3252.
227. Goldberg, D., Sutton-Smith, M., Paulson, J., Dell, A., *Proteomics* **2005**, *5*, 865.
228. Apte, A., Meitei, N. S., *Methods Mol. Biol.* **2010**, *600*, 269.
229. Loss, A., Bunsmann, P., Bohne, A., Loss, A., Schwarzer, E., Lang, E., von der Lieth, C. W., *Nucleic Acids Res.* **2002**, *30*, 405.
230. Lohmann, K. K., von der Lieth, C. W., *Nucleic Acids Res.* **2004**, *32*, W261.
231. Maass, K., Ranzinger, R., Geyer, H., von der Lieth, C. W., Geyer, R., *Proteomics* **2007**, *7*, 4435.
232. Cooper, C. A., Gasteiger, E., Packer, N. H., *Proteomics* **2001**, *1*, 340.
233. Irungu, J., Go, E. P., Dalpathado, D. S., Desaire, H., *Anal. Chem.* **2007**, *79*, 3065.
234. Joenvaara, S., Ritamo, I., Peltoniemi, H., Renkonen, R., *Glycobiology* **2008**, *18*, 339.
235. Bern, M., Cai, Y., Goldberg, D., *Anal. Chem.* **2007**, *79*, 1393.
236. Nasir, W., Toledo, A. G., Noborn, F., Nilsson, J., Wang, M., Bandeira, N., Larson, G., *J. Proteome Res.* **2016**, *15*, 2826.

237. Liu, M. Q., Zeng, W. F., Fang, P., Cao, W. Q., Liu, C., Yan, G. Q., Zhang, Y., Peng, C., Wu, J. Q., Zhang, X. J., Tu, H. J., Chi, H., Sun, R. X., Cao, Y., Dong, M. Q., Jiang, B. Y., Huang, J. M., Shen, H. L., Wong, C. C. L., He, S. M., Yang, P. Y., *Nat Commun* **2017**, *8*, 438.

238. Walsh, I., Zhao, S., Campbell, M., Taron, C. H., Rudd, P. M., *Curr. Opin. Struct. Biol.* **2016**, *40*, 70.

239. von der Lieth, C. W., Freire, A. A., Blank, D., Campbell, M. P., Ceroni, A., Damerell, D. R., Dell, A., Dwek, R. A., Ernst, B., Fogh, R., Frank, M., Geyer, H., Geyer, R., Harrison, M. J., Henrick, K., Herget, S., Hull, W. E., Ionides, J., Joshi, H. J., Kamerling, J. P., Leeflang, B. R., Lutteke, T., Lundborg, M., Maass, K., Merry, A., Ranzinger, R., Rosen, J., Royle, L., Rudd, P. M., Schloissnig, S., Stenutz, R., Vranken, W. F., Widmalm, G., Haslam, S. M., *Glycobiology* **2011**, *21*, 493.

240. Raman, R., Venkataraman, M., Ramakrishnan, S., Lang, W., Raguram, S., Sasisekharan, R., *Glycobiology* **2006**, *16*, 82R.

241. Maeda, M., Fujita, N., Suzuki, Y., Sawaki, H., Shikanai, T., Narimatsu, H., *Methods Mol. Biol.* **2015**, *1273*, 161.

242. Sharon, N., *Trends Biochem. Sci* **1993**, *18*, 221.

243. Drickamer, K., *Nature* **1992**, *360*, 183.

244. Ostasiewicz, P., Zielinska, D. F., Mann, M., Wisniewski, J. R., *J. Proteome Res.* **2010**, *9*, 3688.

245. West, I., Goldring, O., *Methods Mol. Biol.* **2004**, *244*, 159.

246. Ernst-Cabrera, K., Wilchek, M., *Anal. Biochem.* **1986**, *159*, 267.

247. Bedair, M., El Rassi, Z., *J. Chromatogr. A* **2005**, *1079*, 236.

248. Zhang, W., Wang, H., Zhang, L., Yao, J., Yang, P., *Talanta* **2011**, *85*, 499.

249. Fan, X. L., She, Y. M., Bagshaw, R. D., Callahan, J. W., Schachter, H., Mahuran, D. J., *Anal. Biochem.* **2004**, *332*, 178.

250. Ongay, S., Boichenko, A., Govorukhina, N., Bischoff, R., *J. Sep. Sci.* **2012**, *35*, 2341.

251. Monzo, A., Bonn, G. K., Guttman, A., *TrAC, Trends Anal. Chem.* **2007**, *26*, 423.

252. Ohyama, Y., Kasai, K., Nomoto, H., Inoue, Y., *J. Biol. Chem.* **1985**, *260*, 6882.

253. Monsigny, M., Roche, A. C., Sene, C., Magetdana, R., Delmotte, F., *Eur. J. Biochem.* **1980**, *104*, 147.

254. Bhavanandan, V. P., Katlic, A. W., *J. Biol. Chem.* **1979**, *254*, 4000.

255. Saulsbury, F. T., *J. Rheumatol.* **1997**, *24*, 2246.

256. Lee, A., Nakano, M., Hincapie, M., Kolarich, D., Baker, M. S., Hancock, W. S., Packer, N. H., *OMICS* **2010**, *14*, 487.

257. Cummings, R. D., Kornfeld, S., *J. Biol. Chem.* **1982**, *257*, 11235.

258. Yang, Z., Hancock, W. S., *J. Chromatogr. A* **2004**, *1053*, 79.

259. Alpert, A. J., *J. Chromatogr.* **1990**, *499*, 177.

260. Linden, J. C., Lawhead, C. L., *J. Chromatogr.* **1975**, *105*, 125.

261. Palmer, J. K., *Anal. Lett.* **1975**, *8*, 215.

262. Zhang, J., Wang, D. I., *J. Chromatogr. B Biomed. Sci. Appl.* **1998**, *712*, 73.

263. Chen, C. C., Su, W. C., Huang, B. Y., Chen, Y. J., Tai, H. C., Obena, R. P., *Analyst* **2014**, *139*, 688.

264. Zauner, G., Deelder, A. M., Wuhrer, M., *Electrophoresis* **2011**, *32*, 3456.

265. Selman, M. H., McDonnell, L. A., Palmblad, M., Ruhaak, L. R., Deelder, A. M., Wuhrer, M., *Anal. Chem.* **2010**, *82*, 1073.

266. Selman, M. H., Hemayatkar, M., Deelder, A. M., Wuhrer, M., *Anal. Chem.* **2011**, *83*, 2492.

267. Mysling, S., Palmisano, G., Hojrup, P., Thaysen-Andersen, M., *Anal. Chem.* **2010**, *82*, 5598.

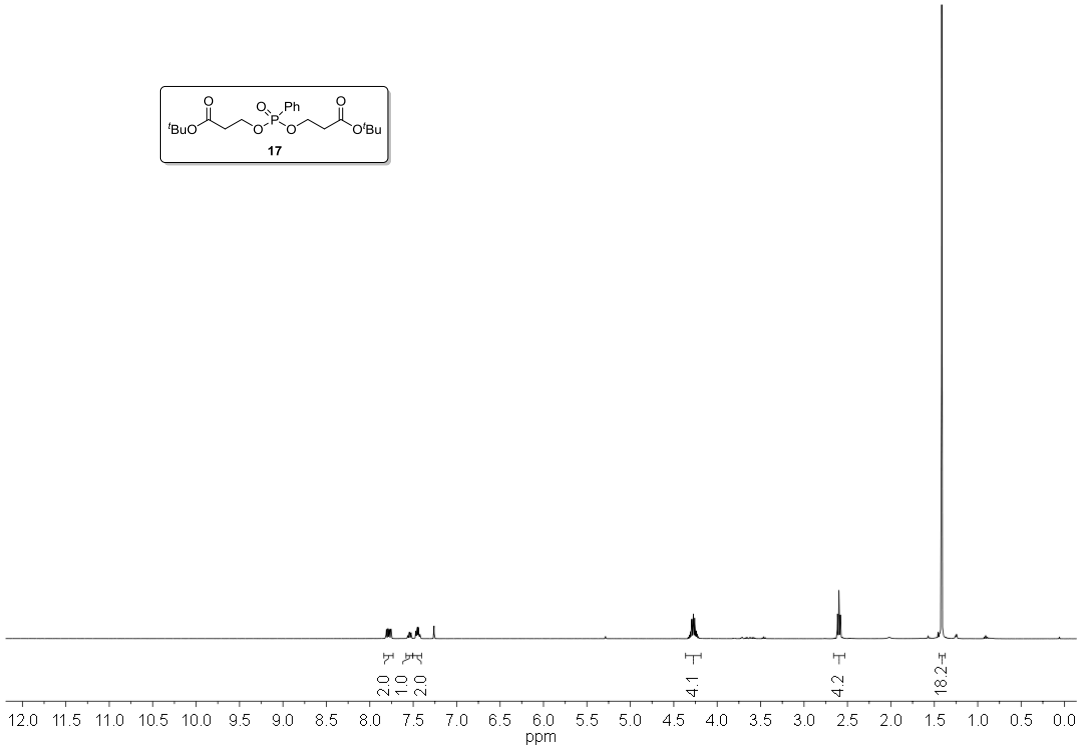
268. Wohlgemuth, J., Karas, M., Eichhorn, T., Hendriks, R., Andrecht, S., *Anal. Biochem.* **2009**, *395*, 178.
269. Takegawa, Y., Deguchi, K., Ito, H., Keira, T., Nakagawa, H., Nishimura, S., *J. Sep. Sci.* **2006**, *29*, 2533.
270. Alpert, A. J., *Anal. Chem.* **2008**, *80*, 62.
271. Lewandrowski, U., Lohrig, K., Zahedi, R. P., Wolters, D., Sickmann, A., *Clin. Proteomics* **2008**, *4*, 25.
272. Hao, P., Guo, T., Sze, S. K., *PLoS One* **2011**, *6*, e16884.
273. Cao, W., Huang, J., Jiang, B., Gao, X., Yang, P., *Sci. Rep.* **2016**, *6*, 29776.
274. Liu, Y. J., Liu, Y. J., Zhang, D., Zhang, R. Q., Li, Z. L., *Anal. Chem.* **2016**, *88*, 1067.
275. Li, X. L., Xiong, Y. T., Qing, G. Y., Jiang, G., Li, X. Q., Sun, T. L., Liang, X. M., *Acs Applied Materials & Interfaces* **2016**, *8*, 13294.
276. Kayser, H., Zeitler, R., Kannicht, C., Grunow, D., Nuck, R., Reutter, W., *J. Biol. Chem.* **1992**, *267*, 16934.
277. Boyce, M., Bertozzi, C. R., *Nat Methods* **2011**, *8*, 638.
278. Prescher, J. A., Dube, D. H., Bertozzi, C. R., *Nature* **2004**, *430*, 873.
279. Sawa, M., Hsu, T. L., Itoh, T., Sugiyama, M., Hanson, S. R., Vogt, P. K., Wong, C. H., *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 12371.
280. Agard, N. J., Prescher, J. A., Bertozzi, C. R., *J. Am. Chem. Soc.* **2004**, *126*, 15046.
281. Laughlin, S. T., Baskin, J. M., Amacher, S. L., Bertozzi, C. R., *Science* **2008**, *320*, 664.
282. Champasa, K., Longwell, S. A., Eldridge, A. M., Stemmler, E. A., Dube, D. H., *Mol. Cell. Proteomics* **2013**, *12*, 2568.
283. Xie, R., Hong, S., Feng, L., Rong, J., Chen, X., *J. Am. Chem. Soc.* **2012**, *134*, 9914.
284. Xie, R., Dong, L., Huang, R., Hong, S., Lei, R., Chen, X., *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 14082.
285. Griffin, M. E., Jensen, E. H., Mason, D. E., Jenkins, C. L., Stone, S. E., Peters, E. C., Hsieh-Wilson, L. C., *Mol Biosyst* **2016**, *12*, 1756.
286. Khidekel, N., Arndt, S., Lamarre-Vincent, N., Lippert, A., Poulin-Kerstien, K. G., Ramakrishnan, B., Qasba, P. K., Hsieh-Wilson, L. C., *J. Am. Chem. Soc.* **2003**, *125*, 16162.
287. Wen, L. Q., Liu, D., Zheng, Y., Huang, K., Cao, X. F., Song, J., Wang, P. G., *Acs Central Science* **2018**, *4*, 451.
288. Rawn, J. D., Lienhard, G. E., *Biochemistry* **1974**, *13*, 3124.
289. Xu, Y., Bailey, U. M., Punyadeera, C., Schulz, B. L., *Rapid Commun. Mass Spectrom.* **2014**, *28*, 471.
290. Krisp, C., Kubutat, C., Kyas, A., Steinstrasser, L., Jacobsen, F., Wolters, D., *J. Proteomics* **2011**, *74*, 502.
291. Chen, W. X., Smeekens, J. M., Wu, R. H., *Mol. Cell. Proteomics* **2014**, *13*, 1563.
292. Xiao, H., Chen, W., Smeekens, J. M., Wu, R., *Nat Commun* **2018**, *9*, 1692.
293. Pinkse, M. W. H., Uitto, P. M., Hilhorst, M. J., Ooms, B., Heck, A. J. R., *Anal. Chem.* **2004**, *76*, 3935.
294. Larsen, M. R., Jensen, S. S., Jakobsen, L. A., Heegaard, N. H., *Mol. Cell. Proteomics* **2007**, *6*, 1778.
295. Thaysen-Andersen, M., Larsen, M. R., Packer, N. H., Palmisano, G., *Rsc Advances* **2013**, *3*, 22683.
296. Bobbitt, J. M., *Adv Carbohydr Chem* **1956**, *48*, 1.

297. Fleury, P. F., Courtois, J. E., Bieder, A., *Bull. Soc. Chim. Fr.* **1952**, *19*, 118.
298. Zhang, H., Li, X. J., Martin, D. B., Aebersold, R., *Nat. Biotechnol.* **2003**, *21*, 660.
299. Tian, Y., Zhou, Y., Elliott, S., Aebersold, R., Zhang, H., *Nat. Protoc.* **2007**, *2*, 334.
300. Zhou, Y., Aebersold, R., Zhang, H., *Anal. Chem.* **2007**, *79*, 5826.
301. Berven, F. S., Ahmad, R., Clauser, K. R., Carr, S. A., *J. Proteome Res.* **2010**, *9*, 1706.
302. Klement, E., Lipinszki, Z., Kupihar, Z., Udvardy, A., Medzihradszky, K. F., *J. Proteome Res.* **2010**, *9*, 2200.
303. Clamp, J. R., Hough, L., *Biochem. J* **1965**, *94*, 17.
304. Huang, J., Qin, H., Sun, Z., Huang, G., Mao, J., Cheng, K., Zhang, Z., Wan, H., Yao, Y., Dong, J., Zhu, J., Wang, F., Ye, M., Zou, H., *Sci. Rep.* **2015**, *5*, 10164.
305. Huang, J. F., Wan, H., Yao, Y. T., Li, J. N., Cheng, K., Mao, J. W., Chen, J., Wang, Y., Qin, H. Q., Zhang, W. B., Ye, M. L., Zou, H. F., *Anal. Chem.* **2015**, *87*, 10199.
306. Nilsson, J., Ruetschi, U., Halim, A., Hesse, C., Carlsohn, E., Brinkmalm, G., Larson, G., *Nat Methods* **2009**, *6*, 809.
307. Kurogochi, M., Amano, M., Fumoto, M., Takimoto, A., Kondo, H., Nishimura, S., *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 8808.
308. Kurogochi, M., Matsushista, T., Amano, M., Furukawa, J., Shinohara, Y., Aoshima, M., Nishimura, S., *Mol. Cell. Proteomics* **2010**, *9*, 2354.
309. Cooper, J. A., Smith, W., Bacila, M., Medina, H., *J. Biol. Chem.* **1959**, *234*, 445.
310. Avigad, G., Amaral, D., Asensio, C., Horecker, B. L., *J. Biol. Chem.* **1962**, *237*, 2736.
311. Taga, Y., Kusubata, M., Ogawa-Goto, K., Hattori, S., *Mol. Cell. Proteomics* **2012**, *11*, M111 010397.
312. Ramya, T. N., Weerapana, E., Cravatt, B. F., Paulson, J. C., *Glycobiology* **2013**, *23*, 211.
313. Zheng, J., Xiao, H., Wu, R., *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 7107.
314. Madera, M., Mann, B., Mechref, Y., Novotny, M. V., *J. Sep. Sci.* **2008**, *31*, 2722.
315. Leon, M. A., *Science* **1967**, *158*, 1325.
316. Gonzaga, H. T., Ribeiro Vda, S., Cunha-Junior, J. P., Ueta, M. T., Costa-Cruz, J. M., *Diagn. Microbiol. Infect. Dis.* **2011**, *70*, 78.
317. Alagesan, K., Khilji, S. K., Kolarich, D., *Analytical and Bioanalytical Chemistry* **2017**, *409*, 529.
318. Karlesky, D., Shelly, D. C., Warner, I., *Anal. Chem.* **1981**, *53*, 2146.
319. Kalia, J., Raines, R. T., *Angew. Chem. Int. Ed. Engl.* **2008**, *47*, 7523.
320. Clayden, J., Greeves, N., Warren, S. G., *Organic chemistry*, Oxford University Press, Oxford; New YorK, **2012**.
321. Hahne, H., Neubert, P., Kuhn, K., Etienne, C., Bomgarden, R., Rogers, J. C., Kuster, B., *Anal. Chem.* **2012**, *84*, 3716.
322. Udenfriend, S., Stein, S., Böhlen, P., Dairman, W., Leimgruber, W., Weigele, M., *Science* **1972**, *178*, 871.
323. Yin, D. T., Urresti, S., Lafond, M., Johnston, E. M., Derikvand, F., Ciano, L., Berrin, J. G., Henrissat, B., Walton, P. H., Davies, G. J., Brumer, H., *Nat Commun* **2015**, *6*, 10197.
324. Humphreys, K. J., Mirica, L. M., Wang, Y., Klinman, J. P., *J. Am. Chem. Soc.* **2009**, *131*, 4657.
325. Pedersen, K. O., *Nature* **1944**, *154*, 575.
326. Windwarder, M., Altmann, F., *J. Proteomics* **2014**, *108*, 258.
327. Demetriou, M., Binkert, C., Sukhu, B., Tenenbaum, H. C., Dennis, J. W., *J. Biol. Chem.* **1996**, *271*, 12755.
328. Hoffmann, M., Marx, K., Reichl, U., Wuhrer, M., Rapp, E., *Mol. Cell. Proteomics* **2016**, *15*, 624.

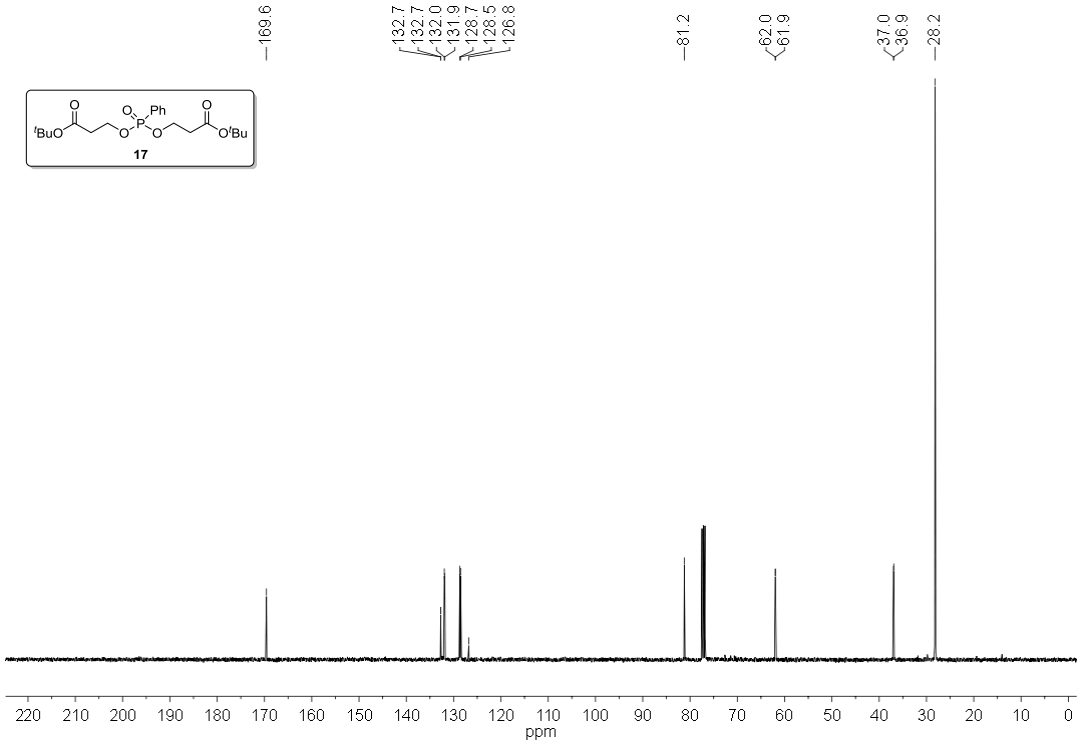
329. Guldbrandsen, A., Vethe, H., Farag, Y., Oveland, E., Garberg, H., Berle, M., Myhr, K. M., Opsahl, J. A., Barsnes, H., Berven, F. S., *Mol. Cell. Proteomics* **2014**, *13*, 3152.
330. Vaudel, M., Barsnes, H., Berven, F. S., Sickmann, A., Martens, L., *Proteomics* **2011**, *11*, 996.
331. Craig, R., Beavis, R. C., *Bioinformatics* **2004**, *20*, 1466.
332. Tabb, D. L., Fernando, C. G., Chambers, M. C., *J. Proteome Res.* **2007**, *6*, 654.
333. Eng, J. K., Jahan, T. A., Hoopmann, M. R., *Proteomics* **2013**, *13*, 22.
334. Vaudel, M., Burkhart, J. M., Zahedi, R. P., Oveland, E., Berven, F. S., Sickmann, A., Martens, L., Barsnes, H., *Nat. Biotechnol.* **2015**, *33*, 22.
335. http://compomics.github.io/projects/reporter.html (accessed August 2018)
336. Vaudel, M., Breiter, D., Beck, F., Rahnenfuhrer, J., Martens, L., Zahedi, R. P., *Proteomics* **2013**, *13*, 1036.
337. Benjamini, Y., Hochberg, Y., *Journal of the Royal Statistical Society Series B-Methodological* **1995**, *57*, 289.
338. Reiner, A., Yekutieli, D., Benjamini, Y., *Bioinformatics* **2003**, *19*, 368.
339. http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_9072.pdf (accessed July 2018)
340. http://wolfson.huji.ac.il/purification/PDF/Thiophilic_Chromatography/AMERSHAM_ThiolSepharose.pdf (accessed July 2018)
341. Kolarich, D., Weber, A., Turecek, P. L., Schwarz, H. P., Altmann, F., *Proteomics* **2006**, *6*, 3369.
342. Liu, T., Qian, W. J., Gritsenko, M. A., Camp, D. G., 2nd, Monroe, M. E., Moore, R. J., Smith, R. D., *J. Proteome Res.* **2005**, *4*, 2070.
343. Chen, R., Jiang, X. N., Sun, D. G., Han, G. H., Wang, F. J., Ye, M. L., Wang, L. M., Zou, H. F., *J. Proteome Res.* **2009**, *8*, 651.
344. Jia, W., Lu, Z., Fu, Y., Wang, H. P., Wang, L. H., Chi, H., Yuan, Z. F., Zheng, Z. B., Song, L. N., Han, H. H., Liang, Y. M., Wang, J. L., Cai, Y., Zhang, Y. K., Deng, Y. L., Ying, W. T., He, S. M., Qian, X. H., *Mol. Cell. Proteomics* **2009**, *8*, 913.
345. Bunkenborg, J., Pilch, B. J., Podtelejnikov, A. V., Wisniewski, J. R., *Proteomics* **2004**, *4*, 454.
346. Rannes, J. B., Ioannou, A., Willies, S. C., Grogan, G., Behrens, C., Flitsch, S. L., Turner, N. J., *J. Am. Chem. Soc.* **2011**, *133*, 8436.
347. Gottlieb, H. E., Kotlyar, V., Nudelman, A., *J. Org. Chem.* **1997**, *62*, 7512.
348. van Kalkeren, H. A., Leenders, S. H., Hommersom, C. R., Rutjes, F. P., van Delft, F. L., *Chemistry* **2011**, *17*, 11290.
349. Jones, D. S., Hammaker, J. R., Tedder, M. E., *Tetrahedron Lett.* **2000**, *41*, 1531.
350. Jones, D. S., Branks, M. J., Campbell, M. A., Cockerill, K. A., Hammaker, J. R., Kessler, C. A., Smith, E. M., Tao, A., Ton-Nu, H. T., Xu, T., *Bioconjug. Chem.* **2003**, *14*, 1067.
351. http://www.thegpm.org/cRAP/index.html

Appendix

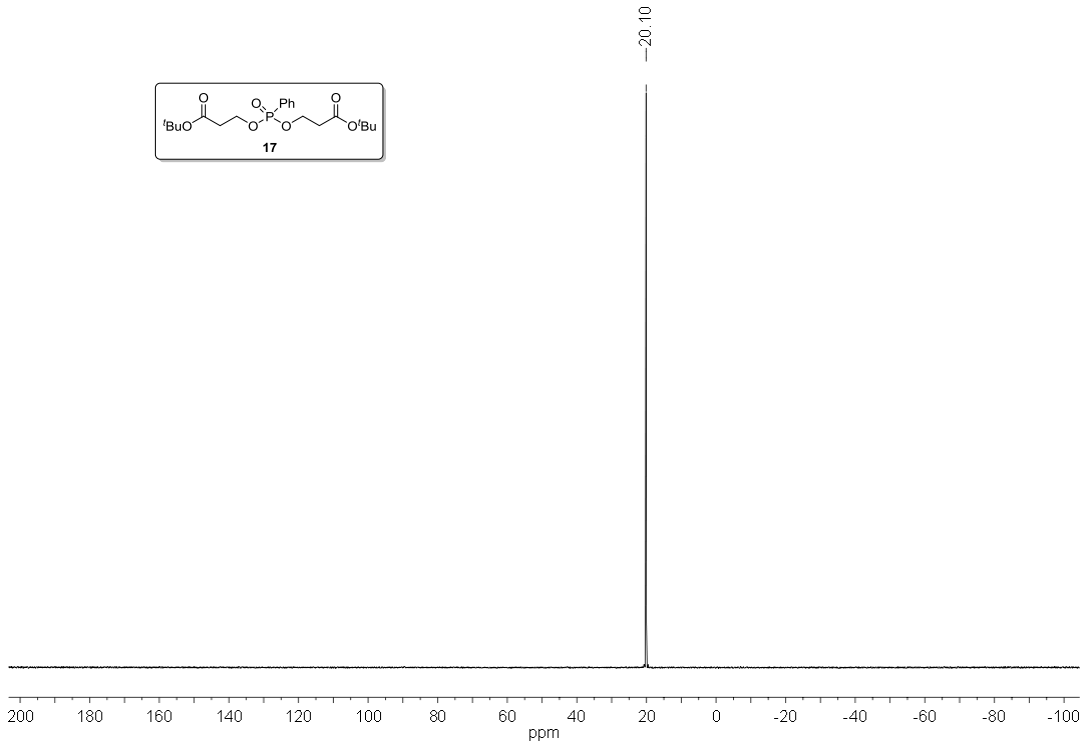
NMR spectra



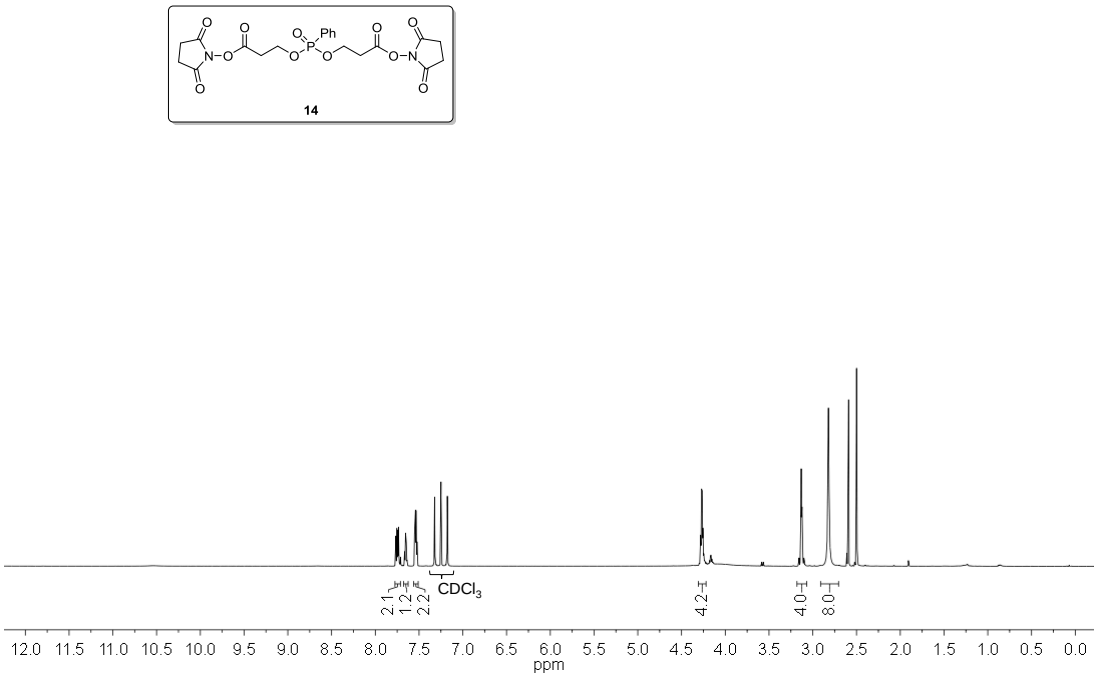
Appendix 1. ^{1}H-NMR of compound **17** in $CDCl_3$ (400MHz).



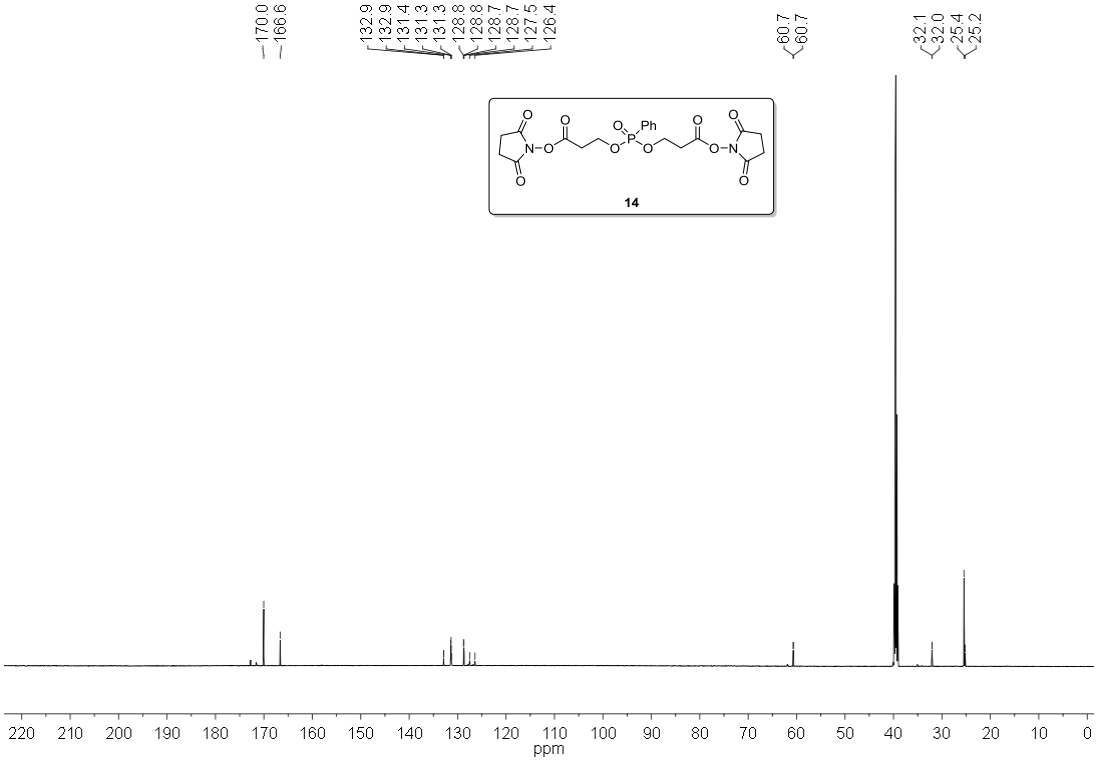
Appendix 2. ^{13}C-NMR of compound **17** in $CDCl_3$ (101MHz).



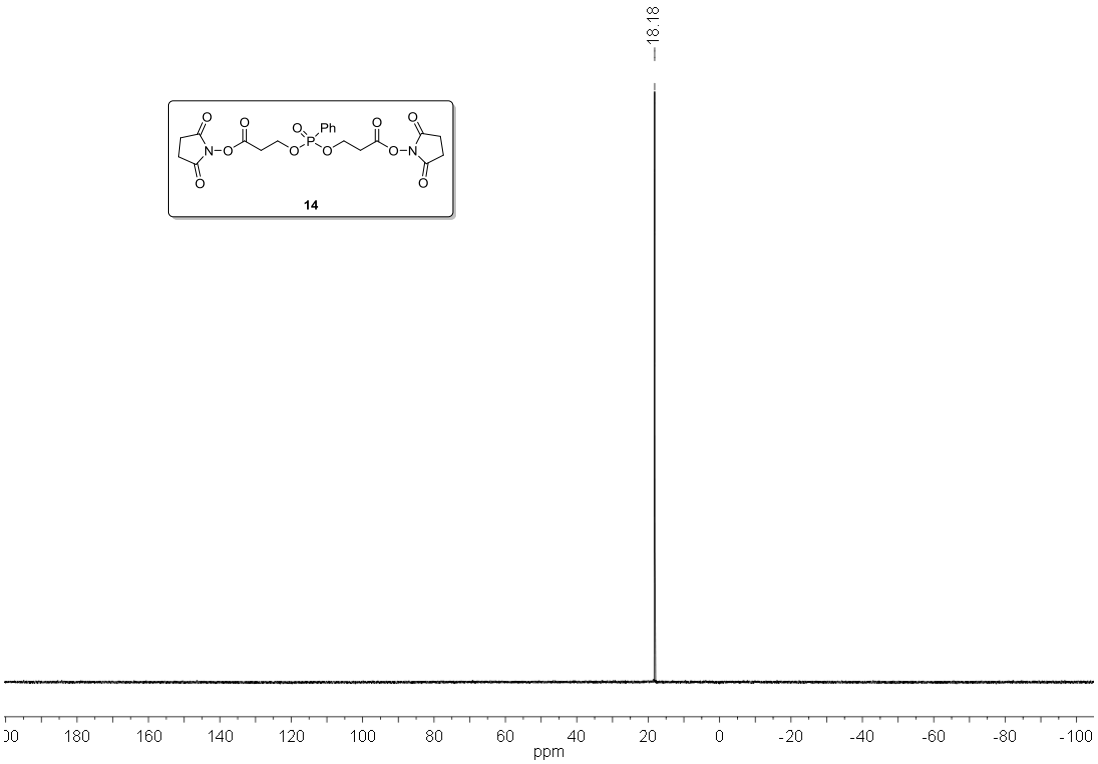
Appendix 3. ^{31}P-NMR of compound **17** in $CDCl_3$ (162MHz).



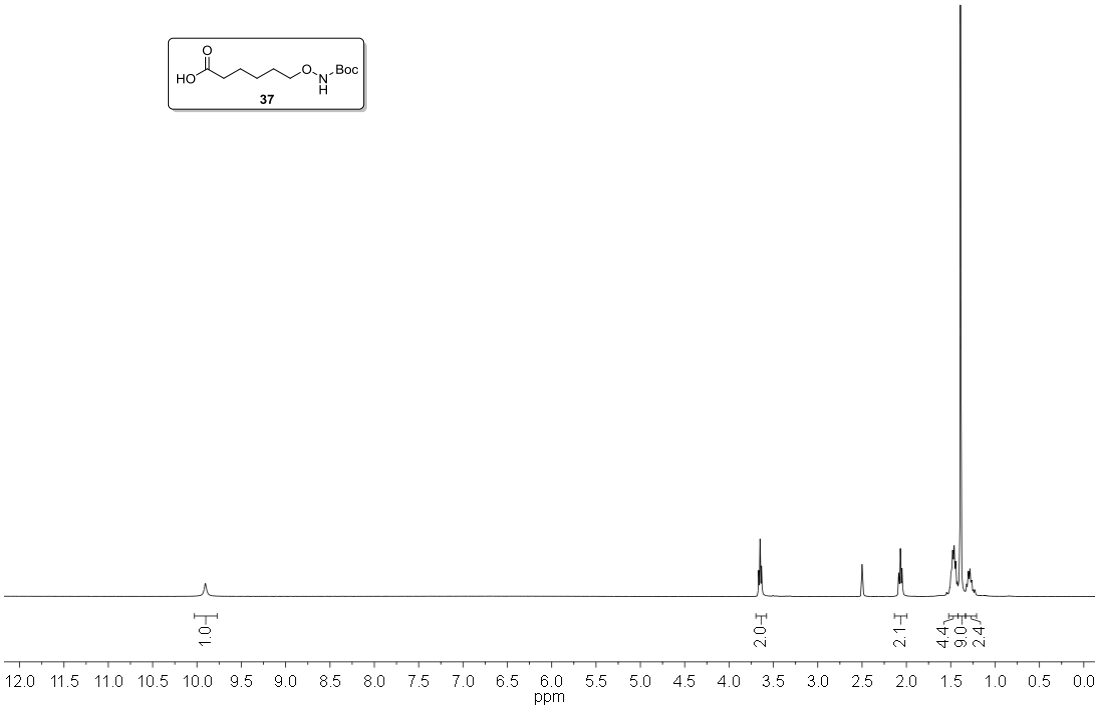
Appendix 4. ^{1}H-NMR of compound **14** in DMSO-d6 (700MHz).



