

Glycans and glycoproteins as specific biomarkers for cancer

Muchena J. Kailemia¹ · Dayoung Park¹ · Carlito B. Lebrilla¹

Received: 23 June 2016 / Revised: 28 July 2016 / Accepted: 12 August 2016 / Published online: 3 September 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract Protein glycosylation and other post-translational modifications are involved in potentially all aspects of human growth and development. Defective glycosylation has adverse effects on human physiological conditions and accompanies many chronic and infectious diseases. Altered glycosylation can occur at the onset and/or during tumor progression. Identifying these changes at early disease stages may aid in making decisions regarding treatments, as early intervention can greatly enhance survival. This review highlights some of the efforts being made to identify N- and O-glycosylation profile shifts in cancer using mass spectrometry. The analysis of single or panels of potential glycoprotein cancer markers are covered. Other emerging technologies such as global glycan release and site-specific glycosylation analysis and quantitation are also discussed.

Keywords Mass spectrometry · Cancer · Disease biomarker · Glycomics · Glycoproteomics · Site-specific glycosylation

Introduction

Cancer is a major cause of death in most parts of the world [1–4]. In the USA alone, more than 500,000 deaths are projected to be associated with cancer and about 1.6 million

new cancer cases are expected to be diagnosed in 2016 [5]. These deleterious effects of cancer continue despite the fact that when cancer is diagnosed early, it can be contained or even cured. Efforts to identify biomarkers that can detect cancer early and discriminate a given type of cancer from other diseases face many challenges. Some of these challenges include the complexity of biological samples from which potential biomarkers are derived, extensive heterogeneity of potential analytes among different diseases, and limitations of the current analytical methods [6].

A biomarker is a biologically important signature that unambiguously identifies a certain physiological condition. It could be a single measured entity or a panel of indicator substances [6]. A biomarker can be used to screen for a given disease condition, monitor patients undergoing therapy, or even identify the reoccurrence of a disease condition. Markers are also useful in identifying people who are at a higher risk of developing certain diseases. The progress being made in cancer research raises hopes for identification of early detection markers.

Screening for biomarkers requires a thorough investigation of the potential indicator analytes to ensure that they meet specific requirements. For example, they should be able to identify when the condition exists (sensitivity) and when it does not exist (specificity). Most potential cancer biomarkers do not pass this test, resulting in false positives and false negatives, which are not favorable for the patients as well as those seeking prognosis [7, 8]. The samples should be amenable to robust and reproducible instrumentation to minimize errors that may lead to incorrect diagnosis. Where possible, the samples should be easily obtainable in a noninvasive manner, and the assay should be affordable so that it can be accessible to a large population of people. It is also important that a biomarker is validated across a broad range of populations and at different sites (laboratories).

Published in the topical collection *Glycomics, Glycoproteomics and Allied Topics* with guest editors Yehia Mechref and David Muddiman.

✉ Carlito B. Lebrilla
cblebrilla@ucdavis.edu

¹ Department of Chemistry, University of California, Davis, CA 95616, USA

The current methods for discovering cancer biomarkers have been reviewed recently [9]. These methods include gene expression measurements using DNA, RNA, and microRNA gene arrays [9, 10] and protein microarray technologies [11, 12]. For proteins, 2D-gel electrophoresis has been widely used because it is readily available. Although it remains a very useful technique, it requires large amounts of samples due to poor sensitivity, and it is a laborious and slow process. Mass spectrometry has essentially replaced gels because of its versatility in profiling biological compounds. The methods for proteomics, peptidomics, metabolomics, proteoglycomics, glycomics, and MS imaging are all generally based on MS [12].

Several biological fluids have been used for cancer biomarker research, including saliva [13], urine [14], nipple aspirate fluid [15], cerebrospinal fluids, and tumor intestinal fluids [16]. However, serum and plasma are the most commonly used human biological fluids for cancer biomarker research partly because obtaining blood is relatively noninvasive compared with other methods like biopsy [9, 17]. Additionally, blood can provide multiple molecular elements of cancer in the form of circulating cells, proteins, peptides, metabolites, and cell-free DNA and RNA [18]. The proteins in serum and plasma have been the major target analytes for cancer diagnosis. Advancements in proteomics technologies have enabled more accurate profiling of proteins [19]. However, the dominance of a few proteins in serum/plasma and the presence of post-translational modifications make their complete characterization more challenging than analysis of the genes [20].

Protein modifications, which occur as translational, post-translational, regulatory, and/or degradation products, increase the amount of useful cancer-related information that can be obtained from the proteins. The common post-translational modifications (PTMs) include but are not limited to glycosylation, phosphorylation, sulfation, and acetylation. These PTMs play important roles in biological processes, including regulation of proper protein folding, host pathogen interactions, immune responses, and also in inflammation. In this review, we highlight the recent efforts in identifying potential glycosylated proteins and released glycans as cancer biomarkers using mass spectrometry (MS). We will begin with a basic introduction to the mass spectrometry technology and its utility in identifying single glycoprotein markers, panels of glycoprotein markers, and released glycan markers. We will then discuss advancements in site-specific glycosylation analysis for cancer diagnosis.

Mass spectrometry

Mass spectrometry is a versatile technique that has increasingly been used in cancer biomarkers discovery because of its

sensitivity, potential for high throughput, and the capability for tolerating diverse biological mixtures [21–27]. MS has been used in the analysis of various analytes in many fields, including but not limited to proteomics, metabolomics, lipidomics, glycoproteomics, glycomics, and proteoglycomics [28–31]. Samples can be introduced to the mass spectrometer via direct infusion; however, most of the potential cancer markers are found in biological mixtures requiring some pre-separation method. Separation techniques such as liquid chromatography (LC), ion mobility, and capillary electrophoresis (CE) are typically used in conjunction with mass spectrometry.

Recent developments in microchip-based separation and ionization methods have enabled analysis of extremely small amounts of samples with high sensitivity and reproducibility [32, 33]. This capability is particularly important because most of the analytes in biological samples are found in minute amounts that are not readily amenable to other analytical techniques. MS analysis generates a tremendous amount of data that requires the development of bioinformatics methods that can aid with data interpretation and expedite the analysis of large data sets [34, 35].

Lone (Glyco)protein markers for cancer

Most of the Food and Drug Administration (FDA) approved cancer biomarkers are single proteins derived from serum [36], and the majority of these proteins are glycosylated. CA 125 antigen, for example, is a well-documented glycoprotein standard marker for ovarian cancer [37–39]. CA 19-9 has been used as a pancreatic cancer biomarker [40]. CA 15-3 is a useful breast cancer biomarker, especially for monitoring patients who have already been diagnosed with the disease in evaluating the recurrence of the disease [41].

Other glycosylated protein markers include prostate-specific antigen (PSA), which is present in elevated amounts in seminal fluids and plasma, for screening and monitoring prostate cancer patients [42, 43]. Carcino embryonic antigen (CEA), which has been found to change in colorectal, bladder, breast, pancreatic, and lung cancer, is highly N-glycosylated [44–46]. Multiple studies have also indicated that haptoglobin, which is also glycosylated, could be used as a cancer biomarker [47–49]. In one study using in-house ELISA, haptoglobin was found to be correlated with the clinical stage of epithelial ovarian cancer and patient survival [50]. Using statistical analysis, the study found that elevated serum haptoglobin levels were associated with poor outcome in overall patient survival. The isoforms of haptoglobin-1 precursor have been found to be significantly upregulated in all

groups of ovarian cancer patients by other researchers [51].

Another study that identified single glycoprotein potential cancer markers utilized 2-D differential gel electrophoresis (2-D DIGE) and MALDI-TOF/TOF to identify depleted serum markers for breast cancer. Proapolipoprotein A1, transferrin, and hemoglobin were found to be elevated in breast cancer compared with control samples while apolipoprotein A1, apolipoprotein C-III, and α -2-haptoglobin were decreased. Immunochemical reactions were used to validate the selected candidate markers. In this work, it was demonstrated that 2-D DIGE can separate protein isoforms that are not amenable to immunochemical separation due to lack of isomer specific antibodies [52].

(Glyco)protein panels as markers for cancer

A single protein biomarker may not provide the specificity and sensitivity necessary for disease identification. In this regard, identifying a number of individual potential markers that, when combined, would identify a certain disease condition is a common approach. Although it is a huge challenge to assemble and validate a panel of biomarker proteins [17], Table 1 shows representative studies that have identified panels of potential cancer protein markers that are all post-translationally modified. Ralhan et al. [53] identified a panel of proteins by multidimensional LC-MS/MS that were differentially expressed when head and neck cancer patient samples were compared with the controls. For the analysis, the samples were labeled with isobaric mass tags for relative and absolute quantitation (iTRAQ) on a QSTAR Pulsar i MS. Using this approach, a panel of three cancer protein markers that are potentially modified on their serine residues were identified that achieved remarkable sensitivity of 0.92 and specificity of 0.91: 14-3-3 protein zeta/delta (phosphorylated), stratifin

(phosphorylated), and S100-A7 (some serine residues are acetylated). These MS results were confirmed using immuno-precipitation. When evaluated individually, the sensitivity and specificity of each of the proteins were poor [53].

In another study combining 2D-DIGE and MALDI-TOF analysis, serum samples from 50 lung cancer patients and 50 controls were screened for protein biomarkers. A panel of four proteins was able to discriminate healthy controls from patients who had lung cancer. This panel of four biomarkers included carcinoembryonic antigen, retinol binding protein, α -1-antitrypsin, and squamous cell carcinoma antigen. The first three proteins are glycosylated whereas the latter is phosphorylated. The authors used an independent blinded validation set of sera from 49 lung cancer patients and 48 matched controls. The protein markers were able to classify most of the cancer patients in the independent validation set with sensitivity of 77.8 % and specificity of 75.4 % [54].

