Chapter from the book *Tandem Mass Spectrometry - Applications and Principles*

Downloaded from: http://www.intechopen.com/books/tandem-mass-spectrometry-applications-and-principles

Interested in publishing with IntechOpen? Contact us at book.department@intechopen.com
Tandem Mass Spectrometry and Glycoproteins

Pavlina Dolashka

Institute of Organic Chemistry with Centre of Phytochemistry, Bulgarian Academy of Sciences, Sofia, Bulgaria

1. Introduction

Glycoproteins and glycopeptides are organic compounds that are composed of both polypeptide and carbohydrate chains bonded together. For many years glycoproteins and glycopeptides have been a subject of interest; however, recently they have aroused the interest of biochemists and biologists from a wide range of fields. This increased interest is partly due to the fact that glycoproteins were discovered to be abundant in living organisms and glycoproteins appear in nearly every biological process studied. Glycoproteins help many systems within the human body to function properly and optimally, and deficiencies can be responsible for a whole spectrum of diseases, conditions and ailments. Examples of glycoproteins include antibodies that interact directly with antigens. Major histocompatibility complex molecules that interact with the T-cells as part of an adaptive immune response and hemocyanins (Hcs) fudge molecules that transport oxygen in mollusks and arthropods.

Detailed knowledge of protein glycosylation at the proteome level is becoming an important aspect of post-genomic research. Moreover, glycobiology seeks to identify the molecular structure of glycopeptides and to further explore the function of such peptides in relation to other cells and molecules in the body. By determining how glycopeptides are structured and in order to better understand how they work, researchers working in the field of glycobiology may be able to produce treatments and therapies that improve health and may prolong life. Therefore, the development and application of different analysis techniques will increase the knowledge of their structure and function. A number of reviews have been published in the last few years on analytical methods, including chromatography, electrophoresis and MS, for the characterization of glycans and glycoproteins (Zaia, 2004; Morelle et al. 2005; Dell and Morris 2010), and on general MS-based proteome and peptide analysis methods (Medzhiradszky, 2005; Domon et al. 2006; Froehlich et al. 2011).

Mass spectrometric (MS) techniques play a key role in glycoprotein and glycan analysis, to study protein glycosylation at the glycopeptide level. Therefore, MS is becoming an increasingly important aspect in proteomics. Current informatics tools are designed for large, high-throughput mass-spectrometry datasets.

One of the analytical instruments used in laboratories a tandem mass spectrometer. This instrument can analyze numerous compounds, such as those in body fluids and in the environment. Enrichment and separation techniques for glycoprotein and glycopeptide from complex (glyco-) protein mixtures and digests are summarized below. In addition to detection by mass spectrometry, the microarray platform has also become an essential tool...
to characterize glycan structures and to study glycosylation-related biological interactions; here one uses probes as a means to interrogate the spotted or captured glycosylated molecules on the arrays.

2. Glycoproteomics

Almost all secreted and membrane-associated proteins are glycosylated through post-translational modification. Protein glycosylation is one of the most important post-translational modifications in eukaryotes, but remains poorly investigated. N-linked glycans, with a common monosaccharide core and specific attachment motif, appear to be amenable to analysis, but significant difficulties remain.

The key structural issues of glycoproteomics are protein identification, glycosylation site determination, and glycan profiling at individual attachment sites (Tadjiri et al. 2005; Harvey, 2006). The structure of glycoproteins and glycopeptides is composed of a peptide chain with one or more carbohydrate moieties which constitute from less than 1% to more than 80% of the total protein mass. Glycoproteins usually exist as complex mixtures of glycosylated variants (glycoforms). Glycosylation occurs in the endoplasmic reticulum (ER) and Golgi compartments of the cell and involves a complex series of reactions catalyzed by membrane-bound glycosyltransferases and glycosidases.

Based on the saccharide chains, referred to as glycans, two main classes of glycoprotein structures are known: N-glycosylation, in which the oligosaccharide is attached to an asparagine residue, and O-glycosylation, in which the oligosaccharide is attached to a serine or threonine residue. The O-linked glycans consist of N-acetylglactosamine attached to the O-terminus of a threonine (Thr) or serine (Ser) residue (Fig. 1). The most common type of O-linked glycans contain an initial GalNAc residue (or Tn epitope); they are commonly referred to as mucin-type glycans. Other O-linked glycans include glucosamine, xylose, galactose, fucose, or mannose as the initial sugar bound to the Ser/Thr residues.

![Fig. 1. O-linked glycan contain terminal N-acetyl neuraminic acid attached to a threonine residue.](image-url)
The other class of glycoproteins are the N-linked glycans. These molecules consist of an N-acetylglucosamine bond to the amide nitrogen of an asparagine molecule, where X can be any amino acid (-X-Asn-X-Thr-) (Fig. 2A, B).

Because carbohydrates and proteins by themselves serve in a vast number of biological functions, such as structure, enzymes, protection, carriers, immunological, defense, inhibitors, reproduction, etc. several techniques and analytical methods including chromatography, electrophoresis and MS-based proteome and peptide analysis are applied for the characterization of glycans and glycoproteins (Wada 2008; Sandra et al. 2004, 2007).

Another of these methods is nuclear magnetic resonance spectroscopy (NMR). The method is used to obtain information about the structure and dynamics of proteins and glycans. However, mass spectrometry has several advantages over NMR with respect to analysis of H/D exchange reactions: much less material is needed, the concentration of protein or glycan can be very low (as low as 0.1 uM), the size limit is much greater, and data can usually be collected and interpreted much more quickly.

