Inflammatory glycoproteins in cardiometabolic disorders, autoimmune diseases and cancer

Margery A. Connelly, Eke G. Gruppen, James D. Otvos, Robin P.F. Dullaart

Contents

1. Introduction .................................................................................................................. 178
2. Glycoprotein biochemistry and rationale for measuring glycoproteins and glycans .......... 178
3. Assays of glycoproteins in biological fluids and development of high-throughput assays for glycan measurement ................................................................. 180
4. Potential clinical utility for inflammatory glycoprotein assays ...................................... 181
4.1. Glycoprotein assays and cardiometabolic disorders .............................................. 181
4.2. Glycoprotein assays and autoimmune diseases .................................................. 182
4.3. Glycoprotein assays and cancer ............................................................................ 183
4.4. Congenital disorders of glycosylation .................................................................... 183
5. Conclusions and future perspectives ............................................................................. 184

Abstract

The physiological function initially attributed to the oligosaccharide moieties or glycans on inflammatory glycoproteins was to improve protein stability. However, it is now clear that glycans play a prominent role in glycoprotein structure and function and in some cases contribute to disease states. In fact, glycan processing contributes to pathogenesis not only in autoimmune disorders but also in atherosclerotic cardiovascular disease, diabetes and malignancy. While most clinical laboratory tests measure circulating levels of inflammatory proteins, newly developed diagnostic and prognostic tests are harvesting the information that can be gleaned by measuring the amount or structure of the attached glycans, which may be unique to individuals as well as various diseases. As such, these newer glycan-based tests may provide future means for more personalized approaches to patient stratification and improved patient care.

Here we will discuss recent progress in high-throughput laboratory methods for glycomics (i.e. the study of glycan structures) and glycoprotein quantification by methods such as mass spectrometry and nuclear magnetic resonance spectroscopy. We will also review the clinical utility of glycoprotein and glycan measurements in the prediction of common low-grade inflammatory disorders including cardiovascular disease, diabetes and cancer, as well as for monitoring autoimmune disease activity.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
1. Introduction

Given the imperfections in the armamentarium of conventional biomarkers for diagnosis, prognosis, or risk prediction and disease prevention at the individual patient level, there is an ongoing effort using novel high-precision laboratory techniques to discover new biomarkers that will increase the sensitivity and specificity above current clinical tests [1–4]. Glycoproteins play key roles in inflammatory and pathological processes [5–9]. Thus, it is not surprising that investigation of the clinical utility of assays that measure inflammatory glycoproteins has received much attention [10–13]. Besides the clinical information that can be gleaned by quantifying circulating levels of glycoproteins, it is now clear that measurements based on the glycan structures of circulating proteins represent another avenue for improving diagnosis, prognosis and risk prediction of common inflammatory disorders [4,7,13–19]. Here we will briefly review the biochemistry and metabolism of glycoproteins, provide insight into the glycoprotein assays that are currently available for clinical use and describe newer high-throughput technologies that are being employed for identifying new glycan-based biomarkers that will add to the current armamentarium and are expected to improve patient care.

2. Glycoprotein biochemistry and rationale for measuring glycoproteins and glycans

Protein glycosylation is the enzyme-mediated post-translational process responsible for the attachment of glycan chains either to the nitrogen of an asparagine residue (N-linkage) or the oxygen of a serine or threonine residue (O-linkage) [8,20]. While most O-linked glycoproteins remain intracellular or are secreted and become part of the extra-cellular matrix, most of the abundant proteins in the circulation are N-linked glycoproteins. N-linked glycosylation is initiated in the endoplasmic reticulum and the oligosaccharide chains are further modified via a set of glycosyltransferases in the Golgi apparatus to form the basic

