Urinary glycan markers for disease

Dominic S. Alonzi*, Ying-Hsiu Su† and Terry D. Butters*

*Oxford Glycobiology Institute, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QX, U.K. and †Department of Microbiology/Immunology, Drexel University College of Medicine, 3805 Old Easton Road, Doylestown, PA 18901, U.S.A.

Abstract

Robust assays for the isolation and characterization of urinary FOS (free oligosaccharides) have been developed to screen patients for altered protein and/or lipid glycosylation. A FOS analysis can therefore identify potential biomarkers for hepatocellular carcinoma, since variations in glycosylation as a result of tumorigenecity should be detectable in the FOS of patients. HCC (hepatocellular carcinoma) accounts for 80–90% of all liver cancers. It occurs more often in men than women and occurs mostly in people 50–60 years old. The disease is more common in parts of Africa and Asia than in North or South America and Europe. Using a combination of solid-phase extraction techniques and affinity chromatography, followed by separation of urinary FOS by NP (normal phase)-HPLC and HIAX (hydrophilic interaction and anion-exchange)-HPLC, more than 200 different species have been identified in patient samples. The high incidence of small sialylated oligosaccharides in HCC patients suggests that pro-inflammatory markers may be detected as early indicators of disease progression. In addition, the methods developed here to isolate and analyse excreted glycoprotein- and glycosphingolipid-bound oligosaccharides have been used to characterize changes in metabolic processes that underlie a number of human genetic disorders. The ability to predict disease status in microlitre amounts of readily available non-invasive urine samples indicates that rapid methods for screening can be developed.

Introduction

Glycosylation is one of the most common post-translational modifications in eukaryotes, and a majority of cellular proteins have covalently attached oligosaccharides in N- or O-linked forms. Secretory proteins in particular are glycosylated, and the physical properties and biological functions of these macromolecules are significantly influenced by the oligosaccharide moiety. N-glycosylation consists of an oligosaccharide chain N-linked to asparagine in the sequence motif Asn-Xaa-Ser/Thr, where Xaa is any amino acid except proline. N-glycosylation requires the production of an oligosaccharide precursor, which is transferred en bloc to nascent proteins in the ER (endoplasmic reticulum). After the transfer of the oligosaccharide precursor structure to the nascent protein, several subsequent processing reactions occur in the ER, including α-glucosidase I and II, which contribute to protein folding via the calnexin/calreticulin cycle. In addition, N-glycan chains can be further diversified/modified in the Golgi apparatus, with the addition of N-acetylhexosamine, fucose and sialic acid residues [1].

O-glycosylation consists of a glycan linked to the hydroxy group of serine or threonine residues. The frequency of O-glycosylation on glycoproteins is high, particularly on secreted or membrane-bound mucins, which are rich in serine and threonine [1].

Another important class of glycoconjugates is the GSLs (glycosphingolipids). GSLs are diverse and highly complex cellular components localized primarily in the plasma membranes [2]. They are distributed with high specificity between mammalian species, organs, cell types and subcellular domains. GSLs mediate a wide variety of cellular functions, including cellular growth, cell–cell interactions and ion transport [3]. The molecules possess two regions, a lipophilic membrane-localized anchor and a hydrophilic carbohydrate moiety ranging from monosaccharides to complex oligosaccharides. The lipophilic ceramide consists of sphingol substituted at the amino group by a fatty acid in amide linkage. The carbohydrate moiety is linked at the primary hydroxy group of the sphingol in a glycosidic linkage. This unusual combination as a glycoconjugate explains, in part, many of the properties of the GSLs described above.

The search for glyco-markers of disease, i.e. oligosaccharides as indicators of protein and lipid glycosylation, has mainly focused on variation during malignancy [4]. Glycosylation changes at the cell surface are widely observed from early stages of disease to metastasis and secondary growth [5,6]. However, the biological mechanism that causes these alterations is less widely studied, and there appears to be little evidence indicating a dependency of initiation or progression on cell-surface glycan architecture. The changes that occur in malignant cells may involve several different glycosylation pathways. Examples have been found of either

Key words: cancer, fluorescence, free oligosaccharide, glycolipid storage, HPLC.
Abbreviations used: 2-AA, 2-anthranilic acid; AAL, concanavalin A; A; ER, endoplasmic reticulum; ERAD, ER-associated degradation; FOS, free oligosaccharide(s); GSL, glycosphingolipid; GU, glucose unit(s); HCC, hepatocellular carcinoma; HIAX, hydrophilic interaction and anion-exchange; HILIC, hydrophilic interaction chromatography; NB-DNJ, N-butyldenoojirimycin; NP, normal-phase; SNA, Sambucus nigra.