Selective capture of glycoproteins using lectins or other glycan-binding antibodies is another complementary method for the discovery of potential glycoprotein biomarkers. Lectins are proteins that recognize and bind specific glycan moieties and can be used to isolate target glycoproteins from a sample mixture. Thus, they provide the advantage of minimizing the number of compounds to be analyzed by MS [55–57]. Lectin-based methods have been used to detect differential glycosylation between cancer and healthy controls [58–61]. In a recent study, a glycoprotein binding lectin array was used to identify glycosylation changes in the sera of patients with benign diseases compared with those with stage III ovarian cancer. Lectins whose binding response differed significantly between benign conditions and cancer were used to pull isotopically tagged glycoproteins for cancer biomarker screening. Quantitative LC-MS/MS on an LTQ-ion trap was used to distinguish between 15 healthy controls, 19

Table 1 Contains a group of proteins that have been identified as potential biomarkers when combined. The columns represent the cancer type, the protein marker panel, the analytical technique that was used for the analysis, and the reference number in the review

Cancer type	Panel of protein markers	Technique	Reference no.
Lung	Carcinoembryonic antigen, retinol binding protein, α -1-antitrypsin and squamous cell carcinoma antigen	2D-DIGE and MALDI-TOF	[54]
Ovarian	Corticosteroid-binding globulin (CBG), serum amyloid p component (SAP) and complement factor B (CFAB)	Lectin array, and quantitative LC-MS/MS	[58]
Head and neck	14-3-3 protein zeta/delta (YWHAZ), stratifin and S100-A7	Multidimensional LC-MS/MS	[53]
Pancreatic cancer	α -1-antichymotrypsin (AACT), thrombospondin-1 (THBS1) and haptoglobin (HPT)	Label-free and TMT strategies, LC-MS/MS	[67]

benign ovarian diseases, 21 stage I/II ovarian cancer, and 30 stage III ovarian cancer cases. A panel of four glycoproteins, including corticosteroid-binding globulin (CBG), serum amyloid p component (SAP), complement factor B (CFAB), and histidine-rich glycoprotein (HRG) were found to differentiate ovarian cancer patients from benign diseases and healthy controls. Combining CBG, SAP, and CA 125 showed improved discrimination of stage III ovarian cancer from benign diseases compared to CA 125 alone (Fig. 1) [58].

N-acetylglucosaminyltransferase-V (GnT-V) catalyzes the formation of an unusual β -1,6-linked *N*-acetylglucosamine, which has been found to be upregulated in cancer cells [62–64]. A breast cancer investigation used the lectin phytohemagglutinin-L4 (L-PHA), which binds specifically to β -(1,6)-branched N-linked structures on glycoproteins [65]. This lectin showed selective binding to breast epithelial cells as the disease progresses to invasive carcinoma, but not to normal healthy cells. Using this lectin, β -1,6-*N*-acetylglucosamine (β -1,6-GlcNAc) branched N-linked glycan-containing glycoproteins were enriched from the tissues of normal healthy controls and invasive carcinoma patients. In contrast to the healthy controls, 12 glycoproteins were enriched in the tissues of all four patients with invasive ductal breast carcinoma tissues [65].

Tandem mass tags (TMT) are chemical labels used to quantify multiple peptides or proteins in biological samples simultaneously [66]. Quantitative proteomics using TMT labeling was performed to identify a panel of three glycoproteins that could potentially distinguish pancreatic cancer and other pancreatic conditions like diabetes, cyst, chronic pancreatitis, and obstructive jaundice from controls [67]. Depletion of the 14 most abundant proteins was first carried out and fucosylated glycoproteins were extracted using agarose-bound *Aleuria aurantia* lectin (AAL). A combination of α -1-antichymotrypsin (AACT), thrombospondin-1 (THBS1), and haptoglobin (HPT) showed good sensitivity in diagnosing pancreatic cancer compared with the following conditions: normal controls (AUC = 0.95), diabetes (AUC = 0.89), cyst (AUC = 0.82), and chronic pancreatitis (AUC = 0.90). Adding CA 19-9 to the panel distinguished pancreatic cancer from patients with and without obstructive jaundice [67].

The studies highlighted in the above discussions are additional proofs that modified proteins have an influence in cancer, and their modifications could potentially be used as cancer biomarkers. Owing to this fact, substantial method development efforts are being made to identify the protein modifications and profile their levels in biological fluids in a high throughput manner. Some of these efforts have begun to bear fruits but a lot remains to be

Fig. 1 ROC curve analyses of significantly changed proteins detected by ELISA. (a) ROC curve for differentiating stage III ovarian cancer from normal healthy controls; (b) ROC curve for differentiating stage III ovarian cancer from benign diseases; (c) ROC curve for distinguishing stage I/II ovarian cancer from normal healthy controls. Reprinted from ref. [58]. Copyright © 2012, American Chemical Society

done. In the subsequent text we highlight some of the progress made to identify and profile protein modifications, especially O- and N-linked carbohydrates in the search for cancer biomarkers.

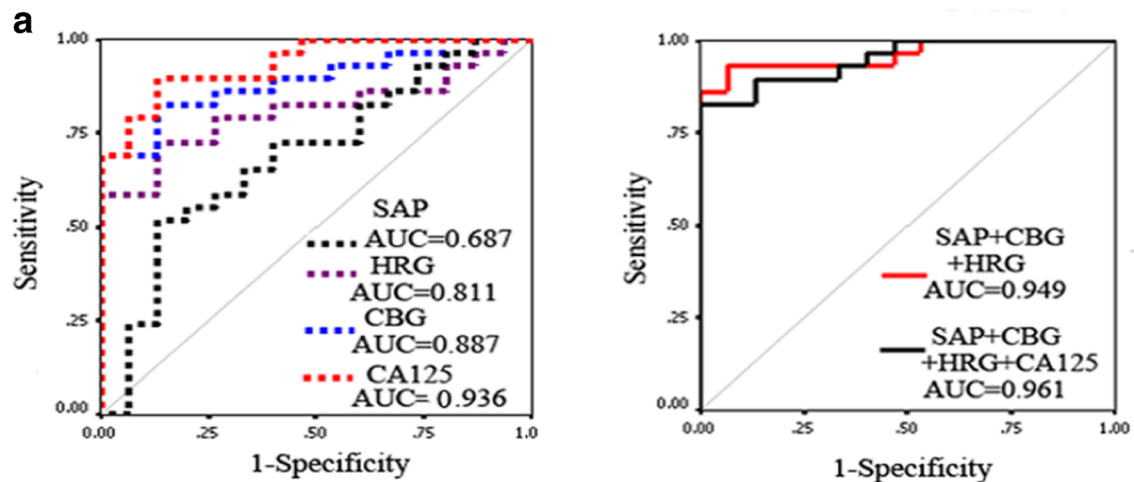
Post-translational modifications of proteins in cancer

PTMs of proteins include the covalent attachment of the molecules to amino acid(s) enzymatically during protein biosynthesis [68]. PTMs provide valuable additional details to the biological functions of proteins. The most common PTMs include the following: glycosylation (N- and O-linked), phosphorylation, acetylation, methylation, and sulfation [68, 69]. Differential expression of modifications between normal and cancerous cells occurs at the onset or during cancer progression [70,71]. These changes could be due to a single modification or multiple modifications.

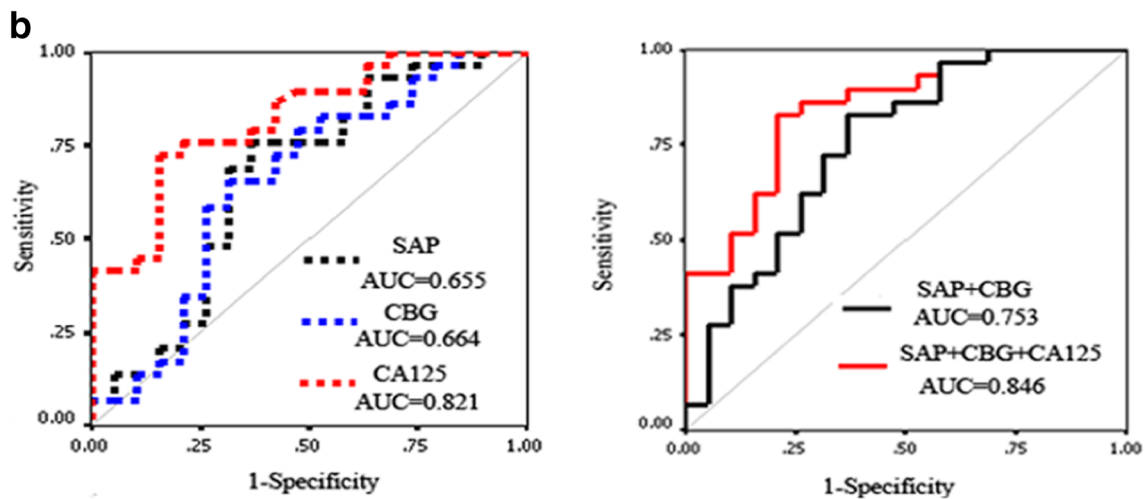
Glycosylation is a substantial PTM with respect to size and occurrence, where more than 50 % of proteins are estimated to be glycosylated [25, 72]. For humans, this value may be larger, and for some fluids such as human milk, glycosylated proteins may be as high as 80 % [73]. Glycans are often found attached to proteins on the cell surface and in the extracellular matrix where they mediate many biological activities [74, 75]. Major types of glycans include N-linked glycans attached to the nitrogen atom in the asparagine side chain within a consensus amino acid sequence Asn-X-Ser/Thr (X should not be proline), and O-linked glycans attached to the oxygen atom of several amino acid residues including serine and threonine. Other types of glycans include glycosaminoglycans usually found attached to the proteins (proteoglycans) and also lipid chains as in glycolipids [76]. There is now a considerable body of work showing that glycans released from cancer cells differ from those in healthy cells [7, 75, 77–80]. Tumors have been found to overexpress certain types of glycoproteins and glycolipids [79, 81]. For example, there is an increase in the size and branching of N-glycans in tumors [75]. High mannose glycans have also been found to increase in other cancers (Fig. 2) [82].

Glycan analysis is more challenging compared with proteins. The non-template nature of their biosynthesis and involvement of a wide variety of essentially competing enzymes during their synthesis leads to more diverse

