Discovering new clinical markers in the field of glycomics

Jayne E. Telford, Margaret A. Doherty, Tharmala Tharmalingam and Pauline M. Rudd
Dublin-Oxford Glycobiology Laboratory, NIBRT-The National Institute for Bioprocessing Research and Training, University College Dublin, Dublin 4, Ireland

Abstract
Glycosylation modifications have been reported in a number of disease states and, as a result, there is significant focus on the discovery and development of glycan-based biomarkers. Glyco-biomarkers have the potential to enhance the efficacy and efficiency of the diagnostic procedures for these diseases.

Introduction
Glycosylation is one of the most abundant and complex post-translational modifications, with its importance underlined by the fact that it has an impact upon protein function. Aberrant glycosylation plays a pivotal role in a multitude of physiological states, including cancer [1,2], immunity [3], arthritis [4,5] and schizophrenia [6]. These disease-associated alterations in glycosylation are exploited in biomarker discovery and can be utilized independently or in combination with currently used protein biomarkers in order to increase their sensitivity and specificity.

Owing to the immense heterogeneity of oligosaccharides, characterization has remained challenging and technological advancements have been mired in difficulties. Conventional methods of glycan analysis employ an orthogonal approach, utilizing HILIC (hydrophilic-interaction chromatography), capillary electrophoresis and MS to identify the constituents of the oligosaccharide pool. The most predominantly used technique is HILIC. Briefly, glycans are cleaved from the protein backbone by PNGaseF (peptide N-glycosidase F), subsequently, the addition of fluorescent label is accomplished via 2-aminobenzamide and, finally, samples are applied to HILIC [7]. A dextran ladder is used to assign glucose units to the oligosaccharide pool. Recent advances in the development of high-throughput HILIC technologies have enabled the glycan analysis of large sample sets such as those investigated by Knezević et al. [8]. This study examined plasma glycans from 1008 individuals and correlated variability, heritability and environmental determinants with variations in their profiles [8]. Another study that was supported by the advent of high-throughput glycan analysis was a genome-wide association study conducted on the plasma glycans from a population of 3000 [9]. An association was found between single-nucleotide polymorphisms and the levels of specific glycans present. These studies highlight the potential of glycoanalytical high-throughput technologies in elucidating the impact of glycosylation on numerous disease states and have proven instrumental in the identification of potential glycan-based targets for the development of biomarkers.

Glycomic biomarkers in the diagnosis of cancer

Breast cancer
Breast cancer is the most common cancer in women in Europe [10] and in the US [11]. The diagnosis of breast cancer can be complex and multi-faceted, and typically includes monitoring of serum CA15–3 (cancer antigen 15–3) and carcinoembryonic antigen levels. However, both markers possess shortfalls, having low specificity and sensitivity, and thus are only employed as a method of ascertaining prognosis and monitoring disease progression [12]. Significant research is being carried out on breast-cancer-specific glycan modifications, with the ultimate aim of generating glycan-based diagnostic tools. Studies on breast cancer serum samples have elucidated glycosylation...
alterations including increased sialylation, higher levels of sLe^a (sialyl-Lewis X) and significant changes in fucosylation in breast cancer [1,13].

Whole-serum glycan analysis elucidated breast-cancer-specific glycosylation alterations, and subsequent multivariate statistical analysis showed potential in separating breast cancer patients from controls (Figure 1) [1]. Further investigations demonstrated that an individual glycan, A3F1G1, may be useful in the diagnosis of breast cancer, as the levels of A3F1G1 correlated better with the progression of the disease and occurrence of metastasis than the levels of CA15–3 [1]. Recent investigations have elucidated other potential glycan markers for use in prognosis determination in breast cancer [14,15]. Glycan analysis of serum from early-stage breast cancer patients who were lymph-node-negative and lymph-node-negative elucidated the existence of glycan differences between the two groups [14]. Further analyses showed that the altered glycans (A3F1G1, A2F1G1 and FA2) could be used to distinguish the lymph-node-positive patients from the lymph-node-negative patients and as such may hold potential as a non-invasive method of prognosis prediction in breast cancer patients [14].