© The Authors Journal compilation © 2011 Biochemical Society
loss of expression or increased expression of certain structures [7,8], the persistence of incomplete or truncated structures [9], the accumulation of precursors and, less commonly, the appearance of novel structures [10]. Most modifications are a result of up- or down-regulation of glycosyltransferases that appear to have a selective influence on tumorigenicity as a result of oncogenic signalling pathways [11–13]. The role of glycosylation in the promotion or inhibition of tumour cell invasion and metastasis has received little attention from cell biologists involved in cancer research, mainly because structural and functional concepts of glycosylation in cancer are more difficult to understand than the functional role of certain proteins and their genes in defining cancer cell phenotypes. Glycosylation appears to be considered as less important than other cancer research avenues, such as the identification of oncogenes and anti-oncogenes [14], the role of apoptosis [15], angiogenesis [16], growth factor receptor binding and integrin and cadherin function [17], despite the fact that aberrant glycosylation profoundly affects all of these processes [18].

Glycosylation changes have been identified in various diseases, ranging from systemic genetic diseases to localized malignancy, such as colorectal cancer or HCC (hepatocellular carcinoma). Because of their diverse structures and the information they carry, glycans provide a valuable source of biomarkers for disease diagnosis and prognosis. In this ‘omics’ era, the concept of glycomics has evolved. In the past, research was restricted to glycosylation analysis of individual glycoproteins, and large-scale glycome screening at the tissue level was not possible because of technical difficulties. With the recent advances in analytical technologies, glycomics and glycoproteomics is gaining momentum in biomarker research [7,19].

The focus of the present review will be mainly on the applications of free N-linked and O-linked glycans [FOS (free oligosaccharides)] and GSLs in biomarker discovery, where the global changes in glycosylation are probed rather than selecting individual glycoconjugates. The technique we have employed is an HPLC-based method with glycans, free or enzymatically released, fluorescently labelled with 2-AA (2-anthranilic acid) [20,21]. The reproducibility and ease of analysis allows this technique to be applied in any experimental or diagnostic laboratory.

**FOS as potential biomarkers**

FOS are formed in the cell as a result of protein degradation either from a protein quality control route ERAD (ER-associated degradation) pathway or from protein degradation in the lysosome. Whatever the route and initial origin of the glycan, FOS are detected in all mammalian tissues, serum and urine and correlate with many diseases, both genetic and epigenetic.

An assay for urinary FOS analysis is required to screen patients for altered glycosylation in disease. This would provide a non-invasive monitoring of both any disease characteristic variations in glycosylation and disease progression. Alterations in urinary FOS have previously been utilized to identify glycosylation changes following α-glucosidase inhibition by the imino sugar, NB-DNJ (N-butyldeoxyxojirimycin). By analysing the unique FOS with the structure Glc3Man3GlcNAc1, pharmacokinetic/bioavailability data on dosing and recovery was determined in mouse [20]. The same free glycans and additional glycosylated structures in plasma and urine from Gaucher disease and Niemann–Pick type C patients taking an oral treatment for these diseases, NB-DNJ (Zavesca), were detected.

The glycosylated FOS produced following NB-DNJ treatment used ion-exchange and a lectin affinity-purification step to focus directly on the glycan of interest. This obviously biases the FOS profile produced and discounts a huge proportion of potentially important glycomarkers, such as sialylated and other charged glycans. A robust urinary complete FOS assay has subsequently been developed. This allows the screening of a number of patients to elucidate potential biomarkers of disease in HCC and colorectal cancer sufferers. The analysis of FOS in urine has been optimized to ensure minimal loss of the free glycans; however, some loss of disaccharides is inevitable using the method employed here. This is in fact an aid to further analysis, especially of N-glycan structures, which are consequently larger than the chitobiose core. The process from extraction to initial HPLC analysis requires three steps. Freeze-dried or frozen urine are both shown to be suitable forms to use for analysis, with recovery of FOS being greater than 95% in both cases. Separate analysis of the same urine samples demonstrated the reproducibility and confidence supporting this method, and a complete qualitative analysis was observed not to be significantly different (P = 0.58).