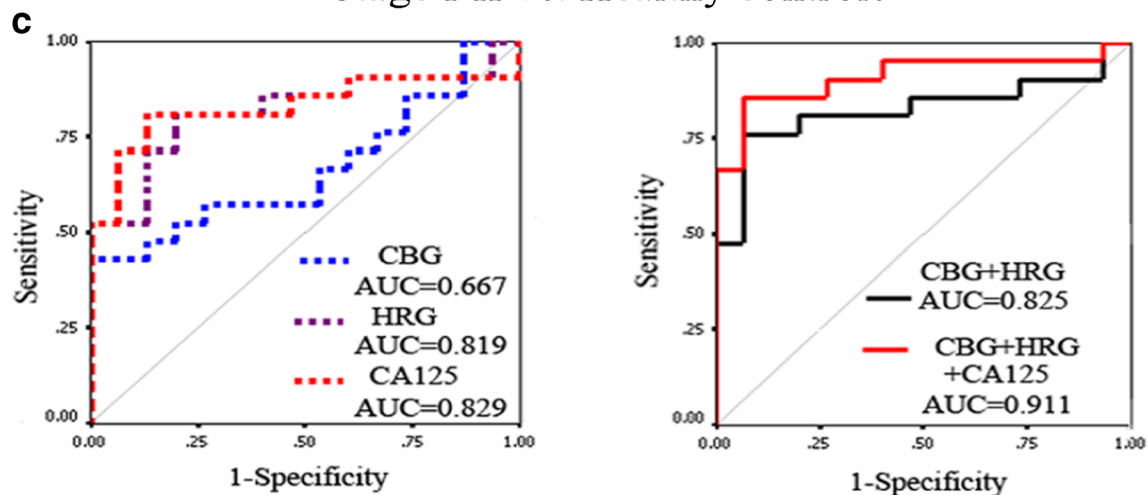
Stage III vs. Healthy controls



Stage III cancer vs. Benign diseases



Stage I/II vs. Healthy controls



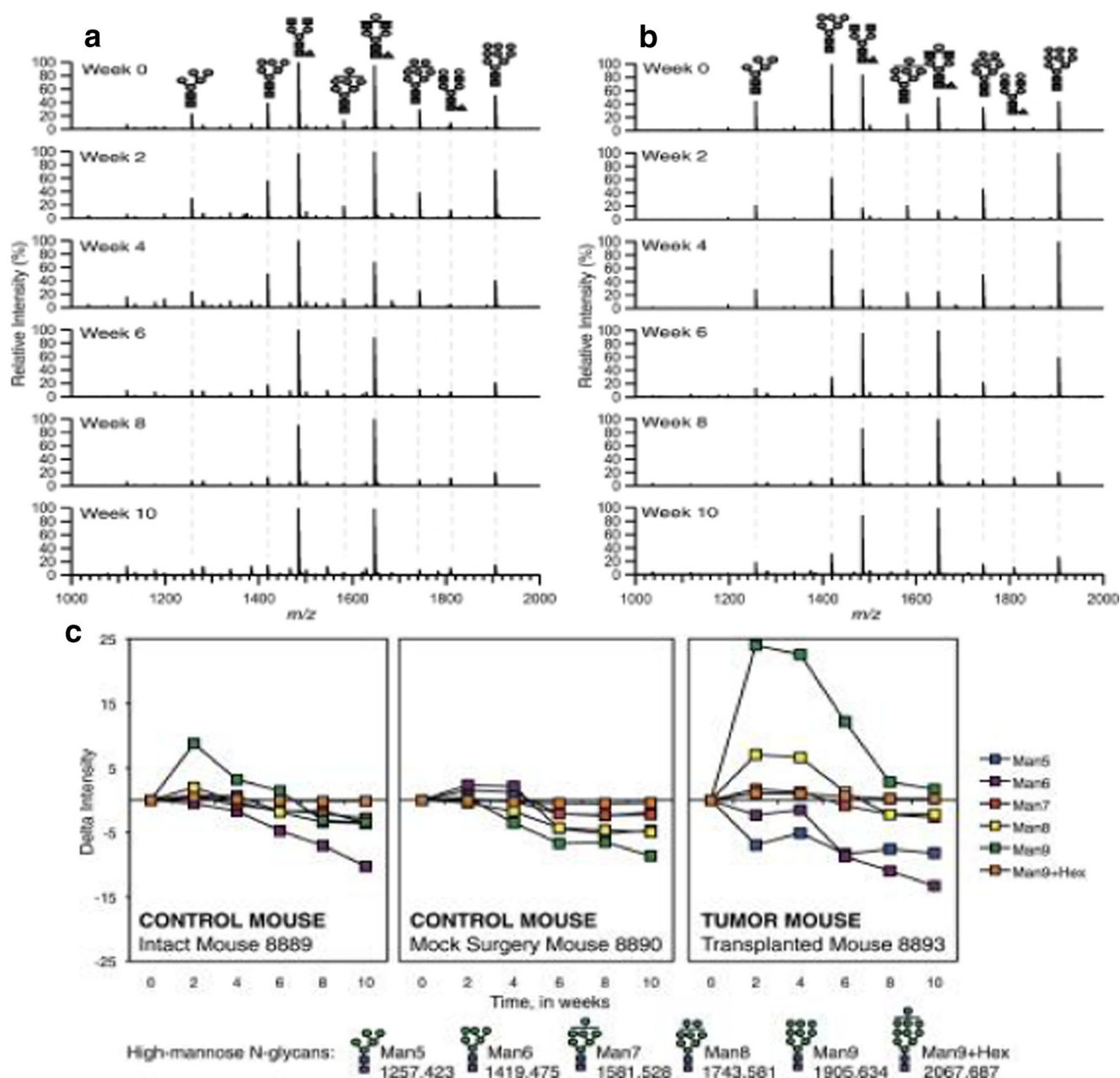


Fig. 2 Representative MALDI FT-ICR mass spectra of **A**, mock surgery control and **B**, tumor-transplanted mouse sera at different points; m/z 1000–2000 region of 10 % acetonitrile fraction taken in the positive ion mode is shown. **C**, Change in intensities from MALDI FT-ICR MS of high-mannose N-linked glycans in sera of an intact control mouse, mock surgery control mouse, and a tumor-transplanted mouse during breast

cancer progression. All values relative to week 0 intensity. Error bars are expressed as standard error of the mean (S.E.) from three spectral scans per mouse sample. Symbol representations of glycans: N-acetylglucosamine, blue square; mannose, green circle; glucose, blue circle. Reprinted from ref. [81]. Copyright © 2011, by the American Society for Biochemistry and Molecular Biology

glycan molecules having minor structural differences. Glycans are, however, easier to quantitate in relative amounts. Recent developments in liquid chromatography separation, ionization, and mass spectrometry methods have enabled extensive characterization of oligosaccharides and high throughput analysis of larger sample sizes [83]. These developments could potentially lead to the discovery of glycan specific cancer biomarkers.

Global N- and O-glycan analysis

Efforts to identify cancer biomarkers from released glycans have been reviewed recently [40, 78, 84, 85]. The common approach is a global analysis of glycosylation involving cleaving the glycans from the proteins using enzymatic or chemical means. Usually, peptide-N-glycosidase F (PNGase F) enzyme is used to cleave the N-glycans from purified proteins or from

biological mixtures. O-Glycans can be cleaved from the proteins using alkaline sodium borohydride (NaBH_4) in a β -elimination reaction, whereas both N- and O-glycans can be cleaved simultaneously using hydrazinolysis (Fig. 3) [78, 86]. This step is usually followed by glycan enrichment using solid phase extraction (SPE) on a porous graphitized carbon (PGC) or hydrophilic interaction chromatography (HILIC) column. The samples are then analyzed using methods such as MALDI-TOF, MALDI-Fourier transform ion cyclotron resonance (FTICR), Chip-TOF, and Chip-Q-TOF mass spectrometry for compositional or structural analysis. The glycans can be analyzed in their native state or chemically reduced to minimize the multiplicity of MS peaks due to different glycan anomeric configurations, especially when a high resolving column (e.g., PGC) is used for separation. Glycans may also be derivatized through permethylation or acetylation to achieve uniform ionization in the mass spectrometer [33, 86, 87]. There are numerous important studies involving glycan analysis in the search for cancer biomarkers [33, 40, 78, 85, 86, 88–91].

MALDI-MS has been widely used to profile glycans as potential cancer biomarkers because of its sensitivity and high ionization efficiency even at higher mass range [83, 88, 92–98]. In a study to identify distinguishing features between ovarian cancer patients and healthy persons, O-glycans were cleaved from conditioned media and from serum and profiled using high resolution MALDI-FTICR-MS. Tandem mass spectrometry was used for further structural identification of glycans using infrared multiphoton dissociation (IRMPD). Several unique glycans observed in ovarian cancer patient serum samples, mostly neutral structures composed only of hexose and *N*-acetylhexosamine residues, were absent in the controls [88]. This work also investigated four ovarian cancer cell lines, OVCAR-3, Caov-3, SK-OV-3, and ES-2, and

obtained similar glycan profiles. Multiple oligosaccharides observed in the supernatants of the cancer cell lines were also observed in the patient sera. This was one of the first studies to show that glycans harvested directly from cells or serum can be used as potential cancer markers without requiring isolation of proteins using antibodies [88]. Recent work has identified stage-specific N-glycans released from case-control plasma samples that could potentially be used as ovarian cancer markers [99].

In a related study, O-glycans were chemically cleaved from breast cancer tumor cell lines, sera from mice transplanted with mammary tumors, sera from breast cancer patients, and those collected from healthy individuals. This was accomplished using β -elimination with sodium borohydride dissolved in sodium hydroxide solution [100]. High resolution MALDI-FTICR-MS and IRMPD were used to characterize the glycan structures. Differences in intensities were found in the O-glycan profiles that distinguished breast cancer patients from the controls. This study also identified changes in specific structures during tumor growth, which may be used as potential indicators of disease onset [100].

Global release methods have also been conducted for the analysis of N-glycans. Using MALDI-MS, permethylated N-glycans from 10 healthy males and 24 prostate cancer patients were compared using statistical methods, principle component analysis (PCA), ANOVA) and receiver operating characteristic (ROC). Out of the 50 N-glycans identified, 12 glycan structures, half of which were fucosylated, were found to vary significantly between disease and control specimens [95]. The relative abundances of high mannose and complex biantennary structures decreased in cancer while the fucosylated complex biantennary and complex tetraantennary N-glycans were elevated in cancer in comparison to the healthy controls. A separate study by the same group noted increased fucosylation of serum N-glycans in breast cancer patients [96].

A recent study investigated the possibility of using permethylated N-glycans from serum and isolated immunoglobulin G (IgG) as ovarian cancer markers. The samples were obtained from healthy controls and from patients with late-stage recurrent ovarian cancer (OC) before they started an experimental drug treatment (baseline) and prior to the second treatment cycle. Compared with the controls, the level of bisecting glycans decreased in late stage recurrent ovarian cancer patients, and this trend was accentuated in patients who underwent the first cycle of the trial treatment. Triantennary and tetraantennary glycans that were sialylated and fucosylated seemed to increase in the baseline OC samples and also showed further increases in the patients' serum after the experimental drug treatment [101]. Given that patients exhibited increased tumor burden post-experimental treatment, these glycan changes were further assessed as potential markers

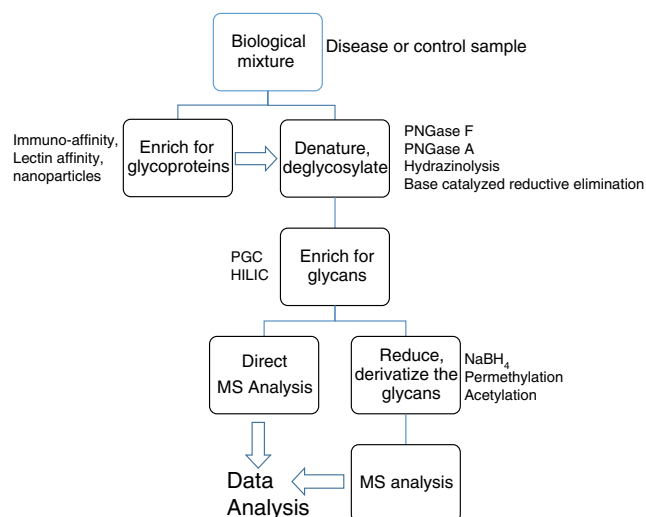


Fig. 3 The general procedure for the released glycan analysis

for disease progression. Contrastingly, when IgG-specific glycans were investigated, fucosylation seemed not to be a strong determinant. Galactosylated glycan levels decreased whereas the agalactosylated glycans increased in patients with OC when compared to the healthy controls [101]. Other studies have reported increased agalactosylated IgG glycans in ovarian cancer [102], lung cancer [103], and non-small cell lung cancer (NSCLC) patients [104].