### Table 1

<table>
<thead>
<tr>
<th>Category</th>
<th>Positive acute phase proteins</th>
<th>Molecular weight (kDa)</th>
<th>Glycosylation sites (#)</th>
<th>UniProt number</th>
<th>Adult concentrations in serum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binding or transport proteins</td>
<td>α1-Acid glycoprotein (AGP, orosomucoid)</td>
<td>41–43</td>
<td>5</td>
<td>P02763</td>
<td>0.5–1.2 mg/mL</td>
</tr>
<tr>
<td></td>
<td>Haptoglobin</td>
<td>100</td>
<td>4</td>
<td>P00738</td>
<td>0.3–3.0 mg/mL</td>
</tr>
<tr>
<td></td>
<td>Ceruloplasmin</td>
<td>151</td>
<td>6</td>
<td>P00450</td>
<td>0.2–0.6 mg/mL</td>
</tr>
<tr>
<td></td>
<td>Mac-2 (or galectin-3) binding protein</td>
<td>85–97</td>
<td>7</td>
<td>Q08380</td>
<td>1.4–16.1 μg/mL</td>
</tr>
<tr>
<td>Antiproteases</td>
<td>α1-Antitrypsin</td>
<td>52</td>
<td>3</td>
<td>P01009</td>
<td>0.9–2.0 mg/mL</td>
</tr>
<tr>
<td></td>
<td>α2-Macroglobulin</td>
<td>179</td>
<td>8</td>
<td>P01023</td>
<td>1.3–3.0 mg/mL</td>
</tr>
<tr>
<td></td>
<td>α1-Antichymotrypsin</td>
<td>68</td>
<td>6</td>
<td>P01001</td>
<td>1.5–3.5 mg/mL</td>
</tr>
<tr>
<td></td>
<td>Kallistatin</td>
<td>58</td>
<td>4</td>
<td>P29622</td>
<td>10 μg/mL</td>
</tr>
<tr>
<td>Complement system</td>
<td>C2</td>
<td>83</td>
<td>8</td>
<td>P06581</td>
<td>0.02–0.4 mg/mL</td>
</tr>
<tr>
<td></td>
<td>C3</td>
<td>185</td>
<td>3</td>
<td>P01024</td>
<td>0.1–1.8 mg/mL</td>
</tr>
<tr>
<td></td>
<td>C5</td>
<td>190</td>
<td>4</td>
<td>P01031</td>
<td>0.02–0.4 mg/mL</td>
</tr>
<tr>
<td></td>
<td>C1 esterase inhibitor</td>
<td>105</td>
<td>7 N-; 8 O-linked</td>
<td>P05155</td>
<td>0.21–0.39 mg/mL</td>
</tr>
<tr>
<td>Coagulation system</td>
<td>Fibrinogen α, β, γ</td>
<td>340</td>
<td>5 N-; 2 O-linked</td>
<td>P02671, -75, -79</td>
<td>1.5–4.0 mg/mL</td>
</tr>
<tr>
<td></td>
<td>Plasminogen</td>
<td>92</td>
<td>1 N-; 2 O-linked</td>
<td>P00747</td>
<td>Plasma 120–200 μg/mL</td>
</tr>
<tr>
<td></td>
<td>Vitronectin</td>
<td>140</td>
<td>3</td>
<td>P04004</td>
<td>Plasma 110–140 μg/mL</td>
</tr>
<tr>
<td></td>
<td>α2-Antiplasmin</td>
<td>70</td>
<td>4</td>
<td>P08697</td>
<td>70 μg/mL in plasma, 47.6 μg/mL in serum</td>
</tr>
<tr>
<td></td>
<td>Prothrombin</td>
<td>72</td>
<td>3</td>
<td>P00734</td>
<td>Detection range 0.031–32 μg/mL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>43</td>
<td>3</td>
<td>P05121</td>
<td>Plasma 5–40 ng/mL</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>Plasminogen activator inhibitor-1 (PAI-1)</td>
<td>72</td>
<td>3</td>
<td>P00750</td>
<td>1–18 ng/mL</td>
</tr>
<tr>
<td></td>
<td>Tissue plasminogen activator (tPA)</td>
<td>72</td>
<td>3 N-; 1 O-linked</td>
<td>P02751</td>
<td>0.3 mg/mL</td>
</tr>
<tr>
<td></td>
<td>Fibrinectin</td>
<td>220–440</td>
<td>7 N-; 3 O-linked</td>
<td>P13093</td>
<td>0.5–100 ng/mL</td>
</tr>
<tr>
<td></td>
<td>Lipoprotein phospholipase A2 (Lp-PLA2)</td>
<td>45</td>
<td>2</td>
<td>P02741</td>
<td>hsCRP = 1.0 μg/mL; ≥3.0 μg/mL risk for CVD</td>
</tr>
<tr>
<td></td>
<td>C-reactive protein (CRP), pentamer</td>
<td>115–120</td>
<td>1†</td>
<td>P00787</td>
<td>2.1–3.8 mg/mL</td>
</tr>
<tr>
<td></td>
<td>Serum amyloid A (SAA)</td>
<td>13.5</td>
<td>0</td>
<td>P00748</td>
<td>Plasma 0.1–100 ng/mL</td>
</tr>
</tbody>
</table>

---

*Reference for adult (age 20–60 years) concentrations: C.A. Burris, E.R. Ashwood, and D.E. Bruns, eds., Tietz Textbook of Clinical Chemistry and Molecular Diagnostics (Fourth edition) Philadelphia, WB Saunders, 2006, Chapter 56 pp. 2251–2302. If no standardized assay is available, a normal detection range was reported from a commercially available ELISA assay.


‡ Reference for adult (age 20–60 years) concentrations: C.A. Burris, E.R. Ashwood, and D.E. Bruns, eds., Tietz Textbook of Clinical Chemistry and Molecular Diagnostics (Fourth edition) Philadelphia, WB Saunders, 2006, Chapter 56 pp. 2251–2302. If no standardized assay is available, a normal detection range was reported from a commercially available ELISA assay.

glycan structure. The sequence of sugar residues and the overall structure of the oligosaccharide chain depend on the cell type-specific glycosyltransferases and glycosidases and the availability of the various sugar nucleotide donors [20]. Given the vast number of known glycosyltransferases, glycosidases and monosaccharides, and the diversity of linkages that can occur, the molecular structures of protein-bound glycans are remarkably diverse, even before the glycoproteins have been released into the circulatory system [21]. Plasma levels of the majority of circulating glycoproteins rise (positive acute phase proteins) or fall (negative acute phase proteins) during the acute phase response, the systemic reaction to the presence of infection, tissue damage, cancer and pregnancy [5,16,22,23]. Table 1 provides

![Diagram of glycan structures](image)

**Fig. 1.** Examples of N-linked glycans showing mannose-rich as well as bi-, tri-, and tetra-antennary glycan structures. Two N-acetyl glucosamine (GlcNAc) residues can be attached via \((\beta1-2), (\beta1-4)\) or \((\beta1-6)\) linkages to mannose residues at the sites of glycan branching. Sialic acid and fucose residues are added or removed during inflammatory processes.