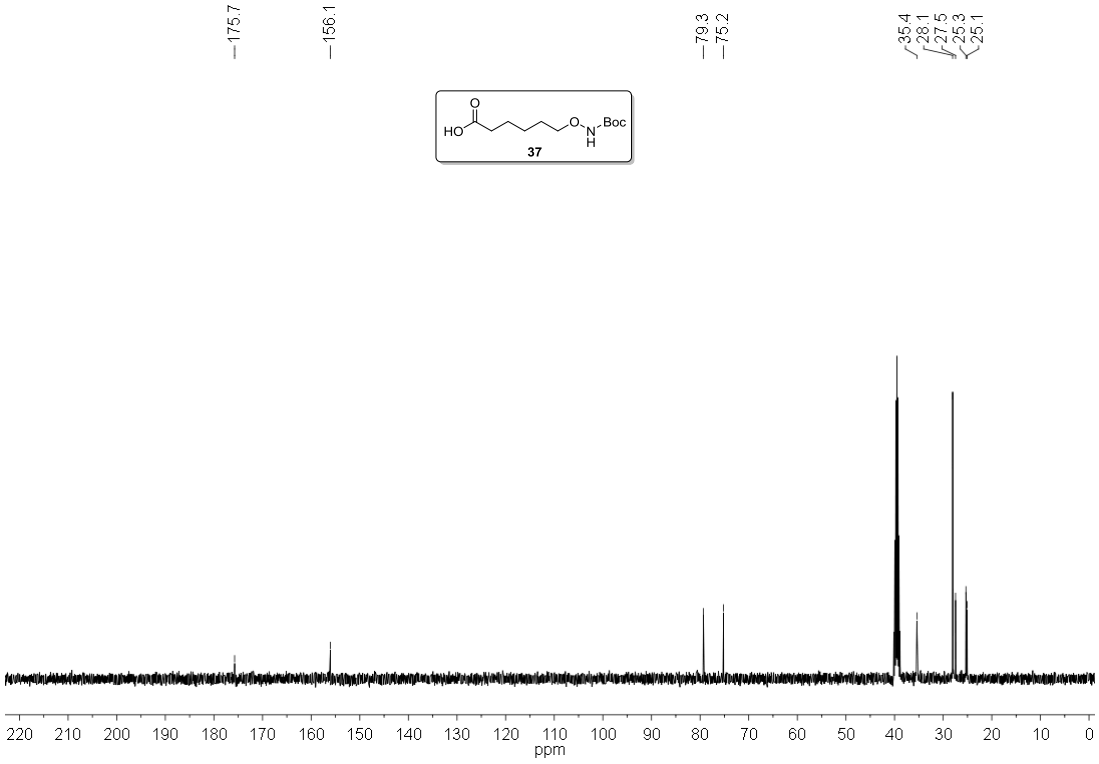
Appendix 5. ^{13}C-NMR of compound **14** in DMSO-d6 (176MHz).



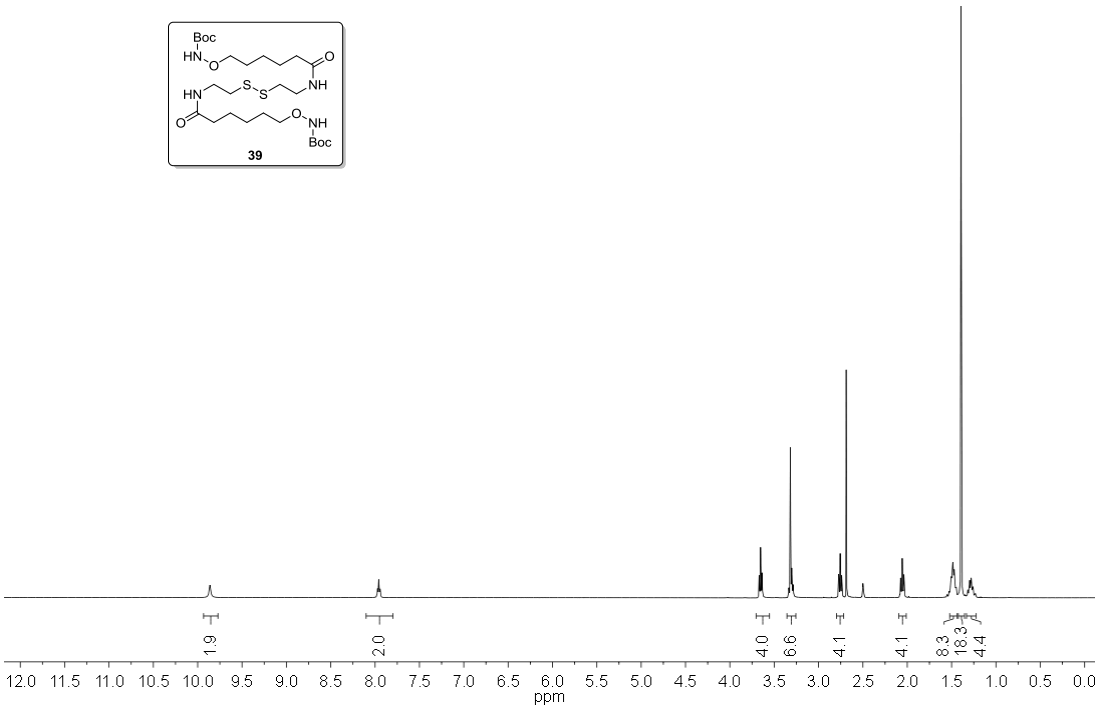
Appendix 6. ^{31}P-NMR of compound **14** in DMSO-d6 (283MHz).



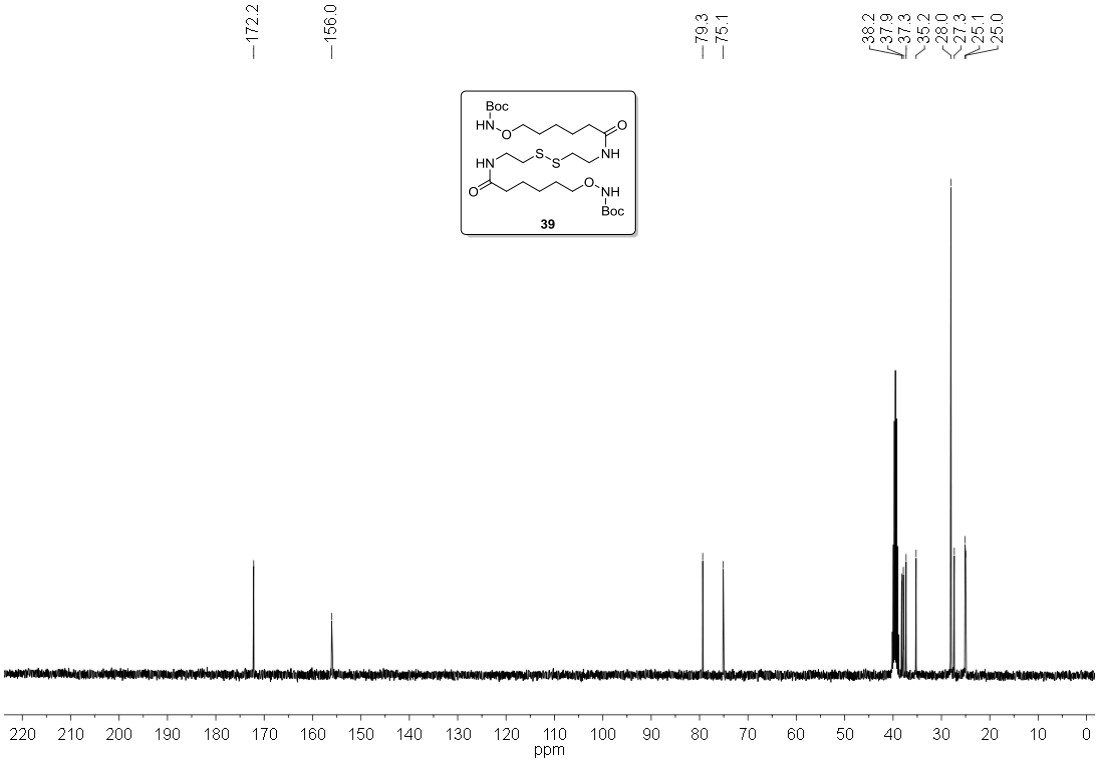
Appendix 7. [1]H-NMR of compound **37** in DMSO-d6 (400MHz).



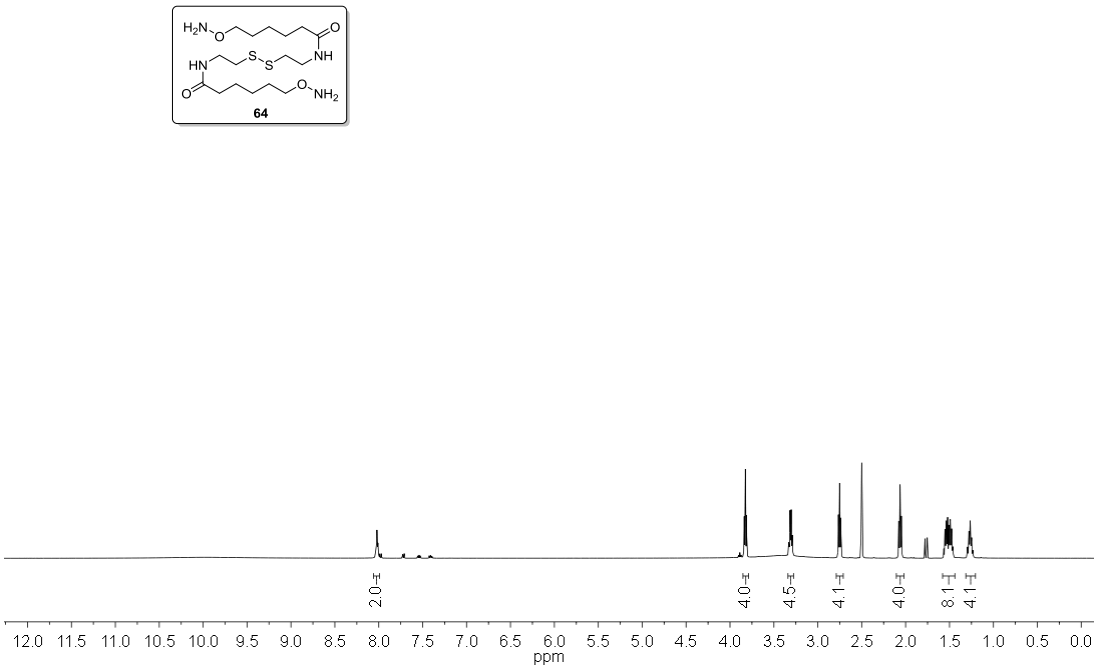
Appendix 8. [13]C-NMR of compound **37** in DMSO-d6 (101MHz).



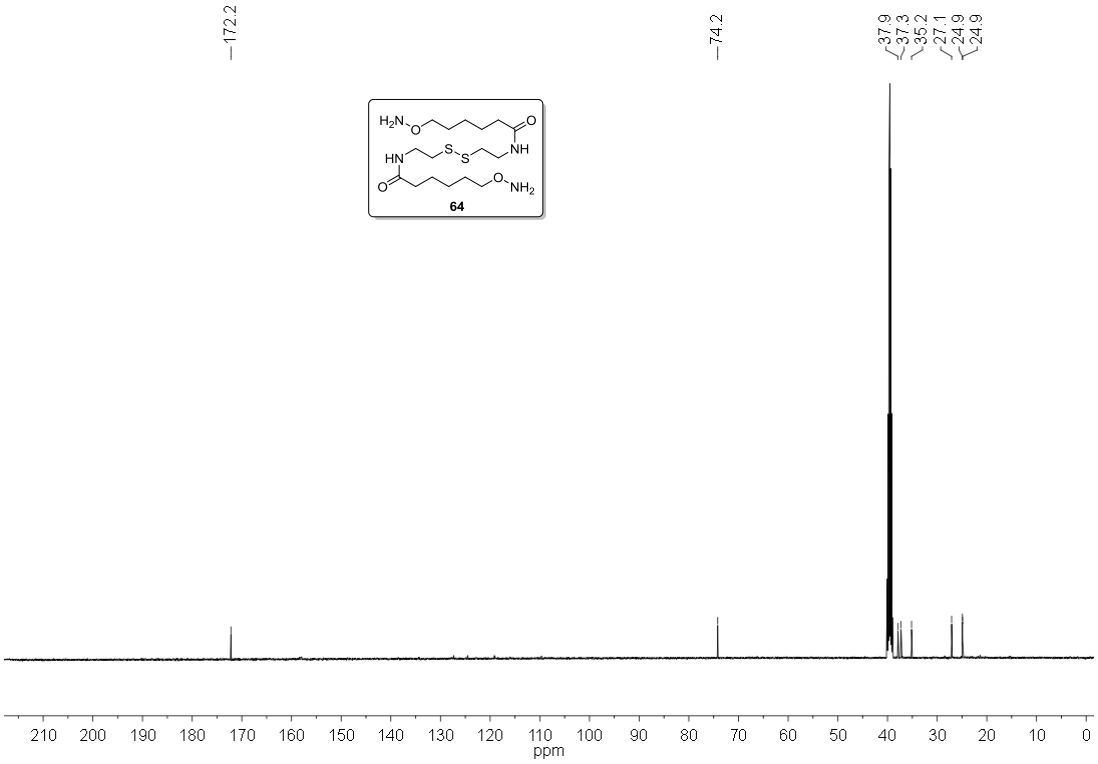
Appendix 9. [1]H-NMR of compound **39** in DMSO-d6 (400MHz).



Appendix 10. [13]C-NMR of compound **39** in DMSO-d6 (101MHz).



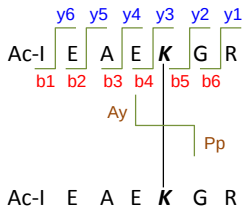
Appendix 11. ^{1}H-NMR of compound **64** in DMSO-d6 (500MHz).



Appendix 12. ^{13}C-NMR of compound **64** in DMSO-d6 (126MHz).

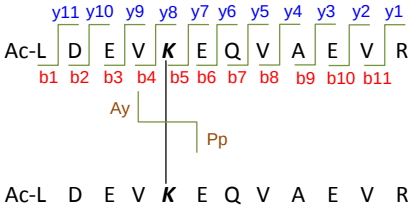
Development of an MS-cleavable cross-linker

Table A 1. Theoretical masses and fragments for the cross-linked peptide Ac-IEAEKGR (**27**) used as database for ions searches.



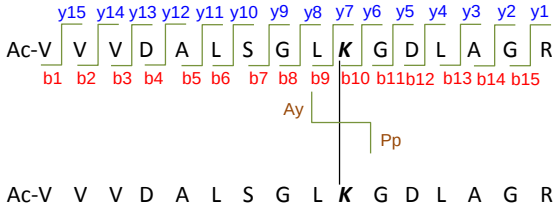
Ac-IEAE***K***GR						
		***m/z* 1+**	***m/z* 2+**	***m/z* 3+**	***m/z* 4+**	**Sequence**
Cross-linked peptide		1953,9323	977,4698	651,9823	489,2390	
Peptide fragment ions	b1	-	-	-		Ac-I
	b2	285,1445	143,0759	95,7197		Ac-IE
	b3	356,1816	178,5944	119,3987		Ac-IEA
	b4	485,2242	243,1157	162,4129		Ac-IEAE
	b5	1722,7992	861,9032	574,9379		Ac-IEAE***K***
	b6	1779,8206	890,414	593,9451		Ac-IEAE***K***G
	y6	1798,8377	899,9225	600,2841		EAE***K***GR
	y5	1669,7951	835,4012	557,2699		AE***K***GR
	y4	1598,7580	799,8826	533,5908		E***K***GR
	y3	1469,7154	735,3613	490,5766		***K***GR
	y2	232,1404	116,5738	78,0517		GR
	y1	175,1190	88,0631	59,0445		R
Linker fragment ions	Acryloyl (Ay)	898,4629	449,7351	300,1591		
	Phosphonate (Pp)	1056,4762	528,7417	352,8302		

Table A 2. Theoretical masses and fragments for the cross-linked peptide Ac-LDEVKEQVAEVR (**28**) used as database for ions searches.



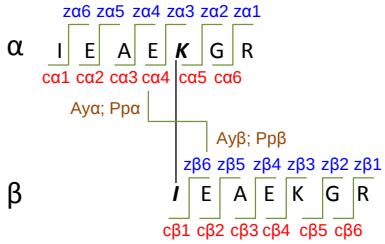
Ac-LDEV***K***EQVAEVR						
		***m/z* 1+**	***m/z* 2+**	***m/z* 3+**	***m/z* 4+**	**Sequence**
Cross-linked peptide		3178,5561	1589,7817	1060,1902	795,3950	
Peptide fragment ions	b1	-	-	-		Ac-L
	b2	271,1288	136,0681	91,0478		Ac-LD
	b3	400,1714	200,5894	134,0620		Ac-LDE
	b4	499,2399	250,1236	167,0848		Ac-LDEV
	b5	2349,1267	1175,0670	783,7138		Ac-LDEV***K***
	b6	2478,1693	1239,5883	826,7280		Ac-LDEV***K***E
	b7	2606,2279	1303,6176	869,4141		Ac-LDEV***K***EQ
	b8	2705,2963	1353,1518	902,4370		Ac-LDEV***K***EQV
	b9	2776,3334	1388,6703	926,1160		Ac-LDEV***K***EQVA
	b10	2905,3760	1453,1916	969,1302		Ac-LDEV***K***EQVAE
	b11	3004,4444	1502,7258	1002,1530		Ac-LDEV***K***EQVAEV
	y11	3023,4615	1512,2344	1008,4920		DEV***K***EQVAEVR
	y10	2908,4345	1454,7209	970,1497		EV***K***EQVAEVR
	y9	2779,3919	1390,1996	927,1355		V***K***EQVAEVR
	y8	2680,3235	1340,6654	894,1127		***K***EQVAEVR
	y7	830,4367	415,7220	277,4837		EQVAEVR
	y6	701,3941	351,2007	234,4695		QVAEVR
	y5	573,3355	287,1714	191,7833		VAEVR
	y4	474,2671	237,6372	158,7605		AEVR
	y3	403,2300	202,1186	135,0815		EVR
	y2	274,1874	137,5973	92,0673		VR
	y1	175,1190	88,0631	59,0445		R
Linker fragment ions	Acryloyl (Ay)	1510,7748	755,8910	504,2631		
	Phosphonate (Pp)	1668,7880	834,8977	556,9342		

Table A 3. Theoretical masses and fragments for the cross-linked peptide Ac-VVVDALSGLKGDLAGR (**29**) used as database for ions searches.



Ac-VVVDALSGL***K***GDLAGR						
		***m/z* 1+**	***m/z* 2+**	***m/z* 3+**	***m/z* 4+**	**Sequence**
Cross-linked peptide		3488,8405	1744,9239	1163,6184	872,9661	
Peptide fragment ions	b1	-	-	-		Ac-V
	b2	241,1547	121,0810	81,0564		Ac-VV
	b3	340,2231	170,6152	114,0792		Ac-VVV
	b4	455,2500	228,1287	152,4215		Ac-VVVD
	b5	526,2871	263,6472	176,1006		Ac-VVVDA
	b6	639,3712	320,1892	213,7953		Ac-VVVDAL
	b7	726,4032	363,7053	242,8059		Ac-VVVDALS
	b8	783,4247	392,2160	261,8131		Ac-VVVDALSG
	b9	896,5088	448,7580	299,5078		Ac-VVVDALSGL
	b10	2901,5378	1451,2725	967,8508		Ac-VVVDALSGL***K***
	b11	2958,5593	1479,7833	986,8579		Ac-VVVDALSGL***K***G
	b12	3073,5862	1537,2968	1025,2003		Ac-VVVDALSGL***K***GD
	b13	3186,6703	1593,8388	1062,8949		Ac-VVVDALSGL***K***GDL
	b14	3257,7074	1629,3573	1086,5740		Ac-VVVDALSGL***K***GDLA
	b15	3314,7289	1657,8681	1105,5811		Ac-VVVDALSGL***K***GDLAG
	y15	3347,7616	1674,3844	1116,5920		VVDALSGL***K***GDLAGR
	y14	3248,6932	1624,8502	1083,5692		VDALSGL***K***GDLAGR
	y13	3149,6247	1575,3160	1050,5464		DALSGL***K***GDLAGR
	y12	3034,5978	1517,8025	1012,2041		ALSGL***K***GDLAGR
	y11	2963,5607	1482,2840	988,5251		LSGL***K***GDLAGR
	y10	2850,4766	1425,7419	950,8304		SGL***K***GDLAGR
	y9	2763,4446	1382,2259	921,8197		GL***K***GDLAGR
	y8	2706,4231	1353,7152	902,8126		L***K***GDLAGR
	y7	2593,3391	1297,1732	865,1179		***K***GDLAGR
	y6	588,3100	294,6586	196,7749		GDLAGR
	y5	531,2885	266,1479	177,7677		DLAGR
	y4	416,2616	208,6344	139,4254		LAGR
	y3	303,1775	152,0924	101,7307		AGR
	y2	232,1404	116,5738	78,0517		GR
	y1	175,1190	88,0631	59,0445		R
Linker fragment ions	Acryloyl (Ay)	1665,9170	833,4621	555,9772		
	Phosphonate (Pp)	1823,9303	912,4688	608,6483		

Table A 4. Theoretical masses and fragments for the cross-linked peptide IEAEKGR (**32**) used as database for ions searches.



IEAEKGR						
		***m/z* 1+**	***m/z* 2+**	***m/z* 3+**	***m/z* 4+**	**Sequence**
Cross-linked peptide		1869,9111	935,4592	623,9752	468,2338	
Peptide fragment ions	cα1	-	-	-		I
	cα2	260,1605	130,5839	87,3917		IE
	cα3	331,1976	166,1025	111,0707		IEA
	cα4	460,2402	230,6238	154,0849		IEAE
	cα5	1655,8046	828,4060	552,6064		IEAE***K***
	cα6	1712,8260	856,9167	571,6135		IEAE***K***G
	cβ1	-	-	-		***I***
	cβ2	1327,6299	664,3186	443,2148		***I***E
	cβ3	1398,6670	699,8372	466,8939		***I***EA
	cβ4	1527,7096	764,3585	509,9081		***I***EAE
	cβ5	1655,8046	828,4060	552,6064		***I***EAEK
	cβ6	1712,8260	856,9167	571,6135		***I***EAEKG
	zα6	1740,8084	870,9079	580,9410		EAE***K***GR
	zα5	1611,7658	806,3866	537,9268		AE***K***GR
	zα4	1540,7286	770,8680	514,2477		E***K***GR
	zα3	1411,6861	706,3467	471,2336		***K***GR
	zα2	216,1217	108,5645	72,7121		GR
	zα1	159,1002	80,0538	53,7049		R
	zβ6	673,3390	337,1732	225,1179		EAEKGR
	zβ5	544,2964	272,6519	182,1037		AEKGR
	zβ4	473,2592	237,1333	158,4246		EKGR
	zβ3	344,2167	172,6120	115,4104		KGR
	zβ2	216,1217	108,5645	72,7121		GR
	zβ1	159,1002	80,0538	53,7049		R
Linker fragment ions	Acryloyl (Ayα)	856,4523	428,7298	286,1556		
	Acryloyl (Ayβ)					
	Phosphonate (Ppα)	1014,4656	507,7364	338,8267		
	Phosphonate (Ppβ)					

Observation: The sequences of this cross-linked peptide were distinguished as α and β due to the linkage in different positions.

Table A 5. Bovine serum albumin (P02769) fasta file from Uniprot KB used for datasearch.

>sp|P02769|ALBU_BOVIN Serum albumin OS=Bos taurus OX=9913 GN=ALB PE=1 SV=4

```
         10         20         30         40         50
MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFKDLGE EHFKGLVLIA
         60         70         80         90        100
FSQYLQQCPF DEHVKLVNEL TEFAKTCVAD ESHAGCEKSL HTLFGDELCK
        110        120        130        140        150
VASLRETYGD MADCCEKQEP ERNECFLSHK DDSPDLPKLK PDPNTLCDEF
        160        170        180        190        200
KADEKKFWGK YLYEIARRHP YFYAPELLYY ANKYNGVFQE CCQAEDKGAC
        210        220        230        240        250
LLPKIETMRE KVLASSARQR LRCASIQKFG ERALKAWSVA RLSQKFPKAE
        260        270        280        290        300
FVEVTKLVTD LTKVHKECCH GDLLECADDR ADLAKYICDN QDTISSKLKE
        310        320        330        340        350
CCDKPLLEKS HCIAEVEKDA IPENLPPLTA DFAEDKDVCK NYQEAKDAFL
        360        370        380        390        400
GSFLYEYSRR HPEYAVSVLL RLAKEYEATL EECCAKDDPH ACYSTVFDKL
        410        420        430        440        450
KHLVDEPQNL IKQNCDQFEK LGEYGFQNAL IVRYTRKVPQ VSTPTLVEVS
        460        470        480        490        500
RSLGKVGTRC CTKPESERMP CTEDYLSLIL NRLCVLHEKT PVSEKVTKCC
        510        520        530        540        550
TESLVNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP DTEKQIKKQT
        560        570        580        590        600
ALVELLKHKP KATEEQLKTV MENFVAFVDK CCAADDKEAC FAVEGPKLVV

STQTALA
```

Table A 6. BSA cross-linked peptides from literature[87] with masses adjusted for PXL cross-linker.

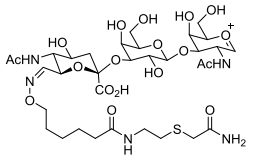
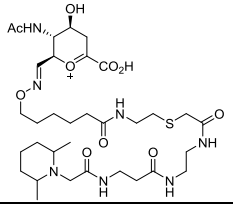
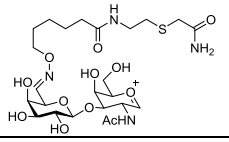
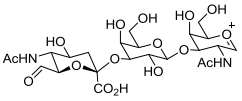
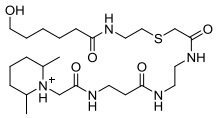
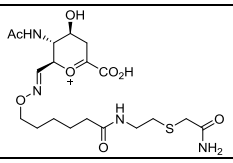
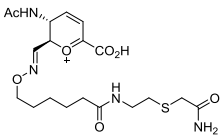
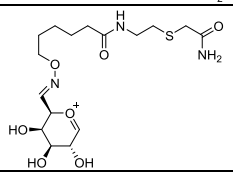
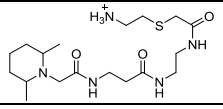
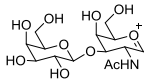
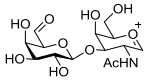
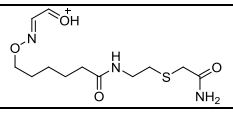
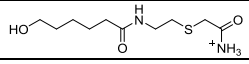
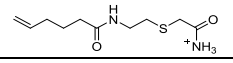
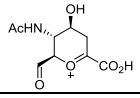
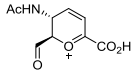
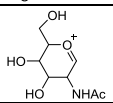
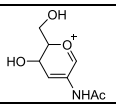
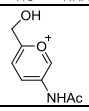
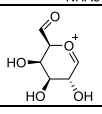
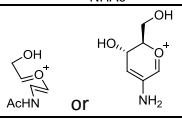
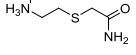
Peptide code	Peptide sequence	Literature mass	Theoretical corrected mass	*m/z* 3+	*m/z* 4+
1XL	$_{130}$K.DDSPDLPKL***K***PDPNTLCDEFK.A$_{152}$ $_{451}$R.SLG***K***VGTR.C$_{460}$	3398	3528	1177,0	883,0
2XL	$_{232}$R.AL***K***AWSVAR.L$_{242}$ $_{263}$K.VH***K***ECCHGDLLECADDR.A$_{281}$	3252	3382	1128,3	846,5
3XL	$_{436}$R.***K***VPQVSTPTLVEVSR.S$_{452}$ $_{557}$K.HKP***K***ATEEQL.K$_{569}$	3085	3215	1072,7	804,8
4XL	$_{232}$R.AL***K***AWSVAR.L$_{242}$ $_{371}$R.LA***K***EYEATLEECCAK.D$_{387}$	2953	3083	1028,7	771,8
5XL	$_{222}$R.CASIQ***K***FGER.A$_{233}$ $_{482}$R.LCVLHE***K***TPVSEK.V$_{496}$	2871	3001	1001,3	751,3
6XL	$_{197}$K.GACLLP***K***IETMR.E$_{210}$ $_{489}$K.TPVSE***K***VTK.C$_{499}$	2513	2643	882,0	661,8
7XL	$_{130}$K.DDSPDLPKL***K***PDPNTLCDEFK.A$_{152}$ $_{451}$R.SLG***K***VGTR.C$_{460}$	2293	2423	808,7	606,8
8XL	$_{232}$R.AL***K***AWSVAR.L$_{242}$ $_{256}$K.LVTDLT***K***VH.K$_{267}$	2291	2421	808,0	606,3
9XL	$_{232}$R.AL***K***AWSVAR.L$_{242}$ $_{263}$K.VH***K***ECCHGDLLECADDR.A$_{281}$	2150	2280	761,0	571,0
10XL	$_{436}$R.***K***VPQVSTPTLVEVSR.S$_{452}$ $_{557}$K.HKP***K***ATEEQL.K$_{569}$	2120	2250	751,0	563,5
11XL	$_{209}$R.E***K***VLASSAR.Q$_{219}$ $_{241}$R.LSQ***K***FPK.A$_{249}$	1974	2104	702,3	527,0
12XL	$_{209}$R.E***K***VLASSAR.Q$_{219}$ $_{241}$R.LSQ***K***FPK.A$_{249}$	1788	1918	640,3	480,5
13XL	$_{401}$K.HLVDEPQNLI***K***QNCDQFE***K***LGEYGFQNALIVR.Y$_{434}$	3953	4083	1362,0	1021,8
14XL	$_{75}$K.TCVADESHAGCE***K***SLHTLFGDELC***K***VASLR.E$_{106}$	3528	3658	1220,3	915,5
15XL	$_{256}$K.LVTDLT***K***VH***K***ECCHGDLLECADDR.A$_{281}$	3021	3151	1051,3	788,8
16XL	$_{130}$K.DDSPDLP***K***L***K***PDPNTLCDEFK.A$_{152}$	2581	2711	904,7	678,8
17XL	$_{489}$K.TPVSE***K***VT***K***CCTESLVNR.R$_{508}$	2245	2375	792,7	594,8
18XL	$_{386}$K.DDPHACYSTVFD***K***L***K***HLVDEPQNLIK.Q$_{413}$	3220	3350	1117,7	838,5
19XL	$_{155}$K.***K***FWG***K***YLYEIAR.R$_{168}$	1711	1841	614,7	461,3

Table A 7. BSA peptides from literature[87] with chemical cleavage from PXL linker.

Chemical modification	Peptide code	Peptide sequence	Peptide exact mass	*m/z* modified peptide 1+	2+
Acryloyl	Ay-1	$_{130}$K.DDSPDLPKL***K***PDPNTLCDEFK.A$_{152}$	2443,1	2498,1	1249,6
	Ay-2	$_{451}$R.SLG***K***VGTR.C$_{460}$	816,5	871,5	436,3
	Ay-3	$_{232}$R.AL***K***AWSVAR.L$_{242}$	1000,6	1055,6	528,3
	Ay-4	$_{263}$K.VH***K***ECCHGDLLECADDR.A$_{281}$	2112,9	2167,9	1084,5
	Ay-5	$_{436}$R.***K***VPQVSTPTLVEVSR.S$_{452}$	1638,9	1693,9	847,5
	Ay-6	$_{557}$K.***H***KPKATEEQL.K$_{569}$	1307,7	1362,7	681,9
	Ay-7	$_{371}$R.LA***K***EYEATLEECCAK.D$_{387}$	1813,8	1868,8	934,9
	Ay-8	$_{222}$R.CASIQ***K***FGER.A$_{233}$	1194,6	1249,6	625,3
	Ay-9	$_{482}$R.LCVLHE***K***TPVSEK.V$_{496}$	1538,8	1593,8	797,4
	Ay-10	$_{197}$K.GACLLP***K***IETMR.E$_{210}$	1387,7	1442,8	721,9
	Ay-11	$_{489}$K.TPVSE***K***VTK.C$_{499}$	987,6	1042,6	521,8
	Ay-12	$_{209}$R.E***K***VLASSAR.Q$_{219}$	1017,6	1072,6	536,8
	Ay-13	$_{459}$R.CCT***K***PESER.M$_{469}$	1165,5	1220,5	610,8
	Ay-14	$_{256}$K.LVTDLT***K***VH.K$_{267}$	1152,7	1207,7	604,4
	Ay-15	$_{241}$R.LSQ***K***FPK.A$_{249}$	861,5	916,5	458,8
	Ay-16	$_{547}$K.***K***QTALVELLK.H$_{558}$	1141,7	1196,7	598,9
	Ay-17	$_{557}$K.H***K***PK.A$_{562}$	518,3	573,3	287,2
Phosphonate	Pp-1	$_{130}$K.DDSPDLPKL***K***PDPNTLCDEFK.A$_{152}$	2443,1	2711,2	1356,1
	Pp-2	$_{451}$R.SLG***K***VGTR.C$_{460}$	816,5	1084,5	542,8
	Pp-3	$_{232}$R.AL***K***AWSVAR.L$_{242}$	1000,6	1268,6	634,8
	Pp-4	$_{263}$K.VH***K***ECCHGDLLECADDR.A$_{281}$	2112,9	2380,9	1191,0
	Pp-5	$_{436}$R.***K***VPQVSTPTLVEVSR.S$_{452}$	1638,9	1907,0	954,0
	Pp-6	$_{557}$K.***H***KPKATEEQL.K$_{569}$	1307,7	1575,8	788,4
	Pp-7	$_{371}$R.LA***K***EYEATLEECCAK.D$_{387}$	1813,8	2081,9	1041,4
	Pp-8	$_{222}$R.CASIQ***K***FGER.A$_{233}$	1194,6	1462,6	731,8
	Pp-9	$_{482}$R.LCVLHE***K***TPVSEK.V$_{496}$	1538,8	1806,9	903,9
	Pp-10	$_{197}$K.GACLLP***K***IETMR.E$_{210}$	1387,7	1655,8	828,4
	Pp-11	$_{489}$K.TPVSE***K***VTK.C$_{499}$	987,6	1255,6	628,3
	Pp-12	$_{209}$R.E***K***VLASSAR.Q$_{219}$	1017,6	1285,6	643,3
	Pp-13	$_{459}$R.CCT***K***PESER.M$_{469}$	1165,5	1433,5	717,3
	Pp-14	$_{256}$K.LVTDLT***K***VH.K$_{267}$	1152,7	1420,7	710,9
	Pp-15	$_{241}$R.LSQ***K***FPK.A$_{249}$	861,5	1129,6	565,3
	Pp-16	$_{547}$K.***K***QTALVELLK.H$_{558}$	1141,7	1409,8	705,4
	Pp-17	$_{557}$K.H***K***PK.A$_{562}$	518,3	786,4	393,7

New approach for glycopeptide enrichment

Table A 8. Description, mass and chemical structure of main oxonium ions and tag fragments.

Description	Mass	Chemical Structure	Description	Mass	Chemical Structure
tag-oxNeu5Ac-Hex-HexNAc	840,3179		TMT tag-oxNeu5Ac	742,3804	
tag-oxGal-HexNAc	609,2436		oxNeu5Ac-Hex-HexNAc	595,1981	
TMT tag fragment	516,3214		tag-oxNeu5Ac (Sialic474)	475,1857	
tag-oxNeu5Ac [-H_2O]	457,1751		tag-oxGal (Gal405)	406,1643	
TMT tag fragment	402,2533		Hex-HexNAc	366,1395	
ox-Gal-HexNAc	364,1238		tag-sugar fragment	304,1326	
tag fragment	249,1267		tag fragment [-H_2O]	231,1162	
oxNeu5Ac (Sialic229)	230,0659		oxNeu5Ac [-H_2O]	212,0553	
HexNAc	204,0866		HexNAc [-H_2O]	186,0761	
HexNAc [-2*H_2O]	168,0655		oxGal (Gal160)	161,0445	
HexNAc fragmentation	144,0655		HexNAc [-2*H_2O,-CH_3O]	138,0550	
tag fragment	135,0587		HexNAc fragmentation [-H_2O]	126,0550	
TMT reporter ion	126,1277		tag fragment	118,0321	

Observation: The stereochemistry of hexoses and hexosamines are here absent and represented in a generic way, as well as glycosidic bonds, unless specific annotations because mass spectrometry does not provide absolute configuration information.

Table A 9. List of *O*-glycan motives containing modificated sialic acid used as glycan database.