A similar approach was employed in the investigation of the relationship between the levels of CTCs (cancelling tumour cells) and the serum glycosylation profiles in metastatic breast cancer. CTCs are cells that have separated from the primary tumour and can migrate in the bloodstream, and patients with CTC counts above a certain threshold have a poorer prognosis than those below the threshold [15]. The results showed that a number of sLe^a-containing glycans (A3F1G1, A4F1G1 and A4F2G2) and total sLe^a content were increased in serum samples from breast cancer patients who had CTC counts above the threshold in comparison with serum samples from breast cancer patients with CTC counts below the threshold [15], and therefore may hold potential in prognosis determination.

Ovarian cancer

Ovarian cancer is the most fatal cancer of all gynaecological cancers in women in Europe [10] and in the US [11]. The diagnosis of ovarian cancer, as with many other cancers, is an arduous process and routinely involves ultrasonography, examination of the level of CA125 (cancer antigen 125) in the serum or a combination of both of these methods. The serum glycoprotein CA125 is the most widely used serum biomarker for ovarian cancer; however, the use of CA125 as a biomarker for ovarian cancer has its limitations, such as its unreliability and non-specific nature, in diagnosing early-stage ovarian cancer [16]. The low specificity of CA125 is evident from observations that its levels are increased in conditions that are distinct from ovarian cancer, such as advanced adenocarcinomas [17], chronic pancreatitis [18] and benign gynaecological conditions, such as endometriosis [19]. Consequently, supplementary biomarkers are essential to enhance the detection of ovarian cancer and glycosylation analyses provide a medium for generating potential targets for diagnostic development.

Ovarian-cancer-specific glycan alterations include higher levels of sLe^a, along with increased core fucosylated agalactosylated biantennary glycans [2]. Supplementary experiments investigated the glycoproteins that may be involved in these alterations. Isolation and analysis of serum IgG showed reduced galactosylation and sialylation on the IgG heavy chain in samples from advanced ovarian cancer patients when compared with controls [2]. In addition to IgG, certain positive acute-phase proteins, haptoglobin, α1-acid glycoprotein and α1-antichymotrypsin had altered glycosylation in ovarian cancer samples, such as higher levels of sLe^a [2].

Prostate cancer

Prostate cancer is the most widespread cancer in males in Europe [10] and in the US [11]. Currently, the detection of prostate cancer is carried out by monitoring the serum levels of PSA (prostate-specific antigen); however, this process can be unspecific and unreliable because of the observation that PSA levels can also be increased in conditions that are distinct from prostate cancer, such as benign prostatic hyperplasia [20]. There are a number of reports of altered glycosylation in prostate cancer. Serum glycome analysis showed increased sLe^a in samples from prostate cancer patients in comparison with control serum [21]. Reduced sialylation of PSA was found in prostate cancer serum when compared with PSA from healthy seminal plasma [22]. Seminal plasma PSA was used as a control because PSA levels in control serum were too low to analyse. Decreased α2,3-linked sialic acid levels and lower fucose content were found on PSA from prostate cancer serum when compared with PSA from healthy seminal plasma [21]. There are five subforms of PSA (F1–F5). F3 (the most abundant form) from prostate cancer serum was found to
have altered levels of α2,3-linked sialic acids and reduced core fucosylated glycans when compared with control PSA [22]. Therefore the glycosylation patterns of the PSA subforms should be considered in the development of biomarkers for prostate cancer, and the addition of glycan alterations as a marker may increase the specificity of currently used diagnostic tests.

**SLeα involvement in the formation of metastases**

As discussed in the present review, sLeα levels are increased in a number of cancers, which is notable considering that sLeα is thought to play a role in the formation of metastases. SLeα is a ligand for the selectin proteins (E-selectin, P-selectin and L-selectin), and these interactions are vital in the process of leucocyte extravasation, which can be exploited by cancer cells and used in the formation of metastases. E-selectins (on endothelial cells) bind to sLeα on tumour cells, bringing about tethering and rolling of the tumour cell, which is subsequently followed by integrin activation and firm adhesion of the tumour cell to the endothelial cell. This process ultimately results in extravasation of the tumour cell from the bloodstream, allowing it to form secondary tumour(s) in other tissues (reviewed in [23]).