Table 2

<table>
<thead>
<tr>
<th>Disease</th>
<th>Serum test</th>
<th>Glycoprotein(s) or sugar residue</th>
<th>Assay type</th>
<th>Clinical application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiometabolic disorders</td>
<td>hsCRP</td>
<td>High-sensitivity C-reactive protein</td>
<td>Nephelometry</td>
<td>Risk for CVD or all-cause mortality and prognosis for recurrent events in patients with coronary disease or ACS Detecting bleeding disorders and ACS has been shown to have associations with CVD and all-cause mortality Risk assessment for CVD or all-cause mortality</td>
</tr>
<tr>
<td></td>
<td>Fibrinogen</td>
<td>Fibrinogen</td>
<td>ELISA or activity assay</td>
<td>Risk assessment for CVD or all-cause mortality</td>
</tr>
<tr>
<td>Total serum sialic acid</td>
<td>N-acetyleneuraminic acid</td>
<td>Colorimetric, enzymatic, chromatographic and fluorescence assays</td>
<td>Risk assessment for CVD or all-cause mortality</td>
<td></td>
</tr>
<tr>
<td>GlycA</td>
<td>N-acetylglucosamine</td>
<td>Lectin binding</td>
<td>Lectin binding</td>
<td>Risk assessment for CVD or all-cause mortality</td>
</tr>
<tr>
<td>Mac-2 BP, Fuc-Hpt</td>
<td>Mac-2 binding protein Fucosylated haptoglobin</td>
<td>ELISA and Lectin-antibody ELISA</td>
<td>Distinguish NASH from fatty liver</td>
<td></td>
</tr>
<tr>
<td>Autoimmune diseases</td>
<td>CRP</td>
<td>Conventional C-reactive protein</td>
<td>Nephelometry</td>
<td>Infection, tissue injury, and inflammatory disorders. Assessment of disease activity in RA</td>
</tr>
<tr>
<td></td>
<td>ESR</td>
<td>Fibrinogen and immunoglobulins</td>
<td>Sedimentation rate of red blood cells per hour</td>
<td></td>
</tr>
<tr>
<td>MBDA</td>
<td>VCAM-1, EGF, VEGF-A, IL-6, TNF-R1, MMP-1, MMP-3, YKL-40, Leptin, Resistin, SAA and CRP</td>
<td>Luminex based assays</td>
<td>Assessment of disease activity in RA</td>
<td></td>
</tr>
<tr>
<td>GlycA</td>
<td>N-acetylglucosamine</td>
<td>MUC16 or cancer antigen 125</td>
<td>ECLIA</td>
<td>Monitoring therapy, detecting recurrence of ovarian cancer</td>
</tr>
<tr>
<td>PSA, Pro2PSA, CA152</td>
<td>MUC1 protein levels</td>
<td>WPD2 or human epididymis protein 4</td>
<td>ELISA</td>
<td>Monitoring therapy, detecting recurrence of ovarian cancer</td>
</tr>
<tr>
<td>CA125</td>
<td>Cell adhesion glycoproteins</td>
<td>Sialylated oligosaccharide on MUC1</td>
<td>CMSI</td>
<td>Monitoring therapy for breast cancer</td>
</tr>
<tr>
<td>CA19-9</td>
<td>CEA</td>
<td>Combined HE4, CA125H, apoa-A-1, prealbumin, transferrin</td>
<td>ELISA and ECLIA</td>
<td>Predicting risk of colorectal cancer</td>
</tr>
<tr>
<td>GlycA</td>
<td>N-acetylglucosamine</td>
<td>OVA1,</td>
<td>Immunoassays</td>
<td>Monitoring therapy detecting recurrence of multiple cancers</td>
</tr>
</tbody>
</table>

More extensive lists of glycoprotein tests and biomarkers can be found in references [12,19]. ACS, acute coronary syndrome; AGP, α1-acid glycoprotein; CEA, Carcinoembryonic antigen; CMSI, chemiluminescent microparticle 2-step sandwich immunoassay; hsCRP, high sensitivity C-reactive protein; CVD, cardiovascular disease; ECLIA, electrochemiluminescence immunoassay; EGF, epidermal growth factor; ELISA, enzyme linked immunosorbent assay; ESR, erythrocyte sedimentation rate; MMP-1, matrix metalloproteinase 1; MMP-3, matrix metalloproteinase 3; NASH, non-alcoholic steatohepatitis; NMR, nuclear magnetic resonance; RA, rheumatoid arthritis; RIA, radioimmunoassay; SAA, serum amyloid A; TNFRI, tumor necrosis factor receptor type 1; VCAM-1, vascular cell adhesion molecule 1; VEGF-A, vascular endothelial growth factor A.
examples of both positive and negative acute phase glycoproteins and illustrates the diverse roles they play during an inflammatory reaction. Inflammatory glycoproteins are predominantly synthesized and secreted by hepatocytes but can be produced by activated macrophages and neutrophils in the periphery [5,15,17]. While IL-6 is the predominant stimulator of overall glycoprotein production during acute and chronic inflammation, other cytokines such as IL-1β, TNFα, interferon γ, TGFβ and IL-8, stimulate the production of subsets of glycoproteins. Because inflammation is the basis for many autoimmune and chronic low grade inflammatory diseases such as cardiovascular disease (CVD), type 2 diabetes (T2DM) and cancer, glycoproteins play an integral part in the physiology and pathophysiology of these diseases. As a result, many current clinical tests utilize circulating levels of inflammatory glycoproteins (e.g. haptoglobin and α-fetoprotein) for diagnostic or prognostic purposes.

Besides changes in circulating protein levels, the glycan structures of acute phase glycoproteins are dynamically altered by glycosidases, glycosyltransferases and sialyltransferases in the circulation [14,15]. Post-translational modifications in glycan structures during inflammation include changes in the number of antennary branches, increased sialylation and fucosylation and decreased galactosylation [14–16]. While the glycos of some proteins remain rich in mannose residues, the carbohydrate structures of many N-linked inflammatory glycoproteins become bi-, tri- and tetra-antennary after inflammation-mediated processing [14–16] (Fig. 1). These branched glycans are rich in N-acetylgalactosamine (GlcNAc), N-acetylgalactosamine, sialic acid and fucose residues in a myriad of different arrangements, contributing to the potential diversity of glycan structures [14–17,20,21] (Fig. 1). Therefore, there are both intracellular and extracellular post-translational processes that contribute to the overall diversity of glycan structures that can occur in any individual. These are also many factors that can influence glycan complexity including: 1) cell-type specific expression of glycosyltransferases, glycosidases, 2) availability of the various monosaccharides, 3) age, 4) gender, 5) epigenetic background, 5) environment (e.g. diet, health, smoking and alcohol consumption) and 6) disease processes (e.g. autoimmune diseases, cancer as well as low-grade inflammatory diseases such as CVD and T2DM) [21,24].

