Rapid and sensitive methods for the analysis and identification of O-glycans from glycoproteins

Kozak, R.P.

Citation

Kozak, R. P. (2017, January 24). *Rapid and sensitive methods for the analysis and identification of O-glycans from glycoproteins*. Retrieved from https://hdl.handle.net/1887/45434

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/45434

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/45434 holds various files of this Leiden University dissertation.

Author: Kozak, R.P.
Title: Rapid and sensitive methods for the analysis and identification of O-glycans from glycoproteins
Issue Date: 2017-01-24

Chapter 2

Protein O-glycosylation analysis

Gerhild Zauner[1], Radoslaw P. Kozak[2], Richard A. Gardner [2], Daryl L. Fernandes [2], André M. Deelder[1] and Manfred Wuhrer[1]

[1]Biomolecular Mass Spectrometry Unit, Department of Parasitology, Leiden University Medical Center, Postbus 9600, 2300 RC, Leiden, The Netherlands

[2]Ludger Ltd., Culham Science Centre, Oxfordshire, OX14 3EB, United Kingdom

Biological Chemistry (2012) 687-708.

Abstract

This review provides an overview on the methods available for analysis of O-glycosylation. Three major themes are addressed: (i) analysis of released O-glycans including different O-glycan liberation, derivatization and detection methods; (ii) analysis of formerly O-glycosylated peptides yielding information on O-glycan attachment sites; (iii) analysis of O-glycopeptides, representing by far the most informative but also most challenging approach for O-glycan analysis. While there are various techniques available for the identification of O-linked oligosaccharides, the focus here is on MS fragmentation techniques such as collision induced fragmentation (CID), electron capture dissociation (ECD) and electron transfer dissociation (ETD). Finally, (iv) the O-glycan analytical challenges that need to be met will be discussed.

Introduction

Protein glycosylation is a very heterogeneous group of post-translational modifications that has been found in essentially all living organisms, ranging from eubacteria to eukaryotes.[1]

Two major types of protein-linked oligosaccharides (N- and O-glycans) are differentiated on the basis of biosynthesis and linkage: while protein N-glycosylation is initiated by the *en bloc* transfer of a 14-monosaccharide building block onto asparagine amide groups of the nascent polypeptide forming an N-glycosidic linkage,[2] O-glycans are built up on the protein by the sequential enzymatic transfer of individual monosaccharides starting with the glycosylation of hydroxyl groups of the side chains of various amino acids including serine and threonine. Protein glycosylation is predominantly found on secreted proteins, the extracellular portions of membrane proteins, and the luminal side of organelles such as endoplasmatic reticulum, Golgi apparatus, and lysosomes.[2]

O-glycosylation in mammals represents a very diverse group of modifications that are often classified on the basis of the innermost monosaccharide. In mucin O-glycans this initiating monosaccharide is an alpha-linked GalNAc, and 20 human GalNAc transferases have been described that differ in polypeptide preference and tissue-specific expression.[3,4] Eight different core structures of GalNAc-linked O-glycans are distinguished dependent on the substitution of GalNAc with beta-linked Gal, beta-linked GlcNAc or alpha-linked GalNAc in the 3-position and/or 6-position (Figure 1).[5] Changes in mucin O-glycosylation appear to be a general phenomenon of malignancies[6,7] and the resulting mucin O-glycopeptide motifs of, for example,

MUC1 have been shown to be targeted by specific autoantibody responses in cancer which may be of diagnostic value.[8]

Another modification, O-mannosylation, has been observed on various proteins from brain and skeletal muscles, and defects in the involved transferases are associated with various congenital diseases.[9] Other O-glycans exhibit an innermost fucose, glucose, galactose, xylose or N-acetylglucosamine.[9-11] A very peculiar case of O-glycosylation is the so-called O-linked N-acetylglucosamine (O-GlcNAc). This modification, termed O-GlcNAcylation, is mainly found on intracellular, nucleocytosolic proteins and is involved in modulating cellular signalling pathways.[12] Even though these diverse types of O-glycosylation are very interesting, a closer description of the specific methodological aspects is out of scope for this review and in the following we will here focus on mucin-type O-glycosylation.

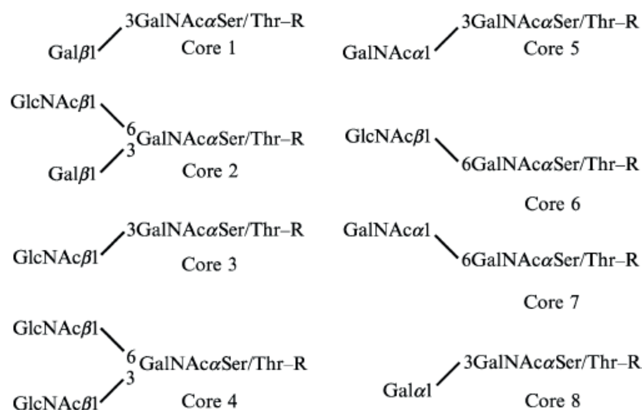

Figure 1. Eight types of O-glycan core structures. Taken from Mormann et al.[5] with permission.

A principle challenge in the field of O-glycosylation analysis still comprises finding the most suitable analytical technique that allows the analysis of the desired glycans and their attachment site(s). In general, there are three main ways to analyse glycosylation: (a) analysis and characterization of glycans on intact glycoproteins; (b) analysis and characterization of glycans covalently attached to (glyco)peptides; or (c) structural analysis of released glycans.

The intact glycoprotein, released glycans and generated glycopeptides can be analysed using various analytical platforms such as HPLC profiling, lectin affinity chromatography, mass spectrometric analysis, capillary electrophoresis[13-15] or nuclear magnetic resonance (NMR). Depending on the method used, there are also other steps required to obtain material suitable for analysis; for example in order to perform an HPLC profiling experiment, the glycans need to be released from the protein first and then modified with a fluorescent tag in order to allow optical detection.[13] More recently, Hülsmeier et al.[16] provided a good overview on N- and O-glycosylation analysis methods currently available within the field.[16] The Human Proteome Organisation Human Disease Glycomics/Proteome Initiative initiated an excellent multi-institutional study aiming to compare methodologies currently used for O-glycan analysis. A variety of mass spectrometric and chromatographic procedures representative of current approaches were used.[17] Two general strategies were found to give the most reliable data. These are direct MS analysis of mixtures of permethylated reduced O-glycans in the positive ion mode and analysis of native reduced glycans in the negative ion mode using LC-MS approaches. In addition, mass spectrometric methodologies to analyze O-glycopeptides were also successful. The ^{1}H NMR studies can provide the entire primary structure of a

released glycan, but the sample needs to be highly purified in large amounts (typically a milligram or more).

Here, we will review current analytical approaches for protein O-glycosylation analysis with a focus on recent methodological advances. Analysis of released glycans, formerly O-glycosylated peptides, and O-glycopeptides in general are covered. Finally, various innovative aspects as well as challenges in the field are discussed.

Analysis of released glycans

Most methods for protein glycosylation analysis rely on the release of the glycans from the glycoprotein.[18-20] Methods for the release, recovery and analysis of N-glycans are well established with the most common method in use being enzymatic release using N-glycosidase F or N-glycosidase A.[13] However, the release and recovery of O-glycans, particularly intact O-glycans, still remains very challenging. One major reason for this is that the enzymatic release of O-glycans is limited to endo-α-N-acetylgalactosamidase, an O-glycanase with a high specificity that only enables the removal of core 1 disaccharides (Galβ1-3GalNAc) from serine or threonine.[21] Therefore, the current best methods for universal removal of O-glycans from glycoproteins rely on chemical release. Over the last 20 years, several methods for the chemical liberation of O-glycans have been developed including techniques that use either reductive or non-reductive β-elimination.

Reductive β–elimination

Reductive β-elimination is one of the most popular techniques for obtaining released O-glycans. The most common conditions have been developed by Carlson et al.:[22] the 'Carlson reductive β-elimination method' typically involves the use of 0.05 M sodium hydroxide and 1.0 M sodium borohydride at a temperature of 45°C with an incubation time of 14-16 h followed by solid-phase extraction of released O-glycan alditols.[22-24] Under the alkaline sodium hydroxide conditions, the O-glycans are cleaved from the glycoprotein and the sodium borohydride immediately reduces the terminal monosaccharide of the released sugars to an alditol. Reduction of the terminal sugar by converting it to an alditol reduces or completely stops a side reaction known as 'peeling' which is a side reaction that may be defined as stepwise degradation of the polysaccharide starting at the reducing end and removing one sugar residue at a time.[25-30] Unfortunately, due to the reductive step employed to suppress peeling, the terminal sugar is no longer available for labelling with a fluorophore or chromophore. Therefore, with the preclusion of fluorescence or UV detection, the released O-glycans are predominantly analysed by mass spectrometry or HPAEC-PAD (high pH anion exchange chromatography with pulsed amperometric detection).[31,32] Despite this limitation, reductive β-elimination is still one of the most widely used methods applied by many researchers.[33-37]

Cooke et al.[38] used reductive β-elimination to release mucin-type O-glycans directly from biopsy tissue of rhesus monkeys. The released oligosaccharides were structurally characterized using positive ion mode MALDI -MS and nano-ESI MS for more sensitive detection of acidic oligosaccharides.

Robijn and co-workers used this method for the analysis of the glycosylation of a subset of a *Schistosoma mansoni* egg glycoproteins (SEA-4D12) bound by an anti-carbohydrate mAb 114-4D12 that is used in a diagnostic circulating egg antigen assay. The pool of released O-glycans was analysed by MALDI-TOF-MS.[39]

Analysis of underivatized glycans using negative ion electrospray mass spectrometry can be used for characterizing both negatively charged and neutral oligosaccharides.[33,36,40,41] Isomer separation (on graphitized carbon) and negative-mode ESI MS/MS characterization of underivatized complex O-glycan mixtures from human MUC5B that were released by reductive β-elimination has been described by Karlsson et al.[42] The authors characterized four separated structural isomers of released oligosaccharides with the composition dHex1Hex2HexNAc2HexNAcol1 (Figure 2).[42] Larsson et al.[43] applied reductive β-elimination to study the O-glycosylation of the colon MUC2 mucin from individual biopsies collected from the sigmoid colon during colonoscopy from 25 patients. More than 100 complex O-linked glycans (mostly mono-, di- and trisialylated) were identified and analyzed by graphitized carbon nanoLC-MS/MS.[43]

Lloyd and co-workers[32] used reductive β-elimination to compare the O-linked carbohydrate chains from MUC-1 immunoprecipitated from [3H]GlcN-labelled breast epithelial cell lines (MMSV1–1, MTSV1–7, and HB-2) derived from cells cultured from human milk, with the material obtained from three breast cancer cell lines (MCF-7, BT-20, and T47D) and analysed them by high pH anion chromatography. The data showed that the primary cells had a higher ratio of GlcN/GalN and more complex oligosaccharide profiles than the cancer cell lines.[32]

A significant increase in the sensitivity in the analysis of released O-glycans by MALDI-MS can be achieved by permethylation of the free hydroxyl groups.[34,37,44-46] Permethylating released O-glycans reduces their polarity, increases the stability of glycans (both sialylated and neutral permethylated glycans) for analysis by MS and may provide linkage information upon MS fragmentation or GC-MS analysis of the monosaccharide constituents, for example as partially methylated alditol acetates.[35,47-52]

Hanisch et al.[51] applied permethylation for structural analysis of neutral O-linked glycans from human skim milk mucins. O-glycans were firstly released by reductive β-elimination, then permethylated and subjected to FAB-MS and EI-MS analysis. The major observed structures were Galβ(1-3)[GlcNAcβ(1-6)]GalNAc core 2 type structures.

Thompson and co-workers[41] applied β-elimination to release and characterize O-glycans of salivary mucin. First the oligosaccharide mixture was analysed by NMR, then the oligosaccharides were fractionated into neutral, sialylated and sulphated pools. Neutral and sialylated species were permethylated and analyzed by MALDI-TOF-MS and GC-MS.[41]

Bleckmann et al.[53] applied the reductive β-elimination technique for O-glycosylation analysis of the CD24 glycoprotein from mouse brain. Following the release, the pool of O-glycans was permethylated and successfully analyzed by MALDI-TOF/TOF-MS and ESI-IT-MS/MS (Figure 3). The results indicated the presence of mucin type glycans as well as O-mannosyl glycans which could be distinguished by tandem mass spectrometry.

Faid et al.[34] also used the reductive elimination and permethylation method for the characterization of O-glycan chains in the diagnosis of defects in glycan biosynthesis. O-glycans were released from the glycoproteins of the serum from a patient suffering from CDG type IIa and analyzed by MALDI-TOF-MS and compared with O-glycans released from the glycoproteins of the control serum. No significant difference between O-glycans released from the control serum and from the serum from the patients suffering from CGG type IIa was observed. The major observed O-glycans were: Neu5Ac1Hex1HexNAcol1 and Neu5Ac2Hex1HexNAcol1 consistent with sialylated and disialylated T-antigen structures.