A very important advantage is that mass spectrometry offers different techniques and approaches. Tandem mass spectrometry of the glycopeptides, isolated from tryptic digests of glycoproteins can be performed in a rapid and sensitive manner. Electrospray ionization (ESI) or matrix-assisted laser desorption/ionization (MALDI) mass spectrometry can be used. In the case of MALDI, profiling of oligosaccharides is achieved more appropriately by the linear time-of-flight (TOF) mode than in the reflection mode. Robust and reliable identification of proteins and the determination of the attachment sites often requires multiple-stage tandem mass spectrometry (Wada 2008).

Combining gel and capillary electrophoresis, nano-LC and mass spectrometry, the elucidation of post-translational modifications of *Trichoderma reesei* cellobiohydrolase I (Sandra et al. 2004 a) of hemocyanin isolated from marine snails *Rapana venosa* (Sandra et al. 2007, Dolashka-Angelova et al. 2004) and of superoxide dismutase (Cu/Zn-SOD) from *K. marxianus* NBIMCC 1984 yeast (Dolashka et al. 2010) were analysed.

The strategy used for the characterisation of RvH1 N-glycosylation combines two main approaches (Fig. 3), where the intact RvH1 was subjected to PNGase F digestion and/or tryptic digestion. The first approach includes reduction and alkylation of RvH1 followed by PNGase F digestion of oligosaccharides from the protein. The N-glycans were analysed by MALDI-TOF and CE-MS and followed by 8-aminopyrene-1,3,6-trisulphonate (APTS) and 3-aminopyrazole (3-AP) labelling. Derivatisation of the oligosaccharides with APTS permits high-resolution CE, allows the simultaneous detection of uncharged and charged glycans and provides easily interpretable spectra.

![Fig. 2. A) Structures of a glycan at m/z 2393[M+Na]+ determined in matrix 2,5-dihydroxybenzoic acid (DHB) B). Structure of an acidic glycan at m/z 1932 [M-H]- determined in matrix 6-aza-2-thiothymine and 2,4,6-trihydroxyacetophenone (THAP).](image)
Fig. 3. Strategy used for the characterisation of the oligosaccharide structure of RvH₁.

The second approach includes reduction and alkylation followed by tryptic digestion of the glycoprotein. The N-glycopeptides were analysed by MALDI-TOF and nano-LC-MS or nano-LC-ESI- techniques. Since the protein sequence of RvH₁ is currently unknown, de novo MS sequencing had to be performed on the glycopeptides. Proteomic techniques, such as HPLC coupled to tandem mass spectrometry (LC-MS/MS), have proven to be useful for the identification of specific glycosylation sites of glycoproteins (glycoproteomics). However, glycosylation sites of glycopeptides produced by trypsinization of complex glycoprotein mixtures are particularly difficult to identify because glycopeptides are usually present in relatively low abundance (2% to 5%) in peptide mixtures compared to the non-glycosylated peptides, and because the sugar fragments in the MS spectrum often dominate the peptide fragments due to glycosidic bonds being more labile than peptide bonds. Some approaches have been developed to overcome this problem, mainly based on multistage MS (Bateman et al. 1998, Demelbauer et al. 2004), electron capture dissociation (ECD) (Håkansson et al. 2001, 2003) or electron transfer dissociation (ETD) (Hogan et al. 2005).

3. Tandem mass spectrometry technique for glycoprotein and glycopeptide analysis

An important application using tandem mass spectrometry is in protein identification. The simplest form of the technique combines two mass spectrometers. One of these instruments
is a Matrix–Assisted–Laser–Desorption/Ionization–Time–of–Flight–Mass–Spectrometer (MALDI-TOF/TOF-MS), where the first mass spectrometer is used to select a single (precursor) mass from the MS spectrum of a mixture of the precursor. After collisional activation (CA) or collision-induced dissociation (CID), the second mass spectrometer is used to separate the fragments ions according to their masses. The resulting “MS/MS” spectrum consists only of production ions from the selected precursor (Fig.4).

![Figure 4. Matrix-Assisted-Laser-Desorption/Ionization–Time-Of-Flight-Mass-Spectrometer (MALDI-TOF/TOF-MS).](image)

There are various methods for fragmenting molecules by tandem MS, including:
- collision-induced dissociation (CID),
- electron capture dissociation (ECD),
- electron transfer dissociation (ETD),
- infrared multiphoton dissociation (IRMPD),
- blackbody infrared radiative dissociation (BIRD),
- electron-detachment dissociation (EDD) and
- surface-induced dissociation (SID).

Tandem mass spectrometry of glycopeptides is known as one of the most important tools in structural glycoproteomics. Various tandem MS (MS/MS) techniques for the analysis of glycopeptides as MALDI, MALDI–TOF/TOF–MS or MALDI–quadrupole–TOF were applied and compared with respect to the information they provide on peptide sequence, glycan attachment site and glycan structure. Glycopeptide ionization was performed mainly by CID or ETD.