The 2-AA fluorescently labelled FOS are then analysed using two HPLC systems (Figure 1). The first system is the NP (normal-phase)-HPLC system using TSK-gel amide-80 column. From analysis of a sample of urine from a control patient over 70 discrete peaks were detected. NP/HILIC (hydrophilic interaction chromatography) is commonly used to separate fluorophore-derivatized oligosaccharides. Although the resolving power of HILIC columns allows good separation between small and large oligosaccharides, broadly based on size and glycosidic linkage, there can be considerable overlap between neutral and charged glycans. Column elution is expressed as GU (glucose units) following calculation of relative retention when compared with an external glucose oligomer standard [22]. The second system uses a Dionex AS11 column and could be used for the separation of 2AA-labelled oligosaccharides into discrete pools of neutral and multiply charged species. This separation greatly increases the number of glycan species detected, with over 100 discrete peaks revealed in a control urine sample. This separation is achieved by a combination of hydrophilic interaction and strong anion-exchange chromatography that we have termed HIAX chromatography [22]. By careful manipulation of the gradient conditions, GU values could be obtained across the entire gradient. This had not been
possible by previous published methods of analysis [22]. Additionally, the eluents used for the analysis were the same as those used for HILIC over a TSKgel Amide-80 column. This greatly simplifies the analytical approach, as a simple column exchange is now all that is required to change separation modes.

The two-column approach gives a ‘FOS fingerprint’ and this has shown to be a robust technique with recovery of urine spiked with free glycans being greater than 95%. The success of this technique is in some ways its Achilles heel with such a complex profile being produced when examining urinary FOS in this global manner. The HIAX column separates on the basis of charge, but confirmation can be obtained by using QAE-Sephadex separation of total oligosaccharide pools which can be performed before HIAX, or HILIC to alleviate the problem of neutral and monosialylated oligosaccharide, if desired. This can be coupled to chemical or enzymatic de-sialylation. Further separation of the urinary FOS has been achieved using a series of lectins, Con A (concanavalin A) (which was previously used to monitor FOS derived from the ERAD pathway following α-glucosidase inhibition) [20], SNA (Sambucus nigra agglutinin) [23] and AAL (Aleuria aurantia lectin) [24]. Con A is one of the most widely used and well-characterized lectins. Con A has broad applicability primarily because it recognizes a commonly occurring sugar structure, α-linked mannose. SNA lectin binds preferentially to sialic acid attached to terminal galactose in α-2,6 and, to a lesser degree, α-2,3 linkage. Binding is also inhibited to some extent by lactose or galactose. This lectin does not appear to bind sialic acid linked to N-acetylgalactosamine. AAL binds preferentially to fucose linked α-1,6 to N-acetylgalactosamine or to fucose linked α-1,3 to N-acetyl-lactosamine-related structures. Con A and SNA gave a recovery of greater than 95%. AAL was less efficient, but the method developed results in a greater than 90% recovery, as observed using standard 2-AA-labelled FOS.

Further analysis of a greater number of samples is in progress to determine potential biomarkers for disease progression can be extracted from the data. Additional separative analysis using lectins with different selectivities and protein-specific affinity-matrices coupled to enzyme digests and MS can be applied to gain more separative and structural analysis. The method developed has been shown to identify variations in FOS species between patients and it remains to be determined, with analysis of a greater number of longitudinal samples, whether these are related to disease progression. This method appears to be a very sensitive tool and may provide a simple non-invasive way of monitoring changes in the cellular processes occurring throughout the course of disease and potential treatment. The ‘FOS fingerprints’ from these samples demonstrated the robustness of the method developed. There was, however, a remarkable similarity between samples, even more so when the FOS were separated on the basis of charge or lectin affinity. The method was successful in detecting FOS, but products of glycosylation found in normal tissue may mask any subtle variations in glycosylation as a result of cancer progression. One major variation observed was in the amount of sialyl-lactose that was significantly increased in 40% of cancer patients. Similar elevations of sialylated structures have
been identified previously [25,26] and may correlate with a pro-inflammatory cascade that is a secondary effect due to disease progression.