Parry et al.[54] applied this technique for characterization of the O-glycans of CD52 by MALDI-TOF-MS. Presented data indicated that the major O-glycans are core 1 and core 2 mucin type structures with and without sialic acid.

As an alternative to the commonly used sodium hydroxide, dimethylamine has been applied for O-glycan removal by Maniatis et al.[55] The protocol describes the use of dimethylamine in combination with microwave radiation for rapid de-O-glycosylation. This technique was evaluated on bovine fetuin and a porcine stomach mucin. The glycoproteins were incubated with 40% dimethylamine for 70 min and the released and reduced glycans were permethylated and analyzed by MALDI-TOF-MS. The results showed that this procedure 'improved' the yields of O-glycans from both glycoproteins when compared to the 'classic' Carlson reductive β-elimination method.[22,23] A point of criticism in this paper might be the choice of the standard compound (O-GlcNAc-modified peptide) used for method development.[55] O-GlcNAc is known to be much more alkali-labile than O-GalNAc and, therefore, the findings may not directly be transferable to the analysis of mucin-type O-glycosylation.[56]

Although a popular technique, as mentioned before, the analysis of O-glycans released by reductive β-elimination method is largely restricted to mass spectrometric detection.

Non-reductive β–elimination

Many researchers have looked at developing non-reductive release methods that can produce O-glycans with free reducing termini allowing further derivatization by either a chromophore or a fluorophore for analysis of glycans by HPLC with UV or fluorescence detection.[17] Over the years a number of different reagents for non-reductive release have been developed.

Chai et al.[57] reported the use of 70% (w/v) aqueous ethylamine and Yamada et al.[58] reported the use of lithium hydroxide in an automatic set-up but the overall reaction yields for both methods were low and the amount of peeling products was relatively high.

Huang et al.[59] developed an ammonia based β-elimination procedure. The authors applied a combination of aqueous ammonium hydroxide (for β-elimination) and ammonium carbonate (for converting reducing glycans to glycosylamines as well as glycosylamine carbonates) to prevent O-glycans from 'peeling'. This method is very appealing as both reagents are volatile and easily removed upon evaporation, thus eliminating the need for an elaborative desalting step. The authors reported that the yields of released glycans were quantitative and that peeling was completely prevented. To test this technique the authors used a maltoheptaose standard as a model for both 1-4-linked glycan cores and fetuin N-glycans with a GlcNAcβ1-4GlcNAc core, next to fetuin O-glycans. Released O-glycans were analyzed by

MALDI-MS. No signals originating from the peeling reaction were observable, however the low molecular structures were found to be difficult to detect because they overlap with peaks that are attributed to matrix.[59] A major weakness of this paper might be that inappropriate compounds (β1-4 linked glucose oligomers instead of β1-3 linked glycans, which are much more prone to degradation by peeling) have been used to establish the method. This issue was addressed by Yu et al.[60] who looked into the stability of released O-glycans under non-reducing ammonia based β-elimination conditions demonstrating that the β1-3-linkage is labile whereas the β1-4-linkage is relatively stable under these conditions.[60] These data demonstrate that a complete prevention of peeling under alkali-catalyzed hydrolysis conditions remains difficult. Investigations indicate that great care has to be taken when employing non-reductive alkaline conditions in glycomic analysis and in obtaining glycoprotein glycans for functional studies.[60]

Further modifications to the ammonia based β-elimination procedure were made by Miura et al.[61] They demonstrated that employing ammonium carbamate under the non-reductive conditions can prevent the unfavourable 'peeling' side reaction[61]. The samples were incubated for 20 h at 60°C with ammonium carbamate and then the solution was subjected to glycoblotting with subsequent methylation of sialic acids and tagging of the oligosaccharides by trans-iminization with O-benzylhydroxylamine hydrochloride. The tagged, methyl esterified O-glycans were eluted with water and analyzed by MALDI-TOF-MS.

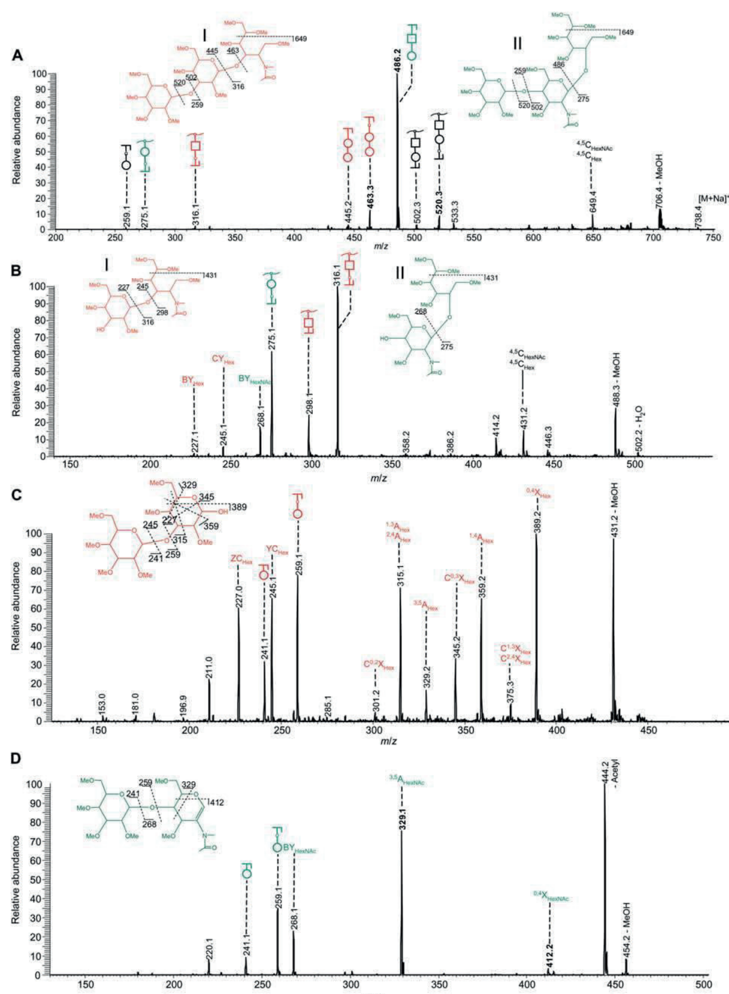

Figure 2. ESI-IT-MSn fragmentation spectra of the permethylated, reduced glycan species Hex2HexNAc1. (A) MS2-spectrum of the sodiated pseudomolecular ion $[M+Na]^+$ at m/z 738.4. Ions selected for further fragmentation are given in bold. (B) MS3-spectrum of the Hex-HexNAcol Y ion at m/z 520.2. (C) MS3-spectrum of the Hex-Hex C ion at m/z 463.3. (D) MS3-spectrum of the Hex-HexNAc B-ion at m/z 486.2. Diagnostically relevant fragment ions are illustrated in insets. Possible isomeric structures are given in different colors and key product ions derived thereof are denoted in the respective color. Non-specific product ions are labelled in black. All fragments observed are sodium adducts. Fragments were assigned according to Domon and Costello.[62] They may be achieved by different fragmentation pathways, only one of which is illustrated in the figure. Open circle, hexose; open square, N-acetylhexosamine. Taken from Bleckmann et al.[53] with permission.

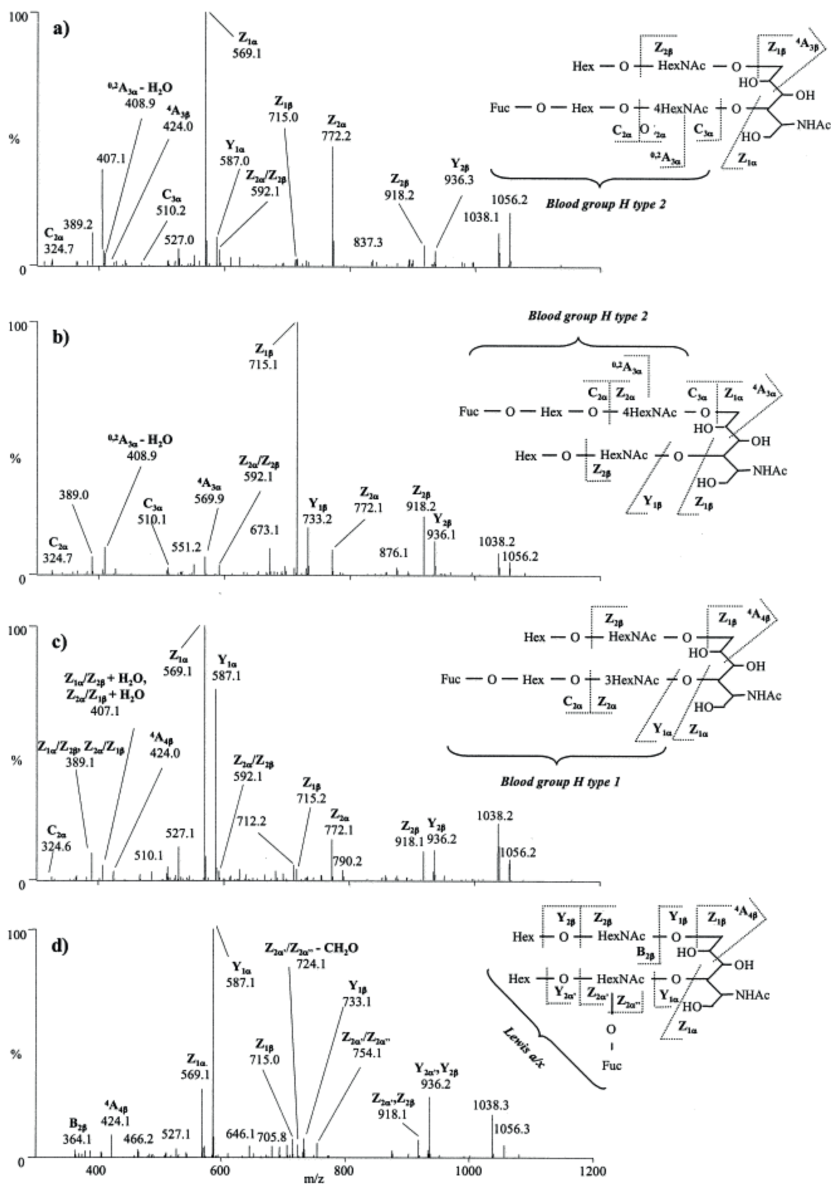

Figure 3. LC-MS^2 spectra and assigned structures of isomeric Core 4 O-linked oligosaccharides with $[M-H]^-$ ion of m/z 1098 corresponding to dHex1Hex2HexNAc2HexNAcol1 isolated from human MUC5B. Isomers were observed at retention time 19.1 min (a), 18.1 min (b), 16.2 min (c), and 14.9 min (d). Taken from Karlsson et al.[42] with permission.

Permethylation of released O-glycans for analysis by mass spectrometry is a technique used by many researchers. Goetz et al.[63] realised that the alkaline conditions used for permethylation were similar to the conditions used for the release of O-glycans and therefore developed a combined method for the release and permethylation of O-glycans. This method is a combination of an enzymatic digestion and a chemical O-glycan release using a solid-phase permethylation procedure. The protocol consists of incubation of the samples for 48 hours with pronase followed directly by release and permethylation using a spin column method.[64,65] As the glycans are released, in their free, non-reduced form by sodium hydroxide, methyl iodide is used as the methylating reagent to immediately convert them to their permethylated derivatives, thereby preventing the 'peeling' side reaction and stopping the oligosaccharides from degrading. In this study the O-glycans of bovine serum fetuin, human IgA and BSSL were successfully analyzed by MALDI-TOF-MS.[63] Although the technique provides good release of O-glycans, the detection of lower molecular weight O-glycans by MALDI-TOF-MS proves difficult as these masses overlap with peaks in the lower mass region that are attributed to matrix.

More recently, Zauner et al.[66] (2011), Wang et al.[67] and Furukawa et al.[68] have introduced new, one-pot combined methods, for O-glycan release and labelling. These methods use dimethylamine,[66] ammonia[67] or sodium hydroxide[68] for β-elimination release and 1-phenyl-3-methyl-5-pyrazolone (PMP) for the concomitant labelling via a Michael addition. The method was applied to analyse O-glycans from bovine submaxillary gland mucin and bile salt-stimulated lipase (BSSL) by reverse phase nanoLC-ESI-MS/MS. The data shows that the technique is quick and provides good release and labelling with an acceptable amount of peeling products.