### 3.1 Collision-induced dissociation

Early experiments with ESI and collision-induced dissociation (CID) on a triple-quadrupole mass spectrometer have already established several of the key features of CID of glycopeptides (Huddleston et al. 1993; Medzihradszky et al. 1996). On the basis of this pioneering work, ESI with CID of glycopeptides has become a key tool in glycoproteomics. The potential of nano-ESI with a quadrupole–TOF mass analyzer for the characterization of N- and O-glycopeptides has been shown to be a sensitive tool that provides information on glycan structure, glycan attachment site, and peptide sequence. This method has been successfully applied to the characterization of O-glycosylated peptides carrying the Tn-
antigen (GalNAcβ1-), the T-antigen disaccharide, or other slightly more elongated O-glycans based on β-linked GalNAc, attached to serine or threonine residues (Chalabi et al. 2006). Another type of glycosylations, analyzed by nano-ESI–quadrupole–TOF are O-fucosylation (Macek et al. 2001), O-linked N-acetylglucosamine (Vosseller et al. 2006), as well as C-mannosylation (Gonzales de Peredo et al. 2002). The C-linked mannose appeared to be very stable in CID, in contrast to O-glycans and N-glycans.

3.2 Electron-transfer dissociation (ETD)
Similar to the peptide structural information obtained from CD, electron-transfer dissociation has recently emerged as an MS/MS technique complementary to CID and RMPD. Peptide fragmentation is generated through gas-phase electron-transfer reactions from singly charged anions to multiply charged protonated peptides. Singly charged anions are used as vehicle for the electron delivery to the multiply protonated peptides. Analogous to ECD, dissociation from electron transfer results in peptide backbone fragmentation into c- and z•-type ions. Therefore this fragmentation is more useful for the analysis of posttranslational modification (PTM) such as phosphorylation (Syka et al. 2004) and glycosylation (Hogan et al. 2005). The feature makes this technique, together with ECD, a very attractive tool for the localization of the PTM attachment.

4. Sample preparation and characterization
The methods to study glycoproteins, glycopeptides or glycans in MS-based analyzes vary according to the specific research question. In many cases, these methods have been developed with the aim to selectively obtain the N-glycoproteome of a particular sample by MS analysis of the corresponding (deglycosylated) tryptic peptides to identify the underlying proteins. In these types of studies, N-glycosylation sites in tryptic glycopeptides usually are identified by conventional LC–MS/MS or MALDI–MS/MS analysis based on the conversion of Asn to Asp upon treatment with N-glycanase (PNGase), or on the localization of the remaining GlcNAc–Asn tag upon treatment with endo-N-acetylglucosaminidases. For the isolation of glycoproteins or glycopeptides by affinity chromatography various lectins have been used so far. Lectin chromatography using concanavalin A (Con A) has been reported for the enrichment of N-glycoproteins from diverse sources (Nasia et al. 2009). Enrichment techniques applied in combination with advanced MS/MS methods for the direct analysis of intact glycopeptides to obtain sequence information of both the glycan and the peptide moiety have been less commonly used. Now approaches using fragmentation techniques in glycopeptide and glycan analyses such as ESI, MALDI and LC/MS/MS–Q-Trap, are very popular to provide information on the peptide and glycan sequences, as well as on the attachment site.

4.1 Characterisation of glycoproteins and glycopeptides
Several methods and techniques have been applied to analyse the oligosaccharide structure of glycoproteins and glycopeptides. Applying both CID and ETD fragmentation techniques sequentially to protonated glycopeptides provides information on the glycan structure (CID) as well as the provide information on the peptide sequence and the glycan attachment site (ETD). The combination of these complementary data sets allows the detailed structural
characterization of glycopeptides species characterization of glycopeptides is performed with different instrumental configurations, such as:
- MALDI-MS/MS
- Liquid chromatography systems compatible with on-line MS of glycopeptides
- ESI-MS/MS
- Capillary electrophoresis with on-line MS

4.1.1 MALDI–MS and MS/MS analyses of glycopeptides
MALDI–MS/MS of glycopeptides has been performed using the following instrumental configurations:
- MALDI–TOF with post source decay (PSD),
- MALDI–TOF/TOF,
- MALDI–quadrupole–TOF, and
- MALDI–IT/TOF MS.

MALDI–TOF/TOF MS of glycopeptides in protonated form has been established using 2,5-dihydroxybenzoic acid (DHB) as a matrix (Uematsu et al. 2005; Wuhrer et al. 2004); the observed fragments result in the MS spectrum. Three different groups of fragment ion signals in MS are observed which provide information on both the peptide and glycan moiety of the glycopeptides. MALDI–TOF/TOF–MS of N-glycopeptides results in all the fragment ions retaining the peptide moiety, as well as in a set of cleavages at or near the innermost N-acetylglucosamine residue. The signals [Mpep+H]+ and [Mpep+H-17]+, usually arising from the cleavage of the side-chain amide bond of the glycosylated asparagine are observed. Fragmentation characteristics are very similar to those observed with MALDI–quadrupole /TOF–MS (Krokhin et al. 2005; Bykova et al. 2006) and MALDI–IT/TOF–MS (Demelbauer et al. 2004; Takemori et al. 2006).

4.1.1.1 MALDI–MS and MS/MS analyses of glycosylation sites and determination of the occupancy at a particular site
Determination of which sites in the glycoprotein are glycosylated and determination of the extent of occupation at each site are generally accomplished by performing trypsinolysis or other degradative reactions, generating a peptide/glycopeptide mixture. Glycopeptides can usually be selectively detected in such mixtures by precursor ion scanning, followed by MS/MS analysis of detected glycopeptides. This approach, however, is only straight forward for proteins with known amino acid sequence. This restriction, which originates from the dominance of the sugar fragments, can be overcome by performing MS/MS or MS3 on the Y1 ion (peptide + GlcNAc) generated via in-source fragmentation or MS/MS, respectively (Sandra et al. 2007).

