Developments in the Identification of Glycan Biomarkers for the Detection of Cancer

L. Renee Ruhaak*, Suzanne Miyamoto‡, and Carlito B. Lebrilla*§

Changes in glycosylation readily occur in cancer and other disease states. Thanks to recent advances in the development of analytical techniques and instrumentation, especially in mass spectrometry, it is now possible to identify blood-derived glycan-based biomarkers using glycomics strategies. This review is an overview of the developments made in the search for glycan-based cancer biomarkers and the technologies currently in use. It is anticipated that the progressing instrumental and bioinformatics developments will allow the identification of relevant glycan biomarkers for the diagnosis, early detection, and monitoring of cancer treatment with sufficient sensitivity and specificity for clinical use. Molecular & Cellular Proteomics 12: 10.1074/mcp.R112.026799, 846–855, 2013.

The ubiquity of glycosylation and the fundamental importance of N-glycans in nearly every biological process give N-glycans high potential as biomarkers. Nearly 70% of the plasma proteome is glycosylated, and most cell membranes and secreted proteins are also highly glycosylated. The N-glycans play important roles in major biological processes that include cell–cell and cell–matrix interactions (1), protein folding (2), receptor binding, and protein clearance (3). There is ample evidence that altered N-glycosylation patterns are present on tumor cells (e.g. Refs. 4–7), and these findings have sparked the search for glycan-based biomarkers for the detection of different types of cancer in biofluids such as serum and plasma (8).

Glycans are highly branched structures built up of monosaccharides that are linked by glycosidic bonds in multiple different ways in a non-template-driven manner. This is in contrast to proteins, which consist of amino acids linked by well-defined peptide bonds in a linear sequence-dependent template-driven fashion. It has been speculated that over 3000 N-glycan structures might exist based on the number of possible monosaccharides and the different linkages that might occur among them (9). The number corresponds well to recent estimates based on a theoretical library of serum glycans in which the actual size of the human serum N-glycome is proposed to be restricted to around 300 to 500 compositions corresponding to fewer than 5000 structures (10).

The pathways involved in N-glycan biosynthesis are generally well understood and have been reviewed extensively (11). Briefly, a dolichol precursor is enzymatically decorated with a glycan precursor (GlcNAc$_2$Man$_9$Glc$_3$), and the dolichol-glycan precursor in the endoplasmic reticulum is then transferred to the target protein. When the protein is folded correctly and transferred to the Golgi, the glycan precursor is modified from a high-mannose structure to the mature N-glycan using several glycosyltransferases and glycosidases. Although this process is well regulated, individual proteins might contain more than one glycosylation site and are often decorated with different glycans on the different sites. Moreover, multiple copies of the same protein are likely to have different glycans attached, leading to large glycan microheterogeneities, further complicating glycan analysis.

Over the years, several approaches have been used for the analysis of glycoproteins and N-glycans. Early studies mostly evaluated lectin binding, showing the presence of certain binding epitopes on specific proteins, cells, or tissues. More structurally extensive methods, including HPLC and, especially, mass spectrometry, have expanded the field of glycomics (12). Several analytical approaches have been used for the profiling of N-glycans, including stand-alone mass spectrometry and separation strategies such as reverse phase chromatography, hydrophilic interaction chromatography (HLIC) (13, 14), porous graphitic carbon (PGC) (15) chromatography, and capillary electrophoresis (16), which have all been reviewed extensively.

Currently, a comprehensive glycomics approach that yields structures as well as quantitative information involves the profiling of serum or plasma glycans by means of mass spectrometry. Because diseases such as cancer alter glycosylation, a systems-wide analysis of the glycome could yield direct changes in health conditions even without consideration of protein identification or abundances. Indeed, a glycomics analysis of serum proteins indicated that many proteins change their glycosylation simultaneously in disease states relative to the control (17, 18). The first global glycomics
approach investigating serum was performed with serum from ovarian cancer patients (19).