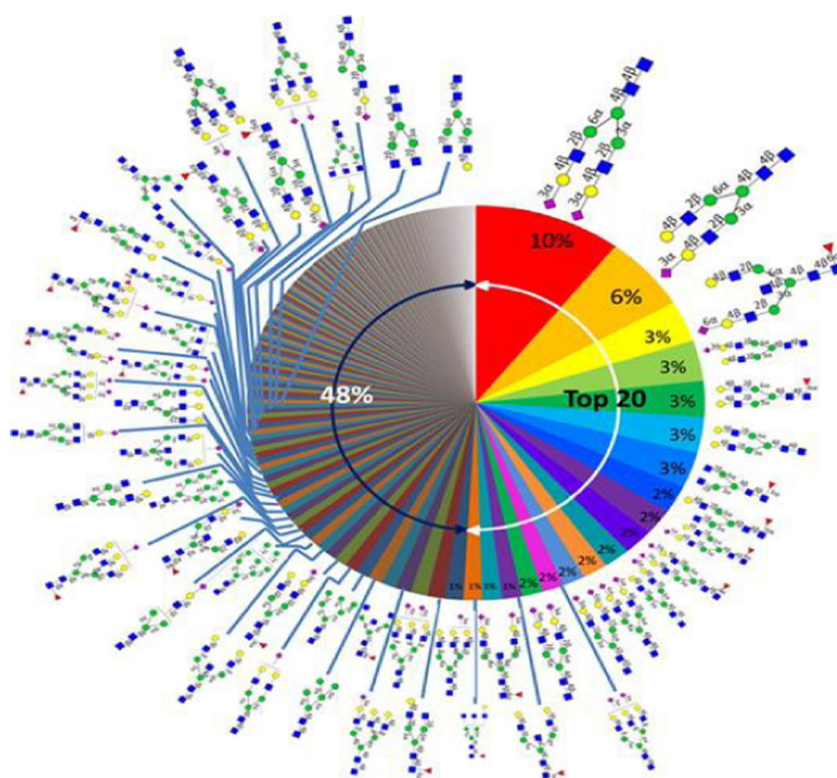
As researchers seek more specific and sensitive markers for cancer, there is a greater need to increase the amount of information obtained from the samples. In this regard, efforts to identify the structures of the glycan biomarkers are gaining interest. Porous graphitized carbon (PGC), which separates glycans based on their size, polarity, and three-dimensional structure, is a good platform for released glycan separation. PGC has high sensitivity, low sample consumption, minimal ion suppression, and large instrumental dynamic range, which are all important for clinical glycomics [105–107].

A novel method that uses rapid throughput nano-flow liquid chromatography on a PGC column to separate isomeric glycans from serum was recently developed. Using this method, more than 100 unique N-glycan compositions having more than 300 glycan species were identified, and their abundances analyzed in two classes of prostate cancer patients. It was found that fucosylated, non-sialylated complex/hybrid glycans were significantly higher in the good prognosis (G) group (four subjects)

compared with the four subjects in the poor prognosis group (P). In addition, sialylated, non-fucosylated complex/hybrid glycans were more abundant in the P group compared with the G group [105]. This strategy increases the level of information obtained for biomarker analysis and may lead to structural specific markers.

Advanced chip-based LC separations have also enabled the analysis of N-linked glycans from serum with very high reproducibility [108]. A library composed of serum N-glycans, including fine details like the residue linkages, was recently created [109]. The library was created by first utilizing PNGase F to release the glycans from whole serum. The glycans were then reduced using NaBH₄ and enriched using graphitized carbon cartridges. The samples were analyzed using nano-LC-chip-TOF MS. The retention times and accurate mass derived from the library enable rapid sequencing of serum N-glycans. This library provides additional structural specific information that can be very useful in cancer biomarker research. The library contains diverse N-glycan groups, including high mannose, sialylated, fucosylated, biantennary, and triantennary complex and hybrid structures (Fig. 4). The library included approximately 50 complete structures and 170 compositions with nearly complete structures were elucidated using this method. This technology provides a deeper level of analysis of the glycan structural species that change with disease, which aids in understanding the mechanisms underlying these biological changes [109].

Fig. 4 Human serum “glycan wheel” based on the relative abundances averaged from nine individual sera. Reprinted with permission from Ting Song, Danielle Aldredge, Carlito B. Lebrilla. ref. [108]. Copyright © 2015, American Chemical Society



Protein-specific glycan markers

Another way to analyze glycosylation is to first isolate the possible glycoprotein targets from biological mixtures, cleave off the glycans, and analyze them using mass spectrometry [89, 90, 110]. The glycans obtained from this method are assumed to be from the glycoproteins isolated. This procedure does not provide site-specific information about the glycoforms. Nonetheless, the information obtained is useful in protein-specific biomarker studies. This method was used to investigate possible glycan stomach cancer markers by first isolating the most abundant serum proteins IgG, haptoglobin, transferrin, and α -1-acid glycoprotein using affinity purification or 2D gel electrophoresis. The study employed LC-MS/MS to identify the proteins and then ultra-performance liquid chromatography (UPLC) with UV detection to separate and identify the glycans on a HILIC stationary phase. Core fucosylated biantennary glycans were found to be elevated in IgG with increased disease progression while sialylated and galactosylated core fucosylated biantennary glycans decreased with disease progression. Contrastingly, sialylated glycans in haptoglobin, transferrin, and α -1-acid glycoprotein increased with disease progression [90].

Another glycan study was conducted to distinguish between non-atrophic gastritis (NAG), duodenal ulcer (DU), and gastric cancer (GC) cases. IgG was isolated from the serum using protein G. The N-glycans were released from IgG using PNGase F, purified by PGC, and analyzed using nHPLC-chip-TOF-MS. A total of 48 IgG glycans were observed in this research. Statistical data analysis showed that among these, eight glycans were significantly altered between GC and NAG. These changes include increased levels of truncated glycans and decreased levels of biantennary galactosylated glycans in the GC group compared with the NAG group. In addition, lower sialylation was observed in GC compared with the DU and NAG [110].

In a pancreatic cancer study, a Western blot using *Aleuria aurantia* lectin (AAL), which preferentially binds to fucosylated structures, was used to identify fucosylated proteins in the serum of cancer patients and healthy controls. Among these proteins, fucosylated haptoglobin was found to be elevated in pancreatic cancer patients compared with the controls. Subsequently, haptoglobin was purified from serum and analyzed using LC-ESI-MS and MALDI-TOF, which showed that α -(1-3)-, α -(1-4)-, and α -(1-6)-fucosyl linkages increased in pancreatic cancer patients compared with the controls [48]. Although fucosylation has been identified to change in other diseases, the elevation of these analytes was more pronounced in pancreatic cancer patients. It has also been observed that haptoglobin- α subunit levels change with the small-cell lung cancer stage and could also be a potential biomarker [89].

From the above discussion, it is clear that global glycan analysis provides useful information about potential cancer biomarkers. Isolation of specific glycoproteins from biological sample mixtures is also gaining interest. From the glycosylation analysis of IgG isolated from serum, similarities in glycan patterns are observed in a protein-specific manner in different cancer types. For instance, increased levels of truncated glycans without galactose have been reported in gastric cancer, ovarian cancer, and lung cancer. Similar studies from other serum glycoproteins may uncover protein-specific glycosylation changes that may serve as potential cancer indicators.

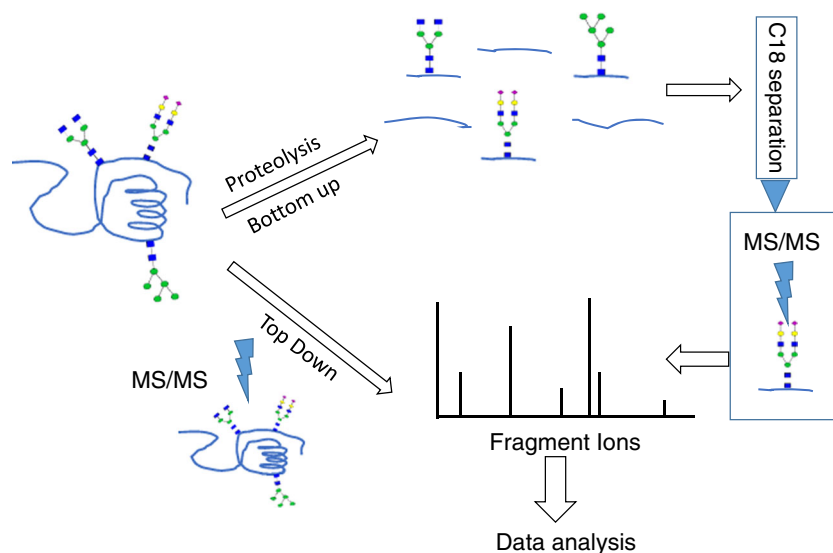
Site-specific glycan markers

Although the analyses of global glycosylation changes provide useful information about cancer and other diseases, one drawback of this method is that the glycan cannot be traced back to the protein site of origin. In order to identify subtle glycosylation differences between disease conditions and gain more knowledge about cancer biology, it is important to not only identify the glycan but also its site of attachment within the protein. This information is obtained by thoroughly investigating the glycosylated peptides (glycopeptides) identified from tandem mass spectrometry of the glycoproteins.

Glycopeptide analysis is more challenging than protein or global glyco-analysis due to several reasons. A single glycan composition in a peptide may contain a large number of isomeric structures because of differences in glycosidic linkages between monosaccharides, branching, and the existence of many monosaccharides bearing the same mass. The presence of multiple unique glycans that share the same peptide backbone causes the MS signal to be split into various glycoforms, thus lowering their individual abundances compared with the peptides that are not glycosylated. It is also challenging to obtain comprehensive glycopeptide fragmentation for both the glycan and the peptide moieties as they have different fragmentation efficiencies [111]. Nonetheless, recent advancements in mass spectrometry technologies have enabled characterization of the glycans with their peptide backbone [112–114].