Since the protein sequence of one subunit of *Rapana venosa* hemocyanin (RvH1) is currently unknown, this approach was not applicable. Therefore, de novo MS sequencing was performed on the glycopeptides, obtained after tryptic digestion of RvH1 (Sandra et al. 2007). The glyco-moiety was removed from the glycopeptide because the sugar fragments often dominate the peptide fragments due to glycosidic bonds being more labile than peptide bonds. This is a phenomenon typically observed when using CID or PSD as fragmentation techniques. To solve this problem electron capture dissociation or electron transfer dissociation were applied. A method of labelling the N-glycosylation sites was used by performing a PNGase F digestion in a buffer containing 50% H218O (Sandra et al. 2004a). The 18O-labelling step can be performed either prior or subsequent to trypsinolysis. When
choosing this approach, it is important to remove the remaining H$_2^{18}$O to prevent unspecific incorporation of the label into all tryptic peptides. Also, trypsin needs to be removed from the hydrolysate. Since H$_2^{18}$O can easily be removed by membrane filtration, labelling the protein prior to the tryptic digestion was preferred. The complexity of the tryptic peptide mixture is reflected in the nano-LC-UV chromatogram. By way of example the MALDI-TOF spectrum corresponding to fraction 54, and an expanded view of the region containing the $^{18}$O-labelling peptide ions at m/z 2406.3296, are presented in Fig. 5. Sequence information can be obtained by performing higher-order MS. The peptides containing the glycosylation sites were detected via the 2 Da spacing between the unlabelled and labelled ions using MS (Fig.5).

![Fig. 5. The nano-LC-MALDI-TOF spectra corresponding to fraction 54 (b) and an expanded view of the regions containing the $^{18}$O-labelling peptide ions at m/z 2406.3296 (inset).](image)

Using this approach, 6 glycopeptides were identified from the enormous complexity of the RvH$_1$ tryptic digest. Nano-LC was used as a preceding separation step, and fractions were directly collected onto a MALDI-target.

### 4.1.2 Characterisation of glycopeptides using MALDI–MS/MS and Q-Trap MS/MS analyses

The structure Man$^\text{GlcNAc}_2$ of the glycan with mass of 1257.4 Da is represented with a specific fragmentation nomenclature in Q-Trap MS/MS spectrum. The MS/MS spectrum of the underivatised singly charged sodium-adduct is shown in Figure 6. The most dominant ions are Y and B that arise from glycosidic cleavages. C and Z ions are also observed. They can be differentiated from the Y and B ions without derivatisation, because of the asymmetrical nature of the molecule. At this stage it is impossible to demonstrate the existence of a branched structure. Many interesting cross-ring cleavage ions are also present as 0,A$^3$, 2,A$^4$, 0,3,A$^3$, and 3,5,A$^3$ ions rendue. They are very informative because they indicate a linkage of a trihexose at carbon 6 of a hexose. When the structure is not methylated, minor information can be extracted on the branching from the cross-ring cleavages.
Fig. 6. Structure of the RNase B. Man ¤ GlcNAc2 N-glycan and fragmentation nomenclature.

MALDI-MS/MS and Q-Trap MS/MS have been performed to analyse the carbohydrate structure in the polypeptide chain of Cu/Zn-SOD from *K. marxianus* NBIMCC 1984 yeast (Dolashka-Angelova et al. 2010). The obtained fractions after treatment of the protein with trypsin were separated by HPLC and their amino acid sequences were determined by MALDI-TOF-TOF. One putative linkage site was observed in the sequence [EVWN(I/L)TGNSPNA(I/L)R] of the peptide with a mass of 1773.51 Da. The orcinol/H2SO4 test was effectively positive, confirming that one glycopeptide is present in a fraction eluted at 19 min by HPLC. Structure of this glycopeptide was analysed by MALDI–MS/MS and Q-Trap MS/MS.

The amino acid sequence of the peptide chain EVWN(I/L)TGNSPNA(I/L)R was determined by MALDI-MS/MS, based on the singly-charged ions (Fig. 7). The difference between two single ions at m/z 1571.01 and m/z 1773.51 is corresponding to one GlcNAc connected to the peptide, which also demonstrates that the peptide of 1773.51 Da is glycosylated. The ion at 1359.23 is the same peptide as y11 [at m/z 1156.74, N(I/L)TGNSPNA(I/L)R], with one GlcNAc being connected to linkage site Asp-Ile/Leu-Thr. Evidently the linkage site (-Asn-Leu/Ile-Thr-) at position 33-35 is the glycosylated one. This linkage site is conserved in several SODs.
Fig. 7. MALDI-TOF-MS spectra of the N-glycan isolated from Cu/Zn-SOD from *K. marxianus* NBIMCC 1984. 1 µl of a 1:1 sugar-matrix mixture was applied onto the MALDI target. A matrix DHB (10 mg/ml dihydroxybenzoic acid solution in 50% AcN) was used.

Fig. 8. MS/MS spectra on a Q-Trap mass spectrometer and structure with fragmentation nomenclature of the single charged [M+Na]^+ of the glycan at m/z 1235.52, isolated from Cu/Zn-SOD from *K. marxianus* NBIMCC 1984.