Glycan number	Glycan motif	Exact Mass
1	HexNAc(1)Hex(1)Sialic474(1)	839.31007
	HexNAc(1)Hex(1)Sialic229(1)	594.191369
2	HexNAc(1)Hex(1)Fuc(1)Sialic474(1)	985.36797
	HexNAc(1)Hex(1)Fuc(1)Sialic229(1)	740.249269
3	HexNAc(2)Hex(1)Sialic474(1)	1042.38947
	HexNAc(2)Hex(1)Sialic229(1)	797.270769
4	HexNAc(1)Hex(1)Sialic474(2)	1313.48794
	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	1068.369239
	HexNAc(1)Hex(1)Sialic229(2)	823.250538
5	HexNAc(2)Hex(1)Fuc(1)Sialic474(1)	1188.44737
	HexNAc(2)Hex(1)Fuc(1)Sialic229(1)	943.328669
6	HexNAc(2)Hex(2)Sialic474(1)	1204.44227
	HexNAc(2)Hex(2)Sialic229(1)	959.323569
7	HexNAc(3)Hex(1)Sialic474(1)	1245.46887
	HexNAc(3)Hex(1)Sialic229(1)	1000.350169
8	HexNAc(2)Hex(1)Sialic474(2)	1516.56734
	HexNAc(2)Hex(1)Sialic474(1)Sialic229(1)	1271.448639
	HexNAc(2)Hex(1)Sialic229(2)	1026.329938
9	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)	1350.50017
	HexNAc(2)Hex(2)Fuc(1)Sialic229(1)	1105.381469
10	HexNAc(2)Hex(3)Sialic474(1)	1366.49507
	HexNAc(2)Hex(3)Sialic229(1)	1121.376369
11	HexNAc(3)Hex(1)Fuc(1)Sialic474(1)	1391.52677
	HexNAc(3)Hex(1)Fuc(1)Sialic229(1)	1146.408069
12	HexNAc(3)Hex(2)Sialic474(1)	1407.52167
	HexNAc(3)Hex(2)Sialic229(1)	1162.402969
13	HexNAc(1)Hex(1)Sialic474(3)	1787.66581
	HexNAc(1)Hex(1)Sialic474(2)Sialic229(1)	1542.547109
	HexNAc(1)Hex(1)Sialic474(1)Sialic229(2)	1297.428408
	HexNAc(1)Hex(1)Sialic229(3)	1052.309707
14	HexNAc(2)Hex(2)Sialic474(2)	1678.62014
	HexNAc(2)Hex(2)Sialic474(1)Sialic229(1)	1433.501439
	HexNAc(2)Hex(2)Sialic229(2)	1188.382738
15	HexNAc(2)Hex(2)Fuc(2)Sialic474(1)	1496.55807
	HexNAc(2)Hex(2)Fuc(2)Sialic229(1)	1251.439369
16	HexNAc(3)Hex(2)Fuc(1)Sialic474(1)	1553.57957
	HexNAc(3)Hex(2)Fuc(1)Sialic229(1)	1308.460869
17	HexNAc(3)Hex(3)Sialic474(1)	1569.57447
	HexNAc(3)Hex(3)Sialic229(1)	1324.455769
18	HexNAc(2)Hex(2)Fuc(1)Sialic474(2)	1824.67804
	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)Sialic229(1)	1579.559339
	HexNAc(2)Hex(2)Fuc(1)Sialic229(2)	1406.440638
19	HexNAc(3)Hex(2)Sialic474(2)	1881.69954
	HexNAc(3)Hex(2)Sialic474(1)Sialic229(1)	1636.580839
	HexNAc(3)Hex(2)Sialic229(2)	1391.462138
20	HexNAc(3)Hex(2)Fuc(2)Sialic474(1)	1699.63747
	HexNAc(3)Hex(2)Fuc(2)Sialic229(1)	1454.518769
21	HexNAc(3)Hex(3)Fuc(1)Sialic474(1)	1715.63237
	HexNAc(3)Hex(3)Fuc(1)Sialic229(1)	1470.513669
22	HexNAc(3)Hex(3)Sialic474(2)	2043.75234
	HexNAc(3)Hex(3)Sialic474(1)Sialic229(1)	1798.633639
	HexNAc(3)Hex(3)Sialic229(2)	1553.514938
23	HexNAc(3)Hex(3)Fuc(1)Sialic474(2)	2189.81024
	HexNAc(3)Hex(3)Fuc(1)Sialic474(1)Sialic229(1)	1944.691539
	HexNAc(3)Hex(3)Fuc(1)Sialic229(2)	1699.572858
24	HexNAc(5)Hex(3)Fuc(1)Sialic474(1)	2121.79117
	HexNAc(5)Hex(3)Fuc(1)Sialic229(1)	1876.672469
25	HexNAc(3)Hex(3)Fuc(2)Sialic474(2)	2335.86814
	HexNAc(3)Hex(3)Fuc(2)Sialic474(1)Sialic229(1)	2090.749439
	HexNAc(3)Hex(3)Fuc(2)Sialic229(2)	1845.630738
26	HexNAc(4)Hex(4)Fuc(2)Sialic474(1)	2226.82247
	HexNAc(4)Hex(4)Fuc(2)Sialic229(1)	1981.703769
27	HexNAc(4)Hex(4)Fuc(3)Sialic474(1)	2372.88037
	HexNAc(4)Hex(4)Fuc(3)Sialic229(1)	2127.761669

Table A 10. List of *N*-glycan motives containing modificated sialic acid used as glycan database.

Glycan number	Glycan motif	Exact Mass	Glycan number	Glycan motif	Exact Mass	Glycan number	Glycan motif	Exact Mass
1	HexNAc(3)Hex(4)Sialic474(1)	1731.627302	30	HexNAc(5)Hex(6)Sialic474(2)	2936.069583	53	HexNAc(7)Hex(7)Fuc(1)Sialic474(2)	3650.339062
	HexNAc(3)Hex(4)Sialic229(1)	1486.508601		HexNAc(5)Hex(6)Sialic474(1)Sialic229(1)	2690.950882		HexNAc(7)Hex(7)Fuc(1)Sialic474(1)Sialic229(1)	3405.220361
2	HexNAc(4)Hex(3)Sialic474(1)	1772.653851		HexNAc(5)Hex(6)Sialic229(2)	2445.832181		HexNAc(7)Hex(7)Fuc(1)Sialic229(2)	3160.10166
	HexNAc(4)Hex(3)Sialic229(1)	1465.498919	31	HexNAc(5)Hex(7)Fuc(1)Sialic474(1)	2770.002429	54	HexNAc(6)Hex(7)Fuc(1)Sialic474(3)	3921.437576
3	HexNAc(3)Hex(4)Fuc(1)Sialic474(1)	1877.685211		HexNAc(5)Hex(7)Fuc(1)Sialic229(1)	2524.883728		HexNAc(6)Hex(7)Fuc(1)Sialic474(2)Sialic229(1)	3676.318875
	HexNAc(3)Hex(4)Fuc(1)Sialic229(1)	1632.56651	32	HexNAc(6)Hex(7)Sialic474(1)	2827.023893		HexNAc(6)Hex(7)Fuc(1)Sialic474(1)Sialic229(2)	3431.200174
4	HexNAc(3)Hex(5)Sialic474(1)	1893.680126		HexNAc(6)Hex(7)Sialic229(1)	2581.905192		HexNAc(6)Hex(7)Fuc(1)Sialic229(3)	3186.081473
	HexNAc(3)Hex(5)Sialic229(1)	1648.561425	33	HexNAc(5)Hex(6)Fuc(1)Sialic474(2)	3082.127492	55	HexNAc(6)Hex(9)Fuc(1)Sialic474(2)	3771.365337
5	HexNAc(4)Hex(4)Sialic474(1)	1934.706675		HexNAc(5)Hex(6)Fuc(1)Sialic474(1)Sialic229(1)	2837.008791		HexNAc(6)Hex(9)Fuc(1)Sialic474(1)Sialic229(1)	3526.246636
	HexNAc(4)Hex(4)Sialic229(1)	1689.587974		HexNAc(5)Hex(6)Fuc(1)Sialic229(2)	2591.89009		HexNAc(6)Hex(9)Fuc(1)Sialic229(2)	3281.127935
6	HexNAc(3)Hex(6)Sialic474(1)	2055.73295	34	HexNAc(6)Hex(6)Sialic474(1)	2664.971069	56	HexNAc(7)Hex(8)Fuc(1)Sialic474(1)	3338.213999
	HexNAc(3)Hex(6)Sialic229(1)	1810.614249		HexNAc(6)Hex(6)Sialic229(1)	2419.852368		HexNAc(7)Hex(8)Fuc(1)Sialic229(1)	3093.095298
7	HexNAc(4)Hex(4)Fuc(1)Sialic474(1)	2080.764584	35	HexNAc(6)Hex(6)Sialic474(2)	3139.148956	57	HexNAc(7)Hex(8)Fuc(1)Sialic474(2)	3812.391886
	HexNAc(4)Hex(4)Fuc(1)Sialic229(1)	1835.645883		HexNAc(6)Hex(6)Sialic474(1)Sialic229(1)	2894.030255		HexNAc(7)Hex(8)Fuc(1)Sialic474(1)Sialic229(1)	3567.273185
8	HexNAc(4)Hex(5)Sialic474(1)	2096.759499		HexNAc(6)Hex(6)Sialic229(2)	2648.911554		HexNAc(7)Hex(8)Fuc(1)Sialic229(2)x	3322.154484
	HexNAc(4)Hex(5)Sialic229(1)	1851.640798	36	HexNAc(6)Hex(7)Fuc(1)Sialic474(1)	2973.081802	58	HexNAc(6)Hex(7)Sialic474(4)	4249.557554
9	HexNAc(5)Hex(3)Fuc(1)Sialic474(1)	2121.791133		HexNAc(6)Hex(7)Fuc(1)Sialic229(1)	2727.963101		HexNAc(6)Hex(7)Sialic474(3)Sialic229(1)	4004.438853
	HexNAc(5)Hex(3)Fuc(1)Sialic229(1)	1876.672432	37	HexNAc(6)Hex(8)Sialic474(1)	2989.076717		HexNAc(6)Hex(7)Sialic474(2)Sialic229(2)	3759.320152
10	HexNAc(5)Hex(4)Sialic474(1)	2137.786048		HexNAc(6)Hex(8)Sialic229(1)	2743.958016		HexNAc(6)Hex(7)Sialic474(1)Sialic229(3)	3514.201451
	HexNAc(5)Hex(4)Sialic229(1)	1892.667347	38	HexNAc(5)Hex(6)Sialic474(3)	3410.24747		HexNAc(6)Hex(7)Sialic229(4)	3269.08275
11	HexNAc(3)Hex(6)Fuc(1)Sialic474(1)	2201.790859		HexNAc(5)Hex(6)Sialic474(2)Sialic229(1)	3165.128769	59	HexNAc(7)Hex(7)Fuc(1)Sialic474(3)	4124.516949
	HexNAc(3)Hex(6)Fuc(1)Sialic229(1)	1956.672158		HexNAc(5)Hex(6)Sialic474(1)Sialic229(2)	2920.010068		HexNAc(7)Hex(7)Fuc(1)Sialic474(2)Sialic229(1)	3879.398248
12	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)	2242.817408		HexNAc(5)Hex(6)Sialic229(3)	2674.891367		HexNAc(7)Hex(7)Fuc(1)Sialic474(1)Sialic229(2)	3634.279547
	HexNAc(4)Hex(5)Fuc(1)Sialic229(1)	1997.698707	39	HexNAc(5)Hex(7)Fuc(1)Sialic474(2)	3244.180316		HexNAc(7)Hex(7)Fuc(1)Sialic229(3)	3389.160846
13	HexNAc(4)Hex(6)Sialic474(1)	2258.812323		HexNAc(5)Hex(7)Fuc(1)Sialic474(1)Sialic229(1)	2999.061615	60	HexNAc(6)Hex(7)Fuc(1)Sialic474(4)	4395.615463
	HexNAc(4)Hex(6)Sialic229(1)	2013.693622		HexNAc(5)Hex(7)Fuc(1)Sialic229(2)	2753.942914		HexNAc(6)Hex(7)Fuc(1)Sialic474(3)Sialic229(1)	4150.496762
14	HexNAc(5)Hex(4)Fuc(1)Sialic474(1)	2283.843957	40	HexNAc(6)Hex(6)Fuc(1)Sialic474(2)	3285.206865		HexNAc(6)Hex(7)Fuc(1)Sialic474(2)Sialic229(2)	3905.378061
	HexNAc(5)Hex(4)Fuc(1)Sialic229(1)	2038.725256		HexNAc(6)Hex(6)Fuc(1)Sialic474(1)Sialic229(1)	3040.088164		HexNAc(6)Hex(7)Fuc(1)Sialic474(1)Sialic229(3)	3660.25936
15	HexNAc(5)Hex(5)Sialic474(1)	2299.838872		HexNAc(6)Hex(6)Fuc(1)Sialic229(2)	2794.969463		HexNAc(6)Hex(7)Fuc(1)Sialic229(4)	3415.140659
	HexNAc(5)Hex(5)Sialic229(1)	2054.720171	41	HexNAc(6)Hex(6)Fuc(1)Sialic474(1)	2811.028978	61	HexNAc(7)Hex(6)Fuc(1)Sialic474(4)	4436.642012
16	HexNAc(6)Hex(3)Fuc(1)Sialic474(1)	2324.870506		HexNAc(6)Hex(6)Fuc(1)Sialic229(1)	2565.910277		HexNAc(7)Hex(6)Fuc(1)Sialic474(3)Sialic229(1)	4191.523311
	HexNAc(6)Hex(3)Fuc(1)Sialic229(1)	2079.751805	42	HexNAc(6)Hex(7)Sialic474(2)	3301.20178		HexNAc(7)Hex(6)Fuc(1)Sialic474(2)Sialic229(2)	3946.40461
17	HexNAc(6)Hex(4)Sialic474(1)	2340.865421		HexNAc(6)Hex(7)Sialic474(1)Sialic229(1)	3056.083079		HexNAc(7)Hex(6)Fuc(1)Sialic474(1)Sialic229(3)	3701.285909
	HexNAc(6)Hex(4)Sialic229(1)	2095.74672		HexNAc(6)Hex(7)Sialic229(2)	2810.964378		HexNAc(7)Hex(6)Fuc(1)Sialic229(4)	3456.167208
18	HexNAc(4)Hex(5)Sialic474(2)	2570.937386	43	HexNAc(6)Hex(8)Fuc(1)Sialic474(1)	3135.134626	62	HexNAc(7)Hex(8)Fuc(1)Sialic474(3)	4286.569773
	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	2325.818685		HexNAc(6)Hex(8)Fuc(1)Sialic229(1)	2890.015925		HexNAc(7)Hex(8)Fuc(1)Sialic474(2)Sialic229(1)	4041.451072
	HexNAc(4)Hex(5)Sialic229(2)	2080.699984	44	HexNAc(5)Hex(6)Fuc(1)Sialic474(3)	3556.305379		HexNAc(7)Hex(8)Fuc(1)Sialic474(1)Sialic229(2)	3796.332371
19	HexNAc(4)Hex(6)Fuc(1)Sialic474(1)	2404.870232		HexNAc(5)Hex(6)Fuc(1)Sialic474(2)Sialic229(1)	3311.186678		HexNAc(7)Hex(8)Fuc(1)Sialic229(3)	3551.21367
	HexNAc(4)Hex(6)Fuc(1)Sialic229(1)	2159.751531		HexNAc(5)Hex(6)Fuc(1)Sialic474(1)Sialic229(2)	3066.067977	63	HexNAc(7)Hex(7)Fuc(1)Sialic474(4)	4598.694836
20	HexNAc(4)Hex(7)Sialic474(1)	2420.865147		HexNAc(5)Hex(6)Fuc(1)Sialic229(3)	2820.949276		HexNAc(7)Hex(7)Fuc(1)Sialic474(3)Sialic229(1)	4353.576135
	HexNAc(4)Hex(7)Sialic229(1)	2175.746446	45	HexNAc(7)Hex(8)Sialic474(1)	3192.15609		HexNAc(7)Hex(7)Fuc(1)Sialic474(2)Sialic229(2)	4108.457434
21	HexNAc(5)Hex(4)Sialic474(2)	2611.963935		HexNAc(7)Hex(8)Sialic229(1)	2947.037389		HexNAc(7)Hex(7)Fuc(1)Sialic474(1)Sialic229(3)	3863.338733
	HexNAc(5)Hex(4)Sialic474(1)Sialic229(1)	2366.845234	46	HexNAc(6)Hex(5)Fuc(1)Sialic474(2)	3123.154041		HexNAc(7)Hex(7)Fuc(1)Sialic229(4)	3618.220032
	HexNAc(5)Hex(4)Sialic229(2)	2121.726533		HexNAc(6)Hex(5)Fuc(1)Sialic474(1)Sialic229(1)	2878.03534	64	HexNAc(7)Hex(8)Fuc(1)Sialic474(4)	4760.74766
22	HexNAc(5)Hex(5)Fuc(1)Sialic474(1)	2445.896781		HexNAc(6)Hex(5)Fuc(1)Sialic229(2)	2632.916639		HexNAc(7)Hex(8)Fuc(1)Sialic474(3)Sialic229(1)	4515.628959
	HexNAc(5)Hex(5)Fuc(1)Sialic229(1)	2200.77808	47	HexNAc(6)Hex(5)Fuc(1)Sialic474(1)	2648.976154		HexNAc(7)Hex(8)Fuc(1)Sialic474(2)Sialic229(2)	4270.510258
23	HexNAc(5)Hex(6)Sialic474(1)	2461.891696		HexNAc(6)Hex(5)Fuc(1)Sialic229(1)	2403.857453		HexNAc(7)Hex(8)Fuc(1)Sialic474(1)Sialic229(3)	4025.391557
	HexNAc(5)Hex(6)Sialic229(1)	2216.772995	48	HexNAc(6)Hex(5)Fuc(1)Sialic474(3)	3597.331928		HexNAc(7)Hex(8)Fuc(1)Sialic229(4)	3780.272856
24	HexNAc(4)Hex(5)Fuc(1)Sialic474(2)	2716.995295		HexNAc(6)Hex(5)Fuc(1)Sialic474(2)Sialic229(1)	3352.213227	65	HexNAc(8)Hex(9)Fuc(1)Sialic474(3)	4651.70197
	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)Sialic229(1)	2471.876594		HexNAc(6)Hex(5)Fuc(1)Sialic474(1)Sialic229(2)	3107.094526		HexNAc(8)Hex(9)Fuc(1)Sialic474(2)Sialic229(1)	4406.583269
	HexNAc(4)Hex(5)Fuc(1)Sialic229(2)	2226.757893		HexNAc(6)Hex(5)Fuc(1)Sialic229(3)	2861.975825		HexNAc(8)Hex(9)Fuc(1)Sialic474(1)Sialic229(2)	4161.464568
25	HexNAc(5)Hex(4)Fuc(1)Sialic474(2)	2758.021844	49	HexNAc(6)Hex(7)Fuc(1)Sialic474(2)	3447.259689		HexNAc(8)Hex(9)Fuc(1)Sialic229(3)	3916.345867
	HexNAc(5)Hex(4)Fuc(1)Sialic474(1)Sialic229(1)	2512.903143		HexNAc(6)Hex(7)Fuc(1)Sialic474(1)Sialic229(1)	3202.140988	66	HexNAc(11)Hex(11)Sialic474(1)	4490.632054
	HexNAc(5)Hex(4)Fuc(1)Sialic229(2)	2267.784442		HexNAc(6)Hex(7)Fuc(1)Sialic229(2)	2957.022287		HexNAc(11)Hex(11)Sialic229(1)	4245.513353
26	HexNAc(5)Hex(5)Sialic474(2)	2774.016759	50	HexNAc(6)Hex(6)Fuc(1)Sialic474(3)	3759.384752	67	HexNAc(8)Hex(9)Fuc(1)Sialic474(4)	5125.879857
	HexNAc(5)Hex(5)Sialic474(1)Sialic229(1)	2528.898058		HexNAc(6)Hex(6)Fuc(1)Sialic474(2)Sialic229(1)	3514.266051		HexNAc(8)Hex(9)Fuc(1)Sialic474(3)Sialic229(1)	4880.761156
	HexNAc(5)Hex(5)Sialic229(2)	2283.779357		HexNAc(6)Hex(6)Fuc(1)Sialic474(1)Sialic229(2)	3269.14735		HexNAc(8)Hex(9)Fuc(1)Sialic474(2)Sialic229(2)	4635.642455
27	HexNAc(5)Hex(6)Fuc(1)Sialic474(1)	2607.949605		HexNAc(6)Hex(6)Fuc(1)Sialic229(3)	3024.028649		HexNAc(8)Hex(9)Fuc(1)Sialic474(1)Sialic229(3)	4390.523754
	HexNAc(5)Hex(6)Fuc(1)Sialic229(1)	2362.830904	51	HexNAc(6)Hex(7)Sialic474(3)	3775.379667		HexNAc(8)Hex(9)Fuc(1)Sialic229(4)	4145.405053
28	HexNAc(6)Hex(3)Fuc(1)Sialic474(2)	2799.048393		HexNAc(6)Hex(7)Sialic474(2)Sialic229(1)	3530.260966	68	HexNAc(9)Hex(10)Fuc(1)Sialic474(4)	5491.012054
	HexNAc(6)Hex(3)Fuc(1)Sialic474(1)Sialic229(1)	2553.929692		HexNAc(6)Hex(7)Sialic474(1)Sialic229(2)	3285.142265		HexNAc(9)Hex(10)Fuc(1)Sialic474(3)Sialic229(1)	5245.893353
	HexNAc(6)Hex(3)Fuc(1)Sialic229(2)	2308.810991		HexNAc(6)Hex(7)Sialic229(3)	3040.023564		HexNAc(9)Hex(10)Fuc(1)Sialic474(2)Sialic229(2)	5000.774652
29	HexNAc(5)Hex(5)Fuc(1)Sialic474(2)	2920.074668	52	HexNAc(6)Hex(10)Fuc(1)Sialic474(1)	3459.240274		HexNAc(9)Hex(10)Fuc(1)Sialic474(1)Sialic229(3)	4755.655951
	HexNAc(5)Hex(5)Fuc(1)Sialic474(1)Sialic229(1)	2674.955967		HexNAc(6)Hex(10)Fuc(1)Sialic229(1)	3214.121573		HexNAc(9)Hex(10)Fuc(1)Sialic229(4)	4510.53725
	HexNAc(5)Hex(5)Fuc(1)Sialic229(2)	2429.837266						

Table A 11. List of *O*-glycan motives containing modificated galactose used as glycan database.

Glycan number	Glycan motif	Exact Mass	Glycan number	Glycan motif	Exact Mass
1	HexNAc(1)Hex(1)	365.1322	16	HexNAc(3)Fuc(1)Gal160(1)Gal405(1)	1320.4902
	HexNAc(1)Gal160(1)	363.11655		HexNAc(3)Fuc(1)Gal405(2)	1565.61
	HexNAc(1)Gal405(1)	608.23635	17	HexNAc(3)Hex(3)	1095.3966
2	HexNAc(2)Hex(1)	568.2116		HexNAc(3)Hex(2)Gal160(1)	1093.38095
	HexNAc(2)Gal160(1)	566.19595		HexNAc(3)Hex(2)Gal405(1)	1338.50075
	HexNAc(2)Gal405(1)	811.31575		HexNAc(3)Hex(1)Gal160(2)	1091.3653
3	HexNAc(1)Hex(2)Fuc(1)	673.2429		HexNAc(3)Hex(1)Gal160(1)Gal405(1)	1336.4851
	HexNAc(1)Hex(1)Fuc(1)Gal160(1)	671.22725		HexNAc(3)Hex(1)Gal405(2)	1581.6049
	HexNAc(1)Hex(1)Fuc(1)Gal405(1)	916.34705		HexNAc(3)Gal160(3)	1089.34965
	HexNAc(1)Fuc(1)Gal160(2)	669.2116		HexNAc(3)Gal160(2)Gal405(1)	1334.46945
	HexNAc(1)Fuc(1)Gal160(1)Gal405(1)	914.3314		HexNAc(3)Gal160(1)Gal405(2)	1579.58925
	HexNAc(1)Fuc(1)Gal405(2)	1159.4512		HexNAc(3)Gal405(3)	1824.70905
4	HexNAc(2)Hex(1)Fuc(1)	714.2695	18	HexNAc(4)Hex(1)Fuc(1)	1120.4282
	HexNAc(2)Fuc(1)Gal160(1)	712.25385		HexNAc(4)Fuc(1)Gal160(1)	1118.41255
	HexNAc(2)Fuc(1)Gal405(1)	957.37365		HexNAc(4)Fuc(1)Gal405(1)	1363.53235
5	HexNAc(2)Hex(2)	730.2644	19	HexNAc(4)Hex(2)	1136.4231
	HexNAc(2)Hex(1)Gal160(1)	728.24875		HexNAc(4)Hex(1)Gal160(1)	1134.40745
	HexNAc(2)Hex(1)Gal405(1)	973.36855		HexNAc(4)Hex(1)Gal405(1)	1379.52725
	HexNAc(2)Gal160(2)	726.2331		HexNAc(4)Gal160(2)	1132.3918
	HexNAc(2)Gal160(1)Gal405(1)	971.3529		HexNAc(4)Gal160(1)Gal405(1)	1377.5116
	HexNAc(2)Gal405(2)	1216.4727		HexNAc(4)Gal405(2)	1622.6314
6	HexNAc(3)Hex(1)	771.2909	20	HexNAc(2)Hex(2)Fuc(1)NeuAc(1)	1167.4177
	HexNAc(3)Gal160(1)	769.27525		HexNAc(2)Hex(1)Fuc(1)NeuAc(1)Gal160(1)	1165.40205
	HexNAc(3)Gal405(1)	1014.39505		HexNAc(2)Hex(1)Fuc(1)NeuAc(1)Gal405(1)	1410.52185
7	HexNAc(2)Hex(2)Fuc(1)	876.3223	21	HexNAc(2)Hex(3)NeuAc(1)	1183.4126
	HexNAc(2)Hex(1)Fuc(1)Gal160(1)	874.30665		HexNAc(2)Hex(2)NeuAc(1)Gal160(1)	1181.39695
	HexNAc(2)Hex(1)Fuc(1)Gal405(1)	1119.42645		HexNAc(2)Hex(2)NeuAc(1)Gal405(1)	1426.51675
	HexNAc(2)Fuc(1)Gal160(2)	872.291		HexNAc(2)Hex(1)NeuAc(1)Gal160(2)	1179.3813
	HexNAc(2)Fuc(1)Gal160(1)Gal405(1)	1117.4108		HexNAc(2)Hex(1)NeuAc(1)Gal160(1)Gal405(1)	1424.5011
	HexNAc(2)Fuc(1)Gal405(2)	1362.5306		HexNAc(2)Hex(1)NeuAc(1)Gal405(2)	1669.6209
8	HexNAc(2)Hex(3)	892.3172	22	HexNAc(3)Hex(2)NeuAc(1)	1224.4392
	HexNAc(2)Hex(2)Gal160(1)	890.30155		HexNAc(3)Hex(1)NeuAc(1)Gal160(1)	1222.42355
	HexNAc(2)Hex(2)Gal405(1)	1135.42135		HexNAc(3)Hex(1)NeuAc(1)Gal405(1)	1467.54335
	HexNAc(2)Hex(1)Gal160(2)	888.2859	23	HexNAc(3)Hex(2)Fuc(2)	1225.4596
	HexNAc(2)Hex(1)Gal160(1)Gal405(1)	1133.4057		HexNAc(3)Hex(1)Fuc(2)Gal160(1)	1223.44395
	HexNAc(2)Hex(1)Gal405(2)	1378.5255		HexNAc(3)Hex(1)Fuc(2)Gal405(1)	1468.56375
	HexNAc(2)Gal160(3)	886.27025	24	HexNAc(3)Hex(3)Fuc(1)	1241.4545
	HexNAc(2)Gal160(2)Gal405(1)	1131.39005		HexNAc(3)Hex(2)Fuc(1)Gal160(1)	1239.43885
	HexNAc(2)Gal160(1)Gal405(2)	1376.50985		HexNAc(3)Hex(2)Fuc(1)Gal405(1)	1484.55865
	HexNAc(2)Gal405(3)	1621.62965		HexNAc(3)Hex(1)Fuc(1)Gal160(2)	1237.4232
9	HexNAc(3)Hex(1)Fuc(1)	917.3489		HexNAc(3)Hex(3)Fuc(1)Gal160(1)Gal405(1)	1482.543
	HexNAc(3)Fuc(1)Gal160(1)	915.33325		HexNAc(3)Hex(3)Fuc(1)Gal160(2)	1727.6628
	HexNAc(3)Fuc(1)Gal405(1)	1160.45305		HexNAc(3)Fuc(1)Gal160(3)	1235.40755
10	HexNAc(3)Hex(2)	933.3438		HexNAc(3)Fuc(1)Gal160(2)Gal405(1)	1480.52735
	HexNAc(3)Hex(1)Gal160(1)	931.32815		HexNAc(3)Fuc(1)Gal160(1)Gal405(2)	1725.64715
	HexNAc(3)Hex(1)Gal405(1)	1176.44795		HexNAc(3)Fuc(1)Gal405(3)	1970.76695
	HexNAc(3)Gal160(2)	929.3125	25	HexNAc(3)Hex(3)NeuAc(1)	1386.492
	HexNAc(3)Gal160(1)Gal405(1)	1174.4323		HexNAc(3)Hex(2)NeuAc(1)Gal160(1)	1384.47635
	HexNAc(3)Gal405(2)	1419.5521		HexNAc(3)Hex(2)NeuAc(1)Gal405(1)	1629.59615
11	HexNAc(4)Hex(1)	974.3703		HexNAc(3)Hex(1)NeuAc(1)Gal160(2)	1378.4294
	HexNAc(4)Gal160(1)	972.35465		HexNAc(3)Hex(1)NeuAc(1)Gal160(1)Gal405(1)	1627.5805
	HexNAc(4)Gal405(1)	1217.47445		HexNAc(3)Hex(1)NeuAc(1)Gal405(2)	1872.7003
12	HexNAc(2)Hex(2)NeuAc(1)	1021.3598	26	HexNAc(3)Hex(3)Fuc(2)	1387.5124
	HexNAc(2)Hex(1)NeuAc(1)Gal160(1)	1019.34415		HexNAc(3)Hex(2)Fuc(2)Gal160(1)	1385.49675
	HexNAc(2)Hex(1)NeuAc(1)Gal405(1)	1264.46395		HexNAc(3)Hex(2)Fuc(2)Gal405(1)	1630.61655
13	HexNAc(2)Hex(2)Fuc(2)	1022.3802	27	HexNAc(4)Hex(2)Fuc(2)	1428.539
	HexNAc(2)Hex(1)Fuc(2)Gal160(1)	1020.36455		HexNAc(4)Hex(1)Fuc(2)Gal160(1)	1426.52335
	HexNAc(2)Hex(1)Fuc(2)Gal405(1)	1265.48435		HexNAc(4)Hex(1)Fuc(2)Gal405(1)	1671.64315
14	HexNAc(2)Hex(3)Fuc(1)	1038.3751	28	HexNAc(4)Hex(3)Fuc(1)	1444.5339
	HexNAc(2)Hex(2)Fuc(1)Gal160(1)	1036.35945		HexNAc(4)Hex(2)Fuc(1)Gal160(1)	1442.51825
	HexNAc(2)Hex(2)Fuc(1)Gal405(1)	1281.47925		HexNAc(4)Hex(2)Fuc(1)Gal405(1)	1687.63805
	HexNAc(2)Hex(1)Fuc(1)Gal160(2)	1034.3438		HexNAc(4)Hex(1)Fuc(1)Gal160(2)	1440.5026
	HexNAc(2)Hex(1)Fuc(1)Gal160(1)Gal405(1)	1279.4636		HexNAc(4)Hex(1)Fuc(1)Gal160(1)Gal405(1)	1685.6224
	HexNAc(2)Hex(1)Fuc(1)Gal405(2)	1524.5834		HexNAc(4)Hex(1)Fuc(1)Gal405(2)	1930.7422
	HexNAc(2)Fuc(1)Gal160(3)	1032.32815		HexNAc(4)Fuc(1)Gal160(3)	1438.48695
	HexNAc(2)Fuc(1)Gal160(2)Gal405(1)	1277.44795		HexNAc(4)Fuc(1)Gal160(2)Gal405(1)	1683.60675
	HexNAc(2)Fuc(1)Gal160(1)Gal405(2)	1522.56775		HexNAc(4)Fuc(1)Gal160(1)Gal405(2)	1928.72655
	HexNAc(2)Fuc(1)Gal405(3)	1767.68755		HexNAc(4)Fuc(1)Gal405(3)	2173.84635
15	HexNAc(3)Hex(1)NeuAc(1)	1062.3864	29	HexNAc(4)Hex(4)	1460.5288
	HexNAc(3)NeuAc(1)Gal160(1)	1060.37075		HexNAc(4)Hex(3)Gal160(1)	1458.51315
	HexNAc(3)NeuAc(1)Gal405(1)	1305.49055		HexNAc(4)Hex(3)Gal405(1)	1703.63295
16	HexNAc(3)Hex(2)Fuc(1)	1079.4017		HexNAc(4)Hex(2)Gal160(2)	1456.4975
	HexNAc(3)Hex(1)Fuc(1)Gal160(1)	1077.38605		HexNAc(4)Hex(2)Gal160(1)Gal405(1)	1701.6173
	HexNAc(3)Hex(1)Fuc(1)Gal405(1)	1322.50585		HexNAc(4)Hex(2)Gal405(2)	1946.7371
	HexNAc(3)Fuc(1)Gal160(2)	1075.3704			

Table A 12. List of N-glycan motives containing modificated galactose used as glycan database.