Another very interesting study was performed in which chromatographic/tagging methodologies used by three laboratories for the analysis of O-glycans released under non-reductive conditions from IgA1 were found to give a diverse body of data. Clearly, peeling of the non-reductive alkaline released O-glycans from glycopeptides exacerbated the frequent occurrence of artifactual chromatographic peaks, severely influencing the reliability of these methodologies for quantitative O-glycomics (Wada et al., 2010a).[17]

Hydrazinolysis

Hydrazinolysis is a method for producing O-glycans in a non-reduced form in high yield,[69,70] forming an attractive alternative to the non-reducing β-elimination procedures described above. This method is very versatile and can be applied to release N- and/or O-glycans depending on the reaction conditions. The removal of N-glycans from glycoproteins by hydrazinolysis typically requires a temperature of 85°C to 100°C and an incubation time of 5-16 h.[70] For the removal of O-glycans a lower temperature is required (60°C, 6 h).[69] The mechanism by which hydrazine removes O-glycans, is not fully understood. It is generally accepted that it proceeds via an initial β-elimination reaction followed by reaction with hydrazine to form hydrazone derivatives. Under the release conditions, the N-acetyl groups of the amino sugars are also removed, therefore it is necessary to perform a re-N-acetylation step with acetic anhydride. As with all the non-reductive β-elimination techniques, the major concern of this method is the occurrence of undesirable peeling. A number of researchers have investigated how to reduce peeling under the hydrazinolysis conditions. In order to try and prevent this degradation, Cooper et al. formed hydrazone intermediates using 50% aqueous hydrazine but the results show

that the degradative peeling had not been significantly reduced.[24] More recent studies have shown that the degree of peeling can be greatly reduced by removal of water from the sample before hydrazinolysis[69] or by buffer exchanging the sample prior to hydrazinolysis with a solution of 0.1% trifluoroacetic acid (TFA) or a low molarity solution of ethylenediaminetetraacetic acid (EDTA).[71] Further experiments showed that samples that were contaminated with $CaCl_2$ and subsequently washed with a solution of EDTA showed a lower amount of peeling and an increased yield of released O-glycans. This data demonstrates the effect of salt removal on the suppression of peeling.

Determination of glycan structures from analysing MS, HPLC or NMR data is one of the major challenges in glycomics studies. Several different types of software tools have been developed for glycomics including GlycoWorkbench,[72] GlycoPeakfinder[73] and GlycoBase.[74] GlycoWorkbench is a tool which offers a graphical interface comprising a tool for building glycan schemes, an exhaustive collection of fragmentation types and a broad list of annotation options. The software is publicly available to download from: http://www.glycoworkbench.org/. GlycoPeakfinder is a web based software system allowing determination of the composition of glycan-derived MS signals independently from the source of spectral data and the fragmentation technique. GlycoBase is a database with MS and HILIC data for over four hundred 2-AB labelled N-linked glycans and seventy five 2-AB labelled O-glycan structures. A number of databases for analysing NMR data are also currently available[75,76] but data analysis still is demanding.

Analysis of formerly O-glycosylated peptides

In most cases the structures of the oligosaccharides are analyzed after releasing O-glycans from peptides by β-elimination. However, there is information that can be obtained from the peptide portion after β-elimination which should not be neglected: whilst no information about the glycan heterogeneity at the different sites can be obtained, the glycosylation sites may be determined.[77]

Initial experiments were performed by Rademaker et al.[78] employing reductive β-elimination with NaOH and $NaBD_4$. This resulted in the released, deutero-labelled O-glycan alditols as well as peptides with a deutero-labelled former O-glycosylation site: for example, in the case of an O-glycosylated threonine, β-elimination of the O-glycan converts the threonine residue into dehydroamino-2-butyric acid which is the reduced to 2-aminobutyric acid deuterated in the C2-position. O-glycosylation sites were elucidated by FAB-MS analysis of the mass-tagged, formerly O-glycosylated peptides providing y- and b-ion series.[78]

A difficulty associated with the β-elimination release of O-glycans from peptides is that the reductant used to prevent peeling of the released O-glycans may degrade the peptide backbone of certain peptides.[79] This was addressed by introducing a β-elimination/Michael addition protocol in which the former site of O-glycan attachment is marked by a tag allowing a mass spectrometric identification of the site of glycosylation.[79] The performance of sodium hydroxide ammonium hydroxide as releasing agents was compared: it was shown that the use of NH_4OH greatly enhances the sensitivity of the β-elimination procedure, as well as obviating the need for sample clean-up. Importantly, the addition of ammonia to the double

bond generated upon glycan elimination was observed, indicating that ammonia had a double function as a basic β-elimination reagent and tagging reagent. The resulting amine-tagged serine and threonine derivatives facilitated the identification of formerly O-glycosylated residues by ESI-quadrupole-CID-MS/MS.[79]

Other researchers have tested other basic chemicals as β-elimination reagents: Hanisch et al.[80] and Mirgorodskaya et al.[81] reported on β-elimination using methylamine as the releasing agent and at the same time as the reagent to modify the formerly glycosylated serine or threonine residue within the peptide prior to MALDI-TOF-MS or ESI-ion trap-MS mapping of the sites.[80,81] Zheng et al.[82] compared the use of various amines (dimethylamine, methylamine and ammonia solution) for the release of O-glycans from glycopeptides via β-elimination.[82] The highest reaction yields were observed when dimethylamine was used to convert the Thr residue of the glycopeptides into its β–elimination/addition derivative with a mass shift of −18 Da or +27Da (addition of dimethylamine) compared to the non-glycosylated Thr. Although the authors successfully identified glyosylation sites, peptide analysis did not give information on the glycan structures.[82]

In two earlier studies removal of O-GlcNAc from O-glycopeptides was described by mild β-elimination followed by Michael addition of dithiothreitol (DTT), deuterated DTT or biotin pentylamine (BAP) to tag O-GlcNAc sites. After chromatographic enrichment of the modified peptides, the former O-GlcNAc sites (as well as phosphorylation sites) could be identified by LC-MS/MS.[83,84]

In a more recent paper, the addition of other reagents during β-elimination, such as 1-phenyl-3-methyl-5-pyrazolone (PMP) in combination with DTT, was tested. It was found that not only were the O-glycans labelled but the formerly glycosylated

peptide portions were also tagged after this reaction, as could be shown by MALDI-TOF-MS. Therefore, after one reaction both the released/labelled O-glycans and the remaining tagged peptide portion can be analysed. This is another demonstration of the feasibility of being able to use the peptide portion for site specific analysis after β-elimination.[68]

Czeszak et al.[85] were likewise interested in analyzing O-glycosylation sites and aimed for a more potent way to chemically modify peptides as they were deglycosylated. They used β-elimination/addition reaction conditions, consisting of dimethylamine and ethanethiol, on formerly glycosylated sites to leave efficiently deglycosylated peptides which were chemically modified to mark the O-glycan attachment sites. After N-terminal derivatization by a phosphonium group, peptide sequencing was carried out by nano-ESI tandem mass spectrometry. [77,85] Depending on the research question, the analysis of deglycosylated peptides can be of value to determine the sites of O-glycosylation, even though the glycan structure and heterogeneity information is lost. [77]

Analysis of O-glycopeptides

By far the most efficient, but obviously analytically challenging, approach to analyse the O-glycan structures plus the attachment sites to the protein is performing O-glycopeptide analysis. As potential O-glycosylation sites are not indicated by specific consensus motifs, as is the case with N-glycosylation, direct analysis of the glycopeptides is difficult even if using newly developed and improved techniques.[86-88] We will describe the major mass spectrometry techniques currently used for the characterization of O-glycopeptides including CID fragmentation and ECD and ETD fragmentation. Generation of O-glycopeptides by proteolytic cleavage procedures employing specific proteinases and unspecific proteinases will be addressed. The study of synthetic model O-glycopeptides will be presented and examples of naturally occurring O-glycopeptides will be given.

CID fragmentation of glycopeptides

High energy CID by FAB-MS has been established for characterizing glycopeptides containing O-linked mono- and disaccharides. Via peptide backbone cleavages both the peptide sequence and the site of attachment of the sugar moiety were obtained from a single high-energy CID spectrum.[89-91]

With the advent of ESI and MALDI ionization, glycopeptide analysis by FAB-MS became gradually replaced by these soft ionization modes: the fragmentation of electrospray-ionized glycopeptides may be performed by CID using a variety of instruments and experimental setups. ESI-ion trap-CID tandem mass spectrometry (MS/MS) of glycopeptides results mainly in fragments arising from glycosidic bond cleavage allowing the determination of the glycan structure with the peptide moieties

staying largely intact.[92-94] When multiple sugar units are attached to a peptide, ion series (appearing often in different charge states) are measured due to sequential carbohydrate losses.[92] Under ion trap-CID conditions O-linked glycopeptides often undergo gas-phase deglycosylation that eliminates the carbohydrate completely, leaving the unmodified, and thus, ‘‘unlabelled’’ Ser or Thr residue(s) behind.[94,95] In the case of ion trap-CID, if the detection of peptide backbone fragments is desired, there is in many cases the necessity to perform MS3 experiments.[94,96] In Figure 4 an example of a Proteinase K generated O-glycopeptide identification from asialofetuin is demonstrated: the upper spectrum (Figure 4.A.) shows the precursor ion (m/z 1120.5, [M+2H]2+; inset), and below the MS2 spectrum is shown. As only a few b- and y-ions could be identified in the MS2, an MS3 was necessary to identify the peptide moiety 291VVVGPSVV298 (m/z 755.4; [M+H]+; Figure 4.B.).[94]

While most of the ESI-CID-MS/MS analyses of glycopeptides are performed with positive-mode ionization, Deguchi et al.[97] also evaluated negative-mode ion trap-CID. In these spectra, they observed C-ions comprising the entire O-glycan chain, representing a gas-phase release of the reducing-end O-glycan. Comparison of the negative-mode fragmentation pattern (MS3) of the gas phase-released O-glycan with the MS2 fragmentation of infused oligosaccharide standards enables the use of spectral matching for the structural assignment of O-glycan structures.[97]

Next to ESI-IT-MS/MS, ESI-quadrupole-TOF MS/MS has been widely used for CID fragment ion analysis of O-glycopeptides. ESI-quadrupole-CID fragmentation spectra of O-glycopeptides may allow the deduction of the O-glycosylation site(s) based on y-type and/or b-type peptide ions comprising the glycan attachment site. This is only true if the glycan moiety is not completely ‘lost’ during the fragmentation

process. Hence, optimization of the fragmentation energy is critical in ESI-quadrupole-CID analysis of O-glycopeptides,[98] similar to the situation in ESI-quadrupole-CID of N-glycopeptides where low fragmentation energies result in glycan fragmentation whilst high energies mainly provide peptide sequence ions lacking the glycan modification.[99]

As an example, nano-ESI low energy quadrupole CID was used in conjunction with other methods to analyze N- as well as O-glycopeptides derived from the M surface protein of human hepatitis B virus. The O-glycosylation site within the O-glycopeptide could be identified due to the excellent coverage of the (glycosylated) sequence ions that was obtained for this particularly large peptide (29 amino acids long).[100]

Alving et al.[101] performed a comparative study by ESI-MS/MS and postsource decay (PSD)-MALDI-MS to analyse O-glycopeptides, whereas another study focused exclusively on ESI-QTOF-MS analysis.[98] More recently, MALDI low energy CID MS/MS experiments on a QTOF instrument[102] have been found suitable for assigning O-glycans to specific O-glycosylation sites.

ECD/ ETD fragmentation of glycopeptides

MS methods using ECD and ETD fragmentation of O-glycopeptides simplify the identification and site assignment of fragmentation-prone side-chain modifications tremendously.