After treatment of the peptide with PNGase F a single peak at 1257.3 [M+Na]^+ was detected which suggests a uniform oligosaccharide chain. The structure of this carbohydrate chain was determined by Q-Trap MS/MS of the singly-charged ion with a mass of 1235.52 Da (M+H)^+ as shown in Figure 8. The sequence can easily be read when considering the Y ions and the combination of B and Y ions (m/z 222.1, 325.1, 425.1, 811.2, 1014.3). The structure,
given as inset in the figure is a classical high mannose type of sugar (GlcNAc2, Man5), with a calculated mass of 1234.4 Da.

### 4.1.3 Electrospray ionization–mass spectrometry (ESI-MS) of glycopeptides

Electrospray is a soft ionization technique and can be performed on solid or liquid samples. Therefore, ESI is typically used to determine the molecular weights of proteins, peptides, and other biological macromolecules. Soft ionization is a useful technique when considering biological molecules of large molecular mass, as it turns the macromolecule being ionized into small fragments.

**Fig. 9.** Fragment spectra of the glycopeptide G1 with determined m/z 1338.16 [M+2H]^{2+}. A) MS/MS analysis of the sugar moiety using a collision energy of 37 eV. Annotation of sugar fragments. The insert shows the complete carbohydrate structure and the cleavage points leading to the respective fragments. * - corresponds to the doubly charged ions B) Pseudo-MS/MS/MS experiment (in-source and collision-induced fragmentation) of the peptide moiety (YEXHAVN*GSTXAAAX) still carrying one HexNAc (m/z 1662.09, [M - H]^{+}), using a cone voltage of 95 V and a collision energy of 80 eV. R-corresponds to the peptide.

ESI and collision-induced dissociation (CID) on a triple-quadrupole mass spectrometer have already established several of the key features of CID of glycopeptides. ESI-MS was applied to analyse the glycopeptides obtained after treatment of Rapana hemocyanin with trypsin and being separated by HPLC (Beck et al. 2007; Sandra et al. 2007).

Six glycopeptides were characterized in detail by analyzing the corresponding HPLC fraction by ESI-MS/MS. The MS/MS spectra of one peptide appearing as a doubly protonated molecule of m/z 1338.16 [M+2H]^{2+} in the MS spectrum itself (not shown) are displayed in Figure 9.
The MS/MS spectrum of the fragment ion at m/z 1661.71, corresponding to the peptide (designated R-, m/z 1459.75), which is N-glycosylated, with a single residual GlcNAc residue, is shown on Figure 9. The spectrum is further dominated by glycan fragments which correspond to a classical mannose type of oligosaccharide, consisting of two N-acetylhexosamines (HexNAc) monomers as well as five hexose (Hex) residues. In a pseudo-MS experiment, the sugar side chain was removed by in-source fragmentation and the peptide with one remaining HexNAc residue was further fragmented in the hexapole collision cell. This allowed the sequence analysis of the peptide and the determination of the site to which the sugar chain was attached (Fig. 9, insert).

The amino acid sequence of the peptide was revealed to be YEXHAVN*GSTXAAX. The glycosylation site N* was identified to be part of the typical N-glycosylation motif N-aa-S/T, where aa can be any amino acid except proline. X represents a leucine or isoleucine residue which could not be distinguished by the applied MS methods.

4.1.4 Characterisation of glycopeptides using a Q-Trap LC/MS/MS system

As was described in 4.1.2 and 4.1.3, MS analysis of glycopeptides may be performed after extensive purification, using HPLC fractions analyzed by flow injection analysis on a Q-TOF or ESI-mass spectrometer. In practice, characterisation of glycopeptides is very difficult, because they are often analyzed from complex peptide/glycopeptide mixtures. When such samples are subjected to LC-ESI-MS/MS analysis with collision-induced fragmentation, data evaluation methods which highlight the relevant glycopeptide MS data within the complex overall data set are required. Strategies applied to achieve this goal comprise the generation of diagnostic fragment ions in the MS-mode (without precursor selection) and/or MS/MS-mode (with precursor selection).

The Q-Trap system with its capabilities to perform typical triple quadrupole scans was additionally used and several glycopeptides from RvH1 were identified and characterized (Sandra et al. 2007). Glycopeptides, selectively detected in a proteolytic mixture by the appearance of collision induced marker oxonium ions, such as m/z 163 (Hex⁺), 204 (HexNAc⁺) or 366 (HexHexNAc⁺), were sequenced. The insert of Figure 9A shows the LC/MS/MS total ion current (TIC) chromatogram of the precursor ion scan (monitoring m/z 204) of the HPLC fraction at time 31.24 min.

The Enhanced Resolution scan (not shown) showed that the glycopeptide with mass 2511.91 eluting at this moment was triply charged at m/z 837.97 [M+3H]⁺. The precursor ion scan at time 31.24 min is presented in Figure 10 (insert). The MS/MS spectrum is dominated by glycan fragmentation series of Y- and B-ions, according to the Domon/Costello nomenclature.

However, peptide fragmentation (Roepstorff/Biemann cleavages) became more dominant when the collision energy was increased, allowing to deduce the peptide sequence MGQYGD(I/L)STNTR from the series of y- and b-ions (Fig. 10). The ion b13 (m/z 1439.6) or y13 (m/z 1457.5) correspond to the peptide which contains two potential linkage sites –D(L/I)S- and –NNT-.