Owing to the above-mentioned problems, it has been challenging to develop algorithms that can automatically identify glycans and their peptide moieties from the tandem MS data. Despite these limitations, progress is being made in our laboratory and elsewhere to develop software that can identify glycopeptides spectra [111, 115–122]. Site-specific glycosylation information can be obtained directly from tandem mass spectrometry of intact proteins (top-down) [123]. Molecular ions from intact glycoproteins are selected and subjected to multiple stage MS (MS^n) analysis (Fig. 5). This method has been limited to few types of MS instruments that can detect

Fig. 5 The glycoprotein and glycopeptide analysis procedure



ions at high mass range and the ones equipped with electron-based fragmentation techniques. These techniques include electron capture dissociation (ECD), electron transfer dissociation (ETD), and negative electron transfer dissociation (NETD) that retain labile protein modifications in the peptide backbone [124].

Alternatively, another common strategy for site-specific glycosylation analysis involves digesting the proteins using specific or nonspecific proteases to produce glycosylated and non-glycosylated peptides (bottom-up). The samples can then be analyzed by MS directly to identify the glycopeptides. However, since the glycopeptides ionize relatively less efficiently compared with the non-glycosylated peptides, a fractionation step can be added to enrich for the glycopeptides [125]. The fractionation step, although a useful step, may affect the levels of analytes in different proportions, which may lead to inaccurate quantitative information. Tandem mass spectrometry (MS/MS or MSⁿ) is used to fragment glycopeptides to help in identifying both the glycan and its peptide moiety. Diagnostic glycan and the peptide backbone fragments are necessary for unambiguous identification of the site of glycosylation [112, 126]. Another aspect of glycosylation analysis is the extent to which the site of glycosylation is occupied with the glycans (site occupancy). The common method to measure the glycosylation site occupancy involves releasing the N-glycans using PNGase F, which deaminates the glycosylated asparagine residues (N) to aspartic acid (D). The mass difference between the N and D (0.984 Da) allows them to be distinguished using MS [127].

The assignment of the glycopeptides is usually carried out using computer algorithms designed to identify the unique fragmentation patterns of the glycopeptides [112, 128]. Glycopeptides produce abundant glycan diagnostic protonated fragments such as HexNAc (204.08 Da), Hex (163.06 Da), Hex + HexNAc (366.14 Da), Neu5Ac-2H₂O (274.09 Da), and

Neu5Ac-H₂O (292.10 Da). We provide a few but not exhaustive examples where site-specific analysis of the glycans provided useful information about potential glycopeptide biomarkers.

In a recent study, the protein clusterin was pulled from the plasma of clear cell renal cell carcinoma (ccRCC) patients before and after curative nephrectomy for localized disease using immunoaffinity purification. Site-specific N-glycan heterogeneity analysis was carried out using LTQ-Orbitrap elite tandem mass spectrometry. Biantennary digalactosylated disialylated oligosaccharides with and without fucose found in glycosite N374 were significantly decreased in disease RCC (+) samples [129].

As mentioned earlier, global analysis of the glycosylation of haptoglobin show elevation of fucosylation in pancreatic cancer patients' serum [48]. Another study was carried out to identify the sites where fucosylation changed. Among the four sites of glycosylation, fucosylation of the biantennary glycans in two sites (N207 and N241) increased, whereas that of the triantennary glycans increased in all the sites [49].

A method for screening and quantifying core fucosylated (CF) peptides in the depleted serum was recently developed [130]. The method uses lens culinaris agglutinin (LCA) to enrich for the core fucosylated peptides. The presence of metal ions Na⁺, Ca²⁺, and Mn²⁺ increased the lectin binding. The same metal ions were added during the elution step to increase the elution of CF. This work identified 630 sites containing core fucosylated structures in 322 proteins using the Orbitrap Elite MS. Additionally, eight potential CF glycopeptide cancer markers that differed between pancreatic cancer and healthy controls or chronic pancreatitis were identified.

Another method to measure the level of core fucosylation in α -2-macroglobulin (A2MG) extracted from serum using immunoprecipitation was recently developed. This method utilizes endoglycosidase F to release a portion of the glycan,

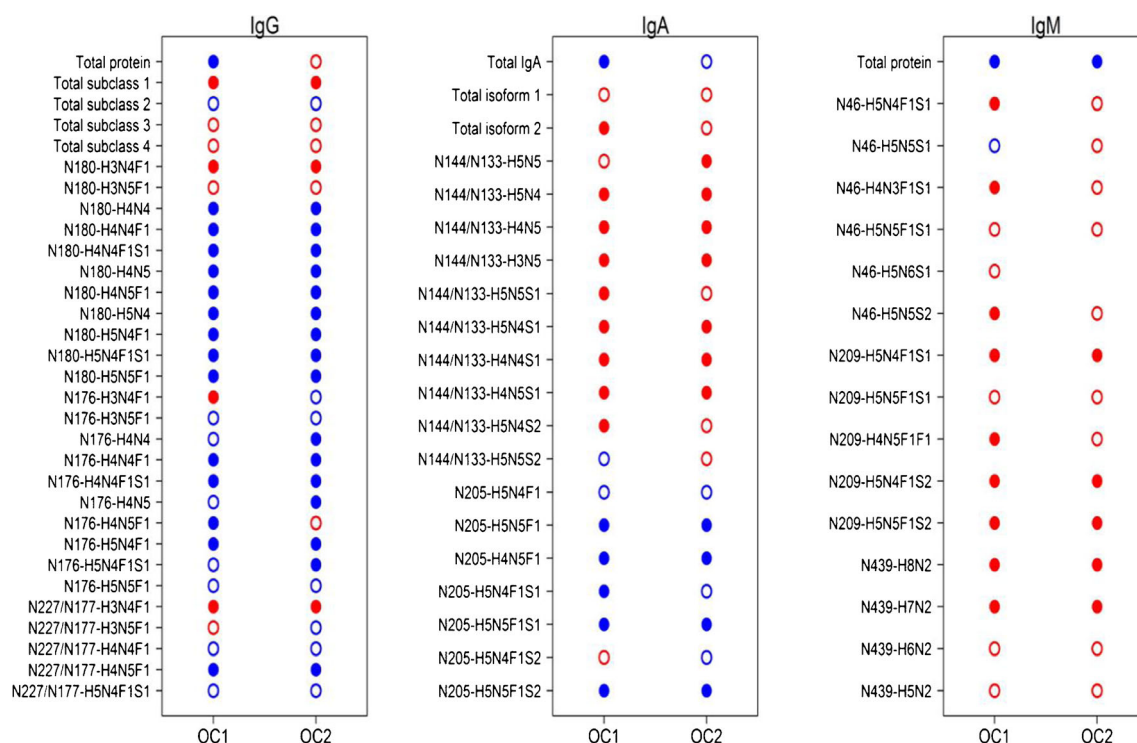


Fig. 6 Differential analysis of peptide and glycopeptide variables from immunoglobulins in epithelial ovarian cancer (EOC). Closed dots indicate significantly different abundance levels between EOC cases and healthy controls, whereas open dots indicate no significance was achieved at the false discovery rate (FDR) < 0.05. Red dots indicate increased levels in EOC cases compared with controls, whereas blue

dots indicate decreased levels in EOC compared with controls. Reprinted with permission from ref. [137]. L. Renee Ruhaak, Kyoungmi Kim, Carol Stroble, Sandra L. Taylor, Qiuting Hong, Suzanne Miyamoto, Carlito B. Lebrilla, and Gary Leiserowitz J. Proteome Res (2016) 15:1002–1010. Copyright © 2016, American Chemical Society

leaving the inner core GlcNAc residue attached to the peptide. The heterogeneity due to multiple glycan structures per peptide was reduced, the ionization efficiency was increased, and the analysis of the data was simplified [131]. Nano-LC-linear trap quadrupole (LTQ)-MS was used to analyze the samples. The study found lower levels of core-fucosylation at sites N396 and N1424 in both chronic pancreatitis and pancreatic cancer compared with normal controls [131]. The method used in this work can be used for researchers targeting core fucosylation, which has been found to change in different types of cancer [48, 132, 133].

Multiple reaction monitoring for glycopeptide analysis

Multiple reaction monitoring (MRM) is a targeted analytical method used for quantitation of biomolecules. MRM is usually performed on a triple quadrupole mass spectrometer (QqQ). The precursor of interest is isolated in the first quadrupole, fragmented in the second, and only specific fragment ions are transmitted through the third quadrupole for detection. This method can quantify analytes from biological samples with high accuracy and precision. MRM, which has been widely applied in quantification of small molecules and also

proteins, is gaining interest in the glycoproteomics field [134–139].

Stable isotope standards and capture by anti-peptide antibodies (SISCAPA) is a useful technique to enrich specific peptides for quantitation [140]. This method was used in combination with the lectin L-PHA (described above) to investigate aberrantly glycosylated metalloproteinase inhibitor 1 (TIMP1) as a marker for colon cancer [141]. First, L-PHA was used to enrich for serum glycoproteins containing β 1,6-GlcNAc N-linked glycans. A monoclonal anti-peptide TIMP1 antibody coupled to magnetic beads was used to enrich for a specific TIMP1 peptide that was subsequently analyzed using MRM with isotopically labeled standards. Using this method, an aberrantly glycosylated TIMP1 isoform was quantified at a level of 0.8 ng/mL in the serum of colorectal cancer patients.

An MRM method to quantify IgG and its subclasses directly from un-depleted serum was recently developed in our laboratory [137]. In this method, trypsin digested IgG glycoprotein standard was used to optimize the QqQ-MS response for glycopeptide analysis. The LC solvent gradient, the MS front end ionization parameters, and the fragmentation energy for the glycopeptides were optimized. The method is able to measure the amount of the protein and how its site-specific glycosylation changes between samples. Robust nonglycosylated peptides found in all the IgG subclasses were used to obtain

the total amount of IgG in the serum. The glycopeptides from the four IgG subclasses (IgG1, IgG2, IgG3, and IgG4) were normalized to the response of their respective protein subclasses. Normalizing the glycopeptide to the level of the protein helps to identify the actual glycosylation changes irrespective of the changes in the level of protein expression between samples [137]. This method has been extended to include other glycoproteins in serum [142] and it is currently applied in our laboratory to screen for disease biomarkers for cancer [143] autoimmune conditions and other diseases. A recent study investigated the immunoglobulins in ovarian cancer patients compared with the controls. Several glycopeptides were found to vary significantly between the ovarian cancer cases and controls (Fig. 6).