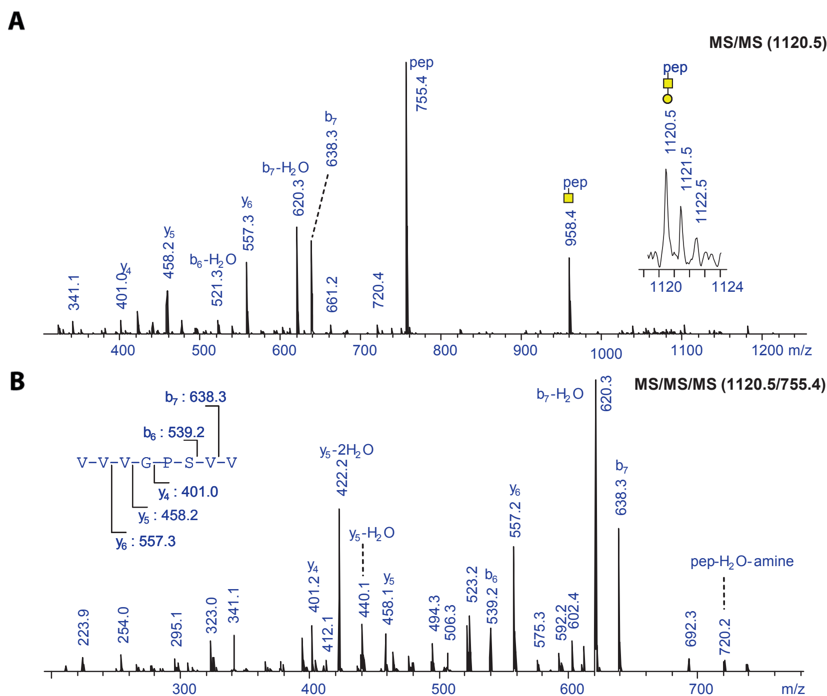

Figure 4. Fragment ion spectra of Proteinase K-generated O-glycopeptides from asialofetuin. A fragment ion spectrum (A; MS2) of the protonated precursor (inset) and a corresponding MS3 spectrum (B) are shown. MS2 was performed on glycopeptide V291–V298 carrying a Hex1HexNAc1 O-glycan ($[M+2H]^{2+}$ at m/z 1120.5). The fragment ion at m/z 755.4 ($[M+H]^+$) was subjected to MS3. Yellow square, N-acetylgalactosamine; yellow circle, galactose. Taken from Zauner et al.[94] with permission.

In ECD, reduced radical cations $[M+nH]^{(n-1)+}$ are generated by irradiation of multiply charged analyte ions, $[M+nH]^{n+}$, produced by electrospray ionization with low-energy electrons (T≤0.2 eV).[103-106] These odd-electron species mainly dissociate by fast and facile fragmentation of the N–Cα bonds of the peptide chain giving rise to c-type and z-type ions. This process has been characterized as ‘non-ergodic’, i.e. bond cleavage occurs prior to randomization of the excess energy gained in the charge-recombination event. As a consequence of the non-ergodic nature of the ECD process, the elimination of side chain modifications, such as glycans from the

peptide backbone, cannot compete with the backbone cleavages within the radical cations.[105] Therefore, ECD can be used for an unambiguous assignment of the attachment sites of glycans within the peptide chain.[105,107] In several studies, ECD fragmentation has been proven very useful for the identification of O-glycan structures and their attachment sites within O-glycopeptides.[5,96,105,108-111] For example, Mormann et al.[105] established a protocol for the analysis and characterization of both standard and synthetic O-glycosylated peptides in which a novel pulse sequence improved the efficiency for electron capture dissociation (ECD) of an unmodified Fourier Transform Ion Cyclotron Resonance (FTICR) mass spectrometer by more than an order of magnitude. This approach showed that as well as obtaining the amino acid sequence, information on the attachment site of the labile glycan moiety along with radical-site-induced fragmentation of the glycosidic bonds can be obtained from such an experiment.[105] Renfrow et al.[110,111] made use of ECD and activated ion-electron capture dissociation (AI-ECD) FTICR-MS2 to analyse the O-glycosylation of IgA. Previously established methods relied on the depression or absence of expected signal at particular cycles in the amino acid sequencing, whereas an AI-ECD product ion spectrum directly identifies sites with and without attached glycans.[110] This study led to the first direct identification of multiple sites of O-glycan attachment in IgA1 hinge region by mass spectrometry, thereby enabling future characterization at the molecular level of aberrant glycosylation of IgA1 in diseases. In a follow up study, both FT-ICR MS and liquid chromatography FT-ICR MS were used to analyse the O-glycan heterogeneity of IgA1 proteins. The techniques demonstrated that one could obtain unbiased accurate mass profiles of IgA1 HR glycopeptides from three different IgA1 myeloma proteins.[111] For the first time an individual IgA1 O-glycopeptide species from an IgA1

hinge region preparation that is reproducible for each IgA1 myeloma protein could be identified by AI-ECD fragmentation. Excitation of the precursor ions employing a CO2 laser prior to electron irradiation was found to be valuable for obtaining high-quality ECD spectra of O-glycosylated peptides (in this case IgA O-glycopeptides).[110,111]

In ETD fragmentation, the electron induced fragmentation results in cleavage of the N-Cα bonds of the peptide producing c- and z-type fragment ions.[112,113] Electron transfer triggers a fragmentation almost exclusively along the peptide backbone leaving the glycan mass intact and therefore allows the identification of the peptide sequence. A good quality ETD spectrum permits the identification of the peptide sequence, the glycan mass, and the unambiguous assignment of the modification site (Figure 5).[92]

An advantage of ETD compared to ECD is that the electrons are transferred to the analyte using an electron carrier, which makes the technique compatible with a wider (and more accessible) range of mass spectrometers. Most studies using ETD MS2 as a tool to analyse O-glycopeptides focused on peptides with a single glycan attached (for example (Christiansen et al.,[114] Perdivara et al.[93]) and Table 1), but recently studies have been published on the site-specific characterization of highly clustered glycosylation sites.[113,115]

There are a few disadvantages of ETD such as the requirement of a precursor ion with sufficient charge density for a high quality ETD fragmentation spectrum[92] and the inefficient fragmentation for peptides with a positive net charge <3+.[116] However, most database searching algorithms do not consider fragment ions bearing charge states higher than 2+; they will therefore not assign such ions, and

thus the peptide fragment assignment will be incomplete when highly charged precursors are fragmented.[117]

Another more general problem is the lack of availability of this fragmentation mode on many commonly used MS instruments. Christiansen et al.[114] tried to analyse the mucin-like O-glycosylation domains of a model protein. The protease digests produced many glycopeptides as determined by CID-MS/MS, but ETD fragmentation of these resulted in only a few interpretable spectra, suggesting that the use of ETD for determining the heterogeneous O-glycosylation at specific sites in regions of multiple occupancies is still in its infancy.[114]

In general, it is believed that the combination of CID and ETD fragmentation, along with improved data analysis software should greatly facilitate O-linked glycopeptides analysis in the future.[92]

Glycopeptides generated by tryptic digestion

The most commonly used enzyme for the proteolytic generation of glycopeptides is trypsin which cleaves the protein at well-defined sites. Specific cleavage makes the analysis of glycopeptides easier as the expected peptide masses can be predicted, if the amino acid sequence of the protein is available.[94]

Steentoft et al.[118] showed with their impressive work that enzymatically glycoengineered human cell lines with simplified O-glycosylation can be used for total O-glycoproteome analysis. Stable 'SimpleCell' lines with homogenous O-glycosylation were generated, which had only truncated GalNAcα or NeuAcα2-6GalNAcα O-glycans attached. Total cell lysates were digested, GalNAc O-glycopeptides isolated and used to analyse the O-glycoproteome by nLC-ESI-LIT-

FTICR-MS. A combination of higher-energy collision-induced dissociation (HCD) and ETD-MS2 mode allowed for peptide sequence analysis with and without retention of

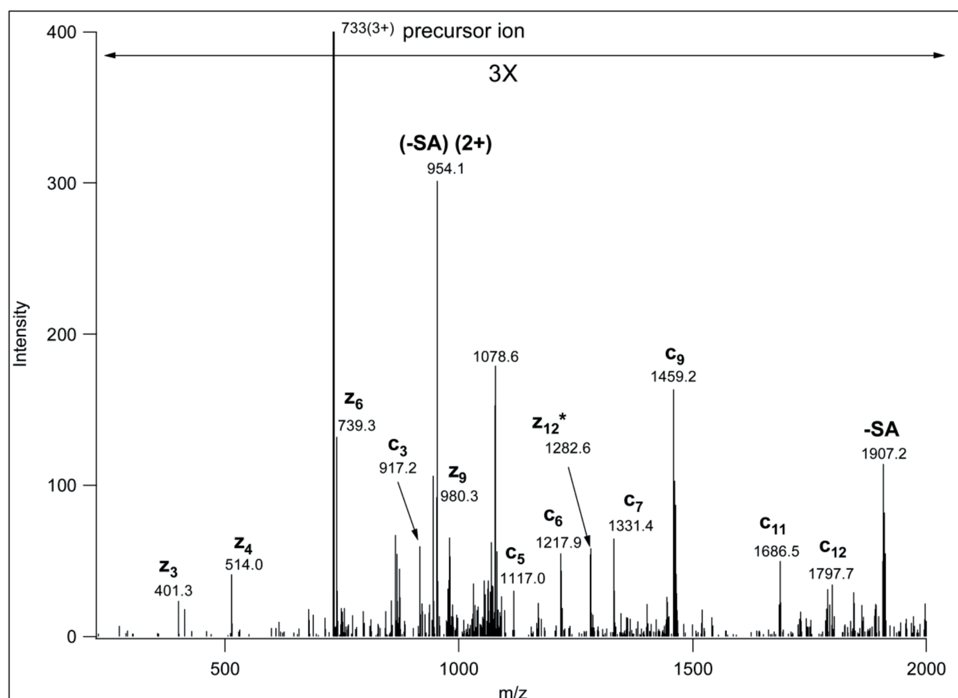

Figure 5. ETD spectrum of AAT(GalNAc1Gal1NeuAc1)LSTLAGQPLLER at m/z 733.0369 ($[M+2H]^{3+}$). This spectrum provides sufficient information for confident sequence identification as well as site assignment. Sialic acid loss from charge-reduced versions of the precursor ion was also detected. Fragments labelled with asterisk indicate hydrogen transfer, i.e. z+1 ions. Taken from Darula et al.[92] with permission.

glycan site-specific fragments. The raw data were searched against a database with a precursor mass tolerance of 10 ppm and a fragment ion mass tolerance of 50 millimass units. Methionine oxidation and HexNAc attachment to serine and threonine were used as variable modifications. Additionally, HCD data showing fragment ions at m/z 204.08 were extracted and further processed. The information obtained from ETD as well as HCD spectra was combined for confident glycopeptide

identifications and minimizing the number of false positive hits. In total, >100 O-glycoproteins with >350 O-glycan sites (the great majority previously unidentified), including a GalNAc O-glycan linkage to a tyrosine residue, were identified. By doing the analysis of three human simplified cell lines derived from different organs, the list of known O-glycoproteins could be already doubled while additional sites of O-glycosylation on several previously characterized O-glycoproteins could be identified. This strategy is expected to be amenable to other types of protein O-glycosylation.[118]

In another report, it was shown that a combination of CID and ETD techniques in LC-MS could prove to be a powerful tool for de novo identification of O-glycosylation at unknown modification sites in proteins.[93] More specifically, amyloid precursor protein (APP) was investigated in terms of its glycosylation. There are two previously reported N-glycosylation sites on APP but the authors focused on O-glycosylation plus the identification of their attachment sites: for the first time three threonines were found to be modified with core 1 type O-glycans. This discovery was possible due to combining information obtained from nanoHPLC-MS CID spectra (sugar oxonium ions), and the highly valuable ETD fragmentation pattern of the glycopeptides. In ETD mainly peptide fragmentation occurred with little or no cleavage of the oligosaccharide portion. Clearly, this is also a demonstration that using nanoHPLC in combination with sophisticated fragmentation techniques such as ETD has great potential for mapping unknown glycosylation sites in proteins.[93]

A method to enrich for glycoproteins from proteomic samples was developed recently.[96] In this study N- and O-glycosylation of proteins from human cerebrospinal fluid (CSF) were investigated. For this particular purpose, the sialic acids within the oligosaccharides linked to glycoproteins from CSF were mildly oxidised by periodate,

further captured onto hydrazide beads via the sialic acid residue of the glycans portion, trypsin digested such that only the glycopeptides remain attached to the beads, and finally the captured glycopeptides can be selectively released by acid hydrolysis. After release from the beads, the glycopeptides were analysed by reverse phase (RP)-LC-ESI-FTICR-MS. Low energy CID fragmentation MS2 and MS3 spectra were acquired, resulting in the identification of 36 N-linked and 44 O-linked glycosylation sites on glycoproteins from CSF. In addition, complementary analyses were conducted in which the precursor ions were guided to the ICR cell and fragmented using ECD. This study led to the discovery of five new O-linked glycosylation sites. Glycopeptide fragment spectra were searched against the Mascot database after entering the mass of the peptide portion as a precursor, computed by subtracting the monoisotopic mass of the glycan from the monoisotopic mass of the glycopeptide precursor. The following settings were applied: precursor mass tolerance, 10 ppm; fragment ion mass tolerance, 0.6 Da; enzyme, trypsin; one missed cleavage allowed; fixed carbamidomethyl modification of cysteine, variable oxidation of methionine. Even though this particular method has its limitation as a sialic acid is required for the capturing, this could also be seen as an advantage in terms of mass spectrometry fragmentation patterns which are simplified due to the removal of the terminal sialic acid residue(s). For complex biological samples, this covalent capturing method with subsequent MS analysis is advantageous as the vast majority of non-glycosylated peptides is removed.[96] For further references on the analysis of tryptic O-glycopeptides see Table 1.