Normally, D(L/I)S is not expected to be a linkage site, but, in the unlikely event that deglycosylation of a glycan linked asparagine had occurred during sample preparation, we thoroughly investigated this option. The ion y7 at m/z 806.4 corresponds to the C-terminal fragment of the peptide -(I/L)STNTR, and the ion at m/z 1009.6 represents the same fragment still containing one GlcNAc moiety. This suggests that only the linkage site –NNT- is glycosylated, most likely via a high mannose like structure.
Fig. 10. Precursor ion scan (insert) at time 31.24 min of the chromatographic separation, and enhanced product ion (EPI) scan of the ion (B) at m/z 837.97.

The ion at m/z 1660.9 corresponds to the intact peptide, represented as ion y13 (m/z 1457.5), which is N-glycosylated with a single GlcNAc residue. The glycan structure of this peptide was determined following the typical ions as Y and B in MS/MS spectrum at m/z 204.1 (GlcNAc), m/z 366.2 (HexGlcNAc), m/z 528.3 (Hex2GlcNAc), 690.5 (Hex3GlcNAc), and 852.4 (Hex4GlcNAc) (Fig. 10).

Combining all data, the carbohydrate structure of the glycan with mass 1054.0 Da (Hex2Man3GlcNAc2) was suggested and the peptide sequence was determined MGQYGD(I/L)STNNTR.

4.2 Structural characterization of glycans by MS/MS analyses

While CID MS/MS techniques do routinely provide information on the glycan moieties of glycopeptides, they rarely reveal information on the peptide sequence and glycan attachment site(s). For a more detailed characterization of protein glycosylation, these techniques may be combined with other types of experiments. Glycopeptides may be treated with exoglycosidases revealing the nature and anomericity of terminal monosaccharide residues.

Alternatively, by treatment of glycopeptides with peptide-N-glycosidase F or A, N-glycans can be released and deglycosylated peptide moieties can be obtained. The peptides may then be subjected to mass spectrometric characterization.

Some glycans, obtained by us after digestion of proteins with PNGase F were identified by Q-Trap analysis. Analysis of the purified N-glycans by Q-Trap is a very sensitive and useful method and several glycans of the high mannose or complex types were identified following the sequence of B and Y ions in the MS/MS spectrum.

Two approaches were applied to analyse the isolated glycans after treatment of RvH1 with PNGase F (Dolashka-Angelova et al. 2010). The first approach included sequencing of the glycans by specific glycosidases and analysis of the fragments via MS before and after
treatment with the enzymes. This approach provided only preliminary results about the structures of the glycans. Therefore, in the second approach, tandem mass spectrometry was applied. The glycan structure was derived from the MS/MS spectra, obtained on a hybrid quadrupole-linear ion trap mass spectrometer.

The same method and techniques were applied and the configurations of 15 N-glycans released from *Haliothix tuberculata* hemocyanin (HtH1) after PNGase F-treatment were performed by Q-Trap tandem mass spectrometry (Velkova et al. 2011). Using both the MALDI-TOF-MS analysis after treatment of the glycans with the specific α1-6(>2,3,4) fucosidase and the Q-Trap-MS/MS analysis, core-structures containing Fuc(α1-6) GlcNAc, were detected.

Fig. 11. MS/MS spectra and structures with fragmentation nomenclature of the double charged [M+2Na]$^{2+}$ of the glycan at m/z 1002.8 isolated from HtH1.

In the MS/MS spectrum the most dominant Y and B ions provided information on the sequence and the branching, but the positions of the monosaccharides were confirmed by the C, Z, X and A ions. The interesting structure of one glycan was determined after sequencing the doubly-charged ion ([M+2Na]$^{2+}$, m/z 1002.8, (Figure 11). As is shown in the spectrum, two deoxyhexose and two terminal methyl-Hex residues are linked to the internal GlcNAc of a molecule with the composition MeHex$_2$HexMan$_3$GlcNAc$_4$Fuc$_2$. Three ions, Z$_2$ at m/z 534.8, Z$_4$ at m/z 575.2 and Y$_2$ at m/z 593.0, demonstrate the presence of a core-linked Fuc(α1-6) residue. Confirmation of additional Fuc(α1-3) branching to GlcNAc are the ions B$_{2a}$ at m/z 548.0 and B$_{3a}$ at m/z 710.0, as well as the ions B$_3$ at m/z 564.1 and B$_5$Y$_5$ at m/z 1469.9, supporting the branching of two terminal MeHex residues. The observed cross-ring fragment ions X$_4$B$_4$ at m/z 767.8, as well as δ$_{3A_2a}$, are additional evidence of the suggested structure Fuc(α1-3)GlcNAc. The ion Y$_4$B$_4$ at m/z 1034.1 corresponds to the composition Man$_3$GlcNAcFucMeHex (carrying a (α1-3)-fucose terminal linkage at GlcNAc).
Consequently, cross-ring fragment ion \( X_{4a} Y_{4b} B_{4} \), at m/z 767.8, is resulting from the ion \( Y_{4b} B_{4} \) at m/z 1034.1 without \( o_{3}A_{2a} \) fragment, containing a MeHex moiety. The evidence of MeHex being linked to GlcNAc, as observed in most molluscan hemocyanins, is derived from the ions \( C_{1a} \) and \( C_{1b} \) at m/z 217.1, ions \( B_{3b} \) and \( B_{2a} Y_{5Y} \) at m/z 402.1, as well as from the ions \( B_{3b} \) and \( B_{2a} Y_{5Y} \) at m/z 564.1. Two alternatively interpretations of the signal at m/z 1294 as \( Y_{5Y} Y_{5Y} B_{5} \) and \( Z_{4a} Y_{1} \) confirm the positions of two fucoses.