Glycan number	Glycan motif	Exact Mass	Glycan number	Glycan motif	Exact Mass	Glycan number	Glycan motif	Exact Mass
1	HexNAc(4)Hex(4)	1460.528788	15	HexNAc(5)Hex(4)NeuAc(1)Gal405(1)	2359.861	24	HexNAc(6)Hex(4)NeuAc(1)Gal160(2)Gal405(1)	2883.014
	HexNAc(4)Hex(3)Gal160(1)	1458.513138	16	HexNAc(5)Hex(6)Fuc(1)	2133.772		HexNAc(6)Hex(4)NeuAc(1)Gal160(1)Gal405(2)	3128.134
	HexNAc(4)Hex(3)Gal405(1)	1703.632938		HexNAc(5)Hex(3)Fuc(1)Gal160(3)	2127.725		HexNAc(6)Hex(4)NeuAc(1)Gal405(3)	3373.254
2	HexNAc(4)Hex(4)Fuc(1)	1606.586697		HexNAc(5)Hex(3)Fuc(1)Gal160(2)Gal405(1)	2372.845		HexNAc(6)Hex(5)NeuAc(1)Gal160(2)	2639.91
	HexNAc(4)Hex(3)Fuc(1)Gal160(1)	1604.571047		HexNAc(5)Hex(3)Fuc(1)Gal160(1)Gal405(2)	2617.964		HexNAc(6)Hex(5)NeuAc(1)Gal160(1)Gal405(1)	2885.03
	HexNAc(4)Hex(3)Fuc(1)Gal405(1)	1849.690847		HexNAc(5)Hex(3)Fuc(1)Gal405(3)	2863.084		HexNAc(6)Hex(5)NeuAc(1)Gal405(2)	3130.15
3	HexNAc(4)Hex(5)	1622.581612		HexNAc(5)Hex(4)Fuc(1)Gal160(2)	2129.74		HexNAc(6)Hex(6)NeuAc(1)Gal160(1)	2641.926
	HexNAc(4)Hex(3)Gal160(2)	1618.550312		HexNAc(5)Hex(4)Fuc(1)Gal160(1)Gal405(1)	2374.86		HexNAc(6)Hex(6)NeuAc(1)Gal405(1)	2887.046
	HexNAc(4)Hex(3)Gal160(1)Gal405(1)	1863.670112		HexNAc(5)Hex(4)Fuc(1)Gal405(2)	2619.98	25	HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	2715.963
	HexNAc(4)Hex(3)Gal405(2)	2108.789912		HexNAc(5)Hex(5)Fuc(1)Gal160(1)	2131.756		HexNAc(5)Hex(5)Fuc(1)NeuAc(2)Gal160(1)	2713.947
	HexNAc(4)Hex(4)Gal160(1)	1620.565962		HexNAc(5)Hex(5)Fuc(1)Gal405(1)	2376.876		HexNAc(5)Hex(5)Fuc(1)NeuAc(2)Gal405(1)	2959.067
	HexNAc(4)Hex(4)Gal405(1)	1865.685762	17	HexNAc(6)Hex(5)Fuc(1)	2174.798	26	HexNAc(6)Hex(6)NeuAc(1)	2481.889
4	HexNAc(4)Hex(5)Fuc(1)	1768.639521		HexNAc(6)Hex(3)Fuc(1)Gal160(2)	2170.767		HexNAc(6)Hex(4)NeuAc(1)Gal160(2)	2477.857
	HexNAc(4)Hex(3)Fuc(1)Gal160(2)	1764.608221		HexNAc(6)Hex(3)Fuc(1)Gal160(1)Gal405(1)	2415.887		HexNAc(6)Hex(4)NeuAc(1)Gal160(1)Gal405(1)	2722.977
	HexNAc(4)Hex(3)Fuc(1)Gal160(1)Gal405(1)	2009.728021		HexNAc(6)Hex(3)Fuc(1)Gal405(2)	2661.007		HexNAc(6)Hex(4)NeuAc(1)Gal405(2)	2968.097
	HexNAc(4)Hex(3)Fuc(1)Gal405(2)	2254.847821		HexNAc(6)Hex(4)Fuc(1)Gal160(1)	2172.783		HexNAc(6)Hex(5)NeuAc(1)Gal160(1)	2479.873
	HexNAc(4)Hex(4)Fuc(1)Gal160(1)	1766.623871		HexNAc(6)Hex(4)Fuc(1)Gal405(1)	2417.902		HexNAc(6)Hex(5)NeuAc(1)Gal405(1)	2724.993
	HexNAc(4)Hex(4)Fuc(1)Gal405(1)	2011.743671	18	HexNAc(6)Hex(6)	2190.793	27	HexNAc(6)Hex(6)NeuAc(2)	2772.984
5	HexNAc(5)Hex(4)	1663.608161		HexNAc(6)Hex(3)Gal160(3)	2184.746		HexNAc(6)Hex(5)NeuAc(2)Gal160(1)	2770.968
	HexNAc(5)Hex(3)Gal160(1)	1661.592511		HexNAc(6)Hex(3)Gal160(2)Gal405(1)	2429.866		HexNAc(6)Hex(5)NeuAc(2)Gal405(1)	3016.088
	HexNAc(5)Hex(3)Gal405(1)	1906.712311		HexNAc(6)Hex(3)Gal160(1)Gal405(2)	2674.986	28	HexNAc(6)Hex(7)Fuc(1)NeuAc(1)	2789.999
6	HexNAc(5)Hex(4)Fuc(1)	1809.66607		HexNAc(6)Hex(3)Gal405(3)	2920.106		HexNAc(6)Hex(4)Fuc(1)NeuAc(1)Gal160(3)	2783.952
	HexNAc(5)Hex(3)Fuc(1)Gal160(1)	1807.65042		HexNAc(6)Hex(4)Gal160(2)	2186.762		HexNAc(6)Hex(4)Fuc(1)NeuAc(1)Gal160(2)Gal405(1)	3029.072
	HexNAc(5)Hex(3)Fuc(1)Gal405(1)	2052.77022		HexNAc(6)Hex(4)Gal160(1)Gal405(1)	2431.882		HexNAc(6)Hex(4)Fuc(1)NeuAc(1)Gal160(1)Gal405(2)	3274.192
7	HexNAc(5)Hex(5)	1825.660985		HexNAc(6)Hex(4)Gal405(1)	2677.001		HexNAc(6)Hex(4)Fuc(1)NeuAc(1)Gal405(3)	3519.312
	HexNAc(5)Hex(3)Gal160(2)	1821.629685		HexNAc(6)Hex(5)Gal160(1)	2188.778		HexNAc(6)Hex(5)Fuc(1)NeuAc(1)Gal160(2)	2785.968
	HexNAc(5)Hex(3)Gal160(1)Gal405(1)	2066.749485		HexNAc(6)Hex(5)Gal405(1)	2433.897		HexNAc(6)Hex(5)Fuc(1)NeuAc(1)Gal160(1)Gal405(1)	3031.088
	HexNAc(5)Hex(3)Gal405(2)	2311.869285	19	HexNAc(5)Hex(5)Fuc(1)NeuAc(1)	2262.814		HexNAc(6)Hex(5)Fuc(1)NeuAc(1)Gal160(2)	3276.208
	HexNAc(5)Hex(4)Gal160(1)	1823.645335		HexNAc(5)Hex(4)Fuc(1)NeuAc(1)Gal160(1)	2260.799		HexNAc(6)Hex(6)Fuc(1)NeuAc(1)Gal160(1)	2787.984
	HexNAc(5)Hex(4)Gal405(1)	2068.765135		HexNAc(5)Hex(4)Fuc(1)NeuAc(1)Gal405(1)	2505.918		HexNAc(6)Hex(6)Fuc(1)NeuAc(1)Gal405(1)	3033.103
8	HexNAc(6)Hex(4)	1866.687534	20	HexNAc(5)Hex(6)NeuAc(1)	2278.809	29	HexNAc(6)Hex(6)Fuc(1)NeuAc(2)	2919.042
	HexNAc(6)Hex(3)Gal160(1)	1864.671884		HexNAc(5)Hex(4)NeuAc(1)Gal160(2)	2274.778		HexNAc(6)Hex(5)Fuc(1)NeuAc(2)Gal160(1)	2917.026
	HexNAc(6)Hex(3)Gal405(1)	2109.791684		HexNAc(5)Hex(4)NeuAc(1)Gal160(1)Gal405(1)	2519.898		HexNAc(6)Hex(5)Fuc(1)NeuAc(2)Gal405(1)	3162.146
9	HexNAc(4)Hex(5)NeuAc(1)	1913.677029		HexNAc(5)Hex(4)NeuAc(1)Gal405(2)	2765.018	30	HexNAc(6)Hex(6)Fuc(1)NeuAc(1)	2627.947
	HexNAc(4)Hex(4)NeuAc(1)Gal160(1)	1911.661379		HexNAc(5)Hex(5)NeuAc(1)Gal160(1)	2276.794		HexNAc(6)Hex(4)Fuc(1)NeuAc(1)Gal160(2)	2623.915
	HexNAc(4)Hex(4)NeuAc(1)Gal405(1)	2156.781179		HexNAc(5)Hex(5)NeuAc(1)Gal405(1)	2521.913		HexNAc(6)Hex(4)Fuc(1)NeuAc(1)Gal160(1)Gal405(1)	2869.035
10	HexNAc(5)Hex(5)Fuc(1)	1971.718894	21	HexNAc(6)Hex(6)Fuc(1)	2336.851		HexNAc(6)Hex(4)Fuc(1)NeuAc(1)Gal405(2)	3114.155
	HexNAc(5)Hex(3)Fuc(1)Gal160(2)	1967.687594		HexNAc(6)Hex(3)Fuc(1)Gal160(3)	2330.804		HexNAc(6)Hex(5)Fuc(1)NeuAc(1)Gal160(1)	2625.931
	HexNAc(5)Hex(3)Fuc(1)Gal160(1)Gal405(1)	2212.807394		HexNAc(6)Hex(3)Fuc(1)Gal160(2)Gal405(1)	2575.924		HexNAc(6)Hex(5)Fuc(1)NeuAc(1)Gal405(1)	2871.051
	HexNAc(5)Hex(3)Fuc(1)Gal405(2)	2457.927194		HexNAc(6)Hex(3)Fuc(1)Gal160(1)Gal405(2)	2821.044	31	HexNAc(6)Hex(7)NeuAc(2)	2935.037
	HexNAc(5)Hex(4)Fuc(1)Gal160(1)	1969.703244		HexNAc(6)Hex(3)Fuc(1)Gal405(3)	3066.164		HexNAc(6)Hex(5)NeuAc(2)Gal160(2)	2931.006
	HexNAc(5)Hex(4)Fuc(1)Gal405(1)	2214.823044		HexNAc(6)Hex(4)Fuc(1)Gal160(2)	2332.82		HexNAc(6)Hex(5)NeuAc(2)Gal160(1)Gal405(1)	3176.125
11	HexNAc(5)Hex(6)	1987.713809		HexNAc(6)Hex(4)Fuc(1)Gal160(1)Gal405(1)	2577.94		HexNAc(6)Hex(5)NeuAc(2)Gal405(2)	3421.245
	HexNAc(5)Hex(3)Gal160(3)	1981.666859		HexNAc(6)Hex(4)Fuc(1)Gal405(2)	2823.059		HexNAc(6)Hex(6)NeuAc(2)Gal160(1)	2933.021
	HexNAc(5)Hex(3)Gal160(2)Gal405(1)	2226.786659		HexNAc(6)Hex(5)Fuc(1)Gal160(1)	2334.835		HexNAc(6)Hex(6)NeuAc(2)Gal405(1)	3178.141
	HexNAc(5)Hex(3)Gal160(1)Gal405(2)	2471.906459		HexNAc(6)Hex(5)Fuc(1)Gal405(1)	2579.955	32	HexNAc(6)Hex(5)Fuc(1)NeuAc(2)	2756.989
	HexNAc(5)Hex(3)Gal405(3)	2717.026259	22	HexNAc(6)Hex(7)Fuc(1)	2498.904		HexNAc(6)Hex(4)Fuc(1)NeuAc(2)Gal160(1)	2754.973
	HexNAc(5)Hex(4)Gal160(2)	1983.682509		HexNAc(6)Hex(3)Fuc(1)Gal160(4)	2490.841		HexNAc(6)Hex(4)Fuc(1)NeuAc(2)Gal405(1)	3000.093
	HexNAc(5)Hex(4)Gal160(1)Gal405(1)	2228.802309		HexNAc(6)Hex(3)Fuc(1)Gal160(3)Gal405(1)	2735.961	33	HexNAc(6)Hex(5)Fuc(1)NeuAc(1)	2465.894
	HexNAc(5)Hex(4)Gal405(2)	2473.922109		HexNAc(6)Hex(3)Fuc(1)Gal160(2)Gal405(2)	2981.081		HexNAc(6)Hex(3)Fuc(1)NeuAc(1)Gal160(2)	2461.862
	HexNAc(5)Hex(5)Gal160(1)	1985.698159		HexNAc(6)Hex(3)Fuc(1)Gal160(1)Gal405(3)	3226.201		HexNAc(6)Hex(3)Fuc(1)NeuAc(1)Gal160(1)Gal405(1)	2706.982
	HexNAc(5)Hex(5)Gal405(1)	2230.817959		HexNAc(6)Hex(3)Fuc(1)Gal405(4)	3471.321		HexNAc(6)Hex(3)Fuc(1)NeuAc(1)Gal405(2)	2952.102
12	HexNAc(6)Hex(4)Fuc(1)	2012.745443		HexNAc(6)Hex(4)Fuc(1)Gal160(3)	2492.857		HexNAc(6)Hex(4)Fuc(1)NeuAc(1)Gal160(1)	2463.878
	HexNAc(6)Hex(3)Fuc(1)Gal160(1)	2010.729793		HexNAc(6)Hex(4)Fuc(1)Gal160(2)Gal405(1)	2737.977		HexNAc(6)Hex(4)Fuc(1)NeuAc(1)Gal405(1)	2708.998
	HexNAc(6)Hex(3)Fuc(1)Gal405(1)	2255.849593		HexNAc(6)Hex(4)Fuc(1)Gal160(1)Gal405(2)	2983.097	34	HexNAc(6)Hex(7)Fuc(1)NeuAc(2)	3081.095
13	HexNAc(6)Hex(5)	2028.740358		HexNAc(6)Hex(4)Fuc(1)Gal405(3)	3228.216		HexNAc(6)Hex(5)Fuc(1)NeuAc(2)Gal160(2)	3077.063
	HexNAc(6)Hex(3)Gal160(2)	2024.709058		HexNAc(6)Hex(5)Fuc(1)Gal160(2)	2494.873		HexNAc(6)Hex(5)Fuc(1)NeuAc(2)Gal160(1)Gal405(1)	3322.183
	HexNAc(6)Hex(3)Gal160(1)Gal405(1)	2269.828858		HexNAc(6)Hex(5)Fuc(1)Gal160(1)Gal405(1)	2739.992		HexNAc(6)Hex(5)Fuc(1)NeuAc(2)Gal405(2)	3567.303
	HexNAc(6)Hex(3)Gal405(2)	2514.948658		HexNAc(6)Hex(5)Fuc(1)Gal405(2)	2985.112		HexNAc(6)Hex(6)Fuc(1)NeuAc(2)Gal160(1)	3079.079
	HexNAc(6)Hex(4)Gal160(1)	2026.724708		HexNAc(6)Hex(6)Fuc(1)Gal160(1)	2496.888		HexNAc(6)Hex(6)Fuc(1)NeuAc(2)Gal405(1)	3324.199
	HexNAc(6)Hex(4)Gal405(1)	2271.844508		HexNAc(6)Hex(6)Fuc(1)Gal405(1)	2742.008	35	HexNAc(6)Hex(7)NeuAc(3)	3226.132
14	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	2059.734938	23	HexNAc(5)Hex(6)NeuAc(2)	2569.905		HexNAc(6)Hex(6)NeuAc(3)Gal160(1)	3224.117
	HexNAc(4)Hex(4)Fuc(1)NeuAc(1)Gal160(1)	2057.719288		HexNAc(5)Hex(5)NeuAc(2)Gal160(1)	2567.889		HexNAc(6)Hex(6)NeuAc(3)Gal405(1)	3469.236
	HexNAc(4)Hex(4)Fuc(1)NeuAc(1)Gal405(1)	2302.839088		HexNAc(5)Hex(5)NeuAc(2)Gal405(1)	2813.009	36	HexNAc(6)Hex(7)Fuc(1)NeuAc(3)	3372.19
15	HexNAc(5)Hex(5)NeuAc(1)	2116.756402	24	HexNAc(6)Hex(7)NeuAc(1)	2643.941		HexNAc(6)Hex(6)Fuc(1)NeuAc(3)Gal160(1)	3370.175
	HexNAc(5)Hex(4)NeuAc(1)Gal160(1)	2114.740752		HexNAc(6)Hex(4)NeuAc(1)Gal160(3)	2637.894		HexNAc(6)Hex(6)Fuc(1)NeuAc(3)Gal405(1)	3615.294

Table A 13. Detailes from identified glycopeptides from serum sample oxidized with periodate and enriched using hydroxylamine beads. Peptides in black were identified as *O*-glycopeptides and peptides in blue were identified as *N*-glycopeptides.

Protein number	Protein	Accession	Peptide number	Peptide	Glycans NHFAGNa	Glycosylation site	Observed m/z	z	Score	Delta Mod	|Log Prob|	Confidence glycosylation site	Site previously reported
1	Ig lambda-7 chain C region	A0M8Q6	1	R.VT[+594.191]HEGSTVEK.T	HexNAc(1)Hex(1)Sialic229(1)	T90	840,862	2	569,4	28,5	9,88	high	no
			1	R.VT[+839.310]HEGSTVEK.T	HexNAc(1)Hex(1)Sialic474(1)	T90	963,421	2	230,6	138,4	3,46	high	no
			2	R.VTHEGS[+839.310]TVEK.T	HexNAc(1)Hex(1)Sialic474(1)	S94	963,421	2	437,6	7,3	9,07	poor	no
2	Ceruloplasmin	P00450	3	K.EHEGAIYPDN[+1689.588]TTDFQR.A	HexNAc(4)Hex(4)Sialic229(1)	N138	1194,816	3	347,6	204,3	9,65	good	yes
3	Haptoglobin	P00738	4	K.MVSHHN[+2471.877]LTTGATLINEQWLLTTAK.N	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)Sialic229(1)	N184	1288,321	4	280,8	144,7	7,56	good	yes
			5	K.VVLHPN[+1689.588]YSQVDIGLIK.L	HexNAc(4)Hex(4)Sialic229(1)	N241	871,905	4	512,1	378,8	12,09	good	yes
			5	K.VVLHPN[+1934.707]YSQVDIGLIK.L	HexNAc(4)Hex(4)Sialic474(1)	N241	1243,911	3	401,8	316,6	11,92	medium	yes
			6	K.VVLHPN[+1851.641]YSQVDIGLIK.L	HexNAc(4)Hex(5)Sialic229(1)	N241	1216,223	3	585,2	387,2	14,59	medium	yes
			6	K.VVLHPN[+2096.760]YSQVDIGLIK.L	HexNAc(4)Hex(5)Sialic474(1)	N241	973,698	4	486,8	329,1	11,72	medium	yes
			7	K.VVLHPN[+2080.700]YSQVDIGLIK.L	HexNAc(4)Hex(5)Sialic229(2)	N241	969,682	4	617,9	514,9	13,97	good	yes
			7	K.VVLHPN[+2325.819]YSQVDIGLIK.L	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N241	1030,712	4	442,5	265,5	10,06	good	yes
			7	K.VVLHPN[+2570.938]YSQVDIGLIK.L	HexNAc(4)Hex(5)Sialic474(2)	N241	1092,242	4	206,1	171,1	6,44	medium	yes
			8	K.VVLHPN[+3066.068]YSQVDIGLIK.L	HexNAc(5)Hex(6)Fuc(1)Sialic474(1)Sialic229(2)	N241	1216,028	4	144,9	27,6	1,55	medium	yes
4	Plasminogen	P00747	9	I.PSC[+57.021]DSSPVSTEQLAPTAPPELT[+1407.522]PVVQDC[+57.021]YHGDGQSYR.G	HexNAc(3)Hex(2)Sialic474(1)	T371	1364,088	4	386,7	0,0	5,41	poor	no
			10	K.IPSC[+57.021]DSSPVSTEQLAPT[+594.191]APPELTPVVQDC[+57.021]YHGDGQSYR.G	HexNAc(1)Hex(1)Sialic229(1)	T365	1189,028	4	449,8	4,1	10,39	poor	yes
			11	R.C[+57.021]TT[+594.191]PPPSSGPTYQC[+57.021]LK.G	HexNAc(1)Hex(1)Sialic229(1)	T264	796,675	3	431,6	15,2	11,43	high	no
			12	R.C[+57.021]TTPPPS[+594.191]SGPTYQC[+57.021]LK.G	HexNAc(1)Hex(1)Sialic229(1)	S268	796,675	3	459,3	0,0	10,11	medium	yes
			13	S.SPVSTEQLAPTAPPELT[+1579.559]PVVQDC[+57.021]YHGDGQSYR.G	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)Sialic229(1)	T371	1270,303	4	677,7	4,9	13,46	medium	no
5	Coagulation factor XII	P00748	14	R.T[+594.191]PPQSQTPGALPAK.R	HexNAc(1)Hex(1)Sialic229(1)	T331	662,987	3	401,2	7,9	7,46	good	no
			15	R.TPPQSQT[+594.191]PGALPAK.R	HexNAc(1)Hex(1)Sialic229(1)	T337	662,987	3	207,0	24,9	2,24	high	yes
6	Alpha-1-antitrypsin	P01009	16	K.YLGN[+1689.588]ATAIFFLPDEGK.L	HexNAc(4)Hex(4)Sialic229(1)	N271	1149,166	3	448,8	334,4	11,47	good	yes
			16	K.YLGN[+1934.707]ATAIFFLPDEGK.L	HexNAc(4)Hex(4)Sialic474(1)	N271	1230,873	3	443,0	272,5	11,10	medium	yes
			17	K.YLGN[+1851.641]ATAIFFLPDEGK.L	HexNAc(4)Hex(5)Sialic229(1)	N271	1203,183	3	493,7	308,1	11,63	medium	yes
			17	K.YLGN[+2096.760]ATAIFFLPDEGK.L	HexNAc(4)Hex(5)Sialic474(1)	N271	1284,892	3	267,9	104,9	6,19	medium	yes
			18	K.YLGN[+2080.700]ATAIFFLPDEGK.L	HexNAc(4)Hex(5)Sialic229(2)	N271	1279,534	3	307,3	226,4	8,11	good	yes
			18	K.YLGN[+2325.819]ATAIFFLPDEGK.L	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N271	1360,910	3	189,0	112,6	4,52	medium	yes
7	Kininogen-1	P01042	19	K.DFVQPPT[+594.191]K.I	HexNAc(1)Hex(1)Sialic229(1)	T261	763,343	2	421,6	171,3	9,21	good	no
			20	K.FSVATQTC[+57.021]QITPAEGPVVT[+594.191]AQYDC[+57.021]LGC[+57.021]VHPISTQSPDLEPILR.H	HexNAc(1)Hex(1)Sialic229(1)	T137	1338,379	4	1062,2	78,2	23,30	medium	no
			20	K.FSVATQTC[+57.021]QITPAEGPVVT[+839.310]AQYDC[+57.021]LGC[+57.021]VHPISTQSPDLEPILR.H	HexNAc(1)Hex(1)Sialic474(1)	T137	1399,659	4	805,7	57,5	18,30	medium	no
			21	R.DIPT[+594.191]NSPELEETLTHTITK.L	HexNAc(1)Hex(1)Sialic229(1)	T273	911,763	3	238,2	47,8	7,51	high	no
			22	R.DIPTNSPELEET[+594.191]LTHTITK.L	HexNAc(1)Hex(1)Sialic229(1)	T281	911,763	3	453,1	2,3	12,00	poor	no
			23	R.IGEIKEETT[+1579.559]VSPPHTSMAPAQDEER.D	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)Sialic229(1)	T401	1083,472	4	324,1	14,2	4,16	medium	yes
			24	R.IGEIKEETTVS[+1579.559]PPHTSMAPAQDEER.D	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)Sialic229(1)	T403	1083,472	4	450,2	0,0	7,46	poor	no
8	Ig gamma-1 chain C region	P01857	25	A.PELLGGPS[+1026.330]VFLFPPKPK.D	HexNAc(2)Hex(1)Sialic229(2)	S122	950,471	3	315,2	170,8	6,40	medium	no
			26	P.PC[+57.021]PAPELLGGPS[+797.271]VFLFPPKPK.D	HexNAc(2)Hex(1)Sialic229(1)	S122	1015,844	3	172,2	80,3	2,28	good	no
			27	R.EEQYN[+2080.765]STYR.V	HexNAc(4)Hex(4)Fuc(1)Sialic474(1)	N180	1090,763	3	244,0	62,4	5,70	medium	yes
			28	R.EEQYN[+2242.817]STYR.V	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)	N180	1144,781	3	245,3	64,0	5,70	medium	yes
			29	R.TPEVTC[+57.021]VVVDVS[+1042.389]HEDPEVK.F	HexNAc(2)Hex(1)Sialic474(1)	S150	796,113	4	509,5	6,3	7,41	poor	no
9	Ig gamma-3 chain C region	P01860	30	K.SC[+57.021]DT[+839.310]PPPC[+57.021]PR.C	HexNAc(1)Hex(1)Sialic474(1)	T137	675,941	3	404,8	15,6	10,27	good	no
10	Ig alpha-1 chain C region	P01876	31	R.LAGKPTHVN[+2471.877]VSVVMAEVDGTC[+57.021]Y.-	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)Sialic229(1)	N340	1205,258	4	99,1	15,8	1,73	medium	yes
			32	T.C[+57.021]T[+797.271]AAYPESK.T	HexNAc(2)Hex(1)Sialic229(1)	T205	912,367	2	129,2	18,7	1,98	medium	no
11	Ig alpha-2 chain C	P01877	33	K.TPLTAN[+1689.588]ITK.S	HexNAc(4)Hex(4)Sialic229(1)	N205	883,387	3	271,6	35,8	4,08	medium	yes

	region		34	K.TPLTAN[+1851.641]ITK.S	HexNAc(4)Hex(5)Sialic229(1)	N205	937,404	3	435,2	211,9	10,18	medium	yes
			35	K.TPLTAN[+1997.699]ITK.S	HexNAc(4)Hex(5)Fuc(1)Sialic229(1)	N205	986,090	3	507,1	214,4	10,55	medium	yes
			36	K.TPLTAN[+2013.694]ITK.S	HexNAc(4)Hex(6)Sialic229(1)	N205	991,421	3	228,7	74,9	5,29	good	yes
			37	K.TPLTAN[+2038.725]ITK.S	HexNAc(5)Hex(4)Fuc(1)Sialic229(1)	N205	999,763	3	180,6	43,2	1,83	medium	yes
12	Ig delta chain C region	P01880	38	K.AQASS[+594.191]VPTAQPQAEGSLAK.A	HexNAc(1)Hex(1)Sialic229(1)	S110	812,382	3	190,2	7,4	5,57	poor	yes
			39	K.AQASSVPTAQPQAEGS[+594.191]LAK.A	HexNAc(1)Hex(1)Sialic229(1)	S121	812,382	3	273,4	2,4	6,36	good	no
13	Apolipoprotein E	P02649	40	K.VQAAVGTS[+594.191]AAPVPSDNH.-	HexNAc(1)Hex(1)Sialic229(1)	S208	739,001	3	487,5	24,2	12,16	high	yes
			41	K.VQAAVGTSAAPVPS[+594.191]DNH.-	HexNAc(1)Hex(1)Sialic229(1)	S314	739,001	3	496,2	17,2	10,48	medium	yes
			42	R.AAT[+594.191]VGSLAGQPLQER.A	HexNAc(1)Hex(1)Sialic229(1)	T212	698,002	3	349,4	48,6	10,01	medium	yes
			42	R.AAT[+839.310]VGSLAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(1)	T212	779,709	3	324,4	55,7	7,64	medium	yes
			43	R.AATVGS[+594.191]LAGQPLQER.A	HexNAc(1)Hex(1)Sialic229(1)	S215	698,002	3	505,7	55,3	11,57	medium	yes
			43	R.AATVGS[+839.310]LAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(1)	S215	779,709	3	388,1	27,3	8,64	good	yes
14	Apolipoprotein A-II	P02652	44	K.EPC[+57.021]VESLVS[+594.191]QYFQTVTDYGK.D	HexNAc(1)Hex(1)Sialic229(1)	S35	982,096	3	517,1	4,1	8,30	poor	no
			45	K.EPC[+57.021]VESLVSQYFQT[+594.191]VTDYGK.D	HexNAc(1)Hex(1)Sialic229(1)	T40	982,096	3	393,6	6,8	9,35	good	no
15	Apolipoprotein C-III	P02656	46	D.PEVRPT[+594.191]SAVAA.-	HexNAc(1)Hex(1)Sialic229(1)	T94	846,396	2	294,1	1,6	5,85	poor	yes
			47	K.DKFSEFWDLDPEVRPT[+594.191]SAVA.A	HexNAc(1)Hex(1)Sialic229(1)	T94	968,437	3	191,7	1,5	3,41	poor	yes
			48	K.DKFSEFWDLDPEVRPT[+594.191]SAVAA.-	HexNAc(1)Hex(1)Sialic229(1)	T94	992,117	3	359,1	4,0	7,00	poor	yes
			48	K.DKFSEFWDLDPEVRPT[+839.310]SAVAA.-	HexNAc(1)Hex(1)Sialic474(1)	T94	1073,823	3	310,0	4,0	5,87	poor	yes
			49	K.DKFSEFWDLDPEVRPTS[+594.191]AVA.A	HexNAc(1)Hex(1)Sialic229(1)	S95	968,437	3	462,0	4,1	11,57	poor	no
			50	K.DKFSEFWDLDPEVRPTS[+594.191]AVAA.-	HexNAc(1)Hex(1)Sialic229(1)	S95	992,117	3	266,8	5,3	7,23	poor	no
			51	K.FSEFWDLDPEVRPT[+594.191]SAVAA.-	HexNAc(1)Hex(1)Sialic229(1)	T94	911,077	3	259,2	14,9	6,81	high	yes
			51	K.FSEFWDLDPEVRPT[+839.310]SAVAA.-	HexNAc(1)Hex(1)Sialic474(1)	T94	992,783	3	388,7	13,3	9,65	medium	yes
			52	K.FSEFWDLDPEVRPTS[+594.191]AVAA.-	HexNAc(1)Hex(1)Sialic229(1)	S95	911,076	3	365,8	1,7	9,43	poor	no
			53	W.DLDPEVRPT[+594.191]SAVAA.-	HexNAc(1)Hex(1)Sialic229(1)	T94	1017,965	2	191,5	1,6	2,19	poor	yes
16	Complement component C9	P02748	54	Q.YTTSYDPELT[+1308.461]ESSGSASHIDC[+57.021]R.M	HexNAc(3)Hex(2)Fuc(1)Sialic229(1)	T32	1262,178	3	282,1	0,0	6,79	poor	no
			55	T.TSYDPELTESS[+740.249]GSASHIDC[+57.021]R.M	HexNAc(1)Hex(1)Fuc(1)Sialic229(1)	S35	984,746	3	338,6	0,0	5,74	poor	no
17	Beta-2-glycoprotein 1	P02749	56	R.VYKPSAGN[+1689.588]NSLYR.D	HexNAc(4)Hex(4)Sialic229(1)	N162	1053,451	3	42,0	18,0	1,98	high	yes
			57	R.VYKPSAGN[+2080.700]NSLYR.D	HexNAc(4)Hex(5)Sialic229(2)	N162	888,119	4	196,8	119,7	4,00	good	yes
			58	Y.EPGEEITYSC[+57.021]KPGYVS[+2335.868]R.G	HexNAc(3)Hex(3)Fuc(2)Sialic474(2)	S57	862,569	5	379,0	5,5	3,82	good	no
18	Fibronectin	P02751	59	R.HTS[+594.191]VQTTSSGSGPFTDVR.A	HexNAc(1)Hex(1)Sialic229(1)	S276	820,030	3	432,2	0,5	10,06	good	no
			60	R.HTSVQTT[+594.191]SSGSGPFTDVR.A	HexNAc(1)Hex(1)Sialic229(1)	T279	820,029	3	251,3	6,6	7,57	poor	yes
			61	R.HTSVQTTSSGSGPFT[+594.191]DVR.A	HexNAc(1)Hex(1)Sialic229(1)	T287	820,029	3	577,1	5,2	13,26	poor	no
			62	R.RTT[+839.310]PPTTATPIR.H	HexNAc(1)Hex(1)Sialic474(1)	T2061	717,687	3	235,9	21,9	5,62	high	no
			63	R.RTTPPT[+594.191]TATPIR.H	HexNAc(1)Hex(1)Sialic229(1)	T2064	635,981	3	158,2	0,6	2,28	good	yes
			64	R.RTTPPT[+839.310]TATPIR.H	HexNAc(1)Hex(1)Sialic474(1)	T2064	717,687	3	411,1	0,0	11,51	poor	yes
			65	R.RTTPPTTAT[+594.191]PIR.H	HexNAc(1)Hex(1)Sialic229(1)	T2067	635,981	3	468,7	0,0	11,86	poor	no
			66	R.T[+594.191]TPPTTATPIR.H	HexNAc(1)Hex(1)Sialic229(1)	T2060	583,947	3	185,5	1,0	1,97	good	no
			67	R.TT[+839.310]PPTTATPIR.H	HexNAc(1)Hex(1)Sialic474(1)	T2061	665,653	3	412,8	2,4	8,90	good	no
			68	R.TTPPT[+594.191]TATPIR.H	HexNAc(1)Hex(1)Sialic229(1)	T2064	583,948	3	153,2	11,2	2,18	high	yes
			69	R.TTPPTTAT[+839.310]PIR.H	HexNAc(1)Hex(1)Sialic474(1)	T2067	665,654	3	422,4	0,0	8,19	poor	no
19	Alpha-2-HS-glycoprotein	P02765	70	A.ALAAFNAQNN[+2524.884]GSNFQLEEISR.A	HexNAc(5)Hex(7)Fuc(1)Sialic229(1)	N176	1205,763	4	234,6	95,9	2,32	medium	yes
			71	K.AALAAFNAQNN[+2080.700]GSNFQLEEISR.A	HexNAc(4)Hex(5)Sialic229(2)	N176	1112,218	4	441,2	340,6	10,09	medium	yes
			72	R.T[+594.191]VVQPSVGAAAGPVVPPC[+57.021]PGR.I	HexNAc(1)Hex(1)Sialic229(1)	T341	870,759	3	231,9	10,9	4,56	high	no
			73	R.T[+839.310]VVQPSVGAAAGPVVPPC[+57.021]PGR.I	HexNAc(1)Hex(1)Sialic474(1)	T341	952,464	3	566,5	4,6	12,32	good	no
			74	R.TVVQPS[+594.191]VGAAAGPVVPPC[+57.021]PGR.I	HexNAc(1)Hex(1)Sialic229(1)	S346	870,758	3	500,3	56,2	12,17	medium	yes
			74	R.TVVQPS[+839.310]VGAAAGPVVPPC[+57.021]PGR.I	HexNAc(1)Hex(1)Sialic474(1)	S346	952,464	3	424,9	41,3	11,34	high	yes
20	Serum albumin	P02768	75	A.DLPS[+1569.574]LAADFVESK.D	HexNAc(3)Hex(3)Sialic474(1)	S328	987,774	3	389,7	1,9	6,19	poor	no
			76	A.DLPSLAADFVES[+1569.574]K.D	HexNAc(3)Hex(3)Sialic474(1)	S336	987,774	3	458,2	14,3	6,00	medium	no
			77	A.EVENDEMPADLPS[+959.324]LAADFVESK.D	HexNAc(2)Hex(2)Sialic229(1)	S328	842,116	4	217,2	143,3	1,86	medium	no
			78	C.DKS[+1579.559]LHTLFGDK.L	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)Sialic229(1)	S89	947,410	3	329,7	16,6	2,75	medium	no
			79	D.LPS[+1699.637]LAADFVESK.D	HexNAc(3)Hex(2)Fuc(2)Sialic474(1)	S328	992,770	3	166,1	71,5	1,15	good	no
			80	D.LPSLAADFVES[+1699.637]K.D	HexNAc(3)Hex(2)Fuc(2)Sialic474(1)	S336	992,774	3	395,6	7,0	3,32	medium	no
			81	D.LPSLAADFVES[+1715.632]K.D	HexNAc(3)Hex(3)Fuc(1)Sialic474(1)	S336	998,110	3	289,4	19,7	2,01	high	no
			82	E.MPADLPS[+1496.558]LAADFVESK.D	HexNAc(2)Hex(2)Fuc(2)Sialic474(1)	S328	1063,133	3	191,2	52,7	1,45	good	no

			83	M.PADLPS[+1496.558]LAADFVESK.D	HexNAc(2)Hex(2)Fuc(2)Sialic474(1)	S328	1019,466	3	225,3	50,8	2,23	medium	no
			84	M.PADLPS[+1553.580]LAADFVESK.D	HexNAc(3)Hex(2)Fuc(1)Sialic474(1)	S328	1038,471	3	260,6	9,5	2,19	poor	no
			85	M.PADLPS[+1678.620]LAADFVESK.D	HexNAc(2)Hex(2)Sialic474(2)	S328	1080,150	3	265,6	104,9	3,39	medium	no
			86	M.PADLPSLAADFVES[+1350.500]K.D	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)	S336	970,777	3	300,5	2,0	2,27	poor	no
			87	M.PADLPSLAADFVES[+1496.558]K.D	HexNAc(2)Hex(2)Fuc(2)Sialic474(1)	S336	1019,466	3	466,5	3,4	7,37	good	no
			88	M.PADLPSLAADFVES[+1553.580]K.D	HexNAc(3)Hex(2)Fuc(1)Sialic474(1)	S336	1038,473	3	395,2	4,2	6,30	poor	no
			89	M.PADLPSLAADFVES[+1678.620]K.D	HexNAc(2)Hex(2)Sialic474(2)	S336	1080,151	3	427,6	6,3	6,41	poor	no
			90	M.PADLPSLAADFVES[+1699.637]K.D	HexNAc(3)Hex(2)Fuc(2)Sialic474(1)	S336	1087,157	3	303,4	9,1	2,01	poor	no
			91	M.PADLPSLAADFVES[+1824.678]K.D	HexNAc(2)Hex(2)Fuc(1)Sialic474(2)	S336	1128,829	3	277,1	10,1	3,76	medium	no
			92	N.AETFTFHADIC[+57.021]TLS[+1496.558]EK.E	HexNAc(2)Hex(2)Fuc(2)Sialic474(1)	S541	1122,820	3	359,5	0,0	2,70	poor	no
			93	S.HC[+57.021]IAEVENDEMPADLPS[+1569.574]LAADFVESK.D	HexNAc(3)Hex(3)Sialic474(1)	S328	1115,228	4	176,4	38,1	1,94	medium	no
			94	S.HC[+57.021]IAEVENDEMPADLPSLAADFVES[+1454.519]K.D	HexNAc(3)Hex(2)Fuc(2)Sialic229(1)	S336	1086,465	4	318,7	2,7	2,59	poor	no
			95	V.ENDEMPADLPSLAADFVES[+1042.389]K.D	HexNAc(2)Hex(1)Sialic474(1)	S336	1074,140	3	382,5	3,2	2,71	medium	no
21	Serotransferrin	P02787	96	K.C[+57.021]GLVPVLAENYN[+1689.588]K.S	HexNAc(4)Hex(4)Sialic229(1)	N432	1056,118	3	537,2	370,3	11,75	medium	yes
			96	K.C[+57.021]GLVPVLAENYN[+1934.707]K.S	HexNAc(4)Hex(4)Sialic474(1)	N432	1137,824	3	468,2	285,4	10,78	good	yes
			97	K.C[+57.021]GLVPVLAENYN[+1851.641]K.S	HexNAc(4)Hex(5)Sialic229(1)	N432	1110,136	3	444,2	305,5	10,90	medium	yes
			98	K.C[+57.021]GLVPVLAENYN[+2080.700]K.S	HexNAc(4)Hex(5)Sialic229(2)	N432	1186,488	3	364,7	187,3	9,84	medium	yes
			98	K.C[+57.021]GLVPVLAENYN[+2325.819]K.S	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N432	951,397	4	301,5	50,0	5,34	medium	yes
			99	R.QQQHLFGSN[+1689.588]VTDC[+57.021]SGNFC[+57.021]LFR.S	HexNAc(4)Hex(4)Sialic229(1)	N630	1051,932	4	464,7	306,1	10,48	medium	yes
			99	R.QQQHLFGSN[+1934.707]VTDC[+57.021]SGNFC[+57.021]LFR.S	HexNAc(4)Hex(4)Sialic474(1)	N630	1113,212	4	342,0	255,9	8,59	high	yes
			100	R.QQQHLFGSN[+1851.641]VTDC[+57.021]SGNFC[+57.021]LFR.S	HexNAc(4)Hex(5)Sialic229(1)	N630	1092,446	4	466,0	320,2	11,44	high	yes
			100	R.QQQHLFGSN[+2096.760]VTDC[+57.021]SGNFC[+57.021]LFR.S	HexNAc(4)Hex(5)Sialic474(1)	N630	1153,728	4	369,2	35,4	9,48	good	yes
22	Hemopexin	P02790	101	A.T[+1188.383]PLPPTSAHGNVAEGETKPDPDVTER.C	HexNAc(2)Hex(2)Sialic229(2)	T24	976,681	4	522,5	12,9	9,93	medium	yes
			102	A.T[+594.191]PLPPTSAHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic229(1)	T24	828,134	4	658,1	49,3	16,42	good	yes
			102	A.T[+839.310]PLPPTSAHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)	T24	889,413	4	627,0	44,7	15,42	good	yes
			103	A.T[+959.324]PLPPTSAHGNVAEGETKPDPDVTER.C	HexNAc(2)Hex(2)Sialic229(1)	T24	919,416	4	451,1	26,7	9,64	medium	yes
			104	A.TPLPPT[+1188.383]SAHGNVAEGETKPDPDVTER.C	HexNAc(2)Hex(2)Sialic229(2)	T29	976,681	4	155,3	0,8	2,11	poor	yes
			105	A.TPLPPTSAHGNVAEGET[+1068.369]KPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T40	946,678	4	730,9	1,8	17,42	poor	no
			106	A.TPLPPTSAHGNVAEGET[+1188.383]KPDPDVTER.C	HexNAc(2)Hex(2)Sialic229(2)	T40	976,681	4	633,4	17,9	15,42	medium	no
			106	A.TPLPPTSAHGNVAEGET[+1433.501]KPDPDVTER.C	HexNAc(2)Hex(2)Sialic474(1)Sialic229(1)	T40	1037,960	4	624,4	8,6	15,42	poor	no
			107	A.TPLPPTSAHGNVAEGET[+1579.559]KPDPDVTER.C	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)Sialic229(1)	T40	1074,729	4	780,7	8,4	17,24	poor	no
			108	A.TPLPPTSAHGNVAEGET[+594.191]KPDPDVTER.C	HexNAc(1)Hex(1)Sialic229(1)	T40	828,133	4	559,4	19,1	12,60	medium	no
			108	A.TPLPPTSAHGNVAEGET[+839.310]KPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)	T40	889,413	4	731,3	15,4	17,42	medium	no
			109	A.TPLPPTSAHGNVAEGET[+959.324]KPDPDVTER.C	HexNAc(2)Hex(2)Sialic229(1)	T40	919,416	4	736,0	12,0	17,42	good	no
			110	A.TPLPPTSAHGNVAEGETKPDPDVT[+1068.369]ER.C	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T47	946,678	4	753,4	7,5	18,42	medium	no
			111	A.TPLPPTSAHGNVAEGETKPDPDVT[+594.191]ER.C	HexNAc(1)Hex(1)Sialic229(1)	T47	828,133	4	747,8	16,7	17,42	medium	no
			111	A.TPLPPTSAHGNVAEGETKPDPDVT[+839.310]ER.C	HexNAc(1)Hex(1)Sialic474(1)	T47	889,413	4	810,9	1,4	19,24	good	no
			112	G.PNSC[+57.021]SANGPGLYLIHGPNLYC[+57.021]YS[+1678.620]DVEK.L	HexNAc(2)Hex(2)Sialic474(2)	S437	1176,505	4	334,4	3,2	2,78	poor	no
			113	K.ALPQPQN[+1689.588]VTSLLGC[+57.021]TH.-	HexNAc(4)Hex(4)Sialic229(1)	N453	1142,493	3	383,3	308,8	10,38	medium	yes
			113	K.ALPQPQN[+1934.707]VTSLLGC[+57.021]TH.-	HexNAc(4)Hex(4)Sialic474(1)	N453	1224,203	3	340,0	226,2	8,17	medium	yes
			114	K.ALPQPQN[+1851.641]VTSLLGC[+57.021]TH.-	HexNAc(4)Hex(5)Sialic229(1)	N453	1196,511	3	397,8	230,9	10,28	medium	yes
			115	K.ALPQPQN[+2080.700]VTSLLGC[+57.021]TH.-	HexNAc(4)Hex(5)Sialic229(2)	N453	1272,865	3	366,9	256,9	10,32	high	yes
			115	K.ALPQPQN[+2325.819]VTSLLGC[+57.021]TH.-	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N453	1354,569	3	322,0	121,8	8,99	medium	yes
			116	L.PPT[+1516.567]SAHGNVAEGETKPDPDVTER.C	HexNAc(2)Hex(1)Sialic474(2)	T29	980,932	4	353,2	22,4	3,79	medium	yes
			117	L.PPTSAHGNVAEGETKPDPDVT[+1516.567]ER.C	HexNAc(2)Hex(1)Sialic474(2)	T47	980,930	4	882,6	11,6	19,54	medium	no
			118	P.LPPT[+594.191]SAHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic229(1)	T29	778,609	4	266,3	27,4	6,93	medium	yes
			118	P.LPPT[+839.310]SAHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)	T29	839,888	4	433,1	20,7	9,28	high	yes
			119	P.LPPTS[+839.310]AHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)	S30	839,888	4	625,1	0,6	14,50	poor	no
			120	P.LPPTSAHGNVAEGETKPDPDVT[+839.310]ER.C	HexNAc(1)Hex(1)Sialic474(1)	T47	839,888	4	787,7	0,8	18,66	good	no
			121	R.VLGAPVALGLWS[+1366.495]LC[+57.021]WSLAIATPLPPTSAHGNVAEGETKPDPDVTER.C	HexNAc(2)Hex(3)Sialic474(1)	S15	770,744	8	619,9	4,4	9,90	medium	no
23	Vitronectin	P04004	122	K.N[+1648.561]GSLFAFR.G	HexNAc(3)Hex(5)Sialic229(1)	N169	854,017	3	143,8	61,5	2,10	medium	yes
			122	K.N[+1893.680]GSLFAFR.G	HexNAc(3)Hex(5)Sialic474(1)	N169	935,724	3	74,8	27,6	1,72	medium	yes
			123	K.N[+1689.588]GSLFAFR.G	HexNAc(4)Hex(4)Sialic229(1)	N169	867,690	3	151,7	45,8	1,83	good	yes

