

## BioSimilar 2013: Current analytical techniques for characterization of carbohydrate biosimilars - Parastoo Azadi - Complex Carbohydrate Research Centre

**Parastoo Azadi**

Complex Carbohydrate Research Centre

Production of high-quality pharmaceutical recombinant therapeutic glycoprotein with consistency in glycan quality is still challenging. Since glycans are responsible for bioactivity, solubility, immunogenicity, and clearance rate from circulation, it is vital to have detailed map of glycans in therapeutic glycoproteins. However, due to the enormous diversity of carbohydrate structures and their heterogeneity, this still remains one of the bottlenecks of full structural characterization. Detailed glycoprotein structural analysis has to be able to identify the peptide sequence where the glycans are attached, as well as the structure of the glycan portion, including oligosaccharide sequence and glycosyl linkages. We will detail methods for mass spectrometry (MS) experiments on both released glycans (“glycomics”), as well as on intact glycopeptides (“glycoproteomics”) using EDT, HCD and CID fragmentation pathways that are needed for quantitation and full elucidation of the structure of glycoproteins. Additional data will be shown where a combination of 2D-NMR, glycosyl composition and glycosyl linkage analysis, will provide information on the glycan topology as well as detection methods for potential non-human modifications that could arise from mammalian expression systems such as Gal $\alpha$ 1-3Gal and N-glycolylneuraminic acid (NeuGc). Our consolidated experiments will outline all the necessary information pertaining to the glycoprotein, including glycan fine structure, attachment site, and glycosylation degree to be obtained pharmaceutical recombinant glycoproteins.

### Introduction

Glycoprotein glycans are essential for and regulate many physiological processes during development and disease progression (Hakomori 1989; Muramatsu 1993; Olden 1993 Fukuda 1996). Identifying specific glycan structures, deciphering the proteins that express each glycan, and understanding in detail how these structures change, e.g., as cells differentiate or as tumor cells progress, are components of the emerging field of glycoproteomics. A large number of proteins are involved in regulating glycan expression and function, including glycosyltransferases, glycosidases, other enzymes involved in sugar nucleotide metabolism and transport, as well as carbohydrate-binding proteins known as lectins (Coutinho et al. 2003; Breton et al. 2006). The genes that encode many of these enzymes have been isolated, expressed, and characterized extensively by functional studies, including generating null mice (Taniguchi et al. 2002).

It is estimated that the murine glycome, for example, encodes over 650 genes (Lowe and Marth 2003). A major challenge, therefore, is to determine how glycan structures change during progression, how transcripts of genes in the glycome change as cells initiate differentiation programs, and then to synthesize an understanding of how transcript changes can be used to identify and predict changes in glycan expression. A sensitive, quantitative technique for glycoprotein glycan analysis, therefore, is a critical component, just as quantitative methods have had a significant impact on comparative proteomics.

The requirements of a methodology for quantitative glycan expression comparison are its capability of detecting even subtle changes in structure with high sensitivity and high resolution. Mass spectrometric techniques based on electrospray ionization-mass spectrometry (ESI-MS) or matrix assisted laser desorption ionization with time of flight detection (MALDI-TOF-MS) have found important applications in high-throughput proteomic analyses due to substantial improvements in the instrumentation and the development of computer algorithms that allow the analysis of large amounts of data (Mann et al. 2001). Studies have shown that it is often possible to detect glycans released from glycoproteins using similar MS techniques without derivatization (Keck et al. 2005). However, the composition of native oligosaccharides are varied due to the presence of N-acetylated and acidic residues, and this variation affects the ionization capabilities of the oligosaccharides, sometimes preventing them from being detected by MS. For this reason, derivatization of oligosaccharides by chemical means such as permethylation is often performed before MS analysis, because this chemical modification stabilizes the sialic acid residues in acidic oligosaccharides. Additionally, the addition of methyl groups to an oligosaccharide population also enables them to become more uniformly ionized by reducing the variation of the chemical properties among the glycans, caused by the presence of highly polar –OH and –NH<sub>2</sub> groups in the molecule; therefore, facilitating relative quantitation of the individual oligosaccharides in a mixture. Methylated glycans ionize more efficiently than their native counterparts, and due to their hydrophobic nature, are easily separated from salts and other impurities that may affect the MS analysis (Dell et al. 1994; Kang et al. 2005). Additionally, the fragmentation of methylated glycans is more predictable than that of their native counterparts, leading to accurate structural assignments when MS/MS analysis is

performed (Dell et al. 1994; Viseux et al. 1997; Reinhold and Sheeley, 1998; Sheeley and Reinhold, 1998; Morelle et al. 2004; Mechref et al. 2006; Wuhler and Deelder, 2006).

Several MS procedures, including MALDI-TOF-MS and ESI-MS have been used successfully to analyze methylated glycans with high sensitivity and resolution (Dell et al. 1994; Reinhold and Sheeley 1998; Sheeley and Reinhold 1998). A significant drawback of these techniques, however, is their restricted ability to provide quantitative information due to differences in ionization efficiencies that still remain among the various components present in the sample even after methylation (Chen and He 2005), as well as the possible presence of different amounts of interfering ions among the samples. Although there are some reports on the use of MALDI-TOF MS for quantitative studies of oligosaccharides and protein glycosylation (Pitt and Gorman 1997), there are still important limitations on the use of MS for quantitative purposes.

In the field of proteomics, one means of obtaining quantitative proteomic data from mass spectrometric analyses is through the incorporation of specific isotopic label into a population of molecules, either by metabolic incorporation or chemical derivatization (Goshe and Smith 2003; Tao and Aebersold 2003). In this approach, the sample containing the “heavy” isotope is mixed with that of the “light” isotope, followed by MS analysis of the resulting mixture. The mass analyzer resolves the isotopically labeled species, permitting their relative abundances to be determined from the ratio of the light- and heavy-labeled molecular ions. Numerous isotopic labeling procedures have been established for the study of protein mixtures and are widely used in high throughput proteomic studies (Goshe and Smith 2003).

The use of isotopic labels for the quantitative analysis of glycans offers promise for the detection and measurement of changes in the abundance of specific oligosaccharide structures present in complex glycoprotein mixtures such as those obtained from cells or tissues. To demonstrate this approach, we investigated the use of [ $^{13}\text{C}$ ] methyl iodide ( $^{13}\text{CH}_3\text{-I}$ ) as an isotopic label in permethylation reactions for oligosaccharide analysis by MALDI-TOF MS. Various neutral human milk oligosaccharides, as well as N-linked glycans released from several purified glycoproteins were permethylated, either with  $^{13}\text{CH}_3\text{-I}$  or with  $^{12}\text{CH}_3\text{-I}$ , different proportions of isotopically labeled and nonlabeled glycans were mixed, and each mixture analyzed by MALDI-TOF MS or ESI-MS on a hybrid linear ion trap Fourier transform mass spectrometer (LTQ-FT). Quantitative data were obtained from the peak heights of the different derivatives and analyzed for linearity and reproducibility. The results obtained demonstrated that isotopic labeling with  $^{13}\text{CH}_3\text{-I}$  provides a linear response, with good reproducibility. This technique of labeling oligosaccharides with  $^{13}\text{CH}_3\text{I}$  during permethylation offers the potential for rapid and accurate detection of the relative difference in the abundance of glycan structures during the analysis of glycans from distinct populations such as those from non-diseased and diseased cells or cells at different stages of differentiation