Based on these data the structures of the isolated N-glycans from HtH1 were proposed (Velkova et al. 2011). It was found that most of the glycans have a common structural feature, one \( a_{1-6} \) linked fucose being attached to the trimannosyl core. Some of the structures have being partially modified by the methyl group, and in a few glycans a second or a third fucosyl residue in the Fuc(\( a_{1-3} \))GlcNAc motif were identified.

4.3 Capillary electrophoresis and mass spectrometry

4.3.1 Characterization of glycopeptides by CE

Capillary electrophoresis (CE) and capillary gel electrophoresis have been widely used for complex carbohydrate separation (Sandra et al. 2004b) because of enhanced separation efficiency and shorter analysis times. CE-based carbohydrate analysis can be applied easily to determine the molar ratio, and the degree of polymerisation of oligosaccharides, and to detect changes in the extent or nature of the oligosaccharide distribution (fingerprinting).

Using capillary electrophoresis, MALDI-MS and ESI-MS in combination with glycosidase digestion we studied the N-terminal functional unit (FU) RvH1-a of the structural subunit RvH1 (Dolashka-Angelova et al. 2004)(Fig. 12).

![Fig. 12. ESI-MS and capillary electrophoresis of Glp1 following glycosidase digestion.](www.intechopen.com)
Oligosaccharide fragments were released from the glycoprotein by Smith degradation and separated by means of a Superdex 300 column. Glycopeptide fragments, giving a positive reaction for the orcinol/H$_2$SO$_4$ method, were separated by HPLC. The glycopeptide 1 with mass 2786 Da, was analyzed by MALDI-MS before and after treatment with PNGase-F. In the MALDI mass spectrum, acquired before PNGase-F treatment only one molecular ion at 2786 m/z was detected, generated by the intact Glp1 sodium adduct ion [Glp1+Na]$^+$. In contrast, three signals were observed after PNGas-F treatment. The signal at 2786 m/z is related to the intact Glp1, whereas the signal at m/z 1609 [M+Na]$^+$ is attributed to an oligosaccharide. The first signal (1177 m/z) corresponds to the deglycosylated peptide. To determine the oligosaccharide sequence, ESI-MS and capillary electrophoresis of Glp1 were used, following glycosidase digestion (Fig. 12).

Different glycolytic enzymes were added to glycopeptide 1: $\beta$1-2,3,4,6-GlcNAcase; $\alpha$1-2,3-mannosidase, $\alpha$1-2,3,6- mannosidase, $\beta$1-3,4,6-galactosidase. The mass of this carbohydrate moiety (1586 Da, Table 1) would account for an oligosaccharidic chain containing (SO$_4$)$^-$ MeGal GlcNAc$^-$ Man$^-$ being connected to the peptide.

<table>
<thead>
<tr>
<th>No</th>
<th>Enzymes</th>
<th>[M+H]$^+$ (m/z)</th>
<th>Structure of Glycopeptide 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$\alpha$1-2,3- and $\alpha$1-2,3,6-mannosidase and $\beta$1-2,3,4,6-N-GlcNAcase</td>
<td>1747</td>
<td>$\leftarrow$ Man((\beta1)-4) GlcNAc((\beta1)-4) GlcNAc →-P</td>
</tr>
<tr>
<td>B</td>
<td>$\alpha$1-2,3-mannosidase and $\beta$1-2,3,4,6-N-GlcNAcase</td>
<td>1908</td>
<td>Man((\alpha1)-6) $\Leftarrow$ Man((\beta1)-4) GlcNAc((\beta1)-4) GlcNAc →-P</td>
</tr>
<tr>
<td>C</td>
<td>$\alpha$1-2,3,6-mannosidase and $\alpha$1-2,3-mannosidase</td>
<td>1953</td>
<td>3MeGlcNAc((\beta1)-2)$\Leftarrow$Man((\beta1)-4)GlcNAc((\beta1)-4)GlcNAc→-P</td>
</tr>
<tr>
<td>D</td>
<td>$\beta$1-2,3,4,6-N-GlcNAcase</td>
<td>1583</td>
<td>-GlcNAc((\beta1)-4) GlcNAc→-P</td>
</tr>
</tbody>
</table>

Table 1. Proposed structures of the carbohydrate chains of Glp1. The carbohydrate chains were calculated on the basis of observed [M +H]$^+$ signals in the ESI mass spectra from Fig.12. P represents the peptide with sequence FANATSIDGPNA (1177 Da).