			124	K.N[+1851.641]GSLFAFR.G	HexNAc(4)Hex(5)Sialic229(1)	N169	921,710	3	162,4	48,5	2,46	medium	yes
			124	K.N[+2096.760]GSLFAFR.G	HexNAc(4)Hex(5)Sialic474(1)	N169	1003,415	3	254,2	30,3	3,96	medium	yes
			125	K.N[+2095.747]GSLFAFR.G	HexNAc(6)Hex(4)Sialic229(1)	N169	1003,083	3	233,2	32,6	2,00	medium	yes
			126	R.N[+2471.877]ISDGFDGIPDNVDAALALPAHSYSGR.E	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)Sialic229(1)	N242	1311,813	4	389,4	65,3	6,29	good	yes
			127	Y.PRN[+2216.773]ISDGFDGIPDNVDAALALPAHSYSGR.E	HexNAc(5)Hex(6)Sialic229(1)	N242	1311,310	4	330,3	10,1	1,80	medium	yes
24	Apolipoprotein B-100	P04114	128	K.FVEGSHN[+1934.707]STVSLTTK.N	HexNAc(4)Hex(4)Sialic474(1)	N3411	1181,175	3	301,3	169,2	9,44	medium	yes
			129	R.NNALDFVT[+594.191]K.S	HexNAc(1)Hex(1)Sialic229(1)	T3220	808,365	2	538,7	136,0	11,12	good	yes
25	Histidine-rich glycoprotein	P04196	130	R.SS[+594.191]TTKPPFKPHGSR.D	HexNAc(1)Hex(1)Sialic229(1)	S272	707,671	3	320,7	0,1	6,17	poor	no
			131	R.SST[+594.191]TKPPFKPHGSR.D	HexNAc(1)Hex(1)Sialic229(1)	T273	531,005	4	396,1	10,0	8,74	medium	no
			131	R.SST[+839.310]TKPPFKPHGSR.D	HexNAc(1)Hex(1)Sialic474(1)	T273	592,285	4	207,4	2,5	2,55	poor	no
			132	R.SSTT[+594.191]KPPFKPHGSR.D	HexNAc(1)Hex(1)Sialic229(1)	T274	531,005	4	451,4	33,4	9,27	medium	no
			133	R.SSTTKPPFKPHGS[+594.191]R.D	HexNAc(1)Hex(1)Sialic229(1)	S283	531,004	4	325,2	13,4	7,45	medium	no
			134	R.VIDFN[+1689.588]C[+57.021]TTSSVSSALANTK.D	HexNAc(4)Hex(4)Sialic229(1)	N125	926,895	4	236,6	72,7	5,30	high	yes
26	Plasma protease C1 inhibitor	P05155	135	K.VAT[+839.310]TVISK.M	HexNAc(1)Hex(1)Sialic474(1)	T48	829,408	2	260,5	2,1	4,83	poor	yes
			136	K.VATT[+594.191]VISK.M	HexNAc(1)Hex(1)Sialic229(1)	T49	706,848	2	336,2	8,4	7,01	poor	no
			137	K.VATTVIS[+594.191]K.M	HexNAc(1)Hex(1)Sialic229(1)	S51	706,847	2	259,5	10,1	2,76	medium	no
			137	K.VATTVIS[+839.310]K.M	HexNAc(1)Hex(1)Sialic474(1)	S51	829,408	2	470,0	0,0	10,13	poor	no
27	Complement factor H	P08603	138	K.IPC[+57.021]SQPPQIEHGTIN[+1689.588]SSR.S	HexNAc(4)Hex(4)Sialic229(1)	N882	928,398	4	379,4	323,8	9,10	high	yes
			139	K.IPC[+57.021]SQPPQIEHGTIN[+1851.641]SSR.S	HexNAc(4)Hex(5)Sialic229(1)	N882	968,912	4	530,8	417,8	13,13	medium	yes
			140	K.IPC[+57.021]SQPPQIEHGTIN[+2080.700]SSR.S	HexNAc(4)Hex(5)Sialic229(2)	N882	1026,177	4	287,2	222,3	8,28	good	yes
28	Complement C4-A	P0C0L4	141	R.GLN[+1689.588]VTLSSTGR.N	HexNAc(4)Hex(4)Sialic229(1)	N1328	932,067	3	221,9	57,8	4,33	good	yes
			142	R.PVAFS[+1944.692]VVPTAAAAVSLK.V	HexNAc(3)Hex(3)Fuc(1)Sialic474(1)Sialic229(1)	S900	893,913	4	211,4	13,5	2,70	good	no
29	Coagulation factor V	P12259	143	K.APSHQQAT[+594.191]TAGSPLR.H	HexNAc(1)Hex(1)Sialic229(1)	T804	705,994	3	516,9	0,0	12,15	poor	yes
			143	K.APSHQQAT[+839.310]TAGSPLR.H	HexNAc(1)Hex(1)Sialic474(1)	T804	787,700	3	522,3	0,0	13,72	poor	yes
			144	K.APSHQQATT[+594.191]AGSPLR.H	HexNAc(1)Hex(1)Sialic229(1)	T805	705,994	3	100,2	7,2	2,24	good	yes
			144	K.APSHQQATT[+839.310]AGSPLR.H	HexNAc(1)Hex(1)Sialic474(1)	T805	787,700	3	158,0	19,3	2,78	medium	yes
			145	K.APSHQQATTAGS[+594.191]PLR.H	HexNAc(1)Hex(1)Sialic229(1)	S808	705,993	3	117,3	1,2	2,33	poor	no
30	Inter-alpha-trypsin inhibitor heavy chain H1	P19827	146	F.VLSALQPS[+594.191]PTHSSSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	S651	835,393	3	771,7	1,8	18,82	poor	no
			147	F.VLSALQPSPT[+594.191]HSSSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	T653	835,393	3	241,6	30,8	6,88	high	yes
			147	F.VLSALQPSPT[+839.310]HSSSNTQR.L	HexNAc(1)Hex(1)Sialic474(1)	T653	917,099	3	201,6	13,8	2,85	medium	yes
			148	R.RTFVLSALQPS[+594.191]PTHSSSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	S651	727,850	4	496,0	13,6	13,74	medium	no
			149	R.RTFVLSALQPSPT[+839.310]HSSSNTQR.L	HexNAc(1)Hex(1)Sialic474(1)	T653	789,131	4	546,3	22,5	13,90	medium	yes
			150	R.RTFVLSALQPSPTHS[+594.191]SSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	S655	727,850	4	630,0	1,8	14,64	medium	no
			151	R.TFVLSALQPS[+594.191]PTHSSSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	S651	918,098	3	468,1	20,2	12,20	good	no
			151	R.TFVLSALQPS[+839.310]PTHSSSNTQR.L	HexNAc(1)Hex(1)Sialic474(1)	S651	999,804	3	738,4	0,0	18,17	poor	no
			152	R.TFVLSALQPSPT[+594.191]HSSSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	T653	918,098	3	560,5	20,4	14,25	medium	yes
			152	R.TFVLSALQPSPT[+839.310]HSSSNTQR.L	HexNAc(1)Hex(1)Sialic474(1)	T653	999,804	3	497,2	22,0	12,37	medium	yes
			153	R.TFVLSALQPSPTHSS[+839.310]SNTQR.L	HexNAc(1)Hex(1)Sialic474(1)	S656	999,804	3	877,4	2,0	21,00	poor	no
31	Hepatocyte growth factor activator	Q04756	154	R.AVPS[+594.191]SSSPQAQALTEDGRPC[+57.021]R.F	HexNAc(1)Hex(1)Sialic229(1)	S92	936,754	3	195,5	17,6	3,92	medium	no
32	Apolipoprotein F	Q13790	155	K.DANISQPET[+594.191]TKEGLR.A	HexNAc(1)Hex(1)Sialic229(1)	T273	751,680	3	366,2	35,1	9,21	high	no
			155	K.DANISQPET[+839.310]TKEGLR.A	HexNAc(1)Hex(1)Sialic474(1)	T273	833,387	3	197,5	15,6	2,30	high	no
			156	K.DANISQPETT[+594.191]KEGLR.A	HexNAc(1)Hex(1)Sialic229(1)	T274	751,680	3	284,1	26,8	6,86	high	yes
33	Inter-alpha-trypsin inhibitor heavy chain H4	Q14624	157	K.IPKPEAS[+594.191]FSPR.R	HexNAc(1)Hex(1)Sialic229(1)	S640	608,291	3	347,0	55,6	7,91	high	no
			157	K.IPKPEAS[+839.310]FSPR.R	HexNAc(1)Hex(1)Sialic474(1)	S640	689,998	3	322,9	59,2	8,35	high	no
			158	R.LAILPAS[+1188.383]APPATSNPDPAVSR.V	HexNAc(2)Hex(2)Sialic229(2)	S696	1078,499	3	424,3	3,6	9,56	poor	no
			159	R.LAILPASAPPAT[+1188.383]SNPDPAVSR.V	HexNAc(2)Hex(2)Sialic229(2)	T701	1078,499	3	236,1	0,3	6,34	poor	no
			160	T.QTPAPIQAPS[+1516.567]AILPLPGQSVER.L	HexNAc(2)Hex(1)Sialic474(2)	T733	947,464	4	261,4	3,8	3,52	poor	no
34	Proteoglycan 4	Q92954	161	K.SPDESTPELS[+839.310]AEPTPK.A	HexNAc(1)Hex(1)Sialic474(1)	S881	842,038	3	596,0	9,1	13,12	poor	no
			162	K.SPDESTPELSAEPT[+594.191]PK.A	HexNAc(1)Hex(1)Sialic229(1)	T885	760,332	3	431,6	57,9	11,51	good	no
			162	K.SPDESTPELSAEPT[+839.310]PK.A	HexNAc(1)Hex(1)Sialic474(1)	T885	842,038	3	282,7	6,6	7,85	medium	no

Table A 14. Detailes from identified glycopeptides from CSF sample oxidized with periodate and enriched using hydroxylamine beads. Peptides in black were identified as *O*-glycopeptides and peptides in blue were identified as *N*-glycopeptides.

Protein number	Protein Name	Protein Name	Peptide number	Peptide	Glycosylation	Glycosylation site	Observed *m/z*	z	Score	Delta Mod	|Log Prob|	Confidence glycosylation site	Site previously reported
1	Ig lambda-7 chain C region	A0M8Q6	1	R.VT[+839.310]HEGSTVEK.T	HexNAc(1)Hex(1)Sialic474(1)	T90	963.4215	2	224.7	119.29	4.54	high	no
			2	R.VTHEGS[+839.310]TVEK.T	HexNAc(1)Hex(1)Sialic474(1)	S94	963.4222	2	438.14	0	7.15	poor	no
			3	R.VTHEGST[+839.310]VEK.T	HexNAc(1)Hex(1)Sialic474(1)	T95	963.4225	2	397.38	0	7.38	poor	no
2	Neural cell adhesion molecule L1-like protein	O00533	4	R.N[+2999.062]ATK.L	HexNAc(5)Hex(7)Fuc(1)Sialic474(1)Sialic229(1)	N513	1144.7793	3	135.96	2.34	2.52	poor	no
3	Neurosecretory protein VGF	O15240	5	P.ERAPLPPPAPS[+594.191]QFQAR.M	HexNAc(1)Hex(1)Sialic229(1)	S229	786.0484	3	275.47	165.26	7.70	medium	no
			6	R.APLPPPAPS[+594.191]QFQAR.M	HexNAc(1)Hex(1)Sialic229(1)	S229	691.0006	3	571.04	373.64	11.95	medium	no
			6	R.APLPPPAPS[+839.310]QFQAR.M	HexNAc(1)Hex(1)Sialic474(1)	S229	772.7074	3	412.21	282.74	9.77	good	no
4	Prothrombin	P00734	7	R.GHVN[+2096.760]ITR.S	HexNAc(4)Hex(5)Sialic474(1)	N121	965.071	3	52.33	18.4	2.87	medium	yes
5	Haptoglobin	P00738	8	K.VVLHPN[+1851.641]YSQVDIGLIK.L	HexNAc(4)Hex(5)Sialic229(1)	N241	1216.2225	3	87.8	87.8	6.31	medium	yes
6	Plasminogen	P00747	9	R.C[+57.021]TT[+594.191]PPPSSGPTYQC[+57.021]LK.G	HexNAc(1)Hex(1)Sialic229(1)	T264	796.6748	3	63.29	1.14	2.34	medium	no
			10	R.C[+57.021]TTPPPSS[+594.191]GPTYQC[+57.021]LK.G	HexNAc(1)Hex(1)Sialic229(1)	S269	796.6748	3	314.98	0	8.11	poor	no
			11	S.SPVSTEQLAPTAPPELT[+1271.449]PVVQDC[+57.021]YHGDGQSYR.G	HexNAc(2)Hex(1)Sialic474(1)Sialic229(1)	T371	1193.2823	4	373.11	6.6	2.58	medium	no
			12	S.SPVSTEQLAPTAPPELT[+1334.441]PVVQDC[+57.021]YHGDGQSYR.G	HexNAc(2)Hex(2)Fuc(1)Sialic229(2)	T371	1209.0212	4	597.66	6.6	9.38	poor	no
7	Alpha-1-antitrypsin	P01009	13	K.YLGN[+1689.588]ATAIFFLPDEGK.L	HexNAc(4)Hex(4)Sialic229(1)	N271	1149.163	3	275.28	143.73	7.48	good	yes
8	Complement C3	P01024	14	T.PATNHMGNVTFT[+1824.678]IPANR.E	HexNAc(2)Hex(2)Fuc(1)Sialic474(2)	T89	1222.5464	3	114.23	0	2.55	poor	no
9	Kininogen-1	P01042	15	K.FSVATQTC[+57.021]QITPAEGPVVT[+594.191]AQYDC[+57.021]LGC[+57.021]VHPISTQSPDLEPILR.H	HexNAc(1)Hex(1)Sialic229(1)	T137	1338.3794	4	690.64	24.5	13.95	medium	no
			15	K.FSVATQTC[+57.021]QITPAEGPVVT[+839.310]AQYDC[+57.021]LGC[+57.021]VHPISTQSPDLEPILR.H	HexNAc(1)Hex(1)Sialic474(1)	T137	1399.907	4	407.28	3.6	6.89	poor	no
			16	R.DIPT[+594.191]NSPELEETLTHTITK.L	HexNAc(1)Hex(1)Sialic229(1)	T273	911.7619	3	179.15	29.61	6.35	high	no
			17	R.DIPTNSPELEETLTHTIT[+594.191]K.L	HexNAc(1)Hex(1)Sialic229(1)	T287	911.7619	3	459.66	0.02	11.95	poor	no
10	Ig gamma-3 chain C region	P01860	18	K.SC[+57.021]DT[+839.310]PPPC[+57.021]PR.C	HexNAc(1)Hex(1)Sialic474(1)	T137	675.9415	3	322.66	13.27	7.67	good	no
11	Apolipoprotein E	P02649	19	K.VQAAVGT[+594.191]SAAPVPSDNH.-	HexNAc(1)Hex(1)Sialic229(1)	T307	739.0008	3	469.64	15.68	11.77	medium	yes
			20	K.VQAAVGT[+823.251]SAAPVPSDNH.-	HexNAc(1)Hex(1)Sialic229(2)	T307	815.3531	3	126.78	17.93	4.26	high	yes
			21	K.VQAAVGTS[+1068.369]AAPVPSDNH.-	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S308	897.0604	3	468.02	31.17	11.32	medium	yes
			21	K.VQAAVGTS[+823.251]AAPVPSDNH.-	HexNAc(1)Hex(1)Sialic229(2)	S308	815.3535	3	533.55	4.44	11.68	poor	yes
			22	K.VQAAVGTS[+594.191]AAPVPSDNH.-	HexNAc(1)Hex(1)Sialic229(1)	S308	739.001	3	575.19	45.56	13.03	high	yes
			22	K.VQAAVGTS[+839.310]AAPVPSDNH.-	HexNAc(1)Hex(1)Sialic474(1)	S308	820.7087	3	380.48	17.24	9.96	good	yes
			23	K.VQAAVGTSAAPVPS[+1068.369]DNH.-	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S314	897.0604	3	512.64	2.68	11.29	poor	yes
			24	K.VQAAVGTSAAPVPS[+594.191]DNH.-	HexNAc(1)Hex(1)Sialic229(1)	S314	739.0006	3	484.38	1.91	10.12	poor	yes
			25	K.VQAAVGTSAAPVPS[+823.251]DNH.-	HexNAc(1)Hex(1)Sialic229(2)	S314	815.3531	3	336.51	2.69	7.27	poor	yes
			26	R.AAT[+1068.369]VGSLAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T212	856.0627	3	350.13	55.83	8.51	good	yes
			26	R.AAT[+823.251]VGSLAGQPLQER.A	HexNAc(1)Hex(1)Sialic229(2)	T212	774.3559	3	290.01	46.05	7.56	medium	yes
			27	R.AAT[+1204.442]VGSLAGQPLQER.A	HexNAc(2)Hex(2)Sialic474(1)	T212	901.4203	3	181.48	7.44	2.34	good	yes
			27	R.AAT[+959.324]VGSLAGQPLQER.A	HexNAc(2)Hex(2)Sialic229(1)	T212	819.7133	3	474.05	49.57	10.90	high	yes
			28	R.AAT[+594.191]VGSLAGQPLQER.A	HexNAc(1)Hex(1)Sialic229(1)	T212	698.0025	3	482.2	71.22	11.22	medium	yes
			28	R.AAT[+839.310]VGSLAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(1)	T212	779.7086	3	569.65	77.71	11.75	good	yes
			29	R.AATVGS[+1068.369]LAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S215	856.0627	3	700.56	49.26	16.74	medium	yes
			29	R.AATVGS[+823.251]LAGQPLQER.A	HexNAc(1)Hex(1)Sialic229(2)	S215	774.3559	3	310.95	50.58	8.00	medium	yes
			30	R.AATVGS[+1204.442]LAGQPLQER.A	HexNAc(2)Hex(2)Sialic474(1)	S215	901.4203	3	466.01	54.52	11.14	high	yes
			30	R.AATVGS[+959.324]LAGQPLQER.A	HexNAc(2)Hex(2)Sialic229(1)	S215	819.7133	3	601.49	49.84	14.74	medium	yes
			31	R.AATVGS[+594.191]LAGQPLQER.A	HexNAc(1)Hex(1)Sialic229(1)	S215	698.0006	3	325.74	56.46	6.17	medium	yes
			31	R.AATVGS[+839.310]LAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(1)	S215	779.7086	3	563.48	56.06	12.14	medium	yes

			32	R.AATVGS[+797.271]LAGQPLQER.A	HexNAc(2)Hex(1)Sialic229(1)	S215	765.6934	3	291.35	30.5	5.99	medium	yes
			33	R.VRAAT[+594.191]VGSLAGQPLQER.A	HexNAc(1)Hex(1)Sialic229(1)	T212	783.0595	3	185.21	8.88	5.84	good	yes
12	Apolipoprotein C-III	P02656	34	K.DKFSEFWDLDPEVRPT[+594.191]SAVAA.-	HexNAc(1)Hex(1)Sialic229(1)	T94	992.1164	3	337.22	3.96	7.40	poor	yes
			35	K.DKFSEFWDLDPEVRPTS[+839.310]AVAA.-	HexNAc(1)Hex(1)Sialic474(1)	S95	1073.8239	3	368.59	5.22	8.72	poor	no
13	Fibrinogen alpha chain	P02671	36	K.ESSS[+594.191]HHPGIAEFPSR.G	HexNAc(1)Hex(1)Sialic229(1)	S562	558.7452	4	230.14	2.34	5.99	poor	no
			37	K.ESSSHHPGIAEFPS[+594.191]R.G	HexNAc(1)Hex(1)Sialic229(1)	S572	558.7441	4	270.18	3.49	6.46	poor	no
			38	R.HPDEAAFFDT[+594.191]ASTGK.T	HexNAc(1)Hex(1)Sialic229(1)	T522	729.9745	3	329.9	3.12	7.07	poor	no
14	Complement component C9	P02748	39	T.TSYDPELTESS[+740.249]GSASHIDC[+57.021]R.M	HexNAc(1)Hex(1)Fuc(1)Sialic229(1)	S35	984.7442	3	385.68	0	9.58	poor	no
15	Beta-2-glycoprotein 1	P02749	40	R.VYKPSAGN[+2096.760]NSLYR.D	HexNAc(4)Hex(5)Sialic474(1)	N162	1189.1758	3	99.53	35.29	3.83	medium	yes
16	Fibronectin	P02751	41	R.HTSVQTT[+594.191]SSGSGPFTDVR.A	HexNAc(1)Hex(1)Sialic229(1)	T279	820.029	3	119.47	11.14	3.52	high	no
			42	R.HTSVQTTS[+594.191]SGSGPFTDVR.A	HexNAc(1)Hex(1)Sialic229(1)	S280	820.029	3	494.08	0	12.06	poor	no
			43	R.RTTPPT[+839.310]TATPIR.H	HexNAc(1)Hex(1)Sialic474(1)	T2064	717.6872	3	307.57	0	7.97	poor	yes
			44	R.TTPPT[+594.191]TATPIR.H	HexNAc(1)Hex(1)Sialic229(1)	T2064	583.9472	3	174.32	13.78	2.56	medium	yes
			45	R.TTPPTTAT[+594.191]PIR.H	HexNAc(1)Hex(1)Sialic229(1)	T2067	583.947	3	301.4	0	6.61	poor	yes
17	Alpha-2-HS-glycoprotein	P02765	46	K.AALAAFNAQNN[+1689.588]GSNFQLEEISR.A	HexNAc(4)Hex(4)Sialic229(1)	N176	1014.4405	4	327.35	221.94	9.85	medium	yes
			46	K.AALAAFNAQNN[+1934.707]GSNFQLEEISR.A	HexNAc(4)Hex(4)Sialic474(1)	N176	1075.4676	4	87.47	74.22	2.43	medium	yes
			47	K.AALAAFNAQNN[+1851.641]GSNFQLEEISR.A	HexNAc(4)Hex(5)Sialic229(1)	N176	1054.9537	4	443.36	363.81	12.06	medium	yes
			47	K.AALAAFNAQNN[+2096.760]GSNFQLEEISR.A	HexNAc(4)Hex(5)Sialic474(1)	N176	1116.2326	4	379.62	50.52	9.49	medium	yes
			48	K.AALAAFNAQNN[+2325.819]GSNFQLEEISR.A	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N176	1173.2465	4	367.77	261.49	8.74	good	yes
			49	R.T[+594.191]VVQPSVGAAAGPVVPPC[+57.021]PGR.I	HexNAc(1)Hex(1)Sialic229(1)	T341	870.7578	3	453.56	5.78	11.61	poor	no
			49	R.T[+839.310]VVQPSVGAAAGPVVPPC[+57.021]PGR.I	HexNAc(1)Hex(1)Sialic474(1)	T341	952.4646	3	456.25	10.6	11.22	medium	no
			50	R.TVVQPS[+594.191]VGAAAGPVVPPC[+57.021]PGR.I	HexNAc(1)Hex(1)Sialic229(1)	S346	870.758	3	426.12	54.99	10.39	high	yes
			50	R.TVVQPS[+839.310]VGAAAGPVVPPC[+57.021]PGR.I	HexNAc(1)Hex(1)Sialic474(1)	S346	952.4649	3	481.77	44.02	11.67	high	yes
18	Serum albumin	P02768	51	C.DKS[+1579.559]LHTLFGDK.L	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)Sialic229(1)	S89	947.4104	3	305.54	14.37	4.03	medium	no
			52	C.DKSLHT[+1579.559]LFGDK.L	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)Sialic229(1)	T92	947.4095	3	307.73	3.3	4.47	poor	no
			53	D.EMPADLPS[+1407.522]LAADFVESK.D	HexNAc(3)Hex(2)Sialic474(1)	S328	807.6092	4	186.58	89.79	1.07	medium	no
			54	D.LPSLAADFVES[+1699.637]K.D	HexNAc(3)Hex(2)Fuc(2)Sialic474(1)	S336	992.7784	3	273.36	1.18	2.06	poor	no
			55	E.MPADLPS[+1496.558]LAADFVESKDVC[+57.021]K.N	HexNAc(2)Hex(2)Fuc(2)Sialic474(1)	S328	1230.5516	3	212.71	4.98	2.45	poor	no
			56	E.VENDEMPADLPS[+943.329]LAADFVESKDVC[+57.021]K.N	HexNAc(2)Hex(1)Fuc(1)Sialic229(1)	S328	931.4131	4	204.98	43.55	1.26	good	no
			57	K.SLHT[+1313.488]LFGDK.L	HexNAc(1)Hex(1)Sialic474(2)	T92	777.6786	3	219.13	29.56	2.60	medium	no
			58	L.PSLAADFVES[+2372.880]K.D	HexNAc(4)Hex(4)Fuc(3)Sialic474(1)	S336	1179.503	3	380.04	35.66	6.41	medium	no
			59	M.PADLPS[+1678.620]LAADFVESK.D	HexNAc(2)Hex(2)Sialic474(2)	S328	1080.1488	3	245.39	1.6	4.16	poor	no
			60	S.HC[+57.021]IAEVENDEMPADLPS[+1881.700]LAADFVESK.D	HexNAc(3)Hex(2)Sialic474(2)	S328	1192.7507	4	286.64	13.44	2.29	medium	no
			61	V.ADES[+594.191]AENC[+57.021]DK.S	HexNAc(1)Hex(1)Sialic229(1)	S82	866.8168	2	325	126.82	6.65	high	no
19	Serotransferrin	P02787	62	I.PSDGPS[+1715.632]VAC[+57.021]VK.K	HexNAc(3)Hex(3)Fuc(1)Sialic474(1)	S56	708.8021	4	355.88	26.22	6.23	medium	no
			63	R.QQQHLFGSN[+1689.588]VTDC[+57.021]SGNFC[+57.021]LFR.S	HexNAc(4)Hex(4)Sialic229(1)	N630	1051.934	4	198.74	156.9	6.81	medium	yes
			64	R.QQQHLFGSN[+1851.641]VTDC[+57.021]SGNFC[+57.021]LFR.S	HexNAc(4)Hex(5)Sialic229(1)	N630	1092.4473	4	142.77	84.95	5.96	medium	yes
20	Hemopexin	P02790	65	A.KALPQPQN[+2957.022]VTSLLGC[+57.021]TH.-	HexNAc(6)Hex(7)Fuc(1)Sialic229(2)	N453	1206.2624	4	85.93	9.96	1.38	medium	yes
			66	A.T[+1188.383]PLPPTSAHGNVAEGETKPDPDVTER.C	HexNAc(2)Hex(2)Sialic229(2)	T24	976.6808	4	588.73	19.05	11.89	medium	yes
			67	A.T[+594.191]PLPPTSAHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic229(1)	T24	828.134	4	108.92	43.15	3.81	high	yes
			67	A.T[+839.310]PLPPTSAHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)	T24	889.4102	4	474.87	46.91	8.75	high	yes
			68	A.T[+823.251]PLPPTSAHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic229(2)	T24	885.3973	4	47.89	28.51	2.65	high	yes
			69	A.TPLPPT[+1188.383]SAHGNVAEGETKPDPDVTER.C	HexNAc(2)Hex(2)Sialic229(2)	T29	976.6814	4	571.45	2.33	10.02	poor	yes
			70	A.TPLPPT[+959.324]SAHGNVAEGETKPDPDVTER.C	HexNAc(2)Hex(2)Sialic229(1)	T29	919.4165	4	631.85	0.43	15.20	poor	yes
			71	A.TPLPPTS[+1553.515]AHGNVAEGETKPDPDVTER.C	HexNAc(3)Hex(3)Sialic229(2)	S30	1067.9655	4	109.99	1.68	5.51	poor	no
			72	A.TPLPPTS[+594.191]AHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic229(1)	S30	828.1336	4	542.84	0.72	11.37	poor	no
			73	A.TPLPPTSAHGNVAEGET[+1068.369]KPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T40	946.6788	4	657.18	5.67	16.01	poor	no
			74	A.TPLPPTSAHGNVAEGET[+1188.383]KPDPDVTER.C	HexNAc(2)Hex(2)Sialic229(2)	T40	976.683	4	562.3	12.45	12.33	medium	no
			74	A.TPLPPTSAHGNVAEGET[+1433.501]KPDPDVTER.C	HexNAc(2)Hex(2)Sialic474(1)Sialic229(1)	T40	1037.9604	4	792.42	8.64	18.20	medium	no
			74	A.TPLPPTSAHGNVAEGET[+1678.620]KPDPDVTER.C	HexNAc(2)Hex(2)Sialic474(2)	T40	1099.2398	4	496.39	8.4	11.50	poor	no
			75	A.TPLPPTSAHGNVAEGET[+1204.442]KPDPDVTER.C	HexNAc(2)Hex(2)Sialic474(1)	T40	980.6962	4	510.41	1.89	13.04	medium	no
			75	A.TPLPPTSAHGNVAEGET[+959.324]KPDPDVTER.C	HexNAc(2)Hex(2)Sialic229(1)	T40	919.4128	4	178.53	11.76	4.31	medium	no

			76	A.TPLPPTSAHGNVAEGET[+1313.488]KPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(2)	T40	1007.9587	4	287.55	8.4	7.53	poor	no
			76	A.TPLPPTSAHGNVAEGET[+823.251]KPDPDVTER.C	HexNAc(1)Hex(1)Sialic229(2)	T40	885.3973	4	322.64	15.29	8.30	medium	no
			77	A.TPLPPTSAHGNVAEGET[+1553.515]KPDPDVTER.C	HexNAc(3)Hex(3)Sialic229(2)	T40	1067.9638	4	133.9	8.4	5.66	poor	no
			78	A.TPLPPTSAHGNVAEGET[+1579.559]KPDPDVTER.C	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)Sialic229(1)	T40	1074.7293	4	355.43	8.4	7.70	poor	no
			79	A.TPLPPTSAHGNVAEGET[+1798.634]KPDPDVTER.C	HexNAc(3)Hex(3)Sialic474(1)Sialic229(1)	T40	1129.2455	4	200.27	8.4	5.92	poor	no
			80	A.TPLPPTSAHGNVAEGET[+594.191]KPDPDVTER.C	HexNAc(1)Hex(1)Sialic229(1)	T40	828.1341	4	535.46	16.8	12.90	medium	no
			80	A.TPLPPTSAHGNVAEGET[+839.310]KPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)	T40	1185.5486	3	603.2	20.16	14.57	high	no
			81	A.TPLPPTSAHGNVAEGET[+797.271]KPDPDVTER.C	HexNAc(2)Hex(1)Sialic229(1)	T40	878.9041	4	310.89	7.02	9.04	poor	no
			82	A.TPLPPTSAHGNVAEGETKPDPDVT[+594.191]ER.C	HexNAc(1)Hex(1)Sialic229(1)	T47	828.133	4	769.47	2.48	18.20	poor	no
			83	K.ALPQPQN[+1689.588]VTSLLGC[+57.021]TH.-	HexNAc(4)Hex(4)Sialic229(1)	N453	1142.4937	3	389.72	258.03	10.62	medium	yes
			84	K.ALPQPQN[+1851.641]VTSLLGC[+57.021]TH.-	HexNAc(4)Hex(5)Sialic229(1)	N453	1196.5114	3	396.76	267.42	11.55	good	yes
			84	K.ALPQPQN[+2096.760]VTSLLGC[+57.021]TH.-	HexNAc(4)Hex(5)Sialic474(1)	N453	1278.2186	3	293.5	196.41	8.84	medium	yes
			85	K.ALPQPQN[+2080.700]VTSLLGC[+57.021]TH.-	HexNAc(4)Hex(5)Sialic229(2)	N453	954.8995	4	290.52	185.65	8.61	medium	yes
			86	K.LNAAKALPQPQN[+2591.890]VTSLLGC[+57.021]TH.-	HexNAc(5)Hex(6)Fuc(1)Sialic229(2)	N453	1207.2658	4	248.53	37.5	3.92	medium	yes
			87	L.PQPQN[+2674.956]VTSLLGC[+57.021]TH.-	HexNAc(5)Hex(5)Fuc(1)Sialic474(1)Sialic229(1)	N453	1057.4307	4	324.35	102.18	4.56	medium	yes
			88	N.AAKALPQPQN[+2827.024]VTSLLGC[+57.021]TH.-	HexNAc(6)Hex(7)Sialic474(1)	N453	1209.0259	4	44.98	0.74	1.69	poor	yes
			89	P.LPPT[+594.191]SAHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic229(1)	T29	778.6083	4	410.82	29.17	9.99	medium	yes
			89	P.LPPT[+839.310]SAHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)	T29	839.8881	4	376.39	24.41	9.19	medium	yes
			90	P.LPPTS[+839.310]AHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)	S30	839.8877	4	751.72	0.6	18.49	poor	no
			91	R.SWPAVGN[+1851.641]C[+57.021]SSALR.W	HexNAc(4)Hex(5)Sialic229(1)	N187	1086.1083	3	250.68	98.92	7.42	medium	yes
21	Vitronectin	P04004	92	K.N[+2743.958]GSLFAFR.G	HexNAc(6)Hex(8)Sialic229(1)	N169	914.6205	4	77.21	21.54	1.10	medium	yes
22	Histidine-rich glycoprotein	P04196	93	R.VIDFN[+1851.641]C[+57.021]TTSSVSSALANTK.D	HexNAc(4)Hex(5)Sialic229(1)	N125	1289.5455	3	242.21	159.16	7.78	good	yes
23	Complement component C7	P10643	94	K.ENPLT[+839.310]QAVPK.C	HexNAc(1)Hex(1)Sialic474(1)	T696	968.4586	2	402.28	170.99	10.64	medium	yes
24	Chromogranin-A	P10645	95	E.EAEAGEEAVPEEEGPT[+1798.634]VVLNPHPSLGYK.E	HexNAc(3)Hex(3)Sialic474(1)Sialic229(1)	T251	1187.7673	4	179.62	12.48	2.47	high	yes
			96	E.EEEAEAGEEAVPEEEGPT[+2127.762]VVLNPHPSLGYK.E	HexNAc(4)Hex(4)Fuc(3)Sialic229(1)	T251	1334.3242	4	291.49	11.52	2.51	medium	yes
			97	E.EEEEEAEAGEEAVPEEEGPT[+1146.408]VVLNPHPSLGYK.E	HexNAc(3)Hex(1)Fuc(1)Sialic229(1)	T251	1153.2486	4	420.42	11.88	6.32	medium	yes
			98	E.EEEEEAEAGEEAVPEEEGPT[+1579.559]VVLNPHPSLGYK.E	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)Sialic229(1)	T251	1262.03	4	398.82	10.56	7.79	medium	yes
			99	E.EEEEEEEAEAGEEAVPEEEGPT[+1454.519]VVLNPHPSLGYK.E	HexNAc(3)Hex(2)Fuc(2)Sialic229(1)	T251	1295.3041	4	446.62	9.6	8.19	poor	yes
			100	E.EEEEEEEAEAGEEAVPEEEGPTVVLNPHPS[+1454.519]LGYK.E	HexNAc(3)Hex(2)Fuc(2)Sialic229(1)	S259	1294.8015	4	359.93	0.12	3.70	poor	no
			101	E.EEEEEEEEAEAGEEAVPEEEGPT[+740.249]VVLNPHPSLGYK.E	HexNAc(1)Hex(1)Fuc(1)Sialic229(1)	T251	1148.748	4	151.97	55.02	1.71	medium	yes
			102	K.REEEEEEEEEAEAGEEAVPEEEGPT[+1068.369]VVLNPHPSLGYK.E	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T251	1302.0638	4	122.85	12	2.76	medium	yes
			103	R.EEEEEEEEEAEAGEEAVPEEEGPT[+1068.369]VVLNPHPSLGYK.E	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T251	1263.0387	4	496.95	12	8.90	medium	yes
			103	R.EEEEEEEEEAEAGEEAVPEEEGPT[+823.251]VVLNPHPSLGYK.E	HexNAc(1)Hex(1)Sialic229(2)	T251	1201.7561	4	373.96	4.25	7.25	poor	yes
			104	R.EEEEEEEEEAEAGEEAVPEEEGPT[+594.191]VVLNPHPSLGYK.E	HexNAc(1)Hex(1)Sialic229(1)	T251	1144.496	4	331.41	16.8	6.29	medium	yes
			104	R.EEEEEEEEEAEAGEEAVPEEEGPT[+839.310]VVLNPHPSLGYK.E	HexNAc(1)Hex(1)Sialic474(1)	T251	1205.772	4	658.61	66.07	14.72	medium	yes
			105	R.EEEEEEEEEAEAGEEAVPEEEGPTVVLNPHPS[+1068.369]LGYK.E	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S259	1263.0387	4	242.6	12.04	4.45	good	no
			106	V.PEEEGPTVVLNPHPS[+1391.462]LGYK.E	HexNAc(3)Hex(2)Sialic229(2)	S259	1152.1672	3	400.22	3.86	8.35	poor	no
25	Inter-alpha-trypsin inhibitor heavy chain H1	P19827	107	R.RT[+594.191]FVLSALQPSPTHSSSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	T642	727.8502	4	687.21	5.37	16.73	medium	no
			108	R.RTFVLS[+594.191]ALQPSPTHSSSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	S646	727.8506	4	697.1	2.45	16.54	medium	no
			109	R.RTFVLSALQPS[+594.191]PTHSSSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	S651	727.8502	4	492.82	24.05	10.82	medium	no
			109	R.RTFVLSALQPS[+839.310]PTHSSSNTQR.L	HexNAc(1)Hex(1)Sialic474(1)	S651	789.1306	4	720.71	0	17.54	poor	no
			110	R.RTFVLSALQPSPT[+594.191]HSSSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	T653	727.8502	4	587.14	17.62	14.73	good	yes
			110	R.RTFVLSALQPSPT[+839.310]HSSSNTQR.L	HexNAc(1)Hex(1)Sialic474(1)	T653	789.1307	4	665.12	30.92	16.54	good	yes
			111	R.TFVLSALQPS[+594.191]PTHSSSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	S651	918.0977	3	449.77	0	10.54	poor	no
			111	R.TFVLSALQPS[+839.310]PTHSSSNTQR.L	HexNAc(1)Hex(1)Sialic474(1)	S651	999.8051	3	153.02	7.29	3.53	medium	no
			112	R.TFVLSALQPSPT[+594.191]HSSSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	T653	918.0974	3	307.89	8.82	9.34	poor	yes
			112	R.TFVLSALQPSPT[+839.310]HSSSNTQR.L	HexNAc(1)Hex(1)Sialic474(1)	T653	999.8049	3	144.08	6.55	3.56	poor	yes
			113	R.TFVLSALQPSPTHS[+594.191]SSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	S655	918.0974	3	879.68	2.04	20.62	medium	no
26	Prostaglandin-H2 D-isomerase	P41222	114	A.APEAQVS[+594.191]VQPNFQQDK.F	HexNAc(1)Hex(1)Sialic229(1)	S29	794.0281	3	201.28	34.01	2.87	high	yes
			115	K.SVVAPATDGGLN[+1892.667]LTSTFLR.K	HexNAc(5)Hex(4)Sialic229(1)	N78	1271.2357	3	223.97	30.96	6.03	medium	yes
			116	K.SVVAPATDGGLN[+2038.725]LTSTFLR.K	HexNAc(5)Hex(4)Fuc(1)Sialic229(1)	N78	990.1925	4	103.71	44.17	3.72	medium	yes
			117	K.SVVAPATDGGLN[+2054.720]LTSTFLR.K	HexNAc(5)Hex(5)Sialic229(1)	N78	994.1964	4	219.37	175.87	3.97	medium	yes