Although it was once thought that the only purpose for having carbohydrate side-chains on glycoproteins was to aid in protein stability, it has become increasingly clear that glycans play a much more active role in glycoprotein structure and function. Glycans participate in many key biological processes including ligand binding, transport and clearance, cell adhesion, receptor binding and activation and signal transduction [4,7–9,14,15,20]. Inflammation-induced glycan modifications affect protein folding by masking sites for protease cleavage, preventing proteolysis and extending the circulating half-life of serum proteins [4,8,9,20,25]. Moreover, they alter a protein’s tertiary or quaternary structure, redirecting it to different cell membrane receptors and changing its downstream cellular effects [4,8,9,15,20]. These functional alterations may lead to modulation of the immune response or, if modified aberrantly, can lead to autoimmunity disease. For example, glycans are a fundamental part of self- versus non-self-recognition and alterations in immunoglobulin G (IgG) glycosylation have been reported in various immune diseases including rheumatoid arthritis (RA) [8,20]. Therefore, glycans are often casual in the disease process and monitoring these changes may provide pertinent information regarding disease stages. In effect, both desirable and undesirable changes in glycan structure may be exploited for risk assessment, patient stratification, diagnostic or prognostic purposes [14,13,18,24,26,27].

Alpha1-acid glycoprotein (AGP), also known as orosomucoid, provides a good example of how changes in glycan structure can affect glycoprotein function and be exploited for diagnostic or prognostic purposes [15]. Normal circulating concentrations of AGP range from 0.6–1.2 mg/mL, and its plasma level is increased up to 50-fold during an acute inflammatory response, making AGP the second most abundant circulating protein (1–3% of plasma protein) [4,15]. AGP contains 5 sites for N-linked glycosylation and is therefore very high in carbohydrate content (~40%) [4,15]. During an acute phase response, the lengths of the oligosaccharide chains on AGP increase and are modified from bi- to tri- and tetra-antennary branches, accompanied by an increase in fucosylation and sialylation [4,15]. Both the immunomodulatory and the binding properties of AGP are strongly dependent on its carbohydrate composition; therefore, inflammation-mediated alterations in glycan structure have a profound effect on AGP function [15].

Increased fucosylation of AGP has been reported in some diseases, allowing measurement of AGP fucosylation to be useful for diagnostic purposes. For example, fucosylated AGP was significantly higher in patients with liver cirrhosis compared to steatosis of the liver, non-alcoholic steatohepatitis (NASH) and fibrosis due to chronic viral-induced hepatitis, suggesting that this glycan marker may be useful for detecting liver cirrhosis [15]. Interestingly, AGP glycan modification appears to occur in some inflammatory diseases, but not others. For example, increased AGP glycan branching has been observed in patients with asthma and RA but not in patients with ulcerative colitis [15]. Moreover, glycan structure modifications on AGP led to reduced collagenase-3 activity and collagen binding, which could exacerbate the disease process in RA patients [15]. This may be true for many other circulating inflammatory glycoproteins (Table 1). Given the diversity in the numbers of glycoproteins in biological fluids as well as the unique changes that may occur in some diseases and not others, there is likely a wealth of information yet to be mined from glycoproteins as well as their glycans for clinical use [24].

3. Assays of glycoproteins in biological fluids and development of high-throughput assays for glycan measurement

Currently, concentrations of individual inflammatory glycoproteins are determined using immunochemical methods such as enzyme-linked immunosorbent assays (ELISAs), electrochemiluminescence immunosassay (ECLIA), luminescence based assays, radioimmunoassays (RIA) and nephelometric assays that quantify the amount of protein present in biological samples (Table 2). Such assays are employed to determine protein levels of many of the inflammatory glycoproteins including AGP, haptoglobin, α1-antitrypsin, α2-macroglobulin, α1-microglobulin and β2-microglobulin. While quantifying protein levels remains the mainstay for measurement of inflammatory glycoprotein levels, measuring the glycan portion of inflammatory proteins is becoming increasingly useful for diagnostic purposes. This can be accomplished using lectin-binding ELISAs (Table 2) as well as some of the newer high-throughput technologies such as mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR) which have recently been introduced to the clinical laboratory.

MS techniques are becoming more common place in clinical laboratories. However, effective analysis of protein-derived circulating glycans is still difficult to accomplish due to the high complexity that is caused by variations in glycan linkage and branching, macro- and micro-heterogeneity. Currently, a combination of methods is often used. Here we describe some of the major MS-based approaches used in glycomics research which may eventually identify new tests for clinical use. Methods for O-linked structures are less well developed compared to methods for N-linked structures and will not be discussed in this review.

Normal phase high performance liquid chromatography (HPLC) is a well-known separation technique that has been used in laboratories for years. In addition, ultra performance liquid chromatography (UPLC) involves HPLC with very high pressure and is one of the newest chromatography technologies in the field of glycomics. UPLC allows high efficiency separations and reduced analysis times [28]. UPLC has the ability to separate glycan isomers. Until recently, UPLC was not widely used in the field of glycan profiling due to the lack of appropriate stationary phases [29,30]. Hydrophilic interaction liquid chromatography (HILIC) is a separation technique which is related to normal phase
HILIC. HILIC columns were originally used for analysis of highly polar analytes and later also for other types of substances including peptides [31] and glycans [32]. A limitation of HILIC-based analyses is the amount of time required per chromatographic run. However, since the introduction of sub-2-μm stationary phases, HPLC or UPLC in combination with HILIC have been used for analyzing glycans [29,32]. Separation of structural isomers is often achieved which makes HILIC in combination with HPLC or UPLC a valuable tool for structural analysis of oligosaccharides.

Fluorescence detection is a glycan analysis method for quantifying fluorescently labeled glycans. The labeled glycans can be separated by, for example, HILIC and detected by sensitive fluorescence detectors or by MS in some cases. The use of a fluorescence detector enables quantification of even minor glycans. Tagging the glycans with a fluorescent label such as 2-aminobenzamide (2-AB) allows the glycan to be detected even at femtomole levels [33]. Besides 2-AB other fluorescent tags are commercially available. The advantages of 2-AB is that it is compatible with multiple analytical methods including MS which makes it possible to obtain mass and structural information [34].