Conclusions and overview

Glycosylation is a post-translational modification of proteins that affects the biochemical activities of the glycoproteins, and plays a substantial role in cancer biology. Glycomics and glycoproteomics methods are less developed than proteomics partly because of the inherent challenges associated with their analysis. However, as the awareness of the roles of glycosylation and other PTMs in cancer biology continues to increase, multidisciplinary efforts are likely to expand the field. Advanced analytical instruments will enable scientists to develop new and robust methods that will uncover potential cancer biomarkers. The complex and laborious glycosylation data analysis process will be improved by applying new or improved bioinformatics methods. These methods are likely to unravel hidden information that would otherwise be lost and also enable fast data analysis to increase the MS obtainable glycan and glycopeptide structural details. Global glycan analysis and protein-specific analysis from human fluids would ultimately add more specific potential cancer markers. A relatively more challenging but useful site-specific glycosylation analysis is likely to provide crucial and detailed information about protein glycosylation patterns. Targeted methods like MRM are gaining interest in the glycoproteomics field. These methods are expected to provide precise quantitative information about the glycosylation site heterogeneity in cancerous cells, tissues, or bio-fluids compared with non-cancerous ones.

Despite the advances in glycan analysis, there remain challenges in developing glycan markers. Further progress in analysis will be made to yield exact structures with extensive site-specific heterogeneity in a quantitative manner. The major limitation is that these methods require considerable expertise as well as expensive instrumentation. Their implementation may be limited to selective groups for some time. To obtain truly relevant clinical markers, extensive multidisciplinary and collaborative efforts between those with analytical tools

and clinicians are necessary today. Furthermore, considerable efforts should also be placed in making these methods more widely available and easier to use.

Compliance with ethical standards

Conflict of interest The authors have declared no conflict of interest.

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin*. 2011;61(2):69–90.
2. Ferlay J, Autier P, Boniol M, Heanue M, Colombet M, Boyle P. Estimates of the cancer incidence and mortality in Europe in 2006. *Ann Oncol*. 2007;18(3):581–92.
3. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010;127(12):2893–917.
4. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, et al. Cancer statistics, 2008. *CA Cancer J Clin*. 2008;58(2):71–96.
5. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin*. 2016;66(1):7–30.
6. Sawyers CL. The cancer biomarker problem. *Nature*. 2008;452(7187):548–52.
7. Kohn EC, Azad N, Annunziata C, Dhamoon AS, Whiteley G. Proteomics as a tool for biomarker discovery. *Dis Markers*. 2007;23(5/6):411–7.
8. Diamandis EP, van der Merwe D-E. Plasma protein profiling by mass spectrometry for cancer diagnosis: opportunities and limitations. *Clin Cancer Res*. 2005;11(3):963–5.
9. Kulasingam V, Diamandis EP. Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. *Nat Clin Pract Oncol*. 2008;5(10):588–99.
10. van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AAM, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. 2002;415(6871):530–6.
11. Espina V, Woodhouse EC, Wulfkühle J, Asmussen HD, Petricoin Iii EF, Liotta LA. Protein microarray detection strategies: focus on direct detection technologies. *J Immunol Methods*. 2004;290(1/2):121–33.
12. Wulfkühle JD, Liotta LA, Petricoin EF. Proteomic applications for the early detection of cancer. *Nat Rev Cancer*. 2003;3(4):267–75.
13. Nagler RM. Saliva as a tool for oral cancer diagnosis and prognosis. *Oral Oncol*. 2009;45(12):1006–10.
14. Laxman B, Morris DS, Yu J, Siddiqui J, Cao J, Mehra R, et al. A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer. *Cancer Res*. 2008;68(3):645–9.
15. Alexander H, Stegner AL, Wagner-Mann C, Du Bois GC, Alexander S, Sauter ER. Proteomic analysis to identify breast cancer biomarkers in nipple aspirate fluid. *Clin Cancer Res*. 2004;10(22):7500–10.
16. Teng P-N, Bateman NW, Hood BL, Conrads TP. Advances in proximal fluid proteomics for disease biomarker discovery. *J Proteome Res*. 2010;9(12):6091–100.
17. Hanash SM, Pitteri SJ, Faca VM. Mining the plasma proteome for cancer biomarkers. *Nature*. 2008;452(7187):571–9.
18. Hassanein M, Callison JC, Callaway-Lane C, Aldrich MC, Grogan EL, Massion PP. The state of molecular biomarkers for the early detection of lung cancer. *Cancer Prev Res*. 2012;5(8):992–1006.

19. Polanski M, Anderson NL. A list of candidate cancer biomarkers for targeted proteomics. *Biomarker Insights*. 2006;1:1.
20. Chen X, Shan Q, Jiang L, Zhu B, Xi X. Quantitative proteomic analysis by iTRAQ for identification of candidate biomarkers in plasma from acute respiratory distress syndrome patients. *Biochem Biophys Res Commun*. 2013;441(1):1–6.
21. Lam YW, Mobley JA, Evans JE, Carmody JF, Ho SM. Mass profiling-directed isolation and identification of a stage-specific serologic protein biomarker of advanced prostate cancer. *Proteomics*. 2005;5(11):2927–38.
22. Palmblad M, Tiss A, Cramer R. Mass spectrometry in clinical proteomics—from the present to the future. *Proteom Clin Appl*. 2009;3(1):6–17.
23. Miller RA, Spellman DS. Mass spectrometry-based biomarkers in drug development. In: Woods AG, Darie CC (Ed.) *Advancements of mass spectrometry in biomedical research*, vol 806, pp 341–359. *Advances in Experimental Medicine and Biology*. Switzerland: Springer International Publishing; 2014.
24. Pusch W, Flocco MT, Leung SM, Thiele H, Kostrzewa M. Mass spectrometry-based clinical proteomics. *Pharmacogenomics*. 2003;4(4):463–76.
25. Zhang Y, Jiao J, Yang P, Lu H. Mass spectrometry-based N-glycoproteomics for cancer biomarker discovery. *Clin Proteom*. 2014;11(1):18.
26. Jimenez CR, Verheul HMW. Mass spectrometry-based proteomics: from cancer biology to protein biomarkers, drug targets, and clinical applications. *Am Soc Clin Oncol educational book/ASCO Am Soc Clin Oncol Meeting*. e504–510. 2014.
27. Yin H, Lin Z, Nie S, Wu J, Tan Z, Zhu J, et al. Mass-selected site-specific core-fucosylation of ceruloplasmin in alcohol-related hepatocellular carcinoma. *J Proteome Res*. 2014;13(6):2887–96.
28. Griffiths WJ, Wang Y. Mass spectrometry: from proteomics to metabolomics and lipidomics. *Chem Soc Rev*. 2009;38(7):1882–96.
29. Ly M, Laremore TN, Linhardt RJ. Proteoglycomics: recent progress and future challenges. *OMICS J Integrative Biol*. 2010;14(4):389–99.
30. Wührer M, Catalina MI, Deelder AM, Hokke CH. Glycoproteomics based on tandem mass spectrometry of glycopeptides. *J Chromatogr B*. 2007;849(1):115–28.
31. Zaia J. Mass spectrometry and the emerging field of glycomics. *Chem Biol*. 2008;15(9):881–92.
32. Hua S, Lebrilla C, An HJ. Application of nano-LC-based glycomics towards biomarker discovery. *Bioanalysis*. 2011;3(22):2573–85.
33. Alley WR, Madera M, Mechref Y, Novotny MV. Chip-based reversed-phase liquid chromatography-mass spectrometry of permethylated N-linked glycans: a potential methodology for cancer-biomarker discovery. *Anal Chem*. 2010;82(12):5095–106.
34. Eng JK, McCormack AL, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom*. 1994;5(11):976–89.
35. Kronewitter SR, De Leoz MLA, Strum JS, An HJ, Dimapasoc LM, Guerrero A, et al. The glycozyler: automated glycan annotation software for high performance mass spectrometry and its application to ovarian cancer glycan biomarker discovery. *Proteomics*. 2012;12(15/16):2523–38.
36. Ludwig JA, Weinstein JN. Biomarkers in cancer staging, prognosis, and treatment selection. *Nat Rev Cancer*. 2005;5(11):845–56.
37. Yin BWT, Lloyd KO. Molecular Cloning of the CA125 Ovarian Cancer Antigen: identification as a new mucin, MUC16. *J Biol Chem*. 2001;276(29):27371–5.
38. Jacobs I, Bast RC. The CA 125 tumor-associated antigen: a review of the literature. *Hum Reprod*. 1989;4(1):1–12.
39. Miralles C, Orea M, España P, Provencio M, Sánchez A, Cantos B, et al. Cancer antigen 125 associated with multiple benign and malignant pathologies. *Ann Surg Oncol*. 2003;10(2):150–4.
40. Adamczyk B, Tharmalingam T, Rudd PM. Glycans as cancer biomarkers. *Biochim Biophys Acta - General Subjects*. 2012;1820(9):1347–53.
41. Duffy M, Shering S, Sherry F, McDermott E, O'higgins N. CA 15-3: a prognostic marker in breast cancer. *Int J Biol Markers*. 1999;15(4):330–3.
42. Drake RR, Jones EE, Powers TW, Nyalwidhe JO. Chapter 10—altered glycosylation in prostate cancer. *Adv Cancer Res*. 2015;126:345–82.
43. Catalona W, Richie J, Ahmann F, Hudson M, Scardino P, Flanigan R, et al. Comparison of digital rectal examination and serum prostate specific antigen in the early detection of prostate cancer: results of a multicenter clinical trial of 6,630 men. *J Urol*. 1994;151(5):1283–90.
44. Hammarström S. The carcinoembryonic antigen (CEA) family: structures, suggested functions, and expression in normal and malignant tissues. *Semin Cancer Biol*. 1999;9(2):67–81.
45. Benchimol S, Fuks A, Jothy S, Beauchemin N, Shiota K, Stanners CP. Carcinoembryonic antigen, a human tumor marker, functions as an intercellular adhesion molecule. *Cell*. 1989;57(2):327–34.
46. Moertel CG, Fleming TR, Macdonald JS, Haller DG, Laurie JA, Tangen C. An evaluation of the carcinoembryonic antigen (CEA) test for monitoring patients with resected colon cancer. *JAMA*. 1993;270(8):943–7.
47. Thompson S, Turner G. Elevated levels of abnormally-fucosylated haptoglobins in cancer sera. *Br J Cancer*. 1987;56(5):605.
48. Okuyama N, Ide Y, Nakano M, Nakagawa T, Yamanaka K, Moriwaki K, et al. Fucosylated haptoglobin is a novel marker for pancreatic cancer: a detailed analysis of the oligosaccharide structure and a possible mechanism for fucosylation. *Int J Cancer*. 2006;118(11):2803–8.
49. Miyoshi E, Nakano M. Fucosylated haptoglobin is a novel marker for pancreatic cancer: detailed analyses of oligosaccharide structures. *Proteomics*. 2008;8(16):3257–62.
50. Zhao C, Annamalai L, Guo C, Kothandaraman N, Koh SCL, Zhang H, et al. Circulating haptoglobin is an independent prognostic factor in the sera of patients with epithelial ovarian cancer. *Neoplasia (New York, NY)*. 2007;9(1):1–7.
51. Ahmed N, Barker G, Oliva KT, Hoffmann P, Riley C, Reeve S, et al. Proteomic-based identification of haptoglobin-1 precursor as a novel circulating biomarker of ovarian cancer. *Br J Cancer*. 2004;91(1):129–40.
52. Huang HL, Stasyk T, Morandell S, Dieplinger H, Falkensammer G, Griesmacher A, et al. Biomarker discovery in breast cancer serum using 2-D differential gel electrophoresis/MALDI-TOF/TOF and data validation by routine clinical assays. *Electrophoresis*. 2006;27(8):1641–50.
53. Ralhan R, DeSouza LV, Matta A, Tripathi SC, Ghanny S, Gupta SD, et al. Discovery and verification of head-and-neck cancer biomarkers by differential protein expression analysis using iTRAQ labeling, multidimensional liquid chromatography, and tandem mass spectrometry. *Mol Cell Proteom*. 2008;7(6):1162–73.
54. Patz EF, Campa MJ, Gottlin EB, Kusmartseva I, Guan XR, Herndon JE. Panel of serum biomarkers for the diagnosis of lung cancer. *J Clin Oncol*. 2007;25(35):5578–83.
55. Drake RR, Schwegler EE, Malik G, Diaz J, Block T, Mehta A, et al. Lectin capture strategies combined with mass spectrometry for the discovery of serum glycoprotein biomarkers. *Mol Cell Proteom*. 2006;5(10):1957–67.
56. Zhao J, Simeone DM, Heidt D, Anderson MA, Lubman DM. Comparative serum glycoproteomics using lectin selected sialic