The intrinsic complexity of glycosylation allows the discovery of biomarkers at various levels of structural analysis. The first types of biomarkers were based solely on compositions (the monosaccharides present in a specific glycan), which can be accurately monitored via stand-alone mass spectrometry (7, 20). Compositional analysis can yield the numbers of hexoses, N-acetylated hexosamines, fucoses, and sialic acids. Variations in the number of antennae, the number of terminal galactoses, and the degree of sialylation and fucosylation can all be inferred solely from mass spectral data (e.g. Refs. 21–24). A combination of techniques allowing isomer separation with mass spectrometry improves the analysis. Each N-glycan composition can yield a number of isomers that have the same composition but different arrangements of monosaccharide linkages. Nanoflow liquid chromatography (nLC) employing PGC stationary phase coupled with mass spectrometry allows the effective separation of native glycan structures with relative and potentially absolute quantitation (25–28). Using this method, biomarkers can now be discovered on a compound specific level. The potentially most specific, but also technically most difficult, are protein-specific and site-specific glycosylation. The analytical approach employs tools for protein-specific analysis by isolating proteins for glycan release from 16 h or overnight to 10 min. An alternative method that could be used to shorten the N-glycan release time might be pressure-cycling technology (35).

Upon release, the N-glycans are purified using a cartridge containing PGC in which the glycans are eluted in three fractions (32) of varying proportions of acetonitrile in water to yield (i) a fraction containing neutral glycans, with enrichment of high-mannose-type glycans, (ii) mostly neutral hybrid and complex type glycans, and (iii) mostly hybrid and complex type glycans containing N-acetyl neuraminic acid residues (Fig. 1). The high accuracy of the FTICR allows the unambiguous identification of glycan compositions based on the accurate mass in combination with a retrosynthetic glycan composition library (10). Using the strategy described, 64 glycan compositions can be detected consistently from serum with an average coefficient of variation (cv) of less than 10% (33).

The application of this method to a smaller set of sera from mice transplanted with human tumor tissue and sera from breast cancer patients compared with that of healthy controls showed increased levels of high-mannose-type structures with cancer in both mouse and human sera (7). A second sample set containing serum from ovarian cancer patients compared with serum from matched controls recently revealed increased levels of sialylated glycans, whereas several neutral glycans, including high-mannose-type glycans, were decreased. Importantly, levels of a small group of truncated glycans (Hex3HexNAc4, Hex3HexNAc4Fuc1, Hex3HexNAc5, and Hex3HexNAc5Fuc1) were increased (20). A similar approach has been used by Kim et al. (36), who applied MALDI-TOF-MS for the detection of biomarkers for ovarian cancer in desialylated serum samples and found increased levels of the fucosylated bi- and tri-antennary glycans Hex3,HexNAc4,Fuc1, and Hex3,HexNAc5,Fuc1.
The permethylation of glycans before analysis eliminates the need for fractionation as performed in the analysis of native glycans and allows the whole mixture to be examined simultaneously, typically with MALDI-TOF-MS (37). Although permethylation is a relatively harsh chemical reaction, its products, and especially the N-glycans containing sialic acids, are known to be more stable in the MALDI ionization process. Because oligosaccharides containing sialic acid are the most abundant species in serum, these studies tend to focus more on sialylated components. The application of this method to a set of breast cancer samples revealed a significant increase in fucosylated and sialylated glycans (38). Moreover, in a separate study, three glycans were shown to provide good sensitivity and specificity for the separation of serum samples from patients with hepatocellular carcinoma and controls (39). Application of the method to a set of ovarian cancer samples showed that the levels of tri- and tetra-antennary N-glycans were increased, independent of their levels of fucosylation and sialylation (40). Moreover, levels of glycans containing a bisecting GlcNAc were shown to be decreased (40).