**Glycomic biomarkers for RA (rheumatoid arthritis) and schizophrenia**

**RA**

RA is a chronic autoimmune inflammatory disease of unknown aetiology affecting approx. 1% of the world’s population [24]. The pro-inflammatory cytokines, TNF-α (tumour necrosis factor α) and IL (interleukin)-17, are implicated in the pathogenesis of RA and are present at higher levels in serum and joints in patients suffering from the disease [25]. Disrupting the IL-17 signalling pathway is an effective strategy in controlling the symptoms of this debilitating illness [24]. Traditionally, serological diagnosis of RA has relied on the measurement of the rheumatoid factor which recognizes the CH2 and CH3 domains of the Fc region of human IgG [26]. However, rheumatoid factor is not specific for RA and is found in additional conditions [27]. Other auto-antibodies that are targeted as potential biomarkers for RA that have a higher specificity for RA than rheumatoid factor, include antibodies against citrullinated fibrinogen [28] and antibodies against cyclic citrullinated protein [29]. IgG molecules possess complex biantennary N-glycans linked to Asn^39 on the Fc region. IgGs are characterized according to the number of terminal galactose (G) residues present, namely, G0, G1 and G2. Modifications in IgG galactosylation were first reported in RA serum in 1985 [30]. Agalactosylated IgG levels are augmented in the sera of patients with RA, the terminal GlcNAc residue binds to mannose-binding lectins and activates the lectin pathway [31]. Ercan et al. [5] observed a significant increase in the G0/G1 ratio in serum samples of RA patients when compared with controls. Evidently, monitoring the changes in IgG galactosylation represents an important step in discovering a diagnostic biomarker.

**Schizophrenia**

Schizophrenia is a chronic neuropsychiatric illness for which no diagnostic biomarkers currently exist and consequently glycan-based biomarkers may prove valuable in the diagnosis of the disease. Glycan analysis studies on serum and CSF (cerebrospinal fluid) from controls and first-onset schizophrenia patients identified several N-glycans that may possess potential as targets for biomarker development [6]. Serum samples were separated into HAS (high-abundant protein) and LAS (low-abundant protein) fractions, and significant disease-related alterations were found in each fraction by analysis with HILIC [6]. One peak in the HAS HILIC profile was significantly higher in serum from male schizophrenia patients when compared with controls, and was significantly lower in samples from female schizophrenia patients compared with controls [6]. The glycans contained in this peak were A3F1G3S3 (predominant), A4G4S2, A4G4S3, FASG3S3 and FA3BG3S3. Two peaks were altered in the LAS fraction, one accounted for by glycans A3F1G3S3 and the other by A4G4LacS4 [6]. A number of schizophrenia-associated alterations in peaks in the CSF HILIC profile were identified and multivariate statistical analysis incorporating these alterations showed promising results in separating schizophrenia patients from controls [6].

A3F1G3S1, which was increased in LAS and HAS samples from male schizophrenia patients, contains the sLeα epitope, described above as being increased in certain cancers and is also reportedly increased in chronic inflammation [32]. SLeα levels may be linked to inflammation in schizophrenia. There are reports of inflammation, evident from increased C-reactive protein levels and from increased levels of pro-inflammatory cytokines such as IL-2, IL-6 and IL-8 in schizophrenia [33–35].

**Conclusions**

Glycosylation of proteins has a significant impact in the pathogenesis of numerous diseases, and specific changes in glycans are often a hallmark of certain disease states. Therefore glycan-based diagnostic tools hold remarkable potential in improving the specificity and efficiency of disease detection systems, in addition to providing invaluable knowledge on disease biochemistry and pathogenesis.

**Funding**

Work in our laboratory is supported by the European Commission under the Seventh Framework Programme (FP7) (SchizDX [grant number 223427], EuroGlycoArrays [grant number 215536]), the Sixth Framework Programme (FP6) (Gly4dis [Ref No 37661] and EuroCarb DB [Ref No 11952]), Enterprise Ireland [grant number PC/2008/0022] and Science Foundation Ireland (Reproductive Biology Research Cluster (RBRC) [grant number 07/SRC/B1156].

©The Authors Journal compilation ©2011 Biochemical Society
References