			118	K.SVVAPATDGGLN[+2200.778]LTSTFLR.K	HexNAc(5)Hex(5)Fuc(1)Sialic229(1)	N78	1030.7061	4	196.67	27.53	4.65	medium	yes
			119	K.SVVAPATDGGLN[+2201.791]LTSTFLR.K	HexNAc(3)Hex(6)Fuc(1)Sialic474(1)	N78	1030.7061	4	212.92	0.54	3.79	good	yes
27	Fractalkine	P78423	120	R.LGVLIT[+594.191]PVPDAQAATRR.Q	HexNAc(1)Hex(1)Sialic229(1)	T329	791.4107	3	331.63	24.05	9.50	medium	yes
			121	R.LGVLITPVPDAQAAT[+839.310]RR.Q	HexNAc(1)Hex(1)Sialic474(1)	T338	873.117	3	505.58	9.58	12.51	medium	no
28	Hepatocyte growth factor activator	Q04756	122	R.AVPSSSS[+594.191]PQAQALTEDGRPC[+57.021]R.F	HexNAc(1)Hex(1)Sialic229(1)	S95	936.7544	3	242.21	2	7.44	poor	no
29	Apolipoprotein F	Q13790	123	K.DANIS[+594.191]QPETTKEGLR.A	HexNAc(1)Hex(1)Sialic229(1)	S269	751.6804	3	127.96	0.76	2.37	good	no
			124	K.DANISQPET[+594.191]TKEGLR.A	HexNAc(1)Hex(1)Sialic229(1)	T273	751.6804	3	198.08	19.81	5.92	high	no
30	Inter-alpha-trypsin inhibitor heavy chain H4	Q14624	125	K.IPKPEAS[+594.191]FSPR.R	HexNAc(1)Hex(1)Sialic229(1)	S640	608.2911	3	354.55	62.04	8.68	high	no
			125	K.IPKPEAS[+839.310]FSPR.R	HexNAc(1)Hex(1)Sialic474(1)	S640	689.9977	3	279.77	71.32	6.31	high	no
			126	K.IPKPEASFS[+594.191]PR.R	HexNAc(1)Hex(1)Sialic229(1)	S642	608.2912	3	380.99	3.27	9.13	poor	no
31	Dickkopf-related protein 3	Q9UBP4	127	R.EIHKITNN[+2159.751]QTGQMVFSETVITSVGDEEGR.R	HexNAc(4)Hex(6)Fuc(1)Sialic229(1)	N121	897.2232	6	78.32	14.27	1.09	medium	yes
32	ProSAAS	Q9UHG2	128	K.RLET[+594.191]PAPQVPAR.R	HexNAc(1)Hex(1)Sialic229(1)	T247	643.6533	3	291.72	160.32	7.24	medium	yes
			128	K.RLET[+839.310]PAPQVPAR.R	HexNAc(1)Hex(1)Sialic474(1)	T247	725.3588	3	280.6	128.03	7.62	medium	yes
			129	R.LET[+839.310]PAPQVPAR.R	HexNAc(1)Hex(1)Sialic474(1)	T247	673.326	3	320.71	131.89	7.03	medium	yes

Table A 15. Detailes from identified glycopeptides of 5-replicate CSF samples enriched using hydroxylamine beads. Peptides in black were identified as *O*-glycopeptides and peptides in blue were identified as *N*-glycopeptides.

Protein number	Protein Name	Assession	Peptide number	Peptide	Glycosylation	Glycosylation site	Site previously reported	Observed *m/z*	z	Score*	Delta Mod*	|Log Prob|*	Confidence glycosylation site**	Replicate number
1	Ig lambda-7 chain C region	A0M8Q6	1	R.VT[+839.310]HEGSTVEK.T	HexNAc(1)Hex(1)Sialic474(1)	T90	no	963.422	2	436.3	32.8	6.4	high	1, 3, 4, 5
2	CCR4-NOT transcription complex subunit 1	A5YKK6	2	K.VQEVIFGLALLN[+2216.773]SSSSDLR.G	HexNAc(5)Hex(6)Sialic229(1)	N135	no	1422.639	3	200.3	16.7	1.66	medium	5
3	Phosphatidylinositol 3-kinase regulatory subunit beta	O00459	3	K.IN[+2325.819]RTQAEEMLSGK.R	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N628	no	1268.194	3	359.1	60.0	6.86	medium	4
4	Neural cell adhesion molecule L1-like protein	O00533	4	R.N[+2973.082]ATKLR.V	HexNAc(6)Hex(7)Fuc(1)Sialic474(1)	N513	no	1225.852	3	280.4	12.4	1.61	medium	2
5	Neurosecretory protein VGF	O15240	5	R.APLPPPAPS[+1068.369]QFQAR.M	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S229	no	849.061	3	557.3	275.2	11.79	medium	3
			6	R.APLPPPAPS[+594.191]QFQAR.M	HexNAc(1)Hex(1)Sialic229(1)	S229	no	691.001	3	465.5	230.5	11.31	medium	3
			6	R.APLPPPAPS[+839.310]QFQAR.M	HexNAc(1)Hex(1)Sialic474(1)	S229	no	772.707	3	438.9	218.8	9.7	medium	1, 2, 3, 4
6	Protein CutA	O60888	7	R.QVTES[+1068.369]VSDSITVLP.-	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S170	no	848.383	3	406.9	3.9	6.71	poor	1
7	Serine/threonine-protein kinase 10	O94804	8	K.S[+839.310]DPPTLLTPSK.W	HexNAc(1)Hex(1)Sialic474(1)	S252	no	997.972	2	402.4	4.6	7.82	poor	2
			9	K.SDPPT[+594.191]LLTPSK.W	HexNAc(1)Hex(1)Sialic229(1)	T256	no	875.412	2	275.6	10.6	5.95	high	4
			10	K.SDPPTLLT[+839.310]PSK.W	HexNAc(1)Hex(1)Sialic474(1)	T259	no	997.972	2	368.3	4.3	6.2	poor	2, 4, 5
			11	K.SDPPTLLTPS[+594.191]K.W	HexNAc(1)Hex(1)Sialic229(1)	S261	no	875.412	2	331.0	0.0	4.32	poor	5
			11	K.SDPPTLLTPS[+839.310]K.W	HexNAc(1)Hex(1)Sialic474(1)	S261	no	997.973	2	323.1	22.1	4.9	medium	1, 4
8	Ceruloplasmin	P00450	12	K.EN[+2570.938]LTAPGSDSAVFFEQGTTR.I	HexNAc(4)Hex(5)Sialic474(2)	N397	yes	1175.491	4	306.3	147.0	8.07	good	1
9	Prothrombin	P00734	13	R.GHVN[+2570.938]ITR.S	HexNAc(4)Hex(5)Sialic474(2)	N121	yes	1123.132	3	278.1	91.8	5.7	medium	1, 2
10	Haptoglobin	P00738	14	K.VVLHPN[+1934.707]YSQVDIGLIK.L	HexNAc(4)Hex(4)Sialic474(1)	N241	yes	1243.909	3	138.2	20.7	2.41	medium	1
			15	K.VVLHPN[+2080.700]YSQVDIGLIK.L	HexNAc(4)Hex(5)Sialic229(2)	N241	yes	969.685	4	245.0	60.5	3.87	medium	1, 3, 5
			15	K.VVLHPN[+2325.819]YSQVDIGLIK.L	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N241	yes	1374.623	3	239.6	77.9	4.02	medium	1, 2
			15	K.VVLHPN[+2570.938]YSQVDIGLIK.L	HexNAc(4)Hex(5)Sialic474(2)	N241	yes	1092.244	4	265.9	102.4	5.86	medium	1, 2, 3, 5
			16	K.VVLHPN[+2096.760]YSQVDIGLIK.L	HexNAc(4)Hex(5)Sialic474(1)	N241	yes	1297.928	3	273.5	132.0	6.21	medium	1, 2, 3, 4, 5
			17	K.VVLHPN[+2461.892]YSQVDIGLIK.L	HexNAc(5)Hex(6)Sialic474(1)	N241	yes	1064.983	4	322.8	102.3	6.12	medium	1, 2
11	Plasminogen	P00747	18	K.IPSC[+57.021]DSSPVST[+839.310]EQLAPTAPPELTPVVQDC[+57.021]YHGDGQSYR.G	HexNAc(1)Hex(1)Sialic474(1)	T359	no	1250.308	4	388.1	1.2	8.01	poor	1
12	Alpha-1-antitrypsin	P01009	19	G.N[+2095.747]ATAIFFLPDEGK.L	HexNAc(6)Hex(4)Sialic229(1)	N271	yes	880.381	4	286.0	47.3	3.21	medium	1
			20	G.N[+2920.075]ATAIFFLPDEGK.L	HexNAc(5)Hex(5)Fuc(1)Sialic474(2)	N271	yes	1448.620	3	253.4	56.8	1.82	medium	2
			21	K.YLGN[+1689.588]ATAIFFLPDEGK.L	HexNAc(4)Hex(4)Sialic229(1)	N271	yes	1149.167	3	359.4	165.2	8.38	medium	1, 2, 4
			21	K.YLGN[+1934.707]ATAIFFLPDEGK.L	HexNAc(4)Hex(4)Sialic474(1)	N271	yes	1230.873	3	409.9	254.4	9.14	medium	1, 2, 4, 5
			22	K.YLGN[+1851.641]ATAIFFLPDEGK.L	HexNAc(4)Hex(5)Sialic229(1)	N271	yes	1203.184	3	368.7	209.6	9.36	medium	1, 2, 3, 4, 5
			22	K.YLGN[+2096.760]ATAIFFLPDEGK.L	HexNAc(4)Hex(5)Sialic474(1)	N271	yes	1284.891	3	296.0	128.3	6.60	medium	1, 2, 3, 4, 5
			23	K.YLGN[+2080.700]ATAIFFLPDEGK.L	HexNAc(4)Hex(5)Sialic229(2)	N271	yes	959.905	4	397.6	227.0	9.64	medium	1, 2, 3, 4, 5
			23	K.YLGN[+2325.819]ATAIFFLPDEGK.L	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N271	yes	1361.242	3	414.7	191.0	8.94	medium	1, 2, 3, 4, 5
			23	K.YLGN[+2570.938]ATAIFFLPDEGK.L	HexNAc(4)Hex(5)Sialic474(2)	N271	yes	1442.954	3	339.6	172.4	6.93	medium	2, 3, 4, 5
			24	K.YLGN[+2095.747]ATAIFFLPDEGK.L	HexNAc(6)Hex(4)Sialic229(1)	N271	yes	1284.554	3	307.4	113.8	6.51	medium	1
			25	Y.LGN[+2524.884]ATAIFFLPDEGK.L	HexNAc(5)Hex(7)Fuc(1)Sialic229(1)	N271	yes	1372.919	3	311.0	113.3	7.35	medium	1
13	Kininogen-1	P01042	26	R.DIPT[+594.191]NSPELEETLTHTITK.L	HexNAc(1)Hex(1)Sialic229(1)	T273	no	911.763	3	373.1	9.7	8.94	medium	1, 4, 5
			27	R.DIPTNS[+594.191]PELEETLTHTITK.L	HexNAc(1)Hex(1)Sialic229(1)	S275	no	911.763	3	230.8	9.8	3.05	medium	2
			27	R.DIPTNS[+839.310]PELEETLTHTITK.L	HexNAc(1)Hex(1)Sialic474(1)	S275	no	993.470	3	267.3	14.5	3.12	medium	2
			28	R.DIPTNSPELEET[+839.310]LTHTITK.L	HexNAc(1)Hex(1)Sialic474(1)	T281	no	993.470	3	454.9	4.2	9.06	poor	1
			29	R.DIPTNSPELEETLT[+594.191]HTITK.L	HexNAc(1)Hex(1)Sialic229(1)	T283	no	911.763	3	530.3	3.2	10.51	poor	5
			30	R.DIPTNSPELEETLTHT[+594.191]ITK.L	HexNAc(1)Hex(1)Sialic229(1)	T285	no	911.762	3	520.8	0.7	10.90	poor	1, 3
14	Insulin-like growth factor II	P01344	31	R.DVST[+1068.369]PPTVLPDNFPR.Y	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T96	yes	908.410	3	303.8	6.0	7.03	poor	2, 4, 5
			31	R.DVST[+1313.488]PPTVLPDNFPR.Y	HexNAc(1)Hex(1)Sialic474(2)	T96	yes	990.116	3	412.7	7.0	7.80	medium	3
			31	R.DVST[+823.251]PPTVLPDNFPR.Y	HexNAc(1)Hex(1)Sialic229(2)	T96	yes	826.702	3	425.6	6.9	10.85	poor	1, 3, 4
			32	R.DVST[+839.310]PPTVLPDNFPR.Y	HexNAc(1)Hex(1)Sialic474(1)	T96	yes	832.056	3	392.1	1.0	6.18	poor	1, 5

			33	R.DVSTPPT[+1068.369]VLPDNFPR.Y	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T99	yes	908.410	3	453.2	6.1	9.63	poor	1, 2, 3, 4, 5
			33	R.DVSTPPT[+1313.488]VLPDNFPR.Y	HexNAc(1)Hex(1)Sialic474(2)	T99	yes	990.115	3	407.7	2.3	6.56	poor	2
			33	R.DVSTPPT[+823.251]VLPDNFPR.Y	HexNAc(1)Hex(1)Sialic229(2)	T99	yes	826.702	3	498.8	6.4	11.16	poor	4
			34	R.DVSTPPT[+839.310]VLPDNFPR.Y	HexNAc(1)Hex(1)Sialic474(1)	T99	yes	832.057	3	234.9	7.0	2.14	poor	3
15	Ig gamma-1 chain C region	P01857	35	K.TKPREEQYN[+1835.646]STYR.V	HexNAc(4)Hex(4)Fuc(1)Sialic229(1)	N180	yes	1169.821	3	103.1	14.3	1.31	medium	5
			35	K.TKPREEQYN[+2080.765]STYR.V	HexNAc(4)Hex(4)Fuc(1)Sialic474(1)	N180	yes	938.899	4	351.9	54.3	7.38	medium	1, 2, 4, 5
			36	K.TKPREEQYN[+1934.707]STYR.V	HexNAc(4)Hex(4)Sialic474(1)	N180	yes	902.384	4	156.0	40.6	2.41	medium	5
			37	K.TKPREEQYN[+1997.699]STYR.V	HexNAc(4)Hex(5)Fuc(1)Sialic229(1)	N180	yes	1223.840	3	180.7	14.9	2.43	poor	2, 4, 5
			37	K.TKPREEQYN[+2242.817]STYR.V	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)	N180	yes	979.412	4	329.3	64.9	7.04	medium	2, 3, 4, 5
			38	K.TKPREEQYN[+2096.760]STYR.V	HexNAc(4)Hex(5)Sialic474(1)	N180	yes	942.898	4	238.9	121.5	6.15	medium	1, 3, 4, 5
			39	K.TKPREEQYN[+2445.897]STYR.V	HexNAc(5)Hex(5)Fuc(1)Sialic474(1)	N180	yes	1030.182	4	308.5	37.1	5.82	medium	2
			40	R.EEQYN[+1934.707]STYR.V	HexNAc(4)Hex(4)Sialic474(1)	N180	yes	1042.083	3	223.0	6.3	2.49	poor	1
			41	R.EEQYN[+2080.765]STYR.V	HexNAc(4)Hex(4)Fuc(1)Sialic474(1)	N180	yes	1090.764	3	243.7	46.3	4.32	medium	1, 2, 3
			42	R.EEQYN[+2096.760]STYR.V	HexNAc(4)Hex(5)Sialic474(1)	N180	yes	1096.095	3	293.8	118.4	6.42	medium	1
			43	R.EEQYN[+2242.817]STYR.V	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)	N180	yes	1144.781	3	283.0	44.9	6.13	medium	1
16	Ig gamma-2 chain C region	P01859	44	K.TKPREEQFN[+1851.641]STFR.V	HexNAc(4)Hex(5)Sialic229(1)	N176	yes	873.628	4	323.7	30.6	5.15	medium	4
			45	K.TKPREEQFN[+2013.694]STFR.V	HexNAc(4)Hex(6)Sialic229(1)	N176	yes	914.138	4	355.5	38.5	5.66	good	3
			46	K.TKPREEQFN[+2080.765]STFR.V	HexNAc(4)Hex(4)Fuc(1)Sialic474(1)	N176	yes	1240.865	3	386.5	262.3	8.96	medium	1, 2, 4, 5
			47	K.TKPREEQFN[+2242.817]STFR.V	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)	N176	yes	971.414	4	363.3	238.5	8.88	medium	1, 3, 4, 5
			48	R.EEQFN[+1835.646]STFR.V	HexNAc(4)Hex(4)Fuc(1)Sialic229(1)	N176	yes	998.394	3	267.1	54.3	5.21	medium	1, 2, 3
			48	R.EEQFN[+2080.765]STFR.V	HexNAc(4)Hex(4)Fuc(1)Sialic474(1)	N176	yes	1080.101	3	318.4	68.2	6.75	medium	1, 2, 3, 4, 5
			49	R.EEQFN[+1934.707]STFR.V	HexNAc(4)Hex(4)Sialic474(1)	N176	yes	1031.415	3	221.2	39.1	3.48	medium	1
			50	R.EEQFN[+2242.817]STFR.V	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)	N176	yes	1134.119	3	319.4	75.3	6.99	medium	1, 2, 3, 4, 5
17	Ig gamma-3 chain C region	P01860	51	K.S[+839.310]C[+57.021]DTPPPC[+57.021]PR.C	HexNAc(1)Hex(1)Sialic474(1)	S119	no	675.941	3	373.6	18.5	9.36	good	1, 3, 4
			52	K.SC[+57.021]DT[+839.310]PPPC[+57.021]PR.C	HexNAc(1)Hex(1)Sialic474(1)	T122	no	675.941	3	438.8	14.3	8.70	medium	1, 3, 4, 5
			53	K.TKPREEQYN[+1835.646]STFR.V	HexNAc(4)Hex(4)Fuc(1)Sialic229(1)	N227	yes	873.628	4	329.0	15.3	3.88	medium	5
18	Ig gamma-4 chain C region	P01861	53	K.TKPREEQFN[+1835.646]STYR.V	HexNAc(4)Hex(4)Fuc(1)Sialic229(1)	N177	yes	873.628	4	287.6	43.2	4.06	good	2
			53	K.TKPREEQFN[+2080.765]STYR.V	HexNAc(4)Hex(4)Fuc(1)Sialic474(1)	N177	yes	934.901	4	185.2	2.5	1.86	poor	2
			54	K.TKPREEQFN[+2242.817]STYR.V	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)	N177	yes	975.414	4	372.3	72.8	8.21	medium	3
19	Ig alpha-1 chain C region	P01876	55	R.LSLHRPALEDLLLGSEAN[+1934.707]LTC[+57.021]TLTGLR.D	HexNAc(4)Hex(4)Sialic474(1)	N144	yes	1225.585	4	204.7	5.4	1.25	poor	3
			56	R.PALEDLLLGSEAN[+1731.627]LTC[+57.021]TLTGLR.D	HexNAc(3)Hex(4)Sialic474(1)	N144	yes	1022.715	4	212.4	39.1	1.72	medium	2
20	Apolipoprotein E	P02649	57	A.KVEQAVET[+839.310]EPEPELR.Q	HexNAc(1)Hex(1)Sialic474(1)	T26	yes	865.074	3	418.5	190.7	9.84	medium	1, 2, 3, 4
			58	K.VQAAVGT[+1068.369]SAAPVPSDNH.-	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T307	yes	897.060	3	368.1	3.4	6.38	poor	2, 5
			58	K.VQAAVGT[+1313.488]SAAPVPSDNH.-	HexNAc(1)Hex(1)Sialic474(2)	T307	yes	978.767	3	392.9	9.2	9.14	medium	4, 5
			58	K.VQAAVGT[+823.251]SAAPVPSDNH.-	HexNAc(1)Hex(1)Sialic229(2)	T307	yes	815.354	3	368.9	1.8	10.22	poor	3
			59	K.VQAAVGT[+839.310]SAAPVPSDNH.-	HexNAc(1)Hex(1)Sialic474(1)	T307	yes	820.708	3	396.1	18.4	9.87	good	1, 4
			60	K.VQAAVGTS[+1068.369]AAPVPSDNH.-	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S308	yes	897.060	3	391.3	22.7	9.50	medium	1, 2, 3, 4, 5
			60	K.VQAAVGTS[+1313.488]AAPVPSDNH.-	HexNAc(1)Hex(1)Sialic474(2)	S308	yes	978.768	3	381.8	12.2	9.40	medium	1, 2, 3, 4, 5
			60	K.VQAAVGTS[+823.251]AAPVPSDNH.-	HexNAc(1)Hex(1)Sialic229(2)	S308	yes	815.355	3	473.2	20.1	9.75	good	1, 2, 3, 4, 5
			61	K.VQAAVGTS[+594.191]AAPVPSDNH.-	HexNAc(1)Hex(1)Sialic229(1)	S308	yes	739.001	3	470.0	11.3	10.76	good	1, 2, 3, 4, 5
			61	K.VQAAVGTS[+839.310]AAPVPSDNH.-	HexNAc(1)Hex(1)Sialic474(1)	S308	yes	820.707	3	478.2	18.0	10.91	good	1, 2, 3, 4, 5
			62	K.VQAAVGTSAAPVPS[+1068.369]DNH.-	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S314	yes	897.060	3	579.5	6.2	12.16	poor	1, 2, 4
			62	K.VQAAVGTSAAPVPS[+1313.488]DNH.-	HexNAc(1)Hex(1)Sialic474(2)	S314	yes	978.767	3	350.4	4.8	6.93	poor	1, 5
			62	K.VQAAVGTSAAPVPS[+823.251]DNH.-	HexNAc(1)Hex(1)Sialic229(2)	S314	yes	815.353	3	389.6	0.7	8.43	medium	1
			63	K.VQAAVGTSAAPVPS[+594.191]DNH.-	HexNAc(1)Hex(1)Sialic229(1)	S314	yes	739.001	3	552.1	3.5	12.20	poor	1, 2, 3, 4
			63	K.VQAAVGTSAAPVPS[+839.310]DNH.-	HexNAc(1)Hex(1)Sialic474(1)	S314	yes	820.708	3	518.9	13.0	10.33	good	1, 2, 4
			64	R.AAT[+1068.369]VGSLAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T212	yes	856.063	3	514.6	51.5	10.76	good	1, 2, 3, 4, 5
			64	R.AAT[+1313.488]VGSLAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(2)	T212	yes	937.769	3	375.9	50.4	8.58	medium	1, 2, 3, 4, 5
			64	R.AAT[+823.251]VGSLAGQPLQER.A	HexNAc(1)Hex(1)Sialic229(2)	T212	yes	774.356	3	348.5	24.5	8.07	good	1, 2, 3, 5
			65	R.AAT[+1204.442]VGSLAGQPLQER.A	HexNAc(2)Hex(2)Sialic474(1)	T212	yes	901.419	3	346.0	15.9	7.95	good	1, 2, 3, 4, 5
			65	R.AAT[+959.324]VGSLAGQPLQER.A	HexNAc(2)Hex(2)Sialic229(1)	T212	yes	819.714	3	539.9	35.3	10.87	good	1, 2, 3, 4, 5
			66	R.AAT[+1350.500]VGSLAGQPLQER.A	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)	T212	yes	950.106	3	345.7	14.6	6.80	good	1
			67	R.AAT[+1433.501]VGSLAGQPLQER.A	HexNAc(2)Hex(2)Sialic474(1)Sialic229(1)	T212	yes	977.773	3	301.8	16.1	5.61	medium	1, 2, 4
			68	R.AAT[+594.191]VGSLAGQPLQER.A	HexNAc(1)Hex(1)Sialic229(1)	T212	yes	698.002	3	446.4	36.8	10.32	good	1, 2, 3, 4, 5

			68	R.AAT[+839.310]VGSLAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(1)	T212	yes	779.709	3	566.2	35.4	11.32	high	1, 2, 3, 4, 5
			69	R.AAT[+797.271]VGSLAGQPLQER.A	HexNAc(2)Hex(1)Sialic229(1)	T212	yes	765.696	3	335.1	28.4	8.30	good	1, 3, 4, 5
			70	R.AATVGS[+1068.369]LAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S215	yes	856.062	3	654.7	44.6	14.14	medium	1, 2, 3, 4, 5
			70	R.AATVGS[+1313.488]LAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(2)	S215	yes	937.768	3	440.9	37.2	9.90	medium	1, 2, 3, 4, 5
			70	R.AATVGS[+823.251]LAGQPLQER.A	HexNAc(1)Hex(1)Sialic229(2)	S215	yes	774.355	3	489.0	37.2	10.51	medium	1, 2, 3, 4, 5
			71	R.AATVGS[+1204.442]LAGQPLQER.A	HexNAc(2)Hex(2)Sialic474(1)	S215	yes	901.419	3	580.4	36.5	12.76	medium	1, 2, 3, 4, 5
			71	R.AATVGS[+959.324]LAGQPLQER.A	HexNAc(2)Hex(2)Sialic229(1)	S215	yes	819.713	3	659.5	41.0	14.66	medium	1, 2, 3, 4, 5
			72	R.AATVGS[+1433.501]LAGQPLQER.A	HexNAc(2)Hex(2)Sialic474(1)Sialic229(1)	S215	yes	977.773	3	362.9	12.1	6.41	medium	1, 2, 3, 4, 5
			73	R.AATVGS[+594.191]LAGQPLQER.A	HexNAc(1)Hex(1)Sialic229(1)	S215	yes	698.002	3	660.2	33.5	14.79	medium	1, 2, 3, 4, 5
			73	R.AATVGS[+839.310]LAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(1)	S215	yes	779.709	3	664.3	46.2	14.86	medium	1, 2, 3, 4, 5
			74	R.AATVGS[+797.271]LAGQPLQER.A	HexNAc(2)Hex(1)Sialic229(1)	S215	yes	765.696	3	455.5	34.0	9.67	medium	1, 2, 3, 4, 5
			75	R.QQT[+839.310]EWQSGQR.W	HexNAc(1)Hex(1)Sialic474(1)	T36	yes	696.301	3	235.4	63.6	3.82	high	1, 2
			76	R.QQTEWQS[+839.310]GQR.W	HexNAc(1)Hex(1)Sialic474(1)	S40	no	696.301	3	374.6	11.2	7.24	medium	1
			77	R.VRAAT[+1068.369]VGSLAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T212	yes	941.119	3	277.2	8.8	5.84	poor	1, 2, 3, 4, 5
			77	R.VRAAT[+1313.488]VGSLAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(2)	T212	yes	1022.825	3	181.5	1.4	2.66	poor	2
			78	R.VRAATVGS[+1068.369]LAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S215	yes	941.119	3	382.9	4.5	10.10	poor	2
			78	R.VRAATVGS[+1313.488]LAGQPLQER.A	HexNAc(1)Hex(1)Sialic474(2)	S215	yes	1022.825	3	377.6	6.2	8.96	poor	2
			79	T.VGS[+1162.403]LAGQPLQER.A	HexNAc(3)Hex(2)Sialic229(1)	S215	yes	806.361	3	346.6	74.2	4.46	medium	4
			80	W.AGLVEKVQAAVGTS[+1068.369]AAPVPSDNH.-	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S308	yes	1096.175	3	412.5	4.8	9.76	poor	1
			81	W.AGLVEKVQAAVGTSAAPVPS[+1068.369]DNH.-	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S314	yes	1096.175	3	442.7	3.9	8.67	poor	1
21	Apolipoprotein C-III	P02656	82	W.DLDPEVRPT[+1068.369]SAVAA.-	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T94	yes	837.039	3	402.5	4.0	7.56	poor	3, 4
			83	W.DLDPEVRPT[+594.191]SAVAA.-	HexNAc(1)Hex(1)Sialic229(1)	T94	yes	1017.966	2	377.8	1.6	8.74	poor	1, 3, 4, 5
			83	W.DLDPEVRPT[+839.310]SAVAA.-	HexNAc(1)Hex(1)Sialic474(1)	T94	yes	760.687	3	355.7	4.0	6.91	poor	1, 3, 5
22	Fibrinogen gamma chain	P02679	84	K.VDKDLQSLEDILHQVEN[+2095.747]K.T	HexNAc(6)Hex(4)Sialic229(1)	N78	yes	1055.721	4	295.2	20.5	5.90	medium	1
			85	K.VDKDLQSLEDILHQVEN[+2096.760]K.T	HexNAc(4)Hex(5)Sialic474(1)	N78	yes	1055.720	4	273.6	23.4	5.79	good	2
23	Beta-2-glycoprotein 1	P02749	86	R.VYKPSAGN[+2096.760]NSLYR.D	HexNAc(4)Hex(5)Sialic474(1)	N162	yes	1189.176	3	203.3	40.5	4.32	medium	1, 4
			87	R.VYKPSAGN[+2570.938]NSLYR.D	HexNAc(4)Hex(5)Sialic474(2)	N162	yes	1010.680	4	340.1	155.9	8.24	medium	1, 2, 3, 4, 5
24	Alpha-1-acid glycoprotein 1	P02763	88	R.NEEYN[+2096.760]K.S	HexNAc(4)Hex(5)Sialic474(1)	N56	yes	965.041	3	204.5	26.7	3.07	medium	1, 4
			89	R.NEEYN[+2570.938]K.S	HexNAc(4)Hex(5)Sialic474(2)	N56	yes	1123.101	3	225.6	41.9	3.72	medium	1, 2, 5
25	Alpha-2-HS-glycoprotein	P02765	90	K.AALAAFNAQNN[+1851.641]GSNFQLEEISR.A	HexNAc(4)Hex(5)Sialic229(1)	N176	yes	1406.605	3	396.5	154.8	8.26	medium	5
			91	K.AALAAFNAQNN[+2325.819]GSNFQLEEISR.A	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N176	yes	1173.751	4	413.4	179.0	9.42	medium	1, 3
			91	K.AALAAFNAQNN[+2570.938]GSNFQLEEISR.A	HexNAc(4)Hex(5)Sialic474(2)	N176	yes	1235.029	4	369.3	123.3	7.39	medium	2, 4, 5
			92	K.VC[+57.021]QDC[+57.021]PLLAPLN[+2096.760]DTR.V	HexNAc(4)Hex(5)Sialic474(1)	N156	yes	1290.208	3	283.0	86.8	6.41	medium	1, 2, 3
			93	K.VC[+57.021]QDC[+57.021]PLLAPLN[+2325.819]DTR.V	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N156	yes	1025.173	4	266.7	67.3	6.06	medium	1
			93	K.VC[+57.021]QDC[+57.021]PLLAPLN[+2570.938]DTR.V	HexNAc(4)Hex(5)Sialic474(2)	N156	yes	1448.599	3	323.1	82.1	6.85	medium	1, 2, 3, 4
			94	R.T[+594.191]VVQPSVGAAAGPVVPPC[+57.021]PGR.I	HexNAc(1)Hex(1)Sialic229(1)	T341	no	870.759	3	609.1	1.1	14.66	medium	2
			95	R.TVVQPS[+1068.369]VGAAAGPVVPPC[+57.021]PGR.I	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S346	yes	1028.818	3	583.1	29.2	12.35	medium	2
			96	R.TVVQPS[+594.191]VGAAAGPVVPPC[+57.021]PGR.I	HexNAc(1)Hex(1)Sialic229(1)	S346	yes	870.758	3	558.6	37.2	11.44	good	1, 2, 3, 4, 5
			96	R.TVVQPS[+839.310]VGAAAGPVVPPC[+57.021]PGR.I	HexNAc(1)Hex(1)Sialic474(1)	S346	yes	952.465	3	646.5	45.8	14.15	high	1, 2, 3, 4, 5
26	Serum albumin	P02768	97	A.DLPS[+1553.580]LAADFVESK.D	HexNAc(3)Hex(2)Fuc(1)Sialic474(1)	S328	no	982.433	3	493.0	13.6	7.74	medium	2
			98	A.DLPSLAADFVES[+1678.620]K.D	HexNAc(2)Hex(2)Sialic474(2)	S336	no	1024.105	3	403.3	14.9	5.15	medium	1
			99	A.DLPSLAADFVES[+1876.672]K.D	HexNAc(5)Hex(3)Fuc(1)Sialic229(1)	S336	no	1090.141	3	433.2	15.2	4.30	medium	1
			100	C.DKS[+1579.559]LHTLFGDK.L	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)Sialic229(1)	S89	no	947.411	3	300.7	6.1	1.91	poor	2
			101	C.DKSLHT[+1579.559]LFGDK.L	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)Sialic229(1)	T92	no	947.411	3	305.0	13.0	2.73	medium	4
			102	E.MPADLPS[+1496.558]LAADFVESK.D	HexNAc(2)Hex(2)Fuc(2)Sialic474(1)	S328	no	1063.134	3	262.6	1.6	1.60	poor	1
			103	E.NDEMPADLPS[+1026.330]LAADFVESK.D	HexNAc(2)Hex(1)Sialic229(2)	S328	no	1025.773	3	363.2	3.2	4.42	poor	4
			104	E.NDEMPADLPS[+1251.439]LAADFVESK.D	HexNAc(2)Hex(2)Fuc(2)Sialic229(1)	S328	no	1100.794	3	350.6	6.1	2.86	poor	4
			105	E.VENDEMPADLPS[+943.329]LAADFVESK.D	HexNAc(2)Hex(1)Fuc(1)Sialic229(1)	S328	no	1074.143	3	412.5	5.9	5.86	poor	2
			106	L.PS[+1845.631]LAADFVESK.D	HexNAc(3)Hex(3)Fuc(2)Sialic229(2)	S328	no	1003.753	3	301.2	83.1	5.21	medium	5
			107	L.PSLAADFVES[+1824.678]K.D	HexNAc(2)Hex(2)Fuc(1)Sialic474(2)	S336	no	996.768	3	431.4	17.9	5.61	medium	4
			108	L.PSLAADFVES[+1845.631]K.D	HexNAc(3)Hex(3)Fuc(2)Sialic229(2)	S336	no	1003.753	3	509.1	48.7	7.63	medium	1, 2, 3, 4, 5
			109	M.PADLPSLAADFVES[+1678.620]K.D	HexNAc(2)Hex(2)Sialic474(2)	S336	no	1080.145	3	389.0	3.2	6.05	poor	3
			110	P.ADLPS[+1496.558]LAADFVESK.D	HexNAc(2)Hex(2)Fuc(2)Sialic474(1)	S328	no	987.103	3	340.1	13.3	5.75	medium	3, 4
			111	P.ADLPSLAADFVES[+1496.558]K.D	HexNAc(2)Hex(2)Fuc(2)Sialic474(1)	S336	no	987.103	3	395.6	9.9	5.33	good	2, 4