- acid glycoproteins with mass spectrometric analysis: application to pancreatic cancer serum. *J Proteome Res.* 2006;5(7):1792–802.
57. Kaji H, Saito H, Yamauchi Y, Shinkawa T, Taoka M, Hirabayashi J, et al. Lectin affinity capture, isotope-coded tagging, and mass spectrometry to identify N-linked glycoproteins. *Nat Biotech.* 2003;21(6):667–72.
 58. Wu J, Xie X, Liu Y, He J, Benitez R, Buckanovich RJ, et al. Identification and confirmation of differentially expressed fucosylated glycoproteins in the serum of ovarian cancer patients using a lectin array and LC-MS/MS. *J Proteome Res.* 2012;11(9):4541–52.
 59. Alvarez-Manilla G, Warren NL, Atwood III J, Orlando R, Dalton S, Pierce M. Glycoproteomic analysis of embryonic stem cells: identification of potential glycobiomarkers using lectin affinity chromatography of glycopeptides. *J Proteome Res.* 2010;9(5):2062–75.
 60. Kuno A, Kato Y, Matsuda A, Kaneko MK, Ito H, Amano K, et al. Focused differential glycan analysis with the platform antibody-assisted lectin profiling for glycan-related biomarker verification. *Mol Cell Proteom.* 2009;8(1):99–108.
 61. Li Y, Wen T, Zhu M, Li L, Wei J, Wu X, et al. Glycoproteomic analysis of tissues from patients with colon cancer using lectin microarrays and nanoLC-MS/MS. *Mol Biosyst.* 2013;9(7):1877–87.
 62. Dennis JW, Laferte S, Waghome C, Breitman ML, Kerbel RS. Beta 1-6 branching of Asn-linked oligosaccharides is directly associated with metastasis. *Science.* 1987;236(4801):582–5.
 63. Ihara S, Miyoshi E, Ko JH, Murata K, Nakahara S, Honke K, et al. Prometastatic effect of N-acetylglucosaminyltransferase V is due to modification and stabilization of active matrilysin by adding β 1–6 GlcNAc branching. *J Biol Chem.* 2002;277(19):16960–7.
 64. Buckhaults P, Chen L, Fregien N, Pierce M. Transcriptional regulation of N-acetylglucosaminyltransferase V by the src oncogene. *J Biol Chem.* 1997;272(31):19575–81.
 65. Abbott KL, Aoki K, Lim J-M, Porterfield M, Johnson R, O'Regan RM, et al. Targeted glycoproteomic identification of biomarkers for human breast carcinoma. *J Proteome Res.* 2008;7(4):1470–80.
 66. Thompson A, Schäfer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem.* 2003;75(8):1895–904.
 67. Nie S, Lo A, Wu J, Zhu J, Tan Z, Simeone DM, et al. Glycoprotein biomarker panel for pancreatic cancer discovered by quantitative proteomics analysis. *J Proteome Res.* 2014;13(4):1873–84.
 68. Mann M, Jensen ON. Proteomic analysis of post-translational modifications. *Nat Biotech.* 2003;21(3):255–61.
 69. Pagel O, Loroch S, Sickmann A, Zahedi RP. Current strategies and findings in clinically relevant post-translational modification-specific proteomics. *Expert Rev Proteom.* 2015;12(3):235–53.
 70. Christiansen MN, Chik J, Lee L, Anugraham M, Abrahams JL, Packer NH. Cell surface protein glycosylation in cancer. *Proteomics.* 2014;14(4/5):525–46.
 71. Holst S, Wuhler M, Rombouts Y (2015) Chapter 6 – Glycosylation characteristics of colorectal cancer. In: Richard RD, Lauren EB (Eds.) *Advances in Cancer Research*, vol 126, pp 203–256. Academic Press.
 72. Walsh G, Jefferis R. Post-translational modifications in the context of therapeutic proteins. *Nat Biotech.* 2006;24(10):1241–52.
 73. An HJ, Froehlich JW, Lebrilla CB. Determination of glycosylation sites and site-specific heterogeneity in glycoproteins. *Curr Opin Chem Biol.* 2009;13(4):421–6.
 74. Shriver Z, Raguram S, Sasisekharan R. Glycomics: a pathway to a class of new and improved therapeutics. *Nat Rev Drug Discov.* 2004;3(10):863–73.
 75. Dube DH, Bertozzi CR. Glycans in cancer and inflammation—potential for therapeutics and diagnostics. *Nat Rev Drug Discov.* 2005;4(6):477–88.
 76. Reis CA, Osorio H, Silva L, Gomes C, David L. Alterations in glycosylation as biomarkers for cancer detection. *J Clin Pathol.* 2010;63(4):322–9.
 77. Varki A, Kannagi R, Toole BP. Glycosylation changes in cancer. 2009.
 78. An HJ, Kronewitter SR, de Leoz MLA, Lebrilla CB. Glycomics and disease markers. *Curr Opin Chem Biol.* 2009;13(5/6):601–7.
 79. Guo H, Abbott KL (2015) Chapter 8 – functional impact of tumor-specific N-linked glycan changes in breast and ovarian cancers. In: Richard RD, Lauren EB (Eds.) *Advances in Cancer Research*, vol 126, pp 281–303. Academic Press.
 80. Drake PM, Cho W, Li B, Prakobphol A, Johansen E, Anderson NL, et al. Sweetening the pot: adding glycosylation to the biomarker discovery equation. *Clin Chem.* 2010;56(2):223–36.
 81. Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. *Nat Rev Cancer.* 2015;15:540–55.
 82. de Leoz MLA, Young LJT, An HJ, Kronewitter SR, Kim J, Miyamoto S, et al. High-mannose glycans are elevated during breast cancer progression. *Mol Cell Proteom.* 2011;10(1).
 83. Kailemia MJ, Ruhaak LR, Lebrilla CB, Amster IJ. Oligosaccharide analysis by mass spectrometry: a review of recent developments. *Anal Chem.* 2013;86(1):196–212.
 84. Ruhaak LR, Miyamoto S, Lebrilla CB. Developments in the identification of glycan biomarkers for the detection of cancer. *Mol Cell Proteom.* 2013;12(4):846–55.
 85. Budnik BA, Lee RS, Steen JAJ. Global methods for protein glycosylation analysis by mass spectrometry. *Biochim Biophys Acta.* 2006;1764(12):1870–80.
 86. An HJ, Lebrilla CB (2010) A glycomics approach to the discovery of potential cancer biomarkers. In: *Functional glycomics*, 199–213. Springer.
 87. JoóáAn H. The prospects of glycan biomarkers for the diagnosis of diseases. *Mol Biosyst.* 2009;5(1):17–20.
 88. An HJ, Miyamoto S, Lancaster KS, Kirmiz C, Li B, Lam KS, et al. Profiling of glycans in serum for the discovery of potential biomarkers for ovarian cancer. *J Proteome Res.* 2006;5(7):1626–35.
 89. Bharti A, Ma PC, Maulik G, Singh R, Khan E, Skarin AT, et al. Haptoglobin α -subunit and hepatocyte growth factor can potentially serve as serum tumor biomarkers in small-cell lung cancer. *Anticancer Res.* 2004;24(2C):1031–8.
 90. Bones J, Mittermayr S, O'Donoghue N, Guttman A, Rudd PM. Ultra performance liquid chromatographic profiling of serum N-glycans for fast and efficient identification of cancer associated alterations in glycosylation. *Anal Chem.* 2010;82(24):10208–15.
 91. Wada Y, Azadi P, Costello CE, Dell A, Dwek RA, Geyer H, et al. Comparison of the methods for profiling glycoprotein glycans—HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study. *Glycobiology.* 2007;17(4):411–22.
 92. Ransom HW, Varghese RS, Goldman L, An Y, Loffredo CA, Abdel-Hamid M, et al. Analysis of MALDI-TOF mass spectrometry data for discovery of peptide and glycan biomarkers of hepatocellular carcinoma. *J Proteome Res.* 2008;7(2):603–10.
 93. Rodrigo MAM, Zitka O, Krizkova S, Moullick A, Adam V, Kizek R. MALDI-TOF MS as evolving cancer diagnostic tool: a review. *J Pharm Biomed Anal.* 2014;95:245–55.
 94. Powers TW, Neely BA, Shao Y, Tang H, Troyer DA, Mehta AS, et al. MALDI imaging mass spectrometry profiling of N-glycans in formalin-fixed paraffin embedded clinical tissue blocks and tissue microarrays. *PLoS One.* 2014;9(9).
 95. Kyselova Z, Mechref Y, Al Bataineh MM, Dobrolecki LE, Hickey RJ, Vinson J, et al. Alterations in the serum glycome due to metastatic prostate cancer. *J Proteome Res.* 2007;6(5):1822–32.
 96. Kyselova Z, Mechref Y, Kang P, Goetz JA, Dobrolecki LE, Sledge GW, et al. Breast cancer diagnosis and prognosis through quantitative measurements of serum glycan profiles. *Clin Chem.* 2008;54(7):1166–75.