Mass spectrometric approach	Sample	Achievements	Reference
CID fragmentation			
ESI-Q-TOF-MS^2	Synthetic MUC2 glycopeptides	Identification of glycan attachment sites	[62,101]
FTICR-MS SORI-CID-MS^2 and MS^3	Urinary O-glycopeptides from Schindler's disease	Negative-mode MS/MS for glycan structural elucidation	[119,120]
LC-ESI-IT-CID-MS^2; MALDI-TOF/TOF-MS^2	Apolipoprotein CIII glycopeptides	Free O-glycopeptides with aberrant glycosylation found in urine of schistosomiasis patients	[121]
ESI-IT-CID-MS^2	IgA1 hinge region tryptic O-glycopeptides	Glycoprofiling, data analysis using the GlycoSpectrumScan online tool	[122]
LC-ESI-IT-CID-MS^2	Human IgG1 monoclonal antibody and CTLA4-Ig fusion protein	Identification of a novel O-glycosylation site on tryptic peptides from CTLA4-Ig fusion protein	[123]
nano-ESI-QQQ-MS and QTOF-MS / CID-MS^2	Glycopeptides from Recombinant F-spondin	Identification of novel O-fucosylation sites	[124]
Nano-ESI-QTOF-MS / CID MS^2	synthetic bisglycosylated MUC1 peptide 25-mer	Comparison of fragmentation spectra on nano-ESI-QTOF-MS/MS and PSD- MALDI-TOF-MS	[98]
ESI-QTOF-CID-MS^2	Tryptic digests with subdigestions of Thrombospondin-1	Identification of three site that carry the O-linked disaccharide Glc-Fuc-O Ser/Thr	[125]
Chip based nano-ESI-QTOF-MS / CID-MS^2	Pronase E glycopeptides from bovine RNAse B, bovine Lactoferrin, bovine κ-casein, human IgG	Non-specifically cleaved glycopeptides can be analysed in a quantitative, isomer-sensitive and site-specific manner	[126]

Table 1. Mass spectrometric analysis of O-glycosylation at the glycopeptides level.

Mass spectrometric approach	Sample	Achievements	Reference
nano-ESI-QTOF-MS / CID-MS^2	Peptides originating from the thrombospondin-1 repeat	Protocol optimized for the identification of O-fucosylation sites	127
MALDI-TOF-MS; ESI-MS / CID-MS^2	O-glycosylated synthetic mucin-derived peptides	Peptide sequencing by CID-MS/MS enables the determination of the formerly O-glycosylated sites	81
ESI-QTOF-MS^2	Glycopeptides and Clostripain-cleaved MUC1 glycopeptides from human breast cancer cell lines	Breast cancer cell line seems to glycosylate the MUC1 tandem repeat peptide at higher density than lactating breast epithelia	128
nanoESI-QTOF-MS/ CID-MS^2	Pronase E derived glycopeptides from a mixture of bovine lactoferrin, bovine kappa casein and bovine fetuin	Protocol for separating and simultaneously characterizing glycopeptides from a mixture	129
UPLC-ESI-QTOF MS/ CID-MS^2	Urine from individuals with urinary tract infections and healthy controls	Identified and structural characterization of a unique C-terminal O-glycopeptide of the human fibrinogen alpha-chain	130
nanoLC-LTQ-MS/ CID-MS^2 and -MS^3	Tryptic α-dystroglycan glycopeptides	Developed a workflow and identified 91 glycopeptides that allowed us to assign 16 specific O-glycosylated residues	131
MALDI-TOF-MS and MALDI low-energy CID MS^2 on a QIT-TOF-MS	O-glycopeptides from recombinant human erythropoietin (rhEPO) from various suppliers obtained by digestion with a cleavage enzyme mix	Comparison of the O-glycopeptides revealed significant differences of the glycosylation patterns between the different rhEPO samples	102

Table 1. (Continued).

Mass spectrometric approach	Sample	Achievements	Reference
ZIC-HILIC-ESI-LIT-TOF-MS/ CID-MS^2	Glu-C digested rhEPO	ZIC-HILIC-MS method for the characterization of N- as well as O-glycosylation	132
LC-LTQ MS / CID-MS^2	Tryptic digests of muscle glycoproteins	Mucin-like domain of DGFc4 bared these modifications: the phosphorylated trisaccharide in conjunction with Hex-HexNAc-Hex, HexNAc-Hex, or Hex	133
nanoLC-ESI-MS/ CID-MS2 and -MS^3	Proteinase K generated glycopeptides of (asialo) fetuin	Protocol for HILIC separation and MS characterization of N- as well as O-glycopeptides	94
ETD/ ECD fragmentation			
nanoLC-ESI-LIT-FTICR-MS/ HCD- and ETD-MS^2	Three different human cell lines	Structural elucidation of the O-glycome within the cell lines	118
nanoLC-ESI-FTICR-MS^2 and -MS^3; combination of external accumulation (XA) with ECD	A 60-residue human-derived peptide and some synthetic peptides	XA-ECD allows shortened data acquisition time and facilitates the analysis of labile molecules	103
FTICR-MS / ECD-MS^2	O-glycosylated synthetic mucin-derived peptides	“Mild” ECD protocol suitable for analysis of such labile species as glycosylated polypeptides	104
FTICR-MS / ECD-MS^2	Standard peptides as well as synthetic glycosylated species	Improved ECD protocol for the analysis and characterization of O-glycosylated peptides	105
nano-ESI-FTICR-MS / ECD-MS^2	O-fucosylated glycopeptides from thrombospondin-1 and properdin	Protocol for characterization of O-fucosylated peptides	109

Table 1. (Continued).

Mass spectrometric approach	Sample	Achievements	Reference
ESI-Q-FTICR-MS / ECD-MS2	Synthetic IgA1 HR peptide and glycopeptide variants and IgA1 HR from a naturally Gal-deficient IgA1 myeloma protein	Successful identification of sites of attachment of O-glycans in the HR of a IgA1 myeloma protein	110
ESI-Q-FTICR-MS / activated ion ECD MS^2	IgA1 myeloma proteins from patient sera	Protocol for characterization of site specific glycosylation using only with small amounts of total protein	111
MALDI-TOF-MS and ESI-QTOF-MS / CID-MS^2	Tryptic (glyco)peptides derived from hepatitis B virus	Characterization of the overall glycosylation of all the subdomains within the virus	100
nanoLC- LTQ-FTICR-MS / ECD-MS^2	Clostripain derived glycopeptides from MUC1-Ig fusion protein	Determination of both the average O-glycosylation site occupancy and the occupancy of individual glycosylation sites in the TR of MUC1 glycoproteins	134
LC-ESI-FTICR-MS; ECD-MS^2 and ETD-MS^2	IgA-specific protease and trypsic digest of polymeric form of IgA1 myeloma protein from patient plasma	A new clinically relevant approach that requires ECD/electron transfer dissociation-type fragmentation to analyze clustered sites of O-glycosylation	115
Direct infusion on ESI-IT-MS/ ETD-MS^2	Tryptic peptide of the Fab region of human IgG	Identification of O-fucose on the antibody	135
MALDI-TOF-MS and ESI-IT-MS / ETD-MS^2	Enzymatic digestion by trypsin and lysylendopeptidase of IgA and hemopexin from sera of RA patients and healthy controls	Identification of an abnormality in the O-glycosylation of IgA associated with rheumatoid arthritis	136

Table 1. (Continued).

Mass spectrometric approach	Sample	Achievements	Reference
CID and ETD/ECD fragmentation			
nanoLC-MS / CID- and ETD-MS^2	Secreted APP695 produced in CHO cells	Identification and structural elucidation of three new O-glycosylation sites on APP695	[93]
ESI-IT-MS / CID- and ETD-MS^2	Synthetic MUC1 glycopeptides carrying different degree of O-glycosylation	Site specific characterisation of mucin-type glycopeptides containing highly densed O-glycan clusters	[113]
MALDI-TOF-MS and ESI-TOF-MS / ETD- and CID-MS^2	Cellobiohydrolase I&II and Endoglucanase I&II from *Trichoderma reesei*	Combining info from different analytical approaches allow to elucidation of heavily glycosylated O-linked glycopeptides	[114]
ESI-LIT-TOF-MS / tandem LITs for CID- and ECD-MS^2 and MS^3	Synthetic O-glycopeptides	Negative-ion CID and positive-ion ECD MSn spectra play complementary roles in the direct structural analysis of O-glycopeptides	[97]
nanoLC-LTQ-Orbitrap-MS / ETD- and CID-MS^2	O-glycopeptides enriched from bovine serum	Advanced protocols combining ETD and CID fragmentation spectra using Protein Prospector for identification	[92,95,117]
nano-LC- ESI-QTOF-MS / CID- and ETD-MS^2	Tryptic glycopeptides of AP 180	Tryptic glycopeptides of AP 180	[137]
nanoLC-ESI-FTICR-MS / CID- and ECD-MS^2 and MS^3	Peptides and glycopeptides from amyloid precursor protein/ amyloid β-peptides in CSF	Identification of novel O-glycosylation sites in threonins and tyrosine 10. Tyr glycosylation might be involved in AD progression	[108]

Table 1. (Continued).

Mass spectrometric approach	Sample	Achievements	Reference
nanoLC-ESI-LTQ-FTICR-MS / CID- and ECD-MS^2 and MS^3	Tryptic glycopeptides from human urine	63 O-linked glycopeptides from 53 glycoproteins were characterized of which 29 novel sites	[138]
FTICR-MS/ CID- and ECD-MS^2	Synthetic mucin derived O-glycopeptides	Protocol for the structural elucidation of mucin derived O-glycopeptides by ECD	[5]
nanoLC- LIT-ETD- and CID-MS^2	Synthetic glycopeptides and proteins of liver in response to traumahemorrhage and resuscitation in a rat model	Protocol using high affinity pan-specific antibodies for exploring the O-GlcNAc proteome	[139]
LC-ESI-FTICR-MS / CID- and ECD-MS2 and MS^3	Human CSF samples	Qualitative information of N- and O-linked glycosylation core structures and their attachment sites in sialylated glycoproteins in a proteomic mixture	[96]
Other MS techniques/ methods			
MALDI-TOF-PSD-MS^2	Synthetic MUC4 glycopeptides	For some glycopeptides observation of glycosylated fragment ions allowing site assignment	[101]
MALDI-ISD-MS	Synthetic glycopeptides, bovine asialofetuin and k-caseine	ISD MALDI MS is suitable for site determination of complex O-glycans at the protein level	[140]
Linear MALDI-TOF-MS and MALDI-TOF-PSD-MS^2	Clostripain-generated glycopeptides from isolated human milk mucin	Characterization of major glycans from human skim milk MUC1	[141]

Table 1. (Continued).

Mass spectrometric approach	Sample	Achievements	Reference
nano-ESI-Q-FTICR-MS2/ positive- and negative-mode IRMPD-MS^2	Pronase E-derived glycopeptides from bovine fetuin A and bovine kappa-casein	Protocol for the IRMPD of the deprotonated form of O-glycopeptides	142
CE-ESI-FTICR-MS	recombinant human chorionic gonadotrophin	Protocol for high-resolution CE-LTQ-FT-MS for the profiling of the intact glycoforms of r-RhCG	143

Table 1. (Continued).

Glycopeptides generated by use of unspecific proteinases

Lately, the use of unspecific proteinases such as Pronase E (proteinase mixture from Streptomyces griseus) or Proteinase K (from Tritirachium album) for glycoprotein cleavage has been shown to be useful for in-depth analysis of the site-specific N- and/or O-glycosylation of individual glycoproteins in many cases.[94,114,129,142,144-149]

Pronase E was immobilized on Sepharose beads for digesting two model proteins (bovine fetuin A and bovine kappa-casein) prior to nanoLC-FTICR-MS applying the infrared multiphoton dissociation (IRMPD) mode, which allows an assortment of O-linked glycopeptide ions encompassing various compositions and charge states.[142] In particular, IRMPD of the deprotonated form of O-glycopeptides has been shown to produce informative side chain cleavages from unoccupied serine and threonine residues, allowing determination of whether a serine or threonine residue is glycosylated. A combination of positive mode and negative mode MS2 was found to provide conclusive assignment of O-glycosites.[142]

In another example, Christiansen et al.[114] showed in their paper that ETD mass spectrometry of the protease-digested and HILIC-enriched highly glycosylated

domains of different glycoproteins allows determining the overall oligosaccharide heterogeneity.[114] Different enzymes have been used for digestion such as trypsin, AspN but also some unspecific proteinases such as proteinase K and the so-called PreTAQ. The main aim of this particular work was to develop complementary strategies to obtain structural information on the glycosylation of these difficult to analyze heterogeneously O-glycosylated domains. However, the O-glycan site occupancy of the mucin-like domains, with the modification intact, requires the determination of the peptide sequence. Clearly, ETD MS/MS has the potential to assign these O-glycosylation sites and at the same time give information on the peptide sequence too. It was shown that by using different types of analytical approaches and putting the information together, many of the structural details of the glycan heterogeneity can be determined on these heavily glycosylated O-linked glycopeptides.[114]

Analysis of synthetic glycopeptides

A more technical approach was undertaken by Thaysen-Anderson et al.[113] who used synthetic glycopeptides from the human mucin-1 (MUC-1) to test the use of ETD for the site-specific characterization of densely glycosylated mucin type O-glycopeptides. High peptide sequence coverage and confident glycosylation site determination were obtained from the infusion of individual glycopeptides, and from glycopeptide mixtures separated using two different online LC approaches (RP and porous graphitized carbon, PGC), using ETD-IT-MS2. Interestingly, highly sialylated glycopeptides showed poorer ETD fragmentation, which limits the applicability of the approach as mucin type O-glycans have a naturally high abundance of sialic acid residues. On the other hand, the capability to characterise the glycosylation sites of

mixtures of densely O-glycosylated mucin-type peptides while keeping the attached glycans structures intact during ETD fragmentation is amenable. Lately, synthetic O-glycopeptides served in some cases as excellent model systems to optimize various MS approaches (see Table 1 for details).