**GSLs as biomarkers of disease**

Tumour progression is associated with alterations in GSL expression, which makes them attractive tumour biomarkers [11]. The detection of any alteration in GSLs resulting from cancer progression may be observed in a non-invasive manner by examining the GSLs found in patient urine. So far, a method has been developed to examine the GSLs found in the urine from a transgenic mouse engineered to phenotypically recapitulate Fabry disease. In this disorder, a genetic deficiency in α-galactosidase leads to lysosomal accumulation of a GSL, Gb3, in epithelial cells that is shed by distal tubules and excreted in the urine [27,28].

The method involves purifying GSLs from urine using a Svennerholm extraction and clean-up using C18 columns before overnight digestion with ceramide glycanase and 2-AA labelling of the released glycan. The separation of Gb3 by the HPLC methods described shows that elevated levels are detected in the urine of both Fabry mouse and a Fabry disease patient carrying the R301Q mutation (Figure 2). In humans, levels of Gb3 are elevated 4–8-fold in urine and Gb3 an excellent non-invasive easily detected metabolite to predict disease and monitor disease progression and response to therapy that can be achieved in any laboratory. As little as 100 μl of urine can be used for analysis. This method is also transferable to the analysis of storage products in other lysosomal storage diseases, complementing genotype and enzyme activity determinations.

The role of GSLs in diseases such as cancer offers the possibility that analysis would provide a viable target for biomarker discovery. To examine whether there are any changes in the GSL complement of cancer patients, including HCC, we have examined the GSLs in primary tissue liver biopsies and extended these studies to the analysis of urinary metabolites.

In primary cancer tissue there appears to be a marked change in lipid glycosylation. The level of Lc3 (amino-CTH GU = 2.68) has increased by 3.8-fold in the tumour cells (Figure 3). The amounts are on average (n = 5) 238 ± 33 pmol/mg in control liver and 906 ± 111 pmol/mg in the tumour samples. Lc3 is the precursor to the neolacto series which generates paragloboside and hence the sialylated paraglobosides. The examination of human urinary GSLs has been more problematic since a lot of non-GSL-derived contaminations are present using the Svennerholm method used for the Fabry murine urine. This has led to the development of a clean-up using amino-propyl columns [29] before ceramide glycanase digestions. This has been effective in producing clean profiles of ganglioside-derived oligosaccharides; however, the non-charged GSLs are more
problematic and the level of Lc3 is not readily separable. Methods are currently in development to improve the detection of lipid metabolites, such as Lc3, in urine samples.

Conclusions and future perspectives
This short review has highlighted the use of carbohydrate analysis to monitor glycosylation techniques in a non-invasive manner. The FOS analysis has highlighted the wide variety of glycans excreted from the body and is effective at diagnosis of genetic diseases such as LSDs (lysosome storage diseases), CDGs (congenital disorders of glycosylation) and monitoring drug treatment [30–32]. The method has all the hallmarks of an effective biomarker discovery technique, the repeatability at every step is good and process variability is well-established, allowing for a robust platform from which biomarker discovery can proceed. However, this is a whole-body glycosylation analysis and it remains to be seen whether a change in glycosylation at the tumour level is detectable at a global level. The method has been shown to detect changes in the response of the body to malignancy in terms of inflammation. The analysis of the GSLs has again been effective in terms of analysing lysosome storage disorders, but also in terms of cancer disease and progression. The complement of GSLs varies in tumour cells and more importantly during tumour progression owing to a number of different factors. Early stages see epigenetic silencing resulting in truncated GSLs typified by the Lc3 structure detected in primary tumours. There is also elevated expression of some glycosyltransferases, e.g. βGalNAc-transferases during malignancy where cells induce hypoxic disturbances which can alter the fatty acid composition of ceramide, not detectable using the methods employed here, and increased levels of modified sialic acids [33], which can be assessed. The initial results shown here demonstrate that there may also be detectable changes in the urine with increased expression levels of enzymes in the GSL pathway, and these can be monitored to evaluate phenotypic changes during the course of disease.

Funding
The authors’ work is supported by the Oxford Glycobiology Institute and the National Cancer Institute [grant number R01CA125642].

References
5 Hakomori, S. (1996) Tumor malignancy defined by aberrant glycosylation and sphingo(glyco)lipid metabolism. Cancer Res. 56, 5309–5318

Received 8 September 2010 doi:10.1042/BST0390393