Biomarker Discovery Employing Separation Methods—Separation methods including LC or capillary electrophoresis with fluorescence or UV detection are useful once structures have been completely or partially elucidated. The most widely applied technique for separation is HILIC-HPLC of 2-AB-labeled glycans (41, 42), which separates oligosaccharides mostly by their size. Using a glucose index, retention times can be matched and glycan structures may be assigned. The resolution of HILIC-HPLC is relatively poor, but increased resolution may be obtained using ultra performance liquid chromatography (UPLC), with ~17 (HPLC) and 45 (UPLC) respective peaks typically observed in serum samples when using these techniques. Although the use of fluorescence allows accurate and robust quantitation, the incidence of overlapping glycans usually requires the use of exoglycosidases for unambiguous identification of the candidate biomarkers.

HILIC-HPLC with fluorescence detection has been widely applied for the identification of candidate biomarkers for a wide range of cancer types, including gastric, lung, ovarian, and breast cancer and those results have been reviewed recently (43). It was observed that in patients suffering from gastric cancer, levels of fucosylated non- and mono-sialylated glycans were decreased, whereas a peak containing disialylated biantennary glycan (Hex5HexNAc4Sia2) showed increased levels (42). In a second study, two peaks, one containing biantennary asialo monogalactosylated glycans and one containing triantennary glycans carrying 2–3-linked sialic acids, were decreased with gastric cancer, and one peak containing triantennary glycans carrying 2–6-linked sialic acids and trisialylated triantennary glycans carrying sialyl Lewis x showed increased levels (44). In a study of serum samples from lung cancer patients and controls, peaks containing mostly tri- and tetra-antennary, highly sialylated glycans, some with antenna and some with core fucosylation, were increased. A significant decrease was observed for peaks containing mostly biantennary glycans, mostly with core fucose (45). When serum profiles of ovarian cancer patients are compared with controls, the most profound changes include the increased levels of peaks containing core fucosylated agalactosyl biantennary glycans and peaks that consist of sialyl-Lewis-x-containing glycans (46). In a separate study, significantly increased levels of peaks

---

1 The abbreviations used are: 2-AB, 2-aminobenzamide; FTICR, Fourier transform ion cyclotron resonance; HILIC, hydrophilic interaction chromatography; MALDI, matrix-assisted laser desorption ionization; nLC, nanoflow liquid chromatography; PGC, porous graphitic carbon; TOF, time of flight; UPLC, ultra performance liquid chromatography.
containing highly sialylated antenna fucosylated glycans (carrying sialyl Lewis x) were also observed in serum samples from advanced-stage breast cancer patients relative to controls (47).

An alternative method of separation that has been applied in glycan profiling is capillary (gel) electrophoresis, which may be performed on DNA sequencing equipment to allow high-throughput multiplexed analysis (48–50). In order for electrophoretic separation to take place, compounds need to be charged, and therefore N-glycans are often labeled with aminopyrene-trisulfonic acid prior to analysis. This label also facilitates fluorescence detection. Clinical application of this method to desialylated serum has shown that the ratio between a bisected fucosylated glycan (Hex5HexNAc5Fuc1) and a triantennary glycan (Hex6HexNAc5) increases significantly in patients suffering liver cirrhosis (51). More recently, the same technology was used to separate hepatocellular carcinoma cases from liver cirrhosis cases (52); it was observed that a fucosylated triantennary glycan (Hex5HexNAc5Fuc1) is significantly increased with hepatocellular carcinoma, and its levels relative to those of a bisected fucosylated glycan (Hex5HexNAc5Fuc1) were shown to increase consistently in cases of hepatocellular carcinoma, independent of the presence of cirrhosis.

**LC/MS Techniques for Biomarker Discovery**—Whereas the use of separation techniques combined with structural neutral detection such as fluorescence spectrophotometry provides very stable quantitation with cv values of less than 5% (53), chromatographic or electrophoretic separation alone suffers from the lack of complete separation between glycan isomers and homologs, as well as difficulties in obtaining analytically reproducible retention times. Nonetheless, glycan analyses at the level of individual compounds will result in even more specific biomarkers for disease detection. The coupling of LC with MS facilitates the distinction of individual glycan structures in complex mixtures with the addition of a complementary second dimension such as mass-based detection. Moreover, the use of LC/MS results in decreased ion suppression, which is typically observed in mixture analyses because of the lower number of glycans that compete for charge during ionization (54). The use of LC/MS in glycomics has been reviewed recently (e.g. 27). Capillary electrophoresis MS is currently less used for clinical applications, mostly because of the difficulty of coupling the two techniques (16, 55, 56).