MS-based detection techniques are promising as enabling methods in the field of glycomics. The glycan can be removed either enzymatically or chemically from the protein. Intact N-linked glycans can be enzymatically split from glycoproteins with an amidasase such as peptide-N-glycosidase F [34]. Alternatively, hydrazinolysis can be used for chemical release. MS provides molecular mass and structural information. A wide variety of MS-based techniques are available for glycoconjugate analysis. However, quantification by MS is not always reliable and for some samples there can be overlap from isobaric glycans (discrete isomeric glycan structures that possess the same mass) [33]. MS can be used alone or coupled to separation methods such as HPLC, UPLC, HILIC or capillary electrophoresis to increase the sensitivity [35–37]. Furthermore, matrix assisted laser desorption–ionization-MS and electrospray ionization-MS are often applied. If there are a variety of possible isomers, each one may be discriminated from the other using multistage analyses. However, MS data can be very complex and interpretation requires expertise.

Although these techniques have been useful for identifying novel glycan moieties on various glycoproteins, and it has been speculated that these novel assays may eventually be useful for diagnostic purposes, none of the MS-based techniques have been routinely employed in the clinical laboratory to date.

Proton (1H) NMR, another high-throughput technological platform that is able to quantify inflammatory glycoproteins based on their glycans, was recently introduced to the clinical laboratory setting [38–43]. Although it is not possible to identify and quantify individual proteins via NMR, it is possible to measure subsets of glycoproteins based on their shared glycan moieties [38,39,42]. Protons on the sugar residues in the oligosaccharide chains emit different signals depending on their structural environment. For example, the N-acetyl methyl group protons emit different NMR signals if they are part of GlcNAc as opposed to N-acetylenuraminic acid (sialic acid), allowing for identification of the various sugar residues based on the chemical shift of their protons, i.e. the position of the signal peak in the NMR spectrum [38,39]. The complex glycan structures of several acute phase proteins including AGP and transferrin have been determined and catalogued using NMR, allowing for easy identification of the NMR signals for a number of the sugar residues found on inflammatory glycoproteins [38,39].

Recently, an NMR-based assay called GlycA was developed that quantifies circulating inflammatory glycoproteins based on a subset of mobile GlcNAc residues [42,44]. In fact, it is only the GlcNAc moieties in β(1 → 2) or β(1 → 6) linkage with a preceding mannose that give rise to the GlycA NMR signal at 2.00 ± 0.01 ppm in the NMR LipidProfile® test spectra of serum or plasma [38,42,45]. It is also possible to quantify the methyl signals from GlcNAc residues at other positions in the bi-, tri-, and tetra-antennary glycans as well as from sialic acid [38,42,44,45]. Therefore, it is possible that there are other NMR signals besides GlycA that, when quantified, may provide useful information for the clinician.

The serum GlycA NMR signal is comprised primarily of contributions from the GlcNac residues on AGP, haptoglobin, α1-antitrypsin, α1-antichymotrypsin and transferrin [42]. Because plasma concentrations of C-reactive protein (CRP) and cytokines are much lower in comparison and they are not heavily glycosylated, they contribute negligibly to the measured GlycA signal [42]. Reduced glycan mobility is another reason why not all proteins with GlcNAc residues produce observable NMR signals, which is the case for fibrinogen and IgG [42]. Haptoglobin, AGP, α1-antitrypsin and α1-antichymotrypsin are positive acute phase proteins that increase in concentration and glycan complexity in inflammatory states [7,14–17], enabling GlycA to be a biomarker of systemic inflammation that is associated with inflammatory markers such as high-sensitivity CRP (hsCRP), fibrinogen, IL-6, serum amyloid A (SAA) and lipoprotein-associated phospholipase A₂ (Lp-PLA₂) [42,46–51] as well as increased neutrophil activity [52]. It has also been reported that GlycA is related to increased mortality risk [1,52,53] [Gruppen et al. unpublished results]. Therefore, despite similarities in disease associations, GlycA, CRP, fibrinogen and other inflammatory markers likely capture different aspects of the inflammatory response [52]. Moreover, it has been reported that hsCRP, but not GlycA, levels were decreased after statin administration [53]. Therefore, it is clear that GlycA and other inflammatory biomarkers may at least be in part independent, and perhaps even additive, in the clinical information they impart. Furthermore, as a composite biomarker that measures both the increased protein levels and enhanced glycosylation states of the most abundant circulating acute phase proteins, GlycA may be a better reflection of a systemic acute phase response than any single glycoprotein component [42]. For example, assays for measuring individual acute phase proteins, such as hsCRP, often exhibit high intra-individual variability [54–57]. One approach to overcome this issue is to measure multiple inflammatory markers at once. For instance, one can compute a low-grade inflammation score, based on the Z-scores of a number of individual inflammation markers, such as hsCRP, TNF-α, IL-6, IL-8, SAA, soluble intercellular adhesion molecule 1 (sICAM-1), ceruloplasmin and haptoglobin [58]. While useful for research purposes, this computation is not convenient for physician use. GlycA, on the other hand, is already a composite biomarker that simultaneously measures multiple markers, giving it the advantage of having low within-subject biological variation [42].