Naturally occurring O-glycopeptides

Recently, it was found that O-glycopeptides can be used to distinguish between healthy and diseased state: for example Balog et al.[121] identified a group of novel aberrantly O-glycosylated peptide derived from the C-terminus of human apolipoprotein C-III (apoC-III) in the urine of *S. mansoni*-infected individuals which was absent for healthy controls. The O-glycan portions identified appeared to represent highly fucosylated core 2 structures.

In another comparative urinary metabolomic study of patients suffering from urinary tract infection (UTI), Pacchiarotta et al.[130] identified a fibrinogen α-chain O-glycopeptide decorated with a di-sialylated T-antigen as the major discriminator between UTI patients and controls. Both studies give new insight into the (glyco)biology that seems to play a role in the diseased state. The identification of such O-glycosylation based discriminators indicates the potential of urinary glycopeptides as biomarkers for infections.

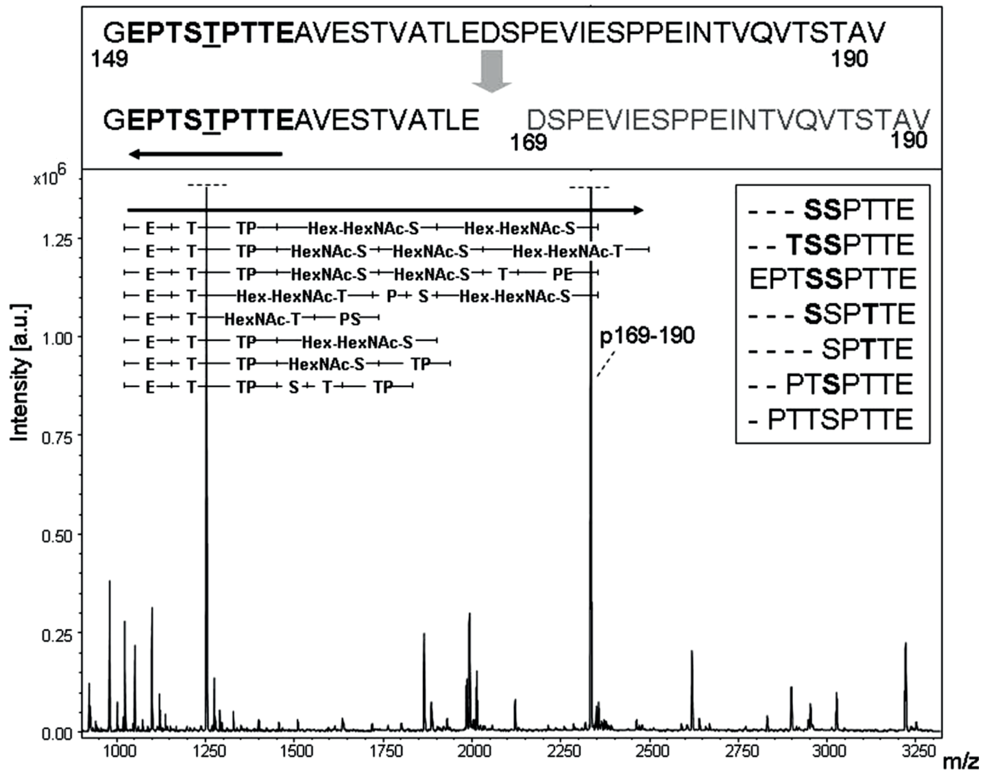

Figure 6. Top-down glycoprotein sequencing of bovine k-caseine. The sample was applied in sDHB matrix and analyzed by ISD-MALDI-MS. Sequences shown on top of the figure refer to the C-terminal domain of the protein, which spontaneously fragments in the ion source to yield p169-190 and a correspondingly truncated k-caseinoglycopeptide. The arrows indicate the reading direction of the amino acid sequence of the z-ion series annotated in the mass spectrum. The insert shows the readable section of clustered Ser/Thr (p150-158). The gray-shaded serine indicates a sequence ambiguity with published data. Glycosylated Ser/Thr residues are highlighted in bold. Taken from Hanisch et al.[140] with permission.

Thirty seven human amyloid precursor protein O-glycopeptides derived from CSF (attached to Thr(−39, −21, −20, and −13), with sialylated core 1 like O-glycans) and, unexpectedly, a series of 27 glycopeptides (the Aβ1-X series, where X was 20 (DAEFRHDSGYEVHHQKLVFF), which were all uniquely glycosylated on Tyr10 by ((Neu5Ac)1-2Hex-(Neu5Ac)-HexNAc-O- structures), were recently identified (Halim et al., 2011). An increase of up to 2.5 times of Tyr10 glycosylated peptides in CSF in

six Alzheimer's disease (AD) patients compared to seven non-AD patients was observed. These sialylated O-glycans, including that of a Tyr residue, the first in a mammalian protein, may modulate human amyloid precursor protein processing, and therefore influence the amyloidogenic pathway associated with AD.

The expected proteolytic resistance of the O-glycosylated peptide stretches may lead to the accumulation of these glycopeptide moieties in degradation pathways, in particular in pathological situations. Hopefully, with the advances in the field of O-glycoproteomics many more O-glycopeptides can be identified within the near future that may be important in (monitoring) disease.[150]

Perspectives

One of the major challenges in the analysis of O-glycans is the unwanted side reaction occurring once the oligosaccharides are released from the protein backbone (named 'peeling'[69]). Hopefully, there will be ways indentified to reduce/abolish this 'peeling' reaction and/or combine the release of O-glycans with a suitable labelling technique. Some recent papers have described first steps towards this goal.[55,67,68,71,151] The efficient blocking of the reducing end of the formerly O-linked oligosaccharide appears to be an important aspect in improving the stability of the O-glycans.

Rather novel techniques for O-glycopeptide analysis such as the radical-mediated mass spectrometric fragmentation methods ETD and ECD are becoming more and more popular for site specific glycoprotein analysis and have expanded the analytical options in glycoproteomics tremendously (see the ECD/ETD Fragmentation of Glycopeptides section). These methods (in particular ETD) will,

probably in combination with CID fragmentation, allow a very detailed analysis of more glycoproteins in the future.

Other important aspects in the O-glycoproteomics field are spectral assignment and data analysis. In this respect, a very interesting collaboration of researchers is the UniCarbDB platform (http://www.unicarb-db.com/) which is currently focused on the mass spectrometric data and structural assignment based on fragmentation data. The aim is to generate a assigned database of assigned spectra and structures supporting structural elucidation.[152]

Recent publications show that combinatorial approaches or glyco-engineering can help to get a picture of the O-glycoproteome of the sample of interest.[96,118] While these approaches are successful for simple or simplified O-glycosylation, they might not be suitable to deal with peptides carrying large and structurally diverse O-glycans, requiring manual interpretation of the corresponding fragmentation spectra.

One very elegant approach has recently been published, in which top-down sequencing on the protein level by MALDI-MS based on the in-source decay (ISD) of intact glycoproteins was induced by hydrogen radical transfer from the matrix.[140] It allows a ladder sequencing from both termini with assignment of O-glycosylation sites based on intense c-, y- and z-type ions. In Figure 6 the top-down glycoprotein sequencing of bovine k-caseine is shown as an example. The feasibility of ISD-MALDI-MS in the localization of O-glycosylation sites was clearly demonstrated for several proteins as fragments are formed under conditions that leave O-glycosidic bonds largely unaffected. This method gives comparable information to that obtained with post source decay (PSD) MALDI or ETD-ESI-MS, which shows its potential for the site determination of complex O-glycans at the protein level. After further

optimization of this approach, it is expected to present an attractive and powerful tool for the site determination of complex O-glycans at the protein level.[140]

The above mentioned technical approaches are already major improvements/advances to study O-glycosylation in greater detail. We believe that also in O-glycoproteomics we have only seen the tip of the glycobiology iceberg,[153] and that there will be major and significant discoveries ahead of us due to the combination/ further improvements of all the latest technical advances.

Reference List:

(1) Spiro, R. G. *Glycobiology* **2002**, *12*, 43R-56R.

(2) Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. *Essential of Glycobiology*, 2 ed.; Cold Spring Harbor Laboratory Press, 2009.

(3) Gill, D. J.; Clausen, H.; Bard, F. *Trends Cell Biol* **2011**, *21*, 149-158.

(4) Ten Hagen, K. G.; Fritz, T. A.; Tabak, L. A. *Glycobiology* **2003**, *13*, 1R-16R.

(5) Mormann, M.; Paulsen, H.; Peter-Katalinic, J. *Eur. J. Mass Spectrom. (Chichester, Eng)* **2005**, *11*, 497-511.

(6) Park, J. H.; Nishidate, T.; Kijima, K.; Ohashi, T.; Takegawa, K.; Fujikane, T.; Hirata, K.; Nakamura, Y.; Katagiri, T. *Cancer Res* **2010**, *70*, 2759-2769.

(7) Taylor-Papadimitriou, J.; Burchell, J.; Miles, D. W.; Daziel, M. *Biochimica et Biophysica Acta* **1999**, *1455*, 301-313.

(8) Pedersen, J. W.; Blixt, O.; Bennett, E. P.; Tarp, M. A.; Dar, I.; Mandel, U.; Poulsen, S. S.; Pedersen, A. E.; Rasmussen, S.; Jess, P.; Clausen, H.; Wandall, H. H. *Int. J. Cancer* **2011**, *128*, 1860-1871.

(9) Stalnaker, S. H.; Stuart, R.; Wells, L. *Curr. Opin. Struct. Biol* **2011**, *21*, 603-609.

(10) Haltiwanger, R. S.; Lowe, J. B. *Annu. Rev. Biochem* **2004**, *73*, 491-537.

(11) Sethi, M. K.; Buettner, F. F.; Ashikov, A.; Krylov, V. B.; Takeuchi, H.; Nifantiev, N. E.; Haltiwanger, R. S.; Gerardy-Schahn, R.; Bakker, H. *J. Biol. Chem* **2012**, *287*, 2739-2748.

(12) Lazarus, B. D.; Love, D. C.; Hanover, J. A. *Int. J. Biochem. Cell Biol* **2009**, *41*, 2134-2146.

(13) Geyer, H.; Geyer, R. *Biochimica et Biophysica Acta* **2006**, *1764*, 1853-1869.

(14) Harvey, D. J. *Expert. Rev. Proteomics* **2005**, *2*, 87-101.

(15) Zaia, J. *OMICS* **2010**, *14*, 401-418.

(16) Hulsmeier, A. J.; Welti, M.; Hennet, T. *Methods Enzymol* **2011**, *491*, 163-182.