In order to allow the combination of good isomer separation with high-sensitivity detection, analytical strategies employing nanoflow technology on a microchip coupled with high-mass-accuracy time-of-flight mass spectrometry have been developed (25). PGC allows for the separation of native, underivatized glycan isomers, thus allowing simpler sample preparation strategies. However, when analyzing native glycans, separation of the anomers at the reducing end occurs, resulting in two signals originating from the same glycan. This is overcome through reduction of the reducing end using NaBH4 to yield the alditol. PGC provides effective separation of structural isomers (15). Analytical separations of N- and O-glycans, as well as of milk oligosaccharides, have recently been reported (25, 26, 28, 57–63).

The application of this analytical method to serum samples allowed the detection of over 300 glycan species, among which were several isomeric structures (26) as shown in Fig. 2A, which is an annotated chromatogram of N-glycans released from serum. Furthermore, the high retention time and peak area repeatability showed it to be highly suitable for biomarker discovery (25, 26). Application of the nLC-PGC-chip-TOF-MS method to a set of serum samples from prostate cancer patients with relatively poor and relatively good prognosis revealed that poor prognosis is associated with decreased levels of fucosylated glycans and increased levels of sialylated compounds. Further analysis showed that levels of 15 glycan structures, originating from nine glycan compositions, were significantly altered between individuals with poor prognoses and those with good prognoses. Analysis of isomers rather than composition yielded significantly greater differentiation between disease and control samples (Figs. 2B and 2C). We have recently developed a library of N-glycan structures based on the 10 most abundant glycoproteins in serum as the core of a functional glycan database for structural identification (58). Using retention time, accurate mass, and fragmentation spectrum, this library will allow the unambiguous identification of specific N-glycan structures in serum and in other human tissues.

Other chip-based chromatographic methods such as reverse phase HPLC for permethylated glycans coupled to ion trap MS have recently been described (64). The methods show very good reproducibility with high cv values, but only limited separation was obtained with 18 glycan structures observed. Also, the separation of structural isomers was not reported. The separation of permethylated glycans on PGC stationary phase has previously been shown to show some isomer separation (65), and perhaps this method could be improved to yield more effective separation. Application of the nano-reverse-phase LC/MS method to a small sample set from late-stage breast cancer patients and controls revealed increased levels of fucosylated glycans, whereas sialylated glycans were decreased with advanced-stage breast cancer. At the level of individual glycans, a biantennary glycan (Hex5HexNAc5Fuc1) showed decreased levels, whereas a fucosylated and sialylated triantennary glycan (Hex9HexNAc5Fuc1) was increased, with late-stage breast cancer (64).

Thus far, compound-specific glycan analysis has been performed on limited numbers of samples. There are specific issues associated with oligosaccharide analysis that make it unique in contrast to classes of biological compounds such as proteins and metabolites. Oligosaccharides are like peptides in that they have sequence; however, unlike that of peptides, their sequence information is insufficient to yield compound identity. Chromatography is important and needed to yield
compound identity as it is in metabolomics. Therefore, effective separation through high-resolution chromatography, high-sensitivity through nanoflow methods, and peak alignment are necessary. However, unlike metabolomics data, chromatographic data, in both positive and negative MS modes, are characterized by multiply charged species, which further complicates peak assignments. Current efforts in overcoming these problems are progressing, but the methods are still not readily available. Nonetheless, compound-specific biomarker discovery employing larger numbers of samples is currently being performed. Compound profiles are expected to yield more specific biomarkers than the compositional biomarkers.