4. Potential clinical utility for inflammatory glycoprotein assays

4.1. Glycoprotein assays and cardiometabolic disorders

Besides serving as biomarkers of acute or chronic inflammation or infection, elevations of glycoproteins such as hsCRP and fibrinogen are of clinical interest as markers of CVD (Table 2). Driving much of this interest is the established role of inflammation in all stages of the atherosclerotic disease process from lesion initiation to progression as well as plaque destabilization [59,60]. Epidemiologic studies have confirmed the link between systemic inflammation and adverse clinical outcomes by demonstrating consistent, independent associations of hsCRP and fibrinogen with both incident CVD and all-cause mortality [61,62]. Among the many inflammatory proteins that could serve as clinical indicators of the risk associated with inflammation, hsCRP has been favored due to its stability in fresh and frozen specimens, wide dynamic range, and availability of relatively inexpensive, standardized, and precise high-sensitivity immunoassays [59,60,63,64]. Glycan moieties themselves, such as sialic acid (N-acetylenuraminic acid), the terminal monosaccharide of glycoconjugates, have also been shown to correlate with CVD [65]. Several types of assays have been deployed for the quantification of total serum sialic acid including colorimetric, enzymatic, chromatographic and fluorescence based assays [65]. Although sialic acid can be found on glycolipids, the majority of serum sialic acid can be found on the glycan chains of AGP, haptoglobin, α1-antitrypsin, α1-antichymotrypsin, ceruloplasmin, fibrinogen and...
transferrin [65]. Sialic acid was shown to be positively associated with TNFα and IL-6 [65] and multiple studies have shown positive associations of total serum sialic acid with CVD, stroke and mortality [65–68]. A recent study reported that sialic acid was an independent risk marker for CVD during 40 years follow-up among Swedish individuals [69]. Taken together, sialic acid is a marker of systemic inflammation that can be used for risk assessment in subjects with CVD, heart failure and T2DM [65–67,69,70].

GlycA, the NMR signal derived from multiple inflammatory glycoproteins, was demonstrated to predict future CVD and T2DM (Table 2) [71–73]. GlycA was shown to be related to the leptin/adiponectin ratio, suggesting that adipose tissue-associated low-grade inflammation could be involved in the regulation of inflammatory glycoproteins [49]. Similar to hsCRP, GlycA was found to be higher in subjects with metabolic syndrome and was positively correlated with body mass index (BMI) and insulin resistance determined by homeostasis model assessment of insulin resistance (HOMA-IR) [48–50]. In the Women’s Health Study (MHS), GlycA was associated with CVD events, independent of traditional risk factors [71]. In the Prevention of Renal and Vascular End-stage Disease (PREVEND) study, GlycA was associated with incident CVD, defined as the combined end-point of CV morbidity and mortality, independent of clinical and lipid measures as well as renal function [72]. Baseline concentrations of GlycA in the Justification for the Use of Statins in Prevention: An Interventional Trial Evaluating Rosuvastatin (JUPITER) trial were significantly associated with incident CVD events, even when adjusting for established risk factors and a family history of premature coronary heart disease [73]. Remarkably, this association was only slightly attenuated by hsCRP, suggesting that the two biomarkers are reflecting somewhat different pathological processes [73]. In addition, GlycA was shown to be associated with future major adverse coronary events and mortality in two different cohorts of patients undergoing coronary angiography [1,52,74]. Of note, the association of GlycA with incident T2DM remained statistically significant both in the WHS and PREVEND even after adjusting for traditional diabetes risk factors and hsCRP [43,75,76]. Thus evidence is accumulating that GlycA may be a useful biomarker for the assessment of CVD and T2DM risk.

A lectin-based assay, called LeCT-Hepa, that exploits the changes in the glycan structure of AGP has been developed to detect liver fibrosis in patients with chronic viral hepatitis and NASH (Table 2) [77]. LeCT-Hepa is a multi-lectin antibody immunoassay that binds glycosylated AGP [77]. First AGP is immunoprecipitated using a high-throughput, automated protein purification system (ED-01), then a fully automated immunoassay analyzer (HISCL-2000i) is employed to acquire the two glycoprotein binding parameters (AOL/DSA and MAL/DSA) that are produced by the binding of glycosylated AGP to three lectins isolated from Aspergillus oryzae (AOL), Maackia amurensis (MAL) and Datura stramonium (DSA) [77]. A formula is then used to calculate a score that was reported to correlate with fibrosis stage as determined by liver biopsy [77]. This assay gave comparable, if not better performance than the Fib-4 index, for the diagnosis of significant fibrosis in chronic hepatitis C patients [78] and comparable performance to FibroScan in hepatitis B infected patients [79]. This assay, however, is not yet available in the clinical laboratory.

Recently it was shown that quantification of two inflammatory glycoproteins quantified by ELISA, fucosylated haptoglobin and Mac-2 binding protein (also known as galectin-3 binding protein), may be useful for the diagnosis of NASH and liver fibrosis (Table 2) [80]. The authors hypothesized that the fucosylation-based sorting machinery is disrupted in ballooning hepatocytes and that hyperfucosylated glycoproteins are secreted from the liver into serum in the diseased liver. Based on this hypothesis they developed a lectin-based ELISA to quantify fucosylated haptoglobin and showed that this assay was useful for the prediction of ballooning hepatocytes in NASH [80]. They also showed that Mac-2 binding protein, quantified by traditional ELISA, was a good biomarker for liver fibrosis. Moreover, the combination of the two glycoproteins was able to distinguish NASH from simple hepatic steatosis [80]. However, additional clinical validation studies are needed to fully understand the clinical usefulness of this combined biomarker test.

4.2. Glycoprotein assays and autoimmune diseases

RA is an autoimmune disease that manifests itself as severe inflammation in multiple joints, leading to erosions of the cartilage and bone and sometimes causing joint deformity, joint pain, swelling, and redness are common symptoms of RA. Tight control of disease activity, including monitoring of acute phase proteins is standard of care in RA management [81,82]. The markers most commonly used to assess RA disease activity are CRP and erythrocyte sedimentation rate (ESR) (Table 2) [81,82]. Both tests have been incorporated into the Disease Activity Score based on 28 joints (DAS28), the core set of measures proposed in the American College of Rheumatology and the American College of Rheumatology/European League Against Rheumatism RA remission criteria [83–85]. However, both CRP and ESR have limitations. For example, ESR is altered by non-inflammatory conditions such as chronic kidney disease, pregnancy, anemia, abnormal red blood cell shape or size, and serum protein concentrations [86]. Because some of these confounding influences are unrelated to RA disease activity, the current treat-to-target recommendations include cautions about the use of ESR for monitoring RA activity [82]. HsCRP exhibits high variability over time, potentially making it less reliable for assessment of RA disease activity at any one time point [55–57]. Moreover, CRP and ESR values are in the normal range in up to half of patients with active disease and they are often discordant with each other. Thus, alternative markers of inflammation whose measurements aren’t affected by these factors would be useful for assessing RA disease activity.