			112	R.RPC[+57.021]FSALEVDET[+1699.637]YVPK.E	HexNAc(3)Hex(2)Fuc(2)Sialic474(1)	T520	no	903.397	4	332.8	34.2	7.68	high	2
27	Serotransferrin	P02787	113	K.C[+57.021]GLVPVLAENYN[+1851.641]K.S	HexNAc(4)Hex(5)Sialic229(1)	N432	yes	1110.135	3	200.8	56.3	4.57	medium	3
			113	K.C[+57.021]GLVPVLAENYN[+2096.760]K.S	HexNAc(4)Hex(5)Sialic474(1)	N432	yes	1191.843	3	321.8	143.1	7.22	medium	2, 3, 4, 5
			114	K.C[+57.021]GLVPVLAENYN[+1934.707]K.S	HexNAc(4)Hex(4)Sialic474(1)	N432	yes	1137.825	3	404.7	171.5	9.66	good	1, 3
			115	K.C[+57.021]GLVPVLAENYN[+2325.819]K.S	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N432	yes	1268.194	3	265.3	115.9	8.08	medium	4
			115	K.C[+57.021]GLVPVLAENYN[+2570.938]K.S	HexNAc(4)Hex(5)Sialic474(2)	N432	yes	1349.902	3	235.3	84.6	4.67	medium	1, 5
			116	Q.QQHLFGSN[+2471.877]VTDC[+57.021]SGNFC[+57.021]LFR.S	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)Sialic229(1)	N630	yes	1215.494	4	237.6	19.6	2.60	medium	5
			117	R.APNHAVVT[+839.310]R.K	HexNAc(1)Hex(1)Sialic474(1)	T608	no	601.953	3	237.9	10.8	1.56	medium	3
			118	R.QQQHLFGSN[+1835.646]VTDC[+57.021]SGNFC[+57.021]LFR.S	HexNAc(4)Hex(4)Fuc(1)Sialic229(1)	N630	yes	1450.581	3	259.9	58.1	5.32	medium	1, 4
			119	R.QQQHLFGSN[+1851.641]VTDC[+57.021]SGNFC[+57.021]LFR.S	HexNAc(4)Hex(5)Sialic229(1)	N630	yes	1092.445	4	274.9	33.9	5.61	medium	1, 3, 4, 5
			119	R.QQQHLFGSN[+2096.760]VTDC[+57.021]SGNFC[+57.021]LFR.S	HexNAc(4)Hex(5)Sialic474(1)	N630	yes	1153.726	4	499.1	23.7	10.38	medium	1, 2, 3, 4
			120	R.QQQHLFGSN[+2038.725]VTDC[+57.021]SGNFC[+57.021]LFR.S	HexNAc(5)Hex(4)Fuc(1)Sialic229(1)	N630	yes	1139.218	4	325.1	57.7	6.01	good	2, 5
			120	R.QQQHLFGSN[+2283.844]VTDC[+57.021]SGNFC[+57.021]LFR.S	HexNAc(5)Hex(4)Fuc(1)Sialic474(1)	N630	yes	1200.503	4	434.5	195.4	8.50	medium	1
			121	R.QQQHLFGSN[+2080.700]VTDC[+57.021]SGNFC[+57.021]LFR.S	HexNAc(4)Hex(5)Sialic229(2)	N630	yes	1149.712	4	309.4	91.1	6.90	medium	2, 3, 4, 5
			121	R.QQQHLFGSN[+2325.819]VTDC[+57.021]SGNFC[+57.021]LFR.S	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N630	yes	1210.990	4	490.7	263.1	10.84	medium	1, 2, 3, 4, 5
			122	R.QQQHLFGSN[+2308.811]VTDC[+57.021]SGNFC[+57.021]LFR.S	HexNAc(6)Hex(3)Fuc(1)Sialic229(2)	N630	yes	1206.734	4	269.1	66.0	4.69	medium	1, 2, 5
28	Hemopexin	P02790	123	A.T[+1068.369]PLPPTSAHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T24	yes	946.679	4	352.6	20.0	5.58	good	3, 4
			124	A.T[+1678.620]PLPPTSAHGNVAEGETKPDPDVTER.C	HexNAc(2)Hex(2)Sialic474(2)	T24	yes	1099.242	4	342.7	4.4	6.50	poor	1
			125	A.T[+839.310]PLPPTSAHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)	T24	yes	889.415	4	403.6	11.3	8.03	good	1, 2, 3, 5
			126	A.T[+959.324]PLPPTSAHGNVAEGETKPDPDVTER.C	HexNAc(2)Hex(2)Sialic229(1)	T24	yes	919.419	4	282.1	4.4	3.47	poor	3
			127	A.TPLPPT[+1204.442]SAHGNVAEGETKPDPDVTER.C	HexNAc(2)Hex(2)Sialic474(1)	T29	yes	980.697	4	481.8	1.8	10.20	medium	2
			128	A.TPLPPT[+594.191]SAHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic229(1)	T29	yes	828.134	4	498.3	4.5	10.32	poor	1, 2, 4, 5
			129	A.TPLPPTS[+594.191]AHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic229(1)	S30	no	828.136	4	253.2	13.3	4.01	medium	1, 3, 5
			129	A.TPLPPTS[+839.310]AHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)	S30	no	889.414	4	483.7	17.7	10.10	medium	5
			130	A.TPLPPTS[+959.324]AHGNVAEGETKPDPDVTER.C	HexNAc(2)Hex(2)Sialic229(1)	S30	no	919.416	4	355.0	5.2	6.60	poor	1, 4, 5
			131	A.TPLPPTSAHGNVAEGET[+1068.369]KPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T40	no	946.680	4	604.3	10.1	11.93	good	1, 2, 4
			131	A.TPLPPTSAHGNVAEGET[+1313.488]KPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(2)	T40	no	1007.958	4	417.8	2.7	8.06	poor	4, 5
			132	A.TPLPPTSAHGNVAEGET[+1188.383]KPDPDVTER.C	HexNAc(2)Hex(2)Sialic229(2)	T40	no	976.681	4	354.4	1.3	8.13	poor	4
			133	A.TPLPPTSAHGNVAEGET[+1204.442]KPDPDVTER.C	HexNAc(2)Hex(2)Sialic474(1)	T40	no	980.697	4	330.3	7.2	6.14	poor	1
			133	A.TPLPPTSAHGNVAEGET[+959.324]KPDPDVTER.C	HexNAc(2)Hex(2)Sialic229(1)	T40	no	919.416	4	290.7	6.4	4.93	good	1, 3, 5
			134	A.TPLPPTSAHGNVAEGET[+594.191]KPDPDVTER.C	HexNAc(1)Hex(1)Sialic229(1)	T40	no	828.134	4	470.1	16.6	9.84	good	1, 2, 3, 4, 5
			134	A.TPLPPTSAHGNVAEGET[+839.310]KPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)	T40	no	889.414	4	486.8	15.1	10.68	medium	1, 2, 3, 4
			135	A.TPLPPTSAHGNVAEGETKPDPDVT[+1433.501]ER.C	HexNAc(2)Hex(2)Sialic474(1)Sialic229(1)	T47	no	1037.962	4	426.4	4.8	8.13	poor	5
			136	K.ALPQPQN[+1851.641]VTSLLGC[+57.021]TH.-	HexNAc(4)Hex(5)Sialic229(1)	N453	yes	1196.509	3	217.0	7.7	3.44	medium	1
			136	K.ALPQPQN[+2096.760]VTSLLGC[+57.021]TH.-	HexNAc(4)Hex(5)Sialic474(1)	N453	yes	1278.214	3	238.6	73.5	5.44	medium	1, 2, 3, 4
			137	K.ALPQPQN[+1934.707]VTSLLGC[+57.021]TH.-	HexNAc(4)Hex(4)Sialic474(1)	N453	yes	1224.203	3	330.8	160.6	6.20	good	1, 4
			138	K.ALPQPQN[+2080.700]VTSLLGC[+57.021]TH.-	HexNAc(4)Hex(5)Sialic229(2)	N453	yes	1272.864	3	278.2	105.1	6.66	medium	1, 2, 3, 4, 5
			138	K.ALPQPQN[+2325.819]VTSLLGC[+57.021]TH.-	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N453	yes	1354.575	3	308.9	120.4	6.46	medium	1, 2, 3, 4, 5
			138	K.ALPQPQN[+2570.938]VTSLLGC[+57.021]TH.-	HexNAc(4)Hex(5)Sialic474(2)	N453	yes	1077.461	4	209.2	77.8	4.65	medium	1, 3
			139	L.PPTSAHGNVAEGETKPDPDVT[+985.368]ER.C	HexNAc(1)Hex(1)Fuc(1)Sialic474(1)	T47	no	1130.496	3	269.7	6.5	2.70	medium	2
			140	P.LPPT[+839.310]SAHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)	T29	yes	839.889	4	380.9	9.2	7.85	medium	2, 3
			141	P.LPPTS[+839.310]AHGNVAEGETKPDPDVTER.C	HexNAc(1)Hex(1)Sialic474(1)	S30	no	839.889	4	490.8	2.9	10.38	poor	1, 2, 3, 5
			142	P.LPPTSAHGNVAEGETKPDPDVT[+839.310]ER.C	HexNAc(1)Hex(1)Sialic474(1)	T47	no	839.889	4	650.1	1.8	13.91	good	5
			143	R.SWPAVGN[+2325.819]C[+57.021]SSALR.W	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N187	yes	1244.167	3	251.7	13.0	4.03	medium	2
			143	R.SWPAVGN[+2570.938]C[+57.021]SSALR.W	HexNAc(4)Hex(5)Sialic474(2)	N187	yes	1325.876	3	284.4	71.3	5.57	medium	2
			144	W.PAVGN[+1851.641]C[+57.021]SSALR.W	HexNAc(4)Hex(5)Sialic229(1)	N187	yes	995.071	3	194.8	5.3	2.78	poor	4

			144	W.PAVGN[+2096.760]C[+57.021]SSALR.W	HexNAc(4)Hex(5)Sialic474(1)	N187	yes	1076.778	3	260.3	62.1	5.14	good	1, 3, 4, 5
			145	W.PAVGN[+1934.707]C[+57.021]SSALR.W	HexNAc(4)Hex(4)Sialic474(1)	N187	yes	1022.760	3	318.9	12.1	4.78	medium	3
			146	W.PAVGN[+2325.819]C[+57.021]SSALR.W	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N187	yes	1153.131	3	309.6	55.2	6.70	medium	1, 3, 4
			146	W.PAVGN[+2570.938]C[+57.021]SSALR.W	HexNAc(4)Hex(5)Sialic474(2)	N187	yes	1234.837	3	288.0	62.4	6.16	medium	1, 3, 4
29	Vitronectin	P04004	147	K.N[+1893.680]GSLFAFR.G	HexNAc(3)Hex(5)Sialic474(1)	N169	yes	935.724	3	250.8	47.7	4.60	medium	1
			148	K.N[+2096.760]GSLFAFR.G	HexNAc(4)Hex(5)Sialic474(1)	N169	yes	1003.417	3	247.6	82.4	5.29	medium	1
30	Histidine-rich glycoprotein	P04196	149	R.S[+594.191]STTKPPFKPHGSR.D	HexNAc(1)Hex(1)Sialic229(1)	S271	no	707.671	3	279.1	8.9	6.43	poor	2
			149	R.S[+839.310]STTKPPFKPHGSR.D	HexNAc(1)Hex(1)Sialic474(1)	S271	no	592.285	4	304.6	6.1	6.95	medium	2, 5
			150	R.SS[+594.191]TTKPPFKPHGSR.D	HexNAc(1)Hex(1)Sialic229(1)	S272	no	707.672	3	205.5	2.4	2.02	poor	5
			151	R.SST[+594.191]TKPPFKPHGSR.D	HexNAc(1)Hex(1)Sialic229(1)	T273	no	707.671	3	173.8	2.5	2.48	poor	4
			151	R.SST[+839.310]TKPPFKPHGSR.D	HexNAc(1)Hex(1)Sialic474(1)	T273	no	592.285	4	324.7	14.4	8.20	medium	3
			152	R.SSTT[+594.191]KPPFKPHGSR.D	HexNAc(1)Hex(1)Sialic229(1)	T274	no	707.671	3	306.2	2.7	5.61	poor	2, 4
			152	R.SSTT[+839.310]KPPFKPHGSR.D	HexNAc(1)Hex(1)Sialic474(1)	T274	no	592.285	4	406.4	5.8	8.87	poor	1, 2, 3, 4
			153	R.SSTTKPPFKPHGS[+594.191]R.D	HexNAc(1)Hex(1)Sialic229(1)	S283	no	707.671	3	174.9	0.0	1.35	poor	3
31	Alpha-1B-glycoprotein	P04217	154	R.SS[+740.249]TSPDR.I	HexNAc(1)Hex(1)Fuc(1)Sialic229(1)	S256	yes	745.300	2	288.6	5.0	1.20	poor	2
32	Calpain small subunit 1	P04632	155	G.GGGGGGGGGGLGGGLGNVLGGLIS[+1068.369]GAGGGGGGGGGGGGGGGGGGTAMR.I	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S34	no	1149.251	4	291.4	50.3	5.64	medium	1
			156	K.GGGGGGGGGGGLGGGLGNVLGGLIS[+1042.389]GAGGGGGGGGGGGGGGGGGGTAMR.I	HexNAc(2)Hex(1)Sialic474(1)	S34	no	1157.018	4	280.3	14.1	5.05	medium	1
			157	K.GGGGGGGGGGGLGGGLGNVLGGLIS[+1052.310]GAGGGGGGGGGGGGGGGGGGTAMR.I	HexNAc(1)Hex(1)Sialic229(3)	S34	no	1159.505	4	282.8	23.2	4.14	medium	1
33	Secretogranin-1	P05060	158	R.GEAGAPGEEDIQGPT[+1068.369]KADTEK.W	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T116	no	1056.786	3	371.9	33.5	8.24	medium	1, 4
			158	R.GEAGAPGEEDIQGPT[+1313.488]KADTEK.W	HexNAc(1)Hex(1)Sialic474(2)	T116	no	1138.493	3	344.2	23.3	6.16	medium	1, 5
			159	R.GEAGAPGEEDIQGPT[+839.310]KADTEK.W	HexNAc(1)Hex(1)Sialic474(1)	T116	no	980.432	3	374.0	13.5	6.09	medium	1, 4
34	Amyloid beta A4 protein	P05067	160	R.GLT[+1068.369]TRPGSGLTNIK.T	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T651	yes	828.396	3	287.0	5.9	5.42	poor	1, 2, 3, 4, 5
			160	R.GLT[+1313.488]TRPGSGLTNIK.T	HexNAc(1)Hex(1)Sialic474(2)	T651	yes	910.102	3	280.3	14.6	5.60	medium	2, 3, 5
			161	R.GLT[+839.310]TRPGSGLTNIK.T	HexNAc(1)Hex(1)Sialic474(1)	T651	yes	752.042	3	373.5	6.4	9.24	poor	1, 3
			162	R.GLTT[+1068.369]RPGSGLTNIK.T	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T652	yes	828.397	3	349.2	2.0	4.80	poor	2
			163	R.GLTT[+594.191]RPGSGLTNIK.T	HexNAc(1)Hex(1)Sialic229(1)	T652	yes	670.336	3	270.1	1.8	3.85	poor	1
			163	R.GLTT[+839.310]RPGSGLTNIK.T	HexNAc(1)Hex(1)Sialic474(1)	T652	yes	752.043	3	284.6	1.7	7.31	poor	2, 3, 5
			164	R.GLTTRPGSGLT[+1313.488]NIK.T	HexNAc(1)Hex(1)Sialic474(2)	T659	yes	910.102	3	200.2	8.8	1.82	medium	1, 2
			165	R.GLTTRPGSGLT[+839.310]NIK.T	HexNAc(1)Hex(1)Sialic474(1)	T659	yes	752.043	3	311.8	7.9	5.88	medium	2, 4, 5
35	Apolipoprotein D	P05090	166	R.ADGTVNQIEGEATPVN[+1851.641]LTEPAKLEVK.F	HexNAc(4)Hex(5)Sialic229(1)	N98	yes	1144.518	4	388.2	159.1	7.97	medium	1, 5
			167	R.ADGTVNQIEGEATPVN[+2080.700]LTEPAKLEVK.F	HexNAc(4)Hex(5)Sialic229(2)	N98	yes	1201.781	4	281.2	104.6	7.11	medium	4
			168	R.ADGTVNQIEGEATPVN[+2242.817]LTEPAKLEVK.F	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)	N98	yes	1242.564	4	298.8	68.2	4.28	medium	1
36	Plasma serine protease inhibitor	P05154	169	R.VEDLHVGAT[+839.310]VAPSSR.R	HexNAc(1)Hex(1)Sialic474(1)	T39	yes	793.041	3	261.8	55.2	6.71	medium	2, 4
			170	R.VEDLHVGATVAPS[+839.310]SR.R	HexNAc(1)Hex(1)Sialic474(1)	S43	no	793.041	3	505.0	2.3	11.36	poor	4
			171	R.VEDLHVGATVAPSS[+839.310]R.R	HexNAc(1)Hex(1)Sialic474(1)	S44	no	793.041	3	396.3	4.6	8.60	poor	1, 3
37	Plasma protease C1 inhibitor	P05155	172	K.VAT[+594.191]TVISK.M	HexNAc(1)Hex(1)Sialic229(1)	T47	yes	706.848	2	372.8	16.7	6.03	medium	2, 3, 4, 5
			172	K.VAT[+839.310]TVISK.M	HexNAc(1)Hex(1)Sialic474(1)	T47	yes	829.408	2	382.0	5.1	6.00	poor	1, 2, 3
			173	K.VATT[+594.191]VISK.M	HexNAc(1)Hex(1)Sialic229(1)	T48	yes	706.847	2	458.4	1.0	7.17	poor	1
			173	K.VATT[+839.310]VISK.M	HexNAc(1)Hex(1)Sialic474(1)	T48	yes	829.408	2	299.2	3.2	3.01	poor	3, 5
38	Tetranectin	P05452	174	T.EPPT[+839.310]QKPK.K	HexNAc(1)Hex(1)Sialic474(1)	T25	yes	588.614	3	424.6	110.4	7.42	medium	1, 2, 3, 4, 5
39	Cathepsin D	P07339	175	K.YSQAVPAVT[+1068.369]EGPIPEVLK.N	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T63	yes	989.470	3	465.9	73.7	9.89	medium	1, 2, 3, 4, 5
			176	K.YSQAVPAVT[+839.310]EGPIPEVLK.N	HexNAc(1)Hex(1)Sialic474(1)	T63	yes	913.118	3	410.5	86.4	8.97	medium	1, 4, 5
40	Complement C4-A	P0C0L4	177	W.GSVTGSQS[+839.310]NAVSPTPAPR.N	HexNAc(1)Hex(1)Sialic474(1)	S1238	no	851.395	3	399.3	4.6	8.74	poor	1, 2, 3, 4, 5
			178	W.GSVTGSQSNAVS[+1068.369]PTPAPR.N	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S1242	no	927.747	3	284.1	12.4	6.02	poor	3, 4, 5
			178	W.GSVTGSQSNAVS[+1313.488]PTPAPR.N	HexNAc(1)Hex(1)Sialic474(2)	S1242	no	1009.453	3	279.9	5.9	4.81	medium	1, 5
			179	W.GSVTGSQSNAVS[+594.191]PTPAPR.N	HexNAc(1)Hex(1)Sialic229(1)	S1242	no	769.688	3	423.9	18.0	9.95	good	1, 3, 5
			179	W.GSVTGSQSNAVS[+839.310]PTPAPR.N	HexNAc(1)Hex(1)Sialic474(1)	S1242	no	851.394	3	384.1	26.9	7.59	medium	2
			180	W.GSVTGSQSNAVSPT[+1068.369]PAPR.N	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T1244	yes	927.747	3	548.7	6.3	11.53	poor	1, 3, 4, 5
			180	W.GSVTGSQSNAVSPT[+1313.488]PAPR.N	HexNAc(1)Hex(1)Sialic474(2)	T1244	yes	1009.453	3	462.4	18.7	10.27	medium	4, 5
			181	W.GSVTGSQSNAVSPT[+594.191]PAPR.N	HexNAc(1)Hex(1)Sialic229(1)	T1244	yes	769.687	3	579.9	12.8	12.09	medium	3, 5
			181	W.GSVTGSQSNAVSPT[+839.310]PAPR.N	HexNAc(1)Hex(1)Sialic474(1)	T1244	yes	851.394	3	586.2	14.2	12.42	medium	1, 2, 3, 4, 5

41	Complement component C7	P10643	182	K.ENPLT[+839.310]QAVPK.C	HexNAc(1)Hex(1)Sialic474(1)	T696	yes	968.459	2	376.2	93.9	7.67	medium	1
42	Chromogranin-A	P10645	183	K.AEGNNQAPGEEEEEEEEAT[+1068.369]NTHPPASLPSQK.Y	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T181	yes	1097.710	4	600.8	2.8	12.27	medium	3
			184	K.AEGNNQAPGEEEEEEEEATNT[+1068.369]HPPASLPSQK.Y	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T183	yes	1097.710	4	449.2	2.9	7.85	poor	3
			185	R.EEEEEEEEEAEAGEEAVPEEEGPT[+839.310]VVLNPHPSLGYK.E	HexNAc(1)Hex(1)Sialic474(1)	T251	yes	1206.022	4	340.6	9.0	6.76	poor	4
43	Clusterin	P10909	186	R.EIRHN[+1835.646]STGC[+57.021]LR.M	HexNAc(4)Hex(4)Fuc(1)Sialic229(1)	N291	yes	1060.110	3	191.5	54.5	3.83	medium	3
			187	R.EIRHN[+1997.699]STGC[+57.021]LR.M	HexNAc(4)Hex(5)Fuc(1)Sialic229(1)	N291	yes	835.846	4	207.1	51.1	4.00	medium	3, 4
			187	R.EIRHN[+2242.817]STGC[+57.021]LR.M	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)	N291	yes	897.126	4	216.9	77.5	4.47	medium	1, 2, 3, 4
			188	R.EIRHN[+2038.725]STGC[+57.021]LR.M	HexNAc(5)Hex(4)Fuc(1)Sialic229(1)	N291	yes	846.102	4	136.7	47.5	4.25	good	4
			188	R.EIRHN[+2283.844]STGC[+57.021]LR.M	HexNAc(5)Hex(4)Fuc(1)Sialic474(1)	N291	yes	907.383	4	285.0	95.3	6.09	good	1, 3
			189	R.EIRHN[+2200.778]STGC[+57.021]LR.M	HexNAc(5)Hex(5)Fuc(1)Sialic229(1)	N291	yes	886.616	4	205.1	103.7	5.06	good	1, 3, 4, 5
			189	R.EIRHN[+2445.897]STGC[+57.021]LR.M	HexNAc(5)Hex(5)Fuc(1)Sialic474(1)	N291	yes	947.897	4	195.4	51.2	3.54	medium	2, 3, 4
			190	R.EIRHN[+2362.831]STGC[+57.021]LR.M	HexNAc(5)Hex(6)Fuc(1)Sialic229(1)	N291	yes	927.130	4	229.0	81.3	5.04	good	1, 2, 3, 5
			190	R.EIRHN[+2607.950]STGC[+57.021]LR.M	HexNAc(5)Hex(6)Fuc(1)Sialic474(1)	N291	yes	988.410	4	233.7	70.5	5.21	medium	1, 2, 3, 4
			191	R.EPQDTYHYLPFS[+1068.369]LPHR.R	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S210	yes	767.839	4	258.3	8.7	4.09	medium	1, 2, 4
			192	R.EPQDTYHYLPFS[+1313.488]LPHR.R	HexNAc(1)Hex(1)Sialic474(2)	S210	yes	829.120	4	332.7	38.1	6.45	high	4, 5
			193	R.EPQDTYHYLPFS[+594.191]LPHR.R	HexNAc(1)Hex(1)Sialic229(1)	S210	yes	649.294	4	328.6	18.3	6.72	good	4
			193	R.EPQDTYHYLPFS[+839.310]LPHR.R	HexNAc(1)Hex(1)Sialic474(1)	S210	yes	710.574	4	357.3	17.4	6.01	medium	1, 4, 5
			194	R.HN[+1835.646]STGC[+57.021]LR.M	HexNAc(4)Hex(4)Fuc(1)Sialic229(1)	N291	yes	927.366	3	163.2	17.5	2.82	medium	1, 2, 3
			195	R.HN[+1997.699]STGC[+57.021]LR.M	HexNAc(4)Hex(5)Fuc(1)Sialic229(1)	N291	yes	981.384	3	196.5	8.8	2.62	poor	3, 4, 5
			196	R.HN[+2096.760]STGC[+57.021]LR.M	HexNAc(4)Hex(5)Sialic474(1)	N291	yes	1014.404	3	231.8	41.0	4.13	medium	1, 2, 4
			197	R.HN[+2325.819]STGC[+57.021]LR.M	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N291	yes	1090.764	3	275.2	14.4	2.64	medium	4
			198	R.HN[+2362.831]STGC[+57.021]LR.M	HexNAc(5)Hex(6)Fuc(1)Sialic229(1)	N291	yes	1103.094	3	119.5	0.2	1.36	poor	5
44	CD99 antigen	P14209	199	P.DGGFDLSDALPDNENKKPT[+1068.369]AIPK.K	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T41	yes	878.403	4	455.4	110.0	8.52	high	1, 2, 4, 5
			200	P.DGGFDLSDALPDNENKKPT[+839.310]AIPK.K	HexNAc(1)Hex(1)Sialic474(1)	T41	yes	821.137	4	394.1	42.5	6.73	medium	1, 3, 5
45	Inter-alpha-trypsin inhibitor heavy chain H1	P19827	201	R.RTFVLSALQPS[+594.191]PTHSSSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	S651	no	727.851	4	559.7	15.3	11.34	medium	1, 2, 3, 4, 5
			201	R.RTFVLSALQPS[+839.310]PTHSSSNTQR.L	HexNAc(1)Hex(1)Sialic474(1)	S651	no	789.131	4	429.9	0.0	9.98	poor	3, 5
			202	R.RTFVLSALQPSPT[+594.191]HSSSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	T653	yes	727.850	4	591.0	0.6	12.53	poor	2
			202	R.RTFVLSALQPSPT[+839.310]HSSSNTQR.L	HexNAc(1)Hex(1)Sialic474(1)	T653	yes	789.131	4	443.3	16.4	9.94	medium	1, 2, 3, 4, 5
			203	R.TFVLS[+594.191]ALQPSPTHSSSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	S646	no	918.099	3	725.6	3.9	15.56	poor	1, 4
			203	R.TFVLS[+839.310]ALQPSPTHSSSNTQR.L	HexNAc(1)Hex(1)Sialic474(1)	S646	no	999.805	3	714.6	12.8	15.91	good	4
			204	R.TFVLSALQPSPT[+594.191]HSSSNTQR.L	HexNAc(1)Hex(1)Sialic229(1)	T653	yes	918.098	3	321.3	12.9	8.13	medium	1, 4
			204	R.TFVLSALQPSPT[+839.310]HSSSNTQR.L	HexNAc(1)Hex(1)Sialic474(1)	T653	yes	999.804	3	381.4	11.3	8.45	poor	1, 2, 3, 4, 5
			205	R.TFVLSALQPSPTHS[+839.310]SSNTQR.L	HexNAc(1)Hex(1)Sialic474(1)	S655	no	999.805	3	670.1	2.0	14.63	poor	2, 3
			206	R.TFVLSALQPSPTHSSS[+839.310]NTQR.L	HexNAc(1)Hex(1)Sialic474(1)	S657	no	999.804	3	601.2	2.0	11.96	poor	5
46	Nebulin	P20929	207	K.N[+3567.273]ATQILNEK.E	HexNAc(7)Hex(8)Fuc(1)Sialic474(1)Sialic229(1)	N6039	no	1149.961	4	338.3	26.9	5.73	medium	5
			208	R.N[+2727.963]ATDIASQIK.Y	HexNAc(6)Hex(7)Fuc(1)Sialic229(1)	N6001	no	947.896	4	209.3	2.3	1.17	poor	5
47	Insulin-like growth factor-binding protein 6	P24592	209	K.ESKPQAGT[+1068.369]ARPQDVNR.R	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T126	yes	706.321	4	333.1	104.7	8.67	high	1
			210	K.ESKPQAGT[+1068.369]ARPQDVNRR.D	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T126	yes	745.347	4	260.7	92.8	5.43	medium	1, 2, 3, 4, 5
			210	K.ESKPQAGT[+1313.488]ARPQDVNRR.D	HexNAc(1)Hex(1)Sialic474(2)	T126	yes	806.626	4	280.3	107.0	5.16	high	1, 3, 4
			211	K.ESKPQAGT[+594.191]ARPQDVNR.R	HexNAc(1)Hex(1)Sialic229(1)	T126	yes	587.777	4	202.2	40.7	2.83	high	4
			211	K.ESKPQAGT[+839.310]ARPQDVNR.R	HexNAc(1)Hex(1)Sialic474(1)	T126	yes	865.073	3	171.2	47.3	3.93	high	3, 4
			212	K.ESKPQAGT[+839.310]ARPQDVNRR.D	HexNAc(1)Hex(1)Sialic474(1)	T126	yes	688.082	4	329.2	167.6	8.32	high	1, 3, 4, 5
48	Zinc-alpha-2-glycoprotein	P25311	213	R.FGC[+57.021]EIENN[+2096.760]R.S	HexNAc(4)Hex(5)Sialic474(1)	N128	yes	1079.090	3	265.1	59.2	6.12	medium	4
49	Hematopoietic progenitor cell antigen CD34	P28906	214	K.ASVNRGAQEN[+3056.083]GTGQATSR.N	HexNAc(6)Hex (7)Sialic474(1)Sialic229(1)	N362	no	1215.491	4	262.3	16.8	6.17	medium	4
50	Pigment epithelium-derived factor	P36955	215	C.QNPASPPEEGSPDPDST[+594.191]GALVEEEDPFFKVPVNK.L	HexNAc(1)Hex(1)Sialic229(1)	T36	no	1406.638	3	582.0	0.0	9.18	poor	4
51	Prostaglandin-H2 D-isomerase	P41222	216	A.APEAQVS[+594.191]VQPNFQQDK.F	HexNAc(1)Hex(1)Sialic229(1)	S29	yes	794.028	3	405.0	173.3	8.59	good	4, 5
			216	A.APEAQVS[+839.310]VQPNFQQDK.F	HexNAc(1)Hex(1)Sialic474(1)	S29	yes	875.734	3	528.2	265.0	11.01	medium	1, 2, 3, 4, 5
			217	A.LSMC[+57.021]KSVVAPATDGGLN[+1810.614]LTSTFLR.K	HexNAc(3)Hex(6)Sialic229(1)	N78	yes	1087.989	4	134.4	54.1	1.15	medium	3
			218	C.KSVVAPATDGGLN[+2200.778]LTSTFLR.K	HexNAc(5)Hex(5)Fuc(1)Sialic229(1)	N78	yes	1062.729	4	289.7	21.6	4.81	good	1, 2, 4, 5
			218	C.KSVVAPATDGGLN[+2445.897]LTSTFLR.K	HexNAc(5)Hex(5)Fuc(1)Sialic474(1)	N78	yes	1498.342	3	209.7	12.7	2.50	medium	1
			219	C.KSVVAPATDGGLN[+2201.791]LTSTFLR.K	HexNAc(3)Hex(6)Fuc(1)Sialic474(1)	N78	yes	1416.640	3	207.4	0.9	1.87	poor	2
			220	C.KSVVAPATDGGLN[+2242.817]LTSTFLR.K	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)	N78	yes	1073.239	4	398.3	140.8	7.72	good	1, 2