**Glycan Biomarkers**—The serum glycome patterns associated with ovarian, gastric, pancreatic, lung, liver, and breast cancer have been studied using the methods described above, and a non-exhaustive overview of the results obtained is provided in Table I. Generally, the individual platforms seem to have some bias toward specific structural features: high-mannose-type glycans were mostly shown to be altered using MALDI-FTICR-MS of fractionated glycans (7, 20), whereas glycan structures containing sialyl Lewis x were most clearly identified as candidate biomarkers using HILIC-HPLC with fluorescence detection (44–47). However, glycans that carry sialyl Lewis x structures and truncated neutral glycans, especially Hex$_3$HexNAc$_5$Fuc$_1$, were shown to increase with several types
The results indicate that these alterations include a more systemic reaction secondary to the malignant process. Truncated neutral glycans are carried by immunoglobulin G (40, 66, 67), whereas the larger compounds carrying sialyl Lewis x have been associated with acute phase proteins such as haptoglobin and alpha-1-antitrypsin (18), suggesting the involvement of the immune response in the altered glycosylation profiles.

The enzymatic trimming of high-mannose glycans and the subsequent build-up of hybrid and complex-type N-glycans

TABLE I
Overview of glycan biomarker studies performed on whole serum or plasma

<table>
<thead>
<tr>
<th>Cancer type</th>
<th>Number of samples</th>
<th>Analytical method</th>
<th>Findings</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast</td>
<td>Cases: 7, Controls: 5</td>
<td>MALDI-FTICR-MS of native glycans</td>
<td>Increased levels of high-mannose-type structures</td>
<td>(7)</td>
</tr>
<tr>
<td>Breast</td>
<td>Cases: 15, Controls: 15</td>
<td>nano-reverse phase LC/MS of permethylated glycans</td>
<td>Increased levels of fucosylated glycans and Decreased levels of sialylated glycans</td>
<td>(64)</td>
</tr>
<tr>
<td>Breast</td>
<td>Cases: 27, Controls: 13</td>
<td>HILIC-HPLC of 2-AB-labeled glycans</td>
<td>Increased levels of highly sialylated antenna fucosylated glycans</td>
<td>(47)</td>
</tr>
<tr>
<td>Breast</td>
<td>Cases: 82, Controls: 27</td>
<td>MALDI-TOF of permethylated glycans</td>
<td>Increased levels of fucosylated and sialylated glycans</td>
<td>(38)</td>
</tr>
<tr>
<td>Gastric</td>
<td>Cases: 80, Controls: 20</td>
<td>HILIC-UPLC of 2-AB-labeled glycans</td>
<td>Decreased levels of fucosylated non- and mono-sialylated glycans</td>
<td>(42)</td>
</tr>
<tr>
<td>Gastric</td>
<td>Cases: 80, Controls: 20</td>
<td>HILIC-HPLC of 2-AB-labeled glycans</td>
<td>Decreased levels of biantennary asialo monogalactosylated glycans and triantennary glycans carrying α2–3-linked sialic acids Increased levels of triantennary glycans carrying α2–6-linked sialic acids and trisialylated triantennary glycans carrying sialyl Lewis X</td>
<td>(44)</td>
</tr>
<tr>
<td>Hepatocellular carcinoma</td>
<td>Cases: 72, Controls: 77</td>
<td>MALDI-TOF of permethylated glycans</td>
<td>Decreased levels of a tri- and a tetra-antennary glycan</td>
<td>(39)</td>
</tr>
<tr>
<td>Hepatocellular carcinoma</td>
<td>Cases: 227, Controls: 80</td>
<td>Capillary gel electrophoresis with laser-induced fluorescence detection of aminopyrenetrisulfonic-acid-labeled glycan after desialylation</td>
<td>Increased levels of the fucosylated triantennary glycan Hex₅HexNAc₅fuc₃</td>
<td>(52)</td>
</tr>
<tr>
<td>Lung</td>
<td>Cases: 100, Controls: 84</td>
<td>HILIC-HPLC of 2-AB-labeled glycans</td>
<td>Increased levels of tri- and tetra-antennary highly sialylated glycans, some with antenna and some with core fucosylation Decreased levels of biantennary glycans, mostly with core fucose</td>
<td>(45)</td>
</tr>
<tr>
<td>Ovarian</td>
<td>Cases: 5, Control: 5</td>
<td>MALDI-TOF-MS of desialylated glycans</td>
<td>Increased levels of Hex₅HexNAc₅Fuc₃ and Hex₅HexNAc₅Fuc₃</td>
<td>(36)</td>
</tr>
<tr>
<td>Ovarian</td>
<td>Cases: 27, Controls: 7</td>
<td>HILIC-HPLC of 2-AB-labeled glycans</td>
<td>Increased levels of core fucosylated agalactosyl biantennary glycans and glycans containing sialyl Lewis X</td>
<td>(46)</td>
</tr>
<tr>
<td>Ovarian</td>
<td>Cases: 19, Controls: 20</td>
<td>MALDI-TOF of permethylated glycans</td>
<td>Increased levels of tri- and tetra-antennary N-glycans Decreased levels of glycans containing a bisecting GlcNAc</td>
<td>(40)</td>
</tr>
<tr>
<td>Ovarian</td>
<td>Cases: 46, Controls: 48</td>
<td>MALDI-FTICR-MS of native glycans</td>
<td>Increased levels of sialylated glycans and a small group of truncated glycans Decreased levels of several neutral glycans, including high-mannose-type glycans</td>
<td>(20)</td>
</tr>
<tr>
<td>Prostate</td>
<td>Good prognosis: 4, Poor prognosis: 4</td>
<td>nLC-PGC-TOF of native glycans</td>
<td>Poor prognosis is associated with decreased levels of fucosylated glycans and increased levels of sialylated compounds</td>
<td>(26)</td>
</tr>
</tbody>
</table>
Glycan Markers for Cancer Detection