Recently a multi-biomarker disease activity (MBDA) blood test was developed to assess disease activity in adult patients with RA (Table 2) [87–90]. The test measures 12 inflammatory biomarkers (VCAM-1, EGF, VEGF-A, IL-6, TNF-R1, MMP-1, MMP-3, YKL-40, leptin, resistin, SAA and CRP), including a number of cytokines and acute phase glycoproteins that play key roles in the underlying pathophysiology of RA [87–90]. The MBDA test is based on an algorithm that uses the concentrations of the 12 biomarkers to generate a score that represents the level of RA disease activity on a scale of 1 (lowest activity) to 100 (greatest). Analytical validation studies have proven the MBDA test to be precise and reproducible [87–90]. The MBDA test was developed to correlate with the 28-joint Disease Activity Score (DAS28) and has been clinically validated by correlations with DAS28 and other disease activity measures in independent RA cohorts, with thresholds established for low, moderate and high disease activity [87–90]. Other studies show that the MBDA test tracks responses to treatment with biologic and non-biologic disease-modifying antirheumatic drugs (DMARDs) and may potentially be an indicator of progressive joint damage in patients with RA [87–90]. The MBDA test, however, has not been validated for diagnosing RA.

GlycA may be useful for assessing disease activity and monitoring anti-inflammatory treatment in patients with autoimmune diseases like RA and SLE (Table 2). GlycA was shown to be higher in RA and systemic lupus erythematosus (SLE) [47,91,92]. In a cross-sectional study that included 166 RA patients and 90 control subjects, GlycA concentrations were higher in RA patients compared to control subjects [47]. Moreover, increased GlycA concentrations were robustly associated with increasing degree of RA disease activity [47]. GlycA was associated with the 28-joint count Disease Activity Score with erythrocyte sedimentation rate (DAS28-ESR) and its components: tender and swollen joint counts, patient-reported global health score, ESR and hsCRP [47]. Additionally, GlycA was significantly correlated with Larsen score, a radiographic scoring of joint disease, whereas hsCRP and ESR were not [47]. GlycA concentrations were not different between rheumatoid factor (RF) positive and negative RA patients, which was expected given
that glycosylated immunoglobulins do not contribute to the GlycA NMR signal [42,47]. Additionally, GlycA was associated with coronary artery atherosclerosis in patients with RA [47]. GlycA levels were also higher in patients with SLE than matched control subjects [91]. In the same study, GlycA levels were positively associated with ESR, hsCRP, E-selectin, sICAM-1 and triglycerides, but not with creatinine, SLE Disease Activity Index (SLEDAI), SLE Collaborating Clinics (SLICC/ACR) Damage Index, or coronary calcium scores [91]. In a separate SLE cohort, mean GlycA levels were somewhat higher in female patients with high disease activity vs. patients with low or no disease activity and non-afflicted women [42,92]. In a longitudinal analysis of SLE activity, GlycA increased significantly along with increases in SLEDAI [92]. Taken together, GlycA may have utility for assessing disease activity in patients with autoimmune diseases such as RA and SLE. Given its ability to predict CVD events and its association with coronary artery atherosclerosis, GlycA may also be useful for assessing CVD risk in patients with autoimmune diseases, for whom traditional CVD risk factors such as low density lipoprotein cholesterol (LDL-C) and total cholesterol lack strong CVD associations [93–98].

4.3. Glycoprotein assays and cancer

Cancer is the second most common cause of death in developed countries, with breast and prostate cancer being the most prevalent in the United States [99]. While early detection has helped reduce cancer-related deaths, many cancers are not discovered until they are at a more advanced stage, when prognosis is often not favorable. Most of the clinically used cancer biomarkers are effective when applied to patients with later stage cancers but are often ineffective at detecting theclinically used cancer biomarkers are effective when applied to patients with autoimmunediseases such as RA and SLE. Given its ability to predict CVD events and its association with coronary artery atherosclerosis, GlycA may also be useful for assessing CVD risk in patients with autoimmunediseases, for whom traditional CVD risk factors such as low density lipoprotein cholesterol (LDL-C) and total cholesterol lack strong CVD associations [93–98].

None of the single glycoprotein tests is considered optimal; therefore, better biomarker tests are needed for early diagnosis, prognosis and personalized medicine in the cancer field [18]. Multivariate algorithms have been developed that increase specificity and/or sensitivity for cancer detection over single biomarker tests (Table 2). The OVA1® test combines the results of s–2 microglobulin, CA125II, apolipoprotein A-I, prealbumin and transferrin into one score of 0–10 [104]. The Risk of Malignancy Algorithm or ROMA™ test combines the results of HE4 enzyme immunoassay (EIA) and CA125 II [105]. Both of these tests measure multiple circulating glycoproteins and have been cleared by the FDA for prediction of malignant ovarian cancer. Additionally, an OVA2® next generation multivariate index assay is currently being evaluated by the FDA for clearance for the same indication. CA19-9 is a cancer associated marker that measures the amount of sialyl Lewis antigen (SLeα) tetrasaccharide on all circulating inflammatory glycoproteins and has been used to monitor response to therapy in patients with an established diagnosis of pancreatic, colorectal, gastric or biliary cancer [18,100].

There are many biomarkers with the potential for improving assay performance when included in a multivariable algorithm. For example, galactosylated, fucosylated and poly N-acetyllactosamine glycoforms of α1-antitrypsin have the potential to distinguish between non-small cell lung carcinoma and benign pulmonary disease [12]. Fucosylated α1-antitrypsin also has the potential to distinguish adenocarcinoma from benign pulmonary disease [12]. Fucosylated haptoglobin combined with CEA may be useful as a prognostic biomarker in colorectal cancer [106] and fucosylated haptoglobin alone may be useful for prostate cancer as it correlates with Gleason scores and biochemical recurrence after radical prostatectomy [107]. In addition, GlycA, the marker of circulating GlcNAc residues, was found to be associated with incident colorectal cancer and colorectal cancer mortality but was not associated with breast cancer or mortality from any other cancer in the WHS [108].