(17) Wada, Y.; Dell, A.; Haslam, S. M.; Tissot, B.; Canis, K.; Azadi, P.; Backstrom, M.; Costello, C. E.; Hansson, G. C.; Hiki, Y.; Ishihara, M.; Ito, H.; Kakehi, K.; Karlsson, N.; Hayes, C. E.; Kato, K.; Kawasaki, N.; Khoo, K. H.; Kobayashi, K.; Kolarich, D.; Kondo, A.; Lebrilla, C.; Nakano, M.; Narimatsu, H.; Novak, J.; Novotny,

M. V.; Ohno, E.; Packer, N. H.; Palaima, E.; Renfrow, M. B.; Tajiri, M.; Thomsson, K. A.; Yagi, H.; Yu, S. Y.; Taniguchi, N. *Mol. Cell Proteomics* **2010**, *9*, 719-727.
(18) Ruhaak, L. R.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A. M.; Wuhrer, M. *Anal. Bioanal. Chem* **2010**, *397*, 3457-3481.
(19) North, S. J.; Hitchen, P. G.; Haslam, S. M.; Dell, A. *Curr. Opin. Struct. Biol* **2009**, *19*, 498-506.
(20) Peter-Katalinic, J. *Methods Enzymol* **2005**, *405*, 139-171.
(21) Dwek, R. A.; Edge, C. J.; Harvey, D. J.; Wormald, M. R.; Parekh, R. B. *Annu. Rev. Biochem* **1993**, *62*, 65-100.
(22) Carlson, D. M. *J. Biol. Chem* **1968**, *243*, 616-626.
(23) Iyer, R. N.; Carlson, D. M. *Arch. Biochem. Biophys* **1971**, *142*, 101-105.
(24) Cooper, C. A.; Packer, N. H.; Redmond, J. W. *Glycoconj. J* **1994**, *11*, 163-167.
(25) Brooks, R. D. *A kinetic study of the Rate of Cleavage of the Glycosidic Bond of Methyl-beta-Glucopyranoside in an Alkaline Medium*. The Institute of Paper Chemistry, Appleton, Wisconsin.1966.
(26) Corbett, W. M.; Kenner, J. *Journal of the Chemical Society* **1955**, 1431-1435.
(27) Greville, M.; Richards, G. N. *Journal of the Chemical Society* **1960**, 1924-1931.
(28) MacLaurin, D. J. *A study of some reaction rates in the homogeneous system water-sodium hydroxide-cellobiose*. The Institute of Paper Chemistry, Appleton, Wisconsin.1969.
(29) Richards, G. N.; Sephton, H. H. *Journal of the Chemical Society* **1957**, 4492-4499.
(30) WHISTLER, R. L.; BEMILLER, J. N. *Adv. Carbohydr. Chem* **1958**, *13*, 289-329.
(31) Kotani, N.; Takasaki, S. *Anal. Biochem* **1997**, *252*, 40-47.
(32) Lloyd, K. O.; Burchell, J.; Kudryashov, V.; Yin, B. W.; Taylor-Papadimitriou, J. *J. Biol. Chem* **1996**, *271*, 33325-33334.
(33) Doohan, R. A.; Hayes, C. A.; Harhen, B.; Karlsson, N. G. *J. Am. Soc. Mass Spectrom* **2011**, *22*, 1052-1062.
(34) Faid, V.; Chirat, F.; Seta, N.; Foulquier, F.; Morelle, W. *Proteomics* **2007**, *7*, 1800-1813.
(35) Ismail, M. N.; Stone, E. L.; Panico, M.; Lee, S. H.; Luu, Y.; Ramirez, K.; Ho, S. B.; Fukuda, M.; Marth, J. D.; Haslam, S. M.; Dell, A. *Glycobiology* **2011**, *21*, 82-98.
(36) Schulz, B. L.; Packer, N. H.; Karlsson, N. G. *Anal. Chem* **2002**, *74*, 6088-6097.

(37) Taylor, A. M.; Holst, O.; Thomas-Oates, J. *Proteomics* **2006**, *6*, 2936-2946.
(38) Cooke, C. L.; An, H. J.; Kim, J.; Solnick, J. V.; Lebrilla, C. B. *Anal. Chem* **2007**, *79*, 8090-8097.
(39) Robijn, M. L.; Koeleman, C. A.; Wuhrer, M.; Royle, L.; Geyer, R.; Dwek, R. A.; Rudd, P. M.; Deelder, A. M.; Hokke, C. H. *Mol. Biochem. Parasitol* **2007**, *151*, 148-161.
(40) Kawasaki, N.; Ohta, M.; Hyuga, S.; Hyuga, M.; Hayakawa, T. *Anal. Biochem* **2000**, *285*, 82-91.
(41) Thomsson, K. A.; Prakobphol, A.; Leffler, H.; Reddy, M. S.; Levine, M. J.; Fisher, S. J.; Hansson, G. C. *Glycobiology* **2002**, *12*, 1-14.
(42) Karlsson, N. G.; Schulz, B. L.; Packer, N. H. *J. Am. Soc. Mass Spectrom* **2004**, *15*, 659-672.
(43) Larsson, J. M.; Karlsson, H.; Sjovall, H.; Hansson, G. C. *Glycobiology* **2009**, *19*, 756-766.
(44) Ciucanu, I.; Costello, C. E. *J. Am. Chem. Soc* **2003**, *125*, 16213-16219.
(45) Jang-Lee, J.; North, S. J.; Sutton-Smith, M.; Goldberg, D.; Panico, M.; Morris, H.; Haslam, S.; Dell, A. *Methods Enzymol* **2006**, *415*, 59-86.
(46) Ciucanu, I.; Kerek, F. *Carbohydrate Research* **1984**, *131*, 209-217.
(47) Geyer, R.; Geyer, H. *Methods Enzymol* **1994**, *230*, 86-108.
(48) Morelle, W.; Michalski, J. C. *Nat. Protoc* **2007**, *2*, 1585-1602.
(49) Morelle, W.; Faid, V.; Chirat, F.; Michalski, J. C. *Methods Mol. Biol* **2009**, *534*, 5-21.
(50) Robbe, C.; Michalski, J. C.; Capon, C. *Methods Mol. Biol* **2006**, *347*, 109-123.
(51) Hanisch, F. G.; Uhlenbruck, G.; Peter-Katalinic, J.; Egge, H.; Dabrowski, J.; Dabrowski, U. *J. Biol. Chem* **1989**, *264*, 872-883.
(52) Egge, H.; Peter-Katalinic, J. *Mass Spectrometry Reviews* **1987**, *6*, 331-393.
(53) Bleckmann, C.; Geyer, H.; Lieberoth, A.; Splittstoesser, F.; Liu, Y.; Feizi, T.; Schachner, M.; Kleene, R.; Reinhold, V.; Geyer, R. *Biol. Chem* **2009**, *390*, 627-645.
(54) Parry, S.; Wong, N. K.; Easton, R. L.; Panico, M.; Haslam, S. M.; Morris, H. R.; Anderson, P.; Klotz, K. L.; Herr, J. C.; Diekman, A. B.; Dell, A. *Glycobiology* **2007**, *17*, 1120-1126.
(55) Maniatis, S.; Zhou, H.; Reinhold, V. *Anal. Chem* **2010**, *82*, 2421-2425.
(56) Hu, P.; Shimoji, S.; Hart, G. W. *FEBS Lett* **2010**, *584*, 2526-2538.

(57) Chai, W.; Feizi, T.; Yuen, C. T.; Lawson, A. M. *Glycobiology* **1997**, *7*, 861-872.
(58) Yamada, K.; Hyodo, S.; Kinoshita, M.; Hayakawa, T.; Kakehi, K. *Anal. Chem* **2010**, *82*, 7436-7443.
(59) Huang, Y.; Mechref, Y.; Novotny, M. V. *Anal. Chem* **2001**, *73*, 6063-6069.
(60) Yu, G.; Zhang, Y.; Zhang, Z.; Song, L.; Wang, P.; Chai, W. *Anal. Chem* **2010**, *82*, 9534-9542.
(61) Miura, Y.; Kato, K.; Takegawa, Y.; Kurogochi, M.; Furukawa, J.; Shinohara, Y.; Nagahori, N.; Amano, M.; Hinou, H.; Nishimura, S. *Anal. Chem* **2010**, *82*, 10021-10029.
(62) Domon, B.; Costello, C. E. *Glycoconjugate journal* **1988**, *5*, 397-409.
(63) Goetz, J. A.; Novotny, M. V.; Mechref, Y. *Anal. Chem* **2009**, *81*, 9546-9552.
(64) Kang, P.; Mechref, Y.; Klouckova, I.; Novotny, M. V. *Rapid Commun. Mass Spectrom* **2005**, *19*, 3421-3428.
(65) Kang, P.; Mechref, Y.; Novotny, M. V. *Rapid Commun. Mass Spectrom* **2008**, *22*, 721-734.
(66) Zauner, G.; Koeleman, C. A.; Deelder, A. M.; Wuhrer, M. *Biochim. Biophys. Acta* **2012**, *1820*, 1420-1428.
(67) Wang, C.; Fan, W.; Zhang, P.; Wang, Z.; Huang, L. *Proteomics* **2011**, *11*, 4229-4242.
(68) Furukawa, J.; Fujitani, N.; Araki, K.; Takegawa, Y.; Kodama, K.; Shinohara, Y. *Anal. Chem* **2011**, *83*, 9060-9067.
(69) Merry, A. H.; Neville, D. C.; Royle, L.; Matthews, B.; Harvey, D. J.; Dwek, R. A.; Rudd, P. M. *Anal. Biochem* **2002**, *304*, 91-99.
(70) Patel, T.; Bruce, J.; Merry, A.; Bigge, C.; Wormald, M.; Jaques, A.; Parekh, R. *Biochemistry* **1993**, *32*, 679-693.
(71) Kozak, R. P.; Royle, L.; Gardner, R. A.; Fernandes, D. L.; Wuhrer, M. *Anal. Biochem* **2012**, *423*, 119-128.
(72) Ceroni, A.; Maass, K.; Geyer, H.; Geyer, R.; Dell, A.; Haslam, S. M. *J. Proteome. Res* **2008**, *7*, 1650-1659.
(73) Maass, K.; Ranzinger, R.; Geyer, H.; von der Lieth, C. W.; Geyer, R. *Proteomics* **2007**, *7*, 4435-4444.
(74) Campbell, M. P.; Royle, L.; Radcliffe, C. M.; Dwek, R. A.; Rudd, P. M. *Bioinformatics* **2008**, *24*, 1214-1216.

(75) Jansson, P. E.; Stenutz, R.; Widmalm, G. *Carbohydr. Res* **2006**, *341*, 1003-1010.
(76) Lutteke, T.; Bohne-Lang, A.; Loss, A.; Goetz, T.; Frank, M.; von der Lieth, C. W. *Glycobiology* **2006**, *16*, 71R-81R.
(77) Jensen, P. H.; Kolarich, D.; Packer, N. H. *FEBS J* **2010**, *277*, 81-94.
(78) Rademaker, G. J.; Haverkamp, J.; Thomas-Oates, J. *Organic Mass Spectrometry* **1993**, *28*, 1536-1541.
(79) Rademaker, G. J.; Pergantis, S. A.; Blok-Tip, L.; Langridge, J. I.; Kleen, A.; Thomas-Oates, J. E. *Anal. Biochem* **1998**, *257*, 149-160.
(80) Hanisch, F. G.; Jovanovic, M.; Peter-Katalinic, J. *Analytical biochemistry* **2001**, *290*, 47-59.
(81) Mirgorodskaya, E.; Hassan, H.; Clausen, H.; Roepstorff, P. *Analytical chemistry* **2001**, *73*, 1263-1269.
(82) Zheng, Y.; Guo, Z.; Cai, Z. *Talanta* **2009**, *78*, 358-363.
(83) Wells, L.; Vosseller, K.; Cole, R. N.; Cronshaw, J. M.; Matunis, M. J.; Hart, G. W. *Molecular & cellular proteomics : MCP* **2002**, *1*, 791-804.
(84) Vosseller, K.; Hansen, K. C.; Chalkley, R. J.; Trinidad, J. C.; Wells, L.; Hart, G. W.; Burlingame, A. L. *Proteomics* **2005**, *5*, 388-398.
(85) Czeszak, X.; Ricart, G.; Tetaert, D.; Michalski, J. C.; Lemoine, J. *Rapid communications in mass spectrometry : RCM* **2002**, *16*, 27-34.
(86) Hansen, J. E.; Lund, O.; Tolstrup, N.; Gooley, A. A.; Williams, K. L.; Brunak, S. *Glycoconjugate journal* **1998**, *15*, 115-130.
(87) Hansen, J. E.; Lund, O.; Nielsen, J. O.; Hansen, J. E.; Brunak, S. *Nucleic acids research* **1996**, *24*, 248-252.
(88) Chen, Y. Z.; Tang, Y. R.; Sheng, Z. Y.; Zhang, Z. *BMC bioinformatics* **2008**, *9*, 101.
(89) Medzihradszky, K. F.; Gillece-Castro, B. L.; Settineri, C. A.; Townsend, R. R.; Masiarz, F. R.; Burlingame, A. L. *Biomedical & environmental mass spectrometry* **1990**, *19*, 777-781.
(90) Medzihradszkyaff, K. F.; Gillece-Castroaff, B. L.; Townsendaff, R. R.; Burlingameaff, A. L.; Hardyaff, M. R. *Journal of the American Society for Mass Spectrometry* **1996**, *7*, 319-328.