97. Balog CIA, Stavenhagen K, Fung WLJ, Koeleman CA, McDonnell LA, Verhoeven A, et al. N-glycosylation of colorectal cancer tissues. *Mol Cell Proteom*. 2012;11(9):571–85.
98. de Leoz MLA, An HJ, Kronewitter S, Kim J, Beecroft S, Vinnall R, et al. Glycomic approach for potential biomarkers on prostate cancer: Profiling of N-linked glycans in human sera and pRNS cell lines. *Dis Markers*. 2008;25(4/5):243–58.
99. Hecht ES, Scholl EH, Walker SH, Taylor AD, Cliby WA, Motsinger-Reif AA, et al. Relative quantification and higher-order modeling of the plasma glycan cancer burden ratio in ovarian cancer case-control samples. *J Proteome Res*. 2015;14(10):4394–401.
100. Kirmiz C, Li B, An HJ, Clowers BH, Chew HK, Lam KS, et al. A serum glycomics approach to breast cancer biomarkers. *Mol Cell Proteom*. 2007;6(1):43–55.
101. Alley WR, Vasseur JA, Goetz JA, Svoboda M, Mann BF, Matei DE, et al. N-linked glycan structures and their expressions change in the blood sera of ovarian cancer patients. *J Proteome Res*. 2012;11(4):2282–300.
102. Saldova R, Royle L, Radcliffe CM, Abd Hamid UM, Evans R, Arnold JN, et al. Ovarian cancer is associated with changes in glycosylation in both acute-phase proteins and IgG. *Glycobiology*. 2007;17(12):1344–56.
103. Kanoh Y, Mashiko T, Danbara M, Takayama Y, Ohtani S, Imasaki T, et al. Analysis of the Oligosaccharide chain of human serum immunoglobulin G in patients with localized or metastatic cancer. *Oncology*. 2004;66(5):365–70.
104. Ruhaak LR, Stroble C, Dai J, Barnett MJ, Taguchi A, Goodman GE, et al. Serum glycans as risk markers for non-small cell lung cancer. *Cancer Prev Res*. 2016. doi:10.1158/1940-6207.capr-15-0033.
105. Hua S, An HJ, Ozcan S, Ro GS, Soares S, DeVere-White R, et al. Comprehensive native glycan profiling with isomer separation and quantitation for the discovery of cancer biomarkers. *Analyst*. 2011;136(18):3663–71.
106. Ruhaak LR, Deelder A, Wührer M. Oligosaccharide analysis by graphitized carbon liquid chromatography-mass spectrometry. *Anal Bioanal Chem*. 2009;394(1):163–74.
107. Stavenhagen K, Kolarich D, Wührer M. Clinical glycomics employing graphitized carbon liquid chromatography-mass spectrometry. *Chromatographia*. 2014;78(5):307–20.
108. Chu CS, Niñonuevo MR, Clowers BH, Perkins PD, An HJ, Yin H, et al. Profile of native N-linked glycan structures from human serum using high performance liquid chromatography on a microfluidic chip and time-of-flight mass spectrometry. *Proteomics*. 2009;9(7):1939–51.
109. Song T, Aldredge D, Lebrilla CB. A method for in-depth structural annotation of human serum glycans that yields biological variations. *Anal Chem*. 2015;87(15):7754–62.
110. Ruhaak LR, Barkauskas DA, Torres J, Cooke CL, Wu LD, Stroble C, et al. The serum immunoglobulin G glycosylation signature of gastric cancer. *EuPA Open Proteom*. 2015;6:1–9.
111. Strum JS, Nwosu CC, Hua S, Kronewitter SR, Seipert RR, Bachelor RJ, et al. Automated Assignments of N- and O-site specific glycosylation with extensive glycan heterogeneity of glycoprotein mixtures. *Anal Chem*. 2013;85(12):5666–75.
112. Nwosu CC, Seipert RR, Strum JS, Hua SS, An HJ, Zivkovic AM, et al. Simultaneous and extensive site-specific N- and O-glycosylation analysis in protein mixtures. *J Proteome Res*. 2011;10(5):2612–24.
113. Ueda K, Takami S, Saichi N, Daigo Y, Ishikawa N, Kohno N, et al. Development of serum glycoproteomic profiling technique; simultaneous identification of glycosylation sites and site-specific quantification of glycan structure changes. *Mol Cell Proteom*. 2010;9(9):1819–28.
114. Ueda K. Glycoproteomic strategies: From discovery to clinical application of cancer carbohydrate biomarkers. *Proteom Clin App*. 2013;7(9/10):607–17.
115. Bem M, Kil YJ, Becker C. Byonic: advanced peptide and protein identification software. *Curr Protoc Bioinform*. 2012;40:13.20.11–13.20.14.
116. Ozohanics O, Krenyacz J, Ludányi K, Pollreis F, Vékey K, Drahoš L. GlycoMiner: a new software tool to elucidate glycopeptide composition. *Rapid Commun Mass Spectrom*. 2008;22(20):3245–54.
117. He L, Xin L, Shan B, Lajoie GA, Ma B. GlycoMaster DB: Software to assist the automated identification of N-linked glycopeptides by tandem mass spectrometry. *J Proteome Res*. 2014;13(9):3881–95.
118. Mayampurath A, Yu C-Y, Song E, Balan J, Mechref Y, Tang H. Computational framework for identification of intact glycopeptides in complex samples. *Anal Chem*. 2014;86(1):453–63.
119. Mayampurath AM, Wu Y, Segu ZM, Mechref Y, Tang H. Improving confidence in detection and characterization of protein N-glycosylation sites and microheterogeneity. *Rapid Commun Mass Spectrom*. 2011;25(14):2007–19.
120. Khatri K, Staples GO, Leymarie N, Leon DR, Turiák L, Huang Y, et al. Confident assignment of site-specific glycosylation in complex glycoproteins in a single step. *J Proteome Res*. 2014;13(10):4347–55.
121. Go EP, Rebecchi KR, Dalpathado DS, Bandu ML, Zhang Y, Desaire H. GlycoPep DB: a tool for glycopeptide analysis using a “Smart Search”. *Anal Chem*. 2007;79(4):1708–13.
122. Woodin CL, Maxon M, Desaire H. Software for automated interpretation of mass spectrometry data from glycans and glycopeptides. *Analyst*. 2013;138(10):2793–803.
123. Reid GE, Stephenson JL, McLuckey SA. Tandem mass spectrometry of ribonuclease A and B: N-linked glycosylation site analysis of whole protein ions. *Anal Chem*. 2002;74(3):577–83.
124. Siuti N, Kelleher NL. Decoding protein modifications using top-down mass spectrometry. *Nat Methods*. 2007;4(10):817–21.
125. Ito S, Hayama K, Hirabayashi J (2009) Enrichment strategies for glycopeptides. In: Packer N, Karlsson N (Eds.) *Glycomics*, vol 534, pp 194–203. *Methods in Molecular Biology*. Humana Press.
126. Wührer M, Deelder AM, Hokke CH. Protein glycosylation analysis by liquid chromatography-mass spectrometry. *J Chromatogr B*. 2005;825(2):124–33.
127. Zhu Z, Go EP, Desaire H. Absolute quantitation of glycosylation site occupancy using isotopically labeled standards and LC-MS. *J Am Soc Mass Spectrom*. 2014;25(6):1012–7.
128. Dalpathado DS, Desaire H. Glycopeptide analysis by mass spectrometry. *Analyst*. 2008;133(6):731–8.
129. Gbormittah FO, Bones J, Hincapie M, Tousi F, Hancock WS, Iliopoulos O. Clusterin glycopeptide variant characterization reveals significant site-specific glycan changes in the plasma of clear cell renal cell carcinoma. *J Proteome Res*. 2015;14(6):2425–36.
130. Tan Z, Yin H, Nie S, Lin Z, Zhu J, Ruffin MT, et al. Large-scale identification of core-fucosylated glycopeptide sites in pancreatic cancer serum using mass spectrometry. *J Proteome Res*. 2015;14(4):1968–78.
131. Lin Z, Yin H, Lo A, Ruffin MT, Anderson MA, Simeone DM, et al. Label-free relative quantification of α -2-macroglobulin site-specific core-fucosylation in pancreatic cancer by LC-MS/MS. *Electrophoresis*. 2014;35(15):2108–15.
132. Saldova R, Fan Y, Fitzpatrick JM, Watson RWG, Rudd PM. Core fucosylation and α 2-3 sialylation in serum N-glycome is significantly increased in prostate cancer comparing to benign prostate hyperplasia. *Glycobiology*. 2011;21(2):195–205.
133. Miyoshi E, Moriwaki K, Nakagawa T. Biological function of fucosylation in cancer biology. *J Biochem*. 2008;143(6):725–9.
134. Ruhaak LR, Lebrilla C. Applications of multiple reaction monitoring to clinical glycomics. *Chromatographia*. 2015;78(5/6):335–42.

135. Picotti P, Aebersold R. Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nat Meth.* 2012;9(6):555–66.
136. Song E, Pyreddy S, Mechref Y. Quantification of glycopeptides by multiple reaction monitoring liquid chromatography/tandem mass spectrometry. *Rapid Commun Mass Spectrom.* 2012;26(17):1941–54.
137. Hong Q, Lebrilla CB, Miyamoto S, Ruhaak LR. Absolute quantitation of immunoglobulin G and its glycoforms using multiple reaction monitoring. *Anal Chem.* 2013;85(18):8585–93.
138. Sanda M, Pompach P, Brnakova Z, Wu J, Makambi K, Goldman R. Quantitative liquid chromatography-mass spectrometry-multiple reaction monitoring (LC-MS-MRM) analysis of site-specific glycoforms of haptoglobin in liver disease. *Mol Cell Proteom.* 2013;12(5):1294–305.
139. Yuan W, Sanda M, Wu J, Koomen J, Goldman R. Quantitative analysis of immunoglobulin subclasses and subclass specific glycosylation by LC-MS-MRM in liver disease. *J Proteom.* 2015;116:24–33.
140. Anderson NL, Anderson NG, Haines LR, Hardie DB, Olafson RW, Pearson TW. Mass spectrometric quantitation of peptides and proteins using stable isotope standards and capture by anti-peptide antibodies (SISCAPA). *J Proteome Res.* 2004;3(2):235–44.
141. Ahn YH, Lee JY, Lee JY, Kim Y-S, Ko JH, Yoo JS. Quantitative analysis of an aberrant glycoform of TIMP1 from colon cancer serum by L-PHA-enrichment and SISCAPA with MRM mass spectrometry. *J Proteome Res.* 2009;8(9):4216–24.
142. Hong Q, Ruhaak LR, Stroble C, Parker E, Huang J, Mavarakis E, et al. A method for comprehensive glycosite-mapping and direct quantitation of serum glycoproteins. *J Proteome Res.* 2015;14(12):5179–92.
143. Ruhaak LR, Kim K, Stroble C, Taylor SL, Hong Q, Miyamoto S, et al. Protein-specific differential glycosylation of immunoglobulins in serum of ovarian cancer patients. *J Proteome Res.* 2016;15(3):1002–10.