The fact that altered glycosylated forms of inflammatory glycoproteins have been associated with acute and chronic inflammatory diseases as well as cancer provokes intriguing questions about the potential links between inflammation and cancer. It has been hypothesized that chronic inflammation plays a role as a causal factor for the development of some cancers. For example, persistent infection with Helicobacter pylori causes chronic atrophic gastritis which may lead to dysplasia and gastric carcinoma [100]. Moreover, there is a well-known connection between colorectal cancer and inflammation; however, it is not yet known whether chronic inflammation exacerbates the progression to colorectal cancer or if colorectal cancer stimulates the secretion of cytokines that then stimulate a chronic inflammatory response [109]. The advent of high-throughput techniques for analyzing glycan structures as well as measuring levels of inflammatory glycoproteins based on both their protein and glycan content should elicit much research to address these questions in the near future.

4.4. Congenital disorders of glycosylation

Further evidence for the importance of glycans in protein function and the potential use of glycan isoforms to increase specificity for disease diagnosis, stems from the study of monogenic disorders in the glycosylation pathway, the congenital disorders of glycosylation (CDG) [19]. Over 100 human genetic disorders have been associated with aberrant glycan metabolism [110]. Because these defective genes affect proteins in a variety of functionally diverse metabolic pathways, the clinical presentation can vary, making differentiation between CDG subtypes quite challenging. Currently, diagnostic tests for CDG are limited to electrophoresis or MS-based tests that characterize the various glycoforms of transferrin [19,111]. CDG–I mutations are diagnosed by the presence of transferrin with unoccupied glycosylation sites, whereas CDG–II defects are characterized by the presence of transferrin with immature, truncated glycans [19]. N-glycan profiling holds promise for identifying additional glycoprotein biomarkers to aid in the diagnosis of the many
CDG that are known to exist [19,37]. Nevertheless, interpretation of glyco-
 alterations is complicated by fact that the immune response can lead
to changes in glycan structure besides those caused by the under-
lying genetic defect. Therefore, global glycan profiling in complex bio-
samples for the purpose of diagnosing CDG holds promise, but is
not yet useful in the clinical laboratory setting.

5. Conclusions and future perspectives

With the implementation of personalized medicine comes the task
of discovering and evaluating new biomarkers that have the potential
to improve the performance characteristics of current tests for clinical
care. Many tests are being developed to date that support the relevance
of high throughput assays for biomarkers presumed to be associated
with chronic cardiometabolic disorders like CVD, T2DM and NASH, as
well as autoimmune disorders and cancer. Among other techniques, NMIR
spectroscopy holds promise to identify subjects at risk for a num-
ber of low grade inflammation–associated diseases, and may also have
value to predict mortality [1,42,53,112,113]. As outlined in this review,
it is increasingly appreciated that knowledge about alterations in the
levels of glycoproteins in biological fluids as such, as well as with respect
to the extent and specificity of the various glycan structures may im-
prove risk stratification and identify novel pathogenic pathways.
On the one hand, abnormalities in the process of glycosylation can be linked
to distinct clinical entities, while on the other hand glycomics will open
new avenues from a systems biology perspective. It is anticipated that a
glycomics approach will also be helpful to forge a link with genomics,
lipidomics, proteomics and metabolomics, especially given the fact that the entities measured in the latter ‘omics’ often contribute to the di-
versity observed in glycomics [21,114]. Of further relevance, although it
has been surmised that glycan levels are to an important extent genet-
ically determined with environmental factors possibly playing a less im-
portant role, it is clear that environmental factors such as smoking and
alcohol consumption often lead to measurable differences in glycan
structure [21,2,4,115]. Among other challenges, results from glycomics
analyses by high-throughput techniques combined with a genome-
wide association study (GWAS) approach are required to underpin po-
tentially important novel causal pathways in disease development
[116].

The complex chemistry of glycans makes detailed analyses of their
structures limited to specialist research laboratories with the most com-
plete structural analyses only being possible using a combination of sev-
eral advanced analytical techniques. From a clinical perspective there is
a quest for technologies to analyze complex samples quickly with min-
nimal need for specialist facilities and technical expertise. However, it
is clear that we are moving on a trajectory toward a time when the wealth
of information that has yet to be mined from glycoproteins and their
glycans will contribute to a more personalized approach to patient care.

Acknowledgements

The authors thank Professor I. P. Kema for reviewing the section
about the assays for glycoprotein measurement and Bruce P. Damiano,
Ph.D. for helpful comments and suggestions.

References

Esko, M.L. Tammsoo, R. Magi, S. Smit, P. Lusso, G. Kaplan, L. Wolfe, M.J. Memoli, M. He,
H. Vega, L.J. Kim, Y. Huang, N. Hussein, E. Nievas, R. Mitchell, M. Garofalo, A.
Losing, D.C. Ireland, C. Grunes, R. Cimbro, V. Patel, G. Holzapfel, D. Salahuddin,
T. Bristol, D. Adams, R.E. Marciano, M. Hedeg, Y. Li, K.R. Calvo, J. Stoddard, J.S.
Boerkoel, J.A. Chicini, G. Di Pasquale, D. Verthelyi, S.D. Rosenzweig, Glycosyl-
arthritis, hypogammaglobulinemia, and resistance to viral infections, N. Engl. J.
012 (Epub ahead of print).
H. Vega, L.J. Kim, Y. Huang, N. Hussein, E. Nievas, R. Mitchell, M. Garofalo, A.
Losing, D.C. Ireland, C. Grunes, R. Cimbro, V. Patel, G. Holzapfel, D. Salahuddin,
T. Bristol, D. Adams, R.E. Marciano, M. Hedeg, Y. Li, K.R. Calvo, J. Stoddard, J.S.
Boerkoel, J.A. Chicini, G. Di Pasquale, D. Verthelyi, S.D. Rosenzweig, Glycosyl-
larthritis, hypogammaglobulinemia, and resistance to viral infections, N. Engl. J.