(91) Settineri, C. A.; Medzihradszky, K. F.; Masiarz, F. R.; Burlingame, A. L.; Chu, C.; George-Nascimento, C. *Biomedical & environmental mass spectrometry* **1990**, *19*, 665-676.
(92) Darula, Z.; Chalkley, R. J.; Baker, P.; Burlingame, A. L.; Medzihradszky, K. F. *European journal of mass spectrometry (Chichester, England)* **2010**, *16*, 421-428.
(93) Perdivara, I.; Petrovich, R.; Allinquant, B.; Deterding, L. J.; Tomer, K. B.; Przybylski, M. *Journal of proteome research* **2009**, *8*, 631-642.
(94) Zauner, G.; Koeleman, C. A.; Deelder, A. M.; Wuhrer, M. *Journal of separation science* **2010**, *33*, 903-910.
(95) Darula, Z.; Chalkley, R. J.; Lynn, A.; Baker, P. R.; Medzihradszky, K. F. *Amino acids* **2011**, *41*, 321-328.
(96) Nilsson, J.; Ruetschi, U.; Halim, A.; Hesse, C.; Carlsohn, E.; Brinkmalm, G.; Larson, G. *Nature methods* **2009**, *6*, 809-811.
(97) Deguchi, K.; Ito, H.; Baba, T.; Hirabayashi, A.; Nakagawa, H.; Fumoto, M.; Hinou, H.; Nishimura, S. *Rapid communications in mass spectrometry : RCM* **2007**, *21*, 691-698.
(98) Hanisch, F. G.; Green, B. N.; Bateman, R.; Peter-Katalinic, J. *Journal of mass spectrometry : JMS* **1998**, *33*, 358-362.
(99) Wuhrer, M.; Catalina, M. I.; Deelder, A. M.; Hokke, C. H. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **2007**, *849*, 115-128.
(100) Schmitt, S.; Glebe, D.; Alving, K.; Tolle, T. K.; Linder, M.; Geyer, H.; Linder, D.; Peter-Katalinic, J.; Gerlich, W. H.; Geyer, R. *The Journal of biological chemistry* **1999**, *274*, 11945-11957.
(101) Alving, K.; Paulsen, H.; Peter-Katalinic, J. *Journal of mass spectrometry : JMS* **1999**, *34*, 395-407.
(102) Stubiger, G.; Marchetti, M.; Nagano, M.; Grimm, R.; Gmeiner, G.; Reichel, C.; Allmaier, G. *Journal of separation science* **2005**, *28*, 1764-1778.
(103) Haselmann, K. F.; Budnik, B. A.; Olsen, J. V.; Nielsen, M. L.; Reis, C. A.; Clausen, H.; Johnsen, A. H.; Zubarev, R. A. *Analytical chemistry* **2001**, *73*, 2998-3005.
(104) Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Analytical chemistry* **1999**, *71*, 4431-4436.

(105) Mormann, M.; Peter-Katalinic, J. *Rapid communications in mass spectrometry : RCM* **2003**, *17*, 2208-2214.
(106) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Analytical chemistry* **2000**, *72*, 563-573.
(107) Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Sorensen, E. S.; Zubarev, R. A. *Analytical chemistry* **2003**, *75*, 2355-2361.
(108) Halim, A.; Brinkmalm, G.; Ruetschi, U.; Westman-Brinkmalm, A.; Portelius, E.; Zetterberg, H.; Blennow, K.; Larson, G.; Nilsson, J. *Proceedings of the National Academy of Sciences of the United States of America* **2011**, *108*, 11848-11853.
(109) Mormann, M.; Macek, B.; Gonzalez de Peredo, A.; Hofsteenge, J.; Peter-Katalinic, J. *International Journal of Mass Spectrometry* **2004**, *234*, 11-21.
(110) Renfrow, M. B.; Cooper, H. J.; Tomana, M.; Kulhavy, R.; Hiki, Y.; Toma, K.; Emmett, M. R.; Mestecky, J.; Marshall, A. G.; Novak, J. *The Journal of biological chemistry* **2005**, *280*, 19136-19145.
(111) Renfrow, M. B.; Mackay, C. L.; Chalmers, M. J.; Julian, B. A.; Mestecky, J.; Kilian, M.; Poulsen, K.; Emmett, M. R.; Marshall, A. G.; Novak, J. *Analytical and bioanalytical chemistry* **2007**, *389*, 1397-1407.
(112) Syka, J. E.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 9528-9533.
(113) Thaysen-Andersen, M.; Wilkinson, B. L.; Payne, R. J.; Packer, N. H. *Electrophoresis* **2011**, *32*, 3536-3545.
(114) Christiansen, M. N.; Kolarich, D.; Nevalainen, H.; Packer, N. H.; Jensen, P. H. *Analytical chemistry* **2010**, *82*, 3500-3509.
(115) Takahashi, K.; Wall, S. B.; Suzuki, H.; Smith, A. D. t.; Hall, S.; Poulsen, K.; Kilian, M.; Mobley, J. A.; Julian, B. A.; Mestecky, J.; Novak, J.; Renfrow, M. B. *Molecular & cellular proteomics : MCP* **2010**, *9*, 2545-2557.
(116) Halfinger, B.; Sarg, B.; Lindner, H. H. *Electrophoresis* **2011**, *32*, 3546-3553.
(117) Darula, Z.; Medzihradszky, K. F. *Molecular & cellular proteomics : MCP* **2009**, *8*, 2515-2526.

(118) 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. *Nature methods* **2011**, *8*, 977-982.
(119) Froesch, M.; Bindila, L.; Zamfir, A.; Peter-Katalinic, J. *Rapid communications in mass spectrometry : RCM* **2003**, *17*, 2822-2832.
(120) Froesch, M.; Bindila, L. M.; Baykut, G.; Allen, M.; Peter-Katalinic, J.; Zamfir, A. D. *Rapid communications in mass spectrometry : RCM* **2004**, *18*, 3084-3092.
(121) Balog, C. I.; Mayboroda, O. A.; Wuhrer, M.; Hokke, C. H.; Deelder, A. M.; Hensbergen, P. J. *Molecular & cellular proteomics : MCP* **2010**, *9*, 667-681.
(122) Deshpande, N.; Jensen, P. H.; Packer, N. H.; Kolarich, D. *Journal of proteome research* **2010**, *9*, 1063-1075.
(123) Bongers, J.; Devincentis, J.; Fu, J.; Huang, P.; Kirkley, D. H.; Leister, K.; Liu, P.; Ludwig, R.; Rumney, K.; Tao, L.; Wu, W.; Russell, R. J. *Journal of chromatography. A* **2011**, *1218*, 8140-8149.
(124) Gonzalez de Peredo, A.; Klein, D.; Macek, B.; Hess, D.; Peter-Katalinic, J.; Hofsteenge, J. *Molecular & cellular proteomics : MCP* **2002**, *1*, 11-18.
(125) Hofsteenge, J.; Huwiler, K. G.; Macek, B.; Hess, D.; Lawler, J.; Mosher, D. F.; Peter-Katalinic, J. *The Journal of biological chemistry* **2001**, *276*, 6485-6498.
(126) Hua, S.; Nwosu, C. C.; Strum, J. S.; Seipert, R. R.; An, H. J.; Zivkovic, A. M.; German, J. B.; Lebrilla, C. B. *Analytical and bioanalytical chemistry* **2012**, *403*, 1291-1302.
(127) Macek, B.; Hofsteenge, J.; Peter-Katalinic, J. *Rapid communications in mass spectrometry : RCM* **2001**, *15*, 771-777.
(128) Muller, S.; Alving, K.; Peter-Katalinic, J.; Zachara, N.; Gooley, A. A.; Hanisch, F. G. *The Journal of biological chemistry* **1999**, *274*, 18165-18172.
(129) Nwosu, C. C.; Seipert, R. R.; Strum, J. S.; Hua, S. S.; An, H. J.; Zivkovic, A. M.; German, B. J.; Lebrilla, C. B. *Journal of proteome research* **2011**, *10*, 2612-2624.
(130) Pacchiarotta, T.; Hensbergen, P. J.; Wuhrer, M.; van Nieuwkoop, C.; Nevedomskaya, E.; Derks, R. J.; Schoenmaker, B.; Koeleman, C. A.; van Dissel, J.; Deelder, A. M.; Mayboroda, O. A. *Journal of proteomics* **2012**, *75*, 1067-1073.
(131) Stalnaker, S. H.; Hashmi, S.; Lim, J. M.; Aoki, K.; Porterfield, M.; Gutierrez-Sanchez, G.; Wheeler, J.; Ervasti, J. M.; Bergmann, C.; Tiemeyer, M.; Wells, L. *The Journal of biological chemistry* **2010**, *285*, 24882-24891.

(132) Takegawa, Y.; Ito, H.; Keira, T.; Deguchi, K.; Nakagawa, H.; Nishimura, S. *Journal of separation science* **2008**, *31*, 1585-1593.
(133) Yoshida-Moriguchi, T.; Yu, L.; Stalnaker, S. H.; Davis, S.; Kunz, S.; Madson, M.; Oldstone, M. B.; Schachter, H.; Wells, L.; Campbell, K. P. *Science (New York, N.Y.)* **2010**, *327*, 88-92.
(134) Sihlbom, C.; van Dijk Hard, I.; Lidell, M. E.; Noll, T.; Hansson, G. C.; Backstrom, M. *Glycobiology* **2009**, *19*, 375-381.
(135) Valliere-Douglass, J. F.; Brady, L. J.; Farnsworth, C.; Pace, D.; Balland, A.; Wallace, A.; Wang, W.; Treuheit, M. J.; Yan, B. *Glycobiology* **2009**, *19*, 144-152.
(136) Wada, Y.; Tajiri, M.; Ohshima, S. *Journal of proteome research* **2010**, *9*, 1367-1373.
(137) Graham, M. E.; Thaysen-Andersen, M.; Bache, N.; Craft, G. E.; Larsen, M. R.; Packer, N. H.; Robinson, P. J. *Journal of proteome research* **2011**, *10*, 2725-2733.
(138) Halim, A.; Nilsson, J.; Ruetschi, U.; Hesse, C.; Larson, G. *Molecular & cellular proteomics : MCP* **2012**, *11*, M111.013649.
(139) Teo, C. F.; Ingale, S.; Wolfert, M. A.; Elsayed, G. A.; Not, L. G.; Chatham, J. C.; Wells, L.; Boons, G. J. *Nature chemical biology* **2010**, *6*, 338-343.
(140) Hanisch, F. G. *Analytical chemistry* **2011**, *83*, 4829-4837.
(141) Muller, S.; Goletz, S.; Packer, N.; Gooley, A.; Lawson, A. M.; Hanisch, F. G. *The Journal of biological chemistry* **1997**, *272*, 24780-24793.
(142) Seipert, R. R.; Dodds, E. D.; Lebrilla, C. B. *Journal of proteome research* **2009**, *8*, 493-501.
(143) Thakur, D.; Rejtar, T.; Karger, B. L.; Washburn, N. J.; Bosques, C. J.; Gunay, N. S.; Shriver, Z.; Venkataraman, G. *Analytical chemistry* **2009**, *81*, 8900-8907.
(144) An, H. J.; Peavy, T. R.; Hedrick, J. L.; Lebrilla, C. B. *Analytical chemistry* **2003**, *75*, 5628-5637.
(145) Clowers, B. H.; Dodds, E. D.; Seipert, R. R.; Lebrilla, C. B. *Journal of proteome research* **2007**, *6*, 4032-4040.
(146) Dodds, E. D.; Seipert, R. R.; Clowers, B. H.; German, J. B.; Lebrilla, C. B. *Journal of proteome research* **2009**, *8*, 502-512.
(147) Froehlich, J. W.; Barboza, M.; Chu, C.; Lerno, L. A., Jr.; Clowers, B. H.; Zivkovic, A. M.; German, J. B.; Lebrilla, C. B. *Analytical chemistry* **2011**, *83*, 5541-5547.

(148) Songsrirote, K.; Li, Z.; Ashford, D.; Bateman, A.; Thomas-Oates, J. *Journal of proteomics* **2010**, *73*, 1479-1490.
(149) Wuhrer, M.; Koeleman, C. A.; Hokke, C. H.; Deelder, A. M. *Analytical chemistry* **2005**, *77*, 886-894.
(150) Cooper, H. J.; Hakansson, K.; Marshall, A. G. *Mass Spectrom Rev* **2005**, *24*, 201-222.
(151) Zauner, G.; Deelder, A. M.; Wuhrer, M. *Electrophoresis* **2011**, *32*, 3456-3466.
(152) Hayes, C. A.; Karlsson, N. G.; Struwe, W. B.; Lisacek, F.; Rudd, P. M.; Packer, N. H.; Campbell, M. P. *Bioinformatics* **2011**, *27*, 1343-1344.
(153) Marino, K.; Bones, J.; Kattla, J. J.; Rudd, P. M. *Nat. Chem. Biol* **2010**, *6*, 713-723.