			221	C.KSVVAPATDGGLN[+2471.877]LTSTFLR.K	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)Sialic229(1)	N78	yes	1130.503	4	352.0	20.1	4.72	medium	4
			222	K.SVVAPATDGGLN[+1689.588]LTSTFLR.K	HexNAc(4)Hex(4)Sialic229(1)	N78	yes	1203.547	3	310.8	151.8	6.74	medium	1
			222	K.SVVAPATDGGLN[+1934.707]LTSTFLR.K	HexNAc(4)Hex(4)Sialic474(1)	N78	yes	1285.249	3	325.3	194.2	7.80	medium	1, 2, 3, 4
			223	K.SVVAPATDGGLN[+1835.646]LTSTFLR.K	HexNAc(4)Hex(4)Fuc(1)Sialic229(1)	N78	yes	1252.226	3	439.2	276.3	10.59	medium	1, 2, 3, 4, 5
			223	K.SVVAPATDGGLN[+2080.765]LTSTFLR.K	HexNAc(4)Hex(4)Fuc(1)Sialic474(1)	N78	yes	1333.934	3	384.3	228.4	9.33	medium	1, 2, 3, 4
			224	K.SVVAPATDGGLN[+1851.641]LTSTFLR.K	HexNAc(4)Hex(5)Sialic229(1)	N78	yes	1257.560	3	467.0	326.5	11.04	good	1, 2, 4, 5
			224	K.SVVAPATDGGLN[+2096.760]LTSTFLR.K	HexNAc(4)Hex(5)Sialic474(1)	N78	yes	1339.265	3	423.8	49.6	10.55	medium	1, 2, 3, 4
			225	K.SVVAPATDGGLN[+1876.672]LTSTFLR.K	HexNAc(5)Hex(3)Fuc(1)Sialic229(1)	N78	yes	1265.903	3	470.3	402.6	11.67	medium	1, 3, 5
			225	K.SVVAPATDGGLN[+2121.791]LTSTFLR.K	HexNAc(5)Hex(3)Fuc(1)Sialic474(1)	N78	yes	1347.610	3	396.1	246.2	9.47	medium	1, 2, 3, 4, 5
			226	K.SVVAPATDGGLN[+1877.685]LTSTFLR.K	HexNAc(3)Hex(4)Fuc(1)Sialic474(1)	N78	yes	949.931	4	351.8	165.2	7.54	medium	2
			227	K.SVVAPATDGGLN[+1892.667]LTSTFLR.K	HexNAc(5)Hex(4)Sialic229(1)	N78	yes	1271.232	3	432.7	281.5	10.30	medium	1, 3, 4, 5
			227	K.SVVAPATDGGLN[+2137.786]LTSTFLR.K	HexNAc(5)Hex(4)Sialic474(1)	N78	yes	1015.209	4	419.1	206.4	8.73	good	1, 2, 4
			228	K.SVVAPATDGGLN[+1893.680]LTSTFLR.K	HexNAc(3)Hex(5)Sialic474(1)	N78	yes	1271.573	3	411.2	220.3	9.99	medium	1, 2
			229	K.SVVAPATDGGLN[+1956.672]LTSTFLR.K	HexNAc(3)Hex(6)Fuc(1)Sialic229(1)	N78	yes	969.685	4	246.3	45.7	4.51	medium	4
			229	K.SVVAPATDGGLN[+2201.791]LTSTFLR.K	HexNAc(3)Hex(6)Fuc(1)Sialic474(1)	N78	yes	1374.604	3	289.8	34.3	6.19	medium	1, 2
			230	K.SVVAPATDGGLN[+1997.699]LTSTFLR.K	HexNAc(4)Hex(5)Fuc(1)Sialic229(1)	N78	yes	1306.246	3	408.4	261.5	10.06	medium	1, 2, 3, 4, 5
			230	K.SVVAPATDGGLN[+2242.817]LTSTFLR.K	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)	N78	yes	1387.953	3	449.9	285.7	10.32	medium	1, 2, 3, 4, 5
			231	K.SVVAPATDGGLN[+1997.699]LTSTFLRK.N	HexNAc(4)Hex(5)Fuc(1)Sialic229(1)	N78	yes	1011.960	4	255.1	57.9	3.99	good	3, 5
			231	K.SVVAPATDGGLN[+2242.817]LTSTFLRK.N	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)	N78	yes	1073.241	4	323.0	127.0	6.88	medium	2, 5
			232	K.SVVAPATDGGLN[+2038.725]LTSTFLR.K	HexNAc(5)Hex(4)Fuc(1)Sialic229(1)	N78	yes	1319.921	3	455.8	350.7	10.62	medium	1, 2, 3, 4, 5
			232	K.SVVAPATDGGLN[+2283.844]LTSTFLR.K	HexNAc(5)Hex(4)Fuc(1)Sialic474(1)	N78	yes	1401.627	3	325.8	222.4	7.51	medium	1, 2, 3, 4, 5
			233	K.SVVAPATDGGLN[+2038.725]LTSTFLRK.N	HexNAc(5)Hex(4)Fuc(1)Sialic229(1)	N78	yes	1362.615	3	330.8	136.8	6.35	medium	2, 5
			233	K.SVVAPATDGGLN[+2283.844]LTSTFLRK.N	HexNAc(5)Hex(4)Fuc(1)Sialic474(1)	N78	yes	1083.496	4	183.9	28.8	3.48	medium	2, 5
			234	K.SVVAPATDGGLN[+2054.720]LTSTFLR.K	HexNAc(5)Hex(5)Sialic229(1)	N78	yes	994.192	4	490.3	421.3	12.47	high	3
			234	K.SVVAPATDGGLN[+2299.839]LTSTFLR.K	HexNAc(5)Hex(5)Sialic474(1)	N78	yes	1055.471	4	433.3	257.1	9.85	medium	1, 2, 3, 4, 5
			235	K.SVVAPATDGGLN[+2055.733]LTSTFLR.K	HexNAc(3)Hex(6)Sialic474(1)	N78	yes	1325.589	3	410.1	186.9	10.04	medium	1, 4
			236	K.SVVAPATDGGLN[+2055.733]LTSTFLRK.N	HexNAc(3)Hex(6)Sialic474(1)	N78	yes	1026.715	4	282.4	99.4	5.30	medium	1
			237	K.SVVAPATDGGLN[+2079.752]LTSTFLR.K	HexNAc(6)Hex(3)Fuc(1)Sialic229(1)	N78	yes	1333.592	3	299.3	88.7	5.78	medium	1, 3
			238	K.SVVAPATDGGLN[+2080.700]LTSTFLR.K	HexNAc(4)Hex(5)Sialic229(2)	N78	yes	1333.915	3	391.1	196.5	9.60	medium	3
			238	K.SVVAPATDGGLN[+2325.819]LTSTFLR.K	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N78	yes	1061.964	4	315.4	92.0	6.30	medium	2, 3, 4
			239	K.SVVAPATDGGLN[+2121.791]LTSTFLRK.N	HexNAc(5)Hex(3)Fuc(1)Sialic474(1)	N78	yes	1390.308	3	225.8	106.2	4.90	medium	2
			240	K.SVVAPATDGGLN[+2159.751]LTSTFLR.K	HexNAc(4)Hex(6)Fuc(1)Sialic229(1)	N78	yes	1360.590	3	235.4	13.4	3.02	medium	1
			240	K.SVVAPATDGGLN[+2404.870]LTSTFLR.K	HexNAc(4)Hex(6)Fuc(1)Sialic474(1)	N78	yes	1081.476	4	227.9	1.9	3.38	poor	1
			241	K.SVVAPATDGGLN[+2200.778]LTSTFLR.K	HexNAc(5)Hex(5)Fuc(1)Sialic229(1)	N78	yes	1030.457	4	495.0	206.1	10.51	medium	1, 2, 3, 4, 5
			241	K.SVVAPATDGGLN[+2445.897]LTSTFLR.K	HexNAc(5)Hex(5)Fuc(1)Sialic474(1)	N78	yes	1091.985	4	472.1	314.4	10.61	medium	1, 2, 3, 4, 5
			242	K.SVVAPATDGGLN[+2200.778]LTSTFLRK.N	HexNAc(5)Hex(5)Fuc(1)Sialic229(1)	N78	yes	1062.729	4	399.5	40.1	7.73	medium	5
			242	K.SVVAPATDGGLN[+2445.897]LTSTFLRK.N	HexNAc(5)Hex(5)Fuc(1)Sialic474(1)	N78	yes	1124.010	4	373.3	182.1	8.71	medium	1, 2, 3, 4, 5
			243	K.SVVAPATDGGLN[+2216.773]LTSTFLR.K	HexNAc(5)Hex(6)Sialic229(1)	N78	yes	1034.714	4	461.3	323.2	9.23	medium	1, 2, 3, 4, 5
			243	K.SVVAPATDGGLN[+2461.892]LTSTFLR.K	HexNAc(5)Hex(6)Sialic474(1)	N78	yes	1460.975	3	383.4	198.8	8.43	medium	1, 2, 3, 4, 5
			244	K.SVVAPATDGGLN[+2267.784]LTSTFLR.K	HexNAc(5)Hex(4)Fuc(1)Sialic229(2)	N78	yes	1047.459	4	374.3	143.0	8.97	medium	3
			244	K.SVVAPATDGGLN[+2512.903]LTSTFLR.K	HexNAc(5)Hex(4)Fuc(1)Sialic474(1)Sialic229(1)	N78	yes	1108.736	4	406.8	208.5	9.49	medium	1, 2, 3, 4, 5
			245	K.SVVAPATDGGLN[+2283.779]LTSTFLR.K	HexNAc(5)Hex(5)Sialic229(2)	N78	yes	1051.709	4	339.3	122.1	7.95	medium	1, 3, 5
			246	K.SVVAPATDGGLN[+2325.819]LTSTFLRK.N	HexNAc(4)Hex(5)Sialic474(1)Sialic229(1)	N78	yes	1093.979	4	282.3	57.0	4.37	medium	5
			247	K.SVVAPATDGGLN[+2362.831]LTSTFLR.K	HexNAc(5)Hex(6)Fuc(1)Sialic229(1)	N78	yes	1071.463	4	389.4	139.0	9.48	medium	3
			248	K.SVVAPATDGGLN[+2362.831]LTSTFLRK.N	HexNAc(5)Hex(6)Fuc(1)Sialic229(1)	N78	yes	1102.987	4	256.8	54.2	5.25	medium	2, 4
			248	K.SVVAPATDGGLN[+2607.950]LTSTFLRK.N	HexNAc(5)Hex(6)Fuc(1)Sialic474(1)	N78	yes	1164.272	4	322.9	37.7	7.43	medium	3
			249	K.SVVAPATDGGLN[+2226.758]LTSTFLR.K	HexNAc(4)Hex(5)Fuc(1)Sialic229(2)	N78	yes	1382.599	3	354.9	185.8	9.27	medium	1, 2, 3, 4, 5
			249	K.SVVAPATDGGLN[+2471.877]LTSTFLR.K	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)Sialic229(1)	N78	yes	1464.305	3	472.5	290.0	10.66	medium	1, 2, 3, 4, 5
			249	K.SVVAPATDGGLN[+2716.995]LTSTFLR.K	HexNAc(4)Hex(5)Fuc(1)Sialic474(2)	N78	yes	1159.761	4	375.4	176.1	10.25	medium	1, 3, 4
			250	K.SVVAPATDGGLN[+2226.758]LTSTFLRK.N	HexNAc(4)Hex(5)Fuc(1)Sialic229(2)	N78	yes	1069.227	4	342.5	79.9	5.85	high	2, 5
			250	K.SVVAPATDGGLN[+2471.877]LTSTFLRK.N	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)Sialic229(1)	N78	yes	1130.502	4	213.9	32.6	2.85	medium	1, 3, 4
			251	K.SVVAPATDGGLN[+2258.812]LTSTFLR.K	HexNAc(4)Hex(6)Sialic474(1)	N78	yes	1045.470	4	409.8	222.0	8.39	medium	1, 2, 3, 4, 5
			252	K.SVVAPATDGGLN[+2366.845]LTSTFLR.K	HexNAc(5)Hex(4)Sialic474(1)Sialic229(1)	N78	yes	1071.963	4	467.2	187.7	8.65	medium	5
			253	K.SVVAPATDGGLN[+2403.857]LTSTFLRK.N	HexNAc(6)Hex(5)Fuc(1)Sialic229(1)	N78	yes	1113.239	4	282.8	33.4	3.36	medium	1

			254	K.SVVAPATDGGLN[+2419.852]LTSTFLRK.N	HexNAc(6)Hex(6)Sialic229(1)	N78	yes	1117.242	4	302.0	49.0	6.10	medium	3
			255	K.SVVAPATDGGLN[+2429.837]LTSTFLR.K	HexNAc(5)Hex(5)Fuc(1)Sialic229(2)	N78	yes	1450.291	3	453.3	266.4	10.67	medium	1, 2, 3, 4
			255	K.SVVAPATDGGLN[+2674.956]LTSTFLR.K	HexNAc(5)Hex(5)Fuc(1)Sialic474(1)Sialic229(1)	N78	yes	1148.998	4	458.4	259.3	10.75	medium	1, 2, 3, 4
			255	K.SVVAPATDGGLN[+2920.075]LTSTFLR.K	HexNAc(5)Hex(5)Fuc(1)Sialic474(2)	N78	yes	1210.281	4	332.3	74.3	8.11	medium	1, 2, 3, 4
			256	K.SVVAPATDGGLN[+2429.837]LTSTFLRK.N	HexNAc(5)Hex(5)Fuc(1)Sialic229(2)	N78	yes	1119.993	4	296.4	38.2	4.74	medium	2
			257	K.SVVAPATDGGLN[+2445.832]LTSTFLR.K	HexNAc(5)Hex(6)Sialic229(2)	N78	yes	1455.968	3	371.2	115.8	7.98	medium	1, 3, 4, 5
			257	K.SVVAPATDGGLN[+2690.951]LTSTFLR.K	HexNAc(5)Hex(6)Sialic474(1)Sialic229(1)	N78	yes	1153.505	4	412.4	248.1	8.31	medium	1, 2
			257	K.SVVAPATDGGLN[+2936.070]LTSTFLR.K	HexNAc(5)Hex(6)Sialic474(2)	N78	yes	1214.783	4	330.9	95.9	7.09	medium	1, 5
			258	K.SVVAPATDGGLN[+2524.884]LTSTFLR.K	HexNAc(5)Hex(7)Fuc(1)Sialic229(1)	N78	yes	1111.724	4	336.3	52.9	6.03	medium	2, 3
			259	K.SVVAPATDGGLN[+2528.898]LTSTFLRK.N	HexNAc(5)Hex(5)Sialic474(1)Sialic229(1)	N78	yes	1144.501	4	298.4	49.4	4.28	medium	1, 5
			260	K.SVVAPATDGGLN[+2553.930]LTSTFLR.K	HexNAc(6)Hex(3)Fuc(1)Sialic474(1)Sialic229(1)	N78	yes	1118.998	4	278.6	24.4	1.79	medium	2
			261	K.SVVAPATDGGLN[+2837.009]LTSTFLRK.N	HexNAc(5)Hex(6)Fuc(1)Sialic474(1)Sialic229(1)	N78	yes	1221.535	4	281.0	41.4	5.82	medium	2
			262	L.N[+3447.260]LTSTFLR.K	HexNAc(6)Hex(7)Fuc(1)Sialic474(2)	N78	yes	1466.611	3	226.6	0.3	1.56	poor	1
			263	M.C[+57.021]KSVVAPATDGGLN[+1956.672]LTSTFLR.K	HexNAc(3)Hex(6)Fuc(1)Sialic229(1)	N78	yes	1388.271	3	124.3	3.6	1.71	poor	4
			264	M.C[+57.021]KSVVAPATDGGLN[+2200.778]LTSTFLR.K	HexNAc(5)Hex(5)Fuc(1)Sialic229(1)	N78	yes	1102.991	4	324.7	22.4	2.57	medium	2
			265	Q.AAPEAQVS[+1715.632]VQPNFQQDK.F	HexNAc(3)Hex(3)Fuc(1)Sialic474(1)	S29	yes	893.897	4	322.3	35.8	1.67	high	1
			266	V.APATDGGLN[+2137.786]LTSTFLR.K	HexNAc(5)Hex(4)Sialic474(1)	N78	yes	1257.897	3	273.1	66.6	4.16	medium	1
			267	V.APATDGGLN[+2362.831]LTSTFLR.K	HexNAc(5)Hex(6)Fuc(1)Sialic229(1)	N78	yes	1332.896	3	234.5	73.5	3.67	medium	1, 3, 5
			268	V.APATDGGLN[+2366.845]LTSTFLR.K	HexNAc(5)Hex(4)Sialic474(1)Sialic229(1)	N78	yes	1333.908	3	203.4	1.0	1.13	poor	3
			268	V.APATDGGLN[+2611.964]LTSTFLR.K	HexNAc(5)Hex(4)Sialic474(2)	N78	yes	1416.281	3	254.8	11.9	3.82	medium	1
			269	V.APATDGGLN[+2403.857]LTSTFLR.K	HexNAc(6)Hex(5)Fuc(1)Sialic229(1)	N78	yes	1346.241	3	263.2	79.2	5.75	medium	1
			270	V.APATDGGLN[+2404.870]LTSTFLR.K	HexNAc(4)Hex(6)Fuc(1)Sialic474(1)	N78	yes	1346.573	3	234.5	3.0	2.69	poor	4
			271	V.APATDGGLN[+2565.910]LTSTFLR.K	HexNAc(6)Hex(6)Fuc(1)Sialic229(1)	N78	yes	1400.260	3	218.8	17.0	2.02	medium	3, 5
			271	V.APATDGGLN[+2811.029]LTSTFLR.K	HexNAc(6)Hex(6)Fuc(1)Sialic474(1)	N78	yes	1482.630	3	164.0	3.1	1.78	poor	5
			272	V.APATDGGLN[+2570.938]LTSTFLR.K	HexNAc(4)Hex(5)Sialic474(2)	N78	yes	1402.612	3	243.5	45.1	2.58	medium	3
			273	V.APATDGGLN[+2581.905]LTSTFLR.K	HexNAc(6)Hex(7)Sialic229(1)	N78	yes	1054.949	4	225.9	48.0	3.29	medium	1, 5
			274	V.APATDGGLN[+2770.002]LTSTFLR.K	HexNAc(5)Hex(7)Fuc(1)Sialic474(1)	N78	yes	1101.725	4	222.3	19.3	2.31	medium	1
			275	V.VAPATDGGLN[+2055.733]LTSTFLR.K	HexNAc(3)Hex(6)Sialic474(1)	N78	yes	1263.563	3	239.7	60.0	3.66	medium	3, 5
			276	V.VAPATDGGLN[+2080.765]LTSTFLR.K	HexNAc(4)Hex(4)Fuc(1)Sialic474(1)	N78	yes	1271.904	3	201.3	127.7	3.67	good	4
			277	V.VAPATDGGLN[+2200.778]LTSTFLR.K	HexNAc(5)Hex(5)Fuc(1)Sialic229(1)	N78	yes	1311.915	3	188.1	19.8	1.73	medium	4
			278	V.VAPATDGGLN[+2201.791]LTSTFLR.K	HexNAc(3)Hex(6)Fuc(1)Sialic474(1)	N78	yes	1312.252	3	181.0	29.1	1.72	medium	1, 3
			279	V.VAPATDGGLN[+2242.817]LTSTFLR.K	HexNAc(4)Hex(5)Fuc(1)Sialic474(1)	N78	yes	1325.926	3	173.2	53.0	2.09	medium	5
			280	V.VAPATDGGLN[+2258.812]LTSTFLR.K	HexNAc(4)Hex(6)Sialic474(1)	N78	yes	1331.258	3	250.7	91.5	3.67	medium	5
			281	V.VAPATDGGLN[+2267.784]LTSTFLR.K	HexNAc(5)Hex(4)Fuc(1)Sialic229(2)	N78	yes	1334.248	3	171.4	11.1	1.74	medium	5
			282	V.VAPATDGGLN[+2420.865]LTSTFLR.K	HexNAc(4)Hex(7)Sialic474(1)	N78	yes	1385.615	3	205.4	98.7	1.11	medium	2
			283	V.VAPATDGGLN[+2429.837]LTSTFLR.K	HexNAc(5)Hex(5)Fuc(1)Sialic229(2)	N78	yes	1388.268	3	222.5	7.9	2.54	poor	1
			284	V.VAPATDGGLN[+2528.898]LTSTFLR.K	HexNAc(5)Hex(5)Sialic474(1)Sialic229(1)	N78	yes	1421.623	3	211.3	30.6	2.47	medium	1
			285	V.VAPATDGGLN[+2648.976]LTSTFLR.K	HexNAc(6)Hex(5)Fuc(1)Sialic474(1)	N78	yes	1461.640	3	155.3	11.4	2.66	medium	1
52	Arylsulfatase D	P51689	286	R.ALQWNAGSGGLPEN[+2096.760]ETTFAR.I	HexNAc(4)Hex(5)Sialic474(1)	N128	yes	1406.267	3	251.0	46.9	6.31	medium	1
53	Amyloid-like protein 1	P51693	287	R.GVEYVC[+57.021]C[+57.021]PPPGT[+1068.369]PDPSGTAVGDPSTR.S	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T215	yes	1247.857	3	338.7	8.7	8.31	medium	1, 4
			288	R.GVEYVC[+57.021]C[+57.021]PPPGT[+839.310]PDPSGTAVGDPSTR.S	HexNAc(1)Hex(1)Sialic474(1)	T215	yes	1171.506	3	403.8	38.0	8.17	medium	2, 3, 5
			289	R.GVEYVC[+57.021]C[+57.021]PPPGTPDPS[+839.310]GTAVGDPSTR.S	HexNAc(1)Hex(1)Sialic474(1)	S219	no	1171.506	3	380.7	4.8	6.41	poor	2
			289	R.GVEYVC[+57.021]C[+57.021]PPPGTPDPSGT[+1068.369]AVGDPSTR.S	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T221	yes	936.147	4	536.8	4.4	9.75	poor	3
54	Fractalkine	P78423	290	R.LGVLIT[+1068.369]PVPDAQAATR.R	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T329	yes	897.438	3	391.4	173.2	10.31	medium	1, 3, 4, 5
			290	R.LGVLIT[+1313.488]PVPDAQAATR.R	HexNAc(1)Hex(1)Sialic474(2)	T329	yes	979.144	3	301.1	80.2	5.88	medium	2, 3, 5
			290	R.LGVLIT[+823.251]PVPDAQAATR.R	HexNAc(1)Hex(1)Sialic229(2)	T329	yes	815.730	3	250.5	56.0	4.76	medium	1, 3, 4, 5
			291	R.LGVLIT[+839.310]PVPDAQAATR.R	HexNAc(1)Hex(1)Sialic474(1)	T329	yes	821.085	3	361.1	141.1	9.22	medium	1, 2, 3, 4, 5
			292	R.LGVLITPVPDAQAAT[+1068.369]R.R	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T338	no	897.438	3	605.4	29.9	12.99	medium	1, 3, 4, 5
			292	R.LGVLITPVPDAQAAT[+1313.488]R.R	HexNAc(1)Hex(1)Sialic474(2)	T338	no	979.144	3	367.7	6.0	8.08	poor	2, 3
			292	R.LGVLITPVPDAQAAT[+823.251]R.R	HexNAc(1)Hex(1)Sialic229(2)	T338	no	815.730	3	452.6	23.9	9.76	medium	1, 3, 4, 5

			293	R.LGVLITPVPDAQAAT[+839.310]R.R	HexNAc(1)Hex(1)Sialic474(1)	T338	no	821.085	3	522.8	40.6	11.80	medium	1, 2, 3, 4, 5
			294	R.LGVLITPVPDAQAAT[+839.310]RR.Q	HexNAc(1)Hex(1)Sialic474(1)	T338	no	873.117	3	396.9	16.3	8.47	medium	2
55	55 kDa erythrocyte membrane protein	Q00013	295	H.PLNTVTEDMYTN[+1934.707]GSPAPGSPAQVK.G	HexNAc(4)Hex(4)Sialic474(1)	N50	no	1469.969	3	273.9	48.3	5.41	medium	1
56	Dihydroorotate dehydrogenase (quinone), mitochondrial	Q02127	296	R.FTS[+1579.559]LGLLPR.A	HexNAc(2)Hex(2)Fuc(1)Sialic474(1)Sialic229(1)	S63	no	861.722	3	309.4	19.7	6.05	medium	3
57	HMG domain-containing protein 3	Q12766	297	K.AIEVSSPLPDVLN[+2038.725]ATEPLSTAQR.E	HexNAc(5)Hex(4)Fuc(1)Sialic229(1)	N918	no	1112.504	4	238.8	2.4	3.81	medium	1
			297	K.AIEVSSPLPDVLN[+2445.897]ATEPLSTAQR.E	HexNAc(5)Hex(5)Fuc(1)Sialic474(1)	N918	no	1214.043	4	251.3	63.5	4.93	medium	1
			298	K.AIEVSSPLPDVLN[+2267.784]ATEPLSTAQR.E	HexNAc(5)Hex(4)Fuc(1)Sialic229(2)	N918	no	1170.019	4	306.9	25.7	4.83	medium	1
58	Ectonucleotide pyrophosphatase/phosphodiesterase family member 2	Q13822	299	K.AIIAN[+1851.641]LTC[+57.021]K.K	HexNAc(4)Hex(5)Sialic229(1)	N411	yes	952.406	3	250.8	51.3	4.84	medium	1, 3
			299	K.AIIAN[+2096.760]LTC[+57.021]K.K	HexNAc(4)Hex(5)Sialic474(1)	N411	yes	1034.112	3	289.2	82.9	6.41	medium	1
59	SPARC-like protein 1	Q14515	300	K.VHENENIGT[+1068.369]TEPGEHQEAK.K	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T398	yes	797.590	4	649.1	1.4	12.28	poor	5
			301	K.VHENENIGT[+839.310]TEPGEHQEAK.K	HexNAc(1)Hex(1)Sialic474(1)	T398	yes	740.326	4	487.2	14.4	10.93	medium	2
			302	K.VHENENIGTT[+1068.369]EPGEHQEAK.K	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T399	no	797.590	4	290.4	22.7	7.69	medium	5
			303	K.VHENENIGTT[+839.310]EPGEHQEAK.K	HexNAc(1)Hex(1)Sialic474(1)	T399	no	740.325	4	552.1	29.4	11.80	medium	1, 2, 3, 4, 5
			304	R.EHANS[+1068.369]KQEEDNTQSDDILEESDQPTQVSK.M	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S231	no	1093.213	4	359.4	4.0	5.96	poor	1, 3
			305	R.EHANS[+839.310]KQEEDNTQSDDILEESDQPTQVSK.M	HexNAc(1)Hex(1)Sialic474(1)	S231	no	1035.948	4	369.5	3.6	5.06	poor	4
			306	R.EHANSKQEEDNTQS[+1068.369]DDILEESDQPTQVSK.M	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S240	no	1093.213	4	357.8	2.3	5.38	poor	5
			306	R.EHANSKQEEDNTQS[+1313.488]DDILEESDQPTQVSK.M	HexNAc(1)Hex(1)Sialic474(2)	S240	no	1154.494	4	289.8	3.0	3.98	poor	2
			307	R.EHANSKQEEDNTQS[+839.310]DDILEESDQPTQVSK.M	HexNAc(1)Hex(1)Sialic474(1)	S240	no	1035.948	4	269.7	2.8	3.28	medium	2, 5
			308	R.EHANSKQEEDNTQSDDILEES[+1068.369]DQPTQVSK.M	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S246	no	1093.213	4	494.4	24.2	8.70	medium	1, 3, 5
			308	R.EHANSKQEEDNTQSDDILEES[+1313.488]DQPTQVSK.M	HexNAc(1)Hex(1)Sialic474(2)	S247	no	1154.494	4	311.5	6.8	4.47	poor	1, 2, 4
			309	R.EHANSKQEEDNTQSDDILEES[+839.310]DQPTQVSK.M	HexNAc(1)Hex(1)Sialic474(1)	S246	no	1035.948	4	358.1	20.5	5.10	medium	1, 2, 4, 5
			310	R.EHANSKQEEDNTQSDDILEESDQPT[+1313.488]QVSK.M	HexNAc(1)Hex(1)Sialic474(2)	T251	no	1154.495	4	314.6	6.8	4.17	poor	1, 4, 5
			311	R.EKVHENENIGT[+839.310]TEPGEHQEAK.K	HexNAc(1)Hex(1)Sialic474(1)	T398	yes	804.610	4	420.6	0.0	7.06	poor	2, 4, 5
			312	R.EKVHENENIGTT[+839.310]EPGEHQEAK.K	HexNAc(1)Hex(1)Sialic474(1)	T399	no	804.610	4	471.8	21.0	10.30	medium	1
60	Inter-alpha-trypsin inhibitor heavy chain H4	Q14624	313	K.IPKPEAS[+839.310]FSPR.R	HexNAc(1)Hex(1)Sialic474(1)	S640	no	689.998	3	408.6	32.1	8.90	high	1, 2, 3, 4, 5
			314	K.IPKPEASFS[+839.310]PR.R	HexNAc(1)Hex(1)Sialic474(1)	S642	no	689.998	3	492.6	1.1	9.75	poor	1
			315	T.QTPAPIQAPS[+1516.567]AILPLPGQSVER.L	HexNAc(2)Hex(1)Sialic474(2)	S733	no	1262.949	3	321.9	3.8	4.16	medium	1
61	MHC class I polypeptide-related sequence A	Q29983	316	R.RTVPPMVN[+1934.707]VTR.S	HexNAc(4)Hex(4)Sialic474(1)	N210	yes	1068.810	3	292.6	91.6	5.38	medium	2
62	Shugoshin-like 1	Q5FBB7	317	R.LS[+594.191]LSPK.K	HexNAc(1)Hex(1)Sialic229(1)	S459	no	619.798	2	404.2	0.1	4.97	poor	5
63	Nuclear pore membrane glycoprotein 210-like	Q5VU65	318	Q.PQSIVHFSISN[+1876.672]QTVAVVNR.R	HexNAc(5)Hex(3) Fuc(1)Sialic229(1)	N1120	no	993.955	4	289.4	78.1	5.85	medium	1
64	39S ribosomal protein L21, mitochondrial	Q7Z2W9	319	R.LEKVLLVGADN[+2674.956]FTLLGK.P	HexNAc(5)Hex(5)Fuc(1)Sialic474(1)Sialic229(1)	N140	no	1127.013	4	307.7	67.7	5.82	medium	1
65	V-set domain-containing T-cell activation inhibitor 1	Q7Z7D3	320	K.TGAFSMPEVNVDYN[+1772.654]ASSETLR.C	HexNAc(4)Hex(3)Sialic474(1)	N160	no	1354.571	3	241.6	58.3	5.84	medium	1
66	V-set and transmembrane domain-containing protein 2A	Q8TAG5	321	R.THS[+839.310]TSSPQVVAK.I	HexNAc(1)Hex(1)Sialic474(1)	S194	no	694.325	3	444.5	0.0	9.12	poor	1
67	E3 ubiquitin-protein ligase RNF133	Q8WVZ7	322	R.N[+3526.247]NTASSWLMK.F	HexNAc(6)Hex(9)Fuc(1)Sialic474(1)Sialic229(1)	N11	no	1169.954	4	416.1	30.8	6.55	medium	5
68	Secretogranin-3	Q8WXD2	323	K.LFPAPS[+594.191]EK.S	HexNAc(1)Hex(1)Sialic229(1)	S359	yes	741.841	2	457.0	109.9	8.16	good	1, 3
			323	K.LFPAPS[+839.310]EK.S	HexNAc(1)Hex(1)Sialic474(1)	S359	yes	864.400	2	397.0	63.6	6.13	medium	1, 3, 5
69	Olfactory receptor 5P2	Q8WZ92	324	R.VILFMIILSGN[+2648.912]LSIIILIR.I	HexNAc(6)Hex(6)Sialic229(2)	N38	no	1198.070	4	240.7	15.4	5.22	medium	1
70	Protein FAM3C	Q92520	325	R.S[+1313.488]ALDTAARSTKPPR.Y	HexNAc(1)Hex(1)Sialic474(2)	S42	no	696.829	4	282.5	10.7	4.85	medium	2, 4, 5
			326	R.SALDTAARS[+1068.369]TKPPR.Y	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S50	no	635.549	4	262.0	2.4	4.31	poor	3, 5
			327	R.SALDTAARS[+839.310]TKPPR.Y	HexNAc(1)Hex(1)Sialic474(1)	S50	no	578.283	4	166.8	0.0	1.33	poor	2, 4
			328	R.SALDTAARST[+1068.369]KPPR.Y	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T51	no	635.549	4	350.2	33.8	7.80	medium	1, 2, 3, 4, 5
			328	R.SALDTAARST[+1313.488]KPPR.Y	HexNAc(1)Hex(1)Sialic474(2)	T51	no	696.829	4	207.7	30.3	3.46	medium	1, 2, 3, 4, 5
			329	R.SALDTAARST[+839.310]KPPR.Y	HexNAc(1)Hex(1)Sialic474(1)	T51	no	578.284	4	333.7	22.4	7.62	medium	1, 2, 3, 4, 5
71	Neuronal cell adhesion molecule	Q92823	330	R.FN[+2200.778]HTQTIQQK.Q	HexNAc(5)Hex(5)Fuc(1)Sialic229(1)	N223	yes	1149.152	3	299.4	79.4	4.29	medium	4
72	Zinc finger protein 354B	Q96LW1	331	K.LAPS[+1068.369]QR.N	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S34	no	870.380	2	377.7	3.5	5.88	poor	3
			332	K.LAPS[+839.310]QR.N	HexNAc(1)Hex(1)Sialic474(1)	S34	no	755.851	2	301.7	19.0	3.36	medium	3
73	Olfactory receptor 6C1	Q96RD1	333	R.N[+2404.870]FSILEISFTTVSIPK.F	HexNAc(4)Hex(6)Fuc(1)Sialic474(1)	N63	no	1050.970	4	290.3	40.5	5.76	medium	5
74	Seizure 6-like protein	Q9BYH1	334	R.DALPEGDAS[+594.191]PLGPYLLPSGAPER.G	HexNAc(1)Hex(1)Sialic229(1)	S40	no	972.789	3	319.3	31.9	6.54	medium	3, 5

			334	R.DALPEGDAS[+839.310]PLGPYLLPSGAPER.G	HexNAc(1)Hex(1)Sialic474(1)	S40	no	1054.497	3	280.9	28.1	3.49	medium	3
			335	R.DALPEGDASPLGPYLLPS[+594.191]GAPER.G	HexNAc(1)Hex(1)Sialic229(1)	S49	yes	972.788	3	484.3	9.8	9.08	good	3, 4, 5
			335	R.DALPEGDASPLGPYLLPS[+839.310]GAPER.G	HexNAc(1)Hex(1)Sialic474(1)	S49	yes	1054.497	3	545.1	12.5	11.43	medium	3
75	Augurin	Q9H1Z8	336	R.EAPVPT[+1313.488]KTK.V	HexNAc(1)Hex(1)Sialic474(2)	T47	no	762.020	3	544.9	9.6	8.62	medium	1, 2, 3
			337	R.EAPVPT[+839.310]KTK.V	HexNAc(1)Hex(1)Sialic474(1)	T47	no	603.961	3	464.2	2.9	7.71	poor	2, 5
			338	R.EAPVPTKT[+839.310]K.V	HexNAc(1)Hex(1)Sialic474(1)	T49	no	603.961	3	530.4	0.0	9.16	poor	1
76	Disco-interacting protein 2 homolog B	Q9P265	339	R.RT[+1944.692]ELTAATGER.H	HexNAc(3)Hex(3)Fuc(1)Sialic474(1)Sialic229(1)	T1444	no	1050.443	3	372.3	9.7	5.85	poor	5
77	DNA dC->dU-editing enzyme APOBEC-3B	Q9UH17	340	K.LAEFLSEHPN[+2607.950]VTLTISAAR.L	HexNAc(5)Hex(6)Fuc(1)Sialic474(1)	N113	no	1170.019	4	209.4	40.7	5.45	medium	1
78	ProSAAS	Q9UHG2	341	K.RLET[+1068.369]PAPQVPAR.R	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T247	yes	801.712	3	346.4	154.0	8.71	medium	1, 2, 3, 4, 5
			341	K.RLET[+1313.488]PAPQVPAR.R	HexNAc(1)Hex(1)Sialic474(2)	T247	yes	883.419	3	245.5	57.1	4.09	medium	1, 2, 4
			342	K.RLET[+594.191]PAPQVPAR.R	HexNAc(1)Hex(1)Sialic229(1)	T247	yes	643.653	3	287.4	110.0	5.42	medium	1, 3
			342	K.RLET[+839.310]PAPQVPAR.R	HexNAc(1)Hex(1)Sialic474(1)	T247	yes	725.360	3	199.7	19.7	2.47	good	4, 5
			343	R.GLS[+1068.369]AASPPLAETGAPR.R	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	S44	no	855.059	3	314.6	9.7	8.14	poor	3
			344	R.GLSAASPPLAET[+1068.369]GAPR.R	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T53	yes	855.060	3	627.3	19.3	11.99	medium	1, 2, 3, 4, 5
			345	R.GLSAASPPLAET[+839.310]GAPR.R	HexNAc(1)Hex(1)Sialic474(1)	T53	yes	778.706	3	463.7	9.0	9.67	good	1, 4
			346	R.LET[+1068.369]PAPQVPAR.R	HexNAc(1)Hex(1)Sialic474(1)Sialic229(1)	T247	yes	749.679	3	507.4	173.1	9.76	medium	1, 2, 3, 4, 5
			346	R.LET[+1313.488]PAPQVPAR.R	HexNAc(1)Hex(1)Sialic474(2)	T247	yes	831.385	3	385.5	102.2	7.17	medium	1, 2, 3, 4, 5
			347	R.LET[+839.310]PAPQVPAR.R	HexNAc(1)Hex(1)Sialic474(1)	T247	yes	673.326	3	503.7	248.0	10.36	good	1, 2, 3, 4, 5
			348	R.LET[+839.310]PAPQVPARR.L	HexNAc(1)Hex(1)Sialic474(1)	T247	yes	725.360	3	304.6	192.3	7.23	good	5
79	A disintegrin and metalloproteinase with thrombospondin motifs 1	Q9UHI8	349	R.LAT[+839.310]AAPGEKPPAPLQFHLLR.R	HexNAc(1)Hex(1)Sialic474(1)	T170	yes	742.385	4	547.0	430.5	12.50	good	1, 2, 3, 4
80	IgGFc-binding protein	Q9Y6R7	350	R.VVTVAALGTN[+2096.760]ISIHKDEIGK.V	HexNAc(4)Hex(5)Sialic474(1)	N1743	yes	1041.237	4	355.2	44.6	8.34	good	2, 5

Acknowledgments

After all these intense years of dedication to my PhD work, now it is time to reflect on people who have supported and helped me so much throughout this period.

Firstly, I would like to express my gratitude to Prof. Dr. Ulrika Westerlind for the opportunity to develop my scientific work, for the guidance, encouragement, and discussions that certainly had a great impact to form my scientific personality. Thank you for trusting in my abilities to learn and work in such a complex field. I am also thankful to Prof. Dr. Herbert Waldmann for providing an interdisciplinary environment for discussions in Chemistry and also for being part of the PhD committee. Further, I would also like to thank Prof. Dr. Albert Sickmann and the ISAS institute for the crucial environment and technologies to develop the work here described.

My research would have been impossible without the aid and support of Dr. René Zahedi who taught me so much in the mass spectrometry and proteomics fields, as well as the entire Protein Dynamics group, in particular, Dr. Verena Tinnerfeld for the help with the cross-linker experiments.

I am especially grateful to Dr. Jonas Nilsson for the collaboration on the new enrichment method for glycopeptides project and for sharing knowledge in the view of glycoproteomics. Marshall Bern gave me also an invaluable help regarding Byonic which was vital for the data analysis here described. My sincere thanks to Prof. Dr. Frode Berven and Dr. Astrid Guldbrandsen not only for the collaboration in the evaluation of our enrichment method but also for being such a kind host during my stay in Bergen. And here I would like to profoundly thank Dr. Marc Vaudel for arranging this collaboration and all scientific advice. In addition, I am thankful to EMBO for the financial support during my stay in the University of Bergen.

I also want to thank the scientific and technical staff of ISAS for measuring HR-MS, for the amino acid analysis, for all technical support on different laboratory equipment, for IT support; the technical service of the TU Dortmund for measuring NMR samples; and Dr. Petra Janning from Max-Plank Institut für Molekulare Physiologie for support on ESI-MS.

Further, I am genuinely thankful to my labmates Manuel, Christian, Jin, Sandra and Hui for being more than work colleagues, but friends, besides sharing knowledge and helping during all the years not only creating a nice atmosphere at work but also outside our working place. Thanks to all the other ISAS members for scientific and cultural exchange that goes beyond the working field.

Thanks a million to Gisele, Christopher, Iris and Oliver for all the critics and suggestions during the review of this thesis.

There are no words to describe how important the support of my family was, no matter the circumstances and how far from them I decided to go. If today I am here, it is due to all the love, freedom and assistance they gave me.

All the friends I made during this journey, independently where they came from, softened the years far from my family and with a challenging work in progress. Thank you to help to clear my mind. Your support and company were so important to keep me on track.

Finally, I would like to thank Christopher for his unconditional support and patience that transcend anything else. Thanks for bringing me peace whenever I needed the most.

Last but not the least, thank you, dear reader of this work.

Eidesstattliche Versicherung (Affidavit)

Caixeta Pereira, Vanessa
Name, Vorname
(Surname, first name)

154729
Matrikel-Nr.
(Enrolment number)

Belehrung:
Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden, § 63 Abs. 5 Hochschulgesetz NRW.

Die Abgabe einer falschen Versicherung an Eides statt ist strafbar.

Wer vorsätzlich eine falsche Versicherung an Eides statt abgibt, kann mit einer Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft werden, § 156 StGB. Die fahrlässige Abgabe einer falschen Versicherung an Eides statt kann mit einer Freiheitsstrafe bis zu einem Jahr oder Geldstrafe bestraft werden, § 161 StGB.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:
Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offence can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offences of this type is the chancellor of the TU Dortmund University. In the case of multiple or other serious attempts at deception, the candidate can also be unenrolled, Section 63, paragraph 5 of the Universities Act of North Rhine-Westphalia.

The submission of a false affidavit is punishable.

Any person who intentionally submits a false affidavit can be punished with a prison sentence of up to three years or a fine, Section 156 of the Criminal Code. The negligent submission of a false affidavit can be punished with a prison sentence of up to one year or a fine, Section 161 of the Criminal Code.

I have taken note of the above official notification.

Ort, Datum
(Place, date)

Unterschrift
(Signature)

Titel der Dissertation:
(Title of the thesis):

Chemical Tools for Protein Analysis: Development of an MS-cleavable Cross-linker and a New Approach for Glycopeptide Enrichment

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel selbstständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht.
Die Arbeit hat in gegenwärtiger oder in einer anderen Fassung weder der TU Dortmund noch einer anderen Hochschule im Zusammenhang mit einer staatlichen oder akademischen Prüfung vorgelegen.

I hereby swear that I have completed the present dissertation independently and without inadmissible external support. I have not used any sources or tools other than those indicated and have identified literal and analogous quotations.

The thesis in its current version or another version has not been presented to the TU Dortmund University or another university in connection with a state or academic examination.*

*Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the PhD thesis is the official and legally binding version.

Ort, Datum
(Place, date)

Unterschrift
(Signature)