Fig. 3. Glycans shown to be altered in ovarian cancer, as grouped by family. This overview shows that glycans altered significantly are closely linked to each other and may be grouped in families that are either increasing (Family 1) or decreasing (Families 2, 3, and 4) with ovarian cancer. For a compositional key, please see Fig. 1.

follows standard biological pathways in protein glycosylation. Because of the substrate specificity of the enzymes, this process usually takes place one monosaccharide at a time. Because each nascent glycan structure functions as the substrate for a glycosyltransferase and/or glycosidase, it is to be expected that levels of certain glycan structures will be closely correlated with other, related structures. The resulting glycan structures can be grouped into homologous families. We exploited this principle retrosynthetically to build an N-glycan library of hypothetical serum glycans (10). Indeed, it has recently been shown that several networks generated by using 34 of the 39 glycans that were altered in ovarian cancer patients yielded strong correlations between structurally related glycans (20). It is this pattern of global glycan changes that will help to define the complexity of glycan changes in a systematic way. The detection of such strongly correlated glycan networks shows that not only one glycan structure might be altered, but a whole family of structures might be affected (see Fig. 3).

Although the studies summarized in this review aimed to identify biomarkers for cancer and provide good evidence in support of candidate biomarkers, most of the studies performed so far have not been forward in a second testing set. Furthermore, in most cases, no cut-off values were determined that would be used in later validation studies. In order to further evaluate the potential of glycan-based biomarkers for the detection of cancer, such studies, as well as further validation studies assessing the specificity and sensitivity in less well-defined datasets, will have to be performed according to study design strategies proposed by the National Institutes of Health Early Detection Research Network (68).

CONCLUSION AND FUTURE PERSPECTIVES

The various levels of complexity associated with protein glycosylation have hindered the field of glycobiology. However, a new and unique paradigm for biomarker assays is now taking shape. Altered patterns of glycosylation are observed in terms of gross compositions of the glycans, with specific glycan compounds, and with glycan-associated proteins. These alterations in glycosylation can be associated with cancer either as a product of the tumor or as a reaction to the disease. Although it is still not apparent how the field of cancer glycan biomarkers will evolve, a likely scenario would be to use the various levels of complexity for disease diagnosis. Glycan compositional analysis may be used to detect the disease, and a more in-depth test—either structure- or protein-site-specific—could provide greater sensitivity and specificity, information about the tissue location, and details of the phenotype allowing one to determine the best potential treatment. Further advancements of mass spectrometric techniques and further maturation of the application of such techniques in the field of glycosylation analysis, and particularly in biomarker discovery and validation, are needed to enable such strategies.

LC/MS analysis of oligosaccharides generates complex data that require specific data-processing methods unique to glycans that have not yet been fully optimized, such as simultaneous deconvolution of the charge states, alignment and peak selection in the LC, and sequence analysis mass spectrometry. Efforts to this end are under way (20, 69, 70). It is expected that bioinformatics solutions specific to glycan analysis using LC/MS will progress toward facilitating the data analysis of large-sample glycomics datasets.

Plasma and serum are complex protein mixtures dominated by a few high-abundance proteins (71). Glycan changes in the global profile may come from some of these proteins, and protein-specific glycosylation studies will link the glycan profile analysis with the large body of research on protein biomarkers. The serum/plasma N-glycan patterns are dominated by glycans derived from the top 5% to 10% of abundant glycoproteins, including, but not limited to, IgG, transferrin, haptoglobin, IgA, alpha-1-antitrypsin, and alpha-2-macroglobulin. More targeted approaches have the potential to improve the specificity of glycan biomarkers. Indeed, several studies have already identified altered glycosylation patterns on IgG with gastric (72) and ovarian (40, 46) cancer. Other studies have focused on the acute phase proteins haptoglobin (45, 46, 73), alpha-1-antitrypsin (74), and alpha-1-acid glycoprotein (46, 73). The recent advances in mass spectrometric techniques and their application to glycomics biomarker research now also allow one to search for protein- and site-specific glycan biomarkers. A method for the site-specific analysis of the glycosylation profile of simple protein mixtures using protease digestion and subsequent analysis using nLC-PGC-qTOF-MS/MS has been developed (29–31). The application of this method to immunopurified glycoproteins would allow the evaluation of site-specific biomarkers. Site-specific glycosylation patterns are likely to provide improved specificity for cancer because of the combination of glycan and glycosylation site.
The proteins targeted so far have mostly been higher abundant proteins, primarily because glycosylation analysis of lower abundant proteins is more challenging. Recently, several methods for the depletion of higher abundant proteins have been applied to serum or plasma. Interestingly, only limited relative changes in the glycosylation profile between the total protein pool and the lower abundant proteins were observed (75, 76). The further improvement and employment of depletion tools in glycan biomarker studies might yield better identification of cancer-specific biomarkers.

Overall, the recent advancements in the identification of glycan-based biomarkers for the detection of cancer would not have been possible without advancements in separation and mass spectrometry techniques. These developments will soon allow the structure-specific and site-specific detection of glycans. It is possible that future diagnostic tests will include every aspect of glycan analysis—glycan compositions for the general detection of the disease, glycan compounds, and site-specific information to provide insight into the location and severity of the disease. Some combination might further yield the patient’s therapeutic phenotype for more effective treatments. In any case, these methods will provide significant insight into the glycobiological aspects of the disease process.

Acknowledgments—The authors are thankful for the funding provided by the National Institutes of Health, the Department of Defense, the Tobacco Related Disease Research Program and the LUNGevity Foundation.

§ To whom correspondence should be addressed: Carlito B. Lebrilla, Department of Chemistry, University of California, Davis, One Shields Avenue, Davis, CA 95616. Tel.: 1-530-752-0504. Fax: 1-530-754-8995, E-mail: cblebrilla@ucdavis.edu.

REFERENCES

Glycan Markers for Cancer Detection

profiling of dried blood spots. Anal. Chem. 84, 396–402