Depending on the specificity of the individual enzymes, the different linkages were identified by recording the pattern of molecular weights resulting from the digestion with the pool of glycosidases. After a 24 h incubation at 37°C, the sample was first analyzed by capillary electrophoresis. Four different peaks were detected, indicating that the treatment causes heterogeneity resulting from different cleavage sites (Fig. 12, insert). On the basis of these results and an amino acid sequence analysis it was concluded that the functional unit RvHi-a contains 7% of oligosaccharides, that the glycans are N-glycosidically attached to Asn262 and Asn401, and that the structure is as follows:

\[
\text{Glp 1} \quad \text{SO Man(\(\alpha1\)-6)} \quad \text{\(\Leftarrow\text{Man(\(\beta1\)-4) GlcNAc(\(\beta1\)-4) GlcNAc \quad \text{\(\Leftarrow\text{FANATSIDGPNA}}\right)

3\text{MeGal(\(\beta1\)-3) GlcNAc(\(\beta1\)-2) Man(\(\alpha1\)-3)}
\]

www.intechopen.com
Using capillary electrophoresis, in combination with MALDI-MS, ESI-MS and Q-Trap-MS/MS (Fig. 13), the carbohydrate structure of several glycoproteins have been identified (Sandra et al. 2004 b; 2007).

Fig. 13. Combination of CE, ESI and Q-Trap MS/MS system.

4.3.2 Characterization of glycans by CE
Labelling of the glycans has recently been proven to be of great value in the CE-MS/MS analysis of glycoproteins (Sandra et al. 2004 b). It permits high-resolution CE, allows the simultaneous detection of uncharged and charged glycans, and provides easily interpretable spectra. The latter label, representing a novel sugar tag, allows CE-MS detection in the positive ion mode and provides complementary information. The intact RvH1 was subjected to peptide-N-glycosidase F (PNGase F) digestion and the obtained glycans were labelled by 8-aminopyrene-1,3,6-trisulfonate (APTS) as well as by 3-aminopyrazole (3-AP). The N-glycans were analyzed by MALDI-TOF and CE-MS/MS.

MS/MS on APTS-derivatised sugars resulted in easily interpretable MS/MS spectra, since Y-frgments predominated. The CE-MS electropherogram of the APTS-labelled RvH1 N-glycans is presented in Figure 14 (insert) and, by way of an example, the on-line MS spectrum of the compound migrating at time 10.0 min is presented. This glycan is detected as a doubly and triply negatively charged ion and the composition corresponds to a fucosylated Man\textsubscript{3}GlcNAc\textsubscript{2} core structure. From the MS/MS data obtained on the [M-2H]\textsuperscript{2-} ion at m/z 748, one can deduce that the fucose is attached to the proximal GlcNAc residue (Fig. 14).

The CE-MS(/MS) is a much more sensitive technique which allows identification of the presence of unusual structures. CE-MS/MS was performed to provide additional information of the oligosaccharide structure of RvH1. A new structure was suggested for the four times negatively charged ion at m/z 555.7 (Sandra et al. 2007). A sulphated oligosaccharide was observed before in RvH\textsubscript{1} \textsuperscript{10} , and for the ion at m/z 555.7 the following sulphated structures can indeed be proposed: SO\textsubscript{4}FucHex\textsubscript{3}Man\textsubscript{3}GlcNAc\textsubscript{2} and (SO\textsubscript{4})\textsubscript{2}FucHexHexNAc\textsubscript{2}Man\textsubscript{3}GlcNAc\textsubscript{2}. 
In order to distinguish these two possibilities, the sample was reinjected for an on-line MS/MS experiment, keeping Q1 fixed to select the ion at m/z 555.7 (Figure 15). Apparently, CE could separate two isomeric compounds (eluted after 10.5 and 10.9 min) with similar tandem mass spectra and, hence, the difference did not reside in the sequence. The spectra do not fit any of the sulfated glycans mentioned earlier. Following MS/MS data, a structure with the HexAHexNACFucGlcNAc oligosaccharide at both the α-1,3 and the α-1,6-arm could be suggested.
5. Conclusion

Protein glycosylation plays an important role in a multitude of biological processes such as cell-cell recognition, growth, differentiation, and cell death. It has been shown that specific glycosylation changes are key in disease progression and can have a diagnostic value for a variety of disease types such as cancer and inflammation. The complexity of carbohydrate structures and their derivatives makes their study a real challenge.

Mass spectrometric (MS) techniques play a key role in glycoprotein and glycan analysis, study protein glycosylation at the glycopeptide level. Tandem mass spectrometry provides both the separation of glycopeptides and the ability to determine the glycan composition and site-specific glycosylation. Current informatics tools are designed for large, high-throughput mass-spectrometry datasets and can analyze numerous compounds, such as those in body fluids. The methods and strategies being developed are compatible with the problems of microheterogeneity commonly found, allowing characterization of even very complex minor components. Therefore, MS is becoming an increasingly important aspect in proteomics of eukaryotic cells.

6. Acknowledgments

This work was supported by a research grant 451 VS.016.009N of FWO (Belgium) and the Bulgarian Academy of Sciences; the Bulgarian National Science Fund TK01-496/2009, HTC01-187 (Ukraine), DFG-01/2008 (Germany) and CNR (Italy). The author is indebted to Prof. Josef Van Beemen (Gent, Belgium) for critically reading and correcting the text.

7. References


Huddleston, M.J.; Bean, M.F.; Carr, S.A. (1993) Collisional fragmentation of glycopeptides by electrospray ionization LC/MS and LC/MS/MS: methods for selective


Tandem Mass Spectrometry - Applications and Principles presents comprehensive coverage of theory, instrumentation and major applications of tandem mass spectrometry. The areas covered range from the analysis of drug metabolites, proteins and complex lipids to clinical diagnosis. This book serves multiple groups of audiences; professional (academic and industry), graduate students and general readers interested in the use of modern mass spectrometry in solving critical questions of chemical and biological sciences.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: